

VOLUME ONE

ENCYCLOPEDIA OF STRESS

SECOND EDITION

EDITOR-IN-CHIEF
GEORGE FINK



DEDICATION

For Ann Elizabeth

EDITOR-IN-CHIEF

George Fink

Professorial Research Fellow (formerly Director)
Mental Health Research Institute of Victoria
Parkville, Melbourne, Victoria
Australia
Formerly Director, MRC Brain Metabolism Unit
Edinburgh, Scotland, UK.

ASSOCIATE EDITORS

Bruce McEwen

The Rockefeller University
Laboratory of Neuroendocrinology
1230 York Avenue, Box 165
New York NY 10021

E. Ronald de Kloet

Universiteit Leiden
Natuurwetenschappen Wiskunde en
LACDR Medical Pharmacology
Einsteinweg 55, kamer HB 902
Leiden 2333 CC
The Netherlands

Robert Rubin

Department of Psychiatry (116A)
VA Greater LA Healthcare System
11301 Wilshire Blvd.
Los Angeles, CA 90073

George Chrousos

First Department of Pediatrics,
Athens University Medical School,
Aghia Sophia Children's Hospital,
115 27 Athens, Greece

Andrew Steptoe

University College London
Department of Epidemiology and Public Health
Psychology
1-19 Torrington Place London WC1E 6BT
UK

Noel Rose

Johns Hopkins University
Center for Autoimmune Disease Research
Molecular Microbiology and Immunology
615 North Wolfe Street, Rm E5014
Baltimore MD 21205

Ian Craig

SGDP Research Centre
Institute of Psychiatry, Molecular Genetics Group
De Crespigny Park, Box P082
Denmark Hill
London SE5 8AF
UK

Giora Feuerstein

Merck Research Laboratories 42-209
Department of Cardiovascular Diseases
770 Sumneytown Pike
West Point PA 19486

CONTENTS

<i>Contents by Subject area</i>	<i>xxiii</i>
<i>Preface to First Edition</i>	<i>xxxi</i>
<i>Preface to the Second Edition</i>	<i>xxxiii</i>
<i>Guide to Encyclopedia</i>	<i>xxxv</i>
<i>Foreword</i>	<i>xxxvii</i>

VOLUME 1

A

Acute Stress Disorder and Posttraumatic Stress Disorder	1	Adrenal Medulla	52
R. Yehuda and C. M. Wong		R. Kvetnansky and R. McCarty	
Acute Stress Response: Experimental	7	Adrenaline	60
K. Pacak and R. McCarty		T. M. Pollard	
Acute Trauma Response	15	Adrenocortical Function, Factors Controlling Development Thereof	64
W. C. Chiu, D. E. Carlson and M. P. Lilly		T. Else and G. D. Hammer	
Adenylyl Cyclases and Stress Response	21	Adrenocorticotrophic Hormone (ACTH)	69
F. A. Antoni		M. E. Rhodes	
Adjustment Disorders	24	Aerobic Exercise and Stress Reduction	73
M. Dascalu and D. Svrakic		E. J. C. de Geus and J. H. Stubbe	
Adolescence	28	Affective Disorders	78
G. N. Swanson		D. F. MacKinnon	
Adolescent Suicide	36	Aggression	84
M. Berk, R. Suddath and M. Devich-Navarro		E. F. Coccaro and E. C. Manning	
Adrenal Cortex	38	Aggressive Behavior	89
G. P. Vinson, B. J. Whitehouse and J. P. Hinson		J. M. Koolhaas	
Adrenal Insufficiency	47	Aging and Adrenocortical Factors	92
H. S. Willenberg, S. R. Bornstein and G. P. Chrousos		C. Lord and J. C. Pruessner	
		Aging and Psychological Stress	96
		B. W. J. Penninx	
		Aging and Stress, Biology of	102
		M. A. Horan, R. N. Barton and G. J. Lithgow	
		AIDS	108
		M. H. Antoni and D. G. Cruess	
		Airline Accidents	114
		G. Li	
		Alarm Phase and General Adaptation Syndrome	119
		R. McCarty and K. Pacak	

Bereavement	317	Caregivers, Stress and	416
P. J. Clayton		S. H. Zarit, K. Bottigi and J. E. Gaugler	
Beta-Adrenergic Blockers	323	Catecholamines	419
M. B. Hamner and G. W. Arana		U. Lundberg	
Beta-Endorphin	332	Central Stress Neurocircuits	424
M. Lee and S. L. Wardlaw		S. Kollack-Walker, H. E. W. Day and H. Akil	
Blood Pressure	335	Cerebral Metabolism, Brain Imaging	432
A. Sherwood and R. A. Carels		K. P. Ebmeier, C. L. Donaghey and N. J. Dougall	
Blood-Brain Barrier, Stress and	342	Chaperone Proteins and Chaperonopathies	438
S. N. Malaeb and B. S. Stonestreet		A. J. L. Macario and E. Conway de Macario	
Borderline Personality Disorder	348	Chaperonopathies	444
H. W. Koenigsberg and L. J. Siever		A. J. L. Macario and E. Conway de Macario	
Brain and Brain Regions	351	Chemical Warfare	449
A. G. Watts		J. Berberich	
Brain Natriuretic Peptide (BNP)	357	Chernobyl, Stress Effects of	452
P. Pervanidou and G. P. Chrousos		A. Tønnessen and L. Weisæth	
Brain Trauma	360	Child Abuse	457
B. Pentland		C. C. Swenson and L. Saldana	
Breast Cancer	364	Child Physical Abuse	460
A. Moyer		C. C. Swenson	
Burnout	368	Child Sexual Abuse	463
C. Maslach and M. P. Leiter		J. A. Cohen	
		Childbirth and Stress	467
		S. Ayers and E. Ford	
		Childhood Stress	472
		S. Sandberg	
		Cholesterol and Lipoproteins	478
		C. M. Stoney	
		Chronic Fatigue Syndrome	484
		A. J. Cleare and S. Wessely	
		Chronic Social Stress: GR Sensitivity in Leukocytes	493
		A. Weizman and B. Rotberg	
		Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress	496
		M. P. Antoch and R. V. Kondratov	
		Circadian Rhythm Effects on Cardiovascular and Other Stress-Related Events	500
		R. Manfredini, B. Boari, R. Salmi, A. M. Malagoni and F. Manfredini	
Calbindin	373		
F. A. Antoni			
Calcium, Role of	374		
F. A. Antoni			
Calcium-Dependent Neurotoxicity	375		
J. S. Kelly			
Cancer	378		
D. Spiegel			
Cancer Treatment	384		
F. I. Fawzy, A. L. Canada and N. W. Fawzy			
Captivity, Adaptation to	388		
R. H. Rahe			
Captivity, Recovery from	392		
R. H. Rahe			
Cardiovascular System and Stress	396		
P. Hjelm Dahl			
Cardiovascular Disease, Stress and	410		
G. P. Chrousos and G. Kaltsas			

C

Circadian Rhythms, Effects of Prenatal Stress in Rodents	505	Corticosteroids and Stress	613
S. Maccari and O. Van Reeth		A. Munck	
Circadian Rhythms, Genetics of	508	Corticotropin Releasing Hormone (CRH)	620
F. W. Turek and M. H. Vitaterna		A. T. Lim	
Cognition and Stress	513	Corticotropin Releasing Factor Receptor Deficiency in Mice	623
M. W. Eysenck		N. J. Justice and K.-F. Lee	
Cognitive-Behavioral Therapy	515	Corticotropin Releasing Factor-Binding Protein	627
G. A. Fava		P. J. Lowry, C. F. Kemp and R. J. Woods	
Combat Reaction, Chronic	518	Corticotropin-Releasing Factor Circuitry in the Brain – Relevance for Affective Disorders and Anxiety	630
R. H. Rahe		D. A. Gutman and C. B. Nemeroff	
Combat Stress Reaction	524	Corticotropin-Releasing Factor (CRF) Family of Neuropeptides – Role in Inflammation	635
M. Dobson		A. Gravanis and A. N. Margioris	
Combat, Acute Reactions to	529	Corticotropin-Releasing Factor Receptors	641
R. H. Rahe		D. E. Grigoriadis	
Common Cold and Stress	533	Cortisol Awakening Response	649
A. Smith		A. Steptoe	
Community Studies	536	C-Reactive Protein	653
C. J. Holahan, R. H. Moos and L. M. Groesz		W. J. Kop and A. A. Weinstein	
Comorbid Disorders and Stress	542	Crime Victims	659
S. Sundram and Avril Pereira		I. Robbins	
Comparative Anatomy and Physiology	549	Crisis Intervention	662
A. G. Watts		D. Hamaoka, D. Benedek, T. Grieger and R. J. Ursano	
Concentration Camp Survivors	553	Critical Thermal Limits	667
J. D. Kinzie		J. Roth	
Congenital Adrenal Hyperplasia (CAH)	556	Crowding Stress	669
A. Solomon and P.-M. G. Bouloux		L. Kovács and P. Csermely	
Conservation of Resources Theory	562	Cultural Factors in Stress	672
S. E. Hobfoll and J. S. Ford		J. W. Berry and B. Ataca	
Control and Stress	568	Cultural Transition	678
A. Steptoe		M. S. Kopp	
Coping and Stress: A Lens and Filter Model	574	Cushing's Syndrome, Medical Aspects	682
R. H. Rahe		S. R. Bornstein, M. Gruber, H. S. Willenberg, C. A. Stratakis and G. P. Chrousos	
Coping Skills	578	Cushing's Syndrome, Neuropsychiatric Aspects	688
A. DeLongis and E. Puterman		M. N. Starkman	
Corticosteroid Receptor Genes: Functional Dissection in Mice	584	Cytokines	692
F. Tronche		G. D. Marshall	
Corticosteroid Receptors	594		
O. C. Meijer, E. R. de Kloet and B. S. McEwen			
Corticosteroid-Binding Globulin (Transcortin)	605		
B. E. P. Murphy			

Cytokines, Chronic Stress, and Fatigue S. Jain and P. J. Mills	698	Diet and Stress, Non-Psychiatric J. Wardle and E. L. Gibson	797
Cytokines, Stress, and Depression B. E. Leonard and C. Song	705	Diet and Stress, Psychiatric V. March and M. H. Fernstrom	806
Cytotoxic Lymphocytes M. A. Fletcher and N. G. Klimas	711	Disaster Syndrome P. Valent	811
D		Disasters and Mass Violence, Public, Effects of G. Stevens, B. Raphael and M. Dobson	814
Death Anxiety R. Kastenbaum	717	Disease, Stress Induced H. S. Willenberg, S. R. Bornstein and G. P. Chrousos	824
Defensive Behaviors D. C. Blanchard, M. Yang, M. Hebert and R. J. Blanchard	722	Dissociation J. R. Maldonado	828
Demand–Control Model T. Theorell	727	Distress G. Matthews	838
Dental Stress T. K. Fábán, G. Fábán and P. Fejérdy	733	Divorce, Children of K. N. Hipke, S. A. Wolchik and I. N. Sandler	844
Depersonalization: Systematic Assessment M. Steinberg	736	Domestic Violence B. Donohue, H. Hill and T. Maier-Paarlberg	848
Depression and Coronary Heart Disease F. Lespérance and N. Frasure-Smith	741	Dopamine, Central G. D. Stanwood	852
Depression and Manic–Depressive Illness R. T. Rubin and B. J. Carroll	744	<i>Drosophila</i> Genes and Anoxia G. G. Haddad	859
Depression and Stress, Role of n-3 and n-6 Fatty Acids C. Song and B. E. Leonard	754	<i>Drosophila</i> Studies M. Allikian and J. Tower	864
Depression Models K. Matthews and C. Stewart	760	Drug Use and Abuse J. R. Mantsch	866
Depression, Immunological Aspects M. R. Irwin	766	E	
Dermatological Conditions M. A. Gupta	773	Earthquakes, Stress Effects of M. Livanou and M. Başoğlu	871
Desensitization F. A. Antoni	778	Eating Disorders and Stress D. C. Jimerson	876
Dexamethasone Suppression Test (DST) R. T. Rubin and B. J. Carroll	780	Eclampsia and Pre-Eclampsia A. Makrigiannakis, G. Petsas and G. P. Chrousos	880
DEX-CRH Test N. C. Schommer and I. Heuser	784	Economic Factors and Stress R. A. Catalano	884
DHEA J. Herbert	788	Education Levels and Stress J. Mirowsky and C. E. Ross	888
Diabetes, Type 1 A. Riazi and C. Bradley	792		

Food Shift Effect F. K. Stephan	88	Glucocorticoid Receptor Mutations and Polymorphisms J. W. Koper	183
Freud, Sigmund D. J. Lynn	90	Glucocorticoids – Adverse Effects on the Nervous System R. M. Sapolsky	185
G		Glucocorticoids, Effects of Stress on J. C. Buckingham	190
GABA (Gamma Aminobutyric Acid) J. D. C. Lambert	97	Glucocorticoids, Overview B. E. Pearson Murphy	198
Gastrointestinal Effects R. Murison and A. M. Milde	109	Glucocorticoids, Role in Stress J. C. Buckingham	210
Gender and Stress R. J. Handa and W. C. J. Chung	115	Glucose Transport A. L. McCall	217
Gene–Environment Interactions in Early Development I. S. P. Davis and R. Plomin	122	Glycobiology of Stress G. Lauc and M. Flögel	222
Genetic Factors and Stress C. A. Koch and C. A. Stratakis	128	Gonadotropin Secretion, Effects of Stress on M. Ferin	228
Genetic Polymorphisms in Stress Response I. W. Craig	135	Graves' Disease (Thyrotoxicosis) W. M. Wiersinga	234
Genetic Predispositions to Stressful Conditions D. Blackwood and H. Knight	141	Grieving A. Ray and H. Prigerson	238
Genetic Testing and Stress V. Senior and M. Cropley	146	Group Therapy N. Wong	242
Genetic Variation of HPA Axis Activity and Function in Farm Animals P. Mormède	150	Gulf War Syndrome, Psychological and Chemical Stressors H. Soreq	248
Ghrelin and Stress Protection T. Brzozowski, M. Pawlik, D. Drozdowicz, Z. Sliwowski, S. J. Konturek, W. W. Pawlik and P. C. Konturek	153	H	
Glia or Neuroglia G. W. Bennett and D. E. Ray	161	Health and Socioeconomic Status T. Chandola and M. Marmot	255
Glucocorticoid Effects on Memory: the Positive and the Negative O. T. Wolf	166	Health Behavior and Stress A. Steptoe	262
Glucocorticoid Negative Feedback M. F. Dallman	172	Heart Disease/Attack G. J. Baker, S. Suchday and D. S. Krantz	266
Glucocorticoid Receptor Mutant Mice as Models for Stress-Induced Affective Disorders P. Gass	176	Heart Failure, Stress Effects M. Alevizaki and G. P. Chrousos	272
		Heart Rate J. R. Jennings	274
		Heat Resistance A. J. L. Macario and E. Conway de Macario	278

Heat Shock Genes, Human	284	Hyperthermia	381
A. J. L. Macario and E. Conway de Macario		J. Roth	
Heat Shock Proteins: HSP60 Family Genes	288	Hyperthyroidism	388
H. Kubota		I. M. Lesser and D. L. Flores	
Heat Shock Response, Overview	292	Hyperventilation	390
A. J. L. Macario and E. Conway de Macario		C. Bass	
Hemostasis and Stress	300	Hypnosis	394
Roland von Känel		W. G. Whitehouse, E. C. Orne and M. T. Orne	
Herpesviruses	305	Hypocortisolism and Stress	400
D. A. Padgett, M. T. Bailey and J. F. Sheridan		C. M. Heim and U. M. Nater	
Hippocampal Neurons	311	Hypoglycemia	408
R. L. Spencer and S. T. Bland		B. M. Frier	
Hippocampus, Corticosteroid Effects on	321	Hypotension, Hypovolemia, and Septic Shock	413
M. Joëls and H. Karst		A. Beishuizen, A. B. Johan Groeneveld and I. Vermes	
Hippocampus, Overview	327	Hypothalamic-Pituitary-Adrenal Axis	421
M. J. Meaney and S. J. Lupien		M. F. Dallman, S. Bhatnagar and V. Viau	
Hiroshima Bombing, Stress Effects of	332	Hypothermia	428
R. J. Lifton		M. J. Taylor	
HIV Infection/AIDS	336	Hypothyroidism	439
B. W. Dixon		R. T. Joffe	
Holocaust Survivors, Experiences of	339	Hysteria	442
P. Valent		K. Pajer	
Holocaust, Stress Effects of	342		
P. Valent			
Homeostasis	347		
B. S. McEwen			
Homosexuality, Stress and	348	Immobilization Stress	445
J. Drescher		R. Kvetnansky and R. McCarty	
Hostility	354	Immune Cell Distribution, Effects of Stress on	449
L. H. Powell and K. Williams		F. S. Dhabhar	
HPA Alterations in PTSD	359	Immune Function, Stress-Induced Enhancement	455
R. Yehuda		F. S. Dhabhar	
Hurricane Katrina Disaster, Stress Effects of	364	Immune Response	462
C. Piotrowski		P. J. Delves	
11β-Hydroxysteroid Dehydrogenases	368	Immune Suppression	470
J. R. Seckl		P. Prolo and F. Chiappelli	
Hyperreactivity (Cardiovascular)	372	Immune Surveillance – Cancer, Effects of Stress on	477
A. Georgiades		D. Spiegel and F. S. Dhabhar	
Hypertension	376	Immune System, Aging	481
A. Steptoe		M. A. Horan	

Immunity	485	L	
F. Chiappelli			
Impact of Terrorism on the Development of Mental Health Symptoms	493	Learned Helplessness	567
R. Yehuda		D. M. Isaacowitz and M. E. P. Seligman	
Impotence, Stress and	497	Learning and Memory, Effects of Stress on	571
M. R. Gignac, G. M. Rooker and J. K. Cohen		M. Lindau, O. Almkvist and A. H. Mohammed	
Impulse Control	500	Left Ventricular Mass	577
I.-M. Blackburn		T. G. Pickering	
Incest	502	<i>Leishmania</i>, Stress Response in	579
A. P. Mannarino		M. Shapira	
Income Levels and Stress	506	Leptin, Adiponectin, Resistin, Ghrelin	584
S. V. Subramanian and I. Kawachi		A. Gavrilu, D. Barb and C. S. Mantzoros	
Indigenous Societies	511	Leukocyte Trafficking and Stress	592
W. W. Dressler		S. Hong, M. U. Goebel and P. J. Mills	
Industrialized Societies	517	Life Events and Health	599
J. Siegrist		P. Surtees and N. Wainwright	
Infection	521	Life Events Scale	603
H. Friedman and S. H. Pross		E. Wethington	
Inflammation	530	Lockerbie Air Crash, Stress Effects of	608
G. Z. Feuerstein, R. R. Ruffolo, C. Coughlin, J. Wang and D. Miller		H. Livingston and M. Livingston	
Instinct Theory	535	Loss Trauma	612
R. Gardner Jr.		A. Bifulco	
Integrative Medicine (Complementary and Alternative Medicine)	537	Lymph Nodes	616
D. Spiegel		T. L. Whiteside	
Interactions Between Stress and Drugs of Abuse	540	Lymphocytes	622
P. V. Piazza and M. Le Moal		N. R. Rose	
J		M	
Job Insecurity: The Health Effects of a Psychosocial Work Stressor	549	Macrophage Antimycobacterial Activity, Effects of Stress on	629
J. E. Ferrie and P. Martikainen		C. S. Boomershine and B. S. Zwilling	
K		Macrophages	634
Kidney Function	557	W. P. Fehder, F. Tuluc, W.-Z. Ho and S. D. Douglas	
W. J. Welch		Major Depressive Disorder	640
Korean Conflict, Stress Effects of	563	A. B. Negrão and P. W. Gold	
C. A. Goguen and M. J. Friedman		Male Partner Violence	645
		M. Ingram, N. P. Yuan and M. P. Koss	
		Marital Conflict	651
		P. T. McFarland and A. Christensen	

Marital Status and Health Problems	653	Mitochondria	754
I. M. A. Joung		I. Manoli, S. Alesci and G. P. Chrousos	
Marriage	660	Monoamine Oxidase	761
A. C. Yoneda and J. Davila		P. Huezo-Diaz and I. W. Craig	
Maternal Deprivation	667	Motor Vehicle Accidents, Stress Effects of	764
R. L. Huot, C. O. Ladd and P. M. Plotsky		T. C. Buckley and E. B. Blanchard	
Medical Profession and Stress	674	Mucosal Secretory Immunity, Stress and	768
K. G. Power and V. Swanson		J. A. Bosch and D. Carroll	
Meditation and Stress	678	Multi Drug Resistance P Glycoprotein and other Transporters	774
J. L. Kristeller		E. C. M. de Lange	
Membrane Glucocorticoid Receptors	686	Multiple Personality Disorder	783
P. J. Gasser and M. Orchinik		R. P. Kluft	
Memory and Stress	693	Multiple Sclerosis	790
S. J. Lupien and F. S. Maheu		A. T. Reder	
Memory Impairment	699	Multiple Trauma	795
A. J. Parkin [†]		P. N. Soucacos and E. O. Johnson	
Menopause and Stress	703	Musculoskeletal Problems and Stress	800
N. E. Avis		S. Svebak	
Menstrual Cycles and Stress	706	Myopathy	807
R. Suri and L. Altshuler		G. A. Small	
Mental Stress Testing	712		
P. G. Saab, K. A. Kline and J. R. McCalla			
Metabolic Syndrome	717	Natural Killer (NK) Cells	815
L. Keltikangas-Järvinen		T. L. Whiteside, M. Boyiadzis and R. B. Herberman	
Metabolic Syndrome and Stress	721	Negative Affect	822
R. Rosmond		A. A. Stone and A. A. Gorin	
Metals, Oxidative Stress, and Brain Biology	724	Neighborhood Stress and Health	825
D. I. Finkelstein, T. Lynch, S. Wilkins, R. A. Cherny and A. I. Bush		A. V. Diez Roux	
Metastasis	726	Nelson's Syndrome	828
M. K. Demetrikopoulos		A. Stathopoulou, K. Dimitriou and G. Kaltsas	
Metyrapone: Basic and Clinical Studies	730	Neural Stem Cells	832
E. A. Young		U. S. Sohur, J. G. Emsley, B. D. Mitchell and J. D. Macklis	
Migraine	733	Neurodegenerative Disorders	840
N. M. Ramadan		M. F. Mendez and A. M. McMurtry	
Mineralocorticoid Receptor Polymorphisms	744	Neurodevelopmental Disorders in Children	844
R. H. DeRijk and E. R. de Kloet		N. J. Rinehart and B. J. Tonge	
Minorities and Stress	748	Neuroendocrine Systems	851
I. Mino, W. E. Profit and C. M. Pierce		G. Fink	
		Neurogenesis	865
		P. Tanapat and E. Gould	

[†]Deceased.

Pheromones	119	Prison	217
A. Kumar, C. A. Dudley, S. Chakravarty and R. L. Moss		D. L. Whitehead and A. Steptoe	
Pituitary Regulation, Role of	127	Prisoners of War	223
F. A. Antoni		C. Tennant	
Police, Stress in	131	Problem-Solving Skills Training	227
R. J. Burke		A. M. Nezu, C. M. Nezu and T. J. D’Zurilla	
Posttraumatic Stress Disorder – Clinical	135	Prolactin and Stress	231
N. C. Feeny, L. R. Stines and E. B. Foa		G. Tolis, G. Rombopoulos, D. Kaltsas, E. Katounda, V. Kaltzidou and N. Angelopoulos	
Posttraumatic Stress Disorder – Neurobiological basis for	140	Pro-opiomelanocortin (POMC)	233
M. Barad		A. B. Bicknell	
Posttraumatic Stress Disorder in Children	145	Prostaglandins	240
A. M. La Greca		S. Moshonov, U. Zor and Z. Naor	
Posttraumatic Stress Disorder, Delayed	150	Proteases in Prokaryotes and Eukaryotic Cell Organelles	247
A. Holen		E. Conway de Macario and A. J. L. Macario	
Posttraumatic Stress Disorder, Neurobiology of	152	Proteases in the Eukaryotic Cell Cytosol	252
J. D. Bremner		E. Conway de Macario and A. J. L. Macario	
Posttraumatic Therapy	157	Protein Synthesis	258
A. M. Rasmusson, C. M. Monson and P. A. Resick		M. A. Brostrom and C. O. Brostrom	
Pregnancy – Maternal and Perinatal Stress – Effects of	165	Proteosome	266
P. Smirnaki and M.-A. Magiakou		M. Maldonado and J. Wang	
Premenstrual Dysphoric Disorder	173	Psoriasis	271
S. Nowakowski, P. Haynes and B. L. Parry		A. B-. Kirschbaum	
Pre-pulse Inhibition	180	Psychoanalysis	274
M. van den Buuse		P. Roazen	
Pressure, Effects of Extreme High and Low	184	Psychological Stressors, Overview	278
R. J. Værnes		S. M. Monroe and G. M. Slavich	
Primate Hierarchies and Personality	194	Psychoneuroimmunology	284
R. M. Sapolsky		R. Dantzer	
Primate Models, Behavioral– Immunological Interactions	199	Psychosocial Factors and Stress	288
M. L. Laudenslager and S. Tiefenbacher		J. Siegrist	
Primate Models, Cardiovascular Disease	204	Psychosomatic Heart Disease: Role of Sympathetic and Sympathoadrenal Processes	292
J. R. Kaplan		G. W. Lambert and T. Dawood	
Primate Models, Overview	211	Psychosomatic Medicine	296
D. M. Lyons		T. Theorell	
Primates: Rearing and Effects of Stress on Primate CNS Function	214	Psychotherapy	302
T. K. Newman and C. S. Barr		G. C. Smith	

Psychotic Disorders J. Ventura	308	Revenge Fantasies M. Horowitz and S. Meffert	391
Q		Rheumatic Disorders A. T. Masi and J. C. Aldag	395
Quality of Life S. M. Skevington	317	S	
R		Salivary Cortisol C. Kirschbaum and D. H. Hellhammer	405
Racial Harassment/Discrimination D. R. Williams and S. A. Mohammed	321	Salt Appetite R. S. Weisinger and N. Chen	409
Recovery from Stress S. Sonnentag and C. Fritz	327	Schizophrenia M. van den Buuse and D. Copolov	418
Reductive Stress J. P. Kehrer	331	School Stress and School Refusal Behavior C. A. Kearney, L. C. Cook and G. Chapman	422
Reenactment Techniques N. Wong	336	School Violence and Bullying J. Juvonen	425
Refugees, Stress in J. D. Kinzie	338	Seasonal Changes in Stress Responses R. J. Nelson and L. B. Martin II	427
Regional Blood Flow, Stress Effects W. W. Blessing	341	Seasonal Rhythms L. M. Romero	432
Relaxation Techniques W. G. Whitehouse, E. C. Orne and M. T. Orne	345	Secretagogue G. Fink	435
Religion and Stress S. Packer	351	Selective Serotonin Reuptake Inhibitors (SSRIs) M. Gitlin	440
9/11, Religion and Stress S. Packer	357	Self-Esteem, Stress, and Emotion D. Roger	443
Remodelling of Neuronal Networks by Stress E. Fuchs and G. Flügge	364	Selye, Hans B. Tuchweber and P. Bois	448
Renal and Adrenocortical Actions of Dopamine B. C. Williams, Y.-C. Lo and S. W. Walker	371	Sepsis, Acute Respiratory Distress Syndrome, and Glucocorticoid Resistance G. U. Meduri	450
Reproduction, Effects of Social Stress On C. A. Shively	374	Serotonin B. C. Williams and Y.-C. Lo	457
Reproductive Dysfunction in Primates, Behaviorally Induced J. L. Cameron	380	Serotonin in Stress S. Kusljic and M. van den Buuse	461
Resistance L. M. Zabarenko	386	Serotonin Transporter Genetic Modifications K. Sugden	465
Restraint Stress R. J. Servatius, G. Salameh, K. M. Coyle and W. P. Paré	389	Sex Differences in Human Stress Response B. M. Kudielka, D. H. Hellhammer and C. Kirschbaum	469

Sex Steroids, Response to Stress and Susceptibility to Depression	474	Startle Response	561
M. V. Seeman and L. E. Ross		C. O. Ladd, P. M. Plotsky and M. Davis	
Sex-Specific Effects of Early Social Stress in Mammals: A Study in Guinea Pigs	479	Steroid Hormone Receptors	568
S. Kaiser and N. Sachser		M. Beato and J. Klug	
Sexual Assault	484	Steroid Hydroxylases	581
N. C. Feeny, T. J. Linares and E. B. Foa		F. H. de Jong	
Sexual Dysfunction	490	Strain Differences in Stress Response in Rodents	585
W. T. O'Donohue and L. Woodward Tolle		P. Mormède	
Sexual Offenders	494	Stress and Anxiety: Treatment with 2nd Generation Antipsychotic Drugs	587
F. M. Saleh and H. M. Malin		J. S. Ballon, J. A. Boyd and D. A. Wirshing	
Sickle Cell Disease and Stress	502	Stress and CNS Arousal: Genomic Contributions	591
K. Midence		A. C. Ribeiro and D. W. Pfaff	
Sleep Loss, Jet Lag, and Shift Work	504	Stress Effect of Assisted Reproduction	596
E. Van Cauter		F. M. Helmerhorst, R. Sibug and E. R. de Kloet	
Sleep, Sleep Disorders, and Stress	506	Stress Effects, Overview	599
A. N. Vgontzas, S. Pejovic and M. Karataraki		A. Steptoe	
Smoking and Stress	515	Stress Generation	601
F. J. McClernon and D. G. Gilbert		J. E. Roberts and J. A. Ciesla	
Social Capital	520	Stress Hyporesponsive Period	606
I. Kawachi		S. Levine, E. R. de Kloet, G. Dent and M. S. Schmidt	
Social Networks and Social Isolation	523	Stress in University Students	612
L. F. Berkman		B. Andrews and J. Hejdenberg	
Social Status and Stress	528	Stress Induced Anovulation	615
D. de Ridder		S. L. Berga and T. L. Loucks	
Social Stress, Animal Models of	533	Stress Management and Cardiovascular Disease	631
J. Bugajski, A. Gądek-Michalska and A. J. Bugajski		D. Lane and D. Carroll	
Social Support	539	Stress Management, CAM Approach	636
T. C. Antonucci, J. E. Lansford and K. J. Ajrouch		K. Krebs	
Social Support in Trauma	542	Stress of Self Esteem	640
K. O'Donnell and A. Steptoe		J. C. Pruessner, S. Wuethrich and M. W. Baldwin	
Somatic Disorders	545	Stress System Balance Hypothesis	646
F. Creed		E. R. de Kloet	
Space, Health Risks of	548	Stress, Beneficial Effects of	650
V. A. Convertino		S. Joseph and P. A. Linley	
Spinal Cord Injury, Physical Stress of	554	Stress, Definitions and Concepts of	653
S. Bhatia and M. Quigley		B. S. McEwen	
Spinal Cord Injury, Psychological Stress of	557		
S. D. Schnakenberg-Ott and M. R. Lovell			

Stress, Insulin Resistance, and Type II Diabetes	654	Torture	749
C. Tsigos and I. Kyrou		I. Genefke, H. Marcussen and O. V. Rasmussen	
Stress, NPY, and Cardiovascular Diseases	660	Transport-Related Stress	756
Z. Zukowska		R. G. Smart	
Suicide Terrorism, Genesis of	667	Trans-sexualism	760
A. Speckhard		R. A. Allison	
Suicide, Biology of	677	Trauma and Memory	765
M. A. Oquendo, L. Giner and J. J. Mann		B. A. van der Kolk	
Suicide, Psychology of	684	Trauma Group Therapy	767
D. Lester and R. L. Walker		J. Dwyer and L. G. Martin	
Suicide, Sociology of	689	Traumatic Stress and Posttraumatic Stress Disorder, the Israeli Experience	771
D. Lester and R. M. Fernquist		E. Klein and J. Zohar	
Surgery and Stress	693	Trier Social Stress Test	776
C. Vögele		B. M. Kudielka, S. Wüst, C. Kirschbaum and D. H. Hellhammer	
Survivor Guilt	695	Type A Personality, Type B Personality	782
P. Valent		W. S. Shaw and J. E. Dimsdale	
Sympathetic Nervous System	697		
D. S. Goldstein			
Synthetic Glucocorticoids	704		
A. M. Karssen and E. R. de Kloet			
Systemic Lupus Erythematosus	708		
D. J. Wallace			
		U	
		Ulceration, Gastric	787
		R. Murison and A. M. Milde	
		Ultradian Rhythms	791
		S. L. Lightman	
		Understimulation/Boredom	794
		V. J. Sutherland	
		Unemployment, Stress, and Health	797
		M. Bartley	
		Urocortins	804
		É. M. Fekete and E. P. Zorrilla	
		V	
		Vaccination	813
		V. E. Burns, A. C. Phillips and K. M. Edwards	
		Vasoactive Peptides	817
		W. K. Samson and M. M. White	
		Vasopressin	824
		L. P. Renaud	
		Vietnam Veterans, Postwar Experiences and Health Outcomes	830
		J. A. Boscarino	
Teaching and Stress	713		
E. R. Greenglass			
Temperature Effects	717		
J. Roth			
Terrorism	719		
P. Bell			
Thermal Stress	723		
C. M. Blatteis			
Thermotolerance, Thermoresistance, and Thermosensitivity	726		
S. Gatti McArthur, D. Alberati and T. Bartfai			
Three Mile Island, Stress Effects of	735		
A. L. Dougall and A. Baum			
Thymus	738		
M. S. Vacchio			
Thyroid Hormones	743		
T. J. Visser and E. Fliers			

Violence	838	War, Suicide and Sacrifice	860
E. K. Englander		V. Hazboun	
Viral Virulence and Stress	842	War-Related Posttraumatic Stress Disorder, Treatment of	865
A. J. L. Macario and E. Conway de Macario		L. B. Slone and M. J. Friedman	
Viruses and Stress	850	Work-Family Balance	868
G. Z. Feuerstein, R. R. Ruffolo and B.-N. David		J. G. Grzywacz and A. B. Butler	
		Workplace Stress	871
		U. Lundberg	

W

Waist-Hip Ratio	853
R. Rosmond	
War Stress in the Former Yugoslavia	855
M. Flögel, S. Šupraha Goreta and G. Lauc	

VOLUME 4

<i>Contributors</i>	1
<i>Subject Index</i>	21

CONTENTS BY SUBJECT AREA

ANIMAL STUDIES AND MODELS

- Animal Models (Nonprimate) for Human Stress
- Circadian Rhythms, Effects of Prenatal Stress in Rodents
- Fish, Stress in
- Immobilization Stress
- Pre-pulse inhibition
- Primate Hierarchies and Personality
- Primate Models, Behavioral-Immunological Interactions
- Primate Models, Cardiovascular Disease
- Primate Models, Overview
- Primates: Rearing and Effects of Stress on Primate CNS Function
- Sex-Specific Effects of Early Social Stress in Mammals: a Study in Guinea Pigs

CONFLICT, WAR, TERRORISM

- Chemical Warfare
- Gulf War Syndrome, Psychological and Chemical Stressors
- Hiroshima Bombing, Stress Effects of
- Holocaust, Stress Effects of
- Impact of Terrorism on the Development of Mental Health Symptoms
- Korean Conflict, Stress Effects of
- Lockerbie Air Crash, Stress Effects of
- Northern Ireland, Post Traumatic Stress Disorder in
- Nuclear Warfare, Threat of
- Oklahoma City Bombing, Stress Effects of
- Persian Gulf War, Stress Effects of
- Prisoners of War
- 9/11, Religion and Stress
- Suicide Terrorism, Genesis of
- Terrorism
- Torture

- Traumatic Stress and Posttraumatic Stress Disorder; the Israeli Experience
- Vietnam Veterans, Postwar Experiences and Health Outcomes
- War Stress in the Former Yugoslavia
- War, Suicide and Sacrifice

DISASTERS

- Airline Accidents
- Chernobyl, Stress Effects of
- Disasters and Mass Violence, Public, Effects of
- Disaster Syndrome
- Earthquakes, Stress Effects of
- Floods, Stress Effects of
- Hurricane Katrina Disaster, Stress Effects of
- Motor Vehicle Accidents, Stress Effects of
- Three Mile Island, Stress Effects of

DIURNAL, SEASONAL AND ULTRADIAN RHYTHMS

- Circadian rhythm effects on cardiovascular and other stress-related events
- Circadian Rhythms, Genetics of
- Night Shiftwork
- Seasonal Changes in Stress Responses
- Seasonal Rhythms
- Sleep Loss, Jet Lag, and Shift Work
- Sleep, Sleep Disorders, and Stress
- Ultradian Rhythms

DRUGS (EFFECTS)

- Alcohol, Alcoholism and Stress:
A Psychobiological Perspective
- Drug Use and Abuse
- Interactions between Stress and Drugs of Abuse
- Opioids

DRUGS (TREATMENT)

- Antipsychotic Drugs and Stress
- Anxiolytics
- Benzodiazepines
- Beta-Adrenergic Blockers
- Pharmacological Treatments of Stress
- Selective Serotonin Reuptake Inhibitors (SSRIs)
- Stress and Anxiety: treatment with 2nd generation antipsychotic drugs

GENERAL CONCEPTS AND MODELS

- Alarm Phase and General Adaptation Syndrome
- Allostasis and Allostatic Load
- Behavior, Overview
- Conservation of Resources Theory
- Control and Stress
- Demand–Control Model
- Effort–Reward Imbalance Model
- Environmental Factors
- Evolutionary Origins and Functions of the Stress Response
- Fight-or-Flight Response
- Instinct Theory
- Life Events Scale
- Psychological Stressors, Overview
- Remodelling of neuronal networks by stress
- Selye, Hans
- Stress, Definitions and Concepts of
- Stress Effects, Overview
- Stress System Balance Hypothesis
- Stress, Beneficial Effects of

GENETICS AND GENOMICS

- Circadian clock genes as modulators of sensitivity to genotoxic stress
- Corticosteroid Receptor Genes: Functional Dissection in Mice
- Corticotropin Releasing Factor Receptor Deficiency in Mice
- Drosophila Genes and Anoxia
- Expression Profiling of Stress Responsive Gene Patterns
- Gene Environment Interactions in Early Development
- Genetic Factors and Stress
- Genetic Polymorphisms in Stress Response
- Genetic Predispositions to Stressful Conditions
- Genetic Testing and Stress
- Genetic variation of HPA axis activity and function in farm animals
- Glucocorticoid Receptor Mutant Mice as Models for Stress-Induced Affective Disorders

- Glucocorticoid Receptor Mutations and Polymorphisms
- Mineralocorticoid Receptor Polymorphisms
- Neuroticism Genetic Mapping of
- Neuroticism Response to Stress, Genetic Mapping of Mice
- Serotonin Transporter Genetic Modifications
- Strain Differences in Stress Response in Rodents
- Stress and CNS Arousal; Genomic Contributions

HUMAN COGNITION, EMOTION AND BEHAVIOR

- Adolescence
- Aggression
- Aggressive Behavior
- Aging and Psychological Stress
- Anger
- Anxiety
- Bereavement
- Burnout
- Captivity, Adaptation to
- Captivity, Recovery from
- Caregivers, Stress and
- Childbirth and Stress
- Cognition and Stress
- Combat, Acute Reactions to
- Combat Reaction, Chronic
- Combat Stress Reaction
- Concentration Camp Survivors
- Coping Skills
- Coping and Stress: a Lens and Filter Model
- Cortisol Awakening Response
- Death Anxiety
- Dental Stress
- Depersonalization: Systematic Assessment
- Diet and Stress, Non-Psychiatric
- Dissociation
- Distress
- Emergency Personnel, Stress in
- Emotional Inhibition
- Emotions: Structure and Adaptive Functions
- Fatigue and Stress
- Fear
- Fear and the Amygdala
- Firefighters, Stress in
- Grieving
- Holocaust Survivors, Experiences of
- Homosexuality, Stress and
- Hostility
- Impotence, Stress and
- Impulse Control
- Industrialized Societies
- Learning and Memory, Effects of Stress on
- Marriage

Medical Profession and Stress
 Memory Impairment
 Menopause and Stress
 Neuroimaging and Emotion
 Nightmares
 Optimism, Pessimism and Stress
 Pain
 Parenting, Stress of
 Peacekeeping
 Police, Stress in
 Prison
 Psychosomatic Medicine
 Recovery from Stress
 Refugees, Stress in
 Religion and Stress
 Self-Esteem, Stress, and Emotion
 Somatic Disorders
 Stress Generation
 Stress in University Students
 Suicide, Psychology of
 Suicide, Sociology of
 Surgery and Stress
 Survivor Guilt
 Teaching and Stress
 Trauma and Memory
 Type A Personality, Type B Personality
 Understimulation/Boredom

HUMAN HEALTH AND PHYSICAL ILLNESS

Amnesia
 Adrenal Insufficiency
 AIDS
 Alzheimer's Disease
 Arthritis
 Asthma
 Atherosclerosis
 Attention-Deficit/Hyperactivity Disorder, Stress and
 Breast Cancer
 Cancer
 Cardiovascular Disease, stress and
 Chaperonopathies
 Chronic Fatigue Syndrome
 Common Cold and Stress
 C-Reactive Protein
 Cytokines, Chronic Stress, and Fatigue
 Cushing's Syndrome, Medical Aspects
 Cushing's Syndrome, Neuropsychiatric Aspects
 Dermatological Conditions
 DHEA
 Diabetes, Type I
 Disease, Stress Induced
 Eclampsia and preeclampsia
 Endometriosis
 Enuresis

Epilepsy
 Fibromyalgia
 Graves' Disease (Thyrotoxicosis)
 Health Behavior and Stress
 Heart Disease/Attack
 Heart Failure, Stress Effects
 Hemostasis and Stress
 Hyperreactivity (Cardiovascular)
 Hypertension
 Hyperthermia
 Hyperthyroidism
 Hyperventilation
 Hypoglycemia
 Hypotension, Hypovolemia, and Septic Shock
 Hypothermia
 Hypothyroidism
 Hysteria
 Metastasis
 Metastasis
 Migraine
 Multiple Trauma
 Nelson's Syndrome
 Neurodegenerative Disorders
 Organ Transplantation, Stress of
 Parkinson's Disease
 Perinatal Dexamethasone
 Pregnancy, Maternal and Perinatal Stress, Effects of
 Premenstrual Dysphoric Syndrome
 Prolactin and stress
 Psychosomatic heart disease: role of sympathetic
 and sympathoadrenal processes
 Rheumatic Disorders
 Sickle Cell Disease and Stress
 Space, Health Risks of
 Spinal Cord Injury, Physical Stress of
 Spinal Cord Injury, Psychological Stress of
 Stress Effect of Assisted Reproduction
 Stress Induced Anovulation
 Stress, Insulin Resistance and Type II Diabetes
 Stress, NPY and Cardiovascular Diseases
 Ulceration, Gastric

HUMAN MENTAL HEALTH AND PSYCHOPATHOLOGY

Acute Stress Disorder and Posttraumatic Stress
 Disorder
 Adjustment Disorders
 Adolescent suicide
 Affective Disorders
 Antisocial Disorders
 Anxiety
 Arthritis – Psychological
 Borderline Personality Disorder
 Child Abuse
 Child Sexual Abuse

Comorbid Disorders and Stress
Corticotropin-Releasing Factor Circuitry in the
Brain – Relevance for Affective Disorders
and Anxiety
Defensive Behaviors
Depression and Coronary Heart Disease
Depression and Manic-Depressive illness
Depression and Stress, Role of n-3 and n-6
Fatty Acids
Depression, Immunological Aspects
Depression Models
Diet and Stress, Psychiatric
Disaster Syndrome
Dissociation
Eating Disorders and Stress
Elder Abuse
Ethnicity, Mental Health
HPA alterations in PTSD
Hypocortisolism and Stress
Incest
Learned Helplessness
Life Events and Health
Loss Trauma
Male Partner Violence
Major Depressive Disorder
Multiple Personality Disorder
Negative Affect
Neurodevelopmental Disorders in Children
Neurosis
Obsessive–Compulsive Disorder
Panic Disorder and Agoraphobia
Paranoia
Posttraumatic Stress Disorder – Clinical
Posttraumatic Stress Disorder, Delayed
Posttraumatic Stress Disorder in Children
Posttraumatic Stress Disorder – neurobiological
basis for
Posttraumatic Stress Disorder,
Neurobiology of
Psychotic Disorders
Revenge Fantasies
Schizophrenia
Sexual Assault
Sexual Offenders
Social support in trauma
Stress of Self Esteem
Suicide, Biology of
Trans-sexualism
Violence

IMMUNOLOGY, INFECTION AND INFLAMMATION

Antibody Response
Autoimmunity

Corticotropin-releasing factor (CRF) family of
neuropeptides – role in inflammation
Cytokines
Cytokines, Stress, and Depression
Cytotoxic Lymphocytes
Herpesviruses
HIV Infection/AIDS
Immune Cell Distribution, Effects of Stress on
Immune Function, Stress-Induced Enhancement of
Immune Response
Immune Suppression
Immune Surveillance – Cancer, Effects of Stress on
Immune System, Aging
Immunity
Infection
Inflammation
Leishmania, Stress Response in
Leukocyte Trafficking and Stress
Lymphocytes
Mucosal Secretory Immunity, Stress and
Multiple Sclerosis
Natural Killer (NK) Cells
Neuroimmunomodulation
Neuroinflammation
Psoriasis
Psychoneuroimmunology
Sepsis, Acute Respiratory Distress Syndrome and
Glucocorticoid Resistance
Systemic Lupus Erythematosus
Vaccination
Viral Virulence and Stress
Viruses and Stress

LABORATORY STUDIES AND TESTS

Antimineralocorticoid Challenge
Dex-CRH Test
Dexamethasone Suppression Test (DST)
Mental Stress Testing
Metyrapone: Basic and Clinical Studies
Trier Social Stress Test
Waist–Hip Ratio

THERAPIES

Aerobics, Exercise and Stress Reduction
Behavior Therapy
Cancer Treatment
Desensitization
Exercise
Family Therapy
Group Therapy
Hypnosis
Integrative Medicine (Complementary and
Alternative Medicine)

Meditation and Stress
 Problem-Solving Skills Training
 Psychosomatic Medicine
 Reenactment Techniques
 Relaxation Techniques
 Stress Management and Cardiovascular Disease
 Stress Management, CAM Approach
 Trauma Group Therapy

PHYSIOLOGICAL, PHARMACOLOGICAL AND BIOCHEMICAL ASPECTS

Acute Stress Response: Experimental
 Acute Trauma Response
 Adenylyl Cyclases and Stress Response
 Adrenal Cortex
 Adrenaline
 Adrenal Medulla
 Adrenocorticotrophic Hormone (ACTH)
 Adrenocortical Function, factors controlling development thereof
 Aging and Stress, Biology of
 Aging and Adrenocortical Factors
 Alcohol and Stress: Social and Psychological Aspects
 Aldosterone and Mineralocorticoid Receptors
 Ambulatory Blood Pressure Monitoring
 Amenorrhea
 Amygdala
 Anatomy of the HPA Axis
 Androgen Action
 Angiotensin
 Angiotensin receptors
 Annexin
 Anti-CRF
 Antidepressant Actions on Glucocorticoid Receptors
 Apoptosis
 Arterial Baroreflex
 Autonomic Nervous System
 Autotolerance
 Avoidance
 Beta-Endorphin
 Blood–Brain Barrier, Stress and
 Blood Pressure
 Brain and Brain Regions
 Brain Natriuretic Peptide (BNP)
 Brain Trauma
 Calbindin
 Calcium-Dependent Neurotoxicity
 Calcium, Role of
 Cardiovascular System and Stress
 Catecholamines
 Central Stress Neurocircuits
 Cerebral Metabolism, Brain Imaging
 Chaperone Proteins and Chaperonopathies
 Cholesterol and Lipoproteins
 Comparative Anatomy and Physiology
 Congenital Adrenal Hyperplasia (CAH)
 Corticosteroid-Binding Globulin (Transcortin)
 Corticosteroid Receptors
 Corticosteroids and Stress
 Corticotropin Releasing Factor (CRF)
 Corticotropin Releasing Factor-Binding Protein
 Corticotropin Releasing Factor Receptors
 Critical Thermal Limits
 Dopamine, Central
 Drosophila Studies
 Electrodermal Activity
 Endocrine Systems
 Estrogen
 Ethanol and Endogenous Opioids
 Excitatory Amino Acids
 Excitotoxins
 Familial Patterns of Stress
 Febrile Response
 Feedback Systems
 Fibrinogen and Clotting Factors
 Feeding Circuitry (and Neurochemistry)
 Fetal Stress
 Food Intake and Stress, Human
 Food Intake and Stress, Non-Human
 Food Shift Effect
 GABA (Gamma Aminobutyric Acid)
 Gastrointestinal Effects
 Gender and Stress
 Ghrelin and Stress Protection
 Glia or Neuroglia
 Glucocorticoid Negative Feedback
 Glucocorticoid Effects on Memory: the Positive and Negative
 Glucocorticoids – Adverse Effects on the Nervous System
 Glucocorticoids, Effects of Stress on
 Glucocorticoids, Overview
 Glucocorticoids, Role in Stress
 Glucose Transport
 Glycobiology of Stress
 Gonadotropin Secretion, Effects of Stress on
 Heart Rate
 Heat Resistance
 Heat Shock Genes, Human
 Heat Shock Proteins: HSP60 Family Genes
 Heat Shock Response, Overview
 Hippocampal Neurons
 Hippocampus, Corticosteroid Effects on
 Hippocampus, Overview
 Homeostasis
 11 β -Hydroxysteroid Dehydrogenases
 Hypothalamic-Pituitary-Adrenal Axis

Instinct Theory
Kidney Function
Left Ventricular Mass
Leptin, Adiponectin, Resistin, Ghrelin
Lymph Nodes
Macrophage Antimycobacterial Activity, Effects of
Stress on
Macrophages
Maternal Deprivation
Membrane Glucocorticoid Receptors
Memory and Stress
Menstrual Cycles and Stress
Metabolic Syndrome
Metals, Oxidative Stress and Brain Biology
Mitochondria
Monoamine Oxidase
Multi Drug Resistance P Glycoprotein and other
Transporters
Musculoskeletal Problems and Stress
Myopathy
Neural Stem Cells
Neuroendocrine Systems
Neurogenesis
Neuropeptide Y
Neuropeptides, Stress-Related
Nitric Oxide
Nutrition
Obesity, Stress and
Orexin
Oxidative Stress
Oxidative Stress and Acidosis, Molecular
Responses to
Oxidative Stress and Aging
Oxytocin
Paraventricular Nucleus
Peptides
Personality Processes
Pheromones
Pituitary Regulation, Role of
Pressure, Effects of Extreme High and Low
Pro-opiomelanocortin (POMC)
Prostaglandins
Proteases in Prokaryotes and Eukaryotic Cell
Organelles
Proteases in the Eukaryotic Cell Cytosol
Protein Synthesis
Proteosome
Reductive Stress
Regional Blood Flow, Stress Effects
Renal and Adrenocortical Effects of
Dopamine
Reproduction, Effects of Social Stress on
Reproductive Dysfunction in Primates,
Behaviorally Induced

Resistance
Restraint Stress
Salivary Cortisol
Salt Appetite
Secretagogue
Serotonin
Serotonin in Stress
Sex Differences in Human Stress Response
Sex Steroids, Response to Stress and Susceptibility
to Depression
Sexual Dysfunction
Smoking and Stress
Startle Response
Steroid Hormone Receptors
Steroid Hydroxylases
Stress Hyporesponsive Period
Sympathetic Nervous System
Synthetic Glucocorticoids
Temperature Effects
Thermal Stress
Thermotolerance, Thermoresistance, and
Thermosensitivity
Thymus
Thyroid Hormones
Urocortin
Vasoactive Peptides
Vasopressin

PSYCHOLOGICAL THERAPY

Cognitive Behavioral Therapy
Freud, Sigmund
Posttraumatic Therapy
Psychoanalysis
Psychotherapy
War-Related Posttraumatic Stress Disorder,
Treatment of

PSYCHOSOCIAL AND SOCIOECONOMIC ASPECTS

Childhood Stress
Child Physical Abuse
Chronic Social Stress: GR Sensitivity
in Leukocytes
Community Studies
Crowding Stress
Crime Victims
Crisis Intervention
Cultural Factors in Stress
Cultural Transition
Divorce, Children of
Domestic Violence
Economic Factors and Stress
Education Levels and Stress

Employee Assistance and Counseling
 Environmental Stress, Effects on Human
 Performance
 Health and Socioeconomic Status
 Income Levels and Stress
 Indigenous Societies
 Job Insecurity; the health effects of a psychosocial
 work stressor
 Marital Conflict
 Marital Status and Health Problems
 Minorities and Stress
 Neighborhood Stress and Health
 Psychosocial Factors and Stress

Quality of Life
 Racial Harassment/Discrimination
 School Violence and Bullying
 School Stress and School Refusal Behavior
 Social Capital
 Social Networks and Social Isolation
 Social Status and Stress
 Social Stress, Animal Models of
 Social Support
 Transport-Related Stress
 Unemployment, Stress and Health
 Work–Family Balance
 Workplace Stress

PREFACE TO FIRST EDITION

“Stress” remains one of the most frequently used but ill-defined words in the English language. Stress is a phenomenon that has quite different meanings for the politician, social scientist, physician, nurse, psychotherapist, physiologist, molecular biologist, and perhaps you and me. This diversity of meanings was one impetus for creating the *Encyclopedia of Stress*, the aim being to derive a definition of stress from a variety of expert descriptions. The second impetus was the obvious need for an up-to-date compendium on one of the most important social, medical, and psychological phenomena of our age. We were fortunate in attracting stars for our Editorial Board and a set of most distinguished contributors for the 400 or so entries—indeed, the list of contributors is a Who’s Who in stress research.

We anticipate that the diversity of our readers will equal the diversity of the topics covered. They will find that the coverage of the Encyclopedia extends well beyond the general adaptation theory of Hans (Janos) Selye and the fight-or-flight response of Walter Cannon. Nevertheless, the general principles annunciated by these two great pioneers in the field still underpin our understanding of the biology of the stress phenomenon. That is, stress is a real or perceived challenge, either endogenous or exogenous, that perturbs body equilibrium or “homeostasis.” The stressor may range from overcrowding, traffic congestion, violence, bereavement, redundancy, or unemployment to physical, chemical, biological, or psychological insults. Whether the person can adapt to or cope with the stress will depend on the nature and severity of the stressor and the person’s physical and mental state, which in turn depends on genetic, experiential, social, and environmental factors. These issues are discussed in depth in the Encyclopedia, as are the mechanisms of coping and the impact of stress on health and predisposition to diseases such as

cancer, infection, rheumatoid arthritis, heart disease, high blood pressure, and mental disorder.

Aggression remains a hallmark of human behavior, even as we move into the third millennium, and so the Encyclopedia covers several topical areas that have only recently been analyzed systematically. These include war and specific wars, posttraumatic stress disorder (PTSD; formerly thought of vaguely as “shell shock”), rape, torture, marital discord and spousal abuse, and the Holocaust. In tackling these topics we accept that our entries may not include all the nuances that are necessary for a full understanding of what these phenomena are all about and described so graphically and sensitively in Tolstoy’s *War and Peace* or Pat Barker’s monumental *Regeneration* trilogy on the horrific psychological traumas of the First World War. Nevertheless, an important start has been made in that we now accept that PTSD is not just a lack of “bottle” (courage or “guts”), but rather a syndrome that needs to be and can be understood within the framework of medicine and psychology.

Biologically, the stress response reflects a set of integrated cascades in the nervous, endocrine, and immune defense systems. As in most areas of biology, molecular genetics has made a significant difference in the precision with which we now understand the physiopathological processes of the stress response. And so the adage formerly applied to diabetes mellitus may now apply equally to stress: “Understand stress and you will understand medicine.”

In summary, we hope that this first *Encyclopedia of Stress* will indeed define the phenomenon and at the same time provide a valuable source of information on a phenomenon that affects us all. In setting out on this adventure we were aware that there is nothing new under the sun and that stress has been around since the first biological particles, bacteria, or even viruses competed for the same mechanisms of replication.

There is a tendency for each generation to imagine that stress and its untoward effects are uniquely harsh for them, but it is not this misconception that underlies this work. Rather, the stress of “stress” itself—the massive accumulation of knowledge—made it seem propitious to bring the information together in a systematic manner that allows ready access to all who need or wish to understand the phenomenon.

The idea of producing this Encyclopedia was conceived at an Academic Press reception in San Diego held in conjunction with the annual meeting of the Society for Neuroscience in 1996. I am deeply indebted to Erika Conner for enabling conversion of the idea to a concept and then a project and to

Jennifer Wrenn, Christopher Morris, and Carolan Gladden, all of Academic Press, for their enthusiasm, encouragement, and herculean efforts that converted the concept into a reality. For giving generously of their intellect, expertise, sound advice, and unstinting work, I am greatly indebted to my friends and colleagues on the Editorial Board, who made the project such a satisfying experience. To all our contributors go our profound thanks for taking time out from wall-to-wall schedules to produce their entries, which together have made this Encyclopedia.

George Fink
August 1999

PREFACE TO THE SECOND EDITION

The popularity of the first edition of the Encyclopedia of Stress, the several major stressful conflicts that have occurred since 1999, and the significant advances in our knowledge of stress prompted the production of this second edition. In addition to more than 140 new articles, most of the original 400 first edition articles have been updated for the second edition to reflect, for example, advances in our understanding of corticotropin releasing factor (CRF) and urocortin receptors, ideas about the role of central CRF in the overall control of the psychological/mental as well as neuroendocrine response to stress, and the novel discoveries of the complex neuropeptide regulation of hunger and satiety. The effects of maternal and perinatal stress on subsequent body development and function in the adult, the adverse effects of stress on brain (and particularly hippocampal memory function), the role of the amygdala in fear and aggression, and the role of stress in the etiology of obesity and the metabolic syndrome all feature prominently in this new edition. 9/11 and the many new conflicts since 1999 have engendered new articles on terrorism, suicide bombers and their effects, as well as a retrospective on previous human apocalypses such as Hiroshima. Coverage of the whole field of post traumatic stress disorder (PTSD) has been expanded, reflecting the greater recognition and understanding of PTSD, and possibly related disorders such as *combat fatigue* and *burnout*. Stress and depression and anxiety, already touched on in the first edition, now feature prominently. On the conceptual side, *allostasis* (modified from Hans Selye's *heterostasis*), or stability through centrally regulated change in physiological set points, receives greater emphasis than it did in the first edition.

Overall, our understanding of stress mechanisms in the human has benefited from two major quantum leaps in technology. First, the advances in molecular genetics and genomics and sequencing of the human

genome (published after the first edition had appeared) have increased the rigor and precision of our understanding of the molecular neurobiology of stress. Thus, new to the second edition are a series of articles on the genetic factors that play a role in susceptibility to stress and in various components of the stress response. Secondly, functional brain imaging has enhanced our understanding of the neurobiology of stress in the human. Many of the articles in the second edition reflect the positive impact of these two advances.

Accessibility to the second edition of the Encyclopedia of Stress will be facilitated by virtue of the fact that it is published online as well as in hard copy. My profound thanks for their selfless work go to our team of distinguished authors and to my friends and colleagues on the Editorial Board. Production of the Encyclopedia would not have been possible without our colleagues at Elsevier, and in particular Bob Donaldson and Hilary Rowe whose dedication, commitment and excellent work I am pleased to acknowledge.

George Fink
November 2006

Note on Terminology of Corticotropin Releasing Factor/Hormone and the Catecholamines

The central nervous regulation of the anterior pituitary gland is mediated by substances, mainly peptides, which are synthesized in the hypothalamus and transported to the gland by the hypophyseal portal vessels. Because these compounds are transported by the blood, the term "hormone" gained acceptance in the neuroendocrine literature. The major hypothalamic peptide involved in the stress response is the

41-amino acid corticotropin releasing factor (CRF). The Endocrine Society (USA), following convention, adopted the term corticotropin releasing hormone (CRH). However, this nomenclature has been challenged. Hauger et al. (*Pharmacol Rev* 55: 21–26, 2003), in liaison with the International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification, argued that the CRF's function extends well beyond the biology of a hormone, and that therefore it should be termed corticotropin releasing factor (CRF) rather than hormone. Since the terminology of CRF versus CRH has yet to be resolved, the two terms and abbreviations are here used interchangeably, depending on author preference.

Adrenaline and *noradrenaline* are catecholamines that play a pivotal role in the stress response. These

terms are synonymous with *epinephrine* and *norepinephrine*, respectively. Both sets of terms are used interchangeably in the endocrine, neuroendocrine and stress literature, and this principle has been adopted for the Encyclopedia. Style has depended on author preference, but wherever possible within-article consistency has been ensured.

George Fink
November 2006

Hauger, R. L., Grigoriadis, D. E., Dallman, M. F., Plotsky, P. M., Vale, W. W., and Dautzenberg, F. M. (2003). International Union of Pharmacology. XXXVI. Current status of the nomenclature for receptors for corticotropin-releasing factor and their ligands. *Pharmacological Reviews* 55(1), 21–26.

FOREWORD

This is the second edition of the *Encyclopedia of Stress*, the three heavy volumes that first appeared in 2000 under the overall editorial guidance of George Fink. From its original coining in reference to rather unusual and unpleasant situations, the word *stress* is now quasi synonymous with *life*, i.e. modern life! This second edition of the *Encyclopedia of Stress* presents several hundreds of articles written by experts in that particular field, especially for this series, discussing, clarifying, explaining to the average-reader as well as to anyone more educated in one field or another, the extraordinary involvement of stress and our normal or abnormal reactions to it, both in our physical body and its accompanying mind. The table of contents is truly amazing, from the stress of being born to the stress of dying and so many circumstances in between, with clear presentations of the pertinent, current knowledge about them. This is truly an encyclopedia where, surely, you will find some explanation and answer to your own stresses (and those of others).



Roger Guillemin, MD, PhD
Nobel laureate in Physiology & Medicine

A

Acute Stress Disorder and Posttraumatic Stress Disorder

R Yehuda and C M Wong

Mt. Sinai School of Medicine and Bronx Veterans Affairs,
New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, Volume 1,
pp 1–7, © 2000, Elsevier Inc.

Introduction

Relationship between Acute Stress Disorder (ASD) and Posttraumatic Stress Disorder (PTSD)

Epidemiology

Risk Factors for ASD and PTSD

Comorbid Disorders

Biological Findings in ASD and PTSD

Treatment of ASD and PTSD

Summary

Glossary

Acute stress disorder

A mental condition that can occur following exposure to extreme stress or trauma but by definition does not last longer than 1 month [i.e., after 1 month, if the symptoms persist, it is appropriate to consider the diagnosis of posttraumatic stress disorder (PTSD)].

Avoidant symptoms

Experiences that represent measures taken by the individual to avoid feeling, thinking, or coming into contact with reminders of the trauma; these experiences can also include having difficulty recalling important events of the trauma, avoiding people and places associated with the trauma, and generally feelings of emotional numbness, detachment from others, and feeling like there is no future.

Depersonaliza- tion

An alteration in the perception or experience of the self so that one feels detached

Derealization

from, and, as if one is an outside observer of, one's mental processes or body (e.g., feeling like one is in a dream).

Dissociation

An alteration in the perception or experience of the external world so that it seems strange or unreal (e.g., people may seem unfamiliar or mechanical).

A disruption in the usually integrated functions of consciousness, memory, identity, or perception of the environment. The disturbance may be sudden or gradual, transient, or chronic.

Dysphoria

An unpleasant mood such as sadness, anxiety, or irritability.

Flashback

A recurrence of a memory, feeling, or perceptual experience from the past.

Hyperarousal symptoms

Experiences that involve more physical reactions to trauma, including difficulty with sleep and concentration, irritability, and anger, needing to be on guard overly much, and having an exaggerated startle response to unexpected noises.

Intrusive symptoms

Experiences in which trauma survivors relive the trauma or become distressed by reminders of the trauma. These include unwanted thoughts and images, nightmares and bad dreams, extreme distress when exposed to a trigger or reminder, and a physiological response to triggers, including palpitations, sweating, and shortness of breath.

Posttraumatic stress disorder

A mental condition that can occur following exposure to extreme stress or trauma and lasts for 1 month or longer. Acute PTSD lasts from 1 to 3 months according to current formulations in the DSM-IV. Chronic PTSD lasts 3 or more months according to current formulations in the DSM-IV. In delayed onset PTSD, the onset of symptoms begins 6 or more months after the traumatic event(s) according to current formulations in the DSM-IV.

<i>Stressor</i>	In this context, it is a life event or life change that is so challenging as to potentially be associated with the onset, occurrence, or exacerbation of a psychological symptom or a mental disorder.
<i>Traumatic event</i>	According to DSM-IV, an experience that involves death or serious injury or another threat to one's physical integrity. A person does not need to actually experience the physical injury to be traumatized – just the knowledge that a person could have been severely injured or had a serious threat of injury and the feelings that accompany this realization are enough to produce a posttraumatic reaction.

Both acute stress disorder (ASD) and posttraumatic stress disorder (PTSD) are conditions that can occur in people who have been exposed to a traumatic life event. Traumatic events are thought to occur in at least 50% of the population. A traumatic event is defined in the *American Psychiatric Association's Diagnostic and Statistical Manual for Mental Disorders*, 4th edn (DSM-IV) as an experience in which a person underwent, witnessed, or was confronted with "an event that involved actual or threatened death or serious injury, or a threat to the physical integrity of self or others" and a subjective response of "intense fear, helplessness, or horror." The definition of trauma is purposely identical for both ASD and PTSD because what differentiates the two conditions is the point in time following the event at which the symptoms are experienced. Not all people who have experienced trauma will develop ASD and/or PTSD. ASD defines the immediate response to trauma for up to 4 weeks posttrauma, whereas PTSD encompasses a more chronic response to trauma beginning 1 month after the trauma and lasting at least 1 month after that. Recent estimates suggest that as much as 14% of persons in the United States will develop these conditions at some point during their lives. Given the chronic nature of PTSD and the extreme disability that can be associated with this condition, this statistic is alarming and points to PTSD as a major public health problem.

Introduction

Posttraumatic stress disorder was originally defined in 1980 to describe long-lasting symptoms that occur in response to trauma. The diagnosis revived psychiatry's long-standing interest in how stress can result in behavioral and biological changes that ultimately lead to disorder. At the time the formal diagnosis of PTSD was being conceptualized, the field was focused heavily on describing the psychological consequences

of combat Vietnam veterans and others who had chronic symptoms following exposure to events that had occurred years and even decades earlier. It was widely assumed that the symptoms that persisted in these survivors were extensions of those that were present in the earlier aftermath of a traumatic event. Indeed, some data from recently traumatized burn victims supported the idea that symptoms such as intrusive thoughts were present in the early aftermath of a trauma. However, at the time the diagnosis was established, there were no longitudinal data – either retrospective or prospective – that formally established the relationship between acute vs chronic posttraumatic symptoms.

Chronic PTSD was not initially conceptualized as being qualitatively different from what might have been observable in trauma survivors in the acute aftermath of the trauma. Rather, chronic PTSD suggested a failure of restitution of the stress response. The implicit assumption behind the diagnosis was that most trauma survivors initially developed symptoms as a direct result from exposure to the event.

As the awareness of the longterm effects of trauma became more widespread, investigators began to explore the reaction of trauma survivors in the immediate aftermath of the event. Studies of the acutely traumatized revealed that survivors do experience symptoms in the immediate aftermath. Clinicians felt it important to provide mental health interventions as early as possible so as to perhaps prevent the development of more chronic conditions. ASD first appeared as a diagnosis in the DSM-IV in 1994 in order to provide a diagnosis for people before they were eligible to receive the diagnosis of PTSD (i.e., having symptoms for less than 1 month's duration). Thus, ASD arose in part out of the need for justification for acute intervention. However, as described later, the syndrome of ASD has features that are not directly associated with the more chronic response of PTSD.

One of the important findings to be generated from research in the last two decades is that PTSD does not occur in everyone exposed to trauma, nor is PTSD the only possible response to trauma. Prospective longitudinal studies have now demonstrated that mood and other anxiety disorders can also occur following a traumatic event, and these may be present even in the absence of PTSD. The question of why some people develop PTSD and others do not has not been fully resolved and is the subject of current investigation in the field. However, insofar as this condition no longer serves to characterize a universal stress response, it has been important to redefine the nature of traumatic stress responses and determine the risk factors for developing these conditions. It has also

been important to consider the relationship between acute and chronic stress response.

Relationship between Acute Stress Disorder (ASD) and Posttraumatic Stress Disorder (PTSD)

Figure 1 shows a time continuum that describes the onsets of ASD and acute, chronic, and delayed PTSD. In ASD, symptoms must last for at least 2 days up to 4 weeks. In PTSD, the symptoms last for a period of at least a month following the traumatic event(s) and can be either acute (symptom duration of 3 months or less) or chronic (symptom duration of more than 3 months). The onset of PTSD can be delayed for months and even years.

Proponents of ASD hoped the diagnosis would help survivors mobilize acute intervention, thereby preventing more chronic disability. The idea of an acute stress disorder was initially considered to be superfluous with the diagnosis of adjustment disorder, which was defined in 1980 to describe early symptoms after

any stressful event. The proponents of ASD, however, were interested in emphasizing that, unlike adjustment disorders that were expected to resolve even without treatment, ASD was expected to develop into PTSD. Furthermore, ASD focused on a different set of symptoms that were precursors of PTSD such as dissociation.

Indeed, the symptom of dissociation is a prominent difference between ASD and adjustment disorder. Interestingly, dissociation per se is not a symptom of PTSD. However, there is much support for the idea that if people dissociate at the time of trauma, they are at an increased risk to develop PTSD (Table 1).

Most studies examining the relationship between dissociation at the time of a trauma and development of PTSD have been retrospective, in some studies, assessing the dissociation 25 years after the traumatic experiences. This means that assessment of dissociation is based largely on subjective recollections of victims. However, in studies that have been prospective, peritraumatic dissociation was still found to be one of the best predictors of PTSD 6 months and

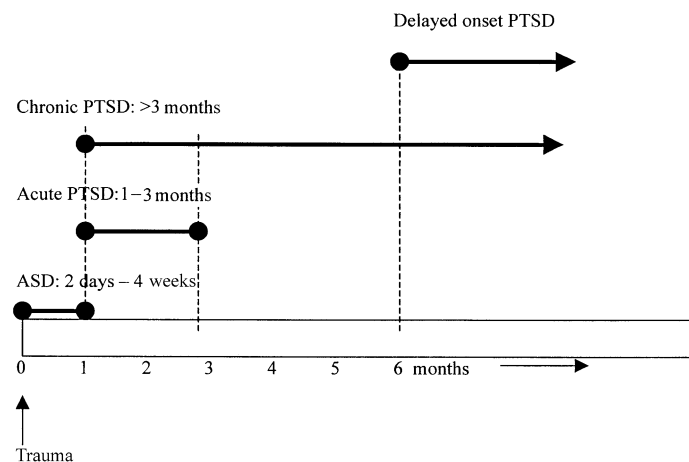


Figure 1 Time line comparing ASD and PTSD.

Table 1 Comparison of symptoms of adjustment disorder, ASD, and PTSD

<i>Adjustment disorder</i>	<i>ASD</i>	<i>PTSD</i>
With anxiety	Intrusive reexperiencing	Intrusive reexperiencing
With depression	Avoidance	Avoidance
With mixed anxiety and depression	Hyperarousal	Hyperarousal
With disturbance of conduct		
With mixed disturbance of emotions and conduct	Dissociation Subjective sense of numbing, detachment, or absence of emotional responsiveness A reduction in awareness of his or her surroundings (e.g., being in a daze) Derealization Depersonalization Dissociative amnesia (i.e., inability to recall an important aspect of the trauma)	

1 year following the trauma in a group of mixed civilian trauma. Posttraumatic depression was also a good predictor for the occurrence of PTSD at 1 year. This would perhaps support the idea of including dissociation as a criterion for ASD given its predictive property for the future development of PTSD.

ASD and PTSD are both characterized by the development of intrusive, avoidant, and hyperarousal symptomatology following a traumatic event. These three symptom clusters represent different expressions of the response to trauma.

Intrusive symptoms describe experiences in which the survivor relives the event in their mind. After a traumatic event, it is common for the survivor to have thoughts of the event appear in one's mind and to think about the event despite not wanting to. Frightening thoughts or nightmares about the event are also common, and the person may feel very distressed. Sometimes a survivor may even have physical symptoms such as heart palpitations, sweating, and an increased breathing rate in response to being reminded of the event.

In addition to experiencing one or more of the just-mentioned intrusive symptoms, the survivor may make one or more attempts to avoid being reminded of the traumatic experience. The individual may go out of her way to avoid being in situations that remind her of the event. The individual may forget details about the traumatic event. The survivor may also experience a more generalized feeling of wanting to shut out the world, which is manifested by a feeling of emotional numbness and an inability to connect with others.

Finally, the trauma survivor may experience a range of more physiological symptoms, such as an inability to sleep, inability to concentrate, and anger and irritability. The individual may constantly feel alert and on guard, scanning the environment for signs of danger. The survivor may be very sensitive to noises and have an exaggerated startle response to unexpected or loud sounds. These symptoms are called hyperarousal symptoms.

The conceptualization of ASD remains in evolution at the writing of this volume. The salient question is whether ASD is really a disorder that indicates a deviation from a normative response to stress. To the extent that ASD is a good predictor of PTSD, it may be that this syndrome captures elements of a maladaptive response.

Epidemiology

In considering the prevalence rates of ASD and PTSD, the main observation is that neither are inevitable responses to stress.

The range of prevalence rates of ASD for various types of trauma, including traumatic brain injury (TBI), disasters, and shootings, ranges from 7.2 to 33%, which interestingly are about the same as those for PTSD (see later).

The range of prevalence rates of PTSD in trauma survivors in DSM-IV is 3–58%. This great variation is thought to reflect mostly a difference in the type and quality of the trauma experienced. However, it may also reflect study methodology (i.e., whether people were evaluated in person or on the phone; whether they were selected randomly or from a convenience sample, and/or whether the study was conducted in the immediate aftermath of the trauma or years later). Studies have found that there were differences in the rates of PTSD depending on the type of trauma experienced. For example, 65% of men and 45.9% of women with PTSD reported rape as their most upsetting trauma. Other types of trauma associated with high probabilities of PTSD development include combat exposure (38.8%), childhood neglect (23.9%) and childhood physical abuse (22.3%) in men, and sexual molestation (26.5%), physical attack (21.3%), being threatened by a weapon (32.6%), and childhood physical abuse (48.5%) in women.

In the aggregate, PTSD is estimated to occur on average in 25% of individuals who have been exposed to traumatic events with rates considerably higher for life-threatening events than for those with lower impact. Lifetime prevalence of exposure to at least one traumatic event was also found to range from 40 to 55% in various studies. In the National Comorbidity Survey (NCS), it was found that about 8% of the population were exposed to four or more types of trauma, some of which involved multiple occurrences. The lifetime prevalence of PTSD has consistently been demonstrated to be twice as high in women as in men.

Risk Factors for ASD and PTSD

Risk factors can be divided into two broad categories: those pertinent to the traumatic event (e.g., severity or type of trauma) and those relevant to individuals who experience the event (e.g., gender, prior experiences, personality characteristics). Although some risk factors for PTSD appear to be related to prior experiences, data have also emerged implicating biological and possibly genetic risk factors for PTSD.

Risk factors for the development of ASD include female gender, high depression scores in the immediate aftermath of the event, an avoidant coping style, and the severity of the traumatic event. These are similar to the risk factors that have been associated with the development of chronic PTSD. However, it

should be noted that this has not been an area of extensive study.

Risk factors for PTSD also include female gender, posttraumatic depression, an avoidant coping style, and the severity of trauma. However, additional risk factors have also been noted, such as a history of stress, abuse, or trauma, a history of behavioral or psychological problems, preexisting psychiatric disorders, a family history of psychopathology, genetic factors, subsequent exposure to reactivating environmental factors, initial psychological reaction to trauma such as emotional numbing, early separation, preexisting anxiety and depressive disorders, and depression at the time of trauma. Parental PTSD has also been found to be a risk factor for the development of PTSD.

Comorbid Disorders

PTSD rarely occurs in the absence of other conditions. Approximately 50–90% of individuals with chronic PTSD have a psychiatric comorbidity. This reinforces the idea that PTSD is not distributed randomly throughout the population, but rather that there are subgroups of people that are more vulnerable to the development of PTSD and other psychiatric disorders. This also raises the question of whether PTSD develops as a separate disorder from other psychiatric comorbidities seen in association from it or whether people who have these constellations of disorders are more likely to get PTSD.

Common comorbidities in PTSD patients include major depressive disorder, alcoholism, drug abuse, personality disorders, anxiety disorders such as panic disorder, and generalized anxiety disorder and dissociative disorder.

Symptoms of depression are observed frequently in trauma survivors with PTSD. A significant correlation between early PTSD symptoms and the occurrence of depression seems to predict the chronicity of PTSD. The sequence in which PTSD and major depressive disorder (MDD) occurs following trauma is of particular interest. Some studies proposed that depression is “secondary” to PTSD because its onset followed that of PTSD. In the NCA, patients with comorbid PTSD and MDD reported that the onset of the mood disorder followed that of PTSD in 78.4% of the sample. Vietnam veterans also reported that the onset of their phobias, MDD, and panic disorder followed that of the PTSD. Retrospective studies were flawed in that patients may not recall accurately which symptom came first after being symptomatic for decades. In contrast, Israeli combat veterans reported a simultaneous onset of PTSD and MDD in 65% of the sample, with 16% reported having the MDD precede the PTSD.

In a groundbreaking prospective, longitudinal study, it was demonstrated that the presence of depression in the early aftermath of a trauma was associated with the subsequent development of PTSD. These data contradict the notion that comorbid conditions in PTSD reflect secondary consequences of the PTSD symptoms and suggest instead that other psychiatric symptoms influence the presence of PTSD symptoms. Indeed, a history of previous psychiatric symptoms is a predictor of PTSD, but posttraumatic symptoms may also be strong predictors.

Biological Findings in ASD and PTSD

One of the most provocative observations in both ASD and PTSD that has emerged is that the biological findings do not conform to classic patterns of biological stress responses that have been described in animal and human models of the stress response. This is particularly true in the area of neuroendocrinology of PTSD. The distinctness in the biology of PTSD and stress further supports the idea that PTSD represents a circumscribed and specific response to a traumatic event that may be different from a normative response in which the reaction does not reach a certain magnitude or, if it does, becomes attenuated quickly.

There have been only a few studies of the biology of ASD; however, the biological alterations in ASD seem to resemble those in PTSD. First and foremost, there has been a demonstration of the enhanced negative feedback inhibition of the hypothalamic-pituitary-adrenal (HPA) axis, similar to findings in PTSD (see later). At 2 weeks posttrauma, rape victims who subsequently recovered from ASD at 3 months showed a “normal” response to dexamethasone administration, whereas rape victims who failed to recover from ASD at 3 months showed a hypersuppression of cortisol on the dexamethasone suppression test (DST).

In another study, women with a prior history of rape or assault had significantly lower cortisol levels in the immediate aftermath of rape compared to those who did not. A past history of rape or assault is generally considered to be a potent predictor for the development of PTSD. In a study of motor vehicle accident victims, a lower cortisol response in the immediate aftermath of the trauma was associated with the presence of PTSD at 6 months, whereas higher cortisol levels at the time of the trauma were associated with the development of depression at 6 months. These data indicate that it might be possible to identify high-risk ASD patients through biological means.

In PTSD, there have been far more studies that have yielded fairly consistent results. The studies demonstrate a different profile of HPA alterations in PTSD compared to that in other psychiatric disorders such

as major depression. In the aggregate, these studies have demonstrated a hypersensitization of the HPA axis to stress that is manifested by a decreased cortisol release and an increased negative feedback inhibition. Hypersecretion of the corticotropin-releasing hormone (CRH) has also been described in PTSD, as has blunting of the adrenocorticotrophic hormone (ACTH) response to CRH. Other studies have demonstrated lower 24-h urinary cortisol excretion levels in PTSD patients compared to other psychiatric groups and normal controls. PTSD subjects have also been shown to have lower plasma cortisol levels in the morning and at several time points throughout the circadian cycle and to have a greater circadian signal-to-noise ratio compared to normal controls.

In addition, studies support the notion that the enhanced negative feedback regulation of cortisol is an important feature of PTSD. PTSD subjects have an increased number of lymphocyte glucocorticoid receptors (which are needed for cortisol to exert its effects) compared to psychiatric and nonpsychiatric control groups and show an enhanced suppression of cortisol in response to dexamethasone. This response is distinctly different from that found in depression and other psychiatric disorders.

There have been numerous other systems studies in PTSD such as the sympathetic nervous system, the serotonergic system, and the immune system. Psychophysiological, neuroanatomical, cognitive, and behavioral alterations have also been noted; however, it is beyond the scope of this volume to encompass and discuss all these findings.

Treatment of ASD and PTSD

There are several types of specialized treatments that have been developed in recent years to address the specific needs of trauma survivors who suffer from both ASD and PTSD. These treatments include cognitive behavioral therapy, psychotherapy, group therapy, and medication therapy. Chronic PTSD appears to be more resistant to treatment than acute PTSD. It is usually necessary to employ a combination of therapies in the treatment for PTSD, but the success rate of specialized trauma-focused treatment is quite encouraging.

Summary

ASD and PTSD represent responses that can occur following exposure to traumatic events. To date, it is useful to conceptualize ASD and PTSD as two distinct conditions that are very much related. ASD often but

not always leads to a development of PTSD. ASD and PTSD seem to share similar symptoms, prevalence, risk factors, and neurobiological alterations. ASD appears to be easier to treat than chronic PTSD, which suggests that it is probably useful to treat traumatic stress symptoms earlier rather than later in the course of adaptation to trauma.

Acknowledgments

This work was supported by NIMH R-02 MH 49555 (RY), Merit Review Funding (RY), and APA/PMRTP 5T32-MH-19126-09/10 (CMW), and VA-Research Career Development award (CMW).

See Also the Following Articles

Avoidance; Brain Trauma; Depersonalization: Systematic Assessment; Dissociation; Posttraumatic Stress Disorder in Children; Posttraumatic Stress Disorder, Delayed; Posttraumatic Stress Disorder, Neurobiology of.

Further Reading

- Andreasen, N. C. (1980). Posttraumatic stress disorder. In: Kaplan, J. L., Freedman, A. M. & Saddock, B. J. (eds.) *Comprehensive Textbook of Psychiatry* (3rd edn. Vol. 2), pp. 1517–1525. Baltimore: Williams & Wilkins.
- Breslau, N., Davis, G. C., Andreski, P., Peterson, E. L. and Schultz, L. R. (1997). Sex differences in posttraumatic stress disorder. *Archives of General Psychiatry* **54**, 1044–1048.
- Foa, E. B. (ed.) (1993). *Posttraumatic Stress Disorder. DSM-IV and Beyond*. Washington, DC: American Psychiatric Press.
- Friedman, M. J., Charney, D. S. and Deutch, A. Y. (eds.) (1995). *Neurobiological and Clinical Consequences of Stress: From Normal Adaptation to Post-traumatic Stress Disorder*. Hagerstown, MD: Lippincott-Raven.
- Kessler, R. C., Sonnega, A., Bromet, E. and Nelson, C. B. (1995). Posttraumatic stress disorder in the National Comorbidity Survey. *Archives of General Psychiatry* **52**, 1048–1060.
- Shalev, A. Y., Freeman, S., Peri, T., Brandes, D., Shahar, T., Orr, S. P. and Pitman, R. I. K. (1998). Prospective study of posttraumatic stress disorder and depression following trauma. *American Journal of Psychiatry* **155**, 630–637.
- Yehuda, R. (ed.) (1998). *Psychological Trauma*. Washington, DC: American Psychiatric Press.
- Yehuda, R. (ed.) (1999). *Risk Factors for Posttraumatic Stress Disorder. Progress in Psychiatry Series*. Washington, DC: American Psychiatric Press.
- Yehuda, R. and McFarlane, A. C. (eds.) (1997). *Psychobiology of Posttraumatic Stress Disorder*. New York: New York Academy of Sciences.

Acute Stress Response: Experimental

K Pacak

National Institute of Neurological Disorders and Stroke and National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA

R McCarty

University of Virginia, Charlottesville, VA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 8–16, © 2000, Elsevier Inc.

Overview of Acute Stress Responses

Stress-Responsive Neuroendocrine Systems

Types of Acute Stressors

Summary and Conclusions

Glossary

<i>Footshock</i>	A laboratory stressor that involves placing a laboratory rat or mouse into a chamber with a floor of metal rods spaced equally apart. Scrambled electric current of known intensity is passed through the metal rods and this results in a mildly painful stimulus to the animal's footpads. This has been a favored stressor in many laboratories because the intensity, duration, and frequency of the footshocks are under the control of the experimenter.
<i>Hemorrhage</i>	An acute decrease in blood volume. In laboratory studies, hemorrhage has been employed as a physiological stressor and blood is often withdrawn via an arterial or venous catheter.
<i>Hypoglycemia</i>	A reduction in blood glucose levels. This physiological change has been employed as a stressor in laboratory studies and insulin is often administered to effect a reduction in blood glucose.
<i>Immobilization</i>	Often employed as a laboratory stressor, immobilization involves restricting the movement of a laboratory animal (typically a rat or mouse) by taping its limbs to padded metal mounts affixed to a frame. A variation on this approach includes restraint, where an animal's movement is restricted by placing it into a plexiglas cylinder.
<i>Neuroendocrine signature</i>	The concept that exposure to a given stressor elicits a pattern of neuroendocrine responses that is specific to that stressor.

Overview of Acute Stress Responses

Humans and animals respond to exposure to a particular stressor by mobilizing a host of neural, neuroendocrine, endocrine, and metabolic systems. Each stressor produces a specific neurochemical signature, with quantitatively if not qualitatively distinct central and peripheral mechanisms involved. These neurochemical changes occur not in isolation but rather in concert with physiological, behavioral, and even experiential changes. In evolutionary terms, natural selection has favored this pattern of coordinated stress responses by enhancing the survival and propagation of genes. This pattern of stress responsiveness is referred to by Goldstein as "primitive specificity." These and other findings are included in the description of acute stress responses that follows.

A stressor may be viewed as a stimulus that disrupts homeostasis. Stressors may be divided into several groups based upon their character, duration, and intensity. Regarding duration, stressors may be characterized as acute (single, intermittent, and time-limited exposure versus continuous exposure) or chronic (intermittent and prolonged exposure versus continuous exposure).

Physical stressors include cold, heat, intense radiation, noise, vibration, and many others. Chemical stressors include ether, all poisons, and insulin. Pain stress may be elicited by many different chemical and physical agents. Psychological stressors profoundly affect emotional processes and may result in behavioral changes such as anxiety, fear, or frustration. Social stressors include an animal's placement into the territory of a dominant animal and in humans unemployment and marital separation among others. Many stressors used in animal research, however, are mixed and act in concert, such as handling, immobilization, anticipation of a painful stimulus, and hemorrhage with hypotension.

In terms of duration, stressors may be divided into two main categories: acute versus chronic stressors. It should be noted that many stressors differ in their intensity. The adaptive responses that are elicited in response to an acute stressor include the physiological and behavioral processes that are essential to reestablish homeostatic balance. During an acute stress response, physiological processes are important to redirect energy utilization among various organs and selectively inhibit or stimulate various organ systems or their components to mobilize energy reserves and to be prepared for exposure to additional, unpredictable challenges. Thus, upon exposure to metabolic

stressors, certain tissues tend to reduce their consumption of energy while others, especially those that are important for locomotor activity, receive sufficient nutrients to function properly. The central nervous system also has priority during metabolic stress responses and preferentially receives a sufficient amount of nutrients from the circulation. The increased supply of energy to crucial organs is achieved preferentially by release of catecholamines and glucocorticoids that in general increase gluconeogenesis and glycogenolysis, inhibit glucose uptake, and enhance proteolysis and lipolysis. The immune system is another essential component of these physiologically adaptive stress responses.

Behavioral adaptation is viewed as a facilitation of adaptive and an inhibition of nonadaptive central and peripheral neural pathways that enable an organism to cope successfully with stressful situations. These behavioral responses include altered cognitive and sensory thresholds, increased alertness, selective memory enhancement, suppression of feeding and reproductive behaviors, and stress-induced analgesia. If physiological or behavioral responses fail to reestablish homeostasis or are excessive and abnormal, diseases of adaptation may be more likely to occur in susceptible individuals. This is apparent upon exposure to chronic or repeated stressors but may happen as well during acute stress reactions in a situation where the intensity of the stressor is abnormally high. Both physiological and behavioral responses are affected by activation of the primary stress effector systems, including the sympathetic nervous system (norepinephrine release), the adrenomedullary system (epinephrine release), the hypothalamic-pituitary-adrenocortical system [adrenocorticotrophic hormone (ACTH) and glucocorticoid release], the parasympathetic nervous system (acetylcholine release), and the renin-angiotensin system (renin release from juxtaglomerular cells in the kidney). Several other systems contribute to a reestablishment of homeostasis, including the hypothalamic-pituitary-thyroid axis (e.g., responsive to cold and heat), the hypothalamic-pituitary-gonadal axis (temporary decrease in reproductive function to shift energy to other more important organ systems), release of growth hormone, and changes in immune function. All of these systems act directly by altering the release or the biological effects of many mediators of acute stress responses (e.g., neurotransmitters, hormones, cytokines, etc.) or they act indirectly by altering levels of monitored variables, with consequent reflexive adjustments determined by internal homeostats.

Stress-Responsive Neuroendocrine Systems

Included below are some examples of ways in which the major stress-responsive physiological systems are altered following exposure to acute stressors.

Hypothalamic-Pituitary-Adrenocortical Axis

The hypothalamic-pituitary-adrenocortical axis is one of the main effector systems that is activated in animals upon exposure to an acute stressor. Generally via several feedback mechanisms at different levels, glucocorticoids released from the adrenal cortex regulate their own synthesis and secretion by actions within the adrenal gland (inhibition of glucocorticoid synthesis), the anterior pituitary (inhibition of corticotrophs), and the hypothalamus [inhibition of corticotropin-releasing hormone (CRH) synthesis and secretion] as well as higher brain centers (e.g., hippocampus). Glucocorticoids readily penetrate the blood-brain barrier and suppress acute stress-induced increases in central neural levels of norepinephrine. Norepinephrine is thought to be the most potent stimulator of CRH neurons within the hypothalamic paraventricular nucleus. Thus, central noradrenergic neurons that terminate in the paraventricular nucleus and synapse on CRH neurons are generally believed to be the site of another negative feedback mechanism by which glucocorticoids attenuate neuroendocrine responses to acute stressors.

During an acute stress response, ACTH and glucocorticoids vary depending on the type and intensity of the stressful stimulus. Exposure of animals to cold stress evokes a very modest increase in plasma ACTH, exposure of animals to hemorrhage or hypoglycemia evokes ACTH responses that reflect the intensity of the stimulus, pain elicits moderate ACTH responses, and immobilization for 120 min evokes the greatest increase in plasma levels of ACTH. The correlation between norepinephrine release in the hypothalamic paraventricular nucleus and plasma ACTH responses may demonstrate the participation of paraventricular norepinephrine in stress-induced activation of the hypothalamic-pituitary-adrenocortical axis ([Figure 1](#)). Although a strong correlation is found between norepinephrine responses in the paraventricular nucleus and plasma ACTH responses during exposure to different stressors, such correlations may not reflect causal interrelationships. During exposure to metabolic stressors such as hypoglycemia or hemorrhage, increments in plasma ACTH are much higher than those of norepinephrine in the

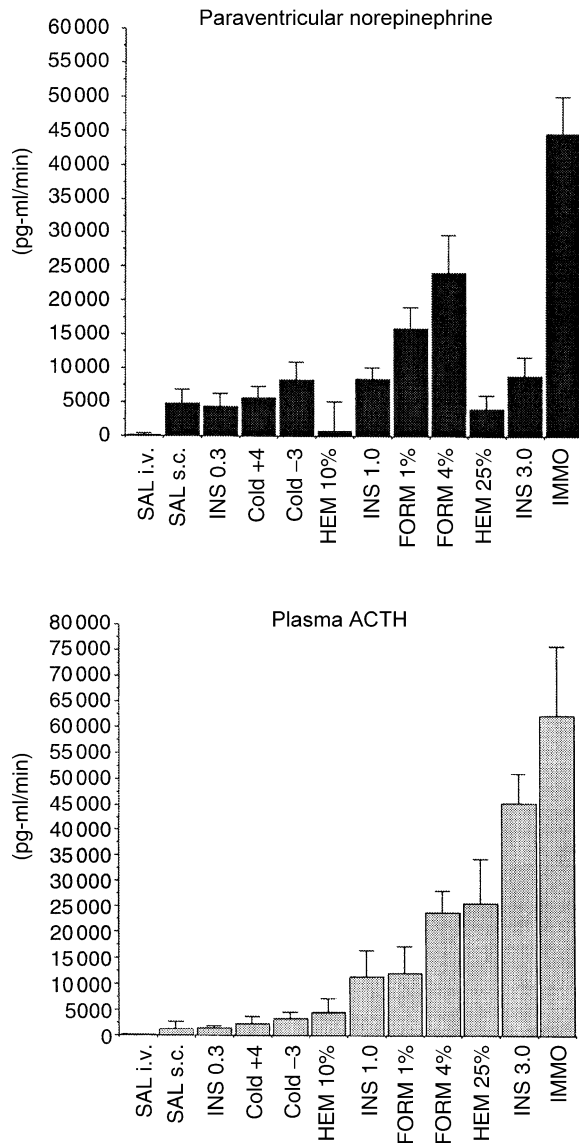


Figure 1 Area under curve measures for microdialysate levels of norepinephrine in the hypothalamic paraventricular nucleus and plasma levels of ACTH from animals exposed to a variety of stressors.

paraventricular nucleus, suggesting that some other paraventricular neurotransmitter or extraparaventricular neuromediators participate in the regulation of the hypothalamic-pituitary-adrenocortical axis following exposure to these stressors. Exposure to pain elicited marked norepinephrine responses in the paraventricular nucleus and moderate plasma ACTH responses, suggesting that paraventricular norepinephrine is responsible for the activation of the hypothalamic-pituitary-adrenocortical axis upon exposure to pain stress. During immobilization stress, the relationship

between paraventricular norepinephrine and plasma ACTH was intermediate, suggesting that both paraventricular norepinephrine and other paraventricular and extraparaventricular neuromediators participate in the regulation of the hypothalamic-pituitary-adrenocortical axis during immobilization stress. Cold stress elicited only minimal paraventricular norepinephrine and plasma ACTH responses, suggesting that other neuroendocrine axes (e.g., thyroid axis) participate in the regulation of homeostasis during acute exposure to cold.

Acute activation of the hypothalamic-pituitary-adrenocortical axis is attended by elevations of plasma glucocorticoids that exert many metabolic effects to mobilize energy necessary for survival of the organism and restoration of homeostasis. Regulatory metabolic effects include gluconeogenesis, glycogenolysis, lipolysis, inhibition of lipogenesis, and proteolysis. Other regulatory effects include anti-inflammatory, anti-allergic, and behavioral effects. Permissive effects of glucocorticoids are important for other hormones or factors to function properly (e.g., catecholamine-induced vascular resistance and cardiac performance are decreased in cases of adrenal insufficiency). Generally, permissive effects of glucocorticoids are responsible for maintenance of a basal or resting state, while regulatory effects are seen when glucocorticoids reach abnormally high concentrations such as when an animal is exposed to an acute and extremely intense stressor.

Hypothalamic-Pituitary-Gonadal Axis

Exposure of animals and humans to acute stress is associated with small and often short-lived increases in plasma luteinizing hormone and androgens. In humans, acute stress-induced increases in androgens appear to be caused by an alteration in plasma volume and a decrease in androgen metabolic clearance rather than from a net increase in androgen release. The mechanisms by which luteinizing hormone is increased under stress conditions are not clear but one possibility is that high levels of ACTH may stimulate gonadotropin hormone-releasing hormone neurons in the hypothalamus. Stress-induced responses are also dependent upon estrogen levels. Thus, in female animals during proestrous or in gonadectomized animals there is no increase in luteinizing hormone upon exposure to acute stress, but instead a rapid decrease in circulating luteinizing hormone may be observed. Norepinephrine and serotonin as well as interleukin-1 also participate in the regulation

of the hypothalamic-pituitary-gonadal axis during acute stress and most likely exert inhibitory effects on gonadotropin-releasing hormone neurons or luteinizing hormone. Follicle-stimulating hormone may follow the same stress response pattern as luteinizing hormone, but often the changes are very small or do not occur at all. The presence of circulating estradiol may provide some protection against the adverse and inhibitory effects of acute stress on the gonadal axis.

In contrast, upon exposure to chronic stress, there are decreases in reproductive function to conserve and redirect energy for other organ systems that are vital for immediate survival. Such inhibition occurs at all levels, including the hypothalamus (inhibition of gonadotropin hormone-releasing hormone secretion), the anterior pituitary (inhibition of luteinizing hormone release), and the gonads (reduced sex hormone secretion). Stress-induced activation of the hypothalamic-pituitary-adrenocortical axis is one of the main factors contributing to this inhibition of gonadal function since corticotropin-releasing hormone and glucocorticoids are known to inhibit gonadotropin hormone-releasing hormone. In addition, glucocorticoids exert inhibitory effects at the level of the anterior pituitary and the gonads.

Sex differences in stress-induced regulation and in responses of primary stress effector systems exist and the presence of circulating estrogen and androgen has been implicated in these differences. It is postulated that estrogens and androgens are responsible for sex hormone-dependent brain organization that reflects differences in responsiveness to stress. For example, estrogen effects on corticotropin-releasing hormone expression may be related to greater stress-induced hypothalamic-pituitary-adrenocortical responses and for a higher incidence of depression, anxiety, and other disorders in human females exposed to stress. In experimental animals, these sex differences are eliminated following surgical removal of the gonads.

Hypothalamic-Pituitary-Growth Axis

It is well documented that acute exposure to various stressors (e.g., hypoglycemia, exercise, hemorrhage, pain) increases secretion of growth hormone from the pituitary gland. In contrast, other stressors such as cold, handling, hypertonic saline, and electric shock produce marked decreases in plasma growth hormone levels. Acute administration of glucocorticoids also increases growth hormone less concentrations in plasma. Thus, it appears that acute stress responses that are associated with elevations in plasma glucocorticoid levels may increase plasma growth hormone levels through stimulation of the growth

hormone gene via glucocorticoid-responsive elements in its promoter region.

In contrast, chronic stress (a state of chronically elevated circulating levels of glucocorticoids) inhibits growth hormone release from the anterior pituitary gland and increases resistance of peripheral tissues to the actions of insulin-like growth factors (e.g., IGF-1). Such inhibition may occur via corticotropin-releasing hormone-induced increases in somatostatin secretion. Thus, dwarfism that occurs in an environment of severe and chronic psychosocial deprivation may reflect a chronically inhibited growth axis.

Little is known about the neural pathways that are involved in stress-induced activation or inhibition of growth hormone secretion. Some evidence suggests that growth hormone responses correlate with changes in catecholamine content in several hypothalamic nuclei, including the paraventricular, dorsomedial, and arcuate nuclei, but more detailed and sophisticated studies are needed to clarify the involvement of catecholamines and other neurotransmitters or neuromodulators in stress-induced growth hormone axis regulation.

Hypothalamic-Pituitary-Thyroid Axis

In contrast to mildly stressful situations, high-intensity stress-induced activation of the hypothalamic-pituitary-adrenocortical axis is associated with inhibition of the thyroid axis. This inhibitory action is reflected by a decrease in release/production of hypothalamic thyrotropin-releasing hormone, decreased production/secretion of thyroid-stimulating hormone, and by inhibition of conversion of thyroxine to the more biologically active triiodothyronine in peripheral tissues. The exception to this pattern is exposure to acute cold stress, forced swimming, or noise stress where profound thyroid axis activation occurs. This is not surprising since during cold exposure, an organism needs to generate significant amounts of energy to maintain body temperature and thyroid hormones are known to increase heat production and therefore core temperature. Upon exposure to cold stress, there is increased secretion of hypothalamic thyrotropin-releasing hormone, which leads to increases in release of thyroxine within 15–20 min and triiodothyronine within 1–2 h. The mechanism by which cold-induced activation of thyrotropin-releasing hormone occurs is not fully understood, although catecholamines are thought to participate in this process. It is well known that in a hypothyroid state, plasma catecholamines are elevated and during a hyperthyroid state plasma catecholamines are either unchanged or decreased. In thyroidectomized rats, larger increments in plasma epinephrine and norepinephrine occur during cold

exposure, suggesting compensatory activation of catecholamine secretion as a result of the loss of the thyroid gland. The organism is faced with difficulty in maintaining body temperature when the hypothalamic-pituitary-thyroid axis is unable to enhance thermogenesis.

The reason that the thyroid axis is inhibited upon exposure to other stressors is unknown but one possible explanation is that an organism conserves energy necessary for survival and restoration of homeostasis. Corticotropin-releasing hormone, somatostatin, and cytokines (e.g., IL-1 and IL-6) most likely contribute to acute stress-induced inhibition of the thyroid axis.

Prolactin

Prolactin is another anterior pituitary hormone that is increased during exposure of animals to various stressors such as exercise, hypoglycemia, hemorrhage, immobilization, and pain stress. Other situations associated with elevated plasma prolactin levels are nipple stimulation during lactation, chest wall lesions (herpes, incisions), and the acute stress of blood withdrawal. The mechanisms by which acute stressors increase prolactin secretion remain unknown. The physiological consequences of prolactin hypersecretion during stress are not clear, although it is hypothesized that prolactin enhances immune function.

Immune System

The immune system is an important component of acute stress responses although interactions between immune system components and other stress effector systems are very complex and still poorly understood. Upon exposure to various acute stressors, inflammatory cytokines such as tumor necrosis factor and interleukin (IL)-1 and -6 are increased. These substances participate in the maintenance of immunologic homeostasis as well as in regulation of the hypothalamic-pituitary-adrenocortical axis. Interleukin-6 appears to be a very potent stimulator of the hypothalamic-pituitary-adrenocortical axis that most likely occurs via activation of paraventricular CRH neurons. Neuronal pathways participating in interleukin-induced CRH activation are still not well characterized. Stress-induced elevated plasma glucocorticoid levels in turn have profound inhibitory effects on various immune functions. For example, production of cytokines, mediators of inflammation, leukocyte function and traffic, cell-mediated immunity, and suppressor T-cell function are all decreased by stress-induced elevations in glucocorticoid levels.

Immunosuppressive and anti-inflammatory effects of glucocorticoids are essential to prevent harmful

infections or unwanted exaggerated immune reactions so as to increase chances for survival. This is well-documented from studies where Lewis rats that are characterized by hyporesponsiveness of the hypothalamic-pituitary-adrenocortical axis (deficient CRH neurons?) have increased susceptibility to inflammatory diseases upon exposure to a variety of stressful stimuli.

Autonomic Nervous System

The sympathoneuronal and adrenomedullary systems are two major stress effector systems. Numerous recent studies have demonstrated differential responses of these two effector systems during exposure to various stressors. In several recent studies, an attempt was made to measure simultaneously release of norepinephrine in the hypothalamic paraventricular nucleus (the most important area in the regulation of the hypothalamic-pituitary-adrenocortical axis and the site where noradrenergic neurons synapse on CRH neurons) using *in vivo* microdialysis and plasma levels of catecholamines and ACTH. These multiple neuroendocrine measures provided an answer to whether stressor-specific noradrenergic activation of the hypothalamic-pituitary-adrenocortical axis and stressor-specific responses of sympathoadrenal and adrenomedullary systems occur. During exposure to any of several stressors, there was a significant correlation between plasma levels of ACTH and paraventricular microdialysate levels of norepinephrine. Although for all stressors levels of norepinephrine in the paraventricular nucleus correlated significantly with plasma ACTH levels, detailed examination of these variables showed differences and not necessarily causal interrelationships. During exposure to metabolic stressors, such as hypoglycemia and hemorrhage, increments of plasma ACTH were much larger than those of norepinephrine in the paraventricular nucleus, suggesting that ACTH release can occur by pathways other than activation of noradrenergic neurons terminating in the paraventricular nucleus. Other intra- or extra-paraventricular non-adrenergic neurons are most likely responsible for stimulating ACTH release during these stressors. In response to pain stress, relatively small ACTH responses were associated with large increases in paraventricular norepinephrine release, suggesting that norepinephrine in this brain region could be a stimulator of CRH neurons. The slope of the relationship between paraventricular norepinephrine release and plasma ACTH responses was intermediate during immobilization stress, suggesting that both noradrenergic and nonadrenergic neurons participate in the regulation of the hypothalamic-pituitary-adrenocortical axis.

Table 1 Selected behavioral and physiological adaptations designed to restore homeostasis during acute exposure to stressful stimulation

<i>Behavioral adaptation</i>	<i>Physiological adaptation</i>
Increased arousal and alertness	Oxygen redirection to CNS and stress-activated tissues
Euphoria or dysphoria	Nutrients redirected to CNS and stress-activated tissues
Suppression of appetite or feeding behavior	Increased blood pressure and heart rate
Suppression of reproductive behavior	Increased respiratory rate
	Increased gluconeogenesis, glycogenolysis, and lipolysis
	Detoxification of endogenous or exogenous toxic products
	Inhibition of reproductive system
	Inhibition of colonic motility

in response to the combined psychological and physical stress of immobilization.

These and other findings can be applied to a new concept of stressor-specific neurocircuits in brain and neuroendocrine signatures that are related to specific stressors. The recognition and description of stressor-specific neurocircuits, including mapping of neurotransmitters participating in these neurocircuits, may lead to the development of new therapies for stress-related disorders.

Types of Acute Stressors

Included below are descriptions of the responses to some of the acute stressors that have been employed frequently in laboratory animal studies of stressful stimulation. In the majority of these studies, the focus of attention was directed at the sympathetic-adrenal medullary system or the hypothalamic-pituitary-adrenocortical system. In the former, measures of plasma levels of norepinephrine and epinephrine were reported and in the latter plasma levels of ACTH and/or corticosterone were made.

Hypoglycemia

Hypoglycemia is characterized by a fall in plasma glucose levels under various conditions, including starvation, hepatic dysfunction, administration of insulin, pancreatic B cell tumor, alcohol use, and adrenal insufficiency. Acute decreases in blood glucose levels are associated with increases in circulating levels of epinephrine, ACTH, and cortisol, with little if any concurrent sympathoneuronal system activation as indicated by circulating norepinephrine levels. The selective adrenomedullary activation during

hypoglycemia constitutes key evidence for the differential regulation of stress effector systems. Hypoglycemia is characterized by many symptoms associated with neuroglycopenia, including tremulousness, sweating, hunger, palpitations, anxiety, psychotic behavior, seizures, and coma.

The central neuronal circuitry responsible for the elicitation of hunger and eating in response to hypoglycemia remains poorly understood. Stimulation of catecholamine release by insulin-induced hypoglycemia is thought to involve both neurogenic and nonneurogenic mechanisms. Neurogenic mechanisms participate in the initial acute phase (the first 30 min after insulin administration) in which plasma epinephrine levels increase but plasma norepinephrine levels remain unchanged. Glucose-sensitive receptors in the hypothalamus and preganglionic cholinergic nerves that innervate the adrenal medulla are involved in this phase when an absolute fall in plasma glucose levels rather than actual hypoglycemia is an important mechanism to trigger catecholamine responses. Nonneurogenic mechanisms participate in the second phase (30–50 min after insulin administration) and produce increases in both plasma epinephrine and norepinephrine. These responses are independent of a functionally intact nerve supply to the adrenal medulla and can be reversed by glucose administration, suggesting that hypoglycemia is the mechanism that triggers catecholamine responses. However, it should be mentioned that several other studies have not found a biphasic pattern of hypoglycemia-induced catecholamine responses, which may be explained by differing experimental conditions, including the severity of hypoglycemia.

Hemorrhage

Depending upon whether hypotensive or nonhypotensive hemorrhage occurs, various physiological responses are elicited, including activation of the hypothalamic-pituitary-adrenocortical axis, increased vasopressin release, and elevated plasma catecholamine levels. Similar to hypoglycemia, hemorrhage elicits relatively small increases in paraventricular levels of norepinephrine. Activation of all of these systems is an important counterregulatory mechanism to maintain homeostasis.

In clinical studies, acute hypotensive hemorrhage is an extremely dangerous situation since it may lead to shock. Acute hemorrhagic shock is characterized by an activation of all of the main stress effector systems. Clinical symptoms of the former include low energy level, fatigue, a feeling of being cold, dizziness, and sleepiness; clinical signs of the latter include cool or mottled extremities, increased heart rate, low blood pressure, pallor, and altered mental status, ranging

from restlessness and agitation to coma. Restoration of blood volume is a high priority, although fluid replacement that is too rapid may be harmful since it may shut down the level of activity of the sympathoadrenal system in the critical period when survival depends on the activity of these stress effector systems.

Hypoxia

Asphyxiation produces the biochemical triad of hypoxemia, hypercarbia, and acute respiratory acidosis. Chemoreceptors, located especially in the carotid bodies adjacent to the carotid sinuses, respond to decreased arterial oxygen concentrations, increased arterial carbon dioxide concentrations, and decreased arterial blood pH. Chemosensitive cells near the ventral surface of the medulla respond to these same stimuli. Acute exposure to hypoxemia and hypercarbic acidosis is associated with increased or unchanged plasma levels of norepinephrine and epinephrine and with increased sympathetic nerve traffic as measured by microneurography. Findings of unchanged plasma norepinephrine levels during acute hypoxemia may be explained by increased nor-epinephrine clearance. The activation of sympathoneuronal and adrenomedullary systems is associated with increases in heart rate, respiration, and vasoconstriction. Sympathectomy decreases and adrenalectomy eliminates increases in heart rate and cardiac output that result from severe hypoxia.

Acute mountain sickness caused by sudden exposure to hypoxia is a frequent cause of morbidity and mortality in people who travel to high altitudes. It is estimated that 30 million people are at risk for altitude-related diseases in the western United States every year. Approximately 20% of tourists who travel to ski resorts in Colorado experience acute mountain sickness. Symptoms of acute mountain sickness include headache, nausea, anorexia, insomnia, cough, difficulties in physical activity, mental changes, seizures, hallucinations, and, finally, in severe cases, coma. All of these symptoms may occur suddenly after a few hours in a high-altitude environment and require immediate medical attention.

Pain

Following exposure of laboratory animals to acute pain induced by formalin administration, only small increments in plasma ACTH levels occur despite large increments in plasma norepinephrine and epinephrine levels. Brainstem centers are the recipients of ascending pathways carrying nociceptive information and are perhaps the most important brain areas in pain mechanisms. Glutamate, aspartate, substance P, and

calcitonin gene-related peptide are neurotransmitters that participate in acute pain responses. It is known that behavioral analgesia or inhibition of spinal neurons that respond to painful stimuli may be eliminated by electric stimulation or opiate microinjection into several brainstem regions.

Cold

Cold exposure has been employed as a stressful stimulus in laboratory studies. Several experimental approaches have been taken in these studies. In one such approach, animals are placed in their home cages in a cold environment maintained at a constant temperature and remain there for varying periods of time. Typically the stress session can extend from several hours to multiple days. Alternatively, animals are immersed in water at a given temperature for less than an hour. Exposure to a cold environment at or below 0°C results in dramatic increases in plasma norepinephrine but only modest increases in plasma levels of epinephrine and ACTH. In contrast, placement of laboratory rats in cold water (20–25°C) for as little as 15 min is attended by substantial increases in plasma levels of epinephrine and norepinephrine. The increases in plasma levels of both catecholamines in response to cold swim stress are negatively correlated with water temperature over the range of 20–34°C, with the highest levels of plasma catecholamines at the lowest water temperatures. Indeed, plasma levels of epinephrine during swim stress at 18–20°C are among the highest ever measured for a laboratory stressor.

Footshock

Footshock has been especially valuable as a stressor in studies of neural and endocrine responsiveness. Footshock allows investigators exquisite control over the frequency, intensity, and duration of the stressful stimulus. In addition, footshock is often employed as an aversive stimulus in studies of learning and memory, allowing for studies that link physiological and behavioral responses. Footshock stress activates the sympathetic-adrenal medullary and hypothalamic-pituitary-adrenocortical systems, especially at moderate to high intensity footshock levels (1–2 mA).

Immobilization

For more than 30 years, Kvetnansky and his colleagues have employed immobilization as a consistent feature of their studies on physiological adaptations to stress. Immobilization is a potent stressor and includes physical as well as psychological dimensions. Typically,

laboratory rats are fixed to a holder by taping their limbs to padded metal mounts. The stress sessions usually last from 30 to 120 min and may be repeated each day for up to 1 month to assess responses to acute versus repeated stress. A single session of immobilization results in extremely high levels of circulating norepinephrine, epinephrine, ACTH, and corticosterone when compared to other laboratory stressors. Levels of norepinephrine, epinephrine, and ACTH usually peak within the first 30 min of immobilization and then decrease to a stable but highly elevated level for the remainder of the stress session. In contrast, plasma levels of corticosterone peak later in the stress session and remain high for the remainder of the period of immobilization.

Summary and Conclusions

Since its inception, the field of stress research has lacked a strong theoretical foundation upon which investigators could formulate questions, design experiments, and interpret their findings and the work of others. From the beginning of his career, Hans Selye offered a definition of stress that was all encompassing and he advanced a view that exposure to any stressor results in a consistent pattern of physiological responses. Although these concepts are no longer at the forefront of stress research, they have left an indelible imprint.

In the absence of a strong theoretical foundation, much of the effort in the field of stress research has been linked to the development of techniques for measuring biologically active molecules or visualizing substances within various tissues, especially neural tissue. The nature of the stressor in a given study was for a time almost an afterthought. Thus, many laboratories developed their own unique stressors and focused much of their creative energies on assay methodologies. Without a set of consistent stress paradigms, comparisons of research results across laboratories were difficult.

In this overview of acute stress responses, we have provided a brief background on the major stress-responsive neuroendocrine systems. Then we examined the pattern of neuroendocrine responses across some of the commonly used laboratory stressors. After distilling the results of thousands of studies of neuroendocrine activation during acute stress, the case for specificity of physiological responses to different stressors is overwhelming. These neuroendocrine signatures allow an animal to respond to a given stressor with an optimal recruitment of physiological responses

to meet the immediate demands of the situation. It is important to keep in mind that most acute stress paradigms present the subject with a great deal of uncertainty and loss of control. The subject has never experienced the stressor before and has no knowledge of how long the stressor will persist, if escape is possible, if stressor intensity will remain constant, and so on. These psychological variables affect central neural circuits that coordinate neural and endocrine activation and must be considered in interpreting the findings of a given acute stress experiment.

See Also the Following Articles

Animal Models (Nonprimate) for Human Stress; Homeostasis; Hypoglycemia.

Further Reading

- Buckingham, J. C., Cowell, A.-M., Gillies, G. E., Herbison, A. E. and Stell, J. H. (1997). The neuroendocrine system: Anatomy, physiology and responses to stress. In: Buckingham, J. C., Gillies, G. E. & Cowell, A.-M. (eds.) *Stress, Stress Hormones and the Immune System*, pp. 9–47. New York: Wiley.
- Chrousos, G. P. (1998). Stressors, stress, and neuroendocrine integration of the adaptive response. *Annals of the New York Academy of Sciences* 851, 311–335.
- Goldstein, D. S. (1995). *Stress, Catecholamines, and Cardiovascular Disease*. New York: Oxford University Press.
- Johnson, E. O., Kamilaris, T. C., Chrousos, G. P. and Gold, P. W. (1992). Mechanisms of stress: A dynamic overview of hormonal and behavioral homeostasis. *Neuroscience and Biobehavioral Reviews* 16, 115–130.
- McCarty, R. and Gold, P. E. (1996). Catecholamines, stress and disease: A psychobiological perspective. *Psychosomatic Medicine* 58, 590–597.
- Pacak, K., Palkovits, M., Kopin, I. J. and Goldstein, D. S. (1995). Stress-induced norepinephrine release in the hypothalamic paraventricular nucleus and pituitary-adrenocortical and sympathoadrenal activity: *In vivo* microdialysis studies. *Frontiers Neuroendocrinology* 16, 89–150.
- Pacak, K., Palkovits, M., Kvetnansky, R., Yadid, G., Kopin, I. J. and Goldstein, D. S. (1995). Effects of various stress on *in vivo* norepinephrine release in the hypothalamic paraventricular nucleus and on the pituitary-adrenocortical axis. *Annals of the New York Academy of Sciences* 771, 115–130.
- Rivier, C. and Rivest, S. (1991). Effect of stress on the activity of the hypothalamic-pituitary-gonadal axis: Peripheral central mechanisms. *Biology of Reproduction* 45, 523–532.

Acute Trauma Response

W C Chiu, D E Carlson and M P Lilly

University of Maryland School of Medicine, Baltimore, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by W C Chiu and M P Lilly, volume 1, pp 17–24, © 2000, Elsevier Inc.

Neuroendocrine Response to Injury
Immune Response to Injury
Metabolic Response to Injury
Interactions of the Biological Responses
Therapy and Modification of Responses

Glossary

<i>Apoptosis</i>	Programmed cell death; deletion of cells by fragmentation into membrane-bound particles that are phagocytosed by other cells.
<i>Cytokine</i>	A protein that acts as an intercellular mediator in the generation of the immune response.
<i>Glucocorticoids</i>	A class of steroid hormones, secreted by the adrenal cortex, that influence intermediary metabolism and exert an anti-inflammatory effect.
<i>Reflex</i>	A reaction in response to a stimulus applied to the periphery and transmitted to the nervous centers in the brain or spinal cord.
<i>Sodium pump</i>	Sodium-potassium adenosine triphosphatase, an enzyme responsible for maintaining the sodium and potassium electrical gradient between the intracellular and interstitial compartments.

The acute trauma response to injury encompasses a complex interplay of multiple host homeostatic processes. Although the overall response to trauma is most easily conceptualized by considering the individual components of the neuroendocrine, metabolic, and immune systems, it is important to recognize that many molecular, cellular, humoral, and physiological mechanisms and cascades develop interactively. The ultimate ability of the injured patient to recover from the insult depends on the dynamic relationship between this host response and timely medical or surgical intervention ([Figure 1](#)).

Neuroendocrine Response to Injury

The classic neuroendocrine reflex is comprised of afferent stimuli, central nervous system integration and modulation, and efferent output. Following traumatic injury, multiple discrete chemical and physiological perturbations are detected by specific sensors. After local transduction of stimuli, these signals are transmitted to the central nervous system via specific neural pathways and directed to highly specialized cellular nuclei for processing and integration with other signals. The resultant efferent output is the secretion of various hormonal substances by individual organ systems. Modulation of the intensity and pattern of release of these hormones is affected by short reflex arcs and long regulatory feedback loops, whether stimulatory or inhibitory.

Afferent Input

After a severe injury, tissue destruction and hemorrhage appear to be two of the most important early alterations. From the site of injury, peripheral sensory receptors of primary nociceptive afferent neurons detect mechanical deformation and thermal changes. Damaged cells release chemicals, such as histamine, potassium, serotonin, thrombin, and others, which also provide nociceptive input. The cell bodies of these neurons reside in the dorsal root ganglia and transmit signals to the spinal dorsal horn, where secondary neurons and collateral pathways project to tertiary neurons in the brain stem and thalamus.

Hypovolemia is the most common cause of circulatory shock after injury. Significant blood loss leads to a decrease in effective circulating volume and blood pressure. Tissue injury and fluid sequestration may contribute to hypovolemia. These changes are sensed by pressure-sensitive baroreceptors in the aorta and carotid arteries and volume-sensitive stretch receptors in the atria and vena cava. Fibers from these primary neurons ascend in the vagus nerve to the nucleus of the solitary tract. Secondary neurons project to the medullary reticular formation and cardiovascular pressor and depressor areas of the brain stem.

Many other alterations, including fear and anxiety or changes in oxygen concentration, temperature, and energy substrates, provide afferent input to the neuroendocrine system. The perception of danger or threat of injury stimulates the central limbic system with fiber projections to the hypothalamus. Optic, olfactory, or auditory information may provide substantial

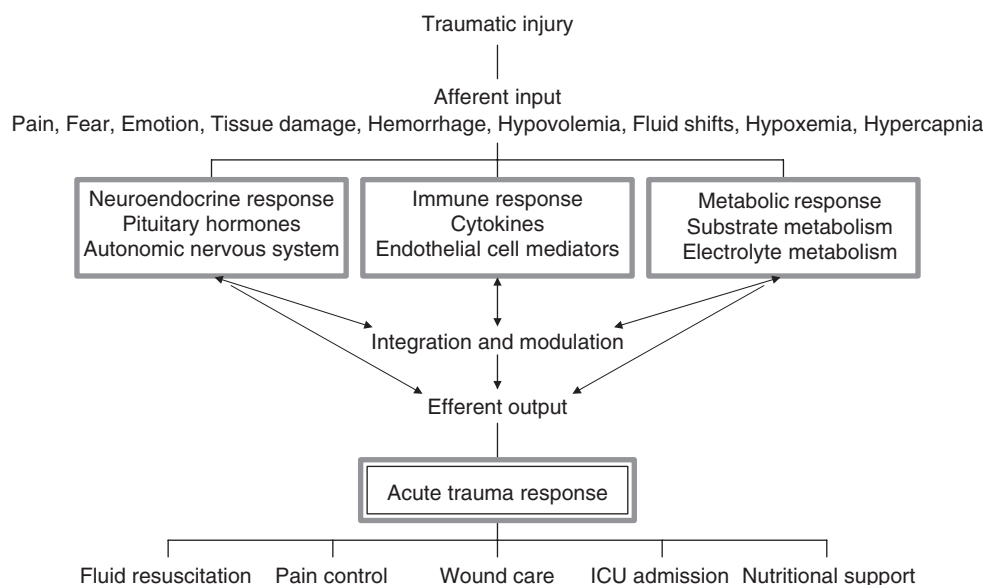


Figure 1 Schematic diagram depicting the interactions of the neuroendocrine, immune, and metabolic systems after traumatic injury. Processing and integration of multiple stimuli, and modulation through the interactions of various mediators, form the overall biological response. Modification of this response by therapeutic interventions facilitates achieving and maintaining homeostasis.

signals. Injuries that affect airway, breathing, or circulation may result in hypoxia, hypercarbia, or acidosis. Chemoreceptors located in the carotid and aortic bodies and in the central nervous system detect decreases in the plasma concentration of oxygen and increases in the concentration of carbon dioxide or hydrogen ions. Fibers from these neurons project to both the cardiovascular and respiratory centers of the medulla. Hemorrhage and exposure after trauma leading to hypothermia and core temperature changes are sensed in the hypothalamic preoptic area.

Integration and Modulation

Multiple signals from different pathways are modulated at several levels. The primary input depends on the nature, intensity, and duration of the stimulus. Sequential or repetitive stimulation, as well as the status and responsiveness of the receptor, all influence the response. The sensory signal arrives at the medullary center for integration with other ascending signals, interneurons, and descending inhibitory inputs. The various signals from nociceptors, baroreceptors, chemoreceptors, somatosensory afferents, and limbic system project from the brain stem nuclei to converge at the hypothalamus.

Efferent Output

Pituitary regulation There are two main components to the efferent limb of the neuroendocrine response: hormones under hypothalamic-pituitary axis control and hormones under autonomic nervous

system control. The hypothalamic-pituitary-adrenal (HPA) axis has been extensively studied and represents a model neuroendocrine response system to stress. Although corticotropin-releasing hormone (CRH), adrenocorticotropic hormone (ACTH), and cortisol are secreted sequentially, the fine control of the system is achieved through neural integration, feedback loop regulation, and intermediary substances. Multiple inputs are integrated at the hypothalamic paraventricular nucleus (PVN) to generate the efferent arc.

CRH is the primary stimulant of the synthesis and secretion of ACTH by corticotrophs in the anterior pituitary gland. ACTH is released as a cleavage product of proopiomelanocortin and subsequently acts to stimulate the synthesis and release of glucocorticoids from the adrenal cortex.

Cortisol is the primary glucocorticoid in the response to stress and has widespread effects on host metabolism. Among its many metabolic actions are potentiation of glucagon and epinephrine leading to hyperglycemia, stimulation of gluconeogenesis in the liver, stimulation of amino acid release in skeletal muscle leading to proteolysis, and increase in lipolysis and fatty acid release. Cortisol is also essential in the compensatory hyperosmolar response necessary to achieve complete blood volume restitution. The magnitude of injury influences the intensity of the CRH-ACTH-cortisol response.

While earlier hemorrhage potentiates the ACTH response, circadian factors may mask this effect. Along with the elevation in plasma cortisol levels, the normal diurnal variation is lost. This altered

pattern of adrenal cortisol secretion is associated with a similar loss of diurnal variation in both ACTH secretion and the adrenocortical sensitivity to ACTH stimulation. In pure hypovolemia, plasma cortisol levels rapidly normalize when blood volume has been restored. Persistently elevated cortisol suggests ongoing stimulation from tissue hypoxia or supervening infection and is associated with reduced survival.

Hypothalamic thyrotropin-releasing hormone (TRH) stimulates the release of thyroid-stimulating hormone (TSH) from the anterior pituitary. TSH is responsible for the release of primarily thyroxine (T_4) from the thyroid gland, which is converted to triiodothyronine (T_3) peripherally. Thyroid hormones also have multiple metabolic effects, the most important of which are increases in oxygen consumption and heat production. After injury, peripheral conversion of T_4 to T_3 is impaired and there is no compensatory rise in TSH release.

The role of the endogenous opioids after injury and hemorrhage is an area undergoing active investigation. The opioids, including the endorphins and enkephalins, are secreted by cells in the central nervous system, intestinal wall, and adrenal medulla. They act upon various receptors to produce different effects. In the central nervous system, opioid peptides act as neurotransmitters to attenuate pain by inhibiting nociception. The excess release of endogenous opioid peptides has been linked to the cardiovascular collapse of shock. One suggested mechanism for opioid-induced hypotension after acute hemorrhage is a decrease in sympathetic outflow, mediated by serotonergic pathways. After hemorrhagic or septic shock, the administration of the opiate receptor antagonist naloxone has been shown to improve cardiovascular function, systemic blood pressure, and survival time.

Growth hormone (GH), gonadotropins, and prolactin are other anterior pituitary hormones involved in the neuroendocrine response. GH promotes protein synthesis and is elevated following injury. Gonadotropin-releasing hormone (GnRH) stimulates the release of follicle-stimulating hormone (FSH) and luteinizing hormone (LH), both of which are suppressed after injury. Although the hypothalamus exerts tonic suppression of pituitary prolactin release, elevated levels have been reported after injury in adults.

Arginine vasopressin (AVP), or antidiuretic hormone, is synthesized in the hypothalamus and transported to the posterior pituitary gland for storage. Secretion of AVP is increased immediately after injury. Stimuli for AVP release include increased plasma osmolality and changes in effective circulating volume. Actions of AVP include reabsorption of free water in the renal distal tubules and collecting ducts,

peripheral and splanchnic vasoconstriction, and hepatic glycogenolysis and gluconeogenesis.

Autonomic regulation The catecholamines are essential components of the physiological response to injury. Norepinephrine is primarily released from the axon terminals of postganglionic neurons with subsequent overflow into plasma. The rate of norepinephrine release correlates with the intensity of sympathetic nervous system activity. Epinephrine is directly secreted into the circulation from stimulation of the adrenal medulla. Like other hormones, the catecholamines exert both hemodynamic and metabolic actions, including enhanced cardiovascular function, vasoconstriction, hepatic glycogenolysis and gluconeogenesis, and peripheral lipolysis.

The renin-angiotensin-aldosterone system serves to maintain intravascular volume. Renin catalyzes the conversion of angiotensinogen to angiotensin I within the kidney, which in turn is converted into angiotensin II in the pulmonary circulation. Angiotensin II and ACTH are stimuli for the secretion of aldosterone from the adrenal cortex. The predominant action of aldosterone is enhancement of the Na^+/K^+ -ATPase enzyme in the renal distal convoluted tubules and collecting ducts to increase reabsorption of sodium and elimination of potassium.

Insulin and glucagon play opposing roles in the maintenance of substrate homeostasis. Insulin is the primary anabolic hormone and promotes the storage of carbohydrate, protein, and lipid. After injury, sympathetic nervous system activity inhibits insulin secretion from pancreatic beta islet cells and cortisol inhibits glucose uptake with the net result of hyperglycemia and catabolism. Glucagon is synthesized in the pancreatic alpha islet cells and increases protein and lipid catabolism by increasing hepatic glycogenolysis and gluconeogenesis.

Immune Response to Injury

Immune cells and inflammatory mediators activated by tissue injury are important components in the response to trauma. Among the substances released by activated inflammatory cells are cytokines, endothelial cell factors, and eicosanoids. The actions of these varied mediators in hemorrhage, sepsis, and injury have been the subject of intense investigation.

Cytokines

Cytokines are a class of small proteins produced by the full range of inflammatory cells. They are active at low concentrations and may have paracrine, autocrine, or endocrine effects. Cytokines regulate the production of immune cells and other cytokines by

either potentiating or attenuating the inflammatory response.

Manifestations of cytokine response include fever and alterations in cardiovascular and respiratory function. Excessive cytokine release has been implicated in the ongoing inflammation and injury associated with sepsis and multiple organ failure. The number of cytokines involved is too numerous to list here, but the best characterized include the interleukins (ILs), tumor necrosis factor- α (TNF- α), the interferons, and granulocyte/macrophage-colony stimulating factor (GM-CSF).

After injury, TNF- α , or cachectin, is one of the earliest cytokines released by monocytes, macrophages, T cells, and endothelial cells. TNF- α has been shown to act synergistically with IL-1 to induce hypotension, tissue injury, and death in animals. Its metabolic effects include proteolysis in skeletal muscle and hepatic gluconeogenesis.

The interleukins are a family of cytokines also derived from monocytes, macrophages, T cells, and endothelial cells. The best characterized of these cytokines in trauma is IL-1. The two species IL-1 α and IL-1 β have differential actions, but IL-1 β is believed to be more biologically relevant. Among its many systemic effects are mediating fever by stimulating the synthesis of E series prostaglandins in the thermoregulatory center of the hypothalamus, promoting synthesis of hepatic acute phase proteins, and promoting skeletal muscle proteolysis.

Interferon- γ is a glycoprotein released by stimulated helper T cells. It activates circulating and tissue macrophages to release TNF- α and other interleukins. GM-CSF is a growth factor cytokine that plays a prominent role in wound healing after injury. It augments maturation of functional leukocytes and delays the apoptosis of immune cells.

Endothelial Cell Mediators

Endothelial cell-derived substances have predominantly local actions of vasomotor regulation and modulation of coagulation, but may have additional systemic effects on cardiovascular and neuroendocrine systems, as well as cytokine stimulation. Nitric oxide (NO) was formerly known as endothelium-derived relaxing factor. It has been shown that a basal release of NO induces a tonic state of arteriolar vasodilatation. Cellular injury, endotoxin, and hypoxia stimulate increased release of NO, and its effects are primarily exacerbation of the vasodilatory state. Conversely, the endothelins exert potent vasoconstrictor properties. Endothelin-1 is the most biologically active member of this family and is estimated to have 10 times more vasoconstrictor potency than angiotensin II. Increased serum levels of endothelin-1

correlate with the severity of injury. Platelet-activating factor stimulates production of thromboxane (TXA₂) and promotes platelet aggregation. Its actions include increasing vascular permeability by altering the shape of endothelial cells.

Eicosanoids

The eicosanoids are phospholipid mediators derived from arachidonic acid through the action of phospholipase A and include the prostaglandins, thromboxanes, and leukotrienes. Actions of prostaglandins and prostacyclin include increased vascular permeability, leukocyte migration, and vasodilation. TXA₂ is synthesized in platelets, acts as a potent vasoconstrictor, and promotes platelet aggregation. The leukotrienes mediate anaphylaxis by promoting capillary leakage, bronchoconstriction, and vasoconstriction.

Metabolic Response to Injury

Substrate Metabolism

Altered substrate metabolism after injury is often compared and contrasted with that during fasting. The adaptive responses to fasting of decreased substrate utilization and decreased metabolic rate are modified by the stress responses to pain, tissue damage, hemorrhage, and inflammation. The early phase is characterized by generalized catabolism, protein wasting, negative nitrogen balance, and hyperglycemia. Fever, increased myocardial work, and early wound healing all increase energy expenditure and metabolic rate. After existing hepatic glycogen stores are depleted, lipids become the primary source of fuel. Lipolysis is enhanced by catecholamines, cortisol, and glucagon. Despite increased insulin secretion, hyperglycemia results from increased hepatic glycogenolysis and gluconeogenesis, mediated by cortisol and catecholamines, and in the periphery, from inhibition of insulin-mediated glucose uptake.

The increased urinary excretion of nitrogen reflects net proteolysis. The severity of injury correlates with the degree of total body protein turnover. Following trauma, the major source of amino acids is skeletal muscle, the primary carriers of nitrogen being alanine and glutamine. Although the total amino acid concentration is unchanged, changes in alanine and glutamine synthesis and release depend on the extent of overall injury and oxygen consumption.

Electrolyte Metabolism

Decreased effective circulating volume is exacerbated by tissue damage and obligatory sequestration of fluid into extravascular compartments. After trauma, not only is there interstitial accumulation of fluid from the capillary leak of injured and inflamed tissue,

but there is also mobilization of salt and water into the intracellular compartment. Previous studies have shown that the cellular membrane sodium-potassium adenosine triphosphatase (Na^+ , K^+ -ATPase), or sodium pump, is impaired in shock leading to muscle cell membrane depolarization. An influx of sodium and chloride ions, followed by water, results in cell swelling and further depletion of extracellular volume, and may lead to irreversible shock.

The renal conservation of salt and water may also be affected by increased glomerular filtration fraction, increased proximal reabsorption of sodium, and increased blood flow to juxtamedullary nephrons through aldosterone and glucocorticoid effects. Increased levels of AVP contribute to water retention and may manifest as oliguria. In the presence of hypovolemia or hypotension, there is increased risk of acute tubular necrosis. Adequate resuscitation with balanced salt solutions helps to maintain medullary osmolar gradient and tubular fluid flow.

Interactions of the Biological Responses

Although the effects of individual hormonal and immune substances have been well characterized, their interrelationships in the overall biological response have yet to be well elucidated. Multiple stimuli certainly interact and converge in the central nervous system. This concept is supported by studies demonstrating potentiated responses of ACTH to hemorrhage after simultaneous or repetitive application of subthreshold stimuli.

One focus of interest has been the interactions of immune mediators with the HPA system. Impaired host wound healing and decreased resistance to infection in some critically injured patients have popularized the concept of injury-induced immunosuppression. Although it is known that pharmacological doses of corticosteroids produce immunosuppression, it is unclear whether the physiological increase in endogenous cortisol directly causes immunosuppression. Instead, the immune dysfunction after injury may also be mediated by the stimulation of the autonomic nervous system or other hormones.

There is evidence that IL-1 may activate the HPA axis by a central action on the hypothalamus. IL-1 increases the release of CRH from isolated rat hypothalamus, and other studies have shown that IL-1-stimulated ACTH release is associated with decreased CRH content in the median eminence. Other investigators have reported increased levels of circulating ACTH and corticosterone in rats after intravenous injection of IL-1.

It is still unclear whether IL-1 has a direct or indirect effect on the HPA axis. There is some evidence that these actions of IL-1 may be mediated by

prostaglandins. Microinjections of prostaglandin antagonist into the preoptic area block the ACTH response to intravenous IL-1 β . In addition, the effect of IL-1 β on ACTH release may not be a result of changes in CRH. Reports have suggested that IL-1 β stimulates norepinephrine and dopamine release in the hypothalamus and that the ACTH response to IL-1 injection in the median eminence can be blocked by adrenergic antagonists. Similarly, both IL-6 and TNF- α have also been associated with increased levels of circulating ACTH. Like IL-1, this stimulatory effect may also be related to activation of the HPA system through a prostaglandin-mediated increase in CRH release.

Studies have documented the immunosuppressive effects of androgens, citing differences between male and female immune function. Androgen-deficient mice have enhanced IL-2 and interferon- γ production by peripheral T cells. In addition, it has been shown that castrated male mice subjected to soft tissue trauma and hemorrhage had elevated plasma corticosterone levels and splenocyte proliferation compared to sham animals.

Local production of various hormones from immune cells may lead to paracrine effects. Lymphocytes have been shown to produce ACTH, a TSH-like peptide, and a GH-like peptide, among others. Immune cells also possess various hormone receptors. Receptors for CRH, TRH, growth hormone releasing hormone, ACTH, and opioids have been found on leukocytes. These findings certainly point to a communication between the immune and neuroendocrine systems. Considerably more research is needed to further understand these complex interactions.

Therapy and Modification of Responses

Fluid Resuscitation

Initial management after severe injury involves control of hemorrhage, administration of intravenous isotonic crystalloid fluids, and blood transfusion, if needed, which immediately improves effective circulating volume. The replenishment of intravascular volume results in an increase in capillary hydrostatic pressure, reversing the flux of water out of the interstitial compartment. Reflex sympathetic vasoconstriction and catecholamine release are diminished as increased blood pressure is sensed by baroreceptors. The result is improved capillary perfusion and organ function, and the prevention of the adverse consequences of tissue ischemia and cellular hypoxia.

Recent evidence suggests that when severe hemorrhage accompanies trauma, delaying intravenous fluid resuscitation until bleeding is controlled can reduce morbidity and mortality. Aggressive and rapid fluid resuscitation may accentuate blood loss

when hemorrhage is not yet controlled. Furthermore, rapid fluid resuscitation may accelerate reperfusion organ injury so that sufficient fluid should be infused at only a moderate rate to restore optimal intravascular volume.

Pain Management

Administration of narcotic analgesics provides pain control, reduces sympathetic nervous system activity, reduces the release of endogenous opioids, and decreases cortical and limbic inputs. Studies have shown that preoperative administration of morphine diminishes the usual rise in plasma ACTH, cortisol, GH, and glucose. Furthermore, studies have shown that denervation of the wound or regional anesthesia ablates afferent signals and suppresses adrenal stimulation.

Local Wound Care

Tissue damage induces the local release of cytokines, immune complexes, and complement split products from phagocytes and lymphocytes. An exaggerated local inflammatory response from the wound may amplify and cause adverse systemic effects. Therefore, early removal of debris and debridement of devitalized tissues will attenuate this inflammation. Stabilization of fractured limbs prevents repetitive soft tissue trauma, and fasciotomy relieves compartment syndrome to improve tissue perfusion.

Intensive Care Unit Admission

One of the goals of resuscitation is to restore and maintain optimal perfusion and prevent tissue ischemia. Vasomotor changes, myocardial dysfunction, shock, hypothermia, and respiratory insufficiency all contribute to the difficulty of gauging the adequacy of therapy without intensive care unit (ICU) admission. In an ICU environment, all physiological functions may be closely monitored for evidence of deterioration and to assess response to therapy. A pulmonary artery catheter allows the monitoring of cardiac filling pressures and cardiac output. Systemic oxygen delivery and consumption estimates can be derived from these measurements to facilitate resuscitation.

Nutritional Support

Most victims of mild injuries easily withstand the increased metabolic demand and limited catabolism without nutritional supplementation. Those with severe injuries require replenishment of nutritional substrates to avoid excessive energy depletion and protein wasting. Nutritional formulas are provided by either the enteral or parenteral route. The beneficial effects of enteral nutrition include maintenance of gut mucosal barrier function and enhanced production of trophic gut hormones.

Although glutamine is the most abundant amino acid in the body, its concentration in plasma and tissues is depleted after injury. Studies have shown that glutamine-enriched nutrition can improve muscle protein synthesis and improve nitrogen balance. Arginine, a nonessential amino acid, has been shown to augment macrophage and natural killer cell activity. In critical illness, supplemental arginine improves nitrogen retention and wound healing. The majority of the experimental data on these immune-enhancing formulas is still too inconclusive to justify their use.

Summary

The acute trauma response consists of multiple biological components reacting to various signals from local wounded tissue and systemic alterations. In severe injury, exaggerated and unchecked release of mediators result in detrimental systemic effects and may lead to multiple organ failure. In addition to host endogenous modulation of these mediators, timely therapeutic interventions may modify local inflammation, neuroendocrine signals, and systemic alterations. Recovery from trauma ultimately depends upon the coordination of the neuroendocrine, immune, and metabolic systems toward restoring homeostasis and survival. Future research in this area may elucidate new therapeutic options for the critically injured trauma patient.

Acknowledgments

This article was supported in part by National Institute of General Medical Sciences Grant R01 GM 063050 (D.E. Carlson) and National Institute of Diabetes and Digestive and Kidney Diseases Grant K08 DK 02181 (Clinical Investigator Award, M.P. Lilly).

See Also the Following Articles

Endocrine Systems; Hypotension, Hypovolemia, and Septic Shock; Immune Response; Immune Suppression; Neuroendocrine Systems; Inflammation.

Further Reading

- Carlson, D. E., Chiu, W. C. and Scalea, T. M. (2006). Cecal ligation and puncture in rats interrupts the circadian rhythms of corticosterone and adrenocortical responsiveness to adrenocorticotrophic hormone. *Critical Care Medicine* 34, 1178–1184.
- Chaudry, I. H., Ayala, A., Ertel, W. and Stephan, R. N. (1990). Hemorrhage and resuscitation: immunological aspects. *American Journal of Physiology* 259 (Regulatory Integrative Comp. Physiol. 28), R663–R678.

- Faist, E., Schinkel, C. and Zimmer, S. (1996). Update on the mechanisms of immune suppression of injury and immune modulation. *World Journal of Surgery* 20, 454–459.
- Gann, D. S. and Foster, A. H. (1994). Endocrine and metabolic responses to injury. In: Schwartz, S. I., Shires, G. T. & Spencer, F. C. (eds.) *Principles of surgery* (6th edn., pp. 3–59). New York: McGraw-Hill, Inc.
- Gann, D. S. and Lilly, M. P. (1984). The endocrine response to injury. In: Wilder, R. J. (vol. ed.) *Multiple trauma*, Massion, W. H. (series ed.) *Progress in critical care medicine*, pp 15–47 (Vol. 1). Basel, Switzerland: S. Karger AG.
- Gann, D. S. and Wright, P. A. (1995). Shock – The final common pathway? *Advances in Trauma and Critical Care* 10, 43–59.
- Guirao, X. and Lowry, S. F. (1996). Biologic control of injury and inflammation: much more than too little or too late. *World Journal of Surgery* 20, 437–446.
- Koretz, R. L. (1995). Nutritional supplementation in the ICU: how critical is nutrition for the critically ill? *American Journal of Respiratory and Critical Care Medicine* 151, 570–573.
- Lilly, M. P. and Gann, D. S. (1990). The neuroendocrine response to injury. In: Border, J. R., Allgöwer, M., Hansen, S. T. Jr. & Rüedi, T. P. (eds.) *Blunt multiple trauma: comprehensive pathophysiology and care*, pp. 167–190. New York: Marcel Dekker, Inc.
- Lilly, M. P. and Gann, D. S. (1992). The hypothalamic-pituitary-adrenal-immune axis: a critical assessment. *Archives of Surgery* 127, 1463–1474.
- Lilly, M. P., Jones, R. O., Putney, D. J. and Carlson, D. E. (2000). Post-surgical recovery and time-of-day mask potentiated responses of ACTH to repeated moderate hemorrhage in conscious rats. *Journal of Endocrinology* 167, 205–217.
- Lin, E., Lowry, S. F. and Calvano, S. E. (1999). The systemic response to injury. In: Schwartz, S. I., Shires, G. T. & Spencer, F. C. (eds.) *Principles of surgery* (7th edn., pp. 3–51). New York: McGraw-Hill, Inc.
- Madden, K. S., Sanders, V. M. and Felten, D. L. (1995). Catecholamine influences and sympathetic neural modulation of immune responsiveness. *Annual Review of Pharmacology and Toxicology* 35, 417–448.
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* 338, 171–179.
- Monk, D. N., Plank, L. D., Franch-Arcas, G., Finn, P. J., Streat, S. J. and Hill, G. L. (1996). Sequential changes in the metabolic response in critically injured patients during the first 25 days after blunt trauma. *Annals of Surgery* 223, 395–405.
- Schadt, J. C. (1989). Sympathetic and hemodynamic adjustments to hemorrhage: a possible role for endogenous opioid peptides. *Resuscitation* 18, 219–228.
- Schlag, G. and Redl, H. (1996). Mediators of injury and inflammation. *World Journal of Surgery* 20, 406–410.
- Shah, K. J., Chiu, W. C., Scalea, T. M. and Carlson, D. E. (2002). Detrimental effects of rapid fluid resuscitation on hepatocellular function and survival after hemorrhagic shock. *Shock* 18, 242–247.

Addison's Disease See: Adrenal Insufficiency.

Adenylyl Cyclases and Stress Response

F A Antoni

University of Edinburgh, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by F A Antoni, volume 1, pp 25–26, © 2000, Elsevier Inc.

Properties of Adenylyl Cyclase

Adenylyl Cyclase in Stress

Glossary

Adenylyl cyclases Enzymes (EC 4.6.1.1) that generate the second messenger adenosine 3'5' monophosphate (cAMP) from ATP.

Heterotrimeric G-proteins

A group of proteins that consist of three subunits (α , β , γ) and couple cell surface receptors to their effector enzymes. On activation of cell surface receptors by their ligands, heterotrimeric G-proteins dissociate into the α subunit and the $\beta\gamma$ complex, both of which regulate the activity of effector enzymes such as adenylyl cyclase in the cell membrane.

Protein kinases

Enzymes that tag proteins with phosphoryl residues derived from ATP. The activity of protein kinases may be regulated by a variety of intracellular messengers, protein–protein interactions, or phosphorylation by other protein kinases.

Second messengers

Chemicals generated by extracellular stimuli arriving at the cell surface that transduce the extracellular signal toward the interior of the cell.

Properties of Adenylyl Cyclase

Signaling through cAMP is fundamentally important in the central nervous system. Genetic and pharmacological studies have revealed that alterations of both the synthetic (adenylyl cyclase, AC) and the catabolic (cyclic nucleotide phosphodiesterase, PDE) limbs of the cAMP cascade may cause a profound disruption of learning and memory. Further work has implicated cAMP signaling in the development of opiate dependence, alcoholism, the disorders of mood and affect, and neuronal death.

There are 10 genes encoding AC known in the human genome, and orthologs of all of these are found in mouse and rat. One of these genes encodes a soluble adenylyl cyclase that is a sensor for bicarbonate ions and is highly abundant in the choroid plexus but is otherwise absent from brain parenchyma. Nine of the genes encode enzymes that are integral membrane proteins. These ACs are differentially controlled by heterotrimeric G-proteins, Ca^{2+} , and protein phosphorylation (Figure 1). In turn, heterotrimeric G-proteins are regulated by seven transmembrane domain receptors on the cell surface. Several isoforms of adenylyl cyclase are expressed in the brain and show unique topographical distribution.

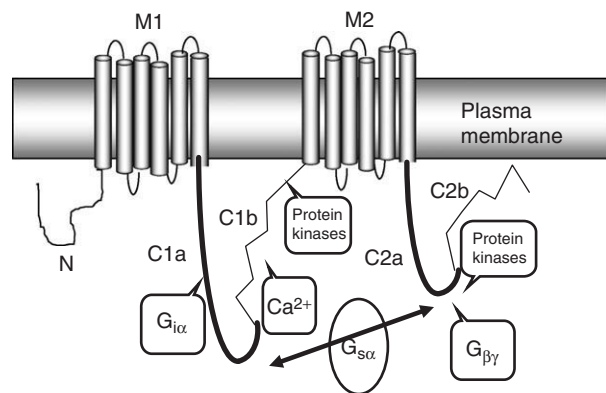


Figure 1 Schematic representation of mammalian adenylyl cyclases and their control. C1a and C2a, homologous catalytic domains (thick line); C1b and C2b, nonhomologous regulatory domains (thin lines); $\text{G}_{\beta\gamma}$, $\text{G}_{i\alpha}$, and $\text{G}_{s\alpha}$, G-proteins; M1 and M2, intramembrane domains; N, N-terminal cytoplasmic loop. Approximate sites of isotype-specific modulation reported for various cyclases to date are indicated in boxes along the polypeptide chain. The C1a and C2a domains must come into physical contact to form the catalytic core. This is facilitated by $\text{G}_{s\alpha}$, which binds to the C2a domain. Most adenylyl cyclases are inhibited by $\text{G}_{i\alpha}$, which binds to a site on the C1a domain.

Colocalization of ACs with distinct functional properties within the same cell has been demonstrated. This raises the possibility of the formation of cellular domains, for example, neuronal growth cones, where cAMP-dependent signaling has functional characteristics distinct from other parts of the cell such as the cell body. Further, the formation of subplasmalemmal microdomains with different types of ACs and PDEs also seems possible.

Some ACs such as AC types I and VIII are stimulated by intracellular Ca^{2+} and calmodulin. These enzymes are prominently expressed in the cerebral cortex and are important for learning and memory. Type I AC shows synergistic stimulation by Ca^{2+} /calmodulin and G_s -coupled receptors. Therefore, this enzyme may function as a co-incidence detector for the activation of ionotropic glutamate receptors and G_s -coupled receptors for a wide range of neuro-effectors. Another type of AC, isotype II, is synergistically stimulated by $\text{G}_{s\alpha}$ and $\text{G}_{\beta\gamma}$, or $\text{G}_{s\alpha}$ and protein kinase C-dependent phosphorylation. These properties suggest that the type II enzyme is a co-incidence detector for neurotransmitter receptors coupled to G_s and $\text{G}_{i/o}$, or G_s and G_q . These pairings are remarkable because they suggest that the direction as well as the magnitude of the effects of neurotransmitters on cAMP synthesis are dependent on the context of the stimulus applied. A third group of ACs is inhibited by Ca^{2+} ions. These enzymes are likely to participate in signaling circuits in which periodic surges of cAMP levels alternate with increases of intracellular Ca^{2+} (i.e., intracellular oscillations). Typically, in neuroendocrine cells, cAMP produces depolarization of the membrane potential and a rise in intracellular Ca^{2+} . The rise in intracellular Ca^{2+} triggers various effector mechanisms (e.g., secretion or gene expression) and will also provide negative feedback to the AC system to limit the rise of cAMP levels. Depending on the kinetics of cAMP synthesis and degradation, the cell may repolarize and return to the unstimulated state or enter a phase of oscillatory behavior.

Adenylyl Cyclase in Stress

Stressful stimuli increase the level of cAMP in the neocortex and the adenohypophysis. In the cerebral cortex, this is thought to be the consequence of the activation of AC through G_s -coupled receptors such as β_1 -adrenergic and corticotropin releasing hormone (CRH) type 1 and type 2 receptors. Mice overexpressing the CRH receptor present with an increased anxiety phenotype, indicating that cAMP could be an intracellular mediator of stressor activation. However, targeted deletion of the genes for AC I and VIII produced an anxious phenotype. Thus,

it is not warranted to generalize that cAMP is a mediator of stress and anxiety. In the anterior pituitary gland, cAMP is the major intracellular messenger activated by CRH. Here, the AC system operates as a co-incidence detector for activation by $G_{s\alpha}$, and protein kinase C-dependent phosphorylation.

Adrenal corticosteroid hormones interfere with the regulation of the cAMP pathway both in the short and the long term. Examples of long-term influence include control of the levels of AC in the neocortex and hippocampus. In these tissues, Ca^{2+} /calmodulin-stimulated AC activity is reduced by adrenalectomy and repeated stress. By contrast, the level of type II AC mRNA is increased by repeated stress in the hippocampus. An intriguing possibility is the induction of a protein called activator of G-protein signaling (AGS1, also known as Dex-ras1) by glucocorticoids. This intracellular protein activates G_i and thus reduces the biosynthesis of cAMP by ACs that are sensitive to $G_{i\alpha}$. Significantly, there are also ACs (isoforms II, IV, and VII) that are activated by $G_{\beta\gamma}$ subunits derived from G_i . Thus, by regulating the levels of AGS1, glucocorticoids could alter the characteristics of the cAMP response, essentially eliminating the action of receptors coupled to inhibitory G-proteins. With respect to short-term effects, in pituitary corticotrophs Ca^{2+} -mediated feedback inhibition of cAMP accumulation is essential for the early glucocorticoid feedback inhibition of adrenocorticotrophic hormone (ACTH) release stimulated by CRH. If cAMP levels rise above the range controlled by Ca^{2+} inhibition, glucocorticoid feedback is markedly diminished. This is the basis of the escape from glucocorticoid feedback induced by arginine vasopressin (AVP) at the pituitary level.

In peripheral tissues, cAMP is also a major intracellular mediator of the stress response in the adrenal cortex, liver, and fat tissue. The levels of cAMP degrading PDEs are increased by glucocorticoids in

adipose tissue; the effects on AC expression have not been studied so far.

Further Reading

- Antoni, F. A. (1997). Calcium regulation of adenylyl cyclase – relevance for endocrine control. *Trends in Endocrinology and Metabolism* 8, 7–13.
- Antoni, F. A. (2000). Molecular diversity of cyclic AMP signaling. *Frontiers in Neuroendocrinology* 21, 103–132.
- Antoni, F. A., Sosunov, A. A., Haunso, A., et al. (2003). Short-term plasticity of cyclic adenosine 3', 5'-monophosphate signaling in anterior pituitary corticotrope cells: the role of adenylyl cyclase isotypes. *Molecular Endocrinology* 17, 692–703.
- Cooper, D. M. F., Karpen, J. W., Fagan, J. W., et al. (1998). Ca^{2+} -sensitive adenylyl cyclases. *Advances in Second Messenger and Phosphoprotein Research* 32, 23–52.
- Gannon, M. N., Akompong, T., Billingsley, M. L., et al. (1994). Adrenalectomy-induced alterations of calmodulin-dependent hippocampal adenylate cyclase activity: role of guanine nucleotide-binding proteins. *Endocrinology* 134, 853–857.
- Jurevicius, J. and Fischmeister, R. (1996). cAMP compartmentation is responsible for a focal activation of cardiac Ca^{2+} channels by β -adrenergic agonists. *Proceedings of the National Academy of Sciences USA* 93, 295–299.
- Mons, N. and Cooper, D. M. F. (1995). Adenylate cyclases – critical foci in neuronal signaling. *Trends in Neuroscience* 18, 536–542.
- Orth, D. N. and Kovacs, W. J. (1998). The adrenal cortex. In: Wilson, J. D., Foster, D. W., Kronenberg, H. M. & Larsen, P. R. (eds.) *Williams Textbook of Endocrinology*, pp. 517–664. Philadelphia: W. B. Saunders.
- Sunahara, R. K., Dessauer, C. W. and Gilman, A. G. (1996). Complexity and diversity of mammalian adenylyl cyclases. *Annual Review of Biochemistry* 36, 461–480.
- Wolfgang, D., Chen, I. and Wand, G. S. (1994). Effects of restraint stress on components of adenylyl cyclase signal transduction in the rat hippocampus. *Neuropsychopharmacology* 11, 187–193.

Adjustment Disorders

M Dascalu and D Svrakic

Washington University School of Medicine, St. Louis,
MO, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1,
pp 27–31, © 2000, Elsevier Inc.

Definition
Epidemiology
Etiology
Diagnosis
Course and Prognosis
Treatment
Conclusions

Glossary

<i>Bereavement</i>	Loss through death.
<i>Bereavement process</i>	The umbrella term for bereavement reactions over time.
<i>Bereavement reactions</i>	Psychological, physiological, or behavioral responses to bereavement.
<i>Crisis</i>	A self-limited response to hazardous events and experienced as a painful state. It can last from a few hours to weeks.
<i>Diagnostic and Statistical Manual of Mental Disorders DSM-IV</i>	The latest and most up-to-date classification of mental disorders. It was published in 1994.
<i>Temperament and character inventory (TCI)</i>	A self-report questionnaire that allows diagnosis of personality disorders using a quantitative rating of seven factors, four dimensions of temperament (harm avoidance, novelty seeking, reward dependence, and persistence) and three dimensions of character (self-directedness, cooperativeness, and self-transcendence).

Definition

Adjustment disorder is characterized in DSM-IV by the development of emotional or behavioral symptoms in response to one or more identifiable psychosocial stressors. The symptoms are considered clinically significant because they either impair the individual's functioning or are subjectively perceived to cause distress in excess of what would be expected

from exposure to the stressor(s). These symptoms neither fulfill criteria for an Axis I diagnosis nor represent the exacerbation of a preexisting Axis I or Axis II disorder. They are not caused by normal bereavement. The symptoms occur within 3 months of the onset of a stressor and in response to it and do not last longer than 6 months after the end of the stressor or its consequences. The nature and severity of the stressors is not specified in DSM-IV. They include interpersonal, occupational, or medical problems. The stress may be single or multiple, such as loss of job associated with divorce and physical illness. The stress may happen once or it may be chronic, lasting more than 6 months (such as living in poverty), or recurrent (such as seasonal business difficulties). It can affect only one person or many (as in a flood or an earthquake). Whatever the nature of the stressor the person is overwhelmed by it. The individuals may experience anxiety, depression or behavioral symptoms (e.g., erratic actions) or various combinations. If the symptoms progress to meet criteria for another Axis I disorder (e.g., major depression), a diagnosis of adjustment disorder can no longer be used.

Historically a diagnosis similar to adjustment disorder first appeared in DSM-II. Transient situational disorder was understood in a developmental context and its subtypes were defined accordingly; for example, "adjustment reaction" of childhood, adolescence, or adulthood. It required an unusually severe stress. The term adjustment disorder first appeared in DSM-III. A link to severe and unusual stress was no longer required, and the subtypes were recategorized according to their symptomatic presentation. In DSM-II and DSM-III the duration of adjustment disorder was not specified. In DSM-III-R, the duration of the illness was restricted to 6 months after the cessation of the stressor, and the subtype adjustment disorder with physical complaints was added.

Epidemiology

The adjustment disorders are considered to be very common. One epidemiological study conducted in children and adolescents found a prevalence rate of 7.6%. Precise data are not available for adults because the structured interviews used in general population do not include adjustment disorder. Several studies of the prevalence of adjustment disorders were done in clinical samples including patients admitted in the hospital for medical or surgical problems. The reported rates were between 5 and 13%

in adults and up to 42% in children and adolescents. In a recent study of 1039 consecutive referrals to consultation–liaison psychiatry services, a diagnosis of adjustment disorder was made in 125 patients (12.0%). It was the sole diagnosis in 81 patients (7.8%) and in 44 (4.2%) it was diagnosed comorbidly with other Axis I and II diagnoses, most frequently with personality disorders and organic mental disorders. It had been considered as a rule-out diagnosis in a further 110 patients (10.6%).

Several studies found that this disorder is more common in women (sex ratio, 2:1).

Etiology

Adjustment disorder is caused by one or more stressors. There is a significant individual variation in response to stress so the severity of the stressors is not always predictive of the development of the illness.

Individual Factors

It is not clear why there is so much individual variation in the development of psychopathology in persons who experience similar stressors and also why when a reaction occurs the symptoms are so variable. Psychoanalytic researchers have highlighted the importance of development of adequate defense mechanisms as a child. Those who were able to develop mature defense mechanisms seem to be less vulnerable to and recover more quickly from the stressor. In this context, the roles of the mother and the rearing environment in a person's capability to respond adequately to stress are critical.

More recently, it has been shown that immature defense mechanisms correlate highly with poor character development as defined by the TCI (Temperament and Character Inventory). Poor character development is considered to be the common feature for the whole group of personality disorders as defined by DSM-IV. This implies that people who have a personality disorder are at a high risk for developing abnormal reaction to stress and then adjustment disorder. Conform to the definition both disorders can be diagnosed in the same time.

Several studies showed that children and adolescents are at a higher risk of developing an adjustment disorder than adults and that the illness is more frequent in females than in males.

Other Factors

The stressor severity is a complex function of multiple factors; for example, degree, quantity, duration, reversibility, environment, and timing. The severity of

the stressors is not always predictive of the occurrence and severity of adjustment disorder, but without a doubt, serious and/or chronic stressors are more likely to cause an adjustment reaction. It is not clear if all persons are likely to develop symptoms if stress levels are increased enough. The nosological and phenomenological distinction between crisis and stress is critical for the understanding of individual variability in response to stress.

Diagnosis

Diagnostic Clues

The chief complaint may be a nervous breakdown, inability to manage problems of life, or anxiety or depression associated with a specific stressor; the patient's history reveals normal functioning before the onset of the stressor; and the patient's mental status examination shows symptoms of anxiety, depression, or disturbed conduct.

Clinical Features

Several studies reported that between 50 and 87% of patients present with depressed mood. Other common symptoms encountered in more than 25% of the patients include insomnia, other vegetative symptoms (e.g., palpitations), behavioral problems, social withdrawal, and suicidal ideation or gesture.

The development of depressive symptoms alone is less characteristic of children and adolescents. Two recent studies found that among children and adolescents with adjustment disorder, the majority presented with mixed emotional or mixed emotional and behavioral syndromes. In one study, 77% of adolescents but only 25% of adults had behavioral symptoms (disturbance of conduct).

Studies of adjustment disorder that used structured diagnostic instruments have reported a high level of comorbidity. In a mixed group of children, adolescents, and adults, approximately 70% of patients with adjustment disorder had at least one additional Axis I diagnosis.

Several studies have reported a significant association of adjustment disorder with suicidal behavior in adolescents and young adults. A couple of studies done in patients hospitalized after a suicidal gesture showed that more than 50% of the patients met criteria for adjustment disorder with depressed mood. Three retrospective studies of suicide completers below age 30 found that between 9 and 19% of the cases met criteria for adjustment disorder. This data suggest the seriousness of the illness in a subset of persons.

Subtypes of Adjustment Disorder

The adjustment disorders are classified according to the predominant clinical symptoms. Six subtypes are identified and coded in DSM-IV. Three define discrete symptomatic presentations (depressed mood, anxious mood, disturbance of conduct) and two describe mixed clinical presentations (mixed disturbance of anxiety and depressed mood, mixed disturbance of emotions and conduct). The final subtype, adjustment disorder unspecified, is a residual category for presentations not accurately described by one of the other subtypes. One additional subtype, adjustment disorder with suicidal behavior, was proposed for inclusion in DSM-IV. It was ultimately rejected because of concerns that it would discourage a more systematic assessment of symptoms in patients presenting with suicidal behavior.

1. **Adjustment Disorder with Depressed Mood:** The predominant manifestations are depressed mood, tearfulness, and hopelessness. This type must be distinguished from major depressive disorder and uncomplicated bereavement.

2. **Adjustment Disorder with Anxiety.** Symptoms of anxiety – such as palpitations, jitteriness, and agitation – are present in adjustment disorder with anxiety, which must be differentiated from anxiety disorders. The patient is usually nervous, fearful, and worried.

3. **Adjustment Disorder with Mixed Anxiety and Depressed Mood:** Patients exhibit features of both anxiety and depression that do not meet the criteria for an already established anxiety disorder or depressive disorder.

4. **Adjustment Disorder with Disturbance of Conduct:** The predominant manifestation involves conduct in which the rights of others are violated or age-appropriate societal norms and rules are disregarded. Examples of behavior in this category are truancy, vandalism, reckless driving, and fighting. The category must be differentiated from conduct disorder and antisocial personality disorder.

5. **Adjustment Disorder with Mixed Disturbance of Emotions and Conduct:** The combination of disturbance of emotions (anxiety and/or depression) and conduct sometimes occurs.

6. **Adjustment Disorder Unspecified:** Adjustment disorder unspecified is a residual category for atypical maladaptive reactions to stress. Examples include inappropriate responses to the diagnosis of physical illness, such as massive denial and severe noncompliance with treatment and social withdrawal without significant depressed or anxious mood.

Adjustment disorders are also classified according to the clinical course. They are acute if the symptoms last less than 6 months and chronic if the symptoms last longer than 6 months.

Differential Diagnosis

Adjustment disorders must be differentiated from a normal reaction to stress or from other psychiatric disorders that occur following a stress.

1. In acute stress disorder and posttraumatic stress disorder, the stress needs to be severe and it is more clearly specified. The stressors are psychologically traumatizing events outside the range of normal human experience so they are expected to produce the syndromes in the average human being. Both acute stress disorder and posttraumatic stress disorder are characterized by a specific constellation of affective and autonomic symptoms, which is not encountered in adjustment disorders.

2. In normal bereavement, despite difficulties in social and occupational functioning, the person's impairment remains within the expectable bounds of a reaction to a loss of a loved one.

3. Other disorders include major depressive disorder, brief psychotic disorder, generalized anxiety disorder, somatization disorder, conduct disorder, drug abuse, academic problem, occupational problem, or identity problem.

4. If the criteria for any of the above disorders are met, that diagnosis should be used instead of adjustment disorder, even if a stressor was present.

Course and Prognosis

Course

Adjustment disorder begins shortly after a significant stressor. By definition, the symptoms have to commence within the first 3 months after the onset of the stressor. Usually they start in the first couple of weeks. In most cases when the stressor ceases the symptoms remit quickly.

Prognosis

By definition, adjustment disorder is not an enduring diagnosis. The symptoms either resolve or progress to a more serious illness. In most cases, the prognosis is favorable with appropriate treatment, and most patients return to their previous level of functioning within 3 months. It appears that children and adolescents have a poor prognosis. Several studies, which surveyed adolescents up to 10 years after a diagnosis of adjustment disorder, found that only 44 to 60% of

the original sample were well at follow-up. Some of the patients developed major mental disorders, such as schizophrenia, bipolar disorder, major depressive disorder, or substance-related disorders. It is not clear whether the poor outcome is due to adjustment disorder itself or to the frequent existence of comorbidity. At least one study that matched two groups for comorbidity (one with adjustment disorder and the other one without) found that adjustment disorder added no additional risk for poor outcome.

Treatment

Little is known about the proper treatment of adjustment disorder. There are few, if any, treatment studies in patients with adjustment disorder, and systematic clinical trials are necessary because adjustment disorder is one of the most common psychiatric diagnoses. The typical pharmacological intervention is always symptomatic. When given, medication is aimed to alleviate a specific symptom, but should always be in addition to psychosocial strategies.

Stressors

Whenever possible the etiologic stressors should be removed or ameliorated. Interventions designed to minimize the impact of the stressors on daily functioning should be considered.

Psychotherapy

This remains the treatment of choice for adjustment disorder. It can help the patient adapt to the stressor if it is not reversible or time limited. It can also serve a preventive role if the stressor does remit.

1. Crisis intervention is a brief type of therapy that may be useful to decrease stress and facilitate the development of external support. It is designed to resolve the situation quickly by supportive techniques, suggestions, reassurance, and environmental manipulation. The frequency and length of encounters with the therapist is tailored to the patients needs, and sometimes even short-term hospitalization might be warranted.

2. Individual psychotherapy offers the patient the opportunity to understand the personal meaning of the stressor and to develop coping skills.

3. Group psychotherapy can be remarkably useful for patients who experienced similar stresses; for example, a group of renal dialysis patients.

4. Family therapy can sometimes help in some patients, especially those with adjustment disorder with disturbance of conduct who may have difficulties with the school, authorities, or the law.

Pharmacotherapy

This may be useful in some patients when prescribed for brief periods. Depending of the subtype of adjustment disorder, an antianxiety agent or an antidepressant may help. Some patients with severe symptoms may also benefit from short course psychostimulant or antipsychotic medication. There are no studies regarding the use of specific medical agents in adjustment disorder. Few, if any, patients can be adequately treated by medication alone, and psychotherapy should be added to the treatment.

Conclusions

Adjustment disorders are an important and prevalent cause of personal discomfort, absenteeism, addiction, and suicide. This diagnosis should not be used to avoid stigmatization of a patient. Short psychotherapy is the treatment of choice, but pharmacotherapy is often used as an acute way to stabilize the patient and to prevent potentially disruptive behavior.

See Also the Following Articles

Adolescence; Anxiety; Bereavement; Childhood Stress.

Further Reading

- Cloninger, C. R. and Svrakic, D. M. (2000). Personality disorders. In: Kaplan, H. I. & Sadock, B. J. (eds.) *Comprehensive Textbook of Psychiatry* (7th ed.). Baltimore: Lippincott Williams & Wilkins.
- Kaplan, H. I., Sadock, B. J. and Grebb, J. A. (1994). *Synopsis of Psychiatry* (7th ed.). Baltimore: Williams & Wilkins.
- Morrison, J. (1995). *DSM-IV Made Easy*. New York: Guilford.
- Newcorn, J. H. and Strain, J. (1995). Adjustment Disorders. In: Kaplan, H. I. & Sadock, B. J. (eds.) *Comprehensive Textbook of Psychiatry*. Baltimore: Williams & Wilkins.
- Popkin, M. K., Callies, A. L., Colon, E. A. and Stiebel, V. (1990). Adjustment disorders in medically ill inpatients referred for consultation in a university hospital. *Psychosomatics* 31(4), 410–414.
- Snyder, S., Strain, J. J. and Wolf, D. (1990). Differentiating major depression from adjustment disorder with depressed mood in the medical setting. *General Hospital Psychiatry* 12(3), 159–165.
- Strain, J. J., Smith, G. C., Hammer, J. S., et al. (1998). Adjustment disorder: A multisite study of its utilization and interventions in the consultation-liaison psychiatry setting. *General Hospital Psychiatry* 20(3), 139–149.

Adolescence

G N Swanson

Allegheny General Hospital, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G N Swanson, volume 1, pp 32–41, © 2000, Elsevier Inc.

Introduction

Normal Adolescence

Common Stressors

Unusual Stressors

Glossary

<i>Adolescence</i>	A stage of life from puberty to adulthood usually thought of as occurring between the ages of 12 and 19 years. It is characterized by marked physical, psychological, and social change. The developmental task of adolescence is the change from dependence to independence.
<i>Identity</i>	A central aspect of the healthy personality, consisting of an inner awareness of continuity of self and an ability to identify with others, share in their goals, and participate in society.
<i>Puberty</i>	A normal growth process that begins in early adolescence, lasts 2–4 years, and leads to sexual and physical maturity.
<i>Resiliency</i>	The ability to overcome or adapt to stress by maintaining developmental progress and adequate social and academic functioning.

Introduction

Adolescence has been characterized as a challenging stage of life, defined by the psychological task of identity formation. Adolescents develop better coping skills as they mature both as a consequence of their cognitive and emotional development and because of the changes they experience. Teenagers generally respond successfully to many common stressors, such as puberty, school demands, family changes, and peer relations. Unusual stressors may lead to some difficulties, especially in those who are more vulnerable.

Traumatic stressors, although not universally harmful, often impinge on the emotional wellbeing of teenagers and may leave lasting scars. An adolescent's ability to cope with stressors depends on a

number of factors, including family support, cognitive development, previous experiences, and their own unique temperament.

Normal Adolescence

The discrete stage of adolescence has been recognized for only a short period of time. It seems to have come about in Western society as a consequence of the Industrial Revolution, when social changes necessitated the prolongation of childhood. Children, who had previously been able to enter the adult world at an early age, did not have the skills or abilities to perform adult work. Child labor laws were passed, and public schools were established. Families were also better able to financially provide for children for a longer period of time. The transition from childhood to adulthood lengthened, encompassing most of the second decade of life. Indeed, many teenagers nowadays maintain their dependency on their parents during college, thereby extending this transition even further. Puberty is often identified as the starting point of adolescence, although this can normally start as early as 9 or 10 years of age. The endpoint has been less well defined, as there are differences in the legal age of adulthood (18–21 years), and the criteria of independence, the completion of school, starting a new family, and getting a job are even more variable.

Adolescence is often thought of as a period of rebellion, marked by conflicts with parents as a child grows to be an adult. Anna Freud, G. Stanley Hall, and others characterized this stage of life as a time of storm and stress. Psychoanalysts believed that it was an indication of pathology if these stormy relations did not occur. Erik Erikson believed that the psychological task of this stage of life is the development of identity. Specifically, he thought that the adolescent would commit to a core set of values and assume a sex role and career plan. Failure to complete this task would result in identity diffusion, leading to further difficulties in adulthood. Adolescence was the period of an identity crisis, which led to considerable emotional turmoil.

However, Offer and associates have shown that the vast majority (~80%) of adolescents do not experience significant emotional distress. Instead, these teenagers manage the transition to adulthood smoothly, without experiencing a severe identity crisis. They have a positive self-image and are not afraid of the physical changes associated with puberty. They do not

report significant conflicts with their parents and have positive feelings toward their families. They are confident, optimistic about the future, and are willing to work hard in order to reach a goal.

However, a significant minority (~20%) of teenagers reports difficulties in these areas. They report that they feel emotionally empty and overwhelmed by life's problems. They feel much less able to control the world around them. These adolescents also report that they are much less able to talk with their parents about their problems. They appear to be more vulnerable to stressors and may exhibit more behavioral and emotional problems as a result. It seems that the stereotype of the moody, confused, and rebellious teenager is more applicable to this smaller group than to most adolescents.

Considerable cognitive development occurs during adolescence. Early adolescents become very self-conscious, which leads to a tendency to be more egocentric. They tend to believe that the thoughts and feelings that they have are unique. Furthermore, they do not seem to recognize that their peers are going through the same situation. In time, older teenagers become better able to empathize, to think abstractly, and to be introspective. They are able to consider possibilities and to imagine other realities. This leads them to become more philosophical, and to question the rules and realities of their lives in an attempt to gain a clearer understanding of the world. These cognitive abilities lead them to question their identity and to see themselves as individuals with some control over who and what they will be.

These changes become manifest in the adolescents' social environment. In particular, the family becomes less influential, while peer groups become more so. Adolescents begin to establish and exercise independence from parental controls. They learn to negotiate, although some conflicts arise. They spend more time with their friends than they did in childhood. They often imitate influential peers and adults. They try on different hats to see what they might look like, in order to establish an identity of their own. However, although they may dress like their friends or listen to music that their parents abhor, they often retain similar values and beliefs as their parents.

Adolescents tend to get better at coping with stress as they get older. They learn skills over time, utilizing their individual temperament and modeling parental coping strategies. Early in adolescence, teenagers tend to react defensively, denying problems, and thinking wishfully. They misinterpret situations, making mis-attributions to others, and overlooking information that would help them to see things more clearly. They also tend to either under- or overestimate both potential risks and their ability to handle them.

They may vent their frustration or anxiety in open emotional outbursts or they may withdraw and try to keep their feelings hidden. They seek solace and support from others, but do not readily ask for or listen to advice. As they develop, they become better able to appraise problems themselves. They try to change the stressful situation and to negotiate solutions. They learn that communication skills, the art of compromise, and assertiveness are more effective responses than are emotional outbursts or withdrawal. They learn how much of their feelings they need to reveal and become more adept at managing their emotions. They also become better at reading social cues and reacting appropriately. They think about ways to deal with problems and discuss potential solutions with others. Successful responses to stress lead to more success, although at times they may revert to previous, less effective strategies, usually in an impulsive fashion. Over time, most adolescents develop the skills necessary to cope with stress effectively, although some do not.

Common Stressors

Adolescents face a variety of challenges as a normal part of their lives. These include the physical and sexual changes associated with puberty; the demands of school; the desire to initiate and maintain friendships, both platonic and romantic; the need to start working and to make a career choice; and the gradual development of independence from the family. All of these changes can be stressful, but most teenagers report that they do not feel overwhelmed or unable to deal with them. Regardless of gender, socioeconomic status, or race, teenagers identify their most important concerns as career, school performance, and college plans. Many worry about violence, theft, and work. Some report concerns about peer conflicts and parental expectations. Teens are least worried about drugs and alcohol. Boys are more concerned about sexuality and extracurricular activities, while girls are more concerned about their appearance.

Puberty

Puberty, as the starting point for adolescence, is often thought of as very stressful. The physical changes are obvious, as is an increased sexual interest. Parents and teachers often attribute all behavioral changes to raging hormones, but the fact is that there are other social, and psychological factors that must be considered as well. Nevertheless, there is a biological effect of puberty on cognitive, social, and emotional development as well. In particular, the timing of puberty can be stressful and does seem to have an impact on academic and psychosocial functioning.

Boys appear to benefit academically and socially if they reach puberty early, although they often end up having a shorter stature than their later-developing peers. Early in adolescence, girls seem to have a better body image and are more popular if they develop early. However, as a group, they do not seem to do as well academically. Furthermore, by the time adolescence ends, later developing girls have a better body image than girls who reached puberty earlier do.

This may be related to the fact that girls with a later onset of puberty are more slender, which more closely fits with current cultural standards of attractiveness. More importantly, girls who reach puberty later have had a longer time to anticipate and adapt to these changes and as a result seem better able to cognitively process and cope with pubertal changes when they do occur. Finally, puberty may be most stressful when it is very early or very late, as this leads the adolescent to be markedly different from other peers. This information coincides with Offer's reports that most adolescents are comfortable with, and not distressed by, the changes of puberty, as most adolescents will enter puberty at about the same time as their peers and experience it as a normative process.

Peer Relationships

Peer relationships take on much more importance during adolescence. Teenagers expand their social relationships, looking to increase their connections with others. Girls in general seem to develop a capacity for intimacy sooner than boys do. Most teenagers believe they can make friends easily and that they can tell their friends intimate details about themselves. Early adolescents are very interested in being popular and want to fit in with a popular same-sex peer group. There is a degree of stress associated with this, especially if a young teenager wants to belong to a group, yet does not. Most are able to find a group, however, and do not find establishing friendships difficult. These peer groups are often very supportive and discourage deviant behavior.

In middle adolescence, boys and girls begin to mix. Romance and dating become extremely important. Once again, teenagers report some stress during this phase, as they grapple with the new social skills required to establish romantic relationships. However, most report that they believe they are interesting to members of the opposite sex and believe that they are capable of finding a romantic partner.

Dating usually begins between the ages of 12 and 16 years. Adolescents may face some peer pressure to date and may be dropped from a particular group if they do not do so. Similarly, they may feel pressure to engage in a similar level of sexual activity as their

peers. The pressure to conform to a perceived peer norm, which is in conflict with cultural, religious, family, or individual values, seems to be the most stressful aspect of romantic relationships. There are many similarities between adult and adolescent love relationships.

However, adolescent relationships are notably more transient, with strong feelings that do not usually lead to enduring intimacy or self-disclosure. Cognitive development as adolescence progresses helps teenagers to better cope with relationship changes. Gay adolescents are especially vulnerable to stress, due to feelings of isolation from peers and family and to difficulty integrating homosexuality into their identity.

Parents of teenagers frequently describe family life as stressful, due to the ongoing development of adolescent independence. This reflects a change in family roles, which parents may perceive as more difficult than adolescents do. Early adolescence is marked by considerable variation in roles, as the adolescent vacillates between asserting independence and maintaining dependency. There is some anxiety associated with becoming independent, felt by both parent and child. As they become older, most teenagers are confident about their ability to make decisions and try to demonstrate this to their parents. However, although they are in the process of individuating, they neither sever their emotional attachments to their parents nor become free of their parents' influences. Adolescents appear to be less distressed when parents utilize an authoritative parenting style which encourages and supports independence. Parents provide limits and controls with explanations, while allowing the adolescent to express their views. Nevertheless, tensions are still present as negotiations proceed and opposing needs are balanced.

Race and Culture

Racial and cultural issues pose interesting challenges, particularly for the adolescent. A child forms both gender and ethnic identities around the ages of 3–5 years. Children become familiar with cultural differences and history long before puberty begins. However, it is in adolescence that teenagers make a conscious commitment to be a member of their culture. They will generally embrace the values of their culture, which may not be the same as the dominant culture. This may lead to stress, especially if they have frequent interactions with peers of other cultures. They also can understand the abstract concepts of racism and inequality. They may be more likely to experience these as they leave their family and have more contact with peers and adults from other walks of life. Nonwhite adolescents are twice as

likely as their white peers to be funneled into the juvenile justice system rather than the mental health system when they have a problem with the law. They are also more likely to identify their main stressors as environmental ones (such as living in a dangerous neighborhood) rather than more personal ones. Most teenagers come to terms with their ethnic identity and with negative stereotypes and prejudices. They are able to accept themselves and their place within both society and their own ethnic culture. Those that do not do so have more difficulty with self-image and psychological adjustment. Adolescents who immigrate to a new country face additional stress, as they become more dependent on their families at a time when they are trying to develop independence. If they do push to be more independent, they may be more likely to join an inappropriate peer group.

Academics and School

School situations and academic demands are yet other common stressors that teenagers confront. The transition to middle school and the transition to high school are both major life events. Most teenagers report a combination of eagerness and apprehension as they move up to a bigger school with more challenging assignments, more complicated schedules, and more competent upper classmen. Middle schools represent a challenge as they are more impersonal and may require independence than early adolescents are capable of. Most middle-school students have not mastered the social skills needed for this setting, although they may learn them quickly if provided with opportunities to succeed in small groups.

High schools offer more extracurricular activities, which provide opportunities to join in new peer groups, but the many choices may be overwhelming and confusing, and some adolescents may end up feeling excluded. Grades become more meaningful, especially for those planning on college. Parental expectations, Scholastic Amplitude Test (SAT) scores, and college applications create tension, as does the search for a job after graduation.

Work

Most teenagers express a desire to work and feel satisfaction in a job well done. Earning money helps older adolescents become more independent and enables them to practice some of the skills they will need as adults. However, trying to balance a part-time job with school demands, sports, and social activities is often difficult. Adolescents who work are less invested in school, spending less time on homework and missing classes more often than peers who do not

work. They usually work in low-paying jobs and have little authority or opportunity to advance. Those who work the most hours tend to have the lowest grades. However, part-time work has also been associated with better self-esteem and a sense of responsibility in adolescents. Teenagers cope with the stress of work best if they can balance the time they work with their other priorities. Working raises other issues, as adolescents may be overwhelmed with the choices they have.

Establishing a life goal is a daunting task for adolescents. Many are reluctant to commit to a specific career path, as this means eliminating other possibilities. This lack of commitment may be interpreted as a lack of motivation by parents, leading to conflicts. Adolescents also have many misconceptions and cognitive distortions about themselves and the careers they are considering. Accurate information and a frank discussion of their strengths and weaknesses are important. The support of a mentoring adult is often very helpful, as is providing the perspective that any choice is not etched in stone. This issue is also one that tends to extend adolescence most often, as it is commonly the last one resolved.

Unusual Stressors

Some teenagers face more unusual challenges. These can include family problems, such as mental or physical illness; drug or alcohol abuse; parental separation or divorce; social problems, such as poverty and violence; and individual problems, such as pregnancy, serious illness, and school failure. In some instances, teens may experience traumas such as abuse or the death of a loved one. Adolescents with adequate cognitive abilities, emotional development, and supportive families seem better able to cope with these problems successfully. Teenagers who have experienced multiple stressors, have a previous history of psychopathology, or have little parental support are much more vulnerable.

Parent with Medical/Psychiatric Illness

Adolescents who are raised in homes where a parent has a serious medical or psychiatric illness have the ability to understand the illness better than younger children and may have more questions as a result. They are also more likely to exhibit anger and acting out behaviors. This may occur because of the conflict involved between the family's needs for help from the adolescent and the adolescent's needs to become more independent. Younger children may be more likely to avoid discussion of the parent's illness and to experience intrusive thoughts and feelings, while adolescents are more likely to exhibit symptoms of anxiety and depression. This is particularly true of teenage

girls whose mothers have cancer. Parental coping skills play a part in how adolescents respond to parental illness. Parents who are less anxious and depressed have a positive effect on their children. There are some important differences to be considered when a family member has a severe mental illness as opposed to a physical illness. Parents with mental illness are much less likely to be in treatment, which means that professional education and advice are much less forthcoming. There is considerable stigma associated with mental illnesses (although some physical illnesses carry a stigma, such as HIV infection). Adolescents may therefore be much less likely to seek peer or adult support as a result. Teenagers are at a higher risk for depression and other mood disorders when their parents have a chronic mental illness, but studies have not been able to separate genetic influences from psychosocial effects on children and adolescents. Finally, parental conflicts, divorce, and inconsistent parenting practices often are present in families where a parent has a severe mental illness. These multiple risk factors compound the stress on an adolescent, more so than in families where a parent has a severe physical illness. Adolescents who cope best when parents have a serious medical or psychiatric illness share several characteristics. They tend to be actively involved in school, church, work, and other outside activities. They have a close relationship with a supportive adult. They also understand that they are not responsible for their parents' illness. Their families are more cohesive, flexible, and are able to maintain family rituals. Finally, their families effectively communicate information and feelings about the parental illness and are able to make plans for the future.

Parental Alcohol Abuse

The effects of parental alcohol abuse have also been studied, although more attention has been paid to younger children in alcoholic families. Adolescents raised in alcoholic families have been found to have a higher risk of conduct disorder, substance abuse, sexual acting-out, physical and sexual abuse, and academic problems. They also have more difficulties in their peer relationships and are more likely to have romantic relationships with adolescents with substance abuse problems themselves. Some adolescents may take on a more adult role within the family in an attempt to maintain family functioning, while others may disengage. Boys seem to have a higher risk of problems than girls do. Teenagers who cope effectively with parental alcoholism tend to have healthy and supportive relationships with adults outside of the family. Families that are able to maintain family rituals (e.g., such as holiday celebrations,

family dinners, and church attendance) have teenagers that are more resilient in the face of parental substance abuse.

Adolescents with an easy temperament, at least average intelligence, and an internal locus of control are also more resilient. Not all teenagers from alcoholic homes require treatment, although parents with alcoholism should be referred for treatment. However, it is not clear what effect, if any, parental recovery has on adolescents who are already exhibiting problems. Alateen, a self-help program for adolescents from alcoholic families in the USA and Canada, has been effective in providing information and improving mood and self-esteem for teenagers. Group, individual, and family therapy have been helpful. All interventions should provide education regarding the adolescent's increased risk of substance abuse in an attempt at prevention.

Parental Marital Conflict and Divorce

Parental conflict and divorce is a fairly common, but nonetheless major, stressor in today's culture. Most research on separation and divorce has focused on younger children. Studies with adolescents suggest that they too may have difficulty adjusting to this stressor. In addition, children who experience divorce may not manifest problems until adolescence. It is important to remember that most children and adolescents do adapt to divorce and are able to function effectively as adults. Younger adolescents whose parents divorce are more likely to be noncompliant and aggressive and to have problems with substance abuse than are adolescents in nondivorced families. Gender differences have been reported. Adolescent girls from divorced families have had more problems with self-esteem and promiscuity, while adolescent boys have a higher incidence of substance abuse. Both boys and girls have lower academic achievement, but teenage boys are also likely to drop out of school if their parents have divorced and they are living with their mother. Girls are at a greater risk of dropping out if they live with their remarried mother and stepfather. These findings appear to hold even when factors such as race and socioeconomic status are controlled for. Conflict between divorced mothers and their adolescent daughters is common, possibly connected to the teenage girls' tendency to increased sexual acting out. Adolescent boys have a higher risk of disengaging from their family and to engage in delinquent behavior with peers. This may be related to the absence and lack of influence of the noncustodial father, which is all too common in divorce. Many factors mediate adolescents' responses to divorce. An authoritative custodial parent seems to be the most important factor, as the parent is able to provide

the understanding and support needed while maintaining an authority position in the family. Generally, teenagers are better able to cognitively process and understand the reasons for parental divorce than younger children are. However, those who have limited insight, poor problem solving skills, and a history of temperamental difficulty are more vulnerable to experiencing problems with the divorce. The degree of conflict between parents and the amount of contact with each parent are also important. Depression and delinquency have been connected to prolonged parental conflict.

Poverty and Violence

Sociocultural stressors such as poverty and violence have long been recognized as stressful for children and adolescents. However, it is very difficult to disentangle the specific effects these stressors have from each other as well as from other associated risk factors, such as parenting styles and other environmental stressors. Approximately 20% of American children live in poverty – the highest rate for any Western industrialized country. Black children are twice as likely to experience poverty as white children are. Length of time spent in poverty varies, although around 90% of poor children spend less than 5 years in poverty. Contrary to popular belief, poverty is more common in rural rather than urban areas.

Adolescents who experience poverty have a higher rate of delinquency, depression, and poor self-image. They may also incorporate the idea of poverty into their identity rather than to see it as a temporary external condition. Again, parental responses and styles have a marked effect on adolescent coping strategies. Families living in poverty are more likely to utilize inconsistent, punitive, and authoritarian parenting styles, which leads to increased stress and conflict. In contrast, poor but supportive parents, who have a positive outlook on the future, have less distressed teenagers. Teenagers are also able to recognize when they live in less desirable neighborhoods and are well aware of the risks of violence. Living with this chronic stress is difficult and can have a marked effect on their outlook on the future. They may be more fatalistic and experience more posttraumatic stress disorder (PTSD) symptoms as a result of their exposure to violence. Anecdotes abound, both describing those who have problems and those who have been resilient. However, at this time there are no systematic studies that have assessed these issues.

Pregnancy

Some teenagers face significant individual stressors, such as pregnancy, serious illness, or school failure.

These teenagers already tend to be at a higher risk for other stressors, as outlined above. However, considerable study has been given to these issues. Teenage girls who get pregnant are less prepared than their adult counterparts to raise children. They know less about infants and are more distressed by the pregnancy. Around 40% choose abortion, while 45% choose to keep their child. The other 15% either miscarry or choose adoption. Teenage mothers are less responsive to the needs of their babies than are adults. However, longitudinal studies have shown that a majority of teenage mothers complete high school and hold regular employment thereafter. Most support themselves and their children, although they are on welfare at times. They do not end up having more children than peers who have children later. Most importantly, the majority seems to cope effectively over the long run.

In contrast, teenage girls who ultimately choose abortion report considerable psychological distress during the time of pregnancy. This seems to fade somewhat after the abortion. However, studies of women who have undergone abortion indicate that most negative reactions and distress occur in those who are young, unmarried, previously nulliparous, and who delay the procedure until the second trimester. Adolescents are much more likely to fit into this profile and so appear to be at a higher risk for psychological sequelae as a result of the abortion. The debate about the psychological effects of abortion is unresolved, as this particular group remains difficult to study due to the many other stressors that they face.

Serious Illness

It is estimated that 5–10% of teenagers face a serious illness during adolescence. Adolescents who suffer from serious illnesses struggle with a variety of issues. Early adolescents tend to try to deny the existence of their illness, using avoidant coping strategies. These problems are more likely to occur in families where there is little cohesiveness. This often leads to treatment noncompliance and more problems with the illness itself. Conversely, chronic or serious illnesses may impair normal adolescent development.

This seems to be due to the fact that the illness and treatment requirements make the adolescent more dependent on his parents. They in turn may be more unwilling to let the teenager be more independent. Adolescents may also incorporate their illness into their identity. In some situations, such as diabetes, these may be unavoidable due to the nature of the illness. For those who have cancer, however, this may be more problematic. Adolescents with a history of cancer are not more likely to be depressed, but they

are more likely to have somatic complaints and preoccupations and to be distrusting of their bodies. They also tend to have greater difficulties in romantic relationships. Misattributions and misunderstandings about their illness and its prognosis should be addressed when present. Parents and adolescents often need much more education and support than they receive. Efforts should be made to help adolescents understand both their illness and treatment, and information should be made readily available to them. This should occur in a stepwise fashion, allowing teenagers some time to adjust to changes and react to them emotionally and intellectually.

Dropping Out of School

Approximately 10–15% of adolescents drop out of school. Often, they have parents or siblings who have done the same. Parental attitudes about education have a great impact on academic performance, even in adolescence. Adolescents whose home lives are already distressed or whose fathers are absent are much more likely to drop out. Those who have already failed a class or who have been held back a grade are at greater risk as well. Many are working, and many have children. Teenagers find dropping out is very stressful and would recommend against it. Job opportunities are limited and they are often unable to function independently. Successful prevention requires strong cooperation between parents and schools, providing the support and guidance necessary to vulnerable teenagers.

Peer Victimization

Several forms of adolescent victimization have been recognized, including bullying, sexual harassment, and interpersonal violence and emotional abuse in dating relationships. Research is limited, but most adolescents have had some limited experiences of victimization. There may be significant differences based on gender, socioeconomic status, and racial and ethnic groups. Those teenagers who complain of significant levels of psychological distress tend to have been targeted on multiple occasions, and in several forms of victimization. They also feel less of a sense of school belonging. In addition, adolescents who have experienced bullying are more likely to have self-blaming attributions than those who have not. Clinicians and school personnel should be carefully assessing for the occurrence of victimization in multiple areas if a teenager reports experiencing any one form of victimization.

Trauma Trauma often leads to emotional and behavioral changes. Most studies on trauma have

looked at younger children, children and adolescents together, or at adolescents who were traumatized as children. However, adolescents can be abused physically, sexually, emotionally, or in combination. Due in part to their increasing independence and to their cognitive and emotional development, adolescents are less likely than are children to be abused by family members. However, they are more likely to be the victims of rape, assault, or robbery than younger children or adults. There are some gender differences in adolescents, as girls are more likely to experience sexual trauma, while boys are more likely to witness the injury or death of another. Catastrophic life events also may occur, including the death of a parent, sibling, or peer due to illness, accident, violence, disaster, or terrorism. Although age-based differences in stress reaction have received research attention, findings have been inconsistent, due in large part to the lack of normative and pretrauma psychological functioning data.

Teenagers who have been abused may come to attention for incidents that have just recently occurred and been disclosed or may have occurred many years earlier. The length of time since the occurrence of abuse or trauma does not mitigate the severity of symptoms or the need for treatment. It is also not clear if abuse or trauma at an earlier or later age leads to more problems. Adolescents are less likely to be abused than are young children. However, many adolescent victims of abuse have been abused as children. Therefore, teenagers who present with physical or sexual abuse may have a lengthy history of abuse and may have more difficulty as a result.

Multiple neurotransmitter systems are involved in the response to traumatic stress, and chronic stress has been associated with long-term changes in neuronal function in structure. A comprehensive study by De Bellis et al. demonstrated alterations in biological stress systems and adverse influences on brain development in maltreated children and adolescents with PTSD. Increased levels of catecholaminergic neurotransmitters and steroid hormones during traumatic experiences could negatively affect brain development. Causal relationships have not yet been established, however, and it is unclear how traumatic stress specifically affects the still developing adolescent brain.

Abuse Teenagers who have been sexually abused are less likely to show unusual sexual behaviors or preoccupations than are young children. Abused teens also are more likely than children to disclose purposefully, usually out of anger toward the perpetrator. Adolescents are also more likely to have their reports substantiated. However, because they are older, victimized adolescents may be blamed, just as

women who are victims of domestic violence are often blamed for not seeking help or leaving an abusive situation. In addition, adolescents who are physically abused may be seen as provoking a parent, thereby earning physical discipline. This is particularly true in teenagers who have a history of physical aggression, threats, delinquent behavior, noncompliance, and/or defiance. The juvenile justice system is more likely to be involved in these situations and may be less cognizant of the possibility of abuse than either mental health or child protective services.

When evaluating adolescents who have been abused, it is important to be sensitive and to allow them some control over what, when, and how much they disclose. Teenagers will often want to avoid talking about the issue and to minimize the effects it has had on them. Recognition and support for their attempts to cope with the abuse should be provided, in keeping with their psychological development, especially their need to be independent and successful. In addition, it may be helpful to separate the problem from their identity as a person. Hecht et al. suggest a statement such as, "I've talked with a lot of people your age who have been through sexual abuse. We talk about the effects it has had on them. Everyone's different, but I want to find out if some of the things that have bothered other young people have bothered you. I'd also like to find out what sorts of things you have done that seem to help the most." It is also important to understand the adolescent's attributions about the abuse. As adolescents are more independent and society holds them more responsible for their actions, they may be more likely to blame themselves.

Furthermore, teens may want to maintain the idea that they have control over what happens to them, in keeping with their developmental stage. At the same time, adolescents will often fear disclosure to their peers, perhaps believing that they will be stigmatized or that their peers will hold them responsible.

Nevertheless, having the support and understanding of a peer is often very helpful. In assessing an adolescent who has been abused, it is important to obtain a complete history as well as to address possible PTSD symptoms. In addition, any problems with aggression, impulsivity, social skills, attention span, academic performance, depression, anxiety, substance abuse, and delinquency should be explored. Abused adolescents are also at a higher risk for eating disorders, sleep problems, and self-injurious behavior.

Treatment interventions should be individualized, but may include sexual education, information about PTSD symptoms and abuse, and, perhaps most importantly, an attempt to work out with the adolescent an explanation as to why the abuse happened. Group

and individual therapies have both been utilized with success. Treatment interventions should address both the needs of parents and the adolescent. Social skills, problem-solving, and cognitive treatments have been effective with adolescents, but family interventions in particular seem to help maintain progress over the long term. Court-mandated treatment helps to keep abusive parents in treatment.

Death Adolescents rarely experience the death of a parent, sibling, or of a close friend. Furthermore, studies of children who have lost a parent usually group younger children with adolescents. As a result, unique characteristics of adolescent bereavement are difficult to identify. Adolescents are able to conceptualize death abstractly. This allows them to consider religious and philosophical issues and may lead to more uncertainty than in younger children. Adolescents who have never experienced a death have a difficult time adjusting to this change. Adolescents who have experienced the death of a parent are at a higher risk for delinquency, anxiety and depressive symptoms, somatic complaints, and PTSD symptoms. Adolescents who have experienced a disaster are more likely to experience PTSD symptoms if they lost a loved one (parent, friend, or classmate) in the event. However, displacement from home, community, and school also contributes to these problems.

Teenage boys who lose a father are at a higher risk for behavioral and emotional problems as are early adolescents. Sudden deaths are clearly more difficult to cope with than deaths that occur after a protracted illness, as adolescents have a chance to prepare emotionally and intellectually for the event. However, a strong and supportive surviving parent can offer some protection. Once again, teens that can confide in another adult (such as a relative, teacher, or neighbor) seem to cope better, as they are able to talk about their dead parent openly. It is important to remember that subsequent events, such as graduation, making a sports team, dating, or getting a job, may reopen grief feelings and should be considered as additional stressors. Nevertheless, the majority of adolescents appear to cope effectively with the death of a parent. The death of a peer is also a rare event. Studies in this area have been limited and have primarily focused on peers who commit suicide. In these instances, adolescents have exhibited depressive symptoms, but have not had an increase in suicide attempts.

See Also the Following Articles

Adolescent suicide; Alcohol and Stress: Social and Psychological Aspects; Childhood Stress; Divorce,

Children of; Economic Factors and Stress; Familial Patterns of Stress; School Stress and School Refusal Behavior.

Further Reading

- Cobb, N. (1995). *Adolescence – Continuity, Conformity and Change* (2nd edn.). Mountain View, CA: Mayfield.
- Cohen, J., Mannarino, A. and Deblinger, E. (2006). *Treating traumatic stress and grief in children: a clinician's guide*. New York: Guilford Press.
- De Bellis, M., Baum, A. S., Birmaher, B., et al. (1999). Developmental traumatology. Part I Biological stress systems and Part II brain development. *Biological Psychiatry* 45, 1259–1284.
- Eth, S. and Pynoos, R. (1985). Developmental perspectives on psychic trauma in childhood. In: Figley, C. (ed.) *Trauma and its wake*. New York: Brunner/Mazel.
- Haggerty, R., Sherrod, L. R., Garmezy, N., et al. (eds.) (1996). *Stress, risk and resilience in children and adolescents*. Cambridge: Cambridge University Press.
- Hecht, D., Chaffin, M., Bonner, B. L., et al. (2002). Treating sexually abused adolescents. In: Myers, J., Berliner, L., Briere, J., et al. (eds.) *The APSAC handbook on child maltreatment*. Thousand Oaks, CA: Sage.
- Hersen, M., Thomas, J. and Ammerman, R. (2006). *Comprehensive handbook of personality and psychopathology – child psychopathology* (Vol. 3). New York: Wiley.
- Kendall-Tackett, K. and Giacconi, S. (2005). *Child victimization*. Kingston, NJ: Civic Research Institute.
- LaGreca, A., Silverman, W. K., Venberg, E. M., et al. (2002). *Helping children cope with disasters and terrorism*. Washington DC: American Psychological Association.
- Lewis, M. (2002). *Child and adolescent psychiatry – a comprehensive textbook*. Baltimore, MD: Williams & Wilkins.
- Offer, D., Ostrov, E., Howard, K. and Atkinson, R. (1990). Normality and adolescence. *Psychiatric Clinics of North America* 13, 377–388.
- Saigh, P. and Bremner, J. D. (1999). *Posttraumatic stress disorder: a comprehensive textbook*. Boston, MA: Allyn and Bacon.
- Youngblade, L. and Belsky, J. (1990). Social and emotional consequences of child maltreatment. In: Ammerman, R. & Hersen, M. (eds.) *Children at risk*. New York: Plenum.

Adolescent Suicide

M Berk, R Suddath and M Devich-Navarro
University of California, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Risk Factors for Suicide in Adolescents
Neuroscientific Findings in Adolescent Suicide
Treatments and Possible Effects of Antidepressants

Glossary

Serotonin (or 5-hydroxytryptamine; 5-HT) An indole amine that is a key chemical neurotransmitter in the nervous system. Disordered serotonin transmission is implicated in depression, suicide, schizophrenia, anxiety and other mental disorders.

Serotonin receptors Receptors are docking sites for bioactive molecules such as serotonin. There are more than 15 serotonin receptor subtypes all located in the cell membrane. With the exception of the 5-HT₃ receptor, a ligand gated ion channel, all other 5-HT receptors are G protein coupled seven transmembrane (or *heptahelical*) receptors that activate an intracellular second messenger cascade. The different

subtypes subserve different functions and are targets for different drugs. Thus, for example, the 5-HT_{2A} receptor is the docking target for the hallucinogen, Lysergic acid diethylamide (LSD). The serotonin transporter (SERT) is a 12-transmembrane protein that mediates serotonin uptake (back into the neuron) and thus reduces extracellular serotonin levels in the brain and especially at synapses. SERT is a major target for therapeutic intervention, and the selective serotonin reuptake inhibitors (SSRIs) are the most frequently prescribed antidepressants worldwide. Increased synaptic serotonin concentrations generated by SSRIs are thought to alleviate depression.

Serotonin transporter

Adolescent suicide is a serious public health problem. According to the most recent statistics, suicide is the third leading cause of death among 10- to 24-year-olds in the United States, accounting for 11.7% of all deaths in this age group. Nonfatal suicide attempts are also a significant concern in their own right. In 2002, approximately 124 409 visits to U.S. emergency departments were made after attempted suicide or

other self-harm incidents among people ages 10–24 years old. According to data obtained as part of the Youth Risk Behavior Survey, a national survey administered to high school-age youth in 2003, approximately 17% reported having seriously considered attempting suicide, 16.5% reported having made a plan for suicide, 8.5% reported having attempted suicide, and 2.9% reported having made a suicide attempt that required treatment by medical professionals. Suicide attempts can lead to injury, place youth at increased risk for death by suicide and future suicide attempts, place burden on the health-care system, and are associated with a range of adverse outcomes including substance abuse, high-risk sexual acts, school drop-out, poor academic performance, and delinquency.

Risk Factors for Suicide in Adolescents

Prior suicide attempts are the biggest risk factor for a subsequent suicide attempt or completed suicide. Psychopathology is a risk factor for suicide attempts. Approximately 90% of suicide victims have a psychiatric illness at the time of their death. Significant increases in the risk of suicidal behavior are associated not only with depression but also with psychosis, substance abuse, disorders that involve impaired impulse control, and a variety of other psychiatric disorders.

Epidemiological studies of adolescent suicide have demonstrated an increased risk of suicidal behavior with age and significant gender effects that depend entirely on the distinction between completed suicide and suicide attempts. Suicide is relatively uncommon before age 12, and incidence rates rise in the late teens and early twenties. Suicidal ideation and suicide attempts are more common among females than males. In contrast, completed suicide is three to five times more common among males than females. The higher rate of completed suicide in males has been attributed to a variety of factors, including the use of more lethal methods (such as firearms) and higher rates of aggression and substance abuse. The most common way that females attempt suicide is to overdose. In some countries where more lethal drugs are accessible or where emergency medical care is less effective, the rate of completed suicide in females exceeds that of males. Race has been demonstrated as a risk factor for completed suicide. Completed suicide is more common in Whites than African Americans in the United States. The highest rates are among Native Americans and the lowest are among Asian/Pacific Islanders. Similar rates have been described for suicide attempts and recent data suggest that some

groups, such as Hispanic females, may be exhibiting more frequent suicide attempts over recent years.

When evaluating an adolescent's risk for suicide, late adolescent age, male gender, and Native American race suggest a higher risk. The presence of significant psychiatric illness, substance abuse/intoxication, and particularly past suicide attempts also increase an individual's risk. In addition, the presence of significant family conflict, dysfunction with peers, or hopelessness increases the risk of suicide.

Neuroscientific Findings in Adolescent Suicide

Several psychiatric disorders now have a significant and growing body of evidence supporting genetic susceptibility to the disorder and underlying neurological or neurochemical abnormalities. Distinct findings in suicidal patients independent of a specific psychiatric diagnosis have been described. For example, structural brain abnormalities identified as white matter hyperintensities on magnetic resonance imaging (MRI) have been associated with suicide attempts in depressed adolescents. Several studies have described abnormalities present in the serotonin neurotransmitter systems in both suicide attempters and completers. Specifically, decreased serotonin transporter binding and increased serotonin receptor density have been reported in a number of studies and may represent a possible mechanism for a heritable risk for suicidal behavior.

Treatments and Possible Effects of Antidepressants

The U.S. Food and Drug Administration has issued warnings of a possible increase in suicidal behavior in individuals being treated with antidepressant medications. These warnings originated as a result of analyses of published and unpublished data on the use of antidepressants in adolescents that indicated approximately a 2% increase in the rate of suicidal behaviors (but not completed suicides) in treated versus placebo groups. These warnings contrast with other findings, including a clear decrease in the overall suicide rate beginning with the widespread use of antidepressants, the absence of antidepressant medication in the majority of the postmortem analyses of completed suicides, and the demonstrated efficacy of antidepressants in adolescent depression. Current practice guidelines include enhanced monitoring for suicidal behaviors during the initial weeks of antidepressant therapy.

Research into psychotherapeutic treatments targeting suicidal behavior in adolescents has demonstrated

benefits for family- and community-based interventions and for cognitive-behavioral interventions that target preventing repeat suicidal behavior by improving adolescents' problem-solving skills, decreasing hopelessness and negative cognitions, and improving their ability to regulate negative emotions. Common factors in treatment approaches to suicidal adolescents involve identifying and treating risk factors such as underlying depression, substance abuse, and family stressors. Acutely suicidal patients are treated first for any injuries they have sustained and next with inpatient hospitalization or some equivalent level of supervision to assure their safety. When patients are no longer in danger, medication and psychotherapy can be provided on an outpatient basis, with the goal of preventing a future suicide attempt.

See Also the Following Articles

Antidepressant Actions on Glucocorticoid Receptors; Cytokines, Stress, and Depression; Depression and Manic-Depressive Illness; Depression Models; Major Depressive Disorder; Psychotherapy; Sex Steroids, Response to Stress and Susceptibility to Depression; Stress Generation; Suicide, Biology of; Suicide, Psychology of;

Suicide, Sociology of; War, Suicide and Sacrifice; Suicide Terrorism, Genesis of.

Further Reading

- Centers for Disease Control (2004). Methods of suicide among persons aged 10–19 years – United States, 1992–2001. *Morbidity and Mortality Weekly Report* 53, 473–474.
- Centers for Disease Control (2004). School-associated suicides – United States, 1994–1999. *Morbidity and Mortality Weekly Report* 53, 476–477.
- Centers for Disease Control (2004). Suicide attempts and physical fighting among high school students – United States, 2001. *Morbidity and Mortality Weekly Report* 53, 474–475.
- Gould, M. S., Greenberg, T., Velting, D. M., et al. (2003). Youth suicide risk and preventive interventions: a review of the past 10 years. *Journal of the American Academy of Child and Adolescent Psychiatry* 42(4), 386–405.
- Mann, J. J., Brent, D. A. and Arango, V. (2001). The neurobiology and genetics of suicide and attempted suicide: a focus on the serotonergic system. *Neuropsychopharmacology* 24, 467–477.
- Mann, J. J., Emslie, G., Baldessarini, R., et al. (2006). ACNP task force report on SSRIs and suicidal behavior in youth. *Neuropsychopharmacology* 31, 473–492.

Adrenal Cortex

G P Vinson

Queen Mary, University of London, London, UK

B J Whitehouse

King's College London, London, UK

J P Hinson

Queen Mary, University of London, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G P Vinson, B J Whitehouse, and J P Hinson, volume 1, pp 42–51, © 2000, Elsevier Inc.

Structure of the Adrenal Cortex
Hormones of the Adrenal Cortex
Control of Adrenocortical Secretion
Actions of Corticosteroids

Glossary

Aldosterone The principal mineralocorticoid secreted by the adrenal cortex.

Angiotensin II An octapeptide that provides the major hormonal support of the growth and function of the zona glomerulosa of the adrenal cortex and of aldosterone secretion.

Cortex The outer, steroid-forming part of the adrenal gland.

Corticosteroids The collective name for the steroid hormones (glucocorticoid and mineralocorticoid) that are secreted by the adrenal cortex.

Corticosteroid-binding globulin (CBG) The carrier protein to which cortisol binds in the circulation.

Corticotropin (corticotropin, adrenocorticotrophic hormone, ACTH) The pituitary hormone that largely controls growth and activity of the adrenal cortex, particularly the zona fasciculata and cortisol secretion.

Cortisol The principal glucocorticoid secreted by the adrenal cortex.

Cytochrome-P450 A family of enzymes associated with the synthesis of steroid hormones that utilize molecular oxygen to form hydroxyl groups on the steroid molecule.

<i>Dehydroepiandrosterone (DHEA, or DHA)</i>	By amount the major androgen-like hormone secreted by the adrenal cortex.
<i>Glucocorticoids</i>	Steroid hormones (especially cortisol) from the zona fasciculata of the adrenal cortex that affect carbohydrate metabolism.
<i>Medulla</i>	The inner, catecholamine (adrenaline)-secreting core of the adrenal gland.
<i>Mineralocorticoids</i>	Steroid hormones (especially aldosterone) from the zona glomerulosa of the adrenal cortex that affect mineral (especially sodium/potassium) homeostasis.
<i>Renin-angiotensin system (RAS)</i>	The blood-borne complement of enzymes and substrates that gives rise to angiotensin II.
<i>Zona fasciculata</i>	The middle part of the adrenal cortex, secreting glucocorticoids, especially cortisol.
<i>Zona glomerulosa</i>	The outer, mineralocorticoid (especially aldosterone)-secreting part of the adrenal cortex.
<i>Zona reticularis</i>	The inner part of the adrenal cortex, secreting both glucocorticoids and androgens, but also containing apoptotic (dying) cells.

Structure of the Adrenal Cortex

The adrenals are paired glands lying close to the anterior part of the kidney. Their combined wet weight is about 0.01–0.02% of the total body weight, and thus they weigh about 8 g in the adult human in both males and females.

Morphology

The gland has two main parts, the medulla at the center and the cortex surrounding it. These two parts have different developmental origins; the medulla forms part of the sympathetic nervous system, whereas the cortex, like the gonads, has a distinct embryological origin in the germinal ridge. The medulla and cortex also have different functions, though both are associated with the response to stress. Their functional differences lie both in the nature of the hormones they produce – catecholamines such as adrenaline and noradrenaline from the medulla and steroids from the cortex – and the ways in which these secretions are controlled – by neural stimulation in the case of the medulla, but by other hormones circulating in the blood in the case of the cortex.

Histologically, the cells of the adrenal cortex are arranged as three major layers, or zones, organized as concentric shells. The cells of the different zones are

generally distinguished by their shape and size and by their ultrastructure as well as by their arrangement and position within the gland. The three zones are called the *zona glomerulosa*, the *zona fasciculata*, and the *zona reticularis*. Broadly, the zona glomerulosa lies just below the connective tissue capsule and consists of whorl-like arrangements of cells. Cells of the zona fasciculata are arranged as a series of cords running centripetally, which extend to the less clearly organized network of cells of the zona reticularis (see [Figures 1 and 2](#)).

Other zones are recognized in different species, although these are usually smaller and may be transient. In humans, the most prominent of these is the fetal

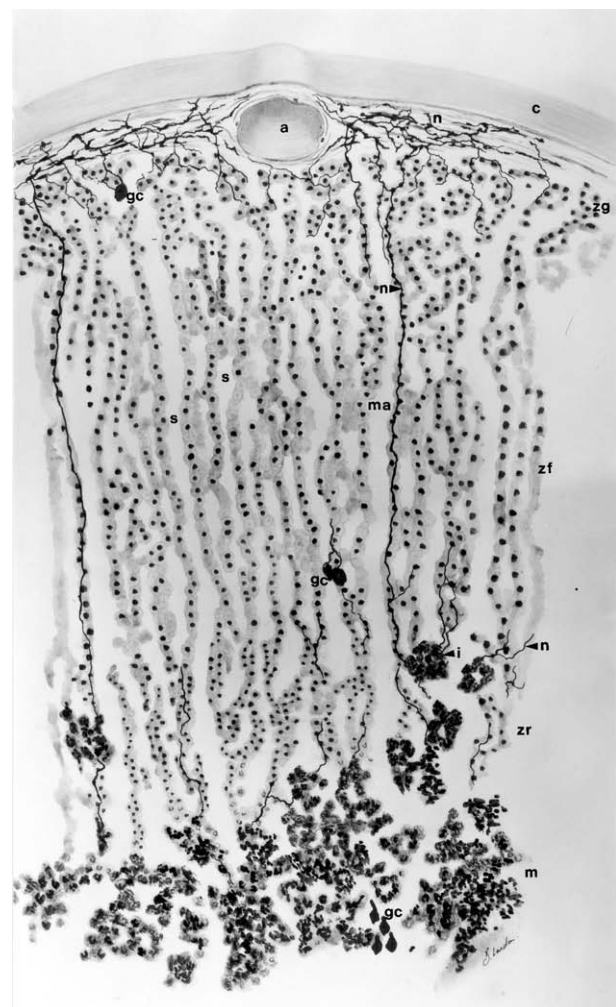


Figure 1 General structure of the mammalian adrenal cortex showing the positions of the zona glomerulosa (zg), zona fasciculata (zf), and zona reticularis (zr) and their relationship to the capsule (c), nerve cells (n), arteriolar supply (a), medullary arteries (ma), sinuses (s), medulla (m), and isolated islets (i) of chromaffin (medullary) cells lying in the inner cortex. Drawing by Bridget Landon. Figure reproduced with permission from G. P. Vinson, J. P. Hinson, and I. E. Toth. (1994). *The neuroendocrinology of the adrenal cortex*. *J. Neuroendocr.* 6, 235–246, Blackwell, Oxford.

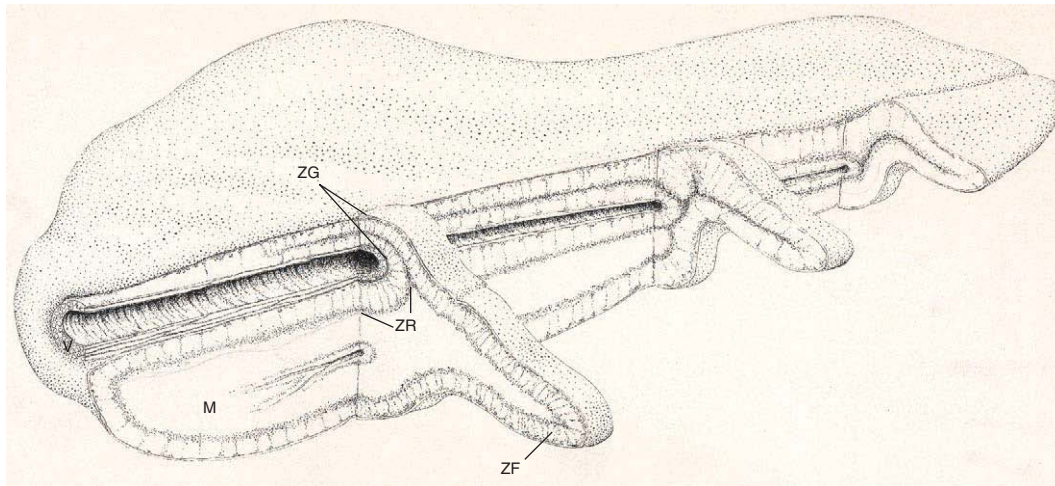


Figure 2 General structure of the human adrenal gland. Arrangement of adrenal cortex in the region of the cortical cuff. Notice the folding results in a doubling of the cortical thickness and the juxtaposition here of the inner zona glomerulosa and the central vein in cortical cuff region. M = medulla; ZG = zona glomerulosa; ZF = zona fasciculata; ZR = zona reticularis; V = central vein. Drawing by B. Landon. Figure reproduced with permission from G. P. Vinson and J. P. Hinson, *The Adrenal Gland* (V. H. T. James ed.), pp. 71–85, Raven Press. ©1992 Lippincott Williams & Wilkins.

cortex (see below). The cells of the cortex are renewed from a region just below the glomerulosa; this may be the role of the cells of the zona intermedia, which has been described as a small band of cells in some species. The cells migrate centripetally over a period of time, and as they do so, they acquire the functional and morphological characteristics of the cells of the zones that they pass through.

In humans, the zona glomerulosa seldom occupies more than 5% of the total cortical volume and its cells are gathered into more or less isolated islets. The cells are relatively small and round, with a high nuclear-to-cytoplasmic ratio. Ultrastructurally, the mitochondria of zona glomerulosa cells are characterized by their lamelliform or shelflike cristae. The smooth endoplasmic reticulum is sparse; ribosomes and polysomes are visible throughout the cytoplasm.

The zona fasciculata cells are larger, with abundant cytoplasm, and the mitochondria contain tubulovesicular cristae, although these may vary between the outer and inner parts of the zone. In conventional stains, these cells are lighter in appearance than zona reticularis cells and hence have been termed *clear cells* as opposed to the *compact cells* of the zona reticularis, which stain more deeply.

The zona reticularis occupies one-third of the cortex, and its cells are intermediate in size between those of the zona glomerulosa and zona fasciculata. The mitochondria of zona reticularis cells are elongated with tubulovesicular cristae; microbodies (lipofuchsin granules) and lipid droplets also occur.

In humans the adrenal is prominent in fetal life, when it reaches a maximum of about 0.3% of body

weight. This declines slightly in the third trimester and precipitously in the neonate. The large size of the gland is attributable to the presence of an additional zone – the fetal zone – which lies between the definitive cortex and the medulla. The cells of the fetal zone are larger than those of other zones and stain only palely in conventional histology. In the fetus, the cells of the definitive cortex contain lamelliform cristae (like the zona glomerulosa of the adult gland), whereas the mitochondria of the fetal zone cells are tubulovesicular. Studies on the origin and ultimate fate of the cells of the fetal zone are inconclusive. Some have held that the fetal zone cells arise from a second discrete proliferation, although this has been contested, and they may arise, like the reticularis, from the outer zones. Similarly, there are two explanations for its disappearance after birth; one that the zone involutes, the other that it is transformed into elements of the definitive cortex.

Circulation and Innervation

The intraglandular circulatory system deserves some attention, particularly because of its responses during stimulation. The adrenal is one of the most highly vascularized organs within the body, and detailed studies have shown that virtually every cell borders the thin and attenuated endothelial cells which line the vascular space (see [Figure 1](#)). The general pattern in mammals is as follows: a variable number of arteries supply the gland which divide in the capsular or subcapsular region to form an arteriolar plexus. Two types of vessels arise from this plexus: first, there

are the thinly walled capillaries, frequently called sinuses, which run centripetally through the cortex. These are the most abundant of the cortical vessels and may be assumed to supply all of the cortical cells. Second, there are the much more sparsely distributed medullary arteries, with thicker walls, which run through the cortex to supply the medulla directly. It is likely that blood from the cortical sinuses and the medullary arteries mixes in the medulla, and, on the basis of the relative abundance of the vessels, it is reasonable to assume that the medulla receives most of its supply from the sinuses. Accordingly the cortex exerts an endocrine control over the medulla, and, in particular, glucocorticoids (see below) are required for the development in the medulla of a key enzyme required for the synthesis of adrenaline. It is in the medulla that the sinuses are drained into the single common vein for the whole gland (see [Figure 1](#)). In the human gland there are also arteriovenous loops in which blood is apparently carried from the subcapsular arterioles in vessels which sweep down through the zona glomerulosa and the zona fasciculata, then loop back to the exterior of the gland to sites close to their origins. Also, as the vein leaves the gland, a portion of the cortex is introverted from the exterior to surround the vessel as the cortical cuff ([Figure 2](#)).

Most early studies reported that nerve fibers which entered the gland passed straight to the medulla without branching and with no evidence of nerve endings. In fact much evidence now shows that nerve fibers reach the cortex, arising partly from the medulla, but also partly through direct innervation. In general, neural arborization is most dense in the capsular and subcapsular regions and nerves may contact both adrenocortical cells and blood vessels. Both catecholaminergic and peptidergic fibers have been identified.

Hormones of the Adrenal Cortex

The hormones of the adrenal cortex are steroids: compounds that contain the *perhydrocyclopentanophenanthrene* structure consisting of four linked rings, three of six carbon atoms and one of five. There are four major groups of these compounds, classified according to the number of carbon atoms they contain.

1. C₂₇ compounds. These are the sterols, including cholesterol.

2. C₂₁ steroids. These include progesterone, the hormone of pregnancy and of female cyclic reproductive function, as well as the compounds produced more or less exclusively by the adrenal cortex. The adrenal C₂₁ steroids, collectively termed corticosteroids, are of two major types according to their

activities, the mineralocorticoids and the glucocorticoids. The activity of the mineralocorticoids is primarily concerned with the regulation of sodium and potassium homeostasis. The most potent mineralocorticoid in all mammals is aldosterone, although other corticosteroids possess some mineralocorticoid activity. The glucocorticoids were originally defined on the basis of their actions on carbohydrate metabolism but are known to have a wide variety of functions (see below). In humans and many other species (including canine, ovine, bovine) cortisol (hydrocortisone) is the most prominent secreted glucocorticoid, and smaller amounts of cortisone and corticosterone are also produced. In the rat and some other rodents (and in many nonmammalian vertebrates), adrenal 17 α -hydroxylase activity (see below) is weak, and corticosterone is the major secreted adrenocortical product.

3. C₁₉ steroids. These include the androgens, which support male structure and reproductive function. Of these, testosterone is the most potent but is normally produced in only small amounts by the adrenal. However, androstenedione, dehydroepiandrosterone (DHEA), and its sulfate (DHEAS) are prominent adrenal secretory products which may be converted to testosterone peripherally.

4. C₁₈ steroids. These include the estrogens, female sex hormones, which are not usually secreted by the adrenal. However, androstenedione and DHEA(S) from the adrenal may be converted to estrogens elsewhere and this is the major source of circulating estrogen in postmenopausal women. In the fetus, DHEAS from the fetal adrenal is converted first to 16 α -hydroxy-DHEAS and then to estriol sulfate by placental enzymes.

The circulating concentrations of the major adrenal steroids are shown in [Table 1](#).

Table 1 Examples of plasma concentrations of adrenal hormones in normal subjects^a

Steroid	Concentration
Aldosterone (pmol per L)	100–400
Androstenedione (nmol per L)	
Adults	2–13
Children	<2
Cortisol (nmol per L)	
07.00–09.00 h	280–720
21.00–24.00 h	60–340
Dehydroepiandrosterone sulfate (μ mol per L)	
Adults	2–11
Children	<2

^aValues may show considerable variations as all adrenal steroids show a diurnal rhythm, and steroid secretion may also be affected by age, nutritional status, and stage of the menstrual cycle.

Biosynthesis of the Corticosteroids

Although their biological activities may be very different (see below), the close structural similarities of the corticosteroids reflect their biosynthetic relationship. Several groups of enzymes are involved in their synthesis within the adrenal cortex, and these are summarized in [Figure 3](#) and in [Table 2](#).

Steroid hormones are formed from cholesterol. Cholesterol may be formed *de novo* from acetate in the adrenal or it may be imported into the gland from circulating low-density lipoprotein (LDL) through a process of receptor-mediated endocytosis. Cholesterol is stored as esters in the lipid droplets within the adrenocortical cell.

The rate-limiting step in these sequences is thought to be (1) the delivery of cholesterol to the inner mitochondrial membrane and CYP11A. Current evidence suggests that this is primarily facilitated by a specific protein, designated the steroidogenic acute regulatory, or StAR, protein. (2) Much evidence suggests that a second “late pathway” rate-limiting step lies between 18-hydroxydeoxycorticosterone or corticosterone and aldosterone.

However, since CYP11B2 (aldosterone synthase) can catalyze all of the steps between deoxycorticosterone and aldosterone, this could only be achieved by the regulation of the supply of the alternative precursors.

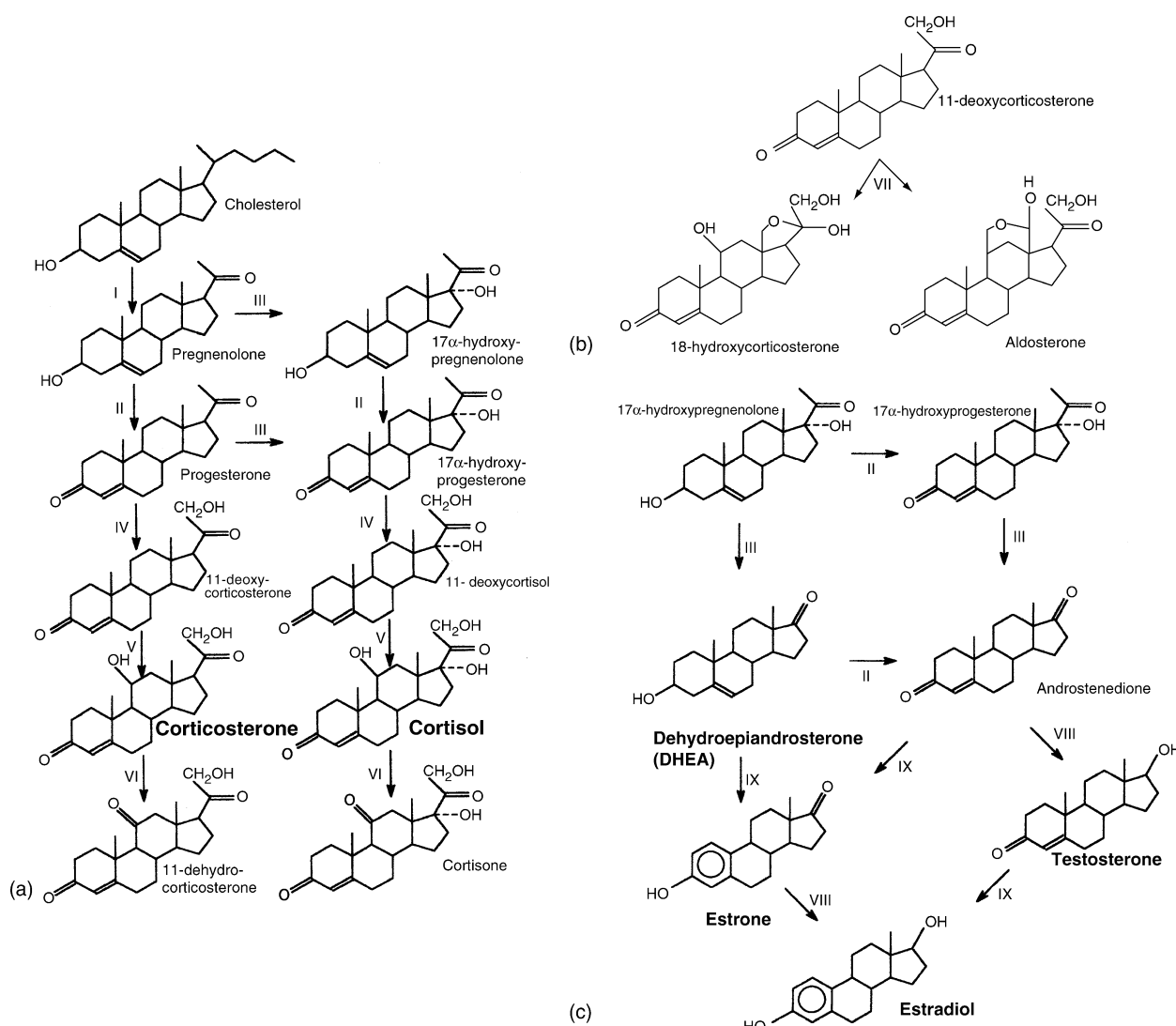


Figure 3 Pathways for biosynthesis of (a) glucocorticoids, (b) aldosterone, and (c) androgens and estrogens. The formation of aldosterone from 11-deoxycorticosterone is catalyzed by a single enzyme, CYP11B2 (aldosterone synthase). This may include the formation of other intermediates, corticosterone and/or 18-hydroxydeoxycorticosterone, that are not shown. Cortisol and, to a lesser extent, corticosterone are the major end-products of the glucocorticoid pathway, though cortisone and 11-dehydrocorticosterone (which are biologically inactive) may also be secreted. All of the dehydrogenase-catalyzed reactions may be reversible, however, i.e., cortisol may be formed from cortisone and androstenedione from testosterone.

Table 2 Enzymes involved in steroid biosynthesis

Number ^a	Type	Systematic designation	Other names	Activity
I	Cytochrome-P450	CYP11A	P-450 _{SCC}	Cholesterol side-chain cleavage
II	Dehydrogenase	3 β -Hydroxysteroid-dehydrogenase, isomerase		Conversion pregnenolone to progesterone
III	Cytochrome-P450	CYP17	P450 _{17α} , 17-hydroxylase, 17-lyase	Hydroxylation at C-17, 20 lyase (side-chain cleavage)
IV	Cytochrome-P450	CYP21	P450 ₂₁ , 21-hydroxylase	Hydroxylation at C-21
V	Cytochrome-P450	CYP11B1	P450 _{11β/18} , 11 β -hydroxylase type II	Hydroxylation at C-11 (and C-18)
VI	Dehydrogenase	11 β -Hydroxysteroid-dehydrogenase		Conversion cortisol to cortisone
VII	Cytochrome-P450	CYP11B2	Aldosterone synthase	Conversion 11-deoxycorticosterone to aldosterone (and 18-hydroxycorticosterone)
VIII	Dehydrogenase	17-Hydroxysteroid dehydrogenase		Conversion androstenedione to testosterone, estrone to estradiol and reverse
IX	Cytochrome-P450	CYP18	Aromatase	Conversion c-19 steroids (e.g., testosterone, DHEA) to estrogens

^aSee Figure 3.

Inborn errors of steroid metabolism exist which may result from congenital defects in the enzymes involved in steroid biosynthesis. A feature of all of these is that the end-product of biosynthesis (for example, cortisol) is produced in inappropriately low amounts. Accompanying this may be the excessive production of the steroid intermediate that immediately precedes the enzyme block. For example, 21-hydroxylase deficiency, the most common form, results in excessive secretion of 17 α -hydroxyprogesterone (cf. Figure 3). This has little intrinsic biological activity, though it can be converted to androgens, resulting in virilization and hirsutism. Another defect, of 11 β -hydroxylase, may result in the secretion of large amounts of 11-deoxycorticosterone, which can have profound mineralocorticoid effects, with excessive sodium retention resulting in hypertension. The lack of cortisol in these patients results in enhanced secretion of ACTH, leading both to the excessive secretion of steroid intermediates and also to the growth (hyperplasia) of the gland – hence the general term *congenital adrenal hyperplasia* or CAH. These fortunately rare conditions can usually be treated by administration of cortisol.

Knowledge of the mechanisms for corticosteroid biosynthesis have led to the development of numerous and more or less specific inhibitors of adrenal steroidogenesis. These are listed in Table 3.

Transport and Metabolism of Corticosteroids

Under normal conditions around 90% of the circulating cortisol is bound to a specific plasma binding protein, called corticosteroid binding globulin (CBG). There is no comparable plasma binding protein for aldosterone, which may be loosely bound to plasma albumin. It is generally thought that only the unbound steroid fraction is biologically active. This view has,

Table 3 Inhibitors of steroidogenesis in the adrenal cortex

Drug	Principle enzymic site of action ^a
Aminoglutethimide	CYP11A, cholesterol side-chain cleavage
Metyrapone	CYP11B1, 11 β -hydroxylase
Etomidate	CYP11B1, 11 β -hydroxylase
Trilostane	3 β -hydroxysteroid-dehydrogenase/isomerase
Ketoconazole	CYP17, 17 α -hydroxylase

^aCf. Table 2.

however, been questioned, but the conditions required for bioactivity of the bound fraction remain unclear.

The liver is the major site for corticosteroid metabolism, and in humans metabolites are generally excreted in the urine rather than by the biliary/fecal route, which predominates in some other species. The main reactions involved are reduction of ring A, hydroxylation, and conjugation as sulfates or glucuronides.

Control of Adrenocortical Secretion

Control of Cortisol Secretion

Cortisol (together with cortisone and corticosterone) is produced exclusively by the zona fasciculata and the zona reticularis of the adrenal cortex. Its secretion is almost entirely related to the degree of stimulation by circulating adrenocorticotrophic hormone (ACTH), which is secreted by the anterior pituitary gland. In the absence of ACTH (e.g., after removal of the pituitary, or hypophysectomy), glucocorticoid levels in the blood fall to about 5% of normal values within 2 h, and this fall may be reversed by administration of ACTH. Under normal conditions, cortisol secretion has a clear diurnal rhythm, paralleling the pattern of

ACTH secretion, with an early morning peak and an evening nadir. However, the secretion of ACTH is itself stimulated by all forms of stress, whatever the time of day.

Physiologically, the acute actions of ACTH are as follows: (1) to increase secretion of corticosteroids and (2) to increase blood flow through the gland. Chronically, as well as maintaining elevated glucocorticoid levels, ACTH brings about increased gland weight, which is attributable both to an increase in adrenocortical cell size and to increased blood content. Though also having some effect on cell number *in vivo*, other related pituitary peptides may stimulate adrenal mitogenesis. ACTH is also required to maintain steroid synthetic enzymes, which decline after a few days in the hypophysectomized animal. Adrenocorticotrophic hormone is also required to maintain steroid synthetic enzymes which decline after a few days in the hypophysectomized animal. Morphologically, the effects of hypophysectomy on adrenal weight are attributable to the degeneration of the zona fasciculata and zona reticularis, while the zona glomerulosa is comparatively unaffected. In the inner zones, cell numbers decline markedly (through *apoptosis*), and the remaining cells are shrunken because of major decreases in mitochondrial volume and smooth endoplasmic reticulum.

Control of Glomerulosa Function and Aldosterone Secretion

Physiologically, the most important stimuli for aldosterone secretion are altered electrolyte and fluid status (particularly sodium loss) or extracellular fluid volume reduction by hemorrhage, the use of diuretics, and postural changes: aldosterone levels are increased by upright posture. Restriction of dietary sodium intake causes hypertrophy of the zona glomerulosa and elevation of aldosterone secretion, with no concomitant effect on glucocorticoids.

The action of sodium depletion and other physiological stimulation of the zona glomerulosa is attributable to the actions of several stimulants which may act in concert. The renin-angiotensin system is especially important. This mechanism, which becomes more active when sodium intake is decreased (or sodium loss increased) or when blood volume decreases, has as its end-product the formation of a peptide hormone, called angiotensin II, which acts on zona glomerulosa cells to stimulate the secretion of aldosterone. Administered chronically, it can also reproduce the morphological changes in the zona glomerulosa which are associated with sodium depletion.

Other stimulants of aldosterone secretion include ACTH (though only transitorily) and potassium ions,

whereas the atrial natriuretic peptides (ANP) cause a decrease in circulating aldosterone concentrations.

There are numerous other factors that stimulate aldosterone secretion and zona glomerulosa function under experimental conditions, although the physiological importance of their actions remains to be clarified. These include serotonin, catecholamines, acetylcholine, vasoactive intestinal peptide, vasopressin, oxytocin, other neuropeptides, and prostaglandin E. Other inhibitors include somatostatin and dopamine.

Control of Adrenal Androgen Secretion

Adrenal androgen secretion [androstenedione and DHEA(S)] rises throughout childhood, paralleling the increase in adrenal size, and peaks just before the onset of puberty in a period that has been termed adrenarche. Secretion of these hormones remains high until 40–50 years of age, when they begin to decline (the *adrenopause*). The mechanism for regulation of adrenal androgen secretion is not completely clear, and although ACTH is a potent stimulant, the secretion of cortisol and the adrenal androgens may be dissociated under some conditions, during adrenarche, for example, when adrenal androgen secretion rises while cortisol is unchanged.

Actions of Corticosteroids

Early in the study of adrenal physiology it was discovered that this gland is essential for life because of the actions of the adrenocortical hormonal products, not those of the medulla. Among the effects of loss of adrenocortical hormones are depletion of sodium ions from the body, decreased blood volume, hypotension, increased plasma potassium levels, muscular weakness, reduced cardiac contractility, and depletion of liver and muscle glycogen together with a generally impaired capacity to respond to physical stress. It was from these beginnings that it became clear that two types of corticosteroid activity were involved. Table 4 shows the relative glucocorticoid and mineralocorticoid activity *in vivo* of some common steroids, and it can be seen that, in practice,

Table 4 Relative glucocorticoid and mineralocorticoid activity of corticosteroids

<i>Steroid</i>	<i>Mineralocorticoid activity</i>	<i>Glucocorticoid activity</i>
Aldosterone	1	0.15
Cortisol	0.001	1
Deoxycorticosterone	0.03	0.02
Corticosterone	0.004	0.4

few corticosteroids are completely specific in their effects. Thus cortisol, which normally shows about 1/1000th the mineralocorticoid activity of aldosterone *in vivo*, may have appreciable effects on electrolyte metabolism when plasma levels are elevated as in Cushing's syndrome.

Although the hormones were first classified on the basis of their physiological actions, they may also be differentiated on the basis of their binding affinity for tissue receptors. There are two main types of receptor: type I, originally detected in the kidney, was considered to be the mineralocorticoid receptor, and the more widely distributed type II receptor was thought to be responsible for glucocorticoid actions. More recently, with the cloning and characterization of the receptors, an unexpected finding has been that the recombinant mineralocorticoid (type I) receptor does not show the expected steroid specificity *in vitro* since it binds aldosterone and cortisol with equal affinity. As the levels of glucocorticoids in plasma may be up to 1000 times higher than those of aldosterone (Table 1), this raises the problem of how a specific mineralocorticoid effect can be exerted in the intact animal. Currently, it is believed that the presence of 11 β -hydroxysteroid dehydrogenase type II (cf. Figure 3a and Table 2) in mineralocorticoid target tissues explains this paradox. The enzyme converts glucocorticoids to their 11-keto derivatives (cortisone and 11-dehydrocorticosterone), which do not bind effectively to the receptor. In this way, the receptor is protected, and cortisol and corticosterone do not cause sodium retention at normal physiological levels. Conversely, the type I enzyme, which catalyses the reverse reaction, thus converting cortisone to cortisol, may enhance local availability of cortisol.

Mineralocorticoid Actions

The main actions of hormones which bind to type I (mineralocorticoid) receptors are to increase the reabsorption of sodium in the kidney and at other secretory epithelial sites, thus reducing the sodium content of urine, sweat, saliva, and so on. Effectively, sodium ions are exchanged for hydrogen and potassium ions, leading to decreased sodium and increased potassium excretion and increased urine acidity. With prolonged mineralocorticoid treatment, extracellular fluid volume expansion occurs, but after a few days the kidney escapes from the mineralocorticoid effect on sodium but not potassium excretion, and thus a major effect of prolonged mineralocorticoid exposure is that body potassium levels are reduced. Hypertension (high blood pressure) is a well-known consequence of excess corticosteroid secretion and is seen in conditions with elevated aldosterone secretion such

as Conn's syndrome as well as in conditions of elevated cortisol secretion such as Cushing's syndrome. The mechanisms by which steroids increase blood pressure is not fully understood. In both old age and chronic illness, there is also a dissociation between circulating cortisol and DHEA, with cortisol remaining unchanged, or decreasing, while DHEA secretion declines.

Glucocorticoid Actions

The effects of glucocorticoid receptor activation are varied and complex. Almost all tissues of the body have been found to contain type II receptors and a very large number of systems and processes are affected by these steroids. They are essential for the maintenance of cardiovascular and metabolic homeostasis, particularly during physically stressful conditions. In addition, they play a role in modulating inflammatory and immune responses and affect central nervous system function. In the perinatal period, glucocorticoids are particularly important in controlling lung development and maturation, which may present particular problems in premature infants who have not been sufficiently exposed to glucocorticoid.

The major metabolic effects of glucocorticoids are the stimulation of protein catabolism and hepatic glycogen synthesis. The actions essentially oppose those of insulin and are most evident in the unfed state. General inhibition of protein synthesis may be partly responsible for the inhibition of growth caused by corticosteroid administration to young subjects.

In general, glucocorticoids exert powerful suppressive effects on the immune and inflammatory systems, and synthetic glucocorticoids are widely used to control the symptoms of many inflammatory and autoimmune diseases. These effects seem to be produced through the ability of the hormones and drugs to inhibit the production and actions of the cytokines that mediate the responses. Paradoxically, however, it has recently been recognized that endogenous corticosteroids may also augment some aspects of the immune response.

Corticosteroids can freely cross the blood-brain barrier and exert complex effects on behavior. Approximately 70% of patients with Cushing's syndrome are reported to be depressed; in contrast, treatment with exogenous glucocorticoid is commonly associated with improved mental state.

Stress

The relationship between increased adrenocortical activity and stresses of many different types has been known since the 1930s, and levels of cortisol in plasma are frequently measured as an index of stress.

However, the mechanisms through which glucocorticoids confer resistance to stress is still debated. There is general agreement that the actions of glucocorticoids in the cardiovascular system which increase vascular reactivity and improve cardiac performance are beneficial under acutely stressful conditions. However, in the past it was often considered that the suppressive effects of glucocorticoids on the inflammatory and immune systems could not be rationalized in the context of protection against stress. More recently it has been suggested that glucocorticoids, by virtue of their ability to suppress the response to infection and injury, may protect against the overreaction of these systems and the risk of excessive vasodilatation, edema, and other damaging effects. The corollary of this is that an inadequate adrenocortical response to stress may be harmful in that defence reactions will proceed unchecked. This may be true of some patients; for example, in active rheumatoid arthritis or, possibly, in chronic fatigue syndrome. In some senses, glucocorticoids appear to regulate defence mechanisms in two opposing ways. In their permissive mode (at levels found in the normal diurnal cycle) the hormones act to prime the body for normal activity, whereas the suppressive effects (at levels found during severe stress) prevent overreaction.

Chronic excess of glucocorticoids also has deleterious effects on the organism. This is most obvious in Cushing's syndrome, the symptoms of which include depression, hypertension and cardiovascular disease, osteoporosis, immunosuppression, visceral obesity, insulin resistance, and metabolic abnormalities. It has also been argued that more subtle forms of exposure to increased levels of corticosteroids such as might be experienced in chronic psychosocial stress could increase susceptibility to infection and tumors and be associated with severe depression and aging-related memory impairment.

See Also the Following Articles

Adrenal Medulla; Adrenocorticotrophic Hormone (ACTH); Aldosterone and Mineralocorticoid Receptors; Androgen Action; Angiotensin; Corticosteroids and Stress; Cushing's Syndrome, Medical Aspects; Cushing's Syndrome, Neuropsychiatric Aspects.

Further Reading

- Barnes, P. J. and Adcock, I. (1993). Anti-inflammatory actions of steroid: molecular mechanisms. *Trends in Physiological Science* **14**, 438–441.
- Belloni, A. S., Mazzochi, G., Mantero, F. and Nussdorfer, G. (1987). The human adrenal cortex: ultrastructure and base-line morphometric data. *Journal of Submicroscopic Cytology* **19**, 657–658.
- Bravo, E. L. (1986). Aldosterone and other adrenal steroids. In: Zanchetti, A. & Tarzi, R. C. (eds.) *Handbook of hypertension: pathophysiology of hypertension-regulatory mechanisms*, (Vol. 8), pp. 603–625. Amsterdam: Elsevier.
- Chrousos, G. P. (1995). The hypothalamic-pituitary-adrenal axis and immune-mediated inflammation. *New England Journal of Medicine* **332**(20), 1351–1362.
- Chrousos, G. P. and Gold, P. W. (1998). A healthy body in a healthy mind and vice versa: the damaging power of uncontrollable stress. *Journal of Clinical Endocrinology and Metabolism* **83**, 1842–1845.
- Flower, R. J. and Rothwell, N. J. (1994). Lipocortin-1: cellular mechanisms and clinical relevance. *Trends in Physiological Science* **15**, 71–76.
- Gaillard, R. C. and Al-Damluji, S. (1987). Stress and the pituitary-adrenal axis. *Baillière's Clinical Endocrinology and Metabolism* **1**(2), 319–354.
- Holmes, M. C., Yau, J. L., Kotelevtsev, Y., Mullins, J. J. and Seckl, J. R. (2003). 11 Beta-hydroxysteroid dehydrogenases in the brain: two enzymes two roles. *Annals of the New York Academy of Sciences* **1007**, 357–366.
- Kino, T. and Chrousos, G. P. (2004). Glucocorticoid and mineralocorticoid receptors and associated diseases. *Essays in Biochemistry* **40**, 137–155.
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* **338**(3), 171–179.
- Miller, W. L. and Tyrell, J. B. (1999). The adrenal cortex. In: Felig, P., Baxter, J. D. & Frohman, L. A. (eds.) *Endocrinology and metabolism* (3rd edn., pp. 555–711). New York: McGraw Hill.
- Payne, A. H. and Hales, D. B. (2004). Overview of steroidogenic enzymes in the pathway from cholesterol to active steroid hormones. *Endocrine Review* **25**, 947–970.
- Seckl, J. R. (2004). Prenatal glucocorticoids and long-term programming. *European Journal of Endocrinology* **3**(Supplement), U49–U62.
- Stocco, D. M. (2001). Tracking the role of a star in the sky of the new millennium. *Molecular Endocrinology* **15**, 1245–1254.

Adrenal Insufficiency

H S Willenberg

University Hospital Duesseldorf, Duesseldorf, Germany

S R Bornstein

Technical University of Dresden, Dresden, Germany

G P Chrousos

Athens University Medical School, Greece

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by H S Willenberg, S R Bornstein, and G P Chrousos, volume 1, pp 58–62, © 2000, Elsevier Inc.

Definition and Etiology of Adrenal Insufficiency

Clinical Presentation and Diagnosis

Treatment

Additional Aspects and Implications for Stress

Glossary

<i>Adrenal crisis</i>	Acute manifestation of adrenal insufficiency with symptoms including hypotension, hypovolemic shock, fever, and/or hypoglycemia. This condition is a medical emergency.
<i>Adrenal insufficiency</i>	Characterized by hypocortisolism and/or hypoaldosteronism. See also Primary adrenal insufficiency, Secondary adrenal insufficiency, and Tertiary adrenal insufficiency.
<i>AIDS</i>	Acquired immunodeficiency syndrome.
<i>Cushing's disease</i>	State of chronic hypercortisolism due to increased secretion of corticotropin from pituitary adenomas. See also Cushing's syndrome.
<i>Cushing's syndrome</i>	State of chronic hypercortisolism in general with typical clinical features. See also Cushing's disease.
<i>Primary adrenal insufficiency</i>	State of hypocortisolism and/or hypoaldosteronism due to a pathologic process that involves the adrenal glands in the first place. See also Secondary adrenal insufficiency and Tertiary adrenal insufficiency.
<i>Secondary adrenal insufficiency</i>	State of hypocortisolism and/or hypoaldosteronism due to a diminished capacity of the pituitary gland to secrete corticotropin. See also Primary adrenal insufficiency and Tertiary adrenal insufficiency.
<i>Sheehan's syndrome</i>	Pathologic condition of a newborn characterized by intrapituitary postpartum hemorrhage.

Tertiary adrenal insufficiency

State of hypocortisolism and/or hypoaldosteronism due to a pathologic process that involves the hypothalamus in the first place. See also Primary adrenal insufficiency and Secondary adrenal insufficiency.

Definition and Etiology of Adrenal Insufficiency

Adrenal insufficiency (AI) is a heterogeneous group of diseases leading to a functional impairment of the hypothalamic-pituitary-adrenal (HPA) axis ([Table 1](#)). Adrenal insufficiency is caused by a disease process located in the adrenal glands (primary), the pituitary (secondary), or the hypothalamus (tertiary) and describes a state of glucocorticoid deficiency, mineralocorticoid deficiency, or deficiency in adrenal androgen production. Depending on the lesion that causes the adrenal hypofunction, isolated or combined impairment of adrenal hormone secretion may ensue.

Primary adrenal insufficiency was first described by Thomas Addison one and half a centuries ago; therefore, all primary forms of this condition are also termed Addison's disease. Because clinical symptoms usually do not occur until more than 90% of the adrenal cortex is destroyed or nonfunctional, both organs have to be affected.

The autoimmune destruction of the adrenal glands is the most common cause of primary adrenal insufficiency in the industrialized world, accounting for the majority of cases. Autoimmune adrenalitis can occur isolated or in combination with other autoimmune diseases ([Table 2](#)). Combinations are termed autoimmune polyendocrinopathy syndrome (APS). The monogenetic form of this syndrome, autoimmune polyendocrinopathy–candidiasis–ectodermal dystrophy (APECED, APS type I), is due to mutations in the autoimmune regulatory (AIRE) gene located on chromosome 21q22.3 and usually develops early in life. APS type II, which is also called Schmidt's syndrome when Addison's disease and hypothyroidism are present, is a complex genetic disorder and often occurs spontaneously. Susceptibility to developing this disease is conferred by genes in the human leukocyte antigen (HLA) region on the short arm of chromosome 6, and different susceptibility and resistance alleles of the major histocompatibility complex (MHC) class II

Table 1 Etiology of adrenal insufficiency^a

Primary adrenal insufficiency	Autoimmune (idiopathic) disease	
	Expansive disease	Infection (tuberculosis, fungi) Hemorrhage (sepsis, anticoagulants, trauma) Secondary tumor (metastasis)
	Congenital-hereditary diseases	Congenital adrenal hyperplasia Congenital adrenal hypoplasia Congenital unresponsiveness to ACTH X-linked adrenoleukodystrophy / adrenomyeloneuropathy
Secondary and tertiary adrenal insufficiency	Drugs	Inhibitors of steroidogenesis Adrenolytic agents Glucocorticoid antagonists
	Glucocorticoid withdrawal syndrome	Following glucocorticoid therapy
	Expansive disease	Following neurosurgery Primary or secondary tumor, infection (tuberculosis) Inflammation (autoimmune, sarcoidosis)
	Iatrogenic hypopituitarism	following surgery, following radiation
	Other	Postpartum hemorrhage (Sheehan's syndrome) Trauma Apoplexy
Isolated hypoadosteronism		Hyporeninism Surgery Congenital-inherited biosynthetic defect
		Central diseases Prolonged heparin treatment
		Tuberculosis
Isolated adrenomedullary insufficiency		Infiltration by tumors Surgery Congenital-inherited dysautonomia

^aACTH, adrenocorticotrophic hormone.**Table 2** Clinical manifestations of autoimmune polyendocrinopathy syndromes

Autoimmune endocrine diseases	Autoimmune adrenalitis, Autoimmune thyroid disease Autoimmune hypergonadotropic hypogonadism
	Type 1 diabetes mellitus Chronic hypoparathyroidism Autoimmune hypophysitis
Autoimmune diseases of other tissues	Skin/ectodermal manifestations: chronic mucocutaneous candidiasis, vitiligo, alopecia, nail dystrophy, keratoconjunctivitis, enamel dysplasia
	Chronic atrophic gastritis with: pernicious anemia, hypergastrinemia with benign carcinoids
	Celiac disease with malabsorption
	Autoimmune hepatitis

have been identified so far. Other candidate genes include the cytotoxic T-lymphocyte antigen 4 (CTLA4) coding region of chromosome 2 on q33.

Hereditary or congenital adrenal dysfunction is particularly important in the pediatric patient population and is in the majority of cases due to

deficiencies of enzymes involved in adrenal steroidogenesis, of which defects in the CYP21 enzyme gene are the most frequent. Depending on the site and number of hits to the alleles that encode for the steroidogenic cytochromes, different enzyme activities with different phenotypes result. Of these, the salt-wasting forms are the most severe and life-threatening conditions of congenital adrenal hyperplasia (CAH). In puberty and adults, hirsutism, irregularities in the menstrual cycle, acne, or hair loss may point to a heterozygous form of this condition.

Adrenal hypoplasia congenita (AHC) can occur in an X-linked trait due to a mutation of the DAX-1 gene. Hypogonadotropic hypogonadism or premature puberty can be combined with this disorder or represent the only clinical presentation. Isolated glucocorticoid deficiency, triple A syndrome, and the demyelinating X-linked lipid metabolism disorders adrenoleukodystrophy and adrenomyeloneuropathy are also forms to be considered in the workup of Addison's disease.

The destruction of the adrenal cortex can also be the consequence of chronic infectious diseases. Tests of adrenocortical reserve are frequently impaired in chronic systemic diseases, including tuberculosis, acquired immunodeficiency syndrome

(AIDS), toxoplasmosis, histoplasmosis, sarcoidosis, and amyloidosis. Bilateral metastatic infiltration or intraadrenal hemorrhage due to sepsis (*Meningococcus*) or anticoagulation are rare. However, postprimary tuberculosis still accounts for approximately 10% of the cases in the industrialized world and the majority of cases with a low standard of living. This condition is usually complicated by the involvement of other organ systems. Nonetheless, autoimmune adrenalitis may occur in cases with tuberculosis too. Of the fungal infections that are also known to cause adrenocortical dysfunction, histoplasmosis is common. Other causes of primary adrenal insufficiency have a rather low prevalence.

Iatrogenic forms of primary adrenal insufficiency include bilateral adrenalectomy and treatment with glucocorticoid receptor antagonists, such as RU486 (mifepristone); adrenolytic agents, such as mitotane and suramin; and drugs that inhibit steroidogenesis, such as ethomidate, aminoglutethimide, metyrapone, ketoconazole, and trilostane.

Secondary adrenal insufficiency due to a diminished secretion of corticotropin is frequent. Iatrogenic suppression of the HPA axis or adrenal suppression in the course of glucocorticoid treatment and subsequent discontinuation is the most common cause. Also, adrenal suppression is observed in patients with endogenous Cushing's syndrome after correction of the hypercortisolism by curative surgery. Other iatrogenic causes of secondary adrenal insufficiency include hypopituitarism following surgical intervention for pituitary and hypothalamic tumors and radiation. Space-occupying lesions in the hypothalamic-pituitary area (e.g., pituitary adenomas, craniopharyngiomas, meningiomas, or even metastatic infiltration and vascular lesions) may result in hypopituitarism. In addition, trauma, Sheehan's syndrome (intrapituitary postpartum hemorrhage) and apoplexy are also possible causes of impaired pituitary-adrenocortical function. Rare conditions include hypophysitis and granulomatous disorders affecting the pituitary gland.

Usually, secondary forms of adrenal insufficiency are limited to glucocorticoid deficiency in both sexes and androgen deficiency in women.

Tertiary adrenal insufficiency is a rare form of adrenal insufficiency. Main causes include hemorrhage due to trauma, tumorous formations, and ischemia in the hypothalamus. In addition, the hypothalamus is also a site for the endocrine feedback regulation by glucocorticoids and, therefore, is involved in iatrogenic adrenal insufficiency, as already described.

Some disease states of unknown pathophysiology, such as posttraumatic stress disorder, chronic fatigue syndrome, and fibromyalgia, may also be associated with dysfunction of the HPA axis resulting in

hypo- or hypercortisolism. Guidelines for the diagnosis and therapy of these controversially discussed conditions are under development. Psychological care, introduction into self-help groups, and symptomatic treatment are current therapeutic approaches.

In addition, there are disorders that cannot be classified within this scheme and are named relative adrenal insufficiency. In critical illness (e.g., sepsis and shock), it has been found that patients demonstrated high levels of corticotropin and cortisol. However, it has been shown that glucocorticoid therapy is beneficial in the survival for this subgroup of patients. Therefore, it has been suggested that the elevation of corticotropin in critical illness reflects a state of relative hypocortisolism. The elevation of inflammatory cytokines and presence of bacterial toxins cause a profound activation of the HPA axis and are believed to contribute to the development of relative adrenal insufficiency.

Clinical Presentation and Diagnosis

The main clinical manifestations of adrenal insufficiency are caused by hypocortisolism and hypoaldosteronism. All forms of adrenal insufficiency may present as an acute adrenal crisis or as chronic state with exacerbations. Acute symptoms include anorexia and/or vomiting, somnolence, orthostatic or persistent hypotension with dizziness, hypovolemic shock, fever, and hypoglycemia. Adrenal crisis is a medical emergency.

Patients with chronic adrenal insufficiency may present with weakness, failure to thrive (children), weight loss, anorexia, fatigue, nausea and vomiting, orthostatic hypotension, and abdominal pain. In primary adrenal insufficiency, lack of negative pituitary and hypothalamic feedback by glucocorticoids causes increased secretion of adrenocorticotrophic hormone (ACTH), which has melanocyte-stimulating activity causing hyperpigmentation (creases of palms, buccal mucosa, breast areolas and nipples, and scars). In most forms of congenital adrenal hyperplasia, hyperandrogenism (virilization of girls and premature puberty of boys and girls) is observed. In combined glucocorticoid and mineralocorticoid deficiency, relative hypovolemia leads to compensatory secretion of antidiuretic hormone with resulting hyponatremia, whereas the lack of aldosterone leads to salt loss and retention of potassium and protons. The rare forms of defective 17 α -hydroxylation lead to glucocorticoid and androgen deficiency in combination with mineralocorticoid excess.

An adrenal crisis can be diagnosed on the basis of a single cortisol measurement at the time of crisis.

Low serum sodium concentrations in combination with hyperkalemia and elevated plasma renin and ACTH concentrations strongly suggest primary adrenal insufficiency, whereas elevated serum potassium levels are not typical for secondary forms of hypocortisolism. Hypoglycemia may occur in children, women, and very thin men. Lymphocytosis with eosinophilia, azotemia, and metabolic acidosis may be present, and the first results in the diagnostic workup. Routine laboratory parameters may demonstrate no abnormalities in early adrenal hypofunction. Specific study methods to ensure and further differentiate the diagnosis of adrenal insufficiency should not delay therapeutic intervention during an adrenal crisis.

Adrenal stimulation tests provide evidence of a dysfunction of the HPA axis. Among them the rapid (60-min) corticotropin test is useful for the diagnosis of adrenal insufficiency, whereas the corticotropin releasing hormone (CRH) test and the long 48-h ACTH stimulation tests are useful for the differentiation of primary and secondary adrenal insufficiency (Table 3). Although basal cortisol production may be normal, an impaired increase in steroid output may be present and difficult to diagnose. Therefore, these tests, particularly the fast one, are valuable tools in establishing the diagnosis of adrenal hypofunction in cases of decreased adrenal reserve and in monitoring the recovery of the HPA axis after the discontinuation of glucocorticoid therapy or correction of hypercortisolism.

In several countries, screening for congenital adrenal hyperplasia and genetic workup of heterozygous patients have been shown to decrease the risk of early death and to be cost-effective in health programs.

Autoantibodies can develop and are detectable in most cases of autoimmune adrenalitis. They are directed against CYP21, CYP11 β , CYP17, or steroidogenic cells and more frequent in APS type I or II than in isolated autoimmune adrenalitis. In autoimmune polyendocrinopathy syndromes without adrenal insufficiency, they allow the prediction of a possible manifestation of this complication in 85% of patients.

Treatment

The therapeutic goal is an optimal replacement of glucocorticoids and, if necessary, mineralocorticoids. In addition, there is need of an additional glucocorticoid substitution therapy in certain stress-related circumstances. Therefore, all patients with this condition and their close relatives and friends need careful education about the management of their disease and should be handed out an Emergency Health Card, containing information on disease, regimen of hormone substitution and instructions how to avoid or deal with an adrenal crisis.

Usually, patients do very well with oral glucocorticoids. The treatment of choice is hydrocortisone. Doses of 10–20 mg in the early morning may be sufficient, especially in cases with mild secondary adrenal insufficiency. However, in primary forms or complete secondary hypocortisolism, 10–15 mg/m²/day may be necessary. To simulate the circadian rhythm of glucocorticoid secretion split, doses are given: two-thirds in the morning and one-third in the afternoon (Table 4). Before surgery or dental extraction or during periods of intercurrent illness, steroid dosage has to be doubled or even tripled. During severe stress (such as

Table 3 Differentiation among primary, secondary, and tertiary adrenal insufficiency^a

	Adrenal insufficiency		
	Primary	Secondary	Tertiary
Plasma ACTH, baseline	↑	↓	↓
Plasma ACTH after CRH	↑	↓ (response)	↑ (delayed)
Plasma cortisol, baseline during crisis	↓	↓	↓
Plasma cortisol after ACTH (60 min)	↓ ($<20 \mu\text{g dl}^{-1}$)	↑ ($<20 \mu\text{g dl}^{-1}$)	↑ ($<20 \mu\text{g dl}^{-1}$)
Plasma cortisol after CRH	↓ (no elevation)	↓ (no elevation)	↑ or ↔
Urinary free cortisol or 17-hydroxysteroid excretion after 48-h ACTH infusion	↓ (no elevation)	↑ (two- to threefold elevation)	↑ (two- to threefold elevation)
Aldosterone	↓	↔	↔
Renin	↑	↔	↔

^aACTH, adrenocorticotrophic hormone; CRH, corticotropin releasing hormone.

Table 4 Examples of glucocorticoid replacement therapy

Glucocorticoid	Morning	Noon	Late afternoon
Hydrocortisone ^a	15 mg	10 mg	5 mg
Cortisone acetate	25 mg	—	12.5 mg
Prednisone	5 mg	—	—

^aHydrocortisone is recommended.

car accidents, sepsis, or major surgery), an individual's intact adrenal gland will increase its cortisol secretion approximately fivefold. Therefore, during such situations, replacement doses higher than 200 mg day⁻¹ are not required in patients with adrenal insufficiency. Following an operation, provided that there are no complications, the hydrocortisone doses should be returned to the normal replacement regimen within a few days. Parenteral administration may be required. Although cortisol and cortisone have mineralocorticoid effects, a dose of 30 mg day⁻¹ covers only approximately 50% of the daily requirement. Therefore, 0.05–0.1 mg fludrocortisone, a synthetic mineralocorticoid, is given to Addisonian patients. Furthermore, these patients should be put on a sodium-enriched diet because they are frequently not able to conserve this electrolyte in spite of replacement therapy.

Adrenal insufficiency may lead to a deficit in androgen secretion in women. The substitution of dehydroepiandrosterone in women with this disorder has been proven beneficial and is being considered for inclusion in the treatment regimen. Women with secondary or tertiary forms of adrenal insufficiency may be put on hormone replacement therapy with estrogens and a gestagen, which has a partial androgenic action.

In some cases of secondary and tertiary adrenal insufficiency, such as in dermatomyositis and multiple sclerosis, treatment with corticotropin may be more effective than glucocorticoid therapy. In fact, the administration of corticotropin leads to the secretion of adrenal androgens, glucocorticoids, and mineralocorticoids and generally prevents the adrenal glands from hypotrophy but not the hypothalamic-pituitary unit from suppression. Oral glucocorticoid replacement is generally sufficient and easier to adjust to the individual's requirements in the great majority of the patients.

In acute adrenocortical insufficiency, therapy should be initiated immediately after taking blood for the determination of cortisol and corticotropin levels. Hypovolemic shock requires the rapid replacement of sodium and water deficits. This is accomplished by the intravenous infusion of 5%

glucose in normal saline solutions lacking potassium. Cortisol should be administered, starting with a bolus of 100 mg. An adrenal crisis is often precipitated by acute stress. Therefore, an appropriate increase in glucocorticoid replacement serves to restore glucocorticoid function. After successful treatment and improvement of the patient's condition, glucocorticoids can be administered orally and the doses can be lowered to maintenance levels. (The treatment of the different forms of congenital adrenal hyperplasia are not discussed in this article.)

Additional Aspects and Implications for Stress

Glucocorticoids do not participate only in endocrine feedback loops with corticotropin and CRH at the levels of the pituitary gland and the hypothalamus, respectively, but also in feedback loops with other peptides derived from the immune system. Of those mediators, interleukin-6 is a key interface between the immune system and the HPA axis. This cytokine is found to be elevated in patients with hypocortisolism and believed to play a role in the steroid withdrawal syndrome.

Glucocorticoids are associated with a switch in production from T helper cell (Th)1 to Th2 cytokines. The absence of glucocorticoids leads to an enhanced Th1 profile, promoting autoimmune and allergic phenomena, including autoimmune thyroid disease, rheumatoid arthritis, and asthma.

Adrenal insufficiency is a state of hypocortisolism and, thus, is a decreased effectiveness of the adaptive response. The profound dysregulation of the HPA axis with loss of drive (secondary and tertiary forms) or loss of feedback (primary forms) makes a balanced homeostasis impossible. Therefore, frequently, manifestations of adrenal insufficiency are observed under stressful conditions, such as inapparent infection, physical challenge, and surgery. Once the diagnosis of adrenal insufficiency has been established, careful diagnosis and causal therapy of a possible underlying disease are also pivotal.

When Brown-Sequard and Thomas Addison recognized that the adrenal glands were organs necessary for life, they did not have knowledge of Hans Selye's concept of stress. Adrenal dysfunction practically means the disintegration of the powerful stress system, which is necessary for maintenance of homeostasis. Cases of adrenal insufficiency document the importance of such a system. Its dysfunction may result in severe life-threatening disorders.

Further Reading

- Bornstein, S. R. (1996). *The adrenal functional unit*. Stuttgart: Georg Thieme Verlag.
- Chrousos, G. P. (1998). Stressors, stress, and neuroendocrine integration of the adaptive response: 1997 Hans Selye memorial lecture. *Annals of the New York Academy of Sciences* 851, 311–335.
- Eisenbarth, G. S. and Gottlieb, P. A. (2004). Autoimmune polyendocrine syndromes. *New England Journal of Medicine* 350, 2068–2079.

Papanicolaou, D. A. (1997). Cytokines and adrenal insufficiency. *Current Opinion in Endocrinology and Diabetes* 4, 194–198.

Stratakis, C. A. and Chrousos, G. P. (1998). Adrenal diseases. In: Jameson, J. L. (ed.) *Principles of molecular medicine*, pp. 495–503. Totowa, NJ: Humana Press.

Adrenal Medulla

R Kvetnansky

Slovak Academy of Sciences, Slovak Republic

R McCarty

Vanderbilt University, Nashville, TN, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R Kvetnansky and R McCarty, volume 1, pp 63–70, © 2000, Elsevier Inc.

Biosynthesis and Release of Adrenal Medullary

Catecholamines

Stress Effects on the Biosynthetic Capacity of the Adrenal Medulla

Factors Influencing Gene Transcription in the Adrenal Medulla during Stress

Summary and Conclusion

Glossary

Adrenal medulla

The inner portion of the adrenal gland that is part of the sympathetic-adrenal medullary system. The adrenal medulla receives neural input from sympathetic preganglionic fibers arising from the intermediolateral cell column of the thoracic and lumbar regions of the spinal cord. It is known primarily for its ability to synthesize and secrete the stress hormone epinephrine.

Chromaffin cell

The cell type that makes up the bulk of the adrenal medulla, named for its capacity to take up and stain with chromium salts. In most mammalian species, there are two types of chromaffin cells: those that synthesize, store, and secrete norepinephrine and those that synthesize, store, and secrete epinephrine. These cells also synthesize, store, and

secrete a range of biologically active peptides.

Dopamine-β-hydroxylase (DBH)

An enzyme, contained within vesicles, that catalyzes the conversion of dopamine to norepinephrine.

Neuropeptide Y (NPY)

A 36-amino-acid peptide found in cell bodies and nerve terminals of the central and peripheral nervous systems. This peptide is also localized within chromaffin cells of the adrenal medulla as well as other nonneuronal cells. It has important effects on the cardiovascular system and is also a smooth muscle cell mitogen.

Phenylethanolamine N-methyltransferase (PNMT)

An enzyme localized within epinephrine-containing chromaffin cells of the adrenal medulla, adrenergic neurons of the central nervous system, and some other peripheral tissues (e.g., the heart). This enzyme catalyzes the conversion of norepinephrine to epinephrine.

Tyrosine hydroxylase (TH)

An enzyme localized to catecholamine-synthesizing cells that catalyzes the conversion of tyrosine to dihydroxyphenylalanine. The hydroxylation of tyrosine is the rate-limiting step in catecholamine biosynthesis.

The adrenal medulla played a pivotal role in studies of the physiology of stress for most of the twentieth century. As the nineteenth century came to a close, the pharmacologist John Jacob Abel of Johns Hopkins University identified epinephrine (EPI) as the endogenous substance from the adrenal gland with pronounced effects on the cardiovascular system. Soon thereafter, Takamine identified the chemical structure and the synthesis of EPI was described. EPI was the first hormone with a known chemical structure. The physiologist Walter Branford Cannon of Harvard University conducted a series of seminal experiments

on the importance of the adrenal medullary secretion of epinephrine during various types of stressful stimulation in animals and humans. Based on these studies, he introduced the terms homeostasis and fight-or-flight into the scientific literature, and these terms remain in use to the present. Another catecholamine, norepinephrine (NE), was identified in the middle of the twentieth century by Ulf S. von Euler of Karolinska University in Sweden as the neurotransmitter in sympathetic nerves.

The adrenal medulla of many mammalian species, including rats and humans, has both NE-containing and EPI-containing secretory cells. These two types of adrenal chromaffin cells are innervated by sympathetic preganglionic fibers of the splanchnic nerve. Each chromaffin cell contains approximately 10–30 thousand chromaffin granules in which catecholamines (CAs) are stored. Human adrenals contain approximately 6 mg of CAs (the ratio EPI : NE is approximately 4 : 1). The secretion of CAs from adrenal medullary chromaffin cells is controlled in part by the sympathetic preganglionic nerve fibers, and the CAs, secreted via exocytosis, are carried from the site of release by blood that drains from the adrenal medulla and into the adrenal vein. NE and EPI are then carried throughout the body (with the exception of those portions of the central nervous system where the blood–brain barrier is intact) and exert effects on most of the major organ systems of the body.

Biosynthesis and Release of Adrenal Medullary Catecholamines

The CA biosynthetic pathway is summarized in [Table 1](#). CAs are synthesized from the amino acid precursor tyrosine. There are two primary sources of tyrosine: the diet and hydroxylation of the amino acid phenylalanine in the liver.

On entry into an adrenal chromaffin cell, tyrosine is converted to dihydroxyphenylalanine (DOPA) by the soluble cytoplasmic enzyme tyrosine hydroxylase (TH). TH activity is regulated in part by concentrations of Fe²⁺, O₂, Mg²⁺, adenosine triphosphate

(ATP), and the cofactor tetrahydrobiopterin. Under most conditions, TH is the rate-limiting step in CA biosynthesis. DOPA is converted into dopamine (DA) by a nonspecific enzyme, aromatic amino acid decarboxylase (AADC). AADC activity depends on levels of its cofactor pyridoxal phosphate. DA is taken up into storage vesicles and converted to NE by dopamine-β-hydroxylase (DBH), an enzyme found in soluble and membrane-bound forms within storage vesicles. DBH activity is influenced by levels of copper, ascorbic acid, and O₂. NE is then converted to EPI by the soluble cytoplasmic enzyme phenylethanolamine N-methyltransferase (PNMT), which uses S-adenosyl-methionine as the cofactor.

Separate populations of adrenal chromaffin cells contain NE and EPI as the final products of CA biosynthesis. The latter are distinguished from the former by the presence of PNMT. Under resting conditions, low levels of EPI (but very little NE) are released into blood from the adrenal medulla. During some types of stressful stimulation, however, significant amounts of NE may be released from the adrenal medulla and may compose up to 35% of the total circulating NE. The remaining 65% of NE is released from sympathetic nerve terminals and enters blood from the site of release at the neuroeffector junction.

As already mentioned, NE and EPI are stored in chromaffin granules that are formed from the Golgi apparatus and enclosed by a single membrane. The chromaffin granules contain water, various proteins, ATP, and lipid, in addition to either NE or EPI. The most prominent proteins include the chromogranins. The release of CAs together with other compounds from adrenal chromaffin cells is via exocytosis.

Until the mid-1980s, the view of the adrenal medullary system was relatively simplistic. Acetylcholine was released from sympathetic preganglionic nerve terminals and bound to nicotinic cholinergic receptors on the adrenal chromaffin cells. The interaction of acetylcholine with its nicotinic receptor led to a depolarization of the chromaffin cell membrane, resulting in an increase in membrane permeability to sodium. Increased intracellular sodium initiated a series of events that led to an increase in the influx of calcium. Amine storage vesicles fused with the chromaffin cell membrane and released their contents of EPI or NE via exocytosis. Although it was known in the mid-1980s that chromogranins, ATP, and a fraction of the soluble DBH were released during exocytosis, the focus of attention was on EPI and its biological effects.

Since the mid-1980s, a wealth of new information has accumulated and the adrenal medulla can no longer be viewed as a tissue that is under the exclusive control of acetylcholine and concerned primarily with the synthesis, storage, and secretion of CAs ([Table 2](#)).

Table 1 Overview of the biosynthetic pathway of catecholamines in the adrenal medulla

Epinephrine
↑ Phenylethanolamine- <i>N</i> -methyltransferase
Norepinephrine
↑ Dopamine-β-hydroxylase
Dopamine
↑ Aromatic amino acid decarboxylase
Dihydroxyphenylalanine
↑ Tyrosine hydroxylase
Tyrosine

Table 2 Partial listing of neuropeptides in adrenal medullary chromaffin cells and in preganglionic nerve terminals and medullary ganglion cells and of neurotransmitter and neuropeptide receptors on adrenal medullary chromaffin cells^a

<i>Chromaffin cell peptides</i>	<i>Nerve terminal peptides</i>	<i>Receptors</i>
Calcitonin gene-related peptide	Calcitonin gene-related peptide	Nicotinic
Enkephalins	Enkephalins	Muscarinic
Galanin	Galanin	Substance P
Substance P	Substance P	Bradykinin
Neuropeptide Y	Neuropeptide Y	Vasopressin
Vasoactive intestinal polypeptide	Vasoactive intestinal polypeptide	Neurotensin
Neurotensin	Neurotensin	Angiotensin II
Somatostatin	Somatostatin	Atrial natriuretic factor
Dynorphin	Dynorphin	Cholecystokinin
Adrenomedullin	Cholecystokinin	Neurokinins
Vasopressin	PACAP	Endothelins
Guanylin		
CRH		
ACTH		
Renin		
Chromogranins A, B, and C		
Chromostatins		
Secretogranins		
Secretoneurin		
Synaptophysin		
Interleukin-1		
Natriuretic peptides		
Endothelins		
Erythropoietin		
Endorphins		
Cerebellin		
Parathyroid hormone-related protein		

^aACTH, adrenocorticotrophic hormone; CRH, corticotropin releasing hormone; PACAP, pituitary adenylate cyclase activating peptide.

Indeed, more than 30 biologically active substances have been localized to adrenal chromaffin cells, and a number of them have been shown to be released following depolarization of the chromaffin cell membrane. In addition, various biologically active neuropeptides are colocalized with acetylcholine within sympathetic preganglionic nerve terminals. Several of these neuropeptides, including substance P, NPY, vasoactive intestinal peptide (VIP), and pituitary adenylate cyclase-activating polypeptide (PACAP), appear to act as neuromodulators of cholinergic transmission. Finally, chromaffin cells have a range of membrane-bound receptors for various biologically active molecules that are released from preganglionic nerve terminals or are delivered to chromaffin cells via the circulation. These substances may also interact in complex ways to regulate secretory patterns of adrenal chromaffin cells. Each chromaffin cell may also have its own secretory phenotype based on the complement of peptide and classical neurotransmitter receptors contained on its outer cell membrane.

In some instances, peptides released from adrenal chromaffin cells may exert negative feedback effects on continued chromaffin cell secretory activity by stimulating chromaffin cell receptors. This type of intramedullary feedback loop introduces a new layer

of complexity to the regulation of adrenal medullary secretion.

Stress Effects on the Biosynthetic Capacity of the Adrenal Medulla

Plasma Epinephrine and Stress

Many different approaches have been taken to assess the functional status of the adrenal medulla. Beginning in the 1950s, fluorometric assays of CAs were employed to study the effects of stress on whole gland content of EPI and NE. In addition, fluorometric assays allowed for quantification of urinary excretion rates of CAs in laboratory animals and in humans. The earliest studies of plasma levels of CAs also made use of fluorometric assays. However, studies of plasma CAs in laboratory animals and in humans were greatly facilitated by the introduction of highly sensitive radioenzymatic thin-layer chromatographic techniques in the 1970s and high-pressure liquid chromatography with electrochemical detection (HPLC-EC) beginning in the 1980s.

Studies of plasma levels of EPI revealed a number of facets of adrenal medullary regulation. The adrenal medulla is for all intents and purposes the sole source of circulating EPI in all mammalian species that have

been studied, including humans. Gene expression of PNMT and levels of PNMT protein have also been detected in several other organs such as the heart, spleen, thymus, and some brain-stem nuclei. However, EPI synthesized in these tissues has mainly a local function and practically does not influence the EPI concentration in circulation. In all studies dealing with levels of circulating EPI and NE, the subjects must be surgically prepared with indwelling arterial or venous catheters to reduce to the greatest extent possible any stress associated with blood collection. Basal plasma levels of EPI in laboratory rats and in humans are often less than 50 pg ml^{-1} . Plasma EPI levels increase rapidly in response to even minor stressors, and there is a monotonic relationship between the intensity of a stressor and the increment above the baseline in plasma levels of EPI.

Prior stress history exerts a profound effect on the secretory response of the adrenal medulla to subsequent exposure to a stressor. For example, if a laboratory rat is exposed to the same stressor each day for several weeks, the adrenal medullary response to the next exposure to that same stressor is reduced significantly compared to first-time-stressed controls. In contrast, if these same chronically stressed animals are instead exposed to a novel stressor, the adrenal medullary response is significantly greater compared to a first-time-stressed control group. The former pattern has been referred to as habituation, and the latter response has been referred to as sensitization. The mechanisms explaining these effects are not fully understood even if changes in gene expression of CA-synthesizing enzymes are in a good correlation with changes in CA secretion. However, it has been assumed that splanchnic stimulation of the adrenal medulla is reduced during the development of habituation and is enhanced during a sensitizing experience.

An important research question since the late 1960s has been the processes by which the adrenal medulla responds to acute versus chronic intermittent exposure to stressful stimulation. This research question has been subject to ongoing refinement as new and more sensitive neurochemical and molecular biological techniques have become available. Most of the studies in this field of inquiry have focused on regulatory processes affecting CA biosynthesis and release. However, although the adrenal medulla secretes a complex array of biologically active peptides, many fewer studies have examined how the biosynthesis and release of these neuropeptides are affected by chronic intermittent stress and exposure to novel stressors.

Exposure to a single episode of immobilization stress for up to 2 h is attended by a significant decrease of 15–20% in the total adrenal medullary content of EPI. An animal or human experiencing this

sustained exposure to an acute stressor does not know when the stressor will be terminated and, once terminated, if and when the stressor will begin again. In the face of this uncertainty, how does the adrenal medulla meet the immediate needs of replenishing supplies of EPI in adrenal chromaffin cells and what, if any, provisions are made for an uncertain future?

Given that the synthesis of EPI within the adrenal medulla involves a complex interplay of several different enzymes, studies have focused on three key enzymes: TH, DBH, and PNMT. Recently, these enzymes were intensively studied at the level of their gene transcription and translation to protein, as well as the regulation of these processes by transcription factors. The most extensive literature is available for TH because of its role as the rate-limiting step in CA biosynthesis; however, the other two enzymes have also received considerable attention. The discussion here focuses in large part on the effects of acute versus repeated immobilization stress in laboratory rats because there is an extensive literature relating to this stressor and the studies have been performed in a systematic fashion.

Effects of Acute Immobilization Stress on Adrenal Catecholamine Synthesizing Enzymes

Following a single 2-h period of immobilization stress, there are profound changes in mRNA levels of adrenal biosynthetic enzymes but very little change in enzyme activity and protein levels. Compared to values in unstressed controls, TH mRNA levels increase by approximately 600%, DBH mRNA levels increase by approximately 100%, and PNMT mRNA levels increase by approximately 200% (Table 3). In contrast, following a single 2-h period of immobilization stress TH, DBH, and PNMT enzyme activities and protein levels do not change significantly. Note here the dissociation between the levels of message for the CA biosynthetic enzymes versus the actual levels of enzyme activity. Note, too, that increases in DBH mRNA levels are just as likely to not be increased significantly following a single episode of immobilization stress.

Table 3 Adrenal mRNA levels (in arbitrary units) of TH, DBH, and PNMT under basal conditions and immediately and 24 h following one 2-h period of immobilization stress

Enzyme	Baseline	Post-IMMO	24 h post-IMMO
TH	1.00 ± 0.21	$7.50 \pm 0.50^{**}$	1.60 ± 0.25
DBH	1.00 ± 0.10	$1.88 \pm 0.49^{*}$	1.42 ± 0.50
PNMT	1.00 ± 0.32	$3.00 \pm 0.33^{**}$	1.52 ± 0.20

Values are means $\pm$ SEM. * indicates $P < 0.05$ and ** indicates $P < 0.01$ compared to unstressed controls. DBH, dopamine- β -hydroxylase; IMMO, immobilization stress; PNMT, phenylethanolamine-N-methyltransferase; TH, tyrosine hydroxylase.

These changes in mRNA levels for adrenal CA-synthesizing enzymes following a single period of immobilization stress are transient and reach the peak levels 3 h after the immobilization is completed. Twenty-four hours following a single 2-h period of immobilization stress, mRNA levels for TH, DBH, and PNMT return toward baseline levels and are no longer significantly different from unstressed control levels (Table 3).

Recently, methods for quantitative evaluation of gene expression of CA biosynthetic enzymes became available. Kvetnansky and coworkers for the first time quantified basal and stress-induced concentrations of TH, DBH, and PNMT mRNA in rat adrenal medulla, sympathetic ganglia, and various brain nuclei by the reverse transcription (RT)-competitive polymerase chain reaction (PCR) method using corresponding competitors of known concentration.

In the adrenal medulla, the basal concentration of TH mRNA was approximately 0.5 amol ng^{-1} of total RNA. DBH mRNA was approximately 12 times higher (6.0 amol ng^{-1} RNA) and PNMT mRNA was approximately 56 times higher ($28.0 \text{ amol ng}^{-1}$ RNA) than TH mRNA. In the stellate ganglia, the basal concentration of TH mRNA was approximately 30 times lower than in the adrenal medulla, but the concentration of DBH mRNA in the ganglia was similar to the adrenal medulla. Unexpectedly, the presence of a low level of PNMT mRNA has been detected for the first time in the sympathetic ganglia of rats.

Immobilization stress produced approximately a sixfold increase of TH mRNA (3.0 amol ng^{-1} RNA), a 2.5-fold increase of DBH mRNA ($15.0 \text{ amol ng}^{-1}$ RNA), and approximately a fivefold increase in PNMT mRNA ($140.0 \text{ amol ng}^{-1}$ RNA), in spite of the fact that basal levels of PNMT mRNA were already very high. Twenty-four hours after the end of immobilization exposure, the mRNA concentrations of all three adrenomedullary enzymes returned to control values (in the sympathetic ganglia, in contrast, the highest concentrations were detected during this time interval). Thus, the lowest concentration of TH mRNA in the adrenal medulla and sympathetic ganglia supports the hypothesis that TH is the rate-limiting step in the biosynthesis of CAs.

Effects of Repeated Immobilization Stress on Adrenal Catecholamine-Synthesizing Enzymes

The adrenal response to repeated immobilization stress is summarized in Table 4. Compared to unstressed control rats, rats exposed to immobilization stress for 2 h per day for 6 consecutive days had significant elevations in adrenal medullary mRNA levels of TH, DBH, and PNMT even 24 h after the last stress exposure. In addition, TH, DBH, and PNMT enzyme activities and protein levels were already increased significantly, and each of these changes persisted for at least 24 h. A similar pattern of results was obtained following as few as two consecutive daily bouts of immobilization stress. However, repeated immobilization does not produce further increases in mRNA concentrations compared to the elevated values found in the adapted control group (measured 24 h after six daily immobilization periods). The levels of enzyme immunoproteins both in the adrenal medulla and sympathetic ganglia were significantly increased only after repeated immobilization.

These findings reveal that the adrenal medulla exhibits a transient increase in TH, DBH, and PNMT transcriptional activities following a single 2-h period of immobilization stress. However, these increases in transcriptional activities are not accompanied by increases in enzyme proteins and activities. If animals are exposed daily to immobilization stress for 2–6 days, increases in TH, DBH, and PNMT transcriptional activities, as well as levels of enzyme protein and activity, are enhanced and these increases are maintained for at least 24 h. Thus, the adrenal medulla of repeatedly stressed animals exhibits a significant and long-lasting enhancement of CA biosynthetic capacity already at the level of gene expression. However, as previously noted, these increases in CA biosynthetic capacity develop as the adrenal medulla habituates to repeated immobilization stress and actually releases significantly less EPI with each repeated daily exposure to stress. The importance of the increase in CA biosynthetic capacity is unmasked when a repeatedly stressed animal is exposed to a novel stressor and EPI release from the adrenal medulla is amplified over and above that of a first-time-stressed control.

Table 4 Adrenal mRNA levels and activities of TH, DBH, and PNMT under basal conditions and immediately following six daily 2-h periods of immobilization stress^a

Enzyme	Baseline	mRNA levels post-IMMO ^b	Enzyme activity post-IMMO ^b
TH	100 ± 25	375 ± 20**	320 ± 25**
DBH	100 ± 15	385 ± 40**	255 ± 25**
PNMT	100 ± 10	330 ± 25**	160 ± 10*

^aValues are means ± SEM. * indicates $P < 0.05$ and ** indicates $P < 0.01$ compared to unstressed controls. DBH, dopamine-β-hydroxylase; IMMO, immobilization stress; PNMT, phenylethanolamine-N-methyltransferase; TH, tyrosine hydroxylase.

^bExpressed as percentage of unstressed controls.

The pattern of transcriptional activation and translational control of CA-synthesizing enzymes differs in other models of stress. For example, in animals continuously exposed to a cold environment, TH mRNA levels increased by approximately threefold above baseline within 3–6 h after the start of cold exposure and remained elevated throughout. Levels of TH protein peaked after 3 days of continuous cold exposure and TH activity peaked after 6–7 days of continuous cold exposure. Other stressors such as insulin-induced hypoglycemia, administration of 2-deoxyglucose, and hypoxia have also been shown to affect increases in mRNA, protein levels, and activity of these enzymes in the adrenal medulla. Even social hierarchy affects the gene expression of these enzymes in rats. Gene expression and protein levels were higher in subordinate than in dominant rats.

Finally, it is important to note that stress-induced increases in CA biosynthetic enzymes also occur in tissues other than the adrenal medulla. For example, such changes have been described in various sympathetic ganglia, the heart (PNMT only), several hypothalamic nuclei, and noradrenergic cell bodies of the nucleus locus ceruleus and other brain-stem nuclei.

Factors Influencing Gene Transcription in the Adrenal Medulla during Stress

The Splanchnic Nerve

The chromaffin cells of the adrenal medulla are innervated by terminals arising from the splanchnic nerve. It has been hypothesized that splanchnic input is critical for the enhanced transcription of genes for TH, DBH, PNMT, and various neuropeptides during stressful stimulation. Kvetnansky, Sabban, and their colleagues tested this hypothesis by examining the effects of adrenal denervation on the transcriptional activation of TH, PNMT, and NPY by various stressors.

In their first experiment, the right adrenal gland was intact and the left adrenal gland was denervated in a group of laboratory rats. These animals were then immobilized for 2 h and were sacrificed 3 h later. The results indicated that the immobilization-induced increases in mRNAs for PNMT and TH were not dependent on an intact splanchnic nerve. In contrast, splanchnectomy abolished the stress-induced increase in NPY mRNA (Table 5). In other experiments, it has been shown that an intact splanchnic nerve is required for the increase in enzyme activities for TH and PNMT following immobilization stress. Taken together, these results demonstrate that the splanchnic nerve does play a critical role in the immobilization-stress-induced increase in NPY mRNA and in the increase in TH and PNMT enzyme activities. However, other as yet unidentified factors are responsible for the immobilization-stress-induced increases

in TH mRNA. It should be noted that glucocorticoids are an important regulator of PNMT gene expression.

In another set of experiments, Kvetnansky, Sabban, and their coworkers examined the role of splanchnic input on increases in TH mRNA following three different stressors: cold exposure, insulin-induced hypoglycemia, and immobilization (Table 6). With chronic cold exposure and with insulin-induced hypoglycemia, TH mRNA increased several-fold above basal levels in the intact adrenal, but this increase was completely abolished in the denervated adrenal. In contrast, with immobilization stress there was a significant increase in TH mRNA levels in both the intact and the denervated adrenals. Additional

Table 5 Influence of the splanchnic nerve on transcriptional activation of TH, PNMT, and NPY under basal conditions and 3 h following one 2-h period of immobilization stress

Treatment	TH mRNA	PNMT mRNA	NPY mRNA
Sham	100 ± 33	100 ± 17	100 ± 29
Sham + IMMO	479 ± 76**	863 ± 82**	259 ± 35*
Denervated	100 ± 21	110 ± 21	88 ± 35
Denervated + IMMO	426 ± 72**	776 ± 132**	79 ± 14

Reproduced from Sabban, E. L., Nankova, B., Hiremagalur, B., et al. (1996), Molecular mechanisms in immobilization stress elicited rise in expression of genes for adrenal catecholamine biosynthetic enzymes, neuropeptide Y and proenkephalin. In: McCarty, R., Aguilera, G., Sabban, E. L. & Kvetnansky, R. (eds.) *Stress: molecular genetic and neurobiological advances*, pp. 611–628. New York: Gordon and Breach. Values are means ± SEM and are expressed as a percentage of the sham-operated right adrenal gland. * indicates $P < 0.05$ and ** indicates $P < 0.01$ compared to the appropriate sham or denervated control group. IMMO, immobilization stress; NPY, neuropeptide Y; PNMT, phenylethanolamine-N-methyltransferase; TH, tyrosine hydroxylase.

Table 6 Influence of the splanchnic nerve on transcriptional activation of TH under basal conditions and after exposure to three stressors^a

Treatment	Cold stress ^b	Insulin ^c	IMMO ^d
Sham control	100 ± 5	100 ± 23	100 ± 30
Sham + stress	302 ± 36**	375 ± 68**	478 ± 77**
Denervated control	106 ± 12	101 ± 34	100 ± 19
Denervated + stress	104 ± 13	112 ± 37	425 ± 70**

^aAdapted from Kvetnansky, R. and Sabban, E. L. (1998). Stress and molecular biology of neurotransmitter-related enzymes. In Csermely, P. (ed.) *Special issue on stress of life: from molecules to man. Annals of the New York Academy of Sciences* **851**, 342–356. Values are means ± SEM and are expressed in arbitrary units. ** indicates $P < 0.01$ compared to the appropriate sham or denervated control group. IMMO, immobilization stress; TH, tyrosine hydroxylase.

^bCold stress for 5 days at 5°C.

^cInsulin-induced hypoglycemia (5 h after administration of 5.0 IU/kg).

^dSingle 2-h period of immobilization stress with sacrifice 3 h later.

experiments demonstrated that the immobilization-induced increase in adrenal medullary TH mRNA levels is also not dependent on the pituitary gland. The administration of adrenocorticotrophic hormone (ACTH) elicited the induction of TH and DBH gene expression in the rat sympathetic ganglia but not in the adrenal medulla. The ganglia (not the adrenal medulla) were found to express mRNA for melanocortin (MC-2) receptors, which increased after immobilization. Thus, ACTH via the MC-2 receptors may be directly involved in the stress-elicited regulation of NE biosynthesis in the sympathetic ganglia but not in the adrenal medulla. The factors that regulate the increase in TH mRNA in the adrenal medulla of immobilized animals are as yet unknown.

Modulation of Transcription

The expression of genes that encode CA biosynthetic enzymes in the adrenal medulla is mediated by transcriptional mechanisms. The persistence of transcription depends on the duration and repetition of stress stimuli. Recent work identified various transcription factors that are associated with brief or intermediate duration of a single or repeated stressors.

A detailed discussion of the transcription factors and promoters involved in stress-induced increases in mRNAs for TH, DBH, PNMT, and various neuropeptides is beyond the scope of this article. However, it has been shown that the genes coding for the CA biosynthetic enzymes contain several regulatory genomic elements in their promoters. These include the CRE, GRE, API, AP2, POU/Oct, and SP1 sites. The signaling molecules that may influence the transcription rates of the genes for TH, DBH, and PNMT include P-cAMP response element binding protein (-CREB), c-Jun N-terminal kinase (JNK), c-fos, Egr-1, Fra-1, Fra-2, and others. These and other promoters and signaling molecules interact in a complex fashion to regulate mRNA levels for the various adrenal medullary biosynthetic enzymes and neuropeptides under basal conditions and during exposure to stress. The involvement of transcription factors in regulation of gene expression of CA biosynthetic enzymes during stress is not clear yet. However, the present status in the adrenal medulla is as follows. A single episode of stress response-triggered the elevation of c-fos and Fra-2, activation of the mitogen-activated protein kinase (MAPK) JNK, and induction of AP-1 factors, Egr-1, and phosphorylation of CREB. With repeated daily stress exposure, Fra-2 was a major AP-1 factor induced in the adrenal medulla but Egr-1 levels were also markedly elevated. For PNMT gene expression, two transcription factors are especially important: glucocorticoid receptor and Egr-1. The crucial role of glucocorticoids in PNMT gene regulation has

also been confirmed in corticotropin releasing hormone (CRH) knockout mice.

Thus, different transcription pathways are involved in the mediation of the elevated gene expression of CA biosynthetic enzymes during acute or chronic/repeated stress. Recently, the genetic determinants of the sympathoadrenal response to stress have been mapped.

Summary and Conclusion

The adrenal medulla and the concept of stress are inextricably intertwined. The adrenal medulla may function in concert with the sympathetic nervous system, or it may function somewhat independently in meeting the homeostatic demands presented by a variety of stressful stimuli. Although EPI secretion has been the primary focus of physiological studies of the adrenal medulla, it is now apparent that the adrenal medulla secretes a complex array of neuropeptides that exerts effects in the adrenal medulla and throughout the body. The adrenal medulla is innervated by preganglionic sympathetic nerve terminals that release acetylcholine and various colocalized peptides.

A major focus of studies beginning in the early 1970s concerned the ways in which the adrenal medulla responds to acute versus repeated exposure to stress. With the advent of sensitive and specific assay techniques, it has been shown that a complex cascade of biosynthetic changes occurs within the adrenal medulla as animals adapt to stressful stimuli.

Finally, an important lesson that has been learned relates to the stressor-specific pattern of molecular responses of the adrenal medulla. Although many stressors result in increases in mRNA levels for TH and other biosynthetic enzymes, these changes may be dependent on an intact splanchnic nerve or on an as yet unidentified mechanism that is independent of the splanchnic nerve.

In spite of thousands of experiments on the adrenal medulla and stress, there are many important questions that await answers. What mechanism explains the reduced secretion of EPI from the adrenal medulla of animals exposed to chronic intermittent stress and the enhanced secretion of EPI following the exposure of chronically stressed animals to a novel stressor? What factors regulate the increased levels of mRNAs for TH, DBH, and PNMT in the adrenal medulla versus other organs following exposure to various stressors? What are the roles played by the various neuropeptides that are released from splanchnic nerve terminals and from adrenal medullary chromaffin cells, and how do these substances interact with EPI? What is the exact mechanism of the transcription process and involvement of transcription factors

during a single and chronic stress exposure? These and other questions are being addressed by many research groups throughout the world. As answers become available, our knowledge of the physiological and molecular aspects of stress will be enhanced.

See Also the Following Articles

Adrenal Cortex; Adrenaline; Catecholamines; Dopamine, Central.

Further Reading

- Goldstein, D. S. (2001). *The autonomic nervous system in health and disease*. New York: Marcel Dekker.
- Goldstein, D. S. (2003). Catecholamines and stress. *Endocrine Regulations* 37, 69–80.
- Johnson, E. O., Kamilaris, T. C., Chrousos, G. P., et al. (1992). Mechanisms of stress: a dynamic overview of hormonal and behavioral homeostasis. *Neuroscience and Biobehavioral Reviews* 16, 115–130.
- Klimes, I., Weston, K., Gasperikova, D., et al. (2005). Mapping of genetic determinants of the sympathoneural response to stress. *Physiological Genomics* 20, 183–187.
- Kubovcakova, L., Tybitanclova, K., Sabban, E. L., et al. (2004). Catecholamine synthesizing enzymes and their modulation by immobilization stress in knockout mice. *Annals of the New York Academy of Sciences* 1018, 458–465.
- Kvetnansky, R. (2004). Stressor specificity and effect of prior experience on catecholamine biosynthetic enzyme phenylethanolamine N-methyltransferase. *Annals of the New York Academy of Sciences* 1032, 117–129.
- Kvetnansky, R., Jelokova, J., Rusnak, M., et al. (2002). Novel stressors exaggerate tyrosine hydroxylase gene expression in the adrenal medulla of rats exposed to long-term cold stress. In: McCarty, R., Aguilera, G., Sabban, E. L. & Kvetnansky, R. (eds.) *Stress: neural, endocrine and molecular studies*, pp. 121–128. London: Taylor and Francis.
- Kvetnansky, R., Micutkova, L., Kubovcakova, L., et al. (2004). Localization and regulation of phenylethanolamine N-methyltransferase gene expression in the heart of rats and mice during stress. *Annals of the New York Academy of Sciences* 1018, 405–417.
- Kvetnansky, R., Micutkova, L., Rychkova, N., et al. (2004). Quantitative evaluation of catecholamine enzymes gene expression in adrenal medulla and sympathetic ganglia of stressed rats. *Annals of the New York Academy of Sciences* 1018, 356–369.
- Kvetnansky, R., Nankova, B. B., Rusnak, M., et al. (2002). Differential gene expression of tyrosine hydroxylase in rats exposed long-term to various stressors. In: Nagatsu, T., Nabeshima, T., McCarty, R. & Goldstein, D. S. (eds.) *Catecholamine research: from molecular insights to clinical medicine*, pp. 317–320. New York: Kluwer Academic/Plenum.
- Kvetnansky, R. and Sabban, E. L. (1993). Stress-induced changes in tyrosine hydroxylase and other catecholamine biosynthetic enzymes. In: Naoi, M. & Parvez, S. H. (eds.) *Tyrosine hydroxylase: from discovery to cloning*, pp. 253–281. Utrecht: VSP Press.
- Kvetnansky, R. and Sabban, E. L. (1998). Stress and molecular biology of neurotransmitter-related enzymes. In: Csermely, P. (ed.) *Special issue on stress of life: from molecules to man. Annals of the New York Academy of Sciences* 851, 342–356.
- McCarty, R. and Gold, P. E. (1996). Catecholamines, stress and disease: a psychobiological perspective. *Psychosomatic Medicine* 58, 590–597.
- Pohorecky, L. A., Blakley, G. G., Kubovcakova, L., et al. (2004). Social hierarchy affects gene expression for catecholamine biosynthetic enzymes in rat adrenal glands. *Neuroendocrinology* 80, 42–51.
- Sabban, E. L., Hebert, M. A., Liu, X., et al. (2004). Differential effects of stress on gene transcription factors in catecholaminergic systems. *Annals of the New York Academy of Sciences* 1032, 130–140.
- Sabban, E. L. and Kvetnansky, R. (2001). Stress-triggered activation of gene expression in catecholaminergic systems: dynamics of transcriptional events. *Trends in Neurosciences* 24, 91–98.
- Sabban, E. L., Nankova, B., Hiremagalur, B., et al. (1996). Molecular mechanisms in immobilization stress elicited rise in expression of genes for adrenal catecholamine biosynthetic enzymes, neuropeptide Y and proenkephalin. In: McCarty, R., Aguilera, G., Sabban, E. L. & Kvetnansky, R. (eds.) *Stress: molecular genetic and neurobiological advances*, pp. 611–628. New York: Gordon and Breach.
- Sabban, E. L., Nankova, B. B., Serova, L. I., et al. (2004). Molecular regulation of gene expression of catecholamine biosynthetic enzymes by stress: sympathetic ganglia versus adrenal medulla. *Annals of the New York Academy of Sciences* 1018, 370–377.
- Sabban, E. L., Nankova, B. B., Serova, L. I., et al. (2002). Dynamics of transcriptional events in stress-triggered regulation of adrenal catecholamine biosynthetic enzymes. In: Nagatsu, T., Nabeshima, T., McCarty, R. & Goldstein, D. S. (eds.) *Catecholamine research: from molecular insights to clinical medicine*, pp. 321–324. New York: Kluwer Academic/Plenum.
- Tank, A. W., Augonis, C. A., Best, J. A., et al. (1996). Regulation of tyrosine hydroxylase gene expression by multiple receptor-mediated pathways during stress or nicotine treatment. In: McCarty, R., Aguilera, G., Sabban, E. L. & Kvetnansky, R. (eds.) *Stress: molecular genetic and neurobiological advances*, pp. 647–662. New York: Gordon and Breach.
- Wong, D. L., Ebert, S. N. and Morita, K. (1996). Glucocorticoid control of phenylethanolamine N-methyltransferase gene expression: implications for stress and disorders of the stress axis. In: McCarty, R., Aguilera, G., Sabban, E. L. & Kvetnansky, R. (eds.) *Stress: molecular genetic and neurobiological advances*, pp. 677–693. New York: Gordon and Breach.
- Wong, D. L., Tai, T. C., Wong-Faull, D. C., et al. (2004). Genetic mechanisms for adrenergic control during stress. *Annals of the New York Academy of Sciences* 1018, 387–397.

Adrenalectomy See: Adrenal Insufficiency; Feedback Systems; Glucocorticoids – Adverse Effects on the Nervous System; Glucocorticoids, Overview; Neuroendocrine Systems.

Adrenaline

T M Pollard

Durham University, Durham, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by T M Pollard, volume 1, pp 52–57, © 2000, Elsevier Inc.

The Effects of Adrenaline

Measuring Adrenaline Levels

Psychosocial Determinants of Adrenaline Variation

Individual Differences in SAM Reactivity

Conclusion

Glossary

<i>Adrenal glands</i>	Endocrine glands situated above the kidneys, composed of two distinct parts, the adrenal medulla, which secretes adrenaline and noradrenaline, and the adrenal cortex, which secretes corticosteroids.
<i>Adrenergic receptors</i>	Receptors that mediate the action of adrenaline. There are two types, α and β receptors.
<i>Atherogenic</i>	Encouraging the process of atherosclerosis, which is a change in the lining of arteries consisting of an accumulation of lipids, complex carbohydrates, blood and blood products, fibrous tissue, and calcium deposits.
<i>Catecholamine</i>	A group of hormones, including adrenaline, norepinephrine (noradrenaline), and dopamine, which are synthesized from the amino acid tyrosine.
<i>Job strain</i>	A construct used to describe work that makes heavy demands but offers little decision latitude for the worker.
<i>Sympathetic tone</i>	The basal rate of activity of the sympathetic nervous system.

Adrenaline, also known as epinephrine, is the main catecholamine produced by the adrenal medulla. Its secretion into the adrenal vein is under the control of

sympathetic nerves and is the most important indirect pathway of sympathetic activation. Adrenaline plays a major role in the body's physiological response to stress and affects most systems of the body. The axis as a whole is referred to here as the sympathetic adrenal medullary (SAM) axis.

The Effects of Adrenaline

The effects of adrenaline are similar to those of direct sympathetic stimulation, but they last considerably longer. Once in the bloodstream, adrenaline stimulates α and β adrenergic receptors. Through stimulation of cardiac muscle β receptors it causes an increased heart rate, force of contraction, and cardiac output. Its actions on vascular smooth muscle adrenergic receptors cause the shunting of blood away from the skin, mucosa, and kidneys to the coronary arteries, skeletal muscle, and brain. Adrenaline also enhances blood platelet adhesiveness and reduces clotting time. It promotes the role of glucagon in converting glycogen to glucose, leading to an increased secretion of glucose into the bloodstream and production of glucose in the liver. Experimental studies have shown a highly significant increase in glucose levels during infusion of adrenaline. In addition, lipolytic activity in adipose tissues is stimulated, with a resultant increase in levels of plasma free fatty acids, providing another source of energy. There is also an increase in plasma cholesterol levels, including both low-density and high-density fractions, but this effect may be at least partially explained by a reduction in plasma volume. Oxygen supply is improved, as a function of both vascular changes and bronchodilation. As a result, the body is prepared for immediate physical and mental activity, and the response is often referred to as the fight or flight reaction. This emergency function of the adrenal medulla was first identified by Cannon in 1914. At least over a few hours of secretion there appears to be no regulatory mechanism to counteract the effects of adrenaline; that is, these effects are sustained for as long as secretion of the hormone is maintained.

It is thought that repeated adrenaline secretion, or chronically raised levels of adrenaline, are likely to increase risk of cardiovascular disease, partly by changing sympathetic tone and partly by inducing changes that accelerate the process of atherosclerosis. Infusion of adrenaline leads to a sustained rise in blood pressure, and repeated elevation of blood pressure may eventually lead to hypertension via enhanced sympathetic tone and damage to the blood vessels. It is likely that chronic elevation of adrenaline levels may lead to long-term elevation of lipid levels. The free fatty acids whose release is stimulated by adrenaline are used by the liver for the synthesis of very low density lipoproteins, which then become modified to low-density lipoproteins, which are strongly implicated in the atherogenic process. Adrenaline may have a direct role to play in myocardial infarction or cerebrovascular accident through its effects on platelets and clotting time and by the precipitation of cardiac arrhythmias.

Adrenaline can alter a number of aspects of immune function and may well mediate some of the apparent effects of stress on susceptibility to infectious disease and cancers. Studies show that individuals who show large increases in adrenaline in response to laboratory stressors also show the largest immune changes. Furthermore, adrenergic blockers have been demonstrated to prevent stress-induced increases in natural killer cell number and activity. Adrenaline is thought to act on the immune system by altering lymphocyte migration from lymphoid organs such as the spleen.

Measuring Adrenaline Levels

It is possible to determine adrenaline levels in plasma or urine. The main methods used are high-performance liquid chromatography and radioenzymatic assay. The former uses high pressure to separate catecholamines, and this is usually followed by electrochemical detection. The latter relies on radioactive labeling – a radioactive substrate is added together with an enzyme, and the amount of radioactive metabolite formed is measured. Both techniques have been validated and are known to be reliable, but neither is cheap.

Levels of hormone in plasma and urine reflect neuroendocrine activity over different time periods, and it is important to use the appropriate medium. The SAM axis responds to stimuli within a few seconds of their impact, and plasma adrenaline has a half-life of 1 to 2 min, so that levels of hormone in plasma reflect the SAM activity of only the preceding minute or so. Thus, assessment of plasma adrenaline levels is only generally useful for laboratory studies of acute stressors, when an indwelling venous catheter can be used (since venipuncture itself is known to cause an

increase in adrenaline levels). Samples must be spun and frozen very shortly after collection.

Levels of excreted free adrenaline or its metabolites can be measured in urine and have been shown to provide an accurate measure of circulating levels. Hormone assayed in urine usually represents a pooling of levels over a period of hours. This is partly because hormone measured in urine is an integration of the hormone excreted into the urine produced by the body since the previous urination, and partly because excreted hormone reflects hormone secreted some time before. Information on the time lag between secretion and excretion is sparse, although some laboratory studies have produced results indicating that urinary adrenaline excretion rate reflects plasma adrenaline levels within an hour. When samples are collected over a few hours, it is important to make some correction for variation in the amount of urine produced during that time. It is also possible to collect all the urine produced in a given 24-h period in order to look at the overall level produced during a day. However, it is not always easy for study participants to comply with the demands of 24-h collection, and some checking for completeness of the sample is often needed, by assessing either total volume or creatinine clearance. Again, degradation and oxidation is a problem, and an antioxidant such as sodium metabisulphite is usually added to urine samples collected for the purpose of catecholamine determination.

Psychosocial Determinants of Adrenaline Variation

Researchers interested in understanding psychosocial determinants of adrenaline variation have investigated changes in hormone levels within individuals as well as differences between individuals. In general, the former approach has been more successful. As we have seen, adrenaline levels respond quickly to changing circumstances, and a lot of information can be gained from examination of changes in hormone levels over time. Comparisons between individuals usually focus on longer-term measures, such as 24-h urinary excretion rates, and thus only make use of averaged information for each subject. Data on within-individual variability is lost, and it is known that adrenaline excretion rates vary as much or more within individuals as they do between them. Furthermore, absolute levels of hormone are influenced by a large number of factors other than stress or the psychosocial environment, so that it is much more useful if each person can be used as his or her own control.

The main findings of studies examining the reasons for variation in adrenaline levels are summarized

below, looking first at laboratory studies and then at those that have been conducted in people going about their normal lives.

Laboratory Studies of the Determinants of Adrenaline Secretion

The largest body of research on adrenaline responses in humans has been undertaken in laboratories. In general, the focus of these studies has been measuring the adrenaline response to situations and stimuli that are assumed to be stressful. Mason reviewed the early work in an important article published in 1968. He described situations in which adrenaline levels had been shown to rise in humans. These included watching engaging movies, taking examinations, and performing set tasks such as intelligence tests or mental arithmetic. He concluded from these findings that adrenaline can rise in emotionally arousing situations that are pleasant as well as in those that are unpleasant. Since the 1960s, other workers have reported results that can be interpreted in the same way. As a result of these findings, Frankenhaeuser put forward a model in which the SAM axis was considered to be activated in response to mental or physical effort, regardless of emotional content. These studies provided some indication that adrenaline responses in women are less marked than those in men.

Adrenaline Variation in Everyday Life

Influences of habitual health-related behaviors Adrenaline levels respond to a number of features of human behavior, not just the experience of psychosocial stress. Those that have received the most attention and that are probably most important are the intake of caffeine, nicotine, and alcohol and taking exercise. All have been shown to stimulate adrenaline secretion. Caffeine also appears to potentiate adrenaline responses to stress. In laboratory studies it is possible either to eliminate these influences or to control them. Researchers hoping to isolate the effects of psychosocial stressors outside the controlled laboratory environment have a more difficult task. They have sometimes approached the problem by asking subjects to avoid these influences. However, such an approach can bring problems of its own, perhaps, for example, inducing stress in smokers asked to abstain from cigarettes. Thus, others have asked subjects to keep detailed records of their behavior and included this information in subsequent analyses.

Adrenaline variation across societies Biological anthropologists have examined cross-cultural variation

in adrenaline levels and have tried to relate the observed variation to differences in lifestyle. Such cross-sectional data suffer from the problems described previously, but some of the findings are nevertheless suggestive. In general, people living in more urbanized environments appear to excrete higher levels of urinary adrenaline than those engaged in traditional subsistence farming work, despite the lower levels of physical activity characteristic of urbanites. For example, James and colleagues compared urinary adrenaline excretion rates in young Western Samoan men living in villages or the capital, and showed that men living in the city, whether nonmanual workers, students, or laborers, excreted higher levels of urinary adrenaline than villagers on average. Similarly, Tokelauans living on Fakaofu had higher levels of urinary adrenaline than residents of Nukunonu, a nearby island less influenced by the wage economy. Still higher levels were excreted by Tokelauan migrants to New Zealand living in the urban center of Wellington. In both these studies the results were explained largely with respect to working behavior, with the suggestion that engagement in paid work, as opposed to traditional subsistence work, was more stressful.

Adrenaline variation in everyday working life In addition to their laboratory work, Frankenhaeuser and her team also pioneered investigations of the ways in which work-related stress might affect adrenaline levels. Two main kinds of comparisons have been made. The first contrasts hormone levels in people at work and at leisure, usually using a within-subject design; the second compares the effects of different types of work on hormone variation. The comparison of work and leisure addresses the question of whether nonmanual work *per se* causes an increase in adrenaline secretion. It is then relevant to ask what particular aspects of work are associated with elevated adrenaline levels.

There is good evidence that adrenaline levels are higher in men at work than in the same men at leisure. In a particularly large study, researchers in Oxford also found that in a diverse sample of British men, who provided urine samples on a normal working day and on Sunday, a rest day, urinary adrenaline levels were higher on the working day. Women who gave samples as part of the same study also showed significantly raised adrenaline levels on work days. However, there is some suggestion that the effect is not as consistent in women. In another study conducted in the same area, it was found that while men's urinary adrenaline levels were raised on working days, those for women were not, despite the fact that men and women reported experiencing very similar levels of demand.

Taken together, the results of a series of studies can be interpreted to show that workload and time pressure are the key features leading to elevated adrenaline levels at work. For example, increased urinary adrenaline levels in Swedish bus drivers when traffic was heavier were attributed to greater perceived time pressure created by the demands of remaining on schedule. Elevated adrenaline levels have also been seen in sawmill workers operating under time pressure and in people working on conveyor belts.

Efforts have been made to relate variation in adrenaline levels at work to the model of job strain proposed by Karasek and widely tested and applied in studies of workplace stress. Karasek and Theorell suggested that endocrine mechanisms explain the apparent link between job strain and ill health, particularly cardiovascular disease. They proposed that adrenaline is secreted at higher levels in those experiencing job strain due to a heavy workload and lack of decision latitude. However, Karasek acknowledged that adrenaline can also rise with increased workload, whether decision latitude is high or low. There is, then, a general consensus that adrenaline levels are most strongly related to arousal or effort, as in Frankenhaeuser's original model.

Individual Differences in SAM Reactivity

Not all people show the same SAM responsivity to stressors. It has been proposed that more reactive individuals are at greater risk of cardiovascular disease.

Sex Differences

Most early studies of the effects of stress on adrenaline in humans were conducted with only male participants. When researchers began to include women, they found that their responses to stressors were not always the same as men's, as noted previously. It is possible that the high circulating levels of estrogen seen in premenopausal women modify the way in which adrenaline is secreted in response to stress. Evidence for such an effect comes from comparisons of post- and premenopausal women and from studies of women given exogenous estrogen, as, for example, in hormone replacement therapy for postmenopausal women. There is little evidence to support an alternative explanation for increased adrenaline responses in men – that they find exposure to stressors more stressful.

Age

Little is known about the reactivity of the SAM system in normal children. There is evidence that in adults adrenaline secretion in response to acute

laboratory stressors declines with age, for example, in 60- to 75-year-old men compared to 20- to 30-year-old men. It is also the case that basal adrenaline secretion from the adrenal medulla declines with age in adults, but because plasma clearance is reduced too, plasma adrenaline concentrations do not decline.

Individuals Experiencing Chronic Stress

In contrast to initial suggestions, recent findings indicate that those experiencing chronic stress (at work, in relationships, or as caregivers) do not show greater adrenaline reactivity to acute laboratory stressors compared to those not experiencing such chronic stressors.

Conclusion

Adrenaline has widespread and important effects on the body, and these effects have clear-cut implications for cardiovascular disease. It is the hormone most consistently shown to respond to stress both in laboratory studies and in people going about their everyday lives. However, adrenaline also responds to experiences not usually considered stressful, since it rises with mental arousal or effort, whether of a negative or a positive connotation. Furthermore, the adrenaline response may not be consistent across all individuals or all stages of the life course.

See Also the Following Articles

Adrenal Medulla; Beta-Adrenergic Blockers; Cardiovascular System and Stress; Catecholamines; Sympathetic Nervous System; Workplace Stress.

Further Reading

- Baum, A. and Grunberg, N. (1995). Measurement of stress hormones. In: Cohen, S., Kessler, R. C. & Gordon, L. U. (eds.) *Measuring stress: a guide for health and social scientists*, pp. 175–182. New York: Oxford University Press.
- Charmandari, E., Tsigos, C. and Chrousos, G. (2005). Endocrinology of the stress response. *Annual Review of Physiology* 67, 259–284.
- Dimsdale, J. E. and Ziegler, M. G. (1991). What do plasma and urinary measures of catecholamines tell us about human responses to stressors? *Circulation* 22, II36–II42.
- Frankenhaeuser, M. (1989). A biopsychosocial approach to work life issues. *International Journal of Health Services* 19, 747–758.
- Gump, B. B. (1999). Do background stressors influence reactivity to and recovery from acute stressors? *Journal of Applied Social Psychology* 29, 469–494.
- James, G. D. and Brown, D. E. (1997). The biological stress response and lifestyle: catecholamines and blood pressure. *Annual Review of Anthropology* 26, 313–335.

- Mason, J. W. (1968). A review of psychoendocrine research on the sympathetic-adrenal medullary system. *Psychosomatic Medicine* 30, 631–653.
- McKinnon, W., Baum, A. and Morokoff, P. (1988). Neuroendocrine measures of stress. In: Wagner, H. (ed.) *Social psychophysiology theory and emotion: theory and clinical applications*, pp. 43–64. Chichester, UK: Wiley.
- Pollard, T. M., Ungpakorn, G., Harrison, G. A., et al. (1996). Epinephrine and cortisol responses to work: a test of the models of Frankenhaeuser and Karasek. *Annals of Behavioral Medicine* 18, 229–237.
- Seals, D. R. and Esler, M. D. (2000). Human ageing and the sympathoadrenal system. *Journal of Physiology* 528(3), 407–417.
- Stoney, C., Davis, M. and Matthews, K. A. (1987). Sex differences in physiological responses to stress and in coronary heart disease. *Psychophysiology* 24, 127–131.
- Ziegler, M. G. (1989). Catecholamine measurement in behavioral research. In: Schneiderman, N., Weiss, S. M. & Kaufman, P. G. (eds.) *Handbook of research methods in cardiovascular behavioral medicine*, pp. 167–184. New York: Plenum Press.

Adrenergic Blockers See: Beta-Adrenergic Blockers.

Adrenocortical Function, Factors Controlling Development Thereof

T Else and G D Hammer

University of Michigan, Ann Arbor, MI, USA

© 2007 Elsevier Inc. All rights reserved.

Development of the Adrenal Cortex: From the Urogenital Ridge to the Definitive Adrenal Gland
Transcription Factors Involved in Adrenocortical Development
Transcriptional Establishment of Adult Adrenocortical Function

Glossary

Adrenal steroidogenesis Process by which steroid hormones are synthesized from the cholesterol precursor. This multistep process involves several enzymes that together form the adrenal steroidogenic pathways of mineralocorticoid (aldosterone), glucocorticoid (cortisol) and androgen (dehydroepiandrosterone [DHEA]) synthesis.

Adrenocortical primordium Structure within the craniomedial part of the urogenital ridge that harbors the precursor cells for the adrenal cortices. Arises from the adrenogonadal primordium.

Genetically engineered animals Knockout animals are engineered for loss of a gene. Transgenic animals harbor genetically engineered multiple copies of one gene in their genome. Tissue-specific

knockout/transgenic animals are engineered animals that display loss or transgenic gain of a gene in a tissue-specific manner.

Hypothalamus-pituitary-adrenal (HPA) axis

Endocrine stress axis. Higher brain centers control hypothalamic CRH (corticotropin-releasing hormone) release, which in turn leads to the release of ACTH (adrenocorticotropin) from the pituitary corticotroph cells. ACTH stimulates cortisol synthesis and release in the adrenal gland.

Nuclear hormone receptor (NHR)

Family of transcription factors, usually activated by hormone/ligand binding (e.g., estrogen receptor, glucocorticoid receptor). Further members of the NHR family referred to as orphan NHRs, for which a ligand is unknown to date, have been discovered through homology screening.

POMC, ACTH

ACTH is the main tropic pituitary hormone for stimulating adrenal glucocorticoid release. It is derived from a peptide precursor, proopiomelanocortin (POMC), which also gives rise to several other peptides with diverse function, e.g., N-POMC (N-terminal POMC fragment) stimulating adrenocortical cell growth and MSH (melanocyte-stimulating hormone) stimulating skin pigmentation and regulating satiety circuits.

<i>Renin-angiotensin-aldosterone (RAA) system</i>	System to regulate body sodium homeostasis and blood pressure. Aldosterone release from the zona glomerulosa cells is regulated by angiotensin II and potassium. Angiotensin II is generated by conversion of angiotensinogen to angiotensin I by renin and subsequent conversion of angiotensin I to angiotensin II by angiotensin-converting enzyme.
<i>Transcription factor</i>	Protein factor that influences ([trans] represses or [trans]activates) gene expression through binding to certain response elements within the DNA of a gene promoter or enhancer. Transcription factors can interact with each other and mutually modify their ability to activate or repress gene transcription.
<i>Urogenital ridge</i>	Embryonic precursor that is visible as a ridgelike structure on both sides of the dorsal developing body cavity (coelom). The urogenital ridge gives rise to the adrenal cortex, the gonad (testis or ovary), and the kidney.

To serve as the functional end organ of the hypothalamus-pituitary-adrenal (HPA) axis and the renin-angiotensin-aldosterone (RAA) system, the adrenal gland needs to develop into an organ that is capable of responding to specific extrinsic stimuli with the release of different steroid hormones. While the ortholog of the adrenal cortex in lower species, the interrenal organ, consists of intermingled steroidogenic cells within the renal organ, the mammalian adrenal gland is organized into a distinct organ, which is morphologically and functionally zoned. The definitive adult human adrenal cortex comprises three functional zones: the outermost mineralocorticoid-producing zona glomerulosa, the glucocorticoid-secreting zona fasciculata, and the inner zona reticularis, which mainly secretes adrenal androgens.

The regulation of the embryonic development of the adrenal cortex can be divided into three consecutive steps. The first of these is a spatially and temporally regulated program of expression of developmental factors, mainly transcription factors, which induce differentiation and morphological shaping of the adrenogonadal primordium and the later adrenocortical primordium within the urogenital ridge. Second, during late pregnancy functional differentiation becomes increasingly controlled by endocrine factors, secreted by the pituitary (adrenocorticotropin [ACTH] and other proopiomelanocortin [POMC] peptides). The importance of this step is evident in autopsy findings from anencephalic fetuses, in which the adrenal cortex is usually hypoplastic and lacks a sufficient functional zonation.

Last, in adult life the adrenal cortex is a dynamic organ, undergoing a constant cell renewal. Steroid output and tissue growth are regulated dynamically to fit the homeostatic needs of the organism.

Development of the Adrenal Cortex: From the Urogenital Ridge to the Definitive Adrenal Gland

The adrenal gland is of mesodermal origin and is a derivative of the urogenital ridge. The ridgelike impression of this precursor organ becomes evident on both sides of the coelomic cavity in the second month of human embryonic development and is mainly formed by the developing mesonephros. Though the actual origin of the earliest adrenal precursor cells is still enigmatic, with recent observations suggesting a contribution of the intermediate mesoderm, classical studies describe a proliferation and condensation of coelomic epithelial cells lining the mesonephros as the earliest event of adrenocortogenesis. The adrenogonadal primordium can be visualized within the urogenital ridge by positive staining for the common marker steroidogenic factor-1 (SF1). This common structure further gives rise to the adrenocortical and gonadal primordium. At approximately 8 weeks, two different zones can be distinguished in the adrenocortical primordium, the outer definitive and the inner fetal zone. It is not yet established whether this morphological change is caused by a differentiation event of preexisting cells or a distinct migratory event of precursor cells. By the end of the ninth week, the developing adrenal gland becomes encapsulated, and migrating cells of neural crest origin form the precursor cells of the adrenal medulla. The clear demarcation of the two organ structures, cortex and medulla, becomes evident around term, when most of the fetal zone cells undergo apoptosis. Therefore, at birth the human adrenal cortex consists of the derivatives of the definitive zone (zona glomerulosa and zona fasciculata) and a regressing fetal zone. Later in life, at the time of adrenarche, a third zone appears between the zona fasciculata and adrenal medulla, the androgen producing zona reticularis. This zone is unique to primate species.

Notably, defects occurring at particular stages of adrenocortical development can additionally manifest in defective development of other urogenital ridge derived organs, namely, the gonad or kidney.

Transcription Factors Involved in Adrenocortical Development

There is emerging evidence of a myriad of signaling pathways and transcription factors involved in

adrenocortical development. As a result of the elucidation of the molecular basis of the genetic disorder cytomegalic adrenal hypoplasia congenita (AHC) and the investigation of several mouse models of disturbed adrenal development, it has become clear that the two orphan nuclear hormone receptors (NHRs), DAX1 (dosage-sensitive sex reversal-adrenal hypoplasia region on the X chromosome) and SF1, can be considered to comprise a central signaling network of adrenal development and differentiation. On the molecular level, there are several functional interactions between these factors that mutually influence their potential for transcriptional regulation as well as their own expression levels.

Steroidogenic Factor-1

SF1 belongs to the NHR superfamily and shares an overall domain homology with other members of this transcription factor family. SF1 harbors a classical zinc finger domain, which enables it to bind to its cognate response element in the promoter of SF1-regulated genes. The two activation function domains (AF-1, AF-2) of SF1 direct interaction with other transcription factors (e.g., WT1, SRY-related HMG box gene 9 [SOX9], GATA-binding protein 4 [GATA-4]), coactivators (e.g., CREB-binding protein [CBP/p300], steroid receptor coactivator 1), and corepressors (e.g., nuclear receptor corepressor [NCoR]). These interactions are dependent on posttranslational modifications such as phosphorylation, through the mitogen-activated protein kinase (MAP kinase) signaling pathway, and sumoylation. In the case of steroidogenic genes, SF1 has been shown to enhance transcription via the cAMP/PKA (cyclic adenosine monophosphate/protein kinase A) and the MAP kinase pathway. The molecular bases of these mechanisms are not well understood.

The recent description of the crystal structure of Sf1 revealed phospholipids in the pocket of the ligand-binding domain. Future experiments need to determine whether binding of these lipid moieties serves as a regulatory mechanism or as a constitutional part of the SF1 complex.

SF1 is a critical determinant of adrenocortical growth and differentiation. SF1 is the earliest marker of the adrenogonadal primordium in the urogenital ridge and is an essential factor for expression of several steroidogenic genes. Mice deficient in Sf1 die perinatally due to adrenal insufficiency and exhibit complete absence of adrenal glands and gonads, loss of pituitary gonadotropes, and absence of the ventromedial hypothalamus. Increased apoptosis can be detected within the adrenal primordium and the genital ridge of *Sf1*^{-/-} mice, outlining the role of SF1 and

SF1-induced genes as crucial survival factors for the adrenogonadal primordium.

The importance of Sf1 in murine adrenocortical growth and development has been further underscored by the finding that *Sf1* haploinsufficient mice have a decreased cell number in the adrenal primordium and adult adrenal gland, while the degree of steroidogenic potential per cell is elevated.

Recently, several human patients with mutations in *SF1* have been described. Partially resembling the *Sf1* knockout model, and in accordance with a crucial role of *SF1* in adrenogonadal primordium differentiation, these patients display varying degrees of XY sex reversal and adrenal insufficiency, confirming the essential role of SF1 in growth regulation of the adrenal cortex.

DAX1

DAX1 was initially cloned as the gene responsible for the most common variant of congenital AHC, cytomegalic AHC. Male XY patients with *DAX1* mutations present with hypogonadotropic hypogonadism and adrenal insufficiency. AHC is defined as a hypofunctional adrenal gland composed of a retained cytomegalic (non-regressed) fetal zone and dysfunctional definitive zone. In the adrenogonadal primordium, *Dax1* is detectable shortly after *Sf1*, which induces *DAX1* expression through binding to the *DAX* promoter.

Like SF1, DAX1 belongs to the family of NHR, but represents a rather atypical member of this family, as it contains the conserved ligand-binding domain but lacks the typical zinc finger DNA-binding motif. DAX1 contains three copies of a novel 67- to 69-amino-acid repeat that alternately has been reported to bind DNA or polyadenylated RNA and can therefore regulate gene expression on the transcriptional as well as translational level. DAX1 negatively regulates SF1-mediated gene expression via two main mechanisms. First, DAX1 can directly bind SF1 through one of its N-terminal LXXLL motifs; second, it binds to particular DNA structures in the promoter regions of certain genes, e.g., steroid acute regulatory protein gene (*Star*). Interaction with the corepressor Alien and NCoR may be partly responsible for DAX1-mediated gene repression.

DAX1 function provides a feedback mechanism for SF1-mediated transcription, because SF1 induces its own repressor (DAX1) and successively inhibits SF1 function. Though there is a significant overlap of *DAX1* and *SF1* expression, not all *DAX1*-expressing cells stain positive for SF1, giving evidence for functions independent of repression of SF1-mediated transactivation. Indeed, DAX1 has been shown to

repress the function of several other ligand-activated NHR.

Additional complexity arises from the description of a novel *DAX1* splice variant in human tissue. As a result of splicing to two alternative exons, the *DAX1* gene gives rise to two isoforms: *DAX1* and *DAX1A*. Though *DAX1A* does not show a high degree of phylogenetic conservancy and its physiological significance has yet to be proven, the occurrence of a C-terminal truncated *DAX1* is of special interest, because this isoform has been shown to transcriptionally synergize with SF1 in contrast to the mutual inhibitory effect of the classical *DAX1* and SF1.

Other Factors

Several other signaling pathways and transcription factors have been implicated in adrenocortical development. Most of our knowledge about these factors is derived from genetically engineered mouse models and rare human disorders. Depending on the temporal expression of a given transcription factor during adrenocortical development from the urogenital ridge, the formation of additional organs can be compromised. For example, disruption of *Wt1* (Wilms tumor 1) results in a renal and gonadal phenotype, and, on certain mouse strain backgrounds, the absence of the adrenal gland is evident.

Other common developmental pathways, such as the Wnt- or hedgehog-signaling networks, are involved in adrenocortical development as well. Disturbed hedgehog signaling, as in Pallister–Hall syndrome, causes not only skeletal deformation but also a varying degree of adrenal hypoplasia or complete absence of the organ, demonstrating the pleiotropy of these signaling pathways. The Wnt-signaling pathway is of special interest, as its downstream activated transcription factor β -catenin and SF1 can act synergistically to activate gene transcription.

Recent descriptions of knockout mouse models for the transcription factor *Pbx* and the transcriptional cofactor *Cited2* added two animal models displaying absence of the adrenal gland.

Future tissue-specific (adrenal-specific, e.g., *Sf1-Cre*-driven) knockouts of developmental factors will help clarify the complex transcriptional control of adrenocortical development.

Transcriptional Establishment of Adult Adrenocortical Function

The establishment of a fully functional adult organ requires the regulated expression of the various steroidogenic enzymes that generate steroid hormones

from the precursor, cholesterol. This is achieved by the activation of tissue-specific transcription factors, e.g., SF1, that orchestrate the activation of steroidogenesis in response to the stress-induced release of tropic hormones, e.g., ACTH or angiotensin. While acute ACTH-mediated steroid output is not dependent on acute transcriptional activation (transcriptional effects occur within hours, while cortisol release occurs within minutes), ACTH signaling induces steroidogenic gene transcription to ensure sufficient steroidogenic potential and ultimate steroid secretion. Expression levels of critical steroidogenic enzymes match the functional needs of the endocrine response. The main ACTH-signaling pathway is the cAMP/PKA pathway, but others have also been shown to be important (e.g., mitogen-activated protein kinase pathways). The transcriptional regulation of steroidogenic genes is outlined with two examples: *StAR*, which encodes a non-zonal-specific mitochondrial cholesterol transporter, and the zone-specific genes encoding CYP11B1, which converts deoxycortisol to cortisol and is predominantly expressed in the zona fasciculata, and CYP11B2, which is specifically expressed in the zona glomerulosa and catalyzes the last step in aldosterone synthesis.

StAR

StAR-mediated cholesterol transport is a rate-limiting event in steroidogenesis and is in great part regulated posttranslationally, e.g., direct protein phosphorylation by PKA. In addition, the transcriptional regulation of *StAR* is an example of tissue-specific as well as dynamically regulated (stress-induced and endocrine-mediated, e.g., ACTH stimulation) gene expression. Tissue specificity is achieved by transcriptional regulation through SF1 and DAX1. In the human promoter, two binding sites for SF1 regulate basal expression of *StAR*. Through the mutual antagonism of DAX1 and SF1, *StAR* expression can be downregulated directly by DAX1 interference of SF1-mediated transactivation of the *StAR* promoter. Furthermore, DAX1 has been shown to directly bind to a hairpin loop and repress transcription independent of SF1. Other basal transcription factors, which in part may play a role in tissue specificity as well as regulating basal transcription, are SP1 and GATA4. Most of the acute induction of human *StAR* promoter activity is provided by SF1, underlining the importance of SF1 not only for tissue specificity, but also for regulated transactivation upon stimulation with tropic hormones/cAMP. For the mouse *StAR* promoter, interaction of phosphorylated CREB (cAMP-responsive element binding protein), the classical downstream transcription factor of the cAMP/PKA

pathway, with a DNA element in the StAR promoter has also been recently reported.

CYP11B1 and CYP11B2

These two enzymes differ considerably in their intra-adrenal expression patterns, in accordance with their functions. *CYP11B1*, an enzyme mainly of the cortisol synthesis pathway, is expressed in the glucocorticoid-secreting zona fasciculata and is, like the *StAR* gene, regulated by SF1. *CYP11B2*, however, as the final enzyme in aldosterone synthesis, is expressed in the mineralocorticoid-secreting zona glomerulosa. Therefore, transcription of this gene needs to be regulated in a zone-specific pattern and needs to respond to angiotensin II and potassium as the main stimuli for mineralocorticoid production. Recently, nerve growth factor-induced clone B (NGFIB) and Nur-related factor 1 (NURR1) have been shown to selectively induce *CYP11B2* transcription. *NURR1* has further been shown to be expressed upon angiotensin and potassium stimulation. Moreover, the expression level of *NURR1* is highest in the zona glomerulosa and decreases in a centripetal pattern. Therefore, *NURR1* is an illustrative example of a distinct spatial- and stimuli-dependent transcription factor specific for zona glomerulosa function and mineralocorticoid synthesis within the adrenal gland.

See Also the Following Articles

Adenylyl Cyclases and Stress Response; Adrenal Cortex; Adrenal Insufficiency; Adrenocorticotrophic Hormone (ACTH); Aldosterone and Mineralocorticoid Receptors; Angiotensin; Cholesterol and Lipoproteins; Comparative Anatomy and Physiology; Corticosteroids and Stress; Corticotropin Releasing Factor (CRF); Glucocorticoids, Overview; Hypothalamic-Pituitary-Adrenal; Anatomy of the HPA Axis; Pro-opiomelanocortin (POMC); Protein Synthesis; Steroid Hydroxylases.

Further Reading

Achermann, J. C., Ito, M., et al. (1999). A mutation in the gene encoding steroidogenic factor-1 causes XY sex reversal and adrenal failure in humans. *Nature and Genetics* 22(2), 125–126.

- Bardoni, B., Zanaria, E., Guioli, S., et al. (1994). A dosage sensitive locus at chromosome Xp21 is involved in male to female sex reversal. *Nature and Genetics* 7(4), 497–501.
- Clipsham, R. and McCabe, E. R. (2003). DAX1 and its network partners: exploring complexity in development. *Molecular Genetics and Metabolism* 80(1–2), 81–120.
- Else, T. and Hammer, G. D. (2005). Genetic analysis of adrenal absence: agenesis and aplasia. *Trends in Endocrinology and Metabolism* 16(10), 458–468.
- Gallo-Payet, N. and Payet, M. D. (2003). Mechanism of action of ACTH: beyond cAMP. *Microscopy Research and Technique* 61(3), 275–287.
- Hammer, G. D. and Ingraham, H. A. (1999). Steroidogenic factor-1: its role in endocrine organ development and differentiation. *Frontiers in Neuroendocrinology* 20(3), 199–223.
- Hammer, G. D., Parker, K. L. and Schimmer, B. P. (2005). Minireview: transcriptional regulation of adrenocortical development. *Endocrinology* 146(3), 1018–1024.
- Keegan, C. E. and Hammer, G. D. (2002). Recent insights into organogenesis of the adrenal cortex. *Trends in Endocrinology and Metabolism* 13(5), 200–208.
- Lalli, E. and Sassone-Corsi, P. (2003). DAX-1, an unusual orphan receptor at the crossroads of steroidogenic function and sexual differentiation. *Molecular Endocrinology* 17(8), 1445–1453.
- Ludbrook, L. M. and Harley, V. R. (2004). Sex determination: a ‘window’ of DAX1 activity. *Trends in Endocrinology and Metabolism* 15(3), 116–121.
- Luo, X., Ikeda, Y. and Parker, K. L. (1994). A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. *Cell* 77(4), 481–490.
- McCabe, E. R. (2001). Adrenal hypoplasias and aplasias. In: Scriver, C. R., Beaudet, A. L., Sly, W. S., et al. (eds.) *The metabolic and molecular bases of inherited disease* (Vol. III), pp. 4263–4274. New York: McGraw-Hill.
- Mesiano, S. and Jaffe, R. B. (1997). Developmental and functional biology of the primate fetal adrenal cortex. *Endocrinology Reviews* 18(3), 378–403.
- Wrobel, K. H. and Suss, F. (1999). On the origin and prenatal development of the bovine adrenal gland. *Anatomy and Embryology (Berlin)* 199(4), 301–318.
- Zazopoulos, E., Lalli, E., Stocco, D. M. and Sassone-Corsi, P. (1997). DNA binding and transcriptional repression by DAX-1 blocks steroidogenesis. *Nature* 390(6657), 311–315.

Adrenocorticotrophic Hormone (ACTH)

M E Rhodes

Saint Vincent College, Latrobe, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M E Rhodes, volume 1, pp 71–74, © 2000, Elsevier Inc.

Introduction

Control of Secretion

Negative Feedback Mechanisms

Glossary

<i>Adrenal androgens</i>	Hormones produced by the adrenal cortex that have weak male sex hormone-like effects.
<i>Amygdala</i>	Bilateral nerve cell nuclei in the temporal lobes of the forebrain that are concerned primarily with fear processing and which stimulate the secretion of corticotropin-releasing hormone and, in turn, the rest of the hypothalamic-pituitary-adrenocortical axis.
<i>Catecholamine</i>	Any of a group of amines derived from catechol that has important physiological effects as neurotransmitters and hormones. Examples of catecholamines include epinephrine (adrenaline), norepinephrine (noradrenaline), and dopamine.
<i>Circadian rhythms (diurnal rhythms)</i>	A daily rhythmic activity cycle, based on 24-h intervals, that is exhibited in the physiological functions, including hormone secretions, of many organisms.
<i>Glucocorticoids</i>	Hormones produced by the adrenal cortex that increase glucose production in the liver, inhibit glucose metabolism by body tissues, and promote lipid breakdown in fat tissue. The principal glucocorticoid in humans is cortisol (hydrocortisone). When administered in high, therapeutic doses, glucocorticoids suppress immunological function, reduce inflammation, and decrease connective tissue and new bone formation.
<i>Hippocampus</i>	An area of the brain containing nerve cells that inhibit the secretion of corticotropin releasing hormone and, in turn, the rest of the hypothalamic-pituitary-adrenocortical axis.
<i>Hypothalamic-pituitary-adrenocortical (HPA) axis</i>	A hormone axis consisting of cells in the hypothalamus of the brain that secrete corticotropin-releasing hormone (CRH) and vasopressin (AVP), which

Mineralocorticoids

Proopiomelanocortin (POMC)

Vasopressin (AVP)

stimulate the secretion of adrenocorticotrophic hormone (ACTH) from the anterior pituitary gland into the bloodstream. In turn, ACTH stimulates the adrenal cortex to secrete glucocorticoids, mineralocorticoids, and adrenal androgens into the bloodstream.

Hormones produced by the adrenal cortex that reduce the excretion of sodium and enhance the excretion of potassium and hydrogen ions by the kidney. The principal mineralocorticoid in humans is aldosterone.

Precursor polypeptide that is synthesized within anterior and intermediate lobes of the pituitary gland. Products from POMC include adrenocorticotrophic hormone (ACTH) and β -endorphins.

A hormone produced by cells in the hypothalamus that is transported down the pituitary stalk to (1) the anterior pituitary gland, where, along with CRH, it stimulates the secretion of ACTH, and (2) the posterior pituitary gland, where it is carried by the bloodstream to the kidneys, where it reduces the excretion of water. Also known as antidiuretic hormone (ADH).

Introduction

Advances in the measurement of adrenocorticotrophic hormone (ACTH) and its related peptides have elucidated their extensive distribution in the body outside the pituitary gland, with differential processing in different tissues. ACTH and other peptides, including β -endorphin, a peptide known for its analgesic and euphoric effects in the brain, are produced in the pituitary from the enzymatic cleavage (posttranslational processing) of a large precursor protein, proopiomelanocortin (POMC) (Figure 1). POMC is processed in the anterior lobe to yield an N-terminal peptide, whose function is unclear, and the peptides ACTH and β -lipotropin (β -LPH). ACTH and β -LPH are secreted by the anterior pituitary. α -Melanocyte-stimulating hormone (α -MSH) and corticotropin-like intermediate lobe peptide are contained within the ACTH molecule. These peptides are found in species with a distinct intermediate lobe (e.g., amphibians, reptiles, rats); however, they are not secreted as separate hormones in humans. Within the β -LPH molecule exists the amino acid sequence for γ -LPH, β -endorphin, β -MSH, and metenkephalin.

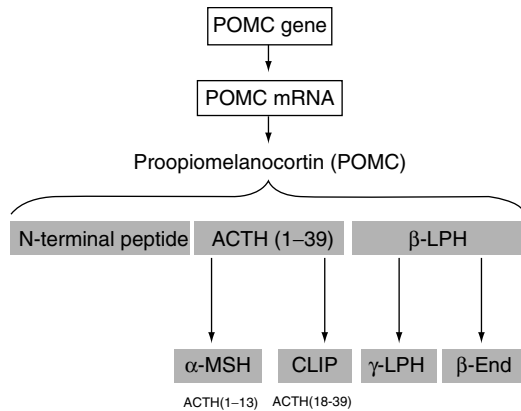


Figure 1 Proopiomelanocortin (POMC) and its biologically active peptide hormones following processing. POMC products are cleaved in a tissue- and species-specific manner. Bioactive products include adrenocorticotrophic hormone (ACTH), lipotropin (LPH), α -melanocyte-stimulating hormone (α -MSH), corticotropin-like intermediate lobe peptide (CLIP), and β -endorphin (β -End).

Evidence from light microscopic studies has suggested that in the anterior pituitary ACTH is secreted by basophilic corticotropes that represent 15–20% of the anterior pituitary cell population. With the electron microscope, ACTH-producing cells appear irregularly shaped and full of secretory granules. The complete structure of ACTH is known and the hormone has been synthesized. The entire molecule is not needed for biological activity. The first 16 amino acids beginning with the N-terminal amino acid are all that is required for minimal biological activity. Progressive increases in activity occur as the length of the chain increases until full biological activity is present with a polypeptide over 22 amino acids long. ACTH, a 39-amino-acid peptide, has a circulating half-life of 7–12 min. The development of immunoradiometric assays specific for intact human ACTH_{1–39} has improved the reliability of ACTH measurement. Radioimmunoassay of plasma ACTH and its precursors also is an accepted method for evaluating ACTH secretion, although the antibodies used also recognize some ACTH fragments. Glucocorticoid concentrations in the blood also can be measured as an index of ACTH secretion.

Control of Secretion

The regulation of ACTH secretion primarily involves the stimulatory effect of corticotropin-releasing hormone (CRH) and vasopressin (AVP), hypothalamic hormones released directly into the portal blood supplying the anterior pituitary, and the inhibitory effect of glucocorticoids. The participation of other regulators in addition to CRH and AVP, modulating ACTH secretion in the absence of or during stress,

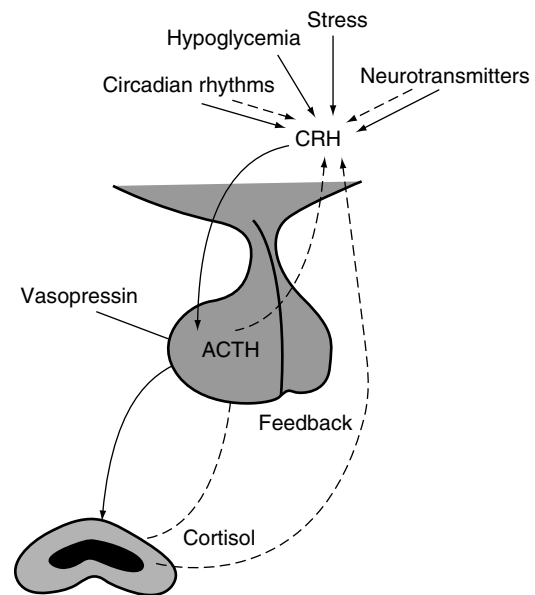


Figure 2 Factors influencing the ACTH secretion from the anterior pituitary. Peripheral vasopressin and cortisol feedback directly modulate ACTH secretion, while other factors such as stressors, circadian rhythms, cortisol feedback, and neurotransmitters modulate its release by influencing the portal secretion of CRH and vasopressin from the paraventricular nucleus of the hypothalamus. Dashed lines represent negative feedback or inhibition; solid lines represent stimulation.

can not be discounted (Figure 2). A number of other factors, such as angiotensin II, catecholamines, and immune factors, have been shown to stimulate ACTH secretion.

Areas of the brain, including the amygdala, hippocampus, and hypothalamus, stimulate and inhibit the hypothalamic-pituitary-adrenocortical (HPA) axis to different degrees depending on the time of day, season of the year, and physical and environmental stressors. The HPA axis has a normal circadian (24-h) rhythm: The secretion of CRH, ACTH, and adrenal hormones is greatest between 7 and 8 a.m., an hour or so after awakening, and lowest in early morning between 2 and 3 a.m.

HPA axis function also is intimately linked with immune function. Interleukins are chemical messengers secreted by cells of the immune system that affect the behavior of the rest of the immune system. In addition to their effects on immune function, interleukin-1 (IL-1), IL-2, and IL-6 appear to stimulate HPA axis activity. IL-1 enhances ACTH release, perhaps due to enhancement of CRH release, or by modulating the actions of other ACTH secretagogues. IL-2 augments POMC gene expression in the anterior pituitary and enhances ACTH release. The potent ACTH release by IL-6 may derive from stimulation of AVP release.

The HPA axis is highly stress responsive. Both physical and psychological stressors can cause increased activity of this hormone axis. In particular, novel stressors (stressors that are new experiences) cause HPA axis activation; with a person's repeated encountering of the stressor, there is less and less HPA axis response. This is particularly true of demanding tasks, which initially provoke an HPA axis response, but not after the person achieves mastery of the task through training and experience. Public speaking in front of a new audience can be a stressful experience to the novice but causes little or no stress, and therefore activation of the HPA axis, in the experienced speaker.

Increased ACTH secretion is also observed in pregnancy and at different periods during the menstrual cycle, because of the effects of estrogen in the circulation. Pathological (abnormal) function of the HPA axis can occur from several causes. These include repeated, uncontrollable environmental stressors, functional psychiatric illnesses such as major depression and schizophrenia, and disorders of the HPA axis itself, such as hormone-secreting tumors of the pituitary gland and adrenal cortex. Major depression is the best-studied psychiatric illness in this regard: 30–50% of major depressives have mildly to moderately increased HPA axis activity, consisting of increased blood concentrations of ACTH and cortisol at all times of the day and night, with preservation of the circadian rhythm of these hormones.

Negative Feedback Mechanisms

The mechanisms regulating ACTH secretion during stress are multifactorial, with the stimulatory effect of the hypothalamus, mainly from CRH and AVP, and the inhibitory influence of glucocorticoids. The HPA axis is regulated by two primary processes. One is closed-loop negative feedback of cortisol to hormone receptors in the hippocampus, hypothalamus, and pituitary gland. There are three types of closed-loop negative feedback systems. In the long-loop system, glucocorticoids exert a negative feedback on the anterior pituitary, hypothalamus, and hippocampus. In the short-loop system, ACTH itself feeds back on the hypothalamus, exerting negative feedback. In the ultrashort-loop system, the releasing hormone (CRH) acts directly on the hypothalamus to control its own secretion. All of these systems act to suppress the secretion of CRH, ACTH, and cortisol itself (Figure 3). The second primary process of ACTH regulation is open-loop driving of the HPA axis by the central nervous system (CNS) and by metabolic control of CRH secretion independent of glucocorticoids.

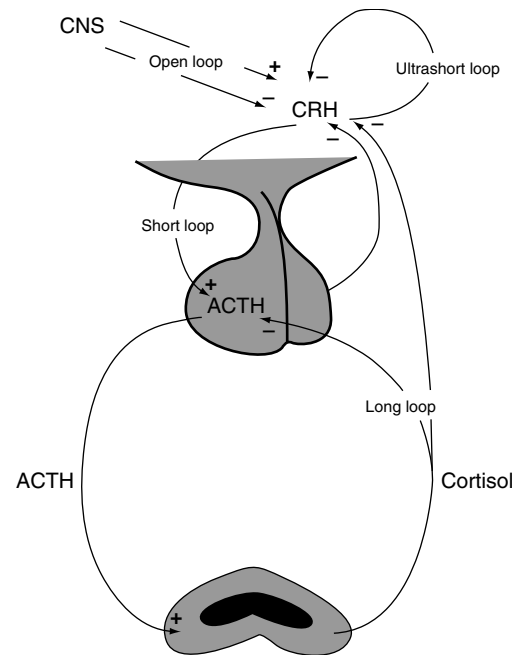


Figure 3 Examples of endocrine feedback mechanisms responsible for regulation of the HPA axis. The closed-loop negative feedback system is composed of the long, short, and ultrashort loops. The open-loop system includes both stimulatory and inhibitory inputs from the CNS that drive the HPA axis independently of closed-loop systems. The open-loop system also includes metabolic control of CRH secretion independent of cortisol. +, positive feedback or stimulation; –, negative feedback or inhibition.

Cushing's syndrome is a clinical condition resulting from chronic elevation of circulating glucocorticoids. Clinical signs and symptoms of Cushing's syndrome include obesity of the face and trunk, weakness and atrophy of limb muscles, increased blood pressure, imbalance of glucose metabolism, and psychological changes. There are two main types of Cushing's syndrome: ACTH dependent and ACTH independent. ACTH-dependent Cushing's syndrome results from increased pituitary secretion of ACTH, usually from a pituitary tumor (Cushing's disease), inappropriate ACTH secretion by nonpituitary tumors, often in the lungs, and inappropriate CRH secretion by non-hypothalamic tumors, in turn stimulating excessive pituitary ACTH secretion. These conditions, all involving excess ACTH production, cause enlargement of the adrenal glands and excessive cortisol secretion from continuous stimulation.

ACTH-independent Cushing's syndrome is caused by primary tumors or abnormalities of the adrenal cortex itself, resulting in excessive cortisol secretion and suppression of ACTH production by the pituitary. Prolonged administration of glucocorticoids for the treatment of certain illnesses also may cause ACTH-independent Cushing's syndrome.

The excessive pituitary production of ACTH and adrenal production of cortisol in Cushing's disease are only partially suppressible by low-dose dexamethasone, a synthetic and potent glucocorticoid, but are more suppressible by higher doses. In contrast, patients with nonpituitary sources of ACTH production will rarely show suppression of ACTH and cortisol by dexamethasone.

ACTH is released from the anterior pituitary in response to various stimuli, and its release is sustained or inhibited by intricate feedback systems, with ACTH at the center of this dynamic homeostatic system of loops. As part of the HPA axis, ACTH is considered one of the major stress hormones. It is now apparent that AVP and CRH cooperate as the major factors involved in the control of ACTH release.

Acknowledgments

This work was supported by NIH grant MH28380.

See Also the Following Articles

Adrenal Cortex; Corticotropin Releasing Factor (CRF); Depression and Manic-Depressive Illness; Glucocorticoid Negative Feedback; Hypothalamic-Pituitary-Adrenal; Neuroendocrine Systems.

Further Reading

- Aguilera, G. (1994). Regulation of pituitary ACTH secretion during chronic stress. *Frontiers in Neuroendocrinology* 15, 321–350.
- Aron, D. C., Findling, J. W. and Tyrrell, J. B. (2004). Hypothalamus and pituitary gland. In: Greenspan, F. S. & Gardner, D. G. (eds.) *Basic and clinical endocrinology* (7th edn., pp. 106–175). New York: McGraw-Hill.
- Charmandari, E., Tsigos, C. and Chrousos, G. (2005). Endocrinology of the stress response. *Annual Reviews in Physiology* 67, 259–284.
- Dallman, M. F., Akana, S. F., Laugero, K. D., et al. (2003). A spoonful of sugar: feedback signals of energy stores and

corticosterone regulate responses to chronic stress. *Physiology and Behavior* 79, 3–12.

- McCann, S. M. (1982). The role of brain peptides in the control of anterior pituitary hormone secretion. In: Muller, E. E. & MacLeod, R. M. (eds.) *Neuroendocrine perspectives*, (Vol. 1), pp. 1–22. Amsterdam: Elsevier Biomedical Press.
- Orth, D. N., Kovacs, W. J. and DeBold, C. R. (1992). The adrenal cortex. In: Wilson, J. D. & Foster, D. W. (eds.) *Williams textbook of endocrinology* (8th edn., pp. 489–619). Philadelphia: W. B. Saunders.
- Porterfield, S. P. (2001). Anterior pituitary gland. In: *Endocrine physiology* (2nd edn., pp. 21–47). St. Louis: Mosby.
- Porterfield, S. P. (2001). Posterior pituitary gland. In: *Endocrine physiology* (2nd edn., pp. 49–58). St. Louis: Mosby.
- Reichlin, S. (1992). Neuroendocrinology. In: Wilson, J. D. & Foster, D. W. (eds.) *Williams textbook of endocrinology* (8th edn., pp. 135–219). Philadelphia: W. B. Saunders.
- Rhodes, M. E. and Rubin, R. T. (1999). Functional sex differences (“sexual diergism”) of central nervous system cholinergic systems, vasopressin, and hypothalamic-pituitary-adrenal axis activity in mammals: a selective review. *Brain Research Reviews* 30, 135–152.
- Rubin, R. T. (1994). Neuroendocrine aspects of stress in major depression. In: Liberman, R. P. & Yager, J. (eds.) *Stress in psychiatric disorders*, pp. 37–52. New York: Springer.
- Schimmer, B. P. and Parker, K. L. (2001). Adrenocorticotrophic hormone; adrenocortical steroids and their synthetic analogs; inhibitors of the synthesis and actions of adrenocortical hormones. In: Hardman, J. G. & Limbird, L. E. (eds.) *Goodman & Gilman's the pharmacological basis of therapeutics* (10th edn., pp. 1459–1485). New York: McGraw-Hill.
- Thorner, M. O., Vance, M. L., Horvath, E. and Kovacs, K. (1992). The anterior pituitary. In: Wilson, J. D. & Foster, D. W. (eds.) *Williams textbook of endocrinology* (8th edn., pp. 221–222). Philadelphia: W. B. Saunders.
- White, A. and Ray, D. W. (2001). Adrenocorticotrophic hormone. In: DeGroot, L. J. & Jameson, J. L. (eds.) *Endocrinology* (4th edn., pp. 221–233). Philadelphia: W. B. Saunders.

Adrenogenital Syndrome See: Congenital Adrenal Hyperplasia (CAH).

Adrenomedullin See: Vasoactive Peptides.

Aerobic Exercise and Stress Reduction

E J C de Geus and J H Stubbe

Vrije University, Amsterdam, The Netherlands

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by E J C de Geus, volume 1, pp 81–86,

© 2000, Elsevier Inc.

Sites of Interaction between Stress and Aerobic Exercise
 Psychological Coping with Stress
 Psychophysiological Stress Reactivity
 Acute Effects of Exercise
 Determinants of Exercise Behavior
 Exercise and Stress: Opposing Effects on Disease
 Risk Factors

Glossary

<i>Aerobic exercise</i>	Prolonged dynamic muscular work energized by the oxidation of food (fat, carbohydrates, and protein).
<i>Endogenous opioids</i>	A class of neuromodulators that can reduce pain and induce a positive (euphoric) mood.
<i>Heritability</i>	The heritability of a trait reflects the proportion of individual variation that can be attributed to genetic factors in a given population. Heritability of a trait can be obtained by comparing trait resemblance between monozygotic (MZ) twins that are genetically identical and dizygotic (DZ) twins that share on average only half of their segregating genes. A greater intrapair resemblance between MZ twins than between DZ twins in exercise behavior, for instance, reflects the contribution of genetic factors to exercise behavior.
<i>Insulin resistance</i>	Inability of tissue to respond to insulin mediated take-up of additional glucose. Insulin resistance is associated with increased risk for cardiovascular disease.
<i>Monoamines</i>	A class of neurotransmitters that are known to influence mood. Antidepressive medication is based on boosting the action of two monoamines (norepinephrine and serotonin) in the central nervous system.
<i>Stress reactivity</i>	The bodily response to the perception of threat or challenge. There is increased activity of the sympathetic nervous system

and decreased activity of the parasympathetic nervous system that becomes manifest as increases in levels of heart rate, blood pressure, and levels of stress hormones.

Sites of Interaction between Stress and Aerobic Exercise

This article reviews the evidence that regular aerobic exercise counters the detrimental effects of psychological stress. The hypothesized psychological, psychophysiological, and physiological effects of exercise are summarized in [Figure 1](#). In order to protect against psychological stress and stress-related disease, exercise is hypothesized to have the following effects: (1) improve psychological coping with stress, (2) reduce physiological reactivity to stress, and (3) counteract the dysregulatory effects of stress reactivity on disease risk factors.

Psychological Coping with Stress

Most of the research on the beneficial psychological effects of regular aerobic exercise has concentrated on improvements in well-being, often defined as the absence of negative affect, or on improvements in self-esteem and self-efficacy. Reducing negative affect, e.g., anxiety, hostility, or depression, decreases the chance that an event will be perceived as stressful (primary appraisal). Increasing feelings of self-efficacy increases the chance that coping skills are perceived to be adequate to deal with a stressful situation (secondary appraisal). Well-being and self-esteem further modify interpersonal behavior such that less conflict is generated, and they favor adaptive problem-oriented coping and social support seeking over detrimental coping strategies such as denial and social isolation. Thus, exercise-induced improvements in psychological coping with stress may decrease the number of situations that are perceived to be stressful. This will decrease the frequency and duration of physiological stress reactions mediated by the autonomic nervous system. That is important because such stress reactivity, when often repeated, is thought to give rise to physiological dysregulation and ultimately disease (see [Figure 1](#)).

Regular exercisers have consistently been found to show the iceberg profile: they have higher vigor (the peak of the iceberg) and lower levels of anxiety,

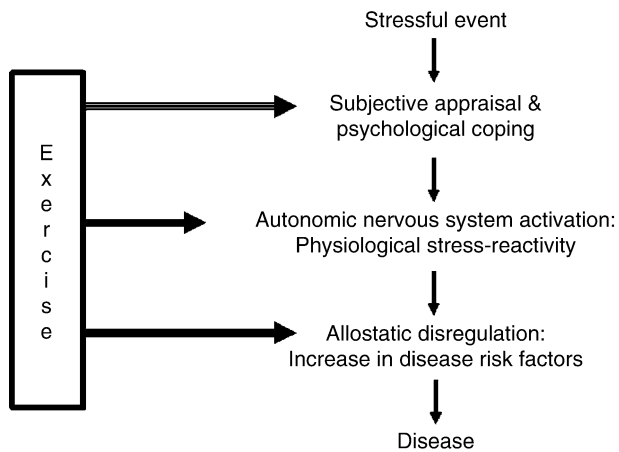


Figure 1 Hypothesized influences of aerobic exercise on the psychological and physiological effects of stressful events.

tension, apprehension, depression, and fatigue compared to nonexercisers. They are also found to be more outgoing, have greater self-esteem and self-efficacy, and score better on tests of coping skills. However, the coexistence of a healthy mind and a healthy body does not prove that the healthy mind is a consequence of the healthy body (such causality was never suggested by Juvenal, the Roman satirist who originated the adage of *mens sana in corpore sano* – he just prayed for both). The beneficial psychological makeup of exercisers in cross-sectional studies may be based on common underlying influences on both exercise behavior and psychological distress. These may be of a genetic origin, or they may reflect environmental influences such as low socioeconomic status and poor social support networks that affect both exercise behavior and psychological well-being. Most importantly, the association may reflect a reversed causality, namely, that emotionally well-adjusted, agreeable, and self-confident individuals with low levels of stress are simply more attracted to sports and exercise, or that only such persons have the necessary energy and self-discipline to maintain an exercise regime. In short, the favorable psychological profile of exercisers may reflect self-selection.

To prove that exercise can be used as a prophylaxis against stress and stress-induced disease, longitudinal training studies are needed. Such training studies must be carefully designed. Subjective expectations, distraction, social attention/interaction, and trainer enthusiasm must be separated from the effects of aerobic exercise per se. Therefore, the effectiveness of aerobic training must be compared to a placebo treatment, which may consist of stress management training, meditation, muscle relaxation, or a nonaerobic exercise program that aims to increase strength or flexibility rather than aerobic fitness. The subjects

must not be allowed to choose these treatments, but are randomly assigned to either aerobic training or control treatment to prevent self-selection of subjects with a specific psychological makeup or a relatively high aerobic endowment into the training group.

Several well-designed longitudinal studies have reported improved mood or reductions in coping behavior, depression, and anxiety after a program of aerobic exercise training in comparison to control manipulations. Unfortunately, many others have failed to replicate these training effects. The most promising evidence comes from studies in clinically depressed patients. These studies tend to report beneficial psychological effects of exercise that match or even exceed those of pharmacological treatment. Even in non-clinical populations, the training-induced changes in anxiety, depression, or self-esteem – when found – were generally largest in the subjects with the least favorable psychological profile before the start of the training. Without denying the potential of exercise as a therapy in subsets of subjects, most of the many reviews on this topic express only cautious optimism about the use of exercise for stress prevention in the population at large.

Psychophysiological Stress Reactivity

The various training-induced adaptations in the organization of the autonomic nervous system and its target organs, e.g., dominance of the parasympathetic over the sympathetic nervous system, greater sensitivity for hormonal cues (β -adrenergic receptors, insulin sensitivity), and better vascularization of muscle tissue, may all critically influence the pattern and intensity of physiological responses to stress, even if psychological coping with stress is unchanged. Although peak reactivity may not be influenced or even larger in exercisers, habituation during stress and recovery after stress may be faster, analogous with their physiological adaptation to bouts of exercise. An extensive set of studies has compared stress reactivity of well-trained exercisers with that of untrained nonexercisers. In about half of these studies, reduced stress reactivity and/or enhanced recovery was found in exercisers. The reverse was found only twice. Although these studies support the idea of lower stress reactivity in exercisers, they were mostly cross-sectional in nature. Consequently, self-selection factors and differences in endowment for fitness and/or psychological makeup, rather than exercise behavior, may explain the differences in reactivity.

Again, training studies yield the most meaningful answers. **Figure 2** provides an example of a typical design of a study to assess exercise training effects on physiological stress reactivity. In this study, four

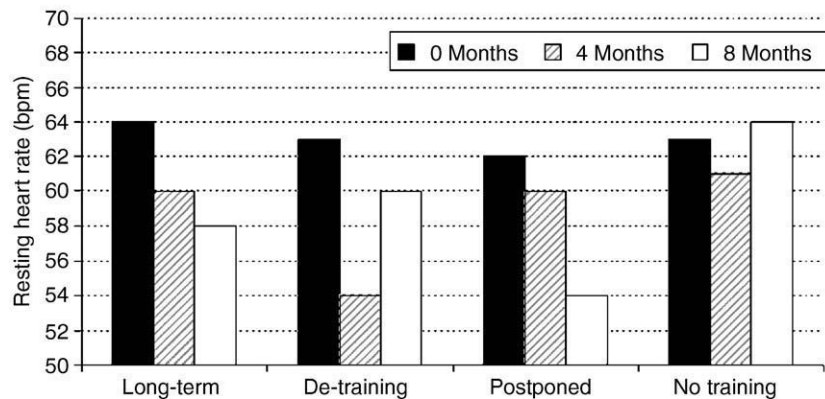


Figure 2 A design for a training study (de Geus et al. (1993) *Psychosomatic Medicine* 55, 347–363) on physiological stress reactivity. The effects of the experimental manipulation are illustrated by the effects on resting heart rate.

groups of adult males took part in an aerobic training program consisting of two to three sessions weekly over periods of either 4 or 8 months. Control subjects were put on a waiting list for 4 months or remained untrained entirely. Some of the trained subjects were de-trained; after an initial 4-month training program, they had to abstain from intensive exercise for 4 months. The figure illustrates how the various manipulations were closely followed by changes in resting heart rate. The effectiveness of aerobic training was further measured by the peak oxygen consumption during a supramaximal exercise test, expressed as milliliters oxygen consumed per kilogram body weight (VO_{2max}). VO_{2max} is generally held to be the best indicator of aerobic fitness and is highly correlated to actual endurance performance (e.g., Cooper test or shuttle run). The average improvement in VO_{2max} amounted to 14% after 4 months and 16% after 8 months, which is the gain usually attained in long-term and intensive training programs. In spite of improved fitness, this study found no evidence of an effect of aerobic exercise training on an extensive set of physiological responses that included heart rate and blood pressure, stress hormones, indices of cardiac sympathetic and parasympathetic drive, and indices of peripheral vascular versus cardiac output responding. Stress reactivity tended to decrease over repeated measurements, but the decreases in reactivity were not affected by aerobic training or ensuing de-training. These results correspond fully with those of many other studies showing that aerobic training does not systematically reduce reactivity to stress or enhance recovery afterward.

Acute Effects of Exercise

Taken the meager experimental support from well-designed training studies, how can we explain the

commonly held belief in the stress-alleviating effects of exercise and the fact that many exercisers state to be motivated by stress reduction and increased psychological well-being? The source of this belief may be largely found in the acute psychological effects of exercise. In the training studies described previously, psychological and psychophysiological testing was preferably done when subjects had not been exercising for a few days. This was done to prevent confounding of the results by possible acute effects of exercise. Studies that have specifically addressed these acute exercise effects have provided evidence of improved mood, reduced feelings of tension, anxiety, and anger, and increased feelings of vigor directly after exercise. This acute feeling better effect of exercise is paralleled by various physiological effects. Although heart rate generally remains elevated above the base level, postexercise blood pressure is seen to decrease below base level for several hours. Neuro-muscular activity and peripheral pulse volume are decreased, and brain alpha activity, an indicator of a relaxed but alert mental state, is increased. Most importantly, there is evidence of reduced heart rate and blood pressure responses to stress in the hours directly after the exercise.

Among the hypothesized mechanisms for the acute psychological and physiological effects of exercise, those related to the endogenous opioids have been most often cited. An endogenous opioid, beta-endorphin, has been held responsible for the euphoric (after) effect of running, which has been ominously termed runner's high, and for a benign endorphin-mediated exercise addiction. However, elevations in peripheral beta-endorphin (a typical distress hormone often co-released with adrenocorticotropin) are seen only during strenuous exercise, not during mild to moderate exercise. More importantly, beta-endorphin levels are associated with negative mood, i.e., expectation of

future pain and exhaustion, rather than positive mood, even during exercise. Finally, the opiate antagonist naloxone does not prevent the exercise-induced elevation in the pain threshold or mood improvement, suggesting that these are largely independent of endorphin action. Since monoaminergic systems in the brain are involved in the pathogenesis of depression and anxiety, a more plausible explanation for exercise-induced mood effects may reside in its effect on central monoamines. Animal models suggest that exercising increases noradrenalin and dopamine metabolism in the brain and that it also increases brain serotonin content. It shares these effects with antidepressive like MAO inhibitors and serotonin-reuptake blockers. Beneficial effects on monoaminergic neurotransmission may also explain why improved mood after training programs is most often found in groups of highly anxious or depressed subjects. An improvement in monoaminergic neurotransmission may be effective only if it is severely compromised, as is the case in anxiety and depression.

Whatever the physiological and psychological causes, the tight contingency between exercise and the mood improvement directly afterward forms a solid basis for the popular belief that exercise relieves stress. Exercising may be an excellent short-term coping strategy that helps people to unwind more rapidly from daily pressures experienced in the school, job, or home environment, even if effects on mood are short-lasting. Likewise, the accumulation of episodes of blunted cardiovascular responsiveness may reduce the health risks of stress, even if reductions in stress reactivity are confined to stressors encountered in the postexercise periods.

For exercisers, exercising may become a major coping strategy, up to a point where stopping exercise could lead to loss of well-being and self-esteem. In fact, most of the popular ideas on the psychological benefits of exercise may stem from exercisers themselves. It must be noted, however, that in most Western countries only 10 to 15% of the population are engaged in exercise with high enough intensity (>70% maximal heart rate), duration (minimally 20 min), and frequency (three times a week) to maintain aerobic fitness level. About 60% have irregular exercise habits or perform mild exercise only, e.g., regular walking or bicycling. The remaining 25% are nearly completely sedentary, and this figure appears to be rather resistant to 50 years of campaigning to increase physical activity. A major future research question is why exercisers exercise and why nonexercisers do not.

Determinants of Exercise Behavior

Most of the research on the determinants of exercise behavior has focused on personality and social and

environmental characteristics, but increasingly innate biological mechanisms are considered as additional factors influencing the choice for exercise behavior. Recent studies on genetically related subjects have confirmed a large genetic influence on the choice for leisure-time exercise behavior in adult subjects, with heritability estimates as high as 85%. The influence of the exercise genes starts to appear in late adolescence. Two potential pathways for genetic influences to affect adult exercise behavior may be through personality characteristics (e.g., high sensation seeking or extraversion and low neuroticism) or physical fitness parameters (e.g., muscle strength, muscle endurance, and aerobic power) that are known to be highly heritable. We add the third possibility – which is not necessarily independent of the other two pathways – that exercisers, based on their genetic receptiveness, extract a larger benefit from exercise in terms of mood improvements and reductions in psychological stress reactivity. Earlier we tacitly assumed that such effects are identical in all subjects, whereas in fact gene–environment interactions are highly likely to occur. Subjects who are genetically more favored by psychological and stress reactivity benefits may find it easier to become lifetime regular exercisers than those who do not experience similar benefits or even suffer from the added burden of having to do something that mainly feels unpleasant. This gene–environment interaction means that the group differences seen in cross-sectional comparisons of long-time exercisers with persistent sedentaries will not be recaptured by the pre- to posttraining differences found in experimental training studies. The group of long-time exercisers will be biased to include only the most receptive subjects. *Mutatis mutandis*, it also means that true beneficial effects for genetically receptive subjects in the studies on psychological coping and psychophysiological stress reactivity cited previously may have been snowed under by null (or negative) effects in those subjects who are genetically less receptive (or even vulnerable).

Exercise and Stress: Opposing Effects on Disease Risk Factors

A final interaction between stress and exercise can be found at the level of disease risk factors, where exercise may systematically oppose the physiological impact of repeated stress reactivity. Evidence for such an interaction comes from the field of cardiovascular disease. Chronic stress is thought to negatively influence a set of atherogenic factors such as high low-density lipoprotein (LDL) (bad) cholesterol, low high-density lipoprotein (HDL) (good) cholesterol, high triglycerides, high insulin, high heart rate and

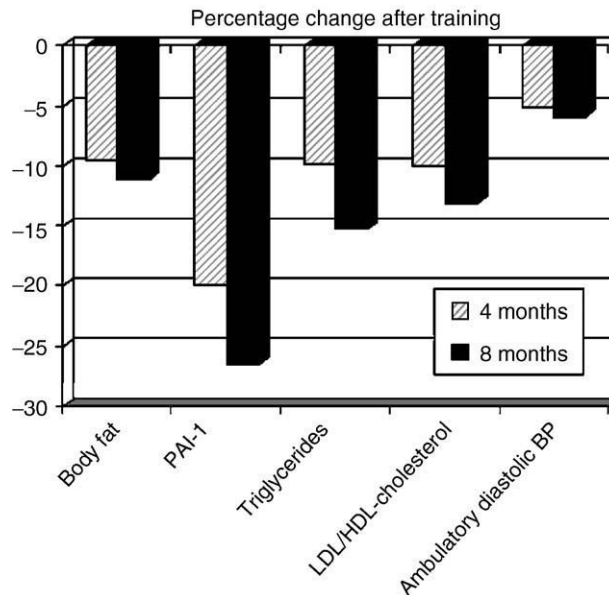


Figure 3 Beneficial effects of 4 and 8 months of aerobic exercising on a cluster of risk factors that are suspected to be affected by psychophysiological stress reactivity: percentage body fat; plasminogen activator inhibitor (PAI-1) activity, a risk indicator for disturbances in the coagulation–fibrinolysis balance; triglycerides; ratio of LDL (bad) to HDL (good) cholesterol; and ambulatory diastolic blood pressure.

blood pressure, and a disturbed balance in the coagulation and fibrinolysis of the blood. This cluster of risk factors is a common clinical finding that is not constituted at random but is the multifaceted symptom of a basic abnormality, called insulin resistance syndrome, rooted in overweight and inefficiency of the (muscular) tissue to take up glucose. These risk factors affect the integrity of the arterial walls in synergy, such that stress-induced increases in levels in one of these parameters increase the pathogenic effect on all of the others.

A few hours of exercise per week is known to be able to counter all these detrimental effects. Although no effects on stress reactivity (task minus baseline) are found, virtually all training studies show a decrease in absolute levels of heart rate and diastolic blood pressure during baseline and stress conditions. This effect is completely reverted after de-training, which strongly argues for a causal effect of exercise. Ambulatory studies have confirmed that the significant decrease in heart rate and blood pressure remains intact in naturalistic situations. In addition, regular exercise reduces the LDL/HDL cholesterol ratio and triglyceride levels, increases fibrinolytic potential of the blood, and does so most likely through preventive maintenance of muscular insulin sensitivity. [Figure 3](#) shows some of the simultaneous beneficial effects on various risk factors using data from the 4- and 8-month training programs described in [Figure 2](#).

In conclusion, aerobic exercise should be seen as an excellent way to compensate for or prevent

stress-related deterioration of cardiovascular health, even if psychological makeup or the acute reactivity to stress is unchanged after months of training. These health benefits suffice to strongly encourage the promotion of aerobic fitness programs to compensate for the aversive effects of daily stress.

See Also the Following Articles

Beta-Endorphin; Exercise; Opioids.

Further Reading

- Babyak, M., Blumenthal, J. A., Herman, S., et al. (2000). Exercise treatment for major depression: maintenance of therapeutic benefit at 10 months. *Psychosomatic Medicine* 62(5), 633–638.
- Bouchard, C., Shepard, R. J. and Stephens, T. (1994). *Physical activity, fitness and health*. Champaign, IL: Human Kinetics.
- De Geus, E. J. C., van Doornen, L. J. P. and Orlebeke, J. F. (1993). Effects of aerobic fitness and regular exercise on psychological make-up and physiological stress-reactivity. *Psychosomatic Medicine* 55, 347–363.
- De Geus, E. J. C., Boomsma, D. I. and Snieder, H. (2003). Genetic correlation of exercise with heart rate and respiratory sinus arrhythmia. *Medicine and Science in Sports and Exercise* 35(8), 1287–1295.
- Dishman, R. K. (1997). Brain monoamines, exercise, and behavioral stress: animal models. *Medicine and Science in Sports and Exercise* 29(1), 63–74.

- Dunn, A. L., Trivedi, M. H., Kampert, J. B., Clark, C. G. and Chambliss, H. O. (2005). Exercise treatment for depression: efficacy and dose response. *American Journal of Preventive Medicine* 28(1), 1–8.
- Gauvin, L. and Spence, J. C. (1996). Physical activity and psychological well-being: knowledge base, current issues, and caveats. *Nutrition Reviews* 54(4), S53–S65.
- Gauvin, L., Rejeski, W. J. and Norris, J. L. (1996). A naturalistic study of the impact of acute physical activity on feeling states and affect in women. *Health Psychology* 15(5), 391–397.
- Lawlor, D. A. and Hopker, S. W. (2001). The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomized controlled trials. *British Medical Journal* 322, 763–767.
- Moore, K. A. and Blumenthal, J. A. (1998). Exercise training as an alternative treatment for depression among older adults. *Alternative Therapies in Health and Medicine* 4(1), 48–56.
- Salmon, P. (2001). Effects of physical exercise on anxiety, depression, and sensitivity to stress: a unifying theory. *Clinical Psychology Review* 21(1), 33–61.
- Soenstroem, R. J. and Morgan, W. P. (1989). Exercise and self-esteem: rationale and model. *Medicine and Science in Sports and Exercise* 21(3), 329–337.
- Stubbe, J. H., Boomsma, D. I. and De Geus, E. J. (2005). Sports participation during adolescence: a shift from environmental to genetic factors. *Medicine and Science in Sports and Exercise* 37(4), 563–570.
- Yeung, R. R. (1996). The acute effects of exercise on mood state. *Journal of Psychosomatic Research* 40(2), 123–141.

Affective Disorders

D F MacKinnon

Johns Hopkins University School of Medicine,
Baltimore, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition
article by D F MacKinnon, volume 1, pp 87–94,

© 2000, Elsevier Inc.

Phenomenology of Affective Disorders
The Experience of Affective Disorders
Behavior in Affective Disorders
Affective Disorders and Personality
Varieties of Affective Disorder
History of the Concept of Affective Disorders
Affective Disorders as Diseases
Affective versus Other Psychiatric Disorders
A Note on Nomenclature

Glossary

<i>Affect</i>	Objective manifestations of emotional state, including facial expression, demeanor, neurovegetative functioning, and behavior, but also verbal statements reflecting mood, self-attitude, and vital sense.
<i>Disease</i>	A clinical syndrome with characteristic signs and symptoms, related to dysfunction in a bodily organ, caused at least in part by biological forces.
<i>Mood</i>	Sustained subjective emotional state.

Neuro-vegetative functioning
Psychosis

Motivated behavior essential to biological functioning, including sleep, appetite, and libido.

Severe symptoms of mental illness involving dysfunction in rational thought and reality testing, including hallucinations and delusions.

Self-attitude

Feelings of personal worth, ranging from grandiosity to self-loathing.

Vital sense

Feelings of physical well-being, including energy, clarity and speed of thought.

Affective disorders are psychiatric syndromes characterized by significant disturbances in mood, the sense of self-worth, and the feelings and functions of physical well-being. Affective disorders can lead to occupational and interpersonal impairment and often (perhaps not often enough) the pursuit of treatment. Suicide is an all-too-common outcome of a severe affective disorder. Major depression, according to the United Nations World Health Organization, is the second leading cause worldwide of lost years of health and productivity. Mood swings routinely wreck marriages and careers and trigger and sustain substance abuse. The clinical features, management, and biology of mania and major depression are described elsewhere in this work; therefore, the present article examines the concept of affective disorders from the multiple perspectives of phenomenology, meaning, behavior, and temperament to complement the growing biological understanding of these disorders.

Phenomenology of Affective Disorders

The clinical syndrome of all affective disorders is a persistent and pervasive disturbance in mood, self-attitude, vital sense, and neurovegetative functioning. In depression, the varieties of mood disturbance include not only melancholia or sadness, but also apathy, anxiety, and irritability. Some patients describe depression as the absence of any feeling. The mood of a manic patient may be elated, but also irascible to the point of belligerence, expansive without a sense of euphoria, or simply excitable.

Self-attitude, the feeling of personal worth, is insidiously disturbed in affective disorders. The disturbance in self-valuation often impedes treatment. Depressed patients may delay treatment because the suffering they experience seems somehow deserved; ultimately suicide may be desired because the patient feels that he or she is such a burden on others. Manic patients may feel smarter than the doctor and others, and may feel misunderstood or insulted by the suggestion they are ill.

Affectively ill patients have a disturbance in the sense of physical well-being, or vital sense. In depressed patients, physical well-being is lacking; patients feel tired, weak, slow, and inept; cognition is halting or uncertain. Manic patients tend to feel energized, quick, tireless, and powerful. Thoughts move swiftly; the patient may have the sense that he or she is thinking circles around others, while others may find the patient's ideas loose, connected by the thinnest of associations – clanging speech, riddled with puns and rhymes, is seen in severe cases – or at the most severe, incomprehensible. Neurovegetative functioning – the body's internal regulation of activity and metabolism – is typically disturbed in both syndromes. Depressed patients may have insomnia or may sleep to excess without feeling rested. Manic patients may feel little to no need to sleep and may arise rested after a few hours and carry on in this manner to the marvel (or annoyance) of others. Appetites tend to diminish with depression (though some will overeat or binge); in mania the drive for food is less often disturbed, but the drives for sex, acquisition, and excitement are enhanced.

Affective disorders are diagnosed from clinical history and observation of the patient's demeanor and behavior. The clinician is more certain of the diagnosis upon hearing the patient's own description of troubling mental phenomena and personally observing signs (e.g., emaciation, lethargy, agitation, rapid speech) consistent with the clinical syndrome. There are no laboratory tests to reveal affective disorder.

The Experience of Affective Disorders

Literature abounds with poignant and gritty descriptions of what it is like to be manic or depressed. A pervasive theme in these descriptions is the relative inability of a person in the grip of a severe mania or depression to perceive that the accompanying attitudes, responsiveness, and emotional states are pathologically disconnected from or disproportionate to the true circumstances. A depressed mood, sleepless nights, lack of appetite, feelings of worthlessness, and wishes for death often seem to a depressed patient to follow inevitably an unwanted change in life such as a divorce or job loss. A profound sense of meaninglessness is attributed to the world, rather than to the patient's altered perception of the world. The elated mood in mania seems to the inexperienced patient to be the natural consequence of possessing superhuman energy and confidence. The heightened sense that everything is possible seems a quality of the world, in the eyes of the manic patient, and not a distortion driven by an abnormal mood.

Insofar as depressive and some manic states are unpleasant, patients generally understand them to be a response to stress, with the implied assumption that they have had a nervous breakdown because they were unable to handle the stress. On being diagnosed with an affective disorder, many patients are relieved to hear their distress attributed to a medical illness or chemical imbalance in the brain, especially when told of the successful record of biological treatment. Patients with repeated episodes of mania and depression eventually come to see the disease as a primary problem, rather than a secondary result of life circumstances. It begins to seem more plausible that, while some episodes may certainly have followed a significant change in life circumstances, others were unprovoked, or even preceded a change in circumstances; indeed, it may occur to the experienced patient that the changes were in some cases caused by the affective disorder. Was a patient's emotional state so inconstant that romantic partners found the relationship intolerable? Did a drop in energy and motivation result in an unfavorable change in job status, which heralded a suicidal episode? Did a patient feel elated because of a vacation in the tropics, or did the motivation to travel stem from a markedly heightened capacity for excitement, pleasure, and financial largesse?

Thus, the general question of whether stress causes affective instability or affective instability causes stress cannot be answered definitively by asking the patient, as the illness distorts the patient's ability to tell the difference. This is one fundamental difference between

an affective disorder and a disorder of the visceral or peripheral organs: in the latter case, the intact mind, seated in the brain, receives unambiguous information from another organ about a disease process, e.g., pain, shortness of breath, weakness. With affective disorders, the major symptoms are in the mind, where they impede perceptions and judgments.

Behavior in Affective Disorders

From the perspective of a family member or treating clinician, the most salient features of an affective disorder are the behavioral difficulties it drives. Most alarmingly, affective disorders often drive self-destructive behavior. The vast majority of suicide victims can be diagnosed, in retrospect, with affective disorder. When the suicidal behavior does not result in death, it nevertheless may permanently strain relationships with those the patient intended to leave behind.

Suicide is the direst behavioral outcome of affective disorder, but not the only significant maladaptive behavior. At the core of maladaptive, affectively driven behaviors is the change in hedonic capacity that typically accompanies them. Where reward is absent, as in the anhedonic, depressed patient, experiences may seem flat and meaningless, and in many cases painful and enervating. A depressed patient may withdraw to bed, avoid family, friends, and work, and seek solace in alcohol or drugs. When reward comes too easily or powerfully, as in mania, the impulse to spend, drive, and love recklessly, to flout the law and social mores, and to abuse alcohol and drugs is nearly irresistible. For the manic patient there is an indiscriminate sense of excitement over a world of newly perceived opportunities for enjoyment, resulting in costly, uninhibited, social and financial decisions.

It is important to note that affective disorders are not essential in such behaviors. Many people who attempt suicide, avoid work, spend recklessly, or drink heavily lack the signs and symptoms of affective disorders. Many patients with affective disorders (perhaps most) manage, somehow, to constrain their behavior. Patients who change their behavior in the midst of an episode of affective disorder may benefit from confrontation and counseling about the consequences of their behavior. Some therapists think of the self-defeating and self-deprecating thoughts and attitudes that arise during depression as behaviors amenable to confrontation and redirection; this idea is the basis of cognitive therapy.

The degree of behavioral impairment associated with affective disorder can be a function not only of the severity of the symptoms, but also of the vulnerability of the patient to behavioral excess or instability.

A libertine when stable is more likely than a saint to engage in reckless acts when affectively ill, if only because the means to such acts are more familiar. An academically and economically marginal college student knocked out of school by affective illness may have a harder time getting back on track than a middle-class mid-career middle manager in a solid marriage. A student not only lacks the benefit of a structured life upon which to recover, but may also have the option unavailable to the older person of re-entering a state of dependency on parents, a state that may reinforce functional debilitation. Patients and families are wisely counseled not to make major changes in career, family, or housing in the midst of an episode of mania or depression. The fewer the changes, the less the risk of serious impairment, and the easier the rehabilitation once the episode clears.

Affective Disorders and Personality

Greater symptom severity may drive one patient's behavior to go awry while a less afflicted patient remains relatively steady, but personality and life circumstances also contribute. Some affectively ill patients prone to dysfunctional behavior may have personality disorders at baseline. Personality disorder refers, in diagnostic nomenclature, to enduring patterns of temperamental vulnerability and problematic behavior. These disorders may themselves be defined by their affective components, e.g., the labile affect of the borderline patient or the constricted affect of the schizoid patient. Thus, it can be hard to tell if a patient who seems to meet diagnostic criteria for a personality disorder really has an affective disorder. Patients in a manic state often display labile affect and indifference to the feelings of others (and other borderline or narcissistic traits), and depressed patients often appear cold and constricted. Thus, it is generally prudent to wait to assign a definitive personality disorder diagnosis to an affectively ill patient until after the affective syndrome clears up and the patient can be observed in the well state.

The concept of affective disorder also becomes salient when considering individuals who seem characterologically gloomy, moody, or excitable. Although they are counted as having mild forms of affective disorder, the existence of chronically gloomy (dysthymic) and chronically moody (cyclothymic) individuals illustrates the permeable boundaries between entities considered as medical disorders versus those understood as features of character or temperament. Indeed, affective and personality disorders may not be fully independent in their origins. It seems plausible to assume that childhood or adolescent onset of affective disorder is likely to result in persistent – even

permanent – changes in the direction of personality development. Whereas affectively stable individuals learn through experience to properly associate their behaviors with desired social and interpersonal consequences, the experience of affective disorder, which warps these affective connections during critical stages of adult development, may lead to the sort of immature emotionality and self-defeating behavior that characterize the personality disorders.

Varieties of Affective Disorder

There are many ways to categorize affective disorders. The most salient clinical distinction is between manic-depressive and melancholic (depressive) disorders. Modern nosology refers to these disorders as bipolar disorder and major depressive disorder. Beyond the differentiation by polarity, affective disorders are classified according to severity and chronicity. [Figure 1](#) illustrates the spectrum of affective disorders, ranging from mild, transient normal mood fluctuations to severe, rapid-cycling bipolar disorder type I.

In clinical reality, the differentiation of mania from depression is not always so clear. Many patients experience either a mixture or the rapid alternation of manic and depressive symptoms. One way to understand these mixed states (following Kraepelin) is to postulate that there are different components of affect (we might call them mood, cognition, and energy) that tend to cycle in synchrony during typical manic (high mood, rapid and copious thoughts, high energy) and depressive (low mood, sluggish and paltry thoughts, diminished energy) episodes, but that synchrony is lost in mixed states ([Figure 2](#)). A patient in a common presentation of mixed state will tend to have a dysphoric (irritable, nihilistic) mood but such racing thoughts and restless energy that the patient will appear irascible and feel highly uncomfortable. In practice, the mixed state is seen as a high-risk state for suicide. In contrast, a patient with agitated depression combines dysphoric mood with restless energy and a poverty of thought, generally manifested as a tendency to ruminate on a single, depressive theme. Other combinations, in particular those involving euphoric mood without high

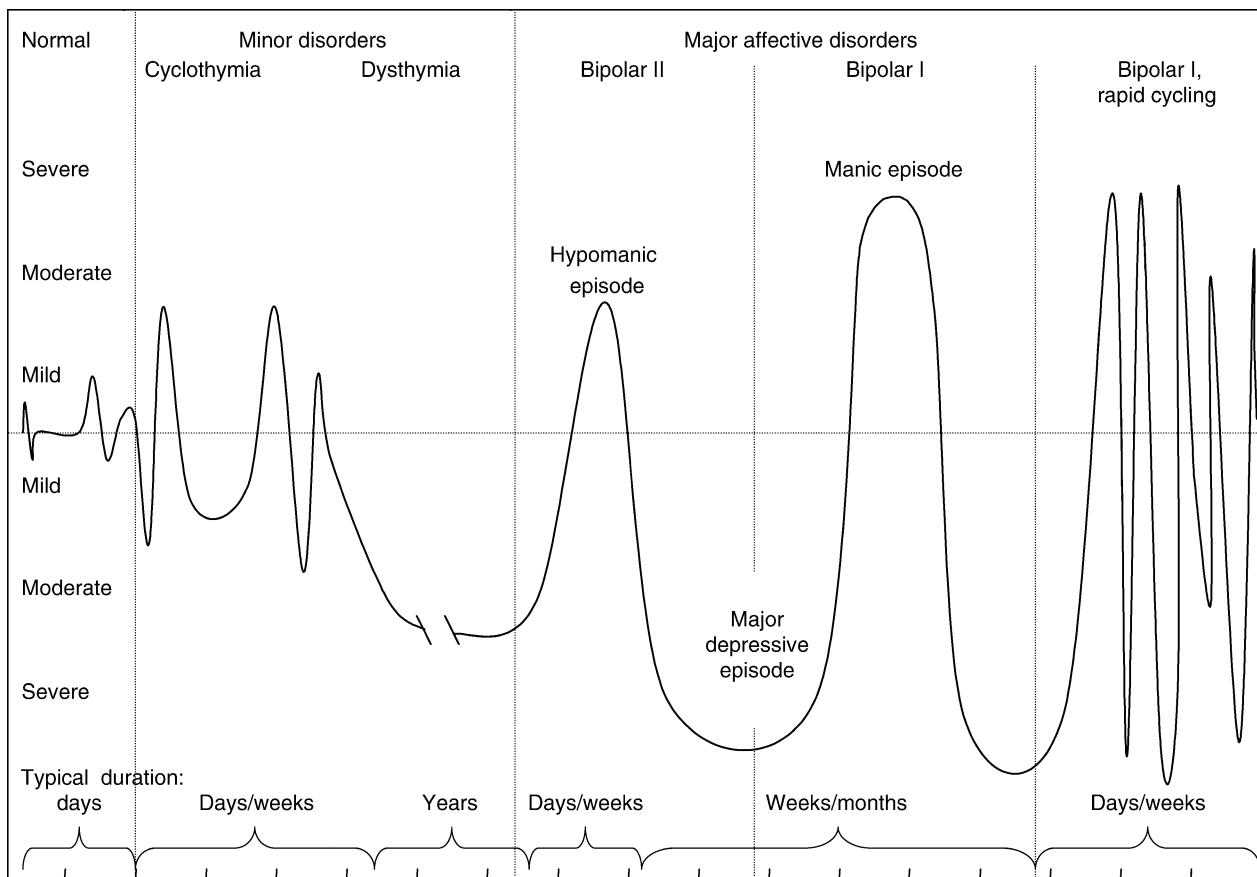


Figure 1 Affective disorder diagnoses as a function of the frequency and amplitude of mood variability. Note that normal mood changes may be frequent but are mild and transient, while the more severe mood swings tend to last weeks, months, and in the case of dysthymia, years.

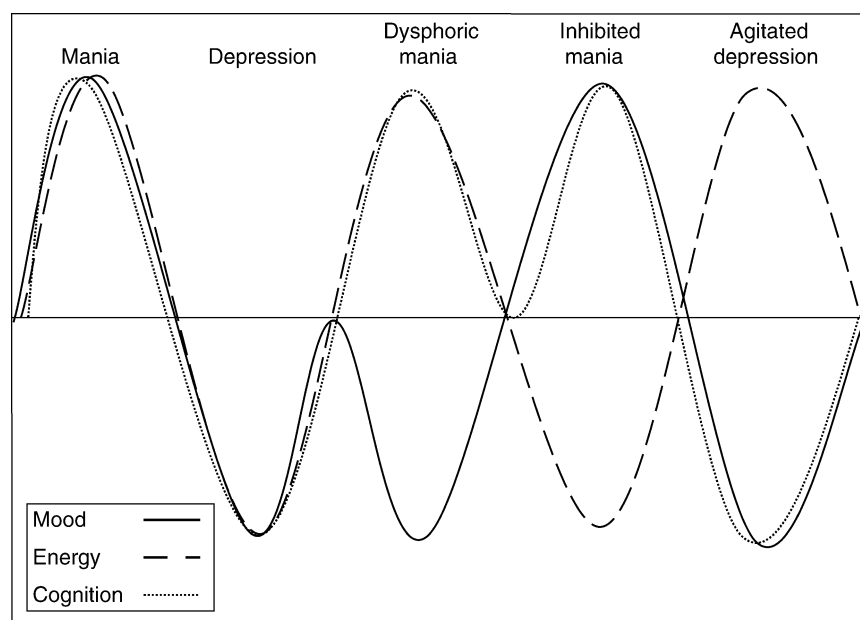


Figure 2 Kraepelin's schema of mixed affective states. The coexistence of a morose mood (low emotion) with physical restlessness or agitation (high volition) and racing thoughts (high intellect) creates a typical mixed state.

energy, seem more rare, though it may be that these patients are rarely brought to treatment.

History of the Concept of Affective Disorders

Descriptions of melancholia and mania date back to Hippocrates, and many medical writers throughout history were able to describe the manic-depressive syndrome in a way familiar to modern psychiatrists. Until the nineteenth century, the prevailing explanations for affective disorders relied on the humoral theory. Abnormal balances of bodily fluids such as black bile (the roots for the term melancholia) or yellow bile (an excess of which was associated with the choleric, or irritable, temperament) were felt to be the cause of affective imbalance.

By the nineteenth century, ideas about the organic basis of disease had become more sophisticated, and observations of psychopathology were made systematically with the view that affective disorders had their basis in an as-yet-unknown disease of the brain. An exemplar of this trend was Emil Kraepelin, who delineated the patterns of symptoms and course in the institutionalized severely mentally ill under his care. Kraepelin noted a general difference in the course and affective symptomatology in the two groups of patients: those having a remitting/relapsing forms of psychosis marked by prominent signs of affective disturbance, and those having a chronic psychosis with affective flattening. Kraepelin called the remitting/relapsing disorder manic-depressive insanity and the chronic

psychotic conditions paranoia (including dementia praecox, now better known as schizophrenia).

Psychiatric epidemiologists have established that affective disorders are not only found in the severely, institutionalized mentally ill, but also are fairly common in the general population. About 1 in 100 individuals in population surveys meet diagnostic criteria for bipolar disorder, and 5–10% have a major depressive episode at some time in life. Many patients with affective disorders do not develop psychosis or require hospitalization; a rather large percentage never enters treatment.

Historical accounts of the obviously manic and melancholic are not too hard to find; it is harder to account for the large number of those who, it is supposed, experienced milder forms of mania and depression but understood these experiences in non-medical terms. Traces of these experiences perhaps are found in the annals of art, literature, politics, business, philosophy, exploration, science, and religion, where a mild depression in a receptive mind may produce enduring insights about the fragility of existence or tenderness to the suffering of others. Mild manic states may have contributed to expansive works of art or ambitious, risky exploration and entrepreneurship. Affectively colored cultural contributions are not necessarily pathological, though the experience of the passionate, creative, or daring historical figure might resemble in detail the experience of a typical patient with an affective disorder. It is just as possible that the availability of effective anti-depressive and antimanic treatments in times past

might have robbed the world of some works of genius as it is possible that the geniuses themselves might have been spared much suffering and contributed even more had treatment been available.

Affective Disorders as Diseases

Efforts to link affective disorders to external circumstances have yielded distinctions between mourning and melancholia (Freud), and reactive versus endogenous depression. But if it is unclear whether stress causes depression or vice versa, then a simpler, assumption-free explanatory model is required if we are to understand the normal dejection that results from, say, a romantic setback, versus the prolonged sadness that accompanies the grieving process, versus a pathological affective disorder. Unlike emotional reactions, affective disorders are clinical syndromes with stereotypical patterns of sign, symptoms, course, and treatment response. Unlike grieving, affective disorders are associated in some cases with known brain disease. Affective disorders are best understood as medical diseases, i.e., as syndromal entities emerging from biological dysfunction in the brain and body.

Like other medical diseases, affective disorders are not limited by culture, intelligence, education, socioeconomic status, or experience. Nor do the symptoms of affective disorder vary much across cultures, although one might expect to see culturally and individually specific manifestations of these symptoms. The typical symptoms of affective disorder cluster over a similar course and respond to the same biological treatments anywhere in the world. Affective disorders, moreover, may arise from known medical causes. Huntington's disease, a genetic disorder of progressive dementia and abnormal movement, carries a high risk for manic-depressive disorder. Some patients develop their first depression after a stroke (particularly a stroke in the left frontal part of the brain). Certain medications can cause a clear depressive syndrome; knowledge of the pharmacology of such drugs spawned research into medications to reverse the depressive syndrome. Cortisol-like drugs (e.g., prednisone) frequently induce manic symptoms in people with no history of affective disorder. If such known changes to the anatomy and biochemistry of the brain can trigger the characteristic signs and symptoms of affective disorder, then it seems reasonable to conclude that the commonly observed affective disorders arise from biological mechanisms as yet unknown, but knowable.

In contrast, clinical problems related to mood alone are strongly culture bound and uniquely expressed, generally have few if any symptoms other than a bad mood, and cannot be mimicked by drugs

or other medical diseases. Unpleasant or unmanageably intense mood reactions of this sort are best viewed primarily as discouragement, demoralization, the blues, or in the official nomenclature, adjustment disorders. Grief, on the other hand, has a syndromal quality to it, with characteristic sensations: the welling-up feeling, the false recognition of the departed in a crowd. Uncomplicated grief is not a disease, of course, but a normal part of human experience: painful, but rich with personal meaning. In real clinical situations, it is not always simple to distinguish grieving from moping from depression, but it is essential to try.

Affective versus Other Psychiatric Disorders

Patients with affective disorder often seem superficially to have other sorts of psychiatric illness, and vice versa. While affectively ill patients may demonstrate cognitive dysfunction (poor attention and concentration in depression, disorganization and distractibility in mania), mania and depression are not a cognitive disorder, like dementia, since the cognitive difficulties tend to clear up with normalization of the mood. In contrast to a chronic psychotic disorder such as schizophrenia, psychotic symptoms in patients with affective disorder tend to occur only when the patient is in the grip of severe mania or depression. In practice, however, the boundaries of affective and psychotic disorders are not always well defined. Some believe that schizophrenia and bipolar disorder have the same cause, with schizoaffective disorder as an intermediate type; it seems more likely that both disorders have complex causes and that some of these causes affect the presentation and course of illness in either disorder. Anxiety disorders, too, may be related to affective disorders; indeed, anxious moods, phobic avoidance, panic disorder, and obsessive-compulsive disorder are seen frequently in patients with affective disorder. Anxiety disorders share some of the characteristic symptoms of affective disorders: disturbances of mood and physical well-being, but miss the typical affective disturbances in self-worth, sleep, and appetite.

In summary, affective disorders are fairly common and often severe medical diseases afflicting mood, self-attitude, vital sense, and the bodily functions of sleep, appetite, and libido. Some patients exhibit marked changes in behavior, while others maintain the appearance of normality despite their symptoms. Although a patient with an affective disorder (and loved ones and caregivers) will often search for a meaningful explanation for the symptoms, the source of the problem is likely to be in the functioning of the brain. Stress may trigger the onset of an affective

disorder in a vulnerable individual; affective disorders certainly cause stress. Patients with affective disorders require not only biological treatment of the disease, but also management of self-destructive behavior and rehabilitation once the episode clears.

A Note on Nomenclature

Strictly speaking, there is no longer a universally accepted set of “affective disorders”. Earlier generations of psychiatrists spoke of manic-depressive reactions, depressive neuroses, and affective psychoses, which melded into the category of affective disorders in the *International Classification of Diseases*, 9th edition (ICD-9), in 1975, and in the *Diagnostic and Statistical Manual*, 3rd edition (DSM-III) in 1983. Subsequently, although the list of disorders included in the category remained more or less constant, the labeling in both classification schemes changed from “affective” to “mood” disorders in their next editions. The significance of this change in terminology is hard to estimate, as the definitions of affect and mood vary from source to source. Taking the prevailing view that affect encompasses mood along with its various associated qualities, the term affective disorders is preferred in a discussion of the conceptual foundations of these disorders, as it seems more descriptive. All of the disorders counted as mood disorders in current diagnostic schemes are disorders of much more than the prevailing emotional state, while in any given patient, change in the mood itself may be a minor

problem compared to changes in energy, motivation, and neurovegetative functioning.

See Also the Following Articles

Depression and Manic-Depressive Illness; Depression Models; Negative Affect; Suicide, Psychology of; Suicide, Sociology of.

Further Reading

- DePaulo, J. R. and Horvitz, L. A. (2002). *Understanding depression: what we know and what you can do about it*. Hoboken, NJ: John Wiley & Sons.
- Goodwin, F. K. and Jamison, K. R. (1990). *Manic depressive illness*. New York: Oxford University Press.
- Jamison, K. R. (1993). *Touched with fire: manic-depressive illness and the artistic temperament*. New York: Free Press.
- Jamison, K. R. (1995). *An unquiet mind*. New York: Vintage Books.
- Kraepelin, E. (1990). *Manic depressive insanity and paranoia* (reprint edn.). Salem, NH: Ayer Company.
- McHugh, P. R. and Slavney, P. R. (1998). *The perspectives of psychiatry* (2nd edn.). Baltimore, MD: Johns Hopkins University Press.
- Mondimore, F. M. (1999). *Bipolar disorder: a guide for patients and families*. Baltimore, MD: Johns Hopkins University Press.
- Robins, L. N. and Regier, D. A. (1991). *Psychiatric disorders in America: the epidemiologic catchment area study*. New York: Free Press.
- Styron, W. (1992). *Darkness visible*. New York: Vintage Books.

Aggression

E F Coccaro and E C Manning

University of Chicago, Chicago, IL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by E F Coccaro, volume 1, pp 95–99, © 2000, Elsevier Inc.

Types of Aggression

Epidemiology of Aggression

Genetics of Aggression

Psychosocial Factors in Aggression

Biology of Aggression

Psychopharmacology of Aggression

Glossary

Antisocial behavior

Behavior in which the rights of others are violated; includes behavior that provides grounds for arrest, deceitfulness, failure to honor obligations, lack of remorse, and irritability and aggressiveness against people and property (i.e., vandalism).

Criminal behavior

Behavior in which a crime is committed, usually behavior that is identified by the criminal justice system.

Delinquent behavior

Antisocial behavior seen in children and adolescents, including aggression and other dysfunctional behaviors (e.g., school truancy).

<i>Epidemiological survey</i>	A study in which a representative sample of a population of individuals is studied to estimate the number of individuals from that general population that have a certain condition or diagnosis.
<i>Functional magnetic resonance imaging (fMRI)</i>	fMRI is the abbreviation for functional magnetic resonance imaging. fMRI allows imaging of brain-related activity in response to human subjects involved in cognitive or emotional tasks.
<i>Genetic studies</i>	Studies that are designed to estimate whether a condition or disorder has an underlying genetic component. These include family studies, which can show whether a condition runs in families; twin studies, which can show whether a condition is more likely to occur in identical than in fraternal twins and therefore must have a genetic component; and adoption studies, which can show the separate and interactive contributions of environmental and genetic factors.
<i>Heritability estimate</i>	A mathematical estimate of the magnitude of the importance of the role of genetic factors in a condition or disorder.
<i>Neurochemistry</i>	The aspect of biological functioning related to chemicals that transmit neuronal signals from one nerve to another.
<i>Norepinephrine (NE)</i>	NE is the abbreviation for norepinephrine, one of the main monoamines in the brain. NE is widely distributed in the brain and is involved in mediating a variety of brain and behavioral functions including, but not limited to, arousal.
<i>Problem solving</i>	The ability to find solutions to everyday types of problems.
<i>Psychosocial influences</i>	Potential influences on behavior (or on conditions/disorders) that are not apparently due to genetic or biological factors.
<i>Serotonin (5-hydroxytryptamine, 5-HT)</i>	5-HT is the abbreviation for serotonin, one of the main monoamines in the brain. 5-HT is widely distributed and is involved in mediating a variety of brain and behavioral functions including, but not limited to, behavioral inhibition.
<i>Transgenerational studies</i>	Studies comparing different generations (e.g., grandparents, parents, and offspring).
<i>Violent behavior</i>	Behavior in which physical aggression in some form is present.

Types of Aggression

Aggression can be defined as intentional act committed by an individual that has the potential to result in the physical or emotional harm of a person or of an object. Although aggression is typically thought of

as constituting a physical assault, acts of aggression also include verbal outbursts in which the individual's voice is used to convey anger in an uncontrolled manner. Broadly speaking, aggression can be subdivided into two types: impulsive aggression and premeditated or nonimpulsive aggression. Aggression is impulsive when the aggressive act is unplanned or in response to an aversive stimulus, generally uncontrolled, and committed without the expectation of achieving a tangible objective. By contrast, premeditated aggression is controlled, involves planning, and is associated with an expectation of reaching a specific goal such as, for example, the intimidation of the victim. Although acts of impulsive aggression can sometimes be criminal in nature, premeditated aggression is more likely to occur in the context of other criminal behavior. Because impulsive, rather than premeditated, aggression appears to be associated with specific biological and pharmacological response characteristics, this review focuses on impulsive aggression.

Epidemiology of Aggression

Despite the clinical importance of impulsive aggression, there is very little epidemiological data specifically regarding impulsive aggressive behavior. Existing epidemiological data exclusively refer to homicide and physical assault (i.e., aggression in general). Currently, the age-adjusted rate for homicide in the United States, based on vital statistics records, is 0.01%. Epidemiologic survey data find that approximately 25% of all males report a history of some physical fighting since 18 years of age, a rate that is twice as high as is found in adult females. Accordingly, approximately 10–15% of the general population report that they have engaged in physical fighting as an adult. This proportion translates into at least 25.0–37.5 million individuals in the United States alone. Although we cannot know what proportion of homicides or physical assault were impulsive in nature, it is likely that a substantial proportion represent impulsive aggressive acts.

Genetics of Aggression

Data from twin, adoption, and family studies support the hypothesis that aggression is under significant, if variable, genetic influence. The differences seen in these results are due, most likely, to the fact that different studies examined aggression in different ways using differing assessment measures. Specifically, aggression has been examined in these types of studies by using categories that more or less reflect aggression (such as delinquent behavior, antisocial behavior, criminal behavior, and violent behavior)

in populations with psychiatric diagnoses or by using scales that more or less reflect the severity of aggression in the general population or in university students. In clinical populations, genetic influences appear to underlie both delinquent behavior and antisocial behavior. In general populations, heritability estimates for measures of aggression range from substantial in children to moderately substantial in adults (44–72%) to apparently nonexistent in adolescents. Most noteworthy is the observation that the aggression measures with the greatest heritability estimates reflect anger and hostility and/or anger, impulsiveness, and irritability. These are the type of behavioral traits that appear to be associated with impulsive aggressive behaviors and that, in turn, have been shown to be strongly correlated with biological factors and with psychopharmacological responses to treatment with anti-aggressive agents.

Psychosocial Factors in Aggression

The most important psychosocial factors involved in the development of aggression appear to be low socioeconomic status (SES), ineffective parenting style, physical punishment during childhood, and exposure to aggression in- and outside the family. Factors involved in maintaining aggression appear to be the poor development of problem-solving skills, a bias to attribute hostile intent to others, and a deficiency in appropriately detecting environmental and social cues.

Low Socioeconomic Status and Parenting Style

Low SES has been linked to the development of aggression in numerous studies. However, it is likely that low SES exerts its effect on aggression through more relevant behavioral factors such as being raised in a criminally violent family and exposure to gang activity and violence. The stress associated with poverty is associated with negative parental interactions and, consequently, parental irritability, a factor that was associated with aggressiveness in children. More specifically, marital problems in low-SES families has been linked to aggression in children. In addition, an association between lack of maternal responsiveness (and poor child rearing) in low-SES homes and aggressiveness has also been reported, a finding consistent with earlier reports showing that nonnurturing parents tend to have aggressive children. A critical predictor in this regard appears to be the lack of favorable warmth toward the children. This factor appears to be associated with the impairment of effective parenting so that children in these families do not fully internalize the difference between right and wrong.

Exposure to Violence

Exposure to violence both inside and outside the home has also been linked to aggression. Physical punishment and child abuse has also been associated with the development of aggression in children. In transgenerational studies, parents who behaved aggressively toward their children had aggressive children who, in turn, raised similarly aggressive children. Harsh discipline and child abuse (regardless of SES) appears to specifically predict the development of impulsive, but not nonimpulsive, aggressive behavior in children. Witnessing aggression has also been associated with the development and instigation of aggression, and it is likely that the presence of aggressive environmental cues may increase the immediate risk of aggression in susceptible individuals.

Deficiency in Problem Solving

A deficiency in problem-solving skills has been noted to be involved in the maintenance of aggression in impulsive aggressive children in particular. Impulsive aggressive children tend to have interpersonal problems and often get rejected by their peers. Although aggressiveness may appear to be a prime reason for peer rejection, cognitive factors may also play a role in this regard. Specifically, aggressive children tend to attribute hostile intent to others with whom they are interacting. It is important to note that the attribution of hostile intent to others may act as the provocation for an aggressive encounter. This hostile attribution is specific to their own interactions; these children do not attribute hostile intentions to other children with whom they are not interacting. The tendency of aggressive children to attribute hostile intent to others may be related to the observation that aggressive children, specifically impulsive aggressive children, have difficulty reading relevant social cues. For example, impulsive, but not nonimpulsive, aggressive children have been found to incorrectly label a visual cue (i.e., a sad face is identified as reflecting another emotion).

It is important to recognize, however, that these various psychosocial factors demonstrate a variable association with aggressive behavior. Like genetic factors, psychosocial factors can only predict aggressiveness as a probabilistic function of who may become aggressive given the presence of other factors. Clues to understanding this finding may be found in examining the role of biology in aggressive behavior.

Biology of Aggression

Among the constitutional/developmental biological factors possibly involved in aggression, the most

studied factors relate to brain neurochemistry, specifically to monoamines such as serotonin (5-HT) and other centrally acting neurotransmitters.

Serotonin

The evidence for a role of brain 5-HT in human aggression is especially strong. In fact, the inverse relationship between impulsive aggressive behavior is the most consistent of all findings in biological psychiatry. Various measures reflecting brain 5-HT function have been shown to correlate inversely with life history, self-report questionnaire, and laboratory measures of aggression. Most notable is the replicated observation that evidence of reduced central 5-HT is present in impulsive, but not premeditated, aggressive subjects. This indicates that impulsive aggressive behavior can be distinguished biologically from nonimpulsive aggression and, accordingly, is a valid subtype of human aggressive behavior. An inverse relationship between 5-HT and aggression is reported in most, although not all, studies. However, when studies are at variance with this basic finding, issues related to sample populations probably account for the differences in reported findings.

Nonserotonin Systems

Catecholamines and vasopressin Although there are less data available to support the role of non-5-HT brain systems and modulators in human impulsive aggression, there are limited data suggesting a facilitatory role for catecholamines (i.e., dopamine and NE) and vasopressin in impulsive aggressive behavior. The relationship of catecholamines and vasopressin to aggression and serotonin, specifically, is noteworthy. In the case of the catecholamines, the typically inverse relationship between aggression and 5-HT may not be seen when catecholamine system function is reduced. Accordingly, correlations between measures of 5-HT and aggression are generally absent in depressed subjects, who typically demonstrate diminished NE system function. In the case of central vasopressin and aggression, 5-HT appears to be inversely related to both central vasopressin and aggression in both animal and human subjects. In human subjects, the relationship between central vasopressin and aggression is present even after vasopressin's relationship with 5-HT is accounted for. In animal studies, both central vasopressin activity and aggression can be suppressed by treatment with selective serotonin reuptake inhibitor (SSRI) agents.

Testosterone and cholesterol In addition to non-5-HT central neurotransmitters, at least two peripheral substances appear to have a relationship with

aggression. Based on data from both animal and human studies, testosterone appears to have a facilitatory relationship with aggression. The relationship among testosterone, aggression, and 5-HT is less clear, although chronic testosterone exposure appears to downregulate central 5-HT receptors in some studies.

Another peripheral, although centrally active, substance that appears to influence aggressive behavior is cholesterol. Over the past decade, data from a variety of studies suggest that a reduction in circulating levels of cholesterol may be associated with an increase in aggressive behavior in both human and animal subjects. Although a link to 5-HT has not been conclusively demonstrated in human subjects, studies in nonhuman primates clearly demonstrate that a reduction in circulating cholesterol levels is associated with a reduction in measures that reflect central 5-HT function.

Localization of Biological Abnormalities

Reductions in cerebrospinal fluid (CSF) levels of neurotransmitter metabolites (e.g., 5-HIAA) suggest a central location for alterations in biological systems. However, CSF metabolites cannot reflect abnormalities in specific locations. Abnormalities in hormonal responses to pharmacological challenge agents do suggest an abnormality in the hypothalamus (a part of the limbic system involved in the regulation of emotion and behavior) of impulsive aggressive individuals. Recent clinical neuroscience and brain-imaging studies have reported data suggesting the hypofunction of parts of the prefrontal cortex and relative hyperfunction of deeper subcortical structures. For example, fluorodeoxyglucose positron emission tomography (FDG-PET) scan data demonstrate lower metabolic rates bilaterally in the lateral and medial areas of the prefrontal cortex of impulsive, but not premeditated, murderers than in normal controls. In contrast, the metabolic rates of subcortical structures (e.g., hippocampus, amygdala, thalamus, and midbrain) in the right hemisphere were found to be elevated in both impulsive and premeditated murderers compared with normal controls. More recent data from fMRI studies suggest a dysfunction of the amygdala-orbital prefrontal cortex such that there is greater activation of the amygdala and less activation of the orbital frontal prefrontal cortex. Because prefrontal structures are thought to regulate inhibition and subcortical structures, such as amygdala, are thought to regulate arousal, these data suggest that impulsive aggressive individuals have, compared to normal controls, less prefrontal inhibition in the context of greater subcortical emotional drive. Premeditated aggressive individuals, by contrast have

normal prefrontal inhibition but greater subcortical emotional drive. These data, in addition to the neurotransmitter-specific data already discussed, suggest that aggression in impulsive aggressive individuals can be treated by strategies that increase prefrontal inhibitory mechanisms and reduce subcortical facilitatory mechanisms.

Psychopharmacology of Aggression

As might be expected from the results of the biological study of aggression, several psychopharmacological agents appear to have anti- or pro-aggressive effects. The classes of agents shown to have anti-aggressive effects in double-blind, placebo-controlled trials include mood stabilizers (e.g., lithium), 5-HT uptake inhibitors (e.g., fluoxetine), anticonvulsants (e.g., diphenhydantoin, carbamazepine, and depakote), and NE beta-blockers (e.g., propranolol and nadolol). The classes of agents that may have pro-aggressive effects include tricyclic antidepressants (e.g., amitriptyline), benzodiazepines (e.g., alprazolam), and stimulant and hallucinatory drugs of abuse (e.g., amphetamines, cocaine, and phencyclidine). The most notable findings from the relatively small literature of the double-blind, placebo-controlled, clinical trials of anti-aggressive agents is that their anti-aggressive efficacy is limited to the impulsive, and not the nonimpulsive, aggressive individual. This indicates that psychopharmacology at this time may have little to offer society in the management of criminally motivated aggression. It remains to be seen, however, what the effect of treatment using agents that possibly dampen the function of deeper subcortical structures might be in a group of non-impulsively aggressive individuals. A second potentially important finding is emerging evidence of a differential psychopharmacology in this area so that some subjects may respond to some agents but not others. In our recent studies with fluoxetine, we noted an inverse relationship between pretreatment 5-HT function and anti-aggressive response to fluoxetine. Moreover, most fluoxetine nonresponders appeared to respond to treatment with divalproate, and the latter drug has also been shown to have anti-aggressive activity in selected impulsive aggressive patients with personality disorder. Given, the differential biology of impulsive aggression, it is likely that multiple strategies may be effective in impulsive aggressive individuals, depending on their own individual neurobiological substrate.

See Also the Following Articles

Antidepressant Actions on Glucocorticoid Receptors; Domestic Violence; Fear and the Amygdala.

Further Reading

- Barratt, E. S., Stanford, M. S., Felthous, A. R., et al. (1997). The effects of phenytoin on impulsive and premeditated aggression: a controlled study. *Journal of Clinical Psychopharmacology* 17, 341–349.
- Best, M., Williams, J. M. and Coccaro, E. F. (2002). Evidence for a dysfunctional prefrontal circuit in patients with an impulsive aggressive disorder. *Proceedings of the National Academy of Sciences USA* 99, 8448–8453.
- Coccaro, E. F. (1998). Central neurotransmitter function in human aggression. In: Maes, M. & Coccaro, E. (eds.) *Neurobiology and clinical views on aggression and impulsivity*, pp. 143–168. New York: John Wiley & Sons.
- Coccaro, E. F. (2003). Intermittent explosive disorder. In: Coccaro, E. F. (ed.) *Aggression, psychiatric treatment and assessment*, pp. 149–166. New York: Marcel Dekker.
- Coccaro, E. F. and Kavoussi, R. J. (1997). Fluoxetine and impulsive aggressive behavior in personality disordered subjects. *Archives of General Psychiatry* 54, 1081–1088.
- Coccaro, E. F., Siever, L. J., Klar, H. M., et al. (1989). Serotonergic studies in affective and personality disorder: correlates with suicidal and impulsive aggressive behavior. *Archives of General Psychiatry* 46, 587–599.
- Cowdry, R. W. and Gardner, D. L. (1988). Pharmacotherapy of borderline personality disorder: alprazolam, carbamazepine, trifluoperazine, and trancypromine. *Archives of General Psychiatry* 45, 111–119.
- Dodge, K. A., Lochman, J. E., Harnish, J. D., et al. (1997). Reactive and proactive aggression in school children and psychiatrically impaired chronically assaultive youth. *Journal of Abnormal Psychology* 106, 37–51.
- Hollander, E., Tracy, K. A., Swann, A. C., et al. (2003). Divalproex in the treatment of impulsive aggression: efficacy in Cluster B personality disorders. *Neuropsychopharmacology* 28, 1186–1197.
- Huesmann, L. R., Leonard, E., Lefkowitz, M., et al. (1984). Stability of aggression over time and generations. *Developmental Psychopathology* 20, 1120–1134.
- Kavoussi, R. K. and Coccaro, E. F. (1998). Divalproex sodium for impulsive aggressive behavior in patients with personality disorder. *Journal of Clinical Psychiatry* 59, 676–680.
- Linnoila, M., Virkkunen, M., Scheinin, M., et al. (1983). Low cerebrospinal fluid 5-hydroxyindolacetic acid concentration differentiates impulsive from nonimpulsive violent behavior. *Life Sciences* 33, 2609–2614.
- Miles, C. G. (1997). Genetic and environmental architecture of human aggression. *Journal of Personality and Social Psychology* 72, 207–217.
- Raine, A., Meloy, J. R., Bihle, S., et al. (1998). Reduced prefrontal and increased subcortical brain functioning assessed using positron emission tomography in predatory and affective murderers. *Behavioral Sciences & the Law* 16, 319–332.
- Virkkunen, M., Rawlings, R., Tokola, R., et al. (1994). CSF biochemistries, glucose metabolism, and diurnal activity rhythms in alcoholic, violent offenders, fire setters, and healthy volunteers. *Archives of General Psychiatry* 51, 20–27.

Aggressive Behavior

J M Koolhaas

University of Groningen, Haren, The Netherlands

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J M Koolhaas, volume 1, pp 100–103, © 2000, Elsevier Inc.

Aggressive Behavior: Form and Function

Physiology of Aggressive Behavior

Aggressive Behavior and Stress

Aggressive Behavior and Coping

Glossary

<i>Aggressive behavior</i>	A set of social behaviors closely associated in time with overt fighting.
<i>Defensive aggression</i>	A set of behaviors performed in defense to an attack by a conspecific or a potential predator.
<i>Dominant</i>	The member of a social group that is the winner of the social interactions with all other group members.
<i>Offensive aggression</i>	A set of aggressive behaviors performed against a conspecific aimed to protect resources.
<i>Subordinate</i>	Those members of a social group who lose in interactions with the dominant and never challenge the dominant.
<i>Violence</i>	A form of aggressive behavior which has lost its original biological function and which may cause serious damage to the opponent.

Aggressive Behavior: Form and Function

The existence of different kinds of aggression has long been recognized mainly on the basis of animal research. Two of these, namely instrumental and irritable aggression, are particularly relevant to bridging the gap between human and animal aggressive behavior. Instrumental aggression is a form of aggressive behavior that is instrumentally used to obtain desiderata, whereas irritable aggression is performed in reaction to aversive (social) events. This distinction seems to be analogous to the distinction between offense and defense made on the basis of animal experiments. Offensive behavior is displayed by a dominant resident against an unfamiliar male conspecific intruder of the home territory. Defensive behavior is performed by an animal that is attacked by either a dominant male or a predator. The offense–defense distinction

plays a prominent role in understanding the biology and physiology of animal aggression.

In addition to this distinction between offense and defense, other forms of aggressive behavior can be distinguished as well, such as infanticide, predator aggression, and maternal aggression, which can be observed in females in the late stages of pregnancy and in the early phases of nursing. This article concentrates on intraspecific aggressive behavior in males.

Aggressive behavior has a wide variety of functions depending on the type of aggression. It may serve to protect the home range and its resources against intruders and to maintain the social hierarchy of a group; it may function as a defense against predators, and it can be used instrumentally to obtain desiderata. Overt aggressive behavior such as fighting is usually preceded by an extensive repertoire of introductory threatening behaviors that may be highly ritualized in some species. Aggressive behavior always involves at least two participants, and the interaction usually ends with flight or submission of one of the competitors. Although aggressive behavior is highly functional, it is potentially harmful. Therefore, throughout the animal kingdom, strong reconciliation and appeasement mechanisms have developed to keep aggression in control and to prevent its adverse consequences. From a biological point of view, aggressive behavior is a highly functional form of social communication. However, much of the scientific and public interest in aggression is motivated by the violent, hostile, and presumably less adaptive forms of aggression observed in everyday human life and clinically across a wide spectrum of psychiatric and neurological disorders. The relationship between the functional and maladaptive extremes of the aggression spectrum is still far from clear and forms a major obstacle in the integration of animal research with data on human aggression. For the sake of clarity, the violent aspects of aggression are not discussed.

Physiology of Aggressive Behavior

Aggressive behavior or the threat of aggression is accompanied by strong neuroendocrine and autonomic stress responses in the two participants of the social interaction. The magnitude and the nature of this response depend on the outcome of the aggressive interaction, that is, winning or losing. Because the formation and maintenance of social hierarchies is one of the most important functions of aggressive behavior, much of the stress of everyday life has its

origin in this social structure. The relationship between aggression and stress obviously relates to the physiology of aggression. Therefore, the most important physiological mechanisms underlying aggressive behavior are briefly summarized. A complicating factor is the fact that various forms of aggressive behavior have a different physiological basis. Moreover, the causal physiological mechanisms underlying aggression cannot always be distinguished from the physiological consequences of the aggressive interaction. These consequences are in turn dependent on the outcome of the aggressive interaction, that is, victory or defeat.

Neuroendocrine Factors

Traditionally, high levels of offensive aggressive behavior have been linked to high levels of plasma testosterone. However, the direct causal relationship between testosterone and the expression of offensive behavior is still being discussed. It seems that various kinds of aggressive behavior may be differentially dependent on testosterone, whereas experiential factors strongly interfere with the degree in which aggression depends on plasma testosterone levels. Winning experience results in a temporary increase in plasma testosterone levels, whereas defeat results in a long-lasting decrease. At the level of the brain, testosterone is metabolized into estrogen and dihydrotestosterone. Evidence in rodents shows the existence of separate estrogen-dependent and androgen-dependent neuronal mechanisms of offense.

Due to its high emotional character and the high amount of physical activity, aggression is usually accompanied by a high activation of the pituitary adrenocortical system and the sympathetic adrenomedullary system. However, the reactivity of these neuroendocrine systems appears to differ as an important trait characteristic between highly aggressive and nonaggressive individuals. Aggressive males are characterized by a high reactivity of the sympathetic nervous system and the adrenal medulla, whereas nonaggressive males show a higher reactivity of the pituitary of the adrenocortical axis and the parasympathetic branch of the autonomic nervous system. Recently, hyporesponsiveness of the hypothalamic-pituitary-adrenocortical (HPA) axis has been associated with the development of extreme and presumably pathological levels of aggressive behavior in rats.

Genetic Factors

The role of genetic factors in aggression has long been recognized. Bidirectional selection in rodents usually leads to highly significant differences in aggressive

behavior within three or four generations. Extensive studies in human twins confirm a considerable genetic contribution to the individual level of aggression and violence. The initial studies in humans focused on the XYY syndrome in relation to criminal violence. Several studies in mice addressed the question of to what extent biological variation in aggressive behavior might be due to genetic variation at the Y chromosome. It appeared that the pairing region of the Y chromosome includes the steroid sulfatase gene and the sex-determining region, which might be responsible for at least some of the individual variation in aggression. Behavioral analyses of the rapidly increasing number of genetically manipulated mouse strains show the involvement of more than 50 other genes involved in elevated levels of aggression. Some of these genes code for important components of the neurobiology of aggression such as serotonergic and monoaminergic neurotransmission and gonadal steroid receptors and metabolism. However, for most of these genes it is unknown how they relate to the existing knowledge of the physiology of aggression. Studies aimed at genetic polymorphisms related to individual variation in aggression are mainly performed in humans. So far, the most consistent findings show polymorphism at the promoter regions of the genes coding for the serotonin transporter, for monoamine oxidase A, and for catechol-O-methyltransferase. However, due to the involvement of these neurobiological systems in a wide variety of other behaviors, it is unlikely that this genetic variation is exclusively correlated with aggression per se.

Central nervous organization There is a large body of literature on the central-nervous-system organization of aggressive behavior in animals. The specific involvement of various limbic and midbrain structures depends on the type of aggressive behavior. Offensive aggression involves specific hypothalamic, premamillary, and preoptic areas; the medial and central nuclei of the amygdala; and the prefrontal cortex. Defensive aggression involves the medial hypothalamus, the septal area, and the periaqueductal gray. Evidence suggests that the neuronal systems involved in these types of aggressive behavior are remarkably similar across species and may be homologous to those in humans. High levels of aggression or violence are often associated with functional disturbances in one or more of these brain regions. In the human literature, much attention has been paid to the dyscontrol syndrome, or episodic rage, which is probably attributed to seizure activity in certain brain areas. Brain imaging studies using functional magnetic resonance imaging (fMRI) and positron emission

tomography (PET) in humans with a history of violence and antisocial behavior show a reduced blood flow and glucose utilization in temporal and prefrontal cortical regions.

The most consistent findings of central nervous system functioning in aggressive behavior in both animals and humans, however, involve serotonin (5-HT). Low concentrations of 5-hydroxyindoleacetic acid, a metabolite of serotonin, in the cerebrospinal fluid have been consistently found as a trait characteristic in highly violent, aggressive, or suicidal patients. These findings are supported by many animal studies showing the anti-aggressive effects of selective 5-HT_{1A} or 5-HT_{1B} receptor agonists. Various other neurotransmitters including noradrenalin, dopamine, and γ -aminobutyric acid (GABA) and neuropeptides such as vasopressin are shown to be involved in the control of aggressive behavior in animals as well. However, their role in human aggression and violence is far from clear.

Aggressive Behavior and Stress

An aggressive interaction elicits a strong stress response in both the winner and the loser of the interaction. Evidence shows that winning and losing (or offense and defense) elicits a differential stress response both in terms of the magnitude of responses and of the neuroendocrine stress systems involved. Because winning and losing social interactions form the basis of the social structure in groups of animals, this differential response might underlie the relationship between the position in the social structure and the vulnerability for stress pathology.

Winner

Relatively few studies consider the physiological changes during aggressive behavior in the victor of the social interaction. However, the available literature clearly shows that winning elicits a moderate activation of the HPA axis and of the sympathetic adrenomedullary system. Both human and animal studies also show an activation of the pituitary-gonadal axis in males. More chronically, dominant males in a social group may suffer from sympathetic nervous system-mediated forms of stress pathology. For example, dominant males have a higher incidence of cardiovascular abnormalities such as arteriosclerosis and hypertension, in particular in unstable social groups.

Loser

Most social stress studies concentrate on defense, that is, the loser of a social interaction, or the subordinate

males in a social structure. A comparison of the stress response elicited by various stress procedures reveals that social defeat can be considered as one of the most severe stressors. Both the corticosterone response and the response of epinephrine and norepinephrine are by far the largest during social defeat stress. Recent evidence reveals that stressors show not only quantitative differences but may also differ qualitatively in the pattern of activated physiological systems. For example, a comparison of the electrocardiovascular response to restraint stress and social defeat reveals that the two types of stressors differ strongly in the balance of activation of the two branches of the autonomic nervous system. Social defeat seems to activate exclusively the sympathetic branch, whereas during restraint stress the parasympathetic nervous system shows a considerable activation as well. This differential autonomic balance may explain the high incidence of electrocardiovascular abnormalities observed during social defeat. Social defeat by a male conspecific induces not only strong acute increases in plasma catecholamines, corticosterone, heart rate, blood pressure, body temperature, prolactin, and testosterone but also increases in a variety of central nervous system neurotransmitters including serotonin, norepinephrine, dopamine, and endorphins. There is a general tendency to extrapolate these acute effects observed during and shortly after a defeat to a lasting pathology when the stressor is present chronically, for example, in the subordinate males in a social group. Indeed chronic subordination stress has been reported to induce a downregulation of serotonergic and noradrenergic neurotransmission, lower plasma testosterone levels, and increased activation of the HPA axis. Most of these changes are consistent with the idea that in a variety of animals species subordinates develop depressionlike symptoms. The subordinate position is also related to an increased sensitivity to addictive substances and immunosuppression, resulting in an enhanced vulnerability to infectious disease. Social defeat can be considered as a traumatic life event. Indeed, a single social defeat induces changes in physiology and behavior that can last from days to weeks and months. In fact, many of the behavioral and physiological changes observed after chronic subordination stress can also be observed weeks after a single social defeat.

Aggressive Behavior and Coping

An important concept in aggression research that relates to stress concerns the relationship between aggression and coping. A number of studies in a variety of species shows a strong positive correlation

between the individual level of offensiveness and the tendency to actively cope with any kind of environmental stressor. Aggressive males have a general tendency to actively deal with different kinds of environmental challenges. The antipode of this active problem-focused coping is a reactive or emotion-focused coping style. This is expressed as the absence of aggressive behavior in a social situation and as passivity in other challenging environmental conditions. These coping styles, as observed in several animal species, are also characterized by differential neuroendocrine and neurobiological profiles. They can be considered as important trait characteristics determining the individual adaptive capacity.

Personality factors or trait characteristics have long been recognized as playing a role in human stress psychophysiology as well. For example, the distinction between proactive and reactive coping styles in relation to individual levels of aggression is also made in human beings. The different coping styles as found in animal studies seem to be analogous to the distinction between the human type A and type B personalities used in cardiovascular psychophysiology. The type A personality is described as aggressive, hostile, and competitive and is physiologically characterized by a high sympathetic reactivity. Although the A-B typology has been seriously questioned in the literature, factors related to aggression such as anger and hostility have been repeatedly shown to play a role in the individual capacity of human beings to cope with environmental challenges. Recent studies in feral populations of various animal species show the evolutionary basis of individual variation in aggression and coping or personality as an epigenetic suit of traits.

See Also the Following Articles

Aggression ; Defensive Behaviors; Impulse Control; Disasters and Mass Violence, Public, Effects of.

Further Reading

- de Boer, S. F., Van Der Vegt, B. J. and Koolhaas, J. M. (2003). Individual variation in aggression of feral rodent strains: a standard for the genetics of aggression and violence? *Behavior Genetics* 33, 485–501.
- Haller, J., Halasz, J., Mikics, E. and Kruk, M. R. (2004). Chronic glucocorticoid deficiency-induced abnormal aggression, autonomic hypoarousal, and social deficit in rats. *Journal of Neuroendocrinology* 16, 550–557.
- Koolhaas, J. M., Korte, S. M., De Boer, S. F., et al. (1999). Coping styles in animals: current status in behavior and stress-physiology. *Neuroscience and Biobehavioral Reviews* 23, 925–935.
- Korte, S. M., Koolhaas, J. M., Wingfield, J. C., et al. (2005). The Darwinian concept of stress: benefits of allostasis and costs of allostatic load and the trade-offs in health and disease. *Neuroscience and Biobehavioral Reviews* 29, 3–38.
- Nelson, R. J. and Chiavegatto, S. (2001). Molecular basis of aggression. *Trends in Neuroscience* 24, 713–719.
- Sih, A., Bell, A. M., Johnson, J. C., et al. (2004). Behavioral syndromes: an integrative overview. *Quarterly Review of Biology* 79, 241–277.
- Simon, N. G., McKenna, S. E., Lu, S. F., et al. (1996). Development and expression of hormonal systems regulating aggression. *Annals of the New York Academy of Sciences* 794, 8–17.
- Veenema, A. H., Meijer, O. C., de Kloet, E. R., et al. (2003). Genetic selection for coping style predicts stressor susceptibility. *Journal of Neuroendocrinology* 15, 256–267.

Aging and Adrenocortical Factors

C Lord and J C Pruessner

McGill University, Montreal, Quebec, Canada

© 2007 Elsevier Inc. All rights reserved.

The Hypothalamic-Pituitary-Adrenocortical (HPA) Axis
HPA Axis Changes with Normal Aging
HPA Axis Regulation Changes and the Hippocampus
Individual Differences in Changes of the HPA Axis with
Aging
Conclusion

Glossary

<i>Acrophase</i>	The time of occurrence of the peak, or overall highest values, in recurrent (physiological) cycles.
<i>Circadian rhythm</i>	Roughly 24-h cycle shown by physiological processes in living organisms; from Latin <i>circadian</i> , meaning “around a day.”
<i>Glucocorticoids</i>	A class of steroid hormones characterized by an ability to bind with the cortisol receptor and trigger similar effects;

	the most representative example is the hormone cortisol.
<i>Homeostasis</i>	The ability of a system, especially living organisms, to regulate its internal environment to maintain a stable, constant condition.
<i>Nadir</i>	The time of occurrence of the overall lowest values in recurrent (physiological) cycles.

The Hypothalamic-Pituitary-Adrenocortical (HPA) Axis

The hypothalamic-pituitary-adrenocortical (HPA) axis is one of the most important hormonal systems in the body. It is activated when the homeostasis of the organism is threatened, and is thus commonly referred to as the stress axis. Threats to the homeostasis include physical, pharmacological, and psychosocial stressors. The major components of the axis are both inside and outside of the brain. Upon stimulation, the hypothalamus releases corticotropin-releasing hormone (CRH). CRH is transported to the pituitary gland by way of the hypophyseal portal vessels. In the pituitary gland, CRH triggers the release of adrenocorticotropic hormone (ACTH) into the systemic circulation whereby it is transported to the adrenal cortex and stimulates the synthesis and release of cortisol. These hormones have a variety of effects on the organism, which can be differentiated into permissive, stimulatory, suppressive, and preparative actions. Systems that are sensitive to cortisol presence include immunity, inflammation, metabolism, neural function, behavior, and reproduction. For example, cortisol inhibits the glucose metabolism in non-essential peripheral tissue, while the glucose supply to essential tissues is maintained. Gluconeogenesis and insulin resistance are augmented, while lipolysis, proteolysis, and inflammation are inhibited. Whether cortisol takes stimulatory or suppressive action in the organism depends on the concentration, the time course, and the target area. Cortisol has the ability to cross the blood-brain barrier. Within the central nervous system, it reduces further activity of the HPA axis by occupying mineralocorticoid (MR) and glucocorticoid (GR) receptors in the hippocampus and frontal lobes, which then serve to inhibit subsequent hypothalamus activity and CRH release.

Apart from responding to external stimulation, the activity of the HPA follows a circadian rhythm. The highest levels of cortisol (the acrophase) are usually found in the morning around the time of awakening. Recent studies suggest that awakening might be a stimulus to the HPA axis itself, yielding the highest

levels of cortisol in the first hour after awakening. Cortisol levels then gradually decline until around noon, when a second peak (usually not as high as in the morning) can be observed around lunchtime. During the afternoon and evening, cortisol levels continuously decline, with the lowest levels (the nadir) around midnight. In addition to this circadian rhythm, circavigintan, circatrigintan, and circaannual rhythms of cortisol, reflecting seasonal effects, and menstrual phase influences on HPA axis activity in women have been described. Furthermore, the secretion of cortisol follows a pulsatile pattern, with up to 15 distinct pulses per hour. These constant variations in hormone levels make exact measurement of HPA axis regulation difficult. Special strategies have thus been developed for precise determination, including repeated measurements of cortisol levels with up to 10-min intervals throughout the 24-h period, challenging the HPA axis with pharmacological agents or psychosocial stimuli to assess cortisol reactivity, or assessing activity of the axis at times with high reliability (early in the morning around the acrophase of the axis, or late in the evening when the nadir is approached).

HPA Axis Changes with Normal Aging

Normal aging is accompanied by a number of subtle but consistent changes in the regulation and activity of the HPA axis. In older age, it appears that the amplitude of the circadian rhythm of basal cortisol is reduced, caused by either lower levels in the morning, during the time of the acrophase, or higher levels in the evening, around the time of the nadir. Studies evaluating the stress response in elderly populations concluded that older subjects can still mount a significant endocrine response to a stressful situation, indicating that the psychoendocrine response is maintained into old age. Furthermore, an interaction of aging with gender has been described: while in younger subjects, men usually show the strongest responses to psychosocial stress, recent studies demonstrate that women tend to show stronger responses in older age. For basal levels, it has been reported that women show stronger increases in the morning during the time of the acrophase with aging.

In addition, the feedback sensitivity of the HPA axis is reduced in older age, which might be directly related to the finding of higher basal cortisol levels. It is speculated that with compromised feedback, inhibition of subsequent HPA axis activity is less effective and thus might cause higher cortisol levels. **Figure 1** depicts the global changes associated with normal aging schematically.

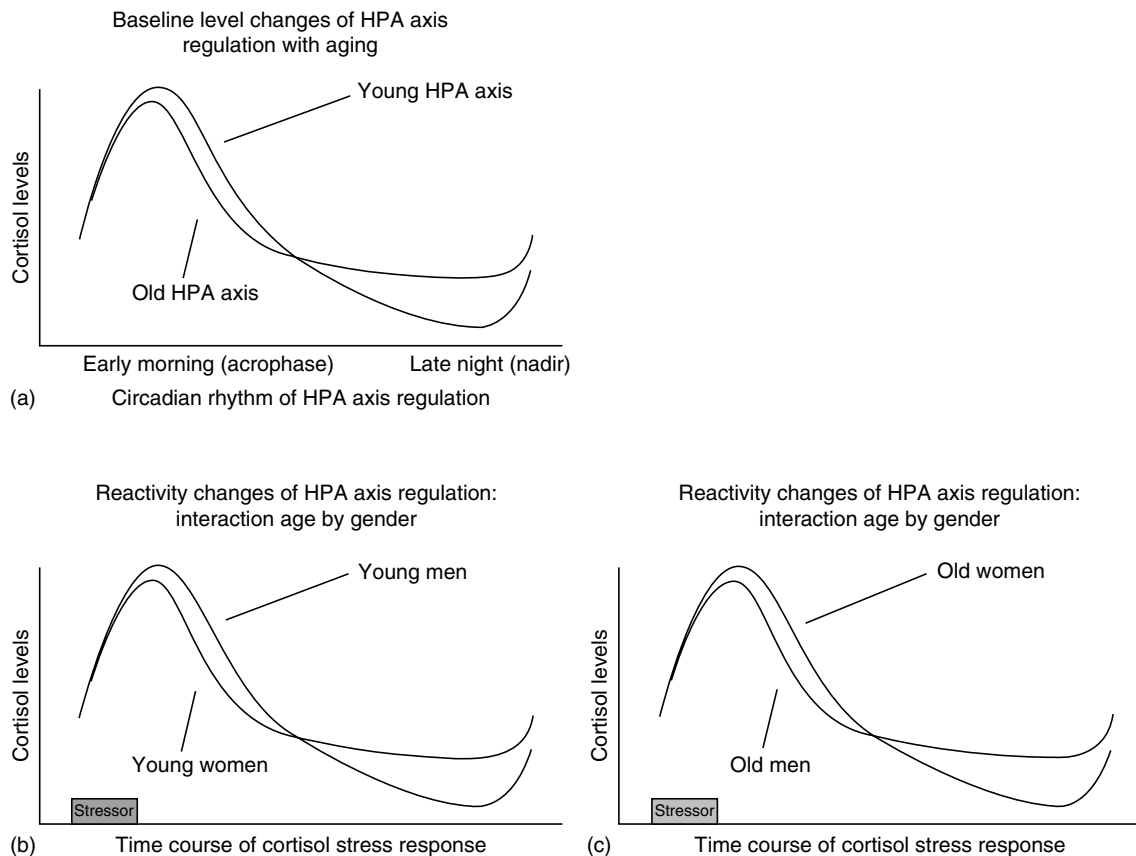


Figure 1 Schematic description of age related changes of HPA axis regulation with regard to circadian rhythm (a) and reactivity to stress in young (b) and elderly (c) men and women.

HPA Axis Regulation Changes and the Hippocampus

The hippocampus (HC), an important structure for memory and learning, is implicated in the changes of the regulation of the HPA axis that occur with aging. The HC normally exerts a regulatory function on HPA axis activity through MR and GR receptors, which inhibit the subsequent release of CRH from the hypothalamus.

Since glucocorticoids have been found to be neurotoxic in both rodents and humans, excessive activation of the HPA axis or cumulative activation of the axis throughout the lifetime can damage hippocampal tissue. In animals, it has been shown that hypersecretion of glucocorticoids leads to dendritic atrophy, especially in the CA3 region of the HC. The impaired integrity of the HC is believed to result in decreased inhibition of CRH release from the hypothalamus, and thus a less effective inhibition of subsequent HPA axis activity. This model of impairment that subsequently promotes further damage has been termed the glucocorticoid cascade hypothesis and is one mechanism to explain age-related changes of HPA axis regulation in humans.

Direct evidence for the link between HPA axis regulation and hippocampal integrity in humans stems from recent studies showing that basal cortisol and ACTH levels are inversely associated with hippocampal volume. Earlier, it was shown that acute or chronic increases in cortisol levels either through psychosocial stress or through pharmacological stimulation lead to declined memory performance in cognitive tasks, suggesting that acute or chronic effects of occupation of GR receptors in the HC can impair memory processing in the HC.

Individual Differences in Changes of the HPA Axis with Aging

Recently, evidence emerged for a more dynamic association between changes of the HPA axis regulation and aging. It seems that age-related changes do not occur uniformly, but show significant variation in interaction with other variables. One study uncovered considerable variations in current as well as accumulated cortisol levels and gave clear evidence of subgroups in a healthy aged population. Subjects demonstrated a progressive year-to-year increase in

cortisol levels with currently high cortisol levels, a progressive year-to-year increase in cortisol levels with currently moderate cortisol levels, or a progressive decrease in cortisol levels with currently moderate cortisol levels. Thus, increasing cortisol levels do not appear to be a universal correlate of aging. It seems that age-related changes do not occur uniformly, but may exhibit significant variation in interaction with other variables. This observation prompts the question of whether physical, psychological, or environmental conditions exist when these changes are minimized or augmented. Endocrinologically, recent studies investigating the effects of dehydroepiandrosterone (DHEA) are a good example of the search for variables that potentially modify age-related changes. In the same context, recent studies investigating the effects of estrogen therapy on HPA axis regulation start to show how other hormones interact with aging to affect the regulation of the HPA axis. In genetic research, the identification of risk genes that promote specific age-related diseases, like the APO-E4 gene for the development of Alzheimer's disease, is another good example for this line of research. Similarly, an increasing number of studies investigate the effects of mental illness such as depression or posttraumatic stress disorder on age-related changes on the HPA axis. These studies start to point out the importance of considering lifelong influences on age-related changes in the HPA axis.

Psychologically, it appears that personality variables such as locus of control, self-concept, and self-esteem might contribute to individual differences in age-related changes in humans. These changes were linked to systematic differences in hippocampal volume as well, replicating and extending previous studies showing an interaction between hippocampal volume and HPA axis regulation. Finally, recent efforts to investigate the impact of early-life events (like the quality of parental caretaking) on the HPA axis regulation open a new chapter in HPA axis research, formulating models in which peri- or post-natal variables might program lifelong sensitivity or reactivity variations in HPA axis regulation.

Conclusion

In humans, age is associated with global changes on different levels. The HPA axis is one system that is affected by these changes. In general, it appears that advancing age is associated with a flattening of the circadian rhythm, as indicated by increased basal and

ACTH levels in the evening and decreased levels in the morning. Moreover, it has been shown that the fast feedback inhibition of glucocorticoids on ACTH secretion is reduced with advancing age, suggesting that age might have an impact on HPA axis regulation at different levels. Furthermore, a global age by gender interaction has been reported, with men showing stronger responses to psychosocial stress in young adulthood and women demonstrating stronger responses to stimulation in older age. Unstimulated, there seems to be some evidence that women show stronger increases in the morning during the time of the acrophase with increased age.

In addition to these global changes, recent studies have documented the impact of interindividual differences on age-related changes in the regulation of the HPA. Evidence is now emerging that links age-related changes of HPA axis regulation to variations in personality variables, hippocampal volume, other hormonal systems such as estrogen, and mental illnesses such as depression or posttraumatic stress disorder.

Future studies will have to take into account the effects of behavioral and physiological systems that might exert a lifelong impact on the HPA axis to determine individual changes of the HPA axis with age.

See Also the Following Articles

Adrenocortical Function, Factors Controlling Development thereof; Adrenocorticotrophic Hormone (ACTH); Circadian Rhythms, Genetics of; Hippocampus, Overview; Cortisol Awakening Response; Hypothalamic-Pituitary-Adrenal.

Further Reading

- Horan, M. A. and Little, R. A. (1998). *Injury in the aging*. Cambridge, UK: Cambridge University Press.
- Haus, E. and Touitou, Y. (1994). Principles of clinical chronobiology. In: Touitou, Y. & Haus, E. (eds.) *Biologic rhythms in clinical and laboratory medicine*, pp. 6–33. Berlin: Springer Verlag.
- Lupien, S. J., de Leon, M., de Santi, S., Convit, A., Tarshish, C., Nair, N. P., et al. (1998). Cortisol levels during human aging predict hippocampal atrophy and memory deficits. *Nature Neuroscience* 1(1), 69–73.
- Morrison, M. F. (2000). *Hormones, gender and the aging brain*. Cambridge, UK: Cambridge University Press.
- Pruessner, J. C., Baldwin, M. W., Dedovic, K., Renwick, R., Mahani, N. K., Lord, C., et al. (2005). Self-esteem, locus of control, hippocampal volume, and cortisol regulation in young and old adulthood. *Neuroimage* 28, 815–826.
- Watson, R. R. (1999). *Health promotion and aging*. New York: Taylor & Francis.

Aging and Psychological Stress

B W J Penninx

VU University Medical Center, Amsterdam,
The Netherlands

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
B W J Penninx and D J H Deeg, volume 1, pp 104–110, © 2000,
Elsevier Inc.

Aging

Psychological Stress: Prevalence of Depression and Anxiety
Disorders in Old Age

Determinants of Psychological Stress in Old Age

Consequences of Psychological Stress in Old Age

Conclusion

Glossary

<i>Major depressive disorder</i>	Severe depressed mood and/or loss of interest in activities in combination with additional symptoms such as loss of appetite, sleep disturbance, fatigue, feelings of worthlessness, thoughts about death, or concentration problems.
<i>Generalized anxiety disorder</i>	Chronic, persistent, and excessive feelings of anxiety and worry.
<i>Obsessive compulsive disorder</i>	Repetitive thoughts, images, or impulses that are experienced as intrusive, or repetitive behaviors that are performed in response to an obsession or certain rules.
<i>Panic disorder</i>	Recurring sudden episodes of intense apprehension, palpitations, shortness of breath, or chest pain.
<i>Phobia</i>	Fear and avoidance out of proportion to the danger.
<i>Subthreshold depression</i>	Clinically relevant depressed mood not fulfilling the diagnostic severity criteria for major depressive disorder.

Aging

In the Western world, the oldest of the old, i.e., people aged 85 years and older, comprise the fastest growing segment of the aged. This is due to a decreasing trend in the number of children born as well as to dramatic changes in mortality leading to increased life expectancy. Life expectancy has consistently been higher for women than for men. Consequently, the elderly population is composed of more women than men. So, in absolute terms, aging is affecting women more than men.

Aging has profound consequences for the individual. For the most part, individual aging is associated

with many adverse changes in human anatomy and physiology. Consequences of these changes are, for instance, losses in the sense of balance and movement, poorer hearing and vision, slower reactions, and weaker muscles. For a large part of adult life, people are normally provided with biological reserves. In later life, these reserves are reduced, which can cause a weakening of one or another biological function essential to life. Consequently, conditions such as heart disease, cancer, respiratory infection, and kidney failure may arise. The biological age-related changes and the consequent development of degenerative and chronic conditions, have a large impact on the physical functioning of older people. The number of older people with difficulties in mobility and activities of daily living increases dramatically with increasing age.

In addition to the impact on biological and physical areas, other aspects of life are also affected. Individual aging is accompanied by a series of social transitions, some entered into voluntarily, some imposed by circumstances. In general, aging tends to be associated with relationship losses, either because of retirement, widowhood, or death of age-peers such as siblings or friends. Increasing age also brings changes in relationship needs, for example, as the result of increasing impairment. Older adults may become more dependent on others, when they lose the ability to fulfill certain social or instrumental tasks themselves. The existing balance in their relationships may be disrupted, introducing strain and discomfort.

Most of the age-related changes, biological, physical, or social, can be expected to constitute losses rather than gains. Several of these changes are potentially stressful to older people. Chronic conditions, for instance, can introduce pain, disability, despair, and fear for treatment and pending death. Widowhood, the need for instrumental support, and the loss of personal relationships might cause feelings of loneliness, dependency, and helplessness. When adjustment to these burdening circumstances is inadequate, an older person might experience substantial psychological stress. In the next section, the prevalence of psychological stress in old age is described, indicated by two prevalent affective disorders: depression and anxiety.

Psychological Stress: Prevalence of Depression and Anxiety Disorders in Old Age

Depression

The most prevalent affective disorder in old age is depression. Prevalence rates of depression vary

considerably depending on the sample studied and methods used. Studies in clinical settings generally find much higher prevalences of depression than studies in community settings. Depression can be assessed using either diagnostic criteria or symptom checklists. Diagnostic criteria are described in the *Diagnostic and Statistical Manual of Mental Disorders (DSM)* published by the American Psychiatric Association and are used throughout the world. According to these criteria, a major depressive disorder is present when a total of five or more out of the following nine symptoms are present: depressed mood; lack of interest; feelings of worthlessness or inappropriate guilt; diminished ability to concentrate or make decisions; fatigue; psychomotor agitation or retardation; insomnia or hypersomnia; significant decrease or increase in weight or appetite; and recurrent thoughts of death or suicidal ideation. The identified symptoms should at least include one or both of the two core symptoms (depressed mood and lack of interest), should be present for 2 weeks or more, and should generally be severe enough to cause disruptions in a person's daily functioning. Overall, the symptoms of moderate-to-severe depression presented to the clinician are rather similar across older people and people in midlife. However, there have been described some subtle differences in symptom experience across age groups. Melancholia (symptoms of noninteractiveness) appears to be a little more frequent in later age than in younger age, with psychomotor disturbances being more obvious in older people.

Symptom checklists ask for presence, intensity, or frequency of a series of symptoms during the preceding 1 or 2 weeks in order to measure the extent of depressive symptomatology. Some examples of commonly used symptom checklists are the Center for Epidemiologic Studies–Depression Scale, the Geriatric Depression Scale, and the Zung Depression Scale. These instruments are well validated and have been proven to be valid and reliable instruments in older populations. It is possible to score relatively high on a symptom checklist without meeting diagnostic criteria for major depressive disorder. In fact, depressive symptom checklists identify for the largest part people with a significant high level of depressive symptoms who do not fulfill the diagnostic severity threshold of major depressive disorder. This condition is often referred to as depressed mood or subthreshold depression.

As confirmed in several aging studies, major depressive disorder is relatively rare among older community-dwelling people, affecting only about 1–2% of the community-dwelling population. Major depressive disorder appears to be less prevalent among older adults than among middle-aged adults. **Table 1** shows the prevalence rates of major depressive

Table 1 Prevalence rates for depression in older men and women in different age groups^a

	Age		
	55–65 years	65–75 years	75–85 years
Major depressive disorder ^b			
Men	1.2%	0.9%	0.5%
Women	3.4%	3.7%	2.2%
Subthreshold depression ^c			
Men	8.5%	8.0%	13.7%
Women	9.5%	14.0%	21.1%

^aData are from the Longitudinal Aging Study Amsterdam.

^bBased on diagnostic DSM-criteria using the Diagnostic Interview Schedule.

^cIndicated by a score on the Center for Epidemiologic Studies Depression Scale of 16 or more.

disorder in a community-based random sample of older men and women aged 55 years and older who participated in the Longitudinal Aging Study Amsterdam (LASA) in The Netherlands. This table shows that among older people, the prevalence of major depressive disorder shows a slightly declining trend with increasing age. Some have suggested that part of the declining prevalence of major depression with older age could be due to bias since those with depression may die earlier and therefore not survive to old age, or due to the fact that older people may underreport depression compared to younger people. Nevertheless, research findings do indicate that the prevalence of major depression in old age is at least not higher than in earlier age. In addition, it is important to realize that a large proportion of older people with a major depressive disorder have had depressive episodes during earlier phases of their lives. Consequently, a personal history of depressive disorder is one of the strongest risk factors for a major depressive disorder in old age. Thus, major depressive disorders in old age most often represent recurring episodes of early-onset depressive disorders. At all ages, older women generally show a higher rate of major depressive disorder than men.

When using a depressive symptom checklist it becomes obvious that a much higher proportion of the population suffers from significant symptoms of depression. Prevalences of subthreshold depression in older community-based populations are in the range of 12–20%. Data from several community studies converge in suggesting that there is a curvilinear relationship between age and depression scores over the entire adult life span, with the highest scores among younger adults and those aged over 75 years. **Table 1** illustrates the positive association between a

significant level of depressive symptoms and age in older participants of the LASA study. In line with findings for major depressive disorders, women generally show higher rates of significant depressive symptoms than men. The elevation in scores among the very old has been shown not to be an artifact of greater endorsement of somatic symptoms. In sum, the prevalence of minor depression increases across old age, whereas the prevalence of clinical major depression generally does not.

Anxiety

A variety of anxiety disorders are described in the DSM, including (1) panic disorder, which refers to recurring sudden episodes of intense apprehension, palpitations, shortness of breath, and chest pain; (2) phobias, which are fears and avoidance out of proportion to the danger; (3) obsessive–compulsive disorder, consisting of repetitive thoughts, images, or impulses experienced as intrusive or of repetitive behaviors that are performed according to certain rules or in a stereotyped fashion; and (4) generalized anxiety disorder, entailing chronic, persistent, and excessive anxiety and worry. The prevalence of anxiety disorders tends to decline over the lifetime. Nevertheless, anxiety disorders are still common in older adults. Table 2 reports the prevalence rates of various anxiety disorders in the LASA study. In this older population, anxiety disorders are more common among women than men, but are not associated with age. Generalized anxiety disorders are more common than panic disorders and phobic disorders. Obsessive–compulsive disorder was the least prevalent.

Like in younger ages, comorbidity of anxiety disorders is not uncommon in older age. In the LASA study, about 10% of the older people with an anxiety diagnosis suffered from two or more anxiety disorders. Also, anxiety and depression disorders frequently co-occur at any age. Thirty-six percent of the older people with a major depressive disorder had one or more anxiety disorders. Vice versa, 13% of those with an anxiety disorder had a major depressive disorder.

Determinants of Psychological Stress in Old Age

Genetic and Biological Vulnerability

Some research studies have indicated that major depressive disorder and subthreshold depression in older age have distinct underlying etiologies. In general, major depressive disorder appears to be more strongly associated with longstanding vulnerability

Table 2 Prevalence rates for anxiety disorders in older men and women in different age groups^a

Disorder	Age		
	55–65 years	65–75 years	75–85 years
Panic disorder ^b			
Men	0.3%	0%	0.3%
Women	1.4%	3.0%	2.0%
Obsessive compulsive disorder ^b			
Men	0%	0%	0%
Women	0.3%	2.7%	0.8%
Social phobia ^b			
Men	1.2%	0.6%	2.3%
Women	4.2%	4.4%	4.9%
Generalized anxiety disorder ^b			
Men	2.8%	10.3%	5.4%
Women	5.2%	12.6%	8.5%

^aData are from the Longitudinal Aging Study Amsterdam.

^bBased on diagnostic DSM-criteria using the Diagnostic Interview Schedule.

factors such as personality and heredity than subthreshold depressed mood. A large proportion of older people with a major depressive disorder have had depressive episodes during earlier phases of their lives. Consequently, a personal history of depressive disorder is one of the strongest risk factors for a major depressive disorder in old age. Other heritability and personality characteristics, such as a family history of depressive disorder or neuroticism have also been found to be predictive. Also anxiety disorders in late life are often associated with longstanding vulnerability factors such as personality and heredity. Thus, both major depressive disorders and anxiety disorders in old age most often represent recurring episodes of early-onset disorders.

Furthermore, some biological vulnerabilities may predispose an older individual for affective disorders. Age-related physiological and brain structure changes have been suggested as plausible risk factors for late-onset depression because of parallels between these changes and pathological changes observed in depression, such as dysregulation of the monoamine system, increased monoamine oxidase activity, deficits in norepinephrine functioning, downregulation of serotonin receptors, and deep white- and gray-matter disease. Recently, much attention has been directed to vascular risk for late-life depression. Several studies have suggested that vascular lesions in selected regions of the brain may contribute to a unique variety of late-life depression. The vascular depression impairments resemble impairments exhibited

in frontal lobe syndromes. Magnetic resonance imaging of depressed patients has revealed structural abnormalities in areas related to the cortical-striatal-pallidal-thalamus-cortical pathway. Endocrine changes have also been associated with late-life depression. For instance, major depression and panic disorders have been associated with elevated basal cortisol levels, elevated awakening curves of cortisol, and nonsuppression of cortisol, which all suggests hyperactivity of the hypothalamus-pituitary-adrenal (HPA) axis. Although these disturbances have been found among older people as well, there are some recent studies that indicate that physical frailty could actually exhaust the body's responses to stress resulting in hypoactivity of the HPA axis. This may explain why some studies among older frail people have actually observed hypoactivity of the HPA axis among depressed people. For other trauma-related stress disorders (e.g., post-traumatic stress disorders), a higher occurrence of hypoactivity of the HPA-axis has also been confirmed. How other types of anxiety disorders (e.g., social phobia and panic disorders) relate to HPA-axis activity has not been examined in detail yet. Finally, low levels of sex steroid hormones, such as testosterone, estradiol, and dehydroepiandrosterone sulfate (DHEAS), which are especially apparent in older women, have been indicated to result in the development of late-life depressive symptoms and maybe even depressive disorders.

Psychosocial Determinants

Whereas biological vulnerability generally seems to increase with age, in terms of psychological predisposition older adults may be less vulnerable than younger adults. Although stressful life events still appear to increase the risk of depression in old age, the relationship between life events and depression may actually be less strong in older than in younger adults. Older individuals have had the opportunity to learn how to cope with stressful circumstances and how to adjust their expectations so as to have fewer feelings of failure. On the basis of age and experience, older people have developed more effective skills with which to manage stressful life events and to reduce emotional distress. Specific stressors, such as loss of partner or other intimates, are more normative in old age and usual in that part of the lifecycle than in younger age, and therefore probably less disruptive. Although their impact might be smaller than in younger age, social circumstances have still been found to be predictive for depressive as well as anxiety symptomatology. The absence of a partner, a smaller social network size, and lower exchange of emotional support are all important factors that increase the chance of experiencing more feelings of depression

and anxiety in old age. Also, depression and anxiety are more common in older people with unfavorable socioeconomic circumstances, such as low income or education.

Somatic Health Determinants

The importance of physical health for the presence of psychological stress is undisputed. Physical illnesses are consistently shown to be included among the strongest risk factors. For instance, depressive symptomatology appears to be higher in the presence of the following diseases: lung disease, arthritis, cancer, diabetes, stroke and coronary heart disease, and other cardiac illnesses. In addition to specific disease, depressive symptoms are also more prevalent in the presence of certain impairments, such as hearing and vision problems and cognitive impairment. Several explanations can be given for the link between physical health and affective disorders. First, affective disorders can occur as an outcome of certain somatic illnesses or medication, reflecting a biologically mediated process. For example, the structural and neurochemical changes involved in stroke and parkinsonism can lead to depression. Second, physical illness can act as a stressor and has several psychosocial consequences. For example, loss of function, role, and independence, negative body image, and sense of identity, pain, and promoted sense of helplessness can be a reaction of being ill and consequently cause increased feelings of depression or anxiety. Finally, it should be mentioned that physical illnesses and depression or anxiety may exhibit similar symptoms, causing artificially raised correlations. For example, silent myocardial infarctions can produce anxietylike symptoms and some symptoms of depression, especially the somatic symptoms such as the presence of a low energy level and sleeping problems, may partly be a manifestation of the disease. Among older people, it is sometimes rather difficult to distinguish pure emotional consequences of a somatic condition from depression. According to psychiatric criteria, a major depressive disorder is only present when not all the depressive symptoms are clearly the consequence of one somatic condition or the use of a certain pharmacological drug. So, this distinction needs to be considered when evaluating the presence of a depression.

Both the biological age-related changes and the consequent development of chronic conditions and impairments have a large impact on the physical functioning of older people. The extent of physical disability is an important factor in the development of psychological stress.

Several studies among older people have shown that symptoms of depression and anxiety are more strongly predicted by problems with activities of daily

living (disability) than by number of chronic conditions. This illustrates that a disability measure identifies the comorbidity, severity and consequences of chronic conditions better than a simple count of the number of chronic conditions.

As mentioned previously anxiety and major depressive disorders are not more prevalent in old age than in young age. By contrast, subthreshold depressed mood shows an increasing trend with increasing age. Nevertheless, the risk factor profile for subthreshold depressed mood over the lifespan is rather stable in terms of the domains represented. Risk factors in younger people have also been found in old age (e.g., physical health problems, negative life events, and lack of social support). However, the importance of risk factors relative to each other may vary across the lifespan, since the distribution of these risk factors is related to age. For instance, the oldest-old report more social isolation and impoverishment, and have more physical health problems and disability. The relative increases in the prevalence of these risk factors, especially in physical health problems, have been found to completely explain the increased prevalence of subthreshold depressed mood in old age. Thus, what seems to be an association between age and depressive symptoms is in fact the result of other risk factors. Age *per se* is not a risk factor for subthreshold depressed mood when other factors are taken into account.

Consequences of Psychological Stress in Old Age

The presence of psychological stress in old age has several adverse health consequences. These consequences have mainly been studied for depression. Since there has been a striking lack of attention to the health consequences of late-life anxiety, this section will mainly focus on the described adverse health outcomes of depression.

Depressed older people have been found to have a higher overall mortality risk during a subsequent follow-up period than nondepressed older people. These increased mortality risks have been described for older people with major depressive disorders as well as for those with subthreshold depression, although risks were generally smaller for the latter. Although not as consistently as for depression, there are a few recent studies that confirm a similar association for anxiety problems: both assessments of anxiety symptoms as well as anxiety disorders have shown to increase the risk for subsequent overall mortality.

Consistent associations with depression have also been found for various other indicators of somatic health decline. For instance, depression increases the

risk for subsequent onset of physical disabilities in mobility or in activities of daily living. Researchers have further detailed the evidence for health decline associated with depression by documenting increased risk of developing hypertension, stroke, and cardiovascular disease among older depressed people. In addition, there is evidence that high levels of depressive symptoms in cardiovascular patients predict an increased future risk for recurrent myocardial infarction and mortality. However, depression appears to be also associated with noncardiovascular mortality so, although less explored, it is likely that depression may impact mortality through various different, and not exclusively cardiovascular, mechanisms.

It is unclear whether the effect of major depressive disorders or depressive symptoms on physical health decline is different in young and old people. For mortality, a few studies described that the association between depression and mortality is weaker among older people, but other studies did not find a difference between young and old people. For other health outcomes, the possible differentiating effect of age in the association between depression and declining health has not been much examined.

How can the adverse health consequences of depression in older people be explained? First, psychological stressors have physiological accompaniments, including altered autonomic balance, and increased HPA axis function, both of which are biologically plausible contributors to pathogenesis. When anger or depression is induced in laboratory studies, people exhibit larger sympathetic nervous system-mediated cardiovascular responses, as documented by larger increases in daytime urinary epinephrine excretion and downregulation of lymphocyte β -adrenoceptors. There is also evidence that parasympathetic nervous system function is reduced in both anxious and depressed people. In addition, depressed people have been shown to exhibit hypersecretion of the adrenal steroid cortisol, adrenal hypertrophy, and an increased cortisol response to adrenocorticotrophic hormone. The resulting hyperactivity of the HPA-axis could result in activation of inflammatory processes (inflammation). Increased levels of inflammatory markers (e.g., interleukin (IL)-6, tumor necrosis factor (TNF)- α , and C-reactive protein) have been demonstrated among patients with depression. Overall, these physiological processes, especially when persistent in time, could result in elevated resting heart rate and blood pressure, decreased heart rate variability, and increased ventricular arrhythmias and myocardial ischemia, and may explain why depressed older people are more vulnerable for subsequent health deterioration or for onset of some conditions.

Second, increased behavioral risk profiles in depressed people may explain their higher risk for adverse health consequences. Behavioral risk factors appear to cluster in the same individuals. Increased smoking and alcohol consumption are well documented in depression. Depressed people not only smoke more often, they are found to be less likely to quit smoking and might inhale more deeply and smoke more of the cigarette than nondepressed smokers. Also greater 24-hour caloric intake, and cholesterol/high-density lipoprotein cholesterol (HDL/C) ratio have been found. Depressed people engage less in physical activities such as walking, gardening, and vigorous exercise activities such as sports.

Third, related somatic symptoms of depression such as fatigue or sleeplessness may worsen the health status of older people. This explanation is in line with findings that vital exhaustion (which has shared common characteristics with depression such as listlessness, loss of energy, irritability, and sleep problems) increases the risk of developing cardiovascular disease. Another mechanism linking depression to subsequent health deterioration may be via psychological mechanisms. Depressed mood may impede recovery processes by discouraging people from obtaining adequate medical attention and rehabilitation and following treatment regimens.

In summary, depression has several adverse health consequences ranging from increased risks of mortality, decline in physical function, and onset of cardiovascular disease. Subsequent health deterioration can be due to physiological changes induced by depression, the unhealthier behavioral risk profile of depressed older people, the somatic accompaniments of depression, or the decreased motivation of depressed older people for self-care.

Conclusion

This article has described that, although the prevalence of major depressive disorder in old age is rather low, anxiety disorders and significant depressive feelings (minor depression) are common in old age. The prevalence of minor depression increases across older age. Major depressive disorders and anxiety disorders are strongly associated with long-standing vulnerability factors such as personality and heredity, whereas minor depression is more often a reaction to the stressors encountered in later life such as worsened health status, economic stress, and the changing social context. Several adverse health

consequences of depression have been described ranging from mortality, declined physical function, and cardiovascular disease.

Depression and anxiety are potentially modifiable conditions. Unfortunately, appropriate care is often not provided for mental problems in older people. This is partly due to the poor recognition of these problems by physicians. Significant feelings of anxiety or depression in older people more often go unrecognized and untreated than in younger people. Older people are more likely to attribute depressive and anxiety symptoms to physical problems and, therefore, tend to present their problems as somatic complaints, rather than as a psychological problem. Furthermore, the fear of stigmatization in older people may cause underreporting, and bias of clinicians (ageism or age-related prejudice) causes underrecognition of psychopathological disorders in late life. Consequently, recognition, prevention, and treatment are important points for intervention of affective disorders in older people.

See Also the Following Articles

Aging and Stress, Biology of; Anxiety; Depression and Manic-Depressive Illness; Glucocorticoids – Adverse Effects on the Nervous System; Depression, Immunological Aspects; Depression and Coronary Heart Disease.

Further Reading

- Birren, J. E., Sloane, R. B. and Cohen, G. D. (1992). *Handbook of mental health and aging* (2nd ed.). San Diego, CA: Academic Press.
- Blazer, D. G. (2003). Depression in late life: review and commentary. *Journals of Gerontology. Series A, Biological Sciences and Medical Sciences* 58, 249–265.
- Fielding, R. (1991). Depression and acute myocardial infarction: a review and reinterpretation. *Social Science and Medicine* 32, 1017–1027.
- Gatz, M., Kasl-Godley, J. E. and Karel, M. J. (1996). Aging and mental disorders. In: Birren, J. E. & Warner Schaie, K. (eds.) *Handbook of the psychology of aging*, pp. 365–382. San Diego, CA: Academic Press.
- Salzman, C. and Leibowitz, B. D. (1991). *Anxiety in the elderly: treatment and research*. New York: Springer Publishing.
- Schneider, L. S., Reynolds, C. F., Lebowitz, B. D., et al. (1994). *Diagnosis and treatment of late life depression*. Washington, D.C.: American Psychiatric Press.
- Schulz, R., Drayer, R. A. and Rollman, B. L. (2002). Depression as a risk factor for non-suicide mortality in the elderly. *Biological Psychiatry* 52, 205–225.

Aging and Stress, Biology of

M A Horan

University of Manchester Medical School, Salford, UK

R N Barton and G J Lithgow

University of Manchester, Manchester, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 111–117, © 2000, Elsevier Inc.

Aging and Cellular Stress Resistance
Stress Responses in Intact Organisms

Glossary

Biological aging

The progressive changes that occur during adult life, many of which are deleterious. Aging rate cannot be measured in individuals. For populations, aging rate is estimated by the slope of the line relating the logarithm of mortality rate to age (Gompertz acceleration parameter); it can also be expressed by the mortality rate doubling time (MRDT; closely related to maximum life span) – approximately 8 years for humans and 3 months for laboratory mice. However, in extreme old age in many species, this relationship breaks down.

Epistasis

The canceling of the phenotypic effects of one gene by the expression of another.

Population aging

Time-related changes in the age structure of populations, in which, due to a combination of increased life expectancy and reduced birth rates, the relative and absolute numbers of old people present in many human populations has greatly increased.

Psychological aging

Changes in behavior and mental processes that arise as individuals progress through life.

Senescence

Aging; sometimes used to refer only to those changes that are deleterious. Cellular senescence occurs when a phenotype of cells that are usually capable of mitosis lose the ability to replicate after a number of cycles of cell division but are not dead (Hayflick limit).

Aging *in vivo* is often related to damage done to molecules, cells, and organs by a variety of toxic factors, either produced endogenously or derived from the environment. For our purposes, potentially (or

actually) harmful conditions are considered as stressors and the psychological and biological reactions thereto as the stress response. In the discussion that follows, we address two important issues concerning the relationships between stress and aging: do stressors affect the rate of aging and does aging affect responses to stressors?

Aging and Cellular Stress Resistance

Stressors for cells include the nonenzymatic glycation of proteins and DNA, as well as oxidative damage to diverse molecules that could be caused by free radicals generated mainly as a by-product of metabolism, by solar radiation, by a variety of stressors that elicit the heat shock response, and by numerous other toxic factors in the environment. Normal function, and even survival, is dependent on the ability of cells to sustain homeostasis by being able to adapt to and/or resist stressors, and to repair or replace damaged molecules and organelles.

Oxidative Stress

Oxygen radicals are reactive molecular species produced by the one-electron reduction of oxygen. It is estimated that approximately 4% of the oxygen consumed by mitochondria is converted to the reactive oxygen species (ROS) hydrogen peroxide and superoxide during oxidative phosphorylation. These ROS can be converted by a number of distinct processes to highly reactive hydroxyl radicals, which can damage proteins, lipids, and DNA. Not surprisingly, organisms have important defenses against oxidative stress, including the small molecules uric acid, glutathione, ascorbic acid, and vitamin E. Oxidative stress induces the synthesis of several antioxidant enzymes including superoxide dismutase (SOD), catalase, and glutathione peroxidase. These enzymes protect cells and tissues against oxidative damage by converting ROS to nonreactive species.

What is the importance of defenses against oxidative damage during aging? The continued effectiveness of inducible responses to environmental damage are certainly thought to be major factors in resistance to diseases and even in the rate of aging. Indeed, it has been proposed that a capacity for increased life span is latent within most (if not all) species and might be uncovered by treatments that enhance cellular stress resistance. This is based on the observation that all known laboratory-based strategies that extend organism life span, whether by genetic

intervention or environmental manipulation, are associated with an increased ability to respond to environmental stressors and decreased susceptibility to stress-induced damage.

Systemic Stress Responses and Aging

One of the most exciting concepts relating stress and aging is that the rate of aging is governed by the rate of accumulation of macromolecular damage and, therefore, that processes determining this rate also define life span. Damage accumulates when repair or degradation of damaged molecules cannot keep pace with the rate of damage, which itself is determined by the production of damaging agents (e.g., ROS) and their subsequent detoxification (e.g., by antioxidant enzymes). This idea is well developed in the free-radical theory of aging and its subsequent modifications. The modern oxygen radical theory states that the decline in function and increasing mortality rate associated with aging are due to ROS production and their subsequent reaction with cellular components, causing irreparable damage. Despite considerable interest in this theory, there is virtually no evidence that it explains the natural aging rate. However, we are now developing a more general picture of how cellular stress resistance and aging rates might be related, mainly from studies of model genetic systems (*Saccharomyces cerevisiae*, *Caenorhabditis elegans*, and *Drosophila melanogaster*).

Fungal Systems and Aging

S. cerevisiae divides by asymmetric budding, leaving a mother cell and a smaller daughter cell. The mother cell exhibits limited division potential and consequently, a limited life span. At the end of the yeast life span, the cell becomes granulated and dies. Yeasts provide excellent systems for the dissection of complex biological problems, and aging is no exception. Many genes are now known to influence yeast life span. The first mutations isolated that conferred an increased life span came from a genetic screen for mutations that also conferred stress resistance, in this case to starvation. These mutations clustered in four genes, *UTH1*, *UTH3*, *UTH4*, and *SIR4*.

An important biomarker of aging, observed in species as diverse as primates and nematodes, is the accumulation of damaged mitochondrial DNA (mtDNA). It is thought that the cause of these altered forms is oxygen radical damage. The filamentous fungi *Podospora anserina* and *Neurospora crassa* have provided interesting information on at least one form of aging-related mtDNA mutation. In *Podospora*, aging is associated with the release of a circular DNA molecule (α senDNA) from the cytochrome

oxidase subunit I gene (COI) in the mitochondrial genome. Mutations in the nuclear genes *gr* (*greisea*) or *i viv* result in the repression of the production of this α senDNA. Combinations of these mutations result in indefinite life spans. The *gr* gene has been cloned and encodes a protein homologous to the angiotensin-converting enzyme (ACE)1 transcription factor, known to regulate the superoxide dismutase (*SOD1*) and metallothionein (*CUP1*) genes in budding yeast. Therefore, the alterations in stress-associated transcription factors cause dramatic alterations in the rate of accumulation of aging-associated damage, which correlates with life span.

Caenorhabditis elegans

C. elegans is a small (1.2-mm) soil-dwelling roundworm. It has a 3-day life cycle with four larval stages (L1–L4) before the final molt to the adult form. Most worms are hermaphrodites (XX), although males (X0) do exist at a ratio of approximately 1:500. The hermaphrodites reproduce by self-fertilization and live for approximately 20 days. This organism has proved to be very useful in aging research because single-gene mutations have arisen that both extend life span (*age* mutations) and increase resistance to environmental stressors.

The *age* mutations can bring about an up to three-fold increases in life span, attributable to a smaller acceleration of mortality rate with advancing age. Further, the mutations also induce increased resistance to several stressors: heat, ultraviolet (UV) radiation, oxidation, and heavy metals. Many, but not all, mutations that enhance resistance to such stressors are associated with increased life span, which suggests a direct relationship between longevity and stress resistance.

The epistasis analysis of *age* mutations has established two genetic pathways that determine life span. The cloning of genes in the first pathway, *age-1*, *daf-2*, and *daf-16* has shown that an insulin-like signaling pathway governs both aging and responses to stressors in this species. The *age-1* gene encodes a homolog of the mammalian phosphatidylinositol-3-kinase catalytic subunit (PI3K). The *daf-2* gene encodes a protein that is 35% identical to human insulin receptor and 34% identical to the insulin-like growth factor 1 receptor (IGF-1R). The *daf-16* gene is a Fork head transcription factor.

With the free-radical theory in mind, worms carrying either *age-1* (*bx546*) or *daf-2* (*e1370*) have been assessed for antioxidant enzyme activities, Cu/Zn SOD, and catalase. The net effect of these two enzymes is to prevent the production of the highly reactive hydroxyl radical, which is thought to be

responsible for age-associated damage to proteins, lipids, and nucleic acids. The activities of both of these enzymes are elevated in *age-1* and *daf-2* mutant worms in mid- and late life. This is consistent with the idea that *age* mutations result in resistance to intrinsic oxidative stressors. Worms mutant in *age-1* are certainly resistant to extrinsic oxidative stress, as measured by survival during exposure to hydrogen peroxide (H_2O_2) and the O_2 generator paraquat. These results suggest that the insulin signaling pathway negatively regulates antioxidant enzyme activity and that extended life span results from enhanced antioxidant defenses. This strongly suggests that over expression of antioxidant enzymes is sufficient to confer extended life span.

Oxidative stress is also thought to be one mechanism by which insulin action may be impaired. For example, it has been shown that whole-body glucose uptake (reflecting metabolic clearance rate, MCR) during a euglycemic glucose clamp declines with age in humans. On the basis of parallel increases in plasma superoxide production, red blood cell membrane microviscosity, and oxidation state of glutathione, it was postulated that insulin action is impaired by free-radical production. However, any such hypothesis based purely on correlation studies must be regarded as very tentative.

Drosophila melanogaster

D. melanogaster is an insect that has been widely used in both stress and aging studies. Transgenic *Drosophila* lines containing additional copies of the Cu/ZnSOD and catalase genes have been generated, and most lines that have elevated antioxidant-enzyme levels are also long-lived. This demonstrates the importance of these two enzymes in determining life span. More recent experiments suggest that the expression of MnSOD in neurons alone is sufficient to extend life span.

The transgenic *Drosophila* experimental system has also been used to test the role of molecular chaperones in determining aging rates. In these experiments, it was shown that additional copies of heat shock protein (HSP)-70 could reduce age-specific mortality rates in *Drosophila*. Just how HSP-70 retards aging is unknown, but these findings are consistent with other experiments in both *C. elegans* and *Drosophila* in which animals given a mild thermal stress and allowed to recover go on to develop acquired thermotolerance and exhibit extended life span. In these experiments, HSPs are induced and may be the principal cause of slowed aging.

Drosophila has also been used for studying the effects of multiple genes on stress responses and

aging. Selection has been undertaken for *Drosophila* lines that differ in their levels of stress tolerance, and at least one laboratory has demonstrated that selection for stress-resistance produces lines with increased longevity. In addition, some long-lived *Drosophila* lines, selected for late reproduction, are stress resistant.

The Heat Shock Response

The most extensively studied response to acute stressors is the expression of HSPs. HSPs are the products of a family of highly conserved genes, found from bacteria to mammalian cells, that are rapidly transcribed and translated in cells and tissues exposed to a variety of stressors. Recent studies have provided evidence that in the intact animal even a behavioral or psychological stress can elicit the response. HSP genes are regulated by heat shock transcription factors (HSF), which are themselves sensitive to a variety of stressors. The HSPs facilitate the disassembly and disposal of damaged proteins and the folding, transport, and insertion of newly synthesized proteins into cellular organelles. Thus, they facilitate the removal of damaged structures as well as their replacement.

There is very good evidence that the expression of HSP genes is sufficient to increase resistance to thermal stress (thermotolerance) in many circumstances. Across different species, distinct HSP families are important in thermal tolerance. HSP-104 is responsible for tolerance to several environmental stressors in the yeast *S. cerevisiae*, and inducible HSP-70 confers thermotolerance in *Drosophila* embryos and in a range of cultured mammalian cells. The addition of antibodies to HSP-70 compromises survival after thermal stress in fibroblasts.

Stress Proteins and Aging in Mammals

There is good evidence that stress proteins (e.g., HSPs and antioxidants) directly influence aging. It is also likely that the converse is true – aging seems to influence the extent to which stress proteins are synthesized in response to a variety of stressors. The expression of HSP-70 has been studied in inbred strains of animals at different ages and also in cells cultured to senescence. These studies show that there is a striking reduction in the magnitude of the induced response with both age and senescence.

Considerable specificity has been observed for responses to different stressors. For example, the physical restraint of an animal induces HSP-70 expression selectively in the adrenal cortex and in the media of blood vessels. These responses are dependent on the activation of the hypothalamic-pituitary-adrenal (HPA) axis and the sympathetic

nervous system, respectively. Old rats show a reduced HSP-70 response at both sites. Other work has shown a reduced HSP-70 response in the myocardium of old rats associated with reduced protection from ischemia in the Langendorff isolated-heart model. Interestingly, an HSP response may be selective for the nature of the stressor. For example, the hepatic HSP-70 response of old rats to passive heat stress is attenuated, whereas that to exertional heat stress is not.

HSPs are also important in the cellular actions of glucocorticoids, the archetypical stress hormones. High-affinity intracellular binding of glucocorticoids (e.g., estrogens and gestagens) requires that their receptors be present as a heteromeric complex containing a HSP-90 dimer; other proteins such as HSP-70, p59/HSP-56, and p23 may also be present, although HSP-70 appears to be involved in the assembly of the complex rather than a component of it. This complex also serves to prevent the receptor from homodimerization and binding to chromatin and possibly serves to inhibit its degradation.

Stress Responses in Intact Organisms

There is a diverse literature from which information can be gleaned on how complex body systems respond when an organism is subjected to stressful stimuli, but we have chosen to concentrate mainly on those aspects to which we have contributed directly. Experimental studies show that old rats and mice are dramatically more sensitive than the young to the lethal effects of bacterial lipopolysaccharide (endotoxin). Similarly, old, multiply injured humans experience a very high death rate from injuries of a severity that would cause only sporadic deaths in the young.

Laboratory Studies

It has long been known that old age is a risk factor for developing malignant systemic inflammation during infections and after tissue injury (systemic inflammatory response syndrome, SIRS). However, there have been few systematic studies and consequently little information on aging and the inflammatory response, except for a number of *in vitro* studies on neutrophil function.

Methodological problems obscure firm conclusions, but the majority of *in vitro* studies have reported impaired neutrophil responses to a variety of stimuli. In contrast, other *in vivo* studies suggest that inflammation may be exaggerated or dysregulated. We have shown old rats to be dramatically sensitive to the lethal effects of endotoxins and that

this was associated with early and marked neutrophil infiltration of the lungs and liver. Similarly, in a model of pneumonia in mice, the old animals had much more marked inflammatory infiltrates. Our studies of experimental human cutaneous wound healing also showed greater and prolonged neutrophil infiltration in the old, together with lesser and delayed macrophage infiltration.

Although the endocrine and metabolic consequences of physical injury or experimental infections have not been studied in the context of aging, the responses to the lesser stress of physical restraint have. This study reported a delayed return to baseline corticosterone concentrations in the old rats. It was hypothesized that, because glucocorticoids are toxic to hippocampal neurons and the hippocampus exerts a tonic restraint on activation of the HPA axis, repeated stressful episodes promote hippocampal cell loss, reduce tonic restraint, and predispose to exaggerated HPA axis activation in the old animals.

Endocrine and Metabolic Responses

A characteristic pattern of change in hormone secretion and the flux of metabolites is seen in a stressed animal. Typically, the sympathoadrenal and HPA systems are stimulated, together with the secretion of glucagon and the pituitary hormones growth hormone (GH), prolactin, and vasopressin. In contrast, there is a suppression of adrenal androgen production, and complex changes take place in the hypothalamic-pituitary-gonadal and hypothalamic-pituitary-thyroid axes that lead to lower concentrations of at least some of their end hormones. As regards metabolites, glycogen and fat are broken down so that the concentrations of glucose, lactate, nonesterified fatty acids (NEFA), glycerol, and ketone bodies increase in the plasma.

In muscle the persistence of stressful stimuli cause net protein breakdown, whereas in the liver the synthesis of a set of plasma proteins known as the acute-phase reactants (APR) is induced.

Any effects of stressors is superimposed on a background of normal age-related changes. In brief, there is little change in plasma adrenocorticotrophic hormone (ACTH) or cortisol during aging in humans. The concentration of norepinephrine, which is derived from sympathetic nerve terminals as well as the adrenal medulla, tends to rise whereas that of adrenaline does not. There is decreased secretion of GH, which takes place mainly at night, but probably not of prolactin or glucagon. Aging is accompanied by a decline in the turnover rate of glucose and a small increase in its concentration, reflecting insulin resistance. There is a tendency for the concentrations of

fat metabolites also to increase, but the changes in their turnover rates are more controversial. Protein turnover decreases, as we expect from the decline in lean body mass with age.

Responses to hypoglycemia The effect of aging on the endocrine responses to hypoglycemia has been of considerable interest because of the possibility of hypoglycemic episodes in treated elderly diabetics. Earlier studies, using a standard insulin tolerance test (0.1 U kg^{-1}), mostly showed no difference between old and young people, although there are sporadic findings of decreased cortisol, GH, glucagon, and adrenaline responses and an exaggerated rise in norepinephrine in the elderly.

More controlled experiments with the infusion of insulin have confirmed this lack of consistent change in endocrine responses to hypoglycemia, but have shown a tendency for fewer signs of sympathoadrenal activation (e.g., increased heart rate) in the elderly that is not always reflected in diminished responses of plasma catecholamines. There is no unanimity over their sensitivity to the symptoms of neuroglycopenia.

Responses to psychological stress A variety of procedures have been used to study the effect of age on responses to psychological stress in the laboratory, nearly all of which have shown reduced heart rate responses in the old. A number of authors have measured catecholamines, finding a tendency for the norepinephrine, but not the adrenaline, response to increase with age as well as an increase in baseline values. There have been few studies of HPA responses to psychological stress in the elderly, and these demonstrate no failure to activate the system.

In only one study of the effects of chronic stress do different age groups appear to have been compared directly. Urinary free cortisol excretion has been measured in subjects ages 47–78 whose spouses had life-threatening illnesses. There was a correlation with age only in those who were identified as suffering from severe depressive illness. Others have studied elderly people caring for spouses with Alzheimer-type dementia and compared them with spouses not needing to provide care; the caregiving spouses were classified as high or low stress. Plasma norepinephrine was higher in the high-stress group than the others; there were no differences in plasma adrenaline, cortisol, or the sensitivity or receptor density of lymphocyte β -adrenergic receptors. Although younger subjects were not studied, the results were contrasted with the β -adrenergic downregulation observed in young men experiencing the chronic stress of homelessness. A further study reported a tendency for

higher insulin (but not glucose) concentrations in those caring for demented spouses.

Response to trauma

The general response The acute rises in cortisol and norepinephrine after elective surgery tend to be exaggerated in old people, whereas the opposite is seen for pituitary hormones. During the first few hours after moderate to major surgery, plasma cortisol in elderly patients was either similar to or significantly higher than its value in younger ones. There was no difference in plasma ACTH, whereas the GH response was lower in the elderly. However, in men ages 43–77 undergoing colectomy, there was no correlation between plasma cortisol and age. Cortisol and ACTH concentrations also tended to be higher in old than in young patients after cholecystectomy. After a minor procedure (inguinal herniorrhaphy), too, plasma cortisol was higher in the elderly, whereas there was no difference in the response of norepinephrine, adrenaline, or aldosterone.

Over the next few days after injury, plasma cortisol and, perhaps, norepinephrine tend to decline to normal levels more slowly in the elderly than in the young. After moderate to major surgery, cortisol continues to be higher in the old; no corresponding change in ACTH is detectable. There is also no difference in plasma GH during this period. The rate of 17-ketogenic steroid excretion (a crude index of cortisol production rate) was raised for a longer period after surgery in older patients, even though the absolute values were not different between young and old. In apparently the only study of the effect of aging responses very soon after accidental injury, we found no difference in plasma cortisol among patients ages 17–40, 41–65, and 66–92 with a wide range of injury severities.

Hypothalamic-pituitary-adrenal axis activation after hip fracture We have studied hip-fracture patients, who are often frail and in poor health. Cortisol concentrations remain elevated 2 weeks later (unlike younger people with similar severity injuries). The same was true of the cortisol concentrations after they took dexamethasone overnight (suggesting impaired feedback inhibition by cortisol). These differences persisted when further measurements were made 8 weeks after the injury. Such a prolonged period of raised cortisol concentrations could have deleterious systemic effects.

Most reported cortisol concentrations have been measured in the morning, and comparisons would be invalidated if there were changes in its nychthemeral rhythm. Reliable indices of HPA function can be based on urinary measurements made over a 24-h

period. We found that hip-fracture patients had a higher cortisol production rate and urinary free-cortisol excretion rate than healthy elderly women. The latter index, which is particularly useful because it usually reflects the integrated concentration of free, biologically active cortisol in plasma, rose on average approximately threefold. The cortisol production rate rose proportionately less because the metabolic clearance rate of cortisol was lower in the injured patients, so only a modest rise in production rate was necessary to account for the observed increases in plasma cortisol concentration and free-cortisol excretion.

A puzzling aspect of our studies is that hip-fracture patients did not show an increase in plasma ACTH. In fact, when concentrations were measured at the end of a 2-h baseline period leading up to injection of corticotropin releasing hormone (CRH), decreased values were seen compared with the healthy elderly. The total areas under the cortisol- and ACTH-time curves after the injection of CRH reflected this disparity, with that for cortisol being greater and that for ACTH smaller in the patients, although the incremental areas were both decreased. The correlation between cortisol and ACTH seen in the control subjects was completely lost in the hip-fracture patients, and their apparent sensitization to ACTH was not reflected in the slope of the dose-response curve. Combined with a lack of change in the response to small doses of exogenous ACTH¹⁻²⁴, this suggests the presence of a stimulus acting independently of ACTH. However, it is very difficult to know what such a stimulus might be.

Myocardial infarction There have been two comparisons of responses to acute myocardial infarction

in elderly and younger patients. One study found no difference in plasma cortisol or ACTH, measured over a period of 12 days after the event, between patients ages 65–81 and 33–55; there was also no difference in the response to a standard ACTH stimulation test on the twelfth day. The other study reported similar peak C-reactive protein (CRP) levels in patients younger than 60 and older than 70. This study is valuable because it appears to be the only attempt to compare acute-phase protein concentrations between age groups after a standard insult, although it is well known that an acute-phase response occurs in old patients. Comparability is important to establish because CRP levels have been used to distinguish the effects of aging from those of underlying inflammatory disease.

See Also the Following Articles

Aging and Psychological Stress; Heat Shock Response, Overview; Hypoglycemia; Immune System, Aging.

Further Reading

- Horan, M. A. and Little, R. A. (1998). *Injury in the aging*. Cambridge, UK: Cambridge University Press.
- Johnson, T. E., Lithgow, G. J. and Murakami, S. (1996). Hypothesis: interventions that increase the response to stress offer the potential for effective life prolongation and increased health. *Journal of Gerontology* 51, B392–B395.
- Lithgow, G. J. (1996). Invertebrate gerontology: the age mutations of *Caenorhabditis elegans*. *Bioessays* 18, 809–815.
- Ricklefs, R. E. and Finch, C. E. (1995). *Aging: a natural history*. New York: Scientific American Library.

Agoraphobia See: Panic Disorder and Agoraphobia.

AIDS

M H Antoni

University of Miami, Coral Gables, FL, USA

D G Cruess

University of Connecticut, Storrs, CT, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M H Antoni and D G Cruess, volume 1, pp 118–125, © 2000, Elsevier Inc.

Stress and AIDS

Stress, Immunity, and Health in HIV Infection

Stress Management for HIV-Infected Individuals

Glossary

<i>Acquired immune deficiency syndrome (AIDS)</i>	A global disease of epic proportions that overwhelmingly affects young people, rendering them unable to defend against common pathogens.
<i>Antibody</i>	A molecule, produced by B lymphocytes in response to an antigen, that has the particular property of combining specifically with the antigen that induces its formation.
<i>Antigen</i>	A molecule that induces the formation of an antibody.
<i>Cytokine</i>	A generic term for soluble molecules that mediate the interactions between cells of the immune system.
<i>Cytotoxicity</i>	The ability to kill cells.
<i>Human immunodeficiency virus (HIV)</i>	The causative agent of AIDS; it is a retrovirus that works by systematically disarming multiple components of the immune system, leaving the infected person susceptible to many of the lethal infections and cancers that encompass AIDS.
<i>Interleukins</i>	A class of molecules involved in signaling between cells of the immune system.
<i>Lymphocytes</i>	A class of immune cells, central to all adaptive immune responses, that specifically recognize individual pathogens, whether they are inside host cells or outside in the tissue fluids or blood. Lymphocytes fall into two basic categories: T lymphocytes (T cells) and B lymphocytes (B cells).
<i>Pathogen</i>	An organism that causes disease.
<i>Phagocytosis</i>	The process by which cells engulf material and enclose it within a vacuole in the cytoplasm.

Proliferation

The clonal expansion of lymphocytes resulting from mitogen or antigen challenge.

T-helper/inducer (CD4) cells

A functional subclass of T cells that can help to generate cytotoxic T cells and that cooperate with B cells in the production of antibodies.

Stress and AIDS

Acquired immune deficiency syndrome (AIDS) is the leading cause of death among men and women ages 15–44 in the United States. Human immunodeficiency virus (HIV) infection can be viewed as a chronic disease whose clinical course is dependent, in part, on an infected individual's ability to maintain optimal cell-mediated immunological control over extant and encountered pathogens. There is mounting evidence linking psychosocial stressors with HIV infection and AIDS. Many psychosocial and behavioral phenomena, including stressful events, stress responses, and distress states, have been associated with impairments in the immune system, which could potentially exacerbate the immunological abnormalities and health consequences observed among HIV-infected individuals. Such associations may be mediated, in part, by stress-related alterations in neuroendocrine functioning (e.g., adrenal hormone regulation), on the one hand, or health behaviors (e.g., substance use, self-care), on the other. Specific immunological components important in contributing to the course of HIV infection have also been shown to be associated with stressors and distress states.

How HIV Infection Affects Immune System Functioning

T-helper/inducer lymphocyte distribution HIV binds with the CD4 receptor, which is most prevalent on a type of T lymphocyte called a T helper/inducer, or CD4, cell. As the infection progresses, there are decrements in CD4 cell counts and impairments in the ability of T lymphocytes to respond to antigenic challenge, opening the door to a wide range of opportunistic infections and cancers. HIV epidemiological and clinical immunology research indicates that precipitous drops in CD4 cell counts, particularly when they fall below 200 cells mm⁻³, are highly predictive of the onset of physical symptoms and general clinical decline. Because CD4 cells are depleted in the

advancing stages of HIV infection, increases in functional aspects of lymphocytes might be important in predicting those HIV-infected individuals (with low CD4 counts) who develop opportunistic infections quickly. Some of these functions include the ability to kill tumors and viruses, a process called cytotoxicity.

Cytotoxic functions One type of psychologically sensitive immune function believed to offer primary protection from replicating HIV, as well as surveillance over other pathogens, is cytotoxicity. High-impact psychosocial stressors such as bereavement, which is quite prevalent among HIV-infected individuals and those who are at risk, appear to be associated with decreases in a specific type of cytotoxicity called natural killer cell cytotoxicity (NKCC). Impairments in NKCC may be especially important during HIV infection because CD4 cells are depleted both quantitatively and qualitatively. Natural killer cells may compensate to some degree for HIV-induced CD4 cell deficiencies. Although other cells such as CD8+ lymphocytes may also perform key cytotoxic functions in HIV+ people, less is known about their associations with psychosocial factors.

Herpesvirus infections Some immune measures may reflect the potential contribution of reputed cofactors of HIV disease progression, including rising antibody titers to herpesviruses such as Epstein-Barr virus (EBV), reflecting viral reactivation. Herpesvirus reactivation may transactivate HIV-infected cells, contributing to increases in HIV viral load. HIV-infected individuals who experience elevated life stressors, distress, depression, an inability to cope with these stressors, and/or a sense of social isolation may be at greater risk for reactivation of specific herpesviruses. Both NKCC and T-cell-mediated cytotoxic functions are generally responsible for keeping these viruses in check. Conversely, stress-associated impairments in cytotoxicity may facilitate herpesvirus reactivation in HIV-infected individuals with subsequent HIV replication and progression to AIDS.

Cytokine dysregulation Maintaining adequate cytotoxic capabilities may depend on retaining T helper cell type 1 (TH1) predominance and production of cytokines, or immunohormones, such as interleukin-2 (IL-2) and IL-12. However, a shift to TH2 predominance and the production of cytokines such as IL-4 and IL-10 occurs in HIV infection. TH1s stimulate both specific (T-cell-associated) and nonspecific (NK-cell-associated) cytotoxicity by producing IL-2 and γ -interferon (γ -IFN). IL-12, a recently discovered cytokine produced by monocytes, may also be important

because of its ability to stimulate NK cells and T cells to release γ -IFN. In addition to affecting cytotoxicity, stressors and distress states have been associated with decrements in other key immune system functions, such as phagocytosis and lymphocyte proliferation, which may also be important in HIV infection.

Stress, Psychoneuroimmunology, and HIV Infection

The field of psychoneuroimmunology (PNI) examines how stressors, stress responses, and distress states relate to the immune system by way of neural and endocrine processes. Our perception of the nature of a stressor and the availability of a coping response to that stressor accompany a series of physiological events leading to specific autonomic, neuroendocrine, and neuropeptide changes. Although neuroendocrine and nervous system mediators of stressor-associated changes in immune system functioning have been identified in humans, most work has occurred outside the HIV arena.

Endocrine-immune interactions Many different neurotransmitters, neurohormones, and neuropeptides within the central nervous system (CNS) have been studied as potential mediators of stressor-induced immunomodulation. Two well-delineated stress response pathways, the hypothalamic-pituitary-adrenal (HPA) axis and the sympathoadrenomedullary (SAM) system, have received the most attention. Elevations in glucocorticoids, a product of the HPA axis, enhance HIV replication, are immunosuppressive, and also accompany affective-disorder and stressful experiences. Once glucocorticoids such as cortisol bind to cytoplasmic receptors on immune cells, the complex travels to the nucleus and binds to DNA sequences that control glucocorticoid-regulated transcription genes for cytokine synthesis. Elevated levels of cortisol are associated with impaired immune-system functioning and accompanying decreases in TH1-like cytokine production. Current PNI research examines how distress- or depression-associated cortisol elevations impact immune function (e.g., NKCC) in HIV-infected individuals and how these changes might relate to disease progression by way of impaired surveillance of latent pathogens such as the herpesviruses.

Elevations in peripheral catecholamines, such as norepinephrine (NE), a product of SAM activation, also depress immune functioning via β -adrenergic receptors on lymphocytes. Sympathetic noradrenergic fibers innervate both the vasculature and parenchymal regions of lymphocytes and associated cells in several lymphoid organs, where NE terminals are generally directed into T-lymphocyte zones. Alterations in electrical activity and NE metabolism in the

hypothalamus occur following injection of antigen or cytokine products of T lymphocytes stimulated *in vitro*, suggesting bidirectional communication between the nervous and immune systems.

Lymphocytes are known to have receptors for many neurohormones including serotonin and cholinergic and β -adrenergic agonists, with most evidence for receptors involving β -adrenergic sites. Administration of β -adrenergic agonists has been associated with decreases in mouse and human NKCC and decreased T-lymphocyte proliferation, and these effects appear to be mediated by increases in intracellular cyclic AMP (cAMP). Patients treated with glucocorticoids show increased β -adrenergic responsiveness due to an increased receptor number and enhanced receptor-adenyl cyclase coupling, evidencing a synergy between SAM and HPA axis-generated stress hormones. *In vivo* human studies involving infusion of adrenergic agents have also shown a decrease in leukocyte number and decreased lymphocyte proliferative response to mitogens following epinephrine administration. These immunomodulatory effects are also most likely mediated by cyclic nucleotide action. In addition, agents that increase lymphocyte cAMP (e.g., isoproterenol and prostaglandin E) can impair lymphocyte proliferation and cytotoxicity, whereas agents that increase lymphocyte cGMP levels (e.g., carbamylcholine) enhance lymphocyte proliferative and cytotoxic abilities.

Stress hormones may have profound effects on monocyte/macrophage cell lines and related cytokines (e.g., IL-1). Such effects could have implications for upstream activities such as antigen processing and presentation to CD4 cells. In addition to catecholamines and corticosteroids, pituitary and adrenal peptide stress hormones, such as met-enkephalin, β -endorphin, and substance P, have been shown to stimulate T-cell cytotoxicity, NKCC, and γ -interferon production as well as macrophage functioning.

Stress, endocrine regulation, and HIV infection The regulation of adrenal hormones may have direct relevance for the pathogenesis of HIV infection. One study found that the ability of HIV to infect nonnal human lymphocytes was enhanced by supplementing the cell culture medium with corticosteroids. Viral susceptibility may be related to the efficacy of lymphocyte or NK cytotoxic abilities, which may be compromised by glucocorticoid secretions during stress. Recent work suggests that HIV-infected people have HPA axis abnormalities, which may increase with the progression of disease, favoring a shift from TH1-like to TH2-like cytokine production. HIV-infected people display elevated resting levels of

cortisol, which are associated with lower CD4 cell counts and decreased NKCC. Cortisol specifically promotes decreased production of TH1-like cytokines such as IL-2 while promoting increased production of TH2-like cytokines such as IL-4 and IL-10. Because cortisol can induce programmed cell death in mature lymphocytes, the HPA axis may also play a role in the progression of HIV infection by facilitating the destruction of viable cells.

Through repeated exposure to multiple stressors, HIV-infected people may become particularly vigilant, physically tense, and sympathetically activated for extended periods of time. Those undergoing chronic sympathetic nervous system arousal with subsequent release of NE from sympathetic nerve terminals may experience health decrements because these endocrines may downregulate the proliferation of naive (uncommitted) T lymphocytes. These cells are already significantly depleted in HIV infection, even after successful antiretroviral treatment. Diminished naive T-lymphocyte proliferation may render HIV-infected people less able to respond to novel pathogens, opening the door to opportunistic infections. In sum, the literature supports the notion that elevated levels of cortisol and catecholamines are associated with impaired immune system functioning. Distress or depression-associated neuroendocrine changes impact immune function (e.g., NKCC) in HIV-infected individuals, and such functional changes might relate directly to HIV disease progression by way of impaired surveillance of latent and novel pathogens, which, once reactivated, can stimulate HIV proliferation.

Stress, Immunity, and Health in HIV Infection

HIV-infected people encounter a large number of major stressors, which, based on the way they perceive them, may result in different emotional responses. These affective responses may relate to altered levels of certain neuroendocrine substances (e.g., cortisol and catecholamines) that have been shown to have depressive effects on lymphocytes. These stress responses may also be accompanied by behaviors (e.g., substance use, unprotected sex, and poor medication adherence) that could have negative implications for the health of HIV-infected people. We now summarize the evidence relating depression and health behaviors to immune function and the evidence relating immune alterations to stress-intervening variables (stressor appraisals, coping responses, and social supports) that may influence physical health status.

Depression and Immunity in HIV Infection

Depression and depressed affect are associated with immunosuppression in healthy individuals. Among HIV-infected people, some studies have found that a depressive state was unrelated to changes in CD4 counts or HIV-related symptoms over 6- and 12-month periods, respectively. Other studies examining immunological and health changes over longer periods found interesting but conflicting results. One found that gay men scoring above the criterion for depression showed a faster rate of decline in CD4 cell counts over a 66-month follow-up period than those scoring below this value. This association was strongest in men who were at the earliest stages of disease. In a separate cohort showing more progressed disease at the time of the depression evaluation, these findings were not evident.

Another study showed that chronically depressed affect (sustained severe depressed mood over a 2-year period) predicted declines in CD4 counts among HIV+ gay men without AIDS as compared to a group of nonchronically depressed men matched on age and CD4 levels. Subsequent work, incorporating more sophisticated statistical procedures and time-dependent covariates (including medication regimen), showed that depression, in combination with elevated stressful life events, does in fact predict a faster rate of disease progression in HIV+ men for up to 9 years. Depressive symptoms among women with HIV have also been shown to predict accelerated rates of CD4 cell decline and HIV-related mortality over a 7-year period. Resolution of major depression is paralleled by increases in NK cell cytotoxicity among HIV+ women. Depressed affect may be associated with an accelerated decline in CD4 counts and a faster rate of clinical disease progression among HIV-infected men and women. This association appears stronger when the potential confounding somatic effects of the disease are controlled by either sampling (excluding progressed subjects), questionnaire refinement (excluding somatically based depression items from the test battery), or statistical controls. Beyond studying the influence of depressive symptoms, it is also important to consider the role of anxiety-related symptoms (e.g., tension and rumination) in the quality of life, immune status, and physical health of HIV-infected individuals.

Health Behaviors and Immunity in HIV Infection

Drug and alcohol use is associated with a wide variety of immunological effects that may have direct health implications for HIV-infected individuals. Alcohol use is related to decreased lymphocyte number and

proliferative responses. Alcoholism and related liver disease is associated with impaired NKCC and decreased NK cell numbers. Evidence suggests that alcohol abuse may amplify the suppressive effects of depression on NKCC. The use of substances ranging from injected drugs to cigarettes has been associated with faster disease progression in HIV-infected people. Thus, it is plausible that substance abuse may contribute to a decline in immunological status and possibly faster disease progression.

Little is known about the direct immunomodulatory effects of different forms of sexual behaviors, but it is plausible that repeated exposure to novel strains of HIV and other sexually transmitted pathogens, such as herpes simplex virus (HSV) and human papillomavirus (HPV), may contribute directly to the development of opportunistic infections (e.g., systemic HSV-2 infections) and cancers (e.g., HPV-associated cervical and anal intraepithelial neoplasias). In addition, exposure to these pathogens may indirectly contribute to accelerated disease course by promoting HIV replication.

Maintaining positive health behaviors such as adequate sleep, nutrition, exercise, and, perhaps most important, adherence to prescribed medications are a central part of managing HIV infection. Triple combination antiretroviral therapy with protease inhibitors offers some hope that viral replication can be slowed more efficiently and with a lower likelihood of resistance or undue side effects than prior antiretroviral regimens. Failure to adhere to the rigid and demanding medication regimens can substantially increase the risk of developing drug-resistant strains of the virus. The resulting treatment philosophy, now more than ever, views HIV infection as a chronic disease in which patient management is critical. Incomplete adherence cannot only compromise the effects of the particular protease inhibitor being used, but can also reduce the efficacy of other related compounds to which cross-resistance has developed. The key is consistent adherence to a demanding medication schedule in the context of an already stressful daily existence. Because depressed affect, substance use, risky sexual behaviors, and poor adherence to HIV medications may all contribute to adverse health outcomes among HIV-infected individuals, it is also important to understand the factors that intervene on these phenomena.

Intervening Variables

The mounting stress of HIV infection may cause infected people to appraise their world as uncontrollable, resulting in their inability to deal with

additional challenges and in a greater likelihood of depression, substance use, risky sexual behaviors, and poor self-care. This sequence is more likely among those with inadequate coping skills and sustained social support losses. Several psychosocial factors, such as perceived loss of control, maladaptive coping strategies, and social isolation, have been related to impaired immune functioning in a wide variety of healthy populations. These processes (and intervention-associated changes in these processes) may be associated with immunological alterations in HIV-infected people. Recently, studies have demonstrated a possible link between these psychosocial processes and the physical course of the infection.

Perceived loss of control Various stressors, especially those viewed as uncontrollable, have been associated with decrements in immune function in healthy people. Some of these are particularly prevalent among HIV-infected people attempting to cope with their illness and include bereavement, the stress of being a caregiver for a loved one with a terminal disease, and divorce or break-up. Among HIV+ men, those with more cumulative stress had faster disease progression over a 9-year period. Some work suggests that stress may affect immune declines in HIV+ people by disrupting sleep.

Coping strategies Coping strategies, such as active coping, active confrontation, “fighting spirit,” and denial, have been related to changes in physical health during HIV-infection. The use of active coping was associated with decreased symptom development in HIV-infected gay men. A “fighting spirit” coping strategy also predicted less symptom development, whereas denial coping predicted a greater number of HIV-related symptoms over a 12-month period. Other coping strategies such as avoidance and denial have also been associated with declines in immune function in HIV+ men. During the stressful period of HIV antibody testing, greater avoidance of distressing AIDS-related thoughts predicts greater anxiety, depression, and confusion as well as lower lymphocyte proliferative responses and NKCC after notification of positive serostatus. Increases in the use of denial during the postnotification period also predict a faster progression to AIDS over a 2-year follow-up period. In another cohort of HIV+ men, greater use of denial predicted faster progression to AIDS over a 7.5-year period. Gay men with AIDS who used a strategy called realistic acceptance showed a shorter survival time than those not endorsing this strategy. Realistic acceptance appears to reflect excessive rumination and focus on HIV-related matters, almost to

the exclusion of many other facets of life. It may be that while denial is maladaptive and predictive of poorer health outcomes, obsessing and ruminating about the disease may also have negative health ramifications. Newer work has produced evidence that greater use of distraction as a coping strategy predicts a slower subsequent course of disease over a 7-year follow-up period. Distraction here reflects an effort to focus on living and pursuing activities that are meaningful rather than denying the existence of the infection or refusing to focus on it excessively. This sort of balance between acceptance and nonrumination has also been identified as a common trait of those who are long-term survivors of HIV infection and AIDS.

Social isolation HIV-infected people face a variety of stressors and challenges that they must continually negotiate in an effort to deal with the potentially overwhelming psychological and physical health consequences of their illness. Stigmatization, alienation from friends and family, increased reliance on medical personnel, social disconnectedness, and gradual deterioration of physical health status and immune functioning may leave these individuals feeling socially isolated. Social resources have been found to buffer the negative physical and psychological effects of major life-threatening events. Low social support availability predicts a faster decline in CD4 counts over a 5-year period among HIV-infected men with hemophilia. Larger social network size and greater informational support provided by that network predict greater survival time in HIV-infected men but only among those with AIDS. Greater perceived social support also predicts a slower progression of disease in initially asymptomatic HIV+ gay men over a 9-year period. It remains to be demonstrated whether actively intervening to increase the quality of HIV-infected people’s social support influences their long-term psychological adjustment, immune functioning, physical health status, and subsequent disease progression.

To summarize, HIV infection is a chronic disease often associated with severe psychosocial and physical stressors, which can potentially overwhelm those infected in a number of domains. The stressors that encompass HIV and AIDS impact the way in which those infected appraise day-to-day situations, choose coping strategies, acquire support, and deal emotionally with stressful events in their lives, which in turn can alter stress hormones as well as influence their ability to engage in healthy behaviors. These processes can then lead to alterations in immune system functioning, which may have repercussions for the development of symptoms and increased HIV viral

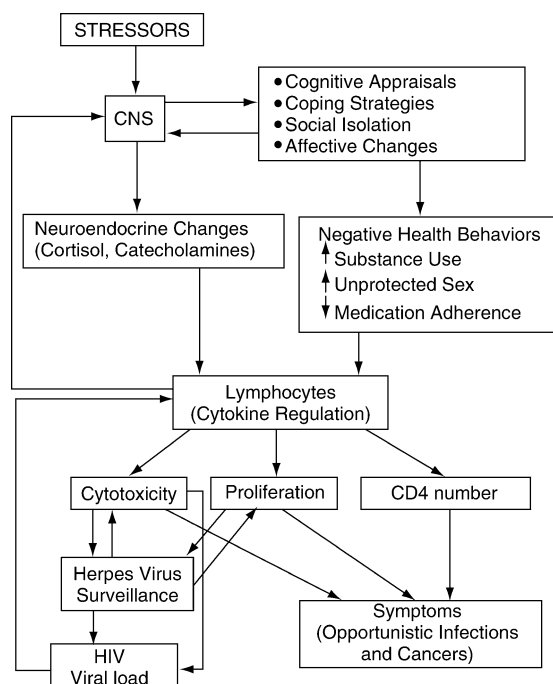


Figure 1 Model relating stressors, psychosocial factors, neuroendocrine changes, health behaviors, immune status, and disease in HIV infection. CNS, central nervous system; CD4, T-helper lymphocytes; HIV, human immunodeficiency virus.

load. **Figure 1** displays the interconnectedness of these multiple factors in HIV disease.

Stress Management for HIV-Infected Individuals

Stressors, stress responses, coping strategies, and social resources are related to psychological adjustment, immune system function, and the physical course of HIV infection. Stress-management interventions that modify how individuals process and deal with stressors help to reduce distress and facilitate adjustment and buffer the stressors on the immune system. We now know from randomized clinical trial that HIV+ people assigned to a time-limited group-based stress management intervention show reductions in negative mood, improved coping (e.g., less denial), and greater social support as well as decreases in stress hormones such as cortisol and NE, greater antiviral immune function against herpes viruses, and greater signs of immune system reconstitution. Greater improvements in mood and reductions in stress hormones predicted the greatest gains in immunity, thus providing support for a PNI model in HIV infection.

Some of these gains appear to be brought about by the ability of these interventions to modify stressor appraisals, improve coping strategies, and enhance

social support in HIV-infected people. Psychosocial interventions that provide information, skills, and support to infected people can help normalize neuroendocrine levels and can also facilitate the adoption and maintenance of positive health behaviors, including better adherence to demanding antiretroviral medication protocols. Because interventions may also decrease denial and depressed affect, they may, in turn, decrease negative health behaviors (e.g., substance use and risky sex) that are associated with immune system decrements. The improvements in immune function that have been associated with these psychosocial interventions raises the interesting question of whether stress management can help reconstitute the immune system once antiretroviral agents have contained HIV. Although we are now learning much about how stressors and stress management interventions influence HIV disease in gay men, far less is known about other HIV-infected populations that are growing in numbers (e.g., poor minority women and injection-drug users). Understanding the mechanisms through which stressful experiences and stress management interventions contribute to quality of life and the physical course of HIV infection may help the growing populations of infected people who are attempting to manage this chronic disease.

See Also the Following Articles

Autoimmunity; Avoidance; Herpesviruses; Immune Suppression.

Further Reading

- Ader, R., Felton, D. and Cohen, N. (2001). *Psychoneuroimmunology* (3rd edn.). New York: Academic Press.
- Antoni, M. H. (2003). Stress management effects in psychological endocrinological and immune function in men with HIV: empirical support for a psychoneuroimmunological model. *Stress* 6, 173–188.
- Antoni, M. H., Esterling, B., Lutgendorf, S., et al. (1995). Psychosocial stressors, herpesvirus reactivation and HIV-I infection. In: Stein, M. & Baum, A. (eds.) *AIDS and oncology: perspectives in behavioral medicine*, pp. 135–168. Hillsdale, NJ: Lawrence Erlbaum.
- Antoni, M. H., Schneiderman, N., Fletcher, M. A., et al. (1990). Psychoneuroimmunology and HIV-I. *Journal of Consulting and Clinical Psychology* 58, 38–49.
- Antoni, M.H., Caricco, A., Duran, R., et al. (in press). Randomized clinical trial of cognitive behavioral stress management on HIV viral load in gay men treated with HAART. *Psychosomatic Medicine*.
- Chrousos, G. (1995). The hypothalamic pituitary adrenal axis and immune mediated inflammation. *New England Journal of Medicine* 332, 1351–1362.

- Cruess, D. G., Douglas, S. D., Petitto, J. M., et al. (2005). Association of resolution of major depression with increased natural killer cell activity among HIV-seropositive women. *American Journal of Psychiatry* **162**, 2125–2130.
- Glaser, R. and Kiecolt-Glaser, J. (1987). Stress-associated depression in cellular immunity: implications for acquired immune deficiency syndrome (AIDS). *Brain, Behavior & Immunity* **1**, 107–112.
- Glaser, R. and Kiecolt-Glaser, J. (1994). *Handbook of human stress and immunity*. New York: Academic Press.
- Herbert, T. and Cohen, S. (1993). Stress and immunity in humans: a meta-analytic review. *Psychosomatic Medicine* **55**, 472–486.
- Ickovics, J. R., Hamburger, M. E., Vlahov, D., et al. (2001). Mortality, CD4 cell count decline, and depressive symptoms among HIV-seropositive women: longitudinal analysis from the HIV Epidemiology Research Study. *Journal of the American Medical Association* **285**, 1466–1474.
- Ironson, G., Solomon, G., Cruess, D., et al. (1995). Psychosocial factors related to long-term survival with HIV/AIDS. *Clinical Psychology and Psychotherapy* **2**, 249–266.
- Klecolt-Glaser, J. and Glaser, R. (1995). Psychoneuroimmunology and health consequences: data and shared mechanisms. *Psychosomatic Medicine* **57**, 269–274.
- Leserman, J., Petitto, J., Gu, H., et al. (2002). Progression to AIDS, a clinical AIDS condition and mortality: psychosocial and physiological predictors. *Psychological Medicine* **32**, 1059–1073.
- Maier, S. F., Watkins, L. R. and Fleshner, M. (1994). Psychoneuroimmunology: the interface between behavior, brain, and immunity. *American Psychologist* **49**, 1004–1017.
- McCabe, P. M. and Schneiderman, N. (1985). Psychophysiologic reactions to stress. In: Schneiderman, N. & Tapp, J. T. (eds.) *Behavioral medicine: the biopsychosocial approach*, pp. 99–131, Hillsdale, NJ: Lawrence.
- McEwen, B. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* **338**, 171–179.
- Pantaleo, G., Gruiosi, C. and Fauci, A. S. (1993). The immunopathogenesis of human immunodeficiency virus infection. *New England Journal of Medicine* **328**, 327–335.
- Roitt, I., Brostoff, J. and Male, D. (1993). *Immunology* (3rd edn.). London: Mosby.

Air Travel See: Anxiety.

Airline Accidents

G Li

Johns Hopkins University, Baltimore, MD, USA

© 2007 Elsevier Inc. All rights reserved.

Accident Rates
Stress Model
Pilot Error
Risk Factors
Policy Interventions

Glossary

Accident An occurrence related to the operation of an aircraft that causes fatal or serious injury to any person or substantial damage to the aircraft.

Air carriers Commercial air transportation companies that transport people and/or goods;

include both major airlines and commuter airlines.

Incident An adverse event with less serious consequences than an accident.

Human factors A scientific discipline aimed at optimizing the human-machine interface.

Pilot error An act by the pilot that is deemed inappropriate and detrimental to flight safety.

Aviation is a complex system. It includes aircraft, airport facilities, regulations, and operation procedures as well as personnel, such as flight crew, air traffic controllers, aircraft dispatchers, and maintenance workers. Civil aviation can be divided into two groups: commercial flights that transport people and goods to generate revenues and noncommercial flights that are operated by private pilots for pleasure or personal business. Commercial flights are operated by air carriers. Noncommercial flights are usually

called general aviation. Air carriers may include major airlines, commuter airlines, and air taxis. These subgroups of commercial aviation are classified according to aircraft capacities and other variables, and they operate under different regulations. Unless specified otherwise, the term airline in this article refers to major airlines, that is, aircraft operations with 10 or more passenger seats or 7500 payload-pounds of capacity. There are approximately 8200 airline aircraft in the United States, transporting approximately 600 million passengers and 14 million tons of goods each year. An aviation accident is defined as an occurrence associated with the operation of an aircraft that results in fatal or serious injury to any person or aircraft damage requiring major repair or replacement. A serious injury is operationally defined as any injury that requires hospitalization for more than 48 h, that involves a fracture, or that causes severe hemorrhage. Aviation incidents refer to adverse events causing only minor personal injury or minor aircraft damage.

Accident Rates

Aviation is the safest mode of transportation. The rate of deaths per 100 million passenger miles is approximately 0.03 for airlines, compared with 0.9 for motor vehicles and 0.06 for trains. In 2004, air carriers in the United States made a total of 10.8 million

departures and flew a total of 1.8 million hours, with 28 accidents and 14 fatalities. The rate of fatal accidents was 0.19 per 100 000 aircraft departures and 0.011 per 100 000 flight hours. Accident rates for airlines have decreased by 94% in the past four decades and have remained fairly stable since 1980 (Figure 1). General aviation and air taxis have much higher accident rates than airlines; approximately 94% of aviation accidents recorded in the United States involve non-commercial flights. In 2004, the rate of fatal accidents per 100 000 flight hours was 1.20 for general aviation and 0.78 for air taxis, which are 108 times and 71 times, respectively, the rate for airlines.

Stress Model

Aviation safety has been studied from various perspectives. Human factors, which aim to optimize the relationship between people and their activities through engineering, technology, training, and resource management, has played a critical role in aviation safety. Human factor researchers consider aviation to be a system consisting of four elements: hardware, software, liveware, and environment. To maintain aviation safety is to keep the aviation system intact and functioning. The aviation system includes many subsystems. Human factors researchers are especially interested in liveware and its interface with other components of the aviation system. The

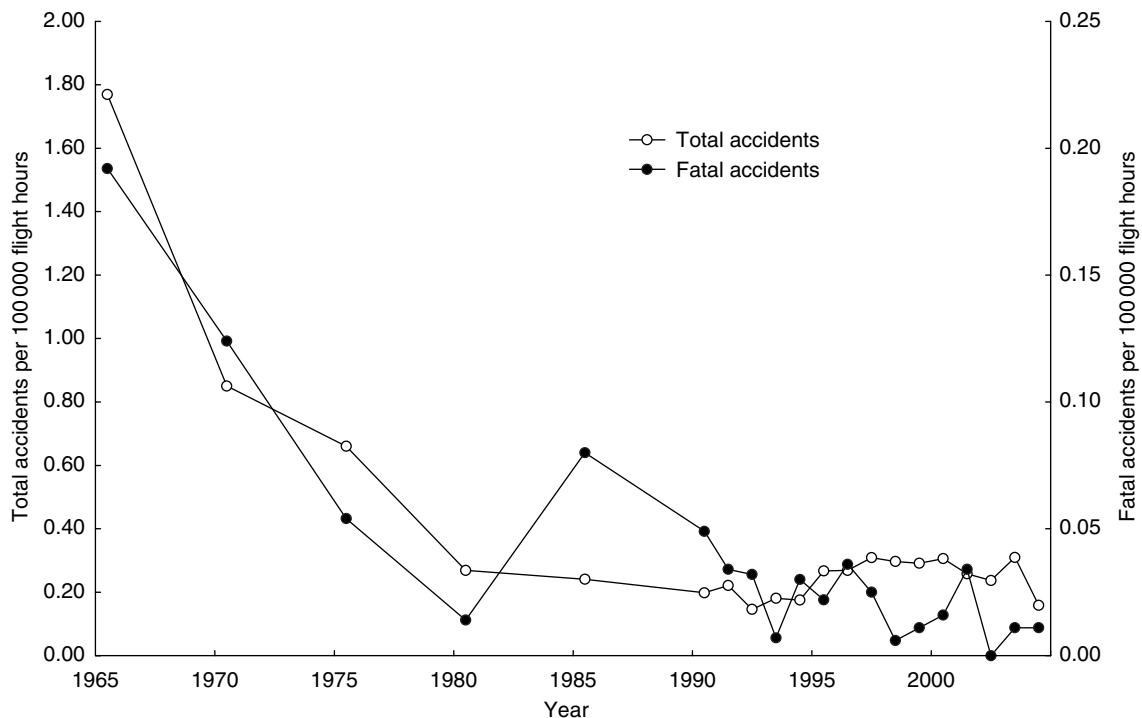


Figure 1 Rates of total accidents and fatal accidents per 100 000 flight hours for air carriers, United States, 1965–2004. Data from U.S. Department of Transportation, Bureau of Transportation Statistics.

information processing model that elucidates the psychological mechanisms of decision making has been used by human factors researchers to understand and enhance pilot performance. For instance, studies of pilot error in operating instruments and interpreting instrument readings led to the improved design and display of flight instruments.

Selye considered stress to be a normal phenomenon and the reaction to stressors to be a process of adaptation. Based on Selye's seminal work, Haakonson proposed a simple stress model to explain aviation accident causation, in which safety is the state in which performance ability matches performance demand and accidents result from the imbalance between performance ability and performance demand, particularly when performance demand exceeds performance ability. It is assumed that every pilot has a certain level of performance ability that is susceptible to gradual decline during the course of a flight and that every flight operation involves a certain level of performance demand. Pilots face four types of stressors during flight: environmental (e.g., noise, vibration, and turbulence), physiological (e.g., sleep deprivation and fatigue), emergency (e.g., fuel shortage and engine malfunction), and life events (e.g., financial strain and marital problems). Some of these stressors can be detrimental to performance ability and others can increase the performance demand. The stress model posits that pilot-error accidents are in essence the consequences of inadequate coping or maladaptation to stressors.

The stress model culminated a great deal of research in the field of aviation safety during the 1980s, especially in military aviation. Questionnaires measuring pilot personality and life experience were used in various pilot population groups in order to determine the role of personal life stress in aviation accidents. Studies conducted in pilots who were involved in aviation accidents found that the prevalence of adverse life events, such as financial and interpersonal difficulties, was significantly higher in those culpable than those not culpable for their accidents. The implications of these findings, however, are ambiguous. The retrospectively reported data and quasi-experimental design make it difficult to infer causality between personal life stress and aviation accident. The general consensus is that even if there is a causal relationship between personal life stress and aviation accident, pilot personality and recent life experience contribute only to a small fraction of all aviation accidents. Furthermore, these studies are of limited practical significance because personal attributes such as personality and life events are hardly modifiable. Consequently, researchers have shifted their focus from personal life events to the environmental and occupational stressors pilots

must confront while flying, such as adverse weather, mechanical malfunction, medical emergency, fatigue, aging, and alcohol abuse.

Pilot Error

Pilot error is regarded as the last frontier in aviation safety. Overall, approximately 80% of aviation accidents and 50% of aviation incidents are attributed to pilot error. The role of pilot error in aviation accidents, however, varies greatly with the type of flight operation. Pilot error is a probable cause in approximately 35% of airline accidents, 74% of commuter/air taxi accidents, and 85% of general aviation accidents. The most common pilot error identified in airline accidents is inattentiveness (e.g., failing to see and avoid, and improper preflight planning), which constitutes 26% of all pilot errors. Other major categories of pilot error are flawed decision (e.g., misjudging visibility and weather conditions, and flying aircraft with known defect) and mishandled aircraft kinetics (e.g., failure to recover from a stall or improper response to bounce), each accounting for 22% of all pilot errors. The remaining pilot errors in airline accidents are poor crew interaction (11%), mishandled wind/runway conditions (7%), improper instrument flight rules approach (3%), and other (9%). Although the pattern of pilot error is similar between airline accidents and commuter/air taxi accidents, there is a notable difference between accidents of commercial aviation and general aviation (Table 1). Specifically, mishandled aircraft kinetics is more common in general aviation accidents than in airline and commuter/air taxi accidents. The overrepresentation

Table 1 Type of pilot error in aviation accidents by aviation category (%)

Type of pilot error	Airlines	Commuters/ air taxis	General aviation
Inattentiveness	26	23	25
Flawed decision	22	20	22
Mishandled aircraft kinetics	22	18	38
Mishandled wind/runway conditions	7	18	11
Poor crew interaction	11	10	0
Other	12	11	4
Total	100	100	100

Data from Baker, S. P., Lamb, M. W., Grabowski, J. G., et al. (2001), Characteristics of general aviation crashes involving mature male and female pilots, *Aviation, Space and Environmental Medicine* **72**, 447–452; Li, G., Baker, S. P., Lamb, M. W., et al. (2002), Human factors in aviation crashes involving older pilots, *Aviation, Space and Environmental Medicine* **73**, 134–138; Li, G., Grabowski, J. G., Baker, S. P., et al. (2006), Pilot error in air carrier accidents: does age matter? *Aviation, Space and Environmental Medicine* **77**, 737–741.

of mishandled aircraft kinetics in general aviation accidents is probably due to the fact that general aviation pilots on average are far less trained than commercial pilots.

It is noteworthy that the prevalence of pilot error has decreased markedly in airline accidents, from 46% in the 1980s to 26% in the late 1990s and early 2000s. The use of advanced technologies (e.g., ground proximity warning systems) and implementation of intervention programs (e.g., cockpit resource management) might have contributed to the decline of pilot error in airline accidents. A similar downward trend has not been observed in commuter/air taxi and general aviation accidents.

Risk Factors

Risk factors are personal or environmental characteristics that are associated with an increased probability of accident occurrence. Over the years, researchers have identified several risk factors for aviation accidents. The pilot's health status is one of the characteristics most scrutinized. Early in the 1930s, studies of student pilots indicated that flight performance, measured by certification tests and accident involvement, was inversely associated with the severity of physical defects. Subsequent research linked selected visual problems and cardiovascular disease to increased accident propensity.

Alcohol use is another risk factor that has been studied extensively. Studies in the early 1960s revealed that over one-third of the pilots who were fatally injured in general aviation accidents had positive blood alcohol concentrations (BACs) at the time of death; and among those tested positive, 57% had $\text{BACs} \geq 100 \text{ mg dl}^{-1}$. Following intensive educational intervention and restrictive regulations, alcohol involvement in fatal general aviation accidents decreased to approximately 8%. Since 1985, the United States Federal Aviation Administration has adopted a zero-tolerance policy regarding alcohol-impaired flying. Pilots are prohibited from flying for 8 h after consuming any alcoholic beverage or with a BAC of 40 mg dl^{-1} or greater. Mandatory alcohol testing programs for airline pilots started in 1995. Each year, at least 10% of airline pilots in the United States are randomly selected for alcohol testing. Those who are involved in an accident are required to submit to alcohol testing within 2 h of the accident. Data from the random testing program suggest that the prevalence of alcohol-impaired flying among airline pilots is approximately 0.06%. The odds of accident involvement for airline pilots under the influence of alcohol (i.e., with a $\text{BAC} \geq 40 \text{ mg dl}^{-1}$) is estimated to be 2.6 times the odds for their sober counterparts. In addition to acute alcohol use, chronic drinking

problems, such as a driving-while-intoxicated history, are also found to be associated with a significantly increased risk of aviation accident in both airline pilots and general aviation pilots.

The relationship between pilot age and flight safety has been the subject of intense research and continuing controversy, mainly because of the federal regulation that requires mandatory retirement by their sixtieth birthday for all airline pilots. This policy, commonly known as the Age-60 Rule, was promulgated in 1960 out of concerns about the potential detrimental effects on safety performance of age-related declines in health status and cognitive functions. Experimental studies using flight simulators suggest that age-related declines are mostly limited to domain-independent cognitive functions; among the few domain-dependent measures that are found to decline with advancing age are time-sharing efficiency and the ability to respond to fast-spoken air traffic control commands. These age-related deficits are due to a decreased working memory capacity in older pilots and can be compensated for to a certain degree by flight experience. The fundamental question concerning the controversy of the Age-60 Rule is whether pilots' risk of accident involvement increases as age rises. The results from epidemiological studies aimed at answering this question are contradictory and inconclusive. To determine the effect of aging on accident risk, researchers must confront both conceptual and practical challenges. Theoretically, aging has two divergent effects on safety performance: the risk due to declines in health and cognitive functions and the benefit from increased flight experience and enhanced safety behavior. Because the adverse and beneficial effects take place simultaneously during the process of aging, it is hard to separate one from the other. Research efforts are further compounded by several practical issues, including the lack of a widely accepted measurement of accident risk, the rarity of airline accidents, the healthy-worker effect induced by the rigorous medical certification process, and the truncated work-life span of airline pilots caused by the Age-60 Rule.

Past safety performance is predictive of future performance. Epidemiological studies indicate that pilots who were involved in an aviation accident or incident in the past 3 years were 70% more likely than those without such a history to be involved in an aviation accident within 3 years. Deviant behavior, such as violating aviation regulations, has a greater effect on accident risk than past experience of accident involvement. The risk of accident involvement for pilots with a prior violation record is more than twice that for their counterparts without a violation history.

Fatigue remains to be a safety concern for airline pilots because they often work extended schedules,

night shifts, and across multiple time zones. It is evident that accident risk increases with duty hours. Airline accidents appear to occur disproportionately when pilots are on flight duty for 10 hours or longer. Current regulations in the United States limit flight time for each airline pilot to 30 h week⁻¹ and 1000 h year⁻¹.

Environmental factors, which play a dominant role in airline accidents, can be easily identified. Of all airline accidents, 25% are directly caused by turbulence and 21% by mechanical failure. Bad weather is also strongly associated with pilot error. In fact, instrument meteorological conditions are recognized as the most important risk factor for pilot error.

Policy Interventions

Airline pilots confront a variety of stressors in their work environment. Although aviation safety has improved remarkably in the past 4 decades, airline pilots have much higher job-related mortality than most other occupational groups. The stress model provides a comprehensive conceptual framework for understanding the causation and prevention of aviation accidents. Pilot error, which is a contributing factor in 35% of airline accidents and 85% of general aviation accidents, can be regarded as the result of inadequate coping with endogenous and exogenous stressors. Factors, such as alcohol use, prior violation history, fatigue, and bad weather, influence accident risk either by impairing pilot's performance ability or by increasing performance demand. Policy interventions, including medical standards, zero-tolerance toward alcohol and drug use, flight proficiency requirements, and work-hour rules, have definitely helped to reduce pilot error and aviation accidents. Programs aimed at enhancing pilot performance ability, however, have their limits. Airline accidents are not merely a function of pilot error or individual risk factors. Rather, they are the results of complex interactions of personal factors, aircraft, and the environment in the context of the aviation system. Advances in technology, which have substantially increased the reliability of the aviation system and decreased pilot performance demand, are largely responsible for the remarkable achievement in aviation safety and will further reduce pilot error and airline accidents.

Acknowledgments

The author's work on aviation safety was funded by Grants R01AG13642 and R01AA09963 from the National Institutes of Health and Grant CCR302486 from the Centers for Disease Control and Prevention.

Preparation of this manuscript was supported in part by a fellowship grant from the John Simon Guggenheim Memorial Foundation. The author thanks Dr. Yandong Qiang for technical assistance and Profs. Susan P. Baker and George W. Rebok for helpful comment.

See Also the Following Articles

Aging and Psychological Stress; Aging and Stress, Biology of; Alcohol and Stress; Social and Psychological Aspects; Cognition and Stress; Fatigue and Stress; Sleep Loss, Jet Lag, and Shift Work; Workplace Stress; Life Events and Health.

Further Reading

- Alkov, R. A., Gaynor, J. A. and Borowsky, M. S. (1985). Pilot error as a symptom of inadequate stress coping. *Aviation, Space and Environmental Medicine* 56, 244–247.
- Baker, S. P., Lamb, M. W., Grabowski, J. G., et al. (2001). Characteristics of general aviation crashes involving mature male and female pilots. *Aviation, Space and Environmental Medicine* 72, 447–452.
- Goode, J. H. (2003). Are pilots at risk of accidents due to fatigue? *Journal of Safety Research* 34, 309–313.
- Haakonson, N. H. (1980). Investigation of life change as a contributing factor in aircraft accidents: a prospectus. *Aviation, Space and Environmental Medicine* 51, 981–988.
- Li, G. (1994). Pilot-related factors in aircraft crashes: a review of epidemiologic studies. *Aviation, Space and Environmental Medicine* 65, 944–952.
- Li, G., Baker, S. P., Grabowski, J. G., et al. (2001). Factors associated with pilot error in aviation crashes. *Aviation, Space and Environmental Medicine* 72, 52–58.
- Li, G., Baker, S. P., Grabowski, J. G., et al. (2003). Age, flight experience, and risk of crash involvement in a cohort of professional pilots. *American Journal of Epidemiology* 157, 874–880.
- Li, G., Baker, S. P., Lamb, M. W., et al. (2002). Human factors in aviation crashes involving older pilots. *Aviation, Space and Environmental Medicine* 73, 134–138.
- Li, G., Grabowski, J. G., Baker, S. P., et al. (2006). Pilot error in air carrier accidents: does age matter? *Aviation, Space and Environmental Medicine* 77, 737–741.
- Morrow, D. G., Menard, W. E., Stine-Morrow, E. A., et al. (2001). The influence of expertise and task factors on age differences in pilot communication. *Psychology and Aging* 16, 31–46.
- Selye, H. (1975). *Stress without distress*. New York: Lippincott.
- Tsang, P. S. (2002). Assessing cognitive aging in piloting. In: Tsang, P. S. & Vidulich, M. A. (eds.) *Principles and practices of psychology*, pp. 507–546. Hillsdale, NJ: Lawrence Erlbaum.
- Wiener, E. L. and Nagel, D. C. (1988). *Human factors in aviation*. San Diego, CA: Academic Press.

Airplane Crashes See: Airline Accidents; Lockerbie Air Crash, Stress Effects of.

Alarm Phase and General Adaptation Syndrome

R McCarty

Vanderbilt University, Nashville, TN, USA

K Pacak

National Institutes of Health, Atlantic GA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 126–130, © 2000, Elsevier Inc.

Hans Selye and the General Adaptation Syndrome
Shortcomings of the General Adaptation Syndrome
Prior Stress History Affects Future Stress Responses
The Death Knell of Nonspecificity
Summary and Conclusions

Glossary

<i>Dishabituation</i>	The facilitation of a habituated stress response following exposure to an intense or a novel stressor.
<i>General adaptation syndrome (GAS)</i>	A syndrome proposed by Selye that consists of the physiological changes experienced by animals, including humans, during prolonged stress; it consists of three distinct phases: the alarm phase, the resistance phase, and the exhaustion phase.
<i>Habituation</i>	The decrease in the amplitude of the stress response after repeated daily exposure to the same stressor.
<i>Sensitization</i>	The enhancement of a nonhabituated stress response by a stressor of high intensity.

Hans Selye and the General Adaptation Syndrome

Hans Selye deserves much of the credit for popularizing the concept of stress in the scientific and medical literature of the twentieth century. Much of the credit also goes to him for the confusion that has resulted in issues relating to the definition of stress and the conceptualization of how stress plays a role in disease

etiology. From the beginning of his career, Selye emphasized a medical model in advancing his views on stress. In his landmark paper published in *Nature* in 1936, Selye first advanced his notion of stress as a “non-specific response of the body to any demand placed upon it.” He was unwavering in promoting this hallmark feature of his definition of stress until his death in the early 1980s.

Selye also advanced the notion of the stress response unfolding in specific stages, which make up the GAS ([Table 1](#)). Upon the exposure of an animal or human to a given stressor, there is a rapid onset alarm phase. The alarm phase is thought to consist of two distinct phases: shock and countershock. The disruptive effects of the stressor are expressed in a dramatic alteration in homeostatic processes, including regulatory processes affecting blood pressure, circulating levels of glucose, electrolyte balance, distribution of blood flow, and membrane permeability. These shock-related responses to the stressor are counteracted in part by the countershock responses of the adrenal cortex, through the release of corticosteroids, and the adrenal medulla, through the release of epinephrine.

If the stressful stimulus persists over a prolonged period of time, the second stage of the GAS, the resistance phase, develops. During the resistance phase, the organism achieves a state of increased adaptation to the untoward effects of the stressor but is more susceptible to the deleterious effects of other homeostatic

Table 1 Stages of the general adaptation syndrome

Stage	Description
Stage 1: Alarm phase	Shock Countershock
Stage 2: Resistance phase	Syntoxic hormones (anti-inflammatory) Catatoxic hormones (pro-inflammatory)
Stage 3: Exhaustion phase	Depletion of adaptation energy Increase in ulceration of the gastrointestinal tract Increased susceptibility to infectious agents Typically ends in death

challenges. In this phase of the GAS, Selye argued, the adrenal cortex plays a key role through a balanced release of syntoxic and catatoxic steroid hormones. Syntoxic glucocorticoid hormones inhibit inflammatory responses, whereas catatoxic hormones promote inflammatory responses to pathogens. Selye was never able to isolate and identify catatoxic adrenal hormones.

At some undefined point, if the stressful stimulus continues and perhaps increases in intensity, the organism enters the third and final stage of the GAS, the phase of exhaustion. Selye suggested that the onset of this third stage is triggered by depletion of adaptation energy stores. This depletion of adaptation energy is attended by enhanced activity of the hypothalamic-pituitary-adrenocortical axis and the onset of pathophysiological changes in the immune system and gastrointestinal tract, resulting ultimately in the death of the organism. Two critical features of stress for many years, gastrointestinal ulcers and increased susceptibility to infectious agents, are reflected in the changes that are thought to accompany the phase of exhaustion. As for adaptation energy, Selye was never able to measure it; he could only infer its depletion when the organism died from lack of it.

Shortcomings of the General Adaptation Syndrome

Not surprisingly, several aspects of Selye's conception of stress and the specific features of the GAS have been the subject of criticism over the years. It is safe to say that Selye's attempts to popularize the concept of stress as it relates to physical and mental health were extremely successful. In contrast, his conceptualization of stress and the details of the GAS have not stood the test of time. We address a few of these issues below.

Selye's definition of stress has been a major source of concern for investigators in this field for many years. His emphasis on the nonspecificity of the stress response has been especially troubling and some have argued that his writings on this subject were contradictory. For example, although Selye emphasized the nonspecificity of the stress response across many different types and intensities of stressors, he did allow room for individual differences in the stress response. Indeed, he was forced to do this to explain the fact that there are many different pathophysiological changes associated with stress. If the stress response were indeed invariant between individuals, there should be only one type of stress-related disorder. However, he addressed this problem by positing that conditioning factors acted to accentuate or inhibit

particular components of the stress response, leading to individual differences in the expression of the deleterious effects of stress. These conditioning factors could be related to genetic, maturational, or environmental influences. Still, he argued that once the effects of these conditioning factors were stripped away, a constellation of nonspecific responses remained. Few if any investigators support this suggestion today.

Selye's work also placed the adrenal cortex at the center of the stress universe. Indeed, an elevation in secretion of steroid hormones from the adrenal cortex became for some a necessary condition for stress to occur. Beginning with the work of John Mason and his colleagues in the 1960s, the centrality of adrenal steroids in the stress response was supplanted by an emphasis on the participation of multiple neural and neuroendocrine pathways in the stress response, with no single system assuming a position of preeminence.

Prior Stress History Affects Future Stress Responses

Work from our laboratory has focused on one area of Selye's GAS that has received little attention. Specifically, we looked at the effects of prior stress history on the pattern of responses to subsequent exposure to a stressor. In this work, we focused our attention on sympathetic-adrenal medullary responses to stress and we measured plasma levels of norepinephrine (NE) and epinephrine (EPI) in laboratory rats. Plasma levels of NE and EPI have been shown to provide an accurate assessment of the activity of the sympathetic nerves and the adrenal medulla, respectively. For all these studies, rats were surgically prepared with indwelling tail artery catheters that permitted remote sampling of blood.

Three types of nonassociative learning have been studied within a framework of quantifying hormonal responses to stressful stimulation: (1) habituation, the decrease in the amplitude of the stress response after repeated exposure to the same stressor; (2) sensitization, the enhancement of a nonhabituated stress response by an intense stressor; and (3) dishabituation, the facilitation of a habituated stress response following exposure to an intense or novel stressor.

Habituation

Chronic intermittent exposure of laboratory rats to a stressor of low or moderate intensity leads to a significant reduction in plasma levels of NE and EPI compared to control animals exposed to the same stressor for the first time. Several different stressors have been

evaluated, including forced swimming, restraint, immobilization, footshock, and exercise. It is important to note that these chronic intermittent stress paradigms provide animals with a high degree of predictability regarding such parameters as type of stressor, stressor intensity, duration of each stress session, and the time of onset of the stressor each day. When provided with this information after several stress sessions, animals are able to activate stress-responsive hormonal systems to the minimum extent necessary to maintain cardiovascular and metabolic homeostasis during each subsequent exposure to the same stressor. This progressive dampening of plasma catecholamine responses to a familiar and highly predictable stressor provides for a significant conservation of energy expenditure.

Sensitization

When animals are exposed to a high-intensity stressor, there are recurring disruptions in cardiovascular and metabolic processes. One such stressor that has been studied is immersion of laboratory rats in water maintained at 18 or 24°C for 15 min per day for 27 consecutive days. When compared to controls exposed to cold swim stress for the first time, animals exposed to chronic intermittent swim stress had significantly greater elevations in plasma levels of NE and EPI. This sensitized response of the sympathetic-adrenal medullary system to chronic intermittent cold swim stress is dependent on stressor intensity. Chronic intermittent swim stress in water maintained at higher temperatures (30°C) results in habituated responses.

Dishabituation

Dishabituation involves an enhancement of the physiological response to a novel stressor in animals exposed repeatedly or continuously to an unrelated stressor. For example, if laboratory rats are exposed to a brief period of footshock stress each day for several weeks, the plasma catecholamine response to footshock gradually decreases over time compared to the response of a naïve control. However, if these same rats are exposed to a novel stressor (e.g., restraint stress), their plasma catecholamine response is amplified compared to the response of naïve controls to restraint stress. Data from our laboratory and the work of others suggest that exposing chronically stressed animals to an unfamiliar stressor presents a much greater challenge to behavioral and physiological homeostasis than the presentation of the same stressor to a naïve control. In this instance, the departure from the expected appears to be the critical

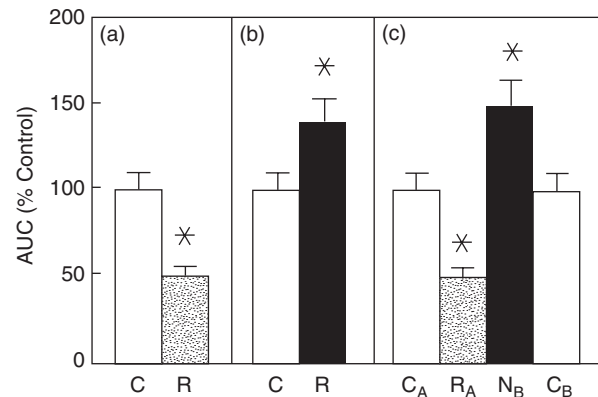


Figure 1 Nonassociative properties of plasma catecholamine responses (expressed as a percentage of control) in rats exposed to chronic intermittent stress. a, Habituation of plasma catecholamine responses of repeatedly stressed rats to a mild stressor compared to first-time stressed controls; b, Sensitization of plasma catecholamine responses of repeatedly stressed rats to an intense stressor compared to first-time stressed controls; c, Dishabituation of plasma catecholamine responses of rats repeatedly exposed to stressor a when presented with a novel stressor b. The plasma catecholamine responses of the a and b controls are presented for comparison. Note that in the same group of animals, the response to a familiar stressor (R_A) may be diminished, whereas the response to a novel stressor (N_B) may be enhanced. AUC, area under the curve; c, first-time stressed controls; C_A , stressor a controls; C_B , stressor b controls; N_B , rats presented with a novel stressor b; R, rats repeatedly stressed; R_A , rats repeatedly exposed to stressor a.

variable in eliciting an enhanced physiological response to stress. Chronically stressed animals live in a moderately stressful but highly predictable environment. The experimenter arrives at approximately the same time each day, the stress session begins at approximately the same time each day, the intensity of the stressor and the duration of the stress session are held constant, and most of the day is free from disturbance. Against this backdrop of consistency, a novel stressor represents an abrupt and at times a dramatic departure from the consistency of the past (Figure 1).

The Death Knell of Nonspecificity

Although the issue of the nonspecificity of stress responses had been discussed by a host of investigators, no study had attempted to compare neuroendocrine response profiles across a range of different stressors. Recently, such a study was conducted and the measures of the stress response included plasma levels of adrenocorticotrophic hormone (ACTH), corticosterone, NE, and EPI, as well as microdialysate levels of NE within the hypothalamic paraventricular nucleus (PVN). The PVN contains cell bodies

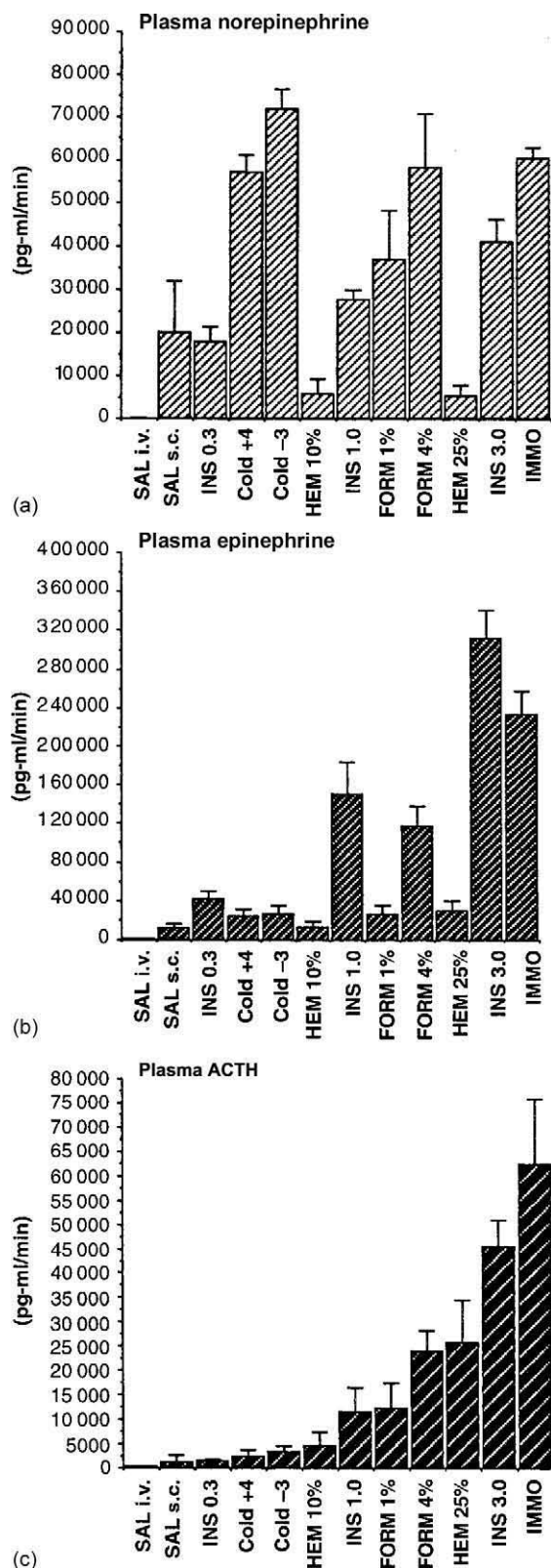


Figure 2 Plasma hormone levels during exposure to various stressors in conscious rats. a, epinephrine; b, norepinephrine; c, ACTH. Each bar represents the mean value for the net areas under the curve (AUCs) for that stressor, where the net AUC for each animal was calculated from the baseline AUC subtracted

of neurons that release corticotropin-releasing hormone (CRH) into the pituitary portal system and NE-containing nerve terminals within this nucleus provide a potent stimulatory signal for release of CRH. The acute stressors that were employed in this study included the following: hemorrhage (10 or 25% of estimated blood volume), intravenous insulin (0.1, 1.0, or 3.0 IU/kg), subcutaneous administration of formaldehyde (1 or 4% solution), exposure to a cold environment for 180 min (4°C or -3°C), and immobilization for 120 min.

The results of this study indicated that the neuroendocrine signatures varied dramatically across stressors (Figure 2). In addition, for those stressors that varied in intensity (hemorrhage, insulin, formaldehyde, and cold exposure), some increments in a given neuroendocrine response were greater from one intensity to the next compared to another neuroendocrine response. For example, the ACTH response to a 25% hemorrhage exceeded by five times the plasma ACTH response to a 10% hemorrhage. In contrast, the plasma EPI response to a 25% hemorrhage exceeded by only twice the plasma EPI response to a 10% hemorrhage. Similarly, the plasma ACTH response to a 4% formaldehyde solution was twice that of the response to a 1% solution. However, the plasma EPI response to a 4% formaldehyde solution was four times greater than the response to a 1% solution.

These two categories of findings (e.g., unique neuroendocrine signatures combined with the absence of proportionate increases across all neuroendocrine responses with increases in stressor intensity) are in striking contrast to the predictions associated with Selye's doctrine of nonspecificity. The results of this study are consistent with a primitive form of specificity whereby each stressor provokes a neuroendocrine response pattern that allows, within some limits, for a maximal homeostatic balance to be maintained.

Summary and Conclusions

It has been more than 65 years since Selye published his often-cited paper in the journal *Nature*. The interdisciplinary field of stress research has matured to a significant degree since that time, and thousands of papers on this topic have appeared in the scientific literature. Many of Selye's original concepts of stress have not stood the test of time; improved analytical

from the total AUC. Stressors shown are saline injection, intravenously (SAL IV) or subcutaneously (SAL SC); insulin (INS) 0.3, 1.0, or 3.0 IU/kg; cold exposure at 4°C or -3°C; hemorrhage (HEM) of 10% or 25% of total estimated blood volume; formalin injection (FORM), 1% or 4% solution; and immobilization (IMMO).

techniques and new discoveries in neuroanatomy, neuroendocrinology, and molecular biology have occurred at a remarkable pace. However, it is important to recognize that Selye did establish a connection between stress and disease that continues to occupy basic and clinical researchers in their laboratories and clinics throughout the world.

See Also the Following Articles

Adrenal Cortex; Adrenal Medulla; Disease, Stress Induced; Selye, Hans.

Further Reading

- Chrousos, G. P. and Gold, P. W. (1992). The concept of stress and stress system disorders. *Journal of the American Medical Association* 267, 1244–1252.
- Goldstein, D. S. (1995). *Stress, catecholamines, and cardiovascular disease*. New York: Oxford University Press.
- McCarty, R. and Gold, P. E. (1996). Catecholamines, stress and disease: a psychobiological perspective. *Psychosomatic Medicine* 58, 590–597.
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* 338, 171–179.

Alcohol and Stress: Social and Psychological Aspects

M A Sayette

University of Pittsburgh, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M A Sayette, volume 1, pp 136–140, © 2000, Elsevier Inc.

Introduction

Effects of Alcohol on Stress

Effects of Stress on Alcohol Consumption

Glossary

<i>Anxiolytic</i>	Anxiety or stress reducing.
<i>Stress</i>	The perception of an event as aversive, in that it concerns harm, loss, or threat.
<i>Stressor</i>	An event that elicits stress.

Introduction

Nearly everyone, from clinicians and researchers, to social and problem drinkers, believes that drinking alcohol can reduce stress. Moreover, this effect of alcohol has been suggested as a reason why people begin and maintain their drinking, despite its abuse potential. These two propositions – that alcohol reduces stress, and that people drink in order to obtain this stress reducing effect – were initially formalized by Conger in the mid-1950s in what has become known as the tension reduction hypothesis. Although the term tension has been replaced over the years by terms such as stress or anxiety, Conger's two basic propositions continue to generate a tremendous amount of investigation. Research includes both

field studies and laboratory experimentation and has involved both human and animal subjects.

Effects of Alcohol on Stress

Historically, the term stress has been variously used to describe the stimuli or events that produce a state of anxiety or apprehension, as well as to describe the complex physiological response to an aversive stimulus. Because people respond to the same stimuli in different ways, however, it has been suggested that stress be defined as the appraisal, or interpretation, of an event as signaling harm, loss, or threat. This approach, adopted here, recognizes that an event may be construed as stressful by one individual, but benign by another. The perception of stress is posited to activate a multidimensional response that may involve a wide range of biological responses (e.g., psychophysiological and neurochemical), behaviors (e.g., escape or avoidance behavior), and in the case of humans, subjective awareness of distress. Not surprisingly, studies of the relationship between alcohol and stress have varied dramatically in terms of both the stressors that are used and stress responses that are measured. Such variability likely has contributed to the complex pattern of findings that characterize this research area.

Animal Studies

Conger's original formulation was based on studies using animals, including rats and cats. It is, therefore, not surprising that extensive animal research (most often using rodents) has been conducted testing the alcohol–stress relationship. Typical stressors include a novel environment, food and water deprivation, experimental conflict, in which conflicting motives

and responses are activated, and noxious stimuli (e.g., electric shock administration, exposure to (aversive) bright environments, and placing an intruder into a resident's cage). These studies tend to show alcohol producing an anxiolytic effect, though the relationships are complex.

Data supporting an anxiolytic effect of alcohol appear strongest in studies using experimental-conflict paradigms. These studies begin by training an animal to run down an alley to obtain food reward. Next a shock is administered to punish this approach behavior. The level of conflict is inferred by the suppression of the approach to the food. Typically alcohol ingestion, through its fear-reduction properties, facilitates the approach behavior.

Whether alcohol reduces stress in animals may depend in part on how stress responding is assessed and the dose of alcohol administered. Pohorecky, who has long chronicled progress in this area, concludes that moderate doses of alcohol have modified both anxiety-related behaviors (such as aggression and conflict), and stress-induced neurochemical changes. These latter responses are present in both the brain and in plasma. The former include levels or turnover of norepinephrine (norepinephrine), serotonin, β -endorphins and γ -aminobutyric acid (GABA), while the latter includes catecholamines, corticosterone, testosterone, nonesterified fatty acids, and amino acids. Moderate doses of alcohol consumption also appear to alter stress-induced changes in the functional activity of the hypothalamo-pituitary-adrenocortical axis. These findings suggest neurobiological mechanisms underlying potential anxiolytic effects of moderate (but not high) doses of alcohol.

Human Studies

The effects of alcohol consumption on human stress have proven to be far more complex than originally thought. Although some would argue that the effects are sufficiently unreliable that the notion of alcohol reducing stress should be rejected altogether, most still would conclude that alcohol can reduce stress. The challenge has been to understand the conditions under which the anxiolytic effects of alcohol should be most apparent, and the possible mechanisms underlying this effect.

A number of factors have been examined that may affect the effect of alcohol on stress. These include individual difference factors and situational factors.

Individual difference factors Individual difference factors include gender, personality traits, drinking history, and a family history of alcoholism. Until recently, the vast majority of studies in this area included only male participants. Despite early reports to the contrary, the consensus drawn from the largest

studies that have included both genders is that effects of alcohol on stress seem to be comparable for men and women. These studies have assessed stress using psychophysiological (e.g., heart rate) and self-report measures, and it is possible that other types of measures of emotional response might reveal gender differences. Even when women appeared more responsive than men to a stressor in sober conditions, this gender difference was not altered by alcohol. There is some evidence, however, to suggest that men and women differ in their beliefs about how alcohol consumption affects them. It is, therefore, possible that at low doses of alcohol, men and women may experience different effects, due to nonpharmacological effects of drinking experiences.

Individuals characterized by personality traits such as sensation seeking or impulsivity have been found to be at increased risk for developing alcohol-related problems. Some have suggested that these individuals may derive enhanced anxiolytic effects from alcohol consumption. Presumably, this would increase the reinforcement value of drinking. Although the studies have not been uniform in their conclusions, there is some support for this hypothesis.

Another individual difference factor that has been studied is drinking history. Specifically, some investigations have found that heavier drinkers, due to their development of tolerance, require greater quantities of alcohol than lighter drinkers to obtain the anxiolytic effects of alcohol. These findings converge with the aforementioned animal data to emphasize the importance of dose in determining the effects of alcohol-stress relationship. It may be that the anxiolytic effects of alcohol are restricted to a relatively narrow range of doses.

In contrast to studies that have subjected participants to a single dose of alcohol during one session, there have been studies that permit evaluation of drinkers over extended time periods. During the mid-1960s through the mid-1970s, a number of experiments examined the effects of alcohol use on stress among alcoholic participants. Although these studies often were limited by small sample sizes, they were impressive for their intensive monitoring of participants for weeks at a time. The methodologically strongest studies of this group indicate an association between alcohol consumption and improved emotional states among these alcoholic participants.

The presence of a family history of alcoholism represents an established risk factor for alcoholism. Several experiments have examined young adults with such a family history. Although at this point in their lives these individuals are not problem drinkers, some of these studies have found that offspring of alcoholic fathers nevertheless experience an increased anxiolytic effect of alcohol. These offspring may be

more responsive to a wide range of stimuli while sober, and, therefore, may be especially sensitive to the sedating effects of alcohol. There is some evidence, however, that some of these enhanced anxiolytic findings actually may be due to family history differences on variables other than stress.

Situational factors Situational factors have been proposed that may influence the effects of alcohol on stress. Alcohol is more likely to reduce stress if it is consumed in the presence of pleasantly distracting activity (e.g., drinking at a party or while watching television) than if consumed without distraction. A second factor that has been suggested to moderate the effects of alcohol on stress is the temporal relationship between drinking and the stressor administration. When alcohol is consumed following a stressor (e.g., after a bad day at work) it may be less effective in reducing stress than if intoxication precedes the stress (e.g., drinking just prior to attending a party). While this latter prediction has been supported by laboratory studies, it remains to be seen if such a pattern also holds in the natural environment.

Mechanisms

Investigators have proposed a number of mechanisms underlying the putative anxiolytic effect of alcohol. These include the view that stress reduction is mediated by a direct pharmacologic influence on the nervous system, as well as the notion that stress reduction is mediated by the effects of alcohol on information processing. Both perspectives recognize that the anxiolytic effects of alcohol are potentially reinforcing and can contribute to alcohol use.

Among those who argue that the anxiolytic effect of alcohol results from a direct effect on the nervous system, there has been debate over whether stress reduction is centrally or peripherally mediated. Support for peripheral mediation comes from research showing alcohol to increase peripheral vasoconstriction while dampening heart rate during a stressor administration. Such work suggests that stress reduction could be restricted to those cardiovascular functions mediated primarily by β -adrenergic action. Other studies have not supported the view that the anxiolytic effects of alcohol are primarily a function of β -adrenergic blocking action. Data from a variety of sources instead seem to favor the role of central nervous system functioning in mediating the putative anxiolytic effect of alcohol. More research is needed to better establish the pharmacologic mechanism for stress reduction.

In contrast to theories positing a direct effect of alcohol on stress response systems are models that view an anxiolytic effect to be a result of the pharmacologic effects of alcohol on information processing.

These models propose that, at moderate and heavy doses, alcohol will impair cognitive processing. This cognitive impairment may or may not alleviate stress, depending on one's circumstances. There are several different models that relate to this indirect explanation. One suggests that alcohol's narrowing of perception to immediate cues and reduction of cognitive abstracting capacity restrict attention to the most salient, immediate aspects of experience. Accordingly, the concurrent activity in which an intoxicated person engages helps to determine the effects of alcohol. Intoxication in the presence of concurrent distraction is predicted to attenuate stress responding, whereas, without a neutral or pleasantly distracting activity, intoxication is not predicted to prove anxiolytic.

Other approaches suggest that alcohol serves to disrupt the appraisal or encoding of new information. Depending on the situation, this cognitive impairment may be anxiolytic. For example, alcohol may be more likely to reduce stress if intoxication precedes the stressor, than if it follows a stressor, as in the latter case, the stressful information would already be encoded prior to drinking. It has also been suggested that alcohol may be particularly effective in disrupting the encoding of information in terms of its self-relevance. The inhibition of encoding processes presumably leads to a reduction in performance-based self-evaluation, which in situations where such evaluation is unpleasant, may reduce stress responses, thus increasing the probability of drinking.

Effects of Stress on Alcohol Consumption

Animal Studies

Animal studies of the effects of stress on drinking behavior have used three types of stressors: shock, restraint, and psychological stressors such as isolation and dominance, sleep deprivation, and early weaning and handling. Alcohol ingestion has been measured during the stressor administration, on the day of the stressor, or on the days following the stressor.

Studies using laboratory animals have found that both physical and psychological stressors are associated with an increased tendency to develop a pattern of drug or alcohol use. One explanation for this relationship is that stress increases the reinforcement properties of drugs, perhaps by increasing glucocorticoid secretion, which in turn may enhance the drug-induced release of dopamine.

Across a number of studies, it appears that the probability of alcohol ingestion is especially enhanced following the termination of a stressor. This may be accounted for in part by the ability of alcohol ingestion to replenish endogenous opiate levels that may become depleted during periods of stress. Studies also suggest that a number of factors may affect the

relationship between stress and drinking in animals. Individual differences in preference for alcohol may affect the stress–alcohol interaction, with stress particularly likely to increase subsequent drinking among animals that have low initial preference for alcohol. Another difference that has been observed concerns neonatal experience, with rats that were weaned earlier showing increased alcohol consumption. Genetic factors have also been implicated as moderators of the effects of stress on alcohol ingestion, with various strains of mice ingesting different amounts of alcohol following stress.

Human Studies

In the 1940s, cross-cultural studies led to the proposal that societies that experienced the most stress were also the societies with the highest per capita levels of drinking. To examine whether stress precipitates alcohol consumption in humans, both survey and experimental studies have been conducted.

A number of retrospective studies have examined the relationship between stress and drinking. These studies indicate a modest increase in drinking (as well as a modest increase in heavy drinking) following stress. This effect may be influenced by a host of individual difference factors, including ethnicity, gender, and age.

Prospective studies have accumulated that assess both drinking and stress over specified time periods. These studies generally reveal a tendency for stress levels to correlate with drinking levels. Further, in many instances, individuals appear to drink in order to cope with stress.

In addition to survey research, experiments have been designed to test the causal relationship between stress and drinking. Often these studies first expose participants to some form of stressor (e.g., threat of shock, a stressful social interaction, or false performance feedback). Next, their alcohol intake is assessed while they participate in what is ostensibly a wine or beer taste-rating task. The critical dependent measure, unbeknown to participants, is the quantity of alcohol consumed. Data from these tasks have been inconclusive. Some studies have found clear effects, such that individuals who were stressed consumed more alcohol than their nonstressed counterparts. Other studies have failed to observe this pattern, and in some cases the reverse has been found, with stress leading to a reduction in drinking. As previously noted with respect to animal research, it may be that drinking is especially likely to increase after the stressor has terminated. Another issue that may obfuscate the findings with human participants is the varied approaches to inducing stress. In some cases, for

example, it is unclear just how effective the stressors were. In summary, while laboratory studies have been equivocal, findings from survey research suggest that increases in stress are associated with increased alcohol consumption.

Further Reading

- Greeley, J. and Oei, T. (1999). Alcohol and tension reduction. In: Leonard, K. & Blane, H. (eds.) *Psychological theories of drinking and alcoholism* (2nd ed., pp. 14–53). New York: Guilford Press.
- Conger, J. (1956). Reinforcement theory and the dynamics of alcoholism. *Quarterly Journal of Studies on Alcohol* 17, 296–305.
- Cooper, M. L., Russell, M., Skinner, J. B., et al. (1992). Stress and alcohol use: Moderating effects of gender, coping, and alcohol expectancies. *Journal of Abnormal Psychology* 101, 139–152.
- Curtin, J. J., Patrick, C. J., Lang, A. R., et al. (2001). Alcohol affects emotion through cognition. *Psychological Science* 12, 527–531.
- Hull, J. G. (1987). Self-awareness model. In: Blane, H. & Leonard, K. (eds.) *Psychological theories of drinking and alcoholism*, pp. 272–304. New York: Guilford Press.
- Langenbucher, J. W. and Nathan, P. E. (1990). Alcohol, affect, and the tension reduction hypothesis: the re-analysis of some crucial early data. In: Cox, M. A. (ed.) *Why people drink: the parameters of alcohol as a reinforcer*, pp. 151–168. New York: Gardner.
- Piazza, P. V. and Le Moal, M. (1998). The role of stress in drug self-administration. *Trends in Pharmacological Science* 19, 67–74.
- Pohorecky, L. A. (1990). Interaction of ethanol and stress: research with experimental animals – an update. *Alcohol and Alcoholism* 25, 263–276.
- Pohorecky, L. A. (1991). Stress and alcohol interaction: an update of human research. *Alcoholism: Clinical and Experimental Research* 15, 438–459.
- Sayette, M. A. (1993). An appraisal-disruption model of alcohol's effects on stress responses in social drinkers. *Psychological Bulletin* 114, 459–476.
- Sayette, M. A., Martin, C. S., Perrott, M. A., et al. (2001). A test of the appraisal-disruption model of alcohol and stress. *Journal of Studies on Alcohol* 62, 247–256.
- Sher, K. (1987). Stress response dampening. In: Blane, H. & Leonard, K. (eds.) *Psychological theories of drinking and alcoholism*, pp. 227–271. New York: Guilford Press.
- Sher, K. (1991). *Children of alcoholics: a critical appraisal*. Chicago: University of Chicago Press.
- Sher, K. J., Wood, M. D., Richardson, A. E., et al. (2005). Subjective effects of alcohol I. In: Earleywine, M. (ed.) *Mind-altering drugs: the science of subjective experience*, pp. 86–134. New York: Oxford University Press.
- Steele, C. M. and Josephs, R. A. (1990). Alcohol myopia: its prized and dangerous effects. *American Psychologist* 45, 921–933.
- Volpicelli, J. R. (1987). Uncontrollable events and alcohol drinking. *British Journal of Addiction* 82, 381–392.

Alcohol, Alcoholism, and Stress: A Psychobiological Perspective

A N Taylor and P Prolo

University of California, Los Angeles, and Veterans Administration Greater Los Angeles Health Care System, Los Angeles, CA, USA

M L Pilati

Rio Hondo College, Whittier, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A N Taylor and M L Pilati, volume 1, pp 131–135, © 2000, Elsevier Inc.

Relapse prevention model

A model that proposes that the goal of treatment for addictive behaviors is to develop the means for coping with stressful situations that put the addict at risk for relapse.

Stress-response-dampening model

A model that proposes that alcohol use is reinforced by its ability to ameliorate responses to stress.

Introduction

Alcohol and the Stress Response

Effects of Stress on Alcohol Consumption

Modifiers of the Stress–Alcohol Interaction

Conclusion

Glossary

Alcohol abuse

The use of alcohol in a manner harmful to the individual and/or others, such as contributing to the failure to fulfill familial or work obligations but not a severe enough problem to meet criteria for alcohol dependence. *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (DSM-IV-TR), makes a distinction between substance abuse and dependence; once the criteria for dependence have been met, a diagnosis of substance abuse is not possible.

Alcohol dependence

According to DSM IV-TR criteria, a failure to control alcohol use despite repeated attempts, continued use despite the knowledge of adverse consequences, and, in general, compulsive use of alcohol. Alcohol dependence may or may not be accompanied by physiological dependence.

Alcoholism

Excessive alcohol use. Although there is no formally accepted definition of alcoholism, the criteria for alcohol dependence are most frequently used by researchers.

Circadian rhythms

Biological rhythms that have a cycle of about 24 h; derived from the Latin *circa diem* “about a day.”

Endorphin compensation hypothesis

A hypothesis that proposes that the use of alcohol following stress is an attempt to compensate for a decrease in the levels of β -endorphin resulting from the termination of the stressor.

Introduction

The effects of alcohol on the mind, brain, and body are as diverse as the effects of stress. However conceptualized, stress is consistently associated with psychological manifestations, including anxiety, irritability and anger, sad and depressed moods, tension, fatigue, and with certain bodily manifestations, including perspiration, blushing or blanching of the face, increased heart rate, decreased blood pressure, and intestinal cramps and discomfort. Sleep duration and quality are also severely impaired by stress. In the field of substance addiction, both acute and chronic alcohol intake have a disruptive effect on a variety of stressors, both at physiological and psychosocial levels. Three fundamental issues arise: (1) the consequences of stress on normal subjects and alcoholics are not uniform, and the psychopathological and physiopathological impact of stress may be significantly greater in some people than in others; (2) the impact of stress is dynamic and multifaceted, such that the same person may exhibit a variety of manifestations of the psychoneuroendocrine-immune stress response with varying degrees of severity in different situations; and (3) the outcome of stress can be variable in the sense that normal subjects and alcohol abusers may position themselves along the spectrum of allostatic regulation, somewhere between the allostatic state (i.e., toward regaining physiological balance) and allostatic overload (i.e., toward physiological collapse, with associated potential onset of varied pathologies). In brief, subjects with different levels of chronic stress (i.e., major life events or minor hassles) may be expected to respond dramatically differently. Finally, inherited changes in response to alcohol, such as an increase in the ability of alcohol to lessen the impact of stressful events, may themselves confer susceptibility to alcohol abuse. Our understanding of the relationships between alcohol

and stress is based on different lines of research both on animal and human subjects.

Alcohol and the Stress Response

Alcohol and the Neuroendocrine-Immune Response to Stress

Acute alcohol administration has a stimulatory action on the hypothalamic-pituitary-adrenal (HPA) axis and related brain systems in humans and rodents, that is, the brain and pituitary β -endorphin (β -EP) systems and the sympathetic nervous system. Alcohol-induced neuroendocrine effects are characterized by dose-dependent elevations in plasma glucocorticoids, adrenocorticotrophic hormone (ACTH), β -EP, and catecholamines. Ethanol activation of the HPA axis is abolished by hypophysectomy or treatment with corticotropin-releasing hormone (CRH) antiserum in rodents, suggesting that it is primarily modulated by ACTH secretion and that CRH is an essential intermediate in the stimulation of ACTH by ethanol. Acute ethanol exposure generally activates the HPA axis when blood alcohol concentrations (BACs) exceed intoxicating levels of 100 mg dl⁻¹, although HPA activation can also occur at lower BACs. In contrast, the ethanol-induced HPA response of some individuals may be either unaffected or blunted. There are also contributions from environmental and genetic factors to this variability.

Chronic exposure to ethanol has a more variable effect on the HPA axis and the β -EP system, ranging from stimulation to either response attenuation or tolerance. For example, a small percentage of alcoholic individuals (<5%) develop clinical features of hypercortisolism or Cushing's syndrome. However, animal studies suggest that chronic ethanol exposure can result in tolerance to the stimulatory effect of ethanol on CRH release in vitro and can impair the ability of the HPA axis to respond to stress. Indeed, alcoholic men have been reported to show blunted ACTH and cortisol responses to CRH and ACTH challenges, as well as other signs of HPA axis dysfunction (i.e., impaired feedback regulation).

Alcohol consumption acts on the autonomic nervous system as well as the neuroendocrine system, in particular the HPA axis. The latter plays a pivotal role in regulating immune surveillance mechanisms, including the production of cytokines that control the inflammatory process as well as events responsible for healing. Given the multidirectional communication among the neuroendocrine, immune, and nervous systems, alcohol can be expected to impact all three interacting systems, in both feed-forward and feedback directions.

Alcohol and the Cardiovascular Response to Stress

The impact of alcohol on cardiovascular responses varies with the state of the subject at the time of alcohol exposure (i.e., baseline or post-stress) and with the cardiovascular measures employed. Although alcohol has been shown to increase resting heart rate, it also is known to produce a paradoxical decrease in blood pressure in both humans and animals. Thus, alcohol's effect on heart rate is comparable to that of stress, whereas its impact on blood pressure is inconsistent. Heart rate, as a psychophysiological measure of the stress response, has provided evidence for the general assumption that alcohol is often used as a means of relief from the physical and psychological effects of stress. However, human research designed to assess the stress-response-dampening model of alcohol use has had mixed results. There is some indication that contradictions in experimental findings may be due to varying alcohol doses, differences in the subjects employed, and differences in the timing of the stress exposure and the alcohol presentation. Making the issue even more complex is the evidence indicating that altered responses to alcohol may have a genetic basis, possibly even conferring susceptibility to alcohol abuse on those with a family history of such substance-abuse problems. Studies failing to see a stress-response dampening effect of alcohol have often employed rather low doses, providing a logical explanation for differing results; it has also been established that alcohol increases baseline levels of physiological arousal. Thus, when looking at the effects of alcohol on the response to a given stressor and calculating change from baseline, results will differ depending on whether the baseline used is a pre- or postethanol measurement.

Effects of Stress on Alcohol Consumption

Stress and Alcohol Use

Animal studies indicate that stress increases alcohol consumption when the stress, whether physical or social, is chronic, unavoidable, and uncontrollable. Evidence from human and animal studies indicates that alcohol has profound anxiolytic effects. In rodents, tolerance to the anxiolytic effects of low or moderate doses of ethanol was shown to develop rapidly and, as such, may explain the augmented consumption of alcohol required to maintain its tension-reducing effects. In addition, adverse experiences early in life, such as maternal separation, have been found to be associated with excessive alcohol consumption in monkeys as well as in humans. Furthermore, prenatal factors, such as maternal stress and drugs or alcohol consumption during pregnancy,

can dysregulate physiological and behavioral stress responses in adulthood. Such findings suggest that early life experiences can result in long-lasting psychobiological changes that may manifest as enhanced alcohol consumption in order to achieve tension reduction.

Consistent with the stress-response dampening model, it is a common and long-held assumption that alcohol is often used as a means of coping with stress. Recently, however, with the observation that alcohol may or may not dampen physiological responses to stress, new theories have evolved. One that is consistent with many observations and the effectiveness of new pharmacological treatments for alcoholism is the endorphin compensation hypothesis. This hypothesis proposes that alcohol is consumed following the termination of stress because the presence of stress increases levels of the endogenous opioid β -EP. Alcohol is sought for its ability to increase β -EP levels and compensate for the deficit that results when the stress has been removed. Consistent with this hypothesis is the use of naltrexone, an opiate antagonist, in the treatment of alcoholism.

Sleep and Alcohol Use

Sleep disturbance is a common complaint of alcohol-dependent patients. Various studies have demonstrated a strong association between poor sleep and psychological factors, especially in relation to perceived stress and associated increases in CRH and cortisol, the hypothalamic and adrenal products of the HPA axis. Both hormones are known to lead to arousal and sleeplessness in human subjects and experimental animals.

The daily sleep-wakefulness cycle is synchronous with circadian patterns of body temperature in humans and animals such that maximum body temperatures occur during the active portion of each day and minimum body temperatures occur during the inactive or sleep portion. In rodents, the daily active period corresponds with the dark phase of the daily light-dark cycle while sleep occurs during the light phase. Acute administration of ethanol was found to alter the rodent's circadian body temperature and activity rhythms in a dose- and time-dependent manner. Chronic exposure to ethanol has also been shown to alter the amplitude of the rat's circadian body temperature and activity rhythms. Confirmation of these findings in translational studies on human subjects could provide a further basis for alcohol-induced sleep disturbances.

Trauma and Alcohol Use

Despite the controversy as to whether or not alcohol is used to minimize the impact of stress or to

compensate for some post-stress condition, there is an abundance of correlational data linking increases in alcohol use with exposure to or the experience of traumatic or uncontrollable stressful events. For example, studies have demonstrated that there is an increase in social indices of alcohol use (such as arrests for driving while intoxicated) following disasters.

Not only are changes in a population's alcohol consumption associated with traumatic events that affect a community (e.g., natural disasters), but a high percentage of those seeking treatment for alcoholism are survivors of abuse or suffering from post-traumatic stress disorder or other anxiety disorders. Although it may be difficult to determine a direct relationship between victimization and alcohol misuse, attempts to define whether or not alcohol abuse preceded, coincided with, or followed trauma have consistently found that trauma did, in fact, precede alcohol misuse. Furthermore, trauma severity is positively correlated with alcohol consumption and the severity of alcohol-related problems. There is reason to believe that, whereas stress plays a role in the etiology of all drug dependence, the drug selected by men is more likely to be alcohol. Thus, there appears to be a gender difference in the drug of choice for the individual struggling with the aftermath or presence of stressful life conditions.

Traumatic Brain Injury and Alcohol Use

Traumatic brain injury (TBI) and its sequelae are of pivotal importance in the young generation (e.g., in victims of car accidents and in deployed soldiers and war veterans) with consequences ranging from physical disabilities to long-term behavioral and social deficits. Close to 25% of TBI cases are associated with alcohol abuse, and alcoholism may, in and of itself, complicate long-term TBI recovery. A growing body of evidence raises the possibility that alcohol intoxication or dependence in combination with TBI may lead to increased deleterious effects on cognitive functioning. Moreover, serious alterations in neuroendocrine and immune regulation have been noted following chronic alcohol use and TBI in animal models and humans that can impair recovery and rehabilitation efforts after brain injury. However, the role of alcohol as an interactive or additive agent in these findings has not been fully assessed.

Alcoholism, Stress, and Relapse

Stressful life events have long been thought to play a role in the return to alcohol use following cessation of problematic drinking behavior. The role of stress in relapse has received much attention, with more recent

reviews concluding that less severe psychosocial stress may not increase the likelihood of relapse, whereas severe acute stressors and chronic stressors perceived as highly threatening may contribute to an increased risk of relapse. In considering the role of stress in relapse, however, the capabilities of the individual must be considered. The impact of even severe stress and whether this leads to relapse are mediated by both protective and risk factors that render the individual more or less likely to return to problematic alcohol use. Risk factors that may increase the likelihood of relapse to substance abuse include the persisting physiological symptoms of abstinence (e.g., HPA axis dysfunction, characterized by a blunted ACTH response to a challenge with CRH).

The role of stress has been the focus of many alcoholism treatments. Most notably, the relapse prevention model introduced by Marlatt and George in 1984 proposes that the goal of treatment for addictive behaviors must be to develop the means for coping with stressful situations. In order to change addictive patterns, situations that are high-risk for relapse need to be identified and methods for coping with such situations developed. According to the model, cessation of addictive behavioral patterns can be achieved through the mastery of handling such high-risk situations. The relapse prevention model has been the focus of much research, although empirical support for its effectiveness is difficult to obtain. Despite research suggesting that relapse does occur in situations identified as being high-risk, the subjective nature of such retrospective reports calls any definitive conclusions into question.

Modifiers of the Stress–Alcohol Interaction

Genetics

Various inbred murine and rodent strains differ in their sensitivities to alcohol at the endocrine and behavioral levels. In the past 2 decades, with the application of selective breeding techniques, rodent models of alcoholism have been derived with high and low oral alcohol preference. Such animal models are particularly useful in the identification of genetic traits associated with the physiological and behavioral bases of alcohol drinking. For example, results indicate that the β -EP system of alcohol-preferring rats is hyperresponsive to alcohol. Thus, the activation of the endogenous opioid system may either enhance the reinforcing effects of low-dose alcohol or attenuate the aversive effects of high-dose alcohol and thereby affect alcohol consumption,

consistent with the previously noted endorphin compensation hypothesis.

The genetic basis for susceptibility to alcohol abuse and dependence has long been accepted, and numerous studies have been conducted to determine how those with a family history of alcoholism (family history positive, FHP) might differ from those without such an inherited risk (family history negative, FHN). There is increasing support that altered neuroendocrine responses to alcohol in the sons of alcoholic men provide a possible basis for the genetic component in the development of alcoholism. For example, nonalcoholic male offspring of alcoholic fathers (known to be at three- or fourfold higher risk for alcoholism than sons of nonalcoholics) demonstrate, in comparison to the sons of non-alcoholics, lower cortisol responses to alcohol, an impaired ACTH response to CRH, increased alcohol-induced plasma β -EP levels, and an enhanced ACTH response to the opioid receptor antagonist, naloxone. These findings indicate that FHP men have altered hypothalamic CRH neuronal activity. Although it is not clear how these alterations may contribute to the development of abusive patterns of alcohol use, these studies suggest that the CRH-mediated response of the HPA axis or the β -EP system to alcohol may serve as a marker to distinguish individuals at risk for alcoholism. Interestingly, these markers do not effectively distinguish FHP women, supporting the notion that there are gender differences in the role of genetic factors in the etiology of alcoholism.

Expectancy

Adding an additional level of complexity to the determination of the effects of alcohol on the human stress response is the vast literature demonstrating that expectancies have a significant impact on the response to alcohol ingestion. It may be that many reported differences in the literature are created by the influence of cognition on reactivity to alcohol and are due to the timing of stress presentation with respect to alcohol use. Indeed, it appears that intoxication prior to stress exposure may alter the cognitive reaction to the stimulus and, as a result, the response to the stressful stimulus may be minimized. Thus, it may be that alcohol does not merely dampen the stress response but that alcohol may also act to reduce the physiological response to a stressor by altering the appraisal of the stressful stimulus. Such a conclusion is supported by studies that find that more complex paradigms are more likely to result in alcohol stress-response dampening.

Conclusion

There is clearly a relationship between alcohol and stress, although the nature of this association is complex and not easily described. A brief look at the relevant literature demonstrates that the interaction of alcohol and stress varies with characteristics of the subject studied, the environment, the alcohol dose, the nature of the stressor imposed, and the timing of the exposure to alcohol and the stressor. Thus, stress, alcohol, and alcoholism are interrelated, but the nature of this relationship is highly variable. The ineffectiveness of many strategies for prevention and treatment of alcohol abuse and the variation in clinical history represent a challenge and an opportunity to better understand the relationship between alcohol and stress.

See Also the Following Articles

Alcohol and Stress: Social and Psychological Aspects; Beta-Endorphin; Drug Use and Abuse; Ethanol and Endogenous Opioids.

Further Reading

- Aston-Jones, G. and Harris, G. C. (2004). Brain substrates for increased drug seeking during protracted withdrawal. *Neuropharmacology* 47(supplement 1), 167–179.
- Besedovsky, H. O. and Del Rey, A. (2001). Cytokines as mediators of central and peripheral immune-neuroendocrine interactions. In: Ader, R., Felten, D. L. & Cohen, N. (eds.) *Psychoneuroimmunology*, (3rd edn. Vol. 1), pp. 1–17. San Diego, CA: Academic Press.
- Cloninger, C. R. (1987). Neurogenetic adaptive mechanisms in alcoholism. *Science* 236, 410–416.
- Corrigan, J. D., Whitneck, G. and Mellick, D. (2004). Perceived needs following traumatic brain injury. *Journal of Head and Trauma Rehabilitation* 19, 205–216.
- Crum, R. M., Ford, D. E., Storr, C. L., et al. (2004). Association of sleep disturbance with chronicity and remission of alcohol dependence: data from a population-based prospective study. *Alcoholism: Clinical and Experimental Research* 10, 1533–1540.
- Froelich, J. C. (1993). Interactions between alcohol and the endogenous opioid system. In: Zakhari, S. (ed.) *Alcohol and the endocrine system*, NIAAA research monograph, 23, pp. 21–36. Bethesda, MD: National Institutes of Health.
- Higley, J. D., Hasert, M. F., Suomi, B. J., et al. (1991). Nonhuman primate model of alcohol abuse: effects of early experience, personality, and stress on alcohol consumption. *Proceedings of the National Academy of Sciences USA* 88, 7261–7265.
- Marlatt, G. A. and George, F. R. (1984). Relapse prevention: introduction and overview of the model. *British Journal of Addiction* 79, 261–273.
- Pohorecky, L. A. (1991). Stress and alcohol interaction: an update of human research. *Alcoholism: Clinical and Experimental Research* 15, 438–459.
- Schuckit, M. (1998). Biological, psychological and environmental predictors of the alcoholism risk: a longitudinal study. *Journal of Studies on Alcohol* 59, 485–494.
- Stewart, S. H. (1996). Alcohol abuse in individuals exposed to trauma: a critical review. *Psychology Bulletin* 120, 83–112.
- Taylor, A. N., Chiappelli, F. and Yirmiya, R. (2001). Fetal alcohol syndrome and immunity. In: Ader, R., Felten, D. L. & Cohen, N. (eds.) *Psychoneuroimmunology* (3rd edn. Vol. 2), pp. 49–71. San Diego, CA: Academic Press.
- Taylor, A. N., Tio, D. L., Heng, N. S., et al. (2002). Alcohol consumption attenuates febrile responses to lipopolysaccharide and interleukin-1 β in male rats. *Alcoholism: Clinical and Experimental Research* 26, 44–52.
- Vgontzas, A. N., Zoumakis, M., Bixler, E. O., et al. (2003). Impaired nighttime sleep in healthy old vs. young adults is associated with elevated plasma IL-6 and cortisol levels: physiologic and therapeutic implications. *Journal of Clinical Endocrinology and Metabolism* 88, 2087–2095.
- Volpicelli, J. R., Tiven, J. and Kimmel, S. C. (1987). Uncontrollable events and alcohol drinking. *British Journal of Addiction* 82, 381–392.
- Waltman, C., McCaul, M. E. and Wand, G. S. (1994). Adrenocorticotropin responses following administration of ethanol and ovine corticotropin-releasing hormone in the sons of alcoholics and control subjects. *Alcoholism: Clinical and Experimental Research* 18, 826–830.
- Wand, G. S., Mangold, D. and Mahmood, M. (1999). Adrenocorticotropin responses to naloxone in sons of alcohol-dependent men. *Journal of Clinical Endocrinology and Metabolism* 84, 64–68.
- Zimmermann, U., Spring, K., Kunz-Ebrecht, S. R., et al. (2004). Effect of ethanol on hypothalamic-pituitary-adrenal system response to psychosocial stress in sons of alcohol-dependent fathers. *Neuropsychopharmacology* 29, 1156–1165.

Aldosterone and Mineralocorticoid Receptors

J W Funder

Prince Henry's Institute of Medical Research,
Clayton, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition
article by J W Funder, volume 1, pp 141–144,
© 2000, Elsevier Inc.

Background

Aldosterone Secretion: Adrenal and
Extra-adrenal

Regulation of Aldosterone Secretion

Mineralocorticoid Receptors: Cloning

Mineralocorticoid Receptors: Characterization

Aldosterone Specificity-Confering Mechanisms:
Transcortin, 11 β HSD2

Apparent Mineralocorticoid Excess

Extraepithelial Actions of Aldosterone

Aldosterone, Cardiac Hypertrophy, and
Cardiac Fibrosis

Aldosterone: A Stress Hormone?

MR Activation and Stress

Glossary

<i>Aldosterone</i>	The classic salt-retaining steroid hormone secreted from the zona glomerulosa of the adrenal cortex.
<i>Aldosterone synthase/ CYP11B2</i>	The enzyme abundantly expressed in adrenal zona glomerulosa and reported in PCR studies on heart, blood vessels, and brain, which, by a sequential three-step process, converts the methyl (-CH ₃) group at carbon 18 to an aldehyde (-CHO) group.
<i>Apparent mineralocorticoid excess (AME)</i>	An autosomal recessive syndrome of inactivating mutations in the gene coding for 11 β -hydroxysteroid dehydrogenase type 2, producing juvenile hypertension, marked salt retention reflecting inappropriate cortisol activation of epithelial mineralocorticoid receptors, and premature death.
<i>Carbenoxolone</i>	The hemisuccinate of glycyrrhetic acid, previously used as a treatment for peptic ulcer.
<i>Glycyrrhetic acid</i>	The active principle of licorice, which can block 11 β -hydroxysteroid dehydrogenase, leading to sodium retention and blood pressure elevation.

*11 β -Hydroxysteroid
dehydrogenase
type
2 (11 β HSD2)*

*Mineralocorticoid
receptors
(MR)*

The enzyme expressed in very high abundance (3–4 million copies per cell) in epithelial aldosterone target cells that converts cortisol and corticosterone (which have affinity for mineralocorticoid receptors similar to that of aldosterone) to receptor-inactive metabolites, and NAD to NADH, thus allowing aldosterone to activate the relatively non-selective mineralocorticoid receptors.

The intracellular receptors that have equal high affinity for aldosterone, corticosterone, and cortisol and via which aldosterone acts in epithelial tissues to promote unidirectional transepithelial sodium transport.

Background

Classically, aldosterone is the physiological mineralocorticoid hormone uniquely secreted from the zona glomerulosa of the adrenal cortex, and acting via intracellular mineralocorticoid receptors in epithelial target tissues to promote unidirectional transepithelial sodium transport. While this remains the case, over the past decade there has been mounting evidence for additional complexity in terms of aldosterone secretion and action, and of MR activation by glucocorticoids. For aldosterone, this ranges from extra-adrenal aldosterone synthesis, to major pathophysiological actions via nonepithelial mineralocorticoid receptors, to rapid nongenomic actions via a range of receptor mechanisms. There remain major residual areas of ignorance in terms of mechanism. The first aldosterone-induced protein was identified in 1999, a serine-threonine kinase (sgk) that directly or indirectly activates epithelial sodium channels. The mechanism of SGK action, and potential roles for other aldosterone-induced proteins in both epithelial and nonepithelial tissue, remain to be established.

Aldosterone Secretion: Adrenal and Extra-adrenal

Aldosterone was isolated in 1953. It was initially called electrocortin, reflecting its effects on electrolyte transport, but was given the definitive name of aldosterone in recognition of its unique aldehyde (-CHO) group at carbon 18, in contrast with the methyl group in other steroids. Formation of this aldehyde group, which is crucial for the ability of aldosterone to access

mineralocorticoid receptors in epithelia (*vide infra*), is achieved by a three-step process catalyzed by the enzyme aldosterone synthase (CYP11B2), expression of which in the adrenal cortex is confined to the outermost cell layer, the zona glomerulosa. By polymerase chain reaction (PCR) techniques, aldosterone synthase expression has recently been described in blood vessels, heart, and brain, with demonstration of enzyme activity reported for the two former tissues. While possible extra-adrenal synthesis does not appear to contribute to circulating hormone levels, the extent to which it may support local tissue effects of aldosterone *in vivo* remains to be determined.

Regulation of Aldosterone Secretion

In vivo aldosterone secretion from the zona glomerulosa is predominantly and independently regulated by angiotensin II and plasma potassium concentrations. Adrenocorticotrophic hormone (ACTH) can also raise aldosterone secretion, but unlike the other two secretagogues, its action is not sustained. While a myriad of other possible modulators of aldosterone secretion have been reported, including endothelin, nitric oxide, and uncharacterized factors from the pituitary and adipose tissue, the relative importance of such inputs to pathophysiological control *in vivo* has not been established. Both experimentally and clinically, changes in angiotensin II and/or plasma potassium appear sufficient to explain changes in aldosterone secretion rates from nonneoplastic adrenals.

Mineralocorticoid Receptors: Cloning

Classical mineralocorticoid receptors binding aldosterone with high affinity and blocked by the antagonist spironolactone have been studied since the mid-1970s, with the human MR cloned and expressed by the late 1980s. The human MR comprises 984 amino acids and, like other members of the steroid/thyroid/retinoid/orphan receptor superfamily, has a domain structure that includes a ligand-binding domain (LBD) and a DNA-binding domain (DBD). Like other members of the receptor superfamily, MR act by regulating gene expression; sgk expression in A6 cells is actinomycin inhibitable, and aldosterone-induced sodium transport is both actinomycin and puromycin inhibitable in toad bladder models of mineralocorticoid action. Mineralocorticoid receptors are part of a subfamily with the receptors for glucocorticoids, progestins and androgens, sharing $\geq 50\%$ amino acid identity in the LBD and $\geq 90\%$ in the DBD.

Mineralocorticoid Receptors: Characterization

The initial MR cloning and expression studies also highlighted two questions regarding aldosterone action, questions that had been previously addressed by studies on MR in tissue extracts. First, when mRNA from a variety of rat tissues was probed with MR cDNA, the highest levels of MR expression were found in hippocampus rather than classic epithelial aldosterone target tissues. Second, when the affinity of various steroids was determined by their ability to compete for [3 H]aldosterone binding to expressed human mineralocorticoid receptors, cortisol was shown to have high affinity for MR, equivalent to that of aldosterone, and corticosterone (the physiological glucocorticoid in rat and mouse) slightly higher affinity. There are thus at least two operational questions, of which the first is the physiological role(s) of MR in nonepithelial tissues, such as hippocampus (and other brain areas and heart, for example). The second is how aldosterone can selectively activate MR, in epithelia or in nonepithelial tissues, in the face of at least equivalent receptor affinity and the much higher circulating levels of the physiological glucocorticoids.

Aldosterone Specificity-Conferring Mechanisms: Transcortin, 11 β HSD2

Differential plasma binding of steroids has long been recognized as one factor favoring aldosterone occupancy of MR. The physiological glucocorticoids cortisol and corticosterone circulate at least 95% protein bound, predominantly to the specific globulin transcortin but also to albumin; aldosterone has negligible affinity for transcortin and is only ~50% albumin bound in plasma, giving it an order of magnitude advantage in terms of plasma free to total concentrations. Second, the enzyme 11 β HSD2 is abundantly expressed in epithelial aldosterone target tissues. 11 β HSD2 has a very low K_m (high affinity) for cortisol and corticosterone, acting essentially unidirectionally to convert them to cortisone and 11-dehydrocorticosterone. These 11-keto congeners have very low affinity for MR, and when they do occupy MR, act as antagonists. Aldosterone, in contrast, is not so metabolized; its 11 β -OH group cyclizes in solution with the unique and very active aldehyde group at C18, forming an 11,18-hemiacetal and rendering aldosterone not a substrate for 11 β HSD2.

Apparent Mineralocorticoid Excess

The crucial role of 11 β HSD2 in conferring aldosterone specificity in otherwise nonselective epithelial

MR is underlined by the clinical syndrome of apparent mineralocorticoid excess (AME). In this rare autosomal recessive disorder, patients present with high blood pressure and marked sodium retention, despite low plasma levels of renin and aldosterone. The syndrome can be diagnosed by an abnormally high ratio of urinary free cortisol to cortisone, reflecting loss-of-function mutation of the sequence coding for 11 β HSD2. Under such circumstances the mutant enzyme shows poor or absent ability to convert cortisol to cortisone, and cortisol inappropriately activates MR, leading to uncontrolled sodium retention and hypertension. A milder version of the syndrome can be produced by licorice abuse, as the active principle of licorice (glycyrrhetic acid) blocks 11 β HSD2 activity, as does the erstwhile antipeptic ulcer drug carbenoxolone, the hemisuccinate of glycyrrhetic acid.

Extraepithelial Actions of Aldosterone

The physiological roles of extraepithelial MR and the extraepithelial effects of aldosterone have been elucidated only in part. Hippocampal MR appear involved in integration of baseline circadian ACTH secretion in response to stress; given the high reflection coefficient of aldosterone at the blood–brain barrier, it appears improbable that these receptors, unprotected by 11 β HSD2, are ever occupied by aldosterone. Other brain MR, in contrast, are clearly occupied and at least to some extent activated by aldosterone, being protected by 11 β HSD2. The amygdala, for example, express both MR and 11 β HSD2, and have been shown to be responsible for the aldosterone-induced increase in salt appetite in sodium deficiency or after sodium loss; interestingly, the amygdala appear within the blood–brain barrier, with no evidence for facilitated aldosterone access to MR. Activation of MR in the anteroventral third ventricle (AV3V) region (which is outside the blood–brain barrier) is necessary for mineralocorticoid-salt hypertension: intracerebroventricular infusion of MR antagonists abolishes the elevation of blood pressure following peripheral infusion of aldosterone to rats, or feeding a high-salt diet to salt-sensitive strains of rats. This has generally been interpreted as a direct role for aldosterone in modulating blood pressure via MR activation in the AV3V region; physiologically, however, this appears unlikely, and the MR activation putatively reflects glucocorticoid occupancy of the receptors under particular conditions.

Aldosterone, Cardiac Hypertrophy, and Cardiac Fibrosis

Aldosterone occupancy of cardiac MR has similarly been shown in a variety of experimental models to

produce cardiac hypertrophy and fibrosis. These effects are independent of hemodynamic changes, as clamping blood pressure by the concomitant infusion of intracerebroventricular MR antagonist does not alter the effects of peripherally infused aldosterone on cardiac hypertrophy and fibrosis. They are also crucially dependent on an aldosterone–salt imbalance; infusion of the same dose of aldosterone to rats fed a low-salt diet is not followed by increases in blood pressure or in cardiac hypertrophy and cardiac fibrosis. Cardiac MR, like those in the CNS, are normally unprotected by 11 β HSD2; the cardiac effects seem to require only a minority of MR to be occupied by mineralocorticoid, as shown in studies of transgenic mice expressing 11 β HSD2 specifically in cardiomyocytes. Importantly, and in contrast with epithelia, occupancy of extraepithelial MR by glucocorticoids does not mimic the effect of aldosterone. Also in contrast with the epithelial effects of aldosterone via MR, which have a time course of hours, the extraepithelial effects of aldosterone take days or weeks to be manifest: vascular inflammatory responses are seen after 1–2 days, increases in collagen type I and type III mRNA are seen by the third week of continuous aldosterone infusion, and plateau levels of blood pressure elevation are seen at 4–6 weeks. The mechanisms underlying these epithelial/extraepithelial differences in the action of aldosterone via identical MR are not understood, but presumably reflect at least in part tissue-specific differences in the coactivators, corepressors, and response elements involved.

Aldosterone: A Stress Hormone?

Aldosterone is a stress hormone depending on the definition of stress. Psychological or physical stressors that raise ACTH and adrenal catecholamine output do not lead to major or sustained elevations in aldosterone levels. Physiologically, aldosterone is raised in response to upright posture to defend intravascular volume; longer-term stressors, in terms of fluid and electrolyte balance (e.g., diarrhea, florid sweating, pregnancy, and lactation on a low-sodium diet) raise aldosterone to act in a tight feedback loop to retain sodium and water. Our renin-angiotensin-aldosterone system appears to have evolved in a context of sodium deficiency, which is often more of an issue for herbivores than for omnivores, and we have much less well developed mechanisms for suppressing aldosterone secretion than for elevating it. On contemporary salt intakes, our inability to lower aldosterone secretion commensurately may thus in a counterintuitive sense be in itself a stressor for the organism, which requires 2–5% of the normal modern salt intake except in response to the stress of

reproduction, gastrointestinal infection, or environmental extremes. The evolutionary craving for salt, and its routine addition to a range of foodstuffs, may thus constitute a stress in terms of the cardiovascular system, most noticeably hypertension, but with increasing evidence for direct effects of MR activation on the heart. The RALES trial, in which adding low-dose spironolactone to conventional ACE inhibitor/loop diuretic/digoxin therapy for heart failure substantially lowered mortality, would also seem to constitute evidence for MR activation operating as a stress mechanism in this sense.

MR Activation and Stress

As noted previously, 11 β HSD2 acts in epithelial tissues to allow aldosterone to selectively activate MR. In such tissues, however, the enzyme debulks intracellular glucocorticoid levels by 90%, rather than completely removes cortisol/corticosterone, so that their levels are still ~10 times those of aldosterone – and thus most MR in such tissues are normally occupied but not activated by glucocorticoids when the enzyme is operating. The key to this distinction appears to be the stoichiometric generation of NADH, which alters the redox state of the cell so that MR–glucocorticoid complexes are inactive: when 11 β HSD2 is blocked, levels of NADH fall, and the MR–glucocorticoid complexes are activated. This bivalent effect of cortisol can be reproduced in cardiomyocyte patch clamp studies, in which normally cortisol is an MR antagonist, but which activates MR when the redox state is altered by

installation of oxidized glutathione (GSSG). In pathophysiological terms, this suggests that the MR in nonepithelial tissues (for example, cardiomyocytes and most neurons) that are always essentially occupied by glucocorticoids may be activated by the metabolic or redox state of the cell, in particular, the generation of reactive oxygen species. If this is the case, then metabolic stress may be mediated via MR through a novel mechanism – that of activating a hitherto quiescent steroid–receptor complex, rather than by increasing the secretion rate and blood levels of a stress hormone such as cortisol.

See Also the Following Articles

11 β -Hydroxysteroid Dehydrogenases; Adrenal Cortex; Glucocorticoids, Overview; Mineralocorticoid Receptor Polymorphisms.

Further Reading

- Brilla, C. G. and Weber, K. T. (1992). Mineralocorticoid excess, dietary sodium, and myocardial fibrosis. *Journal of Laboratory and Clinical Medicine* 120, 893–901.
- Curnow, K. M. (1998). Steroidogenesis, enzyme deficiencies and the pathogenesis of human hypertension. *Current Opinion in Endocrinology and Diabetes* 5, 223–228.
- Funder, J. W. (2004). Is aldosterone bad for the heart? *Trends in Endocrinology and Metabolism* 15, 139–142.
- Funder, J. W. (2004). Cardiac synthesis of aldosterone: going, going, gone...? *Endocrinology* 145, 4793–4795.
- Funder, J. W. (2005). How do central mineralocorticoid receptors modulate blood pressure? *American Journal of Physiology* 288, R356–R357.

Allostasis and Allostatic Load

B S McEwen

Rockefeller University, New York, NY, USA

J C Wingfield

University of Washington, Seattle, WA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by B S McEwen, volume 1, pp 145–149, © 2000, Elsevier Inc.

Turning On and Turning Off Allostasis

Causes of Allostatic Overload

Modulation of Allostatic Responses to Allostatic Load at Different Points in the Life Cycle or When Exposed to Different Social Signals

Role of Behavior and Timing in Allostatic Overload

Examples of Allostatic Overload

Implications of Allostatic Overload in Human Society

Conclusions

Glossary

Allostasis

Achieving stability through change. A process that maintains homeostasis, defined as those physiological parameters essential for life, even though the set points and other boundaries of control may change with environmental conditions. There are primary mediators

of allostasis such as, but not confined to, hormones of the hypothalamic-pituitary-adrenal (HPA) axis, catecholamines, and cytokines. Allostasis also clarifies an inherent ambiguity in the term homeostasis by distinguishing between the systems that are essential for life (homeostatic) and those that maintain these systems in balance (allostatic). In contrast to homeostatic systems such as blood oxygen, blood pH, and body temperature, which must be maintained within a narrow range, allostatic (adaptive) systems have much broader boundaries of operation. Allostatic systems enable an organism to respond to its physical state (e.g., awake, asleep, supine, standing, exercising) and to cope with noise, crowding, isolation, hunger, extremes of temperature, physical danger, psychosocial stress, and microbial or parasitic infections. For example, blood pressure rises when an individual gets up in the morning and fluctuates with exercise and other demands of the waking period; these fluctuations make it possible for the individual to function in response to a changing social and physical environment.

Allostatic load and allostatic overload

The cumulative result of an allostatic state (e.g., fat deposition in a bear preparing for the winter, a bird preparing to migrate, or a fish preparing to spawn) is allostatic load. It can be considered the result of the daily and seasonal routines organisms have to obtain food and survive, and extra energy needed to migrate, molt, breed, etc. Within limits, they are adaptive responses to seasonal and other demands. However, if one superimposes on these additional loads of unpredictable events in the environment, disease, human disturbance (for animals in the wild), and social interactions, then allostatic load can increase dramatically. There are two distinctly different outcomes. First, if energy demands exceed energy income and what can be mobilized from stores, then type 1 allostatic overload occurs. For example, breeding birds use increasing food abundance in the spring to raise young. If inclement weather then increases the costs of maintaining homeostasis and the allostatic load of breeding, and at the same time reduces food available to fuel that allostatic load, then negative energy balance results in loss of body mass and suppression of reproduction. Another example is the mass movement of seabirds to

islands in the face of a severe storm that limits access to food. Characteristics of allostatic overload type 1 are that it results in acute increases in glucocorticoids (and possibly other mediators of allostasis) that then trigger facultative physiological and behavioral responses to redirect the individual from its daily or seasonal routine into survival mode. This reduces allostatic load to a level that can be fueled by resources available. Second, if energy demands are not exceeded and the organism continues to take in or store as much or even more energy than it needs, perhaps as a result of stress-related food consumption, choice of a fat-rich diet, or metabolic imbalances (prediabetic state) that favor fat deposition, then type 2 allostatic overload occurs. Besides fat deposition, there are other cumulative changes in other systems, for example, neuronal remodeling or loss in hippocampus, atherosclerotic plaques, left ventricular hypertrophy of the heart, glycosylated hemoglobin and other proteins by advanced glycosylation end products as a measure of sustained hyperglycemia, high cholesterol with low HDL, and chronic pain and fatigue, for example, in arthritis or psoriasis, associated with imbalance of immune mediators. Characteristics of allostatic overload type 2 are intermediate high levels of glucocorticoids (and other mediators of allostasis) that do not trigger facultative physiological and behavioral responses so that the individual remains stuck in a situation of high allostatic load with the deleterious events that follow. Allostatic overload type 2 can be alleviated by a change in environment, as well as by learning to avoid the source of allostatic load, for example, by a change in lifestyle. Medication may be the best recourse to redress the chronic imbalances in type 2 allostatic overload.

Allostatic state

An allostatic state is the sustained elevation or reduction of mediators of allostasis, such as glucocorticoids, catecholamines, or inflammatory cytokines. An allostatic state results in an imbalance of the primary mediators, reflecting excessive production of some and inadequate production of others. Examples are hypertension, a perturbed cortisol rhythm in major depression or after chronic sleep deprivation, chronic elevation of inflammatory cytokines and low cortisol in chronic fatigue syndrome,

imbalance of cortisol, corticotropin-releasing factor (CRF), and cytokines in the Lewis rat that increases risk for autoimmune and inflammatory disorders. Allostatic states can be sustained for limited periods if food intake and/or stored energy such as fat can fuel homeostatic mechanisms (e.g., a flock of birds flying away from a storm, or small mammals forced into a suboptimal habitat by an influx of predators). If imbalance continues for longer periods and becomes independent of maintaining adequate energy reserves, then symptoms of allostatic overload appear. Abdominal obesity is an example of this condition.

Turning On and Turning Off Allostasis

The body responds to challenge, whether exposure to danger, an infection, or a crowded and noisy neighborhood, by turning on an allostatic response, thus initiating a complex pathway for adaptation and coping, and then shuts off this response when the challenge has passed. The primary allostatic responses involve the sympathetic nervous system and hypothalamic-pituitary-adrenal (HPA) axis. Turning on these systems leads to release of catecholamines from sympathetic nerves and the adrenal medulla and the secretion of adrenocorticotrophic hormone (ACTH) from the anterior pituitary gland, which stimulates release of glucocorticoids from the adrenal cortex.

Turning off these responses leads to a return to baseline concentrations of glucocorticoids and catecholamines; this normally happens when the danger is past, the infection is contained, or the living environment is improved. As long as the allostatic response is limited to the period of challenge, protection via adaptation predominates over adverse consequences. However, over weeks, months, or even years, exposure to elevated levels of stress hormones (an allostatic state) can result in chronically high allostatic load and overload, types 1 and 2, with resultant pathophysiological consequences.

Causes of Allostatic Overload

Four situations can lead to allostatic overload, resulting in overexposure to stress hormones over time. The first is frequent stress. For example, blood pressure surges can trigger myocardial infarction in susceptible persons, and experiments in primates have shown that repeated blood pressure elevations over weeks and months accelerate atherosclerosis, which increases the risk of myocardial infarction.

Thus, the frequency of stressful events determines the degree of allostatic load, leading to increased exposure to hormones and other allostatic mediators.

A second cause of allostatic overload is the failure to habituate to repeated challenges, such as in some individuals' cortisol response during a repeated public speaking challenge.

The third cause of allostatic overload is an inability to shut off allostatic responses. This pertains to the termination of stress and also to the decline of these same hormones during the natural diurnal cycle. Blood pressure in some persons fails to recover after acute mental arithmetic stress, and hypertension accelerates atherosclerosis. Intense athletic training also induces allostatic load in the form of elevated sympathetic and HPA activity, which results in reduced body weight, amenorrhea, and the often-related condition of anorexia nervosa. Women with a history of depressive illness have decreased bone mineral density because the allostatic load of chronic, moderately elevated glucocorticoid levels that are associated with depression causes chronic, reduced levels of osteocalcin.

The fourth cause of allostatic overload is inadequate allostatic responses that trigger compensatory increases in other allostatic systems. When one system does not respond adequately to a stressful stimulus, the activity of other systems is increased because the underactive system is not providing the usual counterregulation. For example, if adrenal steroid secretion does not increase appropriately in response to stress, secretion of inflammatory cytokines (which are counterregulated by adrenal steroids) increases. In Lewis rats, the failure to mount an adequate HPA response results in increased vulnerability to autoimmune and inflammatory disturbances; thus, an infection or wound that causes an inflammation will lead to an increased inflammatory response.

Modulation of Allostatic Responses to Allostatic Load at Different Points in the Life Cycle or When Exposed to Different Social Signals

There is growing evidence that allostatic responses to standardized stressors are not constant throughout the life cycle. In vertebrates, the adrenocortical responses to stress may be elevated at one time of year and suppressed at others. Such modulation of allostasis is particularly evident when animals are reproducing. The reasons for this include very short breeding season, severe habitat, increasing age, and uncertain social status. Even though allostatic load increases in the same way or is even more severe, allostatic responses are changed temporarily.

In another model, rats that become subordinate in a psychosocial living situation called the visible burrow system have a stress-induced state of HPA hyporesponsiveness. In these rats, the HPA response to experimenter-applied stressors is very limited, and hypothalamic corticotropin releasing factor messenger RNA levels are abnormally low. Although specific health consequences are not yet known, these stress nonresponsive rats are believed to be the most likely to die over longer times in the visible burrow. Human counterparts of the state of HPA hyporesponsiveness include individuals with fibromyalgia and chronic fatigue syndrome.

Role of Behavior and Timing in Allostatic Overload

Anticipation, or lack of it, and worry can also contribute to allostatic overload. Anticipation is involved in the reflex that prevents us from blacking out when we get out of bed in the morning and is also a component of worry, anxiety, and cognitive preparation for a threat. Anticipatory anxiety can drive the output of mediators such as ACTH, cortisol, and adrenalin; thus, prolonged anxiety and anticipation is likely to result in allostatic load. For example, salivary cortisol levels increase within 30 min after waking in individuals who are under considerable psychological stress due to work or family matters. Moreover, intrusive memories from a traumatic event (e.g., as in post-traumatic stress disorder) can produce a form of chronic stress and can drive physiological responses.

Allostasis, allostatic states, and allostatic overload are also affected by health-damaging or health-promoting behaviors such as smoking, drinking, diet, and amount of exercise. These behaviors are integral to the overall notion of allostasis – how individuals cope with a challenge – and also contribute to allostatic overload by known pathways (e.g., a rich diet accelerates atherosclerosis and progression to non-insulin-dependent diabetes by increasing cortisol, leading to fat deposition and insulin resistance; smoking elevates blood pressure and atherogenesis; exercise protects against cardiovascular disease). Another example in nature is increased corticosterone production of songbirds in response to predator density, which could lead to long-term debilitation if the predator pressure remains constant.

Examples of Allostatic Overload

Cardiovascular and Metabolic Systems

In infrahuman primates, the incidence of atherosclerosis is increased among the dominant males of

unstable social hierarchies and in socially subordinate females. In humans, low job control predicts increased risk for coronary heart disease, and job strain (high psychological demands and low control) results in elevated ambulatory blood pressure at home and increased left ventricular mass index, as well as increased progression of atherosclerosis.

Quantifying allostatic overload, a major challenge, has been attempted with measures of metabolic and cardiovascular pathophysiology. In one study, 10 measures of activity of allostatic systems were assessed in aging subjects in 1988 and again in 1991. Allostatic overload was scored by determining the number of measures for which an individual was in the most extreme quartile, for example, the highest quartile in systolic blood pressure, overnight urinary cortisol, catecholamines, waist-hip ratio (WHR), glycosylated hemoglobin, the ratio of HDL to total cholesterol, or the lowest quartile in DHEA-sulfate or HDL cholesterol. For baseline data in 1988, subjects with higher levels of physical and mental functioning had lower allostatic load scores; these individuals also had less prevalent cardiovascular disease, hypertension, and diabetes. Moreover, among this higher functioning group, those with higher baseline allostatic load scores were more likely to experience incident cardiovascular disease between 1988 and 1991 and were significantly more likely to show declines in cognitive and physical functioning.

The Brain

Repeated stress has consequences for brain function, especially for the hippocampus, which has high concentrations of adrenal steroid receptors. The hippocampus participates in episodic and declarative memory and is particularly important for the memory of context, the time and place of events that have a strong emotional bias, and glucocorticoids are involved in consolidation of contextual fear conditioning. Impairment of hippocampal function decreases the reliability and accuracy of context memories, and this may lead an individual to miss important information and get into a more stressful situation. The hippocampus is also a regulator of the stress response and acts to inhibit shut-off of the HPA stress response.

The mechanism for stress-induced hippocampal dysfunction and memory impairment is twofold. First, acute stress elevates adrenal steroids, which suppresses the mechanisms in the hippocampus and temporal lobe that subserve short-term memory. Stress can impair memory acutely, but fortunately these effects are reversible and relatively short-lived. Second, repeated stress causes an atrophy of dendrites of pyramidal neurons in the CA3 region of the

hippocampus through a mechanism involving both glucocorticoids and excitatory amino acid neurotransmitters released during and in the aftermath of stress. This atrophy is reversible if the stress is short-lived, but stress lasting many months or years can kill hippocampal neurons. Stress-related disorders such as recurrent depressive illness, posttraumatic stress disorder, and Cushing's syndrome are associated with atrophy of the human hippocampus as measured by magnetic resonance imaging. It is not yet clear to what extent this atrophy is reversible or permanent.

Failure to turn off the HPA axis and sympathetic activity efficiently after stress is a feature of age-related functional decline in experimental animals and humans. Stress-induced cortisol and catecholamine secretion returns to baseline more slowly in some aging animals that at the same time show other signs of accelerated aging. One other sign of age-related impairment in rats is that the hippocampus fails to turn off the release of excitatory amino acids after stress; this may accelerate progressive structural damage and functional impairment. Long-term stress also accelerates the appearance of several biological markers of aging in rats, including the excitability of CA1 pyramidal neurons by a calcium-dependent mechanism and loss of hippocampal pyramidal neurons. An important factor may be the enhancement by glucocorticoids of calcium currents in hippocampus in view of the key role of calcium ions in destructive as well as plastic processes in hippocampal neurons.

One speculation is that allostatic overload over a lifetime may cause a loss of resilience in which allostatic systems wear out or become exhausted, leading to either failure to shut off or failure to respond. A vulnerable link in HPA regulation and cognition is the hippocampal region of the brain; according to the glucocorticoid cascade hypothesis, wear and tear on this brain region leads to dysregulation of the HPA axis as well as cognitive impairment. Supporting evidence comes from the finding that some, but not all, aging rats have impairments of both episodic, declarative, and spatial memory and HPA axis hyperactivity, both of which can be traced to hippocampal damage. Recent data in humans, showing hippocampal atrophy and cognitive impairment accompanying progressive elevations of cortisol levels over 4 years, suggests that similar events may occur in our own species.

Early stress and neonatal handling influence the course of aging and age-related cognitive impairment in animals. Early experiences are believed to set the level of responsiveness of the HPA axis and

autonomic nervous system. These systems overreact in animals subject to early unpredictable stress and underreact in animals exposed to the neonatal handling procedure. In the former condition, brain aging is accelerated, whereas in the latter, brain aging is reduced. Life span is also influenced by early-life differences in HPA reactivity: rats that show increased HPA reactivity to novelty as neonates have shorter life spans compared to rats with lesser HPA reactivity to novelty.

The Immune System

The immune system responds to pathogens or other antigens with its own form of allostasis that may include an acute phase response as well as the formation of an immunological memory. At the same time, the immune system reacts to other allostatic systems such as the HPA axis and the autonomic nervous system; these systems tend to contain acute phase responses and dampen cellular immunity. However, not all of the effects are suppressive. Acute stress causes lymphocytes and macrophages to redistribute throughout the body and to marginate on blood vessel walls and within certain compartments such as the skin, lymph nodes, and bone marrow. This trafficking is mediated in part by adrenal steroids. If an immune challenge is not encountered and the hormonal stress signal ceases, immune cells return to the bloodstream. However, when a challenge occurs, as in delayed-type hypersensitivity, acute stress enhances the traffic of immune cells to the site of acute challenge. The immune-enhancing effects of acute stress last for a number of days and depend on adrenal steroid secretion. Thus, acute stress has the effect of calling immune cells to their battle stations, and this form of allostasis enhances responses for which there is an established immunologic memory.

When allostatic overload is increased by repeated stress, the outcome is completely different; the delayed hypersensitivity response is substantially inhibited rather than enhanced. Such suppression of cellular immunity resulting from chronic stress leads to consequences such as increased severity of the common cold, accompanied by increased cold virus antibody titers.

Implications of Allostatic Overload in Human Society

The gradients of health across the range of socioeconomic status relate to a complex array of risk factors that are differentially distributed in human society.

For example, a study of elderly people who had a lifetime of sustained economic hardship pointed to a more rapid decline of physical and mental functioning.

Undoubtedly the best example is the Whitehall studies of the British Civil Service, in which stepwise gradients of mortality and morbidity were found across all six grades of the British Civil Service. Hypertension was a sensitive index of job stress, particularly in factory workers, in other workers with repetitive jobs and time pressures, and in workers whose jobs were unstable due to departmental privatization. Plasma fibrinogen, which predicts increased risk of death from coronary heart disease, is elevated in men in lower British Civil Service grades. In less stable societies, conflict and social instability have been found to accelerate pathophysiological processes and increase morbidity and mortality. For example, cardiovascular disease is a major contributor to the almost 40% increase in the death rate among Russian males in the social collapse following the fall of Communism. Blood pressure surges and sustained elevations are linked to accelerated atherosclerosis as well as increased risk for myocardial infarction.

Another stress-linked change is abdominal obesity, measured as increased WHR. WHR is increased at the lower end of the socioeconomic status gradient in Swedish males, and WHR also increases with decreasing civil service grade in the Whitehall studies. Immune system function is also a likely target of psychosocial stress, increasing vulnerability to infections such as the common cold, as noted previously.

Conclusions

Allostasis and allostatic load and overload are concepts that complement the concept of homeostasis by highlighting the paradoxical protective and damaging effects of the mediators of the stress response. These effects range from the short-term adaptation to danger or injury to the longer-term adaptations to adverse conditions associated with seasonal change, altered food supply, or crowding. They also include the diseases of modern life that are categorized as allostatic overload and represent the maladaptive extremes of responses such as fat deposition that have an adaptive role under conditions of varied food supply. By emphasizing the protective and damaging aspects of the same mediators, it is possible to highlight common mechanisms and interactions among the same mediators that underlie disorders such as diabetes, atherosclerosis, age-related cognitive impairment, autoimmune and inflammatory disorders, and immune dysfunction.

See Also the Following Articles

Homeostasis; Hypothalamic-Pituitary-Adrenal; Stress Effects, Overview.

Further Reading

- Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., et al. (1994). Socioeconomic status and health: the challenge of the gradient. *American Psychologist* 49, 15–24.
- Clinchy, M., Zanette, L., Boonstra, R., Wingfield, J. C. and Smith, J. N. M. (2004). Balancing food and predator pressure induces chronic stress in songbirds. *Proceedings of the Royal Society London Series B* 271, 2473–2479.
- Landfield, P. W. and Eldridge, J. C. (1994). Evolving aspects of the glucocorticoid hypothesis of brain aging: hormonal modulation of neuronal calcium homeostasis. *Neurobiology of Aging* 15, 579–588.
- McEwen, B. S. (1992). Re-examination of the glucocorticoid cascade hypothesis of stress and aging. In: Swaab, D., Hoffman, M., Mirmiran, R., et al. (eds.) *Progress in Brain Research*, pp. 365–383. Amsterdam: Elsevier.
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* 338, 171–179.
- McEwen, B. S. and Sapolsky, R. M. (1995). Stress and cognitive function. *Current Opinion in Neurobiology* 5, 205–216.
- McEwen, B. S. and Stellar, E. (1993). Stress and the individual: mechanisms leading to disease. *Archives of Internal Medicine* 153, 2093–2101.
- McEwen, B. S. and Wingfield, J. C. (2003). The concept of allostasis in biology and biomedicine. *Hormones and Behavior* 43, 2–15.
- McEwen, B. S., Biron, C. A., Brunson, K. W., et al. (1997). Neural-endocrine-immune interactions: the role of adrenocorticoids as modulators of immune function in health and disease. *Brain Research Reviews* 23, 79–133.
- McEwen, B. S., DeLeon, M. J., Lupien, S. J. and Meaney, M. J. (1999). Corticosteroids, the aging brain and cognition. *Trends in Endocrinology and Metabolism* 10, 92–96.
- Meaney, M. J., Tannenbaum, B., Francis, D., et al. (1994). Early environmental programming hypothalamic-pituitary-adrenal responses to stress. *Seminars in the Neurosciences* 6, 247–259.
- Sapolsky, R. M. (1996). Why stress is bad for your brain. *Science* 273, 749–750.
- Sapolsky, R., Krey, L. and McEwen, B. S. (1986). The neuroendocrinology of stress and aging: the glucocorticoid cascade hypothesis. *Endocrine Reviews* 7, 284–301.
- Schulkin, J., McEwen, B. S. and Gold, P. W. (1994). Allostasis, amygdala, and anticipatory angst. *Neuroscience and Biobehavioral Reviews* 18, 385–396.
- Seeman, T. E. and Robbins, R. J. (1994). Aging and hypothalamic-pituitary-adrenal response to challenge in humans. *Endocrine Review* 15, 233–260.

Sterling, P. and Eyer, J. (1988). Allostasis: a new paradigm to explain arousal pathology. In: Fisher, S. & Reason, J. (eds.) *Handbook of Life Stress, Cognition and Health*, pp. 629–649. New York: John Wiley and Sons.

Wingfield, J. C. and Romero, L. M. (2000). Adrenocortical responses to stress and their modulation in free-living vertebrates. In: McEwen, B. S. (ed.) *Handbook of*

Physiology. Section 7: The Endocrine System, Vol. 4: Coping with the Environment: Neural and Endocrine Mechanisms, pp. 211–236. Oxford, UK: Oxford University Press.

Wingfield, J. C. and Sapolsky, R. M. (2003). Reproduction and resistance to stress: when and how. *Journal of Neuroendocrinology* 15, 711–724.

Altitude, High, Stress Effects See: Pressures, Effects of Extreme High and Low.

Alzheimer's Disease

A E Roth

Nathan S. Kline Institute for Psychiatric Research, Orangeburg, NY, USA

W M Greenberg and N Pomara

Nathan S. Kline Institute for Psychiatric Research, Orangeburg, NY, and New York University School of Medicine, New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

The Clinical Picture of Alzheimer's Disease
 Risk Factors for Alzheimer's Disease
 The Pathophysiology of Alzheimer's Disease
 The Hypothalamic-Pituitary-Adrenal Axis in Normal Brain Functioning
 Stress and the Hypothalamic-Pituitary-Adrenal Axis
 Stress and Alzheimer's Disease
 A Possible Treatment Approach
 Conclusion

Glossary

Amyloid precursor protein (APP) A protein that is normally cut into smaller lengths by enzymes, resulting in the peptides amyloid beta (A-β)-40 and A-β-42. A-β-42 may circulate in small clumps (oligomers) and is believed to be neurotoxic and to be involved in the causation of Alzheimer's disease. The A-β peptides may be found later in Alzheimer's disease in larger clumps deposited in the brain, known as amyloid plaques.

Corticosteroid hormones Hormones produced by the adrenal gland as part of the stress response (principally cortisol in humans and corticosterone in rodents).

Declarative memory

Memories of facts and words (semantic memory) and personal events (episodic memory) that can be consciously recalled, as opposed to nondeclarative, implicitly learned procedural skills.

Delirium

A disturbance in consciousness, manifested by difficulties in focusing, sustaining, or shifting attention, with cognitive deficits. A delirium usually develops over a brief period of time and tends to fluctuate over the course of a day. It is directly caused by an active medical condition, such as an infection with a high fever, electrolyte disturbance, or liver or kidney failure.

Dementia

A syndrome of significant memory impairment compared with previous functioning, occurring with other cognitive deficits and not existing only during a period of delirium.

Hypothalamic-pituitary-adrenal (HPA) axis

The system that orchestrates the body's response to stress using hormones, with negative feedback mechanisms to allow the organism to return to baseline activation.

The Clinical Picture of Alzheimer's Disease

Dementia is a syndrome defined as the development of significant memory deficits together with other cognitive impairments, not occurring exclusively during a delirium. Of those diagnosed with dementia, between 50–60% have dementia of the Alzheimer's type, also called Alzheimer's disease (AD), a disorder characterized by symptoms with gradual onset but a

relentlessly progressive course. Early symptoms most prominently involve difficulties remembering recent events and forming new memories, and these are often accompanied by visuospatial and language problems. As the disease progresses, individuals slowly lose the ability to perform the activities of daily living, such as managing finances and driving a car. Eventually, attention, verbal ability, problem solving, reasoning, and all forms of memory become seriously impaired. Personality changes associated with the progression of AD may include increased apathy, anger, dependency, aggressiveness, and occasionally inappropriate sexual behavior; paranoid thinking is also not uncommon. In the latter stages of this disorder, individuals may be mute, completely confused, and bedridden. At this time, available treatments usually offer, at most, a temporary slowing of the symptomatic deterioration.

Early-onset AD is defined as a case in which the onset of dementia occurs at or prior to age 65; when symptoms first appear later in life it is referred to as late-onset AD. Somewhat complicating diagnostic specificity, AD may often co-exist with dementia or cognitive impairment due to another cause, most frequently vascular dementia (i.e., related to small or large strokes). Although making a probable diagnosis of AD is often not very difficult, it is a diagnosis that is only confirmed on autopsy. The prevalence of AD increases with age, from less than 1% of those age 65 to conservatively 13% of those age 85 and 23% of those age 90. As more individuals live to older ages, putting them at higher risk for this serious disorder, the economic burden of providing nursing home and other medical care to afflicted individuals, coupled with the tragic toll it takes on lives of the sufferers, their families, and other caretakers, constitutes a problem of ever-increasing magnitude.

Risk Factors for Alzheimer's Disease

Early-onset AD is infrequent, making up approximately 10% of AD cases seen in clinical practice. Most cases of early-onset AD are familial; they have been linked to specific identifiable mutations and have an autosomal-dominant pattern of inheritance with nearly full penetrance. Mutations in the Alzheimer precursor protein (APP) gene on chromosome 21, the Presenilin-1 gene on chromosome 14, and the Presenilin-2 gene on chromosome 1 have been identified, and they account for approximately 50% of all cases of early-onset familial AD. Also, most individuals who have trisomy-21 (Down syndrome), and therefore three APP genes instead of two, develop AD at relatively early ages, supporting the theory that the overexpression of the APP protein may be involved in causing AD.

In contrast, only a small number of late-onset AD cases are familial and no deterministic genetic mutations have yet been identified for late-onset sporadic AD, which accounts for more than 90% of AD cases. Clearly, genetic factors are not the sole cause of late-onset sporadic AD because the concordance rate for AD in monozygotic twins ranges from 31 to 59%. Instead, it appears that a number of genetic and environmental factors may increase the risk for this disorder. To date, the only genetic factor that has been associated with increased risk for late-onset AD is a form of the apolipoprotein E gene known as the APOE $\epsilon 4$ allele, which has a prevalence of approximately 25% in the general Caucasian population. Having one $\epsilon 4$ allele increases the risk of developing late-onset AD, and having two alleles (homozygosity) further increases the risk. In nondemented individuals, the APOE $\epsilon 4$ allele has been associated with poorer performance on delayed recall tasks and greater hippocampal atrophy compared with controls. The environmental risk factors for late-onset AD are diverse and incompletely characterized. Significant head trauma (i.e., with a loss of consciousness for an extended period) is a predisposing risk factor; higher education is an example of a protective factor. Women appear to be at a slightly higher risk of developing AD.

The Pathophysiology of Alzheimer's Disease

The characteristic features associated with AD include brain atrophy, amyloid plaques, and neurofibrillary tangles, and these abnormalities generally affect the medial temporal lobes the earliest and most intensely. Currently, there is compelling evidence that all types of AD, irrespective of age of onset, result from the neurotoxic effects of increased brain levels of soluble oligomeric forms of A- β peptides, derived from the amyloid precursor protein (APP), and from its aggregated forms (amyloid) as present in senile plaques and in the walls of small cerebral vessels. In cases of early-onset familial AD caused by genetic mutations, the increase in A- β is secondary to the increased production of the peptides A- β -40 and, especially, A- β -42 through the increased activity of β and γ secretases, two cleavage enzymes for APP. In contrast, in sporadic late-onset AD, it has been generally thought that decreased clearance of A- β is the proximate cause. In support of this hypothesis, brain regions associated with increased amyloid deposits in sporadic AD show lower levels of VPS26 and VPS35, two retromer proteins involved in the intracellular compartmental trafficking of the A- β peptides. Interestingly, there is recent evidence that the APOE $\epsilon 4$ allele may also increase A- β peptide

production, raising the possibility that increased A- β production contributes to the development of sporadic AD in elderly individuals with this allele. There is evidence that increased A- β peptide levels facilitate the abnormal phosphorylation of the τ protein, which in turn leads to its forming intracellular neurofibrillary tangles. This supports the general belief that the neurofibrillary tangles are a secondary phenomenon; however, memory deficits correlate more with the presence of neurofibrillary tangles than with plaques. The earliest pathological alterations of AD are usually seen in the entorhinal cortex and nearby CA1 subfield of the hippocampus, both structures in the medial temporal lobe critical for the formation of new memories, thereby explaining the prominent memory symptoms that are typically present as an initial symptom of AD.

AD is eventually accompanied by the pronounced degeneration of cholinergic neurons in the nucleus basalis with associated reduced levels of acetylcholine, and this loss seems to contribute significantly to memory dysfunction. The degeneration of dopaminergic neurons in the noradrenergic nucleus locus ceruleus, the serotonergic dorsal raphe nucleus, and the dopaminergic substantia nigra and the significant reductions in somatostatin-like immunoreactivity and corticotropin releasing hormone (CRH) immunoreactivity in the neocortical brain regions have also been demonstrated in AD and might contribute to the noncognitive symptoms associated with this disease.

Paradoxically, there is evidence, albeit indirect, that AD may be accompanied by overactivity of neuronal synapses using the neurotransmitter glutamate, possibly secondary to the loss of the high-affinity glutamate transporters that has been reported in AD. The overactivation of glutamate *N*-methyl-D-aspartate (NMDA) receptors produces excitotoxic effects, especially in areas such as the hippocampus and entorhinal cortex, in which the density of these receptors is high. These brain regions have been implicated in learning and memory and show profound abnormalities, including atrophy even in the earliest stages of AD, supporting a possible role for disturbances in glutamatergic neurotransmission in the cognitive and degenerative alterations associated with this disorder. There is also some empirical support for oxidative stress and inflammation having etiological roles in the pathogenesis of AD.

The Hypothalamic-Pituitary-Adrenal Axis in Normal Brain Functioning

The HPA manages the body's response to stress and is necessary for normal survival. When this system is activated, the paraventricular nucleus in the

hypothalamus releases CRH; this both activates the sympathetic nervous system and stimulates cells in the anterior pituitary to secrete adrenocorticotrophic hormone (ACTH), which stimulates cells in the adrenal cortex to synthesize and release corticosteroid hormones into the bloodstream. This complex orchestrated response acutely inhibits anabolic activity (growth-promoting functions), increases the opposite catabolic activities, and initiates the fight-or-flight reaction. This fight-or-flight reaction includes increases in heart rate and blood pressure, inhibition of the immune system, and experiential and behavioral activation. The corticosteroid secretion by the adrenal glands (principally cortisol in humans and corticosterone in rodents) provides negative feedback at the levels of both the hypothalamus and pituitary. Another important influence keeping this system in homeostatic balance is modulation provided by the hippocampus in the medial temporal lobe, a structure also critically responsible for acquiring and consolidating new declarative memories (semantic memories of facts and words and episodic memories of personal events, as opposed to nondeclarative, implicitly learned procedural skills).

Two types of corticosteroid receptors exist on the surface of neurons: type I or mineralocorticoid receptors (MRs), which have a higher affinity for corticosteroids, and type II or glucocorticoid receptors (GRs), which have a lower affinity. Under normal, non-stressful conditions, with low levels of circulating corticosteroids, mostly the higher-affinity MRs are stimulated, with GRs remaining mainly unaffected.

Stress and the Hypothalamic-Pituitary-Adrenal Axis

In the hippocampus, when the MRs are predominantly activated, the formation of new declarative memories is supported by the mechanism of long-term potentiation. Under low levels of acute stress, there may even be some improvement in this function. However, when stress levels and the associated increases in circulating corticosteroid levels cause more activation of the GRs, there is interference with forming new memories. This pattern of low levels of stress improving performance and higher levels of stress disrupting performance has sometimes been descriptively referred to as an inverted-U relationship. Slight elevations in stress hormones are also associated with enhanced attention and motivation, both important for performing well on memory and higher-order cognitive tasks but impairing the recall of previous memories. These latter actions appear to be mediated by both β -adrenergic and cholinergic receptors in the basolateral complex of the amygdala.

Stress and the Hypothalamic-Pituitary-Adrenal Axis in Animal Studies

Prolonged HPA overactivity is also associated with cognitive decline and hippocampal atrophy in rats. Rodents exposed to stress paradigms or administered large doses of corticosterone have been reported to have atrophy of the hippocampal neurons and increased sensitivity to toxins, but this could not be replicated in pigtailed macaque monkeys exposed to high doses of cortisone for 12 months. Species differences may be important because these primates proportionately have far fewer hippocampal GRs than rodents.

Stress and Hypothalamic-Pituitary-Adrenal Axis in Normal Aging

The hippocampus plays a major role in learning and memory performance, in particular in declarative (conscious and voluntary) memory. It is susceptible to damage from intense prolonged stress. Cortisol production increases during stress, and long-term exposure to high levels of cortisol causes atrophy and impaired neurogenesis in the hippocampus. There is also evidence of hippocampal atrophy as a function of aging, and HPA overactivity is associated with cognitive decline and hippocampal atrophy in normal elderly.

There are a variety of measures of HPA activity: simple plasma cortisol and salivary cortisol measures have been used, but there is much variation in these over the course of a day. Urinary cortisol may be a more reliable measure of cortisol secretion over time, and some studies have determined that cerebrospinal fluid (CSF) cortisol more accurately reflects cortisol levels in the central nervous system. Another widely used measure of HPA activity is the dexamethasone (DST) test (DST is a synthetic glucocorticoid). DST is typically administered to individuals in the evening, and plasma cortisol levels are measured the following day in the morning or afternoon. Ordinarily, DST suppresses the HPA axis, and plasma cortisol levels will be low the day after DST administration. Failure to suppress the HPA axis (nonsuppression) is evidence of a loss of normal feedback control and homeostasis.

There is some evidence suggesting gender differences in the role of cortisol overproduction in cognitive and memory dysfunction in normal elderly individuals. One study has found that increases in urinary cortisol excretion are significantly associated with declines in memory performance over 2.5 years in healthy elderly women but not in men. Furthermore, women who showed a decrease in urinary cortisol excretion showed significant improvement in memory performance over time. These findings

highlight the possibility that damage to the structure and functioning of the brain from hypercortisolemia might be reversible, at least for healthy elderly women, although it is believed that prolonged and repeated exposure to glucocorticoids increases the risk of irreversible damage.

Indeed, longitudinal examinations of cortisol levels and declarative memory performance reveal that elderly individuals with high and increasing cortisol levels over periods ranging from 1 to 3 years show significantly greater impairments in declarative memory performance and significantly greater hippocampal atrophy than do individuals who show decreases or low and smaller increases in cortisol levels over time. The amount of hippocampal atrophy in elderly individuals with high and increasing levels of cortisol was found to be equivalent to that of older adults exhibiting mild cognitive impairment (MCI). Although cortisol levels, on average, increase with aging, there is much interindividual variability in cortisol production in the elderly, with some individuals actually showing decreases in cortisol levels over time.

Stress and Alzheimer's Disease

Because older age, hypercortisolemia, and hippocampal atrophy are all associated with AD, the role of stress in AD pathology and memory dysfunction has logically attracted some attention. According to the glucocorticoid cascade hypothesis, hippocampal degeneration in AD might result in increased cortisol secretion, due to the loss of normal negative hippocampal feedback to the HPA axis. In turn, the overactive HPA system evident in AD could lead to additional neuronal degeneration, thereby inducing progressive cognitive impairment – potentially a vicious circle.

Memory deficits have also been found in the presence of other disorders that are associated with persistent HPA activation and high cortisol levels, such as Cushing's disease and major depressive disorder. One investigator found that, compared to matched controls, women with histories of recurrent major depression had smaller hippocampal volumes, the size of which were correlated with their lifetime duration of depression, and did more poorly on tests of verbal memory.

Most studies of individuals with AD have found them to have elevated post-DST cortisol levels compared to healthy controls. Increased HPA activity, as evidenced by elevated basal cortisol levels, elevated glucocorticoid activity, increased CSF cortisol levels, and hippocampal atrophy, have also been usually found to be positively associated with the degree of cognitive impairment in elderly individuals with AD.

A relevant variable also appears to be the presence of agitation, found to be more frequent in DST non-suppressors. In one small study that followed AD patients for up to 3 years, increased basal plasma cortisol levels were found to linearly correlate with decline in cognitive functioning.

A recent study demonstrated a relationship between APOE genotype and CSF cortisol levels in individuals with AD. Specifically, in analyses controlling for age, it was found that individuals with AD who had the APOE $\epsilon 4$ genotype had significantly higher cortisol levels than did individuals without the $\epsilon 4$ allele. Similar analyses performed with nondemented elderly individuals revealed only a trend for individuals with the $\epsilon 4$ allele to have higher CSF cortisol levels. These findings suggest that the APOE $\epsilon 4$ genotype might contribute to HPA axis dysfunction. However, CSF cortisol levels were not significantly correlated with measures of global cognitive status, although it is possible that CSF cortisol levels might be associated with more subtle changes in cognitive functioning.

Psychological Stress and the Hypothalamic-Pituitary-Adrenal Axis in Alzheimer's Disease

The evidence of a direct role of stress in AD onset remains limited and inconclusive. As previously noted, there are many studies reporting a significant association between acute and chronic stress with memory impairment, yet few studies have considered the role of elevated stress levels in the development of AD. In a longitudinal analysis of subjects' proneness to psychological distress and AD onset, investigators found that in analyses that accounted for depressive symptoms, for each 1-point increase on the distress-proneness scale there was a corresponding 6% greater risk of AD onset. Proneness to psychological distress was measured with a well-established index of neuroticism. Furthermore, greater proneness to psychological distress was associated with poorer levels of, and a greater decline in, global cognitive functioning. Additional analyses revealed an association between greater psychological distress and greater decline in episodic memory functioning. Specifically, decline was 16% greater for each 1-point increase on the distress proneness scale in episodic memory functioning. However, proneness to psychological distress was not associated with pathological AD as examined postmortem.

In a similar analysis examining distress-proneness and AD onset in a community sample of Caucasian and African American (49.8%) elderly individuals, it was found that individuals who scored in the ninetieth percentile or higher on a neuroticism index were

2.4 times more likely to develop AD over the course of 3–6 years. Additional analyses accounting for depressive symptoms or excluding individuals with depression at baseline did not alter the results. Although the researchers did not have postmortem data for AD pathology, the APOE genotype (a potential risk factor for AD) was available; and the researchers found no significant association between APOE genotype and distress-proneness in the full sample. Interestingly, the magnitude of the association between distress-proneness and AD onset was significantly different in Caucasian versus African American individuals. In the Caucasian sample, with each 1-point increase on the distress-proneness scale the odds of developing AD increased 12%. In contrast, the odds of developing AD in the African American sample increased only 2% for every 1-point increase on the distress-prone scale. These findings highlight the possibility of sociocultural differences in the association between stress and AD onset.

Other studies examining psychosocial stressors and AD onset have reported similar findings. For example, in a longitudinal analysis of psychosocial stressors and AD onset, when accounting for demographic factors, including education, individuals who reported psychosocial stressors as far back as childhood were found to have significantly higher rates of AD onset relative to individuals with fewer or no psychosocial stressors. Specifically, the incidence rate of AD over 9 years in individuals reporting no psychosocial stressors was 1.7%, and this rose to 5% in individuals reporting one or two and to 8.3% for individuals reporting three or more psychosocial stressors. The psychosocial stressors or risk factors that were considered included events from childhood (i.e., death of a parent and poverty), adulthood (i.e., death of a spouse, arduous manual labor, and serious illness in a child), and old age (i.e., deterioration in financial status and physical illness). Psychosocial stressors that were significantly associated with AD onset included the death of a parent during childhood, arduous manual labor, physical illness in a spouse during old age, and serious illness in a child (offspring). Possibly relevant to this finding is an earlier study reporting that, when accounting for education and alcohol consumption, men who had occupations involving manual labor were 5.3 times more likely to be diagnosed with late-onset AD than were men in nonarduous occupations. The finding that psychosocial stressors earlier in the life span might contribute to dysfunction in later life is also congruent with research using animal models in which earlier experiences altered noradrenergic and serotonergic functions in adulthood and the rate of aging.

A Possible Treatment Approach

An interesting avenue in exploring the role of stress hormones in AD is the potential therapeutic utility of corticosteroid receptor antagonists such as mifepristone (RU-486). In a pilot study in which participants were randomized to placebo or treatment with RU-486 for 6 weeks, participants taking RU-486 improved on measures of memory, whereas participants in the placebo condition declined. Other studies have also revealed that short-term use of RU-486 improves memory performance in bipolar patients but not for individuals with schizophrenia. Indeed, using a corticosteroid receptor antagonist to try to delay the onset or slow the progression of cognitive impairment and AD deserves further consideration.

Conclusion

Studies examining HPA dysfunction, cortisol levels, and indices of psychological distress have found significant associations with memory impairment, memory decline, AD, or AD onset. Although it appears that psychological distress, independent of depression, may play a role in AD onset, research on the biological mechanisms linking stress to AD onset remains inconclusive. HPA overactivity and high or increasing levels of cortisol seem to be associated with cognitive dysfunction and brain changes associated with AD, but it is unclear whether these impairments precede AD onset or are a function of the illness. Longitudinal research examining the biological mechanisms of stress and AD onset would be fruitful for further understanding the link between psychological stress and AD.

See Also the Following Articles

Aging and Psychological Stress; Aging and Stress, Biology of; Glucocorticoids – Adverse Effects on the Nervous System; Aging and Adrenocortical Factors.

Further Reading

Davis, K. L., Davis, B. M., Greenwald, B. S., et al. (1986). Cortisol and Alzheimer's disease. I: Basal studies. *American Journal of Psychiatry* 143, 442–446.

- Fratiglioni, L., Ahlbom, A., Viitanen, M., et al. (1993). Risk factors for late-onset Alzheimer's disease: a population-based, case-control study. *Annals of Neurology* 33, 258–266.
- Li, S., Mallory, M., Alford, M., et al. (1997). Glutamate transporter alterations in Alzheimer's disease are possibly associated with abnormal APP expression. *Journal of Neuropathology and Experimental Neurology* 56(8), 901–911.
- Lupien, S. J., DeLeon, M., DeSanti, S., et al. (1998). Longitudinal increase in cortisol during human aging predicts hippocampal atrophy and memory deficits. *Nature Neuroscience* 1, 69–73.
- Lupien, S. J., Nair, N. P. V., Brière, S., et al. (1999). Increased cortisol levels and impaired cognitive cognition in human aging: implication for depression and dementia in later life. *Reviews in Neuroscience* 10, 117–139.
- Nemeroff, C. B., Kizer, J. S., Reynolds, G. P., et al. (1989). Neuropeptides in Alzheimer's disease: a post-mortem study. *Regulatory Peptides* 25(1), 123–130.
- Peskind, E. R., Wilkinson, C. W., Petrie, E. C., et al. (2001). Increased CSF cortisol in AD is a function of APOE genotype. *Neurology* 56, 1094–1098.
- Pomara, N., Singh, R., Deptula, D., et al. (1992). Glutamate and other CSF amino acids in Alzheimer's disease. *American Journal of Psychiatry* 149(2), 251–254.
- Pomara, N., Doraiswamy, P. M., Tun, H., et al. (2002). Mifepristone (RU 486) for Alzheimer's disease. *Neurology* 58(1), 1436–1437.
- Pomara, N., Greenberg, W. M., Branford, M. D., et al. (2003). Therapeutic implications of HPA axis abnormalities in Alzheimer's disease: review and update. *Psychopharmacology Bulletin* 37(2), 120–134.
- Sapolsky, R. M., Krey, L. and McEwen, B. S. (1986). The neuroendocrinology of stress and aging: the glucocorticoid cascade hypothesis. *Endocrine Reviews* 7, 284–301.
- Seeman, T. E., McEwen, B. S., Singer, B. H., et al. (1997). Increase in urinary cortisol excretion and memory declines: MacArthur studies of successful aging. *Journal of Clinical Endocrinology and Metabolism* 82(8), 2458–2465.
- Sheline, Y. I., Sanghyavi, M., Mintun, M. A., et al. (1999). Depression duration but not age predicts hippocampal volume loss in medically healthy women with recurrent major depression. *Journal of Neuroscience* 19, 5034–5043.
- Small, S. A., Kent, K., Pierce, A., et al. (2005). Model-guided microarray implicates the retromer complex in Alzheimer's disease. *Annals of Neurology* 58, 909–919.
- Wilson, R. S., Barnes, L. L., Bennett, D. A., et al. (2005). Proneness to psychological distress and risk of Alzheimer's disease in a biracial community. *Neurology* 64, 380–382.

Ambulatory Blood Pressure Monitoring

T G Pickering

Columbian Presbyterian Medical Center, New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 164–167, © 2000, Elsevier Inc.

Technique

Advantages of Ambulatory Blood Pressure Monitoring

White Coat Hypertension

The Influence of Personality and Mood

Effects of Acute Stress

Effects of Chronic Stress

Glossary

Ambulatory blood pressure monitoring (ABPM)

A technique for measuring blood pressure over prolonged periods of time (typically 24 h) in free-living subjects. Most devices measure the blood pressure from a cuff placed on the upper arm inflated at regular intervals (usually every 15–20 min) by the monitor, which is a battery-operated device a little bigger than a portable radio.

Dippers and nondippers

Blood pressure normally decreases by at least 10% during the night in comparison with the average daytime level. The difference between the average daytime and nighttime blood pressure has been used to classify people as dippers (>10% difference) or nondippers (<10% difference).

White coat effect

The increase of blood pressure associated with the physician's office setting, usually expressed as the difference between the clinic and average daytime pressures.

White coat hypertension

Hypertension defined by a persistently elevated clinic or office blood pressure together with a normal blood pressure at other times (usually measured with ambulatory blood pressure monitoring).

Technique

Ambulatory blood pressure monitoring (ABPM) is performed using small fully automatic battery-operated monitors that weigh less than one pound and can be worn on a belt while subjects go about their normal daily activities. The monitor typically

contains a pump that inflates the blood pressure cuff at preset intervals, a pressure detection system (usually operating by the oscillometric method), and a memory chip for storing the readings. The cuff is usually placed on the nondominant arm above the elbow and is connected to the monitor by a thin tube. Subjects are asked to keep their arm still and by their side while a reading is taken. The monitors do not work reliably during vigorous activity (e.g., exercise) or in environments where there is a lot of noise or vibration. Subjects may be asked to keep a diary describing their activity, mood, and location at the time of each reading. Blood pressure during sleep can also be recorded. At the end of 24 h the subject returns and the readings are downloaded into a computer for analysis. Several types of monitor are available, but not all have been validated for accuracy according to standard protocols.

Advantages of Ambulatory Blood Pressure Monitoring

ABPM has revolutionized the study of the relationships between stress and blood pressure. Previously, reliance was made on blood pressure measured either in an office or in a laboratory setting. In the former case, a very limited number of readings is usually made, and the blood pressure that is recorded may represent the white coat effect rather than the effects of everyday stress. And while reactivity testing in the laboratory permits multiple measurements to be made during exposure to standardized stressors, the situation is artificial, and generalizability to blood pressure measured in real life is very poor. ABPM can give three different types of information relating to blood pressure: the average level, short-term changes, and the diurnal rhythm. Normally the blood pressure falls by about 10% during the night, but in about 25% of people the changes are smaller than this – the so-called nondipping pattern. There is now substantial evidence that the average level of blood pressure as measured by ABPM gives a better predictor of cardiovascular risk than conventional clinic measurements. In addition, nondippers have been reported to be at higher risk than dippers. The pathological significance of short-term blood pressure variability is unclear, although it is hypothesized that increased variability may be an independent risk factor.

White Coat Hypertension

One of the important phenomena identified by ABPM is white coat hypertension, which occurs in about 20% of patients with mild hypertension. Such patients usually have less target organ damage than patients with the same level of clinic pressure but whose pressure remains high outside the clinic, and they also appear to be at lower risk of cardiovascular morbidity. The origins of white coat hypertension are unexplained, but it is usually attributed to the stress associated with an encounter with a physician. It has been suggested that the phenomenon may be a conditioned response.

The Influence of Personality and Mood

ABPM can be used to measure blood pressure during normal daily activities, including work. By having subjects keep a diary of their activities and emotions it is possible to correlate behavioral factors and blood pressure. Thus, while anger is the mood that is associated with the greatest increase of blood pressure, blood pressure also increases when subjects report that they are happy as compared to having a neutral mood. There are also reports of associations between negative moods and ambulatory blood pressure. The analysis of such interactions requires that there is statistical control for other influences such as posture, location, and activity. A number of reports have found that subjects who score high on hostility report more negative moods and have higher blood pressures in association with them than subjects who are less hostile. Hostile subjects also have higher nocturnal blood pressure and show a more rapid increase on waking.

Effects of Acute Stress

Blood pressure is typically highest during the hours of work, which may be the result of both physical and mental activity being greater than at home. In men, blood pressure usually falls when they get home in the evening, but in women it may not do so, particularly if they have small children. The effects of commonly occurring activities on blood pressure are shown in [Table 1](#). A study of men working in machine shops found a significant correlation between the frequency of somatic complaints such as palpitations, fatigue, and dizziness and ambulatory systolic pressure. Paramedics show 10 mmHg higher blood pressures when they are in an ambulance and 7 mmHg elevations at the scene of an accident than during nonworkday activities.

Table 1 Average changes of blood pressure associated with 15 commonly occurring activities

<i>Activity</i>	<i>Systolic (mmHg)</i>	<i>Diastolic (mmHg)</i>
Meetings	+20.2	+15.0
Work	+16.0	+13.0
Transportation	+14.0	+9.2
Walking	+12.0	+5.5
Dressing	+11.5	+9.7
Chores	+10.7	+6.7
Telephone	+9.5	+7.2
Eating	+8.8	+9.6
Talking	+6.7	+6.7
Desk work	+5.9	+5.3
Reading	+1.9	+2.2
Business (at home)	+1.6	+3.2
Television	+0.3	+1.1
Relaxing	0	0
Sleeping	−10.0	7.6

Effects of Chronic Stress

Occupational stress has been investigated in several studies. In a study of 50 normotensive working women doing technical and clerical jobs, who also kept diaries describing their activities, emotions, and perceptions of stress, it was found that the average pressures were 116/78 mmHg at work, 113/74 mmHg at home, and 102/63 mmHg during sleep. The most powerful behavioral predictor of systolic pressure was the perception that one's job is stressful, which was associated with higher pressures in all three situations – at work, at home, and during sleep. The perception that there was more stress at work than at home on the day of the study was associated with higher systolic pressures at work, but not at home or during sleep. Potential sources of domestic stress were also related to blood pressure. Being married was associated with higher work diastolic pressures, while having children was associated with higher systolic and diastolic pressures both at work and at home. Women who reported higher levels of stress at home than at work also had relatively higher diastolic pressures at home than at work. Thus, single women whose main source of stress was at work showed a pattern similar to the one seen in men, while married women with children, whose level of stress may be higher at home, did not show the typical decrease of blood pressure when they returned home.

Young physicians, when on call for the emergency room, show an elevation of ambulatory blood pressure in comparison to a weekend day, which exceeded 10 mmHg for systolic pressure in 40% of cases. Blood pressure while asleep was not affected, however.

In this case the acute work stress raised the blood pressure during the exposure to stress, but to what extent this was due to mental as to opposed to physical activity is not clear. A very different picture is obtained with more chronic stress: subjects whose jobs are characterized as high strain (defined by a combination of high demands and low control) have elevated ambulatory blood pressures in comparison with subjects in less stressful jobs, not only during the hours of work, but also, more surprisingly, while they are at home and asleep. That this is a causal relationship is suggested by the finding that over a 3 year period those who remain in high-strain jobs show a persistent elevation of blood pressure, and those whose jobs become less stressful show a decrease of pressure. These results provide the strongest evidence to date that exposure to chronic stress can elevate the blood pressure throughout the 24 h period. Interestingly, there appears to be no relation between blood pressure and job strain in women, despite the fact that women are more likely than men to be in high-strain jobs. This might be because women react to different types of stress than men: another study found that lack of social support predicted ambulatory blood pressure in women, while hostility was a better predictor in men.

See Also the Following Articles

Allostasis and Allostatic Load; Hypertension.

Further Reading

- Goldstein, I. B., Jamner, L. D. and Shapiro, D. (1992). Ambulatory blood pressure and heart rate in healthy male paramedics during a workday and a nonworkday. *Health Psychology* **11**, 48–54.
- James, G. D., Moucha, O. P. and Pickering, T. G. (1991). The normal hourly variation of blood pressure in women: Average patterns and the effect of work stress. *Journal of Human Hypertension* **5**, 505–509.
- Linden, W., Chambers, L., Maurice, J. and Lenz, J. W. (1993). Sex differences in social support, self-deception, hostility, and ambulatory cardiovascular activity. *Health Psychology* **12**, 376–380.
- Pickering, T. G. (1994). Psychosocial stress and hypertension B: clinical and experimental evidence. In: Swales, J. D. (ed.) *Textbook of hypertension*, pp. 640–654. London: Blackwell Scientific.
- Pickering, T. G., James, G. D., Boddie, C., et al. (1988). How common is white coat hypertension? *Journal of the American Medical Association* **259**(2), 225–228.
- Pickering, T. G., Devereux, R. B., James, G. D., et al. (1996). Environmental influences on blood pressure and the role of job strain. *Journal of Hypertension* **14**(Supplement 5), S179–S186.
- Schnall, P. L., Schwartz, J. E., Lansbergis, P. A., Warren, K. and Pickering, T. G. (1999). A longitudinal study of job strain and ambulatory blood pressure: results from a three-year follow-up. *Psychosomatic Medicine* **60**, 697–706.
- Verdecchia, P., Porcellati, C., Schillaci, G., et al. (1994). Ambulatory blood pressure: an independent predictor of prognosis in essential hypertension. *Hypertension* **24**, 793–801.

Amenorrhea

V E Beshay and B R Carr

University of Texas Southwestern Medical Center at Dallas, Dallas, TX, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by B R Carr, volume 1, pp 168–175, © 2000, Elsevier Inc.

Introduction

The Menstrual Cycle

Basic Embryology

Classification of Amenorrhea

Clinical Evaluation of Amenorrhea

Management of Amenorrhea

Glossary

Androgen insensitivity syndrome

A condition characterized by resistance to androgens, which subsequently leads males to develop into normal-appearing females. These patients are reared as females; they lack a uterus, fallopian tubes, and a vagina. Formerly known as testicular feminization.

Androgens
Aromatase

Male sex hormones.

Enzyme responsible for the conversion of androgens into estrogens.

Endometrial hyperplasia
Endometriosis

A condition characterized by the abnormal thickening of the endometrium.

A condition characterized by the presence of endometrial tissue outside the uterus, for example, on the bowel surfaces or ovary.

<i>Endometrium</i>	The membrane lining of the uterine cavity.
<i>Eugonadotrophism</i>	A condition characterized by normal levels of gonadotropins.
<i>Galactorrhea</i>	A condition characterized by milk discharge from the nipple.
<i>Hypergonadotropic hypogonadism</i>	A condition characterized by elevated gonadotropins due to ovarian failure and diminished ovarian hormone secretion.
<i>Hypogonadotropic hypogonadism</i>	A condition characterized by low gonadotropins due to alterations of the central nervous system and hypothalamic-pituitary axis, leading to diminished ovarian hormone secretion.
<i>Hysterosalpingogram</i>	An x-ray film of the uterus and fallopian tubes obtained after the instillation of dye in the uterine cavity. It is used to evaluate the contour of the uterine cavity and patency of the fallopian tubes.
<i>Menarche</i>	The first menstrual period.
<i>Müllerian agenesis</i>	A condition characterized by the lack of development of the uterus, fallopian tubes, cervix, and upper vagina.
<i>Polycystic ovary syndrome</i>	A disease characterized by menstrual disturbances, obesity, excessive androgen production, and polycystic ovaries.
<i>Primary amenorrhea</i>	A condition characterized by the absence of menstruation by age 16 in a phenotypic female.
<i>Secondary amenorrhea</i>	A condition characterized by the cessation of menstruation for 3–6 months in a female who has previously menstruated.
<i>Sexual ambiguity</i>	A condition characterized by difficulty in assessing sex based on physical features because the individual was born with genitalia and/or secondary sex characteristics that are determined to be neither exclusively male nor female or that combine features of the male and female sexes.
<i>Virilization</i>	A condition characterized by the abnormal development of masculine features in females such as hair growth, baldness, and deepening of the voice.

Introduction

Amenorrhea is the cessation of the menstrual flow. It can be a physiological finding, such as in pregnant or menopausal patients, or a pathological finding with a possible underlying medical condition. Because amenorrhea may be a sign of other medical illnesses, it is important to evaluate patients presenting with this problem. Amenorrhea may cause stress to patients because they may perceive a loss of their femininity due to the lack of cyclical menstruation, or it may be secondary to stressors the patient may be experiencing.

The Menstrual Cycle

In females, usually the first sign of pubertal development is breast development. This is followed by pubic hair growth, growth spurt, and, finally, menarche (the first menstrual flow). Menarche is usually reached between the ages of 8 and 13 years.

The menstrual cycle is a very complex, hormone-driven process. The hypothalamus first produces gonadotropin-releasing hormone (GnRH), which acts on the pituitary gland. In response to GnRH, the pituitary gland secretes luteinizing hormone (LH) and follicle-stimulating hormone (FSH). LH and FSH act on the different cells of the ovary to cause growth of the ovarian follicles and production of hormones, namely estrogen and progesterone.

The menstrual cycle is divided into two phases: the follicular phase, during which there is recruitment and growth of follicles, and the luteal phase, which occurs after the release of the mature follicle. Estrogen dominates the follicular phase, whereas progesterone dominates the luteal phase. These hormones also play a role in the endometrium; estrogen leads to the growth of the endometrial lining, whereas progesterone leads to its maturation. If pregnancy does not occur, progesterone levels drop, which causes the shedding of the endometrial lining, and thus a new cycle is initiated. This is the basis of the progestin withdrawal challenge.

Basic Embryology

Male and female gonads develop in response to the chromosomal makeup of the developing fetus. In males, the Y chromosome drives the development of the testes. The testes, in turn, produce a substance that inhibits the development of female internal genitalia, namely, the uterus, fallopian tubes, and vagina. The testes also produce testosterone, which is responsible for the development of the male internal genitalia. Testosterone is also converted to another hormone, dihydrotestosterone, which leads to the development of the male external genitalia.

Unlike in males, in females there are no substances identified yet that drive genital development. In the absence of a Y chromosome, the internal and external genitalia are female. Two X chromosomes are required for normal ovarian development.

Classification of Amenorrhea

There are a number of causes of amenorrhea, and some are related to stress. The usual definition of amenorrhea is the cessation of menstrual flow.

Women who fail to menstruate by age 16 are often considered to have primary amenorrhea, whereas those who have previously been menstruating and now exhibit at least 3 months of amenorrhea are considered to have secondary amenorrhea. In this evaluation, women who present with delayed puberty also have amenorrhea and should not be considered as a separate entity. Women who present with sexual ambiguity since birth or significant virilization may also present with amenorrhea, but they should be evaluated as separate disorders. Using this approach, the diagnosis and treatment is somewhat simplified. It should be pointed out that classifications consisting of primary and secondary amenorrhea are not used here because they do not describe the pathophysiology; the classification chosen is presented in [Table 1](#).

The largest category is chronic anovulation, which includes women who have the ability to ovulate but do not. This category is subcategorized into women who are producing estrogen (eugonadotrophism) and those who are not producing estrogen (hypogonadotropic hypogonadism). A second category is ovarian failure (hypergonadotropic hypogonadism), in which the germ cells are usually absent. The third category comprises abnormalities and defects in the development of the female reproductive tract.

Chronic Anovulation

The major cause of amenorrhea is chronic anovulation. In this disorder, women have ovaries and follicles but for various reasons do not ovulate normally. With appropriate therapy, these women can ovulate and later conceive. In women with chronic anovulation, the ovaries do not secrete estrogen and progesterone in the normal cyclic pattern. In order to differentiate between women who produce estrogen from those who do not, a progestin challenge is administered either orally or intramuscularly. Women who exhibit bleeding after receiving an oral or intramuscular progestin challenge are producing estrogen and fall into the category of chronic anovulation with estrogen present; women who fail to exhibit a menstrual bleed after a progestin challenge have low levels of estrogen and are categorized as chronic anovulation with estrogen absent.

Chronic anovulation with estrogen present Women with chronic anovulation who experience withdrawal menstrual bleeding following a progestin challenge and exhibit normal FSH levels are said to be in a state of chronic estrous due to acyclic production of estrogen. The ovarian follicles of women with this disorder do not secrete estrogen but instead secrete androgens such as androstenedione, which is converted in

the peripheral tissues by extraglandular aromatase into the weaker estrogen estrone. This condition may be due to problems primarily in the ovary or in the hypothalamic-pituitary-gonadal feedback loops. The consequence is that these women fail to ovulate and produce estrogen and do not experience cyclic withdrawal bleeding.

The primary cause of chronic anovulation with estrogen present is polycystic ovarian syndrome (PCOS). PCOS is a complex disorder (probably inherited and perhaps related to insulin resistance) characterized by the development of hirsutism, androgen excess, obesity, and menstrual disturbances including amenorrhea, oligomenorrhea, or dysfunctional uterine bleeding at the time of expected puberty. The clinical picture varies and laboratory tests may be helpful but are only supportive in confirming the diagnosis. In the majority, but not in all, women with PCOS, plasma LH levels are elevated while the plasma FSH levels are normal or low. These findings have led some to suggest that a ratio of LH to FSH of greater than 2 to 3 may be a useful laboratory distinction; however, the evaluation of a single sample of gonadotropins may lead to the wrong diagnosis and only 80% of women may exhibit this finding. In addition, the majority of women with PCOS have moderately elevated serum androgen levels. An ultrasound demonstrating multiple superficial small follicles surrounding the surface of the ovary in a ringlike pattern associated with increased stromal density in a woman with amenorrhea appears to be the best test to confirm the diagnosis. However, as stated previously, the diagnosis of PCOS syndrome is not based on pathological changes in the ovaries or plasma hormone abnormalities but is primarily based on the clinical evidence of chronic anovulation with varying degrees of androgen excess and menstrual disturbances. New criteria put forth for the diagnosis of PCOS includes positive findings in two of the following findings: (1) clinical or biochemical evidence of hyperandrogenism, (2) irregular menstrual cycles with oligo- or anovulation, and (3) evidence of polycystic ovaries on sonography. The etiology of this disorder is multifactorial and is not yet clearly understood.

Chronic anovulation with estrogen present, obesity, hirsutism, and even polycystic ovaries can be seen in a variety of other endocrine disorders in addition to PCOS (see [Table 1](#)). These women may present with Cushing's syndrome, hyperthyroidism, hypothyroidism, and/or late adult-onset adrenal hyperplasia due to 21-hydroxylase or 11 β -hydroxylase deficiency. Rarely, patients with chronic anovulation with estrogen present can present with ovarian tumors such as a Brenner tumor or cystic teratomas and ovarian

Table 1 Classification of amenorrhea^a

Chronic anovulation with estrogen present		
PCOS		
Adrenal disease	Cushing syndrome Adult-onset adrenal hyperplasia	
Thyroid disease	Hypothyroidism Hyperthyroidism	
Ovarian tumors	Granulosa-theca cell tumors Brenner tumors Cystic teratomas Mucinous/serous cystadenomas Krukenberg tumors	
Chronic anovulation with estrogen absent (hypogonadotropic hypogonadism)		
Hypothalamic	Tumors	Craniopharyngioma Germinoma Hamartoma Hand-Schüller-Christian disease Teratoma Endodermal sinus tumors Metastatic carcinoma Tuberculosis Syphilis Encephalitis/meningitis Sarcoidosis Kallmann syndrome GnRH receptor mutations Idiopathic hypogonadotropic hypogonadism Chronic debilitating disease
	Infection and other disorders	Stress Weight loss/diet Malnutrition Psychological eating disorders (anorexia nervosa, bulimia) Exercise Pseudocyesis Prolactinomas Other hormone-secreting pituitary tumors (ACTH, TSH, GH, gonadotropin) Nonfunctional tumors (craniopharyngioma) Metastatic carcinoma Empty sella Arterial aneurysm Sheehan syndrome Panhypopituitarism Sarcoidosis Hemochromatosis Lymphocytic hypophysitis
	Functional	
Pituitary	Tumors	
	Space-occupying lesions	
	Necrosis	
	Inflammatory/infiltrative	
	Gonadotropin mutations (FSH)	
Ovarian failure (hypergonadotropic hypogonadism)		
Gonadal agenesis		Turner syndrome 45,X Mosaicism
Gonadal dysgenesis	Abnormal karyotype Normal karyotype	Pure gonadal dysgenesis (46,XX; 46,XY, Swyer syndrome)
Ovarian enzymatic deficiency	17 α -Hydroxylase deficiency 17,20-Lyase deficiency Aromatase deficiency	
Premature ovarian failure	Idiopathic – premature aging Injury	Mumps oophoritis Radiation Chemotherapy

Continued

Table 1 Continued

	Resistant ovary	Idiopathic Mutations of FSH receptor Mutations of LH receptor
	Autoimmune disease Galactosemia	
Anatomic defects (outflow tract)	Labial agglutination/fusion Imperforate hymen Transverse vaginal septum Cervical agenesis – isolated Cervical stenosis – iatrogenic Vaginal agenesis – isolated Müllerian agenesis (Mayer-Rokitansky-Kuster-Hauser syndrome) Complete androgen resistance (testicular feminization) Endometrial hypoplasia or aplasia – congenital Asherman syndrome (intrauterine synechiae)	

^aNot including disorders of congenital sexual ambiguity. ACTH, adrenocorticotrophic hormone; FSH, follicle-stimulating hormone; GH, growth hormone; GnRH, gonadotropin-releasing hormone; LH, luteinizing hormone; PCOS, polycystic ovarian syndrome; TSH, thyroid-stimulating hormone.

cancers in which the ovary and/or tumor secrete increased amounts of androstenedione, which aromatize in extraglandular sites to estrogen. As a result, such patients may present with a pattern similar to that of a woman with PCOS that resolves at the time of the removal of the tumor or the treatment of the primary disorder. It is thus important to rule out these conditions first, prior to making a diagnosis of PCOS.

Chronic anovulation with estrogen absent Women with chronic anovulation with estrogen absent due to low or absent estrogen production fail to experience withdrawal bleeding or experience only vaginal spotting following a progestin challenge. Usually the FSH level is either in the normal or low range. This is an important point because evaluating just the FSH level alone (if within the normal range, for example, 4–8 IU l⁻¹) does not confirm the cause of amenorrhea. Chronic anovulation with estrogen absent is a result of hypogonadotropic hypogonadism, which is secondary to organic or functional disorders of the central nervous system (CNS) and hypothalamic-pituitary axis. It may be clinically helpful, but not always practical, to subdivide these into hypothalamic and pituitary causes (see [Table 1](#)). Hypothalamic tumors and other destructive disorders of the hypothalamus are relatively rare causes of amenorrhea and require radiographic evaluation such as a computerized axial tomography (CAT scan) or magnetic resonance imaging (MRI). In women with tumors, headaches and other neurological symptoms or signs are often present.

The most common cause of chronic anovulation with estrogen absent is a functional disorder of the hypothalamus or CNS in the presence of a normal MRI or CAT scan. In these cases, the cause of chronic anovulation is one of exclusion. A history of stress or a stressful event can often be obtained, which can include the loss of a loved one, entering a stressful work environment, or going off to college. In these cases, the stress reduces the hypothalamic secretion of GnRH, leading to reduced gonadotropin (FSH and LH) secretion followed by reduced ovarian estrogen secretion, resulting in amenorrhea. This can happen even in women with normal body weight.

Other causes related to stress and emotional disorders include weight loss and dieting, which are particularly common in young teenage girls. The most severe form of weight loss-induced amenorrhea is in anorexia nervosa, characterized by distorted attitudes toward eating and weight, self-induced vomiting, extreme emaciation, and distorted body images. Women who excessively exercise, particularly in such activities as marathon running, ballet dancing, or, more recently, working out in the gym with weights and using steroids to reduce their percentage of body fat, may develop amenorrhea. Amenorrhea is more likely to develop in women with a previous history of abnormalities in menstruation prior to the onset of excessive exercise or weight loss. It is known that women who develop amenorrhea associated with stress, exercise, or dieting exhibit alterations in the menstrual cycle that can be associated with normal

cyclic bleeding early in the onset of the increased exercise, weight loss, or dieting. Later, as the disorder progresses, problems in the production of progesterone during the luteal phase occur, followed by anovulatory cycles associated with oligomenorrhea and, finally, the development of amenorrhea. In some of these women a withdrawal bleed after progestin challenge may occur in the anovulatory phase, with some estrogen being produced, but in most women the failure to exhibit a withdrawal bleed following a progestin challenge suggests that the disease is more severe. If untreated, it may lead to problems associated with estrogen deficiency such as osteoporosis. Following a reduction of stress or exercise and increased weight gain, we may see a reversal from amenorrhea to ovulatory menstrual cycles, requiring no further treatment.

Other chronic debilitating diseases such as acquired immunodeficiency syndrome (AIDS), malabsorption, or cancer may result in hypogonadotropic hypogonadism. A relatively rare cause of hypothalamic amenorrhea in women is due to Kallmann's syndrome, which has been shown to be associated with defects in olfactory bulb development. It is known that GnRH neurons develop in this same rostral part of the brain as the olfactory bulbs. Thus, women with Kallmann's syndrome exhibit not only low gonadotropins and amenorrhea but also a lack of the sense of smell, or anosmia. As stated previously, the diagnosis of functional disorders of the CNS and hypothalamic-pituitary axis resulting in amenorrhea may be suggested by the patient's history, but it usually requires some form of imaging using MRI or CAT scan.

Disorders of the pituitary include not only tumors but also Sheehan's syndrome (postpartum necrosis of the pituitary due to hemorrhage), spontaneous necrosis of the pituitary, congenital absence of the pituitary, empty sella, and inflammatory or infiltrative disorders (see Table 1). However, the most common causes of amenorrhea associated with the pituitary are disorders associated with increased prolactin secretion. These women present with amenorrhea plus galactorrhea associated with an increased prolactin level. In women with amenorrhea and elevated prolactin levels, imaging of the pituitary is required and a thyroid-stimulating hormone (TSH) level is obtained as well. Hyperprolactinemia and galactorrhea may also be associated with taking medications, including psychoactive medications such as tricyclic antidepressants, phenothiazines, amphetamine, haloperidol, and chlorpromazine and antihypertensive medications such as verapamil and α -methyldopa. Hyperprolactinemia and galactorrhea may also be seen with recent breastfeeding. The majority of women with hyperprolactinemia do not exhibit demonstrable pituitary

tumors. Thus, there is some controversy about the level of serum prolactin at which head imaging should begin; some have suggested prolactin levels exceeding $50 \mu\text{g l}^{-1}$. Women with prolactin levels greater than $50 \mu\text{g l}^{-1}$ have approximately a 20% chance of presenting with a pituitary tumor, whereas those who have $100 \mu\text{g l}^{-1}$ have a 50% risk; those with over $200 \mu\text{g l}^{-1}$ have an approximately 90–100% chance of harboring a prolactin-secreting tumor.

Ovarian Failure

Ovarian failure in women is associated with increased gonadotropins in response to low or no secretions from the ovary, and hence the term hypergonadotropic hypogonadism is often used. The development of amenorrhea and hypoestrogenism associated with elevated gonadotropins prior to the age of 40 defines ovarian failure. However, it should be pointed out that cessation of ovarian function due to loss of germ cells and follicles in the ovary can occur at any age, even *in utero*, such as in gonadal dysgenesis or ovarian agenesis. When ovarian failure occurs prior to puberty, the presentation is that of a phenotypic female with primary amenorrhea and lack of secondary sexual development. When it occurs after pubertal development, the presentation is secondary amenorrhea with the primary complaint being hot flashes.

As is true for disorders of chronic anovulation, ovarian failure can result from several causes (see Table 1). These include gonadal agenesis or dysgenesis. Gonadal dysgenesis can be associated with normal or abnormal karyotypes. Women harboring Y chromosome material have an increased risk of gonadal tumors, and the gonads should be removed. Rare causes of ovarian failure include 17α -hydroxylase deficiency, resistant ovary syndrome, autoimmune disorders associated with galactosemia, and iatrogenic causes due to chemotherapy or radiation therapy. The diagnosis of ovarian failure should be suspected in all cases of primary amenorrhea and sexual infantilism and in women with secondary amenorrhea who develop hot flashes and other signs of estrogen deficiency. As stated earlier, this is confirmed by documenting an increased FSH at the menopausal range (i.e., $>30\text{--}40 \text{ IU l}^{-1}$).

Female Reproductive Tract Defects

Defects of the female reproductive tract can result in amenorrhea either from obstruction or the absence of endometrial tissue. Some of these conditions may be developmental and some may be due to iatrogenic causes. Women with anatomical defects have normal ovaries and ovarian function and develop secondary sexual characteristics. Ovulation can be proven by changes in the basal body temperature or by elevated

serum progesterone levels in the luteal phase of the menstrual cycle. These women also have a normal female 46,XX karyotype.

The logical approach to the evaluation and classification of anatomic defects is to start from the lowest entry at the opening of the reproductive tract and move upward. Thus, anatomic causes include labial agglutination or labial fusion and imperforate hymen. If the obstruction is farther up into the vagina, it is known as a transverse vaginal septum. Rarely, a complete absence of the cervix may be encountered. Women with all these conditions often present with increasing abdominal pain resulting from the accumulation of blood behind the obstruction. In such instances the pain is cyclic and predictable in nature, and it is associated with the onset of menstruation. If the diagnosis is delayed, endometriosis, adhesions, and infertility may result.

There are two additional disorders that are not associated with obstruction and pain but result from lack of development of the uterus and vagina: Müllerian agenesis or dysgenesis associated with a 46,XX karyotype and androgen insensitivity syndrome (formerly known as testicular feminization, in which there is a testis and a 46,XY karyotype). The diagnosis of androgen insensitivity is easily confirmed by the observation of the absence of pubic and axillary hair due to the androgen resistance and the finding of male-level testosterone on laboratory analysis. Individuals with androgen insensitivity should have their testes removed to prevent tumor development following the completion of breast development, which occurs at ages 12–14.

In rare instances, a woman may have both the absence of the uterus and the lack of female secondary sexual characteristics. In such individuals, a karyotype should be obtained. Individuals with a 46,XY karyotype have androgen deficiency, most commonly due to 17 α -hydroxylase deficiency, testicular regression, or gonadal dysgenesis. Those with the 46,XX karyotype have Müllerian agenesis plus another disorder such as ovarian failure and/or hypogonadotropic hypogonadism.

In addition, reproductive tract defects may also be due to iatrogenic causes such as scarring and stenosis of the cervix following dilation and curettage, conization, laser, or loop electrosurgical excision procedure (LEEP) used to treat cervical dysplasia. The destruction of the endometrium following a vigorous curettage after postpartum hemorrhage or therapeutic abortion or subsequent infection associated with a missed abortion results in scar tissue or uterine synechiae (Asherman's syndrome). The diagnosis of this defect is rare in the absence of a previous surgical procedure or pregnancy.

The diagnosis of female reproductive tract defects is established by history and physical examination. The presence of a labial agglutination, fusion, imperforate hymen, and transverse vaginal septum are easily recognized. An ultrasound can be obtained to confirm the location of the blockage and the presence or absence of the uterus. To evaluate and confirm the diagnosis of Asherman's syndrome, a hysteroscopic procedure or occasional hysterosalpingography may be indicated. Occasionally, laparoscopy may be required to confirm the final diagnosis of the reproductive tract anomaly.

Clinical Evaluation of Amenorrhea

The evaluation of amenorrhea should begin with a thorough medical history and physical examination. If a woman presents with primary amenorrhea and absence of secondary sexual characteristics, this may suggest a primary gonadal problem or ovarian failure. In contrast, if a woman presents with previous regular menstrual periods but develops amenorrhea following a dilation and curettage for hemorrhage, this suggests a diagnosis of uterine adhesions. The evaluation of amenorrhea requires a knowledge of the classification in [Table 1](#). A flowchart to aid in evaluating women with amenorrhea is presented in [Figure 1](#). This evaluation includes women with a female phenotype but excludes women with disorders of virilization or sexual ambiguity. The laboratory tests required include a pregnancy test and tests for levels of FSH, prolactin, and TSH. We find it helpful to evaluate the estrogen status based on the progestin withdrawal test. This is helpful prior to determining therapy for infertility as well. Evaluation of the estrogen status includes two parts: (1) the physical examination of the patient, looking for signs of previous estrogen secretion such as breast development and current estrogen secretion, which includes the presence of a well rugated, moist vagina with abundant clear stretchable cervical mucus (known as spinnbarkeit). The estrogen status can be confirmed by a progestin challenge using medroxyprogesterone acetate (5–10 mg once or twice daily for 5 days) or progesterone in oil (100–200 mg) administered intramuscularly, particularly when patient compliance is an issue. It should be pointed out that Depo-Provera should not be used because this, in fact, causes amenorrhea. The one exception to this overall evaluation is a woman who is known on physical examination to have an absent uterus; in this case it is more prudent and cost effective to perform a few tests such as a chromosomal analysis to differentiate Müllerian agenesis from complete androgen insensitivity syndrome.

Following the initial visit, the patient returns and the physician evaluates the levels of FSH and prolactin and the response to the progestin challenge. If the FSH is elevated, the evaluation is directed to ovarian failure or hypergonadotropic hypogonadism. However, if the FSH is low or normal, then the evaluation of the progestin challenge test is used to differentiate the two categories of chronic anovulation. Finally, if the serum prolactin is elevated, the TSH level should be measured and the pituitary further evaluated by either an MRI or CAT scan.

The majority of women with significant elevations of prolactin do not respond to a progestin challenge and thus are included in the category of chronic anovulation with estrogen absent. Such women should experience withdrawal bleeding following a cycle of estrogen plus progestin, but this test and its results are somewhat confusing because a high percentage of women do not bleed after one cycle of therapy. An evaluation of the outflow tract or reproductive tract is made on physical examination and if needed by hysterosalpingography, ultrasound, or hysteroscopy. During the physical examination, if the patient exhibits a vagina, a patent cervix that can be confirmed by uterine sounding (placing an instrument through the cervix into the uterine cavity), and a uterus without a history of pregnancy or operative procedure, a disorder of the female reproductive tract is unlikely. After disorders of the reproductive tract have been ruled out, a normal or low FSH level indicates a

disorder of the hypothalamus or pituitary. Most of these women have functional disorders of the hypothalamus, although lesions of the hypothalamus and pituitary may be present. The use of an MRI or CAT scan depends on the history, clinical presentation, and prolactin level. Overall, the flowchart presented in [Figure 1](#) is quite useful in the majority of cases, and the diagnosis is simplified and easily obtained.

Management of Amenorrhea

In general, the management of amenorrhea depends on the cause and on the current desires of the patient.

Chronic Anovulation

The treatment of women with chronic anovulation can be subdivided as before into treatment for those who are making estrogen and for those who are not. In women who are producing estrogen but are not ovulating, the treatment depends on the desires of the patient. If the patient is obese, weight loss should be encouraged, and this may in fact improve the overall clinical situation, including insulin resistance, if present, and hypercholesterolemia. Recent evidence has also shown that metformin, an oral hypoglycemic agent, may reduce insulin resistance, improve hirsutism, and cause the resumption of menses. If the woman is not hirsute and does not desire pregnancy, monthly withdrawal menses should be induced by either progesterone therapy or, more simply, using

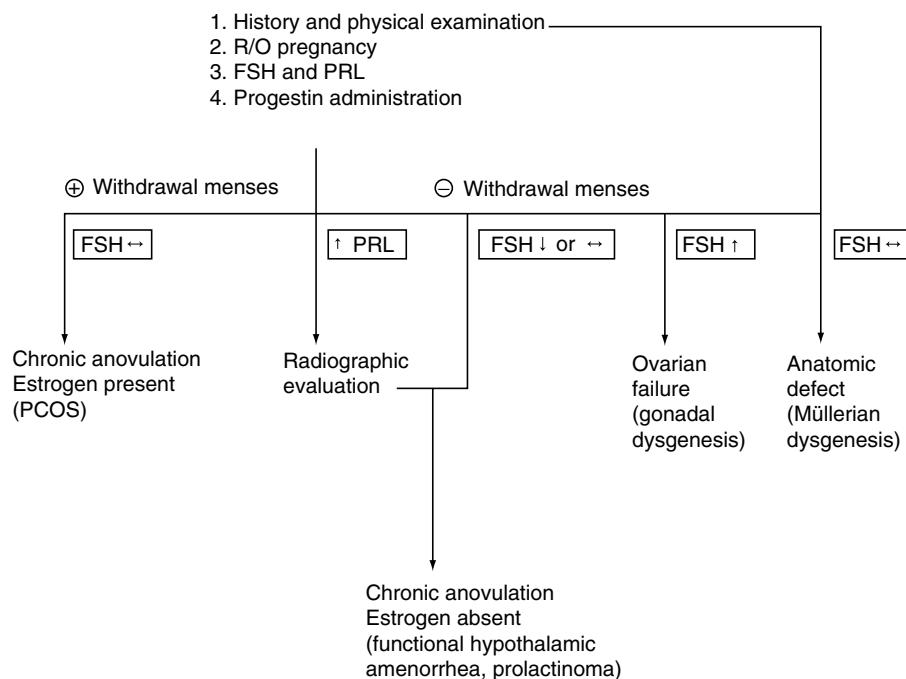


Figure 1 Suggested flow chart to aid in the evaluation of women with amenorrhea. FSH, follicle-stimulating hormone; PCOS, polycystic ovarian syndrome; PRL, prolactin; R/O, rule out.

oral contraceptive pills, which reduces the risk of hemorrhage (i.e., dysfunctional uterine bleeding and endometrial neoplasia). On the other hand, if the woman is hirsute but does not desire pregnancy, excess male hormone production can be suppressed through oral contraceptive pills and/or antiandrogens. Oral contraceptive pills are also indicated if there are disorders in menstruation (i.e., dysfunctional uterine bleeding). Finally, if pregnancy is desired, ovulation can be induced with either clomiphene citrate, metformin, gonadotropins, or, occasionally, laparoscopic ovarian wedge resection or drilling.

In those women who are not making estrogen with chronic anovulation, the treatment should be directed at eliminating the primary cause such as weight loss, exercise, or stress. Treatment in women with amenorrhea and low estrogen should be aggressive because delayed treatment can lead to reduced body mass development and early osteoporosis. In women with elevated prolactin levels in the absence of prolactin-secreting tumors, cyclic estrogen–progestin therapy or, more commonly, oral contraceptive pills can be prescribed. Oral contraceptive pills and hormone replacement therapy not only prevent bone loss but also maintain secondary sexual characteristics. If pregnancy is desired, it is important that body weight and nutritional requirements for the potential developing fetus be returned to normal prior to the induction of ovulation. But if the amenorrhea persists in spite of the reduction in stress and increase in body weight and percentage of body fat, then gonadotropins or pulsatile GnRH therapy is often successful in inducing ovulation and pregnancy. In women who present with pituitary tumors that secrete prolactin, the treatment of choice is pharmacological, using dopamine agonists such as bromocriptine or cabergoline.

Ovarian Failure

Most cases of gonadal or ovarian failure are permanent, and patients should be started on hormone replacement therapy, particularly estrogen, as soon as possible. Again, estrogen maintains secondary characteristics and prevents premature osteoporosis. If the diagnosis of gonadal failure is made in a woman prior to breast development such as in Turner's syndrome, a regimen of low-dose estrogen gradually increased over a period of time may be important in producing normal breast development. Growth hormone may also be used prior to estrogen therapy in achieving greater height, if indicated. In women with gonadal dysgenesis, estrogen treatment begins with low-dose (i.e., conjugated) estrogens (0.3 mg) for 3–6 months, slowly increasing the dosage from

0.625 to 1.25 mg over 1 year to augment breast development. It is necessary to initiate progestin therapy after approximately 1 year of estrogen therapy to induce withdrawal bleeding to prevent endometrial hyperplasia. However, if progestin is started prior to the initiation of breast development, then breast development may be abnormal. Women with disorders such as 17 α -hydroxylase deficiency should be treated with glucocorticoids as well as with hormone therapy. Women with autoimmune ovarian failure have been treated with a variety of medicines, but none appears to be successful, so hormone replacement therapy appears appropriate. Occasionally, the repositioning of the ovary or oophoropexy may be helpful prior to women receiving radiation therapy. In addition, it has been hypothesized, but not proven, that pretreatment with LH-releasing hormone analogs or oral contraceptive pills prior to chemotherapy may be successful in maintaining ovarian function.

Women with amenorrhea due to ovarian failure are rarely able to conceive on their own. In some cases, ovarian follicular depletion may not be complete and spontaneous ovulation and a rare pregnancy may occur. The current treatment for infertility due to ovarian failure is to obtain donor oocytes from normal ovulatory women, followed by *in vitro* fertilization in which the sperm of the patient's husband is used to fertilize the donor eggs. The fertilized embryo is then transferred to the patient, who has been appropriately treated with exogenous estrogen and progesterone synchronized to mimic the normal ovulatory cycle.

Female Reproductive Tract Disorders

The primary treatment of outflow tract disorders is surgical – the incision of labial fusion, imperforate hymen, and vaginal septum, which leads to the return of regular menstrual periods and fertility. With respect to specific disorders, chronic labial adhesions in children can be treated with intermittent estrogen cream; however, a functional vagina in women with Müllerian agenesis or testicular feminization is more difficult to achieve. First, an attempt at nonsurgical dilation of a blind-ending vaginal pouch or perineal dimple is indicated. If this fails, a reconstruction of the vagina using skin grafts is performed. Disorders of cervical obstruction can be treated by the dilation of the cervix, and pregnancy can be achieved by intrauterine insemination. If the cervix is absent, however, the treatment usually requires a hysterectomy because the retained blood behind the obstruction can cause significant pain and infection. Uterine scarring or Asherman's syndrome is best treated by the direct hysteroscopic resection of the adhesions.

Further Reading

- Carr, B. R. (1994). Current evaluation and treatment of amenorrhea. *The American Fertility Society: Guidelines for Practice*, pp. 1–11.
- Carr, B. R., Blackwell, R. E. and Azziz, R. (eds.) (2005). *Essential reproductive medicine*. New York: McGraw-Hill.
- Ferin, M. (1999). Clinical review 105: stress and the reproductive cycle. *Journal of Clinical Endocrinology and Metabolism* 84, 1768–1774.
- Gidwani, G. P. and Rome, E. S. (1997). Eating disorders. *Clinical Obstetrics and Gynecology* 40, 601–615.
- Redman, L. M. and Loucks, A. B. (2005). Menstrual disorders in athletes. *Sports Medicine* 35, 747–755.
- Rock, J. A., Zacur, H. A., Dlugi, A. M., et al. (1982). Pregnancy success following corrections of imperforate hymen and complete transverse vaginal septum. *Obstetrics and Gynecology* 59, 448–451.
- Speroff, L. and Fritz, M. A. (eds.) (2005). *Clinical gynecologic endocrinology and infertility*. Philadelphia: Lippincott Williams and Wilkins.
- Stein, I. F. and Leventhal, M. L. (1935). Amenorrhea associated with bilateral polycystic ovaries. *American Journal of Obstetrics and Gynecology* 29, 181.
- Turner, H. H. (1938). A syndrome of infantilism, congenital webbed neck, and cubitus valgus. *Endocrinology* 23, 66.
- Warren, M. P. and Fried, J. L. (2001). Hypothalamic amenorrhea, the effects of environmental stresses on the reproductive system: a central effect of the central nervous system. *Endocrinology and Metabolism Clinics of North America* 30, 611–629.

Amnesia**J S Simons**

Institute of Cognitive Neuroscience, University College London, London, UK

K S Graham

MRC Cognition and Brain Sciences Unit, Cambridge, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J S Simons and K S Graham, volume 1, pp 176–177, © 2000, Elsevier Inc.

Semantic memory

Our general knowledge about the world, such as information about words and their meanings, facts, concepts, objects, and people; typically retrieved without recollection of where and when the information was initially acquired.

Temporal lobe

The area of the brain generally considered the permanent store of our episodic and semantic memories, part of the neocortex.

Organic Amnesia**Stress-Related and Transient Amnesic States****Glossary****Episodic memory**

The store of personally experienced events from the past, including details of the content of the episode and also of its context — when and where it occurred.

Frontal lobe

The brain area thought to play a role in the strategic control of memory processes necessary for complex recollective memory functions.

Hippocampal complex

Structures located in the medial temporal lobe of the human brain (e.g., the hippocampus, subiculum, parahippocampal gyrus, and entorhinal cortex); implicated in the learning and at least the temporary storage of recent memories.

The term amnesia refers to a partial or total loss of memory that can result from organic brain damage and/or some forms of psychological stress. The loss of memory can affect events or experiences that occur subsequent to the onset of the disorder (anterograde amnesia) or those that took place prior to its appearance (retrograde amnesia). Furthermore, although amnesia most often occurs for personal autobiographical events (episodic memory), some patients also show a loss of their factual knowledge about the world (semantic memory).

Organic Amnesia

The most well-studied case of organic amnesia in the literature is that of HM, who became amnesic in his mid-twenties after an operation to relieve the symptoms of severe epilepsy. HM's operation involved surgery to the hippocampal complex and parts of the temporal lobe in both brain hemispheres. Although

HMs epilepsy improved following surgery, he became profoundly amnesic, unable to remember events just after they had happened and many that had occurred prior to his operation. For example, he was severely impaired at recalling lists of words that had just been read to him and at recognizing those words as the ones he had heard before when they were presented to him again. This was also the case when he was asked to remember other types of stimuli, such as faces, numbers, sounds, or shapes. There are now over 50 reported patients who, like HM, show difficulties with learning after damage to the hippocampal complex. This region in the human brain is, therefore, considered critical for the acquisition of new episodic and semantic memories.

As well as being unable to remember events from moment to moment, patients with damage similar to HM often have difficulty remembering events that occurred a number of years previously. This pattern of memory loss, in which memories from the distant past (e.g., childhood or early adulthood) are typically remembered better than more recently experienced events (e.g., from the day, week, or year before), is a phenomenon that has fascinated memory researchers for over 100 years. The fact that memory loss can be so dramatically affected by the age of the memory has led some researchers to conclude that the system involved in acquiring new human memories, the hippocampal complex, may not necessarily be involved in the permanent storage of all our episodic and semantic memories. Instead, other areas of the brain, such as the anterior temporal lobes, may be the location for our enduring stores of human memory, especially semantic knowledge about the world. Studies of patients with damage that affects the temporal neocortex, but less so the hippocampal complex, highlight the complementary roles played by these regions in long-term memory. These patients often show the opposite pattern of memory impairment to HM and other patients with damage to the hippocampal complex, that is, better recall of recent memories than those from childhood or early adulthood. Although it is currently unclear how the different temporal-lobe regions interact in the acquisition and storage of episodic and semantic memory and whether the effects of memory age evident in patients reflect a true transfer of information between brain structures, most current theories of memory concur that memory consolidation, perhaps occurring during sleep, is a critical aspect of memory storage.

In recent years, the important role in memory played by another brain area, the frontal lobe, has become clearer. Patients with specific damage to this region often exhibit symptoms of impulsiveness and lack of inhibition that affect personality and behavior,

but careful testing can reveal impairments in certain aspects of memory function. For example, such patients often have difficulty remembering details of the context in which previous events were experienced, such as where or when the event occurred. Moreover, they are typically impaired, like patients with hippocampal complex damage such as HM, at recalling lists of words that were previously presented. In contrast to hippocampal complex dysfunction, however, frontal lobe damage may spare the ability to recognize information as having been seen before. Such evidence suggests that the frontal lobe and hippocampal complex work together to perform relatively complex, effortful memory tasks such as recalling words, that require the strategic search of the memory store, as well as organization and manipulation of retrieved material. The idea that interactions between the frontal and temporal lobe areas are important for memory function is supported by data from a patient with damage to the uncinate fascicle, a bundle of fibers that provides an important connection between these brain regions. This patient was impaired at recalling autobiographical events from his past, but performed normally at recognizing and discriminating events that he had experienced from those he had not.

Stress-Related and Transient Amnesic States

In addition to instances of organic amnesia, psychological stresses can result in impairments to and distortions of memory. One of the main types of nonorganic memory loss occurs as part of the syndrome termed posttraumatic stress disorder and in many circumstances can be linked to the experience of a precipitating traumatic event such as a natural disaster, accident, or war. The phenomenon was first recognized during World War I, when soldiers returned from the front affected by shell shock, characterized by recurring nightmares and flashbacks to their traumatic experiences. The disorder can also manifest as the loss of one's identity and of autobiographical memories from around the time of the precipitating event, referred to as psychogenic or dissociative amnesia. For example, studies of soldiers conducted following World War II suggested that up to 5% had no memory for the traumatic combat events they had recently experienced.

One possible explanation for these types of amnesia may be that they reflect a psychological defense mechanism that attempts to repress or inhibit retrieval of the traumatic memory and prevent it from reaching conscious awareness. Evidence suggests that such retrieval inhibition processes are supported by regions in the frontal lobes, emphasizing the important role

for this brain area in memory. Stress-related forgetting may be transient in nature, and patients can sometimes recover their memories after the appropriate treatment or therapy. The recovery of traumatic memories following therapy has proven controversial, however, with some commentators questioning the accuracy of such recovered recollections and suggesting that the techniques sometimes used in therapy sessions might lead patients to remember false traumatic events (ones that had not in fact previously occurred). This debate remains largely unresolved, with evidence suggesting that retrieval inhibition processes do certainly exist but, at the same time, that one's memory of past events can be highly susceptible to suggestibility biases and distortions at the time of recall.

Although one of the diagnostic criteria for psychogenic amnesia is that there is no evidence of structural brain damage, the increasing use of sophisticated neuroimaging techniques that measure levels of blood flow in the brain suggest that there can be altered brain function in some cases, although the reasons for this are currently unclear. For example, one study examined brain activity in a patient with a persistent psychogenic amnesia for the whole of his past life. When the patient was thinking about his autobiographical memories, blood flow was found to be greater in the same brain regions that control subjects engaged during the processing of nonpersonal memories. This suggests that some patients with psychogenic amnesia may treat previously personally salient episodic memories in a neutral, semantic way, as if they no longer belonged to them, perhaps in order to escape the emotional associations of the traumatic episode(s) that triggered the amnesia. Alternatively, it is possible that the brain regions typically involved in processing autobiographical recollections, which are similar to those involved in processing emotional information, may be temporarily disrupted as a result of experiencing the stressful event, leaving the brain

regions used for processing neutral memories as the only means by which autobiographical memories can be processed.

The important link between stress and memory can be further revealed in a syndrome called transient global amnesia (TGA). Although not classified as a psychogenic amnesia (patients are rarely confused about their own identity), the disorder is characterized by abrupt confusion and a complete anterograde amnesia, which eventually shrinks after (typically) 4–6 hours. Although there are known risk factors for TGA, such as a history of epilepsy or migraine, in as many as 30% of cases attacks can be directly linked to precipitating stresses, such as strenuous exertion, pain, immersion in water, and emotional events. Recent neuroimaging studies have revealed physiological changes (low blood flow) in the hippocampal complex during an attack. It is possible, therefore, that extreme stress may result in memory impairment because the hippocampal complex is part of a complicated network in the human brain that mediates fear-related behavior and stress responses, as well as aspects of memory function. Highly stressful situations may overstimulate the hippocampus and related structures, leading to decreased blood flow in this region of the brain and subsequent memory loss.

Further Reading

- Baddeley, A. D., Wilson, B. A. and Watts, F. N. (1995). *Handbook of memory disorders*. New York: John Wiley.
- Cohen, N. J. and Eichenbaum, H. B. (1993). *Memory, amnesia, and the hippocampal system*. Cambridge, MA: MIT Press.
- Parkin, A. J. (1998). *Case studies in the neuropsychology of memory*. Hove, UK: Psychology Press.
- Simons, J. S. and Spiers, H. J. (2003). Prefrontal and medial temporal lobe interactions in long-term memory. *Nature Reviews Neuroscience* 4, 637–648.
- Squire, L. R. and Schacter, D. L. (2002). *Neuropsychology of memory* (3rd edn.). New York: Gilford Press.

Amygdala

M S Fanselow and R Ponnusamy

University of California, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M S Fanselow and G D Dale, volume 1, pp 178–182, © 2000, Elsevier Inc.

Introduction

Multiple Sources of Input: Afferents to Amygdala

Variegated Outputs: Efferents from Amygdala

Conclusions and Functions

Glossary

<i>Central nucleus of the amygdala (CeA)</i>	This striatal-like nucleus is considered the motor interface of the amygdala.
<i>Lateral amygdala (LA) and basolateral amygdala (BLA)</i>	These cortical-like structures are considered the sensory interface for the amygdala's role in fear, anxiety, and stress.
<i>Long-term potentiation (LTP)</i>	A relatively long-lasting form of synaptic plasticity that has been proposed as a mechanism that mediates some forms of learning and memory
<i>Pavlovian fear conditioning</i>	A learning paradigm in which an aversive event is paired with neutral stimuli, resulting in a variety of defensive behavioral and autonomic conditional responses to the previously neutral stimulus. It is the natural way that defensive behavior is controlled but can also contribute to anxiety disorders. In the laboratory it has become a standard way to assess learning and memory.

Introduction

The amygdala, an almond-shaped cluster of interconnected nuclei in the medial temporal lobe, may be the neural structure most intimately tied to emotion. It has long been known that damage to the amygdala produces disturbances in emotional behavior. The linkage is particularly compelling in the case of negative emotions such as fear. In both human and nonhuman animals, fear-related behavior, which is typically analyzed with Pavlovian fear conditioning, depends on the amygdala. Humans with bilateral amygdala damage have difficulty in recognizing the

negative emotions in the facial expression of others and as a result may make inappropriate social advances. Rats with lesions of the amygdala lose their fear of cats. Additionally, the ability of emotion to influence declarative memory storage (flash bulb memories) is mediated by the amygdala. The amygdala is well situated for a role in stress and emotion. Features of the environment, initially processed by several different brain regions, converge on the amygdala. Much of this information relates to external and internal stress. Plasticity within certain nuclei of the amygdala allows this structure to connect these stressors with environmental features that predict stress. Therefore, it may also act as a feedforward device that generates behavior in anticipation of stress. This article first reviews the major afferents to the amygdala that bring it this environmental information. Then it turns to the efferents that project to the brain regions that generate behavioral responses to stress. It concludes with a description of the processing that goes on within the amygdala that enables this feedforward regulation of stress.

Multiple Sources of Input: Afferents to Amygdala

Nucleus Tractus Solitarius

The nucleus tractus solitarius (NTS) projects to the amygdala's central nucleus (CeA), both directly and indirectly, via the parabrachial nucleus. A large body of evidence implicates the NTS in stress-related responses. It receives first-order projections from several cranial nerves, including the vagus. These inputs primarily reflect the contribution of visceral sensory information but also convey general somatic input. In addition, ascending pain-related fibers originating from the spinal cord terminate in the NTS. Thus, the NTS can provide the amygdala with visceral sensory, somatosensory, and pain-related information.

Thalamus

The amygdala receives a wide range of sensory information via monosynaptic projections from the thalamus. Considerable functional importance has been attributed to projections from the posterior thalamus to the amygdala. The moderate projection from the paraventricular nucleus terminates in the medial and ventrolateral divisions of the lateral amygdala (LA). Those arising from the medial geniculate provide the LA with auditory information. Pain information

represented in the posterior intralaminar nucleus of the thalamus also project to several amygdaloid nuclei. Stimulation of these posterior thalamic projections to the LA can support Pavlovian fear conditioning to an auditory cue. Somatosensory regions of the thalamus are an important route of unconditional stimulus transmission to the amygdala in fear conditioning. However, information about the unconditional somatosensory stimulus is also transmitted from other sources that send fibers through (but do not form essential synapses in) the thalamus en route to the amygdala. The CeA may receive direct thalamic inputs about the conditional stimulus as well as nociceptive stimuli. Furthermore, parts of the ventro-posterior medial thalamic nucleus that primarily receive visceral and gustatory information project to the amygdala. These projections may support the amygdala's role in avoidance of novel flavors. In summary, the thalamus is a significant and direct source of information about the internal and external milieu for the amygdala.

Periaqueductal Gray

The midbrain periaqueductal gray (PAG) is the key structure for integrated defensive behavior. While much emphasis has been placed on descending projections from the amygdala to PAG (see later), these projections are largely reciprocal. There are strong projections from PAG to amygdala, particularly to the CeA. In certain situations, lesions of the dorso-lateral PAG will enhance conditional fear responses that depend on the amygdala. This has led to the suggestion that the ascending projections from the dorsolateral PAG may exert an inhibitory influence on the amygdala.

Temporal, Insular, and Parietal Cortex

There is a massive exchange of information between the amygdala and the cortex. Along with the sensory projections from the thalamus, the amygdala receives both modality-specific and multimodal neocortical projections. Unimodal projections coding aspects of visual, auditory, and somatosensory information project primarily to the LA and the basolateral nucleus (BLA). Importantly, these projections tend not to originate from the primary sensory cortex. Rather, unimodal sensory information originates from secondary sensory regions. For example, auditory projections arise from Te2 and Te3, and visual information arrives through the perirhinal cortex. The insular cortex, which receives monosynaptic projections from the primary somatosensory cortex, sends substantial projections to both the BLA and LA. It has been suggested that these projections provide the amygdala

with nociceptive information that can support fear conditioning.

The hippocampal formation receives some of the most highly preprocessed information in the brain. This component of the temporal cortex projects to the amygdala via sparse projections from the entorhinal cortex, CA1, and subiculum. Fibers from the entorhinal cortex and subiculum pass through the ventral angular bundle, and electrical stimulation of this tract results in a monosynaptic response in the BLA. These projections seem to play an important role in providing contextual or place information for amygdala-dependent behaviors, as lesions of this tract produce deficits in contextual fear conditioning. At least as far as simple emotion-related behaviors are concerned, the thalamo-amygdala and cortico-amygdala circuits appear to have some redundant function and act compensatorily because damage in both pathways appears necessary to eliminate amygdala-dependent behaviors. It is possible that the cortical pathways may allow amygdala-dependent stress responses to be more finely contoured to the specifics of environmental stimulation, but such an analysis has not been performed.

Frontal Cortex

The infralimbic (IL) region of the medial prefrontal cortex strongly projects to the intercalated cells in the amygdala. In turn, intercalated cells project to the CeA. Lesions of IL impair extinction, and electrical stimulation of IL reduces the expression of conditional fear. Therefore, it has been proposed that insufficient inhibition of the amygdala by the medial prefrontal cortex could predispose an individual to develop anxiety disorders. Thus, in terms of cortical information, the amygdala predominately receives highly processed sensory representations that likely play important roles in the coordination of appropriate behavioral responses.

Summary

A diverse array of neural pathways brings the amygdala a wide range of environmental information. Inputs from various functional systems terminate in certain amygdaloid nuclei, and, therefore, different nuclear divisions are exposed to different sets of information. Note that the same divisions might receive inputs from various sources. These interdivisional projections provide a route by which the various kinds of information entering the different portions of amygdala can be associated. In addition to convergence, parallel distribution of information from the same source to multiple amygdaloid areas is another typical feature in the organization of amygdaloid

inputs. Most of this sensory information has been highly processed before it arrives at the amygdala. This suggests that the amygdala may provide an information-integrating function in the regulation of stress.

Variegated Outputs: Efferents from Amygdala

In addition to providing reciprocal projections to many of the sensory regions from which it receives input, the amygdala also sends efferent projections to regions responsible for behavioral, autonomic, and endocrine responses to stress. These outputs are organized in a point-to-point manner. It also provides parallel outputs to multifunctional systems, and there are also parallel outputs from various amygdaloid nuclei to the same functional system. These anatomical relationships allow amygdala activity to influence the unconditional response to stress and also to provide substrates by which stimuli predictive of impending stressors can produce anticipatory behavioral adjustments.

Hypothalamus

Physical stress produces a variety of neuroendocrine responses, including increases in plasma adrenocorticotropin (ACTH), corticosterone, norepinephrine, and epinephrine. Control of these responses is mediated in part by a direct action of the hypothalamic paraventricular nucleus on pituitary secretion. The paraventricular hypothalamus has been shown to receive direct projections from the CeA, providing a pathway by which amygdala activity can contribute to endocrine stress responses. A variety of studies have demonstrated a role for the amygdala in the expression of neuroendocrine responses to stress. Lesions of the amygdala attenuate the elevation of plasma epinephrine, norepinephrine, and corticosterone normally observed following footshock, and similarly attenuate the elevation of plasma ACTH produced by immobilization stress but do not completely block these responses. However, given that amygdala lesions do not completely block or alter baseline hormone levels, it is possible that the amygdala lesion-induced attenuation of hormone responses reflects disruption of feedforward amygdala-dependent conditional responses rather than an amygdala control of unconditional endocrine responses.

Projections from the LA and the BLA to the hypothalamus are light and target the lateral hypothalamic area. However, the CeA provides substantial inputs to a larger number of hypothalamic nuclei. Most of the projections originate in the medial division.

Moderate to heavy projections are directed to the medial and lateral preoptic area, lateral hypothalamus, and paraventricular nucleus. Lesions of both the CeA of the amygdala and the lateral hypothalamus block the increase in arterial blood pressure caused by a conditional fear stimulus, leading to the suggestion that projections from the amygdala to the lateral hypothalamus mediate this response. Importantly, the hypertension to the shock unconditional response remains intact following hypothalamic lesions. In addition to these direct projections to the hypothalamus, the CeA has strong projections to the bed nuclei of the stria terminalis, which also innervates hypothalamic nuclei. Thus, the lateral hypothalamus seems to mediate conditional sympathetic autonomic arousal to fear stimuli via input from the CeA.

Periaqueductal Gray

The CeA projects heavily to the caudal half of the midbrain PAG. These projections initially innervate more dorsal regions of the PAG, areas that serve active or circa strike defensive behaviors and autonomic arousal. Stimulation of the medial and central amygdala inhibits the behaviors that are generated by stimulation of these dorsal regions of the PAG.

As one moves caudally in the PAG, the CeA projections gradually rotate to innervate the ventral portions of the PAG. These ventral PAG areas are essential for conditional fear-related defenses such as freezing. This area is also critical for the opioid analgesia that accompanies fear. Consistent with this pattern, cells in the caudal ventral PAG express *c-fos* following fear conditioning. However, this part of the PAG is not critical for the autonomic responses to conditional fear stimuli. The amygdala, predominantly through projection emanating from the CeA, exerts control over species-specific defensive behaviors organized by the PAG.

Brain Stem

The CeA sends projections throughout the brain stem. One function of these projections is to provide control over autonomic function. For example, stimulation of the rabbit's CeA causes a heart rate deceleration mediated by monosynaptic projections to the dorsal motor nucleus of the vagus. This response mimics the bradycardia rabbits show to conditional fear stimuli, which also depends on the CeA. While amygdala lesions block the bradycardiac conditional response to the stimulus paired with electric shock, it is important to note that these lesions do not affect baseline heart rate or the unconditional heart rate orienting response to the conditional stimulus in rabbits. Thus, projections from the CeA to brain stem

autonomic nuclei may mediate parasympathetically controlled conditional responses.

Amygdala stimulation enhances reflexes organized in the brain stem such as eyeblink and startle. There are both mono- and disynaptic projections from the CeA to the nucleus reticularis pontis caudalis, which is the part of the brain stem acoustic startle circuit that is modulated by conditional fear stimuli. It has previously been shown in rats, cats, rabbits, and monkeys, using various anterograde tracers, that CeA neurons project to brain stem areas including the NTS. These results provide conclusive evidence for a GABAergic pathway from the CeA to the NTS. Although lesions of the amygdala block the ability of fear stimuli to modulate reflexes, it is important to note that the reflex itself is left intact.

Thus, projections from the CeA to several brain stem regions allow for widespread modulation of autonomic function and reflexive behaviors.

Cortex

In addition to the amygdala receiving substantial cortical projections, recent work suggests that it is in a position to exert widespread influence over cortical activity. Based on anterograde tracer studies, the LA provides strong, topographically organized inputs to the ventral and dorsal perirhinal cortices. The LA also provides a moderate input to the infralimbic cortex as well as to the ventral agranular insula. Projections from the BLA to the frontal cortex are more widespread than those from any other amygdaloid nucleus. The CeA sends substantial ascending projections to the cholinergic basal forebrain, which in turn projects diffusely to virtually all cortical regions. Basal forebrain stimulations produce desynchronization of cortical EEG, as does CeA stimulation. Basal forebrain stimulation supports associative changes in receptive field properties of auditory cortical neurons. Based on these findings, it has been suggested that cholinergic modulation of cortical activity, mediated by amygdala projections to the basal forebrain, is critical for the induction of the stimulus-specific changes in cortical receptive fields observed during Pavlovian fear conditioning.

Hippocampal Formation

In addition to modulating cortical dynamics, the amygdala also exerts a strong influence over hippocampal activity. The hippocampus is considered to be particularly susceptible to chronic stress because it is heavily enriched with corticosteroid receptors and it participates in the termination of stress responses via the glucocorticoid-mediated negative feedback of the

hypothalamus-pituitary-adrenocortical (HPA) axis. Because this medial temporal lobe structure is crucial for declarative-explicit memory in humans and spatial-relational memory in rodents, evidence is emerging that stress generally impairs hippocampal-dependent memory tasks in both humans and rats. Lesions of the BLA attenuate the induction of long-term potentiation (LTP) in the dentate gyrus, while BLA stimulation significantly facilitates LTP induction. Recent findings indicate that amygdalar activity is crucially involved in the emergence of stress-induced impairments in hippocampal LTP and hippocampal-dependent spatial memory in rats. The critical period of the amygdalar contribution to stress effects on hippocampal functions was determined by applying muscimol either before or immediately after stress. Results indicate that intra-amygdalar muscimol infusions before restraint tailshock stress effectively blocked stress-induced physiological and behavioral effects. Thus, the amygdala plays a critical role in the induction of cortical- and hippocampal-dependent processes thought to underlie learning and memory.

Summary

The CeA, which receives input from most of the amygdala nuclei, projects to regions responsible for autonomic, hormonal, and behavioral responses. In this way, the amygdala is in a position to muster the full arsenal of the body's responses to stress.

Conclusions and Functions

There are numerous brain structures that are essential components of the neural circuitry mediating stress responses. The amygdala is one of the neural hubs integrating these components and stands in a position to both influence and be influenced by the stress response. Animal models of chronic stress reveal that stress-induced structural plasticity in amygdala neurons provides a candidate cellular substrate for affective disorders.

The amygdala receives a plethora of sensory input from both the environment and the body. Much of this input relates to stressful and threatening events. However, the amygdala also receives both simple and highly processed sensory information that can serve a signaling function. This suggests that the amygdala may integrate these sources of information. An early suggestion was that the amygdala recognizes signals for events with emotional significance. This is perhaps most clear in fear conditioning, in which cues that predict an event come to produce a number of fear-related responses. These behaviors are lost in rats, mice, rabbits, and humans, without an

amygdala. There is an LTP of amygdala responses to stimulation of its thalamic, hippocampal, and cortical afferents. Cells in the LA receive both nociceptive somatosensory information and auditory information, and exposure to fear conditioning itself can potentiate the response of these cells. In this manner, stimuli that predict stress can come to activate the amygdala, even if they are not stressful in their own right. Thus, neural plasticity within the amygdala provides the ability for this structure to anticipate stress through learning processes such as Pavlovian fear conditioning. Consistent with such a view, the amygdala projects to most of the brain areas responsible for rapid coping with impending danger.

Under circumstances in which prolonged stress disrupts cognition that is mediated by the hippocampus, it facilitates amygdala-dependent cognition. More specifically, under circumstances in which chronic stress induces dendritic atrophy and debranching in CA3 pyramidal neurons of the hippocampus, it induces hypertrophy of pyramidal and stellate neurons in the LA, the BLA, and the bed nucleus of stria terminalis, an amygdaloid projection site central to anxiety. Chronic unpredictable stress, however, had little effect on pyramidal and stellate neurons and induced atrophy only in BLA bipolar neurons. Thus, stress hormones released as a result of stress-induced amygdala activity can strengthen the excitatory drive within the BLA and thereby influence subsequent information processing by the amygdala and its downstream targets. This suggests that chronic stress could lead to an imbalance in HPA axis function through a gradual gain in excitatory control exerted by the amygdala. In this way, a system that evolved to deal with immediate threats of great evolutionary urgency also has the ability to cause debilitating pathology. The amygdala allows organisms to better cope with future stress by modulating cortical, hippocampal, and striatal circuits involved in declarative memory and motor learning. Through such anticipatory mechanisms, the amygdala provides for the adaptive feedforward regulation of stress. But the amygdala's powerful and adaptive role in short-term emergencies may also leave it open to a role in pathology accompanying chronic stress.

See Also the Following Articles

Brain and Brain Regions; Defensive Behaviors; Emotions: Structure and Adaptive Functions; Fear; Fear and the Amygdala.

Further Reading

- Aggleton, J. P. (2000). *The amygdala: a functional analysis* (2nd edn). New York: Oxford University Press Inc.
- Bechara, A., Tranel, D., Damasio, H., et al. (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science* 269(5227), 1115–1118.
- Cahill, L. and McGaugh, J. L. (1998). Mechanisms of emotional arousal and lasting declarative memory. *Trends in Neuroscience* 21(7), 294–299.
- Fanselow, M. S. (1994). Neural organization of the defensive behavior system responsible for fear. *Psychonomic Bulletin and Review* 1, 429–438.
- Fendt, M. and Fanselow, M. S. (1999). The neuroanatomical and neurochemical basis of conditioned fear. *Neuroscience and Biobehavioral Reviews* 23, 743–760.
- Kim, J. J., Koo, J. W., Lee, H. J. and Han, J. S. (2005). Amygdalar inactivation blocks stress-induced impairments in hippocampal long-term potentiation and spatial memory. *Journal of Neuroscience* 25(6), 1532–1539.
- LeDoux, J. E. (1996). *The emotional brain: the mysterious underpinnings of emotional life*. New York: Simon & Schuster.
- Maren, S. and Fanselow, M. S. (1995). Synaptic plasticity in the basolateral amygdala induced by hippocampal formation stimulation in vivo. *Journal of Neuroscience* 15(11), 7548–7564.
- Maren, S. and Fanselow, M. S. (1996). The amygdala and fear conditioning: has the nut been cracked? *Neuron* 16(2), 237–240.
- Pare, D., Quirk, G. J. and Ledoux, J. E. (2004). New vistas on amygdala networks in conditioned fear. *Journal of Neurophysiology* 92(1), 1–9.
- Rizvi, T. A., Ennis, M., Behbehani, M. M. and Shipley, M. T. (1991). Connections between the central nucleus of the amygdala and the midbrain periaqueductal gray: topography and reciprocity. *Journal of Comparative Neurology* 303(1), 121–131.
- Roozendaal, B., Koolhaas, J. M. and Bohus, B. (1991). Attenuated cardiovascular, neuroendocrine, and behavioral responses after a single footshock in central amygdaloid lesioned male rats. *Physiology Behavior* 50(4), 771–775.
- Shi, C. and Davis, M. (1999). Pain pathways involved in fear conditioning measured with fear-potentiated startle: lesion studies. *Journal of Neuroscience* 19(1), 420–430.
- Vyas, A., Mitra, R., Shankaranarayana Rao, B. S. and Chattarji, S. (2002). Chronic stress induces contrasting patterns of dendritic remodeling in hippocampal and amygdaloid neurons. *Journal of Neuroscience* 22(15), 6810–6818.
- Walker, D. L., Toufexis, D. J. and Davis, M. (2003). Role of the bed nucleus of the stria terminalis versus the amygdala in fear, stress, and anxiety. *European Journal of Pharmacology* 463(1-3), 199–216.

Anatomy of the HPA Axis

A G Watts

University of Southern California, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
A G Watts, volume 2, pp 477–483, © 2000, Elsevier Inc.

Introduction

Hypothalamus: The Neuroendocrine Paraventricular Nucleus

Neurohypophysis: Median Eminence and Infundibular Stalk

Adenohypophysis: Corticotropes in the Pars Distalis

Adrenal Cortex: Zona Fasciculata

Glossary

<i>Acidophils</i>	Cells that stain with acidic (anionic) histochemical stains.
<i>Basophils</i>	Cells that stain with basic (cationic) histochemical stains.
<i>Chromophobes</i>	Cells that do not stain with either acidic or basic histochemical stains.
<i>Fenestrated capillary</i>	A specialized capillary that has transcellular circular openings in the endothelial walls allowing the passage of large molecules or cells in or out of the capillary. They are particularly prevalent in regions where secretory products are released into the vasculature.
<i>Magnocellular neurons</i>	Neurons of relatively large size.
<i>Mesoderm</i>	A principal layer of the mammalian embryo from which derives the vascular tissue.
<i>Neuroectoderm</i>	A principal layer of the mammalian embryo from which derives the nervous tissue.
<i>Parvicellular neurons</i>	Neurons of relatively small size.
<i>Telencephalon</i>	The part of the forebrain containing the cortical regions, basal ganglia, amygdala, and septal nuclei.

Introduction

The hypothalamic-pituitary-adrenal (HPA) axis is a critical endocrine system that contributes to the maintenance of the daily energy balance and forms a fundamental part of the endocrine stress response. It consists of three serially arranged elements connected by vascular links, whose final motor output depends

largely on a feed-forward sequence of humorally directed chemical signals: neuroendocrine (corticotropin releasing hormone, CRH; and in some species arginine vasopressin, AVP) originating from a specific set of neuroendocrine neurons in the hypothalamic paraventricular nucleus (PVN), endocrine (adrenocorticotrophic hormone, ACTH, released from the anterior pituitary corticotropes), and adrenal glucocorticoid (generally cortisol or corticosterone, depending on the species). The impact of various neural and humoral inputs on each of these three elements allows for a complex array of interacting regulatory influences including afferent neural control in the brain, a variety of circulating hormones, and local paracrine interactions within the anterior pituitary and adrenal cortex.

Hypothalamus: The Neuroendocrine Paraventricular Nucleus

Location

The hypothalamic paraventricular nucleus (PVN) is a heterogenous collection of neurons that constitutes one of the principal motor output systems of the hypothalamus. It is a bilateral structure located either side of the third ventricle and is an elaboration of the periventricular zone of the hypothalamus. One set of PVN neurons provides extensive projections to the hindbrain and spinal cord and regulates behavioral and autonomic motor activity. The other set is neuroendocrine in nature and has terminals in the neurohypophysis. In turn, these neuroendocrine neurons can be further divided into two groups. The first are magnocellular and are found in the PVN and supra-optic nuclei. They have terminals in the infundibular process (or neural lobe) and release either AVP (or vasotocin in nonmammalian groups) or oxytocin into the general vasculature. The second are parvicellular and are generally located (at least in the rat) more medially than the magnocellular neuroendocrine neurons in the PVN. They have terminals in the median eminence (ME) and release release-regulating hormones into the hypophyseal portal vasculature to regulate anterior pituitary function.

Subdivisions

The various subdivisions of the PVN of most species are not easily differentiated into discernible cytoarchitectonic compartments. However, in the rat (in which much of the work on neural compartmentalization

has been performed) it is possible to delineate approximately 10 separate subdivisions. Of these, the dorsal aspect of the medial parvicellular (mpd) part contains the CRH neuroendocrine neurons responsible for generating the adrenocortical stress response. CRH stimulates the release of ACTH both *in vivo* and *in vitro*, and lesions of the PVN or its projections to the ME have convincingly demonstrated that these cells are the source of CRH secreted into the hypophyseal portal vasculature.

Corticotropin Releasing Hormone Neuroendocrine Motor Neurons

In the rat, there are approximately 2000 CRH neurons on each side of the PVNmpd (Figure 1a). The cell bodies of these neurons are typically fusiform and are approximately 10 μm in diameter. They have a rather simple dendritic arbor, usually consisting of two, sometimes three, relatively thick dendrites located in a bipolar arrangement (Figure 1b). These dendrites are significantly shorter than those found on neurons with descending projections. The dendrites of CRH neurons generally remain within the nucleus, but in some instances they extend into the magnocellular parts of the PVN. Their axons emerge laterally from the PVN (Figure 1a), eventually curving around the dorsal aspect of the fornix, before turning ventrally and then caudally to terminate in the portal capillary plexus located in the external zone of the ME.

In many species, AVP is also an important regulator of ACTH secretion and is released into the hypophyseal portal vasculature. In the rat, AVP is synthesized in approximately half of the neuroendocrine CRH neurons in the PVN, which are generally more numerous in the dorsal part of the PVNmpd. AVP appears to play a key role in modulating the actions of CRH on anterior pituitary corticotropes, although the extent of its contribution varies among species. Parvicellular AVP is exquisitely sensitive to corticosterone feedback and is copackaged with CRH in dense core vesicles, but it is regulated differentially from CRH by glucocorticoids and a variety of stressors. Although a number of studies have attempted to show that AVP from axons originating in magnocellular cells contribute to ACTH secretagogue release in the external layer of the ME, the precise contribution of this magnocellular system to the regulation of ACTH still remains to be clarified.

Afferent Input

Coordinating the activity of CRH neurons in the PVNmpd in response to stress depends on sets of afferent inputs from many regions of the brain. In general, these inputs can be divided into two groups:

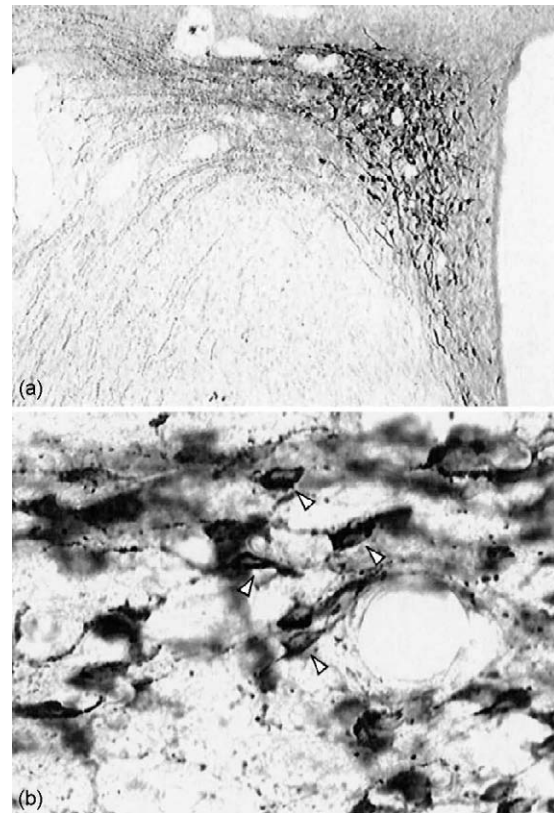


Figure 1 Paraventricular nucleus (PVN) of an adrenalectomized rat and CRH. a, Low-power photomicrograph of PVN stained immunohistochemically for CRH; b, high-power photomicrograph of PVN neuroendocrine CRH neurons. In a, note the densely stained neurons in the medial parvicellular part of the PVN and their axons exiting the nucleus in a lateral direction. The medially located third ventricle is to the right. In b, note the characteristic fusiform cell bodies (white arrowheads) with large unstained nuclei, and the generally bipolar location of the dendrites. Photomicrographs from the author's laboratory.

those that target PVNmpd neurons directly and those that are channeled through intermediate regions such as the bed nuclei of the stria terminalis (BST) and peri-PVN regions.

Many of the inputs that directly target CRH neurons originate from the hindbrain, some regions of the hypothalamus, and circumventricular organs in the forebrain. Direct catecholaminergic inputs into the PVN from the hindbrain travel in the ventral noradrenergic bundle and play a particularly prominent role in conveying visceral sensory information to neuroendocrine neurons. Noradrenergic innervation to the parvicellular neuroendocrine PVN originates predominantly from the A2 cell group in the medial and commissural divisions of the nucleus of the solitary tract. The A2 cell group receives first-order vagal and glossopharyngeal afferents and therefore can also relay visceral sensory information to the PVN. A minor contribution arises from the A1 and A6

(locus ceruleus) cell groups. Most of the adrenergic fibers (which also contain neuropeptide Y, NPY) arise in the C1 cell groups. The hypothalamic cell groups that project directly into the PVNmpd include the dorsomedial nucleus, some preoptic nuclei, and some peri-PVN regions. The neurochemical nature of these projections is still under investigation, but both γ -aminobutyric acid (GABA) and glutamate are likely to be important transmitters in these projections.

The subfornical organ (SFO) and vascular organ of the lamina terminalis (OVLT) are circumventricular organs that lack a blood–brain barrier and can provide an entry point for humoral factors that are ordinarily excluded from the central nervous system (CNS). Both of these midline structures are located in the forebrain and provide efferent projections directly into the PVN. They are thought to be particularly important for mediating the HPA response to increased circulating angiotensin II (particularly the SFO) and increases in vascular osmolality (the OVLT). In addition, direct projections from a third cell group in the rostral hypothalamus, the median preoptic nucleus, may also contribute in this function.

Information from telencephalic regions, including the amygdala, hippocampus, and lateral septal nuclei, probably play an prominent role in mediating the more cognitive influences on the activity of the HPA axis. However, these particular regions do not provide direct projections into the PVN, and their impact on neuroendocrine CRH neurons and ACTH secretion are thought to be channeled more through projections from a variety of nuclei in the BST. There is evidence that GABA plays an important role in this function.

Neurohypophysis: Median Eminence and Infundibular Stalk

The defining feature of neuroendocrine neurons is that they release their chemical signals into fenestrated capillaries in the neurohypophysis rather than into the characteristic synaptic cleft found in neuron-to-neuron transmission. Once in the vasculature, these signals are then carried to their targets by the blood flow and act on targets that are effectively defined by the presence of specific receptors.

Structure

The neurohypophysis is a midline structure that derives from an outgrowth from the floor of the hypothalamus. It consists of the ME, the infundibular stalk, and the infundibular process (the neural lobe or pars nervosa) of the pituitary gland (**Figure 2**). Of these structures, the ME and infundibular stalk

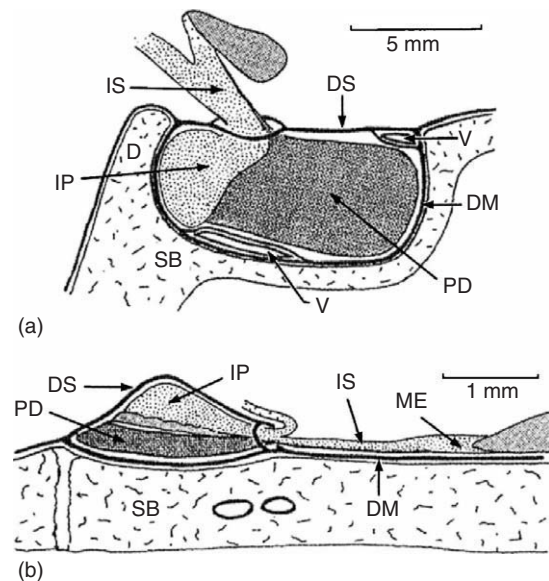


Figure 2 Diagrams based on midsagittal sections of pituitary glands. a, Human; b, rat. Anterior/rostral is to the right. D, dorsum sellae; DM, dura mater; DS, diaphragm sellae; IP, infundibular process of the neurohypophysis; IS, infundibular stem of the neurohypophysis; ME, median eminence of the neurohypophysis; PD, pars distalis; SB, sphenoid bone; V, venous sinus. Adapted from Daniel, P. M. and Pritchard, M. M. L. (1975), *Studies of the hypothalamus and the pituitary gland, Acta Endocrinologia* **30** (supplement 201), p. 45.

are critical parts of the HPA axis. The ME is where axons from neuroendocrine CRH neurons in the PVNmpd have their terminals adjacent to the primary capillaries of hypophysial portal system. It is a highly complex structure consisting of a looped network of fenestrated capillaries, axon terminals, and tanyocytes, which are specialized ependymal cells that have processes extending from the third ventricle to the axon terminals on the capillary plexus. In the case of the CRH neurons that regulate anterior pituitary corticotrope function, secretagogues are released in an activity-dependent manner into the blood flowing through the hypophysial portal vasculature toward the pituitary gland from neuroendocrine terminals in the ME.

Organization of Blood Supply

The organization of the blood supply to the rat pituitary gland is shown in **Figure 3**. The arterial blood supply of the ME and infundibular stalk arises from the internal carotid artery and the circle of Willis, and it flows into the capillary bed of the ME from branches of the anterior hypophysial artery. These vessels form extensive capillary loops in the external part of the ME close to the terminals of the CRH neuroendocrine motor neurons. As they leave the ME, these capillaries drain into the long portal vessels

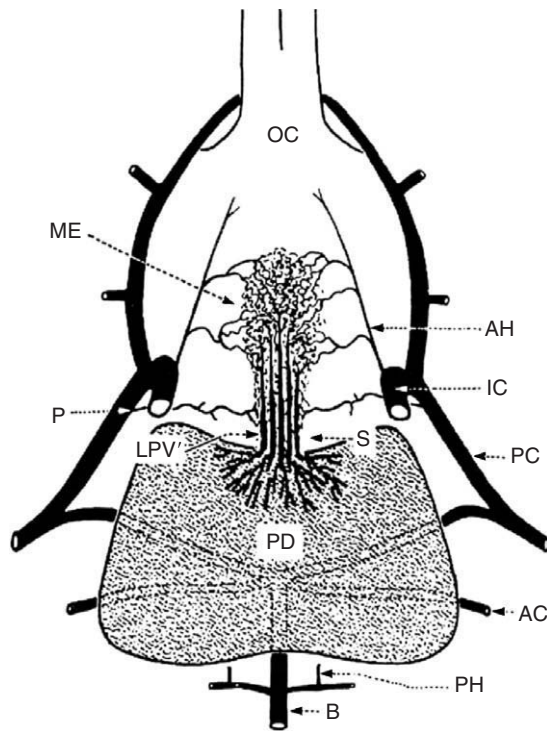


Figure 3 The pituitary gland of the rat showing the main features of the blood supply to its ventral aspect. AC, anterior cerebral artery; AH, anterior hypophyseal artery; B, basilar artery; IC, internal carotid artery; LPV, long portal vessel; ME, median eminence; OC, optic chiasm; P, peduncular artery; PC, posterior communicating artery; PD, pars distalis; PH, posterior hypophyseal artery; S, pituitary stalk. Adapted from Daniel, P. M. and Pritchard, M. M. L. (1975), *Studies of the hypothalamus and the pituitary gland*, *Acta Endocrinologia* 30(supplement 201, p. 56.)

that lie parallel in the infundibular stalk and then course into the pars distalis. On entering the pars distalis, the long portal vessels anastomose, allowing the release-regulating hormones access to its secretory cells. A second capillary bed is found closer the pituitary gland. It is supplied by blood from the artery of the trabecula and produces the short portal vessels, which then enter the pars distalis in parallel to the long portal vessels.

Adenohypophysis: Corticotropes in the Pars Distalis

The adenohypophysis is nonneural tissue that develops from an ectodermal evagination of the roof of the buccal cavity, Rathke's pouch. It has three parts: pars distalis, tuberalis, and intermedia. The pars tuberalis makes up approximately 75% of the adenohypophysis (Figure 2) and contains the corticotropes that synthesize ACTH from its precursor proopiomelanocortin (POMC). Pars distalis cells are arranged in cords or clusters interspersed between capillaries

from the hypophyseal portal vessels. Corticotropes are separated from one another by thin-walled vascular sinusoids and are supported by a skeleton of reticular tissue. Classically, cells of the pars distalis have been characterized into three groups based on their stainability: acidophils, basophils, and chromophobes. Although these categories are not as clear cut as was originally thought, corticotropes are basophilic and show strong periodic acid Schiff (PAS) staining. CRH and AVP released into the hypophyseal portal vasculature eventually bathe corticotropes and stimulate the synthesis and secretion of ACTH. In terms of their response to these secretagogues, there appears to be three type of corticotropes: those responsive to CRH and not AVP; those responsive to CRH or AVP; and those responsive to CRH and AVP. Thus, changes in the CRH/AVP ratio could affect the overall ACTH secretory pattern.

Adrenal Cortex: Zona Fasciculata

Structure

The mammalian adrenal glands are bilateral structures located in perirenal adipose tissue found dorsal to the kidney. They each consist of cortical steroid-synthesizing tissue derived from the mesoderm and medullary catecholamine-synthesizing tissue derived from the neuroectoderm (Figure 4). The thin outermost part of the adrenal cortex, the zona glomerulosa, is the main site of mineralocorticoid production. Glucocorticoids (either cortisol or corticosterone) are primarily produced by the zona fasciculata in response to circulating ACTH (Figure 4), and this constitutes the largest amount of steroid-synthesizing tissue in the adrenal gland. The androgens dehydroepiandrosterone and androstenedione are produced in the innermost zona reticularis. The zona fasciculata consists of long thin cords, each of one or two cells separated by parallel fenestrated venous capillaries. These cells are approximately 20 μm in diameter and contain the large amounts of smooth endoplasmic reticulum, mitochondria, and lipid droplets characteristic of steroid-synthesizing cells. They also possess reasonable amounts of rough endoplasmic reticulum, as well as greater numbers of gap junctions than are found between zona glomerulosa cells.

Blood Supply

The blood supply to the adrenal cortex is ultimately derived from three groups of arteries: the superior (arising from the inferior phrenic artery), the middle (arising from the aorta), and the inferior (arising from the renal artery or aorta) adrenal arteries. These arteries then branch to form a plexus in the adrenal

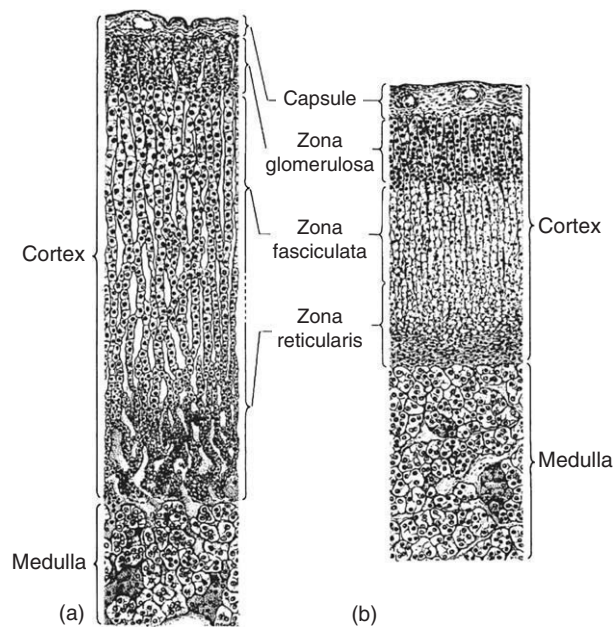


Figure 4 Sections through the adrenal glands of rats. a, Normal rat; b, hypophysectomized rat. Because the functional capacity of the adrenal cortex is maintained by circulating ACTH, hypophysectomy results in tremendous shrinkage of the cortex. The medulla (catecholamine-synthesizing tissue) is not affected by hypophysectomy. Both sections are drawn to scale. Adapted from Turner, C. D. and Bagnara, J. T. (1976), *General endocrinology*, Philadelphia: W. B. Saunders.

capsule (which constitutes the outside of the gland), from which arises the cortical arteries supplying the three layers of the adrenal cortex. Venous outflow from the cortex enters a number of sinuses in the adrenal medulla and eventually join the adrenal vein. In the human, the left adrenal vein joins the renal vein and the right adrenal vein enters the vena cava.

Innervation

The mammalian adrenal cortex is innervated by postganglionic sympathetic fibers in the greater and lesser splanchnic nerves originating in thoracic regions T8 to T11 of the vertebral column. This postganglionic innervation is most likely derived from pre- and para-vertebral ganglia, including the suprarenal ganglion.

Moreover, the cortex also receives input from adrenal medullary ganglion cells, emphasizing the complexity of the local regulation that can occur within the adrenal gland. Nerve terminals have been found adjacent to adrenal parenchyma and blood vessels in rat and human adrenal cortices. Although the function of preganglionic innervation of the cortex is not entirely clear, there are experimental reports documenting that in the rat the splanchnic nerve plays a role in regulating the prefeeding increase in corticosterone secretion. Although vagal sensory innervation of the adrenal is present, parasympathetic motor innervation of the adrenal is not as clearly defined.

See Also the Following Articles

Adrenal Cortex; Hypothalamic-Pituitary-Adrenal; Genetic Variation of HPA Axis Activity and Function in Farm Animals; HPA Alterations in PTSD; Neuroendocrine Systems.

Further Reading

- Black, V. A. (1988). The adrenal gland. In: Weiss, L. (ed.) *Cell and tissue biology. A textbook of histology*, pp. 1035–1066. Baltimore, MD: Urban & Schwarzenberg.
- Daniel, P. M. and Pritchard, M. M. L. (1975). Studies of the hypothalamus and the pituitary gland. *Acta Endocrinologia* 30(supplement 201), pp. 45, 56.
- Halmi, N. S. (1988). The hypophysis. In: Weiss, L. (ed.) *Cell and tissue biology. A textbook of histology*, pp. 973–994. Baltimore, MD: Urban & Schwarzenberg.
- Herman, J. P., Figueiredo, H., Mueller, N. K., et al. (2003). Central mechanisms of stress integration: hierarchical circuitry controlling hypothalamo-pituitary-adrenocortical responsiveness. *Frontiers in Neuroendocrinology* 24, 151–180.
- Swanson, L. W. and Sawchenko, P. E. (1983). Hypothalamic integration: organization of the paraventricular and supraoptic nuclei. *Annual Review of Neuroscience* 6, 275–325.
- Turner, C. D. and Bagnara, J. T. (1976). *General endocrinology*. Philadelphia: W. B. Saunders.
- Watts, A. G. (2005). Glucocorticoid regulation of peptide genes in neuroendocrine CRH neurons: a complexity beyond negative feedback. *Frontiers in Neuroendocrinology* 26, 109–130.

Androgen Action

R J Handa and T R Pak

Colorado State University, Ft. Collins, CO, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R J Handa and R H Price, Jr., volume 1, pp 183–187, © 2000, Elsevier Inc.

Androgens

The Androgen Receptor

Physiological Actions of Androgens

Androgen-Associated Pathology

Interactions between Stress and Androgens

Summary

Glossary

<i>Concentration gradient</i>	A region consisting of a high concentration of a compound adjacent to an area of low concentration of that same compound.
<i>Coregulatory protein</i>	A protein that interacts with a transcription factor and the transcriptional machinery. Coregulatory proteins have a profound effect on transcription rate, either increasing (coactivator) or decreasing (corepressor) it.
<i>Diurnal rhythm</i>	A biological rhythm based on the 24-h day.
<i>Glucocorticoid</i>	A form of steroid hormone that is produced primarily by the adrenal glands and is released following stress.
<i>Lipophilic</i>	Affinity for fatty substances.
<i>Luteinizing hormone (LH)</i>	A peptide hormone that is released from the pituitary gland and promotes testosterone synthesis from within the testes.
<i>Promoter</i>	The regulatory region of a gene. Often promoters are upstream of the protein-encoding regions of a gene.
<i>RNA polymerase holoenzyme</i>	The complex of proteins that are directly involved in transcribing a gene.
<i>Steroid hormones</i>	A class of hormones that are derived from cholesterol.
<i>Transcriptional</i>	Pertaining to the process of synthesizing mRNA from a DNA template.
<i>Transcription initiation site</i>	The region of a gene to which the mRNA polymerase holoenzyme binds to begin transcription.

Androgens

Major Forms of Androgens

The term androgen is derived from the Greek roots *andro* (man) and *gennan* (to produce). The biological definition of an androgen is any substance that specifically promotes the growth of the male gonads. There are four major forms of circulating androgens. Testosterone (T) is the predominant form in males and is analogous to estrogen in females. Other common forms of androgens are dihydrotestosterone (DHT), androstenedione, dehydroepiandrosterone (DHEA), and its sulfonlated derivative, DHEAS. The most potent androgens are T and DHT, while DHEA/S and androstenedione are weak androgens. Both males and females produce all forms of androgens, though in differing amounts. The majority of androgens produced by males is T, which circulates at 10-fold higher levels than DHT. In females, the predominant forms of circulating androgens are DHEA/S followed by androstenedione and finally T. Overall, androgen levels are roughly 10-fold higher in males than in females. Androgen levels do, however, undergo a moderate daily variation in what is called a diurnal rhythm. In men, highest levels of testosterone are found in the bloodstream during the early morning hours, whereas low levels are found in the blood during the evening. In rodents, a nocturnal species, testosterone levels peak in the first few hours after nightfall. In women, cyclic variations in testosterone have been reported, with peak levels occurring during the preovulatory phase of the menstrual cycle.

Origins

In males, testosterone is produced largely by Leydig cells in the testes. In females, T is produced by the interstitial and stromal thecal cells in the ovaries. In addition to the gonadally derived testosterone that circulates at high levels in males, a small proportion is synthesized in target tissues from the testosterone precursors androstenedione and DHEA/S. These often circulate at relatively high levels in males but act only very weakly on their own. Dihydrotestosterone is synthesized to a limited extent in the gonads; the majority is locally converted from T in many androgen target tissues. Androstenedione and DHEA/S are largely produced by the adrenal cortex. Recent studies

have also demonstrated that the brain has the appropriate enzymes to produce androgens *de novo*. The extent to which this synthesis affects physiology remains to be determined. Androgen synthesis and metabolism can occur anywhere the enzymes that catalyze such reactions are present. Though the primary tissues of origin are presented above, precursors and metabolites are produced in many tissues.

Synthesis Pathways

The biosynthetic pathway that produces endogenous androgens is common to many steroid hormones and begins with the conversion of cholesterol to pregnenolone in the mitochondria of steroid-synthesizing cells. Cholesterol is composed of 27 carbon atoms, including a 17-carbon sterol backbone with an 8-carbon side chain. The major androgens including testosterone have no side chain and are composed of only 19 carbon atoms.

The enzymes that catalyze the steroid synthesis pathway belong to a family of iron-coordinating proteins called cytochrome P450 enzymes. The steps in the testosterone synthesis pathway are briefly described here and are detailed in [Figure 1](#). Regulation of testosterone synthesis begins at the level of cholesterol transport from cellular storage sites to the inner mitochondrial membrane by the steroidogenic acute regulatory protein. Once in the mitochondrion, the P450 side chain cleavage enzyme (P450_{scc} or CYP11A) catalyzes the oxidation and removal of the six carbon atoms that comprise the majority of the side chain on cholesterol. The product, pregnenolone, then diffuses passively out of mitochondria to the smooth endoplasmic reticulum. Here the actions of the P450_{c17} (CYP17) enzyme further shorten the side chain, removing all carbons, leaving only a ketone at the C17 position. The final two steps in testosterone biosynthesis are catalyzed by dehydrogenases, which

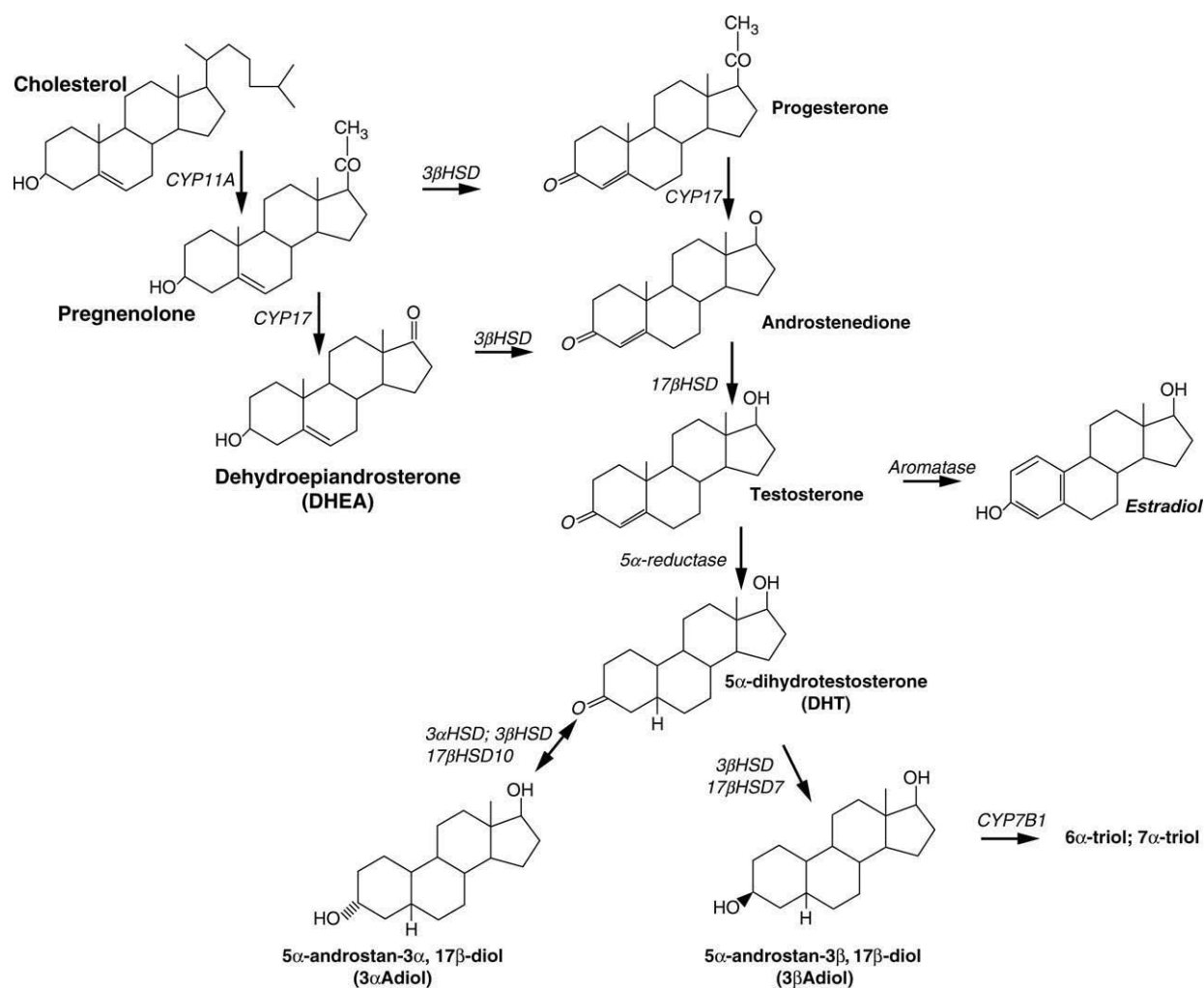


Figure 1 Schematic diagram depicting the biosynthetic pathway of androgens from cholesterol. Enzymes responsible for catalysis are designated in italics. The names of steroid molecules are shown in bold.

have specificity for the C3- and C17-bound oxygens, respectively. The dehydrogenase that converts the C3 ketone to a hydroxyl group also isomerizes the C5 double bond to a C4 double bond.

Metabolites of Testosterone

In target tissues, testosterone can be converted by further enzyme catalysis. Aromatase can convert testosterone to estrogen by oxidoreduction of the C3 ketone and removal of the C10 methyl group, which allows aromatization of the A ring in the sterol backbone. This leads to a product that has very different effects than the parent compound. Testosterone can also be converted by the enzyme 5α -reductase to DHT, which is a more potent androgen than testosterone. The primary site of DHT formation is in peripheral target tissues such as the prostate, liver, and skin. The local metabolism of testosterone to DHT allows the androgenic signal to be amplified in a tissue-specific manner. Excessive production of DHT has been shown to be a factor in prostate abnormalities, such as benign prostatic hyperplasia (BPH). The pharmacological inhibition of 5α -reductase, such as with the compound finasteride, has proven to be an effective therapeutic agent.

DHT concentrations in androgen-dependent tissues are maintained through the direct formation of DHT from circulating testosterone and from the interconversion of DHT and 3α -diol, a primary metabolite of DHT. This bidirectional catalysis is accomplished by the actions of the oxidoreductase 3α -hydroxysteroid dehydrogenase (3α HSD) and, to a lesser extent, 3β HSD and 17β HSD10. A variety of enzymes also convert DHT to 3β -diol, including 3β HSD, 3α HSD, and 17β HSD7; however, this process is irreversible. While DHT is a pure androgen, binding and activating only androgen receptors (AR), the metabolite 3β -diol has very low affinity for the AR and preferentially binds the beta form of estrogen receptors. In the prostate, 3β -diol has been implicated in the regulation of epithelial cell proliferation. 3β -diol is ultimately converted to the inactive compounds 6α - or 7α -triol and excreted in the urine.

Secretion and Transport

Steroid hormones are lipophilic and thus can easily pass through cell membranes. For this reason, they are secreted along a concentration gradient from synthetic cells to the circulating plasma and do not need to be secreted by a vesicular membrane fusion pathway. Consequently, circulating levels of androgens accurately reflect rates of synthesis. There are, however, circulating plasma-binding proteins known as

the steroid hormone-binding globulins (SHBGs) that bind to testosterone and related compounds and transport them throughout the organism. Nonspecific interactions with other plasma proteins such as serum albumin also occur. The SHBGs can protect the steroids from degradative enzymes that shorten their half-life and can prevent renal excretion. Once at a target tissue, steroid hormones are able to easily enter cells by diffusing across the plasma membrane. Once in the cell, steroid hormones are bound by intracellular receptors. In the case of androgens such as T and DHT, the specific receptor has been termed the androgen receptor (AR).

The Androgen Receptor

Molecular Organization

Androgen receptors are dynamic proteins that are members of a superfamily of nuclear receptor proteins. This superfamily has a common structure composed of functional domains, depicted schematically in [Figure 2a](#). The domains include a hormone-binding domain near the C-terminus and a DNA-binding domain near the center of the linearized protein. The other domains include a hinge region to facilitate conformational change and a transactivation domain, which is important for interaction of the steroid receptor with transcriptional machinery. In the absence of hormone, the androgen receptor is maintained in an inactive state by a complex of chaperone proteins called heat shock proteins.

Function

Classically, the AR is a nuclear transcription factor and modulates gene expression in a hormone-dependent manner. In this respect, hormone binding induces a conformational change that dissociates the heat shock protein from the receptor, thus activating the receptor. The activation pathway of ARs by hormone is depicted in [Figure 2b](#). The conformational change sequesters the ligand-binding domain with hormone attached into a hydrophobic pocket and reveals the DNA-binding domain, which has high affinity for certain DNA sequences called hormone response elements (HREs). Upon activation, other regions are revealed that allow dimerization with other hormone-bound androgen receptors. The activated dimers then are able to bind to the HREs, a palindromic sequence of nucleotides in DNA. These HREs are not present in the promoter region of all genes, only those that are androgen responsive. Examples of such genes are the cytokine, interleukin-6, and the androgen receptor itself.

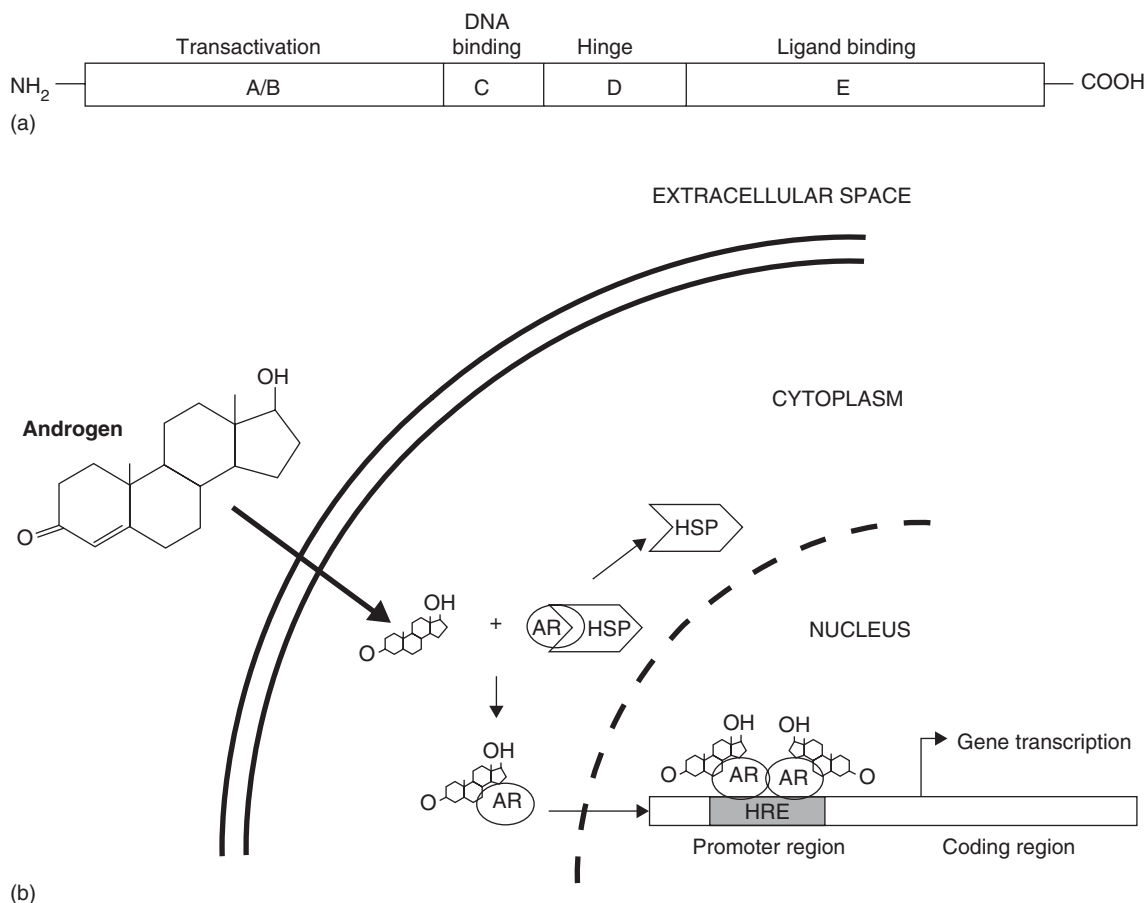


Figure 2 Schematic diagram of (a) androgen receptor (AR) functional domains and (b) pathway for activation of gene transcription. HRE, hormone response element; HSP, heat-shock protein. Cell membrane is depicted as a broad bar separating the cytoplasm from the extracellular space. Dashed line indicates nuclear membrane. DNA of androgen-responsive gene is shown as a rectangle in the nucleus.

Coregulatory Proteins

The control of gene expression by ARs is also dependent on factors such as the appropriate promoter context and the cell-specific presence of coregulatory proteins. Once the activated AR is bound to the HRE, for gene expression to occur it still must communicate on a molecular level with the RNA polymerase holo-enzyme, which occupies the transcription initiation site. It is generally accepted that this contact is not direct, but occurs through coregulatory proteins that do not bind to DNA directly but act through protein-protein contacts.

Nongenomic Androgen Actions

Hormone-dependent transcription is the primary mechanism of action for ARs; however, a growing body of evidence suggests that androgens can act rapidly, within seconds to minutes, to affect cell function, suggesting the presence of a nongenomic pathway. These nongenomic effects involve the activation of second

messenger signaling cascades that result in the phosphorylation of a variety of transcription factors. Historically, ARs were thought to be inactive prior to localization in the nucleus; however, the observed rapid effects of androgens have forced the reevaluation of this hypothesis. One possibility is that the AR, following ligand binding and dissociation of heat shock proteins, is able to associate with the plasma membrane through interactions with G-protein-coupled receptors. In addition, androgens can bind directly to the SHBG receptor, also localized in the plasma membrane, and activate the mitogen-activated kinase second messenger system. However, to date a true membrane androgen receptor has not been identified.

Physiological Actions of Androgens

The specific set of genes that is regulated by androgens leads to diverse physiological functions. In general, growth-associated processes are involved, such as those in prostate, hair, skin, muscle, and

bone. Typical examples of androgen-dependent functions are those associated with puberty in males. Increasing levels of T during the mid-teen years lead to a deepening of the voice, growth of facial and axillary hair, activation of sebaceous glands, coarsening of the skin, and growth of long bones and muscle. Androgens also regulate spermatogenesis, a process that is initiated as testicular size increases during puberty. Spermatogenesis persists throughout the life of the male but begins to decline as circulating testosterone levels decline.

Androgen-Associated Pathology

As noted previously, androgens are involved in many anabolic processes. When these growth processes go awry in the testes or prostate, cancerous pathology can result. One of the most common treatments for prostate cancer is to block androgen receptor function with antagonist drugs.

Some of the developmental defects associated with malfunctioning androgen signaling have led to the discovery of specific androgen functions. The most common syndrome associated with a nonfunctional androgen receptor is called androgen insensitivity syndrome or tfm (testicular feminized male). This rare disorder in genetic males results in a phenotypically female individual with breast development, a lack of pubic hair, and formation of feminized genitalia. Another syndrome showing hermaphroditism is a defect in the 5 α -reductase gene. Individuals with this defect cannot produce DHT and thus have ambiguous genitalia until puberty, when testosterone levels rise and partially compensate for the DHT deficiency.

Interactions between Stress and Androgens

Studies using animal models have shown that stress can affect circulating levels of testosterone, and testosterone can influence responses to stress. Observations of baboons in their natural habitat have shown that in response to a stressful stimulus, testosterone levels respond differently depending on the animal's rank in a dominance hierarchy. Dominant males show an increase in testosterone levels following stress, while subordinate males' testosterone levels

decrease. The reason for this differential response appears to be related to the animal's normal circulating level of glucocorticoids. The subordinate males are in a state of chronic stress based on an elevated basal level of circulating glucocorticoids. In contrast, dominant males have low basal levels of circulating glucocorticoids. In response to an acute stressor, both dominant and subordinate males show rapid increases in circulating glucocorticoid levels and decreases in luteinizing hormone secretion. However, dominant males appear to be less sensitive to the inhibitory effects of glucocorticoids on testosterone production. This differential response of testosterone levels to stressful stimuli promotes the reproductive fitness of the dominant males while reducing that of subordinate males.

Summary

The androgens are best known for their functions in reproduction and anabolic growth, though they have been shown to be important for other physiological functions such as regulating stress responses and neural development. Ongoing research is attempting to more precisely identify the role of androgens in cognitive functioning.

See Also the Following Articles

Aggressive Behavior; Cholesterol and Lipoproteins.

Further Reading

- Brown, T. R. (1995). Androgen receptor dysfunction in human androgen insensitivity. *Trends in Endocrinology and Metabolism* 6, 170–175.
- Hiipakka, R. A. and Liao, S. (1998). Molecular mechanism of androgen action. *Trends in Endocrinology and Metabolism* 9, 317–324.
- Lindzey, J., Kumar, M. V., Grossman, M., Young, C. and Tindall, D. J. (1994). Molecular mechanisms of androgen action. *Vitamins and Hormones* 49, 383–432.
- Sapolsky, R. M. (1990). Stress in the wild. *Scientific American* 262, 106–113.
- Sapolsky, R. M. (1997). *The trouble with testosterone and other essays on the biology of the human predicament*. New York: Scribner.
- Wilson, J. D. (1996). Role of dihydrotestosterone in androgen action. *Prostate Supplement* 6, 88–92.

Anger

R W Novaco

University of California, Irvine, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R W Novaco, volume 1, pp 188–194 © 2000, Elsevier Inc.

Anger and Stress

The Experience and Expression of Anger

Anger Physiology

Cognition, Anger, and Threat

Anger Dyscontrol and Regulation

Anger Treatment

Glossary

<i>Aggression</i>	Behavior intended to cause psychological or physical harm to someone or to a surrogate target. The behavior may be verbal or physical, direct or indirect.
<i>Anger</i>	A negatively toned emotion, subjectively experienced as an aroused state of antagonism toward someone or something perceived to be the source of an aversive event.
<i>Anger reactivity</i>	Responding to aversive, threatening, or other stressful stimuli with anger reactions characterized by automaticity of engagement, high intensity, and short latency.
<i>Cathartic effect</i>	The lowering of the probability of aggression as a function of the direct expression of aggression toward an anger instigator. The lowering of arousal associated with such catharsis is more or less immediate and can be reversed by re-instigation.
<i>Escalation of provocation</i>	Incremental increases in the probability of anger and aggression, occurring as reciprocally heightened antagonism in an interpersonal exchange.
<i>Excitation transfer</i>	The carryover of undissipated arousal, originating from some prior source, to a new situation having a new source of arousal, which then heightens the probability of aggression toward that new and more proximate source.
<i>Hostility</i>	An attitudinal disposition of antagonism toward another person or social system. It represents a predisposition to respond with aggression under conditions of perceived threat.
<i>Inhibition</i>	A restraining influence on anger expression. The restraint may be associated with either external or internal factors.

Anger and Stress

Anger is a negatively toned emotion, subjectively experienced as an aroused state of antagonism toward someone or something perceived to be the source of an aversive event. It is triggered or provoked situationally by events that are perceived to constitute deliberate harm-doing by an instigator toward oneself or toward those to whom one is endeared. Provocations usually take the form of insults, unfair treatments, or intended thwartings. Anger is prototypically experienced as a justified response to some wrong that has been done. While anger is situationally triggered by acute, proximal occurrences, it is shaped and facilitated contextually by conditions affecting the cognitive, arousal, and behavioral systems that comprise anger reactions. Anger activation is centrally linked to threat perceptions and survival responding.

As a normal human emotion, anger has considerable adaptive value, although there are sociocultural variations in the acceptability of its expression and the form that such expression takes. In the face of adversity, it can mobilize psychological resources, energize behaviors for corrective action, and facilitate perseverance. Anger serves as a guardian to self-esteem, operates as a means of communicating negative sentiment, potentiates the ability to redress grievances, and boosts determination to overcome obstacles to our happiness and aspirations. Akin to aggressive behavior, anger has functional value for survival.

Despite having multiple adaptive functions, anger also has maladaptive effects on personal and social well-being. Generally, strong physiological arousal impairs the processing of information and lessens cognitive control of behavior. Because heightened physiological arousal is a core component of anger, people are not cognitively proficient when they become angry. Also, because the activation of anger is accompanied by aggressive impulses, anger can motivate harm toward other people, which in turn can produce undesirable consequences for the angered person, from direct retaliation, loss of supportive relationships, or social censure. An angry person is not optimally alert, thoughtful, empathic, prudent, or physically healthy. Being a turbulent emotion ubiquitous in everyday life, anger is now known to be substantially associated with stress-related cardiovascular disorders, both coronary heart disease and essential hypertension. Anger is also a symptom of posttraumatic stress disorder (PTSD), and it has high relevance to the PTSD derivative of violent crime

victimization and especially to combat or war zone exposure.

The Experience and Expression of Anger

There is a duality of psychosocial images associated with anger experience and anger expression. The emotional state is depicted as eruptive, destructive, unbridled, savage, venomous, burning, and consuming but also as energizing, empowering, justifying, signifying, rectifying, and relieving. The metaphors, on the one hand, connote something pressing for expression and utilization, but on the other, imply something requiring containment and control. This duality in psychosocial imagery reflects conflicting intuitions about anger, its expression, and its consequences that abound in ordinary language and are reflected in both scholarly literature and artistic works from the classical period to contemporary times. This Janus-faced character of anger foils attempts to understand it and to therapeutically intervene with recurrently angry individuals.

The facial and skeletal musculature are strongly affected by anger, mobilized by a mixture of adrenaline and noradrenaline hormonal secretions. The face becomes flushed, and the brow muscles move inward and downward, fixing a hard stare on the target. The nostrils flare, and the jaw tends toward clenching. This is an innate pattern of facial expression that can be observed in toddlers. Tension in the skeletal musculature, including raising of the arms and adopting a squared-off stance, are preparatory actions for attack and defense. The muscle tension provides a sense of strength and self-assurance. An impulse to strike out accompanies this subjective feeling of potency. From an evolutionary perspective, our perceptual system has been shaped to detect angry faces rapidly, especially those of angry males.

When people report anger experiences, they most typically give accounts of things that have happened to them. For the most part, they describe events physically and temporally proximate to their anger arousal. As a rule, they provide accounts of provocations ascribed to events in the immediate situation of the anger experience. This fosters the illusion that anger has a discrete external cause. The provocation sources are ordinarily identified as the aversive and deliberate behavior of others; thus, anger is portrayed in the telling as being something about which anger is quite fitting. People are very much inclined to attribute the causes of their anger to the personal, stable, and controllable aspects of another person's behavior.

However, the response to the question "What has made you angry?" hinges on self-observational

proficiencies and is often based on intuitions. Precisely because getting angry involves a loss in self-monitoring capacity, people are neither good nor objective observers when they are angry. When inspecting any particular episode, the immediate causes of the anger are readily identifiable. Far less commonly do people disaggregate their anger experiences into multi-causal origins, some of which may be prior, remote events and ambient circumstances, rather than acute, proximal events. Anger experiences are embedded or nested within an environmental-temporal context. Disturbances that may not have involved anger at the outset leave residues that are not readily recognized but that operate as a lingering backdrop for focal provocations.

Anger is inherently a disposition to respond aggressively, but aggression is not an automatic consequence of anger because aggressive behavior is regulated by inhibitory control mechanisms, engaged by internal and external cues. In this regard, physical constraints, expectations of punishment or retaliation, empathy, consideration of consequences, and prosocial values operate as regulatory controls on aggression. While the experience of anger creates a readiness to respond with aggression, that disposition may be otherwise directed, suppressed, or reconstituted. Thus, the expression of anger is differentiated from its experience.

One aspect of anger that influences the probability of aggression is its degree of intensity. The higher the level of arousal, the stronger the motivation for aggression and the greater the likelihood that inhibitory controls will be overridden. Strong arousal not only impels action, but also impairs cognitive processing of aggression-mitigating information. A person in a state of high anger arousal is perceptually biased toward the confirmation of threat, is less able to attend to threat-discounting elements of the situation, and is not so capable of re-appraising provocation cues as benign. Because anger and aggression occur in a dynamic interactional context, the occurrence of aggression will, in turn, influence the level of anger. Thus, anger reactivity can be seen as a mode of responding characterized by automaticity, high intensity, and short latency.

Important forms of the dynamic interrelation of anger and aggression are the escalation of provocation and the cathartic effect. Escalation involves increases away from equilibrium, whereby succeeding events intensify their own precursors. In the case of anger and aggression, escalation refers to incremental change in their respective probabilities, occurring as reciprocally heightened antagonism in an interpersonal exchange. Anger-elicited aggression may evoke further anger in response, progressively generating justification for retaliation. In contrast,

when physical aggression is deployed by an angry person against the anger instigator and there is no retaliation, anger arousal and further aggression are diminished. This is called the cathartic effect – its conditions should not be confused (as they often are) with those involving aggression by non-angry people, vicarious or observed aggression, or aggression not received by the anger instigator. However, the arousal-reducing effect of aggression carried out by angry people against those who have made them angry is reinforcing of aggressive behavior. This means that when anger is reinstated by a new provocation, the likelihood of aggressive behavior is increased. The cathartic expression of anger, whether through destructive aggression or through verbal communication intended to be constructive, can be understood as an organismic action to restore equilibrium.

An alternative to the deliberate expression of anger is suppression, which is largely a product of inhibitory controls. Anger suppression can be functional in promoting interpersonal or social conciliation. For example, when there is a high probability of impending violence, it is prudent to restrain anger expression to diminish the likelihood of triggering a physical assault. Whether in a domestic, occupational, or street context, anger is adaptively muffled when physical retaliation can be expected or when a cool head is needed to solve a problem. In the short term, suppressing even the verbalization of anger not only may be beneficial interpersonally, but also may lead to more regulated physiological reactivity levels. However, recurrent deployment of anger suppression as a stress-coping style will most likely have deleterious effects on cardiovascular health, as has been found for essential hypertension.

Because anger and aggression are thought to be differentially socialized for males and females, the question of gender differences in the experience and expression of anger arises. It has generally been found that the anger of women is comparable to that of men from the standpoint of experienced intensity. However, the style of anger expression varies by gender, especially according to the context of anger activation and its anticipated consequences. Males are more likely to be angered in a public place or by impersonal triggers, whereas females are more likely to be angered at home or by being let down by someone close to them. Females are more likely to become angered by verbal aggression and insensitive/condescending behavior, and males are more likely to be angered by behavior causing physical harm. Men, when angered, are more inclined to use physical aggression than women, who in turn are more likely to fear aggressive retaliation.

Anger Physiology

A defining condition of anger is physiological arousal, the activation of which has evolutionary roots. For our prehistoric ancestors, when a threat to survival or survival resources was detected, it was advantageous to be mobilized to respond energetically and to sustain effort. The flight-or-fight response refers to this hard-wired physiological mechanism that gets instantaneously triggered to engage survival behavior, to focus attention on the survival threat, and to enable the organism to not succumb to pain. Anger is the emotional complement of the organismic preparation for attack, which also entails the orchestration of signals of attack readiness so as to ward off opponents or to coerce compliance.

The arousal of anger is marked by physiological activation in the cardiovascular, endocrine, and limbic systems, as well as other autonomic and central nervous system areas, and by tension in the skeletal musculature. The autonomic signature of anger corresponds to a mixture of adrenaline and noradrenaline. Autonomic system arousal, especially cardiovascular, has been commonly observed in conjunction with anger by scholars from the classical age (such as Seneca, Aristotle, and Plutarch) to the early behavioral scientists of the nineteenth and twentieth centuries (especially Charles Darwin, William James, G. Stanley Hall, and Walter B. Cannon). Laboratory research has reliably found anger arousal to entail increases in both systolic and diastolic blood pressure and, to a somewhat lesser extent, heart rate. It is differentiated from fear by a stronger increase in diastolic pressure, in muscle electrical activity (measured by electromyogram recordings), and in total peripheral resistance. Associated with this cardiovascular activation is facial flushing, which is often reported by people reflecting on their anger experience and also is measured by face temperature. Indeed, in terms of psychosocial imagery, there is no better metaphor for anger than hot fluid in a container.

Autonomic arousal is primarily engaged through adrenomedullary and adrenocortical hormonal activity. The secretion by the adrenal medulla of the catecholamines adrenaline and noradrenaline and by the adrenal cortex of glucocorticoids provides a sympathetic system effect that mobilizes the body for immediate action (e.g., the release of glucose, stored in the liver and muscles as glycogen). In anger, the catecholamine activation is more strongly noradrenaline than adrenaline (the reverse being the case for fear). The adrenocortical effects, which have longer duration than the adrenomedullary ones, are

mediated by secretions of the pituitary gland, which also influences testosterone levels. The pituitary-adrenocortical and pituitary-gonadal systems are thought to affect readiness or potentiation for anger responding.

The central nervous system structure that has been identified in anger activation is the amygdala, the almond-shaped limbic system component located deep in the temporal lobe. The amygdala is the key site for aversive motivational system. Activation in the corticomedial amygdala is associated with anger and attack priming. Anger has also been linked with left-prefrontal cortical activity, which has typically been associated with positive affect and approach motivation. The central nervous system neurotransmitter serotonin, which is also present in blood platelets, affects anger potentiation, as low levels of this hormone are associated with irritable mood. Serotonin imbalances are related to deficits in the modulation of emotion. While serotonin and other neurotransmitters (noradrenaline and dopamine) are involved in anger activation, the neural structures and circuitry in anger dysregulation remain to be disentangled.

These various physiological mechanisms thus pertain not only to the intensity of anger arousal but also to its duration. Arousal activation eventually decays to baseline levels, but recovery time may be prolonged by exposure to new arousal sources or by rumination. The potency of a provocation may be heightened by the carryover of undissipated excitation from a prior arousal source, which may not have been anger specific (i.e., an otherwise stressful circumstance, such as exposure to bad news, work pressure, or traffic congestion). This excitation transfer of arousal residues facilitates anger, augments its intensity, amplifies blood pressure, and raises the probability of aggression. Residual arousal from unresolved anger events can transfer to future conflicts and further intensify anger reactivity to instigating events. In turn, unexpressed anger is associated with exaggerated and more prolonged cardiovascular responses to a variety of stressful stimuli.

In this regard, a stress framework is highly useful. It is unquestionably the case that physiological arousal is activated by exposure to commonly identified stressors, such as noise, crowding, difficult tasks, and high-pressure job environments filled with time demands or exposure to abrasive interactions. Both acute and prolonged exposure to such conditions may induce physiological activation that decays slowly. Therefore, when someone experiences an event that pulls for the cognitive label anger and this event occurs concurrently with already elevated arousal, the anger system is then more easily engaged.

Cognition, Anger, and Threat

Humans have elaborate neurocognitive systems for detecting threat. We have a neural architecture (especially the limbic system) specialized for the processing of emotion and emotion-cognition interactions, and the amygdala is centrally involved in detecting events as threats. In parallel, higher level cognitive reasoning elaborates this information, in what are termed appraisal processes. The perception of threat is conjoined with sympathetic nervous system activation of autonomic arousal components, such as heart rate, blood pressure, and respiration increases, that prepare the body for emergency action.

In addition to potentiating action, anger in such states of mobilization has adaptive value as a source of information. To others, it communicates perceived wrongdoing, threat of aggression, or intent of reprisal. Such information exchange prior to aggression can facilitate social and interpersonal negotiations toward conflict resolution. For the self, it serves as information for prioritizing and decision making. The intensity of anger, for example, can help focus and maintain attention on relevant goals and help one estimate progress toward those goals. When pressed to make a decision, anger serves as a summary affective cue that can be processed without need for elaborate analysis.

To get angry about something, one must pay attention to it. Anger is often the result of selective attention to cues having high provocation value. A principal function of cognitive systems is to guide behavior, and attention itself is guided by integrated cognitive structures, known as schemas, that incorporate rules about environment-behavior relationships. What receives attention is a product of the cognitive network that assigns meaning to events and the complex stimuli that configure them. Expectations guide attentional search for cues relevant to particular needs or goals. Once a repertoire of anger schemas has been developed, events (e.g., being asked a question by someone) and their characteristics (e.g., the way the question was asked, when it was asked, or who asked it) are encoded or interpreted as having meaning in accord with the pre-existing schema. Because of their survival function, the threat-sensing aspect of anger schemas carries urgent priority and can pre-empt other information processing.

Since the writings of the Stoic philosophers of the classical period, anger has been understood to be strongly determined by personal interpretations of events. The concept of appraisal is that of interpretation, judgment, or meaning embedded in the perception of something – not as a cognitive

event occurring after that something has happened. The appraisal of provocation is in the seeing or hearing. Appraisal, though, is an ongoing process, so various reappraisals of experience will occur and will correspondingly affect whether or not the probability of aggression is lessened, maintained, or intensified. Rumination about provoking circumstances will of course extend or re-vivify anger reactions. In addition, the occurrence of certain thoughts can prime semantically related ideas that are part of an anger schema.

Perceived malevolence is one of the most common forms of anger-inducing appraisal. When another person's behavior is interpreted as intending to be harmful to oneself, anger and aggression schemas are activated. In turn, receiving information about mitigating circumstances (e.g., learning that the person was fatigued and working overtime) can defuse the appraisal of personal attack and promote a benign reappraisal. Perceiving malevolence pulls for anger by involving the important theme of justification, which includes the externalization of blame. When harm or injustice has been done, the social norms of retaliation and retribution are engaged. Indeed, one view of anger is that it is a socially constituted syndrome or a transitory social role governed by social rules. Thus, its meaning and function would be determined by the social systems in which it occurs and of which it is an integral part.

Justification is a core theme with regard to the activation of anger and aggression, being rooted in ancient religious texts, such as the Bible and the Koran, as well as classical mythologies about deities and historical accounts of the behavior of ancient rulers. Correspondingly, anger and physical aggression are often viewed as ways of applying a legitimate punitive response for transgression or as ways of correcting injustice. Frequently, however, an embellished justification serves the exoneration of blame for destructive outcomes of expressed anger.

As people monitor their physical and social environment for threats to their resources or self-esteem, they operate with expectations about how events and the behavior of others will unfold. When opposition, anger, or antagonism is expected, this can lead to selective perception of situational cues in line with an aggressive script, a cognitive programming that guides behavior. Anger arousal provides energy and justification for aggressive script enactment.

Anger Dyscontrol and Regulation

Anger is a highly functional human emotion, and it is one to be appreciated as a rich part of cultural life, but the survival value of the aggression-enabling function of anger is an archaic remnant with rare

contemporary necessity. The challenges presented by civilized society are predominantly psychological, rather than physical, thus attenuating anger's adaptive worth. Effective coping with the demands of modern life requires understanding complex information, problem solving, and prudent action, not energized rapid responding. Even in emergency situations, anger requires regulation. Contrary to intuitions, anger can be detrimental to survival in a physical threat crisis. It is counterproductive for energy conservation in a prolonged fight, for monitoring additional threat elements and hazards, and for effective strategy selection in circumstances in which survival threat lingers and/or remains obscure. The regulation of the intensity and duration of anger arousal is pivotal to its merit or utility.

While the physiological components of anger, such as increased blood flow, may be adaptive for survival in a short-term danger episode, the by-products of recurrent engagement of anger are hazardous in the long term. Unregulated anger commonly has been found to be associated with physical and psychological health impairments. In the realm of physical health, chronic anger has been established as having detrimental effects most centrally on the cardiovascular system, and these are related to mortality. Persons high in generalized hostility who are reactively angry are at considerable risk for coronary heart disease. When such persons are confronted with a stressful demand, they have strong cardiovascular responses in blood pressure, neurohormonal secretions, and cholesterol. A hostile, cynical, and distrusting outlook also necessitates high vigilance for thwarting and malevolence, resulting in prolonged neurohormonal activation conducive to atherosclerosis.

In addition to these pathogenic effects for a personality style that is overly expressive of anger, the coronary system is also impaired by recurrently suppressed anger and has long been identified as a causal variable in the etiology of essential hypertension. People who have difficulties expressing anger tend to be at risk for chronically elevated blood pressure, as mediated by high plasma renin activity and nor-adrenaline. The suppression of anger has been robustly correlated to elevated blood pressure in laboratory studies and to sustained hypertension in field studies. Also, studies using ambulatory recorders of blood pressure have found that persons who are high in hostility and who have a tendency to inhibit the expression of that hostility have greater cardiovascular reactivity to provocation events, as well as elevated resting blood pressure.

With regard to psychological well-being, anger occurs in conjunction with a wide range of

psychiatrically classified disorders, including a variety of impulse control dysfunctions, mood disorders, personality disorders, and forms of schizophrenia, especially paranoid schizophrenia. In addition, the activation of anger has long been recognized as a feature of clinical disorders that result from trauma, such as dissociative disorders, brain damage syndromes, and, especially, PTSD. Anger also appears in mental state disturbances produced by general medical conditions, such as dementia, substance abuse disorders, and neurological dysfunctions resulting from perinatal difficulties.

Among hospitalized psychiatric patients in long-term care in both civil commitment and forensic institutions, anger is a salient problem, as identified by both clinical staff and the patients themselves. Importantly, it is linked to assaultive behavior by psychiatric patients both inside and outside such facilities. Such patients typically have traumatic life histories, replete with experiences of abandonment and rejection, as well as economic and psychological impoverishment. For them, anger becomes entrenched as a mode of reactance to stressful or aversive experiences, and it is a significant aspect of their resistance to treatment. Chronically angry people are reluctant to surrender the anger-aggression system that they have found useful to engage and because they discount the costs of its engagement. Psychiatric hospital staff, especially those on acute admissions units and in long-term institutions, have very stressful occupations as a result of the anger episodes of the patients in their care. Posttraumatic stress disorder commonly occurs among staff who have been victims of assault by patients.

Anger Treatment

In the treatment of anger disorders, cognitive-behavioral therapy (CBT) approaches have been found to be effective with a wide range of clinical populations. CBT approaches incorporate many elements of behavior therapy, such as training in self-monitoring, relaxation, and social skills, but also centrally seek to modify cognitive structures and the way in which a person processes information about social situations. They strongly emphasize self-regulation, cognitive flexibility in appraising situations, arousal control, and learning prosocial values and scripts. By making extensive use of therapist modeling and client rehearsal, anger proneness is modified by first motivating client engagement and then restructuring cognitive schemas, increasing capacity to regulate arousal and facilitating the use of constructive coping behaviors.

A misleading aspect of the notion of anger management is that it implies that treatment is centrally

about what to do when one gets angry. Instead, a very significant objective of anger treatment is about how to not get angry in the first place – hence the priority given to regulatory controls for anger experience or activation. The parameters or state markers for anger activation that receive attention in CBT anger treatment are reactivity (frequency on onset and how easily anger is triggered), latency (how rapidly activated), intensity (how strongly engaged), and duration (persistence of arousal). Treatment aims to minimize anger reactivity, intensity, and duration and to moderate anger expression to reduce the costs of anger dyscontrol.

To facilitate anger regulation, anger treatment procedures strive to disconnect anger from the threat system. This is done first through the provision of safety, patience, and psychological space for reflection, exploration, and choice. The client's view of anger is normalized, to obviate worries about being a bad or unworthy person. The therapist acknowledges the legitimacy of the client's feelings, affirming his or her self-worth. Building trust in the therapeutic relationship is pivotal. As self-regulation hinges on knowledge, education about anger and discovery of the client's personal anger patterns or anger signature is facilitated. Much is done to augment self-monitoring and to encourage the moderation of anger intensity. As tension or strain may surface in the course of treatment, the therapist models and reinforces non-anger alternative responding so as to build replacements for the automatized angry reactions that had been the client's default coping style.

One CBT approach to anger treatment that has received significant support for its efficacy is called stress inoculation (SI). In this treatment approach, anger provocation is simulated by therapeutically paced exposure to anger incidents created in imaginal visualization and in role play. The progressively graduated exposure, directed by the therapist, involves a hierarchy of anger incidents produced by the collaborative work of client and therapist. This graduated, hierarchical exposure, done in conjunction with the teaching of stress coping skills, is the basis for the inoculation metaphor.

The SI approach to anger treatment involves the following key components: (1) client education about anger, stress, and aggression; (2) self-monitoring of anger frequency, intensity, and situational triggers; (3) construction of a personal anger provocation hierarchy, created from the self-monitoring data and used for the practice and testing of anger coping skills; (4) arousal reduction techniques of progressive muscle relaxation, breathing-focused relaxation, and guided imagery training; (5) cognitive restructuring by altering attentional focus, modifying appraisals,

and using self-instruction; (6) training behavioral coping in communication and respectful assertiveness as modeled and rehearsed with the therapist; and (7) practicing the cognitive, arousal regulatory, and behavioral coping skills while visualizing and role playing progressively more intense anger-arousing scenes from the personal hierarchies.

See Also the Following Articles

Aggression; Autonomic Nervous System; Emotions: Structure and Adaptive Functions; Hostility; Sympathetic Nervous System.

Further Reading

- Averill, J. R. (1982). *Anger and aggression: an essay on emotion*. New York: Springer-Verlag.
- Berkowitz, L. (1993). *Aggression: its causes, consequences, and control*. New York: McGraw Hill.
- Chesney, M. and Rosenman, R. H. (1985). *Anger and hostility in cardiovascular and behavioral disorders*. Washington, D.C.: Hemisphere.

- Follette, V. M., Rusek, J. I. and Abueg, F. R. (1998). *Cognitive behavioral therapies for trauma*. New York: Guilford.
- Friedman, H. (1992). *Hostility, coping, and health*. Washington, D.C.: American Psychological Association.
- Johnson, E. H., Gentry, W. D. and Julius, S. (1992). *Personality, elevated blood pressure, and essential hypertension*. Washington, D.C.: Hemisphere.
- Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford, UK: Oxford University Press.
- Ortony, A., Clore, G. L. and Collins, A. (1988). *The cognitive structure of emotions*. New York: Cambridge University Press.
- Philippot, P. and Feldman, R. S. (2004). *The regulation of emotion*. Mahwah, NJ: Erlbaum.
- Potegal, M. and Knutson, J. F. (1994). *The dynamics of aggression: biological and social processes in dyads and groups*. Hillsdale, NJ: Erlbaum.
- Siegmán, A. W. and Smith, T. W. (eds.) (1994). *Anger, hostility, and the heart*. Hillsdale, NJ: Erlbaum.
- Taylor, J. L. and Novaco, R. W. (2005). *Anger treatment for people with developmental disabilities*. Chichester, UK: Wiley.

Angiotensin

O Baltatu

Max-Delbrück-Center for Molecular Medicine, Berlin-Buch, Germany, and Sultan Qaboos University, Muscat, Sultanate of Oman

M Bader

Max-Delbrück-Center for Molecular Medicine, Berlin-Buch, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by O Baltatu, M Bader, and D Ganten, volume 1, pp 195–199, © 2000, Elsevier Inc.

Components of the Renin–Angiotensin System
Endocrine Plasma Renin–Angiotensin System
Paracrine Tissue Renin–Angiotensin System
Renin–Angiotensin System and Stress

Glossary

Blood–brain barrier

A term for the impermeability of the endothelial cells of cerebral capillaries to components of the blood, due *inter alia* to tight junctions, so that endo- and transcytotic transport across the capillary wall occurs in a highly specific and regulated manner.

Cardiovascular control

Control of cardiovascular homeostasis by neural and local mechanisms.

Cardiovascular homeostasis

The maintenance of the cardiovascular parameters, especially blood pressure and blood volume at levels optimal for blood supply to all organs.

Components of the Renin–Angiotensin System

Angiotensinogen

Angiotensinogen (AOPEN) is the only known precursor molecule to be converted by the renin–angiotensin system (RAS) into active angiotensins. It is a glycoprotein with a molecular weight between 61 and 65 kDa, and it shows close structural similarities to the serine protease inhibitor family. Originally found to be synthesized in the liver, AOPEN is known to be present in several other organs, especially in glial cells of the brain.

Renin

Renin is a very specific aspartyl proteinase whose only known substrate is AOPEN. Renin cleaves the decapeptide angiotensin I (Ang I) from the N-terminus of mature AOPEN. Renin is produced primarily by

the kidney and released into the circulation, but is also produced in other tissues. Renin synthesis is stimulated by the sympathetic nervous system and by loss of volume and sodium.

Angiotensin Converting Enzyme

Angiotensin converting enzyme (ACE) is a dipeptidyl-carboxypeptidase that cleaves the C-terminal histidine-leucine of Ang I to generate angiotensin II (Ang II). It is found in high concentrations in the lung, kidney, and heart and is distributed throughout the brain. High amounts are detected in areas lacking the blood–brain barrier such as the diencephalon, pituitary, and pineal gland. ACE is also able to hydrolyze other substrates, such as bradykinin, opioid peptides, and substance P. Thus, ACE is additionally involved in physiological mechanisms beyond the RAS.

Active Angiotensins

The best studied peptide of the system is Ang II, and the functions of the RAS are largely attributed to this active metabolite. Besides Ang II, additional peptides of the RAS have been found to be biologically active, such as Ang III, Ang IV, and Ang II(1–7). In fact, Ang III may be even more potent than Ang II in the brain when acting on a common receptor. Ang II(1–7) can be formed via an ACE-independent mechanism and has the same potency as Ang II in vasopressin release. Studies on Ang IV suggest a biological role for this peptide as well.

There is evidence that proteases other than renin and ACE can produce bioactive angiotensin species. Such proteases include ACE 2, cathepsin G, tonin, tissue plasminogen activator, chymase, and other neutral peptidases.

Receptors

The use of highly specific ligands has led to the description and cloning of two main Ang II receptors, namely AT₁ and AT₂. Most of the known classical cardiovascular effects of Ang II are attributed to the AT₁ receptor. In contrast to humans, two different subtypes have been found for this receptor in rodents: AT_{1a} and AT_{1b}. Whereas AT_{1a} has been localized in blood vessels, kidney, lung, liver, and specific brain areas that are involved in the control of blood pressure and fluid homeostasis, AT_{1b} receptors were originally described in glandular tissues, such as the anterior pituitary and adrenal gland.

Endocrine Plasma Renin-Angiotensin System

The endocrine RAS is dependent on renin released by the kidney into the blood, so that AOPEN is cleaved

in the circulation to liberate the inactive decapeptide Ang I. ACE then cleaves Ang I to yield the octapeptide Ang II, the active molecule of the system. Circulating Ang II is one of the most potent vasoconstrictors and acts on specific receptors to elicit the following localized effects:

- constriction of the vasculature
- reabsorption of sodium and water in the kidney
- release of aldosterone from the adrenal glands
- induction of inotropic and arrhythmogenic effects in the heart
- incitement of thirst at specific brain areas outside the blood–brain barrier

Figure 1 summarizes the roles of the RAS in cardiovascular and fluid–electrolyte regulation.

Paracrine Tissue Renin-Angiotensin System

In addition to the circulating RAS, a local formation and action of Ang II has been observed at the level of various organs. The components of the RAS have been detected in several tissues including the heart, vasculature, kidney, adrenal glands, and brain. The local formation and action of Ang II in peripheral organs can therefore occur independent from the circulating RAS. However, RAS components required for angiotensin synthesis, such as renin and AOPEN, can be also taken up from the circulating blood. The latter is possible only for organs accessible to blood proteins, such as the kidney, heart, adrenal glands, and brain areas situated outside the blood–brain barrier. Nevertheless, there are areas in the brain that are inaccessible to the circulating Ang II because of the blood–brain barrier but that express functional angiotensin receptors. In these brain regions, only locally formed angiotensins can act.

The tissue RASs seem to regulate long-term organ function and may be responsible for pathological structural changes. **Figure 2** illustrates the functions of the endocrine and tissue RASs.

Renin-Angiotensin System and Stress

Both physical and emotional stressors induce peripheral and central responses designed to preserve cardiovascular homeostasis. As one of the main regulators of cardiovascular homeostasis, the RAS is stimulated by stressors that lead to a lowering in the blood pressure, such as hemorrhage, loss of blood volume and sodium, or increased activity of the sympathetic nervous system. The RAS acts synergistically with the sympathetic nervous system, increasing blood pressure and maintaining organ perfusion, and electrolyte

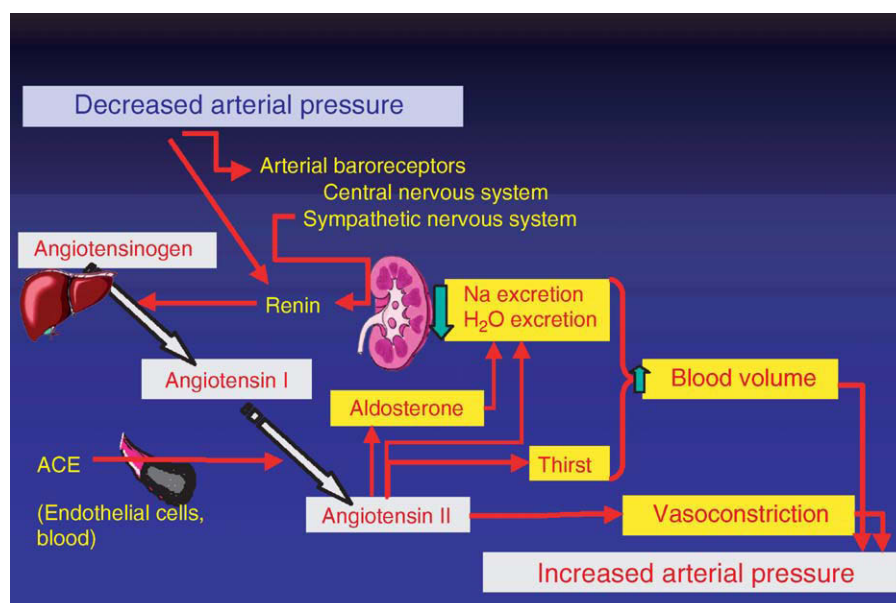


Figure 1 Roles of the renin–angiotensin system in cardiovascular and fluid–electrolyte regulation.

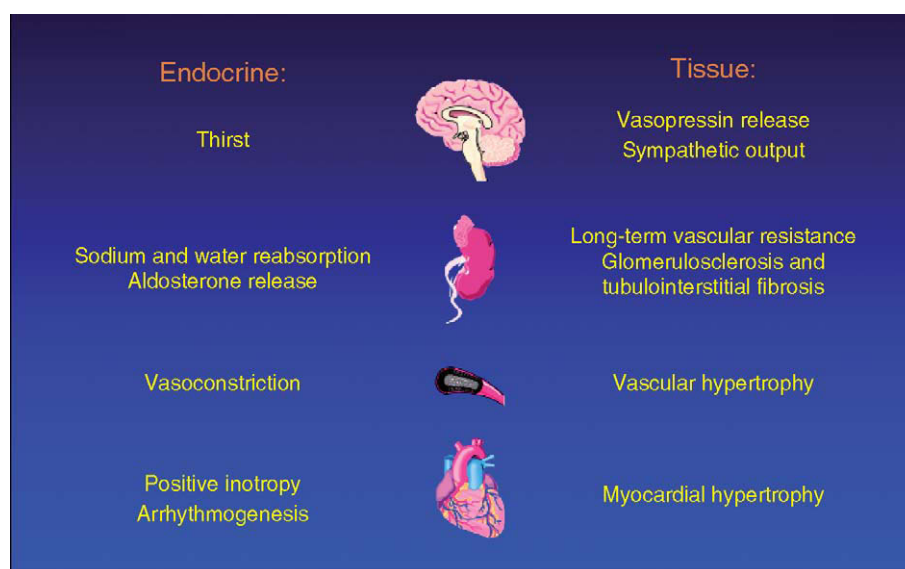


Figure 2 Functions of the endocrine and tissue renin–angiotensin systems.

and volume balance. Recently, it has been demonstrated that Ang II is also an important mediator of inflammation by stimulating oxidative stress.

The physical and behavioral responses to stress are integrated at the level of the central nervous system, including areas known to regulate the homeostasis of the *milieu interieur*. AOGEN and the angiotensin receptors are present in areas responsible for the maintenance of homeostasis, such as the hypothalamic and brain-stem nuclei (Figure 3), and participate in the stress response.

The sympathetic nervous system, the component of the autonomic nervous system that is essentially involved in the fight-or-flight reaction, is modulated by the RAS not only in the periphery but also in central areas of autonomic control. During the reaction to stress, there is an increase in blood levels of renin in addition to epinephrine. In the central nervous system, an interrelationship between the activity of the brain RAS and the sympathetic system and other central autonomic pathways has also been shown. Stress-induced brain RAS activation may

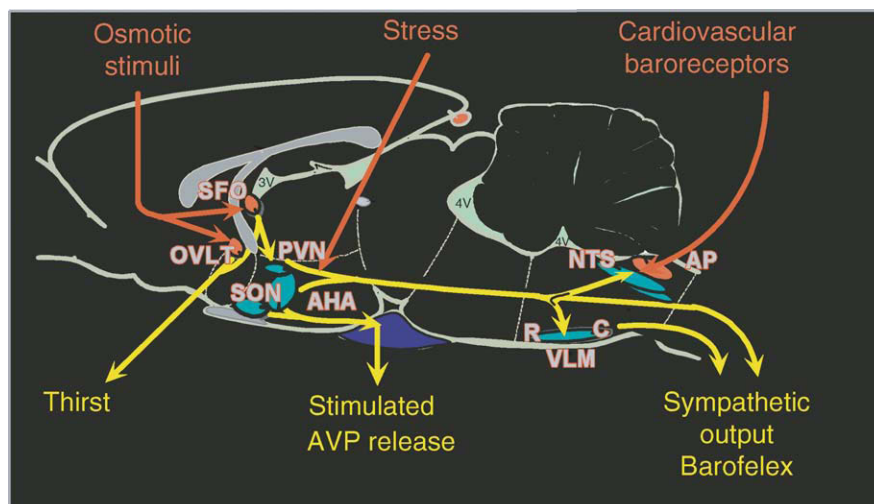


Figure 3 Angiotensin sites involved in the central control of homeostasis.

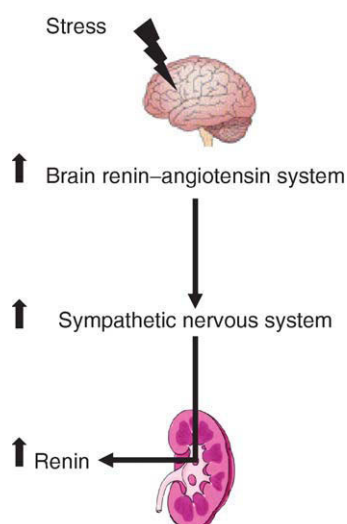


Figure 4 Stress-induced brain RAS activation may induce renin release from the kidney, possibly through sympathetic pathways.

induce renin release from the kidney, possibly through sympathetic pathways (Figure 4).

Various stressor agents activate the hypothalamic-pituitary axis. Ang II, together with other neurotransmitters and neuromodulators, modulates the activity of the hypothalamic-pituitary axis and stimulates the release of vasopressin and corticotropin-releasing factor (CRF). In addition, it has been proposed that the brain RAS is involved in behavioral responses, such as cognition, memory, and central analgesia.

Prolonged overactivation of the RAS may lead to a disease state. Drugs that block the actions of the RAS, such as ACE or renin inhibitors and AT1 antagonists, are widely used for the treatment of various cardiovascular diseases, and their effectiveness may be

partially due to the inhibition of the brain RAS and thereby the attenuation of the sympathetic output and the stress response.

See Also the Following Articles

Angiotensin Receptors; Blood Pressure; Salt Appetite.

Further Reading

- Bader, M., Peters, J., Baltatu, O., et al. (2001). Tissue renin-angiotensin systems: new insights from experimental animal models in hypertension research. *Journal of Molecular Medicine* 79, 76–102.
- Baltatu, O., Bader, M. and Ganten, D. (1998). Brain and renin-angiotensin system. In: Nicholls, M. G., Ikram, H., Brunner, H. & Sweet, C. (eds.) *100 Years of Renin-Angiotensin System*, pp. 167–170. Oxford: Hughes Associates.
- Baltatu, O., Bader, M. and Ganten, D. (1998). Functional testing of components of the brain renin-angiotensin system in transgenic animals. In: Ulfendahl, H. & Aurell, M. (eds.) *Renin-Angiotensin*, pp. 105–114. London: Portland Press.
- Baltatu, O. and Bader, M. (2003). Brain renin-angiotensin system – lessons from functional genomics. *Neuroendocrinology* 78, 253–259.
- Baltatu, O., Campos, L. A. and Bader, M. (2004). Genetic targeting of the brain renin-angiotensin system in transgenic rats: impact on stress-induced renin release. *Acta Physiologica Scandinavica* 181, 579–584.
- de Gasparo, M., Catt, K. J., Inagami, T., et al. (2000). International union of pharmacology, XXIII. The angiotensin II receptors. *Pharmacology Reviews* 52, 415–472.
- Fink, G. D. (1997). Long-term sympatho-excitatory effect of angiotensin II: a mechanism of spontaneous and renovascular hypertension. *Clinical and Experimental Pharmacology and Physiology* 24, 91–95.

- Henry, J. P. (1992). Biological basis of the stress response. *Integrative Physiological and Behavioral Science* **27**, 66–83.
- Leckie, B. J. (2005). Targeting the renin-angiotensin system: what's new? *Current Medical and Chemical Cardiovascular Hematology Agents* **3**, 23–32.
- Morgan, L., Broughton, P. F. and Kalsheker, N. (1996). Angiotensinogen: molecular biology, biochemistry and physiology. *International Journal of Biochemistry and Cell Biology* **28**, 1211–1222.
- Saavedra, J. M., Ando, H., Armando, I., et al. (2004). Brain angiotensin II, an important stress hormone: regulatory sites and therapeutic opportunities. *Annals of the New York Academy of Sciences* **1018**, 76–84.
- Touyz, R. M. (2005). Molecular and cellular mechanisms in vascular injury in hypertension: role of angiotensin II – editorial review. *Current Opinions in Nephrology and Hypertension* **14**, 125–131.

Angiotensin Receptors

T A Jenkins and F A O Mendelsohn

University of Melbourne, Parkville, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

Angiotensin Receptors

Actions of Angiotensin

Pathophysiology of the RAS

The RAS and Stress

Glossary

<i>Angiotensin-converting enzyme (ACE)</i>	Enzyme that converts angiotensin I to angiotensin II.
<i>Angiotensin II receptors</i>	The two classes of Ang II receptors, AT ₁ and AT ₂ , occur in a variety of tissues, including brain, adrenal, heart, vessel wall, and kidney.
<i>Angiotensin II (Ang II)</i>	Octapeptide; a major active mediator of the renin-angiotensin cascade.
<i>Cardiovascular disease</i>	Disease of the heart and blood vessels.
<i>Hypertension</i>	High arterial blood pressure.
<i>Renin-angiotensin system (RAS)</i>	Hormone system that plays an important role in regulating body fluid balance, blood volume, arterial pressure, and cardiac, vascular, and renal function.

The physiological roles of the renin-angiotensin system (RAS) include regulation of blood pressure, fluid and electrolyte homeostasis, and aldosterone and pituitary hormone release. The octapeptide angiotensin II (Ang II) is the major active component of the RAS and is generated by cleavage of the substrate, angiotensinogen, through the successive actions of two proteolytic enzymes, renin and angiotensin-converting enzyme (ACE). Ang II acts via two receptors, AT₁ and AT₂. Although originally viewed solely

as a circulating system, all components of the RAS have been found in a variety of tissues, including brain, adrenal, heart, vessel wall, and kidney, suggesting the presence of tissue-specific systems participating in physiological and pathophysiological effects of the RAS (Figure 1).

Recently, the classical view of the RAS was changed by the discovery of the enzyme ACE2, a human homolog of ACE, and awareness that other angiotensin peptides, such as Ang IV and Ang 1–7, may have physiological actions.

Angiotensin Receptors

AT₁ Receptor

The AT₁ receptor mediates most of the classical actions of Ang II. It displays a seven transmembrane domain topology and is a member of the G-protein-coupled receptor family, with signaling predominantly via G_{q/11}, phosphoinositides, and calcium. In addition, AT₁ activation is linked to vascular and cardiac hypertrophy through transactivation of the epidermal growth factor (EGFR) and its signaling pathways, primarily through mitogen-activated protein (MAP) kinase activation. In rodents, two subtypes exist, AT_{1A} and AT_{1B}, which have 95% homology.

Many of the functions of Ang II are paralleled by the distribution of AT₁ receptors in cardiovascular, nervous, renal, endocrine, and reproductive tissues. In the central nervous system, AT₁ receptors are found in the circumventricular organs, including the subfornical organ (SFO) and vascular organ of the lamina terminalis (OVLT), and regions of the hypothalamus such as the paraventricular nucleus and median eminence, which regulate body fluid homeostasis. AT₁ receptors in the nucleus of the solitary tract and ventrolateral medulla reflect important roles of Ang II in cardiovascular control. In the periphery,

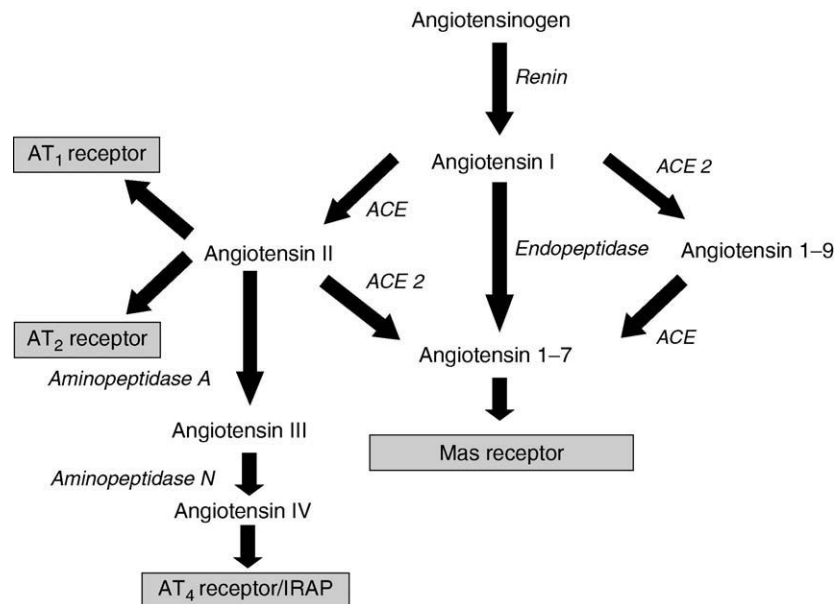


Figure 1 The renin-angiotensin system.

AT₁ receptors occur in high levels in the zona glomerulosa of the adrenal cortex and adrenal medulla. In the kidney, AT₁ receptors occur in glomeruli, proximal tubule, and medullary interstitial cells. AT₁ receptors occur on vascular smooth muscle cells and in the conduction system and vagal ganglia of the heart.

The antagonists losartan and candesartan are AT₁ selective and are used clinically to treat hypertension and cardiovascular disease.

AT₂ Receptor

The function and characteristics of the AT₂ receptor are less well understood. While it is also a seven transmembrane receptor, it shares only 34% homology with the AT₁ receptor. The signaling mechanisms are distinct from the AT₁ receptor and include stimulation of tyrosine or serine/threonine phosphatases, which inactivate MAP kinases, and potassium channel activation. AT₂ receptor stimulation may inhibit growth and promote apoptosis, thereby opposing actions at the AT₁ receptor.

The AT₂ receptor is highly expressed during fetal development and rapidly declines at birth. In the adult brain, AT₂ receptors are constrained to the molecular layer of the cerebellum, thalamus, and superior colliculus. In neurons, AT₂ receptor activation may result in hyperpolarization and decreased excitability via enhanced K⁺ channel activity. In the periphery, it is found within the adrenal medulla, cardiac fibroblasts, and the non-pregnant uterus. Expression of the AT₂ receptor is upregulated in pathological conditions such as heart failure, renal

failure, vascular injury, and wound healing. AT₂ receptor stimulation also induces vasodilation mediated by bradykinin, nitric oxide, and cyclic GMP. The AT₂ receptor also binds directly to the AT₁ receptor, and may antagonize its function via this heterodimerization.

The AT₂ receptor is selectively inhibited by the nonpeptide compound PD 123319 while the peptide analog CGP 42112A acts as a partial agonist.

AT₄ Receptor

The hexapeptide fragment of angiotensin (3–8) (Ang IV) has actions distinct from Ang II and binds to a novel binding site distinct from the AT₁ and AT₂ receptors. The AT₄ receptor is a type II integral membrane protein identical to the insulin-regulated aminopeptidase (IRAP), a zinc metallopeptidase of the M1 family that translocates rapidly to the cell surface after insulin stimulation.

The AT₄/IRAP receptor occurs abundantly in the brain within the neocortex, CA1 to CA3 regions of the hippocampus and motor nuclei, where it is closely associated with cholinergic neurons, consistent with the memory-enhancing properties of Ang IV. In the periphery, it is found in blood vessels, cortex, and outer stripe of the outer medulla of the kidney and adrenal medulla and cortex. Its main peripheral physiological effect is vasodilation.

Two N-terminally modified analogs of Ang IV have high affinities for the AT₄ receptor: Nle¹-Ang IV and norleucinal Ang IV and exhibit agonist properties along with a structurally distinct peptide, LVV-hemorphin 7.

Ang 1–7 Receptor

A possible role for Ang 1–7 as a cardioprotective peptide with vasodilator, diuretic, and antiproliferative effects has recently been suggested. The Ang 1–7 receptor may be the G-protein-coupled receptor Mas, encoded by the Mas1 proto-oncogene.

Actions of Angiotensin**Fluid and Electrolyte Balance**

Ang II rapidly induces drinking, whether administered peripherally, where it acts on AT₁ receptors within the circumventricular organs, or centrally, where the lamina terminalis encompassing the median preoptic nucleus is the principal site of action.

Centrally administered Ang II also stimulates salt appetite. This action is blocked by centrally administered AT₁ and AT₂ receptor antagonists; however, the central sites of action of Ang II-induced salt appetite have yet to be determined. Mineralocorticoids have a major influence over this stimulus and are believed to have a concomitant action with the RAS in initiating salt appetite.

Blood Pressure Regulation

Infusion of Ang II into the periphery or brain raises blood pressure. Peripheral administration of Ang II increases arterial pressure via increased vascular resistance, a consequence of both direct action on vascular smooth muscle and central actions resulting in activation of the sympathetic nervous system and stimulation of vasopressin release.

In a genetic animal model, the spontaneous hypertensive rat (SHR), the brain Ang II system is overactive. Higher levels of Ang II and angiotensin receptors have been observed, and the SHR exhibits an exaggerated pressor response to central Ang II.

Hormone Release

Ang II has several roles as a neuropeptide. Central infusion of Ang II is a potent stimulant for the release of vasopressin. Ang II may also participate in cyclic regulation of reproductive hormones. Plasma Ang II is raised in the luteal phase of the menstrual cycle, and administered Ang II elevates plasma levels of luteinizing hormone during this phase.

Behavior

The RAS, either acting directly or through the modulation of various neurotransmitters, has been implicated in a range of behavioral and cognitive processes. While there is conflicting evidence of the role of Ang II in memory and learning, it is clear that

central infusions of Ang IV facilitate memory retention and retrieval in rats. Ang IV promotes learning in passive avoidance and spatial memory tasks, and reverses memory deficits induced by scopolamine-induced amnesia and bilateral perforant pathway lesions. The mechanisms of how AT₄ ligands mediate their effects are unclear, but may involve interactions with brain cholinergic systems, modulation of neural glucose metabolism, or inhibition of neuropeptide metabolism by IRAP.

Pathophysiology of the RAS

The local tissue RAS has important effects on the pathological changes of organ structure that are mediated by altered modulated gene expression, growth, fibrosis, and the inflammatory response.

Hypertension

The RAS regulates blood pressure, and inhibitors of the system have important therapeutic actions in treating hypertension. ACE inhibitors are widely used in hypertension and cardiovascular and renal disorders. AT₁ antagonists, such as losartan and candesartan, are increasingly used in hypertension and are well tolerated.

Cardiac Hypertrophy

Cardiac hypertrophy is characterized by increased deposition of extracellular matrix (ECM) constituents, proliferation of cardiac fibroblasts, and hypertrophic growth of cardiac myocytes, which alter the contractile properties of the heart.

The RAS is a key mediator of cardiac adaptations to hemodynamic overload. Ang II directly induces cellular responses in myocardial cells. In cultured cardiac fibroblasts, Ang II stimulates fibroblast proliferation, collagen synthesis, and the expression of ECM proteins such as fibronectin via activation of the AT₁ receptor. In cardiac myocytes, Ang II exerts a direct growth-promoting effect on only neonatal cells, and thus may act indirectly by inducing the expression of growth factors such as TGF- β_1 .

Both clinical and experimental studies demonstrate a favorable impact of AT₁ antagonists and ACE inhibitor therapy on the progressive remodeling and ventricular dysfunction of cardiac hypertrophy by reducing the action of Ang II. In addition, the beneficial effect of ACE inhibition might be related to local accumulation of bradykinin.

Microvascular Injury

Kidney Both ACE inhibitors and AT₁ antagonists exert protective effects in chronic renal disease by

improving glomerular filtration, preventing glomerular and tubulointerstitial toxicity of ultrafiltered proteins, and limiting the direct nephrotoxic effects of Ang II, especially in diabetes.

Eye Neovascularization in the retina of diabetic patients is a major cause of visual loss. There is a large body of evidence for a local RAS within the eye that is activated in diabetes. This appears to be directly responsible for an increase in vascular endothelial growth factor (VEGF) that is implicated in the pathogenesis of diabetic retinopathy. Inhibition of ACE appears to reduce concentrations of VEGF, with a concurrent antiproliferative effect independent of systemic VEGF levels or blood pressure. AT₁ receptor blockade has been shown to reduce neovascularization independent of VEGF levels in animal models. This may be due to antagonism of activation of MAP kinase, which is a potent cellular proliferation stimulator, by Ang II.

Brain Circulating and locally formed Ang II controls cerebral blood flow by AT₁ receptor stimulation in cerebral vessels and sympathetic nerves. The Perindopril Protection against Recurrent Stroke Study (PROGRESS) has shown that ACE inhibitor therapy combined with diuretics can decrease recurrence of stroke by 44% in hypertensive patients and by 42% in nonhypertensive patients. Chronic AT₁ blockade inhibits brain and cerebral artery AT₁ receptor and reduces cortical ischemia and tissue swelling produced by middle cerebral artery occlusion with reperfusion.

Atherosclerosis

Ang II plays a critical role in the initiation and progression of atherosclerosis via promotion of superoxide production and acting as a pro-inflammatory molecule, with resultant local inflammation and tissue damage. Treatment with ACE inhibitors or AT₁ antagonists may slow progression of atherosclerosis. This is in part associated with decreased expression of inflammatory mediators and improved endothelial function.

The RAS and Stress

Both brain and peripheral Ang II participate in the stress response. In response to stress, AT₁ receptor expression increases in the hypothalamic paraventricular nucleus and median eminence, key sites controlling the hypothalamic-pituitary-adrenal (HPA) axis. Concurrently, there is increased synthesis and release of corticotropin-releasing factor (CRF) from the median eminence and adrenocorticotrophic hormone (ACTH) from the anterior pituitary and consequently adrenal glucocorticoid responses (Figure 2). Associated with this are central actions that contribute to enhanced peripheral sympathetic activity.

The increase in sympathetic stimulation leads to enhanced renin release, increased circulating Ang II, and, consequently, fluid retention and vasoconstriction, with increased blood pressure and heart rate.

The RAS has been implicated in a number of experimental models of stress. Restraint stress in rats

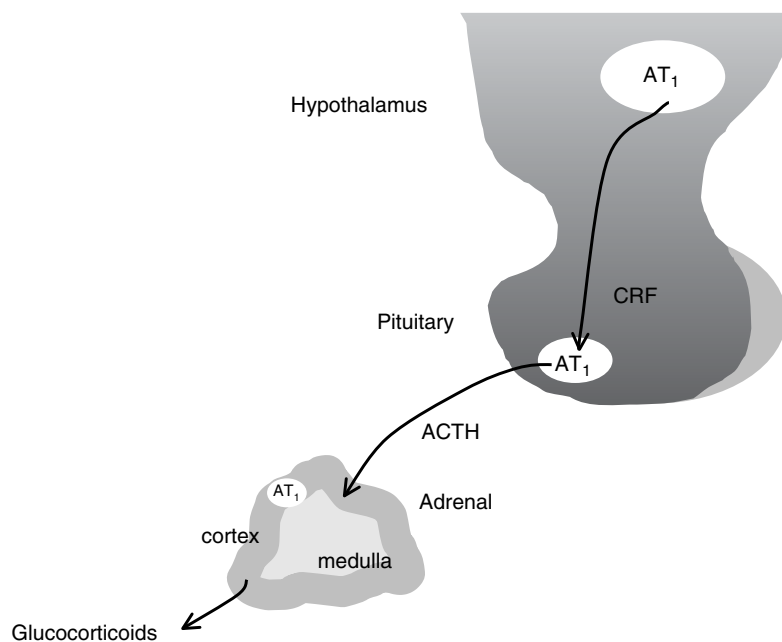


Figure 2 Angiotensin II and the stress response.

increases AT₁ receptor levels in the hypothalamic paraventricular nucleus, SFO, and median eminence, consistent with roles in HPA activation. Concurrent changes in Ang II receptor number in the anterior pituitary and adrenal zona glomerulosa and medulla parallel the peripheral influences of the RAS on the HPA axis. Isolation stress enhances brain Ang II receptor numbers, which can be suppressed by AT₁ receptor antagonists. Conversely, central AT₁ receptor antagonism attenuates stress-induced sympathetic nervous and adrenomedullary responses.

The SHR is more susceptible to stress than normotensive Wistar-Kyoto (WKY) rats, and has an overactive central RAS. Differential expression of AT_{1A}: AT_{1B} receptors is observed in the HPA of SHR, and Ang II produces an exaggerated response of ACTH and corticosterone.

See Also the Following Articles

Adrenal Cortex; Adrenal Medulla; Aldosterone and Mineralocorticoid Receptors; Angiotensin; Corticotropin Releasing Factor (CRF); Hypertension; Hypothalamic-Pituitary-Adrenal; Kidney Function.

Further Reading

- Albiston, A. L., Mustafa, T., McDowall, S. G., et al. (2003). AT₄ receptor is insulin-regulated membrane aminopeptidase: potential mechanisms of memory enhancement. *Trends in Endocrinology and Metabolism* **14**, 72–77.
- Ferrario, C. M. and Chappell, M. C. (2004). Novel angiotensin peptides. *Cellular and Molecular Life Sciences* **61**, 2720–2727.

- Gallinat, S., Busche, S., Raizada, M. K. and Summers, C. (2000). The angiotensin II type 2 receptor: an enigma with multiple variations. *American Journal of Physiology, Endocrinology and Metabolism* **278**, E357–E374.
- Gard, P. R. (2004). Angiotensin as a target for the treatment of Alzheimer's disease, anxiety and depression. *Expert Opinion on Therapeutic Targets* **8**, 1–8.
- Kaschina, E. and Unger, T. (2003). Angiotensin AT₁/AT₂ receptors: regulation, signalling and function. *Blood Pressure* **12**, 70–88.
- Kim, S. and Iwao, H. (2000). Molecular and cellular mechanisms of angiotensin II-mediated cardiovascular and renal diseases. *Pharmacological Reviews* **52**, 11–34.
- McKinley, M. J., Albiston, A. L., Allen, A. M., et al. (2003). The brain renin-angiotensin system: location and physiological roles. *International Journal of Biochemistry and Cell Biology* **35**, 901–918.
- Saavedra, J. M., Ando, H., Armando, I., et al. (2004). Brain angiotensin II, an important stress hormone: regulatory sites and therapeutic opportunities. *Annals of the New York Academy of Sciences* **1018**, 76–84.
- Saavedra, J. M. J., Ando, H., Armando, I., et al. (2005). Anti-stress and anti-anxiety effects of centrally acting angiotensin II AT₁ receptor antagonists. *Regulatory Peptides* **128**, 227–238.
- Unger, T. (2002). The role of the renin-angiotensin system in the development of cardiovascular disease. *American Journal of Cardiology* **89**(2A), 3–9.
- Watanabe, T., Fujioka, T., Hashimoto, M. and Nakamura, S. (1998). Stress and brain angiotensin II receptors. *Critical Reviews of Neurobiology* **12**, 305–317.
- Wright, J. W. and Harding, J. W. (2004). The brain angiotensin system and extracellular matrix molecules in neural plasticity, learning, and memory. *Progress in Neurobiology* **72**, 263–293.

Animal Models (Nonprimate) for Human Stress

J E Ottenweller

East Orange Department of Veterans Affairs Medical Center and University of Medicine and Dentistry of New Jersey, Newark, NJ, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 200–205, © 2000, Elsevier Inc.

Human Stress Disorders
Stress Responses
Animal Models
Summary

Glossary

- Acute stress** Changes in physiology and behavior that occur during and shortly after exposure to a stressor.
- Chronic stress** Changes in physiology and behavior that persist for days or weeks during repeated stressors or after exposure to a stressor or stressors.
- Posttraumatic stress disorder** A psychological disorder that follows exposure to a traumatic event and is characterized by intrusive imagery, sleep disturbances and nightmares, poor concentration and memory, hypervigilance, and numbness.

Stressor A change in the environment that is sensed by an organism, is aversive or potentially harmful to that organism, and elicits acute and/or chronic stress responses.

An organism undergoes a series of physiological changes when exposed to a stressor. In the short term, several brain centers can be activated, including, among others, the locus coeruleus in the brain stem, the paraventricular nucleus in the hypothalamus, and amygdala and hippocampus in the limbic system. The activation of these centers leads to increased sympathetic nervous system (SNS) activity and elevated hypothalamic-pituitary-adrenocortical (HPA) axis hormones. These responses are homeostatic mechanisms that allow an individual to cope with a stressor and eventually turn off the stress response through negative feedback. However, if people are exposed to a very severe or prolonged stressor, more persistent changes characteristic of posttraumatic stress disorder or a general stress reaction can become established. Animal models of stress have been used to study these and other facets of acute and chronic stress responses.

Human Stress Disorders

Posttraumatic stress disorder is the only chronic stress disorder recognized in the *Diagnostic and Statistical Manual of Mental Disorders* published by the American Psychiatric Association, but it is grouped with anxiety and panic disorders, which are characterized by acute stress responses. Stress can also contribute to depressive disorders. In addition, a number of medical conditions result from stress. It is well recognized that stress is linked to cardiovascular disease, particularly hypertension and sudden cardiac death. Stress can also suppress immune function, leading to susceptibility to infections. In some abused children, stress can produce growth hormone deficiency and small stature. Stress and excessive exercise in women can interfere with ovarian cyclicity and fertility. These are only a few examples of the ways stress directly affects health through its actions on diverse systems within the body.

Stress Responses

Neuroendocrine and Immune Responses

Stress affects a number of physiological systems whose regulatory mechanisms interact, and corticotropin-releasing hormone (CRH) appears to have an important integrative function. In the brain stem, increased CRH

is required for normal SNS activation by stress. In the hypothalamus, CRH increases with stress and its secretion leads to release of adrenocorticotropin by the pituitary gland, which in turn stimulates the adrenal cortex to secrete glucocorticoids. Although its exact function remains unclear, CRH is also present in the limbic system. Thus, CRH appears to have a central role in coordinating the major responses to stress.

The catecholamines, epinephrine, and norepinephrine also play a critical role in the stress response. Plasma epinephrine comes from the adrenal medulla and is more responsive to psychological stressors, whereas plasma norepinephrine comes largely from the sympathetic nerve terminals and is more responsive to physical stressors. In addition to increasing the heart rate and blood pressure during stress, the catecholamines stimulate the breakdown of glycogen and fats to increase the availability of energy for the brain and muscles.

Almost all endocrine systems respond to stress in one way or another. Depending on the duration, intensity, and nature of the stressor, plasma prolactin and vasopressin can be elevated and growth hormone can be suppressed or stimulated. The hypothalamic-pituitary-thyroid and hypothalamic-pituitary-reproductive axes are generally suppressed by stress. In addition, plasma insulin usually decreases and glucagon increases.

These endocrine changes can be related to changes in immune function. Infection activates a number of cytokines that induce fever, stimulate the HPA axis, and inhibit the reproductive axes. In addition, the glucocorticoids can inhibit immune function, and a number of the other hormones have immunomodulatory roles. Thus, the interactions between the endocrine and immune systems in response to stress are bidirectional.

Psychological Responses

Exposure to stressors can result in posttraumatic stress disorder, and the stress response is a major component of anxiety disorders and phobias. Stress can also significantly impact cognitive function by impairing or enhancing it depending on the stressor and the type of learning or memory being assessed. In addition to these effects of stress, a number of psychological processes including habituation and sensitization, predictability, controllability, and coping can modify the responses to stressors.

There is a coordinated response to a stressor by most systems in the body that allows an individual to cope with the stressor and restore baseline conditions after the stressor has ended. One of the primary concerns when evaluating animal models of stress is what features of the stress response are important to judge the validity of the models. From the discussion

so far, it is clear that this will depend on the organ system being studied and the stressors being modeled. For example, many researchers use animal models involving psychological stress because they argue that most human stressors are primarily psychological or have a large psychological component, but common physical stressors in people are often overlooked, including, among others, disease and infection, exercise, childbirth, surgery, and physical trauma during even minor accidents (broken bones, cuts, strains, and burns). An example of having to consider the organ system being studied is research into the effects of cardiovascular stress that has often used dogs, rabbits, and pigs rather than rodents because the cardiovascular systems of the former are more like those of humans. Thus, there are many valid animal models of human stress directed toward understanding different facets of the stress response.

Animal Models

The endocrine, immune, and behavioral responses to stressors have been studied most often using rodents in which the neural, endocrine, and immune substrates involved in the stress response have been delineated. However, a lot of applied stress research in animal husbandry is overlooked; the effects of stress on reproduction, weight gain, and health have often been studied in commercially important species. Finally, as previously noted, cardiovascular responses to stress have often been studied in rabbits, dogs, and pigs because of subtle differences between people and rodents in cardiovascular regulation.

Acute Stress

Physical stressors Most studies of acute responses to a variety of stressors have been carried out in rodents. Commonly used acute stressors include restraint, shock, swimming, cold, loud noises, novelty, and social conflict. These stressors can be applied for periods as short as 1–2 min or as long as 6–8 h. The acute response to these stressors may be followed during the stressor and for several hours after the stressor has ended. It is clear that most, if not all, of these stressors have substantial short-term effects on behavior, physiology, and immune function. It is commonly thought that the acute stress response is self-limiting because the responding systems are adaptive and negative feedback loops will terminate the responses. Failure to terminate the acute stress response appropriately can lead to persistent changes that are characteristic of a chronic stress state.

Most of these acute stressors are applied in sessions during which the animals must be removed from their

home cages, which in itself is a mild stressor. Thus, the stressors also include a novelty component involving removal from the home cage and handling. In order to isolate the effects of a particular stressor, sometimes the stressed animals are compared with controls that have been removed from their cages and put in a novel environment, whereas other times all the animals are handled 1–2 weeks before exposure to the stressor in order to habituate them to handling and novelty.

Restraint stressor protocols are described in another entry in this encyclopedia (see **Restraint Stress**). Electrical shock as a stressor has been used mostly in rodents and includes both foot shock and tail shock. One of the advantages of shock is that stressor intensity can be graded monotonically while duration and other parameters are held constant. With the other stressors listed, it is more difficult to manipulate and assess stressor intensity and to prevent the confounding effects of stressor duration. Foot shock is administered in Skinner boxes with electrified grids on the floor, and the animals are freely moving. Tail shock is administered to restrained animals through electrodes attached to the tail. One advantage of foot shock is that animals are freely moving, but this also creates a problem because they can learn to reduce the amount of shock they receive by assuming particular stances on the grid. Tail shock permits us to administer uncontrollable, inescapable shock, to ensure the delivery of the same amount of shock to each rat, and to verify the intensity of each shock.

Swim stressors are now used more frequently in rodents because Animal Welfare Committees are increasingly reluctant to approve shock stressors. The rodents are placed in a container of water for 5–30 min with either running water or a weight to prevent floating. The intensity of stress is manipulated by the duration of the swimming and the temperature of the water; the colder the water, the more intense the stressor. Swimming is a compound stressor with an exercise or exertion component and an aversive component involving cold temperatures and being wet. Exposure to cold alone, without swimming, is also used as a stressor. However, when repeated or long-term cold exposure is used, there are physiological adaptations to low temperatures that may be independent of stress responses and confound the interpretation of findings; for example, the thyroid hormone axis is generally suppressed by stressors but stimulated by long-term exposure to cold. Finally, loud noises around 120 dB have been used as stressors, but some species will rapidly adapt during a session of repeated or continuous exposures to this level of sound.

Psychological stressors There are also acute stressors that are thought to produce more psychological than physical stress, such as novelty and a variety of social stressors. Novelty stressors can entail simply moving the home cage to a different room, transferring the animal to a bare cage or a cage with unused bedding, or exposing the animal to unfamiliar objects. These stressors are listed in the order that is generally thought to reflect increasing stressor intensity. Some acute social stressors that have been used include exposure to predators, exposure to conspecifics in species with dominance hierarchies, and exposure to pheromones from more dominant conspecifics. Finally, the division of stressors into physical and psychological is mostly heuristic in that all physical stressors clearly have a major psychological component and psychological stressors often have a physical component, as when some fighting is allowed between conspecifics in social conflict paradigms. Conversely, it has been shown that anticipation of physical stressors, such as shock or immobilization, imparts a strong psychological component to these stressors.

Nonstressed controls In using animal models of stress, it is important to select appropriate control groups. If the stress effect is small or we are studying complex interactions between stress and other factors, it is important to provide the most stress-free environment for the control animals. Animals in stress studies are often maintained in large vivariums without an appreciation of the stressfulness of this housing. Loud noises in general and particularly during animal care, handling during cage changes, and other experimenters working in the same room often lead to elevated stress hormones in control rats. In addition, data collection must be as stress-free as possible, for example, either very rapid or through indwelling catheters to ensure basal hormone levels in controls. It may also be important to house non-stressed controls apart from stressed animals because stressed rats may bring back stress-specific odors to their home cages.

Acute stress responses and stressor intensity The acute stress response is characterized by the changes in the endocrine systems previously described. One critical question in experimental models is how to evaluate and compare the stressfulness of different stressors. One possibility is to study responses to a graded series of stressor intensities and determine which stress-responsive systems respond in a monotonic way to increasing stressor intensity. For example, one study has shown that acute prolactin responses to foot shock are a better index of stressfulness than

adrenocortical hormones or catecholamines because the responses of the latter are not as well graded; that is, they tend to be all-or-nothing types of responses.

Chronic Stress

Chronic stressors Commonly used chronic stressors include repeated exposures to the acute stressors previously listed, continuous performance tasks, overcrowding, and complex social situations when appropriate for the species. Repeated exposures to the same stressor can lead to habituation, that is, smaller responses over time and eventually no response to the stressor. Two approaches have been used to prevent this habituation. In the first, different acute stressors are alternated on consecutive days. The second approach uses relatively few exposures to a very intense stressor that prevents habituation and may actually produce sensitization, that is, larger responses to the stressor over time. Whether habituation or sensitization occurs can be related to the intensity of the stressor inasmuch as animals will habituate to low-intensity shock stress and sensitize to the same shock stress at higher intensities.

Continuous performance tasks are generally operant tasks that require frequent responses to avoid aversive stimuli or, in a few cases, to obtain positive reinforcement. Frequent motor responses are one aspect of this stressor, but another is the requirement for frequent or continuous attentiveness over days. Increasing the number of animals (three or more) in a cage is stressful, but so is single housing of social species. Finally, several chronic stress models involve the daily changing of cagemates in species with dominance hierarchies, which leads to the continual disruption of and need to reestablish the hierarchies. Most recently, a few laboratories have added naturalistic environments to the disruption of dominance hierarchies and social interactions. In these chronic social stressors, rank in the hierarchy usually determines stressfulness with more dominant individuals generally showing less stress.

Finally, there is an extensive literature on two other specific animal stress models: learned helplessness and fear conditioning. The learned helplessness paradigm was used in the past as a model to understand the physiological underpinnings of depression and to screen compounds for antidepressant activity. In this model, an animal was first exposed to inescapable shock, and then it failed to escape subsequent shock when given the opportunity. Antidepressants will reverse this failure to escape, and so the paradigm was used to screen novel compounds for antidepressant activity. More recently, inescapable stress without the

helpless behavior has been used for animal models of depression.

Another model of psychological stress is fear-conditioned behavior. In these models, an animal is stressed in a particular context and later returned to the context without the stressor. On reexposure to the context, the animal becomes hyperresponsive, which may be manifested in fear-potentiated startle, freezing behavior, or elevated stress hormones (e.g., glucocorticoids).

Chronic stress responses The chronic stress state is characterized by persistent changes in physiology and behavior that last beyond the time when the acute stress response should have dissipated. Posttraumatic stress disorder is such a state whose effects can last for years or decades after the precipitating traumatic stressor. In animals, most persistent stressor effects are followed for days or weeks, although there are neonatal stress regimens whose effects have been reported to last throughout the life of the animal. It remains unclear how physiological systems that constantly adapt to changes in the environment maintain a new set point for a such long time after some stressful situations. For example, the plasma glucocorticoid response to an acute stressor is well understood; a multilevel, negative feedback system returns glucocorticoid levels to baseline 1–3 h after most stressors. However, baseline glucocorticoid levels can remain elevated for 3–5 days after repeated exposure of rats to tail shock. It is unclear whether this is due to persistently increased activity in the stress-mediating areas of the brain leading to increased CRH release or persistent changes within the HPA axis itself that are only reversed after several days.

Most animal models of chronic stress are characterized by persistent neuroendocrine, immunological, and behavioral changes. The entire HPA axis is activated, the thyroid and reproductive axes are suppressed, growth hormone is suppressed, and prolactin is stimulated. Most studies report immune suppression after chronic stress, but there can be enhanced immune function; for example, chronic stress will exacerbate pathology in experimental allergic encephalomyelitis, an autoimmune animal model for multiple sclerosis. The behavioral changes that are usually seen include hyperarousal indicated by increased startle responses and fear of novelty indicated by decreased exploration in an open field. In addition, chronic stress can inhibit spatial memory in the Morris Water Maze, but it can enhance the acquisition of a classically conditioned eye-blink response. Thus, the effects on learning and memory may depend on the kind of stressor and the type of learning being tested.

Psychological modifiers In addition to these behavioral effects of stressors, a number of psychological variables will affect the stress response. In most studies, the ability or perception of the ability to control the occurrence of stressors produces smaller stress responses. In these studies, one animal has control of stressor presentation and a second is yoked so that it receives the same stressors but without control; yoked animals show larger responses. The ability to perform coping or escape responses during stressors also has a major role in determining stressfulness; animals permitted coping reactions or escape responses show smaller stress responses than those that are not. The predictability of a stressor can also be manipulated so an animal can anticipate an upcoming stressor. In most experiments, unpredictable stressors result in greater responses, but one study of rats exposed to stressors in their home environment showed that predictable stress can lead to increased chronic stress responses.

Summary

This short description of animal models for human stress cannot be comprehensive because so many different types of stressors are used, the effects even with the same stressor are often dependent on the particular stress paradigm, and there are important differences among animal species in their responses to particular stressors. Instead, this article illustrates the wide range of stressors that have been used and the diversity of the stress respondents that have been studied. When studying stress, it is important to remember that many, if not all, of the stress-responsive systems are reacting to a stressor even if the experimenter is measuring only one respondent. Finally, animal models of stress have been critical in understanding how the human body responds to stress, what the health consequences of those responses are, and what therapies can be used to treat stress.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Acute Stress Response: Experimental; Posttraumatic Stress Disorder, Neurobiology of; Primate Models, Overview; Restraint Stress.

Further Reading

- Cameron, O. G. (1994). *Adrenergic dysfunction and psychobiology*. Washington, DC: American Psychiatric Press.
- Chrousos, G. P., McCarty, R., Pacak, K., et al. (1995). Stress, basic mechanisms and clinical implications. *Annals of the New York Academy of Sciences* 771.
- Csermely, P. (1998). Stress of life from molecules to man. *Annals of the New York Academy of Sciences* 851.

- Fabris, N., Markovic, B. M., Spector, N. H., et al. (1994). Neuroimmunomodulation, the state of the art. *Annals of the New York Academy of Sciences* 741.
- Goldberger, L. and Breznitz, S. (1982). *Handbook of stress, theoretical and clinical aspects*. New York: Macmillan.
- Hernandez, D. E. and Glavin, G. B. (1990). Neurobiology of stress ulcers. *Annals of the New York Academy of Sciences* 597.
- Moberg, G. P. (1985). *Animal stress*. Bethesda, MD: American Physiological Society.
- Locke, S., Ader, R., Besedovsky, H., Hall, N., Solomon, G. and Strom, T. (1985). *Foundations of psychoneuroimmunology*. New York: Aldine.
- Schneiderman, N., McCabe, P. and Baum, A. (1992). *Perspectives in behavioral medicine, stress and disease processes*. Hillsdale, NJ: Lawrence Erlbaum.
- Selye, H. (1980/1983). *Selye's guide to stress research* ((vols. 1–2). New York: Van Nostrand Reinhold.
- Shapiro, A. P. (1996). *Hypertension and stress, a unified concept*. Mahwah, NJ: Lawrence Erlbaum.
- Tache, Y. and Rivier, C. (1993). Corticotropin-releasing factor and cytokines: role in the stress response. *Annals of the New York Academy of Sciences* 697.
- Yehuda, R. and McFarlane, A. C. (1997). Psychobiology of posttraumatic stress disorder. *Annals of the New York Academy of Sciences* 821.

Annexin

A Mulla and J C Buckingham

Imperial College School of Medicine, London, UK

R J Flower

St. Bartholomew's and the Royal London School of Medicine and Dentistry, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 623–628, © 2000, Elsevier Inc.

Structure of Lipocortin 1

Distribution of Lipocortin 1

Regulation of the Expression and Cellular

Disposition of Lipocortin 1

Functions of Lipocortin 1

Cell Surface-Binding Sites for Lipocortin 1

Glossary

Host defense system

A complex system that enables the organism to defend itself from tissue injury, inflammation, or infection by innate (natural) and acquired (specific) mechanisms.

Signal sequence

An N-terminal hydrophobic amino acid sequence that directs newly translated proteins to the Golgi apparatus for packaging into secretory vesicles, post-translational processing, and release by exocytosis.

Glucocorticoids (GCs) exert diverse actions in the body, which are required for the maintenance of homeostasis. Their complex effects are mediated mainly by mechanisms involving the induction or repression

of key target genes. Lipocortin 1 is the product of one such gene; it fulfills a prominent role in effecting the regulatory actions of the steroids in the host defense and neuroendocrine systems.

Structure of Lipocortin 1

Lipocortin 1 (LC1, annexin 1) is a 37-kDa member of the annexin superfamily of Ca^{2+} and phospholipid-binding proteins. Annexins are found in a wide range of organisms (vertebrate, invertebrate, and plant) and are characterized structurally by a highly homologous core domain that normally comprises a fourfold repeat of 70 amino acids and enables the proteins to bind Ca^{2+} and negatively charged phospholipids. The core is attached to the N-terminal domain, which is unique in length and sequence for each member of the family; this domain confers the biological specificity of the individual annexins and thus enables the proteins to exert distinct and specific actions on their target cells. **Figure 1** illustrates the primary structure of LC1; note that the N-terminal domain includes potential sites for phosphorylation, glycosylation, and protease action. LC1 is encoded by a single gene in mammals and in the chicken, but two LC1 genes have been identified in the pigeon. The rat, mouse, and human genes show a high degree of homology; each comprises 13 exons, the first of which is non-coding (**Figure 1**), but no spliced variants have been described to date. A TATA motif 30 bp upstream of the transcription start site is consistent with an inducible and tissue-specific mode of regulation. However, the promoter region of the mammalian gene is characterized poorly, although potential response elements for GCs, activating enhancer binding

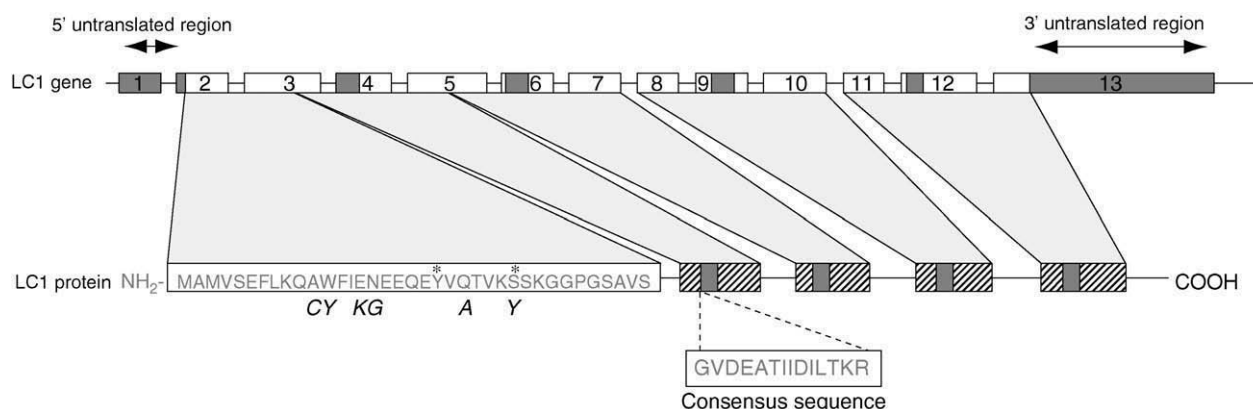


Figure 1 Schematic diagram of the structure of human lipocortin 1 and its encoding gene. *, potential phosphorylation sites; ▨, repeated units with 25–35% sequence homology; ■, consensus sequence in repeated unit. Italics represent the corresponding sequence in rat.

protein (AP)-1, Oct-1, cAMP, and serum have been identified.

Distribution of Lipocortin 1

LC1 is found in many organs and tissues of the body, but is often localized in discrete populations of differentiated cells, particularly epithelia. It is particularly abundant in the lung, thymus, spleen, placenta, and seminal fluid. In addition, many cells involved in the inflammatory response (e.g., macrophages, polymorphonuclear neutrophils, and monocytes) express very substantial amounts of LC1, as do cells of the neuroendocrine system. Within the pituitary gland LC1 is expressed mainly by nonsecretory folliculostellate cells but also by endocrine cells, whereas in the hypothalamus it is expressed most strongly in the median eminence and paraventricular nucleus (i.e., areas associated with the regulation of anterior pituitary function). LC1 is also expressed in lesser amounts elsewhere in the brain where it appears to be confined mainly to microglial cells, although some neuronal expression may also occur.

The subcellular distribution of LC1 is unusual; it is abundant in the cytoplasm, but a small proportion is also found on the external surface of the cell membrane, where it attaches via a Ca^{2+} -dependent linkage to binding proteins. LC1 also binds to the inner leaflet of the plasma membrane, and cell fractionation studies have identified an integral membrane pool of LC1.

Regulation of the Expression and Cellular Disposition of Lipocortin 1

Glucocorticoids

LC1 was first identified as a GC-inducible protein in extracts of conditioned media from rat peritoneal macrophages. Subsequent studies showed that the expression of LC1 mRNA and protein in experimental

animals is reduced substantially by adrenalectomy but that it is restored by maintenance doses of GCs. The administration of the GC-receptor antagonist mifepristone (RU486) also lowers the resting LC1 levels in rodents, thus confirming a role for the steroids in the regulation of LC1 expression.

In addition to inducing the synthesis of LC1, GCs also modify the cellular disposition of the protein in such a way that intracellular (cytoplasmic) LC1 is exported to the outer surface of the cell. This process of exportation, termed externalization, is likely to be an important facet of LC1 biology because it enables the protein to reach specific binding sites on the outer surface of cells that appear to be critical to its action. Externalization is normally evident within 0.5–2 h of the steroid challenge and involves the passage of both preformed and newly synthesized LC1 across the membrane. Intracellular (cytoplasmic) stores of the protein may fall initially, but, ultimately, the positive effects of the steroid on LC1 synthesis exceed the effects on exportation and thus the depleted stores are replenished. The effects of GCs on the synthesis and cellular disposition of LC1 are specific, and other steroids (androgens, estrogens, progestogens, mineralocorticoids, and cholesterol) are without effect.

The mechanism by which LC1 crosses the membrane is unclear, but it is unlikely to involve exocytosis. LC1 is a polar molecule and, although its sequence includes some short hydrophobic regions, it lacks a classical signal sequence required for entry to the vesicular packaging system. Moreover, drugs that block the classical endoplasmic reticulum–Golgi pathway of protein secretion do not affect the cellular exportation of LC1.

Other Factors

Several factors other than GCs are implicated in the regulation of LC1 expression, including interleukin 6 (IL-6) and phorbol esters, which exert positive

influences. The apparently complementary action of IL-6 and the steroids is interesting because there are other examples of these substances acting cooperatively, for example, in the generation of acute-phase proteins in the liver. LC1 in the brain is upregulated readily by tissue injury and excitotoxins, such as the excitatory amino acids glutamate or kainate, but the underlying mechanisms are unknown.

Functions of Lipocortin 1

Control of Inflammation

LC1 exhibits significant anti-inflammatory activity in a number of experimental models of inflammation and is now believed to serve as an endogenous regulator of the inflammatory response. In accord with this view, the passive immunization of animals with neutralizing antisera against LC1 quenches the ability of GCs to suppress edema formation and neutrophil migration in several models of inflammation; it also exacerbates brain damage caused by middle cerebral artery occlusion in a rat model of stroke. Conversely, human recombinant LC1 (hu-r-LC1) and peptides derived from it have beneficial effects in several models of inflammation. Interestingly, autoantibodies to LC1 have been identified in the plasma of some patients with chronic inflammatory disease, such as rheumatoid arthritis and systemic lupus erythematosus. Moreover, there is evidence of a causal relationship between the antibody titers and the incidence of certain types of steroid resistance, particularly in cases in which higher doses of steroid are still able to repress the symptoms of the disease. However, not all the anti-inflammatory actions of GCs can be attributed to LC1 and, in some instances, GCs are effective but LC1 is not; for example, steroids inhibit the generation of both carrageenin- and histamine-induced paw edema in the rat, but LC1 is effective only against carrageenin.

Some of the anti-inflammatory actions of LC1 are almost certainly explained by the ability of the protein to prevent the generation of proinflammatory eicosanoids, but others, such as the inhibition of IL-1- or IL-8-induced leukocyte migration, cannot be accounted for in this way. Early work in the field suggested that LC1 blocks the release of arachidonic acid from membrane phospholipids by inhibiting the enzyme phospholipase A2 (PLA2). Later data, however, refuted this view and, although there is good evidence that the inhibition of arachidonic acid release induced by LC1 is secondary to a reduction in cPLA2 activity, the exact mechanism involved remains to be determined; an effect on a signal transduction cascade upstream of the enzyme is a distinct possibility.

Lipocortin 1 and the Neuroendocrine System

LC1 fulfills a significant role as a mediator of GC action in the hypothalamic-pituitary-adrenocortical (HPA) axis, acting mainly at the levels of the hypothalamus and anterior pituitary gland but also possibly at other sites in the brain (e.g., the hippocampus) to effect the negative feedback actions of the steroids on the axis. *In vivo* central or peripheral administration of anti-LC1 antisera effectively negates the ability of GCs (corticosterone or dexamethasone) to block the overt increases in adrenocorticotrophic hormone (ACTH) secretion induced by IL-1 β , whereas central injections of hu-r-LC1 abolish the rises in GC secretion induced by either IL-1 β or IL-6. Similarly, the inhibitory actions of GCs on the evoked release of the corticotropin releasing hormone (CRH) from the hypothalamus and ACTH from the anterior pituitary gland *in vitro* are mimicked by hu-r-LC1 and are reversed specifically by anti-LC1 antisera and LC1 antisense oligodeoxynucleotides (ODNs) directed against sequences specific to the LC1 gene.

LC1 also contributes to other aspects of GC action within the neuroendocrine system. In particular, it acts at the pituitary level to mediate the inhibitory actions of the steroids on the secretion of prolactin by the lactotrophs; in addition, it acts within the hypothalamus to mediate the acute stimulatory actions of the steroids on the secretion of GH.

Other Actions of Lipocortin 1

In addition to its effects on inflammatory responses and on the neuroendocrine system, LC1 may also contribute to other facets of GC action, including the manifestation of antipyresis, the suppression of growth and differentiation of some cells, and the modulation of the expression of some genes. In addition, like other annexins, LC1 has been implicated in the intracellular processes of vesicular fusion, particularly those concerned with endocytosis.

The suppressive actions of LC1 on cell growth have been studied most widely in A549 cells. This adenocarcinoma cell line requires prostaglandins for growth and, thus, when prostanoid production is inhibited (e.g., by administration of GCs or drugs that inhibit cyclo-oxygenase), cell growth is slowed or even halted. In contrast, cell growth is promoted readily by epidermal growth factor (EGF), which acts via a tyrosine kinase receptor to liberate prostanoids via a complex cascade involving the activation of cPLA2. The inhibitory effects of GCs on EGF-stimulated cell growth are associated with an increase in LC1 expression; moreover, the suppression of cell growth and prostanoid generation induced by the steroids are

mimicked by LC1 but reversed by exposure of the cells to anti-LC1 antisera or anti-LC1 ODNs. LC1 serves as a substrate for and is phosphorylated by the EGF receptor; it may thus produce its effects on cell growth and eicosanoid generation by disrupting the grb-, sos-, and mitogen-activated protein (MAP) kinase signaling cascade used by the EGF receptor to stimulate cPLA2 activity. However, the point at which LC1 disrupts this cascade remains to be defined.

The possibility that LC1 may contribute to the inhibitory effects of GCs on the expression of the inducible forms of cyclo-oxygenase (COX-2) and nitric oxide synthase (iNOS) has been investigated. COX-2 expression appears to be independent of LC1. However, in rat models of endotoxic shock, the ability of GCs to block the fatal hypotension that develops within hours of the administration of lipopolysaccharide is blocked by neutralizing anti-LC1 antisera, as is the expression of iNOS (assessed by western blot analysis or measurement of nitrate release), suggesting that LC1 is at least partially responsible for the downregulation of iNOS induced by the steroids. The molecular basis of this action is obscure, but it almost certainly involves the disruption of a key signal transduction system rather than a direct genomic action of LC1.

Lipocortin 1-Derived Peptides

An interesting and important discovery has been that at least some facets of the biological activity of LC1 can be attributed to sequences in the N-terminal of LC1 of the molecule. For example, a peptide comprising residues 2–26 (LC1_{Ac2-26}) has a profile of anti-inflammatory and antiproliferative activity similar to that of the full-length recombinant molecule, although it is approximately 200 times less potent. LC1_{Ac2-26} also mimics the effects of the full-length molecule on the neuroendocrine system, although again it is less potent. Shorter peptides may also be active. These peptides have the advantage that they are easier to work with than the full-length protein and inexpensive to produce; they thus provide valuable tools for research and may ultimately provide the base from which LC1 peptides or nonpeptide analogs may be developed for therapeutic use. The N-terminal peptides do not, however, mimic the effects of LC1 on vesicular aggregation; this property of LC1 requires the repeater sequences in the core and is thus likely to reflect the traditional Ca²⁺- and phospholipid-aggregating properties of the annexins.

Cell Surface-Binding Sites for Lipocortin 1

The preceding sections show that LC1 and the peptides derived from it mimic many of the diverse

regulatory effects of GCs within the host defense and neuroendocrine systems. However, they provide little insight into the mechanisms involved. Several lines of evidence suggest that at least some of the actions of the peptides may be exerted extracellularly, possibly via cell surface-binding sites. For example, the anti-LC1 antibodies that quench many of the actions of the steroids are unlikely to penetrate cells readily; moreover, the steroids themselves cause the cellular exportation of LC1. High-affinity saturable LC1-binding sites have been identified on the surface of a variety of cell types by fluorescence-activated cell (FAC) analysis, including human and rodent peripheral blood leukocytes, rat anterior pituitary cells, and various cell lines (e.g., the rat pituitary GH3 line). Confocal imaging has shown that the distribution of the surface-binding sites is punctate (discrete clumps) and variable; this is consistent with the widespread phenomenon of ligand-induced receptor clustering, a process that may be essential for internalization. The interaction of LC1 with the membrane cannot be explained merely by the binding of the protein to phospholipid because (1) it is abolished by mild trypsinization of the cells and (2) the ability of trypsinized cells to recover their capacity to bind LC1 in culture is prevented by the inclusion of the protein synthesis inhibitors cycloheximide and puromycin in the medium, thus suggesting a requirement for protein. Two candidate membrane proteins of molecular mass 15–18 kDa have been isolated from human monocytes. Their structure and function remain to be elucidated, but preliminary studies suggest that they are unlikely to be a classic seven-transmembrane-domain receptor and are more likely to be a protein involved in membrane organization and turnover.

See Also the Following Articles

Corticosteroids and Stress; Glucocorticoids, Overview; Immunity.

Further Reading

- Ahluwalia, A., Buckingham, J. C., Croxtall, J. D., Flower, R. J., Goulding, N. J. and Perretti, M. (1996). The biology of annexin 1. In: Seaton, B. A. (ed.) *The annexins*, pp. 161–199. Austin, TX: R. Landes and Co.
- Blackwell, G., Carnuccio, R., DiRosa, M., et al. (1980). Macrocortin: a polypeptide causing the anti-phospholipase effect of glucocorticoids. *Nature* **287**, 147–149.
- Buckingham, J. C. (1996). Stress and the neuroendocrine-immune axis: the pivotal role of glucocorticoids and lipocortin 1. Fifteenth Gaddum memorial lecture. *British Journal of Pharmacology* **118**, 1–19.
- Carey, F., Forder, R., Edge, M., et al. (1990). Lipocortin-1 fragment modifies pyrogenic actions of cytokines in rats. *American Journal of Physiology* **259**, R266–R269.

- Christian, H. C., Taylor, A. D., Flower, R. J., et al. (1997). Characterization and localization of lipocortin 1-binding sites on rat anterior pituitary cells by fluorescence-activated cell analysis/sorting and electron microscopy. *Endocrinology* **138**, 5341–5351.
- Croxtall, J. D. and Flower, R. J. (1994). Antisense oligonucleotides to human lipocortin-1 inhibit glucocorticoid-induced inhibition of A549 cell growth and eicosanoid release. *Biochemical Pharmacology* **48**, 1729–1734.
- Goulding, N. J., Pan, L., Wardwell, K., et al. (1996). Evidence for specific annexin I-binding proteins on human monocytes. *Biochemical Journal* **316**, 593–597.
- Horlick, K., Cheng, K., Wong, W., et al. (1991). Mouse lipocortin 1 gene structure and chromosomal assignment: gene duplication and origins of a gene family. *Genomics* **10**, 365–374.
- Kovacic, R., Tizard, R., Cate, R., et al. (1991). Correlation of gene and protein structure of rat and human lipocortin 1. *Biochemical Journal* **30**, 9015–9021.
- McKanna, J. A. (1993). Lipocortin 1 immunoreactivity identifies microglia in adult rat brain. *Journal of Neuroscience Research* **36**, 491–500.
- Pepinsky, R. and Sinclair, L. (1986). Epidermal growth factor-dependent phosphorylation of lipocortin. *Nature* **321**, 81–84.
- Perretti, M., Ahluwalia, A., Harris, J., et al. (1993). Lipocortin 1 fragments inhibit neutrophil accumulation and neutrophil dependent oedema in the mouse. *Journal of Immunology* **151**, 4306–4314.
- Philip, J. G., Flower, R. J. and Buckingham, J. C. (1998). Blockade of the classical pathway of protein secretion does not affect the cellular exportation of lipocortin 1. *Regulatory Peptides* **73**, 133–139.
- Schlaepfer, D. and Haigler, H. (1990). Expression of annexins as a function of cellular growth state. *Journal of Cell Biology* **111**, 229–238.
- Taylor, A. D., Christian, H. C., Morris, J. F., et al. (1997). An antisense oligodeoxynucleotide to lipocortin 1 reverses the inhibitory actions of dexamethasone on the release of adrenocorticotropin from rat pituitary tissue *in vitro*. *Endocrinology* **138**, 2909–2918.
- Violette, S., King, I., Browning, J., et al. (1990). Role of LC1 in the glucocorticoid induction of the terminal differentiation of a human squamous carcinoma. *Journal of Cellular Physiology* **142**, 70–77.
- Wallner, B., Mattaliano, R., Hessian, C., et al. (1986). Cloning and expression of human lipocortin, a phospholipase A2 inhibitor with potential anti-inflammatory activity. *Nature* **320**, 77–81.

Anorexia Nervosa See: Eating Disorders and Stress.

Antibody Response

R J Booth

University of Auckland, Auckland, New Zealand

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R J Booth, volume 1, pp 206–211,

© 2000, Elsevier Inc.

What Is an Antibody Response?

Stress Proteins as Antigens

Stress and Antibody Response in Animals

Stress and Antibody Response in Humans

Stress, Antibodies, and Human Disease

Conclusions

Glossary

<i>Antibodies</i>	Proteins produced by B lymphocytes with the ability to bind specifically to molecular shapes.
<i>Antigens</i>	Molecular shapes, often on infectious agents or produced during infections, that can be recognized by antibodies and other immune receptors.
<i>Autoantibodies</i>	Antibodies able to bind to normal components of the body.
<i>B Lymphocytes</i>	Cells produced by the bone marrow and exported to secondary lymphoid organs where they are able to respond to antigens by producing antibodies specific to them.

<i>Immuno-globulins</i>	A synonym for antibodies. Different classes of immunoglobulins function in different ways. IgA is important in secretions, IgG and IgM are important in blood and other internal body fluids, and IgE is important in allergy.
<i>Stress proteins</i>	A family of molecules produced by cells and organisms under stress. These are sometimes called heat shock proteins and are highly conserved across many different organisms.
<i>T Lymphocytes</i>	Cells produced by the thymus and exported to secondary lymphoid organs where they are able to respond to antigens. A subset of T lymphocytes called helper cells (Th cells) assists B lymphocytes to produce antibodies.

What Is an Antibody Response?

Antibodies are produced as part of the immune system's behavior toward antigens of foreign material such as infectious agents like bacteria and viruses. They are soluble proteins that have the ability to bind specifically to shapes on particular antigens and to neutralize them or to lead to their inactivation and removal (see [Figure 1](#)).

Antibodies are produced by a class of lymphocytes called B lymphocytes that are derived from the bone marrow and active in secondary lymphoid organs such as lymph nodes, spleen, tonsils, and other gut-associated lymphoid organs. There are many millions of B lymphocytes in the body, but each B lymphocyte has the capacity to make antibodies with only one shape of antigen-binding site. In order to be activated to produce antibodies against a particular antigen, B lymphocytes must interact with antigens through specific cell surface receptors (see [Figure 2](#)) and usually also require assistance from a class of T lymphocytes (derived from the thymus) called helper T lymphocytes (Th cells). Some antigens, such as lipopolysaccharide (LPS) cell wall products of some

bacteria, can stimulate B lymphocytes to make antibodies without much in the way of T cell help and are called T-independent antigens.

Stress Proteins as Antigens

When cells or microorganisms are subjected to physical stresses such as changes in temperature or the composition of the medium in which they exist, they increase synthesis of stress proteins (or heat shock proteins). These stress proteins represent a highly conserved cellular defense mechanism mediated by multiple, distinct gene families that participate in many essential biochemical pathways of protein maturation and function during both times of stress and normal cellular homeostasis. They are thought to operate to protect the stressed cells by scavenging abnormal proteins and protecting newly synthesized proteins through a chaperone function. They are also strongly immunogenic and so play a prominent part in immune responses against microorganisms. For example, several protein antigens involved in antibody and other responses to mycobacteria have been identified as members of highly conserved heat shock protein families, suggesting that immune responses against microorganisms may be dominated by responses against the stress protein family. Furthermore, these highly immunogenic microbial stress proteins have very similar structures to stress proteins in human cells. Through this molecular mimicry, immune cross-reactions between microbial and human stress proteins may be involved in some chronic inflammatory and autoimmune disorders such as rheumatoid arthritis, ankylosing spondylitis, systemic lupus erythematosus, and inflammatory bowel diseases. Liaison between heat shock proteins and the immune response may therefore be both beneficial to the host as antigenic targets and detrimental to the host as promoters of autoimmunity. Psychological stress in infancy, associated with strain in families, for example, can be involved in the induction or progression of diabetes-related autoimmunity.

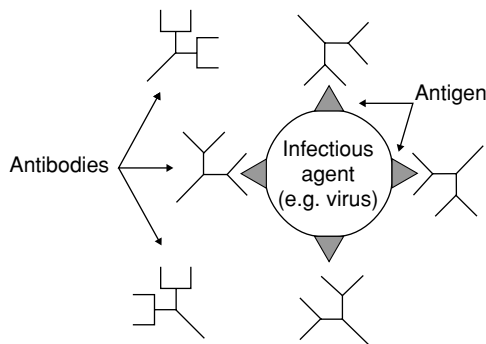


Figure 1 Antibodies interacting with specific antigenic shapes through their antigen-binding site.

Stress and Antibody Response in Animals

A considerable body of research has investigated the effects of stress on antibody responses in animals and, as summarized in [Table 1](#), a variety of factors have been found to be important.

When rats or mice are subjected to physical stressors such as restraint or exposure to coldness, the effect on antibody responses depends on whether the stressor is short- or long-term. For example, short periods of restraint stress (2 h over 2 consecutive days) enhance the primary antibody response to an

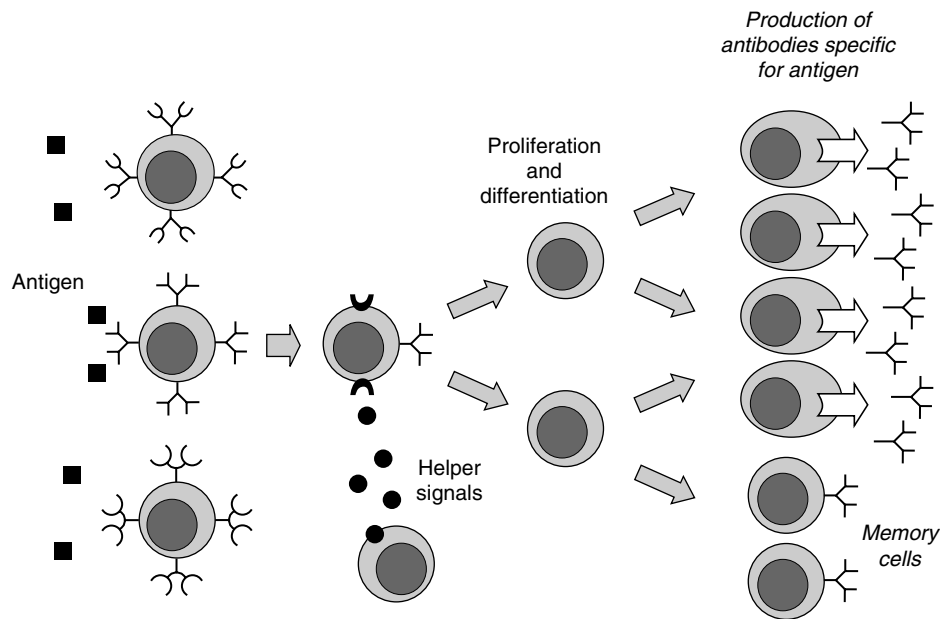


Figure 2 B lymphocytes in secondary lymphoid organs display surface immunoglobulins with antigen-binding specificity of the antibodies they are capable of producing when stimulated. Through these surface immunoglobulin receptors, antigens select appropriate B lymphocytes from the antigen-sensitive repertoire. Further activation signals come from helper T lymphocytes. The result is the proliferation of activated B lymphocytes and the production of antibody-forming cells and memory cells that are capable of responding more rapidly to antigens in future.

Table 1 Factors affecting the relationship between stress and antibody responses in animals

Types of stresses or stressors
Acute versus chronic stress
Controllability of stress
Social or psychological versus physical stressors
Activation of B lymphocytes versus synthesis of antibodies
Primary versus secondary antibody responses
Route of immunization
Timing and duration of stressor relative to antigen
Balance of neuroendocrine hormones
Season of the year

injected antigen in mice, while longer periods of restraint (6 h over 4 days) do not. The augmented response occurs in both immunoglobulin M (IgM) and IgG antibody classes. The effect is mediated at least in part by hormones from the adrenal cortex and can be blocked by α -adrenoceptor agonist drugs such as phentolamine and potentiated by β -adrenoceptor agonists such as propranolol. The temporal relationship between the stressor and antigen challenge is also important. Antibody responses are markedly suppressed in mice restrained before antigen injection but not significantly affected when the animals are stressed after immunization. Moreover, stresses in mice (such as restraint or social confrontation) diminish primary antibody responses but have little effect on secondary (or memory) responses.

An important group of hormones involved in stress-induced inhibition of antibody (and other) immune responses are the corticosteroids. These are produced by the cortex of the adrenal gland in response to stress-induced signals from the hypothalamus via the pituitary gland. Injection of the hypothalamic hormone corticotropin-releasing factor (CRF) into a rat 20 min before antigen administration diminishes both primary and secondary antibody responses, while CRF infusion 24 h after antigen does not affect antibody levels, suggesting that CRF (and corticosteroids) affect the initial induction phase of the antibody response. Corticosteroids have a direct inhibitory effect on B lymphocyte activation but also act on T cells to diminish helper activity. Conditions that lead to inhibition of antibody responses against T helper-dependent antigens often lead to enhanced responses against T-independent antigens, and adrenalectomy blocks both these effects by removing the gland that produces corticosteroids.

In addition to corticosteroids, two other types of hormone mediate stress effects on antibody responses. The pituitary hormone prolactin generally facilitates antibody responses and so acts as a balance to corticosteroid inhibitory effects. In addition, the pineal hormone melatonin counteracts many stress responses including antibody inhibition. In doing this, melatonin appears to act on Th cells to induce the release of opioid peptides with immuno-enhancing

properties. Melatonin and opioid effects therefore may come into play to effect immune recovery following corticosteroid-induced immune depression that results from stressful situations.

We cannot, however, consider the effects of stress on antibody responses simply in terms of enhancement or suppression because response kinetics can be different under stressed and nonstressed conditions. For example, in some studies, repeated cycles of physical restraint stress in mice delayed the development of antibodies to a virus infection but did not affect the ultimate magnitude or class of antibodies produced. Such a delay could nevertheless have consequences for the course of an infection and allow prolonged multiplication of an infectious agent, leading to the development of illness symptoms. A further factor affecting stress effects on antibody responses is the season of the year during which the stress occurs. In rats, unstressed animals display increased antibody responsiveness in spring and summer and decreased responsiveness in autumn and winter. This annual rhythm is accentuated in stressed animals and might therefore contribute to greater stress-induced susceptibility to infection during the winter months.

Although much animal experimentation has focused on physical stressors, social stress can also affect antibody responses in animals. For example, rats exposed to the odors of other stressed rats exhibit altered antibody responses as if they themselves had been stressed. Also, defeat associated with territorial defense affects a rat's ability to generate antibodies in response to an injected antigen. The more time such rats spend in submissive postures, the lower their antibody levels. Social status in monkeys also affects antibody-related factors. Low social status is associated with lower body weight, greater elevated cortisol responses to social reorganizations, and less aggressive behavior. Correspondingly, low-status monkeys have diminished immune responsiveness and a greater probability of being infected by upper respiratory tract viruses.

Clearly, a variety of factors mediate effects of stress on antibody production, but these effects occur as part of a bidirectional network of interactions between the immune system and the nervous system rather than a one-way, cause-effect process. This means that immune activity, or defects in immunity, can also influence the way animals behave. As an example, certain inbred strains of mice have a propensity to develop autoimmunity and produce high concentrations of antibodies against some of their own (self) antigens. In these mice, there is an association between high titers of autoantibodies and emotional behaviors resembling those of stressed animals. Disturbed emotional reactivity in these strains may

reflect the effect of autoimmunity on the hypothalamic-pituitary-adrenal axis, highlighting the bidirectional nature of the communication pathways between immune and nervous systems. The immune systems of animals given low doses of immunomodulatory drugs, such as cyclophosphamide, are subsequently more reactive to stress. This finding has implications for management of immunocompromised patients whose antibody responses may be excessively sensitive to stress-induced modulation.

Stress and Antibody Response in Humans

Investigations into the effects of stress on human antibody responses have tended to focus less on deliberately applied physical stressors (as in the animal experiments) and more on stressful situations in which people find themselves. Two particularly fruitful situations for this area have been students undergoing university examinations and people giving constant care to relatives suffering from chronic dementia. For example, when saliva samples from healthy students are compared a week before, 2 weeks after, and during examinations, antibodies normally present in saliva (the IgA class) are lower during the examination period. Moreover, the better a student's social support during the pre-exam period, the higher his or her salivary IgA during examinations. Although the psychologically perceived stress associated with examinations quickly diminishes after the examinations, the reduction of salivary IgA levels can persist for weeks. The stress of examinations is also associated with increased norepinephrine concentrations in saliva, particularly in students most aroused adrenergically by the examination. Antibody titers against Epstein-Barr virus (EBV) also increase during examination periods, indicating poorer immune control of latent virus reactivation.

IgA is the dominant antibody isotype not only in saliva but also on all mucosal surfaces where it acts as a first line of defense against microbial invasion. Because it is relatively noninvasive to measure, salivary IgA has been assessed in a variety of stressful situations aside from examinations. Similar to the conclusions from animal studies, it seems that in humans, some short-term stresses promote antibody production (at least in the saliva), while others, particularly longer-term stresses, depress it. For example, in one study, salivary IgA in professional soccer coaches was found to increase transiently during a match, while in another, nurses who reported more frequent episodes of anxiety had significantly lowered salivary IgA secretion rates. Such effects can be moderated by other factors, such as a sense of humor. Mild stress, as measured by the frequency of daily

hassles, correlates with diminished salivary IgA concentration, especially in people with low scores on a humor scale.

When salivary IgA is produced in response to challenge with a novel antigen, similar results are found. Following oral immunization, adults who report more desirable events on a day-to-day basis have higher concentrations of antigen-specific salivary IgA than when they report more undesirable events. Undesirable events appear to lower antibody levels primarily by increasing negative mood, whereas desirable events increase antibody levels by decreasing negative mood. These effects can be moderated by using coping skills such as deliberate relaxation. Significant increases in salivary IgA concentration occur in healthy volunteers after relaxation even though no change is observed in salivary corticosteroids.

Not only IgA antibodies but also IgG and IgM antibody classes in the blood are affected by stress. For example, in women immunized with a novel antigen, keyhole limpet hemocyanin, those reporting more stressful events in their lives tended to have lower baseline and 3-week postimmunization serum IgG antibody levels. In contrast, women reporting more good events tended to have higher IgG levels. Psychological distress scores correlated negatively and psychological well-being scores correlated positively with IgG antibody concentrations.

Similar effects are observed when people are immunized with more clinically relevant antigens such as hepatitis B vaccine. In one study, the antibody response of medical students given a series of three hepatitis B vaccinations around the time of major examinations depended on stress and anxiety levels. Those students who produced antihepatitis B antibodies after the first injection were significantly less stressed and anxious than those who did not. In addition, students who reported greater social support demonstrated higher titers of antibodies against hepatitis B antigens. Antibody responses to a variety of different vaccines, such as hepatitis B recombinant vaccine, meningococcal C conjugate vaccine, and influenza multiple subunit vaccine, have all been found to be diminished in people either living in highly stressful situations (e.g., caregivers for chronically ill relatives) or with high levels of daily perceived stress (measured through psychological stress questionnaires). Moreover, the detrimental effect of stress on the development and maintenance of antibody responses following vaccination can be alleviated by social support and by behavioral stress management interventions, suggesting that the context in which stressful events occur for a person is important in determining their effects on immune function. Also, the temporal relationship between perceived stresses

and their effects on antibody responses to vaccination is not straightforward. For example, in one study of healthy young adults given influenza vaccine, stress ratings on the 2 days before and on the day of vaccination were not associated with effects of antibody responsiveness, but stress during the 10 days following vaccine administration appeared to shape the long-term antibody response.

The context in which a stressful event occurs for a person is important in determining its effect on immune function. Events may be perceived as stressful, but if they have additional meaning or significance for the person undergoing them this may qualify their effects. For example, air traffic control is often considered to be a stressful occupation because of the intensity, responsibility, and unpredictability of the workload. Yet when psychophysiological stress reactions were assessed in air traffic controllers before and after each of two working sessions, the working sessions caused a marked increase in the concentration and secretion rate of salivary IgA, in contrast to the expected immunosuppressive effects. Corticosteroid but not salivary IgA levels were correlated with the amount of actual or perceived workload, indicating that positive emotional engagement rather than a negative stress effect may have been responsible for the observed changes.

The importance of context is also apparent from studies of emotional disclosure. When volunteers are asked to write anonymously and confidentially about traumatic events in their lives they generally report the process to be quite upsetting and stressful. If, however, through their writing they create a new relationship with the traumatic events, the effect on immune function appears supportive rather than inhibitory. This is exemplified by a study in which volunteers wrote about emotionally upsetting events every day for 4 days and then were immunized with hepatitis B vaccine. Compared with people in a control group who wrote about mundane topics, participants in the emotional expression group showed significantly higher antibody levels against hepatitis B antigens 4 and 6 months after their initial vaccination. Another study reported significantly lower EBV antibody titers (suggesting better cellular immune control over the latent virus) in people after a stressful emotional disclosure intervention.

By contrast, the stress of caring for a relative with progressive dementia such as Alzheimer's disease is a relatively hopeless chronic situation with little prospect of abatement. Following vaccination with an influenza virus vaccine, such caregivers exhibit a poorer anti-influenza antibody response and virus-specific T cell response compared with the control subjects, although, as discussed earlier, these effects

can be diminished by participation in stress management interventions. Greater impairment in the dementia victim is associated with greater depression and loneliness in caregivers. In 2001, in a critical review of the literature on psychological stress and antibody response to immunization in humans, Cohen and colleagues concluded that the evidence supporting an association between psychological stress and suppression of the antibody response is convincing for secondary responses but less impressive for primary responses. Since then, however, a number of studies have demonstrated reduced primary responses to vaccination, particularly in elderly people suffering high levels of chronic stress. Cohen and colleagues also concluded that health practices do not appear to mediate the relationship between stress and antibody responses to vaccination but that there are indications that elevated cortisol levels in stressed people may play a role.

Stress, Antibodies, and Human Disease

Extrapolating from the effects of stress on responses to vaccination, we might also expect there to be a relationship between stress and infectious disease in humans. A variety of studies show this to be so, particularly in the case of upper respiratory tract infections. Family functioning affects the frequency of influenza B infection, with indices of infection being significantly associated with both family cohesion and adaptability. This raises the possibility that family dysfunction may lead to altered immune responses, which in turn increase susceptibility to infection. A study of a large number of infants, children, adolescents, and young adults in a rural Caribbean village found that abnormal glucocorticoid response profiles, diminished immunity, and frequent illness were associated with unstable family or household conditions. Family relationships and concomitant stress and immunosuppression may therefore be important intermediary links between socioeconomic conditions and child health. In support of this, a study of preschool children observed that a subset of children (labeled as reactive) sustained higher illness rates under conditions of high family stress but lower rates in low-stress conditions.

Non-family-based stresses are also important in susceptibility to infections, as demonstrated by a series of studies from Sheldon Cohen and colleagues in which volunteers were asked to complete questionnaires about perceived stress before being intentionally inoculated intranasally with a cocktail of viruses associated with upper respiratory tract infections. Psychological stress was associated in a dose-response

manner with an increased risk of acute infectious respiratory illness following virus inoculation, and this risk was attributable to increased rates of infection rather than to an increased frequency of symptoms after infection.

Subsequently, Cohen and colleagues repeated the experiment with a new group of volunteers, but this time, in addition to the perceived stress questionnaire, they asked people to answer questions about their interpersonal and social support systems before they sprayed the virus concoction into their noses. Again, they observed a relationship between high levels of perceived stress and increased likelihood of infection, but additionally found the opposite relationship between social support and illness. Those people with more types of social ties were less susceptible to common colds, produced less mucus, were more effective in ciliary clearance of their nasal passages, and shed less virus. The presence of a diverse social network therefore buffers the potentially adverse effects of perceived stress with respect to susceptibility to infection. Further work is required to determine the role of virus-specific antibody responses in these effects.

Finally, autoimmune illnesses typically involve high titers of antibodies against various normal components of the body, and severity is usually correlated with autoantibody titer. There are reports that stress, depression, anxiety, and anger are associated with symptomatology in patients with various autoimmune illnesses, although to date, there has been little work directly linking stress, antibody levels, and illness severity. One study that has done so investigated antinuclear autoantibody titers in women and found a significantly higher incidence in women who had been sexually abused as children. These findings suggest that sexually abused girls may show evidence of an alteration in normal immune homeostatic function.

Conclusions

We have seen from the animal studies that stress affects antibody responses in different ways depending on the animal, the type and timing of the stressor, and its context. Some, but not all, of the effects appear to be mediated by the classical stress hormones of the hypothalamus-pituitary-adrenal axis. In order to make sense of the relationship between stress and antibody responses in humans, however, we must look beyond a purely biological or physical construction of the immune system and link its characteristics to psychological, social, and cultural processes. Kevin Ashbridge and Roger Booth have proposed that the immune system is engaged in the process of self-determination and therefore shares with the

neurological and psychological domains the goal of establishing and maintaining self-identity. From this perspective, the framework in which immune recognition occurs cannot be considered only as the context in which cells, receptors, and antigens operate, but as all the domains of life in which individuals define themselves. As a result, the meaning of events surrounding an antigenic stimulus (e.g., exposure to an infectious agent) is likely to condition the features of any observed antibody response. As discussed previously, this is particularly apparent in humans, in whom the effects of stressful events on antibody responses are dependent upon the context of meaning of those events.

See Also the Following Articles

Acute Stress Response: Experimental; Autoimmunity; Corticosteroids and Stress; Cytokines; Cytotoxic Lymphocytes; Immune Cell Distribution, Effects of Stress on; Immune Function, Stress-Induced Enhancement; Immune Response; Immune Suppression; Infection; Lymphocytes; Neuroimmunomodulation; Pituitary Regulation, Role of; Psychoneuroimmunology; Psychosocial Factors and Stress.

Further Reading

- Booth, R. J. and Ashbridge, K. R. (1993). A fresh look at the relationship between the psyche and immune system: teleological coherence and harmony of purpose. *Advances in Mind-Body Medicine* 9, 4–23.
- Burns, V. E. and Carroll, D. (2003). Life events, perceived stress and antibody response to influenza vaccination in young, healthy adults. *Journal of Psychosomatic Research* 55, 569–572.
- Cohen, S. and Williamson, G. M. (1991). Stress and infectious disease in humans. *Psychological Bulletin* 109, 5–24.
- Cohen, S. and Doyle, W. J. (1997). Social ties and susceptibility to the common cold. *Journal of the American Medical Association* 277, 1940–1944.
- Cohen, S. and Miller, G. E. (2001). Psychological stress and antibody response to immunization: a critical review of the human literature. *Psychosomatic Medicine* 63, 7–18.
- Flinn, M. V. and England, B. G. (1997). Social economics of childhood glucocorticoid stress response and health. *American Journal of Physical Anthropology* 102, 33–53.
- Glaser, R. and Kiecolt-Glaser, J. K. (1998). The influence of psychological stress on the immune response to vaccines. *Annals of the New York Academy of Sciences* 840, 649–655.
- Herbert, T. B. and Cohen, S. (1993). Stress and immunity in humans: a meta-analytic review. *Psychosomatic Medicine* 55, 364–379.
- Kaufmann, S. H. (1992). Heat shock proteins in health and disease. *International Journal of Clinical and Laboratory Research* 21, 221–226.
- Kiecolt-Glaser, J. K. and Glaser, R. (1996). Chronic stress alters the immune response to influenza virus vaccine in older adults. *Proceedings of the National Academy of Sciences of the United States of America* 93, 3043–3047.
- Miller, G. E. and Cohen, S. (2004). Psychological stress and antibody response to influenza vaccination: when is the critical period for stress, and how does it get inside the body? *Psychosomatic Medicine* 66, 215–223.
- Petrie, K. J. and Booth, R. J. (1995). Disclosure of trauma and immune response to a hepatitis B vaccination program. *Journal of Consulting and Clinical Psychology* 63, 787–792.
- Stone, A. A. and Bovbjerg, D. H. (1994). Stress and humoral immunity: a review of the human studies. *Advances in Neuroimmunology* 4, 49–56.
- Vedhara, K. and Cox, N. K. (1999). Chronic stress in elderly carers of dementia patients and antibody response to influenza vaccination. *Lancet* 353, 627–631.
- Vedhara, K. and Bennett, P. D. (2003). Enhancement of antibody responses to influenza vaccination in the elderly following a cognitive-behavioural stress management intervention. *Psychotherapy and Psychosomatics* 72, 245–252.

Anti-CRF

E P Zorrilla, Y Zhao and G F Koob

The Scripps Research Institute, La Jolla, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Pharmacology of Corticotropin-Releasing Factor/Urocortin Receptor Systems

Nonneuroendocrine Distribution of CRF₁ Receptors

Selected Corticotropin-Releasing Factor Receptor Antagonists

Therapeutic Potential of CRF₁ Antagonists

Glossary

Adrenocorticotrophic hormone (ACTH, also known as adrenocorticotropin)
cLogP/cLogD

A hormone released into systemic circulation by the anterior pituitary as part of the hypothalamic-pituitary-adrenal (HPA)-stress response. ACTH induces the release of glucocorticoids from the adrenal cortex.

Calculated partition coefficients that express a compound's lipophilicity. A partition coefficient expresses the differential solubility of a compound in two solvents, with the log ratio of the concentrations of a neutral molecule partitioned between two solvents, classically octanol and water, called LogP. LogD expresses the partition coefficient at a specified pH, thereby accounting for charged forms of the molecule at a physiologically relevant pH (most typically that of blood, which has a pH of 7.4). Druglike substances have lipophilicities within a limited range, associated with favorable absorption, bioavailability, hydrophobic drug-receptor interactions, metabolism, and toxicity.

Corticotropin-releasing factor (CRF)

Isolated by Vale and colleagues in 1981, the first hypothalamic ACTH secretagogue to be identified. In addition to its hypothalamic endocrine function, extra-hypothalamic CRF is important in behavioral and autonomic stress responses.

Hypothalamic-pituitary-adrenal (HPA) axis

The neuroendocrine system which, during stress, amplifies a neural signal of physiological or psychological stress into a systemic, endocrine response. Hormones secreted as part of the HPA stress response cascade equip the organism to mobilize and utilize energy resources more effectively and, more generally, respond to the stressor.

Paralog

A gene (or gene product) related to a second gene (or gene product) by descent from an ancestral genetic duplication event. Although descended from common ancestral deoxyribonucleic acid (DNA) sequences, paralogs generally evolve new, distinct functions, though often related to that of the original sequence/gene product.

Pharmacophore

Term coined by Paul Ehrlich in 1909 as "a molecular framework that carries (*phoros*) the essential features responsible for a drug's (pharmac^{on}'s) biological activity." Updated by Peter Gund in 1977 to describe "a set of structural features in a molecule that is recognized at a receptor site and is responsible for that molecule's biological activity."

Urocortins

Structurally related mammalian paralogs of corticotropin-releasing factor (CRF). Though these three known peptides (urocortins (Ucn) 1, 2, and 3) share sequence identity with CRF, each CRF family peptide is coded by a different gene and has different pharmacological properties, tissue distributions, and regulatory pathways, leading to their diverse roles in stress responses to homeostatic threats. Urocortins, unlike CRF, exhibit high affinity for CRF₂ receptors.

Corticotropin-releasing factor (CRF) receptor antagonists have been sought since the isolation of the eponymous stress-related adrenocorticotropin-releasing peptide in 1981 by Vale and colleagues. Discovered for its hypophysiotropic function in the hypothalamic-pituitary-adrenal (HPA) axis, CRF and, later, urocortins (Ucn)-1, -2, and -3, paralogs of CRF, were subsequently found also to play roles in diverse homeostatic stress responses, via extra-hypothalamic brain sites and peripheral tissues. CRF receptor antagonists have potential therapeutic applications in cases where these responses are inappropriate in magnitude, timing, or context.

Pharmacology of Corticotropin-Releasing Factor/Urocortin Receptor Systems

CRF-related peptides interact with two known mammalian CRF receptor subtypes, CRF₁ and CRF₂, both belonging to the class B1 (secretinlike) subfamily of

G-protein-coupled receptors for polypeptide hormones. The CRF₁ receptor has been cloned in humans, mice, rats, and tree shrews, among other species, and exists in multiple isoforms (e.g., CRF_{1a}–CRF_{1h}), with the best known and functional isoform being the CRF_{1(a)} subtype. The CRF₂ receptor has three known functional membrane-associated subtypes in humans, CRF_{2(a)}, CRF_{2(b)}, and CRF_{2(c)}, as well as a ligand-sequestering, soluble CRF_{2(a)} isoform discovered in mice. CRF₂ and CRF₁ receptors have ~70% sequence identity. Whereas CRF has high, preferential affinity for CRF₁ versus CRF₂ receptors, Ucn 1 is a high-affinity agonist at both receptors. The type 2 Ucns, Ucn 2 and Ucn 3 are more selective for membrane CRF₂ receptors. Because of this pharmacological profile, references to antiCRF treatments most often refer to functional or competitive antagonists of CRF₁ receptor-mediated effects, whereas antitype 2 Ucn treatments might target CRF₂ receptors. The biological actions of CRF, Ucn-1, and Ucn-2 in rodents also are modulated by a CRF-binding protein (CRF-BP), a 37-kD secreted glycoprotein that binds and putatively immunosequesters CRF and Ucn-1 with equal or greater affinity than do CRF receptors. However, structural requirements for binding to CRF receptors and the CRF-BP differ, so many (if not most) CRF receptor antagonists do not interact with the CRF-BP.

Nonneuroendocrine Distribution of CRF₁ Receptors

CRF₁ receptors mediate not only the HPA-axis neuroendocrine response to stress, but also other aspects of organism stress responses. In brain, the distribution of CRF₁ receptors is highly conserved in stress-responsive brain regions, including neocortex,

the extended central amygdala, medial septum, hippocampus, hypothalamus, thalamus, cerebellum, and autonomic midbrain and hindbrain nuclei. This receptor distribution, concordant with that of its natural ligands (CRF, Ucn-1), is consistent with the proposed role for extrahypothalamic CRF₁ receptors in behavioral and autonomic stress responses. Accordingly, excessive CRF₁ signaling putatively has a pathophysiological role in several stress-related disorders, including anxiety and depressive disorders, substance dependence, and irritable bowel syndrome. Hereafter, the identity and potential therapeutic indications for CRF₁ receptor antagonists are summarized.

Selected Corticotropin-Releasing Factor Receptor Antagonists

Nonselective Peptide Corticotropin-Releasing Factor Receptor Antagonists

N-terminally truncated and substituted analogs of CRF can act as competitive partial agonists or full receptor antagonists (Table 1). Examples of these, in chronological order of discovery, include the partial agonist [Met¹⁸,Lys²³,Glu^{27,29,40},Ala^{32,41},Leu^{33,36,38}]r/hCRF₉₋₄₁ (also known as α -helical CRF₉₋₄₁); and the full receptor antagonists [D-Phe¹²,Nle^{21,38},C α MeLeu³⁷]r/hCRF₁₂₋₄₁ (also known as D-Phe-CRF₁₂₋₄₁); cyclo(30-33)[D-Phe¹²,Nle^{21,38},Glu³⁰,Lys³³]r/hCRF₁₂₋₄₁ (also known as astressin); and cyclo(30-33)[D-Phe¹²,Nle²¹,C α MeLeu²⁷,Glu³⁰,Lys³³,Nle³⁸,C α MeLeu⁴⁰]Ac-r/hCRF₉₋₄₁ (also known as astressin-B). These peptide ligands have roughly the same order of binding and antagonist potency at CRF₁ versus CRF₂ receptors and do not cross the blood–brain barrier. Thus, they are subtype nonselective, peripherally acting CRF receptor antagonists.

Table 1 Alphabetical selection of peptide CRF receptor antagonists

Familiar name	CRF receptor binding affinity (IC ₅₀ or K _i)		Chemical structure
	CRF ₁ (nM)	CRF ₂ (nM)	
α -helical CRF ₉₋₄₁	19	1.1	[Met ¹⁸ ,Lys ²³ ,Glu ^{27,29,40} ,Ala ^{32,41} ,Leu ^{33,36,38}]r/hCRF ₉₋₄₁
Antisauvagine-30	400	1.1	[D-Phe ¹¹ ,His ¹²]sauvagine ₁₁₋₄₀
Astressin	13.2	1.5	Cyclo(30-33)[D-Phe ¹² ,Nle ^{21,38} ,Glu ³⁰ ,Lys ³³]r/hCRF ₁₂₋₄₁
Astressin-B	High affinity	Moderate affinity	Cyclo(30-33)[D-Phe ¹² ,Nle ²¹ ,C α MeLeu ²⁷ ,Glu ³⁰ ,Lys ³³ ,Nle ³⁸ ,C α MeLeu ⁴⁰]acetyl-r/hCRF ₉₋₄₁
Astressin ₂ -B	>500	1.3	Cyclo(31-34)[D-Phe ¹¹ ,His ¹² ,C α MeLeu ^{13,39} ,Nle ¹⁷ ,Glu ³¹ ,Lys ³⁴]acetyl-sauvagine ₈₋₄₀
D-Phe-CRF ₁₂₋₄₁	19.2	4.4	[D-Phe ¹² ,Nle ^{21,38} ,C α MeLeu ³⁷]r/hCRF ₁₂₋₄₁
K31440	288	1.5	Acetyl-[D-Tyr ¹¹ ,His ¹² ,Nle ¹⁷]sauvagine ₁₁₋₄₀
K41498	425	0.7	[D-Phe ¹¹ ,His ¹² ,Nle ¹⁷]sauvagine ₁₁₋₄₀

CRF, corticotropin-releasing factor; IC₅₀, concentration required for 50% inhibition.
Data from Zorrilla et al., 2003 (with permission of Elsevier).

Subtype-Selective Peptide Corticotropin-Releasing Factor Receptor Antagonists

Antisauvagine-30 was reported as the first preferential CRF₂ antagonist. Following its development, several subsequent series of CRF₂ receptor peptide antagonists have been identified, each of which is more potent, selective, and enzymatic degradation resistant than antisauvagine-30. These new antagonists included the antisauvagine-30 analogs K31440 and K41498; [Tyr¹¹, His¹², Nle¹⁷]SVG₁₀₋₄₀, which has 2360-fold selectivity and subnanomolar affinity for the CRF₂ receptor; and astressin₂-B, a long-acting, potent, and selective CRF₂ antagonist (Table 1). Although these antagonists do not oppose the classic stress-related effects of CRF reviewed above, they do reverse stress-induced CRF₂-mediated endpoints, some of which are recruited by Ucn3 and some by CRF. Such endpoints include, for example, stress-induced gastric stasis and, possibly a component of stress-induced anorexia.

Selective CRF₁ peptide antagonists also appear to have been identified recently in the course of seeking minimal peptide fragments with CRF receptor antagonist activity. For example, Yamada and colleagues followed upon a patent application (W01/29086) of the Solvay pharmaceutical group which described a peptide comprising the 12 C-terminal residues of astressin as a potent antagonist of CRF receptors. Through amino acid substitution, of Nle38 with the bulkier, lipophilic cyclohexylalanine residue and of Ala31 with an unnatural residue (D-Ala), they identified a metabolically stable, high-affinity ($K_i \sim 3$ nM) CRF₁ antagonist that potently (0.1 mg kg⁻¹ intravenous; IV) reduced adrenocorticotrophic hormone (ACTH) secretion in a rat model of sepsis. This peptide is likely a selective CRF₁ antagonist because the Solvay group concurrently published their findings on the patented 12-residue N-terminal truncated astressin derivative on which Yamada and colleagues began their investigation. The disclosed fragment not only retained CRF₁ affinity, but was inactive at the CRF_{2(a)} receptor. Thus, it appears that lactam bridge-constrained N-terminally truncated astressin derivatives of 12–15 residue length may be selective CRF₁ receptor peptide antagonists. Such compounds could be useful for the treatment of pathologies that involve only peripheral CRF₁ hyperactivation.

Small-Molecule CRF₁ Receptor Antagonists

Pharmacophore and selectivity Small molecules that penetrate the blood–brain barrier with high and selective CRF₁ (versus CRF₂) affinity also have been identified (Table 2) and reviewed in considerable detail recently. With few exceptions, every known CRF₁

antagonist follows a common pharmacophore. Prototypical compounds (see Figure 1) share one or two lipophilic top/side units, a central mono-, bi- or tricyclic ring core, and a conformation-stabilizing di- or trisubstituted aromatic bottom group. Each ring core contains a putative proton-accepting ring nitrogen, and the hydrogen-bond accepting core is thought to interact with histidine-199 of the CRF₁ receptor, a polar amino acid in the third transmembrane domain that is not shared in the CRF₂ receptor or CRF-BP sequences. Therefore, CRF₁ antagonists with this pharmacophore are highly selective (CRF₁ versus CRF₂ and CRF-BP). Most known small-molecule CRF₁ receptor antagonists also show at least 1000-fold binding selectivity relative to other receptors, ion channels, and reuptake sites screened to date. However, this is not universal as one exception, R278995 (also known as CRA0450), actually exhibits 50-fold greater affinity for the σ_1 receptor than for the CRF₁ receptor.

The binding site of the classic pharmacophore also differs from the agonist binding site, whose confirmation is allosterically modified by CRF₁ antagonists, making these small molecules noncompetitive antagonists of peptide ligands. Exceptions to the pharmacophore include oxo-7H-benzo[e]perimidine-4-carboxylic acid derivatives, which are subtype nonselective CRF receptor antagonists discovered by Alanex; CC 2064460, a moderately potent arylamidrazone CRF₁ antagonist that lacks a central ring core with the customary hydrogen-bond accepting nitrogen; and stereospecific N-phenylphenylglycines, which also lack a ring core but were identified through computational screening based on a classic pharmacophore training set.

Lipophilicity An obstacle in the clinical use of CRF₁ antagonists is that most preclinical structures have been excessively lipophilic with poor water solubility, bioavailability, and pharmacokinetic properties. For example, the rank-ordered oral bioavailabilities of antalarmin (19%) < CP-154,526 (27%) < DMP904 (33%) < DMP696 (37–50%) follow the inverse rank of their cLogDs and cLogPs. Accordingly, the oral dose (30 mg kg⁻¹) of CP-154,526 needed to achieve the anxiolytic-like threshold of brain CRF₁ receptor occupancy in rats (~50–60% of receptors bound) is one order higher than that required by the similar affinity, but much less hydrophobic, compounds DMP696 (1–3 mg kg⁻¹), R121919 (2.5 mg kg⁻¹), and SSR125543A (6.5 mg kg⁻¹). Because of their high lipophilicity, many CRF₁ receptor antagonists have very high estimated bioconcentration factors at physiological pHs, a potential index of toxicity liability. Most early CRF₁ antagonists also failed Lipinski's

Table 2 Alphabetical selection of small-molecule CRF₁ receptor antagonists

Familiar name	CAS registry number ^c	CRF receptor binding affinity		Chemical structure
		CRF ₁ (nM)	CRF ₂ (nM)	
Antalarmin	157284-96-3	1	>10 000	<i>N</i> -butyl- <i>N</i> -ethyl[2,5,6-trimethyl-7-(2,4,6-trimethylphenyl)-7 <i>H</i> -pyrrolo[2,3- <i>d</i>]pyrimidin-4-yl]-amine
CC 2064460	438552-89-7	35	ND	Propanimidic acid, <i>N</i> -(3-ethynylphenyl)-2-oxo-, 2-(2-chlorophenyl)hydrazide (9CI)
CP-154,526	157286-86-7	2.7	>10 000	<i>N</i> -butyl- <i>N</i> -[2,5-dimethyl-7-(2,4,6-trimethylphenyl)-7 <i>H</i> -pyrrolo[2,3- <i>d</i>]pyrimidin-4-yl]- <i>N</i> -ethylamine
CRA0165	402751-10-4	10	>10 000	4-[5,6-dihydro-3-(2-methylphenyl)-1(2 <i>H</i>)-pyridinyl]- <i>N</i> -ethyl-6-methyl- <i>N</i> -[4-(1-methylethyl)-2-(methylthio)phenyl]
CRA1000	226948-11-4	16-21	>10 000	<i>N</i> -ethyl-4-[4-(3-fluorophenyl)-1,2,3,6-tetrahydro-1-pyridinyl]- <i>N</i> -[4-isopropyl-2-(methylsulfanyl)phenyl]-6-methylpyrimidin-2-amine
CRA1001	229346-94-5	19-22	>10 000	2-[<i>N</i> -(2-bromo-4-isopropylphenyl)- <i>N</i> -ethylamino]-4-[4-(3-fluorophenyl)-1,2,3,6-tetrahydropyridin-1-yl]-6-methylpyrimidine
DMP695	354994-31-3	3.3	>10 000	<i>N</i> -(2-chloro-4,6-dimethylphenyl)-1-1-methoxymethyl-(2-methoxyethyl)-6-methyl-1 <i>H</i> -1,2,3-triazolo[4,5- <i>c</i>]pyridin-4-amine mesylate
DMP696	202578-52-7	1.7	>10 000	4-(1,3-dimethoxyprop-2-ylamine)-2,7-dimethyl-8-(2,4-dichlorophenyl)-pyrazolo[1,5- <i>a</i>]-1,3,5-triazine
DMP904	202579-74-6	1.0	>10 000	<i>N</i> -(1-ethylpropyl)-3-(4-methoxy-2-methylphenyl)-2,5-dimethyl-pyrazolo[1,5- <i>a</i>]pyrimidin-7-amine
LWH63	276890-57-4	0.7	>10 000	4-ethyl[(2,5,6-trimethyl-7-(2,4,6-trimethylphenyl)-7 <i>H</i> -pyrrolo[2,3- <i>d</i>]pyrimidin-4-yl)amino]-1-butanol
MJL-1-109-2	596107-16-3	1.9	ND	pyrazolo[1,5- <i>a</i>]-1,3,5-triazin-4-amine, 8-[4-(bromo)-2-chlorophenyl]- <i>N</i> , <i>N</i> -bis(2-methoxyethyl)-2,7-dimethyl-(9CI)
NBI-27914	184241-44-9	2	>10 000	5-chloro- <i>N</i> -(cyclopropylmethyl)-2-methyl- <i>N</i> -propyl- <i>N</i> -(2,4,6-trichlorophenyl)-4,6-pyrimidinediamine
NBI-30545	195054-99-0	2.8	ND	pyrazolo[1,5- <i>a</i>]pyrimidin-7-amine, 3-(2,4-dimethoxyphenyl)- <i>N</i> -(2-methoxyethyl)-2,5-dimethyl- <i>N</i> -propyl-(9CI)
NBI-35965	354999-74-9	4	>10 000	ND
R121919 ^a	195055-03-9	3.5	>10 000	pyrazolo[1,5- <i>a</i>]pyrimidin-7-amine, 3-[6-(dimethylamino)-4-methyl-3-pyridinyl]-2,5-dimethyl- <i>N</i> , <i>N</i> -dipropyl-(9CI)
R121920	195055-01-7	4	ND	3-[6-(dimethylamino)-3-pyridinyl]-2,5-dimethyl- <i>N</i> , <i>N</i> -dipropyl-pyrazolo[1,5- <i>a</i>]pyrimidin-7-amine
R278995 ^b	N/A	54	>10 000	1-[8-(2,4-dichlorophenyl)-2-methylquinolin-4-yl]-1,2,3,6-tetrahydropyridine-4-carboxamide benzenesulfonate
SC241	392336-72-0	5	>10 000	3-[(2-bromo-4-isopropylphenyl)-5-methyl-3 <i>H</i> -[1,2,3]triazolo[4,5- <i>d</i>]pyrimidin-7-yl]-bis-(2-methoxy-ethyl)-amine
SN003	197801-88-0	6.8	ND	(+/-)- <i>N</i> -[2-methyl-4-methoxyphenyl]-1-(1-(methoxymethyl)propyl)-6-methyl-1 <i>H</i> -1,2,3-triazolo[4,5- <i>c</i>]pyridin-4-amine
SSR125543A	321839-75-2	2	>10 000	4-(2-chloro-4-methoxy-5-methylphenyl)- <i>N</i> -[2-cyclopropyl-1(<i>S</i>)-(3-fluoro-4-methylphenyl)ethyl]-5-methyl- <i>N</i> -(2-propynyl)thiazol-2-amine hydrochloride

^aAlso known as NBI-30775;^bAlso known as CRA0450;^cCAS, Chemical Abstract Service, a database of chemical substance information; ND, not determined.Table modified from Zorrilla et al., 2003 (with permission from Elsevier) and Zorrilla & Koob, 2004 (with permission from *Expert opinion on Investigational Drugs*).

“rule of 5” criteria for drug candidates, derived from a review of physiochemical properties of compounds within the United States Adopted Names database, due to excessive lipophilicity (cLogP > 5). Due to unfavorable pharmacokinetics, this degree of lipophilicity is rare for effective central nervous system (CNS)-acting agents. Of more than 100 drugs

marketed as of 1992 for CNS indications, not one had a logD greater than 4; most (~85%) had a logD of 0–3. Much preclinical discovery since 2002 has therefore focused on identifying less hydrophobic CRF₁ antagonists, with recent success. Early studies indicate that CRF₁ antagonists as a class do not share some of the drawbacks of benzodiazepines, such as

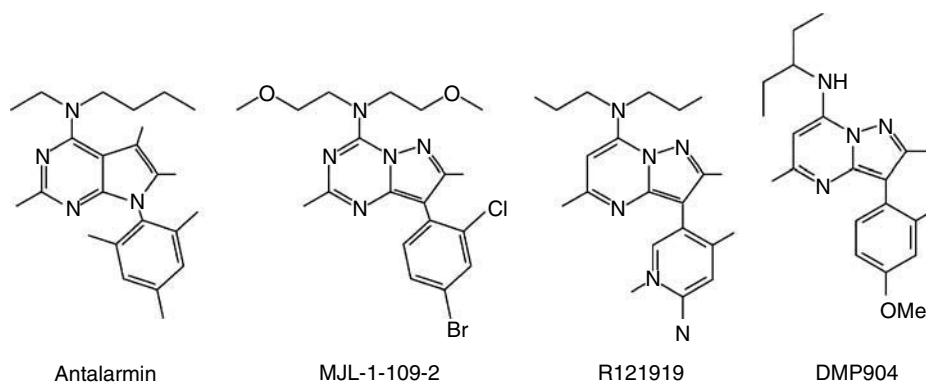


Figure 1 Representative selection of heterocyclic, small-molecule selective CRF₁ receptor antagonists that conform to the classic CRF₁ antagonist pharmacophore. See text for description of pharmacophore structural features and [Table 2](#) for chemical structure, CAS registry numbers, and binding properties of each compound.

tolerance, sedative-hypnotic, or amnesic effects, or abuse liability. They also do not lead to HPA-axis insufficiency at preclinically therapeutic doses, but isolated reports of liver enzyme elevations led to the discontinued clinical development of R121919, which had shown dose-related efficacy in an open-label Phase IIa clinical trial of patients with major depression.

Therapeutic Potential of CRF₁ Antagonists

Endocrine Disorders

CRF released into the portal blood from the median eminence stimulates ACTH release from the anterior pituitary into systemic circulation, which elicits the synthesis and release of glucocorticoids from the adrenal cortex – the classic HPA-axis stress response. HPA-axis dysregulation can result in endocrine disorders of hypercortisolemia, including Cushing's syndrome. Most cases of Cushing's syndrome are iatrogenic (e.g., from synthetic glucocorticoids), or due to ACTH- or cortisol-hypersecreting tumors and hyperplasias. However, central hypersecretion of neuroendocrine CRF leads to a Cushingoid state, exemplified in CRF-overexpressing transgenic mice. Likewise, clinical syndromes with hypothalamic CRF-driven chronic hypercortisolemia, including severe alcohol dependence, visceral obesity, melancholic major depression, and anorexia nervosa, are marked by Cushingoid somatic and endocrine features that remit when cortisol levels normalize. Peripherally acting, long-acting peptide CRF receptor antagonists, such as astressin-B, have been especially effective in opposing CRF-evoked ACTH release and therefore may have a therapeutic role in managing Cushingoid consequences of these disorders, if not also the disorders themselves.

Sympathoadrenomedullary Response

Intracerebroventricular (ICV) CRF administration evokes the sympathoadrenomedullary stress response in animal models, as it elevates blood pressure and heart rate; increases circulating levels of catecholamines, glucagon, and glucose; increases body temperature; and raises energy expenditure. Conversely, in animal models, CRF receptor antagonists, in particular those for the CRF₁ subtype such as CP-154,526, attenuate both CRF- and stress-induced tachycardia, hypertension, and hyperthermia.

Central Nervous System-Related Disorders

Anxiety disorders Consistent with autonomic findings, clinical data suggest that brain CRF is hypersecreted in some pathological anxiety conditions. Cerebrospinal (CSF) CRF levels, which mostly reflect nonneuroendocrine brain CRF secretion, are high in some patients with posttraumatic stress disorder, underweight patients with anorexia nervosa, ethanol withdrawn alcoholics, and some adults with obsessive-compulsive disorder, all conditions marked by high anxiety levels. Conversely, effective treatment of juveniles with obsessive-compulsive disorder with the tricyclic antidepressant clomipramine decreased their CSF CRF levels.

The anxiolytic-like actions of nonselective peptide and small-molecule CRF₁ receptor antagonists, CRF₁ knockout, and CRF₁ antisense knockdown all collectively emphasize an endogenous role for extrahypothalamic, limbic CRF₁ neurotransmission in hyperarousal and anxiety. For example, ICV α -helical CRF₉₋₄₁, a nonselective CRF receptor partial agonist, reversed the ability of restraint, foot shock, or social isolation to wake rodents prematurely from pentobarbital-induced sleep, and small-molecule CRF₁ antagonists like CRA1000 and R121919 also mitigate

insomnia-like effects of stress on sleep. Stress or CRF administration also increase locomotor activity in familiar environments, a sign of behavioral hyperarousal that is blocked by pretreatment with CRF₁ receptor antagonists and absent in CRF₁-deficient mice.

Nonselective or CRF₁ receptor antagonists also reduce the avoidance of unfamiliar, exposed areas in the open field, elevated plus maze, defensive withdrawal, or light/dark box tests by rodents that were stressed by restraint, swim stress, social defeat, conspecific stress, ethanol withdrawal, conditioned effects of psychostimulants, or novelty (but these actions are not observed in unstressed rodents). Knock-out of CRF₁, but not CRF₂, receptors in mice or antisense knockdown of CRF₁, but not CRF₂, receptors in rats likewise produces anxiolytic-like effects in exploration-based models of anxiety. Anxiolytic-like actions of nonselective or CRF₁ receptor antagonists are not limited to exploration-based models as they also were seen in rodent models of foot shock or conditioned freezing, fear- or light-potentiated acoustic startle, shock-induced defensive burying, social interaction, neonatal isolation-induced pup ultrasonic vocalization, and in unconditioned conflict models, such as punished drinking and punished crossing. Peptide nonselective or nonpeptide CRF₁ antagonists also reduced stress-induced anorexia in rats and unconditioned submissive/defensive behaviors in social defeat models. In nonhuman primates, CRF₁ exhibited oral anxiolytic-like action as well. For example, in rhesus monkeys, DMP904 (effective dose inducing 50% reduction (ED₅₀) 21 mg kg⁻¹) reduced stereotypical mouthing of rhesus monkeys in a human intruder test, and antalarmin (20 mg kg⁻¹) attenuated behavioral signs of anxiety, normalized exploratory, and sexual activity, and decreased HPA-axis and sympathoadrenomedullary hormone responses. Finally, depressed patients who participated in the open-label clinical trial with R121919 reported a reduction of anxious symptoms (and improved sleep quality) in relation to escalating doses of the small-molecule CRF₁ antagonist.

Anxiolytic-like efficacy of antagonists is evident once they occupy 50–60% of brain CRF₁ receptors in rats, with further efficacy seen with incremental brain receptor occupancy. Acute anxiolytic-like efficacy of CRF₁ antagonists does not depend on actions at the HPA axis, as at least six compounds (antalarmin, CP-154,526, CRA1000, CRA1001, DMP904, and R121919) reduced anxiety-like behavior at doses that did not reduce ACTH or corticosterone responses to stress. The importance of central as opposed to pituitary CRF₁ receptors in anxiety-like behavior is highlighted by the finding that

limbic-selective deletion of CRF₁ receptors in mice is sufficient to produce anxiolytic-like effects. Anxiogenic-like CRF signaling may be mediated by mitogen-activated protein (MAP) kinase extracellular signal-related kinase 1/2 (ERK1/2) pathways in the forebrain. Consistent with these findings and sites of anxiogenic-like action by exogenous CRF, the locus coeruleus, bed nucleus of the stria terminalis, and amygdala are sensitive to site-specific anxiolytic-like CRF₁ antagonist administration.

Depression Evidence for increased extrahypothalamic CRF activity, perhaps due to tonic, stimulatory glucocorticoid action on CRF synthesis, has also been observed in affective disorders and may contribute to depressive symptomatology. Patients with major depression exhibit signs of chronic hypothalamically driven HPA-axis hyperactivation, including hypercortisolemia, blunted ACTH responses to CRF infusion, nonsuppression in the dexamethasone suppression test, and increased numbers of paraventricular nucleus (PVN) CRF neurons postmortem. The resulting excess glucocorticoid levels exert positive feedback on extrahypothalamic CRF systems in the amygdala and bed nucleus of the stria terminalis, and thereby may lead to behavioral symptoms of depression. Accordingly, abnormally high CSF CRF levels are seen in severely depressed patients, who frequently exhibit comorbid pathological anxiety. Although CSF CRF concentrations are not elevated in patients with bipolar disorder, chronic administration of effective pharmacotherapies for bipolar disorder decreases amygdala CRF₁ receptor synthesis or expression in rats. Finally, very high CSF CRF levels and downregulated cortical CRF receptor binding and mRNA expression also were seen postmortem in patients who committed suicide.

Consistent with the therapeutic potential of CRF₁ antagonists for major depression, subchronic dosing with DMP696 or R121919 reduced mouse forced-swim immobility, and acute antalarmin reduced forced-swim immobility in CRF₂ knockout mice, a genetic model of high-stress responsiveness. Also, antalarmin, SSR125543A, or CRA1000 reduced forced-swim immobility in outbred rats, and chronic SSR125543A treatment increased swimming in Flinders Sensitive Line rats, a putative genetic model of depression. Findings are somewhat equivocal, however, as antalarmin, CP-154,526, DMP904, R121919, and DMP696 had no reliable effect on forced-swim immobility in normal mice, with negative findings seen after acute, subchronic, and even chronic (16 days) dosing, unlike positive antidepressant controls. Each of these antagonists, as well as R278995, also lacked activity in the tail suspension test with acute dosing.

CP-154,526, CRA1000, and R278995 also can reduce the acquisition of learned helplessness. However, despite an initial report with CP-154,526 that showed reversal of shuttlebox escape deficits, subsequent studies with CP-154,526, DMP904, DMP696, or CRA1000 have not found that CRF₁ antagonists could acutely reverse learned helplessness in rats, once established.

R278995 reduced hyperemotionality of olfactory bulbectomized rats, regarded by some as a depression model, though it lacked antidepressant-like activity in the rat differential-reinforcement-of-low-rate 72-sec model. Chronic antalarmin or SSR125543A treatment also improved physical coat appearance in a mouse model of chronic mild stress, a fluoxetine-like effect that correlated with reversal of stress-induced suppression of hippocampal neurogenesis. The chronicity of treatment may be relevant for antidepressant CRF₁ action as repeated administration of CRF₁ antagonists has somewhat more often resulted in antidepressant-like activity. Consistent with this perspective, depressed patients treated with escalating doses of R121919 in an open-label study reported reduced depressive symptoms and showed normalization of sleep electroencephalography. Clinically, CRF₁ gene haplotypes also predict differential treatment response to fluoxetine in anxious, depressed Mexican-American people.

As was observed for anxiolytic-like activity, antidepressant-like activity appears to be dissociable from acute actions on HPA-axis activity, supporting the role of extrahypothalamic CRF in proximately mediating depressive-like symptoms. For example, acute R121919, CP-154,526, and R278995 failed to reduce rat forced-swim immobility, despite the fact that R121919 could reduce concurrent swim-induced ACTH secretion. Conversely, LWH234 (a different CRF₁ antagonist) did reduce forced-swim immobility, but at a dose which had no effect on swim-induced ACTH secretion.

Gastrointestinal Disorders

CRF₁-receptor antagonists also might be promising tools for the treatment of functional gastrointestinal disorders, including irritable bowel syndrome. In animal models, diverse stressors delay gastric emptying and stimulate colonic motor function, and endogenous CRF, through both brain and peripheral receptors, mediates these stress-induced changes in gastrocolonic motility. The colonic-stimulating effects of stress are observed as colonic hypermotility, decreased transit time, defecation, and watery diarrhea and are mimicked by ICV administration of CRF or Ucn-1 in rats, mice, and Mongolian gerbils.

Central or peripheral administration of selective CRF₁, but not CRF₂, agonists stimulates colonic motility, and selective CRF₁, but not CRF₂, antagonists attenuate stress- or CRF/Ucn-1-induced stimulation of colonic motor function. For example, selective CRF₁ receptor antagonists such as CP154,526, CRA1000, NBI-35965, NBI-27914, and antalarmin, injected peripherally or ICV, blunted the acceleration of colonic transit induced by restraint; the defecation induced by water avoidance stress, restraint, or social stress; and the diarrhea induced by morphine withdrawal. Likely substrates for CRF₁-mediated stimulation of colonic motor function include, centrally, the PVN and locus coeruleus/Barrington nuclei that activate sacral parasympathetic nervous system activity, and, peripherally, the colonic myenteric nervous system.

CRF₁ receptor signaling is also involved in visceral hyperalgesia. For example, both stress and ICV CRF induce visceral hyperalgesia to colorectal distention in rats, and these actions of stress can be blocked by ICV α -helical CRF₉₋₄₁. More recently, systemic administration of NBI-35965, a nonpeptide CRF₁ antagonist, was found to block water avoidance stress-induced hyperalgesia to colorectal distention in adult offspring of maternal separation. Antalarmin (by intraperitoneal (IP) route) likewise reduced ICV CRF-induced hypersensitivity to colorectal distention and also blunted the visceral hyperalgesia of high-anxiety strain rats. Finally, α -helical CRF₉₋₄₁ attenuated electrical stimulation-induced visceral symptoms in patients with diarrhea-predominant irritable bowel syndrome.

CRF₁ receptors are involved in the early phase of postoperative gastric ileus. Peripheral CP-154,526 administration prevented abdominal surgery-induced slowing of gastric emptying. Similarly, CRF₁ knockout mice did not show postoperative gastric ileus, whereas abdominal surgery inhibited postoperative gastric emptying by 75% in wild-type mice.

CRF₂ receptor antagonists might be effective treatments to relieve stress-induced gastric stasis and perhaps other adverse gastroparesis. Stress or central infusion of CRF receptor agonists, through vagal effectors, delay gastric emptying, reduce antral gastric motility, inhibit high-amplitude gastric contractions, and shift duodenal activity from fasted to fed motor patterns, actions reversed by central astressin₂-B (CRF₂), but not NBI-27914 treatment. Complementing the central CRF₂ pathway for slowing gastric emptying, peripheral (by IV or IP route) injection of high-affinity and/or selective CRF₂ agonists (Ucn-1, Ucn-2, or Ucn-3) also delays gastric emptying more potently than CRF. Conversely, systemic pretreatment with CRF₂ (e.g., antisauvagine-30 or

astressin₂-B), but not CRF₁ (e.g., CP-154,526, DMP904, and NBI-27914) antagonists blocks the gastromotor inhibiting effects of Ucn₁ as well as restraint stress-induced gastric stasis. Thus, stress releases CRF like peptides that inhibit gastric motility through the brain–gastrointestinal tract CRF₂ receptor systems. Likely substrates include the PVN and dorsal vagal complex in the brain, and myenteric fibers of the enteric nervous system in the periphery.

Urologic Disorders

Epidemiologic and clinical findings indicate comorbidity of anxiety with several urologic problems, including overactive bladder, interstitial cystitis, and incontinence. In feline models of interstitial cystitis, CSF CRF levels are increased and subjects show exaggerated acoustic startle reactivity. This is somewhat expected because of the overlap in neuroanatomical substrates of CRF-related emotional responding and of micturition and pelvic viscerosensation. Shared neural substrates include Barrington's nucleus, amygdala, hippocampus, and prefrontal cortex. CRF also appears to play a role in the control of micturition and bladder sensation via the spinal cord. In particular, CRF immunoreactivity is present in intermediolateral cell column, suggesting a role in urinary motor function, and in sensory processing areas, including the dorsal horn, medial and lateral collateral pathways, dorsal intermediate gray, and laminae VII and X. Early studies indicate that CRF dose-dependently promotes micturition in awake animals, whereas CRF receptor antagonists have an inhibitory effect. Furthermore, CRF receptor antagonists have reversed bladder overactivity in animal models.

Pregnancy and Parturition

Changes in circulating CRF, Ucn-1, and CRF-BP levels help coordinate the events of pregnancy and parturition. CRF is produced in the hypothalamus and placenta during pregnancy, and placental CRF accesses both maternal and fetal circulations. During pregnancy, placental CRF secretion increases exponentially, peaking at birth. Accordingly, maternal plasma CRF levels increase from the second trimester exponentially towards term, rise dramatically during labor, and fall quickly after birth. Unbound (free) CRF levels are regulated by CRF-BP levels, which normally decrease near term (36 weeks' gestation), resulting in a dramatic increase in bioavailable maternal plasma CRF. Women with preterm labor show high CRF and low CRF-BP levels, suggesting an involvement of high bioavailability of maternal CRF in the onset of parturition. Consistent with this perspective, maternal free CRF measurement predicts outcome of

certain at-risk pregnancies, and the regulation of the length of gestation by CRF has been termed a placental clock. For example, the highest levels of maternal CRF at 20 weeks' gestation are observed in high-risk populations that ultimately deliver before 34 weeks. A prospective study of 232 pregnancies likewise observed that elevated maternal plasma CRF levels at 33 weeks' gestation carried a 3.3-fold greater risk of preterm birth, whereas women who delivered postterm had lower circulating CRF levels at that time than mothers who delivered at term. Similarly, a prospective study of 282 pregnancies found that maternal plasma CRF levels at 18–20 weeks and 28–30 weeks' gestation were higher in women who ultimately delivered preterm. Thus, CRF₁ antagonists might delay premature birth. Supporting this hypothesis, CRF₁ receptor antagonists delayed the onset of labor in sheep model studies.

Inflammatory States

CRF is a proinflammatory peptide via direct autocrine/paracrine actions in immune and immune lineage cells. In brain, CRF and/or its receptors are localized in microglia, astrocytes, and infiltrating monocytes/macrophages. In the periphery, CRF receptors are found in macrophages, lymphocytes, and mast cells. Immune CRF and Ucn-1 are expressed by human cord blood-derived cultured mast cells by the 10th gestational week and also are produced by human leukemic mast cell lines. In addition to mast cells, Ucn-1 also is seen in lymphocytes, macrophages, and fibroblasts, a cellular distribution that underlies moderate overall expression in thymus, spleen, and skin.

Supporting proinflammatory properties for CRF and Ucn-1, both peptides provoke resting monocytes to secrete interleukin (IL)-1 β and IL-6 *in vitro* via a CRF₁-mediated mechanism, and CRF triggers lymphocyte proliferation. CRF activates microglia, stimulates astrocytosis and monocyte migration, and augments macrophage production of free radicals and arachidonic acid. CRF receptor ligands also activate mast cells *in vitro*, an action that can be blocked by the CRF₁ antagonist antalarmin, but not the CRF₂ antagonist astressin₂-B and which leads to release of vascular endothelial growth factor, which is elevated in arthritis, psoriasis, and other chronic inflammatory conditions. The findings suggest a potential anti-inflammatory role for CRF₁ antagonists.

Inflammatory disorders Correspondingly, CRF immunoreactivity is elevated in local immune accessory cells in experimental models of arthritis and uveitis, and increased CRF or Ucn-1 immunoreactivity also are evident in human inflamed tissue, including

joints of patients with rheumatoid arthritis, lamina propria macrophages of patients with ulcerative colitis, reproductive tissue of patients with endometriosis, and thyroid glands of patients with Hashimoto's thyroiditis. The number of Ucn-1-positive cells in synovia of arthritic patients correlates strongly with inflammation severity. Conversely, chronic systemic antalarmin treatment reduced joint inflammation in an adjuvant-induced arthritis model in LEW/N rats. Also, CRF₁ agonists (IP) increase intestinal mucosal permeability to macromolecules, and chronic psychosocial stress reduces intestinal host defense and initiates intestinal inflammation through a mast cell CRF₁ receptor-dependent mechanism.

Neurological insults Excitotoxic, ischemic, and hypoxic brain insults upregulate CRF production in topographic and temporal relation to neuronal injury, and CRF potentiates ischemia-induced neuronal apoptosis in the presence of microglia *in vitro*. Conversely, CRF receptor antagonists reduced neuronal death induced by diverse brain injuries independent of the adrenocortical axis. R121920, a small-molecule CRF₁ antagonist, reduced infarct volume at 24 h in subtemporal and intraluminal suture middle cerebral artery occlusion models even when treatment was delayed up to 1 h postocclusion. Astressin, a nonselective peptide CRF antagonist also ameliorated the effects of 15-min ischemia in a Schaffer collateral-CA1 field potential hippocampal slice model. Central astressin treatment, prior to, or following, infantile seizures also decreased damage in hippocampal cell fields by 84%, perhaps also reflecting the antiexcitotoxic and not only anti-inflammatory properties of CRF₁ receptor antagonists.

Atopic/allergic disorders Stress worsens cutaneous diseases, such as psoriasis and atopic dermatitis. Recent evidence suggests that CRF/Ucn-1 crosstalk between mast cells, neurons, and keratinocytes underlie such exacerbation. Skin mast cells, which play a role in allergy and inflammation by releasing histamine, proteases, proteoglycans, prostaglandin D₂, and leukotriene C₄, secrete CRF and Ucn-1 in response to psychosocial stress. Accordingly, acute psychosocial stress increases skin CRF levels and triggers mast cell-dependent vascular permeability, which also results from intradermal CRF application. CRA1000, a CRF₁ antagonist, blocks the actions of repeated stress to thicken the epidermis and increases mast cell numbers in dermis, suggesting a role for CRF₁ receptors in chronic contact dermatitis. Mast cells also release CRF and Ucn-1 in response to immunoglobulin E

receptor cross-linking, suggesting a general role for CRF₁ activity in atopic disorders.

Acknowledgments

Supported by NIH grants DK64871 and DK26741 from the National Institute of Diabetes and Digestive and Kidney Diseases. The authors thank Mike Arends for editorial assistance. This is publication number 18309 from The Scripps Research Institute.

See Also the Following Articles

Anxiety; Blood Pressure; Cardiovascular System and Stress; Central Stress Neurocircuits; Corticotropin Releasing Factor (CRF); Corticotropin Releasing Factor-Binding Protein; Corticotropin-Releasing Factor Receptors; Cytokines; Gastrointestinal Effects; Hypothalamic-Pituitary-Adrenal; Immune Response; Immunity; Macrophages; Peptides; Psychosomatic Medicine; Somatic Disorders; Stress Effects, Overview; Ulceration, Gastric; Urocortins; Cardiovascular Disease, Stress and.

Further Reading

- Chen, C. (2006). Recent advances in small molecule antagonists of the corticotropin-releasing factor type-1 receptor – focus on pharmacology and pharmacokinetics. *Current Medicinal Chemistry* 13, 1261–1282.
- Gravanis, A. and Margioris, A. N. (2005). The corticotropin-releasing factor (CRF) family of neuropeptides in inflammation: potential therapeutic applications. *Current Medicinal Chemistry* 12, 1503–1512.
- Kalantaridou, S. N., Makriganakis, A., Zoumakis, E., et al. (2004). Reproductive functions of corticotropin-releasing factor. Research and potential clinical utility of antalarmins (CRF receptor type 1 antagonists). *American Journal of Reproductive Immunology* 51, 269–274.
- Nemeroff, C. B. and Vale, W. W. (2005). The neurobiology of depression: inroads to treatment and new drug discovery. *Journal of Clinical Psychiatry* 66, 5–13.
- O'Kane, M., Murphy, E. P. and Kirby, B. (2006). The role of corticotropin-releasing factor in immune-mediated cutaneous inflammatory disease. *Experimental Dermatology* 15, 143–153.
- Taché, Y., Million, M., Nelson, A. G., et al. (2005). Role of corticotropin-releasing factor pathways in stress-related alterations of colonic motor function and viscerosensitivity in female rodents. *Gender Medicine* 2, 146–154.
- Zorrilla, E. P. and Koob, G. F. (2004). The therapeutic potential of CRF₁ antagonists for anxiety. *Expert Opinion in Investigational Drugs* 13, 799–828.
- Zorrilla, E. P., Taché, Y., and Koob, G. F. (2003). Nibbling at CRF receptor control of feeding and gastrocolonic motility. *Trends in Pharmacological Science* 24, 421–427.

Antidepressant Actions on Glucocorticoid Receptors

J L W Yau and J R Seckl

University of Edinburgh, Western General Hospital,
Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by J L W Yau and J R Seckl, volume 1, pp 212–221,
© 2000, Elsevier Inc.

Glucocorticoid Receptors

Neuroendocrine Abnormalities in Depressive Disorders

Antidepressant Action

How Antidepressants May Regulate GR

Glossary

<i>Atypical antidepressants</i>	Include compounds that act similarly to tricyclic antidepressants but have a different chemical structure; also compounds with different pharmacological actions.
<i>Dexamethasone suppression test</i>	A clinical test to check the sensitivity of the hypothalamic-pituitary-adrenal stress axis. The synthetic steroid dexamethasone is injected and the fall in plasma cortisol levels is measured. Patients with major depression fail to respond with the normal fall in plasma cortisol seen in control subjects.
<i>Glucocorticoid receptors</i>	~94 kDa intracellular protein that mediates the effects of glucocorticoid hormones. They are also known as ligand-activated transcription factors since in the hormone-bound state, these receptors specifically bind to and modulate the activity of target gene promoters.
<i>Glucocorticoids</i>	A major subclass of steroid hormones produced by the zona fasciculata cells of the adrenal cortex that regulate fuel metabolism and mediate part of the stress response as well as aspects of immunosuppression and inflammation and behavioral functions.
<i>Monoamine oxidase inhibitor</i>	Type of antidepressant drug that inhibits one or both forms of brain monoamine oxidase (MAO), thus increasing the cytosolic stores of norepinephrine, dopamine, and 5-HT in nerve terminals.
<i>Multidrug resistance P-glycoprotein (PGP)</i>	A steroid transporter localized on the blood–brain barrier that limits the access of dexamethasone and cortisol to both mouse and human brain.

Transcription factors

Regulatory agents that are not part of the regulated gene(s). These *trans*-acting factors regulate gene transcription by binding directly or through an intermediate protein to the gene at a particular DNA sequence, called a *cis*-regulatory region, usually located in the 5' flanking promoter region of the gene.

Tricyclic antidepressants

Drugs that act by inhibiting uptake of norepinephrine and/or 5-HT by monoaminergic nerve terminals, thus acutely facilitating transmission.

Following the termination of the stress response, negative feedback exerted by circulating glucocorticoids (GCs; corticosterone in rats and mice, cortisol in humans and most other mammals) via glucocorticoid receptors (GRs) acts upon the hypothalamic-pituitary-adrenal (HPA) axis to restore low basal GC levels. In patients with major depression, the HPA axis is hyperactive as indicated by elevated circulating cortisol levels. This has led to the hypothesis that GR expression and/or function in depression may be abnormal. Successful therapy with antidepressant drugs restores HPA axis feedback sensitivity and lowers basal cortisol levels. One important aspect of antidepressant action may be to facilitate GR-mediated feedback inhibition upon the HPA axis by its actions on GR.

Glucocorticoid Receptors

The glucocorticoid receptor (GR) is a ligand-activated transcription factor that resides primarily in the cytoplasm as part of a multiprotein complex with heat shock proteins (hsps). This complex maintains unbound GRs in a ligand-accessible but inactive protein conformation. GCs are lipophilic substances and readily cross the cell membrane to interact with the intracellular GR. When activated by the binding of glucocorticoids (cortisol, corticosterone), the GR undergoes a conformational change, dissociates from the hsp complex, and translocates to the nucleus, where it binds to specific DNA sequences termed glucocorticoid-responsive elements (GRE) in the promoter region of glucocorticoid-responsive genes to influence gene transcription ([Figure 1](#)).

The hormone-activated GR can also act by interacting with other transcription factors in the absence of specific DNA binding. For example, GR inhibits

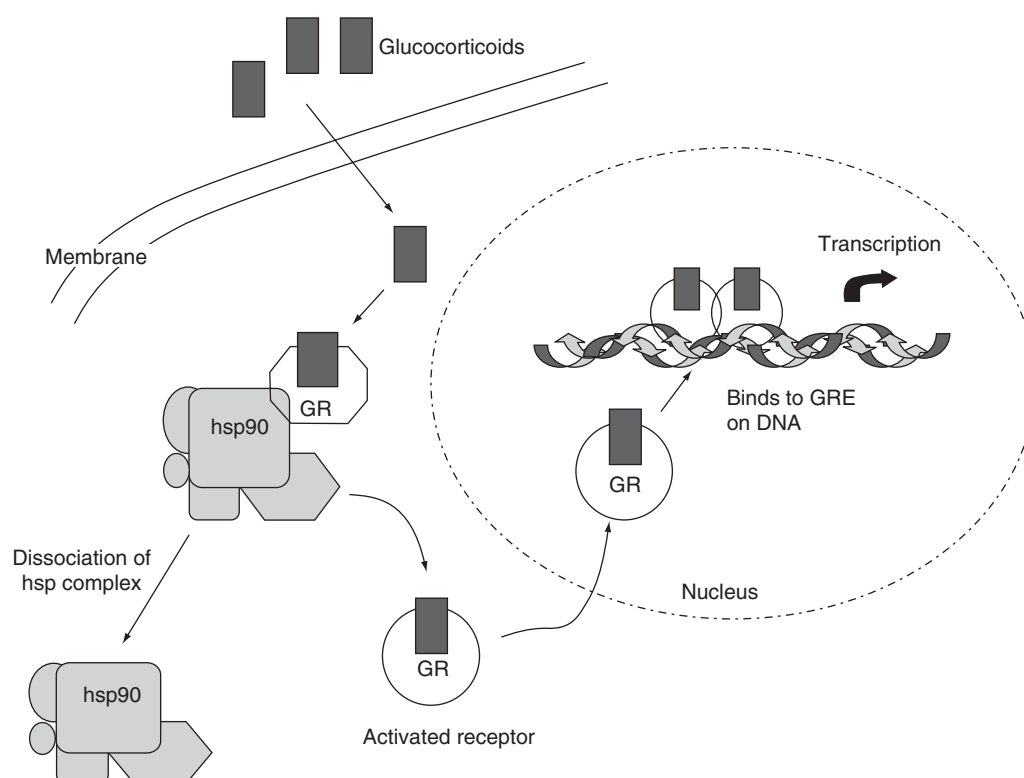


Figure 1 Simplified model of activation of glucocorticoid receptors. Glucocorticoids diffuse across the plasma membrane and bind to the glucocorticoid receptor (GR). Hormone binding causes a conformational change that causes dissociation of the GR–heat shock protein (hsp) complex and nuclear translocation. The activated GR binds as a homo- or heterodimer to glucocorticoid response element sequences (GREs) in the regulatory region of target genes to regulate transcription.

activating protein-1 (AP-1), a transcription factor composed of dimers of Jun and Fos family proteins, and thus represses transcription of AP-1-regulated genes such as collagenase and interleukin-2. Interaction with other transcription factors, such as cAMP response element binding protein (CREB), enhances GR-mediated transcription.

Glucocorticoid Feedback Regulation

Two types of adrenal steroid receptor have been described in the brain, a type 1 receptor indistinguishable from the peripheral mineralocorticoid receptor (MR) and a type 2 receptor indistinguishable from the peripheral GR. MR has a more restricted localization in brain and is most densely localized in hippocampal and septal neurons. GR, in contrast, is more diffusely distributed and is found particularly in areas within the limbic brain, not only in neurons but also in glia and vascular tissues. The hippocampus plays an important role in learning and memory and is also involved in HPA axis control. Much evidence supports an inhibitory role of the hippocampus, which thus suppresses HPA function. For example, lesions to the hippocampus usually lead to hypersecretion of GCs from the adrenals. MR and GR are particularly highly expressed in the hippocampus as shown using

molecular (e.g., *in situ* hybridization; see Figure 2), ligand binding, and immunohistochemical approaches. These two types of receptors bind GCs with differing affinities and act in coordination to regulate HPA axis function, with high-affinity MR controlling basal HPA activity during the nadir of the circadian rhythm of GC secretion, while lower affinity GRs are involved in the termination of the stress response.

Hippocampal MR and GR Density and HPA Activity

Manipulations that increase hippocampal GR, such as neonatal handling, a subtle environmental stress in which rat pups are physically removed from their mothers for 15 min daily during the first 2 weeks of life and which models some of the effects of naturally occurring variations in maternal care, improves sensitivity to GC negative feedback inhibition. In contrast, manipulations that decrease hippocampal GR density without neuronal damage, such as prolonged postnatal maternal separation or prenatal maternal stress or glucocorticoid exposure, reduce GC feedback sensitivity and therefore lead to chronic exposure to elevated GCs. Thus, the density of GR in the hippocampus is of crucial importance to feedback regulation of circulating GC levels. Hippocampal MRs are also important in terms of control of inhibitory

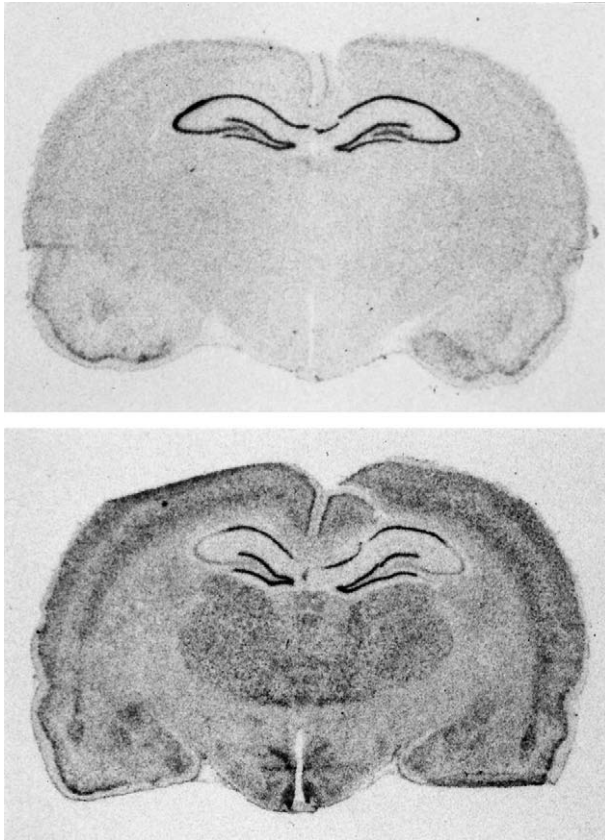


Figure 2 *In situ* hybridization localization of type 1 (mineralocorticoid receptor, top) and type 2 (glucocorticoid receptor, bottom) mRNA expression in the rat brain at the level of the dorsal hippocampus. Autoradiographs show MR expression is almost restricted to the hippocampus, while GR is more widespread but also highly expressed in the hippocampus.

tone over the HPA axis. Rat strains with high levels of hippocampal MR expression, such as the Lewis rat, show lower basal and stress-induced HPA activity compared with Wistar rats from which they are derived. Conversely, aged rats that generally have reduced hippocampal MR (and GR) expression show increased basal HPA activity. Central administration of endotoxin to rats, which impairs MR function, causes chronically elevated basal HPA activity.

Neuroendocrine Abnormalities in Depressive Disorders

Hyperactivity of the HPA axis is frequently found in patients with major depression who show increased concentrations of cortisol in plasma, urine, and cerebrospinal fluid and abnormal 24 h secretory patterns. These patients also fail to respond with the normal fall of plasma cortisol levels following administration of dexamethasone, a synthetic glucocorticoid that binds to GR. This forms the basis of the dexamethasone suppression test, one of the most widely studied

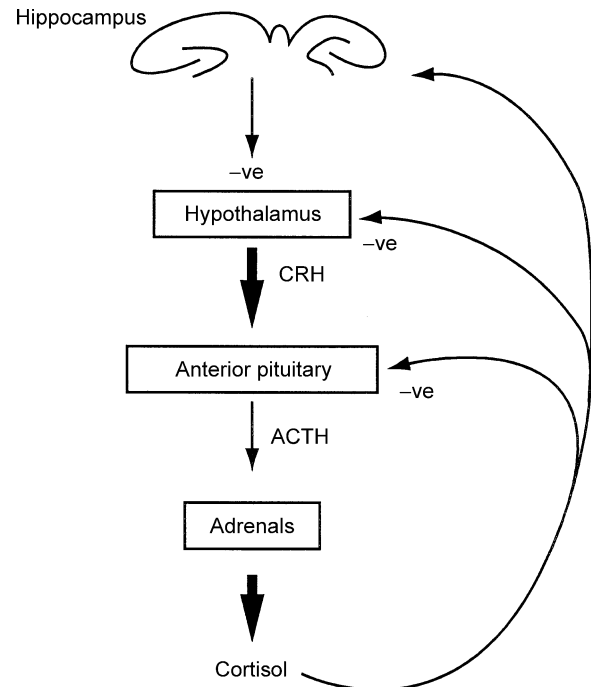


Figure 3 Schematic diagram of feedback control of HPA axis in depression. Cortisol suppresses the axis by binding to GRs in key feedback sites, hippocampus, paraventricular nucleus (PVN) of the hypothalamus, and pituitary gland. In major depression there is a hyperactive HPA axis with excess CRH and cortisol secretion, likely due to a hippocampal-hypothalamic defect.

endocrine tests in the history of psychiatry. A significant percentage (20–50%) of patients with major (melancholic) depression fail to suppress cortisol in response to dexamethasone challenge. This indicates inadequate GC feedback, which could be at any of a number of levels of the HPA axis (Figure 3).

In response to an intravenous dose of corticotrophin releasing hormone (CRH), patients with major depression secrete decreased amounts of ACTH. One possible mechanism to explain the blunted responsiveness to CRH could be pituitary hyporesponsiveness to CRH, itself due to CRH receptor desensitization (a consequence of sustained exposure to high CRH concentrations). Alternatively, another ACTH secretagogue may be implicated in HPA activation in depression (such as vasopressin). Finally, increased cortisol in patients with depressive illnesses may be inhibiting ACTH, but perhaps inadequately for the degree of elevation of plasma cortisol. To test this, a combination of dexamethasone suppression and then CRH stimulation (the dexamethasone/CRH test) has been used in patients with depression. Dexamethasone-pretreated patients show enhanced ACTH and cortisol responses to CRH, in contrast to earlier observations. These findings suggest that in severe depression, hypercortisolemia *per se* attenuates the ACTH response to CRH and that the equivalent basal state of the HPA axis is in fact

activated. The sensitivity (i.e., likelihood to differentiate normal from pathological state) of this DEX/CRH test is approximately 80%, which exceeds that of the standard dexamethasone suppression test. Thus, a limbic-hypothalamic defect resulting in excessive CRH output could explain the increased cortisol secretion in depression (Figure 3). This concept is supported by studies showing increased CRH levels in cerebrospinal fluid, increased postmortem hypothalamic CRH peptide and mRNA, and a lower than normal number of central CRH receptors among depressed patients.

Since GCs bind to MR and GR in brain sites involved in feedback control of HPA axis activity (i.e., MR and GR in the hippocampus, GR [and low levels of MR] in parvocellular neurons of the hypothalamus, anterior pituitary, and other sites), it has been suggested that depression may be associated with a primary imbalance of MR/GR expression and/or function. Although there is no direct evidence for reduced receptor concentrations in the brains of patients with depression, decreased GR mRNA expression have been reported in postmortem brain tissue (hippocampus and cortex) from individuals with major depressive disorder.

Antidepressant Action

Antidepressant drugs have been in use since the 1950s. Although clinically effective, their mechanism of action is still not clearly understood. Before 1980, the major classes of antidepressants in use were the tricyclics, monoamine oxidase inhibitors, and lithium. By the late 1980s, the serotonin reuptake inhibitors, currently the most commonly prescribed antidepressants, were introduced. All classes of antidepressants increase the concentration of biogenic amines (norepinephrine [NE] and/or serotonin [5-HT]) in the synaptic cleft. The ability of antidepressant drugs to facilitate monoaminergic transmission led to the monoamine theory, proposed in 1965, which suggests that depression results from functionally deficient monoaminergic transmission in the CNS. However, studies have not consistently shown a lack of biogenic amines or their metabolites in body fluids of depressed patients. There is also a clear discrepancy between the rapid effect of antidepressants in increasing synaptic concentrations of biogenic amines and the time course for the clinical efficacy of antidepressant treatment, which takes 2–3 weeks. A common, intracellular adaptation beyond the levels of 5-HT and NE and their receptors could account for the actions of the different types of antidepressants. Successful antidepressant treatment leads to normalization of HPA axis hyperactivity. Therefore, one such mechanism of antidepressant action is thought to be through direct

modulation of GR and/or MR in tissues involved in HPA feedback regulation.

Effects on HPA Function in Humans

Longitudinal studies in which depressed patients who were nonsuppressors of dexamethasone were retested weekly suggested that a return to dexamethasone suppressibility of plasma cortisol levels precedes the resolution of depressive psychopathology. Moreover, antidepressants may sometimes reduce CRH levels in human spinal fluid, suggesting that long-term administration of antidepressants may suppress the HPA system and that lowering HPA activity and clinical response are causally related. Finally, antidepressant resistant depression can be improved with GR synthesis inhibitors such as metyrapone.

Effects on HPA Function and Hippocampal Corticosteroid Receptors in Rats

Numerous studies in rats have shown that long-term *in vivo* treatment with a range of tricyclic and non-tricyclic antidepressants can enhance GC feedback inhibition (as indicated by decreased basal and/or stress-induced GC secretion) and/or increase corticosteroid receptor expression, particularly MRs, in the hippocampus (Table 1). An increase in MR and/or GR mRNA or binding in rat hippocampus can be observed irrespective of preferential inhibitory action of the antidepressant on reuptake of monoamines. This effect is dependent on the duration of antidepressant administration and requires weeks rather than days of treatment. For example, administration of the tricyclic antidepressant amitriptyline or moclobemide, a reversible inhibitor of monoamine oxidase type A (MAO-A), increased levels of hippocampal MR transiently by 40–70% between 2 and 5 weeks after the start of treatment. This is comparable to the increases in MR mRNA after long-term treatment with antidepressants. The levels of hypothalamic and hippocampal GR were also significantly increased by 20–25% at 5 weeks of treatment. The antidepressant-induced increases in brain GRs may underlie the observed decrease in basal circulating ACTH and corticosterone levels and in some cases decreased adrenal size. Indeed, the time course of antidepressant actions on GRs in the animal studies coincides with their long-term actions on HPA activity and follows closely that of clinical improvement of depression in humans. It is interesting to note that the antidepressant-induced increases in hippocampal MR are more pronounced than in GR and that the increases in the receptors are not maintained despite continual antidepressant treatment (Table 1). Systemic administration of antimineralocorticoids has been found to

Table 1 Antidepressant drug effects on mineralocorticoid and glucocorticoid receptor mRNA expression and binding sites in rat hippocampus

Antidepressant class	Drug	Dose/route	Duration	Hippocampal		Reference
				MR	GR	
Tricyclic	Amitriptyline (NE = 5-HT uptake)	20 mg/kg/day i.p.	2 weeks	87% incr (mRNA)	56% incr (mRNA)	Seckl, J. R. and Fink, G. (1992)
		10 mg/kg/day i.p.	10 weeks	17–28% incr (mRNA)	No effect	
		10 mg/kg/twice s.c. daily	3 weeks	16–29% incr (mRNA)	24–29% incr (mRNA)	Johansson, I. M. et al. (1998)
		4.5 mg/kg/day, p.o.	2–5 weeks	40–74% incr (binding sites)	0–26% incr (binding sites)	Reul, J. M. H. M. et al. (1993)
	Imipramine (NE > 5-HT uptake)	5 mg/kg/day i.p.	7 weeks	18% incr (binding sites)	No effect	Reul, J. M. H. M. et al. (1993)
			2–8 weeks	0–70% incr (mRNA)	No effect	Brady, L. S. et al. (1991)
	Desipramine (NE uptake)	10 mg/kg/day i.p.	2 weeks	60% incr (mRNA)	42% incr (mRNA)	Seckl, J. R. and Fink, G. (1992)
		10 mg/kg/day p.o.	4 weeks	32–36% incr (mRNA)	No effect	
MAO inhibitor	Moclobemide	4.5 mg/kg/day p.o.	2–5 weeks	65–76% incr (binding sites)	0–10% incr (binding sites)	Reul, J. M. H. M. et al. (1994)
			7 weeks	19% incr (binding sites)	No effect	Reul, J. M. H. M. et al. (1994)
Atypical Citalopram	(5-HT uptake)	20 mg/kg/day	2 weeks	17% incr (mRNA)	No effect	Seckl, J. R. and Fink, G. (1992)
	Fluoxetine (5-HT uptake)	5 mg/kg/day i.p.	2–8 weeks	0–56% incr (mRNA)	No effect	Brady, L. S. et al. (1992)
	Venlafaxine (5-HT > NE uptake)	10 mg/kg/day, p.o.	4 weeks	25% incr (mRNA)	No effect	Yau, J. L. W. and Seckl, J. R. (unpublished)

Abbreviations: i.p., interperitoneal; s.c., subcutaneous; p.o., oral route; incr, increase.

Note: The relative potency of antidepressant drugs in blocking reuptake of norepinephrine (NE) and serotonin (5-HT) shown in brackets. Full reference details in further reading.

impair the therapeutic efficacy of antidepressants, consistent with an important role for MRs in vulnerability to depression. MR controls basal HPA activity and, although not directly involved in the termination of the stress response, may influence the effects of GR on feedback control of the HPA axis. It is unclear how the lower plasma corticosterone levels are maintained long after hippocampal GR expression has returned to control levels and is no longer increased following long-term antidepressant treatment.

Effects in Transgenic Mice with Reduced GR Levels

To investigate the role of GR in HPA axis regulation and depression, several groups have generated mice with targeted GR deficiencies. Transgenic mice with impaired GR function were first created by attenuating mRNA expression with a GR antisense transgene leading to a 50% decrease in GR protein. These animals have numerous characteristics of human depression, including resistance to suppression of GCs by dexamethasone, feeding disturbances, and cognitive deficits. Long-term treatment with the antidepressant desipramine increased hypothalamic GR mRNA and binding, reduced HPA axis activity as shown by the partial reversal of enhanced ACTH and corticosterone secretion, and improved cognitive function. To dissect the role of brain GR from peripheral GR signaling and to rule out developmental effects, transgenic mice have recently been created with GR specifically disrupted throughout the hippocampus and most of the cortex in a time-dependent manner (GR immunoreactivity is lost in 60% hippocampal neurons at 2 months but not before 3 weeks of age and completely deleted by 6 months). These mice develop depression-like changes in adrenal axis regulation and behavior that mimic major depressive disorder in humans, including hyperactivity of the HPA axis (increased nadir and peak corticosterone release), impaired negative feedback inhibition of the HPA axis, and increased depression-like behavior (forced swim and tail suspension tests). Importantly, chronic treatment with the antidepressant imipramine reversed both the HPA axis abnormalities and behavioral phenotype despite the loss of GR action in these mice. This suggests that imipramine's known effects on forebrain GR are not crucial for its antidepressant effects and that the increase in MR mRNA expression in the hippocampus following imipramine may have compensated for the GR deficiency. These data support the hypothesis that a dysfunctional GR system may play a causative role in the etiology of depression.

Behavioral Effects of Antidepressant Actions via GR

There is now increasing evidence to link elevated cortisol levels in a subgroup of elderly humans to impaired memory associated with the hippocampus. Depression shows neuroendocrine abnormalities similar to those of unsuccessful aging, such as elevated GCs and impaired memory (dexamethasone nonsuppressors show significantly more commission errors in a word-learning paradigm [declarative memory] compared to suppressors). Chronically elevated GCs increase the vulnerability of the hippocampus to other insults, which may lead to structural and even neurotoxic tissue damage. Recent magnetic resonance image (MRI) studies have shown smaller hippocampal volumes in recurrent major depression and in healthy elderly subjects in whom cortisol levels are rising with age. Successful antidepressant drug therapy in major depression reverses hippocampal atrophy and cognitive impairments. Long-term treatment of middle-aged rats with the antidepressants amitriptyline or desipramine until 2 years old reduced the occurrence of hippocampal-dependent spatial memory impairments, and this may, in part, be a consequence of the reduced GC levels. Whether antidepressants can have the same effect in elderly humans with elevated cortisol levels and cognitive impairments has yet to be studied.

How Antidepressants May Regulate GR

Recent studies have focused on possible underlying mechanisms whereby antidepressant actions can modulate hippocampal GR. There is evidence to suggest that antidepressant effects, perhaps by increasing monoamines in the synaptic cleft, in particular 5-HT, start the cascade of events leading to increased transcription of GR. Serotonin acting via a ketanserin-sensitive cell surface receptor leads to increased second messenger (e.g., cAMP) levels in the cytosol of the cell, which in turn increases tertiary messengers, cAMP-dependent transcription factors in the nucleus, which might bind to the promoter of the GR gene to increase transcription ([Figure 4](#)).

Serotonin

5-HT increases GR mRNA and binding sites directly in hippocampal neurons in culture. This effect involves the second messenger cAMP and is blocked by ketanserin (mostly a 5-HT₂ receptor antagonist). Thus, 5-HT may act via a ketanserin-sensitive cell surface 5-HT receptor to increase GR. Antidepressant drugs may therefore increase hippocampal GR expression *in vivo* via its actions on potentiating monoaminergic (e.g., 5-HT and NE) neurotransmission.

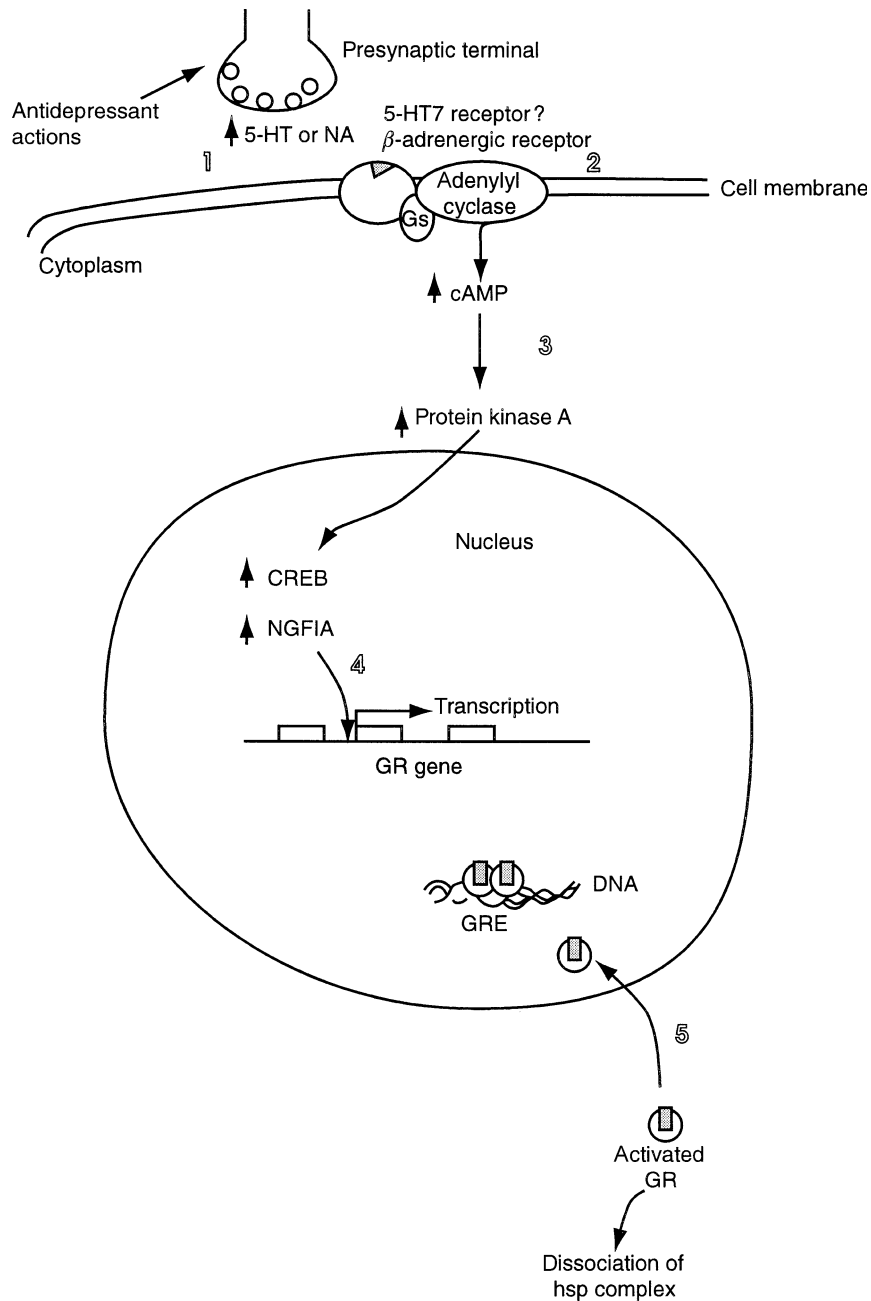


Figure 4 Model of the action of chronic antidepressant treatments on GR. Antidepressants increase monoamines (5-HT or NE) in the synaptic cleft (1), which activate cell surface monoamine receptors coupled to the cAMP-protein kinase A cascade (2). This leads to increases in the second messengers cAMP and protein kinase A (3), leading to an increase in cAMP-dependent transcription factors in the nucleus (4), which may bind to the GR gene promoter to increase transcription. Alternatively, antidepressants may act directly to increase GR activity/DNA binding, possibly by inhibiting the membrane steroid transporter and thus inducing GR translocation from the cytoplasm to the nucleus (5).

Transcription Factors

Chronic antidepressant administration increases the expression of the nuclear transcription factor CREB in the rat hippocampus. Interestingly, temporal cortex CREB concentrations in major depressive disorder patients treated with antidepressants at the

time of death were higher than in untreated patients. CREB could be activated by 5-HT and NE receptors that directly stimulate cAMP production (e.g., 5-HT_{4,6,7} or β -adrenergic receptors) or 5-HT or NE receptors (e.g., 5-HT_{2A,C} and α 1-adrenergic receptors) that lead to stimulation of Ca^{2+} -dependent

protein kinases. Upregulation of CREB appears to be a common action of chronic antidepressant treatments, and since CREB can upregulate GR gene transcription, this may be one aspect whereby antidepressants facilitate GR-mediated feedback inhibition and thereby reverse GC hypersecretion in depression. Chronic amitriptyline treatment in rats has also been shown to increase the expression of another transcription factor, NGFIA, in the hippocampus. This transcription factor may bind to the GR promoter to increase transcription of GR.

Direct Modulation of GR

Treatment of primary cultures of rat brain hypothalamic neurons and hippocampal neurons for up to 14 days with the tricyclic antidepressants desipramine and amitriptyline increases GR mRNA (also GR binding sites and immunopositive cells) despite the lack of monoaminergic terminals in these cultures. A direct effect of antidepressants on GR was also shown when mouse fibroblast LTK cells or Neuro-2A treated with desipramine produced a 50–200% increase in the reporter gene chloramphenicol acetyltransferase (CAT) activity transcribed from a GR gene promoter region. These fibroblast cells do not secrete catecholamines, and therefore the desipramine-induced increase in GR gene activity involves a mechanism independent of NE reuptake inhibition.

Intriguingly, desipramine treatment for 24 h in mouse L929 fibroblast cells induced GR to translocate from the cytoplasm to the nucleus (using GR immunostaining technique) in the absence of exogenous steroid. Desipramine also potentiates dexamethasone-induced GR translocation and dexamethasone-induced GR-mediated gene transcription. One possible explanation for this steroid-independent modulation of GR *in vitro* could be that desipramine directly acts on one or more hsp and facilitates the dissociation of the receptor from the hsp complex.

Membrane Steroid Transporter

Membrane steroid transporters such as the multidrug resistance P-glycoprotein (MDR PGP) can regulate the function of GR and MR by modulating the intracellular access of steroid hormones. The MDR PGP, found in the endothelial cells of the blood–brain barrier, limits the access of dexamethasone and cortisol to both mouse and human brain. Antidepressants may modulate GR by inhibiting steroid transporters, thus increasing intracellular levels of steroids resulting in translocation of GR. MDR PGP inhibitors have been shown to induce partial GR translocation in HeLa (human ovarian cancer) cells, thus providing some evidence for this possibility. Translocation of GR could lead to the reduction in GR expression

(mRNA and binding sites) seen during the early phase of antidepressant actions (shown both *in vitro* and *in vivo* in rat hippocampus) and that precedes the upregulation of GR after 2 weeks or longer depending on antidepressant dose and route of administration. Recent studies have suggested that MDR PGP can regulate corticosterone access to the brain and HPA activity in rats and mice. Transgenic mice with deletion of the MDR PGP genes (two isoforms are found in rodents) show increased access of corticosterone to the brain and increased negative feedback on the HPA axis by corticosterone. Although the antidepressant clomipramine has been shown to regulate the intracellular levels of cortisol by inhibiting membrane steroid transporters in fibroblast cells and rat primary neurons in culture, whether this occurs *in vivo* remains to be determined.

Thus, antidepressants via a variety of mechanisms facilitate GR and MR function and hence reduce HPA overactivity. The importance of the mechanisms in antidepressant and other potential therapeutic effects is the subject of active study.

See Also the Following Articles

Depression and Manic-Depressive Illness; Depression Models; Glucocorticoid Negative Feedback; Hypothalamic-Pituitary-Adrenal.

Further Reading

- Barden, N. (1996). Modulation of glucocorticoid receptor gene expression by antidepressant drugs. *Pharmacopsychiatry* 29, 12–22.
- Barden, N., Reul, J. M. H. M. and Holsboer, F. (1995). Do antidepressants stabilize mood through actions on the hypothalamic-pituitary-adrenocortical system? *Trends in Neuroscience* 18, 6–11.
- Boyle, M. P., Brewer, J. A., Funatsu, M., et al. (2005). Acquired deficit of forebrain glucocorticoid receptor produces depression-like changes in adrenal axis regulation and behavior. *Proceedings of the National Academy of Sciences of the USA* 102, 473–478.
- Brady, L. S., Whitfield, Jr. H. J., Fox, R. J., Gold, P. W. and Herkenham, M. (1991). Long-term antidepressant administration alters corticotropin-releasing hormone, tyrosine hydroxylase and mineralocorticoid receptor gene expression in the rat brain. *Journal of Clinical Investigation* 87, 831–837.
- De Kloet, E. R., Vreugdenhil, E., Oitzl, M. S. and Joëls, M. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Reviews* 19, 269–301.
- Dinan, T. (1994). Glucocorticoids and the genesis of depressive illness: a psychobiological model. *British Journal of Psychiatry* 164, 356–371.
- Holsboer, F. and Barden, N. (1996). Antidepressants and hypothalamic-pituitary-adrenocortical regulation. *Endocrine Reviews* 17, 187–205.

- Jacobson, L. and Sapolsky, R. (1991). The role of the hippocampus in feedback regulation of the hypothalamic-pituitary-adrenal axis. *Endocrine Reviews* **12**, 118–134.
- Johansson, I. M., Bjartmar, L., Marcusson, J., et al. (1998). Chronic amitriptyline treatment induces hippocampal NGFI-A, glucocorticoid receptor and mineralocorticoid receptor mRNA expression in rats. *Brain Research Molecular Brain Research* **62**, 92–95.
- Modell, S., Yassouridis, A., Huber, J. and Holsboer, F. (1997). Corticosteroid receptor function is decreased in depressed patients. *Neuroendocrinology* **65**, 216–222.
- Pariante, C. M., Thomas, S. A., Lovestone, S., Makoff, A. and Kerwin, R. W. (2004). Do antidepressants regulate how cortisol affects the brain? *Psychoneuroendocrinology* **29**, 423–447.
- Reul, J. M. H. M., Stec, I., Söder, M. and Holsboer, F. (1993). Chronic treatment of rats with the antidepressant amitriptyline attenuates the activity of the hypothalamic-pituitary-adrenocortical system. *Endocrinology* **133**, 312–320.
- Reul, J. M. H. M., Labeur, M. S., Grigoriadis, D. E., De Souza, E. B. and Holsboer, F. (1994). Hypothalamic-pituitary-adrenocortical axis changes in the rat after long-term treatment with the reversible monoamine oxidase-A inhibitor moclobemide. *Neuroendocrinology* **60**, 509–519.
- Seckl, J. R. and Fink, G. (1992). Antidepressants increase glucocorticoid and mineralocorticoid receptor mRNA expression in the rat hippocampus *in vivo*. *Neuroendocrinology* **55**, 621–626.

Antihypertensive Drugs See: Beta-Adrenergic Blockers.

Antimineralocorticoid Challenge

M Kellner and K Wiedemann

University Hospital Hamburg-Eppendorf, Hamburg, Germany

© 2007 Elsevier Inc. All rights reserved.

Introduction

Findings in Healthy Subjects

Preliminary Findings on HPA Axis in Psychiatric Disorders

Future Research

Glossary

<i>Canrenoate</i>	Antimineralocorticoid agent for use in humans. Potassium salt of canrenone. Active metabolite of spironolactone.
<i>Hippocampus</i>	Part of the limbic system of the brain. Involved in endocrine, emotional, and cognitive functions.
<i>Hypothalamic-pituitary-adrenocortical (HPA) Axis</i>	Major stress-mediating endocrine system in humans. Activation results in the release of cortisol from the adrenal glands. Regulated by suprahypothalamic neural influences, hypothalamic release of corticotropin-releasing hormone, pituitary release of adrenocorticotrophic

Major depression

Mineralocorticoid receptor

Posttraumatic stress disorder

Spironolactone

hormone, and negative feedback effects of cortisol.

Psychiatric illness with markedly depressed mood, considerable loss of interest or pleasure, and several other psychopathological symptoms, such as insomnia, reduced appetite, fatigue, feelings of guilt or worthlessness, diminished ability to concentrate, and suicidal ideation.

Intracellular receptor molecule with high affinity for cortisol, corticosterone, and aldosterone. Found in the central nervous system, primarily in the hippocampus and, furthermore, peripherally in the kidney.

Psychiatric illness that can occur after severe psychological trauma. Characterized by re-experiencing the traumatic event, avoidance of stimuli associated with the trauma, and persistent hyperarousal symptoms, such as exaggerated startle response, hypervigilance, difficulty concentrating, and reduced sleep.

Antimineralocorticoid agent for use in humans. Spironolactone undergoes extensive metabolism in humans: taken orally, up to 58% of spironolactone is excreted as canrenone, canrenoate ester glucuronide, and several sulfur-containing and polyhydroxylated metabolites.

Introduction

Preclinical findings have established that hippocampal mineralocorticoid receptors (MRs) control the inhibitory tone of this limbic structure on the hypothalamic-pituitary-adrenocortical (HPA) axis in terms of both basal activity and stress-evoked reactivity. Hippocampal MRs (type I corticosteroid receptors) have a considerably higher affinity for cortisol, the physiological ligand in humans, than glucocorticoid receptors (type II corticosteroid receptors); thus, they are highly occupied throughout the circadian variation of HPA axis activity. While glucocorticoid receptors are widespread in the central nervous system, MRs in brain are primarily localized in the hippocampus. Also the prereceptor metabolism of corticosteroids regulated by 11β -hydroxysteroid dehydrogenases (11β -HSD) is important for the saturation of hippocampal MRs by cortisol. While 11β -HSD type 2 in the kidneys results in selectivity of distal nephron MRs for aldosterone by inactivation of cortisol, 11β -HSD type 1 in the brain regenerates active glucocorticoids that bind to MRs in the limbic system.

The hippocampal MR is quickly upregulated by psychological stressors in rodents, such as forced swimming or novelty. This effect can be prevented by intracerebroventricular pretreatment with an antagonist for the corticotropin-releasing hormone (CRH). Thus, the MR participates in adaptive mechanisms in the central nervous system evoked by stressful events. Furthermore, aspects of anxiety-like behavior are modulated by the MR, as shown in different animal models. Our understanding of MR action upon HPA axis regulation in humans may be relevant not only for the physiology of the stress response, but also for the pathophysiology of disorders in which HPA axis alterations have been described. Given the high occupation of the hippocampal MR during both baseline and stress conditions, antimineralocorticoid challenge has been the method applied to assess MR function in humans. In human challenge studies, spironolactone and canrenoate have been used as safe antagonists at the MR that display sufficient specificity and penetrate the blood-brain barrier.

Findings in Healthy Subjects

Basal Activity of the HPA Axis

Most studies have shown elevated basal plasma cortisol concentrations as soon as 1 h after administration of spironolactone (200–400 mg orally) or canrenoate (2×200 mg intravenously). Findings about elevated plasma adrenocorticotrophic hormone (ACTH) are

less consistent, because cortisol released by the anti-MR challenge can mask this effect by an activation of the negative glucocorticoid receptor-mediated feedback at the pituitary. The significant effect of spironolactone on cortisol secretion was demonstrated with no diurnal differences between morning and evening administration despite a circadian variation of MR upregulation.

The stimulatory action of anti-MR challenge upon HPA activity can be blocked by the benzodiazepine receptor agonist alprazolam via hippocampal and hypothalamic γ -aminobutyric acid (GABA) receptors. Furthermore, the physiological inhibition of pituitary-adrenal secretion in humans during early nocturnal epochs of non-REM sleep was reported to disappear after acute antimineralocorticoid blockade.

Activated HPA Axis

Both after CRH stimulation and after stimulation with arginine vasopressin, as well as in a combined dexamethasone suppression/CRH stimulation test and during exercise, significant elevations of plasma cortisol after prior anti-MR treatment vs. placebo have been reported. These evidence a shift in set point for cortisol feedback inhibition of the HPA axis toward higher cortisol levels. Thus, MRs can modulate effects of both pituitary and suprapituitary stimuli upon human HPA axis activity.

Basal Activity of the HPA Axis in Elderly Subjects

In a circadian study, Heuser et al. demonstrated that, compared to young subjects, spironolactone treatment had stronger stimulatory effects on basal cortisol secretion in the elderly. A decrease in hippocampal MR capacity during aging might contribute to feedback alterations of the HPA system described in the elderly.

Preliminary Findings on HPA Axis in Psychiatric Disorders

Major Depression

Young et al. applied spironolactone challenges to patients with major depression. They hypothesized a decreased MR activity, because major depression is often accompanied by increased cortisol secretion and decreased dexamethasone suppression, furthermore, because glucocorticoids are capable of down-regulating MR and glucocorticoid receptors. Instead, in patients with major depression increased functional activity of the MR was found with a more pronounced cortisol and ACTH secretion after spironolactone, which is so far not sufficiently understood.

Posttraumatic Stress Disorder (PTSD)

Kellner et al. investigated whether findings of diminished basal and stress-induced cortisol output reported in PTSD could be caused by an enhanced MR activity. In contrast to expectation, spironolactone effects upon basal and CRH-stimulated HPA axis were indistinguishable in PTSD patients and healthy controls. However, underlying the experimental conditions, other compensatory factors, such as hyperactive pituitary glucocorticoid receptor-mediated negative feedback in PTSD, might have masked the putative mineralocorticoid receptor changes.

Future Research

Future antimineralocorticoid challenge studies in humans might use new substances with higher specificity for the MR than spironolactone and canrenoate, e.g., eplerenone. An additional pretreatment with a glucocorticoid receptor antagonist would help to disentangle secondary glucocorticoid-mediated feedback effects on HPA axis activity after MR blockade. However, antiglucocorticoids, such as mifepristone and onapristone, exert effects also at other steroid receptors, such as progesterone receptors, which would complicate the allocation of steroid effects.

An alternative approach to assessing MR function is a newly developed test that uses, for the first time, not an MR antagonist, but an MR agonist after prior depletion of the endogenous cortisol from the MR using metyrapone pretreatment. Metyrapone inhibits peripherally and centrally the synthesis of cortisol via 11- β -hydroxylase and 11- β -hydroxysteroid dehydrogenase type 1.

Given animal data for the potential role of the MR in depression and anxiety, the neuropsychological and behavioral effects of antimineralocorticoid challenge in humans also require detailed research. Preliminary results by Hundt et al. show that subchronic spironolactone comedication impedes the effect of treatment with the antidepressant amitriptyline in patients suffering from major depression, whereas upregulation of MR by depletion of corticosteroid with metyrapone enhances antidepressive effects. Thus, intact MR function seems to be important for antidepressant drug action, and antimineralocorticoid challenge might be also be useful for the prediction of psychotherapeutic drug effects in patients suffering from psychiatric disorders.

See Also the Following Articles

Aging and Psychological Stress; Aging and Stress, Biology of; Corticosteroid Receptors; Depression and Manic-Depressive Illness; Glucocorticoid Negative Feedback.

Further Reading

- Arvat, E., Maccagno, B., Giordano, R., et al. (2001). Mineralocorticoid receptor blockade by canrenoate increases both spontaneous and stimulated adrenal function in humans. *Journal of Clinical Endocrinology and Metabolism* 86, 3176–3181.
- Born, J., Steinbach, D., Dodt, C. and Fehm, H.-L. (1997). Blocking of central nervous mineralocorticoid receptors counteracts inhibition of pituitary-adrenal activity in human sleep. *Journal of Clinical Endocrinology and Metabolism* 82, 1106–1110.
- Deuschle, M., Weber, B., Colla, M., et al. (1998). Mineralocorticoid receptor also modulated basal activity of hypothalamus-pituitary-adrenocortical system in humans. *Neuroendocrinology* 68, 355–360.
- Gesing, A., Bilang-Bleuel, A., Droste, S. K., et al. (2001). Psychological stress increases hippocampal mineralocorticoid receptor levels: involvement of corticotropin-releasing hormone. *Journal of Neuroscience* 21, 4822–4829.
- Grottoli, S., Giordano, R., Maccagno, B., et al. (2002). The stimulatory effect of canrenoate, a mineralocorticoid antagonist, on the activity of the hypothalamus-pituitary-adrenal axis is abolished by alprazolam, a benzodiazepine, in humans. *Journal of Clinical Endocrinology and Metabolism* 87, 4616–4620.
- Heuser, I., Deuschle, M., Weber, A., et al. (2000). The role of mineralocorticoid receptors in the circadian activity of the human hypothalamus-pituitary-adrenal system: effect of age. *Neurobiology of Aging* 21, 585–589.
- Heuser, I., Deuschle, M., Weber, B., Stalla, G. K. and Holsboer, F. (2000). Increased activity of the hypothalamus-pituitary-adrenal system after treatment with the mineralocorticoid receptor antagonist spironolactone. *Psychoneuroendocrinology* 25, 513–518.
- Jahn, H., Schick, M., Kiefer, F., et al. (2004). Metyrapone as additive treatment in major depression: a double-blind and placebo-controlled trial. *Archives of General Psychiatry* 61, 1235–1344.
- Kellner, M., Baker, D. G., Yassouridis, A., et al. (2002). Mineralocorticoid receptor function in patients with posttraumatic stress disorder. *American Journal of Psychiatry* 159, 1938–1940.
- Otte, C., Jahn, H., Yassouridis, A., et al. (2003). The mineralocorticoid receptor agonist, fludrocortisone, inhibits pituitary-adrenal activity in humans after pretreatment with metyrapone. *Life Sciences* 73, 1835–1845.
- Reul, J. M. H. N., Gesing, A., Droste, S., et al. (2000). The brain mineralocorticoid receptor: greedy for ligand, mysterious in function. *European Journal of Pharmacology* 405, 235–249.
- Seckl, J. R. (1997). 11 β -hydroxysteroid dehydrogenase in the brain: a novel regulator of glucocorticoid action? *Frontiers in Neuroendocrinology* 18, 49–99.
- Weinberger, M. H. (2004). Eplerenone: a new selective aldosterone receptor antagonist. *Drugs Today* 40, 481–485.
- Wellhoener, P., Born, J., Fehm, L. H. and Dodt, C. (2004). Elevated resting and exercise-induced cortisol levels after

mineralocorticoid receptor blockade with canrenoate in healthy humans. *Journal of Clinical Endocrinology and Metabolism* 89, 5048–5052.

Young, E. A., Lopez, J. F., Murphy-Weinberg, V., Watson, S. J. and Akil, H. (1998). The role of mineralocorticoid receptors in HPA axis regulation in humans. *Journal*

of Clinical Endocrinology and Metabolism 83, 3339–3345.

Young, E. A., Lopez, J. F., Murphy-Weinberg, V., Watson, S. J. and Akil, H. (2003). Mineralocorticoid receptor function in major depression. *Archives of General Psychiatry* 60, 24–28.

Antipsychotic Drugs and Stress

S Sundram

Mental Health Research Institute of Victoria, Parkville, Vic, Australia

© 2007 Elsevier Inc. All rights reserved.

Introduction

Mechanism of Action of Antipsychotic Drugs

Theories of Atypical Antipsychotic Drug Action

Limitations of the Dopamine D2 Receptor Hypothesis

Partial Dopamine D2 Receptor Agonism

What Happens After the Dopamine D2 Receptor?

The Clinical Role of Antipsychotic Drugs

Interactions between Antipsychotic Drugs and Stress

Conclusions

Glossary

<i>Antipsychotic drugs (APDs)</i>	A diverse range of compounds central to the treatment of psychosis, mania, agitation, behavioral disturbance and sleep disturbance in psychiatric, neurological, and other medical disorders. They are divided into typical and atypical classes depending on whether they have a high (typical) or low (atypical) propensity to induce extrapyramidal side effects.
<i>Dopamine</i>	A catecholamine neurotransmitter central to reward processes and psychiatric disorders. It binds to five receptors numbered 1 to 5, the dopamine D2 receptor is thought to be most relevant to antipsychotic drug action.
<i>Extrapyramidal symptoms</i>	A varied set of distressing and disabling motor side effects induced by antipsychotic drugs binding to dopamine D2 receptors in the striatum of the brain.
<i>Schizophrenia</i>	A serious psychiatric disorder affecting approximately 1% of the population characterized by psychotic symptoms such as hallucinations and delusions, negative symptoms such as social withdrawal, loss of motivation, loss of emotion and impairments in cognition.

Introduction

Antipsychotic drugs (APDs) are a chemically diverse range of compounds that form the cornerstone of current treatments for schizophrenia and related psychotic disorders; bipolar affective disorder; and psychosis or agitation in any psychiatric, neurological, or general medical disorder. They have transformed the practice of psychiatry allowing the closing of large scale institutions and the return of patients to community based treatment. In the United States in 2005, they were the 20th most dispensed class of agent with 43.8 million prescriptions but were the third largest by sales value.

This prominence belies their entirely serendipitous development. The first antipsychotic, chlorpromazine, was synthesized in 1950 by Paul Charpentier as a preanesthetic sedative based upon the antihistamine, promethazine. Its use in 1951 by the French surgeon, Laborit, resulted in his observation of patients displaying a strong indifference to the environment in the presence of full consciousness which he termed artificial hibernation. Chlorpromazine was introduced into psychiatry in 1952 by Delay and Deniker and it revolutionized the treatment for schizophrenia, other psychotic disorders, and agitation leading to the spawning of numerous chemically related phenothiazine compounds. Paul Janssen discovered the next class of antipsychotic agents, butyrophenones, also serendipitously when developing analgesics derived from meperidine. These agents, which include haloperidol and droperidol, are structurally dissimilar to phenothiazines but possessed similar clinical efficacy despite being much more potent and likely to induce distressing and disabling motor side effects termed extrapyramidal symptoms (EPS).

Clozapine, a dibenzazepine, was synthesized in 1958 and is distinct from other APDs in not inducing EPS. It was also subsequently noted that clozapine appeared to be effective for patients who were otherwise unresponsive to APDs. It was, however,

withdrawn following its association with cases of fatal agranulocytosis in the early 1970s. A re-evaluation of clozapine and in particular its effectiveness in treatment-resistant patients with schizophrenia led to its re-introduction under strict hematological monitoring in the late 1980s.

This capacity of clozapine to be effective in schizophrenia and related psychotic disorders whilst not inducing EPS stimulated the development of a new generation of APDs, the so-called atypical or second-generation APDs. These atypical APDs currently include risperidone, olanzapine, quetiapine, amisulpride, ziprasidone, aripiprazole, sertindole, and zotepine, which in the developed world have very much replaced the first generation or typical phenothiazine, butyrophenone, and similar APDs.

This article will discuss the pharmacology of APDs and their proposed mechanism of action with particular reference to the difference between atypical and typical agents. A brief description of their clinical effects will follow and a section of the interaction of APDs with stress systems will conclude the article.

Mechanism of Action of Antipsychotic Drugs

The ongoing debate about the mechanisms of action of APDs underscores the heterogeneity of their pharmacology. The vast majority of APDs have multiple receptor affinities involving numerous neurotransmitter systems. These systems include dopamine, serotonin, muscarinic acetylcholine, norepinephrine (norepinephrine), and histamine. Disentangling sites of action relevant to clinical outcomes remains an ongoing and central task for the development of better therapeutics.

The original critical observation in understanding how APDs worked was that treatment with APDs in mice resulted in an increased turnover of monoamines in brain tissue. The relevant monoamine was specifically identified as dopamine and this was supported by clinical evidence implicating dopamine as psychotogenic. In particular, drugs that induced dopamine release, including amphetamine and cocaine, or stimulated dopamine production, levodopa, caused psychotic symptoms in control subjects and exacerbated psychotic symptoms in patients with stable schizophrenia. The search for dopamine's germane target led to the identification of the dopamine D2 receptor (DAD2R) and the evidence supporting the DAD2R as the main site of action of APDs is now powerfully robust. These data include: the close correlation between the K_i value for the DAD2R for each drug and its clinically effective dose; the

therapeutically effective APD, cis-flupenthixol, pharmacologically differs to its clinically ineffective trans isomer only in its much greater potency as a DAD2R antagonist; no antipsychotic has been identified to date that does not have significant antagonism for the DAD2R; and neuroimaging studies in people with schizophrenia demonstrate greater than 65% occupancy of DAD2R in the striatum is necessary and sufficient for antipsychotic response. However, blockade or loss of the DAD2R in the striatum is also known to induce EPS. Thus, how atypical APDs like clozapine were effective yet with low or no propensity to induce EPS presented a paradox.

Theories of Atypical Antipsychotic Drug Action

How atypical APDs are effective in the treatment of psychotic symptoms but do not induce EPS has not been definitively established. Neuroimaging studies indicate that the EPS syndromes of Parkinsonism and akathisia are induced by striatal DAD2R occupancies of approximately 78% or higher while therapeutic efficacy as stated above occurs with striatal DAD2R occupancies of approximately 65–70%. Parenthetically, it is unlikely that striatal DAD2R occupancy is directly relevant to antipsychotic action but is used as a proxy measure for DAD2R occupancy in other more relevant brain regions. A number of theories have been proposed to explain why atypical drugs are able to achieve striatal DAD2R occupancy levels commensurate with therapeutic efficacy but not cross the EPS threshold and these will be reviewed briefly below.

The Serotonin 5HT_{2A} Receptor

Using the prototypical atypical APD clozapine, investigators have focused upon its receptor binding profile in an attempt to understand the mechanisms underpinning its atypical properties. Its high affinity for the serotonin 5HT₂ and more specifically the 5HT_{2A} receptor led to the proposal that antagonism at this receptor at therapeutic doses may underlie atypicality by modulating DAD2R occupancy in the striatum and by participating in antipsychotic activity. However, as chlorpromazine and other typical drugs also have high affinities for this receptor, this was refined to a relatively greater affinity for the 5HT₂ compared to the DAD2R expressed as a ratio for each drug. It received support from both *in vivo* and *in vitro* data demonstrating serotonin 5HT₂ receptor-induced increases in dopamine release in the striatum and preliminary reports of clinical efficacy of 5HT_{2A} receptor antagonists.

Much evidence now exists, however, to indicate that 5HT_{2A} receptor antagonism is neither necessary nor sufficient for an APD to exhibit atypical properties. Most critical is that the clinically effective APDs remoxipride and amisulpride do not induce EPS at therapeutic doses but have minimal affinity for the 5HT_{2A} receptor. Using positron emission tomography (PET) neuroimaging, chlorpromazine binds to 65% of cortical 5HT_{2A} receptors at a clinical dose of 500 mg that induces EPS and risperidone at a moderately high clinical dose of 6 mg occupied 95% of 5HT_{2A} receptors but induced EPS in six of seven subjects indicating that 5HT_{2A} receptor occupancy does not ameliorate EPS clinically. This clinical data is also supported by animal data minimizing the relevance of 5HT_{2A} receptor occupancy in reducing EPS.

With respect to the role of 5HT_{2A} receptor antagonism in APD efficacy, it appears similarly unnecessary: as noted above both atypical and many typical APDs do not exhibit affinity for the 5HT_{2A} receptor; 5HT_{2A} receptor antagonists without appreciable affinity for the DAD2R (e.g., ritanserin or MDL 100907) do not have antipsychotic efficacy; PET studies of therapeutic efficacy in schizophrenia demonstrate that the obligatory DAD2R occupancy threshold of 65% is irrespective of 5HT_{2A} receptor occupancy; and high levels of 5HT_{2A} receptor blockade occur at subtherapeutic doses of risperidone, clozapine, olanzapine, and quetiapine.

A further elaboration has been the proposal that inverse agonism at the 5HT_{2A} receptor may contribute to antipsychotic effect. However, both amisulpride and remoxipride do not possess this property and MDL 100907 is a potent 5HT_{2A} receptor inverse agonist but is not an APD suggesting this is also not relevant to the efficacy of APDs.

The Dopamine D3 and D4 Receptors

Similar structurally to the DAD2R, the dopamine D3 and D4 receptors (DAD3R and DAD4R), also bind intracellular inhibitory proteins (Gi/o) thereby inhibiting adenylyl cyclase and thus the three receptors comprise the dopamine D2-like class of dopamine receptors. The high affinity of clozapine for the DAD4R, led to the postulation that binding to this receptor was responsible for its antipsychotic properties while its lower DAD2R occupancy explained its absence of EPS. However, haloperidol and other typical APDs have similar affinities as clozapine for DAD4R and a relatively specific DAD4R antagonist, L-745870, was without antipsychotic effect. Thus, there is little support to date for DAD4R in the action of APDs or in explaining atypical effects.

The pharmacological similarity of DAD3R to the DAD2R and its localization to the mesolimbic

dopamine system, a brain circuit closely implicated in the pathology of schizophrenia have offered it a possible role in the pathology of schizophrenia and action of APDs. The evidence for this includes the high affinity that many typical and atypical APDs exhibit for the DAD3R including the relative specificity of amisulpride as an antagonist for DAD2R and DAD3R implicating DAD3R in explaining its atypical properties. Conversely, clozapine has low affinity (K_i 500 nM) for the DAD3R suggesting that DAD3R antagonism is not necessary for either APD action or in atypicality. The long availability of specific DAD3R antagonists yet the absence of any clinical data in schizophrenia implies the absence of any striking clinical utility, although this awaits publication.

Limitations of the Dopamine D2 Receptor Hypothesis

The weakness of evidence supporting 5HT_{2A}, D3, or D4 receptors in the mechanism of atypical APD action should not entail that the DAD2R hypothesis is flawless. Single photon emission tomography (SPECT) neuroimaging studies demonstrate that some subjects with schizophrenia responded to APD treatment at striatal DAD2R occupancy rates well below 65% and some did not respond with occupancy rates over 90%. Moreover, PET studies of clozapine, and more recently quetiapine, only reached striatal DAD2R occupancies of 0–59% at clinically effective doses indicating that DAD2R occupancy might not be essential for APD efficacy.

Two major hypotheses have been proposed to explain why low striatal DAD2R occupancy rates are congruent with clinically effective doses of clozapine and quetiapine while retaining the centrality of the DAD2R in the mechanism of APD action. One postulates that clozapine and quetiapine at clinically relevant doses preferentially bind to cortical (principally temporal cortex including amygdala and hippocampus) over striatal DAD2R in subjects with schizophrenia. The other proposes that drugs like clozapine and quetiapine dissociate rapidly from the DAD2R. Here, the APD occupies the DAD2R either *in vivo* or *in vitro* immediately it is introduced but within minutes *in vitro* or hours *in vivo* dissociates from the receptor. This contrasts with typical APDs where there is a many fold longer time of binding between drug and DAD2R both *in vivo* and *in vitro*. The rapid dissociation of clozapine and quetiapine from the DAD2R allows the receptor to then be activated by the release of endogenous dopamine ameliorating any EPS.

Neither explanation is incontrovertible, thus the centrality of DAD2R antagonism in APD action is not infallible given the above anomalies in the data. However, it remains the most plausible of hypotheses and at this time there does not seem an overwhelming need to invoke other receptor systems to explain the basic actions of APDs nor the difference between atypical and typical agents. This, of course, does not imply other receptor systems such as the DAD3R, glutamate, γ -aminobutyric acid (GABA), serotonin, acetylcholine, or other neuromodulatory systems may not be involved, and in fact may be involved in other properties of APDs, however, their role in primary antipsychotic efficacy remain to be established.

Partial Dopamine D2 Receptor Agonism

A variant hypothesis to DAD2R antagonism is partial agonism at this receptor based on animal studies of the pharmacology of aripiprazole as the leading exemplar of its class. Both short- and long-term clinical trials of aripiprazole demonstrate it to be an effective APD with low EPS. In animal models where dopamine hyperactivity is present (e.g., apomorphine-induced stereotypy) aripiprazole functions as a DAD2R antagonist, however, in models of dopamine depletion (e.g., reserpine treatment) it exerts a partial agonist activity. Thus it could be effective by functioning as a relative antagonist on DAD2R signaling pathways relevant to psychosis yet permit sufficient dopamine signaling in the striatum to prevent EPS.

What Happens After the Dopamine D2 Receptor?

Although APD binding to the DAD2R appears to be central to their mechanism of action, this can be only a first step in modifying neuronal activity. This is congruent with the time between onset of receptor blockade within 1 hour and clinical efficacy which takes days to weeks. The most plausible sequence of events center on the time required for inhibition of the DAD2R coupled Gi/o protein to result in activation of adenylyl cyclase, increase cyclic AMP levels, and activate protein kinase A (PKA). PKA activation, in addition to phosphorylation of other kinases, receptors, and ion channels, activates transcription factors including the cyclic AMP response element binding protein (CREB). These transcription factors in turn regulate gene expression of receptors, neuropeptides, and synaptic proteins ultimately resulting in *de novo* protein synthesis some days to weeks after initial receptor blockade. It is important to consider, however, that other signaling pathways are also linked to

the DAD2R and their relevance cannot be excluded. Less clearly established is the identity of key proteins regulated through the above signaling pathways and how they effect changes in neuronal function leading to antipsychotic effects. Observations of regional brain differences in activation of immediate early genes between typical and atypical APDs may give direct clues as to which proteins in which brain regions are responsible for EPS and which proteins mediate APD efficacy.

One set of proteins implicated in APD activity involve synaptic plasticity and connectivity. These proteins are decreased in schizophrenia and increased in animals treated with typical and atypical APDs. This raises the possibility that APDs may stimulate synaptic reorganization. It remains to be determined whether such changes occur in humans when responding to the clinical effects of APDs.

The Clinical Role of Antipsychotic Drugs

As noted above, APDs are used in the treatment of a diverse range of psychiatric, neurological, and medical disorders where short term they control psychotic symptoms such as hallucinations and delusions and also agitation and behavioral disturbance. These symptoms are the cardinal features of schizophrenia and related psychotic disorders but may also be observed in bipolar affective disorder, substance-induced states, major depression with psychosis, dementia, delirium, posttraumatic stress disorder, borderline personality disorder, a variety of neurological disorders including Huntington's disease, multiple sclerosis, Parkinson's disease, temporal lobe epilepsy, and medical disorders including endocrine, metabolic, autoimmune, and neoplastic disorders. Furthermore, APDs are effective in long-term prophylaxis of relapse of psychosis in schizophrenia and related disorders and potentially bipolar affective disorder.

As discussed above, the defining difference between typical and atypical APDs is the low or absent propensity of the latter to induce EPS. Nevertheless, the effectiveness of clozapine in treatment-resistant patients with schizophrenia has generated much interest as to other clinical differences between typical and atypical agents. First, there are unequivocal data showing clozapine is superior to typical APDs in treatment-resistant patients but there are now emerging data that clozapine is also superior to other atypical agents. Second, there are less well established reports in patients with schizophrenia demonstrating clozapine to be effective in suicidality; cognitive dysfunction (e.g., working memory, attention, and verbal fluency); negative symptoms (e.g., social withdrawal, lack of motivation and diminished states of feeling),

and substance abuse. The superiority of clozapine in these symptom domains is, however, far from established. Moreover, current evidence would suggest that apart from the effectiveness of clozapine in treatment resistant schizophrenia, there is no clear difference in efficacy between typical and atypical APDs.

Befitting the chemically diverse range of compounds that constitute APDs, the side effect profiles of these agents are varied. The reduced propensity for atypical APDs to induce EPS has been noted and is an advantage for this group over typical APDs. However, the risk is not absent and needs to be offset against other adverse effects. In addition, prolactin and other endocrine effects and cardiac rhythm disturbances are commonly associated with both typical and atypical APDs. A burgeoning area of concern is the association of APDs with the metabolic syndrome, characterized by weight gain and central obesity, hyperlipidemia, hyperglycemia, insulin resistance, and diabetes mellitus type 2 and the increased risk of cardiovascular morbidity and mortality. This has been especially linked to clozapine and olanzapine but may also apply to chlorpromazine and other typical APDs. Furthermore, the association of clozapine with agranulocytosis (1%), epileptic seizures (dose dependent at 5% 600 mg day⁻¹), myocarditis (0.05%), and cardiomyopathy (0.01%) indicates the potential toxicity of these agents. This range and extent of toxicities of APDs highlight the importance of individual monitoring and detailed discussion between physician, patient, and carers when making decisions about APD treatment options.

Interactions between Antipsychotic Drugs and Stress

The interaction between stress hormones and pathways and the major psychiatric disorders such as schizophrenia, bipolar affective disorder, and major depression that are treated with APDs are well described. This has led to explorations into the modulation of stress hormone systems, especially the hypothalamic-pituitary-adrenal (HPA) axis, by APDs and in particular if differences exist between typical and atypical APDs. Hence, it has been proposed that such differential effects may invoke differences in treatment outcomes based upon the effects of HPA dysregulation on the pathology of schizophrenia. Some support for this exists with clozapine, olanzapine, and risperidone but not haloperidol decreasing cortisol and/or adrenocorticotrophic hormone (ACTH) levels and this being associated with changes in psychopathology. Furthermore, studies in healthy subjects found no effect on cortisol levels by the

typical APDs, haloperidol and sulpiride but decreases with olanzapine and quetiapine. In contrast no effect on cortisol levels was observed with amisulpride, an atypical APD with relative specificity for DAD2R and DAD3R.

The regulation of ACTH by monoamines including serotonin and dopamine imply that all APDs should influence ACTH and cortisol levels. However, the decrease in these levels in both control and schizophrenic subjects by atypical APDs has been proposed to be attributable to 5HT_{2A} receptor antagonism. In support of this is the observation that the 5HT_{2A} receptor antagonist, ritanserin, attenuated increases in psychotic symptoms and plasma cortisol levels in schizophrenic patients administered the serotonin agonist *m*-chlorophenylpiperazine. Similarly, the relatively high affinities for the 5HT_{2A} receptor for clozapine, olanzapine, quetiapine, and risperidone but the low affinity of amisulpride is also consistent with this proposal. However, it is also feasible that the attenuation in cortisol and ACTH by atypical APDs could also be attributable to their adrenergic or histaminergic receptor affinities.

The clinical relevance of atypical APDs induced decreases in cortisol and ACTH plasma levels needs further investigation. Two reports have observed a correlation between cortisol levels and negative symptoms in schizophrenia; however, this was not correlated with the dose of risperidone and was not observed in another study. Of greater interest are the relationships of the HPA axis to depressive disorders and the potential role of atypical APDs in these disorders. Here it may be that downregulation of the HPA axis by atypical APDs and the reduction in glucocorticoids may have a beneficial effect on mood and warrants further investigation.

Conclusions

APDs provide the major plank in the treatment of many psychiatric and neurological disorders. The division into atypical and typical classes is based upon their relative propensities to induce unwanted and disabling extrapyramidal motor side effects. Despite their generally widespread monoamine receptor affinities, both antipsychotic efficacy and propensity to induce EPS may be explained by their antagonism at the DAD2R and does not require the invocation of other systems. However, the challenges remain in determining specifically how atypical agents occupy less DAD2R in the striatum while exerting an antipsychotic effect and, moreover, as for all APDs, characterizing what post-DAD2R pathways are essential for antipsychotic efficacy. Furthermore, the unique

clinical profile of clozapine does present etiological challenges to the DAD2R hypothesis and understanding the reasons for its effectiveness in treatment-resistant schizophrenia holds promise for improving the outcomes for this and other disorders. Finally, as the modulation by APDs of systems such as serotonin, acetylcholine, norepinephrine, histamine, the HPA axis, and others is further studied it may become feasible to expand and better target the role of these drugs in the treatment of particular symptoms and disorders.

Further Reading

- Altamura, A. C., Boin, F. and Maes, M. (1999). HPA axis and cytokines dysregulation in schizophrenia: potential implications for the antipsychotic treatment. *European Neuropsychopharmacology* 10, 1–4.
- Cohrs, S., Roher, C., Jordan, W., et al. (2006). The atypical antipsychotics olanzapine and quetiapine, but not haloperidol, reduce ACTH and cortisol secretion in healthy subjects. *Psychopharmacology* 185, 11–18.
- Gardner, D. M., Baldessarini, R. J. and Waraich, P. (2005). Modern antipsychotic drugs: a critical overview. *Canadian Medical Association Journal* 172, 1703–1711.
- Greengard, P. (2001). The neurobiology of slow synaptic transmission. *Science* 294, 1024–1030.
- Kane, J. M., Honigfeld, G., Singer, J., et al. (1988). Clozapine for treatment-resistant schizophrenics. *Archives of General Psychiatry* 45, 789–796.
- Kapur, S. and Remington, G. (2001). Atypical antipsychotics: new directions and new challenges in the treatment of schizophrenia. *Annual Review of Medicine* 52, 503–517.
- Konradi, C. and Heckers, S. (2001). Antipsychotic drugs and neuroplasticity: insights into the treatment and neurobiology of schizophrenia. *Biological Psychiatry* 50, 729–742.
- Lieberman, J. A., Stroup, T. S., McEvoy, J. P., et al. (2005). Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *New England Journal of Medicine* 353, 1209–1223.
- Marder, S. R., Essock, S. M., Miller, A. L., et al. (2002). The Mount Sinai conference on the pharmacotherapy of schizophrenia. *Schizophrenia Bulletin* 28, 5–16.
- McEvoy, J. P., Lieberman, J. A., Stroup, T. S., et al. (2006). Effectiveness of clozapine versus olanzapine, quetiapine, and risperidone in patients with chronic schizophrenia who did not respond to prior atypical antipsychotic treatment. *American Journal of Psychiatry* 163, 600–610.
- Newcomer, J. W. and Haupt, D. W. (2006). The metabolic effects of antipsychotic medications. *Canadian Journal of Psychiatry* 51, 480–491.
- Nyberg, S., Nakashima, Y., Nordstrom, A. L., et al. (1996). Positron emission tomography of in-vivo binding characteristics of atypical antipsychotic drugs. Review of D2 and 5-HT2 receptor occupancy studies and clinical response. *British Journal of Psychiatry* suppl: 40–44.
- Roth, B. L., Sheffler, D. J. and Kroeze, W. K. (2004). Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. *Nature Reviews. Drug Discovery* 3, 353–359.
- Seeman, P., ChauWong, M., Tedesco, J., et al. (1975). Brain receptors for antipsychotic drugs and dopamine: direct binding assays. *Proceedings for the National Academy of Sciences of the United States of America* 72, 4376–4380.
- Sundram, S., Joyce, P. R. and Kennedy, M. A. (2003). Schizophrenia and bipolar affective disorder: perspectives for the development of therapeutics. *Current Molecular Medicine* 3, 393–407.

Antisocial Disorders

K Pajer

The Ohio State University, Columbus, OH, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by K Pajer, volume 1, pp 222–225, © 2000, Elsevier Inc.

Definition

Development and Course

Prevention and Treatment

Glossary

Antisocial personality disorder (ASPD)

Personality disorder characterized by frequent violation of social and legal norms, grandiose thoughts, arrogance, manipulative behavior, superficial charm, lack of attachment to others, little empathy, and an absence of guilt. Synonyms are sociopathy, psychopathy, and dyssocial personality.

Conduct disorder (CD)

Behavior disorder of childhood and adolescence characterized by truancy,

	stealing, running away, fighting, vandalism, fire setting, sexual assault, or other illegal activities.
<i>Continuity</i>	A term used in developmental psychopathology to refer to behaviors that continue from one stage of life to another, e.g., from adolescence to adulthood.
<i>Desistance</i>	A term used by criminologists to describe the cessation of criminal activity in a person who has previously committed illegal acts.
<i>Developmental psychopathology</i>	A field of study that focuses on the development and longitudinal course of psychiatric disorders across the life cycle.
<i>Discontinuity</i>	A term from the field of developmental psychopathology referring to behaviors that do not persist from one life stage to another.
<i>Escalation</i>	A term used in criminology to describe an increase in the frequency of illegal acts or an increase in the severity of such behavior over time.

Definition

Clinical Example

A.W. is a 40-year-old male who was released from prison 6 months ago. This was his second prison term; both terms had been served for burglary. A.W. was first sent to prison for 3 years when he was 25 years old. He had been working as a security system installer. He would install an alarm system and then return later and rob the house. A.W. proudly says that he robbed more than 800 houses in a 3-year period. After he served his first sentence, he learned how to open safes. He did several large jobs in houses and businesses before getting caught and sent back to prison for 10 years.

A.W. has a history of violence, consisting of fights in bars and on the street with other men, and fighting with his wife and girlfriends. His family believes he may have murdered his first wife and successfully hidden the body, but this has never been proven. A.W. openly admits that he initiates violence because he is bored or irritated. He himself was a childhood victim of severe violence perpetrated by his alcoholic father, who would also attack A.W.'s mother. He reports that he was beat many times with his father's hands and fists, as well as with a belt, board, and electrical cord. He suffered multiple episodes of head trauma. A.W. started fighting peers when he was 12 and hit his father for the first time when he was 14.

School was not easy for A.W., and he had numerous learning problems. He left school in the seventh grade at the age of 15 and joined the Army, using

his dead brother's Social Security number. Testing in the Army revealed that he was bright and actually was functioning on a 12th grade level, but he got into so many fights that he eventually received a dishonorable discharge.

A.W.'s health is not good. He suffered his first grand mal seizure while in prison 10 years ago. He now has them regularly, despite medication. He has chronic bronchitis and hypertension. He used to drink heavily, but stopped 10 years ago when put back into prison, and reports that he will not drink again. He denies any drug use other than early experimentation with marijuana.

There have been several long-term female relationships in A.W.'s life, although he has had multiple brief liaisons. He has seven children, born to three different women. His first wife disappeared after the birth of their second child, and it was widely known that A.W. was angry because he thought she was seeing another man. He has had two other important relationships, and both these women have complained that he was physically abusive. He denies being abusive to his children, but has not been home enough to actually parent them. He reports sadness about that, as four of his children are in foster care.

Classification of Antisocial Disorders

The adult syndrome of antisocial personality disorder (ASPD) was first described by Harvey Cleckley in his famous book *The Mask of Insanity*. Cleckley's original characterization of interpersonal, behavioral, and affective deviance has remained the core of subsequent diagnostic criteria. As exemplified by A.W.'s history, people with this disorder are narcissistic, superficially charming, manipulative, often aggressive, and short-tempered; have multiple, unstable relationships; lack empathy and guilt; and display irresponsible behavior, often with overt criminality. The *Diagnostic and Statistical Manual*, 4th edition (DSM-IV) has defined ASPD as a persistent pattern of behavior that does not take into account the rights of others or that seeks to do overt harm to others. Criteria include an inability to maintain a stable relationship or employment, irresponsible behavior such as avoidance of financial obligations or parenting duties, cruel behavior toward animals or humans, and repeated arrests (or illegal activities, even if not caught). The *International Classification of Diseases*, 10th edition (ICD-10) has similar criteria for dis-social personality. Most adults with ASPD began to display deviant behavior in childhood or adolescence. Youths with three or more of the following symptoms by age 13 are classified in the DSM-IV as having

conduct disorder (CD): stealing with or without confrontation with the victim, running away, acts of cruelty or bullying, fighting, truancy, forcing someone into sexual relations, carrying or using a weapon, vandalism, and arrests.

Psychiatric diagnoses are traditionally made through clinical interviews, interviewer-rated symptom questionnaires, or self-report questionnaires. Each of these methods is difficult to use with the person who has antisocial personality features. These patients are rarely suffering or complaining of symptoms. In fact, in a brief interaction such as an examination, they may be quite charming and engaging. They are often puzzled as to why a psychiatric evaluation has been requested and usually blame others for their actions. Self-report inventories are problematic because of the high rate of lying in this population. The two most frequently used methods of assessment are highly structured interviews that follow strict diagnostic criteria such as those found in DSM-IV or the Hare Psychopathy Checklist (PCL). Filling out the PCL requires the use of multiple sources of information, rather than just using a clinical interview.

Epidemiology

The estimated prevalence of CD in prepubertal children is between 2 and 8% for boys and is less than 2% for girls. The gender gap narrows in adolescence and rates increase, with 3–12% of boys meeting criteria for the diagnosis compared to 8–10% of girls. CD is more common in urban areas than in suburban or rural neighborhoods, but poverty is a consistent correlate.

From 5 to 15% of the general adult population may meet criteria for this disorder; in the prison populations, the estimates range from 20 to 80%. ASPD is also much more common in adult males than in females.

Development and Course

In general, most of the studies on the development and course of antisocial disorders have studied males, although more studies are now including antisocial females. The majority of the studies have also used delinquent or criminal populations, so it is often difficult to generalize findings to nonincarcerated populations.

Biological Factors

There are several major lines of research into the biological factors that may underlie antisocial behaviors. One hypothesis is that CD or ASPD is caused by a general hypoarousal of the stress response system,

producing a lack of response to stimuli that would normally provoke stress, fear, or pain. Numerous studies of men and boys with antisocial behavior demonstrate that their autonomic arousal, as measured by changes in skin conductance, heart rate, or blood pressure, is blunted in response to stimuli such as psychological threat, noise, or pain. The few studies conducted in girls have also reported blunted responses. The hypothalamic-pituitary-adrenocortical (HPA) axis also appears to be hypoactive in both the circadian secretion of cortisol and the response to psychological stimuli. These findings have been reported for boys, girls, and men who have antisocial behavior.

This lack of response would lead people with antisocial disorders to pursue excessive stimulation, manifested by sensation-seeking and risk-taking behaviors. Furthermore, they would be unable to learn from aversive consequences or punishment because their stress response would not be robust enough. Studies of heart rate, skin conductance, electroencephalogram (EEG) changes, and basal cortisol levels in subjects with antisocial behavior all support a theory of hypoarousal. Some of these characteristics in childhood and adolescence have been reported to predict adult sociopathy. Indicators of hypoarousal also differentiated between sociopathic and nonsociopathic subjects in a group of men whose fathers were all incarcerated, implying that the abnormalities in arousal may have a genetic basis.

A second line of research on putative biological mechanisms for sociopathy stems from neuropsychological and neuroimaging studies. Both types of data indicate that antisocial behavior is associated with abnormalities in the prefrontal and frontotemporal areas of the brain. These areas of the brain are responsible for the executive functions of the brain: impulse control, evaluation and integration of new data, and regulation of emotional tone and motivation. Abnormalities in these brain areas could lead to a pattern of sociopathic behavior by reducing the ability to delay gratification, decreasing the ability to learn new behaviors in response to past mistakes, and producing intense emotional states such as rage on an unpredictable basis.

Neurotransmitter function and gonadal hormonal influences on aggressive and impulsive behaviors in subjects with sociopathy have also been investigated. The findings in both domains are somewhat contradictory, but this may be due to methodological differences between studies.

In general, decreased serotonin neurotransmission in the central nervous system appears to be associated with aggression, although some researchers suggest that the link may actually be between low serotonin and impulsivity. Testosterone levels appear to be

higher in some groups of violent males, but the relationship seems to be a reflection more of dominance than of pathological aggression.

Stress

The association between stress and antisocial behavior is complex and poorly understood. Many adults and children with antisocial behavior have histories of early exposure to adverse environmental circumstances. Similarly, longitudinal studies of young children demonstrate that many of them who experience obstetrical complications, early abuse or neglect, or repeated traumatizing events develop antisocial behavior by adolescence. This is particularly true for boys.

However, both adults and children with antisocial behavior report far lower subjective stress about these events than most people experience. The stress response system hypoactivity described previously is at odds with what we know from previous clinical and animal studies about exposure to adverse events. In these studies, the resulting stress response system dysfunction is most often characterized by either normal or increased resting function accompanied by hyperreactive and prolonged responses to stimuli. These physiological changes are usually accompanied by a heightened subjective experience of distress in humans. A chronically aroused state such as this can be associated with significant physical and psychiatric morbidity (e.g., hypertension, osteoporosis, dementia, depression). Therefore, one current theory is that stress response system hypoactivity in people with antisocial behavior may be protective, allowing them to survive in very difficult environments.

Childhood Risk Factors

Several risk factors present in childhood are associated with adult sociopathy. One of the most consistent findings, whether measured retrospectively or prospectively, is exposure to inconsistent and harsh parental discipline. The severity of the punishments can range from frequent verbal shaming to overt physical abuse, but the key prognostic factor appears to be the unpredictability of the hostile parental behaviors.

Children who have attentional problems with hyperactivity seem to be at particularly high risk of developing antisocial behavior. These studies have reported the same results for boys and girls. If the child also has a low IQ and one antisocial parent, then early sociopathic behavior is quite likely.

Developmental Paths

The course or path of childhood and adolescent antisocial behavior varies from person to person, but the

continuity between this type of behavior and adult sociopathy is one of the most robust findings in clinical psychiatric research. Children who begin to display antisocial behavior prepubertally appear to be at the highest risk for severe sociopathy and progression to overt adult criminality.

Both childhood and adolescent-onset CD may progress through levels of increasing severity and frequency of the behaviors, a process known as escalation. Escalation is associated with an early onset of antisocial behavior, a history of family sociopathy, violence and aggressive behaviors, social marginalization in childhood, school failure, attention deficit and hyperactivity symptoms, and continued association with deviant peers. The behaviors may also simply cease in adolescence or early adulthood (desistence). Desistence is associated with a later onset of sociopathic behavior, an absence of aggression, a strong family structure, and early school and social success. The course of the disorder may also change when a person becomes involved with drugs or alcohol.

The majority of data on developmental paths have come from studies of males. Although less is known about females, a comprehensive review of the existing literature reported that girls with CD also appear to be at an increased risk for adult sociopathy. Additionally, many of these girls develop other psychiatric disorders and have a great deal of difficulty functioning as mothers or employees. Fewer girls progress to overt criminality when compared to boys, but many remain aggressive and destructive toward family members (e.g., their children or spouses). Researchers are in disagreement about whether a subgroup of girls exists analogous to the subgroup of boys with early-onset antisocial behavior and a particularly severe and intractable course.

Prevention and Treatment

As discussed earlier, antisocial behaviors often begin in childhood or adolescence. Once established, they are very difficult to treat. As a result, researchers in the past two decades have begun to study the effectiveness of prevention. Some innovative prevention programs have reported surprisingly good results.

Although the programs differ somewhat, they have one common feature. All of them provide services to the child in every domain of his or her life. The programs often focus on mothers identified to be at risk for producing antisocial children and work to improve maternal physical and psychological health during pregnancy or following delivery. Didactic parenting training, coupled with close supervision by a

visiting nurse, has also been helpful. Helping the mothers to establish their own support systems and showing them how to negotiate through the intricacies of the medical, welfare, and mental health systems has also been effective. The second type of preventive program focuses on assisting at-risk children in learning social skills in preschool and the community and early recognition and treatment of comorbid disorders such as attentional problems. Prevention programs that focus on whole communities of children or teens at risk have also had some success. These interventions usually occur in selected schools and consist of improving self-esteem through academic achievement or increasing prosocial behavior with game-based activities and rewards.

Standard psychiatric or psychological treatments are not very useful with patients who have antisocial disorders. The most effective treatment strategies are multimodal. Interventions that combine individual cognitive-behavioral psychotherapy, pharmacotherapy for aggression or hyperactivity, parenting training, school supervision, and anger management skills (often for the entire family) have the highest rates of success.

Once a youth has committed an illegal act and is arrested, or adjudicated delinquent, it is much harder for him or her to receive treatment rather than punishment. Data suggest that employing multimodal treatment programs, in combination with a compensation process, to atone for the offense and psychiatric treatment of comorbid disorders works better than incarceration at promoting desistance.

The same treatment dilemma exists for adults with ASPD. Those who do not engage in criminal behavior rarely present for treatment unless they are in the midst of a crisis such as a divorce. The personality characteristics of these patients make them poor candidates for psychotherapy. They may be noncompliant with any type of long-term treatment because they are only temporarily motivated to reduce the discomfort produced by the crisis. Adults with ASPD who have committed crimes do present for treatment while incarcerated, but this is usually because of comorbid psychiatric disorders such as depression. Pharmacological treatment of these comorbid problems or specific problematic behaviors such as aggression is possible, but the underlying pattern of sociopathic behavior is rarely altered.

See Also the Following Articles

Aggression; Aggressive Behavior; Autonomic Nervous System; Child Abuse; Corticosteroids and Stress.

Further Reading

- Cleckley, H. (1976). *The mask of sanity. An attempt to clarify some issues about the so-called psychopathic personality*. St. Louis, MO: The C. V. Mosby Company.
- Ehrensaft, M. K., Moffitt, T. E. and Caspi, A. (2004). Clinically abusive relationships in an unselected birth cohort: men's and women's participation and developmental antecedents. *Journal of Abnormal Psychology* 113, 258–270.
- Farrington, D. P., Loeber, R., Yin, Y., et al. (2002). Are within-individual causes of delinquency the same as between-individual causes? *Criminal Behaviour and Mental Health* 12, 53–68.
- Hare, R. D., Hart, S. D. and Harpur, T. J. (1991). Psychopathy and the DSM-IV criteria for antisocial personality disorder. *Journal of Abnormal Psychology* 100, 391–398.
- Kazdin, A. E. (2002). Family and parenting interventions for conduct disorder and delinquency: a meta-analysis of randomized controlled trials. *Journal of Pediatrics* 141, 738.
- Laub, J. H. and Sampson, R. J. (2003). *Shared beginnings, divergent lives. Delinquent boys to age 70*. Cambridge, MA: Harvard University Press.
- Lilienfeld, S. O. (1998). Methodological advances and developments in the assessment of psychopathy. *Behaviour Research and Therapy* 36, 99–125.
- Loeber, R., Stouthamer-Loeber, M., Farrington, D. P., et al. (2002). Three longitudinal studies of children's development in Pittsburgh: the Developmental Trends Study, the Pittsburgh Youth Study, and the Pittsburgh Girls Study. *Criminal Behaviour and Mental Health* 12, 1–23.
- Morgan, A. B. and Lilienfeld, S. O. (2000). A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. *Clinical Psychology Review* 20, 113–136.
- Pajer, K. (1998). What happens to "bad" girls? A review of the adult outcomes of antisocial adolescent girls. *American Journal of Psychiatry* 155, 862–870.
- Pajer, K., Gardner, W., Rubin, R., et al. (2001). Decreased cortisol levels in adolescent girls with conduct disorder. *Archives of General Psychiatry* 58, 297–302.
- Raine, A. (2002). Biosocial studies of antisocial and violent behavior in children and adults: a review. *Journal of Abnormal Child Psychology* 30, 311–326.
- Silverthorn, P. and Frick, P. J. (1999). Developmental pathways to antisocial behavior: The delayed-onset pathway in girls. *Development and Psychopathology* 11, 101–126.

Anxiety

A Öhman

Karolinska Institutet, Stockholm, Sweden

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A Öhman, volume 1, pp 226–230, © 2000, Elsevier Inc.

Anxiety and Fear

Different Types of Anxiety

Anxiety and Cognition

Anxiety Disorders

The Neuroanatomy of Anxiety

Glossary

<i>Anxiety</i>	Apprehensive anticipation of uncertain, often ill-defined dangers.
<i>Fear</i>	Emotional state associated with attempts to cope with threatening events.
<i>Panic attack</i>	A sudden surge of anxiety associated with physiological symptoms (e.g., heart-rate increases) and catastrophic feelings (e.g., fear of dying) that come suddenly and seemingly unprovoked.
<i>State anxiety</i>	The level of anxiety at a defined point in time.
<i>Trait anxiety</i>	An individual's habitual level of anxiety.

Anxiety is an aversive emotional state associated with the apprehensive anticipation of more or less likely future dangers. It incorporates the somatic symptoms of tension and dysphoric feelings.

Anxiety and Fear

Anxiety is closely related to fear (see **Fear**). The traditional distinction has been that anxiety, in contrast to fear, lacks an obvious eliciting stimulus. According to the definitions given here, anxiety and fear occupy different temporal locations in relation to threatening events, anxiety typically being anticipatory to, and fear elicited by, threatening stimuli. Furthermore, whereas the threat in anxiety often is imagined, in fear it is typically real. These differences promote different phenomenologies in the two states, with anxiety dominated by an unpleasant feeling of foreboding and fear by an urge to escape. However, the perhaps most fruitful distinction can be derived from the controllability of the threat. Fear is predicated on the hope that the situation can be coped with, that it is controllable. Fear, therefore, is a source of motivation

because it supports active attempts to deal with the threat by, for example, escape or avoidance. Anxiety, on the other hand, results when the coping attempt fails, that is, when the situation is perceived as uncertain or uncontrollable. With few or inefficient means to cope with the threat, the option that is left is the apprehensive anticipation of the more or less likely expected disaster.

This analysis gives anxiety and fear a joint origin in an evolved defense system that has served through eons of time to protect organisms from survival threats. The primary emotional counterpart is fear, which supports adaptive coping with threat. However, depending on restrictions in the controllability of the situation, there is a dynamic interplay between the closely related emotional qualities of fear and anxiety. Because fear is the prototype reflected in anxiety, the two emotional states have many similarities, for example, in terms of components (aversive feelings, physiological responses, and behavior; see **Fear**). Thus, the two states may often be hard to distinguish, but this is due to the fact that they are fused in the real world rather than to conceptual obscurity.

Placing anxiety and fear at different regions of a controllability dimension has the added virtue of relating both these states to another related concept dominated by dysphoric feelings – depression. Thus, whereas fear is the emotional concomitant in situations of somewhat controllable threat, anxiety takes over when the uncontrollability of the situation undermines active coping attempts. With further and more lasting uncontrollability, coping becomes fruitless, which makes the organism helpless and replaces anxiety with depression. Thus, the conceptual view advanced here is compatible with the emerging notion of negative affect as a pervasive higher order factor behind both anxiety and mood disorders.

Different Types of Anxiety

Situational versus Free-Floating Anxiety

There are several more or less interdependent bases for distinguishing between different types of anxiety. The controllability analysis puts anxiety on a continuum between fear and depression. There are other related continua that are often invoked in this context. One is that of situational versus free-floating (generalized) anxiety. Situational anxiety is bound to a particular stimulus, such as the stimulus in phobic

anxiety. Whether anxiety of this type should be called anxiety or fear depends on the controllability of the fear stimuli in the situation. In situationally unfocused or free-floating anxiety, the intensity of the anxiety does not vary with the situation, but is always present at a noticeable level.

Panic Attacks versus Generalized Anxiety

Free-floating (generalized) anxiety is correlated with worries (e.g., about finances or academic performance). It is often contrasted with panic attacks, which have been claimed to represent a distinct subtype of anxiety. A panic attack is a sudden surge of often overwhelming anxiety dominated by physical symptoms such as shortness of breath, dizziness, palpitations, trembling, sweating, nausea, hot flashes, and chest pain. It may also involve feelings of depersonalization and fears of dying or going crazy. A panic attack may appear to be spontaneous, coming out of the blue, or to be situationally provoked. For example, a social phobic may experience panic when forced to face a threatening social context.

Some people have occasional panic attacks with no or only little residual anxiety between the attacks. However, because a panic attack is such an overwhelmingly aversive experience, the person may start to develop anticipatory anxiety about further attacks. As a result, occasional attacks will now appear against a fluctuating background of anticipatory anxiety, and the distinction between panic and generalized anxiety will become somewhat blurred.

Factor Analytically Derived Varieties of Anxiety

Factor-analytic work on observed and self-reported anxiety symptoms in normals and patients confirms clinical observations in delineating two clusters of anxiety. Thus, there is one cluster of somatic over-reactivity composed of symptoms such as sweating, flushing, shallow breathing, and reports of heart palpitations, intestinal discomforts, and aches and pains. The other cluster may be called cognitive or psychic anxiety and is composed of intrusive and unwanted thoughts, worrying, ruminations, restlessness, and feelings of muscle tension.

Trait versus State Anxiety

The distinction between panic and generalized (free-floating) anxiety is related to another important distinction. A particular person may suffer from habitually elevated anxiety, never feeling completely free of apprehension and worry. This habitually elevated anxiety is called trait anxiety. Nevertheless, the level of anxiety may still vary somewhat across time and situations. The momentary level of anxiety at a

particular time (e.g., during a panic attack), on the other hand, is called state anxiety. Even though people with high trait anxiety tend also to have high state anxiety in most situations, there may be situations in which the state anxiety of low-trait-anxious people is still higher because of stressful circumstances.

Clinical versus Normal Anxiety

Anxiety is both a phenomenon of everyday life and an important sign of psychopathology. Thus, there is a need to distinguish between normal and clinical anxiety. In general, clinical anxiety is more intense, recurrent, and persistent than normal anxiety. Furthermore, whereas normal anxiety is proportional to the real danger of the situation, the intensity of clinical anxiety is clearly above what is reasonable, given the objective danger involved, tending to paralyze individuals and undermine their coping efforts. Consequently, it results in impeded psychosocial or physiological functioning, which may promote contact with treatment facilities.

Anxiety and Cognition

To be anxious implies a preoccupation with threat and danger. For example, people with high anxiety tend specifically to overestimate the likelihood that they will experience negative events, but they do not differ from people with low anxiety in judging the likelihood that they will experience positive events. Anxiety and cognition, in fact, are intimately related to one another in a two-way communication in which anxiety determines cognition and cognition determines anxiety.

Attentional Bias for Threat

The immobility exhibited by animals that freeze in the face of danger is associated with hypervigilance to the environment. Similarly, there is a large research literature documenting that anxiety is associated with hypervigilance and an attentional bias for threat. For example, when simultaneously exposed to two visual stimuli, one threatening and the other nonthreatening, anxiety patients tend systematically to attend to the threatening stimulus, as revealed by their shortened reaction times to probe stimuli occurring at the location of the threatening stimulus. Low-anxiety people, on the other hand, tend to shy away from threat; their reaction times are faster for probe stimuli at the location of nonthreatening stimuli. This general finding is valid regardless of the threat conveyed by the pictorial (e.g., angry faces) or word (e.g., cancer) stimuli. Similarly, if people high or low in anxiety are asked to name the color in which words are printed

(the Stroop paradigm), the color-naming latency is longer for threatening than for nonthreatening words in anxious (but not in nonanxious) people. This implies that high-anxiety people are distracted by attention to the threatening word content.

The attentional bias for threatening stimuli in highly anxious research participants does not require conscious recognition of the stimuli. Several studies have shown that similar results are obtained when the threatening stimuli are presented masked by other stimuli, thus precluding their conscious recognition. For example, using the Stroop color word paradigm, longer color-naming latencies to threatening words are observed in highly anxious people, even if the word is shown only for 20 ms and immediately masked by illegible letter fragments of the same color. This finding implies that the attentional bias operates at an automatic, preattentive level of information processing. Thus, outside the conscious awareness of individuals, this attentional bias determines that they preferentially attend to threatening events in their surroundings. This preference for focusing attention on threat no doubt helps maintain the anxiety.

Expectancy and Interpretational Bias

Attentional bias provides a case in which anxiety drives cognition. However, it is also abundantly clear that cognition drives anxiety. This is particularly well documented for panic. A compelling argument can be advanced to the effect that catastrophic interpretations provide a necessary condition for the elicitation of panic. According to the dominating model, stress or anxiety results in an autonomic arousal that is perceived by the person. The perception of this bodily change prompts more anxiety and further autonomic arousal in a vicious circle. But it is only when the interpretation of the bodily symptoms becomes catastrophic (e.g., when the person feels that they signal serious illness or imminent death) that the full-blown panic attack emerges.

Several lines of evidence support this model. For example, when panic attacks are pharmacologically promoted in panic patients (e.g., through lactate infusion or CO₂ inhalation), panic is unlikely if patients are informed about the expected effects of the manipulation. If they are not informed, or misinformed, on the other hand, panic patients (but not social phobics) are likely to experience panic. Similarly, if panic patients inhaling CO₂ are led to believe that they can manipulate the CO₂ saturation in the inhaled air by pressing a panic button, this induced sense of illusory control protects them from panic attacks and decreases their self-rated anxiety. Conversely, if panic patients are led to believe through false physiological feedback that their heart is racing,

they become more anxious and show more psychophysiological arousal than control subjects.

Anxiety Disorders

Disorders of anxiety are common in the population, afflicting about one-quarter of adult individuals. One group of the anxiety disorders, phobias, is more closely related to fear than to anxiety (see **Fear**).

Panic Disorder

Recurrent panic attacks have a key position in the diagnosis of anxiety disorders, and they are what the diagnostician first looks for. Panic disorder is a condition characterized by recurring and crippling panic, typically associated with distressing worries about future panic attacks. As a result, the normal social life of the individual is compromised. Typically panic disorder occurs in the context of agoraphobia (see **Fear**), in which the person starts to avoid situations in which he or she will be left helpless and vulnerable in the event of a panic attack.

Generalized Anxiety Disorder

If the anxiety is manifested as uncontrollable and excessive anxiety and worry about events or activities (e.g., finances or school performance), generalized anxiety disorder may be diagnosed. Patients with this disorder worry more or less constantly and their worries are less controllable and more self-driven and ruminative than normal worries. Typically the worry is combined with somatic symptoms or other anxiety symptoms such as restlessness, fatigue, irritability, muscle tension, and concentration and sleep difficulties.

Posttraumatic Stress Disorder

Posttraumatic stress disorder has its origin in intense traumas, that is, in situations whose danger and aver-siveness are outside the range of usual human experience, such as in natural disasters, war, violent crime, or severe accidents. In this disorder, the traumatic event is persistently reexperienced (e.g., in the form of flashbacks), stimuli or events associated with the trauma elicit intense anxiety and are avoided, and the person feels generally numbed with regard to emotions. Common anxiety symptoms experienced by people suffering from posttraumatic stress disorder include difficulties sleeping and concentrating, irritability or outbursts of anger, hypervigilance, and exaggerated startle.

Obsessive-Compulsive Disorder

Obsessions are persistent thoughts, impulses, or images; compulsions are repetitive behaviors such as

hand washing and checking or mental activities such as praying or silently repeating words. The essential features of obsessive-compulsive disorder are recurrent, uncontrollable obsessions or compulsions that are performed with the ostensible purpose of reducing anxiety. The obsessions and compulsions are associated with marked distress, and they interfere with the normal routines, even though the victim, at least at some point, has recognized that they are excessive or unreasonable.

The Genetics of Anxiety Disorders

Statistical modeling based on comorbidity data in twins suggests that there are two independent genetic factors involved in the anxiety disorders. The anxious-misery factor is related to panic disorder, agoraphobia, posttraumatic stress disorder, generalized anxiety disorder, and depression, whereas the fear factor predisposes for specific phobias. The model suggests only a limited role for shared (within-family) environmental factors, but it suggests a substantial role for unique environmental factors (i.e., factors affecting one family member but not others) for all disorders. This genetic difference between the disorders may explain why those related to the anxious-misery factor all respond to the same medication (selective serotonin-uptake inhibitors), which is ineffective in specific phobias. It may also be reflected in psychophysiological data showing a distinct fear response in specific phobics who are exposed to their fear object or situation but normal psychophysiological resting levels. The disorders relating to the anxious-misery factors, on the other hand, show elevated resting levels and little psychophysiological mobilization during stress.

The Neuroanatomy of Anxiety

From the perspective taken here, in which anxiety is seen as a sibling of fear, it is a natural assumption that their neural substrate shows considerable overlap. A neural network for fear activation is centered on the amygdala (see **Fear**). It is an attractive notion that this circuit is sensitized and hyperreactive in anxiety states. Furthermore, there are data suggesting that a brain area closely related to the amygdala, the bed nucleus of the stria terminalis, has a downstream connection, similar to that of the amygdala, to hypothalamic and brain-stem nuclei controlling different components of fear and anxiety, such as autonomic responses and fear behavior. However,

whereas the amygdala mediates short-term fear responses to specific stimuli, the bed nucleus of the stria terminalis is more dependent on long-term contextual stimuli for its activation. This effect appears to be mediated by an abundance of receptors for the corticotropin releasing hormone (CRH; see **Corticotropin Releasing Factor (CRF)**) in the bed nucleus of the stria terminalis, which mediates the activation of the hypothalamic-pituitary-adrenal axis (see **Hypothalamic-Pituitary-Adrenal**). Thus, this opens the possibility of separable but closely related anatomical loci for fear and anxiety. Fear may be primarily dependent on the effects of specific stimuli on the amygdala, and anxiety may be primarily dependent on the effect of contextual stimuli and CRH in the bed nucleus of the stria terminalis.

See Also the Following Articles

Corticotropin Releasing Factor (CRF); Fear; Hypothalamic-Pituitary-Adrenal.

Further Reading

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th edn.). Washington, DC: American Psychiatric Association.
- Barlow, D. H. (2002). *Anxiety and its disorders: the nature and treatment of anxiety and panic* (2nd edn.). New York: Guilford Press.
- Hettema, J. M., Prescott, C. A., Myers, J. M., et al. (2005). The structure of genetic and environmental risk factors for anxiety disorders in men and women. *Archives of General Psychiatry* **62**, 182–189.
- LeDoux, J. E. (1996). *The emotional brain*. New York: Simon & Schuster.
- Mineka, S. and Zinbarg, R. (1996). Conditioning and ethological models of anxiety disorder: stress-in-dynamic-context anxiety models. In: Hope, D. A. (ed.) *Nebraska symposium of motivation. Vol. 43: Perspectives on anxiety, panic, and fear*, pp. 135–210. Lincoln, NE: University of Nebraska Press.
- Mogg, K. and Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. *Behaviour Research and Therapy* **36**, 809–848.
- Öhman, A. (2000). Fear and anxiety: evolutionary, cognitive, and clinical perspectives. In: Lewis, M. & Haviland, J. M. (eds.) *Handbook of emotions* (2nd edn., pp. 573–593). New York: Guilford.
- Walker, D. L., Toufexis, D. J. and Davis, M. (2003). Role of the bed nucleus of the stria terminalis versus the amygdala in fear, stress, and anxiety. *European Journal of Pharmacology* **463**, 199–216.

Anxiety as a Central Aspect of Comorbidity *See: Anxiety.*

Anxiolytics

M Lader

Institute of Psychiatry, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
M Lader, volume 1, pp 231–235, © 2000, Elsevier Inc.

Drugs Used to Treat Anxiety

Selective Serotonin Reuptake Inhibitors

Other Antidepressants

Benzodiazepines

Other Antianxiety Drugs

Glossary

<i>Antidepressant</i>	A medicinal substance used primarily to treat depression, but may have other indications.
<i>Anxiolytic</i>	A medicinal substance that allays anxiety.
<i>Benzo-diazepines</i>	A group of substances used as anxiolytics, hypnotics, muscle relaxants, and anesthetic inducers.
<i>GABA, serotonin, nor-epinephrine</i>	Neurotransmitters involved in anxiety mechanisms.

Drugs Used to Treat Anxiety

The use of anxiety-allaying drugs goes back thousands of years to the discovery that grape juice and grain mash were fermentable and that alcohol possessed psychotropic properties. The nineteenth century saw the introduction of bromides, chloral- and paraldehyde, and the barbiturates. These had the disadvantages of drowsiness, tolerance, danger in overdose, and the likelihood of dependence. The first marketed benzodiazepine was chlordiazepoxide (Librium), but well over 1000 benzodiazepines and related compounds have been synthesized, including diazepam, the most widely used of all.

Since the mid-1990s there has been a marked change in the choice of drugs for treating anxiety

disorders. Official guidelines and practice recommendations in secondary care now recognize the advantages of the selective serotonin reuptake inhibitors (SSRIs) over the benzodiazepines in the pharmacological management of patients with one of the anxiety disorders. Primary care practitioners are now following suit. Thus, the long chain of sedative use – bromides, chloral hydrate, barbiturates, meprobamate, and the benzodiazepines – is being broken, and a new class of drugs has been substituted. This reflects both the superior and the long-term efficacy of the SSRIs and the clear drawbacks of the sedatives, such as a plethora of unwanted side effects, waning of efficacy over time, and dependence and withdrawal problems.

This shift of emphasis from sedatives to drugs licensed primarily as antidepressants leads to definitional problems. Is an SSRI an anxiolytic? An anxiolytic is a medicinal substance that allays anxiety and tension. Examples include primary anxiolytics such as the benzodiazepines and buspirone, whose primary use is to treat anxiety, and secondary anxiolytics such as antihistamines and beta blockers, whose main use lies elsewhere. SSRIs can be fitted into the latter category even though they are now the drugs of choice for treating anxiety. Furthermore, in due course, more prescriptions for SSRIs may be issued for anxiety than for depression.

Selective Serotonin Reuptake Inhibitors

Background

Tricyclic compounds usually inhibit both serotonin reuptake and norepinephrine reuptake, thus being nonselective. Efforts were made to develop compounds that showed selectivity in two ways. First, they blocked the reuptake of serotonin only, and second, they did not bind to other receptors such as the cholinergic and histaminergic receptors, thereby avoiding a range of unwanted effects. Several compounds have been developed. The first was zimeldine, which was withdrawn from the market because of neurotoxicity. The next was fluvoxamine, which has indications on a wide spectrum of disorders. The

most well known of the SSRIs is fluoxetine, a long-acting compound with a half-life of over 40 h. Its demethylated metabolite, norfluoxetine, is also an SSRI and has an even longer half-life, a week or more. This means that it maintains some efficacy for a long period of time and any withdrawal reactions are delayed. Fluoxetine has gained a wide usage, particularly in the United States. Folklore concerning its effects on personality traits and personality of people who are not clinically depressed has also accrued. As a counterpoint, concern has been expressed about its widespread use as a lifestyle compound. In addition, there have been controversies about possible increase in suicidal ideation and hostile and aggressive behavior in those taking fluoxetine. These debates have meandered on for many years without any clear evidence of resolution. Fluoxetine's widespread use in individuals without distinct clinical depression may reflect subtle mood-enhancing effects in mild dysthymic patients, those with chronic low-grade depression.

Sertraline is also a widely used SSRI with a half-life of 26 h, so it is suitable for once daily administration, as are most of the SSRIs. The drawback with sertraline is that the dosage is not a fixed one, having to be attained by titration. Another controversial SSRI is paroxetine. Efficacious and well tolerated, it has been used particularly in anxiety disorders. However, its lack of efficacy in children and adolescents and its reputation for withdrawal reactions has caused concern.

Citalopram has been a successful compound in many countries. It comprises a racemic mixture, the active compound being escitalopram: this enantiomer has enhanced efficacy in both depression and anxiety. It may also have a rapid onset of action. It has therefore been suggested that the R enantiomer interferes with some of the actions of the active S compound.

Indications

Many SSRIs are licensed for one of the anxiety disorders. However, it is important to note that the initial effect of an SSRI in the treatment of anxiety may be to transiently increase the anxiety levels and frequency of panic attacks. The antianxiety effect becomes apparent after a week of treatment. Thus, the patient must be warned about this possibility. It is often the practice with anxious patients to give an initial dose that is half that of the usual initial dose for depression.

Side Effects

SSRIs can have mild sedative or stimulant effects and indeed may have these opposing effects in different patients. Thus, overall in clinical trials, about 10%

of patients complained of somnolence, and an equal number of insomnia. However, some differences can be discerned among the various members of the class: fluoxetine causes the most stimulation, paroxetine and fluvoxamine, the most sedation. Sertraline, citalopram, and escitalopram appear to be relatively neutral. Extensive laboratory assessment of the psychomotor and cognitive effects of SSRIs shows few effects, nor is memory affected. In the elderly, SSRIs can induce a persistent tremor. The SSRIs are much safer than the tricyclics. Thus, death after overdose is much less frequent, reflecting lack of effects on cardiac conduction.

The most common adverse effects of the SSRIs involve the gastrointestinal system and typically affect at least 10% of patients. Systems include nausea, dry mouth, vomiting, abdominal pain, and dyspepsia and constipation or, conversely, diarrhea. Anorexia may occur with fluoxetine, and this has been exploited to treat bulimia. The gastrointestinal side effects of the SSRIs are strongly dose related and can usually be controlled by careful adjustment of the dose.

Weight gain is not common with the SSRIs. However, sexual dysfunction including analgesia in females is well documented.

Clinical Uses

During the last few years, many large-scale, well-controlled multicenter trials comparing SSRIs to placebo in all the anxiety disorders have been published. It is now clear that the SSRIs are quite effective and tolerated in each of these disorders. The efficacy extends across the various compounds, and the SSRIs are now recommended by most expert consensus groups as the treatment of choice in each of the various anxiety disorders.

This consensus is in itself fairly remarkable and represents the clear evolution of the treatment of choice from clomipramine in obsessive-compulsive disorder and the benzodiazepines in generalized anxiety disorder (GAD) and panic disorders. The change to recommending SSRIs in GAD is a most useful change in clinical practice. The SSRIs are better tolerated (although they may have withdrawal symptoms), and they avoid the psychomotor and cognitive impairment, interaction with alcohol, paradoxical disinhibition, and dependence and withdrawal problems that attend the use of the benzodiazepines. The other shift in emphasis has been the acknowledgment that the anxiety disorders are long-term maladies and that the evidence for long-term benzodiazepine effectiveness is nugatory. By contrast, several trials showed that the SSRIs maintain their efficacy over the long term. Indeed, the recommendation for at

least 1–2 years of treatment is often incorporated in official and semi-official guidelines, which itself is a major change in emphasis. It is important that the use of SSRIs in the treatment of anxiety disorders is considered in the context of their wider management, in particular the use of cognitive-behavioral therapy to effect a more lasting improvement. However, further trials to ascertain the best use of these multiple treatments in succession and combination are needed.

Other Antidepressants

Tricyclic antidepressants have been used to treat various anxiety disorders. The most popular has been clomipramine, especially in obsessive-compulsive disorder. Imipramine was also used in panic disorder. The use of the tricyclic compounds has declined greatly because of the superior efficacy and greater tolerability of the SSRIs.

The monoamine oxidase inhibitors (MAOIs) have also shown efficacy in the treatment of various anxiety disorders. However, formal assessments are few. The use of these compounds was never widespread and has also declined with the advent of the SSRIs.

Benzodiazepines

The popularity of the benzodiazepines was related to their relative safety in overdose as compared to their predecessors. After 30 years, misgivings concerning the benzodiazepines were expressed. Their widespread use and the possibility of dependence even at normal therapeutic doses led to recommendations for the restriction of prescribing. Despite this general disapproval, some prescribers still view these drugs quite favorably.

Background

For the prescriber, two aspects of the pharmacokinetics of benzodiazepines are germane: the speed of onset of action and the duration of action. Thus, given orally, diazepam enters the brain rapidly and exerts a prompt effect in anxiety and panic states. The benzodiazepines vary greatly in their duration of action. They potentiate the inhibitory neurotransmitter, γ -aminobutyric acid.

Benzodiazepines do not act directly on GABA receptors, but on their own receptors.

Clinical Pharmacology

In normal individuals, the depressant effects of single higher doses of the benzodiazepines can be detected readily. At lower doses, however, impairment of psychological functioning is quite difficult to quantify,

and subjective effects are usually absent. In the clinical context with anxious patients and with repeated doses, impairment of functioning is more difficult to demonstrate.

Benzodiazepines have marked and quite specific effects on memory functions and interfere with episodic memory, i.e., the system concerned with remembering personal experiences. This effect is independent of any sedation or attentional impairment.

Clinical Uses

The main use of benzodiazepines is in the management of anxiety and stress-related conditions, especially by family physicians. Few differences have been found among the benzodiazepines, although their superiority with respect to placebo has been established, but only in the short term (up to 6–12 weeks).

Benzodiazepines are also effective treatments for panic disorder, preventing the episodes rather than aborting them. Alprazolam can suppress panics, but relapse and even rebound may occur when the benzodiazepine is discontinued, even if tapered off.

Unwanted Effects

The most common unwanted side effects are tiredness, drowsiness, and torpor, so-called oversedation. These effects are dose and time related, being most marked within the first 2 h after large doses. Furthermore, drowsiness is most common during the first week of treatment. Both psychomotor skills, such as driving, and intellectual and cognitive skills are affected. As with most depressant drugs, potentiation of the effects of alcohol can occur. Paradoxical behavioral responses may occur in patients taking benzodiazepines. Such events include increased aggression and hostility, acute rage reactions, uncharacteristic criminal behavior such as shoplifting, and uncontrollable weeping. Other unwanted effects include excessive weight gain, skin rash, impairment of sexual function, menstrual irregularities, and, rarely, agranulocytosis.

In pregnant women, benzodiazepines pass readily into the fetus and have been suspected of producing respiratory depression in the neonate. Finally, benzodiazepines are present in the mother's milk and can oversedate the baby. Although overdose with benzodiazepines is extremely common, deaths due to these drugs alone are uncommon.

Tolerance and Dependence

Clinically, tolerance has not been studied in any detail. However, escalation of dosage with benzodiazepines is comparatively rare. Tolerance can supervene to clinical effects in patients maintained on moderate doses of benzodiazepines. Although many chronically

anxious patients aver that they were helped by benzodiazepines, this may reflect dependence and the need to maintain their drug intake. The percentage of patients taking benzodiazepines chronically who experience withdrawal symptoms on discontinuation of medication ranges between 27 and 45%, depending on the criteria used. The mildest symptoms on withdrawal are anxiety, tension, apprehension, dizziness, insomnia, and anorexia. More severe physical dependence is manifested by withdrawal symptoms of nausea and vomiting, tremor, muscle weakness, and postural hypotension. Occasionally, hyperthermia, muscle twitches, convulsions, and confusional psychoses may occur.

Other Antianxiety Drugs

5-HT_{1A} Partial Agonists

The class of 5-HT_{1A} partial agonist compounds has a complex pharmacology. The first one, buspirone, has been available in many countries for several years. It suppresses activity in presynaptic serotonergic neurons, diminishing serotonin activity and leading to downregulation of 5-HT₂ and perhaps other 5-HT receptors. Buspirone is much less sedative than benzodiazepines, with little propensity to induce psychomotor or cognitive impairment. In formal clinical trials, buspirone proved equieffective and equipotent to diazepam. However, patients on buspirone improved more slowly. Buspirone has fewer and less severe side effects than diazepam. Headache, dizziness, and nausea are the most common complaints; they may be severe and may necessitate lowering the dosage or discontinuing the drug. Discontinuation is not accompanied by either rebound or withdrawal.

Antipsychotic Drugs

Phenothiazines, such as chlorpromazine and trifluoperazine, and a range of other antipsychotic drugs are used by some to treat anxiety. The dosage recommended is quite low, typically less than half the initial antipsychotic dose. Sometimes, even at this dosage, the antipsychotic drug is not well tolerated by the

anxious patient because unwanted autonomic effects such as dry mouth and dizziness too closely resemble their own symptoms. Even more disconcerting but uncommon are extrapyramidal symptoms such as restlessness (mild akathisia) and Parkinsonism. It is not clear if tardive dyskinesia is a real risk, but this possibility should dissuade the practitioner from the indiscriminate use of these drugs as anxiolytics.

Antihistamines

The older antihistamines penetrate to the brain and have sedative effects. Some, such as hydroxyzine, have found some favor as anxiolytics. They are not associated with a withdrawal syndrome.

β-Adrenoceptor Antagonists

Anxiety disorders are associated with a wide range of physical symptoms. In particular, palpitations, tremor, and gastrointestinal symptoms may be assuaged by low doses of a beta blocker.

See Also the Following Articles

Antidepressant Actions on Glucocorticoid Receptors; Anxiety; Depression and Manic-Depressive Illness.

Further Reading

- Bandelow, B., Zohar, J., Hollander, E., et al. (2002). World Federation of Societies of Biological Psychiatry [WFSBP] guidelines for the pharmacological treatment of anxiety, obsessive-compulsive and posttraumatic stress disorders. *World Journal of Biological Psychiatry* 3, 171–199.
- Bourin, M. and Lambert, O. (2002). Pharmacotherapy of anxious disorders. *Human Psychopharmacology* 17, 383–400.
- Grudzinski, A. N. (2001). Considerations in the treatment of anxiety disorders: a pharmacoeconomic review. *Expert Opinion in Pharmacotherapy* 2, 1557–1569.
- National Institute for Clinical Excellence (2004). *Anxiety Clinical Guideline* 22. London: National Institute for Clinical Excellence.
- Sramek, J. J., Zarotsky, V. and Cutler, N. R. (2002). Generalised anxiety disorder: treatment options. *Drugs* 62, 1635–1648.

Apoptosis

J-L Turner and J A Cidlowski

National Institute of Environmental Health Sciences,
Research Triangle Park, NC, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
C L Mann and J A Cidlowski, volume 1, pp 236–239, © 2000,
Elsevier Inc.

Apoptotic Morphology
Glucocorticoid Receptor Signaling
Downstream Signals
Inhibitors of Apoptosis
Conclusion

Glossary

<i>Apoptosis</i>	A type of cell death that involves a programmed sequence of events leading to the elimination of unwanted cells without inducing an inflammatory response.
<i>Caspases</i>	A family of cysteine proteases that cleave substrates amplifying the apoptotic cascade.
<i>Glucocorticoid receptor</i>	A member of the nuclear hormone receptor superfamily, activated upon binding to glucocorticoid hormones, ultimately leading to activation or suppression of genes.
<i>Glucocorticoids</i>	Steroid hormones, produced by the adrenal glands, which are anti-inflammatory, and play a role in apoptosis and metabolism.
<i>Lymphocyte</i>	White blood cell involved in immune and inflammatory responses.
<i>Thymocyte</i>	A lymphocyte produced in the thymus that is a precursor of a T cell.

Apoptosis, also known as cell suicide or programmed cell death, is essential for the development and function of multicellular organisms. It is an energy requiring, highly regulated and ordered process, resulting in the clearance of individual cells in a population and ultimately serving to maintain normal cell homeostasis. Stress can disrupt cell homeostasis and promote or interfere with normal apoptosis, with steroid hormones playing a major role in what is termed the stress response. Persistent stress can lead to sustained levels of steroid hormones, in particular glucocorticoids, which can induce immunosuppression by stimulation of lymphocyte apoptosis. In this way, the immune system is probably the most adversely affected system in the body in response to stress, leading to

susceptibility to infection and disease. These issues prompt the need for a deeper understanding of the relationship between stress on the immune system and steroid hormone-induced apoptosis.

Apoptotic Morphology

There are several distinct morphological features that occur in cells undergoing apoptosis, and are evident in most cell types including lymphocytes. They include cell shrinkage which occurs due to the movement of water and ions out of the cell, membrane blebbing, flipping of phosphatidyl serine to the external side of the plasma membrane, chromatin condensation, and nuclear fragmentation. All of these events occur without loss of plasma membrane integrity, and therefore offset the inflammatory response that occurs in other types of cell death such as necrosis. Membrane blebbing, resulting in apoptotic body formation, and the externalization of phosphatidyl serines both allow for recognition of the apoptotic cell by macrophages. In this way, apoptotic cells are cleared from the body.

Glucocorticoid Receptor Signaling

The stress response ultimately leads to the release of stress hormones such as glucocorticoids, which are anti-inflammatory and immunosuppressive ([Figure 1](#)). Physiologically, the body responds to stress by stimulation of the hypothalamus in the brain leading to the release of corticotropin-releasing hormone (CRH). The anterior pituitary gland is then stimulated by CRH to release adrenocorticotrophic hormone (ACTH, corticotropin), which in turn stimulates release of glucocorticoids from the adrenal glands. The primary glucocorticoid released in mice is corticosterone, while the form in humans is cortisol.

Glucocorticoids, as well as synthetic derivatives such as dexamethasone, elicit their effect through the glucocorticoid receptor (GR), a member of the nuclear hormone receptor superfamily. GR resides in the cytoplasm and, upon binding to glucocorticoid, is shuttled to the nucleus where it induces or represses gene transcription through interactions with DNA and various transcription factors and coactivators ([Figure 2](#)). Cells deficient in GR or with a mutation in the DNA binding domain or nuclear localization sequence will not undergo apoptosis in response to glucocorticoids. This indicates a requirement for the interaction of the GR–glucocorticoid complex with DNA for the activation of apoptosis. Most lymphoid

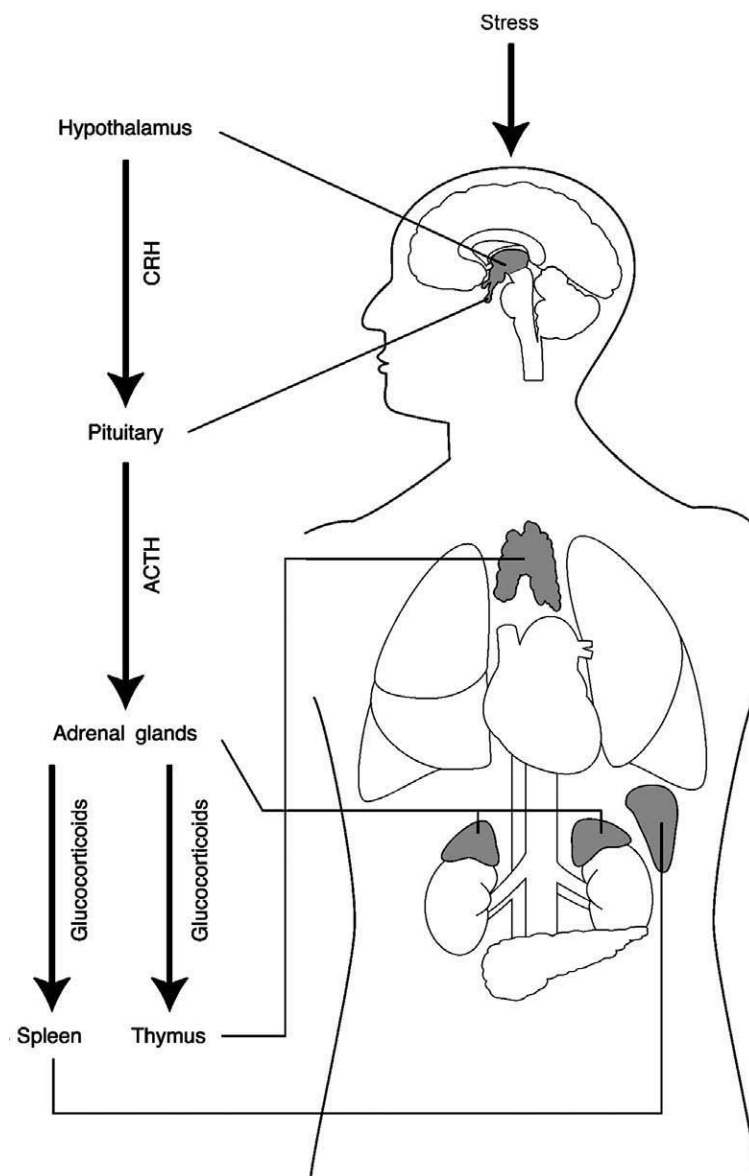


Figure 1 The stress response can lead to immunosuppression. The physiological response to stress begins with the stimulation of the hypothalamus to release corticotropin-releasing hormone (CRH) which leads to the subsequent release of adrenocorticotrophic hormone (ACTH, corticotropin) from the anterior pituitary gland. ACTH signals the adrenal glands to produce and secrete glucocorticoids which can induce apoptosis in thymocytes and lymphocytes in the thymus and spleen.

cells are sensitive to glucocorticoid-induced apoptosis, but there are many cell types which have functional GR but do not undergo apoptosis when exposed to glucocorticoids.

Glucocorticoids can alter the expression of both proapoptotic and prosurvival genes. Although glucocorticoid-induced apoptosis is dependent on *de novo* RNA and protein synthesis, the initiating event(s) that connects to the downstream activation of apoptosis remains unclear. Several candidate proapoptotic genes are reported to be induced by glucocorticoids. Of particular interest, are several of the BH3-only proteins belonging to the proapoptotic Bcl-2 family.

Most notable are Bim (Bcl-2-interacting mediator of cell death) and PUMA (p53 upregulated modulator of apoptosis), which are upregulated by glucocorticoids in a p53-independent manner. These and other BH3-only proteins enhance the activity of the proapoptotic proteins Bax and Bak known to permeabilize the outer mitochondrial membrane leading to the release of cytochrome-c culminating in the activation of effector caspases and finally cell death. Studies show that thymocytes and mature lymphocytes that do not express Bim or PUMA are refractory to glucocorticoid-induced apoptosis. The importance of Bim and PUMA for glucocorticoid-induced lymphocyte

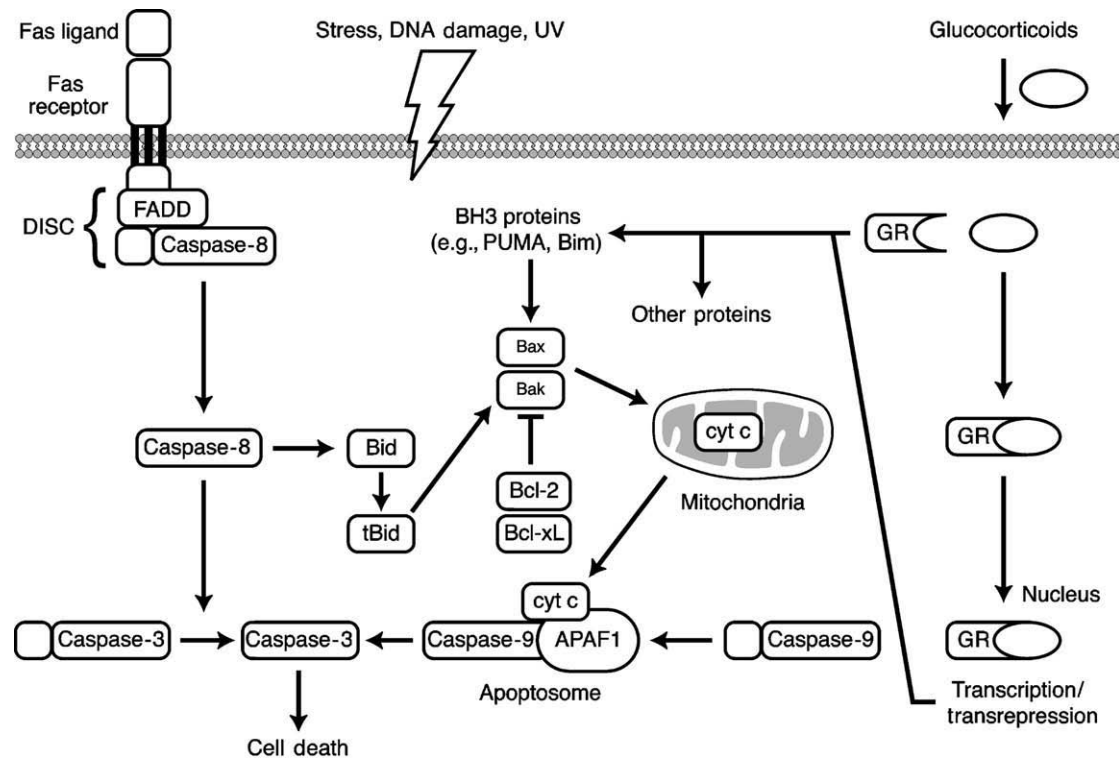


Figure 2 Apoptotic pathways and a model of glucocorticoid-induced apoptosis in lymphocytes. The intrinsic pathway of apoptosis is activated in response to various stimuli including stress, DNA damage, and ultraviolet radiation (UV). The signal activates the proapoptotic Bcl-2 proteins Bax and Bak which complex to form pores in the outer mitochondrial membrane resulting in the release of cytochrome-c (cyt c). This leads to the formation of the apoptosome, which is comprised of cytochrome-c, Apaf-1 and caspase-9. The apoptosome cleaves caspase-3 leading to cell death. Antiapoptotic proteins Bcl-2 and Bcl-xL block the action of Bax and Bak. The extrinsic pathway of apoptosis involves activation of death receptors (e.g., Fas receptor) that become active upon binding ligand (e.g., Fas), signaling the formation of the DISC (death-inducing signaling complex). Cleavage of caspase-8 from the DISC amplifies the apoptotic signal through cleavage of caspase-3, leading to cell death. Cleavage of Bid to tBid (truncated Bid) by caspase-8, connects the extrinsic pathway to the intrinsic apoptotic pathway. Glucocorticoids, released in the body during stress, enter the cell and are bound by the glucocorticoid receptor (GR). This complex localizes to the nucleus where it induces and/or represses a variety of genes that ultimately lead to apoptosis of lymphocytes. It is thought that the BH3-only proteins Bim and PUMA are important in glucocorticoid-induced apoptosis, enhancing the activity of Bax and Bak.

cell death was further demonstrated *in vivo* using PUMA-deficient and *Bim*^{-/-} transgenic mice. Lymphocytes from these PUMA and Bim compromised animals were refractory to dexamethasone-induced apoptosis.

Another proapoptotic gene induced by glucocorticoids is GILZ (glucocorticoid-induced leucine zipper), and is strongly upregulated in the thymus in response to glucocorticoids. GILZ reduces the expression of the antiapoptotic protein Bcl-xL followed by an increase in the activity of caspase-8 and caspase-3. This also suggests that GILZ is an initiating signal of glucocorticoid-induced apoptosis in thymocytes. Other proapoptotic genes upregulated by glucocorticoids that may be important in glucocorticoid-induced apoptosis include the G protein-coupled receptor *tdag8* (T cell death-associated gene 8), the stress response gene *dig2*, and *txnip* (thioredoxin-interacting protein). Although these genes have been identified as being upregulated by glucocorticoids, it

remains unclear as to the exact role they play in the initiation and execution of apoptosis.

Glucocorticoids can also downregulate prosurvival genes leading to enhancement of apoptosis in lymphocytes. They include genes such as *c-myc*, *c-myb*, and *c-Ki-ras*, all of which are involved in cell proliferation and survival. In addition, glucocorticoids can repress NF-κB prosurvival signaling through induction of the inhibitor protein I-κB. It is important to realize that the induction of apoptosis by glucocorticoids involves both upregulation of prodeath genes, working in concert together, and downregulation of survival genes to induce cell death.

Downstream Signals

Apoptosis can be initiated through a variety of stimuli and can proceed through the death receptor (extrinsic) or the mitochondrial (intrinsic) pathway (Figure 2).

However, both pathways have points where they intersect, enhancing the end result that is cell death. The intrinsic pathway is activated by various stimuli such as DNA damage or cellular stress, and involves interplay of both pro- and antiapoptotic members of the Bcl-2 family of proteins. The proapoptotic Bcl-2 proteins Bax and Bak are thought to complex to form pores in the outer mitochondrial membrane leading to the release of cytochrome-c which in concert with Apaf-1 and caspase-9 form the apoptosome and activate the executioner caspase-3. The antiapoptotic proteins Bcl-2 and Bcl-xL normally interfere with the action of Bax and Bak, but are disabled upon activation of apoptosis (possibly through BH3-only Bcl-2 family members). In the extrinsic pathway, death receptors are engaged upon ligand binding (e.g., Fas receptor and Fas ligand) leading to recruitment of FADD (Fas-associated death domain) and the initiator caspase-8 forming the death-inducing signaling complex (DISC), followed by cleavage of the executioner caspase-3. The extrinsic pathway can connect to the intrinsic pathway via the cleavage of the BH3-only Bcl-2 family member Bid to tBid (truncated Bid) by caspase-8. As mentioned in the previous section, the other BH3-only Bcl-2 family members Bim and PUMA enhance the activity of Bax and Bak through the action of glucocorticoids. The later stages of apoptosis include the breakdown of lamins, cytoskeleton proteins, and nuclease activation leading to internucleosomal DNA fragmentation.

It is not clear how glucocorticoids initially signal apoptosis, but it is known that there is a requirement for the activation of Bax and Bak. Bim and PUMA enhance the activity of Bax and Bak in response to glucocorticoids. In studies using Bim-deficient mice, glucocorticoid apoptosis was partially impaired compared to wild-type animals. Thymocytes derived from PUMA knock-out mice also demonstrated partial resistance to glucocorticoid-induced apoptosis. These studies emphasize the importance of Bim and PUMA as candidate proapoptotic proteins in glucocorticoid-induced apoptosis.

Inhibitors of Apoptosis

Apoptosis can be inhibited through a variety of mechanisms, mostly through the action of the antiapoptotic Bcl-2 family members and the inhibitors of apoptosis (IAP) proteins, as well as through growth factor upregulation. Bcl-2 and Bcl-xL block the interaction of the proapoptotic proteins Bax and Bak preventing the formation of pores in the mitochondrial outer membrane, and consequently the release of cytochrome-c and caspase-3 cleavage. Studies show that

the phosphorylation status of the antiapoptotic protein Bcl-2 can dictate its ability to block glucocorticoid apoptosis. IAPs can block apoptosis by directly interacting with caspases and preventing their activation. In thymocytes, glucocorticoid-induced apoptosis can be blocked by overexpression of the IAP known as X-linked inhibitor of apoptosis (XIAP). Many cytokines such as interleukin (IL)-2, -4, and -7, and interferon (IFN)- α , as well as the activation of the T cell receptor can inhibit apoptosis induced by glucocorticoids in thymocytes. In the case of IL-7, the antiapoptotic effects are attributed to its ability to enhance the expression of Bcl-2 proteins in thymocytes and mature T cells.

Conclusion

The immune system is highly prone to the effects of stress. This is primarily due to the increased release of stress hormones, specifically glucocorticoids, which initiate a signaling cascade in immune cells that can lead to cell death by apoptosis. This occurs through the GR, involving both transactivation of proapoptotic genes and transrepression of prosurvival genes. It is still unclear as to how glucocorticoids induce apoptosis, but several candidate genes and mechanisms are emerging. Further study in the area of glucocorticoid signaling is needed to fully understand the impact of stress on the immune system.

Further Reading

- Cidlowski, J. A. and Huang, Se-Te, J. (2002). Phosphorylation status modulates Bcl-2 function during glucocorticoid-induced apoptosis in T lymphocytes. *FASEB Journal* 16, 825–832.
- Conte, D., Liston, P., Wong, J. W., et al. (2001). Thymocyte-targeted overexpression of *xiap* transgene disrupts T lymphoid apoptosis and maturation. *Proceedings of the National Academy of Sciences of the United States of America* 98, 5049–5054.
- Distelhorst, C. W. (2002). Recent insights into the mechanism of glucocorticosteroid-induced apoptosis. *Cell Death and Differentiation* 9, 6–19.
- Erlacher, M., Michalak, E. M., Kelly, P. N., et al. (2005). BH3-only proteins Puma and Bim are rate-limiting for γ -radiation- and glucocorticoid-induced apoptosis of lymphoid cells *in vivo*. *Blood* 106, 4131–4138.
- Herold, M. J., McPherson, K. G. and Reichardt, H. M. (2006). Glucocorticoids in T cell apoptosis and function. *Cellular and Molecular Life Sciences* 63, 60–72.
- Labat-Asselin, M. L., David, M., Vidamment-Biola, A., et al. (2004). GILZ, a new target for the transcription factor FoxO3, protects T lymphocytes from interleukin-2 withdrawal-induced apoptosis. *Blood* 104, 215–223.

- Malone, M. H., Wang, Z. and Distelhorst, C. W. (2004). The glucocorticoid-induced gene *tdag8* encodes a pro-apoptotic G protein-coupled receptor whose activation promotes glucocorticoid-induced apoptosis. *Journal of Biological Chemistry* **279**, 52850–52859.
- Mann, L. M., Hughes, F. M. Jr. and Cidlowski, J. A. (2000). Delineation of the signaling pathways involved in glucocorticoid-induced and spontaneous apoptosis of rat thymocytes. *Endocrinology* **141**, 528–538.
- Schmidt, S., Rainer, J., Ploner, C., et al. (2004). Glucocorticoid-induced apoptosis and glucocorticoid resistance: molecular mechanisms and clinical relevance. *Cell Death Differentiation* **11**, S45–S55.
- Wang, D., Muller, N., McPherson, K. G., et al. (2006). Glucocorticoids engage different signal transduction pathways to induce apoptosis in thymocytes and mature T cells. *Journal of Immunology* **176**, 1695–1702.
- Wang, Z., Rong, Y. P., Malone, M. H., et al. (2005). Thioredoxin-interacting protein (*txnip*) is a glucocorticoid-regulated primary response gene involved in mediating glucocorticoid-induced apoptosis. *Oncogene* **21**, 1–11.

Arginine Vasopressin (AVP) See: Vasopressin.

Arterial Baroreflex

G Parati

IRCCS Istituto Auxologico Italiano, and University of Milano-Bicocca, Milan, Italy

P Castiglioni and M Di Rienzo

IRCCS Fondazione Don C. Gnocchi, Milan, Italy

G Mancia

University of Milano-Bicocca and Ospedale S. Gerardo, Monza, Italy

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G Parati, M Di Rienzo, and G Mancia, volume 1, pp 240–250, © 2000, Elsevier Inc.

Introduction

Anatomy and Physiology

Methods for Assessing Baroreflex Function

Clinical Value of Baroreflex Sensitivity

Conclusion

Glossary

α Coefficient

Frequency-domain estimate of spontaneous baroreflex sensitivity based on the calculation of the square ratio between the spectral powers of heart interval and systolic blood pressure at 0.1 or 0.3 Hz. Models of the cardiovascular system in which the relations between blood pressure and heart interval variability are quantified, in a closed-loop fashion, by a mathematical equation including the

Autoregressive moving average (ARMA) techniques

Arterial baroreceptors

Arterial baroreflex

Black-box model

Fast Fourier transform (FFT)

Frequency-domain methods

effects of one parameter on the other one (moving average component) and the effects of previous values of each considered parameter on its present value (autoregressive component).

Mechanoreceptors located within the arterial wall (mainly in the carotid arteries, renal arteries, and aortic arch) sensitive to changes in arterial wall transmural pressure.

Reflex loop including arterial baroreceptors (peripheral input sensors), afferent neural pathways, central neural integration, efferent neural pathways, and effector organs (e.g., sinus node at the heart level and arteriolar smooth muscle cells at peripheral vascular level).

Mathematical model of the input–output relationship of a physiological system obtained without using any a priori information on the system.

Computationally efficient algorithm for calculating the Fourier transform of a signal, i.e., a representation of amplitude and phase of the individual oscillations composing the signal dynamics as a function of the frequency of the oscillation.

Methods for spontaneous baroreflex modulation and blood pressure/heart rate variability analysis based on the quantification of blood pressure and heart rate changes as a function of their frequency.

<i>Sequence technique</i>	Time-domain method for spontaneous baroreflex analysis, based on the identification of spontaneously occurring beat-by-beat hypertension/bradycardia or hypotension/tachycardia sequences over three or more consecutive beats.
<i>Spectral analysis</i>	Decomposition of a signal (in our context, blood pressure and heart rate time series) into its different oscillatory components as a function of the frequency.
<i>Spontaneous baroreflex modulation of the heart</i>	Reflex changes in heart rate triggered by spontaneous (i.e., not induced by laboratory stimuli) fluctuations in blood pressure.
<i>Time-domain methods</i>	Methods for spontaneous baroreflex modulation and blood pressure/heart rate variability analysis based on the quantification of blood pressure and heart rate changes as a function of time (i.e., in the time domain).

Introduction

The arterial baroreflex represents a mechanism of fundamental importance for blood pressure homeostasis, and an impairment of its cardiac and vascular influences is believed to play a role in several cardiovascular diseases. Baroreflex abnormalities have been described early in the course of diabetes mellitus and hypertension, contributing in the former case to the occurrence of hypotensive episodes and in the latter to the maintenance of the blood pressure elevation. In addition, baroreflex dysfunction is characteristic after myocardial infarction and carries a worse prognosis. The clinical relevance of a baroreflex dysfunction is supported also by studies showing that interventions that improve baroreflex sensitivity (BRS), such as physical training or β -adrenoceptor blockade, may also beneficially affect the patient's prognosis.

The above findings have led to great interest in the assessment of BRS in humans, both in the laboratory and the clinic. This article briefly addresses the role of the arterial baroreflex in cardiovascular physiology, its possible clinical relevance, and the methods most commonly employed for assessing its effectiveness in human cardiovascular regulation. Emphasis is placed on the exciting possibilities offered by modern approaches based on computer analysis of spontaneously occurring blood pressure and heart rate fluctuations.

Anatomy and Physiology

Intravascular pressure is continuously sensed by stretch receptors, located in the systemic and

pulmonary arteries and veins. The two main populations of receptors engaged in this task, known as baroreceptors, can be found at the carotid sinus and the aortic arch (i.e., in two key regions for continuous monitoring of blood pressure and thus of blood supply to vital organs, in particular to the brain). Other important baroreceptor populations involved in regulating perfusion pressure to vital organs are those located in the walls of cerebral and renal arteries, where they appear to play a major role in the local autoregulation. Arterial baroreceptors, in combination with the stretch receptors located in the cardiac walls and in the lungs (so-called cardiopulmonary receptors) also play a role in regulating the secretion of humoral substances involved in blood pressure and volume homeostasis (e.g., renin, vasopressin, catecholamines).

An increase in vascular or cardiac pressure determines an increase in arterial or cardiac transmural pressure, which represents the natural stimulus for stretch receptors located in the vascular or cardiac walls, and thus an increase in firing along their afferent fibers. The opposite occurs for a reduction in cardiac or vascular transmural pressure. The afferent impulses which follow baroreceptor stimulation by a blood pressure rise relay within the nucleus of the solitary tract in the brainstem. This is followed by inhibition of brainstem sympathetic centers and stimulation of brainstem parasympathetic centers. This baroreceptor reflex arc forms a negative feedback loop in which a rise in blood pressure results in the inhibition of central sympathetic outflow to the heart and peripheral circulation. This occurs through a sequential action on the intermediolateral cell column, preganglionic neurons, sympathetic ganglia, and, finally, on postganglionic sympathetic nerves directed to heart, arterioles, veins, and kidneys. The baroreflex negative feedback loop which responds to a rise in blood pressure also includes an increase in the activity of parasympathetic brain stem centers and in the efferent parasympathetic activity directed to the heart. The final result is a reflex reduction in blood pressure and heart rate, aimed at counteracting the initial increase in blood pressure. A brainstem noradrenergic pathway is also involved in this process, interacting with the nucleus of the solitary tract to participate in the suppression of sympathetic outflow. This noradrenergic inhibitory pathway is stimulated by centrally acting α -adrenoceptor agonists, such as clonidine, that may thus potentiate through this mechanism the baroreceptor-mediated vasodepressor response.

In the opposite sense, a fall in blood pressure is responsible for a reduction in baroreceptor transmural pressure and thus for a reduction in

baroreceptor afferent firing. This in turn diminishes central inhibition, resulting in an increase in sympathetic outflow and a reduction in vagal efferent activity and thus a reflex rise in blood pressure and heart rate aimed at counteracting the initial fall in blood pressure.

The arterial baroreflex loop, including the arterial baroreceptors, their afferent fibers, central integrating neurons, efferent fibers, and peripheral effector organs, represents the major regulatory mechanism involved in cardiovascular homeostasis and plays a pivotal role in buffering blood pressure fluctuations and in promoting heart rate variability. Moreover, data are available demonstrating that its derangement in a number of diseases, characterized by autonomic impairment, may not only contribute to the clinical manifestations of these pathological conditions, but may also have a prognostic relevance in increasing the risk of cardiovascular events. This makes the accurate assessment of baroreflex function in humans a key methodological issue.

Methods for Assessing Baroreflex Function

Laboratory Methods for Assessing Baroreflex Sensitivity

Several methods are available to assess the arterial baroreflex function in the laboratory. These techniques have been successfully used to clarify its physiological role in cardiovascular modulation, its alterations in disease, and also how this alteration can be modified by treatment. All of the methods require the application of an external stimulus on the subject under evaluation and offer a spot quantification of BRS obtained in standardized laboratory conditions.

A number of pioneering methods to assess BRS, including carotid sinus massage, electrical stimulation of carotid sinus nerves, anesthesia of carotid sinus nerves and vagi, and occlusion of common carotid arteries, are now only of historical interest. Other laboratory methods currently employed in this field include:

- The Valsalva maneuver (which, however, also triggers alterations in chemoreceptor and cardiopulmonary receptor activity, thus making heart rate responses poorly specific),
- Head-up tilting and lower body negative pressure application (which are also not specific because cardiopulmonary receptor deactivation, due to a reduction in venous return and central blood volume, and vestibular stimulation, in case of tilting, also contribute to the cardiovascular adjustments),
- The intravenous bolus injection of a small dose of a pressor or a depressor agent free from any major

direct effect on the heart (to increase or reduce systolic blood pressure (SBP) and reflexly lengthen or shorten, respectively, R-R interval; the slope of the regression line fitting the SBP and R-R interval changes is taken as a measure of BRS), and

- The use of a neck chamber device (which allows its inner air pressure to be increased or reduced in a graded and long-lasting fashion, thus resulting into graded, long-lasting, and quantifiable reductions and increases in carotid transmural pressure and, as a result, in the input stimulus to carotid baroreceptors).

The key advantage of the neck chamber method is that it allows not only heart rate but also blood pressure modulation by arterial baroreceptors to be specifically investigated (Figure 1). This goes with the disadvantage, however, that only carotid baroreceptor function is addressed and that the effect of carotid baroreceptor involvement is counteracted by the aortic baroreflex, which is deactivated during the blood pressure fall induced by carotid baroreceptor stimulation and stimulated during the blood pressure rise induced by the carotid baroreceptor deactivation.

Despite the useful information they provide on arterial baroreflex function in humans, all the above laboratory methods also carry a number of limitations, including the need of external artificial stimulation on the cardiovascular system, the limited reproducibility of the data so obtained, and the inability to characterize the dynamic features of neural cardiovascular reflex modulation in real-life conditions. This has stimulated the development of new techniques aimed at overcoming these problems.

Techniques for Assessing Spontaneous Baroreflex Modulation of Heart Rate

An important step forward in the investigation of the arterial baroreflex in humans is represented by techniques that analyze the sensitivity of spontaneous baroreflex control of heart rate, i.e., techniques based on the analysis of spontaneously occurring blood pressure and heart rate fluctuations. These techniques do not require any external intervention on the subject under evaluation, and can be used to assess BRS not only in standardized laboratory conditions, but also in daily life. From a methodological point of view, these methods require beat-to-beat monitoring of SBP coupled with an electrocardiogram signal to obtain a precise assessment of R-R interval–SBP relationships. If computer scanning of blood pressure signal is done at sufficiently high sampling frequency (≥ 100 Hz) and if the blood pressure waveform undergoes a parabolic interpolation of its peak, the heart interval can be estimated through the blood pressure signal only by quantifying

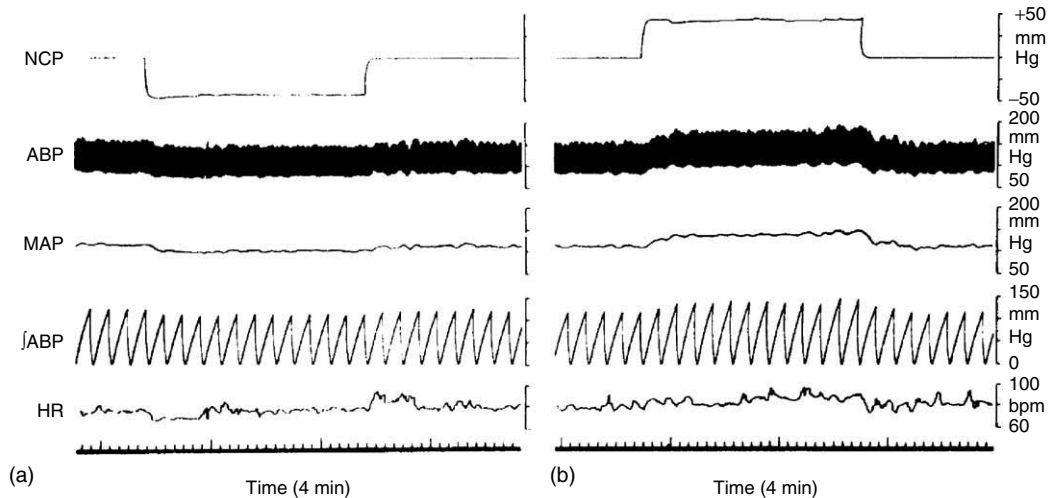


Figure 1 Blood pressure and heart rate effects of a 2-minute application of negative (a) and positive pneumatic pressure (b) within a neck chamber device. ABP, arterial blood pressure recorded through a catheter in the radial artery; $\int$ ABP, blood pressure values integrated every 10 s; bpm, beats per minute; HR, heart rate; MAP, mean arterial pressure; NCP, neck chamber pressure. Reproduced with permission from Ludbrook et al. (1977). The variable neck-chamber method for studying the carotid baroreflex in man. *Clinical Science and Molecular Medicine* 53, 165–171. © The Biochemical Society and the Medical Research Society 1977.

pulse interval (i.e., the interval between consecutive systolic peaks), with a degree of accuracy comparable to that of R-R interval detection from an ECG tracing. A further simplification of the technical requirements for estimating the spontaneous baroreflex in humans is possible because continuous blood pressure recordings can be obtained through noninvasive devices with sufficient accuracy both at rest and in ambulatory conditions.

The sequence method The sequence method is a time-domain analysis, based on computer identification of spontaneously occurring sequences of three or more consecutive beats characterized by either progressive increases in SBP and reflex lengthening in heart interval (R-R interval or pulse interval), or by progressive reductions in SBP and reflex shortening in heart interval (Figure 2). Beat-to-beat changes should be greater than a minimum threshold (usually set at 1 mm Hg for SBP and 5 ms for heart interval) in order to form a sequence. Similarly to what is performed for the laboratory method based on the intravenous bolus injection of vasoactive drugs, the slope of the regression line between spontaneous SBP and heart interval changes is taken as an index of the sensitivity of arterial baroreflex modulation of heart rate (BRS). The experimental demonstration of the specificity of the sequence method (i.e., that R-R interval changes observed in each sequence in response to the spontaneous changes in SBP are actually dependent on the baroreflex) was obtained in conscious cats in which blood pressure and heart rate were continuously recorded for 3–4 h both before and after sino-aortic denervation. In intact animals,

thousands of sequences of either the hypertension/bradycardia or the hypotension/tachycardia type were identified, whereas after sino-aortic denervation the number of both sequences became negligible (Figure 3). The sequence technique can also provide additional pieces of information on the baroreflex function through two parameters: the occurrence rate of spontaneous sequences, usually quantified as number of sequences per hour; and the baroreflex effectiveness index (BEI). BEI is defined as the ratio between the number of sequences occurring in a given period of time, and the number of blood pressure ramps, i.e., progressive increases or progressive reductions of SBP of three or more beats. BEI may range between 0 (when the baroreflex does not generate any sequence to buffer blood pressure ramps) and 1 (when all SBP ramps belong to a hypertension/bradycardia or to a hypotension/tachycardia sequence).

Spectral methods A different approach for estimating the spontaneous BRS on the heart is based on frequency-domain analysis. This means that the dynamics of blood pressure and heart rate time series are analyzed by decomposing the overall variance (or power) as a function of the frequency of the individual oscillations contained in the signals (power spectrum).

α Coefficient Assessment of spontaneous BRS by the α coefficient technique is based on the computation of the power spectrum of SBP and R-R interval beat-to-beat series by either the fast Fourier transform (FFT) algorithm or through an autoregressive

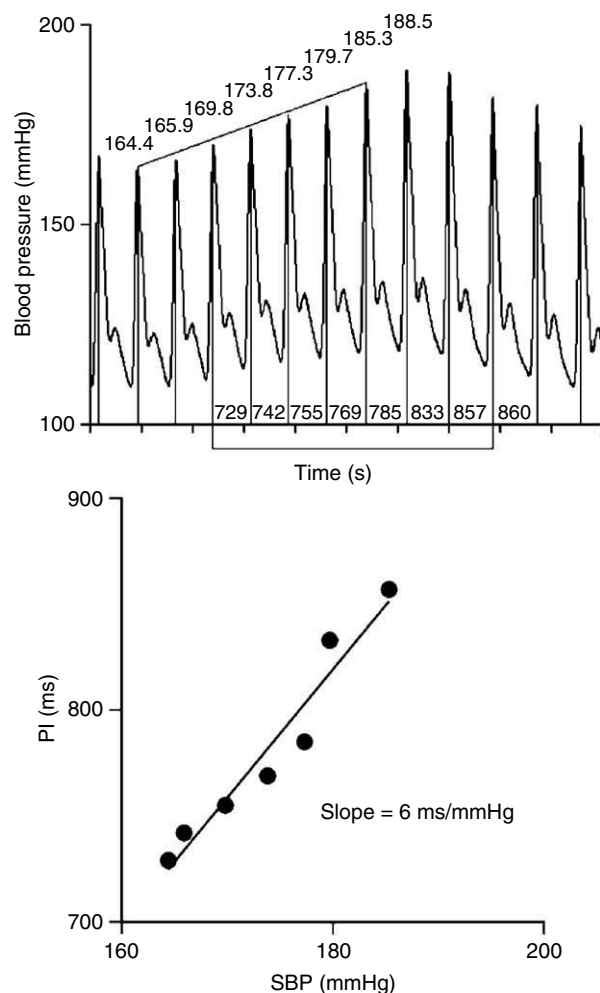


Figure 2 Scheme illustrating the sequence method for assessment of spontaneous baroreflex sensitivity. Blood pressure is scanned in search of a sequence of beats in which a progressive increase of systolic blood pressure (SBP) is followed, after a lag of zero, one, or two beats, by a progressive lengthening of heart interval, measured as R-R interval or pulse interval (PI; hypertension/bradycardia sequence, as shown in the upper panel of the figure); or in which a progressive SBP decrease is followed after a similar lag by a progressive shortening of heart interval (hypotension/tachycardia sequence). Often it is required a minimal beat-to-beat SBP change of 1 mmHg, and a minimal heart interval change of 5 ms to identify a sequence, as in this example. The slope of the regression lines between SBP and PI or R-R interval values forming each sequence (lower panel) gives the estimate of baroreflex sensitivity.

modeling (Figure 4). The SBP and R-R interval spectra so obtained are then integrated obtaining a measure of power (in mm Hg² or ms²) over frequency regions where SBP and heart rate oscillations usually show a high level of linear correlation, which is assumed to be due to the baroreflex coupling. These regions are around 0.1 Hz (the so-called low-frequency [LF] band) and at the respiratory frequency, generally around 0.3 Hz (high-frequency [HF] band). The degree of linear coupling between oscillations in a

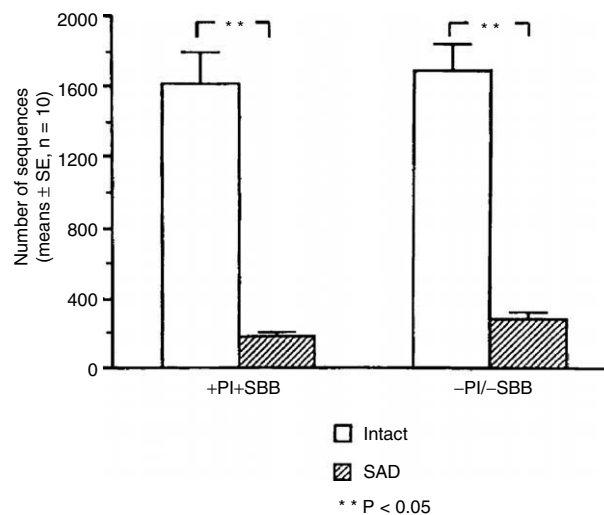


Figure 3 Number of hypertension/bradycardia (+pulse interval (PI)/+systolic blood pressure (SBP)) and hypotension/tachycardia (-PI/-SBP) sequences. Data are shown as average values ($\pm$ standard error of the mean (SE)) derived from the analysis of 3-h intra-arterial recordings for a group of 10 intact cats and for a group of 10 cats in which intra-arterial blood pressure was recorded 3 weeks after sino-aortic denervation (SAD). Asterisks refer to the statistical significance of the differences between the intact and the SAD conditions. Adapted with permission from Bertinieri et al. (1988). Evaluation of baroreceptor reflex by blood pressure monitoring in unanesthetized cats. *Hypertension* 12, 214–222. © 2002.

given frequency band is quantified by the coherence function (which can be considered as a type of linear correlation coefficient in the frequency domain). When the squared modulus of the coherence between SBP and R-R interval is sufficiently high, i.e., greater than 0.5, then the α coefficient is estimated as the root-squared ratio between the R-R interval power and the SBP power, in milliseconds per millimeter of mercury (ms/mm Hg). According to the frequency band considered for spectral integration, an α -LF or an α -HF coefficient is estimated.

Transfer function In part similar to the α coefficient, this method is based on the calculation of the SBP spectrum and of the cross-spectrum between SBP and R-R interval. The transfer function between the output, i.e., heart interval, and the input, i.e., SBP, is estimated as the ratio between the cross-spectrum and the SBP spectrum. The transfer function values in the LF or in the HF bands are taken as an estimate of BRS. If SBP and R-R interval are completely linearly correlated (coherence modulus equal to 1), α coefficient and transfer function estimates are identical. Otherwise, the transfer function method provides estimates lower than the α coefficients, because it tends to remove uncorrelated components from the output (heart interval).

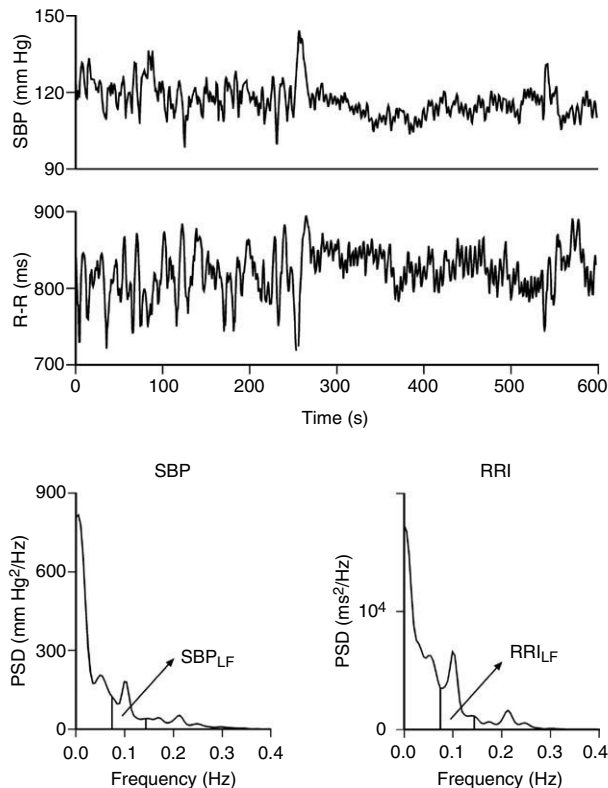


Figure 4 Scheme illustrating baroreflex sensitivity assessment by calculation of the α -coefficient. From a few minutes long (typically, 10 min) beat-to-beat recording of systolic blood pressure (SBP) and R-R interval (RRI) time series (upper panel), SBP and RRI power spectra are derived and integrated over a low-frequency (LF) band between 0.07 Hz and 0.14 Hz (see figure) or over a high-frequency (HF) band around 0.3 Hz; the integral is the area under the curve. From the powers in the LF bands, RRI_{LF} and SBP_{LF} , the α -LF coefficient is obtained as the root square of the RRI_{LF}/SBP_{LF} ratio; the α -HF coefficient is similarly obtained from RRI_{HF} and SBP_{HF} . PSD, power spectral density.

Despite the methodological differences between time-domain and frequency-domain methods, average quantifications of BRS appear to be surprisingly similar. In either case, spontaneous BRS estimates so obtained are quantitatively different from (although closely related to) BRS values derived through the application of laboratory techniques. This emphasizes the complementary nature of the information on baroreflex function provided by laboratory and spontaneous methods, which focus on different aspects of arterial baroreflex modulation of heart rate.

Models

Physiological models Another approach for estimating the BRS is based on the mathematical modeling of the cardiovascular regulatory mechanisms. In this approach, each element of the model represents a specific component of the cardiovascular system.

The model identification is obtained by tuning the model coefficients in order to fit the experimental data. In particular, physiological models may allow to separately estimate a vagal and sympathetic BRS.

Black-box models Black-box models involve a different modeling strategy. It consists in describing individual elements of the cardiovascular system by means of mathematical equations whose coefficients do not have a direct physiological interpretation, but which can be efficiently estimated from the measured heart rate and blood pressure beat-to-beat series. Occasionally, these models also include respiratory signals. This class of models is generally based on a set of autoregressive and moving average equations reciprocally relating SBP and R-R interval beat-to-beat time series (ARMA models). The BRS can then be derived by a mathematical manipulation of the estimated model coefficients. This approach is particularly interesting because it may address aspects of the baroreflex function, in addition to BRS, that are difficult to investigate using experimental laboratory procedures, provided a suitable model is identified.

Other methods

Statistical link (Z-index) The statistical link (Z-index) method is based on the evaluation of the statistical dependence between blood pressure and heart interval as assessed by the Z-coefficient. The method determines the association between two probabilistic events, X and Y, calculating a statistical link coefficient, $Z(X,Y)$, which may range between -1 (when X and Y are mutually exclusive events) and $+1$ (when X is included within Y). Z is equal to 0 when X and Y are independent. Intermediate values between -1 and 0 quantify the partial exclusion of Y by X, while values between 0 and $+1$ quantify the partial dependence of Y on X. Z is estimated by counting the number of times that the X and Y events occurred, and the number of times that they occurred simultaneously, in a total of N observations. To quantify the statistical link between mean blood pressure and heart rate beat-to-beat values, the X and Y events are the occurrence of blood pressure and heart rate values within specific amplitude intervals. By choosing Q intervals for mean blood pressure and M for heart rate, $Z(X,Y)$ is defined over a grid of $Q \times M$ couples of (X,Y) events. The BRS can be estimated by the slope of the regression line in the X-Y plane fitting over-threshold Z values.

Cross-correlation (X-BRS) The cross-correlation (X-BRS) approach is based on the estimation of the cross-correlation between very short (10 s) segments of SBP and heart interval data. For any blood pressure

segment, the cross correlation is repeatedly estimated by considering sliding 10-s segments of heart interval with a time delay with respect to blood pressure up to 5 s. Segments showing the maximal correlation coefficient are selected and the slope of the regression line obtained from the corresponding SBP and heart interval values is the estimated BRS, provided that the probability of being a random regression is lower than 1%.

Clinical Value of Baroreflex Sensitivity

Spontaneous Baroreflex Sensitivity in Health and Disease

Data obtained by methods for assessing spontaneous BRS have provided a large body of information on daily baroreflex function in healthy subjects and on its derangement in disease. This information consists in the detailed evidence on the fast and marked changes in BRS associated with changes in behavioral conditions. Results provided by studies where these

methods were employed have clearly documented the high degree of variability which physiologically characterizes BRS in different conditions. The dynamic nature of the baroreflex assessment offered by techniques like the sequence method, as well as by other methods for spontaneous baroreflex analysis, has offered us the unique possibility to monitor both the slow and fast changes in baroreflex function, the most striking ones being those associated with the sleep-wake cycle. The typical patterns of spontaneous baroreflex modulation of heart rate over 24 h that characterize young and elderly subjects are shown in Figure 5. In addition to the occurrence of minute-to-minute variability, in young individuals BRS showed a clear-cut modulation between the day and night, displaying its lowest values during the daytime, under daily physical and emotional challenges, and its highest values during night sleep. In elderly subjects, 24-h BRS was not only on average smaller than in young individuals, but it also displayed a loss of the physiological day-night modulation, thus suggesting a multifold age-related impairment of baroreflex

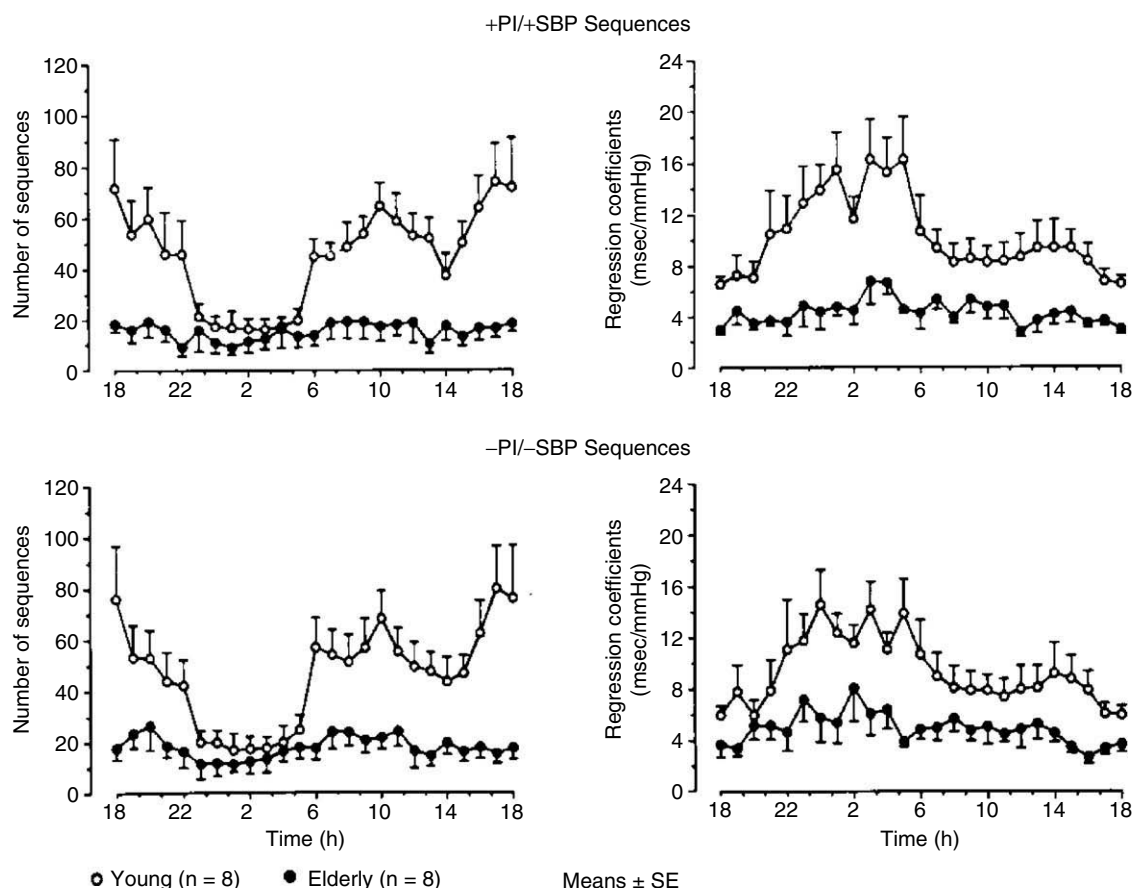


Figure 5 Number of hypertension/bradycardia and hypotension/tachycardia sequences (left) and sequence regression coefficients (right). Data are shown as average hourly values ($\pm$ standard error of the mean (SE)) for a group of young and a group of elderly individuals. Adapted with permission from Parati et al. (1995). Effects of aging on 24-h dynamic baroreceptor control of heart rate in ambulant subjects. *American Journal of Physiology* 268, H1606–H1612. © 1995.

heart rate modulation. These techniques also allowed us to detect impairments of BRS in circumstances such as cigarette smoking, an impairment which escapes recognition when traditional laboratory methods are employed.

Assessment of spontaneous BRS also allowed quantification of changes in hypertensive subjects as compared to normotensive individuals; to quantify impairments occurring during sleep, like in patients with obstructive sleep apnea syndrome; and to study the striking baroreflex dysfunction typical of patients

with primary autonomic failure. Analysis of spontaneous BRS has also been performed in patients with diabetes with or without clinical and laboratory evidence of autonomic dysfunction, as documented by the classic Ewing tests. The results have shown that both time and frequency domain methods can identify impairments of baroreflex control of heart rate when traditional laboratory tests still yield normal responses. In addition, heart rate variability is only marginally reduced (Figure 6) emphasizing the superiority of these techniques as compared to

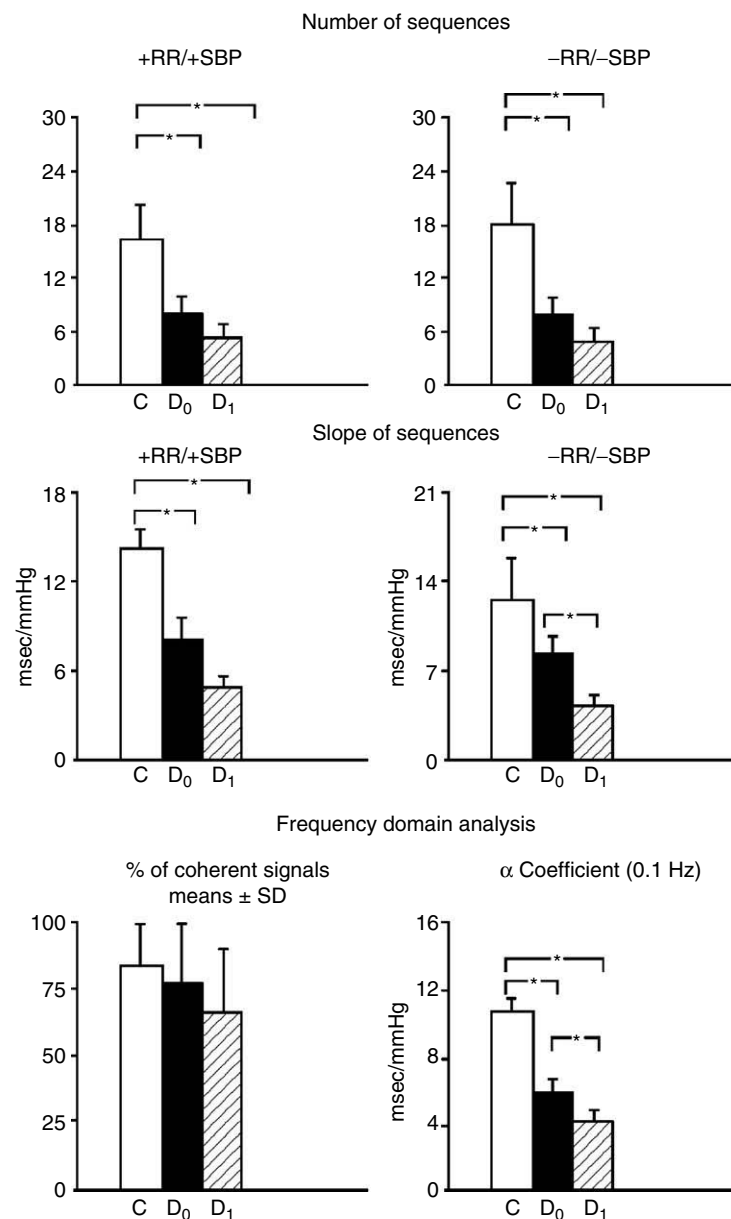


Figure 6 Time and frequency domain estimates of the spontaneous sensitivity of the baroreceptor-heart rate reflex in control healthy subjects (C) and in patients with diabetes without (D₀) or with (D₁) autonomic dysfunction at classical laboratory tests (*P < 0.05). SBP, systolic blood pressure. Reproduced with permission from Frattola et al. (1997). Time and frequency domain estimates of spontaneous baroreflex sensitivity provide early detection of autonomic dysfunction in diabetes mellitus. *Diabetologia* 40, 1470–1475. © 1997.

traditional laboratory procedures in the early detection of autonomic abnormalities that may carry a higher risk of morbidity and mortality. Because of the capability of these methods to provide continuous dynamic measures of BRS, they have been also proposed for the diagnosis of brain death on patients with severe brain injuries allowing the continuous monitoring of brainstem centers included in the baroreflex loop.

Prognostic Value of Baroreflex Sensitivity

The possible clinical relevance of a reduced BRS in patients with coronary artery disease has been suggested by both animal and human studies. The observation that a blunted reflex vagal modulation of the heart may be associated with a higher risk of major arrhythmias has stimulated a number of clinical studies aimed at assessing the possible link between BRS and patient's prognosis after acute myocardial infarction (AMI). Indeed, 2–7 days after an AMI, BRS was found to be markedly depressed by several studies. More consistent evidence on the clinical value of reduced BRS has been provided by a multicenter prospective study aimed at assessing the relative prognostic importance of several markers of neural autonomic cardiac modulation in patients after myocardial infarction (ATRAMI [Autonomic Tone and Reflexes After Myocardial Infarction] Trial). The results of this trial indicate that BRS is significantly related to patients' prognosis, even after accounting for the independent effect of other established prognostic markers such as the degree of heart rate variability and left ventricular function. In fact, BRS seems to represent a prognostic marker which does not provide redundant information as compared to heart rate variability. This is in line with the report that, in patients after AMI, subjects with and without a higher frequency of life-threatening cardiac arrhythmias, respectively, displayed a clear difference in BRS, but not a comparably clear-cut difference in heart rate variability. In spite of the above evidence, the real value of BRS in the evaluation of patients' prognosis after an AMI cannot yet be regarded as unequivocally clarified. Indeed, a clear-cut relationship between impairment of BRS on one side and the infarct size or the degree of left ventricular dysfunction on the other side was not found in a number of papers. The negative results on the relation between BRS and patients' prognosis obtained in some studies might also have depended, however, on the rather crude method which was employed to test the effectiveness of baroreceptor-heart rate control in these studies. This consisted in most instances in the evaluation of the reflex bradycardia which follows the increase in blood pressure

obtained by one or only a few bolus intravenous injections of phenylephrine.

The potential importance of BRS as a clinically relevant marker in cardiac patients has been further supported by a series of studies which showed that pharmacological and nonpharmacological interventions able to reduce sympathetic activity and to increase vagal activity and BRS may be protective against cardiac mortality and, in particular, sudden cardiac death. Administration of β -adrenoceptor blocking agents, which modulate cardiac sympathetic stimulation and increase BRS, was shown to improve morbidity and mortality in myocardial infarction patients even in the case of congestive heart failure. A similar mechanism might explain the beneficial prognostic effect in those patients who use angiotensin-converting enzyme (ACE) inhibitors and who exercise, interventions which have also been reported to increase BRS after AMI. Available data indicate that BRS may have a prognostic value also in two other pathological conditions, namely congestive heart failure and diabetes mellitus. In heart failure patients, the risk of death was significantly higher in patients with lower BRS as assessed by a neck chamber device, although the prognostic impact of a low BRS was no more significant when the influence of an increased sympathetic activity was simultaneously considered. This might depend on the inverse relationship that characterizes sympathetic and parasympathetic cardiovascular modulation, because evidence is available that deterioration of left ventricular function is associated with a decrease in parasympathetic tone. A reduced vagal tone thus may identify not only patients at high risk of sudden death, but also those at high risk of pump failure, both parameters being closely related to prognosis in heart failure. A reduced BRS has also been shown to correlate with an increased frequency of severe arrhythmias. Alternatively, a number of clinical studies have clearly shown the importance of a reduced vagal activity and of a deranged sympathetic activity on the prognosis of patients with diabetes mellitus. Indeed a number of patients with diabetes with clinically evident autonomic dysfunction die suddenly and unexpectedly, probably because of cardiac arrhythmias. Moreover, several studies suggest that even subclinical autonomic dysfunction may be predictive of future mortality in patients with diabetes mellitus. In conclusion, even if we have not yet clearly demonstrated that the prognosis of patients with myocardial infarction, heart failure, or diabetes, and the values of BRS are causally linked to each other, the available evidence is strongly in favor of such a possibility. Further progress in this field is likely to be supported by the recent development of more sensitive and

specific methods to assess spontaneous reflex cardiac autonomic modulation. Indeed, preliminary evidence is available that a reduced spontaneous BRS in patients after stroke is associated with increased mortality.

Conclusion

This article has summarized the role of the arterial baroreflex in human cardiovascular physiology and pathophysiology and the pros and cons of different methods for assessing BRS in humans. Laboratory methods have so far largely contributed to our present knowledge of arterial baroreflex function in humans. Modern techniques for spontaneous baroreflex analysis, however, have more recently provided us with a much deeper insight into the features of daily life baroreflex function as compared to what is offered by traditional laboratory tests. The recent evidence obtained in myocardial infarction patients, in patients with congestive heart failure, and in patients with diabetes through even the relatively crude estimates of BRS offered by laboratory tests has increased the interest in baroreflex analysis in humans. Whether the richer information provided by the new methods for assessing spontaneous BRS might further improve the prognostic evaluations of such patients, however, needs to be tested by large-scale longitudinal studies.

See Also the Following Articles

Beta-Adrenergic Blockers; Cardiovascular System and Stress; Heart Disease/Attack; Hypertension.

Further Reading

- Bertinieri, G., Di Rienzo, M., Cavallazzi, A., et al. (1988). Evaluation of baroreceptor reflex by blood pressure monitoring in unanesthetized cats. *Hypertension* 12, 214–222.
- Di Rienzo, M., Castiglioni, P., Mancia, G., et al. (2001). Advancements in estimating baroreflex function. *IEEE Engineering In Medicine And Biology Magazine* 20, 25–32.
- Di Rienzo, M., Mancia, G., Parati, G., et al. (eds.) (1997). *Frontiers of blood pressure and heart rate analysis*. Amsterdam: IOS Press.
- Frattola, A., Parati, G., Gamba, P., et al. (1997). Time and frequency domain estimates of spontaneous baroreflex sensitivity provide early detection of autonomic dysfunction in diabetes mellitus. *Diabetologia* 40, 1470–1475.
- La Rovere, M. T., Bigger, T. J., Marcus, F. I., et al. (1998). Baroreflex sensitivity and heart-rate variability in prediction of total cardiac mortality after myocardial infarction. ATRAMI (Autonomic Tone and Reflexes After Myocardial Infarction) Investigators. *Lancet* 351, 478–484.
- La Rovere, M. T., Pinna, G. D., Hohnloser, S. H., et al. (2001). Baroreflex sensitivity and heart rate variability in the identification of patients at risk for life-threatening arrhythmias. Implications for clinical trials. *Circulation* 103, 2072–2077.
- Laude, D., Elghozi, J. L., Girard, A., et al. (2004). Comparison of various techniques used to estimate spontaneous baroreflex sensitivity (the EuroBaVar study). *American Journal of Physiology – Regulatory Integrative and Comparative Physiology* 286, R226–R231.
- Ludbrook, J., Mancia, G., Ferrari, A., et al. (1977). The variable neck-chamber method for studying the carotid baroreflex in man. *Clinical Science and Molecular Medicine* 53, 165–171.
- Mancia, G. and Mark, A. L. (1983). Arterial baroreflexes in humans. In: Shepherd, J. T. & Abboud, F. M. (eds.) *Handbook of physiology, sect 2, The cardiovascular system IV, peripheral circulation and organ blood flow*, pp. 755–794. Bethesda, MD: American Physiological Society.
- Mancia, G., Parati, G., Di Rienzo, M., et al. (1997). Blood pressure variability. In: Zanchetti, A. & Mancia, G. (eds.) *Handbook of hypertension*. Amsterdam: Elsevier Science.
- Parati, G., Di Rienzo, M. and Mancia, G. (2000). How to measure baroreflex sensitivity: from the cardiovascular laboratory to daily life. *Journal of Hypertension* 18, 7–19.
- Parati, G., Di Rienzo, M. and Mancia, G. (2001). Dynamic modulation of baroreflex sensitivity in health and disease. *Annals of the New York Academy of Sciences* 940, 469–487.
- Parati, G., Di Rienzo, M., Omboni, S., et al. (2002). Computer analysis of blood pressure and heart rate variability in subjects with normal and abnormal autonomic cardiovascular control. In: Mathias, C. J. & Bannister, R. (eds.) *Autonomic failure*, pp. 211–223. London: Oxford University Press.
- Parati, G., Frattola, A., Di Rienzo, M., et al. (1995). Effects of aging on 24-h dynamic baroreceptor control of heart rate in ambulant subjects. *American Journal of Physiology* 268, H1606–H1612.
- Parati, G., Saul, J. P. and Castiglioni, P. (2004). Assessing arterial baroreflex control of heart rate: new perspectives. *Journal of Hypertension* 22, 1259–1263.
- Parati, G., Saul, J. P., Di Rienzo, M. and Mancia, G. (1995). Spectral analysis of blood pressure and heart rate variability in evaluating cardiovascular regulation. A critical appraisal. *Hypertension* 25, 1276–1286.
- Pitzalis, M., Parati, G., Massari, F., et al. (2003). Enhanced reflex response to baroreceptor deactivation in subjects with tilt-induced syncope. *Journal of the American College of Cardiology* 41, 1167–1173.
- Robinson, T. G., Dawson, S. L., Eames, P. J., et al. (2003). Cardiac baroreceptor sensitivity predicts long-term outcome after acute ischemic stroke. *Stroke* 34, 705–712.
- Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology (1996). Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation* 93, 1043–1065.

Arthritis

R L Wilder

MacroGenics, Inc., Rockville, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 251–254, © 2000, Elsevier Inc.

Types of Arthritis

Rheumatoid Arthritis

Does Stress Play a Role in Rheumatoid Arthritis?

Glossary

<i>Arthritis</i>	A group of disorders characterized by chronic pain and progressive impairment of joints and adjacent soft tissues.
<i>Autoimmunity</i>	An immune response directed at self tissues that can result in inappropriate tissue injury.
<i>Corticosteroids</i>	Natural or synthetic steroid hormones with potent anti-inflammatory actions.
<i>Cytokines</i>	Proteins produced by a variety of cells that mediate intercellular communication or other effector functions such as the generation or control of the inflammatory response.
<i>Hypothalamic-pituitary-adrenal (HPA) axis</i>	A complex set of direct influences and feedback interactions among the hypothalamus, pituitary and adrenal glands.
<i>Hypothalamus</i>	A part of the brain that contains cells that coordinate numerous body functions, including the activity of the stress response system.
<i>Inflammation</i>	A nonspecific response to injury that is intended to neutralize the cause of the injury and to orchestrate repair. It is characterized by heat, pain, redness, and swelling.
<i>Immunity</i>	A specific response to injury or potential injury that is intended to neutralize the cause or potential threat. Macrophages, lymphocytes, antibodies, and cytokines are important components of the immune response.
<i>Lymphocytes</i>	Cells that play key roles in mediating an immune response. There are two general types: B and T lymphocytes. B lymphocytes are primarily involved in producing antibodies; T cells are involved in cellular immune processes and in regulating the activity of other cell types, including B cells.

Macrophages

Major histocompatibility complex

Neuroendocrine-immune loop

Rheumatoid arthritis

Sympatho-adrenal axis

Synovitis

Synovium

Tumor necrosis factor (TNF)- α

Cells that play a key role in inflammatory processes. They engulf foreign materials and secrete an extensive array of biologically active substances such as cytokines. A group of genes located on human chromosome 6. Their name is derived from their role in triggering the rejection of transplanted tissues. They also have major roles in regulating the responsiveness of the immune system.

A theory emphasizing the bidirectional interactions between the brain and the endocrine and immune systems. The immune system stimulates and regulates the activity of the nervous and endocrine systems, and they, in turn, regulate the activity of immune system.

A specific type of arthritis that is characterized by progressive, symmetrical, destructive inflammation of multiple joints, particularly in the hands, wrists, and feet.

The adrenal medullary and autonomic sympathetic nervous systems; the major products produced are the catecholamines, epinephrine, and norepinephrine. The inflammation of the synovium; a major component of multiple types of arthritis.

A thin, delicate membrane that lines the cavity of joints. Its major function is to secrete substances, such as hyaluronic acid, that reduce the friction associated with joint movement.

A cytokine produced by activated macrophages that induces a variety of types of inflammatory injury.

Arthritis is the general term for a large, diverse group of musculoskeletal diseases characterized by chronic pain and progressive impairment of joints and surrounding soft tissues. These illnesses are very common and extremely costly to individuals, families, and society because they seriously impair work ability and decrease individual and household incomes. Affected individuals suffer from the inability to conduct the basic personal tasks, household tasks, and discretionary activities of living. For example, individuals with arthritis go out less frequently, spend less time with hobbies, and have much more difficulty earning a living or maintaining a household. Moreover, some forms of arthritis result in premature mortality in addition to severe disability.

Arthritis, particularly rheumatoid arthritis, is associated with an increased incidence of psychological distress and depression. Some of this distress is an intrinsic component of the disease process itself. In addition, the lack of independence and the inability to participate in various activities lead to frustration and feelings of helplessness. Not surprisingly, arthritis can strain marriages and family relationships.

Types of Arthritis

There are many forms of arthritis. Clinicians generally categorize arthritis into two major groups: monoarticular and polyarticular. Monoarticular disease is pain or swelling in a single joint, whereas polyarticular disease involves multiple joints. Monoarticular arthritis commonly results from joint infection or from gout or other types of crystal depositions in the joint. However, any of the diseases that produce multiple joint involvement can also present with a monoarticular pattern of joint involvement. Polyarticular joint disease is categorized into two subgroups: inflammatory and noninflammatory. For example, osteoarthritis is a very common noninflammatory polyarticular disease, whereas rheumatoid arthritis is a common inflammatory polyarticular disease. The inflammatory and noninflammatory joint diseases differ widely in their etiology and clinical findings. For example, the diseases may vary in regard to their mode of presentation (acute, subacute, or chronic), age at presentation, pattern of joint involvement, x-ray changes, and associated clinical laboratory abnormalities. More than 30 different forms of polyarticular arthritis have been defined. With regard to stress and arthritis, rheumatoid arthritis has been the focus of most research.

Rheumatoid Arthritis

Rheumatoid arthritis is the most prominent member of the polyarticular forms of inflammatory arthritis. Most patients with rheumatoid arthritis have a symmetric arthritis that involves the small joints of the hands and wrists. The disease may also affect the joints of the feet, ankles, knees, hips, elbows, shoulders, neck, and jaw. The lower spine is usually not affected. Prominent symptoms include stiffness upon awakening in the morning, fatigue, and severe joint tenderness and swelling. Later, joint deformity and limitation of motion become prominent. Patients may also develop skin nodules. In severe cases, other parts of the body may be involved, including the eyes, lungs, heart, and nervous system. A hallmark of the disease is an association with an

abnormal immunoglobulin called rheumatoid factor, which is present in the serum of approximately 80% of the patients.

Rheumatoid arthritis is considered a multifactorial autoimmune disease in that no single factor appears to explain its causation or pathogenesis; that is, multiple factors interact to produce the disease process. The disease is much more common in women than in men by approximately 3 or 4 to 1. The most likely period of onset for women is around menopause. Rheumatoid arthritis is uncommon in men younger than age 45, but its incidence approaches that of women in older men. Before menopause, it is common for women to develop rheumatoid arthritis within a year of pregnancy. Interestingly, the period around menopause and after pregnancy is associated with a decreased ability to produce corticosteroids in response to stressful stimuli. Conversely, rheumatoid arthritis rarely develops during a pregnancy, and patients with rheumatoid arthritis frequently improve significantly during a pregnancy. Pregnancy is associated with the production of high levels of the stress hormone cortisol, and many investigators believe that cortisol plays a major role in suppressing the expression of rheumatoid arthritis during pregnancy.

Genetic factors also play a role in determining susceptibility to and the severity of rheumatoid arthritis. The disease clusters in families and is associated with other autoimmune diseases. Current research indicates that multiple genes including genes associated with the major histocompatibility complex on chromosome 6 are involved in regulating susceptibility. The etiology of rheumatoid arthritis is not known with certainty. Many investigators suspect that an infectious agent is involved. Recent evidence supports the possibility that a parvovirus called B19 is involved, but much more investigation is required to confirm or refute this. A sustained autoimmune inflammatory process in the joints mediates the destruction of joints in rheumatoid arthritis. The joints are normally lined by a thin delicate membrane called the synovium. In rheumatoid arthritis, the lining layer becomes massively thickened and infiltrated by inflammatory cells, including macrophages and lymphocytes. The synovial tissue stromal cells also increase markedly in number. The hallmark of the disease, however, is the destructive erosion of bone and cartilage at the margins of the joint where the synovium attaches to bone. This erosive destruction leads to the loss of joint space and deformities of the joint. Current evidence indicates that the destructive process is driven by the sustained production of pro-inflammatory protein mediators called cytokines. $\text{TNF-}\alpha$ is considered to be one of the most important

cytokines in rheumatoid arthritis. Major investigative efforts are currently directed at neutralizing the effects of TNF- α in the joints. Interestingly, the anti-inflammatory corticosteroids are one of the most potent naturally occurring inhibitors of TNF- α production known. High levels of TNF- α production in rheumatoid arthritis has raised questions about the adequacy of corticosteroid production and the integrity of the neuroendocrine-immune stress loop in patients with rheumatoid arthritis.

Does Stress Play a Role in Rheumatoid Arthritis?

Studies of corticosteroids and rheumatoid arthritis are historically tightly linked. Cortisone was first isolated from adrenal tissue in the 1930s, but interest in corticosteroids rocketed when Hench and Kendall and colleagues described the dramatic anti-inflammatory actions of cortisone in patients with rheumatoid arthritis. Their discovery resulted in a Nobel Prize in 1950. Subsequently, it was realized that long-term high-dose therapy with cortisone resulted in devastating side effects. This development led to highly polarized views of the role of corticosteroids in rheumatoid arthritis. These controversies continue today.

Corticosteroids are prime examples of key stress hormones. They serve an essential role in a variety of systemic adaptive responses. We now call these adaptive responses the stress response. This response is necessary to preserve physiological integrity. Significantly, it is now known that components of the stress response system interconnect extensively with components of the immune and inflammatory response systems. This interaction is bidirectional. The production of inflammatory cytokines such as TNF- α normally triggers the production of cortisol as part of a systemic activation of the stress response and the HPA and sympathoadrenal axes; cortisol, in turn, suppresses the production of TNF- α and normally prevents the unchecked overproduction of the toxic inflammatory mediator. Thus, the feedback interconnections of the neuroendocrine-immune loop normally function to restrain the inflammatory system.

The importance of this feedback loop is exemplified by the LEW rat, an inbred strain that has a blunted or sluggish activation of the stress system in response to numerous stressors, including TNF- α . Interestingly, these rats are also highly susceptible to an inflammatory arthritis resembling rheumatoid arthritis and other forms of autoimmune disease, and these diseases are markedly attenuated if the rats are treated with physiological doses of corticosteroids

very early in the course of disease. The origin of the stress-system defect in these rats has not been completely defined, but it clearly involves the brain and the hypothalamus. The LEW rat contrasts strikingly with the F344 inbred rat, which has an overactive stress response system and is resistant to arthritis and a variety of other autoimmune diseases.

The observations in the LEW and F344 rats have led to extensive investigations, still ongoing, investigating the hypothesis that the stress response system is hypoactive in patients with rheumatoid arthritis. The data available to date suggest that the stress system is, in fact, underactive. Whether its underactivity is a primary or acquired pathogenic factor is not known. In any event, most investigators accept the view that stress and abnormalities of the stress system may play a role in rheumatoid arthritis.

See Also the Following Articles

Autoimmunity; Cytokines; Immune Suppression; Menopause and Stress.

Further Reading

- Goronzy, J. J. and Weyand, C. M. (1997). Rheumatoid arthritis. In: Klippel, J. H. (ed.) *Primer on the rheumatic diseases*, pp. 155–161. Atlanta, GA: Arthritis Foundation.
- Papanicolaou, D. A., Wilder, R. L., Manolagas, S. C., et al. (1998). The pathophysiologic roles of interleukin-6 in human disease. *Annals of Internal Medicine* **128**, 127–137.
- Sternberg, E. M., Hill, J. M., Chrousos, G. P., et al. (1989). Inflammatory mediator-induced hypothalamic-pituitary-adrenal axis activation is defective in streptococcal cell wall arthritis-susceptible Lewis rats. *Proceedings of the National Academy of Sciences USA* **86**, 2374–2378.
- Sternberg, E. M., Young, W. S., Bernardini, R., et al. (1989). A central nervous system defect in biosynthesis of corticotropin-releasing hormone is associated with susceptibility to streptococcal cell wall-induced arthritis in Lewis rats. *Proceedings of the National Academy of Sciences USA* **86**, 4771–4775.
- Wilder, R. L. (1995). Neuroendocrine-immune interactions and autoimmunity. *Annual Review of Immunology* **13**, 307–338.
- Wilder, R. L. (1996). Hormones and autoimmunity: animal models of arthritis. *Baillieres Clinical Rheumatology* **10**, 259–271.
- Wilder, R. L. (1997). Corticosteroids. In: Klippel, J. H. (ed.) *Primer on the rheumatic diseases*, pp. 427–431. Atlanta, GA: Arthritis Foundation.
- Wilder, R. L. (1998). Hormones, pregnancy, and autoimmune diseases. *Annals of the New York Academy of Sciences* **840**, 45–50.

Arthritis – Psychological

J W Younger and A J Zautra

Arizona State University, Tempe, AZ, USA

© 2007 Elsevier Inc. All rights reserved.

Definition

Stress and Rheumatoid Arthritis

Stress and Osteoarthritis

A Special Case: Arthritis in Children

Glossary

<i>Adrenocorticotrophic hormone (ACTH)</i>	A hormone secreted by the pituitary that stimulates the adrenal gland to produce cortisol.
<i>Arthritis</i>	An inflammatory condition, primarily involving joints and surrounding tissue, that causes pain and may restrict mobility.
<i>Catecholamines</i>	Chemicals that serve as stress neurotransmitters in the nervous system and that are also secreted as hormones by the adrenal gland.
<i>Cortisol (hydrocortisone)</i>	A primary glucocorticoid that mobilizes energy for stress responses and serves as a potent anti-inflammatory and immunosuppressant.
<i>Epinephrine (adrenaline)</i>	One of the catecholamines responsible for the sympathetic (fight-or-flight) response.
<i>Flare</i>	Sharp but transient increases of inflammation and pain in some types of arthritis.
<i>Hypothalamic-pituitary-adrenal (HPA) axis</i>	A major neuroendocrine system involved in stress responses and responsible for production of catecholamines and cortisol.
<i>Pannus</i>	An abnormal growth of synovial tissue that is associated with joint destruction.
<i>Proinflammatory cytokines</i>	Intercellular immune messengers that increase inflammatory responses.
<i>Synovium</i>	A fluid-filled membrane that surrounds joints and provides joint protection, lubrication, and nutrition.

Definition

Arthritis is a term that covers over 100 inflammatory and degenerative joint conditions. The most common forms include osteoarthritis, rheumatoid arthritis, gout, ankylosing spondylitis, juvenile arthritis, psoriatic arthritis, and systemic lupus erythematosus. Despite having different etiologies, including infection, autoimmune pathology, and physical trauma, these

conditions are tied together by common characteristics of chronic pain and impairment in daily functioning. Stress–arthritis research has focused mainly on the two most common forms: rheumatoid arthritis (RA) and osteoarthritis (OA).

Stress and Rheumatoid Arthritis

RA affects approximately 3 million people in the United States and is three times more likely to occur in women than men. This autoimmune disorder is characterized by the inflammation of the synovium surrounding joints. The disease is typically symmetric, affecting the same joints on both sides of the body. Symptoms include swollen and/or tender joints as well as fatigue, low-grade fever, and general malaise. Many individuals with RA also experience episodic flares, in which symptoms are intensified. RA progresses from initial synovial inflammation to abnormal growth of the synovial tissue (pannus), which overgrows the joint area and breaks down surrounding bone and cartilage through the release of enzymes. RA can be severely disabling because of chronic pain and restricted joint flexibility.

Does Stress Exacerbate Rheumatoid Arthritis Symptoms?

Some evidence supports the view that stress may contribute to the onset of RA, but methodological challenges of retrospection diminish confidence in a number of these findings. A much stronger case has been made, however, for stress as being an aggravating factor in RA disease course. Minor stressors and daily hassles have, in particular, been linked to both self-reported pain and objective markers of inflammation. Both work stress and interpersonal stress predict increases of disease severity, and depression may further elevate stress-related inflammatory processes. Major acute stressors such as the sudden death of a spouse, on the other hand, may actually diminish the inflammatory response during the period of stress adaptation.

Biological support for stress–RA relations have focused mainly on proinflammatory cytokines, intercellular messengers responsible for triggering inflammatory immune responses. Many cytokines, including interleukin (IL)-1, IL-6, and tumor necrosis factor alpha (TNF- α) are stimulated by infection, physical trauma, and psychological stress. These stress-activated cytokines are also implicated in RA pathophysiology

to the extent that many pharmaceutical treatments are designed to suppress either the general immune system or specific cytokines. Biological response modifiers, the newest class of RA drugs, bind to proinflammatory cytokines and prevent their interaction with cell-surface receptors, thus inhibiting inflammatory immune responses.

In healthy individuals, the anti-inflammatory hormone cortisol (hydrocortisone) acts as an immunomodulator, suppressing both the production and activity of proinflammatory cytokines. Cortisol production may be induced through a stress response or triggered by increased levels of IL-6, serving as a modulating feedback loop. When endogenous cortisol levels are low, proinflammatory cytokine levels and disease activity are high. Daily circadian dips of cortisol in the early morning and late evening are associated with higher joint pain. Also, as cortisol levels diminish through late adulthood and old age, RA joint pain worsens. The recognition of hydrocortisone's anti-inflammatory properties led the wide use of externally administered glucocorticoids for managing RA symptoms.

Given that some stress hormones such as cortisol reduce proinflammatory cytokine activity, it may appear contradictory that stress increases pain in RA patients. It is important to note, however, that both pro- and anti-inflammatory agents are released during the stress response. Some stress hormones, such as prolactin, are proinflammatory.

The central pathophysiology of RA, then, may lie in the relative underactivity of anti-inflammatory agents compared with pro-inflammatory agents. This hypothesis is supported by observed dysregulations in the HPA axis. In RA, the HPA axis is probably underactive, leading to insufficient production of cortisol. Specifically, decreased cortisol secretion has been linked to diminished adrenal responsiveness to ACTH stimulation. RA patients may also exhibit abnormal responses to stress-induced catecholamines. In particular, it has been observed that externally administered epinephrine decreases cortisol in RA patients but not in healthy controls. If internal, stress-induced production of epinephrine similarly suppresses cortisol in RA patients, a proinflammatory state may be created, leading to the clinical observations of joint swelling and tenderness.

In addition to direct effects on peripheral sites of pain, stress may influence central processes involved in the affective responses to pain. Under periods of high stress and distress, RA symptoms and pain may be more salient. Conversely, positive emotion may buffer RA patients from the negative affective responses to pain, lowering stress reactivity and thus indirectly diminishing inflammatory processes.

Cognitive and Affective Factors in Rheumatoid Arthritis Stress–Pain Cycles

Given the exacerbating role of stress on RA symptoms, it is important to note that the disease itself may be a significant source of stress for afflicted individuals. In addition to pain, RA patients can suffer extensive physical limitations, which may lead to depressive symptoms and increased dependence on others. Activity limitations due to motion restriction and pain can affect every part of a person's life, impairing his or her ability to work, limiting recreational/social activities, and straining familial relationships. These physical and psychosocial stressors may further exacerbate pain and affect the disease course, potentially leading to a cycle of ever-increasing pain and stress. Not everyone is adversely affected by the disease, however, and many cognitive factors have been found to partially determine the impact of RA on a person's life. Premorbid depression, beliefs about the consequences of arthritis, and coping style all predict a more severe disease course. Resilient personality features and a strong social support network may sustain quality of life and forestall disability even among those with active disease. Positive emotion has also been associated with a reduction in proinflammatory cytokine activity and pain, suggesting that greater attention to emotion regulation may foster resilience.

The discovery of modulating cognitive factors has led to many psychological interventions for RA. Cognitive-behavioral interventions for arthritis management increase self-efficacy, reduce maladaptive coping styles, and decrease helplessness. By changing personal mastery, these therapies reduce negative interpretations of the disease and prevent an accumulation of stress that could further advance disease severity.

Stress and Osteoarthritis

OA is the most common form of arthritis, affecting nearly 30 million individuals in the United States. Symptoms of OA include joint pain, stiffness, and loss of mobility. Affected joints may include the knees, hips, lower back, neck, fingers, and the base of the thumb and big toe. Pain and stiffness is usually the most severe following long periods of inactivity, such as on waking in the morning. The disease is characterized by abnormalities in the cartilage, beginning first with a loss of elasticity. This initial stage is followed by the damage to surrounding bone, including the formation of growths and cysts. These malformations interfere with joint movement and cause pain. In later stages, the synovium may become

inflamed, leading to immune involvement and further cartilage damage and pain.

Although the pathophysiological profile of OA is very different from RA, many OA patients also suffer from considerable pain and physical limitation. OA sufferers complain of many daily stressors: pain, disability, and dependence on others. Fluctuations of OA pain are, like RA, influenced by daily stress. Unlike RA, there is little evidence that stress affects the disease processes.

Cognition and emotion regulation may serve as modulators between physical limitation/pain and consequent distress, however, and a number of factors can increase life satisfaction in those with OA. Greater self-efficacy and strong perceived social support both predict increased life satisfaction. Effective interventions often target these factors. Behavioral and cognitive-behavioral approaches may also help increase activities such as exercise that ease joint pain, increase mobility, and provide psychophysiological stress-buffering effects. Newer emotion-regulation interventions such as mindfulness meditation also show promise as a means of reducing the distress associated with pain episodes. These interventions may enhance psychological well-being, even for those with substantial disease.

A Special Case: Arthritis in Children

Individuals with juvenile arthritis (JA) suffer from complaints very similar to adult RA and OA. As with adult arthritis, there are associations between significant life stress and the development of JA, although a conclusive link cannot yet be drawn. Further, the same cognitive/coping factors seem to play important roles in determining the disease impact on life satisfaction.

Despite similarities, JA may be distinguished from adult arthritis by the added impact of the disease on social and psychological development. Children with JA most often complain of the disease hampering their performance in class, limiting their extracurricular activities, and limiting their activities with peers. Stress resulting from diminished social activities during critical development periods may have an impact on long-term psychological adjustment. Disease severity, however, is not always associated with psychological adjustment, suggesting that some variables play moderating roles. In particular, the child's attitude toward illness (positive or negative) can moderate the relationship between stress and adjustment, buffering the development of depression and strengthening the capacity for resilience. Major tasks for both parents and therapists, then, may be to teach alternative ways to interpret the disease, increase self-efficacy, and further

the development of emotional maturity. These strategies for disease management should be maintained even when the child needs support for the physical tasks in everyday life.

See Also the Following Articles

Arthritis; Autoimmunity; Cytokines; Immune Suppression; Pain.

Further Reading

- Astin, J. A., Beckner, M., Wright, K., et al. (2002). Psychological interventions in rheumatoid arthritis: a meta-analysis of randomized control trials. *Arthritis Care and Research* 47, 291–302.
- Crofford, L. J. (2002). The hypothalamic-pituitary-adrenal axis in the pathogenesis of rheumatic diseases. *Endocrinology and Metabolism Clinics of North America* 31, 1–13.
- Cutolo, M. and Masi, A. T. (2005). Circadian rhythms and arthritis. *Rheumatic Diseases Clinics of North America* 31, 115–129.
- Cutolo, M., Sulli, A., Pizzorni, C., et al. (2003). Hypothalamic-pituitary-adrenocortical and gonadal functions in rheumatoid arthritis. *Annals of the New York Academy of Sciences* 992, 107–117.
- Herrmann, M., Scholmerich, J. and Straub, R. H. (2000). Stress and rheumatic diseases. *Rheumatic Diseases Clinics of North America* 26, 1–27.
- LeBovidge, J. S., Lavigne, J. V., Miller, M. L., et al. (2005). Adjustment to chronic arthritis of childhood: the roles of illness-related stress and attitude toward illness. *Journal of Pediatric Psychology* 30, 273–285.
- McEwen, B. S., Biron, C. A., Brunson, et al. (1997). The role of adrenocorticoids as modulators of immune function in health and disease: neural, endocrine and immune interactions. *Brain Research and Brain Research Review* 23, 79–133.
- Raison, C. L. and Miller, A. H. (2003). When not enough is too much: the role of insufficient glucocorticoid signaling in the pathophysiology of stress-related disorders. *American Journal of Psychiatry* 160, 1554–1565.
- Sharpe, L., Sensky, T. and Allard, S. (2001). The course of depression in recent onset rheumatoid arthritis: the predictive role of disability, illness perceptions, pain and coping. *Journal of Psychosomatic Research* 51, 713–719.
- Straub, R. H., Dhabhar, F. S., Bijlsma, J. W., et al. (2005). How psychological stress via hormones and nerve fibers may exacerbate rheumatoid arthritis. *Arthritis & Rheumatism* 52, 16–26.
- Straub, R. H., Kittner, J. M., Hiejnen, C., et al. (2002). Infusion of epinephrine decrease serum levels of cortisol and 17-hydroxyprogesterone in patients with rheumatoid arthritis. *Journal of Rheumatology* 29, 1659–1664.
- Wilder, R. L. (2002). Neuroimmunoendocrinology of the rheumatic diseases: past, present, and future. *Annals of the New York Academy of Sciences* 966, 13–19.

- Zautra, A. J. (2003). *Emotions, stress, and health*. New York: Oxford University Press.
- Zautra, A. J., Smith, B. W. and Yocum, D. (2002). Psychosocial influences on arthritis-related disease activity. In: Sivik, T., Byne, D., Lipsitt, D. R., Christodoulou, G. N. & Dienstfrey, H. (eds.) *Psycho-neuro-endocrino-immunology (PNEI): a common language for the whole human body*, pp. 201–207. Amsterdam: Elsevier.
- Zautra, A. J., Yocum, D. C., Villanueva, I., et al. (2004). Immune activation and depression in women with rheumatoid arthritis. *Journal of Rheumatology* **31**, 457–463.

Aspergers See: Neurodevelopmental Disorders in Children.

Asthma

A A Kaptein

Leiden University Medical Center, Leiden,
The Netherlands

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
A A Kaptein, volume 1, pp 255–257, © 2000, Elsevier Inc.

Stress as a Cause of Asthma

Stress Brings on Asthma Attacks

Stress Affects the Course of Asthma

Stress Reduction Improves Asthma Outcomes

Glossary

<i>Asthma</i>	Chronic inflammation of the airways, with reversible airway obstruction.
<i>Psycho-analysis</i>	School of thought in psychiatry that emphasizes the role of the unconscious in the etiology and course of psychological and somatic diseases.
<i>Psychology</i>	Study of the behavior of humans.
<i>Self-management</i>	An individual's ability to manage the symptoms, treatment, physical and psychological consequences, and lifestyle changes that are inherent in living with a chronic condition.

Asthma is a respiratory disease, defined as "... a chronic inflammatory disorder of the airways ... the chronic inflammation causes an associated increase in airway hyperresponsiveness that leads to recurrent episodes of wheezing, breathlessness, chest tightness, and coughing. These episodes are usually associated

with widespread but variable airflow obstruction that is often reversible either spontaneously or with treatment." (Clark, 2002). The hallmark of asthma is its reversibility: stimuli, such as stress, elicit episodes of shortness of breath that disappear, either spontaneously or following the administration of medication. Asthma is a highly prevalent disorder: some 10% of the population in industrialized societies is diagnosed with asthma. It now seems that the increase in asthma prevalence has reached its plateau.

Various stimuli lead to asthma attacks in susceptible persons. Stress is one of these stimuli. But there are also others, such as tobacco smoke, cat dander, aspirin, cold air, and exertion. It is the genetic and immunological factors that determine susceptibility to asthma in an individual. Asthma attacks translate into symptoms that lead to limitations in daily functioning, psychological morbidity, and negative social consequences.

Stress has been and still is conceptualized as potentially (1) contributing to the causes of asthma, (2) contributing to the inception of an asthma attack, (3) contributing to the clinical course of asthma, and (4) representing a target for therapeutic interventions in order to reduce the impact of asthma on the daily life of the affected patient. These four topics are discussed below.

Stress as a Cause of Asthma

Asthma is currently viewed as a multicausal disorder; the medical and scientific communities still do not have a perfect understanding of potential causal factors and their interrelations. Patients with asthma report

that stress is one of the major eliciting factors of asthma attacks, and that stress worsens the severity of their asthma attacks. Historically, these clinical observations have often informed past studies on whether stress is a causal factor in the etiology of asthma.

This research took place within a psychoanalytic and psychosomatic paradigm, en vogue in the first decades of the twentieth century. In this paradigm it was assumed that asthma was merely the somatic expression of an underlying source of stress: a psychological conflict within the patient or between the patient and a significant other. Asthma was conceived of as a symptom that symbolically expressed an anxiety neurosis: asthma nervosum could be traced back to the overhearing of the affected child of sexual activities between adults. In psychosomatic views on asthma, specific personality characteristics in combination with a specific interpersonal conflict or stress (centering on dependence vs. autonomy) were seen as the ultimate and main causes of asthma.

Empirical support for these views is not available, and therapeutic strategies following from these views (i.e., psychoanalytic therapy) failed to lead to improvements, with regard to either the disappearance of asthma altogether or reductions in the impact of asthma on the sufferer. However, some of the psychodynamic notions on asthma still seem to have survived. In a recent study on views of general practitioners (GPs) on asthma in various European countries, a GP remarked that "... as children these people very often have overprotective mothers (...) and when they've become adults they often live in circumstances where they do not want to share the air with other people." (Wahlström et al., 2001).

Research on the relation between stress or emotional arousal and changes in pulmonary function continues, but it remains unclear whether stress-induced airway obstruction really exists, let alone whether stress causes asthma. It seems safe to conclude that there is no evidence for the view that stress is an etiological factor in asthma.

Stress Brings on Asthma Attacks

Various stimuli and situations may elicit an attack of asthma in asthma patients. It is difficult for patients themselves, and for their partners and their physicians, to pinpoint exactly the stimulus or stimuli that caused or elicited a particular asthma attack. This, of course, complicates preventive or therapeutic measures. Quite frequently, patients attribute their asthma attacks to stress.

Experimental studies on stress and asthma assign short-term stressful tasks to be performed by healthy controls and people with asthma, in order to examine

their effects on various physiological processes, e.g., pulmonary function. In children with asthma, performing the stressful task of talking into a tape recorder for a 5-min period about a stressful or embarrassing incident led to increases in airway resistance of up to 20%, which is a clinically meaningful change. Observational studies in children and adults with asthma demonstrate that severely negative life events often increase the risk of the occurrence of asthma attacks over the coming few weeks.

Laboratory studies in a small sample of adult patients with asthma show that there is a close connection between the processes of inflammation in asthma and emotional factors, as measured with functional magnetic resonance imaging (fMRI). A recent study demonstrated that neural circuitry underlying stress and emotion can regulate inflammation, and peripheral inflammatory mediators can influence mood and cognition.

Stress Affects the Course of Asthma

As in almost any (chronic) illness, there are great variations in clinical outcome of asthma. Some patients are hardly affected by their asthma; in other patients, asthma appears to completely dominate their lives. A consistent research finding on these observations is that measures of objective severity of asthma are hardly associated with differences in outcome or course of asthma. The stress coping paradigm allows for the explanation of at least some of those differences.

Researchers in Denver, Colorado have developed the psycho-maintenance hypothesis, which pertains to psychological factors maintaining the course and consequences of a (chronic) physical disorder. The researchers examined relationships in hospitalized patients with asthma among objective indicators of asthma severity, psychological characteristics, and outcome measures. Length of hospitalization, frequency of rehospitalization, and severity of medication at discharge turned out not to be determined by objective asthma severity indicators, but by stress-related characteristics. Anxiety (panic-fear) as a personality characteristic, and as a transient state related to dyspnea perception, in combination with various attitudes toward asthma (e.g., optimism, stigma), however, did predict those outcome measures. These findings have been replicated in other settings and other countries.

In addition to stress and attitudes toward asthma, psychiatric morbidity is associated with outcome. Achieving a good control of one's asthma and adequate functional status often appears to be negatively influenced by higher levels of depression,

panic disorder, and social phobia. This relation again is hardly mediated by objective severity of asthma.

Symptom perception in asthma is also related to stress. Patients with higher levels of stress (negative affectivity) associate a wide range of nonspecific symptoms with their asthma, thereby altering the perception of the severity of asthma, which in turn influences their use of medication. These findings have great relevance for the topic discussed in the next paragraph: self-management skills for patients with asthma.

Stress Reduction Improves Asthma Outcomes

Teaching patients how to manage stress associated with asthma has a quite long history. Patients themselves indicate that when they become short of breath, they attempt to relax, which leads to subjective reductions in dyspnea. Behavioral scientists and medical professionals have examined the effects of relaxation training in patients with asthma to a great extent. However, the clinical wisdom of asthma patients has not been supported by empirical research. Relaxation therapies, biofeedback, and breathing retraining do not result in objective improvements in pulmonary function; neither do they affect clinical outcomes such as use of health-care services or medication.

Stress management training, however, appears to produce positive effects. Cognitive-behavioral stress management training leads to clinically meaningful and statistically significant improvements in children and adults with asthma. Cochrane Reviews reveal that medical care that encompasses teaching patients to perceive respiratory symptoms accurately, to respond to impending asthma attacks appropriately, and to cope well with the stress associated with dyspnea episodes and with asthma in daily social circumstances is associated with less disruption of daily activities, a better quality of life, and less use of health-care services.

Recently, two new areas of research pertaining to self-management of asthma have received much attention. The first is studying and examining the effects of expressive writing in adult patients with asthma. Patients have been encouraged to write about the most stressful event of their lives in the experimental condition, and about emotionally neutral topics in the control condition. Clinically meaningful improvements in pulmonary function were observed in the experimental condition. The second new area is research on illness representations, which pertains to studying the cognition of asthma patients of their respiratory condition and its treatment in relation to various outcome measures. Patients' own

ideas, objectively correct or not, turned out to be predictors of the clinical course of asthma. Future research will make clear to what extent these two new topics will be integrated into stress management programs for patients with asthma. Another challenge is to ensure that effective components of self-management training programs will be incorporated into guidelines for health-care providers and applied for the patient's benefit.

See Also the Following Articles

Aerobic Exercise and Stress Reduction; Relaxation Techniques; Somatic Disorders; Stress Management, CAM Approach.

Further Reading

- Chrousos, G. P. (2000). Stress, chronic inflammation, and emotional and physical well-being: concurrent effects and chronic sequelae. *Journal of Allergy and Clinical Immunology* 106, S275–S291.
- Clark, J. H. (chair) (2002). *Global strategy for asthma management and prevention*. Bethesda, MD., NIH, National Heart, Lung, and Blood Institute.
- Gibson, P. G., Coughlan, J., Wilson, A. J., et al. (2000). Self-management education and regular practitioner review for adults with asthma (Cochrane review). *The Cochrane Library* 4. Oxford: Update Software.
- Kaptein, A. A. and Creer, T. L. (eds.) (2002). *Respiratory disorders and behavioral medicine*. London, UK: Martin Dunitz Publishers.
- Kaptein, A. A., Scharloo, M., Helder, D. I., et al. (2003). Representations of chronic illnesses. In: Cameron, L. D. & Leventhal, H. (eds.) *The self-regulation of health and illness behavior*, pp. 97–118. New York and London: Routledge.
- Kinsman, R. A., Dirks, J. F. and Jones, N. F. (1982). Psychomaintenance of chronic physical illness. In: Millon, T., Green, C. & Meagher, R. (eds.) *Handbook of clinical health psychology*, pp. 435–466. New York: Plenum Press.
- Lavoie, K. L., Cartier, A., Labrecque, M., et al. (2005). Are psychiatric disorders associated with worse asthma control and quality of life in asthma patients? *Respiratory Medicine* 99, 1249–1257.
- Main, J., Moss-Morris, R., Booth, R., Kaptein, A. A. and Kolbe, J. (2003). The use of reliever medication in asthma: the role of negative mood and symptom reports. *Journal of Asthma* 40, 357–365.
- McQuaid, E. L., Fritz, G. K., Nassau, J. H., et al. (2000). Stress and airway resistance in children with asthma. *Journal of Psychosomatic Research* 49, 239–245.
- Rietveld, S., Everaerd, W. and Creer, T. L. (2000). Stress-induced asthma: a review of research and potential mechanisms. *Clinical and Experimental Allergy* 30, 1058–1066.

- Ritz, T. and Roth, W. T. (2003). Behavioral interventions in asthma. Breathing retraining. *Behavior Modification* 27, 710–730.
- Rosenkranz, M. A., Busse, W. W., Johnstone, T., et al. (2005). Neural circuitry underlying the interaction between emotion and asthma symptom exacerbation. *Proceedings of the National Academy of Sciences USA* 102, 13319–13324.
- Sandberg, S., Paton, J. Y., Ahola, S., et al. (2000). The role of acute and chronic stress in asthma attacks in children. *Lancet* 356, 982–987.
- Smyth, J. M., Stone, A. A., Hurewitz, A. and Kaell, A. (1999). Effects of writing about stressful experiences on symptom reduction in patients with asthma or rheumatoid arthritis. *Journal of the American Medical Association* 281, 1304–1309.
- Wahlström, R., Lagerlöv, P., Stålsby Lundborg, C., et al. (2001). Variations in general practitioners' views of asthma management in four European countries. *Social Science and Medicine* 53, 507–518.
- Wolf, F. M., Guevera, J. P., Grum, C. M., Clark, N. M. and Cates, C. J. (2003). Educational interventions for asthma in children (Cochrane review). *The Cochrane Library* 2. Oxford: Update Software.
- Wright, R. J., Rodriguez, M. and Cohen, S. (1998). Review of psychosocial stress and asthma: an integrated biopsychosocial approach. *Thorax* 53, 1066–1074.

Atherosclerosis

M W Ketterer

Wayne State University, Detroit, MI, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M W Ketterer and M Randall, volume 1, pp 258–261, © 2000, Elsevier Inc.

Natural History
Technical Limitations
Putative Mechanisms
Epidemiological Data
Treatment Data

Glossary

<i>Adhesion/ aggregation</i>	The process by which active platelets connect with vessel walls and each other to form a thrombus.
<i>Angina</i>	Pain or pressure generally occurring in the chest or abdomen, sometimes radiating to the left arm, believed to be caused by ischemia of the myocardium. Equivalent symptoms may include dyspnea, dizziness, weakness, and palpitations.
<i>Angioplasty</i>	The introduction of a small balloon via a guidewire to an occluded segment of an artery followed by one or more inflations to reduce the narrowing caused by a blockage (also termed percutaneous transluminal coronary angioplasty, or PTCA).
<i>Arrhythmia</i>	Disorders of the normal sequencing of muscle contraction of the heart leading to inefficiency or, in some types, cessation of blood flow.

<i>Claudication</i>	Pain and weakness in the legs when walking caused by atherosclerosis of the arteries feeding the legs.
<i>Collagen fibrils</i>	The smallest structure of collagen (supportive tissue) formed by assembling molecules.
<i>Infarction</i>	Cessation of blood flow by thrombus formation, causing tissue death.
<i>Ischemia</i>	An imbalance of oxygen flow/availability and tissue need; usually due to obstruction by atherosclerosis with a superimposed thrombus.
<i>Perfusion</i>	The passage of a substance (e.g., oxygen) into tissues.
<i>Risk factors</i>	Behaviors or characteristics that have been found to be epidemiologically related to the development of a condition.
<i>Smooth muscle cells</i>	Cells found in the tunica media (middle layer) of arteries that proliferate with injury.
<i>Sudden cardiac death</i>	Death occurring within 1 h of symptom onset. Almost always attributable to atherosclerotic disease complicated by thrombus formation.
<i>Syndrome X</i>	The occurrence of angina-like symptoms in the absence of atherosclerotic disease or ischemia.
<i>Transient ischemic attack</i>	A brief period of dizziness, confusion, or cognitive dysfunction caused by ischemia of the brain.

Natural History

Atherosclerosis is an occlusive disease of the arterial walls that may occur at multiple sites in the body, but is most malignant when it threatens blood flow to critical organs such as the heart or brain. By itself,

atherosclerosis is of no clinical significance. But when atherosclerosis obstructs blood flow either transiently (ischemia) or long enough to kill tissue (infarction), it can cause transient dysfunction or permanent organ damage.

From a population standpoint, the most common clinical manifestation of atherosclerosis occurs when the heart is affected (angina, myocardial infarction, and/or sudden cardiac death). Atherosclerosis of major arteries feeding the brain can cause transient ischemic attack, multi-infarct dementia, or large, localized cerebrovascular accidents, while that of the legs can cause intermittent claudication and necrosis/gangrene and that of the kidney can cause renal dysfunction or failure.

While the initiation of atherosclerotic blockages, or plaques, is believed to occur throughout the life span, in at least some individuals initiation can occur as early as infancy in the form of fatty streaks. By the time a population cohort reaches late adolescence or early adulthood (e.g., young men killed in combat in Korea or Viet Nam), about one-third of young men display some degree of occlusion at postmortem, and a few percent have clinically significant occlusion (luminal blockages of 50% or greater in one of the major epicardial arteries of the heart, or any lesion of the left main segment). In middle age, most individuals have at least some low-level atherosclerotic disease. But women remain less at risk until menopause, when their risk begins to rise, matching men about 10 years later. However, even at advanced ages of 80 or greater, some people have coronary arteries entirely free of atherosclerosis.

In some individuals, at some sites, the myocardium has the ability to grow collateral arteries around occluded sites (perhaps provoked by exercise), which then serve as natural bypasses in the event of abrupt closure of the main trunk. Characteristics epidemiologically associated with the development or aggravation of atherosclerosis are referred to as risk factors. Age (older), sex (male), smoking (more), hypercholesterolemia, family history of early heart disease (early), hypertension, diabetes, and sedentariness are generally accepted risk factors for atherosclerosis.

Initiation of atherosclerosis occurs when the endothelium (a one-cell-thick lining of the arterial wall) is damaged. Multiple mechanisms have been suggested for this event and are not necessarily mutually exclusive. Mechanical strain, chemical insult, and infectious/inflammatory responses are thought to play a role. The cellular repair mechanisms result initially in a fatty streak. Recurrent insult to the site provokes migration of smooth muscle cells and collagen fibrils, resulting in a fibrous plaque. Continued insult and

attempted healing can produce an unstable plaque with a thin cap covering dead cells and a lipid-rich medium. Rupture of such a plaque, followed by platelet adhesion/aggregation and thus thrombus formation, is thought to be the sequence resulting in 85–90% of all myocardial infarctions. Vasomotor events (widespread constriction or localized spasm of arterial segments) may result in increased myocardial workload, plaque rupture, and/or superimposed thrombotic events, causing ischemia, which may trigger life-threatening arrhythmias for other patients.

The biochemical signals provoking each step of this process are still a matter of debate and investigation, but the validation of these theories is always to be found in intervention trials targeting the process but measuring the clinically important outcome. The role of platelet aggregation, for example, found its strongest validation in the powerful effect of antiplatelet substances (e.g., aspirin) in preventing cardiac events.

The significance of a plaque is determined primarily by the size of myocardium dependent upon its blood flow. Thus, plaques more proximal in the vascular tree or bush will produce greater muscle loss if abruptly rupturing and occluding. For 90% of patients (those who are left dominant), the left main coronary segment feeds both the left anterior descending and left circumflex arteries and thus carries higher mortality risk than more distal lesions with less territory at risk. Variants in coronary anatomy (e.g., right versus left dominance) determine the size of muscle affected.

Technical Limitations

The ability to study atherosclerosis is limited by the available technology. Firm measurement is most accurately obtained by postmortem examination of the arteries. Short of this, coronary angiography in living subjects is considered the gold standard. But even angiography has limitations:

- lesions that are low grade or tubular or longitudinal in shape are easily missed by the cardiologist's visual rating of the film;
- the cost and risk of the procedure means that the ethical justification for studying normals is difficult;
- asymptomatic coronary artery disease (CAD) patients will rarely show in cath labs;
- no distinction can be made between fibrous and unstable plaque (unless intravascular ultrasound is used); and
- one-third of cardiac patients whose first symptom is sudden death will not appear in angiographic samples.

Helical computerized tomography (CT) or magnetic resonance imaging (MRI) scanning has shown promise as a noninvasive means of imaging the larger proximal segments of the left anterior descending and left circumflex arteries, but provides less resolution of the right coronary artery. Indirect measures include the provocation of ischemia (by physical exertion, sympathomimetic agents, and/or mental stress) and subsequent measurement by electrocardiographic changes, or fixed (presumable permanently dead tissue) and reversible (transiently ischemic) myocardial wall motion changes (measured on echocardiograms or, more commonly, blood-pool-tagged radionuclide imaging) or radionuclide-tagged perfusion defects of the myocardial walls.

It should be emphasized that symptoms (dizziness, palpitations, dyspnea, or chest pain/pressure) either at rest or on provocation are only weakly linked with all presently available measures of ischemia, even in patients with documented coronary atherosclerosis. In patients with high risk factor loading, chest pain/pressure or dyspnea reliably provoked by exertion and relieved by rest or nitroglycerin may be the only symptom pattern indicative of ischemia. Crushing, substernal chest pain/pressure is considered classic for a myocardial infarction but only occurs in about 50% of such events.

Putative Mechanisms

The means by which psychosocial/emotional stress may cause or aggravate atherosclerosis are multiple. First, stress is clearly a major factor in diminished smoking cessation/abstinence and noncompliance with pill taking, dietary restrictions (cholesterol, sugar, or salt), and exercise. The impact of Zyban/Wellbutrin, a previously established antidepressant, on smoking cessation confirms this. Similar randomly-assigned, controlled clinical intervention trials have not yet been done for the other compliance mechanisms, but epidemiological evidence implicates emotional distress as a strong determinant.

Recent thinking implicates inflammation as a cause of plaque progression and instability. Inflammatory markers (e.g., C-reactive protein, interleukin-6) covary with depression, but we do not yet know if treatment of depression reduces such markers. To the extent that inflammation is caused by chronic, localized, smoldering infections (e.g., *Chlamydia pneumoniae*), depression-induced immunosuppression may enhance plaque growth/instability.

The best studied putative mechanism lies in acute or chronic emotional arousal provoking peripheral vasoconstriction (thus increasing myocardial

workload and oxygen demand) or vasospasm/constriction (and thus endothelial tearing or plaque rupture of susceptible lesions). It is clear that about one-half of patients with previously documented coronary atherosclerosis will demonstrate ischemia when mentally stressed (most reliably by verbal confrontations eliciting anger). Because the overwhelming majority of ischemic episodes on Holter exams occur in the absence of physical exertion (as indicated by heart rate and patient report), it is plausible that mental stress is a common cause of most ischemic episodes.

For the clinician, the question of mechanisms may be entirely moot since the benefits of diagnosis and treatment would accrue regardless of the mechanism, and the quality-of-life benefit alone justifies treatment.

Epidemiological Data

About two dozen studies have attempted to find cross-sectional associations between psychosocial/emotional stress and atherosclerotic severity in angiographic samples. The fatal flaw in these studies has been the failure to appreciate that the controls (i.e., patients whose results are normal or subclinical) used in these studies are physically symptomatic individuals with sufficient risk factor loadings to provoke concern. The occurrence of physical symptoms in the absence of disease implies that most of these control patients have high levels of somatized emotional distress (most notably panic disorder or depression, and referred to as atypical chest pain or syndrome X). Thus, the normals are not representative of a general or healthy population and are not appropriate as controls in studies intended to address etiology.

Only two studies have examined an angiographic sample while obtaining a healthy age/sex/socioeconomic status matched control group. Ketterer et al. recruited 176 males undergoing coronary catheterization with no prior history of revascularization and 56 males who had no history of atherosclerotic disease at any site. Traditional risk factors and multiple measures of psychosocial/emotional distress were quantified. In univariate analyses, several psychosocial measures (more depression as rated by a spouse or friend, more denial of depression as measured by discrepancies in spouse/friend minus patient ratings, more unprovoked nocturnal awakening, but lesser frequency of suicidal ideation) were found to be associated with atherosclerotic severity. This pattern of stigmatized symptoms (frequency of suicidal ideation) having a negative relationship with CAD severity, while spouse/friend-rated depression and a symptom with norms unknown to the layman (frequency of nocturnal awakening) having a positive

relationship with CAD severity suggests that denial is a major confound in measuring distress in this population. Indeed, the strongest correlate of CAD severity was denial of depression (measured by spouse/friend minus self-ratings on parallel versions of the same scale). This pattern unavoidably invalidates results obtained only from patient self-report measures of emotional distress, since the apparent accuracy of self-report data is not only weakened by denial, but also confounded with atherosclerotic disease severity. McDermott et al. have found anger-hostility to covary with disease severity in a similarly controlled study.

Angiographic study of atherosclerotic progression is limited by the ethical constraints imposed by an invasive procedure and costs. Several attempts to study coronary atherosclerotic progression exist in a patient population for whom clinical need (i.e., increased symptoms) existed. These reports found increased progression in patients with more type A behavior/hostility/overcommitment. Because we do not know the status of the arteries of patients in the original sample who did not undergo repeat angiography, nor do we know whether intertest angioplasty occurred, which might affect the second reading, it is doubtful that these studies provide much insight. However, repeat carotid ultrasonography (a noninvasive procedure not requiring clinical justification) also indicates progression in emotionally distressed patients.

Treatment Data

Ornish attempted an aggressive lifestyle modification program in 28 patients who underwent repeat angiography and 20 controls. Patients in the intervention group displayed modest (5%) regression of some plaques and were less likely to display progression than patients in a randomly assigned, usual care control group. Because the intervention had three prongs (vegetarian diet, daily exercise, and stress management), it is not possible to unambiguously ascribe the results to stress management alone. Ornish has not published data regarding hard clinical outcomes (myocardial infarction or death), although at least one report indicates decreased chest pain. Subsequent studies have found that lipid-lowering agents can also cause this sort of modest regression.

Now that noninvasive testing is possible (e.g., helical CT scans), it should be possible to do more and better epidemiological and treatment studies.

See Also the Following Articles

Depression and Coronary Heart Disease; Heart Disease/Attack; Heart Failure, Stress Effects; Cardiovascular Disease, Stress and.

Further Reading

- Cannon, R. O., Camici, P. G. and Epstein, S. E. (1992). Pathophysiological dilemma of "syndrome x". *Circulation* 85(3), 883–892.
- Davies, M. J. and Woolf, N. (1993). Atherosclerosis: what is it and why does it occur? *British Heart Journal* 69(1), 3–11.
- Falk, E. (1992). Why do plaques rupture? *Circulation* 86(Supplement III), 30–42.
- Goodman, M., Quigley, J., Moran, G., Meilman, H. and Sherman, M. (1996). Hostility predicts restenosis after percutaneous transluminal coronary angioplasty. *Mayo Clinic Proceedings* 71(8), 729–734.
- Joksimovic, L., Siegrist, J., Meyer-Hammer, M., et al. (1999). Overcommitment predicts restenosis after coronary angioplasty in cardiac patients. *International Journal of Behavioral Medicine* 6(4), 356–369.
- Julkunen, J., Salonen, R., Kaplan, G. A., Chesney, M. A. and Salonen, J. T. (1994). Hostility and the progression of carotid atherosclerosis. *Psychosomatic Medicine* 56, 519–525.
- Kaplan, J. R., Manuck, S. B., Clarkson, T. B., et al. (1983). Social stress and atherosclerosis in normocholesterlemic monkeys. *Science* 220, 733–735.
- Ketterer, M. W., Brymer, J., Rhoads, K., et al. (1996). Emotional distress among males with "syndrome x". *Journal of Behavioral Medicine* 19(5), 455–466.
- Ketterer, M. W., Kenyon, L., Foley, B. A., et al. (1996). Denial of depression as an independent correlate of coronary artery disease. *Journal of Health and Psychology* 1(1), 93–105.
- Krantz, D. S., Sanmarco, M. I., Selvester, R. H. and Matthews, K. A. (1979). Psychological correlates of progression of atherosclerosis in men. *Psychosomatic Medicine* 41(6), 467–475.
- Mayou, R. (1989). Atypical chest pain. *Journal of Psychosomatic Research* 33(4), 393–406.
- McDermott, M. R., Ramsay, J. M. and Bray, C. (2001). Components of the anger-hostility complex as risk factors for coronary artery disease severity: a multi-measure study. *Journal of Health Psychology* 6(3), 309–319.
- Ornish, D., Brown, S. E., Scherwitz, L., et al. (1990). Can lifestyle changes reverse coronary heart disease? *Lancet* 336(8708), 129–133.
- Pearson, T. A. (1984). Coronary angiography in the study of coronary heart disease. *Epidemiological Reviews* 6, 140–166.
- Pearson, T. A. and Heiss, G. (1993). Atherosclerosis: quantitative imaging, risk factors, prevalence and change. *Circulation* 87(Supplement 3), 1–82.
- Pickering, T. G. (1985). Should studies of patients undergoing coronary angiography be used to evaluate the role of behavioral risk factors for coronary heart disease? *Journal of Behavioral Medicine* 8, 203–211.
- Ross, R. (1999). Atherosclerosis – an inflammatory disease. *New England Journal of Medicine* 340(2), 115–126.
- Whiteman, M. C., Deary, I. J. and Fowkes, G. R. (2000). Personality and social predictors of atherosclerotic progression: Edinburgh artery study. *Psychosomatic Medicine* 62, 703–714.

Atrial Natriuretic Peptide (ANP) *See: Vasoactive Peptides.*

Attention-Deficit/Hyperactivity Disorder, Stress and

L E Arnold

Ohio State University, Columbus, OH, USA

R L Lindsay

Arizona Child Study Center, Phoenix, AZ, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
L Eugene Arnold, volume 1, pp 262–267, © 2000, Elsevier Inc.

Definition and Overview

Aggravation, Complication, or Mimicking of ADHD

Symptoms by Stress

ADHD Symptoms as Stressors

Conclusion

Glossary¹

<i>Amphetamine</i>	Generic name for a stimulant medicine approved by the Food and Drug Administration (FDA) for treatment of attention-deficit/hyperactivity disorder (ADHD). It has two main brands: Adderall and Dexedrine. It is available in immediate-release tablets and extended-release encapsulated beads of different durations (Adderall XR and Dexedrine Spansule).
<i>Atomoxetine</i>	Generic name for a nonstimulant medicine approved by the FDA for treatment of ADHD. The brand name is Strattera. It is available only as an encapsulated powder.
<i>Clonidine</i>	A blood pressure medicine often used in the treatment of ADHD, especially with co-occurring tics or sleep problems. The brand name is Catapres. It is not approved by the FDA for treatment of ADHD but is commonly used, often with one of the FDA-approved drugs. It has been replaced somewhat by guanfacine, a similar drug with longer duration and less sedation.

*Combined type
ADHD*

Comorbid

*Conduct
disorder (CD)*

Desipramine

*Disruptive
behavior
disorders
(DBD)*

Guanfacine

*Hyperkinetic
disorder*

Impairment

*Methyl-
phenidate*

*Operational
criteria*

ADHD with both six or more symptoms of inattentiveness and six or more symptoms of hyperactivity and impulsivity.

Pertaining to other disorders co-occurring with ADHD – i.e., in the same person at the same time

A form of disruptive behavior disorder typified by a repetitive and persistent pattern of behavior in which the basic rights of others or major age-appropriate societal norms and rules are violated. Referred to by the juvenile justice system as delinquency.

Antidepressant that has been shown effective for ADHD in controlled studies. The brand name is Norpramin. Used more in adults than children. Not FDA-approved for ADHD, but accepted by experts.

An array of psychiatric disorders marked by aggression, hostility, rule breaking, defiance of authority, and violation of social norms. The primary forms are oppositional defiant disorder (ODD) and conduct disorder (CD). Some authors include ADHD as a form of DBD.

A blood pressure medicine often used in ADHD, especially with tics. The brand name is Tenex. Although not approved by the FDA for this use, it is supported by a controlled study.

An old term for ADHD in the United States that is still used in Europe. Hyperkinetic literally means overactive.

Interference with normal function or performance.

Generic name for a stimulant medicine approved by the FDA for treatment of ADHD. It has several brands: Ritalin, Concerta, Metadate, and Methylin. It is available in immediate-release tablets and in several long-acting forms, including encapsulated beads of differing durations (Ritalin LA and Metadate CD) and an oral osmotic tablet (Concerta).

Steps that one can go through to see if experts would recognize a patient as having the disorder. If the patient meets all the operational criteria, one can be

¹Adapted with permission from Arnold, L. E. (2004). *A Family Guide to Attention-Deficit/Hyperactivity Disorder*. Newtown, PA: Handbooks in Health Care. Copyright 2004, Handbooks in Health Care.

	sure that clinicians anywhere who are familiar with the disorder would agree with the diagnosis.
<i>Oppositional defiant disorder (ODD)</i>	A form of disruptive behavior disorder typified by a pattern of negativistic, hostile, and defiant behavior.
<i>Pemoline</i>	Generic name for a stimulant medicine approved by the FDA for treatment of ADHD. The brand name is Cylert. It is not recommended as a first or second choice because of rare life-threatening liver damage.
<i>Posttraumatic stress disorder (PTSD)</i>	A form of anxiety disorder in which a person who was exposed to a traumatic event persistently re-experiences the event, avoids things associated with the event, and has persistent arousal that was not present before the trauma.

Definition and Overview

Attention-deficit/hyperactivity disorder (ADHD) is a common disorder of cognition and behavior originating in childhood. According to the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (DSM-IV), diagnosis requires, in addition to sufficient symptoms, that some of the symptoms must have caused impairment before age 7, some impairment is present in two or more settings, and there must be evidence of significant impairment in social, academic, or occupational functioning.

For the inattentive type, there must be six of the following nine symptoms:

1. Often failing to pay attention to details or making careless mistakes
2. Frequent difficulties sustaining attention
3. Often not seeming to listen
4. Often failing to follow instructions or failing to finish things
5. Frequent difficulty organizing
6. Often avoiding tasks that require sustained mental effort
7. Often losing necessary things for tasks
8. Frequent easy distraction
9. Frequent forgetfulness

For diagnosis of the hyperactive-impulsive type, six of the following nine symptoms are necessary:

1. Frequent fidgeting or squirming
2. Often leaving seat
3. Often running or climbing excessively
4. Frequent difficulty playing quietly
5. Often acting as if driven by a motor
6. Often talking excessively
7. Often blurting out answers before questions have been completed

8. Frequent difficulty awaiting turns
9. Often interrupting or intruding

For the combined type, the full-blown syndrome, there must be six symptoms from each cluster.

ADHD itself is not considered a stress disorder. However, it is related to stress in a number of ways. As with most disorders of health, both physical and mental, the symptoms can be aggravated by stress. Many of them can be mimicked by stress. Further, the symptoms of the disorder, especially if untreated, can cause stress both in the individual afflicted and in family members, especially parents, and in teachers, other caretakers, and peers. Later in life, residual symptoms of ADHD can stress spouses or significant others, employers, and co-workers. Finally, adults with residual ADHD can stress their children, whether or not the children also have ADHD. The following description assumes the untreated natural course of the full-blown disorder (combined type).

Aggravation, Complication, or Mimicking of ADHD Symptoms by Stress

Individuals with ADHD have a more tenuous functioning capacity than normal. They are more impulsive, less organized, less attentive to detail, less patient, more easily frustrated. They are also more likely to have additional problems, such as anxiety (about 30%), depression, or learning disorders (20–25%), which further compromise their functional abilities. They often have less adequate verbal memory or other perceptual, associative, or mnemonic ability. Some authorities believe that they are less easily rewarded by ordinary daily activities that normal people find sufficiently rewarding to maintain attention and interest, resulting in frequent boredom and lack of satisfaction. All this adds up to less coping ability. Consequently, the amount of stress needed to induce stress symptoms in these individuals is less than for other people. Since stress by itself can induce concentration and memory problems, it can add an extra increment of impairment to the attention and memory problems of ADHD. Similarly, stress-induced anxiety can add to the restlessness of ADHD, and impulse control often deteriorates further with stress.

Children who are suffering stress symptoms may appear to present with ADHD even if they do not have it. The hypervigilance, anxiety, restless sleep, and preoccupied inattention of posttraumatic stress disorder (PTSD) can mimic the distractibility, restlessness, and inattentiveness of ADHD. In a study of 117 children who appeared before a juvenile/family court secondary to experiencing significant child abuse and/or trauma, children with PTSD demonstrated concurrent ADHD, but not oppositional defiant disorder (ODD)/conduct disorder (CD). Indeed, children presenting

with ADHD symptoms should be routinely assessed for traumatic experiences. There may be a bidirectional association between child maltreatment and ADHD/ODD, with head trauma from abuse aggravating or even causing ADHD, and ADHD symptoms inviting abuse.

The anxiety, misbehavior, withdrawal, depression, preoccupation, easy frustration, and impaired performance of an adjustment disorder can also mimic the inattentiveness, misbehavior, impaired performance, and restlessness of ADHD. The agitation, withdrawal, impaired concentration, and irritability of stress-induced depression can mimic the restlessness, inattentiveness, and impulsiveness of ADHD. Indeed, the final DSM-IV criterion for diagnosing ADHD is that the symptoms are not better explained by another mental disorder.

ADHD Symptoms as Stressors

Untreated symptoms of ADHD stress the persons who have the disorder, their family – parents, siblings, spouse, and children – teachers, employers, and peers.

ADHD as a Stressor of the Person Afflicted

Children or adults with ADHD, though of normal or even superior intelligence, find it difficult to achieve at a level commensurate with their ability. They have to try harder to achieve the same results as peers of the same intelligence. The impulsiveness of the disorder may repeatedly get either children or adults into trouble before they realize what has happened.

For example, a child with inattention may have to struggle for an hour or two to complete the same homework that a peer of normal attentional focus can complete in a half hour. Even then, he or she may forget to turn it in on time and thus not receive full credit for it. Similarly, inattention can interfere with sports performance, a source of the greatest interest and pride for many children, especially boys, who outnumber girls with the disorder two or three to one. Inattentively missing the catch of an easy pop fly, making impulsive mistakes, or failing to hear or carry out the coach's instruction, thus incurring the contempt of peers and coach, or being chosen 18th when sides are chosen, can all stress one's self-esteem. Even at home, children with ADHD may find their achievement and behavior compared unfavorably with those of siblings of equal potential. They may be acutely aware of parental (or teacher) disappointment in their performance or behavior and conclude that they are not loved, at least not as much as siblings or classmates. Needing more prompting by teachers (and parents) for daily function, they may feel picked

on, nagged, and alienated from the usual sources of caregiver support. Thus, the developmental tasks of childhood, including acquisition of academic and social skills, basic skills of civilization, work habits, friendships, secure family relations, autonomous functioning, self-control, and self-confidence, are accomplished – if at all – at a more stressful price in the presence of ADHD.

In adolescence and young adulthood, the inattentiveness and impulsiveness of ADHD lead to poorer driving records than those of age mates (two to four times as many accidents). Parents are thus understandably reluctant to sign for a driving permit as early as the adolescent wishes, and are reluctant to allow use of the family car once a driver's license is obtained. Given the importance of a car for peer esteem and adolescent and young adult social life, this parental reluctance aggravates any family conflict, further cripples peer relations, and adds further increments of psychological stress. The impulsiveness and disorganization of the disorder also lead to a higher rate of teenage pregnancy in girls with ADHD and probably in girlfriends of boys with ADHD, not to mention greater exposure to sexually transmitted diseases. After all, precautions (including abstinence) require some reflection, organization, and impulse control. Thus, the interaction of ADHD with developmental issues of adolescence compound the ADHD-induced stresses from earlier childhood.

Adults with inattention become bored or frustrated with many jobs that others find tolerable or even interesting. They may wallow miserably in a lower-skilled job than their abilities deserve. Conversely, they may have an erratic work history, as their dissatisfaction and short attention span lead them to seek relief in job changes or even to impulsively quit without another job lined up. Employer dissatisfaction with poor attention to detail, restlessness, or failure to follow instructions may also lead to short tenures on many jobs. Thus, occupation, which should be one of the supports of self-esteem and satisfaction, may become an additional stressor economically and psychologically in the presence of ADHD. Similar patterns of frustration, boredom, restlessness, and difficulty sustaining commitment may dog social and intimate relationships of those with ADHD, who have lower rates of successful marriage. Failed love, separation, and divorce are well-documented stressors for anyone.

ADHD as a Stressor of Parents and Teachers

The constant activity, easy frustration, irritability, disorganization, tendency to lose things, and difficulty following instructions found in children with

ADHD can stress any adult attempting to manage their behavior, structure their activities, or teach them. It is particularly frustrating to teachers when they recognize that the child has the intellectual ability but is not doing the work. Many teachers agonize over the wasted potential, or even blame themselves for the child's poor performance. An unsuccessful intelligent pupil is stressful to the teacher's self-image. Erroneous conclusions that the child is lazy or obstinate can poison the teacher-pupil relationship and stress both teacher and child. This stress can possibly spill over into the teacher-parent relationship and affect all three parties.

Parents are stressed by similar disappointments and child performance difficulties at home, which is less intense than at school, but more consistent and enduring (every day, every year). Neglected chores, disorganized messiness, impulsive accidents, poor judgment, hectic overactivity and intrusiveness, conflicts with siblings, undone homework, calls from school, poor report cards, and other stresses are not what most parents anticipate at their child's birth. Such disappointments make it hard for parents to love (or at least hard to like) the child, leading to feeling guilty for being a rejecting parent. Criticisms of the child's behavior and the parents' child-rearing practices by well-meaning friends and relatives may add to the parents' feeling of failure. The child's impulsive destructiveness and accident proneness, as well as chronic need for professional services, round out the picture of a high-maintenance child who stresses the parents economically, physically, and psychologically.

Research has confirmed the profound impact of the problems of the ADHD child on the parents. In a well-controlled study of parent-child interactions, the placebo condition showed the mothers of boys with ADHD to be very directive, commanding, and negative with the child, with little positive interaction; when a dose of stimulant medicine was given to alleviate the child's ADHD symptoms, the same mothers became much friendlier and more respectful of the child, allowed the child more autonomy, and gave many fewer negative messages to the child. This demonstrated that the negative parental attitude often seen toward ADHD children is not part of the parents' personality, but a response to the child's problems. ADHD tends to bring out the worst in those around the afflicted person.

The negativity often spreads into the marital relationship as the frustrated parents disagree about how to handle the child's problems. For example, the father may blame the mother because he can get the child to behave for short times (sometimes by heavy-handed means) and the mother cannot get the child to behave during the much longer intervals she is

responsible for. The mother may feel that the father's disciplinary approach is abusive. Each may blame the other's genes. One parent may wish to medicate the child while the other objects to this. One parent may be enthusiastic about behavior modification but may find the system sabotaged by the other parent's failure to cooperate. In some cases one parent does not even recognize or admit there is a problem, while the other parent struggles with numerous complaints from school. One indicator of the marital stress engendered by a child with ADHD is the higher divorce rates reported for parents of such children. (This may be a partial cause as well as effect: divorce is stressful to children, and stress may nudge a subclinical case over the diagnostic threshold by aggravating otherwise mild symptoms.)

ADHD as a Stressor of Peers

Siblings, classmates, or teammates of children with ADHD are often unsung victims of the afflicted child's symptoms. The afflicted child's (proband's) needs for attention, supervision, and services tend to drain parent, teacher, and coach resources from siblings, classmates, and teammates. This relative neglect can sometimes be carried to the point of actual neglect, as all family efforts are focused on the impaired member. The whole family may have to forego outings that the proband cannot handle, or may cater to the proband, either out of pity or out of fear of impulsive tantrums. Older siblings' social life may be impacted by the need to monitor the proband, by embarrassment, or by the proband's intrusiveness. Younger siblings may be steamrolled by the proband's overactivity and inattentive thoughtlessness, or even be endangered by the proband's reckless impulsiveness.

Classmates or teammates may find their education, projects, activities, or games disrupted by the hyperactive child's impulsive intrusiveness, failure to wait a turn, failure to carry out an assigned part of the team's effort, need for excessive attention, failure to finish the activity, impulsive arguing, irritability, or other failure to be a good team player. It should be no surprise that children with ADHD typically fare poorly in peer sociometric ratings. Tragically, the hyperactive child's normal wish for friendships, implemented by clumsy, intrusive, impatient attempts to impose friendships, often only makes the situation worse. Thus, in the presence of ADHD, the normal need for peer relations stresses both the hyperactive child and available peers.

When ADHD persists into adulthood, the brunt of peer difficulties falls on the significant other, who may have to endure financial uncertainty as well as a disorganized lifestyle and interpersonal inconsistency and

unpredictability. Again, the high rate of failed marriages and other relationships stresses both the direct victim of ADHD and the indirect peer victims. Employers and co-workers also suffer some stress from the ADHD worker's inconsistent performance, restlessness, disruptiveness, failure to listen to or follow instructions, and failure to carry out assigned parts of a project. In some cases the need for the unpleasant task of firing the ADHD worker adds to the employer's stress. Even when the ADHD employee quits voluntarily, there is the stress of finding and training a replacement.

ADHD as Stressor of the Hyperactive Adult's Children

An important but as yet largely unstudied area is the stress on children of having parents with ADHD. Theoretically, parenting characterized by inattentiveness, inconsistency, impatience, impulsiveness, and disorganization should be stressful for any child. Since ADHD is highly heritable, an adult with ADHD has a good chance of having a child with ADHD. Because ADHD children have a special need for parental support and structure, we might expect that they would be especially vulnerable to the inattentive, unstructured disorganization of an ADHD parent. Conversely, what parenting skills the ADHD parent has will be challenged by the additional stress of a child with ADHD, in a vicious cycle. Thus both parent and child would be stressed by their mutual handicap.

ADHD Treatment as Stressor

Despite considerable evidence as to the effectiveness of pharmacotherapy in ADHD, there are considerable misgivings about the increasingly widespread use of medications, which may contribute further to the stresses of living with ADHD or with a child with ADHD. Parents (and children) are bombarded by media messages stating that medications are harmful or unsafe and can lead to abuse (of the medication or of street drugs), and that ADHD is a myth created by intolerant teachers and parents in an effort to place a pharmacologic leash on the child who only requires more discipline.

Any medication strong enough to treat ADHD is also strong enough to harm. One of the stresses of having ADHD or a child with ADHD is the worry about the long-term effect of medications, which is aggravated by media sensationalism. Sudden death due to medications is a recurring concern. In the 1980s this was reported in several patients taking desipramine, a tricyclic antidepressant known to cause cardiac slowing or blocking and that had demonstrated benefit in ADHD. In the 1990s the

combination of methylphenidate and clonidine (an antihypertensive) was reported to have contributed to the deaths of four children, all of whom had other complications such as preexisting cardiac disease (two cases), recent general anesthesia, or other medications. In 2005, Health Canada withdrew Adderall XR (an amphetamine stimulant) from the Canadian market, citing international safety information, including 14 reports of sudden deaths, heart-related deaths, and strokes in children and adults taking recommended doses of the medication. The Food and Drug Administration (FDA) in the United States reviewed the same data but had a very different interpretation and response than Health Canada. The FDA did not conclude that any immediate changes were needed in the approved use of this drug based upon its preliminary understanding of Health Canada's analyses of adverse event reports and the FDA's own knowledge and assessment of the reports received by the agency. However, because the stimulants are adrenergic agents, they can theoretically affect heart function, so a "black box" warning was added to the label as a safety precaution for the minority of patients who have a pre-existing cardiac abnormality. There is no evidence that the reported number of sudden deaths significantly exceeds those expected by chance, and Health Canada has reversed the ban, but the mere discussion is enough to worry many patients and parents.

Potential liver disease is another concern. By 1996, enough evidence had accumulated to confirm the risk of rare irreversible liver necrosis in association with pemoline, which was therefore demoted from a first-line drug to third-line. A recent report concerning a nonstimulant medication, atomoxetine, described liver enzyme elevations in two patients that resolved after the medication was discontinued. In a small, preliminary study, researchers found that after just 3 months, every one of a dozen children treated for ADHD with the drug methylphenidate experienced a threefold increase in levels of chromosome abnormalities – occurrences sometimes associated with increased risks of cancer and other adverse health effects. Even though no epidemiologic data have been found showing a higher rate of cancer in those who took methylphenidate (which has been on the market for 40 years), findings like this are still worrisome, adding to the natural stresses that the disorder brings.

Even nonmedical treatments can generate stress, especially if they are demanding or expensive, and thus stress family resources. For example, a systematic program of behavior modification requires considerable time and effort of the caregiver, who may feel that he or she has already given.

Conclusion

Untreated ADHD – and sometimes even treated ADHD – may be considered a stress generator, with a large ripple effect impacting practically all of society: the 35% of the population who have ADHD, their parents, siblings, teachers, classmates, teammates, mates, children, employers, co-workers, and even unknown other drivers. Thus, effective programs for preventing, curing, or controlling ADHD symptoms could ameliorate the fabric of stress throughout society.

See Also the Following Articles

Caregivers, Stress and; Divorce, Children of; Posttraumatic Stress Disorder in Children; Teaching and Stress.

Further Reading

- American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th edn.). Washington, D.C: American Psychiatric Press.
- Famularo, R., Fenton, T., Kinscherff, R. and Augustyn, M. (1996). Psychiatric comorbidity in childhood post traumatic stress disorder. *Child Abuse and Neglect* 20(10), 953–961.
- Ford, J. D., Racusin, R., Ellis, C. G., et al. (2000). Child maltreatment, other trauma exposure, and posttraumatic symptomatology among children with oppositional defiant and attention deficit hyperactivity disorders. *Child Maltreatment* 5(3), 205–217.
- Weinstein, D., Staffebach, D. and Biaggio, M. (2000). Attention-deficit hyperactivity disorder and posttraumatic stress disorder: differential diagnosis in childhood sexual abuse. *Clinical Psychology Review* 20(3), 359–378.

Autoimmunity

B S Rabin

University of Pittsburgh, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp. 268–274, © 2000, Elsevier Inc.

How Can an Autoimmune Disease Be Initiated and How Can Stress Contribute to the Process?

If Stress Suppresses Immune Function, How Does Stress Predispose to the Development of Autoimmune Disease?

The Relationship of Stress to Specific Autoimmune Diseases
Summation

Glossary

Antibody

A protein that combines with the antigen that caused the antibody to be produced. Antibodies have specificity in that they combine most strongly with the particular antigen that stimulated their production. Antibodies are produced by a type of white blood cell called a B lymphocyte. Each B lymphocyte can only produce a single type of antibody, but the B lymphocyte produces thousands of that particular antibody molecule each second. Antibodies bind to antigens that are located outside of tissue cells.

Antigen

A substance that activates the immune system to produce a specific response that reacts against the antigen that activated the immune response. An antigen is usually something that is foreign to the body such as an infectious agent.

Autoimmunity

An immune reaction that is directed against an individual's own tissue. The immune reaction may damage tissue and cause disease.

Cytokine

One of numerous proteins made by a cell which can influence the activity of other cells. Cytokines bind to cytokine receptors on cells with different receptors binding different cytokines. Lymphocytes and antigen-presenting cells produce cytokines which, depending on the particular cytokine and the cell that it binds to, may either upregulate or downregulate the activity of other immune cells. Cytokines are also involved in the regulation of endocrine hormone production and contribute to the modulation of activity of the central nervous system.

Lymphocytes

White blood cells responsible for the immune response. Different types of lymphocytes originate from a common stem cell in the bone marrow. Lymphocytes that continue their maturation in the bone marrow become B lymphocytes, which function by producing antibody.

Tolerance

Lymphocytes that migrate from the bone marrow to the thymus gland for their maturation become T lymphocytes that function in cell mediated immunity, which is antibody independent.

The specific absence of an immune response to an antigen. This is usually an active process that is related to the deletion of antigen-reactive lymphocytes or a mechanism which suppress the ability of the lymphocytes to respond to antigen. Lymphocytes which have receptors for self-antigen (one's own tissue antigens) are eliminated by the tolerance mechanism(s).

How Can an Autoimmune Disease Be Initiated and How Can Stress Contribute to the Process?

Failure to Eliminate Autoreactive B Lymphocytes

The B lymphocytes are produced and mature in the bone marrow as part of the normal process of B cell maturation. Some B lymphocytes randomly develop immunoglobulin receptors that are reactive to self-antigens. Thus, they have the capability of releasing antibodies that react with self antigens. To prevent the development of autoimmune diseases, the autoreactive B lymphocytes must be eliminated before they can produce harmful reactions. The process of removal of autoreactive lymphocytes is called tolerance. Tolerance may be associated with actual death of the lymphocyte or its functional inactivation. The development of tolerance of newly formed autoreactive B lymphocytes must continue throughout life.

If the process of B cell tolerization were highly effective, no individuals would have autoreactive antibodies in their blood. However, nearly all normal individuals can be found to have low levels of antibodies directed to self-antigens. This suggests that the process of tolerization of newly developed B lymphocytes to self-antigens is not entirely effective.

As there is an epidemiological association of autoimmune disease with stress, it is possible that stress interferes with the tolerization of autoreactive B lymphocytes. Whether this is due to a stress-mediated decrease in the amount of circulating self-antigens which must be present in the marrow for B cell tolerance to develop, a decreased accessibility of self-antigens to the developing B lymphocytes in the marrow, a decreased susceptibility to tolerization of the autoreactive B lymphocytes, or other reasons is undetermined. For example, it is possible that stress hormones alter the kinetics of maturation of B lymphocytes so that there is a shorter window of

opportunity for B lymphocytes to become tolerant to self-antigens, alter blood flow to the bone marrow so that there is a decreased concentration of self-antigens present, or alter the release of self-antigens from tissue so that the concentration of these antigens is decreased in the blood.

Failure to Eliminate Autoreactive T Lymphocytes

The T lymphocytes arise in the bone marrow but mature in the thymus gland. Some of the maturing T lymphocytes have receptors on their surface membrane that will combine with self-antigen. It is in the thymus where the autoreactive T lymphocytes undergo the process of tolerization. Similar to autoreactive B lymphocytes, T cell tolerization may involve cell death or functional inactivation.

If all autoreactive T lymphocytes were eliminated in the thymus, there should be no autoimmune diseases. That autoimmune diseases occur indicates that the process is not totally efficient or that there are ways for T cells to escape from tolerance. If the hormonal response to stress alters mechanisms whereby T cell tolerance occurs, stress may increase the likelihood of an autoimmune disease occurring. It is possible that the hormonal response to stress alters the interactions between the cells that are responsible for the induction of tolerance of T lymphocytes. This may result in increased numbers of autoreactive T cells being present in the peripheral lymphoid tissues.

Alteration of Cell Regulation

In the spleen, lymph nodes, and blood, where many immune reactions are initiated, there are a variety of regulatory interactions which occur between cells of the lymphoid system. The mechanism of regulation is likely through the secretion of soluble factors by a particular type of cell that alters the activity of other cells. The hormonal response to stress may alter the function of the regulatory cells which then may predispose to an autoimmune response. Examples included are listed below.

1. Macrophages produce transforming growth factor- β (TGF- β), which inhibits proliferation of T and B lymphocytes. If stress hormones inhibit TGF- β production, activation of autoreactive T or B lymphocytes may be enhanced.

2. The CD4 Th2 lymphocytes produce the cytokines IL-4 and IL-10. These cytokines suppress CD4 Th1 lymphocyte function. Increased Th2 CD4 lymphocyte function, which is associated with a change in the balance between Th1 and Th2 function, has been associated with activation of latent viral infections. Th1 activity is important as it suppresses latent virus activation. An immune reaction against tissue cells

expressing viral antigens has been associated with the onset of autoimmune disease. Thus, stressor-induced alteration of the balance of immune function may contribute to the development of autoimmunity through the mechanism of latent virus activation.

3. The CD4 Th1 lymphocytes produce interferon- γ , which suppresses CD4 Th2 lymphocyte function. An increase of Th1 relative to Th2 activity has been associated with tissue damage in autoimmune disease. If the hormonal response to stress suppresses Th2 and increases Th1 CD4 lymphocyte function, an alteration of the course of autoimmune disease may occur.

There is a paucity of basic science studies relating to how an alteration of the balance between the various components of the immune system occurs when different concentrations of stress hormones are present. Such information is required before a determination can be made regarding the biological impact of stressor induced alterations of regulatory immune cell function.

Increased Susceptibility to Infection

Often, an acute stress or the experience of chronic stress decreases the function of the immune system. An increased susceptibility to upper respiratory viral infections and the reactivation of latent viral infections are the best documented indications of a biological effect of stressor-induced immune alterations. However, it is likely that resistance to any infectious agent is impaired by stress.

Infectious agents have been associated with the development of autoimmune diseases. Some infectious agents have antigenic sites that are identical to self-antigens. The presence of self and foreign antigenic sites on a single protein molecule frequently results in the breaking of tolerance to the self-antigen. For example, if infected with a bacteria that has an antigenic component on its surface that is identical to an antigen that is present on one's own tissues, tolerance can be broken. The antibody that is produced has the potential of altering the function of tissue or damaging tissue. Rheumatic fever is caused by this immune mechanism.

When a virus infects a tissue, an immune reaction to the virus kills the infected tissue cell. One of the mechanisms of cell killing is related to the release of enzymes from CD8 lymphocytes, neutrophils, and macrophages. The enzymes may degrade (denature) some of the molecules of self-antigen, resulting in a self-molecule that retains an antigenic site characteristic of the self-tissue and, in addition, has some antigenic sites that differ from self. When this molecule is processed by the cells of the immune system tolerance

is broken and an immune response to the self-antigenic determinant results.

If Stress Suppresses Immune Function, How Does Stress Predispose to the Development of Autoimmune Disease?

A logical concern relates to how stress, which is usually associated with a decrease in the function of the immune system, can be associated with diseases which are thought to be caused by increased immune activity. This can occur through alteration of the regulatory components of the immune system, as described above, or through an increased susceptibility to infectious diseases with a subsequent increased risk of an autoimmune response.

Obviously, the relationship between the effects of stress on altering the function of the immune system and the development of an autoimmune disease is complex. Further complications are related to different patients with the same disease having different etiologies to their disease. Eventually, clarification of the factors that contributed to the development of an autoimmune disease in a particular patient depends on identifying the immune pathway that was responsible for the disease in that patient. That allows a more precise characterization of the role of stress in the pathogenesis of autoimmune disease to be identified.

The Relationship of Stress to Specific Autoimmune Diseases

The following presents specific autoimmune diseases that have been shown to have an increased frequency of occurrence or which have an exacerbation at a time of stress. However, there are many unanswered questions in regard to the pathogenesis of autoimmune disease involving onset, remissions, and exacerbations. Many of the studies where stress in a person's life is associated with an autoimmune disease are flawed by the process itself. For example, if an individual who has been diagnosed with an autoimmune disease is asked to recall negative events in their life, that individual may try harder to recall such events or add meaning to events that would seem less important if the individual did not have an autoimmune disease. A person experiencing negative life events may see a physician more frequently just because they are not feeling well emotionally and may be more likely to be diagnosed with a disease at an early stage. It is also possible that early stages of disease that produce metabolic alterations (such as thyroid disease or diabetes) may have behavioral changes due to the altered concentration of endocrine

hormones in their blood which increase their likelihood of experiencing negative life events.

Psoriasis

Psoriasis is a skin disease that affects approximately 2% of the population. The involved areas of the skin are a pink color with white flaky scales. The skin lesions frequently undergo remissions and relapses. Psoriasis involves the infiltration of the skin with T lymphocytes and the production of disease-causing cytokines by the lymphocytes. An immune response to infectious agents in the skin may stimulate cytokine release. A hypothetical model can be constructed which supports the possibility that stress participates in the course of psoriasis by releasing neuropeptides from nerves which terminate in the skin. The neuropeptides, particularly substance P, may enhance the migration of lymphocytes into the skin which stimulate localized disease-enhancing cytokine release.

There are numerous published studies indicating that stress is associated with the exacerbation of symptoms in patients with psoriasis. A summary of the literature evaluating stress effects on the course of psoriasis supports the role of stress as a modulator of disease activity and suggests that the severity of the stressor may be associated with more severe disease. Additional studies have suggested that relaxation procedures which may produce a psychological relaxation (which would have an effect on the immune system if they decreased sympathetic nervous system activation) are capable of ameliorating the clinical appearance of disease in patients with psoriasis. However, there are many factors which make interpretation of the studies difficult. Usually the number of subjects is low and the stage of disease and the actual pathogenesis of the disease in the subjects may vary. Coping skills (including psychosocial support, physical fitness, optimism, sense of humor, and belief systems) which help to buffer the effects of stress on the immune system may vary. Even whether the subject experienced high levels of stress as a youth or the subject's mother experienced high levels of stress during pregnancy may influence an individual's response to stress.

Rheumatoid Arthritis (RA)

Rheumatoid arthritis is associated with a large accumulation of lymphocytes in the joint spaces. There is destruction of the joint spaces and considerable pain. Systemic manifestations include tendonitis, pericarditis, vasculitis, and subcutaneous nodules pleuritis. The disease frequently has a relapsing and remitting course.

It is usually assumed that the disease is produced by activation of the immune system to an antigen. As identical twins are often discordant for RA it is likely that there is an environmental (antigen) involved in the disease pathogenesis. However, there has not been identification of specific antigen(s) to which the immune response is directed.

If the disease is a response to an antigen the functional competency of the immune system and its modification by stress hormones could modify the disease course. If stress does alter the course of RA, it may be through dysregulation of the normal relationship of the various lymphocyte subpopulations to each other. Reports of the effect of stress on the clinical course of RA have been variable and do not provide an unambiguous understanding regarding whether stress alters the onset or course of disease.

Another approach to evaluating whether stress influences the course of disease in patients with RA is to evaluate the effect of stress reduction interventions. Psychological interventions have suggested that slight clinical improvement can be achieved in patients with RA. Thus, regardless of the difficulty in defining an immunologic etiology in RA or a definitive understanding of the pathogenesis of the disease, there is suggestive evidence for an interaction between stress and the disease process.

There are extensive studies which indicate that psychological intervention programs can decrease the pain perceived by patients with RA. Whether the pain reduction produces an alteration in the amount of opioid produced, both within the CNS and peripherally, and whether there is a subsequent change in immune system function which ameliorates the disease process has not been determined.

Recently diagnosed subjects with RA have a lower activation of the sympathetic nervous system following a mental challenge stress task than control subjects without RA. Those subjects with the greatest amount of pain had the lowest sympathetic activation. The lack of an intact stress response in RA subjects may indicate that the immune pathogenesis of RA differs from the immune pathogenesis of other autoimmune diseases or that RA is more likely to occur in subjects who have an impaired hormonal response to stress.

An impaired hypothalamic-pituitary-adrenal (HPA) axis hormonal response to stress may allow the immune system to become overactivated, promoting an autoimmune process. Indeed, experimental animals which have an increased susceptibility to the induction of arthritis have been found to have functional differences of their HPA axes (susceptible animals have a decreased production of corticosterone) in comparison to experimental animals which are resistant to the induction of arthritis. An abnormal HPA response to

immune and inflammatory stimuli has been reported in patients with RA.

Whether humans with RA have an altered response to stress because of modification of CNS function by cytokines or whether there is an innate difference within the CNS that modifies their response to stress has not been defined. However, subjects with another autoimmune disease, multiple sclerosis, have an intact response to the same psychological stressor that produced a diminished response in subjects with RA.

Graves' Thyroid Disease

Graves' hyperthyroidism is associated with weakness, nervousness, weight loss, fatigue, warm and sweaty skin, menstrual irregularities, heat sensitivity, protrusion of the eyes, and an enlarged thyroid gland. The pathogenesis is due to an antibody that binds to a receptor on thyroid epithelial cells that binds thyroid stimulating hormone (TSH). The thyroid stimulating antibody mimics the binding of TSH and stimulates the release of thyroid hormone. However, unlike TSH, which decreases in concentration as the concentration of thyroid hormone increases (reducing the releases of thyroid hormone), there is no inhibition of the physiologic effects of the thyroid stimulating antibody.

There are several studies which found that high levels of stress are more likely to occur in an individual who develops Graves' disease than in individuals who do not develop Graves'. These studies suggest that patients had more negative life events in the months preceding the clinical diagnosis of disease. However, not all studies support an association between prior stress and the onset of Graves'.

It is possible that if thyroid hormones are increased in concentration prior to the clinical diagnosis of disease, behavioral alterations are induced which either contribute to negative life events or the subject's awareness of negative life events. As thyroid hormone measurements are not made in the months preceding the appearance of clinical manifestation of Graves' disease, a possible influence of thyroid hormones on life events cannot be ruled out.

Multiple Sclerosis (MS)

Multiple sclerosis produces weakness, visual dysfunction, and paralysis. Multiple sclerosis lesions develop in the brain and the spinal cord where myelin surrounding nerve fibers is destroyed by lymphocytes and macrophages. There is an infiltrate of CD4 and CD8 lymphocytes, macrophages, and plasma cells. Nerve conduction is impaired. Identical twins have approximately a 50% concordance for MS,

indicating that both genetic and environmental factors are involved with the disease pathogenesis.

The exact pathogenesis of MS is uncertain, but there is considerable evidence suggesting participation of the immune system. This includes characteristic changes in the immunoglobulin content of spinal fluid, which suggests an immune reaction occurring within the CNS; the mitogenic responsive of lymphocytes from patients with MS to antigens derived from the CNS; the clinical response to immunosuppressive therapy; and the existence of an experimental model that is induced by immunization with myelin basic protein, an antigen found in myelin.

The studies of stress as a contributor to the etiology of MS are not definitive. They suffer from the faults of small sample size, heterogeneity of the extent and pathogenesis of disease, various stress buffering capabilities of subjects, and concomitant clinical depression having an effect on immune system function. Definition of the exact immune mechanism(s) which produce demyelination is also a hindrance, as without knowing which immune mechanisms are responsible for the tissue damage it is impossible to know which pathway is being altered by stress.

An increase of Th1 lymphocyte function would be compatible with an immune change that could increase the risk for the development of MS. Interestingly, cerebrospinal fluid (CSF) of patients with MS contains cytokines which promote CD4 Th1 lymphocyte activity. An indirect indication of predominant Th1 activity in MS patients is the paucity of IgE-mediated disease in MS patients, as Th1-derived cytokines suppress the Th2 cells that promote B lymphocytes to synthesize IgE. Stress has been found to increase the production of cytokines from Th1 lymphocytes while sparing alteration of cytokine production by Th2 lymphocytes.

There have been many studies which have looked at the association of stress with the clinical course of patients with MS. The studies found that MS patients had significantly more life stress prior to the onset of disease or disease exacerbation than did controls. The most common life events which were related to exacerbation were those involving financial or interpersonal concerns.

An interesting observation regarding stress and disease activity in MS was made during the Persian Gulf war. This study reported a decreased rate of exacerbations of MS during the war. This observation indicates that an intense psychological stressor can have a profound effect on suppressing immune system function and decrease disease activity. However, this is a unique situation that does not argue against less intense life-event stressors having an influence on the exacerbation of the clinical course of MS.

Overall, the clinical observations and the possible stressor-induced immune changes suggest that stress can alter the course of MS. Studies of stress intervention are needed to confirm or refute this hypothesis. Care will have to be directed to the clinical homogeneity of the patient population including HLA characteristics, social support and other buffering systems available to the patient, use by the patient of medications, and how strongly the patient believes in the potential effectiveness of the intervention therapy.

Insulin-Dependent Diabetes Mellitus (IDDM)

A frequent pathogenesis of IDDM is related to immune-mediated destruction of the insulin-producing β -islet cells. Antibodies to insulin and islet cell antigens are present prior to the clinical onset of disease and may indicate a destructive inflammatory process is occurring. Sensitized T lymphocytes that have a cytotoxic effect on islet cells are also present.

The disease is frequently preceded by a viral infection and may be initiated by an immune response to a virus that specifically infects β -islet cells. Indeed, virus has been isolated from the islets of newly diagnosed diabetic subjects and, when the virus is injected into mice, IDDM has occurred with an associated immune-mediated destruction of the β -islet cells. Active viral infection is suggested because of a rising antibody titer in the patient's serum to the isolated virus.

There are several reports associating life-event stressors with the onset of IDDM. Unlike diseases such as rheumatoid arthritis or multiple sclerosis, which have a relapsing and remitting course, IDDM does not remit. Therefore, studies showing an association of stress with alteration of clinical activity are nonexistent, but studies showing an association of stress with disease onset are numerous. Stressor-induced alterations of immune function may contribute to disease onset by (1) suppressing immune function and increasing susceptibility to viral infection and (2) altering the balance of regulatory lymphocyte populations which then allows an autoimmune response to occur.

Early reports linking stressors in an individual's life with an increased susceptibility to the development of IDDM were not rigorously controlled and presented small sample sizes. However, there are studies which suggest that there may be an association between stress and susceptibility to the onset of IDDM. These studies should be considered to be preliminary, as they have not addressed important questions.

Specifically, does stress increase the risk of IDDM onset in all individuals or only those who are HLA-DR3 and/or HLA-DR4 positive? If this association

only occurs in HLA-susceptible individuals, should the control group of individuals who do not have high levels of stress be restricted to HLA-DR 3- and/or HLA-DR-4-positive individuals?

The only way to control for genetic differences between subjects is to study identical twins. The concordance for IDDM in identical twins is approximately 50%. Environmental factors are usually designated as explaining the lack of 100% concordance. Environmental factors may include viruses, but may also include stress and the ability of a subject to cope with the stressor. A study conducted in identical twins both concordant and discordant for IDDM where factors such as stress, psychosocial support, physical fitness, belief systems, sense of humor, and being an optimist or a pessimist are considered would make a significant contribution to our understanding of the relationship between stress and susceptibility to autoimmune disease. If identical twins discordant for IDDM experience the same stressors, does the twin who develops IDDM have better coping skills?

Studies which have contributed to the suggestion of an association between stress and the onset of IDDM in children have found that severe emotional stress, usually associated with parental loss, especially occurring early in life, increased the risk for the subsequent development of IDDM.

Summation

Characterization of the potential interrelationship between stress and the development or exacerbation of an autoimmune disease requires knowledge of at least two parameters: (1) What components of the immune system are functionally altered by stress? and (2) How does alteration of the immune system initiate immune-mediated damage to self-tissue? Unfortunately, detailed information regarding each of the two questions is still lacking. However, there is abundant evidence from epidemiological studies that the experience of stress is associated with autoimmune disease onset or exacerbation. Continued studies of stressor-induced immune alteration and the mechanism of onset of autoimmune disease will help to clarify their relationship.

See Also the Following Articles

Arthritis; Cytokines; Cytotoxic Lymphocytes; Diabetes, Type I; Immune Cell Distribution, Effects of Stress on; Immune Function, Stress-Induced Enhancement; Immune Suppression; Lymphocytes; Multiple Sclerosis; Thyroid Hormones.

Further Reading

- Anderson, K. O., Bradley, L. A., Young, L. D., McDaniel, L. K. and Wise, C. M. (1985). Rheumatoid arthritis: review of the psychological factors relating to etiology, effects and treatment. *Psychological Bulletin* 98, 358–387.
- Hagglof, B., Blom, L., Dahlquist, G., Lonnberg, G. and Sahlin, B. (1991). The Swedish childhood diabetes study: indications of severe psychological stress as a risk factor for type I (insulin-dependent) diabetes mellitus in childhood. *Diabetologia* 34, 579–583.
- Herbert, T. B. and Cohen, S. (1993). Stress and immunity in humans: a meta-analytic review. *Psychosomatic Medicine* 55, 364–379.
- Kung, A. W. C. (1995). Life events, daily stresses and coping in patients with Graves' disease. *Clinical Endocrinology* 42, 303–308.
- Sternberg, E. M. (1995). Neuroendocrine factors in susceptibility to inflammatory disease: focus on the hypothalamic-pituitary-adrenal axis. *Hormone Research* 43, 159–161.
- Wilder, R. L. (1995). Neuroendocrine-immune system interactions and autoimmunity. *Annual Review of Immunology* 13, 307–338.

Autonomic Nervous System

W R Lovallo

Veterans Affairs Medical Center and University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

J J Sollers III

Department of Psychology, Ohio State University, Columbus, OH, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by W R Lovallo and J J Sollers III, volume 1, pp 275–283, © 2000, Elsevier Inc.

Background

Autonomic Nervous System Compared to the Somatic Motor System

Autonomic Nervous System Has Three Branches

Neurons and Neurotransmitters

Anatomy of the Autonomic Nervous System

Autonomic Regulation of Specific Organs

Local Regulation of Autonomic Function

Physical and Psychological Stress

Summary

Glossary

- Brain stem** The portion of the brain connecting the cerebral hemispheres with the spinal cord, consisting of the midbrain, the pons, and the medulla.
- Endocrine** Pertaining to the secretory glands of the body, which release hormones that circulate in the blood to act on distant glands and target tissues to regulate their function.

Ganglion

A specialized collection of interconnected nerve cell bodies.

Homeostasis

A concept expressing the collective actions of the autonomic nervous system and endocrine system to maintain physiological systems within a range of optimal values consistent with the maintenance of life.

Hypothalamus

The collection of brain nuclei located below the thalamus that act to regulate the endocrine system and the autonomic nervous system to maintain homeostatic processes.

Limbic system

Brain structures of the cortex and basal ganglia that are involved with motivation and the emotions.

Neuro-transmitter

A chemical substance supporting communication across a synapse or neuro-effector junction.

Receptor

A complex protein molecule that is embedded within a cell membrane having chemical affinities for one or more neurotransmitters and that initiates a response by the cell.

In the course of daily life, we must rest or be active, obtain food and digest it, and, occasionally, cope with an exceptional challenge. At all times, during normal activity and during extreme stress, our brain coordinates the actions of our vital organs and our muscles to ensure our survival. The brain uses two systems to do this, the nervous system and the endocrine system. This article describes one branch of the nervous system, called the autonomic nervous system, which plays a central role in the stress response and in normal regulation.

Background

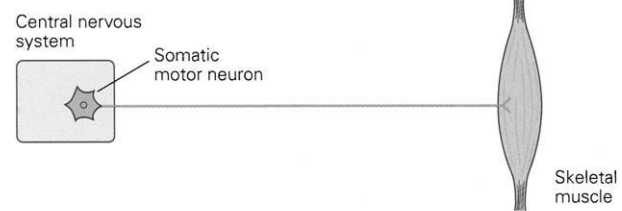
The peripheral nervous system has two branches: the somatic motor system, which controls voluntary movement, and the autonomic nervous system (ANS), which regulates the internal organs and the eyes. There are five main points to consider regarding the ANS and its role in the stress response: (1) the ANS, along with the endocrine system, regulates the vital organs and coordinates their functions; (2) the ANS is part of a layered set of controls over organ function, starting at the local organ and ending with the brain; (3) the ANS is controlled directly by two brain areas, the brain stem and hypothalamus, and indirectly by higher brain centers, including the prefrontal cortex; (4) during normal activity, the vital organs regulate themselves with little outside help, but during stress the ANS is needed to coordinate their actions; and (5) during psychological stress, the bodily responses we experience originate in higher brain centers that act on the brain stem and hypothalamus to cause changes in the body via the ANS.

Two historic figures have shaped our understanding of the functions of the ANS during stress: Claude Bernard (1813–1878), the French physiologist, who showed that living things must respond to changes in their external environment and in the internal environment (the blood and bodily fluids) to maintain proper cellular and organ function, and Walter Cannon (1871–1945), an American physiologist, who coined the term “homeostasis” to describe these processes of self-regulation. He showed that homeostasis was served primarily by actions of the ANS along with the endocrine system. Cannon demonstrated that the ANS reacts to emergencies in a pattern of widespread activation that he called the fight-or-flight response.

Autonomic Nervous System Compared to the Somatic Motor System

There are four points of comparison between the somatic motor nervous system and the ANS: (1) both of them are motor systems that regulate the activity of muscle fibers or secretory cells; (2) the somatic nervous system allows voluntary control over muscles, while the ANS automatically regulates the vital organs and related tissues without voluntary control; (3) somatic nerves always activate contraction of the somatic muscles, while ANS fibers will either activate or inhibit the vital organs; and (4) there are two anatomical differences between the somatic nervous system and the ANS ([Figure 1](#)). In the somatic nervous system, a nerve fiber goes directly from the central nervous system to a skeletal muscle. The ANS

(a) Somatic motor system



(b) Autonomic motor system

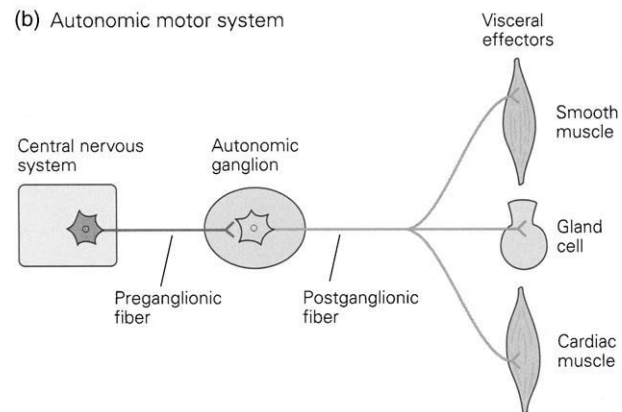


Figure 1 Neurons innervating the skeletal muscles (a) and the vital organs (b). (a) Skeletal motor nerves. Motor neurons that control the skeletal muscles travel from the motor portions of the cerebral cortex, down the spinal cord, and directly to individual muscle fiber bundles without synapsing with other neurons. (b) Autonomic nerves. The autonomic nervous system has two sets of neurons. Preganglionic neurons originate in the pons and medulla of the brain stem. They descend by way of the spinal cord and travel to autonomic ganglia, where they synapse with postganglionic neurons that travel to specific effector organs. There are three types of autonomic effectors: smooth muscle cells, secretory cells, and cardiac muscle. From Iversen, S., Iversen, L. and Saper, C. B. (2000). *The autonomic nervous system and the hypothalamus*. In: Kandel, E. R., Schwartz, J. J. & Jessel, J. H. (eds.) *Principles of neural science* (4th edn., pp. 960–981). New York: McGraw-Hill. Copyright The McGraw-Hill Companies, Inc. Used with permission.

is more complex, with two sets of nerve fibers: preganglionic fibers that travel from the central nervous system to a peripheral autonomic ganglion, where they synapse with postganglionic fibers that travel to a target organ.

Autonomic Nervous System Has Three Branches

The ANS has three branches: the sympathetic, parasympathetic, and enteric nervous systems. The two major branches, the sympathetic and parasympathetic, are diagrammed in [Figure 2](#).

Sympathetic Branch

Sympathetic preganglionic fibers exit from the middle, thoracic, and lumbar segments of the spinal

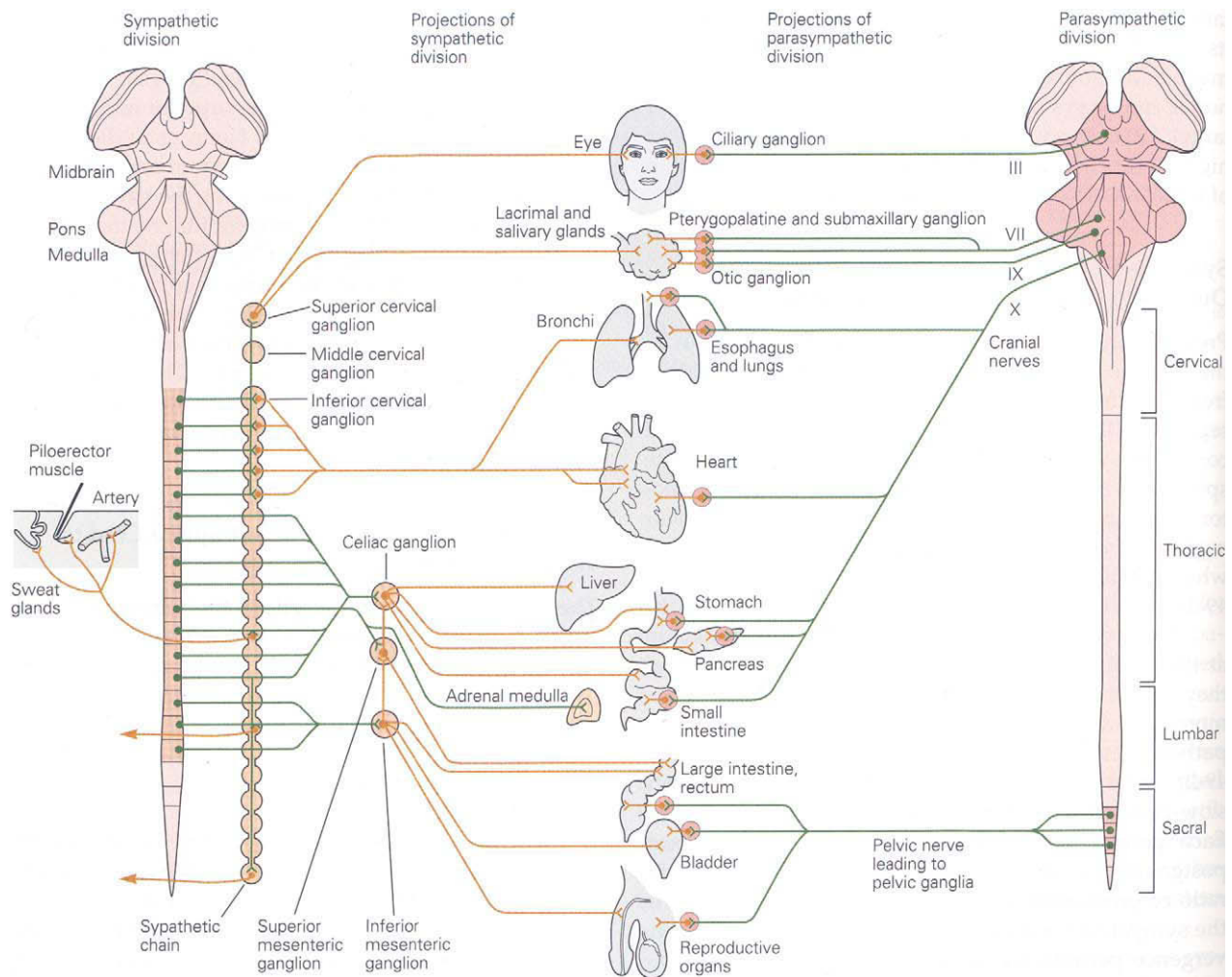


Figure 2 The sympathetic and parasympathetic branches of the autonomic nervous system. The sympathetic branch has preganglionic fibers that exit the spinal cord at the thoracic and lumbar vertebrae. These travel to the sympathetic chain ganglia, located near the spinal cord, or to the celiac or mesenteric ganglia of the gut. The sympathetic postganglionic fibers then exit their respective ganglia and travel to distant target organs. The parasympathetic preganglionic fibers exit the central nervous system either from the pons and medulla or from the sacral level of the spinal cord. The preganglionic fibers travel along several cranial nerves or the splanchnic nerve, and they travel some distance to their ganglia, located near or within the target organs. The postganglionic fibers then travel short distances to their target tissues. From Iversen, S., Iversen, L. and Saper, C. B. (2000). *The autonomic nervous system and the hypothalamus*. In: Kandel, E. R., Schwartz, J. J. & Jessel, J. H. (eds.) *Principles of neural science* (4th edn., pp. 960–981). New York: McGraw-Hill. Copyright The McGraw-Hill Companies, Inc. Used with permission.

cord. The sympathetic branch is so named because it operates in a coordinated fashion and in sympathy with the emotions. In general, the sympathetic branch regulates energy-using functions and is most active during states of fight or flight in order to maintain homeostasis in the face of external threats. It is also active during normal activities, but less so.

Parasympathetic Branch

In the parasympathetic branch, preganglionic fibers leave the central nervous system at points above

and below the sympathetic fibers, hence the name parasympathetic. The upper group of neurons exits the medulla by way of the cranial nerves. These exit the central nervous system directly from the pons and medulla and serve the head, neck, cardiovascular system, and gut. The lower group exits the sacral segments of the spinal column. Parasympathetic activity generally serves energy-conserving functions that Cannon called rest-or-digest. It is active in energy intake, such as feeding and digestion, and during reproductive functions, and it is most active during routine activities.

Enteric Branch

As the name implies, the enteric system pertains exclusively to the organs of digestion, including the pancreas, the gall bladder, and the gastrointestinal tract. Some consider the enteric nervous system to be a branch of the ANS, while others view it as part of the digestive system. All three branches of the ANS are active during normal activities, and they coordinate their actions to maintain homeostasis. However, the sympathetic nervous system predominates, and the other branches are inhibited during states of acute stress, such as periods of fight or flight, when resources must be devoted to the task of immediate survival.

Neurons and Neurotransmitters

There are two points of communication in the ANS. At a ganglion, preganglionic neurons synapse with postganglionic neurons. At a target organ, the postganglionic neuron forms a neuroeffector junction with a muscle cell or a secretory cell.

At both places, neurons communicate using specialized chemicals called neurotransmitters. These are released onto specialized receptors that activate either a postganglionic neuron or an effector cell. There are two ANS neurotransmitters, acetylcholine and norepinephrine (also called noradrenalin).

Preganglionic neurons always use acetylcholine to communicate to their postganglionic neurons. Postganglionic neurons use two primary neurotransmitters to act on their target tissues. In the parasympathetic branch, postganglionic fibers secrete acetylcholine at neuroeffector junctions located at target tissues. In addition, there are parasympathetic postganglionic fibers to the digestive tract that use the purine adenosine triphosphate as their transmitter. These purinergic fibers are involved in relaxation of the esophagus, stomach, and intestines. In the sympathetic branch of the ANS, postganglionic fibers secrete norepinephrine at their neuroeffector junctions. There are two exceptions to this rule. The sympathetic nervous system uses acetylcholine as a transmitter in the neuroeffectors of two tissues: the specialized sweat glands of the skin and the medulla of the adrenal gland.

Autonomic ganglia have an important feature. They are not simply relay stations; instead, they have small neurons that communicate directly between the larger pre- and postganglionic neurons. These small neurons secrete the neurotransmitter dopamine, which inhibits the main ganglionic neurons. The autonomic ganglia use this local inhibition to produce the first level of integrative control over the vital organs.

Anatomy of the Autonomic Nervous System

The brainstem contains several autonomic control nuclei that regulate the actions of the sympathetic and parasympathetic branches. Preganglionic fibers descend from these control centers to their autonomic ganglia, and postganglionic fibers travel to the target organs.

Sympathetic Division

The sympathetic division serves the eye, salivary glands, lungs, heart, liver, spleen, stomach, gut, kidney, genitalia, and all blood vessels except those of the brain. The actions of these sympathetic inputs are summarized in [Table 1](#) for each effector organ.

Sympathetic fibers exit from the spinal cord at thoracic levels 1–12 and lumbar levels 1–3. As shown in [Figure 2](#), there are two sets of sympathetic ganglia. The chain ganglia run parallel to the spinal column, and they are highly interconnected. A second, more distinct group includes the cervical and stellate ganglia, located in the neck, and the celiac, superior, and inferior mesenteric and renal ganglia found in the abdomen.

Sympathetic preganglionic fibers have four patterns of dispersal. Some preganglionic fibers enter the nearest ganglion and their postganglionic fibers exit at the same level of the ganglionic chain; other preganglionic fibers ascend the sympathetic chain ganglion to synapse in the cervical and stellate ganglia before they send postganglionic fibers to the eye, head, neck, mouth, and upper limbs and heart – the stellate ganglion serves the heart particularly. The blood vessels and skin of the lower body are served by postganglionic fibers that receive input from preganglionic fibers that descend from the lumbar and sacral ganglia; and, finally, certain preganglionic fibers pass directly through the chain ganglia without synapsing. They collect to form the three splanchnic nerves that pass through the diaphragm to enter the abdomen before arriving at the celiac, mesenteric, and renal ganglia.

A notable exception to this organization of sympathetic fibers is the innervation of the adrenal gland. In this case, sympathetic preganglionic fibers exit the spinal cord and travel directly to the adrenal medulla without synapsing at a ganglion.

The sympathetic nervous system is especially effective at promoting coordinated action of the bodily organs during states of stress. Its ganglia are close to the spinal cord and far from their target organs, and a single preganglionic fiber may send branches to several of the chain ganglia at different levels of

Table 1 Responses of Effector Organs to Autonomic Nerve Impulses^a

<i>Effector organs</i>	<i>Sympathetic receptors</i>	<i>Sympathetic impulses and responses</i>	<i>Parasympathetic impulses and responses</i>
Eye			
Radial muscle, iris	α	Contraction (pupil enlargement) ++	—
Sphincter muscle, iris		—	Contraction (pupil constriction) +++
Ciliary muscle	β_2	Relaxation for far vision +	Contraction or near vision +++
Heart			
SA node	β_1	Increase in heart rate ++	Decrease in heart rate; vagal arrest ++
Atria	β_1	Increase in contractility and conduction velocity ++	Decrease contractility, and (usually) increase in conduction velocity ++
AV node	β_1	Increase in automaticity and conduction velocity ++	Decrease in conduction velocity; AV block +++
Ventricles	β_1	Increase in contractility, conduction velocity, automaticity, and rate of idioventricular pacemakers +++	Slight decrease in contractility
Arterioles			
Coronary	α, β_2	Constriction +; dilation ++	Dilation +
Skin and mucosa	α	Constriction +++	Dilation
Skeletal muscle	α, β_2	Constriction ++	Dilation +
Cerebral	α	Constriction (slight)	Dilation
Pulmonary	α, β_2	Constriction +; dilation	Dilation
Abdominal viscera			
Kidney	α, β_2	Constriction +++; dilation +	—
Salivary gland	α	Constriction +++	Dilation ++
Veins (systemic)	α, β_2	Constriction ++; dilation ++	—
Lung			
Bronchial muscle	β_2	Relaxation +	Contraction ++
Bronchial glands	?	Inhibition (?)	Stimulation +++
Stomach			
Motility and tone	α, β_2	Decrease (usually) +	Increase +++
Sphincters	α	Contraction (usually) +	Relaxation (usually)
Secretion		Inhibition (?)	Stimulation +++
Intestine			
Motility and tone	α, β_2	Decrease +	Increase +++
Sphincters	α	Contraction (usually) +	Relaxation (usually) +
Secretion		Inhibition (?)	Stimulation ++
Gallbladder and ducts		Relaxation +	Contraction +
Kidney	β_2	Renin secretion ++	—
Urinary bladder			
Detrusor	β_2	Relaxation (usually) +	Contraction +++
Trigone and sphincter	α	Contraction ++	Relaxation ++
Ureter			
Motility and tone	α	Increase (usually)	Increase (?)
Uterus	α, β_2	Pregnant: contraction (a); nonpregnant: Relaxation (B)	Variable
Sex organs, male	α	Ejaculation +++	Erection +++
Skin			
Pilomotor muscles	α	Contraction ++	—
Sweat glands	α	Localized secretion +	Generalized secretion +++
Spleen capsule	α, β_2	Contraction +, relaxation +	—
Adrenal medulla		—	Secretion of epinephrine and norepinephrine
Liver	α, β_2	Glycogenolysis, gluconeogenesis +++	Glycogen synthesis +

Continued

Table 1 Continued

<i>Effector organs</i>	<i>Sympathetic receptors</i>	<i>Sympathetic impulses and responses</i>	<i>Parasympathetic impulses and responses</i>
Pancreas			
Acini	α	Decreased secretion +	Secretion ++
Islets (B cells)	α	Decreased secretion +++	—
	β_2	Increased secretion +	—
Fat cells	α, β_1	Lipolysis +++	—
Salivary glands	α	Potassium and water secretion +	Potassium and water secretion +++
	β_2	Amylase secretion +	—
Lacrimal glands		—	Secretion +++
Nasopharyngeal glands		—	Secretion ++
Pineal gland	β_2	Melatonin synthesis	—

^aReceptor types refer to receptors of the sympathetic nervous system. There are three types of sympathetic receptors: alpha (α), beta-1 (β_1), and beta-2 (β_2). α receptors are most sensitive to norepinephrine released from sympathetic nerve terminals. β_1 receptors are located only at the heart and in fat cells. These are sensitive to norepinephrine from sympathetic nerve terminals and to circulating epinephrine. β_2 receptors are mostly not innervated by sympathetic nerves. They are primarily sensitive to circulating epinephrine. The parasympathetic nervous system acts exclusively by secretion of acetylcholine on cholinergic receptors, regardless of the organ in question. The adrenal medulla is innervated by sympathetic preganglionic fibers that secrete acetylcholine. The sweat glands of the palms of the hands and soles of the feet are innervated by postganglionic sympathetic fibers that release acetylcholine. +, little action; ++, moderate effect; +++, large effect.

the spinal column. There is a ratio of 1 preganglionic fiber to 10 postganglionic fibers. This allows for coordinated activation of organs of the cardiovascular and pulmonary systems that are needed for coping with fight-or-flight situations. Epinephrine secreted by the adrenal medulla circulates throughout the body to act on noninnervated β receptors, further coordinating the organs sustaining fight or flight.

Parasympathetic Division

As shown in [Figure 2](#) and [Table 1](#), parasympathetic ganglia are located close to or within the organs they innervate. Preganglionic fibers arise from the brain stem and sacral region of the spinal cord. The eye, face, and mouth are served by cranial nerves III, VII, and IX. The cardiopulmonary and digestive systems are served by the vagus nerve (cranial nerve X), while the functions of excretion and control of reproductive organs are served by the splanchnic nerves.

The parasympathetic nervous system is designed to promote highly specific regulation of individual organs: (1) parasympathetic preganglionic fibers are long and travel to specific ganglia linked to specific organs; (2) postganglionic fibers are short and are sometimes found entirely within the innervated tissue; and (3) the ratio of pre- to postganglionic fibers is about 1:3, with one preganglionic fiber affecting only a single organ.

Autonomic Regulation of Specific Organs

[Figure 2](#) and [Table 1](#) summarize how the ANS regulates specific organs. The activity of each organ

involves interactions between the organ's intrinsic functioning and inputs from the ANS. In most cases, the sympathetic and parasympathetic branches of the ANS have opposite effects on the organs they innervate, and the balance of sympathetic to parasympathetic outflow determines the organ's final level of activity. This balance shifts constantly as the physiological and behavioral demands of the person change. During states of fight or flight, this balance is predominantly sympathetic in nature.

This regulatory interplay is illustrated in the activity of the heart. With no external controls, the heart beats automatically at a fixed rate of about 105–110 beats per minute. This rate is due to the heart's internal pacemakers, the atrioventricular and sinoatrial nodes. The pacemakers each have alternate phases of excitation and quiescence, and they interact to allow the heart to beat in a coordinated way. The heart's intrinsic force of contraction is determined by how well the ventricles fill with blood between beats and the resulting stretch on the walls of the ventricles. Greater wall tension results in a more effective contraction, with more blood being pumped on the next beat. These intrinsic mechanisms are entirely adequate to sustain life if conditions are constant and no threats to homeostasis are present.

The two branches of the ANS are active at all times, and they coordinate their actions to regulate the heart during all levels of demand on the circulation. During most normal behavioral states, the predominant ANS influence on the heart is through the parasympathetic branch. This branch acts on the heart by way of the vagus nerve, whose firing serves to lower the heart's

rate and force of contraction. The heart's intrinsic rhythm is normally regulated by this parasympathetic input. Up to and including moderate demands on the circulation, the parasympathetic branch can allow the heart to pump more blood simply by firing less often and allowing the heart to increase its rate up to its intrinsic rate.

At times of increased demand, the ANS may cause the heart to change its rate and force of contraction, and it may have to coordinate these changes with changes in the state of the peripheral blood vessels. A rapid change in circulatory demand occurs, for example, when a person suddenly rises from a chair, at which time there is a danger of blood pooling in the legs, causing a loss of circulation to the brain. This danger is averted by a rapid increase in firing by the sympathetic nerves that increase the rate and force of the heart's contractions and also increase the constrictive force of the blood vessels of the lower extremities. During more severe and sustained demands, such as intense physical exercise, the parasympathetic outflow to the heart is almost absent, and the sympathetic influence increases toward a maximum. In addition, endocrine input to the heart by circulating epinephrine greatly augments the heart's contractions. These three influences can cause a sevenfold increase in the volume of blood the heart pumps each minute. In a trained athlete, such increases in circulation can be sustained for hours.

These examples illustrate that the heart can function without outside control if no response is called for. However, the ANS must provide external control to meet rapid or sustained demands on the circulation and to help the heart operate in concert with the blood vessels. Circulating hormones, such as epinephrine, compliment the actions of the ANS to achieve the necessary level and coordination of cardiovascular response during stress. Similar arguments apply to the ANS regulation of other organ systems.

Local Regulation of Autonomic Function

Since ANS controls, especially sympathetic activity, are more or less nonspecific for each organ, there has been considerable interest in local mechanisms that can fine-tune organ function. The previous section illustrated the interplay of intrinsic and extrinsic regulation of the heart. There is still another level of fine-tuning of ANS actions at a given tissue. In all organs, ANS input is regulated by locally released biochemicals classified as peptides, neuromodulators, or metabolites. These act by several mechanisms at either the smooth muscle or the autonomic-neuroeffector junction to modify ANS activity or its effectiveness.

An example of the action of such local modulators is seen in the regulation of blood flow. Blood flow into the tissues is regulated by sympathetic activity at local blood vessels in concert with local metabolic demands. To regulate the level of tonus at the blood vessel wall, sympathetic nerves release norepinephrine at the sympathetic neuroeffector junction, causing contraction of networks of smooth muscle cells, increasing constriction of the blood vessel. This constrictive force can reduce blood flow into a local volume of tissue. If this were the only influence, the cardiovascular system would not adequately regulate blood flow according to the specific needs of local tissues. However, this constrictive force is balanced by local chemical substances that act on the blood vessel wall in relation to the activity of the local tissues and their consequent need for increased blood flow. Metabolites produced by cellular activity can act locally on smooth muscle cells to regulate their activity. For example, carbon dioxide produced by tissue metabolism leads to local production of nitric oxide, and nitric oxide causes relaxation of vascular smooth muscle cells. This results in the relaxation of blood vessels, allowing improved blood flow to a local volume of tissue. In this process of local autoregulation, exercising muscles produce more carbon dioxide when they are working harder, causing greater relaxation in the local resistance vessels and increasing the local blood supply. Neuromodulators can act on the sympathetic nerves themselves to regulate their actions on smooth muscle cells. In the sympathetic nerves, adenosine diphosphate is released along with norepinephrine at the adrenergic nerve terminal. This adenosine diphosphate is rapidly metabolized to adenosine, which then acts on adenosine receptors located on local sympathetic nerve terminals to decrease the release of norepinephrine. The greater the level of sympathetic outflow, the greater the production of adenosine, and the greater the modulation of neurotransmitter release. In the case of the blood vessel wall, autoregulation and neuromodulation work together to balance the effects of sympathetic activation and constrictor tonus, allowing blood flow to be controlled according to local tissues needs. Similar processes occur in the heart muscle.

The balance of intrinsic regulation, ANS inputs, and local regulation occurs in all tissues. This interaction fine-tunes organ function to maintain homeostasis.

Physical and Psychological Stress

Physical and psychological stress result from very different causes, but they both influence the body

by way of the ANS. The earlier example of a person arising from a seated position illustrates a physical challenge to homeostasis that results in a response by the ANS to maintain blood flow to the brain. This ANS response is coordinated by reflex mechanisms in the large blood vessels and brain stem that sense an impending drop in blood pressure and orchestrate the shift toward a sympathetically dominated state until blood flow is normalized. Such ANS responses are referred to as bottom-up because their source is in the body, involving some physical threat to homeostasis.

Psychological stress has a different origin. Our bodies can undergo significant changes in activity when we are anxious, ruminating about past failures, or in an acute state of fear. However, these changes originate in the highest centers in the brain, including the prefrontal cortex and limbic system. These ruminations result in signals to the hypothalamus and brain stem that can cause changes in the state of the body through the ANS and endocrine systems. These brain mechanisms therefore exert a top-down effect on the state of the body. In this manner, psychological stimuli can act as stressors in themselves. These can alter the usual homeostatic balance of the body.

Summary

The ANS and the endocrine systems are the brain's two primary pathways for regulating the vital organs. The ANS innervates all tissues outside of the central nervous system. It regulates contraction of smooth muscle cells and cardiac muscle cells and output by secretory cells. Changes in the internal or external environment of the body are normally countered by the ANS, which serves to maintain homeostasis. The ANS itself is controlled by the brain stem and

hypothalamus. The highest level of control is provided by the brain's limbic system and cerebral cortex. During times of emotional distress, influences descending from these higher centers alter the activity of the hypothalamus and brain stem. These psychologically induced influences can change the activity of the ANS to alter the state of the body and temporarily alter the homeostatic balance.

See Also the Following Articles

Adrenaline; Homeostasis; Regional Blood Flow, Stress Effects; Sympathetic Nervous System.

Further Reading

- Burnstock, G. (1975). Comparative studies of purinergic nerves. *Journal of Experimental Zoology* **194**, 103–133.
- Burnstock, G. and Hoyle, C. H. (eds.) (1992). *Autonomic neuroeffector mechanisms*. Reading, UK: Harwood.
- Iversen, S., Iversen, L. and Saper, C. B. (2000). The autonomic nervous system and the hypothalamus. In: Kandel, E. R., Schwartz, J. J. & Jessel, J. H. (eds.) *Principles of neural science* (4th edn., pp. 960–981). New York: McGraw-Hill.
- Karczmar, A. G., Koketsu, K. and Nishi, S. (eds.) (1986). *Autonomic and enteric ganglia: Transmission and its pharmacology*. New York: Plenum.
- Lovallo, W. R. (2005). *Stress and health: Biological and psychological interactions* (2nd edn.). Thousand Oaks, CA: Sage.
- Nilsson, S. and Holmgren, S. (eds.) (1914). *Comparative physiology and evolution of the autonomic nervous system*. Reading, UK: Harwood.
- Smith, E. E., Guyton, A. C., Manning, R. D. and White, R. J. (1976). Integrated mechanisms of cardiovascular response and control during exercise in the normal human. *Progress in Cardiovascular Disease* **28**, 421–443.

Autotolerance

M Durai

Johns Hopkins University School of Medicine,
Baltimore, MD, USA

K D Moudgil

University of Maryland School of Medicine, Baltimore,
MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
M E F Melo, K D Moudgil, and E E Sercarz, volume 1,
pp 284–290, © 2000, Elsevier Inc.

Self-Tolerance Is Executed through the Deletion and
Regulation of Autoreactive Lymphocytes

Cells Exposed to Stressful Stimuli Upregulate the
Expression of Stress Proteins

Tolerance to Stress Proteins Is Incomplete

Endogenous hsp60 Can Be Processed and Presented to
Antigen-Primed T Cells

Oxidative Stress Influences Several Components of the
Immune System Pertaining to Self-Tolerance and
Autoimmunity

Neural Immune Interactions and Their Relevance to
Autoimmunity

Immunoregulatory and Therapeutic Approaches Based on
Biomolecules and Intracellular Pathways Involved in the
Stress Response

Glossary

Autoimmunity The activity of the immune system targeted against body's own (self) components. In a healthy situation, the immune system protects the body against invading pathogens such as bacteria and viruses. However, as a result of a series of orchestrated events involving genetic and environmental factors, there is a break in self-tolerance and the immune system starts attacking the body's own cellular components, which, if uncontrolled, can lead to pathological damage and autoimmune disease.

Autoimmune diseases Diseases caused by autoimmunity; can involve multiple body tissues (systemic) or be restricted to a particular tissue (organ-specific). Examples are multiple sclerosis, rheumatoid arthritis, systemic lupus erythematosus, and thyroiditis.

Heat-shock proteins (Hsps) A group of intracellular proteins whose expression is induced by a variety of environmental stresses such as heat, cold,

Hypothalamic-pituitary-adrenal (HPA) axis

Oxidative stress

Reactive oxygen species (ROS)

Tolerance/self-tolerance

oxygen deprivation, and other physical and chemical stressors; also called stress proteins. Under normal healthy conditions, Hsps act like chaperones, binding to other cellular proteins and preserving their shape and function, and also participating in their transport within and outside the cell. Under stressful conditions, Hsps protect other cellular proteins from denaturation. Hsps also participate in the display on the cell surface of peptides derived from other proteins for presentation by the MHC molecules to cells of the immune system. The communication and interplay among the three endocrine tissues via specific chemical mediators secreted by these tissues. The hypothalamus secretes a hormone named corticotropin releasing hormone (CRH) in response to stress. CRH acts on the pituitary gland cells that secrete adrenocorticotrophic hormone (ACTH), which in turn acts on the adrenal gland, leading to the secretion of adrenal stress hormones (e.g., cortisol). The adrenal hormones initiate various metabolic events to counter the effect of stress, and they also provide negative regulatory feedback signals to the hypothalamus and the pituitary gland to regulate the secretion of CRH and ACTH. An altered HPA axis activity has been implicated in many diseases, including some autoimmune diseases.

A state of relative excess of reactive oxygen species (ROS) such as oxygen ions, free radicals, and peroxide within the cell. Oxidative stress is linked to disease processes including aging, atherosclerosis, and Alzheimer's disease.

Oxygen species such as oxygen ions, free radicals, and peroxide within the cell. The activity of ROS normally produced during various metabolic processes is controlled by naturally occurring antioxidants; however, in excess ROS are harmful, leading to damage to cellular components such as lipids, DNA, and proteins. Thus, ROS are involved in normal cellular functions, but they also can induce tissue damage leading to mutagenesis and cell death.

A state of immunological unresponsiveness to a self-antigen without compromising the ability of the immune system to

immune tolerance

respond to a foreign antigen. Tolerance is effected by the deletion or inactivation of potentially self-reactive T and B cells during their development in the thymus and bone marrow, respectively (central tolerance) as well as by the control of the activity of self-reactive T and B cells in the periphery (peripheral tolerance) via a variety of mechanisms. A state of tolerance against foreign antigens can also be achieved using the appropriate experimental conditions.

Self-Tolerance Is Executed through the Deletion and Regulation of Autoreactive Lymphocytes

Immunological tolerance is the basic property of the immune system that provides for self/nonself-discrimination so that the immune system can protect the host from external pathogens without reacting against itself. Central tolerance is achieved by the clonal deletion of self-reactive lymphocytes expressing receptors with high avidity for self-ligands. Autoreactive lymphocytes that have escaped negative selection in the central lymphoid organs (the thymus and the bone marrow) are present in the peripheral repertoire, but are kept under control by multiple diverse peripheral tolerance mechanisms. These include anergy (functional inactivation) induction in lymphocytes that recognize self-ligands (first signal) in the absence of appropriate costimulation (second signal) on nonprofessional antigen-presenting cells (APCs), activation-induced cell death (AICD), cytokine-mediated regulation and various types of regulatory cells. The failure of tolerogenic mechanisms either within the central lymphoid organs or in the periphery permits the activation and expansion of self-reactive lymphocytes, which in turn can lead to tissue damage and the manifestation of an autoimmune disease. Autoimmune diseases are broadly categorized into systemic and organ-specific disorders. Systemic lupus erythematosus (SLE) and multiple sclerosis (MS) are examples of systemic and organ-specific autoimmune diseases, respectively.

In the last several years, there has been increasing realization that the expression of a target autoantigen involved in an organ-specific autoimmune disease is not necessarily restricted to that particular tissue. Instead, the autoantigen is either expressed ubiquitously or also expressed within the thymus (ectopic expression). The latter observation represents a paradigm

shift in our understanding of the induction of self-tolerance to tissue-specific antigens. It is now known that many target self-antigens relevant to the pathogenesis of autoimmunity are efficiently expressed within the thymus and that their expression is controlled in part by a transcription factor named autoimmune regulator (*aire*). Furthermore, both humans and mice deficient in the *AIRE* gene exhibit signs of systemic autoimmunity. The human disorder is referred to as the autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy syndrome (APECED). In regard to the mechanisms of peripheral tolerance, a deficiency of a subset of T regulatory cells (Treg, CD4⁺CD25⁺ T cells) also results in autoimmune manifestations. These Treg are characterized by the expression of Foxp3, a unique transcription factor that is a member of the forkhead/winged-helix family of transcriptional regulators. Patients lacking the *FOXP3* gene are deficient in Treg and exhibit features of immune dysregulation, polyendocrinopathy, enteropathy, X-linked (IPEX) syndrome. Thus, in a healthy individual, both central and peripheral tolerogenic mechanisms play critical roles in the induction and maintenance of self-tolerance and, thereby, in the prevention of autoimmunity.

Cells Exposed to Stressful Stimuli Upregulate the Expression of Stress Proteins

Stress proteins (previously known as heat shock proteins; Hsps) are ubiquitous cellular proteins whose expression within cells is upregulated in response to a variety of stressful stimuli, for example, increased temperature, oxidative stress, nutritional deficiency, viral infection, tumor necrosis factor (TNF)- α , endotoxin, adrenergic stimuli, and ultraviolet irradiation. These stimuli may be part of environmental stress or pathological states. The Hsps thus induced protect cells from cellular injury including stress-induced apoptosis. Hsps are highly conserved in prokaryotes as well as eukaryotes, and they are grouped into several families based on their molecular weight ([Table 1](#)). Hsps are localized in various intracellular compartments and function as molecular chaperones, which take part in the assembly, folding, and translocation of cellular proteins ([Table 2](#)). Because Hsps are primarily intracellular proteins, their release from necrotic cells into the extracellular environment indicates nonphysiological tissue damage, and such release of Hsps can induce a range of biological responses.

Table 1 Mammalian heat shock protein (Hsp) families and their cellular location and function^a

<i>Hsp family/members</i>	<i>Localization</i>	<i>Functions</i>
Small Hsps		
Hsp10	Mitochondria	Co-chaperone for Hsp60 activities
α B-crystallin	Cytoplasm	Cytoskeletal stabilization
Hsp27	Nucleus/cytoplasm	Actin dynamics
Hsp32	Cytoplasm	Heme metabolism, antioxidant properties
Hsp40		
Hsp40	Nucleus/cytoplasm	Binds to non-native proteins, regulates the activity of Hsp70
Hsp47	Endoplasmic reticulum	Processing and/or secretion of collagen
Hsp60		
Hsp60	Mitochondria	Both involved in protein folding and assembly
TCP-1	Cytoplasm	
Hsp70		
Hsp70	Nucleus/cytoplasm	Hsp70 regulates HSF-1 activity and Hsp gene transcription
Grp78/Bip	Endoplasmic reticulum	All: Prevent aggregation of unfolded proteins
Grp75	Mitochondria	Bind ATP and show ATPase activity
Hsp90		
Hsp90 (α and β)	Cytoplasm	Bind to other proteins and regulate protein activity
		Assembly and folding of newly synthesized proteins
Grp94/gp96/Hsp100	Endoplasmic reticulum	Prevent aggregation of refolded proteins
Hsp110		
Hsp110 (human)	Cytoplasm/nucleolus	Thermal tolerance
Apg-1 (mouse)	Cytoplasm	Protein refolding
Hsp105	Cytoplasm	

^aApg-1, Autophagy-1, a protein kinase essential for autophagy; ATP, adenosine triphosphate; Bip, immunoglobulin heavy chain binding protein; Grp, glucose-regulated protein; HSF-1, heat-shock factor 1; Hsp, heat shock protein; TCP-1, tailless complex polypeptide.

Table 2 Effects of heat shock proteins on the immune system^a

<i>Immunological effects</i>	<i>Hsp involved</i>	<i>Specific action and cell types</i>
Antigen processing and presentation	Hsp70	Peptide loading and antigen processing of MHC class I molecules
		Interaction with misfolded MHC class II and retention in the ER
	Hsp90	Peptide transport and loading of MHC class I in the ER
	Bip and Calnexin	Assembly of MHC class I and class II molecules in the ER
Expression of intercellular signaling molecules	gp96	Interaction with misfolded MHC class II and retention in the ER, peptide loading of MHC class II molecules
		ICAM-1 and VCAM-1 expression on vascular endothelial cells
	Mycobacterial Hsp60	
	Chlamydial and human Hsp60	E-selectin, ICAM-1, and VCAM-1 expression on vascular endothelial cells
Cytokine production	Hsp60	Production of IL-1 α , IL-6, IL-12, IL-15, nitric oxide, and TNF- α by monocytes and macrophages
		IFN- γ production by T cells
	Hsp60 and Hsp70	IL-6 production by vascular endothelial cells
		IL-6 production by smooth muscle cells
Target cell lysis	Hsp70	Tumor cells
Lymphocyte homing	Ubiquitin	Lymphocyte receptor involved in the interaction with high endothelial venules, thereby in lymphocyte homing

^aBip, immunoglobulin heavy chain binding protein; ER, endoplasmic reticulum; Hsp, heat shock protein; ICAM-1, intercellular adhesion molecule-1; IL, interleukin; INF, interferon; MHC, major histocompatibility complex; TNF, tumor necrosis factor; VCAM-1, vascular cell adhesion molecule-1.

Hsps can influence several functions of the immune system such as antigen processing and presentation, cytokine production, expression of intercellular signaling molecules, lymphocyte trafficking, and target cell lysis (Table 3). Hsps also can serve as

immunodominant antigens, and one hypothesis underlying the immune alterations induced by Hsps invokes immunological cross-recognition between the mammalian and the microbial homologs.

Table 3 Heat shock proteins and their immunoregulatory effect in autoimmune diseases^a

<i>Diseases</i>	<i>Strain/immunogen</i>	<i>Hsp involved</i>
Animal models		
Adjuvant-induced arthritis	Lewis rats/heat-killed <i>M. tuberculosis</i>	Mycobacterial Hsp60, Hsp70 and Hsp10 Self (rat) Hsp60 Human Hsp60
	Lewis rats/DDA, a synthetic adjuvant	Mycobacterial Hsp60 Self (rat) Hsp60
Pristane-induced arthritis	Mice/pristane	Mycobacterial Hsp60 Self (mouse) Hsp60
Collagen-induced arthritis	Lewis rats/collagen type II Mice/collagen type II	Mycobacterial Hsp60 Mycobacterial Hsp60 and Hsp70
Streptococcal cell wall-induced arthritis	Lewis rats/cell walls of streptococci	Mycobacterial Hsp60
Avridine-induced arthritis		Mycobacterial Hsp60 and Hsp70
Insulin-dependent diabetes mellitus	Nonobese diabetic mice/spontaneous	Mycobacterial Hsp60 Self (mouse) Hsp60 Human Hsp60 and Hsp90
Experimental autoimmune encephalomyelitis	Lewis rats/myelin basic protein	Mycobacterial Hsp60 and Hsp10
Human		
Rheumatoid arthritis	—	Mycobacterial Hsp60 Human Hsp60 and Hsp70 Bip
Juvenile chronic arthritis	—	Human Hsp60
Type I diabetes mellitus	—	Human Hsp60, Hsp70, and Hsp90
Multiple sclerosis	—	Human Hsp60 and Hsp70 α B-crystallin
Systemic lupus erythematosus	—	Human Hsp60 and Hsp90
Hashimoto's thyroiditis	—	Human Hsp60

^aBip, immunoglobulin heavy chain binding protein; DDA, dimethyl dioctadecyl ammonium bromide; Hsp, heat shock protein.

Tolerance to Stress Proteins Is Incomplete

Despite the existence of several mechanisms for tolerance induction against self-proteins, the presence of self-Hsp-specific T and B lymphocytes in the mature repertoire has been reported in healthy individuals. The presence of self-hsp60-specific T cells was demonstrated in inbred Lewis rats by immunization with mammalian hsp60 using either vaccinia virus or DNA encoding human hsp60 or using rat hsp60 that is 98% homologous to human hsp60. In this context, it has been shown that Hsps are expressed at high levels in the normal thymus. Moreover, the transgenic overexpression of mouse hsp60, in the thymus failed to eliminate T-cell responses to self-hsp60 following hsp60 immunization. Thus, in contrast to other antigens that are tolerogenic at low levels of thymic expression, hsp60 failed to delete antigen-specific T cells even at high levels of expression within the thymus.

Endogenous hsp60 Can Be Processed and Presented to Antigen-Primed T Cells

Although Hsps are intracellular proteins, they can be efficiently expressed on the cell surface as peptide–

major histocompatibility complex (MHC) complexes. For example, the T cells raised in the Lewis rat against the disease-protective C-terminal epitopes of rat hsp60 can be restimulated *in vitro* in a proliferation assay when cocultured with heat-stressed but not unstressed syngeneic APCs in the absence of exogenous antigen, suggesting that the endogenous hsp60 can both be upregulated and be processed and presented by the APCs to activate antigen-primed T cells. The processing and presentation of hsp60 via class I pathway has also been documented. Cytotoxic T lymphocytes (CTL) specific for mycobacterial hsp60 were shown to recognize stressed macrophages in the absence of exogenously added antigens, suggesting that endogenous self-hsp60 is presented in the context of MHC class I molecules to the CTL. Furthermore, the CTL-mediated lysis of stressed macrophages was suppressed by hsp60-directed antisense oligonucleotides, thereby confirming the specificity of the CTL for hsp60.

The breakdown of self-tolerance leading to the induction of antiself immune response is typically associated with the initiation of autoimmunity. However, studies by different investigators have shown that antiself-hsp60 immunity may be immunosuppressive and beneficial against autoimmunity. Lewis rats

immunized with either self-hsp60 (rat hsp60) or its homologous human hsp60 are protected against subsequent induction of adjuvant arthritis (AA). Furthermore, this protection can be adoptively transferred to naïve Lewis rats through the T cells of protected rats. Similarly, other Hsps in addition to hsp60 possess immunoregulatory activity.

Oxidative Stress Influences Several Components of the Immune System Pertaining to Self-Tolerance and Autoimmunity

Oxidative stress results from the increased production of oxidants, reduced generation of antioxidants, or failure to repair an oxidative damage. The damage to cells/tissues is caused by reactive oxygen species (ROS) comprised of free radicals, peroxides, and so forth (e.g., hydroxyl radical, hydrogen peroxide, and peroxynitrite). ROS can be derived from endogenous (e.g., metabolic reactions within the mitochondria or the liver detoxification pathways) or from exogenous (e.g., environmental pollutants or microbial infections) sources. ROS produced during oxidative stress can modulate a wide variety of cellular functions through the regulation of NF- κ B. Some of the processes influenced by oxidative stress and ROS are summarized in [Figure 1](#).

It has been observed that oxidative stress induces hyporesponsiveness in T lymphocytes, which in turn

correlates with ROS-dependent protein alterations including structural modification of crucial T-cell receptor (TCR) signaling molecules. Also reported is the inhibition of activated/memory T cells by oxidative stress, which is associated with a block of NF- κ B activation. In addition, a predominant decrease in Th1 (but not Th2) cytokine production was observed in the memory/effector subset compared to naïve T cells. ROS can also influence the activity of the major professional APCs such as dendritic cells (DCs). It is known that immature DCs are tolerogenic, whereas mature DCs are immunogenic. In this context, it has been shown that ROS can significantly decrease the tolerogenic activity of DCs. In one study, this inhibitory effect of ROS was shown to be overcome by treatment with soluble CTLA-4. The precise significance of this interrelationship has not yet been fully defined.

Oxidative stress and ROS have been implicated in the pathogenesis of various autoimmune diseases, including arthritis and diabetes mellitus. Oxidative stress can modulate the intracellular levels of G-protein-coupled receptor kinases (GRKs), which in turn influence the activity of G-protein-coupled receptors (GPCRs). It has been observed that the levels of GRK2 and GRK6 are decreased in lymphocytes of patients with rheumatoid arthritis (RA) as well as of rats with AA. Similarly, decreased GRK2 levels have also been found in splenocytes of rats with experimental autoimmune encephalomyelitis (EAE).

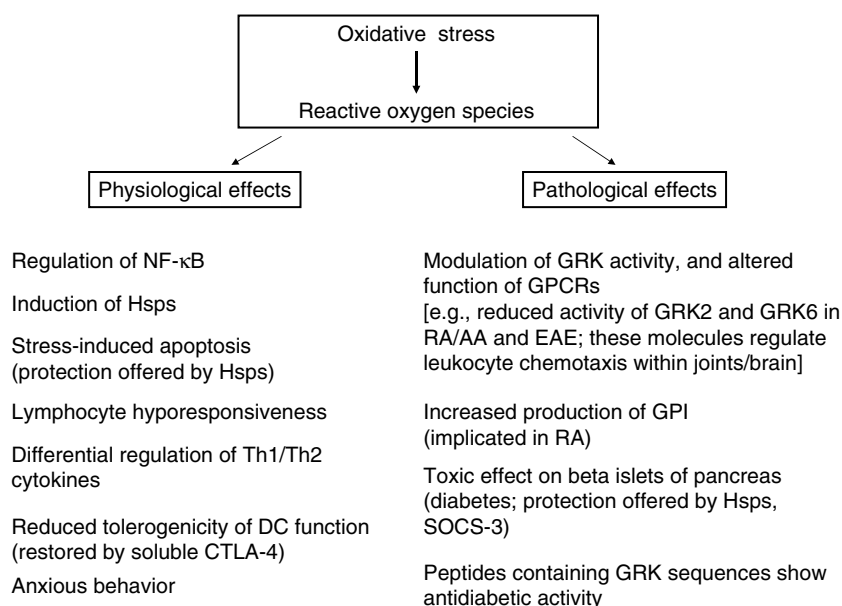


Figure 1 Oxidative stress and its influence on physiological and pathological immune processes. Oxidative stress has been implicated in a large variety of biological reactions. AA, adjuvant arthritis; CTLA-4, cytotoxic T lymphocyte associated-4; DC, dendritic cells; EAE, experimental autoimmune encephalomyelitis; GPCR, G-protein-coupled receptor; GPI, glucose-6-phosphate isomerase; GRK, G-protein-coupled receptor kinase; Hsp, heat-shock protein; NF- κ B, nuclear factor kappa B; RA, rheumatoid arthritis; SOCS, suppressor of cytokine signaling; Th1, T helper 1; Th2, T helper 2.

GRKs are involved in modulating leukocyte chemotaxis into the target organs that constitute the sites of autoimmune inflammation (e.g., the joints in RA and the brain in EAE). Interestingly, peptides containing defined sequences of the kinase-substrate interaction site of GRK2/3 show antidiabetic activity in animal models of type 2 diabetes, supporting the likely therapeutic significance of GRKs in some other diseases as well.

Hypoxia is a characteristic feature of joints observed in both RA patients as well as experimental models of arthritis. Interestingly, it is associated with increased expression of glucose-6-phosphate isomerase (GPI), which has been shown to be one of the target autoantigens in autoimmune arthritis in K/BXN mice. These mice possess a TCR that is transgenic for an epitope within ribonuclease, but it fortuitously cross-reacts with the ubiquitously expressed protein, GPI. An immune response to GPI has been observed in K/BXN mice as well as in RA patients, particularly those with extra-articular complications.

ROS have been shown to have a toxic effect on the β islets of the pancreas, and thereby they have been implicated in the pathogenesis of type 1 diabetes mellitus. Interesting, Hsps can offer protection against ROS-mediated damage to β islets. Another protective pathway in β islets involves the suppressor of cytokine signaling (SOCS)-3. Similarly, SOCS-1 plays a role in the pathogenesis of arthritis, and SOCS-1 deficiency results in severe synovial inflammation and joint destruction. SOCS proteins can be induced by a variety of stimuli, particularly various cytokines, and they regulate cytokine signaling as well as a wide variety of pathways relevant in inflammatory diseases of infectious and autoimmune origin. The pro-inflammatory cytokine interferon (IFN- γ) induces the expression of SOCS-1 and -3.

Oxidative stress may also contribute to an anxious behavior. Furthermore, the levels of two enzymes (glyoxalase 1 and glutathione reductase) that influence oxidative stress also positively correlate with the level of anxiety. Considering the relationship between stress and autoimmunity, this finding provides one of the molecular links between stress and behavior (anxiety) associated with autoimmunity.

Neural Immune Interactions and Their Relevance to Autoimmunity

Several lines of evidence emphasize the interplay between, and cross-regulation of, the activities of the nervous system and the immune system. The nervous system influences the immune system primarily via the hypothalamic-pituitary-adrenal (HPA) axis, the

hypothalamic-pituitary-gonadal (HPG) axis, and the chemical mediators of the autonomic nervous system, whereas the immune system modulates the activities of the neural system via cytokines and the vagus nerve. In experimental models of autoimmune diseases and human subjects, an inadequately functioning or blunted HPA axis has been associated with increased susceptibility and/or enhanced severity of autoimmune diseases such as RA and MS. In some cases, the baseline HPA axis activity in the disease-susceptible versus disease-resistant individuals was comparable, but the dysfunction of the HPA axis was manifest on an exposure to the stressful or inflammatory stimuli. Because the stress response intimately involves the components of the HPA and HPG axes, it is apparent that chronic physical and psychosocial stresses constitute one of the factors that have been implicated in the initiation and aggravation of an autoimmune condition.

The underlying hypothesis involving the HPA axis is that susceptibility to autoimmune diseases may be related to inadequate responsiveness of the HPA axis to stressful stimuli such that there is a decreased secretion of cortisol and other biomolecules that are normally required to downregulate the immune responses. Consequently, unchecked immune reactivity leads to the induction of anti-self-directed T cell and/or antibody responses, which if propagated, may eventually lead to autoimmunity. In support of this proposition, studies in the rat model of AA have revealed that susceptible rat strains (e.g., Lewis and DA rats) have impaired HPA axis compared to resistant rat strains (e.g., Fischer F344 and BN rats) that display an aggravated HPA axis response. Interestingly, chemical/pharmacological or surgical manipulations aimed at altering the HPA axis activity in these rats also resulted in a corresponding change in the pattern of susceptibility/resistance to AA. The complementation of the HPA axis activity by the administration of glucocorticoids or transplantation of the hypothalamus from the resistant rat strain into susceptible rat strain led to acquired resistance to AA, whereas interference with HPA axis activity rendered the resistant rat strain susceptible to AA. Such differences in HPA axis activity in response to stress or inflammatory stimuli have also been observed in male versus female rats, with the HPA axis response impaired in female rats relative to male rats. These results suggest that differences in HPA axis responsiveness might contribute to the higher incidence of autoimmune diseases observed in females compared to males. An additional link between gender-related hormones and susceptibility to autoimmunity is provided by the observation that androgens exert a suppressive, anti-inflammatory effect on T-cell and

antibody responses, whereas estrogens tend to enhance immune reactivity, particularly antibody responses.

The evidence to support a direct association between HPA axis and autoimmune susceptibility in patients with autoimmune diseases is not as clear as in animal models of these diseases. Nevertheless, clinical histories and the examination of the parameters reflecting HPA axis activity support the idea that HPA axis activity is impaired in diseases such as RA and MS. For example, the decreased secretion of cortisol following HPA axis stimulation, reduced number of glucocorticoid receptors on specific cell types, decreased androgen levels, impaired growth hormone response, and inadequately low ratio of steroid hormones to interleukin (IL-)6 and TNF- α have been observed in RA patients. In addition to the HPA axis, the dysregulated autonomic nervous system is also a critical modifier of autoimmune susceptibility and chronicity of the autoreactive immune response. A cholinergic anti-inflammatory pathway involving the vagus nerve is believed to represent a physiological pathway for the detection of inflammatory stimuli via the afferent vagus and the corresponding regulation of the cytokine (TNF- α) response of macrophages via the efferent vagus. The nicotinic acetylcholine receptor α 7 subunit expressed on the macrophage plays a critical role in this pathway. Taken together, it is evident that other factors in addition to the HPA axis and the autonomic nervous system, such as genetic and environmental factors, also play an important role in determining susceptibility/resistance to autoimmunity.

Immunoregulatory and Therapeutic Approaches Based on Biomolecules and Intracellular Pathways Involved in the Stress Response

1. *Altering the cellular levels of Hsps.* Taking into consideration the chaperone function of Hsps in protecting other proteins from damage in stressed cells as well as the finding that anti-Hsp immunity can be immunoregulatory in nature, the expression of cellular Hsps can be induced physiologically or pharmacologically, or by direct administration of a specific Hsp for a beneficial effect in suppressing autoimmunity. For example, the levels of endogenous Hsps can be increased by hyperthermia, geldanamycin, geranylgeranylacetone, cyclopentenone prostanoids, certain heavy metals, salicylates, or dexamethasone. On the other hand, the expression of Hsps can be decreased by retinoic

acid, bioflavonoid compounds, or 15-deoxyspergualin, and this approach may prove useful in augmenting immunity as desired for antitumor immunity.

2. *The use of antioxidants.* A wide variety of synthetic and naturally occurring compounds have been used to counter the cellular effects of oxidative stress and ROS. These compounds possess antioxidant properties. Vitamin E, vitamin C and β -carotenes are well-known antioxidants. The extracts of certain plants have also been found to possess potent antioxidant activity by scavenging free radicals leading to anti-arthritis activity in the rat model of AA. These include pumpkin seed oil, *Semecarpus anacardium*, and *Trewia polycarpa*. In addition, Meloxicam modulates oxidation status and is beneficial in suppressing AA.
3. *Interventions aimed at HPA-HPG axes and the autonomic system.* Pharmacological and surgical procedures have been employed in experimental systems to reconstitute or enhance the activity of the HPA and HPG axes and to alter the activity of the autonomic system and cytokine (TNF- α) secretion by macrophages via the stimulation of the vagus nerve. The administration of androgens to strike a sex hormone balance is another modality that is being exploited for the control of RA, particularly in male patients.
4. *Targeting the stress signaling pathways.* The synovial membrane of arthritic joints is under a variety of stresses: oxidative stress, pro-inflammatory cytokines, heat-stress, and shear stress. Multiple cell surface-resident and intracellular molecules are involved in stress signaling pathways in the milieu of inflammatory synovitis, and these signaling pathways represent novel and promising targets for the control of RA and other autoimmune diseases.

Further Reading

- Anis, Y., Leshem, O., Reuveni, H., et al. (2004). Antidiabetic effect of novel modulating peptides of G-protein-coupled kinase in experimental models of diabetes. *Diabetologia* 47, 1232–1244.
- Durai, M., Gupta, R. S. and Moudgil, K. D. (2004). The T cells specific for the carboxyl-terminal determinants of self (rat) heat-shock protein 65 escape tolerance induction and are involved in regulation of autoimmune arthritis. *Journal of Immunology* 172, 2795–2802.
- Eskandari, F., Webster, J. I. and Sternberg, E. M. (2003). Neural immune pathways and their connection to inflammatory diseases. *Arthritis Research and Therapy* 5, 251–265.

- Griffiths, H. R. (2005). ROS as signalling molecules in T cells – evidence for abnormal redox signalling in the autoimmune disease, rheumatoid arthritis. *Redox Report* 10, 273–280.
- Harbuz, M. (2002). Neuroendocrinology of autoimmunity. *International Review of Neurobiology* 52, 133–161.
- Kavelaars, A., Vroon, A., Raatgever, R. P., et al. (2003). Increased acute inflammation, leukotriene B₄-induced chemotaxis, and signaling in mice deficient for G protein-coupled receptor kinase 6. *Journal of Immunology* 171, 6128–6134.
- Lombardi, M. S., Kavelaars, A. and Heijnen, C. J. (2002). Role and modulation of G protein-coupled receptor signaling in inflammatory processes. *Critical Reviews in Immunology* 22, 141–163.
- Malmberg, K. J., Arulampalam, V., Ichihara, F., et al. (2001). Inhibition of activated/memory (CD45RO(+)) T cells by oxidative stress associated with block of NF-kappaB activation. *Journal of Immunology* 167, 2595–2601.
- Moseley, P. (2000). Stress proteins and the immune response. *Immunopharmacology* 48, 299–302.
- Moudgil, K. D., Chang, T. T., Eradat, H., et al. (1997). Diversification of T cell responses to carboxy-terminal determinants within the 65-kD heat-shock protein is involved in regulation of autoimmune arthritis. *Journal of Experimental Medicine* 185, 1307–1316.
- Piccirillo, C. A. and Shevach, E. M. (2004). Naturally-occurring CD4+CD25+ immunoregulatory T cells: central players in the arena of peripheral tolerance. *Seminars in Immunology* 16, 81–88.
- Pockley, A. G. (2003). Heat shock proteins as regulators of the immune response. *Lancet* 362, 469–476.
- Quintana, F. J. and Cohen, I. R. (2005). DNA vaccines coding for heat-shock proteins (HSPs): tools for the activation of HSP-specific regulatory T cells. *Expert Opinion on Biological Therapy* 5, 545–554.
- Steinhoff, U., Brinkmann, V., Klemm, U., et al. (1999). Autoimmune intestinal pathology induced by hsp60-specific CD8 T cells. *Immunity* 11, 349–358.
- van Eden, W., van der Zee, R. and Prakken, B. (2005). Heat-shock proteins induce T-cell regulation of chronic inflammation. *Nature Reviews Immunology* 5, 318–330.
- Venanzi, E. S., Benoist, C. and Mathis, D. (2004). Good riddance: thymocyte clonal deletion prevents autoimmunity. *Current Opinion in Immunology* 16, 197–202.
- Zugel, U. and Kaufmann, S. H. (1999). Role of heat shock proteins in protection from and pathogenesis of infectious diseases. *Clinical Microbiology Reviews* 12, 19–39.

Avoidance

M S Oitzl

University of Leiden, Leiden, The Netherlands

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by B Bohus, volume 1, pp 291–295, © 2000, Elsevier Inc.

Avoidance Conditioning

Active Avoidance Conditioning

Inhibitory Avoidance Conditioning

Fear Conditioning

Avoidance in Psychiatric Disorders

Closing Comments

Glossary

Active avoidance

A neutral conditioned stimulus (CS; tone, light) is used to announce an aversive unconditioned stimulus (US; electric footshock). In response to the CS, animals learn and perform spatial motor patterns in order to avoid deleterious consequences of the US.

Anxiety

State of intense apprehension, uncertainty, and fear resulting from the anticipation of a threatening event or situation, often to a degree that normal physical and psychological functioning is disrupted.

Approach-avoidance conflict

Situation in which the goal has both positive and negative factors. It is the most typical short term conflict encountered and the most anxiety provoking.

Coping

An individual's behavioral and physiological response repertoire to master a challenge.

Fear

Feeling of agitation and dread caused by the presence or imminence of danger.

Fear conditioning

A neutral conditioned stimulus (CS; tone, light, context) is paired with an aversive unconditioned stimulus (US; electric footshock, air blast), that cannot be avoided. After a single or a few pairings, the CS alone is capable to induce behavioral and physiological fear responses.

Innate (inborn) fear

Needs no previous experience, is often species specific and directed towards actual or potential threats (heights, light, smell, taste, predator).

<i>Inhibitory avoidance (also known as passive avoidance)</i>	The inhibition of innate activities or learned habits, which have led to aversive consequences.
<i>Negative reinforcement</i>	The withdrawing of aversive stimuli (actual or imagined) after the individual has made the appropriate response.

Avoidance Conditioning

Learning about the environment, its physical and social relationships includes the conditioning of stress-induced avoidance behavior, its neuroendocrine and other physiological components. Aversive events are actively avoided. Individuals use distinct behavioral patterns that are determined by genetics, (early) life experiences, and the environment.

Coping strategies serve to reduce or eliminate stress. As an innate defense mechanism, avoidance behavior protects the organism from harm, allows learning of new features of the social and physical environment, and is thus, central for survival and development of the individual. However, when getting pathological, i.e., out of context and magnitude, the price of avoidance is a denial or distortion of reality.

Principles of Avoidance Conditioning

Avoidance and fear are closely related. Placing an animal in a novel environment activates the stress system (surge of corticosteroids and catecholamines), initiates exploratory activity, which is, at the same time, opposed by the fear evoking properties of the situation, e.g., open space, bright light, and another animal. This apparently simple event of novelty initiates an approach–avoidance conflict. Opposite avoidance responses might be observed: immobility (freezing, crouching) or movement (fight-or-flight responses; running, jumping).

Animals and humans learn about the spatial and temporal characteristics of the environment and their emotional consequences. Electric shock is the most frequently used aversive stimulus (stressor) in the laboratory. Avoidance paradigms provide excellent models of specific and successful behavioral coping with stress.

Inhibitory avoidance conditioning requires the inhibition of innate activities or learned habits, which have led to aversive consequences. It is one of the most common modifications of behavior induced by aversive experience. In active avoidance conditioning, individuals learn that signals announce an aversive stimulus. The aversive consequences of this stimulus can not only be terminated with the appropriate active behavior, but also avoided completely by responding with this active behavior already to the

signal. The terminology, inhibitory and active avoidance is somewhat confusing since both paradigms require active processes to avoid and thus reduce the (anticipated) adverse consequences.

Avoidance and Stress Hormones

A prominent hormonal stress response, namely the rise in plasma corticosterone as a consequence of the activation of the hypothalamic-pituitary-adrenal (HPA) axis, is not only a useful marker of the stress experienced by the individual, but also predictive for the strength of the avoidance conditioning and crucial for the extinction of avoidance behavior. Interindividual differences have to be taken into account.

Experiments performed in the 1970s demonstrated that avoidance learning and its extinction are strongly affected by stress hormones. Acquisition as well as extinction of avoidance behavior is facilitated by corticosterone, while the lack of corticosterone (due to removal of the adrenals) inhibits the extinction process of a conditioned avoidance response. Complex interactions between corticosteroids and catecholamines have been described then and are discussed with new interest nowadays.

Active Avoidance Conditioning

The first stage of avoidance learning is usually escape, whereby the reaction of the animal terminates the encounter with the aversive stimulus; i.e., negative reinforcement due to shock reduction. With continued training, anticipatory reactions appear: the animal already responds to the stress-announcing stimulus, and thus, avoids the shock altogether. Due to the total elimination of aversive consequences fast extinction of avoidance behavior might be expected. This is not the case: active avoidance responses are very persistent. It is thought that anticipation of the signal stimulus elicits a motivational state of fear and the subsequent avoidance response decreases fear, i.e., negative reinforcement due to fear reduction. Taken together, negative reinforcement processes are thus, responsible for the acquisition and maintenance of active avoidance behavior.

Spatial paradigms have been developed to study active avoidance in rats and mice. For example, in the shuttle box, the rodent can avoid painful electrical footshock by shuttling from one (aversive, shock) compartment to the other (safe, no shock). Usually, the occurrence of the shock is preceded by a tone or light stimulus, so that the animal can learn to predict the aversive properties of the compartment, and move to the safe side. Since there is no permanently safe area, it takes the animals many trials to learn. Avoidance is acquired faster in the run-way apparatus,

which has fixed safe areas. It is essential to all active avoidance paradigms that the animal learns as soon as possible that the aversive situation can be escaped. Otherwise, maladaptive habits are conditioned.

The acquisition (learning) phase of active avoidance behavior in rats is paralleled by high concentrations of the plasma corticosterone. During perfect avoidance, plasma corticosterone levels decrease and remains low. Evidence suggests that active avoidance behavior reduces the aversive character of the situation. Interesting is the fact that preventing the rats from making the avoidance response leads to a high rise in plasma corticosterone, even if no further footshock is experienced. In the 1970s, observations in rats showed that behavioral controllability and predictability of a stressful situation determined the magnitude of the rise in plasma corticosterone levels. If rats were able to avoid painful electroshocks or to predict their occurrence, plasma corticosterone did not rise as high as in the case of unavoidable or unpredictable shock. The latter condition is used in the so-called learned helplessness paradigm that is considered as an animal model for depression and also posttraumatic stress disorder (PTSD).

Inhibitory Avoidance Conditioning

Innate behavior or learned habits that had been paired with aversive consequences in the past will not be expressed. In the future, inhibition of the innate behavior will keep the unpleasant consequences away, i.e., negative reinforcement. Inhibitory avoidance is a highly emotional learning situation. In fact, it is a conflict situation for the individual, which will display patterns of approach and avoidance of the space or object of preference (approach-avoidance conflict).

In contrast to the extended training period required for active avoidance learning, inhibitory (or passive) avoidance paradigms are usually based on a single learning trial followed by a retention test at defined intervals after training. They have the advantage that neurobiological processes underlying learning and memory are defined within a certain timeframe and thus, can be estimated and modulated accordingly. It is a real pro for studying molecular mechanisms of avoidance.

Experimental paradigms in rodents use the innate preference, for example, dark-to-light (step-through task), high-to-low places (step-down task) followed by an aversive stimulus like footshock. The animal will actively control – inhibit – the performance of the innate pattern in the future. To gain reproducible results, the active behavior must be well defined and the aversive stimulus closely associated in time

with the innate behavior. Immobility behavior is observed, but it is not the typical avoidance response. Latency of avoidance is the most frequently used measurement.

A distinct paradigm of inhibitory avoidance is the so-called conditioned taste aversion. This is based on bait shyness, i.e., limited ingestion of novel foods and the association of their properties with delayed gastrointestinal consequences, and preference for sweet taste. The innate preference of rodents for sweet taste is followed by illness induced by substances like lithium, cyclohexamide, but also amphetamine and morphine. A single association between the sweet-tasting substance and the illness-inducing drug induces avoidance of the preferred taste. Conditioned taste aversion can also be studied in 1-day-old chicks.

Activation of the stress systems takes place in response to the aversive stimulus, affecting memory consolidation. Both corticosteroids and catecholamines are important for optimal memory performance. Specifically increased corticosteroid hormone levels within the context and close-in-time to the learning situation facilitate the memory of the negative event. The result is an active behavioral stress management, the inhibition of the innate/preferred response which is observed during retention test.

Fear Conditioning

Fear-conditioning paradigms are used to investigate the conditioning of emotional processes. During conditioning, the individual cannot avoid the aversive stimulus, which is usually announced by a signal. During retention testing, the signal itself elicits the fear response, which is invariably immobility (freezing), accompanied by high blood pressure and tachycardia.

In fear-conditioning paradigms, a meaningless stimulus (tone or light) is paired with an electric footshock. After one or several pairings, the tone or light becomes the conditioned stimulus. The retention test takes place in the absence of the footshock, at defined times after training. While in so-called classical or Pavlovian conditioning only defined cues are used, fear-conditioning paradigms have incorporated the complexity of the physical or social environment. These latter paradigms are called contextual fear conditioning. Anatomical pathways are well described (emotional response circuits, including amygdala, brainstem nuclei, hippocampus, and prefrontal cortex). Likewise, molecular mechanisms have been reported. Knowledge of neuroendocrine changes include the rise in plasma prolactin, adrenocorticotropin, corticosterone, and catecholamines. Plasma

corticosterone levels in rats are positively correlated with the extent of contextual conditioned fear.

Freezing is of obvious adaptive value. It not only decreases the probability of acoustic or visual detection, but also conserves energy in situations where action is inefficient, like in fear conditioning. Formerly, freezing was considered as a nonfunctional coping response. Individuals may use one or the other coping strategy, depending on genetic make-up and life experiences, influenced by the physical and social environment. Active avoidance paradigms select for active fight-or-flight strategies, while fear-conditioning paradigms select for passive, conservation and withdrawal-oriented approaches.

Avoidance in Psychiatric Disorders

Avoidance is a central feature in anxiety disorders like generalized anxiety, phobias, PTSD, and obsessive-compulsive disorder (OCD).

A person suffering from OCD is subjected to actions and/or thoughts that he/she finds undesirable but is unable to prevent. Simple gestures to complex rituals are carried out to avoid fear or unacceptable impulses. Any behavior that is associated with the reduction or avoidance of fear tends to be highly persistent in the presence of the stimuli involved. Preventing the performance of the obsessive-compulsive action or thoughts, results in acute discomfort or panic, similar to animals, which are prevented to display their conditioned avoidance response.

According to the *Statistical Manual of Mental Disorders – 4th Edition* (DSM-IV), the characteristic symptoms of PTSD, intrusion, avoidance, and hyperarousal result from previous exposure to an extreme trauma and include persistent re-experiencing of the traumatic event. For example: the person commonly makes deliberate efforts to avoid thoughts, feelings, or conversations about the traumatic event and avoiding activities, situations, or people who arouse recollections of it. Interference and deterioration of interpersonal relationships are the consequences.

Animal Models

Each of the three conditioning paradigms, active and inhibitory avoidance, and fear conditioning, provide aspects for potential animal models of various forms of human anxiety disorders. At present, none of the known animal models stands exclusively for one psychiatric disorder as they are currently defined in the DSM. For example, learned helplessness, the response to inescapable and uncontrollable shock, had initially received the greatest attention as animal model for PTSD. Nowadays, it is more accepted as model for depression. The (causal) relation between stress

hormones, PTSD, and the design of animal models is in the center of interest of preclinical science and clinical applications.

The wide variety and complexity of avoidance symptoms in anxiety-related disorders supports the general belief of underlying avoidance and fear-conditioning processes, indicating at the same time the difficulty for efficient behavioral and cognitive therapeutical approaches.

Closing Comments

Avoidance has been discussed mainly as an adaptive response of the individual to environmental challenges. Fear-related avoidance behaviors are of high relevance for physical survival and social competence. While inhibitory avoidance is as fast acquired as fear conditioning (one trial is often enough), active avoidance needs training (repetitions). Conditioned responses are hard to extinguish. Out of a variety of anxiety and other psychiatric disorders, PTSD and OCD were selected as typical examples with symptoms of pathological avoidance. It is most likely that physiologically functional and pathologically maladaptive avoidance use the same brain circuits. Animal models can contribute and elucidate the (molecular) mechanisms of anxiety. Stress and specifically the (re)activity of corticosteroid hormones are critically involved in the acquisition and persistence of conditioned avoidance and fear responses, thus proving their relevance for the development and course of several psychiatric diseases with core symptoms of avoidance.

Acknowledgments

In memory of Bela Bohus (former author of this chapter).

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Animal Models (Nonprimate) for Human Stress; Anxiety; Coping and Stress: A Lens and Filter Model; Fear; Learned Helplessness; Obsessive-Compulsive Disorder; Posttraumatic Therapy.

Further Reading

- Bohus, B. (1998). Neuroendocrinology of stress: behavioral and neurobiological considerations. In: Van de Kar, L. D. (ed.) *Methods in neuroendocrinology*, pp. 163–180. Boca Raton, FL: CRC Press.
- Bureš, J., Burešová, O. and Huston, J. P. (1983). *Techniques and basic experiment for the study of brain and behavior* (2nd edn.). Amsterdam: Elsevier.

- De Kloet, E. R. and Oitzl, M. S. (2006). Cortisol and PTSD: animal experiments and clinical perspective. In: Kato, N., Kawata, M. & Pitman, P. B. (eds.) *PTSD – brain mechanisms and clinical implications*, pp. 13–27. Tokyo: Springer-Verlag.
- Joels, M., Pu, Z., Wiegert, O., et al. (2006). Learning under stress: how does it work? *Trends in Cognitive Sciences* 10, 152–158.
- McGaugh, J. L. (2004). The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annual Review in Neuroscience* 27, 1–28.
- Phelps, E. A. and LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* 48, 175–187.
- Rodrigues, S. M., Schafe, G. E. and LeDoux, J. E. (2004). Molecular mechanisms underlying emotional learning and memory in the lateral amygdala. *Neuron* 44, 75–91.
- Yehuda, R. and Antelman, S. M. (1993). Criteria for rationally evaluating animal models of posttraumatic stress disorder. *Biological Psychiatry* 33, 479–486.

B

Bacterial Infections See: Infection.

Baroreceptors See: Arterial Baroreflex.

Bed Wetting See: Enuresis.

Behavior Therapy

P de Silva

Institute of Psychiatry, University of London,
London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1,
pp 300–303, © 2000, Elsevier Inc.

*Learning
theory*

The body of theory that exists in experimental psychology that attempts to account for how animal and human behavior change takes place.

*Operant
conditioning*

A basic form of learning where a response is strengthened or weakened by systematically manipulating the consequences of the response, such as providing a reward.

Introduction

Historical Development

Techniques of Behavior Therapy

Practice of Behavior Therapy

Conclusions

Glossary

*Behavioral
analysis*

The detailed analysis of a behavior, which is considered essential before treatment is undertaken.

*Classical
conditioning*

A basic form of learning where a neutral stimulus, which is presented repeatedly with a stimulus that elicits a particular response, eventually begins to elicit that response.

Introduction

The term behavior therapy is commonly used to refer to a particular approach to the treatment of psychological disorders. It has several key features. Behavior therapy sees many psychological problems as instances of maladaptive learning or, in some cases, failure to learn. The focus in behavior therapy is the presenting problem itself. It is not assumed that the presenting problem is a manifestation of an underlying primary problem. The aim of behavior therapy is to modify the problem behavior at a behavioral level, using behavioral means. In this endeavour, the principles of learning, and other methods and findings of experimental psychology, are utilized. Specifically, the principles of classical conditioning and those of

operant, or instrumental, conditioning are key cornerstones of this approach.

Other features of behavior therapy include the emphasis on the detailed analysis of the problem behavior (in terms of stimuli, responses, consequences, etc.), the key role given to measurement and quantification, and the use of an experimental approach, where treatment follows a hypothesis-testing paradigm. This also includes a strong emphasis on outcome evaluation.

Historical Development

Historically, behavior therapy developed over a period of time, culminating in its full emergence as a major psychological treatment approach in the 1950s. The basic concept of behavior change via behavioral means is of course very old, and there are examples of this from very early times, including the use of behavioral techniques in early Buddhism two-and-a-half millennia ago in India and in Rome around the beginnings of the Christian era. However, the key developments in the context of modern psychology can be traced to the work of a few key individuals, including John B. Watson, who demonstrated the applicability of learning principles in the understanding of the development of irrational fears. Watson and Rosalie Rayner, in 1920, created new fears in an 11-month-old child named Albert using classical conditioning principles. Four years later, Mary Cover Jones – one of Watson's students – published what might be described as the first formal application of learning principles to the treatment of a problem behavior. A 3-year-old boy, Peter, with a fear of rabbits was treated via the association of the feared stimulus with other positive and pleasant stimuli.

In the 1940s and 1950s much work along these lines was done in South Africa, mainly by Joseph Wolpe. Wolpe, who later emigrated to the United States, undertook serious experimental work – first with animals and then with human subjects – on the modification and elimination of maladaptive responses. His book, *Psychotherapy by Reciprocal Inhibition*, published in 1958 by Stanford University Press, is widely regarded as a landmark publication signaling the emergence of behavior therapy as a major therapeutic approach. Wolpe's work was supplemented by further work, both theoretical and empirical, by other leading figures such as Hans Eysenck, Monte Shapiro, Victor Meyer, and others in London and by Stanley Rachman, an associate of Wolpe who emigrated from South Africa to England in the early 1960s and joined Eysenck.

An equally significant contribution to the development of behavior therapy was made by B. F. Skinner,

whose book *Science and Human Behavior* (1953, Macmillan, New York) is also regarded as a classic in this field. The work of Skinner and associates concentrated on the use of operant conditioning principles, recognizing the key role of consequences in the maintenance and modification of behavior.

The establishment of *Behaviour Research and Therapy* in 1963 as a journal devoted to the field behavior therapy was another major step in the establishment of behavior therapy.

More recent developments have expanded and widened the scope and basis of behavior therapy. In the early stages, behavior therapy, drawing on the principles of classical conditioning, was used for what were then called 'neurotic disorders.' The largest group of patients treated in the early days in this way were those with phobias, or irrational and maladaptive fears. In fact, most of the cases described in Wolpe's 1958 book were individuals suffering from phobias. Those with obsessive-compulsive problems, sexual difficulties, and addictions were also treated. Operant conditioning principles were used to help in the management of behavioral problems in children and to modify problem behaviors in those with pervasive learning disability (mental retardation) and with chronic psychotic illness. These early developments were later supplemented (1) by the ideas of social learning theory, which pointed out that human behavior change takes place not just with direct experience but also through vicarious learning and social transmission and (2) by ideas and techniques of cognitive psychology and cognitive therapy, which showed the importance of cognitions (thoughts, attitudes, beliefs) in determining and/or influencing human behavior. The incorporation of cognitive therapy principles developed by Aaron T. Beck in Philadelphia and others eventually led to what is today referred to by many as cognitive behavior therapy.

Techniques of Behavior Therapy

Behavior therapy is widely applied today for a range of problems. It is considered to be a major, robust approach to helping people to get over problem behaviors and develop more adaptive ones. Some of the major behavior therapy techniques are described briefly in this section.

Exposure Therapy

This is the main approach used for phobias. The principle here is that exposure to the feared – and avoided – stimulus, in therapeutic conditions, leads to a reduction and eventual elimination of fear. This may be done either in real life (*in vivo*) or in imagination or fantasy. The client is exposed in a graded series

of steps to the fear stimulus or situation. As an anxiety-countering strategy, deep muscle relaxation may be used. Clients are usually treated individually, but group treatments are also undertaken. In recent years, clients have also been taught to engage in this type of exposure therapy by themselves, i.e., self-managed exposure therapy, after assessment by a qualified therapist and agreement on a plan of treatment.

In contrast to the graded exposure approach, intensive and/or prolonged exposure is also used. In the early days, this was generally known as flooding. *In vivo* flooding and imaginal flooding are both used. Essentially, the client is exposed to the fear-arousing situation for a prolonged period of time, which is done without a graded, step-by-step approach. Highly anxiety-arousing situations are used from the beginning. A treatment session usually ends when the client's initial high anxiety has subsided to a significant degree.

Modeling

In modeling, the client is exposed to other individuals, or models, engaging in the target behaviors in an appropriate, nonanxious manner. Modeling is used in fear reduction. The active participation of the client is a key aspect. Modeling is also used widely in the treatment of those who have difficulties in assertiveness and other social skills. Again, participant modeling, rather than simply observing the models, is needed to achieve the desired changes.

Response Prevention

This is a technique that is usually employed in conjunction with exposure. The best known application of this approach is in the treatment of compulsive behaviors in obsessive-compulsive disorder. The client is exposed to situations that arouse anxiety and create an upsurge of the urge to engage in the target compulsive behavior. For example, a client with contamination fears and associated compulsive hand-washing rituals may be exposed to items which he or she believes are contaminating, such as dirty table surfaces or door handles in a public convenience. This exposure leads to high anxiety and to a high urge to engage in a cleansing, hand-washing compulsion. However, as part of the prearranged treatment plan, the client refrains from engaging in this compulsive behavior. The compulsive urge and the anxiety both gradually dissipate with time, and the session is ended when a significant reduction is achieved. Repeated sessions lead to progressively lower anxiety and less strong compulsive urges.

Response prevention following exposure is also used with clients with alcohol problems, drug addiction, and compulsive binge eating, among others.

The Use of Reinforcement

Reinforcement is the strengthening of a behavior so that the probability of its occurrence is enhanced. Using reinforcers, or rewards, to strengthen desirable behavior is a major behavior therapy technique. There are two types of reinforcers. A positive reinforcer is something pleasant or welcome that is presented as a consequence of the desired behavior being performed. For example, a child may be given a sweet as a reward each time he puts on his socks by himself. A negative reinforcer is something unpleasant and/or unwelcome that is removed, or avoided, as a consequence of the target behavior being carried out.

It is common in behavior therapy to use social reinforcers as a major means of positive reinforcement of desired behaviors. These include verbal praise, approval, and attention.

Token Economics

A token economy is a whole system of reinforcement used for motivating clients. For example, a group of children in a class or a group of chronic schizophrenic patients in a large hospital ward may be helped in this way. Desired target behaviors are immediately reinforced with tokens: plastic chips, stars on a chart, or points. Maladaptive behavior can lead to the loss of tokens. The tokens are exchangeable for real reinforcers; these are called back up reinforcers in this context. For example, 10 tokens may be exchanged for a bar of chocolate. A token economy program has a set of detailed and explicit rules and procedures that specify how tokens are to be earned and how they are to be exchanged for back up reinforcers.

Aversive Procedures

In the early days of behavior therapy, aversive procedures were commonly used. In aversion therapy, undesirable target behaviors are systematically followed by an aversive consequence, e.g., electric shocks. Clients seeking treatment for deviant sexual interests were, for example, treated with a paradigm where the deviant stimulus (e.g., pictures of fetish object) was accompanied repeatedly by the delivery of a shock. Such aversive procedures are not commonly used today, partly for ethical reasons, but also because these techniques were found to achieve a suppression of the unwanted response rather than its elimination. However, covert aversion (where the unpleasant consequences of an undesirable behavior are presented through powerful imagery) is used with some success. Real physical aversion is still used sometimes to control self-injurious acts in clients with severe problem behaviors.

Extinction

The most effective technique for reducing and/or eliminating an undesirable behavior is to eliminate the reinforcement that follows that behavior. If this can be done systematically, then the behavior gets effectively extinguished. For example, a child's frequent tantrums may be reinforced by the parental attention that such behavior elicits, and the withholding of this reinforcement can lead to the extinction of this behavior. However, sometimes the reinforcer maintaining the target behavior cannot be identified. A further difficulty is that implementation of a treatment program such that the problem behavior is not even occasionally reinforced is often difficult in practice.

Time Out

Time out (short for time out from positive reinforcement) is the temporary withdrawal of a client's (usually a child's) access to positive reinforcers immediately following the performance of the target maladaptive behavior. In child behavior problems (e.g., hitting others, disruptive shouting), the child is taken to a bland time-out area for a limited period of time, following the undesirable behavior.

Practice of Behavior Therapy

In clinical practice, behavioral techniques are used, sometimes in combination, after a careful analysis of the problem. Such detailed behavioral analysis is a necessary first step before therapy is undertaken. While simple, uncomplicated clinical problems may need only simple and straightforward therapy programs, many clinical cases in real life present complex problems that may require the imaginative use of several behavioral techniques, concurrently or sequentially. At each step, a careful analysis is made of the outcome of the intervention and, if necessary, the program is revised. This experimental, or hypothesis-testing, approach is a major feature of the behavioral approach in its application.

In many instances, the behavioral techniques may need to be combined with cognitive techniques. It was noted in an earlier section that the term cognitive behavior therapy is used by many practitioners for this approach. This approach remains essentially within the broad framework of behavior therapy,

but there is emphasis on acknowledging the significance of cognitions. Cognitive techniques are added to the behavioral ones as needed. In the treatment of panic attacks, for example, the behavioral technique of graded exposure to identified panic-provoking situations is combined profitably with helping the client to identify any faulty cognitions (e.g., "My heart is beating fast, this means I may collapse") and providing evidence that these are unfounded.

Conclusions

Behavior therapy has come a long way since its tentative beginnings in the early decades of this century. It is a well-developed, versatile approach to the treatment of a wide range of behavioral problems, and there is much empirical evidence for the efficacy of its various therapeutic techniques. Some of these techniques are also widely used in helping modify the behavior of nonclinical subjects, e.g., in improving classroom behavior. As an approach to the modification of human behavior, behavior therapy remains well grounded in experimental psychology, especially learning theory. Its emphasis on the careful observation and assessment of problems, defining target behaviors, quantification and measurement, and rigorous outcome evaluation makes it a robust scientific approach.

See Also the Following Articles

Behavior, Overview; Cognitive Behavioral Therapy.

Further Reading

- Ellis, A. (1997). Extending the goals of behavior therapy and cognitive behavior therapy. *Behavioral Therapy* 28, 333–339.
- Eysenck, H. J. and Martin, J. (eds.) (1987). *Theoretical foundations of behavior therapy*. New York: Plenum.
- Last, C. G. and Hersen, M. (eds.) (1994). *Adult behavior therapy casebook*. New York: Plenum.
- Plaud, J. J. and Eifert, G. H. (eds.) (1998). *From Behavior Theory to Behavior Therapy*. Boston: Allyn & Bacon.
- Salkovskis, P. M. (1996). *Trends in Cognitive and Behavioural Therapies*. Chichester: Wiley.
- Turner, S. M., Calhoun, K. S. and Adams, H. E. (eds.) (1992). *Handbook of Clinical Behavior Therapy* (2nd ed.). New York: Wiley.
- Wolpe, J. (1990). *The Practice of Behavior Therapy* (4th ed.). Elmsford, NY: Pergamon.

Behavior, Overview

R Dantzer

Integrative Neurobiology, Bordeaux, France

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R Dantzer, volume 1, pp 296–299, © 2000, Elsevier Inc.

Behavioral Responses to Stress
Interactions between Behavioral and Physiological
Components of the Stress Response
Behavioral Sources of Stress

Glossary

<i>Appetitive stimulus</i>	A stimulus that induces approach and exploratory responses.
<i>Aversive stimulus</i>	A stimulus that induces avoidance and escape responses.
<i>Behavioral sequence</i>	A succession of appetitive and consummatory components. The appetitive sequence enables the subject to approach the goal object. Upon reaching the goal, appetitive behavior normally ceases and gives way to consummatory behavior that enables the subject to appropriate its goal.

Behavioral Responses to Stress

The stress response was initially described in the context of catastrophic events such as hemorrhage, surgical injury, burns, septic shock, and intoxications. These events cause a reflex-like response that is non-specific and depends on the importance of the deviation from homeostasis. However, more common stressors such as daily life events do not act this way. The response depends on the evaluation by the subject of the potential threat these factors represent. The transactional model of stress is therefore more appropriate than the stimulus–response of stress to describe what is occurring. Novelty, uncertainty, and the absence or loss of control are important dimensions in the evaluation process. An emotional response occurs if the event is recognized as discrepant, and it involves both physiological and behavioral adjustments.

Novelty occurs when the situation contains no or very few familiar elements. Subjects exposed to novelty explore the situation in order to identify potential sources of danger that need to be avoided and potentially interesting objects that can be safely approached.

When novelty is too intense, fear predominates over exploration, and escape attempts replace investigatory behavior. Uncertainty refers to the inability of predicting what is going to happen when the subject is exposed to an aversive or a potentially rewarding situation. In many cases, harmful or rewarding events do not happen all of a sudden. They are preceded by warning signals that can be used to predict their occurrence, via a process of classical (or Pavlovian) conditioning. A tone that occurs before a painful electric shock becomes a fear signal. Inversely, a bell that rings only when there is no electric shock is a safety signal. In an uncertain situation, the subject tries to find regularities in the succession of events to which he or she is exposed, so as to make the situation more predictable. The behavioral response to a fear signal is an anticipatory response that helps the individual to protect him- or herself from the danger. The response to a safety signal is a decrease in the tension and/or anxiety induced by the possible occurrence of the harmful situation. Behavioral responses to a fear signal depend on the subject's proximity to the danger. When the subject is directly in contact with an uncontrollable threat, he or she usually tries to move away from the danger (active avoidance). When the subject is at a close distance from the danger and cannot move away, he or she usually refrains from getting in contact with it (passive avoidance). This response is associated with immobility and freezing.

When the situation is rewarding rather than aversive, events that regularly precede the occurrence of the reward become positive conditioned stimuli or secondary reinforcers. They induce the development of conditioned behavioral responses that involve approach and other appetitive components of the normal behavioral response to the reward. If a secondary reinforcer is no longer followed by the occurrence of the expected reward, it induces a state of frustration that is characterized by agitation, aggression, and other displacement activities (see below).

In situations that involve physical contact with congeners, behavioral responses to stress form part of the agonistic repertoire. Agonistic behavior refers to the complex of aggression, threat, appeasement, and avoidance behaviors that occurs during encounters between members of the same species. Attack is usually a highly ritualized behavior when it occurs in property-protective fights. The offender attacks its opponent at very specific parts of the body, such as the neck in rats, whereas the defender tries to protect these targets from bites. When defeated, the defender

adopts a submissive posture, exposing to the attacker more vulnerable body sites. Submission usually puts an end to the fight.

Confronted with a potentially harmful situation, the subject usually tries to do something in order to avoid the danger. Coping refers to the efforts undertaken to resolve the situation. Active and passive avoidance responses are examples of coping responses. Behavioral manifestations of coping vary according to the situation and the way the threat can be controlled. In laboratory animals, for instance, coping is typically studied by enabling the subjects to escape or avoid painful electric shocks when they emit a specific response such as a lever press, a wheel turn, or a move from one compartment to the other in a two-compartment cage. While animals in the escape/avoidance group are permitted to terminate electric shocks or delay their occurrence, animals in another group serve as yoked controls since they receive the same electric shocks as the previous animals, but their behavioral response has no effect on the occurrence and duration of electric shocks. As a further control, other animals are placed in the same apparatus as the animals in the other two groups, but they are never exposed to electric shocks. Controllability can also be studied in human subjects by asking them to press a button, usually to terminate an intense noise. As in the animal experiment, the yoked group has no control over the noise.

Helplessness refers to an individual's inability to control the situation when he or she has the opportunity to do so. When this happens after a previous experience of uncontrollable stress, it is referred to as learned helplessness. Learned helplessness has cognitive, motivational, and emotional components.

Coping attempts can be classified according to the strategy that is employed to confront the situation. Active coping refers to all ways of dealing behaviorally directly with the situation, e.g., avoidance or attack. In human subjects, active coping can be cognitive (logical analysis and positive reappraisal) or behavioral. The latter case includes seeking guidance and support and taking concrete actions to deal directly with the situation or its aftermath. In an aversive context, cognitive coping comprises intrapsychic processes aimed at denying or minimizing the seriousness of the situation or its consequences, as well as accepting the situation as it is. Behavioral coping includes seeking alternative strategies to directly confronting the problem.

In general, active coping enables the subject to deal directly with the problem and is therefore characteristic of problem-oriented coping. In cases in which the problem cannot be directly dealt with, it is often preferable to engage in passive coping that focuses mainly on managing the emotions associated with it

(emotion-oriented coping). In the case of passive coping, the behavioral activities that are normally elicited by the stimuli present in the situation are redirected toward more easily accessible targets. A typical example is the redirected aggression that develops toward a subordinate when an animal in a social group is attacked by the dominant member of the group. Whereas redirected activities belong to the same behavioral repertoire as the thwarted behavioral activity, displacement activities belong to a different behavioral repertoire. This is the case for aggressive behavior that is targeted toward peers or other objects in the environment in an animal that does not get the food it was expecting. A well-studied form of displacement activity is schedule-induced behavior. Schedule-induced behaviors, also known as adjunctive activities, typically occur when hungry animals are exposed to intermittent food rewards, delivered in very small amounts independently of the animals' behavior. Immediately after animals have eaten their allocated food, they vigorously engage in other activities depending on what is available in the situation. They drink exaggerated amounts of water if a drinking tube is accessible, attack a congener when present, or gnaw a piece of wood if available. To qualify as an adjunctive behavior, the behavioral activity must occur as an adjunct to the reinforcement schedule, not be directly involved in, or maintained by, the reinforcement contingency, and be persistent and excessive. Independently of the object toward which they are directed, adjunctive behaviors occur at a maximum rate immediately after the reinforcement and decrease in probability when food deprivation is alleviated.

Interactions between Behavioral and Physiological Components of the Stress Response

Compared to uncertainty and lack of control, predictability and controllability have deactivating effects on the physiological arousal evoked by the stress situation. In social conflict situations, activation of stress hormones is in general more pronounced in subordinate animals that behave defensively or passively than in offensive animals that remain active and become dominant. However, the social role is more important than the social status since dominance is associated with a reduced arousal response only when dominant animals are recognized as such by their peers and do not need to reaffirm their status each time they are challenged by subdominant animals.

Schedule-induced polydipsia is associated with lowered plasma levels of stress hormones. This property is shared by other motor activities that involve

sensory overloading from the mouth and tongue (e.g., licking or chewing) or the limbs (e.g., pacing). The de-arousal properties of coping behavior are not always that clear. Coping attempts are associated with increased activation of the sympathetic nervous system, and they result in decreased physiological activation only when they have been successful and the threat is over.

The relationship between behavioral and physiological components of the stress responses is not unidirectional, with behavior influencing physiological responses, but truly bidirectional in the sense that the likelihood of a given behavioral response is dependent on the subject's initial physiology. Individuals with high activity of the sympathetic-adrenal-medullary axis are more likely to try to cope in an active way with a threat than individuals with a low activity of the sympathetic-adrenal-medullary axis. Inversely, individuals with a relatively higher activity of the hypothalamic-pituitary-adrenal axis are more likely to resign and display helplessness than individuals with a normal or a lower activity of this axis.

Behavioral Sources of Stress

Most psychobiological theories of stress emphasize the previously described bidirectional relationships between the behavioral and physiological components of stress in response to stressors. However, they do not necessarily account for the powerful influence of the subject's own behavior on the probability of occurrence of the stressful situation itself.

A typical example for understanding this point is represented by the type A behavior pattern (TABP). TABP is defined as an action-emotion complex that individuals use to confront challenges. The complex involves behavioral dispositions such as aggressiveness, competitiveness, and impatience; specific behaviors such as muscle tenseness, alertness, rapid and emphatic vocal stylistics, and accelerated pace of activities; and emotional responses such as irritation, covert hostility, and above-average potential for anger. Type B individuals are characterized by the relative absence of type A behaviors and confront every challenge with placid nonchalance. Type A individuals have a higher risk of developing coronary heart disease than type B individuals. The deleterious effects of type A behavior appear to be associated with a higher physiological reactivity to external demands, especially when there

are elements of competition. These findings corroborate the previously mentioned influences of behavioral factors on the physiological response to stress. In accordance with the bidirectional nature of hormone-behavior relationships, type A behavior does not occur at will. It is determined at least in part by biological factors and autonomic hyperactivity. All this makes type A behavior a typical psychobiological construct. However, this concept is not sufficient to account for what happens in the real world. Type A individuals actually create their own problems by putting themselves in situations of competition and taking all steps to transform the events they are confronted with in challenges that need to be met in a more hostile way than their peers would be inclined to do.

A similar situation occurs in sensation-seeking individuals. There is evidence that subjects with this personality trait display enhanced responses to novelty and danger. The probability for a given individual to be a sensation seeker is determined by innate and experiential factors. However, what is characteristic of sensation seekers is their strong preference for sensation-inducing situations rather than their increased response to such situations. In sum, individuals are not passively exposed to stress; they actually create their own stress by the way they behave and perceive their environment.

See Also the Following Articles

Aggressive Behavior; Behavior Therapy; Defensive Behaviors; Evolutionary Origins and Functions of the Stress Response; Health Behavior and Stress.

Further Reading

- Carroll, M. E. and Overmier, J. B. (eds.) (2001). *Animal research and human health: advancing human welfare through behavioral science*. Washington, D.C.: American Psychological Association.
- Dantzer, R. (1994). *The psychosomatic delusion*. New York: The Free Press.
- McFarland, D. J. (ed.) (1987). *The Oxford companion to animal behaviour*. Oxford, UK: Oxford University Press.
- Ursin, H. and Eriksen, H. R. (2004). The cognitive activation theory of stress. *Psychoneuroendocrinology* 29, 567-592.

Benzodiazepines

J Bermak

University of California, San Francisco, CA, USA

T Johnstone and K Gee

University of California, Irvine, CA, USA

© 2007 Elsevier Inc. All rights reserved.

History and Introduction

Basic Pharmacology

Clinical Pharmacology

Clinical Uses

Combination of Benzodiazepines with Psychotherapy

Common Adverse Effects

Effects on Cognition

Behavioral Dyscontrol Reactions

Physical Dependence and Withdrawal

Abuse Liability

Tolerance and GABA_A Subunit Specificity

Future Directions and Conclusion

Glossary

<i>Akathisia</i>	Motor restlessness, characterized by the inability to sit still, often occurring as a result of ingesting medications that block dopamine receptors.
<i>Allosteric modulator</i>	A molecule that binds to a protein at a regulatory site, spatially distinct from the primary site of activation. Allosteric modulators alter protein activity by modifying the affinity of the principal molecule to the primary site of activation.
<i>Anterograde amnesia</i>	The inability to form new memories for a limited period of time following a specified event, even though overall consciousness is preserved.
<i>Cognitive-behavioral therapy (CBT)</i>	A type of psychotherapy in which the therapist teaches the patient to restructure his or her maladaptive thoughts (i.e., cognitions) in order to improve overall functioning (i.e., behavior).
<i>Glucuronide conjugation</i>	A form of metabolism, usually in the liver, involving the attachment of glucuronic acid to another molecule, thereby producing a derivative that is more easily eliminated from the body through urine excretion.
<i>Hyperpolarization</i>	The reduction of the electric potential (or voltage) of a neuron to a level below that of the normal resting state. This opposes neuronal excitation by inhibiting or slowing action potential propagation (the flow of positive charge along neurons).

Hypnotic

Ligand-gated ion channel

Acting to induce sleep or drowsiness.

A transmembrane ion channel in which the flow of ions through the channel is increased by the binding of an activating molecule (termed a ligand), such as GABA to the GABA_A receptor.

Limbic system

A circuit of interconnected brain structures that controls and influences emotional behavior and perceptions.

Retrograde amnesia

The loss of memories that have already been formed prior to the inciting event.

Somatic

Pertaining to the body; physical or corporeal. In psychiatry, somatic usually connotes a physical or physiological manifestation of mental anguish.

History and Introduction

Benzodiazepines (BZs) surfaced for broad clinical use in the 1960s as potent anxiolytics (anti-anxiety medications). Prior to this period, barbiturates and meprobamate had been favorites in treating anxiety, but they had variable clinical efficacy and substantial safety risks. The introduction of the BZs, whose safety-efficacy profiles exhibited marked superiority to their predecessors, led to substantial changes in medication-prescribing practices. By 1977, diazepam (brand name, Valium) became the single most commonly prescribed medication in the United States. Then, reports of BZ abuse and addiction spread, invoking widespread concern over their use. The 1980s and 1990s witnessed a decline in BZ popularity coupled with exaggerated beliefs about their side effects and habit-forming potential. In more recent years, however, the accumulation of extensive clinical research has begun to reverse prior misconceptions. This synopsis reviews the basic and clinical pharmacology of BZs, as well as current understandings of their efficacy, safety, and liabilities as anxiolytics.

Basic Pharmacology

The core structure of clinically relevant BZs consists of a benzene ring joined to a seven-member 1,4-diazepine ring (see [Figure 1](#)). BZs bind to GABA_A receptors on the extracellular surface of neurons in the brain. The pentameric (or five-subunit) GABA_A receptor complex is a ligand-activated chloride channel. The binding of the endogenous ligand GABA (γ-aminobutyric acid) to the GABA_A receptor leads to channel opening, allowing chloride ions to flow passively into neurons. The chloride ion flow through

GABA_A channels induces neuronal hyperpolarization. Hyperpolarization then slows nerve transmission by reducing the probability of action potential propagation along neuronal membranes. This physiological action is essential in prohibiting uncontrollable states of brain activation.

BZs exert their therapeutic effects through a mechanism of allosteric modulation. That is, they bind to the GABA_A receptor complex at a location that is distinct from the GABA binding site and facilitate increased chloride ion flow through the channel by enhancing GABA binding (see [Figure 2](#)). Alone, BZs are unable to open GABA_A ligand-gated channels to

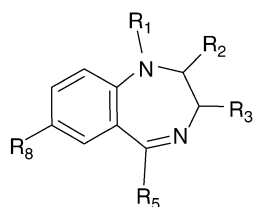


Figure 1 Core structure of 1,4-benzodiazepines. Numbered R groups designate positions of variation that distinguish the individual BZs (see also [Figure 3](#)). Note that 1,4 refers to the position of two nitrogen (or di-aza) atoms.

effect intracellular chloride entry. Such restricted modulatory capacity of BZs contrasts with the direct channel-opening properties of barbiturates and defines the comparative safety of BZs. When taken in overdose, barbiturates are capable of producing massive neuronal inhibition that can lead to respiratory depression, coma, and eventually death. In contrast, maximal BZ effects are limited by the physiological availability of GABA, reducing their potential of generating excessive neuronal inhibition. Although oversedation and some respiratory depression can occur, lethal consequences from BZ overdoses are rare. Lethal inhibitory effects usually require high combination doses of BZs with additional sedatives (e.g., alcohol) to occur in healthy individuals.

Clinical Pharmacology

BZs cause a spectrum of central nervous system (CNS) inhibitory effects: sedation, dampening of memory and cognitive functioning, interruption of seizures, muscle relaxation, and reduction of anxiety. Among the BZs, distinct modes of action are imparted by their individual rates of absorption, delivery to the brain, metabolism, and elimination. For example, diazepam

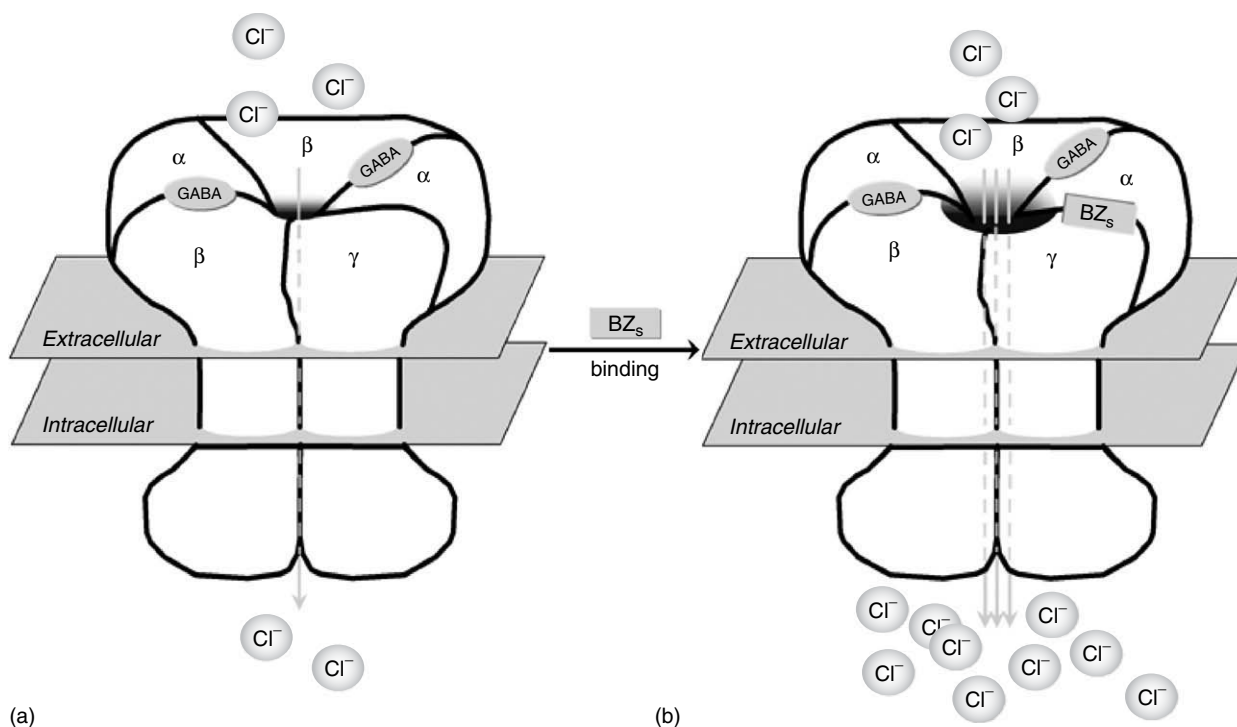


Figure 2 Increased activation of the GABA_A receptor complex by BZs. a, Before BZ binding; b, after BZ binding. The pentameric (or five-subunit) GABA_A ligand-gated channel, consisting of 2 β , 2 α , and 1 γ subunits, is shown conducting chloride intracellularly following GABA binding to the α - β interface. BZ binding at the α - γ interface results in allosteric enhancement of GABA-mediated channel opening, illustrated by increased chloride conductance. In the absence of GABA, BZs are unable to activate the channel, conferring safety in an overdose. In contrast, barbiturates directly activate the receptor complex and are consequently much more dangerous in an overdose.

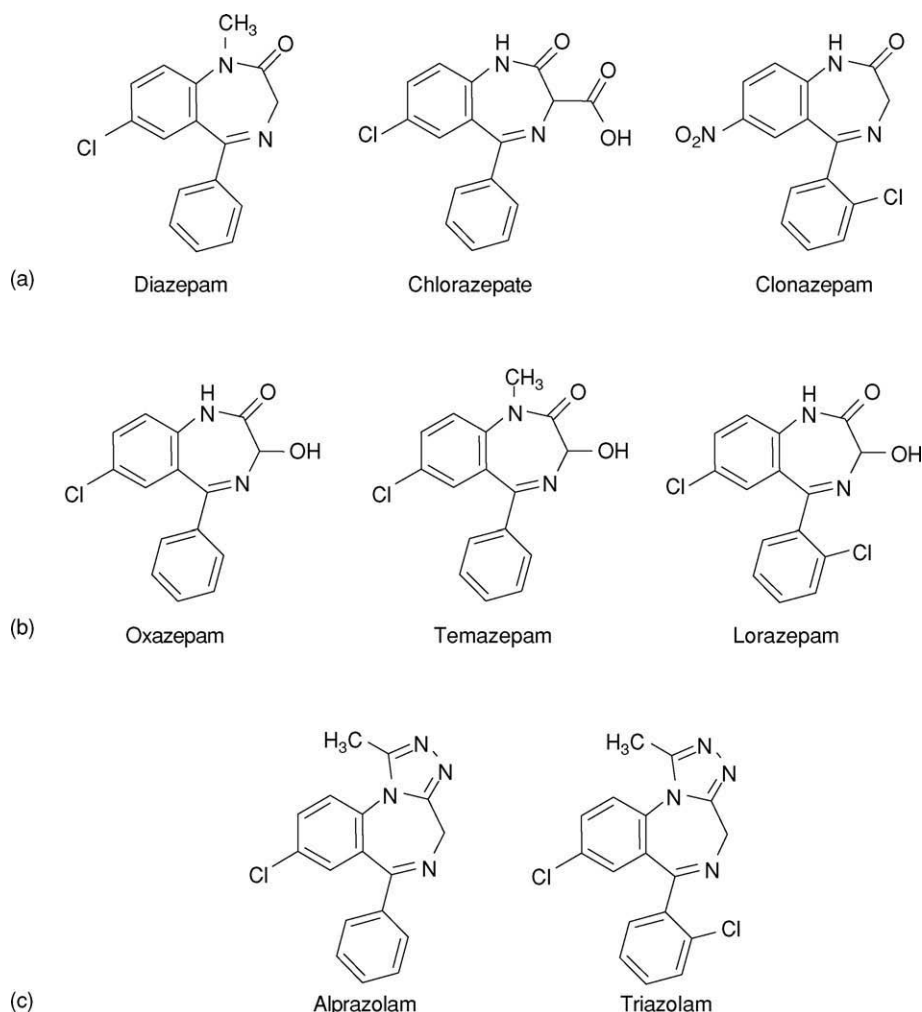


Figure 3 Examples of 1,4-benzodiazepine structures and subclasses. a, 2-Keto; b, 3-hydroxy; c, triazolo.

is very lipid-soluble (lipophilic), rapidly absorbed, and easily distributed into the brain with a quick onset of action. BZs with reduced lipophilicity, lorazepam and clonazepam, are more slowly delivered to their target sites in the brain thereby inducing more gradual onset of CNS effects (particularly with clonazepam).

In addition to such individual differences, categorical features of BZs can also be appreciated by grouping them into three major subclasses (see [Figure 3](#)), particularly regarding their metabolic properties. The metabolism of 2-keto BZs involves slow oxidation in the liver. Also, metabolic products in this subclass are usually active, frequently more active than the parent compounds. As a result, 2-keto BZs exhibit appreciably long-lasting clinical effects. In fact, 50% of their activity (or one half-life) remains between 30 and 60 h after initial dosing. BZs within the 3-hydroxy subclass also undergo liver metabolism, but the process is more rapid (via glucuronide conjugation rather than oxidation), with the inactivation of the parent

compounds. Consequently, 3-hydroxy BZs exhibit shorter half-lives (~ 8 –14 h) than their 2-keto counterparts. Similar to the 2-keto agents, triazolo BZs are also oxidatively metabolized. Yet triazolo metabolites are much less active than the parent compounds. Hence, the overall half-lives of agents within the triazolo subclass range from ~ 3 to 12 h.

Choosing a BZ for any particular clinical indication involves a consideration of various pharmacokinetic factors: desired speed of onset, importance of sustained anxiolytic action, and the influence of a patient's age or other medications on BZ metabolism and physiological accumulation. Consequently, optimal selection and proper dosing of BZs involves patient-specific tailoring.

Clinical Uses

Psychiatric Indications

Relative to other classes of psychiatric medications, BZs have established a remarkably versatile clinical

utility. As a class, BZs are used to treat various stress disorders, either alone or in combination with antidepressants. Most commonly, these include panic disorder and social and generalized anxiety disorders. Somewhat shorter-acting BZs, such as temazepam and triazolam, are frequently prescribed as hypnotics, with insomnia representing a major clinical indication for these compounds.

In combination with other agents, BZs can be quite useful in calming agitation states seen in psychosis and mania and in expediting the speed of antidepressant response in major depression. Symptoms of anxiety and sleep disturbances are observed in many psychiatric conditions, particularly in acute phases of illness, accounting for the benefits that BZs provide in such diverse settings.

Medical Applications

In addition to their uses in psychiatric conditions, BZs are relied on in medical practice to diminish a number of physical symptoms. These include muscle spasms, gastrointestinal cramping, and akathisia (a form of restlessness). Acute and chronic seizures have been treated by diazepam and clonazepam, respectively, with long-standing proven efficacy. These two agents, in addition to lorazepam and chlordiazepoxide, are also commonly used to prevent the symptomatic onset of alcohol withdrawal. With such a range of efficacy in mental and physical conditions, BZs are often touted as the most effective medications available for anxiety states that present with somatic (i.e., physical) symptoms.

Lack of Benefit in Posttraumatic Stress Disorder

Essentially, where there is intense fear or distress, BZs are powerfully and effectively calming. Perhaps the only glaring exception to this rule lies in the treatment of posttraumatic stress disorder (PTSD), a disorder of excessive and prolonged anxiety, often lifelong, following exposure to a severe traumatizing event. In limited studies, BZs have been unable to improve sleep or nightmares in PTSD and have even appeared to worsen the risk of developing chronic PTSD following trauma.

Given the combined presence of anxiety and insomnia in PTSD, it is surprising that BZ efficacy has not extended to this condition. One possible explanation may relate to the neurobiology of PTSD as a powerfully intensified and easily arousable emotional circuit. In PTSD patients, this pathological circuit appears to be so hard-wired in the brain as to be unusually refractory to the calming effects of BZs. In addition, the suppression of this circuit may be uniquely dependent on maximal cognitive functioning,

which can be impaired by acute BZ treatment. In this model, higher cognitive and reasoning abilities are required to stave off unwanted emotional arousal of traumatic memories. Consistent with this conceptualization, intelligence and higher cognitive functioning have been found to reduce the risk of developing PTSD. If such a theory proves valid, short-term trials with BZs are not expected to be beneficial in PTSD, as is generally supported by the current clinical literature. Larger studies designed to comprehensively examine BZ effects in this disorder, however, have not been conducted.

Acute versus Chronic Utility

BZs are usually most effective in relieving acute distress rather than chronic anxiety or despair. Their efficacy in major depressive disorder has been optimally demonstrated during acutely severe, often suicidal, phases of the illness. They are otherwise inferior, or provide no additional benefit, to antidepressants in treating chronic depression. BZs are also not commonly effective in obsessive-compulsive disorder (a stubbornly chronic condition).

However, data are conflicting regarding long-term BZ use. Agents such as alprazolam and clonazepam demonstrate reductions in the frequency of recurrent panic attacks (with or without antidepressant combinations). Also, clonazepam has shown limited, but robust, effects in chronic social anxiety. Hence, the clinical evidence of BZ efficacy in chronic psychiatric disorders is mixed.

Combination of Benzodiazepines with Psychotherapy

BZs can be combined with some forms of psychotherapy, or vice versa, to enhance overall treatment effects when conducted in the proper context. Most studies to date have examined the use of BZs with CBT. CBT produces robust and persisting benefits in a number of anxiety disorders by providing patients with the psychological skills to endure exposure to stressful conditions. However, successful CBT implementation requires several weeks to months of treatment. Some clinical trials have indicated that the addition of BZs to CBT can accelerate the therapeutic process by facilitating fear extinction. While on BZs, CBT patients are capable of exposing themselves more quickly to previously unbearable stress. Yet the subsequent discontinuation of BZs in the course of psychotherapy can reignite prior anxiety, providing no added benefit overall. From these findings, fear extinction appears to be somewhat context-dependent. In other words, severe anxiety may recur easily when

the context of its remedy is altered (e.g., on removal of BZs). Accordingly, in cases in which BZs do hasten the progress of CBT or another psychotherapeutic modality, their discontinuation should be conducted gradually with continued psychotherapy to optimize patient adaptation to newly unmasked stressors.

Conversely, the addition of CBT to BZs reveals consistently successful results in promoting medication discontinuation and can be a particularly useful strategy in cases in which tapering patients off BZs is desired and challenging.

Common Adverse Effects

Overall

The side effects from the acute administration of BZs primarily include sedation, cognitive impairments, ataxia (imbalance), weakness, and motor incoordination. However, many of the acute BZ side effects do not appear to be problematic during long-term use. The most commonly observed side effect, sedation, is dose-dependent and can be well managed with dose reduction until tolerance develops within a few weeks.

Considerations in Elderly

Risks of daytime sedation and ataxia are more worrisome in elderly patients, due to the potential for falls and dire consequences, such as hip fractures. Consequently, a careful consideration of BZ dosing, half-life, and metabolism is given to elderly patients on these medications. The oxidative metabolism of BZs is more likely to be sluggish in the elderly and may produce unpredictably sustained action and the potential for toxic accumulation. BZs with longer half-lives or whose metabolism is impaired by concomitant use of other medications may also produce dangerous effects. The more easily metabolized 3-hydroxy BZs (such as temazepam, oxazepam, and lorazepam), or agents with shorter half-lives, are considered relatively safer options in elderly patients. However, factors of half-life and metabolism are not sufficiently predictive of adverse events. Hence, close monitoring of patient functioning is required when using BZs to address anxiety and/or insomnia in the elderly.

Effects on Cognition

Acute Effects

As with their sedating properties, BZs administered acutely can slow overall cognitive functioning. Impairments are observed in visuomotor (or hand-eye) coordination, attention, working memory (e.g., in following directions), and new memory formation (producing anterograde amnesia).

Chronic Effects

The impact of chronic BZ use on cognitive functioning is much more controversial. Some reviews have illustrated evidence of broad-based cognitive impairments. These retrospective analyses, however, have been unable to control adequately for the burden of underlying psychiatric disease on mental acuity. Recent prospective studies comparing anxious patients with and without chronic BZ treatment do not reveal distinct BZ-induced cognitive deficits, except possibly in the realm of verbal episodic memory.

Behavioral Dyscontrol Reactions

Unexpectedly intense behavioral responses to BZs occur very infrequently (with < 1% incidence). Sensationalized reporting of these effects, however, has produced disproportionate awareness and concern among physicians. Often termed paradoxical, disinhibitory, or dyscontrol reactions, they encompass such symptoms as talkativeness, agitation, irritability, confusion, unseemliness, aggression, and loss of impulse control with potentially violent consequences. When they occur, the pathophysiology is not understood, but reversal by flumazenil (a BZ antagonist) can be effective and supports a BZ causal link to these phenomena. Children, elderly, and alcohol-dependent patients appear to be at greater risk.

Dyscontrol reactions may stem from an acute confusional or disoriented state, such as delirium. If so, individuals with severe illness or impaired cognition (who are at greater risk for the development of delirium) may also be expected to be more vulnerable to the unlikely development of these bizarre, paradoxical BZ reactions.

Physical Dependence and Withdrawal

Debate about chronic clinical use of BZs stems, in part, from their propensity to induce physical dependence (herein defined as the appearance of unpleasant physiological symptoms on sudden BZ withdrawal). Sudden withdrawal from BZs can produce jitteriness, sweating, clamminess, nausea, seizures (less commonly), and the reappearance of pre-BZ anxiety. These symptoms appear within 2–7 days after sudden BZ discontinuation, depending on the half-life of the agent involved. Similarly, rebound refers to the appearance of anxiety symptoms in excess of the pre-BZ state, sometimes seen with the sudden cessation of short-acting BZs (e.g., alprazolam) that eliminate quickly from the bloodstream. Long-acting BZs such as the 2-keto agents (e.g., diazepam and clonazepam) rarely produce rebound and generally tend to produce

fewer withdrawal symptoms than their shorter-acting counterparts (although seizures can still occur).

The daily use of BZs in human studies reveals that physical dependence begins to form after 6–8 weeks, raising ethical objections to their initiation in the treatment of anxiety disorders. In comparison, many other psychotropic agents (including selective serotonin reuptake inhibitors, SSRIs) induce withdrawal symptoms on cessation, leading to physiological dependence. Some experts suggest, then, that these effects from BZs should not be overemphasized when considering poorly treated anxiety as the alternative. Nevertheless, most psychiatrists agree that BZs often necessitate uniquely prolonged and involved dose tapering prior to their discontinuation. This can be more effectively accomplished by converting patients from short-acting BZs (e.g., alprazolam) to long-acting equivalents (e.g., diazepam) in the process of dose tapering (and/or the initiation of CBT, as previously described).

Abuse Liability

Animal versus Human Findings

Much debate has surrounded the addictive potential of BZs. Findings in primate and other animal studies indicate that these medications reinforce self-administration, a preclinical model for abuse liability and a basis for their controlled prescription regulation. Within the class, BZs with rapid onset (e.g., diazepam) and/or short action (e.g., alprazolam and triazolam) are considered more liable to reinforce drug-seeking behavior. However, following initial isolated reports of addiction in humans in the early 1980s, subsequent epidemiological studies have not found this to be a significant problem overall (in contrast to opiate or barbiturate abuse). Approximately 15% of BZ users continue their BZs beyond 1 year, with less than 2% of long-term users exhibiting escalating dosages and/or multiple prescriptions. These low percentages support the notion that early physical dependence is manageable and does not frequently lead to long-term dependence or tolerance to the clinical benefits of BZs. The findings also suggest that the vast majority of patients do not seek BZs for euphoric effects or illicit use.

Vulnerable Patient Populations

The patients who do tend to abuse BZs are usually those with alcohol- or other drug-abuse disorders. In these settings, BZs are abused secondarily to augment the mind-altering effects of other addictive substances. This can be particularly problematic in the clinical setting of alcohol withdrawal when BZ is used as a substitute in alcohol detoxification

(withdrawal). Loading and short-term maintenance with BZs remain the mainstay of ensuring safety during alcohol withdrawal (by preventing withdrawal seizures and fluctuating blood pressures). Yet patients who relapse on alcohol during BZ maintenance may have lethal consequences from the combination. Notwithstanding the latter scenario, BZs can be administered generally without addictive potential in humans, particularly once substance abuse is reassuringly ruled out.

Tolerance and GABA_A Subunit Specificity

Differential Effects

Physiological tolerance refers to the lessening of a drug's effect over time. As previously described, tolerance to the anxiolytic properties of BZs is not commonly observed in long-term use (illustrated by the lack of dose-escalation requirements in routinely treated anxiety disorders). In contrast, the early clinical effects of sedation, ataxia, and cognitive impairment (such as inattention) tend to resolve over time.

Molecular Models

Intensive molecular research efforts have attempted to explain such differential tolerance to BZs. These efforts have uncovered a remarkable diversity of α -, β -, and γ -subunit isoforms making up GABA_A channels throughout the brain (see [Figure 2](#) for GABA_A structure). Of note, $\alpha 1$ subunits are distributed widely and implicated heavily in transmuting the amnestic and sedative effects of BZs, whereas the anxiolytic effects seem to be attributed to the $\alpha 2$ and $\alpha 3$ subunits, whose distribution is concentrated more in the limbic (emotional) brain centers. Observations that $\alpha 1$ (but not $\alpha 2$ and $\alpha 3$) subunits downregulate in response to chronic BZ treatment correlate with clinical findings of tolerance to their amnestic and sedative (but not anxiolytic) effects over time. Functional differentiation of these subunits has also led to the development and clinical availability of $\alpha 1$ -selective nonbenzodiazepine analogs (including zolpidem and zaleplon) for the treatment of insomnia. These $\alpha 1$ -selective hypnotics appear to impart reduced tolerance to their sedative benefits compared with traditional BZs for reasons that are not yet entirely clear.

Future Directions and Conclusion

Our burgeoning molecular understanding of BZ-related tolerance and subunit-specific clinical effects has engendered future hopes of developing powerful anxiolytics that lack sedative, ataxic, and cognitive liabilities. In particular, many pharmaceutical

companies have attempted to produce BZ analogs whose pharmacological action is selective to $\alpha 2$ and $\alpha 3$, but not $\alpha 1$, subunits. A few candidates remain in this space with promising early findings. Yet the engineering of such agents with subunit-selective and pure anxiolytic efficacy has been considerably challenging. Theories about subunit-selective efficacy are also evolving with these efforts and may ultimately prove to be even more complex and interesting than our current understanding.

Until then, BZs remain a time-tested gold standard in the pharmacological treatment of anxiety disorders, particularly for states of acute distress but also for some forms of chronic illness. In summary, the fundamental pharmacology of BZs, as allosteric modulators of the GABA_A receptor complex, delineates a unique balance of safety and efficacy in the therapeutic pursuit of anxiety-free homeostasis.

See Also the Following Articles

Anxiety; Anxiolytics; Cognitive Behavioral Therapy; GABA (Gamma Aminobutyric Acid); Obsessive-Compulsive Disorder; Panic Disorder and Agoraphobia; Sleep, Sleep Disorders, and Stress.

Further Reading

- Cates, M. E., Bishop, M. H., Davis, L. L., et al. (2004). Clonazepam for treatment of sleep disturbances associated with combat-related posttraumatic stress disorder. *Annals of Pharmacotherapy* 38, 1395–1399.
- Davidson, J. R. (2004). Use of benzodiazepines in social anxiety disorder, generalized anxiety disorder, and post-traumatic stress disorder. *Journal of Clinical Psychiatry* 65(supplement 5), 29–33.
- Furukawa, T. A., Streiner, D. L. and Young, L. T. (2002). Antidepressant and benzodiazepine for major depression (Cochrane Review). In: The Cochrane Library (Issue 1). Chichester: Wiley.
- Kindt, M. and Engelhard, I. M. (2005). Trauma processing and the development of posttraumatic stress disorder. *Journal of Behavior Therapy and Experimental Psychiatry* 36, 69–76.
- Korpi, E. R. and Sinkkonen, S. T. (2006). GABA(A) receptor subtypes as targets for neuropsychiatric drug development. *Pharmacology & Therapeutics* 109, 12–32.
- Mancuso, C. E., Tanzi, M. G. and Gabay, M. (2004). Paradoxical reactions to benzodiazepines: literature review and treatment options. *Pharmacotherapy* 24, 1177–1185.
- Möhler, H. and Okada, T. (1977). Benzodiazepine receptor: demonstration in the central nervous system. *Science* 198, 849–851.
- Otto, M. W., Bruce, S. E. and Deckersbach, T. (2005). Benzodiazepine use, cognitive impairment, and cognitive-behavioral therapy for anxiety disorders: issues in the treatment of a patient in need. *Journal of Clinical Psychiatry* 66(supplement 2), 34–38.
- Rosenbaum, J. F. (2005). Attitudes toward benzodiazepines over the years. *Journal of Clinical Psychiatry* 66(supplement 2), 4–8.
- Rudolph, U. and Möhler, H. (2004). Analysis of GABA_A receptor function and dissection of the pharmacology of benzodiazepines and general anesthetics through mouse genetics. *Annual Review of Pharmacology and Toxicology* 44, 475–498.
- Sanger, D. J. (2004). The pharmacology and mechanisms of action of new generation, non-benzodiazepine hypnotic agents. *CNS Drugs* 18(supplement 1), 9–15.
- Schatzberg, A. F., Cole, J. O. and DeBattista, C. (2005). *Manual of clinical psychopharmacology* (5th edn.). Washington, DC: American Psychiatric Publishing.
- Soumerai, S. B., Simoni-Wastila, L., Singer, C., et al. (2003). Lack of relationship between long-term use of benzodiazepines and escalation to high dosages. *Psychiatric Services* 54, 1006–1011.
- Stein, D. J., Ipser, J. C. and Balkom, A. J. (2004). Pharmacotherapy for social phobia (Cochrane Review). In: The Cochrane Library (Issue 4). Chichester: Wiley.
- Vermeeren, A. (2004). Residual effects of hypnotics: epidemiology and clinical implications. *CNS Drugs* 18, 297–328.

Bereavement

P J Clayton

American Foundation for Suicide Prevention,
New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
P J Clayton, volume 1, pp 304–311, © 2000, Elsevier Inc.

Introduction

Depressive Symptoms and Course

Physical Symptoms, Substance Use, and Medical Treatment

Psychiatric Disorders – Depression, Anxiety, and Mania

Complicated or Traumatic Bereavement

Mortality

Pathologic Outcomes and Predictors of Outcome

Treatment

Glossary

<i>Bereavement</i>	The state or fact of being bereaved, suffering the death of a loved one; the loss of a loved one by death.
<i>Disorders</i>	So disturb the regular or normal functions of; an abnormal physical or mental condition: ailment.
<i>Grief</i>	Deep and poignant distress caused by loss, a cause of such suffering.
<i>Mourning</i>	The act of sorrowing; a period of time during which signs of grief are shown.
<i>Outcome</i>	Something that follows as a result or consequence.
<i>Complicated grief</i>	A disordered psychic or behavioral state resulting from a deep and poignant distress caused by or as if by bereavement.

Introduction

Although no-one would argue that the loss by death of someone close is a significant life event, the degree of adversity that the death brings to the individual is variable. The field is just beginning to identify and study people who are most severely affected. In studying bereavement, most investigators have used entry in widowhood as the model, although the response to the death of a child is probably more stressful. This article will concentrate on research on bereavement which by definition is a reaction of a person to loss by death. Other terms in the literature which should not be confused with bereavement are grief, which is the emotional and psychological reaction to any loss but not limited to death, and mourning, which is the

social expression of bereavement or grief, frequently formalized by custom or religion.

At this point there have been numerous longitudinal follow-up studies of people who have been recently bereaved. The majority deal with the widowed although there are excellent studies of children who have lost a parent. Despite the fact that the mode of selection, response rate, gender, age group, time since the death to the first interview, comparison groups, and the selection of instruments and outcome measures differ greatly from study to study, the similarities of outcome outweigh the differences. Our own data highlight the major outcomes, although important additional findings have been identified and there are still areas of controversy.

Bereavement is the ideal condition to study the effect of stress on the organism. It is not species specific so it can be defined identically across species and generations of species. It is a real event, not artificial, imagined, or devised, it is datable, it occurs in present time so it can be studied immediately, not retrospectively, and even prospectively. Finally, the significance of the event would not be disputed. Before reviewing the morbidity and mortality associated with this clear-cut, easily defined stress, it is important to emphasize that by far the most usual outcome following this experience is numbness for a very brief period, depression for a variable time, and then recovery. Recovery is the return of the individual to the level of functioning that he or she had established before the death and/or a change in a positive direction. In some cases this takes months, in others, up to 2 years or more. This is described by Lund et al. (1993: 24) in older adults. Only 10–15% of individuals develop chronic depression, with even fewer experiencing complicated grief. In studying the biology of this particular stress, whole populations have to be studied so that we define the usual, natural outcomes as well as the abnormal responses.

Depressive Symptoms and Course

In studying relatives' reactions to the loss of a significant person, we collected data on three different samples. The first sample was a group of 40 bereaved relatives of people who had died at a general hospital in St. Louis, United States. The second was a sample of widows and widowers chosen from the death certificates and seen at 1, 4, and 13 months after the deaths of their spouses. The third sample was a consecutive series of young (<45 years old) widowed people also seen within 1 month after the death and

again at 1 year. The last two samples were matched through the voter registry with married people of the same age and sex who had not had first degree relatives die, and who were also followed prospectively to ascertain their morbidity and mortality. The findings are exemplified in Table 1 which displays the depressive symptoms that were endorsed at 1 and 13 months. In the first month, a large majority of the sample experienced depressed mood; anorexia and beginning weight loss; insomnia that was initial, middle, and terminal; marked crying; some fatigue and loss of interest in their surroundings, but not necessarily the people around them; restlessness; and guilt. The guilt that these people expressed was different to the guilt seen in depressed patients. For the most part it was guilt of omission surrounding either the terminal illness or the death. Irritability was common but overt anger was uncommon. Suicidal thoughts and ideas were rare. When they did occur they occurred in younger men after the death of a wife or in men and

women after the death of a child. Hallucinations were not uncommon. Many widows and widowers, when asked, admitted that they had felt touched by their dead spouse or had heard his or her voice, had seen his or her vision or had smelled his or her presence. The misidentification of their dead spouses in a crowd was extremely common.

We matched some of our subjects with well-studied hospitalized depressed patients and found, as Freud suggested, symptoms such as feeling hopeless, worthless, being a burden, psychomotor retardation, wishing to be dead, and thinking of suicide differentiated the depressed from the bereaved.

By the end of the first year the somatic symptoms of depression had remarkably improved. Low mood (usually associated with specific events or holidays), restlessness, and poor sleep continued. The few psychological symptoms of depression, although rare, resolved less easily. In our studies, the symptoms were as frequent in men as in women, in people who had lost a spouse suddenly or after a lingering death, in people who had good or bad marriages, and people who were religious or nonreligious. Young people were more likely to have a severer immediate reaction, but by 1 year their outcome was similar. There was no relationship between financial status and outcome.

As Table 2 shows, when the frequency of depressive symptoms were compared to the community controls, all symptoms were more common among the bereaved.

When depressive symptoms reported at 13 months were examined by those with symptoms for the entire year and those who had just developed them, the percent who had new symptoms was similar to those reported by the controls who had had no deaths (Table 3). Those who were disturbed early were more likely to be disturbed later. Delayed grief, like delayed posttraumatic stress disorder (PTSD), may be an interesting ideologic concept but remains unproven. It is also counterintuitive since the usual recognized reaction to a physical or psychological insult is an immediate injury with a gradual resolution. This has been confirmed by others.

Physical Symptoms, Substance Use, and Medical Treatment

We systematically ask about physical symptoms, usually associated with major depression, in both the recently widowed and control subjects (Table 4). In general, physical symptoms were not more frequent in recently bereaved and controls subjects. Nor did they change much over the year. The largest difference was recorded in "other pains" which were mainly in the elderly bereaved and consisted of

Table 1 Frequency of depressive symptoms 1 and 13 months after bereavement

Symptoms	Frequency (%)	
	1 Month (<i>n</i> = 149 ^a)	13 Months (<i>n</i> = 149 ^a)
Crying	89	33 ^d
Sleep disturbance	76	48 ^d
Low mood	75	42 ^d
Loss of appetite	51	16 ^d
Fatigue	44	30 ^c
Poor memory	41	23 ^d
Loss of interest	40	23 ^d
Difficulty concentrating	36	16 ^d
Weight loss of ≥ 2.25 kg	36	20 ^c
Feeling guilty	31	12 ^d
Restlessness (<i>n</i> = 89)	48	45
Reverse diurnal variation	26	22
Irritability	24	20
Feels someone to blame	22	22
Diurnal variation	17	10
Death wishes	16	12
Feeling hopeless	14	13
Hallucinations	12	9
Suicide thoughts	5	3
Fear of losing mind	3	4
Suicide attempts	0	0
Feeling worthless	6	11
Feels angry about death	13	22 ^b
Depressive	42	16 ^d

^a*n* varies from symptom to symptom mostly 148;

^bsignificant by McNemar's chi-square, degrees of freedom (*df*) = 1, $P \leq 0.02$;

^csignificant by McNemar's chi-square, *df* = 1, $P \leq 0.01$;

^dsignificant by McNemar's chi-square, *df* = 1, $P \leq 0.001$. Reproduced with permission from Clayton, P. J. (1982). Bereavement. In: Paykel, E. S. (ed.) *Handbook of affective disorders*. London: Churchill Livingstone.

Table 2 Frequency of depressive symptoms at any time in the first year of bereavement and in controls

Symptoms	Frequency (%)	
	Bereaved (n = 149 ^a)	Controls (n = 131 ^a)
Crying	90	14 ^e
Sleep disturbance	79	35 ^e
Low mood	80	18 ^e
Loss of appetite	53	4 ^e
Fatigue	55	23 ^e
Poor memory	50	22 ^e
Loss of interest	48	11 ^e
Difficulty concentrating	40	13 ^e
Weight loss of ≥ 2.25 kg	47	24 ^e
Feeling guilty	38	11 ^e
Restlessness	63	27 ^e
Irritability	35	21 ^b
Diurnal variation	22	14
Death wishes	22	5 ^e
Feeling hopeless	19	4 ^d
Hallucinations	17	2 ^e
Suicidal thoughts	8	1 ^c
Fear of losing mind	7	5
Suicide attempts	0	0
Worthlessness	14	15
Depressive syndrome	47	8 ^d

^an varies from symptom to symptom;

^bsignificant by chi-square, degrees of freedom (df) = 1, $P \leq 0.05$;

^csignificant by chi-square, df = 1, $P \leq 0.01$;

^dsignificant by chi-square, df = 1, $P \leq 0.0005$;

^esignificant by chi-square, df = 1, $P \leq 0.0001$. Reproduced with permission from Clayton, P. J. (1982). Bereavement. In: Paykel, E. S. (ed.) *Handbook of affective disorders*. London: Churchill Livingstone.

arthritic type pains. In our studies, there were no more physicians' visits or hospitalizations in the year following the bereavement. This is controversial.

The most striking finding of this research is illustrated in Table 4. After a significant stress such as bereavement the use of alcohol, tranquilizers, and hypnotics significantly increased. For the most part these were not new users but people who had used them before. Cigarette smoking also increased. The morbidity and mortality of bereavement may be related to these changes in behavior. It also highlights how people behave after significant stress.

In a separate study, we examined 249 psychiatric inpatients and similarly matched hospital controls and found that there were no significant differences between the groups for the loss of a first degree relative or spouse in the 6 months or 1 year prior to admission to the hospital. In each case only 2% reported such a loss in the 6 months and 3 or 4% in the 7–12 months prior to the hospitalization. In the psychiatric patients, the major diagnosis was a mood disorder. There were diagnoses of alcoholism in

Table 3 Frequency of depressive symptoms at 1 month

Symptom	Those with symptoms at 1 month		Those without symptoms at 1 month (new symptoms)	
	%	n	%	n
Crying	36	132	6	16
Sleep disturbance	59	113	11	35
Low mood	49	111	22	37
Loss of appetite	28	76	4	72
Fatigue	45	65	19	83
Poor memory	34	61	15	88
Loss of interest	40	57	12	84
Difficulty concentrating	33	52	6	94
Weight loss of ≥ 2.25 kg	22	51	17	92
Feeling guilty	25	49	6	97
Restlessness	63	43	28	46
Reverse diurnal variation	39	38	17	109
Irritability	37	35	15	113
Feels someone to blame	63	32	11	111
Diurnal variation	24	25	7	122
Death wishes	33	24	7	124
Feeling hopeless	55	20	6	128
Hallucinations	29	17	6	131
Suicidal thoughts	0	8	3	139
Fear of losing mind	0	5	4	143
Suicide attempts	0	0	0	147
Feeling worthless	56	9	8	139
Feel angry about death	63	19	16	128
Depressive syndrome	27	63	8	86

Reproduced with permission from Clayton, P. J. (1982). Bereavement. In: Paykel, E. S. (ed.) *Handbook of affective disorders*. London: Churchill Livingstone.

both the bereaved psychiatric patients and bereaved hospitalized controls which related to their increased drinking after the death.

Psychiatric Disorders – Depression, Anxiety, and Mania

In our first study of widowhood, 35% of the subjects had a depressive syndrome at 1 month, 25% at 4 months, and 17% at 1 year. Forty-five per cent were depressed at some point during the year and 11% were depressed for the entire year. Adding the second younger sample, 42% were depressed at 1 month and 16% met the criteria at 1 year. Forty-seven per cent were depressed at some time during the year compared to 8% of controls and again, 11% for the entire year. These findings are remarkably similar to those of a study extensively reported by Zisook and Shuchter (1991: 1346; 1993: 157). They identified a large sample of widows and widowers from death certificates and solicited volunteers to be interviewed. The demographics of the final sample were very similar to the first widowed sample reported here.

Table 4 Frequency of physical symptoms in 1 year in bereaved and controls

Symptom	Frequency (%)	
	Probands (n = 149)	Controls (n = 131)
Headaches	36	27
Dysmenorrhea	38	20
Other pains	44	18 ^b
Urinary frequency	30	23
Constipation	27	24
Dyspnea	27	16 ^a
Abdominal pain	26	11 ^c
Blurred vision	22	13
Anxiety attacks	15	8
Alcohol use	19	9 ^a
Tranquilizers	46	8 ^d
Hypnotics	32	2 ^d
Physician visits	79	80
Physician visits ≥ 3	45	48
Physician visits ≥ 6	27	29
Hospitalizations	22	14
General poor health	10	7

^aSignificant by chi-square, degrees of freedom (df) = 1, $P \leq 0.05$;

^bsignificant by chi-square, df = 1, $P \leq 0.005$;

^csignificant by chi-square, df = 1, $P \leq 0.001$;

^dsignificant by chi-square, df = 1, $P \leq 0.0001$; n varies slightly from symptom to symptom. Reproduced with kind permission of Springer Sciences and Business Media. *New results in depression*, 1986, Bereavement and its relation to clinical depression, Clayton, P. J. Berlin-Heidelberg: Spring-Verlag.

They reported that 24% of their sample was depressed at 2 months, 23% at 7 months, 16% at 13 months, and 14% at 25 months. In all of these studies the best predictor of depression at 13 months was depression at 1 month. In the Zisook and Shuchter studies, a past history of depression also predicted depression at 1 year. In our own studies, because we did not have enough depressed patients, it was any psychiatric morbidity prior to the death that predicted depression at 1 year. The other common correlate of depression at 1 year was poor physical health prior to the death of the spouse. So, prior poor physical or mental health predicted poor outcome at 13 months. Although there are methodologic issues in defining both anticipation of a death (e.g., expected versus unexpected) and outcomes (depression, death, remarriage, etc) in our data, unexpected deaths produced more severe immediate responses but the differences disappeared by 13 months, again confirmed by others. The best definitions of unexpected death, is a death in a healthy individual which occurred in less than 2 hours. Using that and comparing it to those bereaved after longer terminal illnesses we found that it has been reported that there was more morbidity especially as measured in physician visits, and long lingering grief symptoms.

In more than one study, about 25% of recently bereaved subjects met criteria for some anxiety disorder immediately following bereavement. More than half, however, had generalized anxiety disorder and that correlated with the severity of the depressive disorder. Others have shown, without clarifying the nature of the anxiety disorder, the major disorders that affect recently bereaved people are depressive disorder alone or concurrent depression and anxiety. New episodes of any anxiety disorder were rare after bereavement. It may be wisest to conclude that these are anxiety symptoms embedded in a depression disorder.

An occasional person, with a bipolar disposition, develops mania following a significant death, probably secondary to the sleep deprivation that the stress induces.

Complicated or Traumatic Bereavement

Lindemann (1944: 24) probably wrote the first papers on bereavement that dealt with prospective observations of 13 bereaved subjects who were hospitalized at Massachusetts General Hospital after the Coconut Grove fire. He combined these observations with those of his own patients who lost a relative during treatment, with relatives whom he saw of patients who died at the hospital, and of relatives of members of the Armed Forces. Although he considered his observations to be of normal grief, many of them were probably related to a pathologic response. More recently, investigators have begun to quantify the abnormal response. Prigerson and colleagues (1995: 616) have developed an inventory of complicated grief and shown it to be related to the severity of depressive symptoms, but with distinct symptoms. Building on earlier inventories such as the Texas Grief Inventory and the Grief Measurement Scale, they took symptoms such as preoccupation with death, crying, searching and yearning for the deceased, disbelief about the death, being stunned, and not accepting the death and added questions about preoccupation with thoughts of the deceased, anger, distrust, detachment, avoidance, replication of symptoms that the deceased experienced, auditory and visual hallucinations, and guilt, loneliness, bitterness, and envy of others who had not lost someone close. Their instrument was shown to be reliable and have criterion validity. Whether it is just defining the most severe depression is not yet clear. Modifying the Grief Measurement Scale in accordance with the above symptoms, Prigerson et al. (1997: 616) looked at a previously ascertained sample that were followed from before the death to until 25 months after the death. They found that if they identified bereaved

at 6 months who scored high on their inventory, at 13 months there was a correlation between high scores and change in smoking and eating, depression, and high systolic blood pressure. At 25 months, these people had an increased risk to develop heart trouble and cancer and more often expressed suicidal ideation as ascertained on a single item. If they tried to separate the sample on complicated grief at 2 months, they could not predict unfavorable outcomes. There are others working on similar instruments, especially in Australia. They measured bereavement phenomena and developed a scale for core bereavement items. This scale also measures images and thoughts around the death, yearning and pining about the death, and some grief symptoms such as sadness, loneliness, longing, tearfulness, and loss of enjoyment. Their inventory was developed using parents of deceased children, bereaved spouses, and bereaved adult children who scored high in descending order on the inventory. However they have not yet correlated the questionnaire with outcome.

Some of the symptoms in all of these inventories are similar to symptoms that occur in PTSD, and it is difficult to justify a new category where the major difference is that in the bereaved, the mood is sadness and loneliness focused on the loss compared to fear or horror in PTSD. It is still an unanswered question. Should we have a diagnosis of complicated bereavement or acknowledge that after certain deaths, most especially those that are sudden and violent, some individuals will develop PTSD which is usually comorbid with major depression? Clearly these investigators are defining an important bereavement response that identifies survivors with a poorer outcome. For other researchers who are studying human reactions to stress, symptoms such as those in the inventory of complicated grief should be identified and studied prospectively. The identification of such a syndrome also makes it possible to study, more specifically, various interventions.

Mortality

More recently the method of study (collecting health data on a large core of men and women and following them prospectively to ascertain point of widowhood and using the Cox hazards regression mode to estimate the effect of bereavement on mortality while controlling for the effects of other covariates), should have greatly improved the validity of the mortality data. Unfortunately, there still are controversies. Three studies have used these methods. De Leon et al. (1993: 519) reported that elderly, young-old (aged 65–74 years), and old-old men (aged > 74 years) showed an increased mortality in the first 6 months of

widowhood. But, because there was an age difference between the future widowed and nonwidowed, when they were age adjusted, the death rate fell to a non-significant level. Among young-old women there was also an increased risk of death in the first 6 months which only decreased slightly after the age adjustment, but enough to make it of borderline significance. The authors suggested that the “true risk period” may be even shorter than 6 months. Schaefer et al. (1995: 1142) also reported that mortality following bereavement was significantly elevated in both men and women after adjusting for age, education, and other predictors of mortality. Here, the highest mortality occurred 7–12 months following the bereavement. Ebrahim et al. (1995: 834) studied 7735 middle-aged men (40–59 years) over 5 years. They included never married, divorced, separated, widowed, and married men and found that there was no increase in mortality in men who became widowed during the follow-up time. There were only 24 men who became widowed in that follow-up period which probably limited the findings. Lusyne et al. (2001: 281) confirmed an excess mortality in the first 6 months of widowhood which was higher for men than women. As others had found, it was highest in the youngest bereaved (<60 years) and dropped with each older group and almost disappeared for the oldest old. There is also a twin study of mortality after spousal bereavement. Without relating whether the twins were mono- or dizygotic, the authors found an increased mortality for young-old (under 70 years) men and women during the first year of bereavement. The death hazard diminished the longer the bereaved were followed. The data were analyzed for smoking status, excessive alcohol consumption, education, body mass, cardiovascular disease, respiratory disease, and other chronic illness and found that these did not affect the relative hazard estimates. Other studies show that remarriage, which is much more likely to occur in men, protects men from death either because the healthier remarry or the remarriage in itself is protective. The death rates of those who remarry are similar to married men of the same age. Taken as a whole, these studies support the fact that there is probably an increased mortality after bereavement in men and women under 70 years of age in the first year.

From these studies there does not seem to be an increased mortality from any one particular cause, such as cancer or cardiovascular disease. Other studies, however, have shown an increase in suicide after a bereavement, although when looking at smaller samples most of those who committed suicide were psychiatrically ill prior to their spouses' deaths. A case control study of 100 elderly suicides compared

to well-matched deaths from other causes, found that the major independent risk factors in those who committed suicide were family discord and mental disorders (mainly depression and substance abuse). Recent bereavement did not distinguish the two groups.

Pathologic Outcomes and Predictors of Outcome

Before reviewing the predictors of pathologic outcomes it should be noted that all studies of widowhood, including those on mortality, emphasize that psychological stability predicts a good outcome, although unfortunately no-one has measured personality traits or variables prior to entry into widowhood. Good health predicts a better outcome. And not having sleep difficulty or having low scores on depression ratings in the first month of bereavement also predicts a better outcome.

There are many ways to assess outcome. Loneliness certainly is the most characteristic outcome following bereavement and this persists in many for years. Under psychological morbidity, chronic depression as measured 1 year after the death and complicated bereavement both lead to a significant morbidity. There is also physical morbidity. The mortality noted postbereavement still needs further study.

There are very few definitive predictors of chronic depression. A past history of psychiatric illness particularly depression predicts it, poor prior physical health may predict it, and depression at 1 month after the bereavement definitely predicts it. Perhaps a bereavement that follows an instantaneous and totally unexpected death predicts it, although it might not be that it is the unexpected aspect of it, but rather whether it is violent in nature. Studies will clarify the predictions as data collection improves. All the other variables, gender, young age, perceived lack of social supports and socioeconomic income do not predict it. Gender is important to note, because major depression is overwhelmingly more common in women than in men but chronic depression after bereavement is not. Stress-related depression affects both genders equally.

Treatment

The vast majority of people who experience a loss will recover gradually without any interventions. It is just that fact that probably contributes to the conclusion in a recent salient review of all qualifying grief counseling studies, which the effect of such counseling is quite weak and in some cases could even be harmful.

In those who become chronically depressed or have complicated grief, an intervention at 6 months after

the bereavement is indicated. In a recent report, 95 volunteers with complicated grief of long duration (>2 years) were randomized to treatment with either interpersonal therapy (IPT) or complicated grief therapy (IPT modified to include cognitive therapy based on techniques for addressing trauma). The results showed, when looking at treatment completers, the complicated grief therapy was superior to IPT especially in those with losses through violent deaths. Still, the response rate was only 51%. Patients also had high rates of comorbid depression and PTSD and 45% were taking antidepressants. It may be the kind of therapy for a specific group of bereaved that Jordan and Neimeyer (2003: 765) were urging researchers to study. Most importantly, the elements of these therapies may be applicable to other stress-related disorders.

Antidepressants have also been used and been shown to be efficacious. For example, an older study of treatment of bereavement-related major depressive episodes in later life with nortriptyline, IPT, the combination, or neither showed that nortriptyline alone or with IPT significantly improved the depressive disorder. IPT alone was not significantly different to the placebo. The outcomes are notable because very few people relapsed when the treatments were stopped and the response rate was relatively high in both nortriptyline treatment groups and the placebo group. Thus it seems to be a good condition to treat with antidepressant medication. It is also important to remember that should the patient be psychotic or suicidal, electroconvulsive therapy (ECT) is probably indicated.

Further Reading

- Clayton, P. J. (1982). Bereavement. In: Paykel, E. S. (ed.) *Handbook of affective disorders*. London: Churchill Livingstone.
- Clayton, P. J. (1986). Bereavement and its relation to clinical depression. In: Hippius, H., Klerman, G. L. & Matussek, N. (eds.) *New results in depression*. Berlin-Heidelberg: Springer-Verlag.
- Clayton, P. (2004). Bereavement and depression. In: Roose, S. & Sackeim, H. (eds.) *Late life depression*. Oxford: Oxford University Press.
- Clayton, P., Demarais, L. and Winokur, G. (1968). A study of normal bereavement. *American Journal of Psychiatry* 125, 168–178.
- Clayton, P. J. and Darvish, H. S. (1979). Course of depressive symptoms following the stress of bereavement. In: Barrett, J. E. (ed.) *Stress and mental disorder*, pp. 121–136. New York: Raven Press.
- Clayton, P. J., Halikas, J. A. and Maurice, W. L. (1971). The bereavement of the widowed. *Disease of the Nervous System* 32, 597–604.

- Clayton, P. J., Herjanic, M., Murphy, G. E., et al. (1974). Mourning and depression, their similarities and differences. *Canadian Psychiatric Association Journal* 19, 309–312.
- Clayton, P. J., Parilla, R. H. Jr., and Bieri, M. D. (1980). Methodological problems in assessing the relationship between acuteness of death and the bereavement outcome. In: Reiffel, J. (ed.) *The psychosocial aspects of cardiovascular disease: the patient, the family and the staff*, pp. 267–275. New York: Columbia University Press.
- De Leon, C. F. M., Kasl, S. F. and Jacobs, S. (1993). Widowhood and mortality risk in a community sample of the elderly: a prospective study. *Journal of Clinical Epidemiology* 46, 519–527.
- Ebrahim, S., Wannamethee, G., McCallum, A., et al. (1995). Marital status, change in marital status, and mortality in middle-aged British men. *American Journal of Epidemiology* 142, 834–842.
- Frost, N. R. and Clayton, P. J. (1977). Bereavement and psychiatric hospitalization. *Archives of General Psychiatry* 34, 1172–1175.
- Jordan, J. R. and Neimeyer, R. A. (2003). Does grief counseling work? *Death Studies* 27, 765–786.
- Lindermann, E. (1944). Symptomatology and management of acute grief. *American Journal of Psychiatry* 101, 141–148.
- Lund, D. A., Caserta, M. S. and Dimond, M. F. (1993). The course of spousal bereavement in later life. In: Stroebe, M. S., Stroebe, W. & Hansson, R. O. (eds.) *Handbook of bereavement*, pp. 24–254. Cambridge: Cambridge University Press.
- Lusyne, P., Page, H. and Lievens, J. (2001). Mortality following conjugal bereavement, Belgium 1991–96: the unexpected effect of education. *Population Studies* 55, 281–289.
- Prigerson, H. G., Bierhals, A. J., Kasl, S. V., et al. (1997). Traumatic grief as a risk factor for mental and physical morbidity. *American Journal of Psychiatry* 154, 616–623.
- Prigerson, H. G., Maciejewski, P. K., Reynolds, III, C. F., et al. (1995). Inventory of Complicated Grief. A scale to measure maladaptive symptoms of loss. *Psychiatry Research* 59, 65–79.
- Rubenowitz, E., Waern, M., Wilhelmson, K., et al. (2001). Life events and psychosocial factors in elderly suicides – a case-control study. *Psychological Medicine* 31, 1193–1202.
- Schaefer, C., Quesenberry, J. C. P. Jr. and Wi, S. (1995). Mortality following conjugal bereavement and the effects of a shared environment. *American Journal of Epidemiology* 141, 1142–1152.
- Shear, K., Frank, E., Houck, P. R., et al. (2005). Treatment of complicated grief. *JAMA* 293, 2601–2608.
- Zisook, S. and Shuchter, S. R. (1991). Depression through the first year after the death of a spouse. *American Journal of Psychiatry* 148, 1346–1352.
- Zisook, S. and Shuchter, S. R. (1993). Uncomplicated bereavement. *Journal of Clinical Psychiatry* 54, 365–372.
- Zisook, S., Chentsova-Dutton, Y. and Schuchter, S. R. (1998). PTSD following bereavement. *Annals of Clinical Psychiatry* 10, 157–163.

Beta-Adrenergic Blockers

M B Hamner and G W Arana

Medical University of South Carolina and VA Medical Center, Charleston, SC, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M B Hamner and G W Arana, volume 1, pp 312–320,

© 2000, Elsevier Inc.

Introduction and Definition

Chemistry

Pharmacology

Role in Stress, Learning, and Memory

Therapeutic Use in Psychiatry

Therapeutic Use in Medical Disorders

Safety and Tolerability

Summary

Glossary

Adrenergic receptors

Specific molecular recognition sites for norepinephrine or epinephrine located on cell membranes in the central or peripheral sympathetic nervous system. These receptors are coupled with G-proteins (proteins that bind guanosine triphosphate and hydrolyze it to guanosine monophosphate) that act as transducers for intracellular second messenger systems such as cyclic adenosine monophosphate. In turn, these second messenger systems may mediate a number of biological responses, ranging from ion channel effects to release of neurotransmitters to regulation of gene expression.

Anxiety

Apprehension of danger and dread accompanied by restlessness, tension,

<i>Autoreceptors</i>	tachycardia, and dyspnea unattached to a clearly identifiable stimulus. Sites on the nerve axon terminal that respond to the neuron's own neurotransmitter.
<i>β-Adrenergic antagonist</i>	A drug that prevents stimulation of the β-adrenergic receptors of the nerves of the sympathetic nervous system and therefore decreases the activity of the heart and reduces the blood pressure, and may have other properties such as reduction of physiological symptoms of anxiety.
<i>Norepinephrine</i>	A hormone, closely related to epinephrine and with similar actions, secreted by the medulla of the adrenal and also released as a neurotransmitter by sympathetic nerve endings and by axonal projections from the neurons in the locus ceruleus in the brain.
<i>Posttraumatic stress disorder</i>	An anxiety disorder developing in persons who have experienced a physically or emotionally traumatic event that is outside of the range of normal human experience, such as combat, physical assault, rape, or disasters such as home fires.
<i>Stress</i>	Any factor that threatens the health of the body or has an adverse effect on its functioning, such as injury, disease, or worry.

Introduction and Definition

The β-adrenergic receptor antagonists have been used since the late 1950s for the treatment of hypertension, cardiac arrhythmias, angina, thyrotoxicosis, glaucoma, migraine, tremor, and, in psychiatry, anxiety, akathisia, and other syndromes. Development of propranolol, the prototype β-blocker, followed Ahlquist's hypothesis that the effects of catecholamines are mediated by α- and β-adrenergic receptors.

β receptors are widely distributed in the vasculature and in neurons of the central and peripheral nervous system. The β receptors are stimulated by norepinephrine. Noradrenergic systems are likely critical in mediating anxiety and stress response. Thus, there has long been interest in β-blockers in psychiatry because of their pharmacological effects in dampening behavioral or physiological responses to noradrenergic stimulation. Agents that have been most frequently studied in various psychiatric disorders include propranolol (Inderal), metoprolol (Lopressor), nadolol (Corgard), acebutolol (Sectral), labetalol (Normodyne Trandate), pindolol (Visken), and atenolol (Tenormin).

Chemistry

The β-blockers share structural similarity with dichloroisoproterenol, first synthesized in the 1950s. However, dichloroisoproterenol is a partial agonist, an action that was thought to preclude its use clinically. The first pure antagonist, pronetholol, was associated with thymic tumors in mice. Propranolol, the next compound to be developed, became the prototype β-blocker. The chemical designation, molecular formula and molecular weight of seven representative β-blockers are shown in [Table 1](#). The molecular structures are illustrated in [Figure 1](#).

Pharmacology

Pharmacodynamics

The β-blockers are competitive antagonists of norepinephrine and epinephrine β-adrenergic receptors. Epinephrine and norepinephrine are released as stress hormones from the adrenal medulla. Norepinephrine is also released from neurons in the peripheral sympathetic nervous system. These catecholamines are

Table 1 Basic chemistry of selected β-blockers

<i>β-Blocker</i>	<i>Chemical designation</i>	<i>Molecular formula</i>	<i>Molecular weight</i>
Acebutolol	± N-[13-acetyl-4-[2-hydroxy-3-1[(methylethyl)amino] propoxy]phenyl]butanamide	C ₁₈ H ₂₈ N ₂ O ₄ HCL	327.90
Atenolol	4-[2'-hydroxy-3'-[(1-methylethyl)amino]propoxy]-phenyl]ethanamide	C ₁₄ H ₂₂ N ₂ O ₃	266.00
Labetalol	5-[1-hydroxy-2-[(1-methyl-3-phenylpropyl)amino]ethyl]salicylamide conochloride	C ₁₉ H ₂₄ N ₂ O ₃	364.9
Metoprolol	1-(isopropylamino)-3-[p-(2-methoxy-ethyl)phenoxy]-2-propanol(2:1) dextro-tartrate salt	C ₃₄ H ₅₆ N ₂ O ₁₂	684.82
Nadolol	1-(tert-butylamino)-3-[(5,6,7,8-tetrahydro-cis-6,7-naphthyl)oxy]-2-propanol	C ₁₇ H ₂₇ NO ₄	309.40
Pindolol	1-(indol-4-yloxy)-3-(isopropylamino)-2-propanol	C ₁₄ H ₂₀ N ₂ O ₂	248.33
Propranolol	1-(isopropylamino)-3-(1-naphthyl)-2-propanol hydrochloride	OCH ₂ CHOHCH ₂ NHCH(CH ₃) ₂	295.81

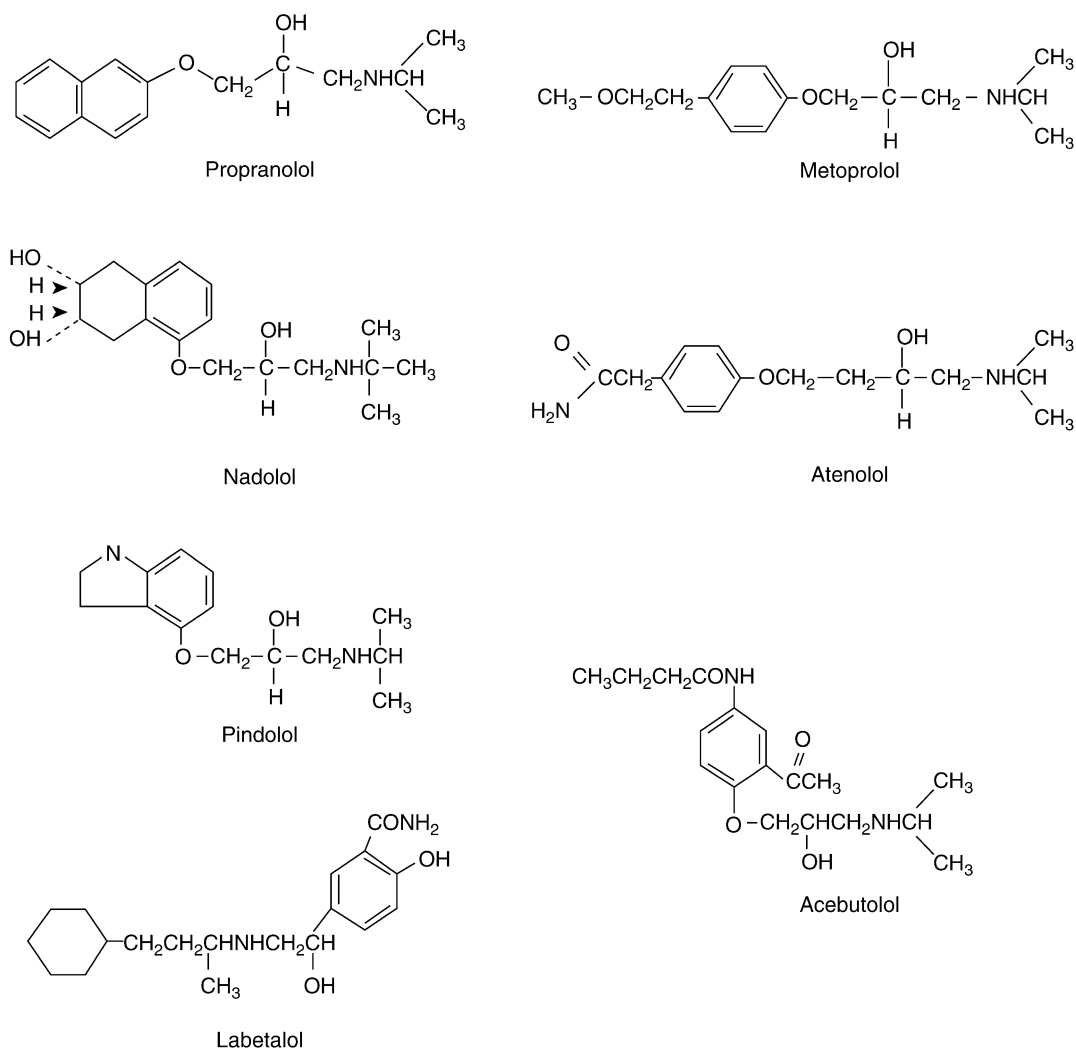


Figure 1 Molecular structures of the β -adrenergic receptor antagonists.

involved in cardiovascular regulation and stress responses. In the central nervous system (CNS), norepinephrine is produced by neurons in the brain stem tegmentum. The locus ceruleus is the primary noradrenergic nucleus, with widespread projections to the spinal cord, cortex, limbic regions, and cerebellum. CNS noradrenergic systems are involved in emotional regulation, learning and memory, pain perception, control of blood pressure, and the neuroendocrine response to stressful stimuli.

The effects of norepinephrine in the brain are mediated by two classes of receptors: α and β receptors. These are further divided into α_1 and α_2 and β_1 and β_2 and β_3 subtypes. Activation of α_1 receptors produces neuronal excitation partly linked to phosphoinositide activation. α_2 receptors are inhibitory autoreceptors that decrease norepinephrine release. Consequently, α_2 receptor activation reduces both central and peripheral noradrenergic activity. α_2

receptors are located in the CNS on locus ceruleus noradrenergic neurons. Clonidine is an example of an α_2 agonist with antihypertensive properties due to reduction in noradrenergic activity. In contrast, yohimbine, an α_2 receptor antagonist, may release norepinephrine, resulting in anxiety due to central noradrenergic activation and vasoconstriction in the periphery.

β receptors are widely distributed in both the CNS and the periphery. β_1 receptors are localized in the cardiac sinoatrial and atrioventricular nodes and myocardium. β_2 receptors are located in the pulmonary bronchi, the eye, peripheral vasculature, bladder and uterus, and liver (where they mediate glucose metabolism). β_2 and β_3 receptors are also involved in lipid metabolism.

The numerous β -blockers vary based on their relative pharmacokinetics, pharmacodynamics, and lipid solubility. Examples of these differences include

relative β_1 versus β_2 antagonism, intrinsic sympathomimetic properties, α -adrenergic receptor affinity, lipid solubility, direct vasodilator properties, certain pharmacokinetic properties, and effects on other neurotransmitter systems (e.g., serotonergic properties). Lipid solubility is important because entry of drug into the brain is dependent on its ability to cross the blood–brain barrier, which is a highly lipid structure. The lipid solubility properties of the seven representative β -blockers are shown in [Table 2](#).

Propranolol is a nonselective β_1 and β_2 antagonist. Atenolol is a relatively selective β_1 antagonist. [Table 2](#) shows the relative β_1 versus β_2 affinities for the commonly used β -blockers.

Pharmacokinetics

Propranolol is almost completely absorbed following oral administration. Nadolol, atenolol, and metoprolol are partially absorbed. Nadolol, in contrast to the other listed agents, has a relatively long plasma half-life that allows once-daily dosing. Propranolol is avidly protein bound (90%) and undergoes extensive hepatic metabolism. In contrast, nadolol and atenolol are not as extensively metabolized and undergo renal excretion.

Nadolol and atenolol are the least lipophilic of the β -blockers and therefore have less penetrance of the blood–brain barrier, yielding more peripheral as opposed to central effects. In contrast, propranolol and metoprolol are more lipophilic and therefore have relatively greater CNS effects. Pindolol is of intermediate lipophilicity – this agent has been of recent interest as a possible antidepressant-potentiating drug due to its intrinsic sympathomimetic activity and serotonergic effects. All of the β -blockers are excreted in human milk and therefore should be used with caution in nursing mothers.

Role in Stress, Learning, and Memory

An extensive preclinical and clinical literature suggests that the CNS noradrenergic system is an important mediator in human as well as in animal response to stress. As such, β -blockers are of interest because of their ability to dampen noradrenergic hyperarousal. There is also emerging evidence from both preclinical and human studies to suggest that β -blockers may attenuate learned responses or memory subsequent to anxiety-provoking stimuli. These latter data suggests that β -adrenergic activation may be required for certain memories associated with emotional arousal.

Animal studies suggest that the β -adrenergic system is involved in enhancing memory associated with

stress. In humans, administration of propranolol selectively impaired memory for an emotionally laden story but did not impair memory for a matched but emotionally neutral story. This latter observation suggests that emotional memory in humans may be affected by β -blockers. Of great interest in recent studies is the utility of these agents in attenuating memories associated with acute stress in emotionally traumatized individuals and whether the long-term sequelae (e.g., posttraumatic stress disorder [PTSD]) could therefore be minimized. Another question for study is whether patients taking β -blockers chronically might have subtle impairment in memory for emotionally significant events.

Therapeutic Use in Psychiatry

Anxiety Disorders

An extensive literature suggests that β -blockers may reduce autonomic symptoms in patients with panic disorder, generalized anxiety disorder, social phobia, and PTSD. They may not be as useful for subjective or cognitive anxiety (e.g., fear, worry, dread, or intrusive memories).

Social phobia Social phobia may include anxiety in specific social situations (e.g., giving a speech [performance anxiety]) or be generalized to a number of different social situations (generalized social phobia). Social phobia can result in significant occupational as well as social difficulties. β -blockers are helpful in specific performance anxiety when used on an as-needed basis but do not appear helpful for generalized social phobia. In performance anxiety, the β -blockers reduce autonomic symptoms associated with the performance situation (e.g., tremor, tachycardia, and diaphoresis). A test dose of the β -blocker should be administered at home before the individual gives the actual performance to ensure tolerability. Customary dose ranges of β -blockers used in social phobia may be 10 to 40 mg daily for propranolol or 50 to 100 mg daily for atenolol. For specific or discrete performance anxiety, the dose should be given 1 to 2 h before the activity.

Two studies have been consistent with the common clinical impression that β -blockers are not effective in generalized social phobia. A trial of atenolol versus placebo or phenelzine in social phobia showed better efficacy for phenelzine as compared with placebo or atenolol. Atenolol was comparable to placebo. Similarly, a placebo-controlled trial of behavior therapy versus atenolol or placebo in the social phobia was consistent with superior efficacy for behavior therapy, with atenolol being comparable to placebo.

Table 2 Comparative pharmacological properties of selected β -blockers

	<i>Relative β receptor selectivity</i>	<i>α_1 antagonist properties</i>	<i>Sympathomimetic properties</i>	<i>Direct vasodilator activity</i>	<i>Metabolism</i>	<i>Excretion</i>	<i>Plasma half-life (h)</i>	<i>Protein binding (%)</i>	<i>Lipid solubility</i>
Acebutolol	$\beta_1 > \beta_2$	—	—	—	Hepatic	Renal 30–40%, nonrenal 50–60%	3–4	26	Low
Atenolol	$\beta_1 > \beta_2$	—	+	—	Minimal	Renal 50%, nonrenal 50%	6–7	6–16	Low
Labetolol	$\beta_1 = \beta_2$	+++	—	—	Hepatic	Renal 30–40%, nonrenal 50–60%	6–8	50	Moderate
Metoprolol	$\beta_1 > \beta_2$	—	—	—	Hepatic	Renal $\leq 5\%$, nonrenal $\geq 95\%$	3–7	12	High
Nadolol	$\beta_1 = \beta_2$	—	—	—	Minimal	Renal 65–70%, nonrenal 30–35%	14–24	30	Low
Pindolol	$\beta_1 > \beta_2$	—	—	—	Hepatic	Renal 35–40%, nonrenal 60–65%	3–4	40	Moderate
Propanolol	$\beta_1 = \beta_2$	—	—	—	Hepatic	Renal 90–95%, nonrenal 5–10%	4	90	High

Posttraumatic stress disorder PTSD is a chronic anxiety disorder that may occur after a severe life-threatening stress (e.g., combat or rape). Symptoms include reexperiencing the trauma via intrusive memories, flashbacks (sudden acting or feeling as if the trauma was recurring), nightmares, avoidance of reminders of the trauma, and symptoms of increased autonomic arousal (e.g., exaggerated startle, hypervigilance, and insomnia). PTSD can result in severe occupational and social impairment. A number of clinical studies have suggested norepinephrine hyperactivity in PTSD. For example, exaggerated norepinephrine or norepinephrine metabolite responses may occur secondary to pharmacological, psychological, or physical stressors.

Despite the evidence for noradrenergic hyperarousal in PTSD, there has been surprisingly little study to date of β -blockers in the disorder. Two of three clinical reports were promising in suggesting potential efficacy. One report suggested that propranolol may reduce intrusion and arousal symptoms in Vietnam veterans with PTSD. In contrast, propranolol was not efficacious in an open trial in Cambodian refugees with PTSD. Another report in children with PTSD who were victims of physical or sexual abuse suggested that propranolol at a dose of 2.5 mg/kg/day could reduce both reexperiencing and hyperarousal symptoms. Controlled studies are needed in both chronic as well as acutely traumatized populations. The potential role of β -blockers in modifying memory related to acute emotional trauma, as discussed in the previous section, is especially intriguing. Of note, several pilot studies have suggested that propranolol may mitigate or perhaps prevent the development of PTSD if administered shortly following acute trauma exposure. More study is needed of these encouraging preliminary reports before this could be considered a standard clinical practice.

Panic disorder and generalized anxiety disorder The β -blockers can reduce autonomic symptoms associated with generalized anxiety disorder (GAD) or panic attacks but do not appear effective for cognitive or psychological aspects of anxiety (e.g., worry, fear, or dread). This is likely because β_1 but not β_2 receptors are located in the brain and β_2 blockade does not appear to affect psychological anxiety. The psychological and cognitive symptoms respond better to benzodiazepines, buspirone, or antidepressants. There are also negative reports on the efficacy of β -blockers in panic and anxiety.

Doses of propranolol used in GAD range from 40 to 320 mg/day in two to four divided doses. A preliminary clinical report suggested that a long-acting

β -blocker, betaxolol, in doses of 5 to 20 mg daily were effective in GAD patients. Betaxolol, a relatively β_1 -selective agent, penetrates the CNS to a much greater extent than other long half-life agents such as nadolol or atenolol. Recent reports have suggested that pindolol may augment the therapeutic response to antidepressants in treatment-refractory panic disorder. For example, a placebo-controlled trial in panic disorder patients refractory to two antidepressant trials and to a trial of fluoxetine responded to the addition of pindolol given at a dose of 2.5 mg three times daily.

Obsessive-compulsive disorder Pindolol may augment the therapeutic effect of antidepressants in patients with treatment-resistant obsessive-compulsive disorder (OCD). For example, a double-blind, placebo-controlled trial of adding pindolol to the selective serotonin reuptake inhibitor paroxetine showed significant improvement in OCD ratings for patients receiving pindolol 2.5 mg orally three times daily (see also antidepressant potentiation section below).

Aggression and Violence

A number of case reports, clinical series, and controlled studies suggest that β -blockers can reduce aggression associated with chronic psychotic disorders, dementia, and head trauma. Behaviors such as impulsivity, largely attributable to orbitofrontal cortical dysfunction, are most likely to respond to β -blockers. This may reflect the extensive frontocortical noradrenergic afferent projections. Relatively lipophilic agents (e.g., propranolol, metoprolol, or nadolol) have been utilized because of their CNS effects. Doses of propranolol or equivalent β -blocker used in treating violent behaviors range from 40 to 520 mg/day in divided doses. Many of the clinical reports involve violent patients refractory to other medications (e.g., antipsychotics or mood stabilizers). For example, nadolol at a dose of 120 mg/day may be more effective than placebo as an adjunct to antipsychotics in aggressive schizophrenic patients. Elderly patients with dementia and associated aggression or agitation may improve with lower doses of propranolol (10–80 mg/day) than are used in younger aggressive patients. An important confound in interpreting the controlled trials of β -blockers in violence is the episodic nature of aggressive behavior. Also, most studies of β -blockers in violence have involved drug combination rather than monotherapy. For patients who have not responded to other psychotropics, a β -blocker trial may be indicated with close monitoring for adverse effects (e.g., hypotension),

especially in light of the high doses needed for therapeutic effect.

Alcohol Withdrawal

Alcohol withdrawal in part involves autonomic instability mediated by significant CNS and peripheral noradrenergic activation. The β -blockers therefore can help reduce certain autonomic symptoms. Unlike the benzodiazepines, β -blockers are not cross-reactive with alcohol, so they are not likely to be effective as a monotherapy for alcohol detoxification. They may be effective as an adjunctive agent to benzodiazepines or other first-line withdrawal agents. A randomized, double-blind trial of atenolol as an adjunct to benzodiazepines found that administration of atenolol based on heart rate parameters improved vital signs and tremor more rapidly than placebo and resulted in less benzodiazepine use and decreased length of hospitalization. Atenolol was administered based on the following heart rate parameters: less than 50 beats/min, no drug; 50 to 79, 50 mg once daily; greater than 80, 100 mg daily. Further study is needed of this potential adjunctive use of β -blockers.

Treatment of Medication Side Effects

Antipsychotic-induced akathisia Akathisia is motor restlessness associated with antipsychotics, in particular the first-generation or traditional agents such as chlorpromazine (Thorazine), fluphenazine (Prolixin), or haloperidol (Haldol). Categorized as an extrapyramidal system side effect (EPS), akathisia can result in substantial noncompliance with antipsychotic medication. Unlike other EPSs such as parkinsonism, akathisia may respond better to β -blockers or benzodiazepines than antiparkinsonism medication such as benztropine (Cogentin). The EPSs in general are thought secondary to blockade of dopamine (D_2) receptors in the nigrostriatal pathway. The D_2 receptor blockade is common to all antipsychotic agents, although the new generation or atypical antipsychotics may have reduced akathisia incidence, perhaps due to their higher ratio of serotonin-to-dopamine receptor blockade. Akathisia differs somewhat from other EPSs in that the mechanism may involve loss of tonic dopamine inhibition of the locus ceruleus due to D_2 receptor blockade. The resulting increased noradrenergic release from the locus ceruleus results in the subjective anxiety and motor restlessness characteristic of akathisia. This mechanism may explain the enhanced efficacy of β -blockers or benzodiazepines, both of which decrease noradrenergic function (β -blockers via postsynaptic receptor blockade and benzodiazepines via decreased locus ceruleus firing rate)

as opposed to antiparkinson agents such as Cogentin or Symmetrel.

For treatment of akathisia, propranolol is given in doses of 10 to 30 mg three times daily (or similar doses of other β -blockers may be considered). β -blockers may be given alone or in combination with benzodiazepines or antiparkinsonism medications. The new generation or atypical antipsychotics have less potential to induce akathisia.

Lithium-induced tremor The low-amplitude action tremor characteristic of lithium may occur at therapeutic or even subtherapeutic doses. In fact, new-onset or worsening of the tremor (higher amplitude, increased frequency) may be a warning sign for lithium toxicity. A stable tremor at therapeutic doses can cause motor dysfunction (e.g., difficulty writing) and therefore may require treatment to enhance lithium tolerability as well as compliance. Lithium should be given at the lowest effective maintenance dose.

Analogous to their use in benign essential or familial tremor, β -blockers (e.g., propranolol in the daily dose range of 20 to 100 mg, in two or three divided doses) can be effective. In addition to the patient being maintained on the lowest effective lithium dose, administration of a single nighttime dose of lithium (which may also reduce lithium-associated renal morphological changes) and reduction of caffeine consumption can help minimize the tremor.

Antidepressant Potentiation

Several reports have suggested that the β -blocker pindolol, which increases forebrain neuronal release of serotonin in animals, may potentiate and possibly reduce the latency of response to selective serotonin reuptake inhibitor antidepressants. Pindolol is a lipophilic and relatively nonselective β_1 - and β_2 -blocker that also has intrinsic sympathomimetic properties. The serotonin augmentation is likely due to its selective serotonin 5-HT_{1A} receptor effects. Whether β -blockers may cause or exacerbate depression remains controversial.

Therapeutic Use in Medical Disorders

The β -blockers, originally developed for the treatment of hypertension and angina, had expanded roles in clinical medicine as antiarrhythmics, prevention of cardiovascular mortality after myocardial infarction, treatment of complications of cardiomyopathies, and various other noncardiovascular as well as psychiatric uses (Table 3).

Table 3 Doses of β -blockers used in psychiatry

<i>Generic name</i>	<i>Trade name</i>	<i>Usual starting dose (mg)</i>	<i>Usual maximum dosage (mg)</i>
Acebutolol	Sectral	400 qd	600 bid
Atenolol	Tenormin	50 bid	75–150 bid
Labetalol	Normodyne, Trandate	100 bid	400–800 bid
Metoprolol	Lopressor	50 bid	75–150 bid
Nadolol	Corgard	40 qd	80–240 qd
Pindolol	Visken	5 bid	30 bid
Propranolol	Inderal	10–20 bid or tid	80–140 bid or tid

Hypertension

The β -blockers are effective as a monotherapy or in combination with other antihypertensives. They may reduce morbidity (e.g., left ventricular hypertrophy) and mortality associated with hypertension.

Angina Pectoris

The β -blockers are useful in monotherapy or in combination with certain antianginals including nitrates and some calcium channel blockers. The reduction in symptoms, especially in exertional angina, is attributable to decreased cardiac workload due to slower heart rate allowing longer diastolic filling time and improved coronary artery perfusion.

Cardiac Arrhythmias

The β -blockers are efficacious in controlling supraventricular tachyarrhythmias and reduction of ventricular arrhythmias. The mechanism includes decreased heart rate by reduced adrenergic stimulation of sinoatrial or atrioventricular node pacemakers and by cardiac membrane stabilization.

Congestive Heart Failure

The β -blockers may reduce symptoms and improve functional status in congestive heart failure (CHF) patients with careful administration. They may be indicated for patients who continue to be symptomatic with standard therapies. However, β -blockers may also precipitate CHF. Their use in dilated cardiomyopathy is based on the theory that they may attenuate the adverse effects on cardiac function of excessive adrenergic stimulation secondary to CHF.

Other Medical Uses

Additional medical diseases in which the reduction of excessive adrenergic stimulation with β -blockers may be useful include acute dissecting aortic aneurysm, dilation of the aorta associated with Marfan's syndrome (propranolol), pheochromocytoma, glaucoma, hyperthyroidism, tremors, migraine prophylaxis, and portal hypertension. Different β -blockers are used

in these various indications depending on their relative pharmacological effects (e.g., relative β_1 versus β_2 antagonism or degree of vasodilating properties).

Safety and Tolerability

Side Effects

The β -blockers are generally well-tolerated. Side effects are listed in [Table 4](#). The more frequent side effects are a consequence of the pharmacodynamic effects (i.e., β_1 and, if applicable, β_2 receptor blockade). These more common side effects include hypotension, dizziness, bradycardia, and sexual dysfunction. The β_1 -selective agents have decreased risk of bronchospasm (asthma exacerbation) than those with β_2 effects. Cardiovascular side effects can be minimized by careful dose titration, monitoring pulse and blood pressure, and patient education. Patients with coronary heart disease can have exacerbation of angina pectoris with discontinuation (especially abrupt withdrawal) of β -blockers. There may also be rebound hypertension in patients with a history of hypertension and worsening of CHF. Diabetic patients may have increased glucose dysregulation (especially hypoglycemia) and may have decreased autonomic symptoms of hypoglycemia (i.e., less subjective awareness of warning symptoms) due to β -blockade. The β_1 agents overall are safer to use in patients with bronchospasms, peripheral vascular disease, Raynaud's disease, and diabetes. Importantly, topically administered β -blockers (e.g., for glaucoma patients) can have systemic effects. Insomnia and nightmares are associated with sleep architecture changes from most β -blockers, although some (e.g., betaxolol) may have fewer adverse sleep effects. Although fatigue or lassitude may occur in some individuals with chronic β -blocker administration, clinical depression is probably rare as a causal association. CNS effects may be minimized by using peripheral-acting (less lipophilic) agents.

Use in Pregnancy

The β -blockers should be used with caution in pregnant or nursing women and only if the benefit clearly

Table 4 Side effects of β -blockers

Cardiovascular
Hypotension
Bradycardia
Lightheadedness or dizziness
Congestive heart failure (with predisposing cardiac disease)
Atrioventricular block
Exacerbation of intermittent claudication
Exacerbation of angina pectoris, primarily with abrupt or rapid discontinuation
Raynaud's phenomenon exacerbation
Respiratory
Dyspnea
Asthma (risk decreased with β_1 -selective agents)
Metabolic
Hypoglycemia in diabetic patients
Gastrointestinal
Nausea
Diarrhea
Abdominal pain
Genitourinary
Sexual dysfunction (erectile dysfunction)
Urinary hesitancy
Neuropsychiatric
Fatigue
Lassitude
Dysphoria
Insomnia
Vivid nightmares
Depression (possible)
Psychosis (rare)

outweighs potential risks. All of the β -blockers are pregnancy category C, which means there are insufficient data in humans to demonstrate safety in pregnancy or to show definite relationship with defects. Embryotoxicity and neonatal toxicity have been observed in animals given β -blockers. There are no adequate studies in pregnant women. There are reports of intrauterine growth retardation and neonatal toxicity (bradycardia, hypoglycemia, and respiratory depression) in neonates of mothers treated with β -blockers. The safety and efficacy of β -blockers in children is not well established.

Drug Interactions

Of primary concern are pharmacodynamic interactions with other antihypertensive agents (e.g., clonidine or reserpine [postural hypotension]), atrioventricular block with other antiarrhythmics (e.g., digoxin or verapamil), or agents that have negative cardiac inotropic effects (e.g., diltiazem). There are several pharmacokinetic interactions: alcohol, cholestyramine, and aluminum hydroxide gels reduce absorption of β -blockers. Phenytoin, rifampin, and phenobarbital and smoking increase hepatic clearance of extensively metabolized β -blockers. Antipyrine, lidocaine, and theophylline have reduced clearance by β -blockers.

Cimetidine decreases hepatic metabolism of propranolol and metoprolol.

Summary

β -blockers have been among the most useful and effective class of compounds used for medical, neurological, and psychiatric conditions. Since their introduction to clinical medicine in the late 1960s, they have been effective in the treatment of cardiovascular, endocrine, neurological, and various psychiatric illnesses, especially anxiety disorders, drug-induced toxicities (e.g., akathisia and tremor), and, recently, potentiation of antidepressant effect in some treatment-resistant disorders.

The possible use of β -blockers to attenuate or reduce the impact of situational trauma or PTSD needs further exploration. Doctors are learning about the importance of the use of selective serotonin reuptake inhibitors to possibly improve the course of affective illness or advance episode frequency. In an analogous fashion, the potential role of β -blockers for prevention and prophylaxis of stress-related disorders needs further study.

See Also the Following Articles

Aggression; Anxiety; Catecholamines; Heart Disease/Attack; Hypertension; Panic Disorder and Agoraphobia.

Further Reading

- Allan, E. R., Alpert, M., Sison, C. E., Citrome, L., Laury, G. and Berman, I. (1996). Adjunctive nadolol in the treatment of acutely aggressive schizophrenic patients. *Journal of Clinical Psychiatry* 57, 455–459.
- Arana, G. W. and Santos, A. B. (1995). β -adrenergic receptor antagonists. In: Kaplan, H. I. & Sadock, B. J. (eds.) *Comprehensive textbook of psychiatry* (6th edn., pp. 1915–1919). Baltimore, MD: Williams & Wilkins.
- Bright, R. A. and Everitt, D. E. (1992). Beta-blockers and depression: evidence against an association. *Journal of the American Medical Association* 267, 1783–1787.
- Cahill, L., Prins, B., Weber, M. and McGaugh, J. L. (1994). Beta-adrenergic activation and memory in anxiety. *Nature* 371, 702–704.
- Cooper, S. J., Kelly, C. B., McGilloway, S. and Gilliland, A. (1990). Beta 2-adrenoreceptor antagonism in anxiety. *European Neuropsychopharmacology* 1, 75–77.
- Dannon, P. N., Sasson, Y., Hirschman, S., et al. (2000). Pindolol augmentation in treatment-resistant obsessive compulsive disorder: a double-blind, placebo controlled trial. *European Neuropsychopharmacology* 20, 165–169.
- Davidson, J. R. T. (1998). Pharmacotherapy of social anxiety disorder. *Journal of Clinical Psychiatry* 59(Supplement 17), 47–51.
- Duman, R. S. and Nestler, E. J. (1995). Signal transduction pathways for catecholamine receptors. In: Bloom, F. E. &

- Kupfer, D. J. (eds.) *Psychopharmacology: the fourth generation of progress*, pp. 303–320. New York: Raven Press.
- Famularo, R., Kinscherff, R. and Fenton, T. (1988). Propranolol treatment for childhood post-traumatic stress disorder, acute type: a pilot study. *American Journal of Diseases of Children* **142**, 1244–1247.
- Hamner, M. B., Diamond, B. I. and Hitri, A. (1994). Plasma norepinephrine and MHPG responses to exercise stress in PTSD. In: Murburg, M. M. (ed.) *Catecholamine function in post-traumatic stress disorder: emerging concepts*, pp. 321–332. Washington, D.C.: American Psychiatric Press.
- Hirschmann, S., Dannon, P. N., Iancu, I., et al. (2000). Pindolol augmentation in patients with treatment-resistant panic disorder: a double-blind, placebo-controlled trial. *Journal of Clinical Psychopharmacology* **20**, 556–559.
- Kinzie, J. D. (1989). Therapeutic approaches to traumatized Cambodian refugees. *Journal of Trauma and Stress* **2**, 75–91.
- Kolb, L. C., Burris, B. C. and Griffiths, S. (1984). Propranolol and clonidine in the treatment of post traumatic stress disorders of war. In: van der Kolk, B. A. (ed.) *Post traumatic stress disorder: psychological and biological sequelae*, pp. 97–105. Washington, D.C.: American Psychiatric Press.
- Kraus, M. L., Gottlieb, L. D., Horwitz, R. I. and Anscher, M. (1985). Randomized clinical trial of atenolol in patients with alcohol withdrawal. *New England Journal of Medicine* **313**, 905.
- Liebowitz, M. R., Schneier, F., Campeas, R., et al. (1992). Phenelzine vs Atenolol in social phobia: a placebo-controlled comparison. *Archives of General Psychiatry* **49**, 290–300.
- McFall, M. E., Murburg, M. M., Ko, G. N. and Ueith, R. C. (1990). Autonomic response to stress in Vietnam combat veterans with post-traumatic stress disorder. *Biological Psychiatry* **27**, 1165–1175.
- Pitman, R. K., Sanders, K. M., Zusman, R. M., et al. (2002). Pilot study of secondary prevention of post-traumatic stress disorder with propranolol. *Biological Psychiatry* **51**, 189–192.
- Ratey, S. J., Sorgi, P., O'Driscoll, G. A., et al. (1992). Nadolol to treat aggression and psychiatric symptomatology in chronic psychiatric inpatients: a double-blind, placebo-controlled study. *Journal of Clinical Psychiatry* **53**, 41–46.
- Southwick, S. M., Krystal, J. H., Morgan, A. C., et al. (1993). Abnormal noradrenergic function in posttraumatic stress disorder. *Archives of General Psychiatry* **50**, 266–274.
- Tome, M. B., Isaac, M. T., Harte, R. and Holland, C. (1997). Paroxetine and pindolol: a randomized trial of serotonergic autoreceptor blockade in the reduction of antidepressant latency. *International Clinical Psychopharmacology* **12**, 81–89.
- Turner, S. M., Beidel, D. C. and Jacob, R. G. (1994). Social phobia: a comparison of behavior therapy and atenolol. *Journal of Consulting and Clinical Psychology* **62**, 350–358.
- Vaiva, G., Ducrocq, F. and Jezequel, K. (2003). Immediate treatment with propranolol decreases posttraumatic stress disorder two months after trauma. *Biological Psychiatry* **54**, 947–949.

Beta-Endorphin

M Lee and S L Wardlaw

Columbia University College of Physicians and Surgeons, New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J J Puder and S L Wardlaw, volume 1, pp 321–324, © 2000, Elsevier Inc.

Structure, Synthesis, and Localization

β-Endorphin and Neuroendocrine Responses

β-Endorphin and Stress-Induced Analgesia

β-Endorphin and the Immune System

β-Endorphin and Cardiovascular Responses

β-Endorphin and Stress-Induced Feeding and Other Behaviors

Glossary

<i>Adrenocorticotropin (ACTH)</i>	The pituitary hormone that regulates adrenal corticosteroid synthesis and release. ACTH, like β-endorphin, is derived from proopiomelanocortin.
<i>Endogenous opioid peptides</i>	Peptides with morphine-like actions that are synthesized in the brain and in peripheral tissues.
<i>Endorphins</i>	A general term that refers to all of the endogenous opioid peptides. β-endorphin refers to a specific peptide.
<i>Naloxone</i>	A compound that binds to opiate receptors and blocks the effects of morphine and the endogenous opioid peptides.
<i>Proopiomelanocortin</i>	The larger precursor protein from which β-endorphin is derived.

β -endorphin is an endogenous opioid peptide that is synthesized primarily in the pituitary gland and in the brain. Stress stimulates the release of β -endorphin in both the pituitary and the brain. β -endorphin and the other endogenous opioid peptides have been implicated in analgesia and in neuroendocrine regulation during times of stress, including the suppression of reproductive function. These peptides also play a role in autonomic regulation, modulation of cardiovascular responses and cerebral blood flow, and the regulation of behavior and immune function.

Structure, Synthesis, and Localization

β -Endorphin is a 31-amino-acid peptide that is derived from the larger molecular weight precursor protein proopiomelanocortin (POMC). β -Endorphin is one of the three major classes of endogenous opioid peptides. The other two classes are composed of the enkephalins and the dynorphins, which are derived from separate precursor proteins. POMC is synthesized independently in the anterior and intermediate lobes of the pituitary gland, in the brain, and in several peripheral tissues, including the gastrointestinal tract, reproductive tract, placenta, and cells of the immune system. The posttranslational processing of POMC is tissue specific and results in the production of peptides with very different biological activities. In the anterior pituitary, POMC is processed predominantly to adrenocorticotropin (ACTH) and β -lipotropin (β -LPH), but some β -endorphin is also produced. ACTH is secreted from the pituitary into the peripheral circulation and is critical for the maintenance of adrenocortical function and cortisol secretion. β -endorphin is also secreted in parallel with ACTH. In the brain, ACTH and β -LPH are further processed to yield α -melanocyte-stimulating hormone (α -MSH) and β -endorphin. The regulation of POMC processing is potentially important, as α -MSH has been shown to antagonize some of the effects of β -endorphin. The regulation of POMC gene expression and peptide release is also tissue specific, but under conditions of stress, β -endorphin is released from both the pituitary and the brain. The hypothalamic releasing factor corticotropin-releasing hormone (CRH), which plays a key role in coordinating the organism's response to stress, has been shown to stimulate pituitary and brain β -endorphin. β -endorphin exerts its effects by binding to specific opioid receptors, particularly the mu and delta receptors, which have been recently cloned and localized throughout the brain.

β -Endorphin is released from the anterior pituitary into the peripheral circulation in parallel with ACTH

in response to a variety of stresses. Numerous studies have documented parallel increases in peripheral blood levels of ACTH and β -endorphin in response to exercise. This appears to be related to the intensity and duration of the exercise. A number of correlations between β -endorphin levels in blood and central nervous system or cardiovascular responses during stress have been reported. One must be cautious, however, in attributing a causal relationship for circulating β -endorphin, because in many instances, brain β -endorphin may also be stimulated at the same time. In addition, peripherally released β -endorphin does not readily cross the blood-brain barrier. In the brain, β -endorphin is synthesized in only two regions, the arcuate nucleus of the hypothalamus and the nucleus of the solitary tract in the brain stem, but β -endorphin fibers project widely to many regions throughout the brain.

β -Endorphin and Neuroendocrine Responses

Endogenous opioid peptides have multiple effects on pituitary function. Effects on gonadotropin (luteinizing hormone [LH] and follicle-stimulating hormone [FSH]), prolactin, ACTH, vasopressin, and oxytocin release have all been demonstrated. There is evidence for a specific role for β -endorphin during stress with respect to the regulation of LH and the hypothalamic-pituitary-gonadal (HPG) axis, of ACTH and the hypothalamic-pituitary-adrenocortical (HPA) axis, and of prolactin release. Stress is known to impair reproductive function in many species. Stress has been shown to suppress the release of LH from the pituitary with subsequent suppression of ovarian and testicular function. Such suppression has been seen with many types of stress ranging from exercise and weight loss to inflammatory and psychological stress. CRH, which is released from the hypothalamus during stress, has been shown to cause a similar suppression of the HPG axis. There is convincing evidence that the suppressive effects of stress and CRH on the HPG axis are mediated by β -endorphin. Thus, both CRH and β -endorphin are important in suppressing reproductive function during times of stress. Exogenous and endogenous opioids also have well-documented effects on the HPA axis. There are some differences, however, depending on the dose and duration of treatment and the species studied. In the human, morphine and β -endorphin have been shown to suppress ACTH and cortisol release in normal subjects and to blunt the HPA response to surgical stress. β -Endorphin has been shown to suppress CRH release from the hypothalamus *in vitro*

and to suppress CRH release into hypophysial portal blood *in vivo*. Conversely, opioid antagonism has been reported to stimulate the HPA axis and to enhance the HPA response to stress. Thus, β -endorphin may be important in limiting the HPA response under conditions of stress.

β -Endorphin and Stress-Induced Analgesia

It has long been known that pain perception is affected by stress. Electrical stimulation of specific brain regions has been shown to induce analgesia that was blocked by treatment with naloxone and other opioid antagonists, implying that this analgesia was mediated by the release of endogenous opioids. Certain forms of stress-induced analgesia are also blocked by naloxone. It is now clear that stress-induced analgesia is mediated by both opioid and nonopioid pathways. There is considerable pharmacological and neuroanatomic evidence to implicate β -endorphin as a mediator of stress-induced analgesia. Recently, a genetic approach has been used to generate mice that are unable to synthesize β -endorphin. These transgenic mice do not exhibit the naloxone-reversible analgesia normally seen after a mild swim stress, supporting a role for β -endorphin in this process. β -Endorphin is also synthesized by immune cells and is known to play a role in mediating analgesia at sites of inflammation. Studies in rats have demonstrated that β -endorphin is released by immune cells at sites of inflammation and can induce analgesia by interaction with receptors on peripheral terminals of sensory nerves. These studies provide convincing evidence that β -endorphin plays an important role within the brain in mediating stress-induced analgesia and can also act locally to affect pain perception at sites of inflammation.

β -Endorphin and the Immune System

Both stress and opioids have suppressive effects on immune function. Naloxone has also been shown to block stress-induced immune suppression, indicating a role for endogenous opioids in this process. β -Endorphin-deficient mice possess enhanced immune responses as measured by increased splenocyte proliferation and inflammatory cytokine production, consistent with an inhibitory effect of endogenous β -endorphin on immune function. β -Endorphin is produced by cells of the immune system and may have local paracrine effects on immune function. Circulating, pituitary-derived β -endorphin, which increases with stress and with inflammation, may also affect immune function. In addition, β -endorphin may affect

immune function indirectly by modulating neuroendocrine and autonomic function.

β -Endorphin and Cardiovascular Responses

The endogenous opioid peptides are localized to regions within the brain that are important in controlling the autonomic nervous system and cardiovascular responses. β -Endorphin is synthesized in the nucleus of the solitary tract in the medulla, an area important in cardiovascular regulation. The effects of opioids on cardiovascular function are complex and depend on the type of opioid injected, the site of injection, and whether the animal is anesthetized or stressed. Although discrepancies exist, it appears that in general, opioids diminish stress-induced cardiovascular effects. Opioid mechanisms appear to facilitate parasympathetic and inhibit sympathetic nervous system reactivity during stress. High-dose infusion of β -endorphin in human subjects reduces blood pressure and circulating norepinephrine levels. Several studies have demonstrated activation of the endogenous opioid system in hemorrhagic shock, sepsis, and trauma and shown that this system plays a complex role both centrally and peripherally in modulating autonomic and cardiovascular responses. Endogenous opioid peptides are also important in regulating cerebral blood flow, especially during periods of hypoxia. β -Endorphin, in particular, has been shown to regulate hypothalamic blood flow.

β -Endorphin and Stress-Induced Feeding and Other Behaviors

Pharmacological studies have shown that opioids generally stimulate food intake. Rats subjected to mild stress exhibit stress-induced eating that can be attenuated by treatment with naloxone, thus implicating a role of endogenous opioids in this process. β -Endorphin administration in rodents has also been shown to stimulate food intake. However, more recently, studies with transgenic mice deficient in β -endorphin indicate that the role of endogenous opioids in modulating feeding and energy homeostasis is quite complex and that β -endorphin may be important in primarily increasing the motivational and reward components of feeding. Finally, there is evidence that endogenous opioids may affect other behavioral and psychological responses to stress, but a specific role for β -endorphin remains to be defined.

See Also the Following Articles

Adrenocorticotrophic Hormone (ACTH); Opioids; Peptides.

Further Reading

- Akil, H., Watson, S. J., Young, E., et al. (1984). Endogenous opioids: biology and function. *Annual Reviews of Neuroscience* 7, 223–255.
- Benyó, Z. and Wahl, M. (1996). Opiate receptor-mediated mechanism in the regulation of cerebral blood flow. *Cerebrovascular Brain Metabolism Reviews* 8, 326–357.
- Bodnar, R. J. (2004). Endogenous opioids and feeding behavior: a 30-year historical perspective. *Peptides* 25, 697–725.
- Castro, M. G. and Morrison, E. (1997). Post-translational processing of proopiomelanocortin in the pituitary and in the brain. *Critical Reviews in Neurobiology* 11, 35–57.
- Cozzolino, D., Sasso, F. C., Cataldo, D., et al. (2005). Acute pressor and hormonal effects of β -endorphin at high doses in healthy and hypertensive subjects: role of opioid receptor agonism. *Journal of Clinical Endocrinology and Metabolism* 90, 5167–5174.
- Ferin, M. (1999). Stress and the reproductive cycle. *Journal of Clinical Endocrinology and Metabolism* 84, 1768–1774.
- Ferin, M., Van Vugt, D. and Wardlaw, S. L. (1984). The hypothalamic control of the menstrual cycle and the role of endogenous opioid peptides. *Recent Progress in Hormone Research* 40, 441–485.
- Goldfarb, A. H. and Jamurtas, A. Z. (1997). β -endorphin response to exercise. *Sports Medicine* 24, 8–16.
- Low, M. J., Hayward, M., Appleyard, S. M. and Rubinstein, M. (2003). State-dependent modulation of feeding behavior by proopiomelanocortin-derived β -endorphin. *Annals of the New York Academy of Sciences* 994, 192–201.
- Mansour, A., Fox, C. A., Akil, H. and Watson, S. J. (1995). Opioid-receptor mRNA expression in the rat CNS: anatomical and functional implications. *Trends in Neuroscience* 18, 22–29.
- McCubbin, J. A. (1993). Stress and endogenous opioids: behavioral and circulatory interactions. *Biological Psychology* 35, 91–122.
- Molina, P. E. (2002). Stress-specific opioid modulation of haemodynamic counter-regulation. *Clinical and Experimental Pharmacology and Physiology* 29, 248–253.
- Morely, J. E. and Levine, A. S. (1980). Stress-induced eating is mediated through endogenous opiates. *Science* 209, 1259–1261.
- Pechnik, R. N. (1993). Effects of opioids on the hypothalamo-pituitary-adrenal axis. *Annual Reviews in Pharmacology and Toxicology* 32, 353–382.
- Refojo, D., Kovalovsky, D., Young, J. I., et al. (2002). Increased splenocyte proliferative response and cytokine production in β -endorphin deficient mice. *Journal of Neuroimmunology* 131, 126–134.
- Rubinstein, M., Mogil, J. S., Japón, M., et al. (1996). Absence of opioid stress-induced analgesia in mice lacking β -endorphin by site-directed mutagenesis. *Proceedings of the National Academy of Sciences USA* 93, 3995–4000.
- Stein, C. R., Hassan, A. H. S., Przewlocki, R., et al. (1990). Opioids from immunocytes interact with receptors on sensory nerves to inhibit nociception in inflammation. *Proceedings of the National Academy of Sciences USA* 87, 5935–5939.

Blood Pressure

A Sherwood and R A Carels

Duke University Medical Center

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A Sherwood and R A Carels, volume 1, pp 331–338, © 2000, Elsevier Inc.

Blood Pressure Measurement

Hypertension

Stress and Blood Pressure

Ambulatory Blood Pressure

Glossary

Ambulatory blood pressure Measurements recorded during the course of normal daily activities, using a wearable blood pressure monitor.

Cardiac output

The volume flow rate (l/min) of blood ejected from the left ventricle of the heart into the systemic circulation.

Diastolic blood pressure

The lowest pressure (mmHg) in the arterial system, occurring during cardiac diastole.

Hemo-dynamics

Assessment of the circulation in terms of blood pressure, cardiac output, and systemic vascular resistance.

Hypertension

Resting blood pressure that is higher than 140/90 mmHg.

Reactivity

A psychophysiological trait expressed as cardiovascular response to one or more psychological stressors.

Sympathetic nervous system

A division of the autonomic nervous system that is activated in association with psychological stress and is involved in cardiovascular regulation.

Systemic vascular resistance

The total peripheral resistance of the systemic vasculature (dynes s cm⁻⁵).

Systolic blood pressure The highest pressure (mmHg) in the arterial system, occurring during cardiac systole.

Blood Pressure Measurement

When the left ventricle of the heart contracts, a bolus of blood is ejected into the systemic circulation, creating a pulse pressure waveform that propagates through the arterial tree. The peak pressure reached in the arterial system is referred to as systolic pressure, since it occurs in association with cardiac systole, or the heart's contraction phase. Once systolic pressure is reached, pulsatile pressure falls as the heart relaxes during diastole, with the lowest pressure, referred to as diastolic pressure, occurring just prior to the onset of the next systolic phase. Pulse pressure is defined as the difference between systolic and diastolic pressure. Mean arterial pressure refers to the average pressure in the arterial system. All blood pressure measurements are recorded in millimeters of mercury (mmHg).

Systolic and diastolic pressure are the most commonly measured aspects of arterial blood pressure. The emphasis on these parameters in medicine is partly a reflection of their relative ease of measurement using noninvasive techniques. Although arterial blood pressure varies according to where it is measured in the arterial tree, it is most commonly assessed in the brachial artery, which runs through the upper region of the arm. Two commonly employed measurement techniques, auscultatory and oscillometric, employ an occlusion cuff that contains an inflatable bladder applying pressure through the arm to first interrupt and then progressively restore blood flow through the brachial artery.

With the auscultatory method, a stethoscope or microphone is positioned over the brachial artery, distal to the occlusion cuff. When the pressure in the occlusion cuff is above systolic pressure, the brachial artery is compressed and no sounds are detected distal to the cuff. When the slowly deflating occlusion cuff drops below the systolic pressure, the compressed artery opens momentarily and the brief episode of turbulent blood flow distal to the cuff is detected as a faint tapping sound; the first phase I Korotkoff sound is used to indicate systolic blood pressure. As the cuff pressure continues to fall, the Korotkoff sounds increase in intensity and quality, followed by a muffling of the sound (phase IV) and ultimately their disappearance (phase V). Diastolic pressure is usually defined as phase V in adults, while in children phase IV may be a closer approximation. With the auscultatory technique, mean arterial pressure cannot be

measured directly, but may be estimated as diastolic pressure plus one-third of the pulse pressure.

The oscillometric method also employs the occlusion cuff in a manner similar to that for auscultatory measurement. However, with the oscillometric technique, cuff pressure oscillations coincident with pulsatile pressure variations in the artery form the basis of blood pressure measurement. As the occlusion cuff is deflated, systolic pressure is indicated by a sharp rise in cuff pressure oscillations. The lowest pressure at which the cuff oscillations are at their greatest amplitude identifies mean arterial pressure. Diastolic pressure is recorded as the point at which the cuff pressure oscillations decrease no further in amplitude.

The auscultatory and oscillometric methods are by far the most widely utilized methods for measuring blood pressure. One of their limitations is that because they use a slowly deflating occlusion cuff, blood pressure measurements are limited in frequency to approximately once per minute. Other techniques include direct arterial pressure recording via a pressure transducer on the tip of an intra-arterial catheter, a method often employed to monitor vital signs during major surgery. One advantage of this invasive method is that it provides a continuous measurement of blood pressure on a beat-by-beat basis. A number of research-orientated techniques, including cuff tracking, vascular unloading, arterial tonometry, and pulse transit time, provide methods to assess blood pressure beat by beat, yet noninvasively, each offering specific strengths and limitations. All of the methods described above have been employed in studies examining the effects of stress on blood pressure.

Hypertension

Hypertension is defined as systolic blood pressure of 140 mmHg or greater, and/or diastolic blood pressure of 90 mmHg or greater. These criteria, based on seated clinic blood pressure measurements, are recognized by the World Health Organization/International Society of Hypertension and the United States Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC). Pressures in excess of 140/90 mmHg are further categorized in terms of severity of hypertension, currently defined as stages 1 and 2 (JNC VII). While blood pressures less than 140/90 mmHg were once considered normal, the JNC VII report now classifies blood pressures in the range 120–139/80–89 mmHg as prehypertension. This new category reflects a growing recognition that the clinical criteria for hypertension are somewhat arbitrary, and that cardiovascular morbidity associated with high blood pressure is more of a

progressive phenomenon than one associated with a specific pressure threshold that engenders a sudden rise in health risk. The term prehypertension also signifies that individuals with pressures in this range are at greater risk of developing hypertension than individuals with lower pressures. Only pressures below 120/80 mmHg are considered optimal.

Prevalence of hypertension is related to age, race, and gender. Hypertension incidence is higher in Blacks than in Whites, particularly in the United States. In young to middle-aged adults, prevalence is much greater in males than females, while incidence increases quite dramatically in women after menopause. Though hypertension is defined by elevation of arterial blood pressure, its clinical significance is derived from morbid events affecting the heart, brain, and kidneys. Complications such as myocardial infarction and stroke are not directly due to elevated pressures, but to the resulting structural changes in the heart and blood vessels. One of the structural consequences of hypertension, hypertrophy of the left ventricle (LVH), is the strongest known predictor, other than advancing age, of cardiovascular morbidity and mortality.

Hypertension is a disease that typically develops over years or decades, and in approximately 90% of cases it is diagnosed as primary (or essential) hypertension, indicating nonspecific etiology. The natural history of hypertension not only is characterized by a progressive increase in blood pressure, but also is accompanied by a progressive rise in systemic vascular resistance (SVR). The structural autoregulation theory of hypertension, originated by Folkow, helps explain why hypertension becomes a disease of elevated SVR. Arterial resistance vessels (arterioles) serve to regulate blood flow in accordance with local tissue demands. Vascular smooth muscle hypertrophy occurs as a physiological adaptation to increased blood pressure, which allows the vessel to maintain its ability to regulate blood flow at higher pressures. However, a consequence of this restructuring of the vessel is that it creates a rise in SVR, which increasingly provides the basis for the hypertension. Thus, structural autoregulation is a cascading pathophysiological adaptation. However, the unresolved question is what causes an individual's pressure to initially rise from normotensive levels, allowing the onset of pathophysiological processes. The answer to this question is likely multiple interacting factors, which vary across individuals, as reflected in Page's Mosaic Theory of hypertension. However, there is overwhelming evidence that the sympathetic nervous system is frequently implicated in the early stages of hypertension. In some, though not all, instances, this may give rise to a hyperkinetic circulatory state in which blood

pressure elevations are due to abnormally high cardiac output (CO) levels. It is the potentially pathophysiological effects of the sympathetic nervous system that suggest that stress may be implicated in the etiology of hypertension.

Stress and Blood Pressure

Psychophysiology of Stress

The impact of emotions and stress on blood pressure were recognized as early as the seventeenth century, as reflected in William Harvey's studies of the circulation. However, it was not until the twentieth century that the psychophysiological mechanisms by which stress influenced blood pressure became better understood. In 1915, Walter Cannon proposed that the sympathetic nervous system served to mobilize the cardiovascular system in preparation for fight or flight by engendering an increase in blood pressure and blood flow to the skeletal muscle vasculature. In the 1940s, the work of Hess identified the defense response, which was characterized by an integrative behavioral and cardiovascular activation pattern in response to electrical stimulation of the hypothalamus. This early animal research set the groundwork for our conception of the stress response as a psychophysiological phenomenon that primes the body in anticipation of recourse to sudden physical exertion.

In the latter half of the twentieth century, human research on the effects of stress on blood pressure was dominated by what has been termed cardiovascular reactivity studies, in which a variety of cardiovascular measures are recorded while volunteers are subjected to psychologically challenging tasks in a research laboratory setting. One of the earliest human studies was reported by Brod and colleagues in 1959, describing detailed cardiovascular response patterns observed while participants performed a stressful mental arithmetic task. Responses typically paralleled those of the animal defense response, with increased blood pressure and circulatory mobilization that occurred in the absence of significant metabolic demands. Essentially the same approach used by Brod has since been utilized extensively, and has led to the documentation of responses to a wide variety of laboratory stressors, including reaction time tasks, Stroop color-word conflict tasks, cold pressor tasks, simulated public speaking, and a variety of psychomotor tasks. One of the consistently striking observations from these studies is the individual differences in the magnitude of blood pressure responses: while at one extreme some individuals may show essentially no change in blood pressure, others may respond with an increase of

50 mmHg or more to exactly the same laboratory stressor. There is compelling evidence that these individual differences in blood pressure reactivity to stress are relatively robust, exhibiting consistency across a variety of laboratory tasks, and showing stability over time.

Blood pressure alone provides a limited picture of the hemodynamic pattern of changes elicited by stress. Hemodynamic studies emphasize the interrelationships among blood pressure, CO, and SVR. Since blood pressure is the product of CO and SVR, the concurrent measurement of blood pressure and CO permits the derivation of SVR, providing a more complete description of the hemodynamic response. Two very early stress studies conducted in the 1940s used invasive procedures to measure CO and, moreover, considered the measurement procedures themselves as presenting the source of psychological stress. These studies demonstrated how the stress response associated with anxiety was characterized by a mobilization of the cardiovascular system, leading to blood pressure increases due to an increase in CO. The same response pattern to a mental arithmetic task was observed over a decade later by Brod. In this classic study, CO was measured invasively using dye dilution, and forearm blood flow was also measured, leading to the demonstration that the increased CO during psychological stress appeared to be distributed to the skeletal muscle vasculature. These studies demonstrated that the fight-or-flight response, previously observed in animals, could be elicited in humans by psychological stressors.

With the development of noninvasive methods to measure cardiac output unobtrusively, in the past two decades there has been a resurgence of laboratory studies of the hemodynamic aspects of the stress response. These studies have consistently illustrated that there are dramatic individual differences in the hemodynamic patterns underlying blood pressure increases during stress. Thus, apparently similar cardiovascular responses to a stressor in terms of blood pressure change may belie very different underlying hemodynamic mechanisms. While some individuals exhibit blood pressure increases during stress due to an increase in CO and a decrease in SVR (fight-or-flight response), others may vasoconstrict, leading to an increase in blood pressure due entirely to increased SVR. Observations of this kind have led to an extension of descriptive categories in cardiovascular reactivity research, from high vs. low or hot vs. cold reactors, referring to the magnitude of blood pressure responses, to myocardial and vascular reactors, referring to the hemodynamic mechanism producing the blood pressure response. The mechanisms and potential health consequences of these individual differences are currently under exploration.

Stress and Hypertension

Although acute stress can indisputably raise blood pressure for brief periods into the range defining hypertension, its role in the etiology of sustained hypertension remains controversial.

Observational studies of stress reactivity There is some evidence that risk of hypertension development may be related to stress reactivity. Hypertensive individuals have been found to show greater blood pressure reactivity than normotensives. Individuals with normal blood pressure who have hypertensive parents also typically exhibit greater blood pressure reactivity to laboratory stressors than normotensives with no family history. Whereas studies describing only observations of the magnitude of blood pressure responses have reported inconsistent findings for the effects of race, a number of studies have reported that Black men, who are at increased risk for hypertension, appear to exhibit abnormally elevated vascular tone, indexed by SVR, during exposure to a variety of laboratory stressors. Gender differences in reactivity have also been observed, with men exhibiting greater systolic blood pressure and SVR increases than premenopausal women. These gender differences in SVR response are consistent with those observed for race in that men are more likely to develop hypertension than premenopausal women.

Since hypertension is a disease whose natural history involves progressively increasing blood pressure, which in turn is secondary to progressively increasing SVR, the tendency for individuals at greater risk for developing hypertension to exhibit pronounced stress-induced blood pressure and/or SVR reactivity is provocative. Reactivity studies have also demonstrated that stress can modify renal function, leading to sodium and fluid retention in some individuals, which may heighten and prolong stress-induced blood pressure elevations; altered renal function may be one of the target organ effects of stress that could impact long-term blood regulation. Additionally, several recent studies have reported an association between blood pressure reactivity and LVH, whereby the magnitude of blood pressure response during stress is directly related to LV mass.

Prospective studies of stress reactivity Although correlational evidence of the kind described previously is provocative, it is inconclusive since it does not rule out the possibility that associations of stress reactivity to hypertension risk and/or end organ damage may simply be a manifestation of underlying pathophysiology; the stress response may be a marker rather than a mechanism of disease. Prospective

studies should help resolve these cause–effect limitations by focusing on the temporal relationship of the correlated observations. As previously noted, hypertensive individuals are characterized by blood pressure hyperactivity, but is blood pressure hyperactivity in normotensives associated with the subsequent development of hypertension? Evidence relating to this question comes both from human and animal studies.

Animal studies In 1969, Forsyth described the effects of prolonged daily exposure to unsignaled shock avoidance schedules in rhesus monkeys. After several months, three of four animals developed hypertension. However, these findings were not replicated in other species. Beginning in the 1970s, a series of elegant studies was reported by Henry and colleagues on the effects of chronic social conflict in mice. When reared in isolation and subsequently transferred to colonies that made social conflict unavoidable, hypertension developed after several months, suggesting that chronic social stress could be a cause of hypertension. However, subsequent studies in other species were again less definitive. The importance of genetic susceptibility was demonstrated by Lawler's 1980s studies of borderline hypertensive rats, which developed hypertension when exposed to prolonged aversive reinforcement contingencies, whereas control animals typically remained normotensive. Other studies have suggested that the development of hypertension in salt-sensitive strains of animals of various species is exacerbated by exposure to stressful behavioral paradigms. In summary, the animal studies suggest that stress may play a role in the development of hypertension, but its impact is only likely to be important in the context of a pre-existing disposition to develop the disease.

Human studies There has been much interest in the possibility that individuals characterized by unusually large magnitude blood pressure reactivity may be at increased risk for the development of hypertension. Several studies have generated indirect support for the so-called stress reactivity hypothesis. In an early study of White males, blood pressure reactivity in response to a hand cold pressor task was examined as a predictor of the development of hypertension at 20- and 36-year follow-up. The study revealed that subjects in the top quartile of systolic blood pressure reactivity had a greater incidence of hypertension from age 40 onward, compared to the less reactive subjects. Other relatively long-term follow-up studies have shown the cold pressor task to predict the development of elevated blood pressure. One criticism of these studies is that the cold pressor may not be a

purely behavioral stressor, since it is a painful stimulus for most individuals; although, in common with other laboratory behavioral stress tasks, its cardiovascular effects are mediated primarily by sympathetic nervous system activation.

Blood pressure reactivity to a reaction time task involving threat of shock was found by Light and colleagues to predict increases in blood pressure 10 to 15 years later. High levels of systolic blood pressure during the reaction time task were associated with higher clinic and ambulatory systolic pressure at follow-up, even after controlling for parental history of hypertension and baseline resting systolic and diastolic blood pressure. Several studies in adolescents and young adults examining reactivity to mental arithmetic as a predictor of hypertension have also found evidence that high blood pressure reactivity and/or delayed recovery of the stress-induced blood pressure response predicts the development of hypertension at follow-up 1–5 years later. Other studies have made similar observations based on initial reactivity assessments using video games, Stropp color-word conflict, and mirror trace tasks. The studies reported to date are inconsistent in whether systolic blood pressure reactivity, diastolic blood pressure reactivity, or both is predictive of hypertension development. This discrepancy may also reflect differences in the population sample being studied, as well as differences in the nature of the blood pressure response associated with the specific stressor employed in a given study. It is also of note that the degree to which reactivity has predicted the development of high pressure or hypertension has been generally modest. As suggested by the animal studies, genetic susceptibility (i.e., family history of hypertension) is likely to be an important determinant of hypertension development, but was not considered in all of the reactivity studies. Furthermore, if stress reactivity is viewed as a disease mechanism, then the predictive value of high reactivity is likely to be largely moderated by the extent to which repeated acute stress responses occur in the time interval from assessment to follow-up. Assessing the frequency of stress responses over the long term presents a methodological challenge that has yet to be met in reactivity studies.

Psychosocial factors and hypertension Relatively few prospective studies have examined the relationship of psychosocial factors to the subsequent development of hypertension. The Framingham Heart Study examined the role of anxiety, anger intensity, and anger suppression in normotensive men. After controlling for biological risk factors, anxiety at baseline remained an independent predictor of hypertension in middle-aged men. The National Health

and Nutrition Examination Survey I Epidemiologic Follow-up Study also followed several thousand initially normotensive individuals for 7 to 16 years to test the hypothesis that symptoms of anxiety and depression increased the risk of developing hypertension. High anxiety and depression were independent predictors of incident hypertension ($\geq 160/95$ mmHg) in White men aged 45 to 64 years and Black men aged 25 to 64 years. In a large prospective study in Alameda County, California, unemployment, job insecurity, and concerns over job performance in men and low-status work in women were associated with increased likelihood of developing hypertension, after controlling for demographic and behavioral risk factors. Finally, several smaller correlational and prospective studies have found a significant relationship between the personality trait defensiveness and hypertension.

Studies of chronic stress Studies of environments that may engender chronic states of stress also provide a test of the effects of stress on blood pressure. In a study of individuals residing near the Three Mile Island nuclear power plant, blood pressure was found to rise shortly after the 1979 accident in which part of the facility was destroyed. Blood pressure remained elevated for 5 years and was associated with increased sympathetic nervous system activity compared to individuals residing further from the nuclear plant. Studies of prisoners also suggest that living in overcrowded conditions is associated with elevated blood pressure. Occupations considered to be associated with chronic stress include air traffic controllers, who have been found to show a much greater incidence of hypertension than workers in comparable, but less stressful, occupations. A relationship between occupational stress and elevated blood pressure and/or hypertension is supported by other studies, including those utilizing ambulatory blood pressure measurement technology to measure blood pressure in the working environment.

Ambulatory Blood Pressure

Methodology

Advances in computer and miniaturization technology over the past 25 years have aided the development of wearable ambulatory blood pressure (ABP) monitors. ABP monitors are able to capture the normal activities and environmental interactions during the course of an individual's day. Utilizing common blood pressure measurement techniques, ABP monitors permit frequent noninvasive blood pressure recordings based upon auscultatory and/or oscillometric measurement principles. ABP monitors are typically programmed to take blood pressure measurements

several times per hour, at variable intervals, throughout the workday, evening, and overnight. Individuals are ordinarily instructed to follow a normal schedule during waking hours and to complete a brief diary immediately following each reading indicating factors such as time, posture, activity, location, and current emotional/cognitive state. Factors such as posture and physical activity have been shown to systematically influence blood pressure. Location and emotional state have also been found to affect blood pressure, but the magnitude of their effects varies greatly from individual to individual. Finally, recent research suggests that the degree of familiarity with a conversant as well as the positiveness or negativeness of the interpersonal interaction affects ABP levels.

White Coat Hypertension

Because morbid events are often associated with hypertension, blood pressure is routinely measured during visits to a physician's office. However, psychological stress associated with visiting a doctor's office can spuriously inflate clinic blood pressures in some patients. Individuals whose blood pressures are in the hypertensive range when measured in the clinic but are normal in other contexts, as indicated by ABP measurements and/or automated home measurements, are classified as white coat hypertensives. Studies suggest that when compared to sustained hypertensives, white coat hypertensives are more likely to be younger female nonsmokers with a lower body mass index. Although most studies have found no association between psychological factors and white coat hypertension, there is some evidence that anxiety may be a related characteristic. Multiple ABP measurements throughout the day are thought to be more representative than clinic measures of the overall load that is placed on the cardiovascular system. Thus, 24-h ABP recordings can provide important supplementation of clinic readings in the accurate diagnosis of hypertension and are more closely correlated with end organ damage, including left ventricular mass, than clinic blood pressure.

Emotional States, Psychosocial Characteristics, and ABP

ABP monitoring is particularly well suited for capturing acute affective states that occur in response to daily stress. Several studies have indicated that emotions, as indicated by diary ratings of moods, influence ABP levels. Both negative and positive moods have been found to be associated with increased ABP. However, a great deal of individual variability in the reporting of diary-rated moods has been observed, suggesting that certain individuals may be more

susceptible to experiencing emotional upset throughout the day. Similarly, the effects of emotions on ABP are more pronounced in some individuals than in others. Nonetheless, studies indicate that for diary ratings of negative emotions, including stress, frustration, and tension, greater emotional intensity is accompanied by higher ABP.

There is some evidence that psychosocial characteristics that are risk factors for cardiovascular disease, including anxiety, depression, and lack of social support, are associated with greater fluctuations in diary ratings of negative moods; these individuals have also been found to exhibit a greater impact of emotion on ABP. There is also evidence to suggest that negative emotional dispositions, such as hostility, may have an impact on ABP, but only in certain social contexts. While recent studies indicate that dispositional characteristics (e.g., defensiveness and pessimism), coping responses (e.g., religious coping), and reactions to racial inequality (e.g., perceived racism) may be associated with ABP levels, these preliminary findings would benefit from further replication.

ABP and Location

Work If a 24-h day is divided into work, home, and overnight periods, most people's blood pressure will be highest during work, lower while at home, and lowest during the overnight period. However, these locational effects on blood pressure vary greatly from one individual to the next. This suggests that location per se is not directly associated with blood pressure, but rather that location is associated with a variety of behavioral factors (e.g., activity level, posture) and psychosocial factors (e.g., mood, social support, negative interpersonal exchanges, stress) that under certain circumstances may influence blood pressure. ABP studies conducted on individuals across a variety of occupations (e.g., paramedics, physicians, clerical personnel) indicate that perceived or actual occupational stress does, in fact, increase ABP levels. These associations have prompted researchers to examine specific job characteristics that might contribute to elevations in blood pressure across a variety of occupations. One model of job characteristics that contributes to work-related stress is Karasek and Theorell's job strain model. Karasek and Theorell maintain that a deleterious combination of high psychological job demands and low decision latitude at work results in a condition of job strain. Numerous studies have found that higher levels of job strain are associated with higher blood pressure, particularly in men. There is also evidence that job strain is associated with increased LV mass, which may be a consequence of ABP effects.

Home Behavioral and psychosocial influences on ABP at home are likely to be very different from behavioral and psychosocial influences at work. The home environment may be influenced by factors such as the quality of familial relationships and their ability to provide social support, the level of domestic and family responsibilities, and the existence of role conflict or dissatisfaction. Although research on the home environment and ABP is sparse, one study that examined blood pressure in working women found that married working women had higher ABP than unmarried working women. These findings suggest that, compared to unmarried working women, the higher ABP levels found in married working women may reflect their greater combined work and home responsibilities, possibly suggesting a total greater workload. Elevated ABP in the home environment may also be associated with home stress, making it more difficult to unwind, both psychologically and physiologically, after work.

Sleep Blood pressure typically falls during nighttime sleep. Most individuals are characterized by >10% fall in nighttime blood pressure and are referred to as dippers, in contrast to non-dippers, who exhibit <10% nighttime blood pressure fall. Sleep blood pressures may be influenced, at least in part, by events that occur during waking hours. Studies have found that the daily experience of negative moods on blood pressure appears to carry over from daytime waking to nighttime sleep. Individuals that frequently report anger during the day tend to have higher levels of blood pressure, especially during sleep. There has been some speculation that unresolved stressful experiences and negative mood states that occur during the day might contribute to disrupted sleep patterns, a diminished quality of sleep, or somatic tension that influences blood pressure levels. In addition, a positive family history of hypertension, low socioeconomic status, and African American ethnicity have been associated with blunted nocturnal blood pressure dipping. Dietary factors, including high sodium consumption and low potassium consumption, also have been related to blunted nighttime blood pressure dipping. Sympathetic nervous system activity also appears to be an important factor, with nighttime catecholamine levels and vascular adrenergic receptor activity also appearing to play a role in nocturnal blood pressure dipping. The magnitude of nocturnal decline in blood pressure has been found to relate to LV mass, with non-dippersexhibiting elevated LV mass and higher likelihood of LVH compared to dippers. This evidence underscores the importance of studying nighttime blood pressures and its possibly complex biopsychosocial determinants.

See Also the Following Articles

Cardiovascular System and Stress; Hypertension; Sympathetic Nervous System; Workplace Stress.

Further Reading

Joint National Committee on Prevention, Detection, Evaluation Treatment of High Blood Pressure. (2003). The seventh report of the joint national committee on prevention, detection, evaluation and treatment of high blood pressure. *Journal of the American Medical Association* 289, 2560–2572.

Lloyd-Jones, D. M., Evans, J. C. and Levy, D. (2005). Hypertension in adults across the age spectrum: current outcomes and control in the community. *JAMA* 294, 466–472.

Manuck, S. B., Kasprowicz, A. L. and Muldoon, M. F. (1990). Behaviorally-evoked cardiovascular reactivity

and hypertension: conceptual issues and potential associations. *Annals of Behavioral Medicine* 12, 17–29.

O'Rourke, M. F., Kelly, R. and Avolio, A. (1992). *The arterial pulse*. Malvern, PA: Lea & Febiger.

Pickering, T. (1997). The effects of occupational stress on blood pressure in men and women. *Acta Physiologica Scandinavica* 640(supplement), 125–128.

Pickering, T. G., Shimbo, D. and Haas, D. (2006). Current concepts: Ambulatory blood-pressure monitoring. *New England Journal of Medicine* 354, 2368–2374.

Shapiro, D., Jamner, L. D., Lane, J. D., et al. (1996). Blood pressure publication guidelines. *Psychophysiology* 33, 1–12.

Steptoe, A. (1997). Behavior and blood pressure: implications for hypertension. In: Zanchettii, A. & Mancia, G. (eds.) *Handbook of hypertension, Vol. 17: Pathophysiology of hypertension*, pp. 674–708. Amsterdam, The Netherlands: Elsevier.

Blood–Brain Barrier, Stress and

S N Malaeb and B S Stonestreet

Brown Medical School, Providence, RI, USA

© 2007 Elsevier Inc. All rights reserved.

Structure and Function

Disturbances in Endogenous Stressors

Disturbances in Exogenous Stressors

Implications and Clinical Correlations

Glossary

<i>Blood–brain barrier</i>	The interface between circulating blood and brain interstitium.
<i>Hypoxia</i>	Deficiency in the amount of oxygen reaching body tissues.
<i>Ischemia</i>	Cessation of blood flow with subsequent deprivation of glucose and oxygen delivery to the tissue.
<i>Neuroinflammation</i>	The activation of the innate inflammatory cascade within the brain parenchyma in response to various insults, cellular damage, and inflammatory molecules.
<i>Neurovascular unit</i>	The histological structure and function forming the blood–brain barrier.
<i>Tight junctions</i>	Protein complexes located at the apical portion of the cell that associate with homologous structures on the adjacent cell to seal the gap between the cells.

Structure and Function

The blood–brain barrier (BBB) serves as an interface between circulating blood and the brain interstitium. Histologically, the BBB corresponds to the neurovascular unit (**Figure 1**), a highly specialized structure composed of specialized brain capillary endothelial cells, astrocytes and astrocyte processes, pericytes, neuronal terminals, and occasional microglial cells. The brain capillary endothelial cells line the interior of the neurovascular unit and differ from systemic endothelial cells by having rare pinocytotic vesicles and absent fenestrations. Thus, the main gap in the barrier is the space between the cells. These endothelial cells have extensive tight junctions that seal the gap to form the barrier and are held together by strong adherence junctions. Tight junctions are located at the apical portion of the cell and are composed of transmembrane proteins such as occludin, claudins and junctional associated proteins, and intracytoplasmic proteins such as zonula occludens and other proteins that bridge the tight junctions to actin and other proteins of the cytoskeleton. The cells of the neurovascular unit express various transporters that regulate the flux of glucose and other nutrients, metabolites, water, sodium, potassium, and other substances into and out of the brain parenchyma. The BBB also isolates the brain tissues from blood constituents and drugs that could alter neuronal function and provides a large surface area by which the brain

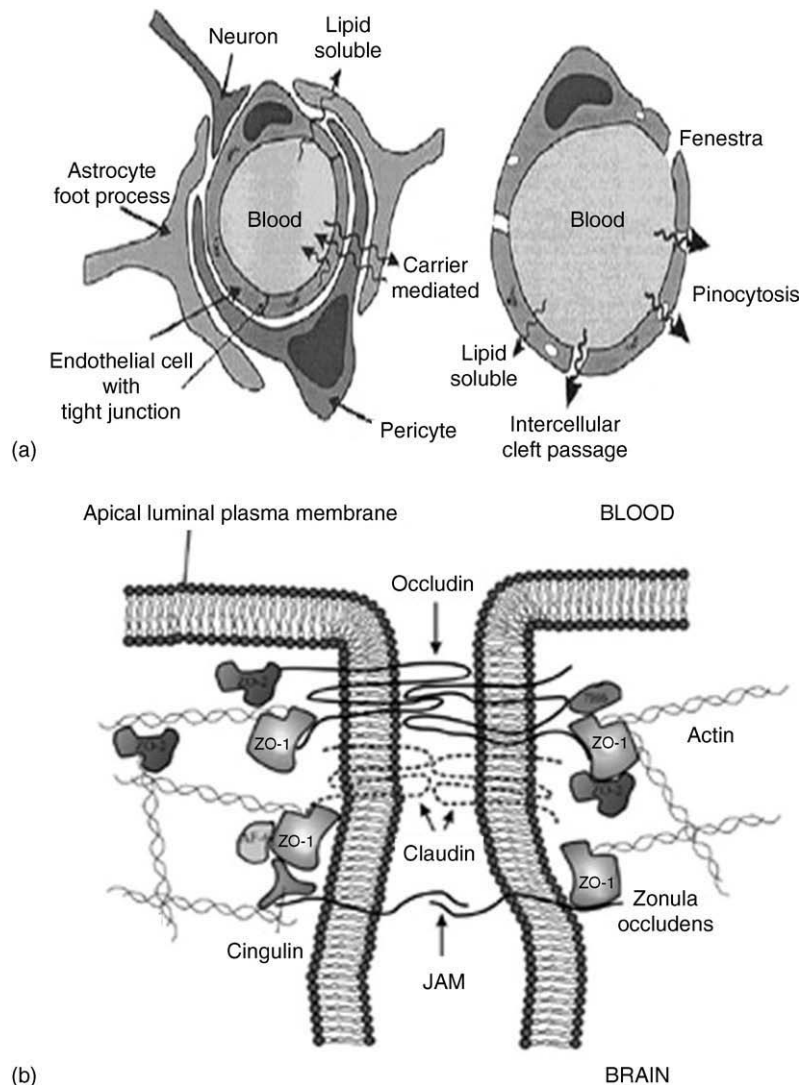


Figure 1 Schematic illustrating specialized features of the neurovascular unit compared to other endothelial vascular beds in the body (a) and the basic structure of tight junctions between endothelial cells at the blood–brain barrier (b). Adapted with permission from Misra (2003) and Hawkins (2005).

can sense alteration in the circulation supplying its various structures. Proper functioning of the BBB is essential to maintain homeostasis of the central nervous system (CNS). Disruption of BBB integrity and alterations in its structure and constitution mediate the pathophysiology of a wide range of CNS disorders and the neurological sequelae of many systemic diseases. The BBB can respond to and be influenced by alterations in the internal and external conditions of the human body as well. This article reviews the pathology of the BBB in various disturbances and stressors intrinsic and extrinsic to the CNS.

Disturbances in Endogenous Stressors

Infection and Inflammation

The BBB integrity and permeability are influenced to a great extent by systemic and local inflammatory mediators acting on the various elements of the neurovascular unit. It has been shown in many experimental models that there is a breakdown in BBB with increased permeability to proteins in conditions of systemic inflammation. These effects are mediated by direct exposure and binding of circulating inflammatory ligands (e.g., lipopolysaccharide) by Toll-like receptors, and of circulating cytokines

(e.g., interleukin-1 and tumor necrosis factor- α) and activated complement proteins to their receptors at the BBB and other specialized structures such as the circumventricular organs and the choroid plexus. In addition, increased local production of these factors within the brain parenchyma causes similar effects.

When molecules of the inflammatory cascade bind to their respective response elements on the BBB, nuclear factor- $\kappa\beta$ and other inflammatory pathways are induced in the endothelial cells. In addition, other cellular constituents of the neurovascular unit are activated, which induce the activation of nitric oxide synthase (iNOS), cyclooxygenases that release prostanooids, free radicals, histamines, bradykinins, vascular endothelial growth factor, and other paracrine factors. These factors in concert serve to further propagate the inflammatory wave across the BBB and into the brain parenchyma. Propagation of these inflammatory waves also upregulates local matrix metalloproteases (e.g., MMP-9), which will further disrupt the extracellular matrix supporting the structure of the neurovascular unit and the blood vessels in the brain. Cell adhesion molecule (e.g., ICAM-1, VCAM, MCP-1) expression is also increased on endothelial cells. These molecules bind to the endothelial cells and facilitate the trafficking of inflammatory cell across the BBB into the brain parenchyma potentiating the ongoing inflammatory cycles. Increases in P-glycoprotein expression and other molecules also control the transport of inflammatory related substances across the BBB.

The expression, localization, and interaction of proteins that form the tight junctions between endothelial cells of the BBB are altered and disrupted during the inflammatory process described previously. This process serves to weaken the integrity of the BBB and increase the permeability of the barrier to cells, proteins, and smaller molecules normally present in the systemic circulation that would otherwise be excluded from the brain parenchyma by an intact BBB under physiological conditions. The resultant disturbances in CNS homeostasis contribute to the pathophysiology of a number of disease processes affecting the CNS.

Chronic local inflammatory pain in peripheral organs has been associated with increased cell adhesion molecule expression, microglial activation, alterations in tight junction proteins, and increased BBB permeability. These alterations in the BBB are thought to mediate some neurological manifestations of chronic fatigue syndrome and other rheumatic conditions.

Microorganisms can cause systemic illness and, consequently, induce the cascade of inflammatory events described previously. Further, many microorganisms have the ability to gain access into the brain parenchyma by translocating through the BBB.

Bacteria (e.g., group B streptococcus, *Escherichia coli*) viruses (HIV), yeast, and prions utilize many of the mechanisms and molecules described previously for inflammation in order to cross the BBB and gain access to the brain. Circumventing the BBB by neurotrophic lentiviral vectors is a major area of interest for gene delivery models in modern neuroscience.

The mechanisms of BBB alteration described for inflammatory conditions are basic not only to inflammatory diseases of the CNS, but also to the pathophysiological processes that underlie brain injury in hypoxia, ischemia, oxidative stress, trauma, tumors, and many other CNS disorders.

Hypoxia/Ischemia Reperfusion Injury

Hypoxia/reoxygenation and ischemia/reperfusion injury alter the permeability and structural properties of the BBB. A large number of *in vitro*, *in vivo*, and magnetic resonance imaging (MRI) studies have documented increased BBB permeability and compromised structural BBB integrity under hypoxic conditions. Both hypoxic and ischemic conditions induce alterations in the expression and localization of a variety of key tight junction proteins, glucose and nutrient transporters, water and electrolyte transporters, and increase in nonspecific transport of blood-borne proteins into the brain parenchyma.

Mechanisms underlying these changes in permeability and structure of the BBB involve the simultaneous activation of a number of second messenger systems that mediate the consequences of hypoxia/reoxygenation, and ischemia/reperfusion injury on the BBB. These include the increases in hypoxia inducible factor, vascular endothelial growth factor, nitric oxide, and the activation of protein kinase C, nuclear factor- $\kappa\beta$, adenylate cyclase, and intracellular calcium pathways. These mechanisms include direct effects of oxygen toxicity and metabolite deprivation, production of oxygen free radicals, cessation of shear blood flow with static leukocytes adhering to the blood vessel walls and invading the neurovascular unit, and the induction of local inflammatory cascades. Local and leukocyte-derived MMP-9 activation mediates BBB breakdown after transient focal cerebral ischemia. During thrombolytic treatment of human stroke and the resulting reperfusion, lipoprotein receptor-mediated induction of MMP by tissue plasminogen activators can predispose a patient to hemorrhagic complications. Higher MMP-9 pretreatment levels have been shown to correlate with more severe hemorrhagic complications after thrombolysis. Biochemical markers of BBB dysfunction and brain ischemia such as elevated S100B values have been shown to be powerful predictors of a poor prognosis for neurological outcomes after cardiac surgery,

particularly when elevated levels are associated with other CNS complications.

Hypercapnia and Acidosis

Markers of BBB permeability are increased in the brain parenchyma and cerebrospinal fluid (CSF) under conditions of severe hypercapnia. Hypercapnia-related increases in bilirubin deposition results from increased BBB permeability for bilirubin and is one of the mechanisms by which hypercapnia can predispose newborn subjects to CNS bilirubin toxicity. Other mechanisms include the accumulation of these markers in the brain because decreases in CSF secretion during prolonged hypercapnia could limit the removal of the substances from the CNS via the CSF route. Elevated carbon dioxide levels and the direct effects of acidosis can increase the barrier permeability to molecules in the systemic circulation, which normally are excluded by the intact BBB from the CNS. At relatively low pHs, the resulting changes in the ionic milieu can alter the negatively charged endothelial cell surface sialoglycoproteins and glycolipids, which are integral parts of the BBB. Changes in blood carbon dioxide concentrations, hydrogen ion, and bicarbonate contents also may alter the expression and activity of H^+ -ATPase, Na^+/H^+ exchanger, a ubiquitous plasma membrane transport system, and other ion exchangers, which normally regulate the cytoplasmic pH in the cells of the BBB. Such changes in the composition of the BBB contribute to the pathophysiology of acid-base-related disturbances in the CNS.

Osmotic Stress

Exposure to hyperosmotic conditions can disrupt the tightness and integrity of the BBB. Osmotic opening of the BBB is potentially reversible upon cessation of the osmotic stress. Reversible osmotic disruption and reconstitution of the BBB may be due not only to simple mechanical shrinkage of the endothelial cells but also to intracellular Ca^{2+} -activated mechanisms. *In vitro* studies have shown that intracellular Ca^{2+} concentrations increase rapidly, and values peak within 10 s after the application of mannitol and return to the basal levels within 200 s. A specific Na^+/Ca^{2+} exchanger mediates the reconstitution of the BBB at the cellular level. Considering the usefulness of reversible opening for delivery of chemotherapeutic agents across the BBB, manipulations of this exchanger by selective inhibitors may have clinical implications in prolonging the opening phase of the capillary endothelial cells for more efficient drug delivery into the CNS.

During chronic hypertonic conditions such as severe hypernatremia, cells accumulate intracellular

organic osmolytes to equilibrate the osmotic gradients. Rapid reduction of extracellular hypertonicity predisposes these cells to excessive swelling and edema. Circulating ions and osmolytes equilibrate slowly across the BBB, and rapid decreases in plasma osmolality can cause major brain fluid shifts resulting in edema and severe cerebral lesions. Rapid correction of chronic hyponatremia can cause osmotic-induced demyelination of the brain. An intact BBB protects oligodendrocytes to a certain extent from the direct exposure to osmotic imbalances. However, disruption of the BBB during an osmotic stress results in massive damage to oligodendrocytes with subsequent development of demyelinating lesions in the brain.

The ability of the brain to exhibit volume regulation when subjected to osmotic stress develops in a region- and age-related fashion. In ovine fetuses, brain volume regulation develops first in the cerebral cortex at 90% of gestation and in the cerebral cortex and cerebellum of newborn lambs, and then it develops in the medulla of the lambs at 20–30 days of age. Therefore, in young subjects, the adverse effects of hyperosmotic stress may vary in an age- and region-specific manner.

Diabetes

The effects of long-standing diabetes on BBB permeability and integrity are region specific and variable. MRI studies in well-controlled diabetics showed increases in signal intensity in the basal ganglia with contrast infusion, indicating increased permeability of the BBB in basal ganglia compared to controls. Other studies suggest that chronic hyperglycemia may induce vascular endothelial growth factor production by glial cells at the blood-retinal barrier. This alters the expression of tight junction proteins such as occludin and ZO-1, and consequently increases capillary endothelial permeability that predisposes to the edema formation in diabetic maculopathy. In contrast, BBB permeability and tight junction protein expression were not altered in the cerebral cortex of diabetic patients. Chronic hyperglycemia decreases the number of glucose transporters available at the BBB. The primary mechanism underlying this downregulation is a posttranscriptional inhibition of glucose transporter mRNA translation. Such an adaptation could decrease the net flux of glucose into the brain during sustained hyperglycemia. The downregulation of glucose transporters during chronic hyperglycemia also contributes to the alarming sensitivity to hypoglycemia in diabetic patients. Diabetic ketoacidosis is associated with a life-threatening development of cerebral edema immediately after the recovery from ketoacidosis. Accumulation of weak acids such as ketoacids and free fatty acids in the cytoplasm of brain cells lowers the intracellular pH,

resulting in activation of the Na^+/H^+ exchanger, leading to cellular swelling and cerebral edema.

Hypertension

Studies on the effects of chronic hypertension on BBB permeability indicate a regional predilection for CNS lesions compared with permeability under normotensive conditions. Focal increases in the rate of BBB transfer for small molecules have been reported to increase up to two to seven times the normal rate of transfer. Particularly susceptible regions include the hippocampus, hypothalamus, blood-retinal barrier, and the blood-CSF barrier at the choroid plexus. Acute hypertensive crises have been associated with disruptions in the BBB integrity, when systolic blood pressures reach a certain threshold, or when the BBB has been weakened by other conditions such as in diabetes. These disruptions in the BBB during acute hypertension are more pronounced to neutrally charged molecules than to negatively charged molecules.

Disturbances in Exogenous Stressors

Smoking

Smoking is one of the major risk factors for stroke in humans. *In vitro* stroke models studying bovine brain microvessel endothelial cell cultures under hypoxic/aglycemic conditions demonstrated that BBB K^+ transport mechanisms by the abluminal Na,K,2Cl -cotransporter are downregulated after exposure to nicotine and cotinine at concentrations comparable to plasma levels seen in smokers, and these changes could influence ischemic brain K^+ homeostasis in crucial times of neuronal recovery after stroke.

Alcohol

Alcohol increases BBB permeability in a dose-dependent fashion. Ethanol-induced activation of myosin light chain kinase (MLCK) in the endothelial cells of the BBB leads to phosphorylation of MLCK, occludin, and claudin-5 with resultant cytoskeletal alterations and tight junction changes that facilitate monocyte migration across the BBB. In addition, alcohol predisposes to oxidative stress in brain endothelial cells, causing BBB dysfunction and aggravation of neuroinflammatory disorders of the CNS.

Immobilization and Restraint Stress

Immobilization and restraint stress have been associated with increased BBB permeability in the hypothalamus, midbrain reticular formation, and the cerebellum. Activation of mast cells that preferentially populate these regions of the brain and possible activation of serotonergic pathways have been speculated

as possible mechanisms. The resultant oxidative stress and BBB breakdown have been associated with increased circulating levels of S100B protein. Interestingly, biting suppresses oxidative stress induced by restraint stress, and the antistress effect of masticatory motor activity movements, such as biting or chewing, may be important for reducing the adverse effects associated with exposure to psychological stressors.

Steroids

Exogenous corticosteroid administration has been shown to decrease BBB permeability in many *in vitro* and *in vivo* studies of adult and fetal models. Other studies suggest that corticosteroid administration promotes tightening of the BBB during adverse events such as focal cerebral ischemia. Glucocorticoids may act by suppressing inflammation and by directly affecting phosphorylation, organization, and upregulation of tight junction proteins in endothelial cells. However, glucocorticoids also can inhibit glucose transport in cultured hippocampal neurons and glia, and may reduce nutrient delivery to the brain during critical periods of hypoxia and ischemia. Pretreatment with antenatal dexamethasone did not reduce the extent of pathological brain injury following ischemia/reperfusion in an ovine fetal model. Estradiols have also been shown to attenuate BBB disruption induced by cerebral ischemia/reperfusion injury in female rats and to enhance brain glucose uptake by upregulation of glucose transporter-1 expression in the BBB. Thus, estradiols may have advantages over corticosteroids in this regard.

Endogenous corticosteroid production by the hypothalamic-pituitary-adrenocortical axis under stressful conditions, such as systemic inflammation and cerebral ischemia, provides an endogenous negative feedback loop to reduce the extent of ongoing neuroinflammation and BBB disruption secondary to these insults. Blocking this mechanism by corticosteroid antagonists markedly exaggerates the damage secondary to ischemic brain injury.

Thermal Stress: Hypothermia, Hyperthermia, and Ultrasound Waves

Hypothermia, when initiated early on, can attenuate the development of the BBB opening to various insults. In an ischemia/reperfusion model, whole body cooling to 31.5–32.5°C attenuated the increase in BBB permeability when cooling was initiated early during reperfusion. Hypothermia also has been effective in attenuating BBB disruption by hyperosmolar infusions. This suggests that there are direct effects of hypothermia on the elements of the neurovascular unit, in addition to its role in the reduction of

inflammatory cascades accompanying ischemia/reperfusion injury.

Hyperthermia has detrimental effects on ischemic infarct volumes both during and after the induction of ischemia. Many studies have found increases in BBB permeability along with disruptions in the integrity of tight junctions and decreased expression of ZO-1 after hyperthermia. The pathophysiology of hyperthermic brain injury and BBB disruption is complex, with many factors such as NOS, histamines, prostaglandins, and other pathways contributing to the hyperthermic-related injury. These changes are more pronounced in younger subjects than in adult subjects and may be implicated in the pathophysiology of febrile seizures in younger children.

Local opening of the BBB by mild hyperthermia induced by ultrasound waves is gaining interest as a novel approach to increase drug uptake in brain microvascular endothelial cells. High-intensity focused ultrasound is capable of reversible BBB disruption in a targeted region of interest by opening capillary endothelial cell tight junctions; this opening reverses after 72 h. Sometimes ultrasound-induced hyperthermia can damage tissues in a small volume within the region of BBB opening. Newer techniques are being studied to achieve local and reversible BBB disruption without damaging neurons by the use of non-invasive focused ultrasound at frequencies suitable for trans-skull sonications applied to cavitate special lipid-coated microbubbles within the cerebral microvasculature for delivery of both low- and high-molecular-weight therapeutic compounds to the brain.

Other Stressors

Effects of long-term mobile communication microwave exposure and electromagnetic radiation on the vascular permeability in the brain has been a public health concern. Current evidence suggest that most of the reported effects are negligible as long as the radiation intensity remains in the nonthermal range. BBB dysfunction has been speculated in the pathophysiology of high altitude cerebral edema and acute mountain sickness and diving and decompression sickness, probably due to circulatory microbubble-induced damage of the BBB.

BBB and Aging

Senescence is associated with changes in BBB function. Conditions associated with aging, such as hypertension and cerebrovascular ischemia, aggravate the age-related alterations in BBB function, and the barrier becomes more susceptible to disruption with aging. Such alterations appear to be region specific.

More profound changes have been associated with Alzheimer's disease, and the BBB is now thought to have an important role in this disorder. Common age-related changes in the neurovascular unit include loss of capillary endothelial cells, elongation of the remaining endothelial cells, decreased capillary diameter, and a decrease in the number of mitochondria in endothelial cells. The aging BBB shows alterations in carrier-mediated transport systems, including a decrease in transport of choline, glucose, butyrate, and triiodothyronine.

Implications and Clinical Correlations

BBB dysfunction plays a major role in the pathogenesis of a large number of neurological disorders including stroke, multiple sclerosis, Alzheimer's disease, Parkinson's disease, cerebral palsy, leukoencephalopathies, and cerebral hemorrhages. Impairment of BBB function contributes significantly to the neurological sequelae resulting from or associated with many systemic stresses and disorders. A thorough understanding of the pathophysiology of the BBB and of the factors influencing its structure, function, and integrity is essential for the development of targeted drug delivery strategies and safer and more effective therapies in the management of many diseases of the CNS.

See Also the Following Articles

Hypertension; Infection; Thermal Stress; Inflammation.

Further Reading

- Abbott, N. J. (2000). Inflammatory mediators and modulation of blood-brain barrier permeability. *Cellular and Molecular Neurobiology* 20(2), 131-147.
- Abbruscato, T. J., Lopez, S. P., Roder, K. and Paulson, J. R. (2004). Regulation of blood-brain barrier Na,K,2Cl-cotransporter through phosphorylation during *in vitro* stroke conditions and nicotine exposure. *Journal of Pharmacological and Experimental Therapy* 310(2), 459-468.
- Haorah, J., Heilman, D., Knipe, B., et al. (2005). Ethanol-induced activation of myosin light chain kinase leads to dysfunction of tight junctions and blood-brain barrier compromise. *Alcoholism: Clinical and Experimental Research* 29(6), 999-1009.
- Hawkins, B. T. and Davis, T. P. (2005). The blood-brain barrier/neurovascular unit in health and disease. *Pharmacology Reviews* 57(2), 173-185.
- Hynynen, K., McDannold, N., Sheikov, N. A., Jolesz, F. A. and Vykhodtseva, N. (2005). Local and reversible blood-brain barrier disruption by noninvasive focused ultrasound at frequencies suitable for trans-skull sonications. *Neuroimage* 24(1), 12-20.
- Jonhsson, P., Backstrom, M., Bergh, C., Jonsson, H., Luhrs, C. and Alling, C. (2003). Increased S100B in blood after

- cardiac surgery is a powerful predictor of late mortality. *Annals of Thoracic Surgery* 75(1), 162–168.
- Misra, A., Ganesh, S., Shahiwal, A. and Shah, S. P. (2003). Drug delivery to the central nervous system: a review. *Journal of Pharmaceutical Sciences* 6(2), 252–273.
- Montaner, J., Molina, C. A., Monasterio, J., et al. (2003). Matrix metalloproteinase-9 pretreatment level predicts intracranial hemorrhagic complications after thrombolysis in human stroke. *Circulation* 107(4), 598–603.
- Rapoport, S. I. (2000). Osmotic opening of the blood–brain barrier: principles, mechanism, and therapeutic applications. *Cellular and Molecular Neurobiology* 20(2), 217–230.
- Rivest, S., Lacroix, S., Vallieres, L., Nadeau, S., Zhang, J. and Laflamme, N. (2000). How the blood talks to the brain parenchyma and the paraventricular nucleus of the hypothalamus during systemic inflammatory and infectious stimuli. *Proceedings of the Society for Experimental Biology and Medicine* 223(1), 22–38.
- Stonestreet, B. S., Sadowska, G. B., McKnight, A. J., Patlak, C. and Petersson, K. H. (2000). Exogenous and endogenous corticosteroids modulate blood–brain barrier development in the ovine fetus. *American Journal of Physiology: Regulatory, Integrative, and Comparative Physiology* 279(2), R468–R477.
- Unger, E. C., Unger, E. C., Porter, T., et al. (2004). Therapeutic applications of lipid-coated microbubbles. *Advanced Drug Delivery Reviews* 56(9), 1291–1314.
- Vannucci, S. J., Seaman, L. B. and Vannucci, R. C. (1996). Effects of hypoxia-ischemia on GLUT1 and GLUT3 glucose transporters in immature rat brain. *Journal of Cerebral Blood Flow and Metabolism* 16(1), 77–81.
- Witt, K. A., Mark, K. S., Hom, S. and Davis, T. P. (2003). Effects of hypoxia-reoxygenation on rat blood–brain barrier permeability and tight junctional protein expression. *American Journal of Physiology: Heart and Circulatory Physiology* 285(6), H2820–H2831.

Borderline Personality Disorder

H W Koenigsberg and L J Siever

Bronx VA Medical Center, Bronx, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by L J Siever and H W Koenigsberg, volume 1, pp 339–341, © 2000, Elsevier Inc.

Phenomenology and Epidemiology of Borderline Personality Disorder

Role of Neurobiology

Role of Life Experience

Psychological Structures in the Borderline Patient

Treatment of Borderline Personality Disorder

Glossary

<i>Neuro-transmitter</i>	One of a number of chemical substances which, when released at the terminal of one neuron and captured by a specialized receptor on the surface of another neuron, transmit a signal from one neuron to another.
<i>Personality</i>	Enduring patterns of perceiving, relating to, and thinking about the environment and oneself.
<i>Personality disorder</i>	The disturbance resulting from maladaptive personality traits which cause functional impairment or substantial distress.

Phenomenology and Epidemiology of Borderline Personality Disorder

Borderline personality disorder (BPD), a disorder characterized by emotional instability and impulsivity, is one of the more serious personality disorders. It is associated with high levels of distress, significant impairments in social and occupational functioning, and a risk of suicide approaching 10%. As a personality disorder, BPD reflects an enduring and inflexible maladaptation in an individual's characteristic pattern of perceiving, relating to, and thinking about the environment and self. In BPD, the individual is highly reactive to psychosocial events such as losses, separations, and frustrations. The individual with BPD may show intense but shifting extremes of emotion, often involving anger and depression. He or she may engage in impulsively self-destructive acts, or may recurrently threaten or attempt suicide. His or her relationships with others are typically intense and unstable, often oscillating between idealization and devaluation of others. By the same token, his/her own sense of identity is unstable and people with BPD are unsure of their goals, values, or sense of self. They may be terrified of real or imagined abandonment and may frantically attempt to control others to prevent this. Under the stress of losses, frustrations, or intense interpersonal relationships, individuals with BPD may regress, becoming more infantile or experiencing transient lapses in reality testing, which can

take the form of hallucinations or delusional ideas. They often experience a sense of boredom and inner emptiness. It is estimated that about 2% of the population meets criteria for BPD. It is more often diagnosed in women and is more common among first degree relatives of those with the disorder.

Role of Neurobiology

BPD appears to be the result of converging biological predispositions and early-life experience. Two biologically mediated dimensions of personality are particularly prominent in BPD, impulsive-aggression and affective instability. In borderline patients, the impulsive-aggression is most often self-directed, but it can also be directed to others. Impulsive-aggression has been shown to be associated with dysregulation in the serotonin neurotransmitter system. Cerebrospinal fluid levels of 5-hydroxyindoleacetic acid (5-HIAA), a breakdown product of serotonin, are inversely correlated with the degree of impulsive-aggressive and suicidal behavior in depressed patients and in patients with personality disorders. Borderline patients show a blunted response to the ingestion of fenfluramine, a serotonin-releasing and potentiating agent which triggers release of the hormone prolactin, and the degree of blunting in prolactin response to fenfluramine correlates with the level of impulsive-aggression manifest in the patient's life history. Studies, which use positron emission tomography (PET) scanning to show regional metabolic activity in the brain, have demonstrated that in BPD patients the response to the fenfluramine is diminished specifically in those brain regions responsible for inhibiting impulsive behaviors. Similar results were found using other serotonergic challenge agents.

The noradrenergic (norepinephrine) neurotransmitter system is also implicated in aggression in patients with personality disorders. Cerebrospinal fluid levels of 3-methoxy-4-hydrophenylglycol (MHPG), a norepinephrine breakdown product, have been shown to correlate positively with impulsive-aggressive behavior. The responsiveness of the noradrenergic system can be assayed by measuring the growth hormone response to clonidine. Growth hormone release in response to clonidine directly correlates with irritability. Since norepinephrine activity is associated with alertness, attention to external stimuli, and extraversion, it may be connected with outward-directed aggression. The interaction between serotonin and noradrenergic systems may determine the nature of aggression in BPD. Low serotonergic activity may predispose to impulsive aggression, which appears to be outward directed when noradrenergic activity

is high and inward directed when noradrenergic activity is low. Studies have suggested that the activity of another neurotransmitter, arginine vasopressin, may directly correlate with levels of aggression in personality disorders.

Affective instability may be associated with dysregulation of the neurotransmitter, acetylcholine. When borderline patients were given an acetylcholine-potentiating agent, physostigmine, they responded by becoming transiently more dysphoric for several minutes to 1 hour, whereas nonBPD patients did not show this reaction. The affective effect of physostigmine was associated with the patients' degree of affective instability, identity confusion, unstable relationships and feelings of emptiness and boredom, suggesting that these symptoms may all be related to affective instability.

Role of Life Experience

Early-life experience also appears to contribute to the development of BPD. Histories of repeated childhood physical or sexual abuse are commonly found among borderline patients, although such abuse experiences are also associated with a variety of other psychiatric disorders as well. Severe physical or emotional neglect may also predispose to BPD. Intense levels of aggressive drive, either resulting from an inborn predisposition or exceptionally powerful anger-inducing life experiences such as abuse, painful illness, or devastating losses, are also believed to contribute to borderline pathology. Some researchers have suggested that individuals who go on to develop borderline personalities have had impairments in their ability to evoke memories of their caretakers for use in soothing themselves in the absence of the caretaker. Other researchers emphasize the interaction of the affectively unstable child with a caretaking environment in which strong feelings are disavowed and the child is made to feel that his or her feelings are invalid.

Psychological Structures in the Borderline Patient

Psychoanalytic investigators have postulated that borderline individuals protect themselves from recognizing the consequences of their unacceptable feelings by developing inflexible and maladaptive defense mechanisms, such as splitting. Splitting is a defense in which the individual partitions his or her psychic experience into compartments which keep aggressive feelings away from images of loved and needed others. The partitioning of memory, which is the essence of splitting, may itself be a consequence

of the impact of affective instability upon emotion-state-dependent memory processes. Thus the use of splitting, with its associated anxiety-controlling effects, may derive from a psychobiological vulnerability to affective dysregulation. The extensive use of the splitting defense fragments the individual's internal images of self and others. This produces a confused sense of self and identity and contradictory pictures of others.

Treatment of Borderline Personality Disorder

Because of their ingrained nature, personality disorders are slow to change, and the impulsiveness and affective instability characteristic of BPD makes this disorder often difficult to treat. While no single treatment has been identified as the treatment of choice for BPD, patients have responded well to a number of psychotherapeutic approaches and to medications. Recent work has extended earlier observations that such anticonvulsant/mood-stabilizing medications as topiramate, divalproex sodium, lamotrigine, and carbamazepine can reduce anger, aggression, irritability, and impulsivity in borderline patients. Borderline patients also show a global positive response to the mood stabilizer, lithium carbonate. Monoamine oxidase (MAO) inhibitors, which affect several neurotransmitter systems, reduce depression, anxiety, hostility, self-destructiveness, and interpersonal sensitivity in BPD. Selective serotonin-reuptake inhibitor (SSRI) antidepressants have been shown to diminish hostility in borderline patients and may also reduce impulsive-aggression and depression. Low doses of antipsychotic medications have reduced the transient psychotic symptoms, hostility, impulsiveness, and, in some studies, depression in BPD. The newer atypical antipsychotic medications have been shown to reduce hostility, suspiciousness, paranoia, anxiety, and depressive symptoms in the disorder. At present, it remains an open question whether the beneficial effects of medication last beyond the 6–12-week periods examined in most studies. For patients with severe self-destructiveness and those who tend to be more action-oriented than introspective, a form of cognitive behavioral therapy, dialectical behavior therapy (DBT), has been shown to diminish self-destructive behavior and reduce the need for inpatient treatment. One recent study demonstrated that DBT in combination with the atypical antipsychotic medication, olanzapine, was more effective than DBT alone. Psychodynamic psychotherapies, which have helped

borderline patients develop an integrated sense of themselves and others and have led to enhanced interpersonal and career functioning, are suitable for patients who can contain dangerous behavior with the structure of an outpatient treatment contract and who are interested in identifying, understanding the origins of, and changing maladaptive interpersonal patterns. For some patients, sequential treatment with medication, DBT, and psychodynamic psychotherapy may be the optimal approach.

See Also the Following Articles

Catecholamines; Multiple Personality Disorder; Serotonin in Stress.

Further Reading

- Coccaro, E. F., Kavoussi, R. J., Hauger, R. L., et al. (1998). Cerebrospinal fluid vasopressin levels correlates with aggression and serotonin function in personality-disordered subject. *Archives of General Psychiatry* 55, 708–714.
- Cowdry, R. W. and Gardner, D. L. (1988). Pharmacotherapy of borderline personality disorder: alprazolam, carbamazepine, trifluoperazine, and tranylcypromine. *Archives in General Psychology* 45, 111–119.
- Kernberg, O. F. (1975). *Borderline conditions and pathological narcissism*. New York: Aaronson.
- Koenigsberg, H. W. (1993). Combining psychotherapy and pharmacotherapy in the treatment of borderline patients. In: Oldham, J. M., Riba, M. B. & Tasman, A. (eds.) *American Psychiatric Press Review of Psychiatry*, Vol 12. Washington DC: American Psychiatric Press.
- Koenigsberg, H. W., Harvey, P. D., Mitropoulou, V., et al. (2002). Characterizing affective instability in borderline personality disorder. *American Journal of Psychiatry* 159, 784–788.
- Koenigsberg, H. W., Woo-Ming, A. M., and Siever, L. J. (In press). Psychopharmacological treatment of personality disorders. In: Nathan, P. R. & Gorman, J. M. (eds.) *A guide to treatments that work*, 3rd edn., New York: Oxford University Press.
- Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. New York: Guilford.
- New, A. S., Hazlett, E. A., Buchsbaum, M. S., et al. (2002). Blunted prefrontal cortical 18fluorodeoxyglucose positron emission tomography response to meta-chlorophenylpiperazine in impulsive aggression. *Archives in General Psychiatry* 59, 621–629.
- Siever, L. J. and Davis, K. L. (1991). A psychobiological perspective on the personality disorders. *American Journal of Psychiatry* 148, 1647–1658.
- Skodol, A. E., Siever, L. J., Livesley, W. J., et al. (2002). The borderline diagnosis ii: biology, genetics, and clinical course. *Biological Psychiatry* 51, 951–963.

Brain and Brain Regions

A G Watts

University of Southern California, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A G Watts, volume 1, pp 342–348, © 2000, Elsevier Inc.

Introduction – Stress and the Brain

Brain Regions Concerned with Processing Stress-Related Information

Glossary

<i>Catecholamines</i>	Chemical signals (epinephrine, dopamine, and norepinephrine) synthesized from the amino acid tyrosine.
<i>Chromaffin cells</i>	Endocrine cells in the adrenal medulla that release adrenaline (primarily) into the blood on stimulation from sympathetic motor system.
<i>Diencephalon</i>	The part of the forebrain containing the thalamus and hypothalamus (including the preoptic area).
<i>Hypoglycemia</i>	Condition in which blood glucose concentrations fall below 60–75 mg dl ⁻¹ .
<i>Hypotension</i>	A reduction in blood pressure.
<i>Immediate early gene</i>	A gene whose product (usually a transcription factor) is rapidly produced as a consequence of receptor activation.
<i>Telencephalon</i>	The part of the forebrain containing the cortical regions (including the hippocampus), basal ganglia, bed nucleus of the stria terminalis, amygdala, and septal complex.

Introduction – Stress and the Brain

A wide variety of stimuli evoke motor events that are considered as stress responses. In simple organisms, these are made up of little more than sets of stereotyped reflexes. If the stressor is sufficiently intense or prolonged, these often consist of avoidance responses accompanied by endocrine-mediated alterations in metabolism. However, there is no evidence to suggest that invertebrates can associate emotional affect such as fear or anxiety with particular stressors. Only with the evolution of more complex central nervous systems (CNS) in general, and the telencephalon in particular, do we begin to see this type of associative processing.

In vertebrate species, stress is typically accompanied by endocrine secretions (e.g., from steroidogenic

tissue in the interrenal/adrenal glands), autonomic motor events (e.g., piloerection, increased secretion from the chromaffin cell/adrenal medulla system, and alterations in cardiovascular tone), sets of behavioral motor actions (e.g., avoidance responses, changes in defensive behavior, and fear-conditioned freezing), and those attendant emotional percepts generally recognized as being the most characteristic human signatures of stress (e.g., fear and anxiety). From a neural-systems perspective, these examples show that virtually all parts of the CNS can be used to varying degrees for processing the information associated with stress. Here we broadly consider which parts of the CNS are (1) concerned with processing the sensory information used to generate stress responses, (2) responsible for mediating the motor aspects of the stress response, and (3) involved with the integrative mechanisms that control stress responses in the face of influences from, for example, learning and memory mechanisms. To do this we explore concepts that are based on data derived from a number of experimental techniques, including lesions, neuroanatomical tract tracing, and the mapping of immediate early gene expression within the brain and spinal cord.

Many physiological stressors originating in the viscera are addressed by evoking simple reflexes to annul their potentially damaging consequences. Virtually no complex central processing is required and control is often exerted by classic negative feedback (homeostatic) mechanisms from the action of the target. An example of reactive homeostatic components of the stress response is the sympathetic activation evoked to counter the immediate visceral effects of many stressors. For example, if stress-related hypotension or hypoglycemia are prolonged or of sufficient intensity, they are rapidly negated by reactive stress responses that are evolutionarily rather old – sympathetically mediated vasoconstriction and catecholamine secretion by chromaffin cells, respectively (see **Comparative Anatomy and Physiology**).

In contrast, however, most stress responses typically have a more proactive (anticipatory) than a reactive (reflex) character, which involves complex neural processing. With stress activation of the hypothalamic-pituitary-adrenal (HPA) axis, for example, although adrenocorticotrophic hormone (ACTH) and corticosterone secretion are obviously triggered by the stressor, they are not motor responses that act to remove the immediate consequences of the stress in a manner similar to the reactive sympathetic responses just described. Instead, the actions of stress-induced

increases in circulating corticosterone levels (mediated in part by interactions with leptin, insulin, and the thyroid hormones) anticipate the possibility that the stressor will lead to a debilitating sequence of events, particularly the catabolic effects of negative energy balance, that require the mobilization of energy stores, particularly protein. Understanding the neural complexity required for processing what are effectively anticipatory events provides the framework for explaining why so many neural systems are involved with modulating stress responses.

Brain Regions Concerned with Processing Stress-Related Information

Sensory Processing

Generally speaking, virtually all extero- and viscerosensory modalities are processed along two types of pathway once they enter the CNS (Figure 1). The first, path A, is a relatively direct and rapidly effective route that requires little central integrative processing. These types of reflex pathways result in rapid and simply organized (stereotyped) motor

responses. Much of the neural processing for these reflexes occurs in the spinal cord (for the behavioral motor components of stress responses), hindbrain (for the autonomic motor aspects of stress responses), and hypothalamus (for the motor output of neuroendocrine stress responses). In path B, the second path, sensory information is also directed to those areas of the brain that provide the additional and often critical modulatory information about arousal state and behavioral planning (timing signals, priorities, etc.), so furnishing the animal with the adaptive plasticity for generating stress responses. Brain regions in the telencephalon (cerebral cortex, amygdala, hippocampus, and septal complex) provide pivotal contributions here, particularly with regard to the neural representations of sensory objects. Collectively, these processes can prioritize stress responses with relation to other behaviors (e.g., placing a higher priority on fleeing a predator than on feeding) and with modulating the organization of stress responses, decreasing or increasing their intensity based on past experience (learning and memory). As a consequence, they are also likely to be involved with generating the detrimental effects of chronic emotional stress. Finally,

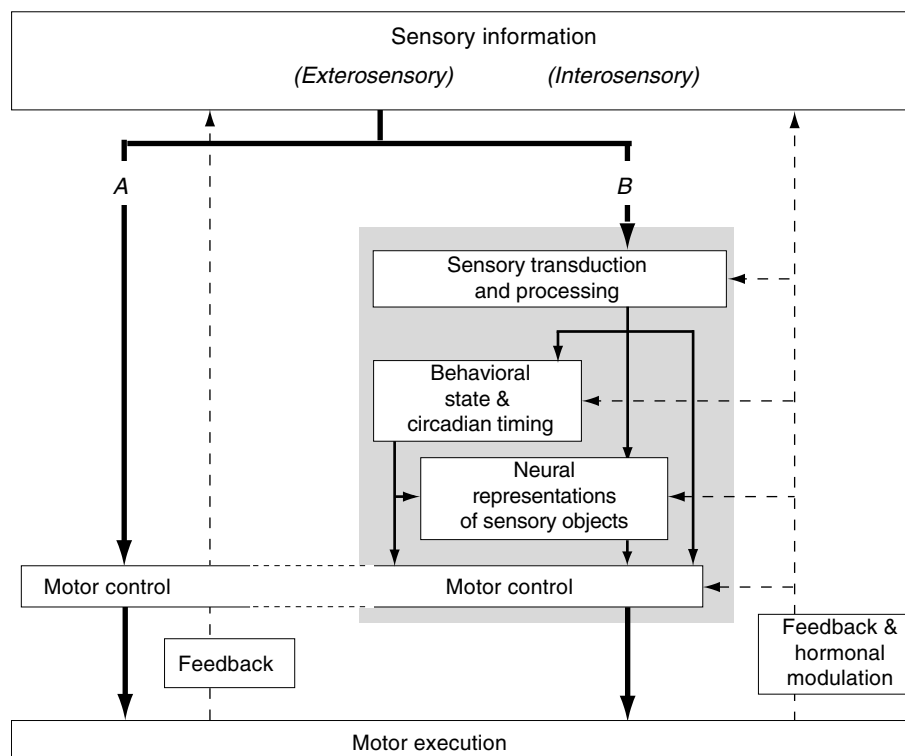


Figure 1 A schematic representation of the two paths along which sensory information is processed in the brain. Path A is a reflex route that does not involve complex central processing and that allows for rapid stereotyped stress responses; path B is a route that allows sensory information to be integrated with other types of information within the brain and that provides stress responses with greater complexity, flexibility, and adaptability. Central neural connections are shown in black, with hormonal and feedback signals shown as dashed lines. The regions that constitute the central control networks in the brain are enclosed within the gray box in path B.

Figure 1 shows that the effects of a stress-associated motor event can serve as feedback signals to modulate the motor output of the system, both by reducing the level of sensory input and motor output (simple negative feedback) and by more indirect effects on central integrative processes.

Spinal cord Neurons in the dorsal horn of the spinal cord are the first recipients of nerve afferents from the skin and viscera that process information related to somatosensory modalities (pressure, temperature, and pain). The spinal cord also receives sensory information from muscle spindles related to body position. Dorsal horn neurons are involved with organizing reflexes related to stress responses, but they are not directly involved with generating the more complex motor events usually thought of as stress responses. However, dorsal horn neurons do project much of this sensory information to the hindbrain and thalamus for the higher-order processing (the sensory transduction and processing part of **Figure 1**, path B). Ultimately, higher-order processing can then influence the type of contextual information responsible for generating the stress responses that are accompanied by fear and anxiety. Clearly, because information related to pain can evoke ACTH release, nociceptive information entering the dorsal horn can affect neuroendocrine motor neurons in the hypothalamic paraventricular nucleus (PVH). However, the pathways responsible for this are not entirely clear.

Hindbrain Parts of the hindbrain receive first-order sensory information from the cranial nerves related to the visual (oculomotor) and vestibular systems, as well as somatosensory (the hypoglossal and trigeminal systems) and gustatory (the glossopharyngeal nerve) information from the oropharynx. In many respects how the hindbrain processes this information in a stress-related manner is similar to how exterosensory information is processed bidirectionally by the dorsal horn of the spinal cord (**Figure 1**) – information can initiate reflex loops and it can be passed further into the brain for more sophisticated processing. Projections from the parabrachial nucleus are probably important for projecting pain-related information received from the dorsal horn and cranial nerves that ultimately impact those parts of the hypothalamus regulating the adrenal gland.

The hindbrain is also critical for processing those sensory modalities originating in the internal environment that can generate stress responses. It receives numerous inputs from sensory receptors in the viscera that in general are conveyed to the nucleus of the solitary tract (NTS) by the vagus nerve and spinal

cord. In turn, the NTS initiates visceral reflexes such as the baroreceptor reflex that regulates cardiovascular function within a normal range. However, neural projections from the NTS along with other hindbrain regions also project this information to the forebrain (including regions in the hypothalamus such as the PVH) and telencephalon that in turn regulate the HPA axis directly. This explains why many viscerosensory stimuli (e.g., hypoglycemia and hypotension) can evoke robust ACTH and corticosterone secretion.

Thalamus With the exception of olfactory information (which enters the cortex directly and most likely impacts stress-related motor circuits by way of the amygdala and bed nuclei of the stria terminalis (BNST), all exterosensory modalities are processed by the thalamus before being projected to the amygdala and cerebral cortex. Consequently, those specific thalamic regions associated with particular sensory modalities have been implicated in processing stress-related sensory information. For example, the medial geniculate region is thought to contribute to the fear-conditioning evoked by pairing an auditory stimulus with foot shock.

Motor Regions

Spinal cord The spinal cord contains two sets of motor neurons that are directly related to expressing stress responses. First, alpha motor neurons in the ventral horn are the mediators of all behavioral motor system stress responses. And second, the intermediolateral column in the thoracic and upper lumbar spinal cord contains the sympathetic preganglionic neurons that innervate the adrenal medulla and are directly responsible for stimulating adrenaline release during stress. These cholinergic neurons also mediate other components of the sympathetic stress responses and are innervated by a variety of afferents including those from the hypothalamus.

Hindbrain Visceral stress responses are also mediated by sets of preganglionic neurons in the hindbrain. For example, stress-induced modifications to cardiovascular and other visceral functions are mediated through the dorsal motor nucleus of the vagus.

Hypothalamus The hypothalamus is a critical premotor region that can modulate most viscerally directed stress responses. It also contains all the motor neurons used for the neuroendocrine component of the stress response. Foremost among these are the

corticotropin releasing hormone (CRH) neurons in the medial parvocellular part of the PVH. The axons of these neurons terminate on the capillaries of hypophysial vasculature in the median eminence and release CRH and arginine vasopressin (AVP) to stimulate ACTH release from corticotropes in the anterior pituitary. The PVH and supraoptic nucleus also contain magnocellular neurons that release oxytocin and AVP during stress events. As with all other motor responses, these neuroendocrine neurons can be activated in relatively simple reflex manner (e.g., by hemodynamic stressors) as well as by inputs from telencephalic regions. In addition, the PVH provides strong projections to sympathetic and parasympathetic preganglionic neurons and is probably a key element in modulating the activity of these systems during stress.

One of the classic functions of the hypothalamus is the generation of motivated behaviors. These include ingestive, reproductive, and defensive behaviors. During the 1940s, Hess demonstrated that electrical stimulation of specific regions of the lateral hypothalamus

(LHA) elicit rage or fearful behavior. It is likely that the often inhibitory influences of stress on, for example, feeding and reproductive behaviors are integrated within the hypothalamus with the relevant control networks of these behaviors. The exact mechanisms used for these interactions are not clear, but again they very likely involve inputs from the telencephalon. Finally, it is important to remember that the hypothalamus contains the suprachiasmatic nucleus (SCH). In mammals, this cell group is the circadian clock, which can strongly influence the way stress responses are generated.

Afferents from those brain regions known to regulate the HPA axis and other stress motor functions are organized as two basic projection patterns (**Figure 2**): those that provide inputs directly to the PVH and those that provide indirect inputs, projecting first to other cell groups that in turn then project to the PVH. Neurons with direct inputs to the PVH can be divided into four afferent sets that provide the rich source of information for integration by neuroendocrine and pre-autonomic cell groups.

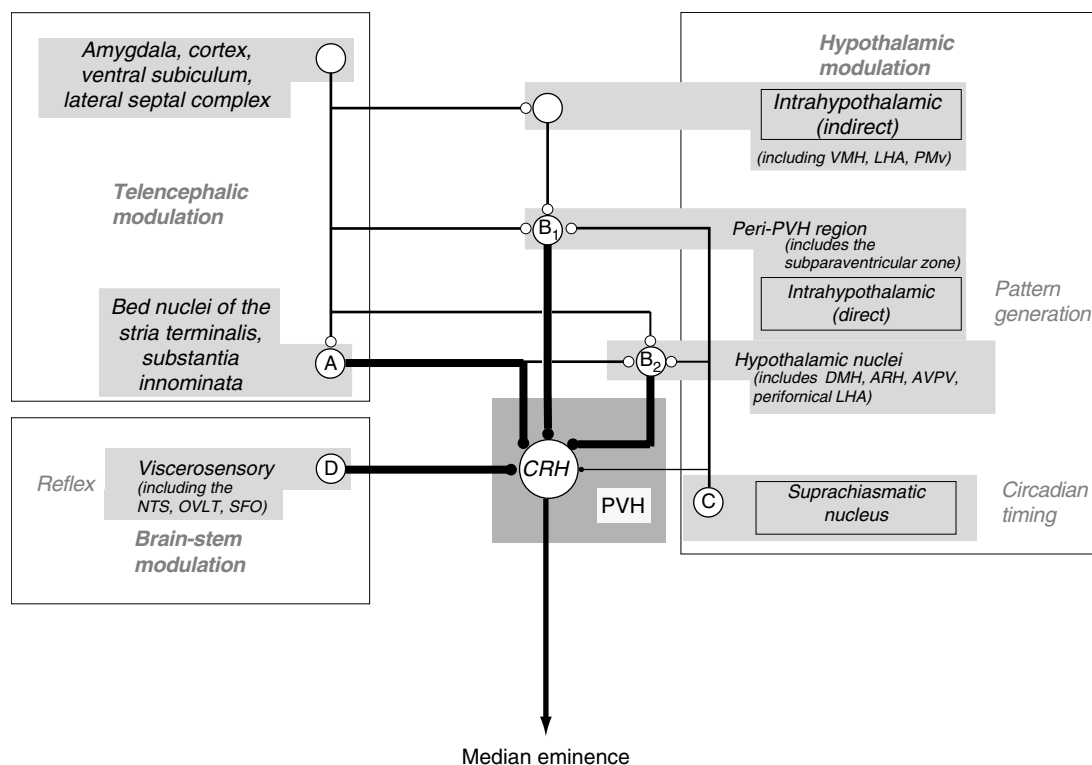


Figure 2 A schematic of the organization of inputs to the paraventricular nucleus of the hypothalamus. This nucleus is a key regulator of neuroendocrine and other autonomic stress responses. However, it receives only a restricted number of direct inputs from the telencephalon, hypothalamus, and hindbrain. Many other hypothalamic and telencephalic regions provide information indirectly by way of these more direct inputs. Finally, circadian timing signals from the suprachiasmatic nucleus impact the PVH directly and, to a greater extent, through other hypothalamic cell nuclei with direct inputs. ARH, arcuate nucleus; AVPV, anteroventral periventricular nucleus; CRH, corticotropin releasing hormone; DMH, dorsomedial nucleus; LHA, lateral hypothalamic area; NTS, nucleus of the solitary tract; OVLT, vascular organ of the lamina terminalis; PMv, ventral premammillary nucleus; PVH, paraventricular nucleus of the hypothalamus; SFO, subfornical organ; VMH, ventromedial nucleus of the hypothalamus.

1. The only direct telencephalic inputs to the PVH comes from some of the BNST and the substantia innominata (Figure 2, group A).

2. Two groups of direct PVH inputs originate intrahypothalamically (Figure 2, group B). First, the PVH is surrounded by local γ -aminobutyric acid (GABA)ergic neurons that are not organized as a distinct hypothalamic nucleus but that electrophysiological and anatomical data suggest form an important group of direct intrahypothalamic inputs (Figure 2, group B₁). They receive a major input from the SCH, from the infralimbic cortex, and from other hypothalamic nuclei such as the ventromedial and ventral premammillary nuclei, which in turn receive massive telencephalic inputs. Second, a well-defined group of projections (Figure 2, group B₂) originate in the arcuate and dorsomedial nuclei, in the preoptic area (in particular, the median preoptic and anteroventral periventricular nuclei), and the LHA. These nuclei also receive inputs from the telencephalon, including the BNST and lateral septal complex (LS). The overall function of these two groups of direct intrahypothalamic inputs is unclear. However, projections from the arcuate and perhaps the dorsomedial nucleus convey information about peripheral energy status derived from circulating leptin, insulin, and ghrelin, which are important for regulating the stress responses derived from changes to energy balance. Some neurons in this group may also constitute a set of neuroendocrine pattern generators similar in principle to those in the hindbrain and spinal cord that control patterning in the behavioral motor system.

The importance of input sets (1) and (2) to stress responses is that they relay information to the PVH from a host of diencephalic and telencephalic regions (Figure 2, groups A, B₁, and B₂). Thus, these areas each receive numerous topographically ordered inputs from the cortex, ventral subiculum (SUBv), amygdala, and LS. Collectively, these inputs are responsible for relaying a great deal of the information critical for controlling stress responses to exterosensory stimuli and under a variety of circumstances that requires complex telencephalic processing.

3. The SCH provides the circadian timing signal that regulates all behaviors and has a critical bearing on the expression of stress responses. Although the SCH provides a small direct PVH projection, its major output (Figure 2, group C) is to the subparaventricular zone in the peri-PVH region and the dorsomedial nucleus. Collectively, its efferents to these regions (Figure 2, group B₁, B₂, and C) are the likely mediators of circadian influence on the PVH-related stress responses.

4. Afferents from a variety of hindbrain nuclei project topographically within the PVH and convey viscerosensory information derived from the spinal cord, vagus, and glossopharyngeal nerves (Figure 2, group B₁). These inputs probably serve two functions; to stimulate patterns of ACTH secretion in response to homeostatic challenges that can generally be described as reflex in nature and to provide the gain control signals required for neuroendocrine CRH neuron function in the basal state. Afferents from the vascular organ of the lamina terminalis and the subfornical organ that provide humoral sensory information about blood osmolality and angiotensin II concentrations, respectively, can also, in terms of function, be included with these direct viscerosensory inputs.

Integrative Regions

Hindbrain The hindbrain contains monoaminergic cell groups that have very wide-ranging connections that in the past have been referred to as the reticular activating system. Included here are the locus ceruleus, the raphe nuclei, parts of the NTS, and pedunculopontine nuclei. Collectively, they receive sensory inputs from the oropharynx and viscera, as well as numerous inputs from the forebrain. They innervate virtually the whole of the neural axis and are thought to regulate behavioral arousal. In some instances (e.g., the locus ceruleus), they are implicated in the stress responses by virtue of their electrophysiological behavior and gene expression. Collectively, however, the precise role of these cell groups in regulating stress responses remains unclear.

Hippocampus It has been known for some time that hippocampal lesions or the severing of hippocampal efferents into the lateral septal complex disrupts ACTH secretion and, thus, glucocorticoid stress responses. The SUBv is the hippocampal subfield most clearly implicated in this context. SUBv lesions increase CRH and AVP gene expression in the neuroendocrine PVH and also elevate plasma ACTH concentrations, suggesting that the SUBv usually provides inhibitory inputs to the PVH. Although specific details about the pathways or the mechanisms involved are not understood, the absence of direct projections from any hippocampal region to the PVH demonstrates that these effects are not mediated directly. The best candidate in this regard is a SUBv-BNST-PVH pathway (Figure 2, group A). The most studied overall function of the hippocampus is in learning and memory, particularly regarding short-term spatial memory and animal navigation. However, how this

type of information relates to the modulation of stress responses is entirely unknown.

Septal complex The septal complex is a collection of forebrain cell groups that lie directly medial to the lateral ventricles. They are characterized by having bidirectional and topographically organized connections with the hippocampus; the LS receives afferents, and the medial septal complex projects back to the hippocampus. As such the LS is in a position to integrate information from all sensory modalities. By virtue of their efferent connections, the parts of the LS that are most closely associated with regulating endocrine stress responses are located in its rostral (LSr) and ventral (LSv) parts. Although the LSv provides sparse inputs directly into the PVH, it does project strongly to at least two regions that can affect stress responses directly: the peri-PVH region (Figure 2, group B₁) and those parts of the hypothalamic preoptic area that then project into the PVH (Figure 2, group B₂). In turn, LSv neurons receive topographically ordered projections from the SUBv. In addition to any effects it might have on the HPA axis, the LS probably plays an important role in the expression of defensive behavior and, as such, may well affect the ordering of behavioral priorities during stress events. Certainly, patterns of immediate early gene expression suggest that the LS is somehow concerned with modulating defensive behavior during fear conditioning.

Amygdala The amygdala has long been considered critical for regulating stress responses and their learned components that have fear as the emotional percept. However, the precise role of amygdala is still unclear with regard to its specific contribution. The amygdala is traditionally subdivided into the lateral, basolateral, medial, central, and cortical regions. However, recent debate has challenged the usefulness of the amygdala as a unitary concept, suggesting that a more viable division is first into cortical (with glutamate as a transmitter) and striatal (with GABA as a transmitter) parts, followed by four functional systems based on the segregation of inputs and outputs. In this scheme, the traditionally defined amygdalar nuclei are included into one of these functional systems. For example, the central nucleus of the amygdala is part of the striatal amygdala and, by virtue of its connections to the BNST and parabrachial nucleus, regulates the autonomic aspects of the stress response. Regarding outputs, the hippocampus and BNST are important components in the output structure of the amygdalar nuclei most closely associated with stress responses. Although again, like other parts of the telencephalon, the complex organization of

stress responses dictates that all parts of the amygdala are probably involved at some level.

Bed nuclei of the stria terminalis The BNST are a set of numerous, often small, and rather poorly differentiated cell groups located immediately rostral and dorsal to the preoptic area of the hypothalamus. Their precise role has proven difficult to define, mostly because experiments have targeted large regions that encompass a number of nuclei. However, neuroanatomical evidence suggests that each BNST cell group has precise and often distinct connections. As a whole, the BNST receive topographically organized projections from the amygdala and hippocampus and, in turn, project to many parts of the hypothalamus. The importance of the BNST to stress responses is that some of these nuclei receive projections from those parts of the cerebral cortex, amygdala and hippocampus (particularly the ventral hippocampal regions) associated with the overall regulation of the stress responses. Consequently, one of their major functions is to integrate and channel processed telencephalic information directly to the PVH and other parts of the hypothalamus (Figure 2, group A).

Cortex Because the cerebral cortex is ultimately involved with formulating and regulating all voluntary behaviors, it must be considered as a major integrative component of stress responses, particularly with respect to behavioral initiation, motor program selection, and prioritization. But as can be imagined, a detailed consideration of its function in this regard is beyond the scope of this article. Regarding cortical influences on neuroendocrine and autonomic motor output, these are generally indirect and are mediated through the LS, BNST, amygdala, and SUBv or through other hypothalamic nuclei that have direct neuroendocrine connections (e.g., Figure 2).

See Also the Following Articles

Amygdala; Comparative Anatomy and Physiology; Hippocampus, Overview; Anatomy of the HPA Axis; Neuroendocrine Systems.

Further Reading

- Blessing, W. (1997). *The lower hindbrain in bodily homeostasis*. New York: Oxford University Press.
- Herman, J. P., Figueiredo, H., Mueller, N. K., et al. (2003). Central mechanisms of stress integration: hierarchical circuitry controlling hypothalamo-pituitary-adrenocortical responsiveness. *Frontiers in Neuroendocrinology* **24**, 151–180.
- LeDoux, J. (1996). *The emotional brain*. New York: Simon & Schuster.

- Swanson, L. W. and Petrovich, G. D. (1998). What is the amygdala? *Trends in Neuroscience* 21, 323–331.
- Swanson, L. W. (2000). Cerebral hemisphere regulation of motivated behavior. *Brain Research* 886, 113–164.
- Swanson, L. W. (1987). The hypothalamus. In: Bjorklund, A., Hökfelt, T. & Swanson, L. W. (eds.) *Handbook of chemical neuroanatomy* (vol. 5), pp. 1–124. Amsterdam: Elsevier.
- Van de Kar, L. D. and Blair, M. L. (1999). Forebrain pathways mediating stress-induced hormone secretion. *Frontiers in Neuroendocrinology* 20, 1–48.
- Watts, A. G. (2005). Glucocorticoid regulation of peptide genes in neuroendocrine CRH neurons: a complexity beyond negative feedback. *Frontiers in Neuroendocrinology* 26, 109–130.

Brain Injury, and Stress *See: Brain Trauma.*

Brain Natriuretic Peptide (BNP)

P Pervanidou and G P Chrousos

University of Athens, Aghia Sophia Children's Hospital, Athens, Greece

© 2007 Elsevier Inc. All rights reserved.

Brain Natriuretic Peptide and Stress Physiology
Brain Natriuretic Peptide and the Exercise-Stressed Myocardium in Healthy Individuals
Brain Natriuretic Peptide and the Failing Heart
Brain Natriuretic Peptide and Myocardial Ischemia
NT-Pro BNP in Relation to Metabolic Risk Factors
Brain Natriuretic Peptide and NT-ProBNP in Pediatrics

Glossary

<i>Amygdala</i>	Brain region involved in the central regulation of the stress response.
<i>Body mass index (BMI)</i>	The weight in kilograms divided by the square of height in meters.
<i>Cardiac myocyte</i>	A cell of myocardial origin.
<i>Dual-energy X-ray absorptiometry (DEXA)</i>	Imaging study assessing body composition and bone mineral density. A useful method for measuring body fat.
<i>Hypothalamic nuclei</i>	Neuron collections in the brain regulating a variety of hormonal secretions and functions.
<i>Prohormone</i>	The initial molecule (peptide in this case) from which the active hormone is derived.
<i>Renin-angiotensin-aldosterone system</i>	Endocrine system that regulates water homeostasis and blood pressure.

Brain natriuretic peptide (BNP) is a hormone of cardiac origin, secreted from the heart ventricles in response to pressure overload. It was isolated originally from porcine brain extracts, but it was soon discovered that heart was its main source. Natriuretic peptides (atrial natriuretic peptide, ANP; BNP; and C-type natriuretic peptide, CNP), deriving from three different genes, comprise a peptide family of high structural homology, as they share a 17-amino acid internal ring. Cardiac ANP and BNP act on the same peripheral receptors, causing diuresis, vasodilatation, and antagonism of the actions of renin and aldosterone. The source of CNP is kidney endothelium and its main action is the regulation of the vascular tone.

BNP is derived from a prohormone, an intracellular 108 amino acid precursor, which is cleaved into two fragments, a 76-amino-acid N-terminal fragment (NT-proBNP) and BNP. Both BNP and the NT-proBNP, have emerged as strong biological markers of cardiac dysfunction: levels of circulating BNP positively correlate with the severity of heart failure (HF) and BNP can improve cardiac disease diagnosis and management. Recombinant human BNP ("Natrecor", Scios Inc) has proved useful in the treatment of some forms of cardiac failure, especially in advance of cardiac surgery. BNP and NT-proBNP have the same diagnostic and prognostic value. Gender, body mass index (BMI) and age (through cardiac and renal function) may influence natriuretic peptide concentrations, but the underlying mechanisms for this are not well known.

Cardiac myocyte stretch, hormones (catecholamines, angiotensin II, and endothelin) and myocardial hypoxia are the main factors causing the observed

changes in the plasma concentrations of these peptides. Stretching of cardiac myocytes is followed by the activation of proBNP gene and *de novo* myocyte peptide synthesis and secretion and there is a positive correlation between the amount of stretch and the resulting increase in BNP and NT-proBNP concentrations. In the normal state, the atrium is the main cardiac production site, but as HF develops, there is a profound activation of ventricular BNP synthesis.

Brain Natriuretic Peptide and Stress Physiology

Besides their peripheral circulation and actions in water and blood pressure regulation, natriuretic peptides and their receptors (NPR-A, NPR-B, and NPR-C), are distributed throughout the central nervous system (CNS) of different species. In the CNS, they represent an important neuromodulatory system, which directly influences the secretion of stress hormones and behavioral manifestations such as arousal and anxiety.

In the CNS of the rat, BNP has a distribution that is different from that of ANP. Furthermore, high-affinity binding sites for BNP and ANP have been found in various brain regions of rats, such as area postrema and certain hypothalamic nuclei. In the CNS of humans and rodents NPR-A is located in the cortex and hippocampus, NPR-B in the amygdala and several brainstem sites, while NPR-C is found in many areas such as the neocortex, limbic cortex, hippocampus and amygdala.

In humans and rodents, ANP has an inhibitory effect on the hypothalamic-pituitary-adrenal (HPA) axis at the pituitary and hypothalamic levels, while CNP induces the opposite effect by stimulating the release of cortisol.

In rats, BNP administration into the cerebral ventricles did not modify the basal plasma corticosterone secretion (which is used as a measure of HPA activity). When rats were subjected to three different stressors (stress, electric shock, and restraint), BNP inhibited stress-induced elevation in plasma corticosterone in a dose-dependent manner. While there is evidence for the inhibitory role of BNP in the HPA axis, the effect of CNP is greater and more stress specific than that of BNP, perhaps due to the different distributions of NPRs in the CNS.

Brain Natriuretic Peptide and the Exercise-Stressed Myocardium in Healthy Individuals

BNP and NT-proBNP are elevated in healthy subjects after prolonged and strenuous exercise. Although these peptides have been established as strong biological

markers of heart disease, the underlying mechanisms that lead to exercise-induced elevation of BNP and NT-proBNP in healthy individuals are not well known.

Myocardial stress, expressed as increased ventricular pressure or cardiac output during exercise, may lead to increased synthesis of BNP and NT-proBNP concentrations in order to lower cardiac afterload by peripheral vasodilatation and diuresis. This relation between exercise and BNP levels seems to be time-dependent.

BNP is a well-known antagonist of the renin-angiotensin-aldosterone system. BNP is elevated in situations of increased central blood volume (fluid overload), such as edematous situations, HF, and exhaustive exercise. Ventricular wall stress is greatly increased in strenuous long-term exercise, as plasma volume increases by up to 22%. BNP is released into the circulation as a counterregulatory homeostatic response, and induces diuresis and vasodilatation and reduces pulmonary and systemic vascular resistance, resulting in reduction of left ventricular (LV) end-diastolic pressure. BNP increases immediately at the onset of, stays elevated during, and decreases with the termination of exercise. BNP is a good marker of volume regulatory response of the hemodynamically stressed myocardium.

Age and cardiac function influence BNP secretion. Thus, older and less trained individuals have higher levels of NT-proBNP. It should be noted that this elevation represents cardiac fatigue, which is different from that resulting from irreversible subclinical cardiac injury.

Brain Natriuretic Peptide and the Failing Heart

The value of cardiac peptides has been recognized by their inclusion in the recent European Guidelines for the diagnosis of chronic HF.

HF is characterized by progressive fatigue and dyspnea and leads to reduced tolerance to efforts. Increased wall stress and neurohormonal (adrenergic, renin-angiotensin-aldosterone, endothelin, and vasopressin) activation stimulate BNP and NT-proBNP secretion and these peptides negatively correlate to LV systolic function.

The majority of studies has proved the relationship between cardiac peptides and cardiac disease: BNP levels are higher in patients with LV hypertrophy and in patients with LV diastolic dysfunction and the specificity of BNP in detecting these situations has been well established. Additionally, BNP correlates with systolic blood pressure (SBP) and pulse pressure (PP) and PP increase has been recognized as an independent risk factor of cardiovascular disease.

BNP and NT-proBNP have become an important addition to the diagnostic tools of heart diseases. They were the first biomarkers to prove their value in screening for LV dysfunction, assessing the prognosis while monitoring patients, decision making during therapy, and predicting adverse cardiac events and readmissions in HF patients.

Brain Natriuretic Peptide and Myocardial Ischemia

BNP concentrations are elevated in individuals with stress-induced myocardial ischemia and the rise in BNP levels are related to the degree of ischemia induced. Additionally, baseline BNP levels before stress testing are higher in those who subsequently develop stress-induced ischemia.

The combination of exercise electrocardiography (ECG) with baseline or peak BNP concentrations increases diagnostic accuracy compared with ECG only. The exact mechanism underlying the association between myocardial ischemia and the release of BNP are not well known. It is possible that the ischemia-induced decrease in LV function and/or ischemia-induced neurohormonal activation may increase ventricular wall stress and BNP secretion.

NT-Pro BNP in Relation to Metabolic Risk Factors

BNP and NT-proBNP levels correlate negatively with BMI and increased expression of clearance receptors by adipose tissue was initially hypothesized to be responsible for the rapid degradation of BNP in obese individuals. In a recent study using dual-energy X-ray absorptiometry (DEXA) instead of BMI as a measure of obesity, lean body mass and not fat mass was responsible for the association between higher BMI and lower BNP and NT-proBNP levels. A substance produced by lean mass that suppresses BNP synthesis or release was then hypothesized to be the underlying cause of this association.

Furthermore, an inverse association between serum NT-proBNP and other risk factors, such as fasting serum insulin, cholesterol and triglycerides and metabolic syndrome itself, has been found independently of age and gender in the adult population. Indeed, NT-proBNP concentrations correlate with PP and the response curve is shifted to the right in adults with the metabolic syndrome.

Brain Natriuretic Peptide and NT-ProBNP in Pediatrics

Reference values of NT-proBNP serum concentrations in umbilical cord and in healthy neonates and

children show a significant increase in the first days of life followed by a rapid decrease during the first year and a gradual decrease throughout infancy. This change may reflect an adaptive mechanism to extra-uterine physiology.

Similarly to adults, NT-proBNP and BNP concentrations are elevated in children with LV dysfunction. Children with congenital or acquired heart diseases, such as dilated cardiomyopathy and myocarditis have elevated NT-proBNP levels, depending on the type and the severity of the disease. Furthermore, NT-proBNP levels are higher in children with HF and are negatively correlated with ejection fraction and positively correlated with New York Heart Association clinical score. There is also a correlation between BNP levels and left-to-right shunt, ventricular volume, and pulmonary-to-systemic pressure ratio in high-pressure lesions.

In conclusion, despite its name brain natriuretic peptide seems to be the least of the three natriuretic peptides involved in CNS regulation of stress. In contrast, it is an established biological marker, extremely useful in the diagnosis and management of acute cardiac incidents as well as the evaluation of exercise-induced myocardial stress in healthy individuals and cardiac patients.

Further Reading

- Das, S. R., Drazner, M. H., Dries, D. L., et al. (2005). Impact of body mass and body composition on circulating levels of natriuretic peptides: results from the Dallas Heart Study. *Circulation* 112, 2163–2168.
- Grandi, A. M., Laurita, E., Selva, E., et al. (2004). Natriuretic peptides as markers of preclinical cardiac disease in obesity. *European Journal of Clinical Investigation* 34, 342–348.
- Hall, C. (2005). NT-ProBNP: the mechanism behind the marker. *Journal of Cardiac Failure* 11, 81–83.
- Horwich, T. B., Fonarow, G. C., Hamilton, M. A., et al. (2001). The relationship between obesity and mortality in patients with heart failure. *Journal of the American College of Cardiology* 38, 789–795.
- Jaszberenyi, M., Bujdoso, E. and Telegdy, G. (2000). Effects of brain natriuretic peptide on pituitary-adrenal activation in rats. *Life Sciences* 66, 1655–1661.
- Lavie, C. J., Osman, A. F., Milani, R. V., et al. (2003). Body composition and prognosis in chronic systolic heart failure: the obesity paradox. *American Journal of Cardiology* 91, 891–894.
- Leers, M. P., Schepers, R. and Baumgarten, R. (2006). Effects of a long distance run on cardiac markers in healthy athletes. *Clinical Chemistry and Laboratory Medicine* 44, 999–1003.
- Maeda, K., Tsutamoto, T., Wada, A., et al. (1998). Plasma brain natriuretic peptide as a biochemical marker of high left ventricular end-diastolic pressure in patients with symptomatic left ventricular dysfunction. *American Heart Journal* 135, 825–832.

- McCord, J., Mundy, B. J., Hudson, M. P., et al. (2004). Relationship between obesity and B-type natriuretic peptide levels. *Archives of Internal Medicine* **164**, 2247–2252.
- Mehra, M. R., Uber, P. A., Park, M. H., et al. (2004). Obesity and suppressed B-type natriuretic peptide levels in heart failure. *Journal of the American College of Cardiology* **43**, 1590–1595.
- Mir, T. S., Flato, M., Falkenberg, J., et al. (2006). Plasma concentrations of N-terminal brain natriuretic peptide in healthy children, adolescents and young adults: effect of age and gender. *Pediatric Cardiology* **27**, 73–77.
- Mosterd, A., Cost, B., Hoes, A. W., et al. (2001). The prognosis of heart failure in the general population: the Rotterdam Study. *European Heart Journal* **22**, 1318–1327.
- Nir, A. and Nasser, N. (2005). Clinical value of NT-ProBNP and BNP in pediatric cardiology. *Journal of Cardiac Failure* **11**, S76–S80.
- Nir, A., Bar-Oz, B., Perles, Z., et al. (2004). N-terminal pro-B-type natriuretic peptide: reference plasma levels from birth to adolescence. Elevated levels at birth and in infants and children with heart diseases. *Acta Paediatrica* **93**, 603–607.
- Olsen, M. H., Hansen, T. W., Christensen, M. K., et al. (2005). N-terminal pro brain natriuretic peptide is inversely related to metabolic cardiovascular risk factors and the metabolic syndrome. *Hypertension* **46**, 660–666.
- Passino, C., Poletti, R., Bramanti, F., et al. (2006). Neurohormonal activation predicts ventilatory response to exercise and functional capacity in patients with heart failure. *European Journal of Heart Failure* **8**, 46–53.
- Passino, C., Severino, S., Poletti, R., et al. (2006). Aerobic training decreases b-type natriuretic peptide expression and adrenergic activation in patients with heart failure. *Journal of the American College of Cardiology* **7**, 1835–1839.
- Sabatine, M. S., Morrow, D. A., de Lemos, J. A., et al. (2004). Acute changes in circulating natriuretic peptide levels in relation to myocardial ischemia. *Journal of the American College of Cardiology* **44**, 1988–1995.
- Sarzani, R., Dessi-Fulgheri, P., Paci, V. M., et al. (1996). Expression of natriuretic peptide receptors in human adipose and other tissues. *Journal of Endocrinological Investigation* **19**, 581–585.
- Scharhag, J., Herrmann, M., Urhausen, A., et al. (2005). Independent elevations of N-terminal pro-brain natriuretic peptide and cardiac troponins in endurance athletes after prolonged strenuous exercise. *American Heart Journal* **150**, 1128–1134.
- Schwachtgen, L., Hermann, M., Georg, T., et al. (2005). Reference values of NT-proBNP serum concentrations in the umbilical cord blood and in healthy neonates and children. *Journal of Cardiology* **94**, 399–404.
- Sengenès, C., Berlan, M., De Glisezinski, I., et al. (2000). Natriuretic peptides: a new lipolytic pathway in human adipocytes. *FASEB Journal* **14**, 1345–1351.
- Silver, M. A., Maisel, A., Yancy, C. W., et al. (2004). BNP Consensus Panel 2004: a clinical approach for the diagnostic, prognostic, screening, treatment monitoring, and therapeutic roles of natriuretic peptides in cardiovascular diseases. *Congestive Heart Failure* **10**, 1–30.
- Staub, M. D., Nusbaumer, C., Zellweger, M. J., et al. (2006). Use of B-type natriuretic peptide in the detection of myocardial ischemia. *American Heart Journal* **151**, 1223–1230.
- Wang, T. J., Larson, M. G., Levy, D., et al. (2004). Impact of obesity on plasma natriuretic peptide levels. *Circulation* **109**, 594–600.
- Wiedemann, K., Jahn, H. and Kellner, M. (2000). Effects of natriuretic peptides upon hypothalamic-pituitary-adrenocortical system activity and anxiety behavior. *Experimental and Clinical Endocrinology & Diabetes* **108**, 5–13.
- Yoshibavashi, M., Kamiva, T., Saito, Y., et al. (1995). Plasma brain natriuretic peptide concentrations in healthy children from birth to adolescence: marked and rapid increase after birth. *European Journal of Endocrinology* **133**, 207–209.

Brain Trauma

B Pentland

Astley Ainslie Hospital, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by B Pentland, volume 1, pp 349–352, © 2000, Elsevier Inc.

Nature and Context of Brain Trauma

Nontraumatic Brain Injuries

Consequences of Brain Injury

Posttraumatic Stress Disorder

Stress in Families

Glossary

Amnesia
Cognitive functions

Metabolism

Loss or impairment of memory.

Mental processes related to thinking, such as memory, orientation, attention, concentration, and perception.

The sum of the chemical processes in the body by which complex substances are

	made and broken down to maintain tissue and organ function.
<i>Posttraumatic amnesia (PTA)</i>	Amnesia relating to events that occurred after the trauma that caused it.
<i>Retrograde amnesia (RA)</i>	Amnesia relating to events that occurred before the trauma that caused it.

Nature and Context of Brain Trauma

In Western countries in peacetime, 1 in 50 of the population will seek medical attention because of head injury each year and about 1 in 5 of these will be admitted to the hospital. Young men ages 15–25 years are particularly liable to suffer such trauma, with the elderly as the next most common group. The human brain lies well protected inside the thick bone of the skull, and it usually takes an external blow of considerable impact to transmit destructive forces to the underlying brain. The most common cause of such injuries is falls, followed by road traffic accidents, assaults, and, less commonly, work-related and sports injuries.

As many as half of those admitted to hospital after brain trauma will have consumed alcohol shortly beforehand. This may cloud the individual's awareness of the otherwise alarming circumstances of the injury. Thus, the young man who is intoxicated may not recall falling from a balcony or being struck by a passing car when the effects of the alcohol wear off. In addition, however, by its very nature brain trauma tends to result in a period of confusion associated with lack of memory, or amnesia. The period of lost memory is divided into that which preceded the trauma, retrograde amnesia (RA), and that which followed it, posttraumatic amnesia (PTA). After injury, the person may not be able to recall events that occurred minutes, hours, days, or months before impact. There is a tendency for the duration of this period of RA to shrink with the passage of time but, in most cases of severe brain trauma, the circumstances immediately preceding the injury are never remembered. PTA is defined as the period of loss of memory between injury and the time the person's confusion resolves and he or she regains continuous memory of day-to-day events. Indeed one of the most useful ways of classifying the severity of brain damage is to measure the duration of PTA. Head injuries are classified as mild, moderate, or severe according to whether the PTA is less than an hour, less than a day, or more than a day, respectively. In some people, the PTA can extend for months.

It might be regarded as a blessing that some victims of serious road accidents, vicious assaults, or terrifying falls have no memory of the event or its

early aftermath and so are spared painful and stressful recollections. Thus, the process of the ambulance journey, emergency room admission, surgery, and intensive care management may all be forgotten. In contrast, if any or all of these events are experienced and remembered they may be perceived as stressful.

Even in a state of altered awareness or reduced consciousness after brain trauma, the individual may be agitated and, to all intents and purposes, anxious and stressed. It is not uncommon to see restless behavior with all the physical accompaniments of a stress reaction during the early postinjury period. Thus, it is incorrect to assume that the absence of memory of stress in the longer term means that stress has not been experienced.

Nontraumatic Brain Injuries

Trauma is not the only form of sudden insult to the brain. Similar acute damage can occur from a number of other causes such as stroke, lack of oxygen, and infection.

- **Stroke:** The term stroke essentially implies an interruption of the blood supply to part of the brain. This may be because of a blockage of the blood flow, resulting in an ischemic stroke, or leakage of blood (i.e., hemorrhage), described as a hemorrhagic stroke. Both come on quickly and usually without warning.

- **Anoxic and metabolic damage:** A number of circumstances can result in the brain being denied oxygen and/or glucose, both of which are essential to normal function, or metabolism. The brain is probably the most sensitive organ in the body in terms of its need for oxygen and glucose. Lack of oxygen is referred to as hypoxia (or anoxia) and may result from cardiac or respiratory arrest, choking, strangulation, drowning, carbon monoxide poisoning, or drug overdose. The commonest cause of lack of glucose (hypoglycemia) is insulin overdose in people with diabetes. Metabolic disturbances other than lack of glucose include sudden disruption to the body chemistry in association with serious disorders of the liver or kidneys. Finally, toxic chemicals may act by interfering with oxygen or glucose supply or poison the brain by another mechanism. For convenience, these uncommon cases are often grouped with hypoxic metabolic damage.

- **Infection:** Infection of the brain (encephalitis or brain abscess) or its surrounding tissues (meningitis) may cause serious disruption to the functioning of the brain. Although this may be short-lived and temporary, some individuals are left with significant residual problems and are regarded as having sustained

permanent brain damage. People whose body defenses are impaired (e.g., those with AIDS or those being treated with immunosuppressant drugs) are particularly liable to these conditions.

Each of these forms of brain injury have characteristics peculiar to them, but, as in the case of trauma, all tend to occur suddenly. Thus, the individuals affected will usually have been functioning normally, going about their everyday lives when they are struck down with a serious event. The subsequent stressful events of hospital admission, intensive care, and monitoring are also similar to trauma cases. In contrast to trauma, these people may recall the onset of symptoms and remember more details of their early treatment.

Clearly in cases of drug overdose, self-hanging, or drowning resulting in hypoxic brain injury, stress may have been a major factor preceding the brain injury and may persist afterward. Although the link is not firmly established, stress may contribute to high blood pressure (hypertension), which is a major risk factor for both ischemic and hemorrhagic stroke.

Consequences of Brain Injury

Irrespective of the exact cause, injury to the brain results in loss of nerve cells, disturbance of connecting pathways within the nervous system, and chemical, including hormonal, changes. Because the brain is essentially the organ that controls all that we think, feel, and do, when it is damaged the effects can be complex. Brain injury may result in impairments of movement (motor function); of sensation (sensory function); of vision, hearing and balance, smell, and taste (special senses); and of heart beat, respiration, and control of bowel and bladder (autonomic function). In addition, how injured people perceive the world and their abilities to remember, concentrate, reason, and judge (cognitive functions) may also be impaired. The injured individual's emotional state may be disturbed, and the attributes that constitute their individual personality, and so how others see them, are also frequently altered.

Often, especially in severe cases, those who survive brain injury present with a complex mixture of physical, cognitive, emotional, and behavioral disorders. Disentangling features of an individual's behavior that may relate to stress from those resulting directly from the damage itself is often difficult or impossible.

People commonly report the phenomenon of frustration after brain damage. There is frustration at the loss or impairment of physical skills but also at cognitive difficulties such as not being able to remember things, find the correct word, or sustain

concentration on a task. The process of rehabilitation involves relearning simple skills, which may assume the form of daunting tasks that the individual may feel incapable of achieving. The physical and psychological impairments limit the activities the person can do successfully, from basic aspects of self-care and domestic chores to managing his or her own finances. The successful and resourceful husband and father may become a dependent member of the family, unable to fulfill these roles. Even those who recover to a greater degree may find both work and leisure pursuits restricted. Loss of employment after brain trauma is common, which in turn may result in financial hardship. Thus, the person may be unable to afford to pay the mortgage or rent, to keep a car, to take a holiday, or to go to a football game. These events and the change of personality that may occur after injury contribute to the high frequency of marital disharmony and breakdown that is well recognized after brain trauma.

It is apparent, therefore, that the individual who survives significant brain injury may experience a number of consequences that are recognized as being major sources of stress in the general population. Unemployment, change or loss of home, divorce, loss of friends, and inability to do normal leisure activities are all common. Added to these may be feelings of guilt, low self-esteem due to disfigurement or disability, and unfavorable reactions from strangers. Those who suffer more minor injuries and lesser degrees of loss may still experience considerable stress at their inability to fulfill tasks previously performed easily.

An additional major source of anxiety and stress in some cases of brain trauma is the legal procedures related to claiming compensation. In general, the process is protracted, involving repeated examinations because of the adversarial nature of most Western legal systems.

Posttraumatic Stress Disorder

Posttraumatic stress disorder (PTSD) is a specific psychiatric diagnosis whose essential feature is the development of characteristic symptoms following an extremely traumatic experience associated with threatened death or serious injury. The individual re-experiences the event in various ways, avoids stimuli associated with the trauma, and shows increased arousal. The symptoms persist for at least a month and are associated with a significant disruption of the individual's lifestyle.

Currently there is a degree of consensus that PTSD is uncommon after brain trauma, but controversy continues as to its actual frequency. Some authorities have

suggested that, because brain injury results in loss of the memory of the injury, PTSD cannot co-exist. Others suggest that it may be more common after milder injuries to the brain, when the loss of memory, both RA and PTA, is shorter. Thus, the individual may recall the circumstances leading to the injury or the stress of emergency treatment, and posttraumatic stressful circumstances are recalled. Finally, it has been suggested that even if the person has no direct memory of the event he or she may develop PTSD as a reaction to the medical sequelae of the injury.

There have been a few reports of PTSD occurring after stroke and other forms of nontraumatic brain injury. Certainly people who recall periods in intensive care units, for whatever cause, may view the experience as life threatening and suffer from disturbing recollections of the circumstances associated with increased arousal. Whether such people strictly fulfill the criteria for PTSD is probably not as important as recognizing the severe stress syndromes that do occur.

Stress in Families

No account of stress associated with traumatic and nontraumatic brain injury would be complete without at least a mention of the effect that an individual's brain injury has on other family members. Individuals who have suffered severe brain damage, in addition to not recalling the event and the early period in hospital, may have little insight into how they have been altered by the experience. Their families, however, will often have spent agonizing days or weeks wondering if their loved ones would survive, followed by anxious weeks or months wondering how much recovery would occur. With the passage of time, they may then face the adjustments to their lives resulting from what is essentially the loss of a spouse, partner, parent, or supportive son or daughter. Contending with changes in roles within a family, coping with behavioral and personality changes, and dealing with the loss of the principal wage earner can impose huge burdens on other family members.

There is increased recognition of, although still inadequate provision for, the effects of stress on carers in these circumstances. Ill health, both physical and mental, is common in the family members of those who survive severe brain injury. Clearly the same major events – loss of income, change of accommodation, loss of friends, and restriction of social life – can affect all family members. Divorce is common. The spouse or partner may have to give up his or her own job or career to become a caregiver or, alternatively, may have to seek employment to maintain an income. Children may also find themselves having to assume the responsibility of caring for a brain-injured parent, disrupting their education or career plans. The spouse, in having to give attention to the injured husband, or parents, in focusing on an injured child, may be unable to provide support to other children in the family, whose mental health and scholastic performance may suffer.

As with other sources of stress, families vary in their ability to cope. There is no direct correlation between the severity of the injury and its consequences and the level of stress experienced. Brain damage that is apparently minor may result in changes in an individual's personality or behavior that other members of the family may find difficult to deal with. There is fortunately a growing awareness of this phenomenon, with professional help available to assist families.

See Also the Following Articles

Amnesia; Brain and Brain Regions; Familial Patterns of Stress; Hypotension, Hypovolemia, and Septic Shock; Posttraumatic Stress Disorder, Neurobiology of.

Further Reading

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th edn., text rev.), Washington, DC: American Psychiatric Association.
Rosenthal, M., Griffith, E. R., Kreutzer, J. S. and Pentland, B. (1999). *Rehabilitation of the adult and child with traumatic brain injury* (3rd edn.). Philadelphia: FA Davis.

Breast Cancer

A Moyer

State University of New York at Stony Brook,
Stony Brook, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
A Moyer, volume 1, pp 353–357, © 2000, Elsevier Inc.

Stresses Associated with Breast Cancer Risk

Stresses Associated with Breast Cancer Diagnosis
and Treatment

The Role of Stress in the Onset and Progression of Cancer

Stress-Reducing Interventions for Women with
Breast Cancer

Glossary

<i>BRCA1, BRCA2</i>	Two genes that, when damaged or mutated, greatly increase the risk of developing breast and/or ovarian cancer; those with a mutation have a 50% chance of passing it on to their children.
<i>Breast-conserving surgery</i>	The removal of a breast tumor and a small margin of normal tissue (also called lumpectomy); radiation therapy is often used in conjunction with breast-conserving surgery.
<i>Breast self-exam</i>	Examining one's breast area visually and by touch for suspicious changes such as lumps or swelling, skin irritation or dimpling, or nipple pain or retraction; most breast cancers are identified by women themselves.
<i>Mammogram</i>	An x-ray of the breast that can identify abnormalities before they can be felt by a woman or her health-care practitioner and before other physical symptoms develop.
<i>Metastasis</i>	The spread of cancer from the site of origin to other regions of the body via the lymph system or the bloodstream.
<i>Recurrence</i>	Cancer that has returned after treatment; breast cancer can recur locally (at the same site), regionally (in the nearby lymph nodes), and distantly (e.g., in the lungs or bone).

Breast cancer is a malignant tumor that develops from the cells of the breast. It is the most common cancer, aside from skin cancer, among women. Although the incidence of breast cancer has been steadily increasing among American women, this rate of

increase began to slow in the 1990s, and mortality rates are now declining. Breast cancer can also occur in men but much more rarely. The likelihood of long-term survival, in part, depends on the extent of spread of the disease. If cancer is localized to the breast, the estimated 5-year survival rate, based on women who were diagnosed in prior years, is 97%; if it has spread to surrounding regions, the rate is 77%; and if it has metastasized to distant sites, the rate is 22%. Due to developments in medical care, the survival estimates for women diagnosed today are presumed to be better.

Stresses Associated with Breast Cancer Risk

All women are at risk for breast cancer. However, particular factors such as a family history of the disease, early menarche or late menopause, and not bearing children or having children after the age of 30 increase the likelihood of developing it. Although more is being discovered about potential predictors, much uncertainty remains. Established risk factors can explain only a small amount of the occurrence of breast cancer, and many women diagnosed with the disease have none of these risk factors. Women concerned about their risk for breast cancer have limited preventive options because factors such as one's family history cannot be altered and recommended modifications in diet, exercise, alcohol consumption, and body weight are no guarantee against contracting the disease.

Not surprisingly, women at higher risk for breast cancer tend to be more distressed. Women with a family history of the disease often overestimate their risk and their concern may be fairly resistant to cancer-risk counseling. Anxiety about their own likelihood of developing breast cancer can be compounded by worry about the health of other affected and unaffected family members.

An important issue is how concern about cancer or being at increased risk affects health behaviors, especially those related to early detection and treatment. Screening procedures such as clinical and breast self-examinations and mammograms can provoke anxiety. Screening mammography has a significant false-positive rate, meaning that many women are subjected to unnecessary worry and the discomfort of diagnostic biopsies. Despite this, having a family history of breast cancer, being at higher risk, having a history of breast problems, and worrying more about breast cancer are associated with a greater

rather than a lower likelihood of seeking a mammogram. However, it has also been documented that anxiety can contribute to delay in seeking medical care once a woman discovers a suspicious breast lump.

A small proportion of breast cancer cases, approximately 5–10%, can be attributed to an inherited susceptibility that greatly increases one's likelihood of getting the disease. A blood test to detect relevant inherited mutations in the genes *BRCA1* and *BRCA2* is now available. However, the stresses associated with obtaining this risk information are complex. In order for results to be interpretable, the testing process often requires extensive documentation of the cancer history in one's family. This can be onerous if there are rifts or a lack of contact among family members. In addition, to identify the particular mutation potentially at work in a family, an affected family member is ideally the first individual tested. This is something that unaffected family members may be reluctant to ask them to do.

Genetic information is especially potent because knowledge about one individual's status will reveal risk information about family members. Relatives may or may not be interested in learning their own risk. Genetic information can also be more complex than most health-related information. In the case of hereditary breast cancer, carrying an identified mutation means that one has only an increased probability of developing the disease. If one is found to carry a mutation, courses of action include more vigilant surveillance, taking the anti-estrogen drug tamoxifen, or prophylactic surgical removal of healthy breasts. Prophylactic surgery is a fairly radical option that is thought to significantly reduce the likelihood of getting breast cancer, but will not eliminate the possibility that it will develop. Thus, decisions about one's options once one has this knowledge can also be a source of turmoil. Although women who test positive and opt for prophylactic surgery have higher distress immediately following the receipt of their test results than those who opt for surveillance or test negative, in the long term, their levels of distress decline, probably due to a reduction in fear of developing cancer. Finally, concerns about employment or insurance discrimination based on the results of genetic tests remain legitimate worries.

Despite these challenges, for some women the advantages of having risk information outweigh the disadvantages, and it can assist them in making decisions about their health care. Early surveys of women at risk for cancer generated concern that those who were more psychologically distressed would be more likely to seek genetic testing. However, research with women actually offered test results have shown that those who opt to learn their status

have more cancer-specific distress but no more global distress than those who opt not to. Learning test results appears to result in decreased distress, especially for noncarriers, but more negative psychological outcomes can occur if test results conflict with one's prior expectations. Long-term follow-up studies of mutation carriers indicate that levels of distress 5 years after learning that one carries a mutation are predicted by factors that existed at pretest such as: worry about cancer, communication difficulties with family members, having young children, having lost a family member to cancer, and greater perception of risk.

Concern about developing breast cancer presents quandaries for women with respect to other health-related decisions. For instance, postmenopausal estrogen replacement therapy can relieve distressing menopausal symptoms and prevent postmenopausal osteoporosis, both important threats to women's health, but it also increases the risk for breast cancer and is no longer considered useful in preventing heart disease. Women are encouraged, with the help of their health-care provider, to determine their own preferred course of action. Women making such decisions about their health care, in general, have been assisted by increased reporting of health-related issues in the media. However, when new reports contradict prior ones, this can be confusing and worrisome.

Stresses Associated with Breast Cancer Diagnosis and Treatment

Breast cancer involves a myriad of stressors. Diagnosis, treatment, recovery, and long-term survival present a series of challenges that can contribute to psychological morbidity. In fact, a small proportion of women who have been treated for breast cancer show symptoms consistent with posttraumatic stress disorder. An initial diagnosis of breast cancer can provoke worries about loss of physical functioning, social value, and financial resources and fears of death. The time between diagnosis and treatment is an extremely stressful period, exacerbated by trying to gather and understand new information.

Treatment for breast cancer usually involves surgical removal of part or all of the breast. Pain, numbness, and arm stiffness and swelling are common. The cultural connection of the breast to women's femininity and sexuality means that surgery for breast cancer can threaten a woman's sense of attractiveness and alter sexual relations. Adjuvant treatment such as radiation therapy, chemotherapy, and hormonal therapy represent considerable physical and psychosocial challenges. Side effects from radiation can include itching, swelling, skin changes, and especially fatigue.

With chemotherapy, there is nausea, hair loss, weight changes, infections, mouth sores, and altered taste sensations. Such physical effects can lead to depression, low self-esteem, family disruptions, sexual difficulties, and fear of losing a partner. Ironically, even the end of treatment can be difficult for patients because the attention from medical personnel is lost and the support of family and friends may diminish. The long-term challenges of living with breast cancer include worry about recurrence and fears about cancer striking family members.

Coping with breast cancer can be especially difficult for certain populations of women. Single women, for instance, can suffer from isolation, inadequate support, and fears about disclosing their disease to new partners. Women may be especially vulnerable to suffering economically when struck with a life-threatening illness. Low-income women who develop breast cancer, or women whose financial situation deteriorates due to their illness, have numerous practical difficulties that compound an already difficult experience.

Of course, breast cancer not only affects the individual dealing with the disease but impacts those in her social network as well. Partners of breast cancer patients often show significant psychological distress as they deal with the demands of added duties and responsibilities in addition to worry about their partner's health. Coping with a life-threatening illness also has numerous ramifications for the family. Breast cancer can disrupt the social, financial, vocational, and educational goals of family members. A key stressor faced by families is dealing with the uncertainty that the illness can bring.

Fortunately, more attention has been devoted to issues of quality of life for women with breast cancer in recent decades. Some of the stressors associated with breast cancer diagnosis and treatment have been alleviated by developments in medical practice and changes in social mores. These include the use of less disfiguring procedures such as breast-conserving surgery, more openness about the once taboo topic of cancer, the availability of more information about the treatment and experience of breast cancer, the opportunity to play a more active role in one's medical treatment, and increased availability of supportive resources. However, some of these developments can also be stressful. For instance, being involved in decisions about treatment can involve assimilating a considerable amount of new information during a stressful period and perhaps feeling more responsibility for the treatment outcome. Fortunately, decision aids to assist physicians in clearly conveying various breast cancer treatment options and eliciting patients' preferences have shown promise in

increasing knowledge and decreasing patients' feelings of conflict about decision-making. Interactive CD-ROMs and audiotapes that patients can use themselves have also been developed.

The Role of Stress in the Onset and Progression of Cancer

The notion that stress might play a role in the occurrence of cancer is not new. The suggestion that stress and grief could predispose women to breast cancer was ventured in the scientific literature in the late 1800s. This idea is also embraced by women in the general public; when queried about the possible causes of breast cancer, stress is a prominently mentioned factor. Some research has focused on the effect of stressful life events, in particular, on the onset of breast cancer. Methodological difficulties in this area, however, include relying on retrospective accounts of stressful life events in women already diagnosed with cancer or a breast abnormality, who may have recall biases due to knowledge of their diagnosis. Some investigators have sought to remedy this problem by making assessments of stress after women have had a diagnostic biopsy for a breast abnormality but before they have learned their results. More methodologically rigorous prospective studies would necessitate following a large number of individuals over an extended period of time to assess the relationship between the occurrence of stressful life events and the subsequent onset of breast cancer. Other methodological problems of studies in this area include lack of comparability of cases and controls on meaningful variables such as age, and the fact that the period during which the occurrence of stressful events is assessed is typically shorter than the several years in which tumors are suspected to develop. A recent meta-analysis by Petticrew examined the most rigorous investigations in this area and concluded that there was no significant relationship between adverse life events such as bereavement and the onset of breast cancer.

Can stress or the reduction of stress play a role in the course of cancer once it is diagnosed? The results of one intervention study have generated much interest and enthusiasm. Women with metastatic breast cancer were assigned either to a supportive group where they were encouraged to express their emotions about cancer and confront their feelings about the disease, which could help alleviate their stress, or to a control group. Although the investigators were only anticipating enhancing quality of life, the group that received the intervention survived an average of almost 18 months longer than the control group. Recent studies of psychosocial interventions,

however, have failed to replicate this survival effect, although beneficial psychosocial outcomes were consistently demonstrated.

One way in which stress might influence the progression of cancer is through alterations in immune functioning. Such changes can result directly from the action of the central nervous system or of stress-elicited hormones, or through changes in behavior such as eating, sleeping, smoking, or adhering to medical care. Women with breast cancer reporting higher levels of stress have shown lowered cellular immune responses, including natural killer cell toxicity and T-cell responses. Cognitive-behavioral stress management intervention, in turn, can improve cellular immune function in breast cancer patients. It is important to determine if such changes could influence tumor growth and metastasis in clinically meaningful ways. For instance, a disrupted diurnal rhythm of the stress hormone cortisol has been associated with a shorter length of survival, mediated through suppressed natural killer cell activity.

Stress-Reducing Interventions for Women with Breast Cancer

Regardless of whether stress reduction strategies will increase the length of survival of women with breast cancer, interventions that help women deal with the challenges of the disease can improve quality of life. A number of supportive interventions for cancer now exist. These include cognitive-behavioral interventions that focus on changing specific thoughts or behaviors or on learning specific coping skills, educational programs that provide information, psychotherapy and counseling by professionals, and support provided by nonprofessionals such as groups of fellow patients. In addition, relaxation and attentional distraction have been used to deal with the nausea and vomiting associated with chemotherapy. Such interventions have proved to be beneficial. A meta-analysis by Meyer and Mark of randomized, controlled-outcome studies concluded that psychosocial interventions for cancer (which have included primarily breast cancer patients) have positive effects on emotional adjustment, functional adjustment, and treatment- and disease-related symptoms. Recently developed interventions have included components such as mindfulness, physical exercise, and written emotional disclosure.

With limited resources for supportive interventions, and the fact that not all cancer patients wish or are able to make use of them, it is important to consider viable alternatives to counseling and support groups. This has brought a focus on means to enhance natural support from those already in an individual's social network and alternatives to face-to-face interactions such as computer-facilitated support systems. The proliferation of self-help books and information available on the Internet related to breast cancer also offer promise.

See Also the Following Articles

Cancer; Cancer Treatment; Estrogen; Genetic Factors and Stress; Metastasis.

Further Reading

- Anderson, B. L., Kiecolt-Glaser, J. K. and Glaser, R. (1994). A biobehavioral model of cancer stress and disease course. *American Psychologist* 49, 389–404.
- Baum, A. and Anderson, B. (eds.) (2001). *Psychosocial interventions for cancer patients*. Washington, DC: American Psychological Association.
- Baum, A., Friedman, A. L. and Zakowski, S. G. (1997). Stress and genetic testing for disease risk. *Health Psychology* 16, 8–19.
- Leucken, L. J. and Compas, B. E. (2002). Stress, coping, and immune function in breast cancer. *Annals of Behavioral Medicine* 24, 336–344.
- McCaul, K. D., Branstetter, A. D., Schroeder, D. M., et al. (1996). What is the relationship between breast cancer risk and mammography screening? A meta-analytic review. *Health Psychology* 15, 423–429.
- Meyer, T. J. and Mark, M. M. (1995). Effects of psychosocial interventions with adult cancer patients: a meta-analysis of randomized experiments. *Health Psychology* 14, 101–108.
- Moyer, A. and Salovey, P. (1996). Psychosocial sequelae of breast cancer and its treatment. *Annals of Behavioral Medicine* 18, 110–125.
- Petticrew, M., Fraser, J. M. and Regan, M. F. (1999). Adverse life events and risk of breast cancer: a meta-analysis. *British Journal of Health Psychology* 4, 1–17.
- Tedder, E. J. (1998). *Understanding and assisting low-income women with cancer*. Binghamton, NY: The Haworth Press.
- Yang, E. V. and Glaser, R. (2003). Stress-induced immunomodulation: implications for tumorigenesis. *Brain, Behavior & Immunity* 17(supplement 1), S37–S40.

Burnout

C Maslach

University of California, Berkeley, CA, USA

M P Leiter

Acadia University, Wolfville, Nova Scotia, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by C Maslach and M P Leiter, volume 1, pp 358–362, © 2000, Elsevier Inc.

Definition and Assessment

A Mediation Model of Burnout and Engagement

Implications for Interventions

Glossary

<i>Burnout</i>	A psychological syndrome of exhaustion, cynicism, and inefficacy, which is experienced in response to chronic job stressors.
<i>Cynicism</i>	A negative, callous, or excessively detached response to various aspects of the job.
<i>Engagement with work</i>	The positive antithesis of burnout, which is characterized by energy, involvement, and efficacy.
<i>Exhaustion</i>	Feelings of being overextended and depleted of one's emotional and physical resources.
<i>Inefficacy</i>	Feelings of incompetence and lack of achievement in work.

Definition and Assessment

Burnout is a psychological syndrome of exhaustion, cynicism, and inefficacy in the workplace. It is considered to be an individual stress experience embedded in a context of complex social relationships, and it involves the person's conception of both self and others on the job. Unlike unidimensional models of stress, this multidimensional model conceptualizes burnout in terms of its three core components.

Burnout Components

Exhaustion refers to feelings of being overextended and depleted of one's emotional and physical resources. Workers feel drained and used up, without any source of replenishment. They lack enough energy to face another day or another person in need. The exhaustion component represents the basic individual stress dimension of burnout.

Cynicism refers to a negative, hostile, or excessively detached response to the job, which often includes a loss of idealism. It usually develops in response to the overload of emotional exhaustion and is self-protective at first – an emotional buffer of detached concern. But the risk is that the detachment can turn into dehumanization. The cynicism component represents the interpersonal dimension of burnout.

Inefficacy refers to a decline in feelings of competence and productivity at work. People experience a growing sense of inadequacy about their ability to do the job well, and this may result in a self-imposed verdict of failure. The inefficacy component represents the self-evaluation dimension of burnout.

What has been distinctive about burnout is the interpersonal framework of the phenomenon. The centrality of relationships at work – whether it be relationships with clients, colleagues, or supervisors – has always been at the heart of descriptions of burnout. These relationships are the source of both emotional strains and rewards, they can be a resource for coping with job stress, and they often bear the brunt of the negative effects of burnout. Thus, if one were to look at burnout out of context and simply focus on the individual exhaustion component, one would lose sight of the phenomenon entirely.

The principal measure of burnout is the Maslach Burnout Inventory (MBI), which provides distinct assessments of each of the three burnout components. Different forms of the MBI have been developed for different types of occupations: the human services survey (MBI-HSS), the educators survey (MBI-ES), and the general survey (MBI-GS). As a result of international interest in burnout research, the MBI has been translated into many languages.

Burnout Correlates

Unlike acute stress reactions, which develop in response to specific critical incidents, burnout is a cumulative stress reaction to ongoing occupational stressors. With burnout, the emphasis has been more on the process of psychological erosion and the psychological and social outcomes of this chronic exposure, rather than just the physical ones. Because burnout is a prolonged response to chronic interpersonal stressors on the job, it tends to be fairly stable over time.

Health Symptoms Of the three burnout components, exhaustion is the closest to an orthodox stress variable, and therefore is more predictive of

stress-related health outcomes than the other two components. Exhaustion is typically correlated with such stress symptoms as headaches, chronic fatigue, gastrointestinal disorders, muscle tension, hypertension, cold/flu episodes, and sleep disturbances. These physiological correlates mirror those found with other indices of prolonged stress. Similarly parallel findings have been found for the link between burnout and various forms of substance abuse.

In terms of mental, as opposed to physical, health, the link with burnout is more complex. It has been assumed that burnout may result in subsequent mental disabilities, and there is some evidence to link burnout with greater anxiety, irritability, and depression. However, an alternative argument is that burnout is itself a form of mental illness, rather than a cause of it. Much of this discussion has focused on depression, and whether or not burnout is a different phenomenon. Research has demonstrated that the two constructs are indeed distinct: burnout is job-related and situation-specific, as opposed to depression, which is general and context-free.

Job Behaviors Burnout has been associated with various forms of job withdrawal – absenteeism, intention to leave the job, and actual turnover. However, for people who stay on the job, burnout leads to lower productivity and effectiveness at work. To the extent that burnout diminishes opportunities for satisfying experiences at work, it is associated with decreased job satisfaction and a reduced commitment to the job or the organization.

People who are experiencing burnout can have a negative impact on their colleagues, both by causing greater personal conflict and by disrupting job tasks. Thus, burnout can be contagious and perpetuate itself through informal interactions on the job. There is also some evidence that burnout has a negative spillover effect on people's home life.

Engagement: The Opposite of Burnout

The opposite of burnout is not a neutral state, but a definite state of mental health within the occupational domain. Engagement with work is a productive and fulfilling state, and is defined in terms of the same three dimensions as burnout, but the positive end of those dimensions rather than the negative. Thus, engagement consists of a state of high energy (rather than exhaustion), strong involvement (rather than cynicism), and a sense of efficacy (rather than inefficacy).

One important implication of the burnout–engagement continuum is that strategies to promote engagement may be just as important for burnout

prevention as strategies to reduce the risk of burnout. A workplace that is designed to support the positive development of the three core qualities of energy, involvement, and effectiveness should be successful in promoting the well-being and productivity of its employees, and thus the health of the entire organization. From this perspective, health is not limited to the physical or emotional well-being of individuals but is evident in enduring patterns of social interactions among people.

A Mediation Model of Burnout and Engagement

Inherent to the fundamental concept of stress is the problematic relationship between the individual and the situation. Thus, prior research has tried to identify both the key personal and job characteristics that put individuals at risk for burnout. In general, far more evidence has been found for the impact of job variables than for personal ones. These job factors fall into six key domains within the workplace: workload, control, reward, community, fairness, and values.

However, more recent theorizing has argued that personal and job characteristics need to be considered jointly within the context of the organizational environment. The degree of fit, or match, between the person and the job within the six areas of work life, will determine the extent to which the person experiences engagement or burnout, which in turn will determine various outcomes, such as personal health, work behaviors, and organizational measures. In other words, the burnout–engagement continuum (with its three dimensions) mediates the impact of the six areas of work life on important personal and organizational outcomes (see [Figure 1](#)).

Job Characteristics: Six Areas of Work Life

An analysis of the research literature on organizational risk factors for burnout has led to the identification of six major domains. Both workload and control are reflected in the demand–control model of job stress, and reward refers to the power of reinforcements to shape behavior. Community captures all of the work on social support and interpersonal conflict, while fairness emerges from the literature on equity and social justice. Finally, the area of values picks up the cognitive-emotional power of job goals and expectations.

Workload Both qualitative and quantitative work overload contribute to burnout by depleting the capacity of people to meet the demands of the job.

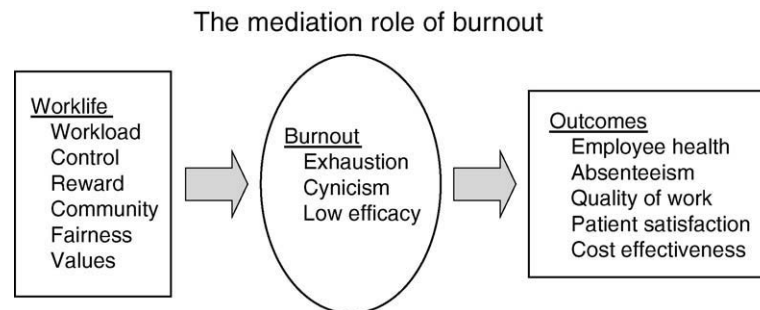


Figure 1 The role of burnout in mediating the impact of work life on personal and organizational outcomes.

When this kind of overload is a chronic job condition, there is little opportunity to rest, recover, and restore balance. A sustainable workload, in contrast, provides opportunities to use and refine existing skills as well as to become effective in new areas of activity.

Control Research has identified a clear link between a lack of control and high levels of stress and burnout. However, when employees have the perceived capacity to influence decisions that affect their work, to exercise professional autonomy, and to gain access to the resources necessary to do an effective job, they are more likely to experience job engagement.

Reward Insufficient recognition and reward (whether financial, institutional, or social) increases people's vulnerability to burnout, because it devalues both the work and the workers, and is closely associated with feelings of inefficacy. In contrast, consistency in the reward dimension between the person and the job means that there are both material rewards and opportunities for intrinsic satisfaction.

Community Community has to do with the ongoing relationships that employees have with other people on the job. When these relationships are characterized by a lack of support and trust and by unresolved conflict, then there is a greater risk of burnout. However, when these job-related relationships are working well, there is a great deal of social support, employees have effective means of working out disagreements, and they are more likely to experience job engagement.

Fairness Fairness is the extent to which decisions at work are perceived as being fair and equitable. People use the quality of the procedures and their own treatment during the decision-making process as an index of their place in the community. Cynicism, anger, and hostility are likely to arise when people feel they are not being treated with the respect that comes from being treated fairly.

Values Values are the ideals and motivations that originally attracted people to their job, and thus they are the motivating connection between the worker and the workplace, which goes beyond the utilitarian exchange of time for money or advancement. When there is a values conflict on the job and thus a gap between individual and organizational values, employees will find themselves making a trade-off between work they want to do and work they have to do, and this can lead to greater burnout.

Personal Characteristics

Although job variables and the organizational context are the prime predictors of burnout and engagement, a few personality variables have shown some consistent correlational patterns. In general, burnout scores are higher for people who have a less hardy personality, who have a more external locus of control, and who score as neurotic on the Five-Factor Model of personality. There is also some evidence that people who exhibit type A behavior (which tends to predict coronary heart disease) are more prone to the exhaustion dimension of burnout.

There are few consistent relationships of burnout with demographic characteristics. Although higher age seems to be associated with lower burnout, it is confounded with both years of experience and with survival bias (i.e., those who survive early job stressors and do not quit). Thus, it is difficult to derive a clear explanation for this age pattern. The only consistent gender difference is a tendency for men to score slightly higher on cynicism. These weak demographic relationships are congruent with the view that the work environment is of greater significance than personal characteristics in the development of burnout.

Implications for Interventions

The personal and organizational costs of burnout have led to the development of various intervention strategies. Some try to treat burnout after it has occurred, while others focus on how to prevent

burnout by promoting engagement. Intervention may occur on the level of the individual, workgroup, or entire organization. At each level, the number of people affected by an intervention and the potential for enduring change increase.

The primary emphasis has been on individual strategies to prevent burnout, rather than social or organizational ones. This is particularly paradoxical given that research has found that situational and organizational factors play a bigger role in burnout than individual ones. Also, individual strategies are relatively ineffective in the workplace, where the person has much less control of stressors than in other domains of his or her life. There are both philosophical and pragmatic reasons underlying the predominant focus on the individual, including notions of individual causality and responsibility and the assumption that it is easier and cheaper to change people instead of organizations.

Recently, the range of options for intervention has expanded, given the recognition of (1) aiming for the positive goal of promoting engagement and (2) using the six areas of work life to better identify the critical intervention targets. These options have now been incorporated into both individual and organizational strategies, all of which focus on the job context and its impact on the people who work within it.

See Also the Following Articles

Anxiety; Depression and Manic-Depressive Illness; Workplace Stress.

Further Reading

- Bakker, A. B., Schaufeli, W. B., Sixma, H. J., Bosveld, W. and Van Dierendonck, D. (2000). Patient demands, lack of reciprocity, and burnout: a five-year longitudinal study among general practitioners. *Journal of Organizational Behavior* 21, 425–441.
- Burke, R. J. and Greenglass, E. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education* 16, 239–253.
- Cherniss, C. (1995). *Beyond burnout*. New York: Routledge.
- Cordes, C. L. and Dougherty, T. W. (1993). A review and integration of research on job burnout. *Academy of Management Review* 18, 621–656.
- Leiter, M. P. and Maslach, C. (2000). *Preventing burnout and building engagement: a complete program for organizational renewal*. San Francisco, CA: Jossey Bass.
- Leiter, M. P. and Maslach, C. (2005). *Banishing burnout: six strategies for improving your relationship with work*. San Francisco, CA: Jossey Bass.
- Maslach, C. and Leiter, M. P. (1997). *The truth about burnout*. San Francisco, CA: Jossey-Bass.
- Maslach, C., Jackson, S. E. and Leiter, M. P. (1996). *Maslach burnout inventory manual* (3rd ed.). Palo Alto, CA: Consulting Psychologists Press.
- Maslach, C., Schaufeli, W. B. and Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology* 52, 397–422.
- Schaufeli, W. and Enzmann, D. (1998). *The burnout companion for research and practice*. Washington, D.C.: Taylor & Francis.
- Schaufeli, W., Maslach, C. and Marek, T. (eds.) (1993). *Professional burnout: recent developments in theory and research*. Washington, D.C.: Taylor & Francis.
- Shirom, A. (2003). Job-related burnout: a review. In: Quick, J. C. & Tetrick, L. E. (eds.) *Handbook of occupational health psychology*, pp. 245–265. Washington, D.C.: APA.
- Toppinen-Tanner, S., Kalimo, R. and Mutanen, P. (2002). The process of burnout in white-collar and blue-collar jobs: eight-year prospective study of exhaustion. *Journal of Organizational Behavior* 23, 555–570.
- van Dierendonck, D., Schaufeli, W. B. and Buunk, B. P. (1998). The evaluation of an individual burnout intervention program: the role of inequity and social support. *Journal of Applied Psychology* 83, 392–407.

C

Calbindin

F A Antoni

University of Edinburgh, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by F A Antoni, volume 1, pp 363, © 2000, Elsevier Inc.

Glossary

EF-hand proteins

A family of diverse proteins that bind Ca^{2+} with high affinity. The EF-hand motif refers to α -helical parts of these proteins that form the Ca^{2+} binding site(s).

Calbindin is one of several EF-hand proteins found in the central nervous system with no clearly established physiological function.

The interest in this protein stems from the fact that it is selectively expressed in certain groups of hippocampal and cerebral cortical nerve cells, which appear to be relatively resistant to Ca^{2+} -mediated neurotoxicity. Furthermore, the levels of calbindin in the brain as well as other tissues are regulated by steroid hormones, including glucocorticoids and estrogens.

Ca^{2+} -mediated toxicity has been proposed to be important for the neurotoxic effects of glucocorticoids. According to the hypothesis put forward by R. M. Sapolsky, corticosteroids make nerve cells vulnerable to increases of intracellular free Ca^{2+} induced by excitatory amino acids. This is particularly evident in the CA1 area of the hippocampus. Sapolsky and coworkers have also postulated that this area of the hippocampus is important for the termination of the stress-induced increase of plasma cortisol. Thus, if CA1 neurons are damaged by high levels of plasma corticosteroids, this may lead to a vicious positive feedback cycle progressively destroying the

hippocampal target neurons of glucocorticoids and thus leading to a decline in cognitive function due to chronic stress.

In addition to the buffering of intracellular Ca^{2+} , overexpression of calbindin also reduces plasma membrane Ca^{2+} currents. Collectively, these actions would have beneficial effects in combating Ca^{2+} -mediated toxicity. Nonetheless, a lot more needs to be learned about the cellular functions of calbindin.

Further Reading

- Baimbridge, K. G., Celio, M. R. and Rogers, J. H. (1992). Calcium-binding proteins in the nervous-system. *Trends in Neuroscience* **15**, 303–308.
- Guo, Q., Christakos, S., Robinson, N., et al. (1998). Calbindin D28k blocks the proapoptotic actions of mutant presenilin 1: reduced oxidative stress and preserved mitochondrial function. *Proceedings of the National Academy of Sciences USA* **95**, 3227–3232.
- Krugers, H. J., Koolhaas, J. M., Medema, R. M., et al. (1996). Prolonged subordination stress increases calbindin-D28k immunoreactivity in the rat hippocampal CA1 area. *Brain Research* **729**, 289–293.
- Landfield, P. W., McEwen, B. S., Sapolsky, R. M., et al. (1996). Hippocampal cell death. *Science* **272**, 1249–1251.
- Lledo, P. M., Somasundaram, B., Morton, A. J., et al. (1992). Stable transfection of calbindin-D28k into the GH3 cell line alters calcium currents and intracellular calcium homeostasis. *Neuron* **9**, 943–954.
- Meier, T. J., Ho, D. Y., Park, T. S., et al. (1998). Gene transfer of calbindin D-28k cDNA via herpes simplex virus amplicon vector decreases cytoplasmic calcium ion response and enhances neuronal survival following glutamatergic challenge but not following cyanide. *Journal of Neurochemistry* **71**, 1013–1023.
- Sapolsky, R. M. (1992). *Stress, the aging brain, and the mechanisms of neuron death*. Cambridge, MA: MIT Press.

Calcium, Role of

F A Antoni

University of Edinburgh, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by F A Antoni, volume 1, pp 366–367, © 2000, Elsevier Inc.

Calcium in the Extracellular Fluid
Intracellular Calcium

Glossary

<i>Heterotrimeric G-proteins</i>	Proteins that consist of three subunits (α , β , γ) and couple cell surface receptors to their effector enzymes. On the activation of cell surface receptors by their ligands, heterotrimeric G-proteins dissociate into the α subunit and the $\beta\gamma$ complex, both of which regulate the activity of effector enzymes such as adenylyl cyclase in the cell membrane.
<i>Second messengers</i>	Chemicals that are generated by extracellular stimuli arriving at the cell surface and that transduce the extracellular signal toward the interior of the cell.

Calcium ions are ubiquitous second messengers in biological systems.

Calcium in the Extracellular Fluid

Calcium is found in several compartments of the body. Extracellular Ca^{2+} is in an ionized as well as a protein-complexed form. There are G-protein-coupled cell surface receptors in several cells of the body that respond to ionized Ca^{2+} concentrations above 1.5 mM. These receptors activate the polyphosphoinositide cycle through the coupling protein G_q and may inhibit adenylyl cyclase through G_i or G_o . Apart from the parathyroid gland, where these receptors control the secretion of parathyroid hormone, the role of the G-protein-coupled Ca^{2+} sensor remains to be explored. It is, however, clearly expressed in the brain as well as the adenohypophysis, where corticotropes appear to express this receptor.

Intracellular Calcium

Inside cells, Ca^{2+} is found in several pools (e.g., smooth endoplasmic reticulum and mitochondria) maintained by various transporter proteins and intracellular ion channels. Intracellular free Ca^{2+} is

fundamental for a multitude of regulatory processes, including intracellular signal transduction.

The action of proteins induced by glucocorticoids in hippocampal neurons and anterior pituitary corticotrophs appears to require intracellular Ca^{2+} . Glucocorticoids augment Ca^{2+} influx through voltage-operated ion channels in the CA1 hippocampal neurons and increase the level of L-type Ca^{2+} channel subunit mRNA in this system. In pituitary corticotroph cells, blockade of the Ca^{2+} -activated protein phosphatase calcineurin markedly reduces the efficiency of glucocorticoid feedback inhibition of adrenocorticotrophic hormone release.

Glucocorticoids enhance the vulnerability of hippocampal neurons to Ca^{2+} -induced neurotoxicity. Thus, prolonged hypercortisolemia, such as that seen in chronic stress, Cushing's disease, or depression, may lead to neuronal loss in the hippocampus. Further, this mechanism may also contribute to the decline of mental function during aging because plasma cortisol levels are known to rise with age. This has led to the glucocorticoid cascade and to the calcium hypotheses of aging.

See Also the Following Articles

Adenylyl Cyclases and Stress Response; Calbindin.

Further Reading

- Antoni, F. A. (1996). Calcium checks cyclic AMP – mechanism of corticosteroid feedback in adenohypophyseal corticotrophs. *Journal of Neuroendocrinology* 8, 659–672.
- Berridge, M. J. (1998). Neuronal calcium signaling. *Neuron* 21, 13–26.
- Brown, E. M., Chattopadhyay, N., Vassilev, P. M., et al. (1998). The calcium-sensing receptor (CaR) permits Ca^{2+} to function as a versatile extracellular first messenger. *Recent Progress in Hormone Research* 53, 257–280.
- Kerr, D. S., Campbell, L. W., Thibault, O., et al. (1992). Hippocampal glucocorticoid receptor activation enhances voltage-dependent Ca^{2+} conductances: relevance to brain aging. *Proceedings of the National Academy of Sciences USA* 89, 8527–8531.
- Landfield, P. W., McEwen, B. S., Sapolsky, R. M., et al. (1996). Hippocampal cell death. *Science* 272, 1249–1251.
- Landfield, P. W., Thibault, O., Mazzanti, M. L., et al. (1992). Mechanisms of neuronal death in brain aging and Alzheimers disease – role of endocrine-mediated calcium dyshomeostasis. *Journal of Neurobiology* 23, 1247–1260.
- Nair, S. M., Werkman, T. R., Craig, J., et al. (1998). Corticosteroid regulation of ion channel conductances and

mRNA levels in individual hippocampal CA1 neurons. *Journal of Neuroscience* 18, 2685–2696.

Sapolsky, R. M. (1994). The physiological relevance of glucocorticoid endangerment of the hippocampus. *Annals of the New York Academy of Sciences* 746, 294–307.

Smith Swintosky, V. L., Pettigrew, L. C., Sapolsky, R. M., et al. (1996). Metyrapone, an inhibitor of glucocorticoid production, reduces brain injury induced by focal and global ischemia and seizures. *Journal of Cerebral Blood Flow & Metabolism* 16, 585–598.

Calcium-Dependent Neurotoxicity

J S Kelly

University of Edinburgh, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J S Kelly, volume 1, pp 364–365,

© 2000, Elsevier Inc.

Excitatory Amino Acid Neurotoxicity
Dysregulation of Calcium Homeostasis
Calcium-Dependent Glial Toxicity
Neuroprotectants

Glossary

Ischemia and/or anoxia

The brain is highly vulnerable to disturbances of the blood supply. Interruption of the blood supply (ischemia) or the delivery of blood containing an inadequate amount of oxygen (anoxia) leads to cell death as an area of the brain is deprived of oxygen, glucose, and other nutrients and the build up of carbon dioxide, lactic acid, free radicals, and other metabolites. Within a few minutes, the most severely damaged cells die by necrosis, and the area of damage is said to be infarcted. This is the process that underlies a stroke when the blood flow to the cerebral cortex and striatum is compromised by reduced blood flow in the middle cerebral artery.

Necrosis

An explosive event caused by exposure to a catastrophic toxic event characterized by almost instantaneous membrane lysis and the release of the intracellular contents into the extracellular space followed by an essentially inflammatory process. In ischemic tissue, glucose and glycogen are depleted within minutes, which leads to a fall in phosphocreatine levels and ATP with an inevitable failure of energy-dependent ion transporters

Apoptosis or programmed cell death

and a loss of ion homeostasis. The passive entry of water into the energy-depleted cell causes intense swelling and osmolytic damage that is sufficient to destroy the cytoplasmic organelles and the nuclear membranes.

In contrast to necrosis, apoptotic cell death is an active and controlled process characterized morphologically by the maintenance of membrane integrity until very late in the death process. The cell membrane shows characteristic blebbing as the cell volume decreases. The nuclear proteins condense to form an easily recognized pattern, the intracellular organelles are well preserved, and the whole still-intact cell is engulfed by phagocytes. The packaging and removal process serves to minimize inflammation. However, under conditions in which the phagocytic step is omitted, cells undergo secondary necrosis and rupture, spilling their contents into the surrounding tissue space.

Free radicals

When the blood supply or oxygen supply is compromised, toxic chemical species are formed that cause cell death either directly or by a process known as lipid peroxidation. The free radicals include a reduction of molecular oxygen to the superoxide anion (O_2^-), the formation of hydrogen peroxide (H_2O_2), hydroperoxy radicals (HOO^*), hydroxyl radicals (OH^*), and singlet oxygen (1O_2) species. Reactive metabolites and free radicals formed under anoxic or ischemic conditions set off a cascade of membrane destruction that also releases cytotoxic species known as lipid peroxy radicals (ROO^*) and hyperoxides ($ROOH$). In normally perfused and oxygenated tissue, lipid peroxides and other reactive species are usually dealt with effectively by specific enzymes that require the presence of vitamin E and a variety of other compounds present in a normal diet known as antioxidants.

Lipid peroxidation

For at least a decade, a rise in intracellular Ca^{2+} concentration ($[\text{Ca}^{2+}]_i$) has been considered one of the most important factors in the final common pathway that leads to cell death, whether by necrosis or apoptosis. Direct activation of calcium-dependent lytic enzymes including proteases, lipases, and nucleases, lipid peroxidation, damage to DNA, and rupture of mitochondria and cell membranes destroy cell structure. Accumulation of calcium phosphate within mitochondria disrupts these organelles, leading to impaired ATP production and, indirectly, further disruption of Na^+ exchangers. Disruption of mitochondria may also trigger apoptotic programs that lead to the death of adjacent groups of cells. The area of damage may be enlarged by the activation of nitric oxide synthetase and free radical production. Thus, calcium-dependent excitotoxicity plays a key role in many acute diseases of the central nervous system (CNS), including ischemic stroke and head injury, and may well be responsible for neuronal death following convulsions and the slow death of neurons in chronic diseases such as Parkinson's disease and motoneuron disease. Indeed, a similar mechanism may be triggered in Alzheimer's by the accumulation of extracellular amyloid.

Excitatory Amino Acid Neurotoxicity

A common cause of calcium-mediated cell death is the excessive release of excitatory amino acids. During normal synaptic activity, the local $[\text{Ca}^{2+}]_i$ is assumed to increase transiently following the episodic activation of excitatory amino acid receptors during synaptic activity. Although in the past this was thought to occur principally due to the entry of Ca^{2+} through *N*-methyl-D-aspartate (NMDA) receptor/ion channels, Ca^{2+} entry is now known to occur through a number of the AMPA (α -amino-3-hydroxy-5-methyl-isoxazole) receptor subtypes and kainate receptors. Recently, relatively modest synaptic stimulation has been shown to lead to local increases in $[\text{Ca}^{2+}]_i$ following the activation of metabotropic receptors and the subsequent release of Ca^{2+} from the endoplasmic reticulum. Following the massive release of glutamate induced by ischemia and/or anoxia, the overactivation of these same excitatory amino acid receptor pathways leads to increase in $[\text{Ca}^{2+}]_i$ enhanced by the opening of voltage-activated Ca^{2+} channels. The substantial depolarization that accompanies the excessive activation of the excitatory amino acid receptors also leads to excessive levels of $[\text{Na}^+]_i$. In turn, high $[\text{Na}^+]_i$ impairs the sodium-calcium exchanger, leading to further increases in $[\text{Ca}^{2+}]_i$. Given the abundance of excitatory amino acids in the CNS, it is easy to imagine that the release of glutamic acid from the cytosol of dead or degenerating

neurons will lead to the death of adjacent tissues. Additional factors may affect the ability of the neurons to resist the rise in $[\text{Ca}^{2+}]_i$, such as aging and the consequent impairment of DNA repair mechanisms, energetic impairment, environmental toxins, or viral infections.

Dysregulation of Calcium Homeostasis

In order for localized rises in intracellular calcium concentrations to efficiently activate enzymes and nearby ion channels in an ordered way, the overall calcium concentration must be kept at a low level of about 100 nM. Thus, the levels and the location of Ca^{2+} are tightly controlled by a complex interplay of inward and outward fluxes, buffering, and internal storage. However, during anoxia and excitotoxicity, these mechanisms either cease to operate or are overwhelmed, and $[\text{Ca}^{2+}]_i$ rises to a level that triggers inappropriate enzymes either directly or through the formation of free radicals and other toxic species. Although Ca^{2+} overloading is unlikely to be the sole mechanism mediating neuronal cell death, several lines of evidence link Ca^{2+} influx and neuronal injury. However, recent work suggests that the route of entry is also important and that Ca^{2+} entry following exposure to glutamate is a more significant cause of cell death than Ca^{2+} entry through voltage-sensitive Ca^{2+} channels. Thus, the possibility exists that the enzymes and/or substrates of the initiators of cell death may be colocalized with NMDA receptors. For instance, there is work to show that suppressing the protein PSD-95 selectively attenuated Ca^{2+} -activated nitric oxide (NO) production and specifically reduced NMDA-mediated neurotoxicity. Thus, targeting PSD-95 could have therapeutic potential.

The failure of several clinical trials to demonstrate a protective role for glutamate receptor antagonists has led to speculation about alternative mechanisms that may promote neurotoxic Ca^{2+} influx during anoxia, perhaps in the absence of glutamate receptor activation. These include the reversal of the Na^+ - Ca^{2+} exchanger, the activation of ATP gated ion channels by a rise in extracellular ADP levels due to increased hydrolysis of ATP, the uptake of Ca^{2+} by mitochondria leading to the release of superoxide free radicals, and, more recently, the activation of acid-sensitive Ca^{2+} permeable ion channels.

Calcium-Dependent Glial Toxicity

Although it has been known for some time that glutamate transporters on astrocytes are responsible for most of the uptake of glutamate from the extracellular space, it is only comparatively recently that the importance of ionotropic glutamate receptors

on astrocytes and oligodendrocytes has been realized. Prolonged activation of these receptors leads to cell death, and the vulnerability of glial cells appears to be equal to that of neurons. Interestingly, glutamate toxicity on oligodendrocytes is in part mediated by a reduction in cysteine uptake leading to glutathione depletion and is blocked by antioxidants. However, the bulk of toxicity is mediated as in neurons by a disruption of $[Ca^{2+}]_i$ homeostasis. Thus, *in vitro*, permanent middle cerebral artery occlusion has been shown to cause massive oligodendrocyte cell death, and in some brain areas the glial cells appear to be more vulnerable than neurons. However, within days there is an increase in the number of oligodendrocytes in the tissue bordering the infarct. It is hoped that the appearance of these new cells is beneficial, if not a harbinger of functional recovery.

Neuroprotectants

Therapeutic attempts to attenuate the damage caused by a rise in $[Ca^{2+}]_i$ focused on excitatory amino acid receptor blockade have failed. In the main, Ca^{2+} channel antagonists have proved disappointing. However, drugs such as riluzole, lamotrigine, and gabapentin are thought to limit glutamate release and/or block Na^+ channels. In more chronic situations, antioxidants may prove useful. Other studies have suggested that neurotrophic factors such as glial-derived growth factor, neurotrophin-3, cardiotrophin, and insulin-like growth factor and immunophilin antagonists may all be neuroprotective. Clearly, the possibility that $[Ca^{2+}]_i$ can also trigger apoptosis may lead to the discovery of new treatments. Although targeting the caspases, the final mediators of cell death, has proved ineffective, intervention at an earlier stage in the apoptotic pathway, at the level of the mitochondria, may promote neuronal and glial cell survival.

See Also the Following Articles

Apoptosis; Excitatory Amino Acids; Cardiovascular Disease, Stress and.

Further Reading

Arundine, M. and Tymianski, M. (2003). Molecular mechanisms of calcium-dependent neurodegeneration in excitotoxicity. *Cell Calcium* 34, 325–337.

- Beal, M. F., Hyman, B. T. and Koroshetz, W. (1993). Do defects in mitochondrial energy metabolism underlie the pathology of neurodegenerative diseases? *Trends in Neurosciences* 16, 125–131.
- Cassarino, D. S. and Bennett, Jr., J. P. (1999). An evaluation of the role of mitochondria in neurodegenerative diseases: mitochondrial mutations and oxidative pathology, protective nuclear responses, and cell death in neurodegeneration. *Brain Research Reviews* 29, 1–25.
- Kinloch, R. A., Treherne, J. M., Furness, L. M., et al. (1999). The pharmacology of apoptosis. *Trends in Pharmacological Sciences* 20, 35–42.
- Lipton, S. A. and Rosenberg, P. A. (1994). Excitatory amino-acids as a final common pathway for neurologic disorders. *New England Journal of Medicine* 330, 613–622.
- Kristián, T., Ouyang, Y. and Siesjö, B. K. (1996). Calcium-induced neuronal cell death in vivo and in vitro: are the pathophysiologic mechanisms different? In: Siesjö, B. K. & Wieloch, T. (eds.) *Molecular mechanisms of ischemic brain damage*, pp. 107–113. Philadelphia, PA: Lippincott-Raven.
- Louvel, E., Hugon, J. and Doble, A. (1997). Therapeutic advances in amyotrophic lateral sclerosis. *Trends in Pharmacological Sciences* 18, 196–203.
- Mattson, M. P. and Chan, S. L. (2003). Calcium orchestrates apoptosis. *Nature Cell Biology* 5, 1041–1043.
- Mattson, M. P. and Magnus, T. (2006). Ageing and neuronal vulnerability. *Nature Reviews in Neuroscience* 7, 278–294.
- Mattson, M. P., Guthrie, P. B. and Kater, S. B. (1989). A role for Na^+ -dependent Ca^{2+} extrusion in protection against neuronal excitotoxicity. *FASEB Journal* 3, 2519–2526.
- Matute, C., Alberdi, E., Ibarretxe, G. and Sanchez-Gomez, M. V. (2002). Excitotoxicity in glial cells. *European Journal of Pharmacology* 447, 239–246.
- Olney, J. W. (1978). Neurotoxicity of excitatory amino acids. In: McGreer, E. G., Olney, J. W. & McGreer, P. L. (eds.) *Kainate acid as a tool in neurobiology*, pp. 95–121. New York: Raven Press.
- Rang, H. P., Dale, M. M., Ritter, J. M., et al. (2003). Chemical transmission and drug action in the central nervous system. In: Rang, H. P., Dale, M. M., Ritter, J. M., et al. (eds.) *Pharmacology*, pp. 464–469. Edinburgh, UK: Churchill Livingstone.
- Rang, H. P., Dale, M. M., Ritter, J. M., et al. (2003). Amino acid transmitters. In: Rang, H. P., Dale, M. M., Ritter, J. M., et al. (eds.) *Pharmacology*, pp. 470–482. Edinburgh, UK: Churchill Livingstone.
- Snyder, S. H., Sabatini, D. M., Lai, M. M., et al. (1998). Neural actions of immunophilin ligands. *Trends in Pharmacological Sciences* 19, 21–26.

Cancer

D Spiegel

Stanford University, Stanford, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D Spiegel, volume 1, pp 368–375, © 2000, Elsevier Inc.

Mind–Body Interactions and Cancer
 Stress and Cancer Incidence and Progression
 Coping Style and Cancer Incidence and Progression
 Social Support as a Stress Buffer
 Physiological Mechanisms of Stress and Support-Related Effects on Cancer Progression
 Stress Reactivity and Disease Progression
 Circadian Rhythm Disruption, Stress, and Cancer
 Psychoendocrinology and Cancer
 Psychoneuroimmunology and Cancer
 Psychosocial Interventions and Immune Function
 Immune Function and Disease Progression
 Stress and Autonomic Responses
 Conclusion

Glossary

<i>Allostatic load</i>	The physiological burden undertaken by the body to maintain allostasis, or internal stability, through environmental change evoked by a series of stressors.
<i>Glucocorticoid</i>	A class of stress hormones that mobilize glucose into the blood.
<i>Hypothalamic-pituitary-adrenal (HPA) axis</i>	One of the major stress response systems in the body; designed to maintain adequate glucose in the blood to help the brain and muscles respond to stressors.
<i>Psycho-neuroimmune connections</i>	The relationship among the mind, brain, and components of the immune system.
<i>Psychosocial factors</i>	Psychological and social factors that affect cancer patients or may be a component of supportive care.

Mind–Body Interactions and Cancer

Cancer research has productively focused on the pathophysiology of the disease, emphasizing aspects of tumor biology as predictors of disease outcome at the expense of studying the role of the body's psychophysiological reactions to tumor invasion. These reactions are mediated by brain–body mechanisms, including the endocrine, neuroimmune, and autonomic

nervous systems. Although a large portion of the variance in any disease outcome is accounted for by the specific local pathophysiology of that disease, some variability must also be explained by host-resistance factors, which include the manner of response to the stress of the illness.

That the stress of cancer is felt psychologically is indicated by the fact that as many as 80% of breast cancer patients report significant distress during initial treatment. Estimates of the prevalence of psychiatric disorders among newly diagnosed cancer patients has ranged from 30 to 44%, and 20–45% of patients exhibit emotional morbidity 1–2 years later. Even though the majority of women diagnosed with breast cancer do not meet the criteria for a psychiatric diagnosis, the vast majority experience the diagnosis of cancer as a major stressor, and 10% have severe maladjustment problems as long as 6 years later. Thus, breast cancer is a disease that causes considerable psychological stress.

Stress and Cancer Incidence and Progression

There is mixed evidence regarding whether social stress, such as divorce, loss of a job, or bereavement, is associated with a greater likelihood of incidence or relapse of cancer. Three recent meta-analyses have reached rather different conclusions. The first, which relied heavily on the European literature, concluded that stress has no effect on cancer incidence. It also reported that bereavement had no effect among 11 studies (odds ratio, OR, 1.06; 95% confidence interval, CI, 0.95–1.18, not significant) but that other stressors had a twofold effect (OR 2.63; 95% CI 2.34–2.96). However, the researchers based their final null conclusion on six studies they deemed to be of higher quality because of the use of population controls, blinding of interviewers to disease status, and other methodological issues. In these six studies, stress had no effect on cancer incidence. Although this meta-analysis clearly raised important questions about the stress–cancer link, it did so only after dismissing the findings of the majority of studies in the field. The second meta-analysis found that both stressful life events, in general, and those that involve loss and separation, in particular, are associated with breast cancer incidence. The authors of the third meta-analysis reported a modest yet statistically significant association between exposure to stress and

the incidence of breast cancer, although they suggested that this association may be an artifact of publication bias.

It may well be that the factors that influence cancer progression are quite different from those affecting incidence. Although this may seem counterintuitive because the tumor burden is much greater as disease advances (by the time of death, a typical cancer patient may have approximately 1 kg of tumors in his or her body), the setting of disease progression is one in which physiological resources are stretched to their limit due to the higher tumor burden and cumulative effects of disease-related distress and disability. Even minor variations in somatic response to stress may have an effect on disease course at this point.

A large prospective study in Finland involving 10 808 subjects found that major life stressors such as divorces or death of a loved one in the preceding 5 years predicted a significantly elevated risk of subsequent breast cancer. However, another study found that reports of subjectively evaluated stress actually predicted lower subsequent breast cancer incidence. One way to interpret this apparent inconsistency in the literature involves the difference between experiencing and reporting stress. Given that the literature associates the repression of distress with a possibly elevated cancer risk, the studies of reported stress classify repressors with those subject who report no stress and therefore as being potentially at a higher cancer risk than if they actually had experienced objective stressors. Clearly more research is needed in this intriguing area.

Coping Style and Cancer Incidence and Progression

Coping-style variables that have been associated with higher cancer incidence include anger suppression and restrained aggression and introversion. Variables associated with slower progression include fighting spirit, expressed distress, assertive uncooperativeness, and assertiveness (being non-type C). Being emotionally reserved has been found to increase substantially the risk of mortality (odds ratio of 3.9) among lung cancer patients as well. Although the solidity of the evidence for these relationships varies, it is of interest that the styles putatively associated with cancer incidence and more rapid progression are similar emotional suppression and nonassertiveness.

Antoni and Goodkin found that a passive coping style is related to the promotion of cervical neoplasia. In another study, they also found that breast cancer patients were more likely to employ a repressive coping style than noncancer patients. Measures of this construct (repression/sensitization as a defensive

style) have been reported as accounting for as much as 44% of the variance in the progression of breast cancer. Although stress may weaken the body's capacity to resist the progression of cancer, coping well with stress may counteract this effect. Work on psychological stress suggests that lack of control is a key factor in the magnitude of the physiological stress responses. People who feel helpless in response to stress may exhibit more exaggerated physiological evidence of stress, whereas those with greater self-efficacy in coping may dampen the effects of stress. In breast cancer patients, the belief that they could control the course of disease was significantly associated with good adjustment. Indeed, several studies have shown that it is not so much the intensity of the stressor as its perceived controllability that determines its effects. Changes in coping styles have been associated with changes in immune function, even over periods of less than a month. For example, in a study of the effects of emotional disclosure, subjects who discontinued their avoidance and engaged more fully in discussions of a stressful or traumatic topic exhibited decreases in antibody titers to the Epstein-Barr virus, indicating more successful surveillance of the latent virus by lymphocytes.

Expression, rather than suppression, of emotion has been postulated to reduce stress by discharging physiological arousal. Stress-related immune fluctuations due to the threat or onset of illness have also been shown to be reduced in those with an increased ability to cope with changes in their life through assertive communication. Active coping through openly expressing concerns and emotions is one factor that may influence the endocrine and immune alterations that may be related to disease progression.

Social Support as a Stress Buffer

Psychosocial treatments appear to buffer the effects of stress on endocrine function. For example, one intervention study found that biofeedback and cognitive therapy were associated with reduced cortisol levels in newly diagnosed breast cancer patients. Social support provided by the intervention may ameliorate the effect of stress on endocrine function. Indeed, providing social support during stress can profoundly alter the endocrine consequences of the stressor. Levine modeled this effect in squirrel monkeys. Although the animals produced large elevations in plasma cortisol concentrations when they were stressed when alone, stress-induced elevations were reduced by 50% when animals had one friend present and they disappeared when five friends were present. Thus, interventions that provide social support have the potential to ameliorate endocrine stress responses.

Thus, the social environment can have a buffering effect on stress. Being socially embedded is associated with less autonomic arousal than social isolation. Indeed, social connection has profound consequences for health. Being well integrated socially reduces all-cause age-adjusted mortality by a factor of two, approximately as much as having low versus high serum cholesterol levels or being a nonsmoker. Furthermore, people's positions in the social hierarchy (including relatively higher status even within the same social class) have health consequences. People are statistically more likely to die after than before their birthdays and important holiday, which are usually important times for social bonding and companionship.

Thus, the presence of adverse emotional events such as traumatic stressors seems to have negative potential health consequences, whereas good social relations seem to be associated with positive health outcomes.

There is evidence from 5 of 11 published randomized trials that psychosocial support is associated with longer survival for patients with breast cancer, malignant melanoma, and lymphoma. One component of the effective interventions is dealing directly with emotional distress associated with fears regarding disease progression.

There is evidence that resilience to stress, including disease-related distress, is associated with emotional style. Indeed, finding meaning in the midst of experiencing a distressing situation has been linked with a positive psychological state. Better adaptation does not involve being persistently upbeat or rigidly maintaining a positive attitude but, rather, dealing directly with negative affect as well. Positive and negative emotions are not merely opposite sides of a single dimension. The suppression of negative emotion tends to reduce the experience of all emotion, positive and negative. There is evidence that breast cancer patients who express more negative emotion (including anger and uncooperativeness) or realistic optimism (fighting spirit) may live longer.

Physiological Mechanisms of Stress and Support-Related Effects on Cancer Progression

Identifying the physiological mechanisms by which emotional expression and social support may influence survival time for patients with cancer is an intriguing problem. Possibilities range from body maintenance activities such as diet, sleep, and exercise to interactions with physicians and better adherence to medical treatment regimens to stress and support-mediated effects on endocrine and immune

function. These systems, which are part of routine somatic control and maintenance functions in the body, are influenced by psychological and social variables and are also involved in host resistance to tumor progression.

There is a long history of research linking stress to increased cancer risk. In a series of classic experiments, crowded living conditions accelerated the rate of tumor growth and mortality in rats. In an authoritative 1998 review of the human stress literature, McEwen documented the adverse health effects of cumulative stressors and the body's failure to adapt the stress response to them. The activation of the HPA axis is an adaptive response to acute stress, but over time, in response to cumulative stress, the system's signal-to-noise ratio can be degraded so that it is partially on all the time; this leads to adverse physiological consequences, including abnormalities in glucose metabolism, hippocampal damage accumulation of abdominal fat, and depression. Depression has been associated with more rapid cancer progression in 19 of 24 recent studies, although it has little relationship to cancer incidence. Abnormalities of HPA axis function, including glucocorticoid receptor hypersensitivity, have also been found to be associated with posttraumatic stress disorder. Thus, adverse events ranging from traumatic stressors to cumulative minor ones and responses to them, including depression, are associated with HPA axis dysregulation. Persistently elevated or relatively invariant levels of cortisol may, in turn, stimulate tumor proliferation. Possible mechanisms include the differential glucocorticoid response of normal and tumor tissue to glucocorticoid signals to secrete glucose into the blood, the activation of hormone receptors in tumors, and immunosuppression. Indeed, glucocorticoids are potentially immunosuppressive, so the effects of acute and chronic stress and hypercortisolemia may include functional immunosuppression as well, as has been shown extensively in animals. There is a growing body of evidence of stress-induced immunosuppression in humans as well. This, in turn, could influence the rate of cancer progression.

Stress Reactivity and Disease Progression

Allostatic load has been defined as the price that the body has to pay for maintaining allostasis (stability through environmental change). Recently, it has been suggested that increased allostatic load, which may occur through the cumulative build-up of stress over a lifespan, may significantly hinder the functioning of different physiological systems. Cancer itself constitutes a series of stressors: the threat to life and health,

the disruption of social activities, and the rigors and side effects of treatment. A given individual's psychological and physiological response pattern to stress in general is likely to be exacerbated by the diagnosis and treatment of cancer, thereby intensifying mind-body interactions mediated by the stress response system.

Reactivity to stressors involves not only the initial response but also the system's ability to reset itself after stress. Older organisms seem to reset themselves after a stressor less readily, leaving stress response systems more tonically upregulated for longer periods of time after cessation of the stressor. Cancer particularly affects older people, and chronic disease-related stressors are likely to further decrease the flexibility of stress response mechanisms. Sapolsky and colleagues found that stressed older animals not only had persistently elevated cortisol levels after the stress was over but also had much more rapid growth of implanted tumors. This finding was confirmed in a Ben-Eliyahu and coworkers, who also identified the suppressive effects of stress hormones on natural killer (NK) cells as one potential mediator of the effect. The magnitude of stress-induced elevation of corticosteroids may not be as important in cancer outcomes as the duration or persistence of the elevation.

Circadian Rhythm Disruption, Stress, and Cancer

Recurrence or metastasis of cancer can be considered a severe chronic stressor. Both short-term and long-term alterations in psychological state (e.g., major depression) are known to be correlated with endocrine changes. Stress is associated with increased sympathetic nervous system (SNS) and HPA axis activity. Catecholamines and cortisol (released during SNS and HPA axis activity) are elevated in cancer patients as well as in bereaved and depressed subjects. Abnormalities in the circadian rhythm of cortisol have been observed in subjects undergoing psychological stress (e.g., depression, unemployment, and posttraumatic stress disorder) and some types of physical stress (e.g., shift work and excessive workloads). Studies in humans by Sephton and coworkers and in animals by Filipski and coworkers have shown that abnormal diurnal patterns of cortisol (which is typically four to five times higher in the morning on waking than in the evening in humans) predict earlier cancer mortality. Furthermore, women who engage in nighttime shift work are at higher risk for developing breast cancer. Thus, the disruption of circadian patterns of sleep and related hormones, including cortisol and melatonin, are associated with an elevated risk of cancer incidence and progression.

Psychoendocrinology and Cancer

There are considerable data indicating that the HPA axis constitutes a likely connection between psychosocial factors and disease progression. Stress is processed through central nervous system (CNS) neurotransmitter systems such as norepinephrine, dopamine, and endorphins, which regulate hypothalamic releasing factors such as corticotropin releasing hormone (CRH), pituitary hormones such as prolactin and adrenocorticotrophic hormone (ACTH), and, through these, adrenocortical steroids and medullary hormones. Cortisol is the classic stress hormone; it is elevated in response to stress for prolonged periods and is immunosuppressive.

The role of endogenous corticosteroids in breast cancer progression is worthy of further exploration because (1) steroid hormone levels are responsive to psychosocial stressors through the HPA axis via CRH, ACTH, and the response of the adrenal in producing them and (2) because breast cancer is a hormone-sensitive tumor, as evidence by the effectiveness of steroid receptor blockers such as tamoxifen in delaying progression. Although the specific receptors blocked involve estrogen and progesterone, cross-reactivity with androgenic steroids and the dedifferentiated state of tumor tissue make corticosteroids of potential interest.

High levels of stress hormones such as cortisol may be associated with worse disease prognosis because breast cancers are often hormone-sensitive. In a study of patients operated on for breast and stomach carcinoma, the failure of morning cortisol levels to decrease within 2 weeks after admission was associated with shorter survival times. Direct evidence for effects of the HPA axis in the development of breast cancer comes from a study in which high serum levels of dehydroepiandrosterone (DHEA), a correlate of HPA axis activity, predicted the subsequent development of breast cancer 9 years later in normal postmenopausal women who had donated blood for a serum bank. Thus, correlational evidence suggests that physiological stress responses associated with poor adjustment to cancer may, indeed, speed disease progression.

Mechanisms have been proposed whereby the neuroendocrine correlates of stress may promote neoplastic growth. Stress hormones may suppress immune resistance to tumors or act via differential effects on gluconeogenesis in healthy versus tumor cells. Tumor cells may become resistant to the catabolic action of cortisol, which inhibits the uptake of glucose in numerous cell types. In such cases, energy is preferentially shunted to the tumor and away from normal cells by cortisol. Several studies found an association between

the stress-related elevation of glucocorticoids and more rapid tumor growth in animals. Another hypothesis suggests that hormones of the HPA axis may actually promote the expression of breast cancer oncogenes. In addition, because stress-related increases in SNS and HPA axis activity are known to have generally suppressive effects on immune function, it is certainly plausible that immune functions important in resistance to breast tumor growth are thereby suppressed. The elevation of glucocorticoids is associated with clinically significant immunosuppression, and the enhanced secretion of norepinephrine during stress has also been associated with suppression of lymphocyte function.

Other hormones that are elevated by stress, such as prolactin, growth hormone, and thyroid hormones, may also influence the course of disease. Prolactin is of special interest because of its role as both a tumor promoter and a stress hormone. There is evidence that a majority of breast carcinomas have prolactin receptors and that the presence of prolactin stimulates the growth of such tumors in tissue culture at physiological levels, perhaps by stimulating the development of estrogen receptors. Prolactin has been found to stimulate prostate carcinoma growth as well. Indeed strong positive correlations between prolactin and both estrogen and progesterone receptor levels in breast tumors have been noted. The surgical stress of radical mastectomy results in transiently elevated serum prolactin levels, perhaps through the removal of target tissue and a disruption in feedback mechanisms. Prolactin levels rise in response to other stressful stimuli as well. Because other pituitary hormones such as ACTH clearly reflect stress, it is possible that prolactin secreted in response to stress can thereby promote tumor growth. Thus, there is evidence for facilitative effects on tumor growth by stress hormones via several different mechanisms. Acute medical illness is associated with the increased secretion of ACTH and cortisol, and surgical stress is associated with lower DHEA/cortisol ratios (i.e., a combination of relatively low DHEA and relatively high cortisol levels) and higher prolactin levels.

Psychoneuroimmunology and Cancer

Many of these hormones are immunosuppressive when elevated. In spousal bereavement, cortisol levels are elevated, whereas NK cytotoxicity is decreased. Furthermore, there is no reason why disease-related stresses, including social isolation, death anxiety, pain, and sleep disturbance, should not affect immune function at the clinically observable level, for example in defenses against the progression of viral

infection. Yet little is known about these mechanisms or their effects on clinical infection. Cohen showed that vulnerability to systemic infection by rhinovirus is significantly mediated by stress and that people with larger social networks are less susceptible to viral infection. This model might well apply as well to vulnerability to cancer progression.

There is a growing body of research that provides evidence of associations between psychological factors and immune function. The work of the Glasers has shown that subjects experiencing such stressors as being a caregiver for an Alzheimer's patient, undergoing divorce or bereavement, or being in a poor relationship have been shown to have greater distress and lower helper/suppressor cell counts and ratios, lower total T-lymphocyte numbers, lower NK cell activity, and poorer response to Epstein-Barr virus.

Further evidence comes from work showing that immune organs such as the spleen are heavily innervated and that there are receptors for a variety of neurotransmitters on the cell membranes of lymphocytes. Stress hormones are known to influence cell adhesion molecules on lymphocytes, providing a possible mechanism by which stress may alter lymphocyte trafficking and therefore immune function. There is good evidence for this hypothesis in work by Dhabhar, who demonstrated that stress and the circadian corticosterone rhythm induce significant changes in leukocyte distribution in animals. Moreover, he showed that acute stress significantly enhances a cutaneous delayed-type hypersensitivity (DTH) response (an *in vivo* measure of cell-mediated immunity). In contrast, chronic stress significantly suppresses DTH responses, and stress-induced trafficking of leukocytes to the skin may mediate these changes of immune function. Most interestingly, he was able to mimic the effects of acute and chronic stress on DTH responses by the administration of corticosterone at doses similar to physiological levels following either acute or chronic stress, suggesting that glucocorticoids may be an important mediator of DTH responses during stress.

Thus, there is growing evidence for links among psychological stress, glucocorticoid responses, and immunosuppression. In rats, for example, stress significantly decreased NK cytotoxicity, increased levels of cortisol and ACTH, and increased the metastatic spread of mammary tumor to lung tissue. In stage I and II breast cancer patients, higher perceived social support and active seeking of social support were related to increased NK cell activity, and lower NK cell activity has been shown to predict disease recurrence. Thus, psychological factors have been associated with changes in immune measures, and these changes are markers of cancer progression.

Psychosocial Interventions and Immune Function

Psychosocial treatments designed to improve coping and assist in stress reduction have been associated with improvements in both immune function and physical health in cancer patients. Supportive relationships seem to modulate stress-induced immunosuppression, for example, among medical students undergoing examination stress. Even the disclosure of traumas in journals has been associated with better immune function. Stress-related immunosuppression may be mitigated by active coping styles or social support. After a 6-week psychiatric intervention for newly diagnosed malignant melanoma patients conducted by Fawzy, α -interferon-augmented cytotoxic activity of NK cells was elevated in the intervention subjects 6 months later and a survival benefit was observed, although there was no direct relationship between changes in NK cell activity and survival time. NK cells are known to kill tumor cells of many different types when tested either *in vitro* or in animal studies and may be particularly prone to stress-induced suppression.

Immune Function and Disease Progression

It is not always clear whether changes in serum immune or endocrine measures are a cause of or result from changes in disease status. Indeed, the immune surveillance theory of cancer development and progression has been questioned. Cancer is not, after all, an immunodeficiency disease, and immunosuppressed patients such as those with acquired immunodeficiency syndrome (AIDS) are at higher risk only for rare cancers such as Kaposi's sarcoma. However, research in animals as well as humans suggests that immunosuppression may be a factor in the rate of disease progression, and *in vitro* measures of immune function that are salient to disease progression are affected positively by social support and adversely by stress.

There is evidence from nonclinical populations that social integration buffers against the stress-induced suppression of NK cell activity. It thus makes sense that the addition of intensive social support in groups for women undergoing substantial medical and other stressors might have a similar buffering effect. Thus, although definitive clinical links have not been made, there is evidence that NK cell activity is associated with the rate of breast cancer progression. This immune parameter, which is affected by stress and social support, also varies with disease progression, making it a reasonable candidate for investigations into the

mediators of the effects of group support on the rate of breast cancer progression.

Stress and Autonomic Responses

Cacioppo and colleagues examined the effects of psychological stress on autonomic responses including heart rate reactivity and cardiac vagal reactivity (as indexed by respiratory sinus arrhythmia reactivity). They found that individuals who demonstrated greater autonomic responses to brief laboratory stress also showed higher stress-related increases in ACTH and cortisol. Thus, autonomic tone may be a good index of the reactivity of the HPA axis to acute stress, and autonomic measures are less invasive than the repeated blood sampling that is required to measure short-term changes in HPA axis activity. The results of psychoneuroimmunological studies have shown that heart rate reactivity after a stressor was also related to changes in NK cell activity, suggesting that the measurement of stress-induced changes in autonomic tone may be of value in defining the possible mechanisms underlying psychosocial effects on immune function. Failures to mount a corticosteroid response to stress may place an undue burden on other stress response mechanisms, such as the autonomic nervous system. The loss of vagal tone is often seen in poor stress responders, leading to cardiovascular burden and possibly to immune effects as well. Autonomic tone has not only been associated with stress, but also with emotion regulation. Indeed, the baseline levels of cardiac vagal tone and vagal tone reactivity have been associated with behavioral measures of reactivity and the expression of emotion; thus, cardiac vagal tone can also serve as an index of emotion, and parasympathetic influences on heart rate may be influenced by emotional expression, for example as a means of coping with stressors.

Conclusion

The literature reviewed here provides evidence that people's manner of responding to disease-related stress may affect the rate of disease progression as well as their psychological adjustment. The stress response system is most effective when it is on when needed and off when it is not. Variance in the rate of cancer progression may be accounted for by more or less adaptive stress response systems. This effect of stress and stress response on cancer progression may be mediated by the endocrine, immune, or autonomic nervous systems or by some combination of the three. There is evidence that stress is a uniform part of the experience of cancer and its treatment, that various coping styles affect stress-mediated physiological

effects, and that these in turn may affect the somatic response to cancer progression. Furthermore, various psychosocial interventions have proven effective in enhancing coping with this stress, and some have demonstrated an effect on disease progression. This evidence suggests that assistance with stress management can contribute to effective cancer treatment.

Further Reading

- Fawzy, F. I., Canada, A. L. and Fawzy, N. W. (2003). Malignant melanoma: effects of a brief, structured psychiatric intervention on survival and recurrence at 10-year follow-up. *Archives of General Psychiatry* 60(1), 100–103.
- Lillberg, K., Verkasalo, P. K., Kaprio, J., et al. (2003). Stressful life events and risk of breast cancer in 10,808

- women: a cohort study. *American Journal of Epidemiology* 157(5), 415–423.
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* 338(3), 171–179.
- Sapolsky, R. M. (1996). Why stress is bad for your brain. *Science* 273(5276), 749–750.
- Sephton, S. E., Sapolsky, R. M., Kraemer, H. C. and Spiegel, D. (2000). Diurnal cortisol rhythm as a predictor of breast cancer survival. *Journal of the National Cancer Institute* 92, 994–1000.
- Sephton, S. and Spiegel, D. (2003). Circadian disruption in cancer: a neuroendocrine-immune pathway from stress to disease? *Brain, Behavior & Immunity* 17(5), 321–328.
- Spiegel, D. and Classen, C. (2000). *Group therapy for cancer patients: a research-based handbook of psychosocial care*. New York: Basic Books.

Cancer Treatment

F I Fawzy and A L Canada

University of California, Los Angeles School of Medicine,
Los Angeles, CA, USA

N W Fawzy

John Wayne Cancer Institute, Saint John's Health
Center, Santa Monica, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 376–380, © 2000, Elsevier Inc.

The stress associated with the diagnosis and treatment of cancer can result in significant reductions in patient quality of life. Fortunately, numerous psychosocial interventions for patients with cancer have been found to be quite effective in assisting with such stress. Interventions, including education, stress management, coping skills training, and support, have resulted in notable improvements in the psychological and physical health of patients with cancer.

The Cancer Problem

Psychosocial Interventions

Intervention Outcomes

Conclusion

Glossary

Active-behavioral coping methods

Actions that an individual does that focus on dealing with problems and tend to make a person feel better in both the short and the long term.

Active-cognitive coping methods

Mental thoughts and techniques that focus on dealing with problems and tend to make a person feel better in both the short and the long term.

Avoidance methods

The behavioral and mental techniques that a person uses to feel better in the short term but that are not directed at solving problems and, therefore, fail to work in the long term.

The Cancer Problem

Cancer remains a significant health problem in the United States with over 1 million new cases reported each year. Although many patients still die from their cancer, major advances in medical care have occurred. Over the past 30 years, many patients have been cured and long-term survival for some types of cancer has increased significantly. Therefore, today's patient with cancer may be placed in one of three basic categories. The first category consists of patients who are cured and yet need to deal with a chronic fear of recurrence. The second category consists of cancer patients and includes those who are not cured but have a long-term prognosis; these patients must face dealing with a chronic disease. The third category consists of those patients who, despite medical advances, are terminal and must deal with death and dying.

Regardless of the category in which a patient finds him- or herself, the adverse psychological effects

associated with the diagnosis and treatment of cancer can be quite severe. Cancer is associated with high rates of emotional distress, which may include major depression, dysthymia, or adjustment disorder with depressed mood. Other possible affective responses to cancer are anxiety, anger, guilt, helplessness, hopelessness, denial, disbelief, and grief. In addition to such psychological effects, the diagnosis and treatment of cancer are often accompanied by adverse physical symptoms. These can include pain, insomnia, fatigue, shortness of breath, nausea and vomiting, and changes in appetite and activity. In addition, the stress associated with diagnosis and treatment is believed to negatively affect immune parameters and other somatic variables in patients with cancer.

Psychosocial Interventions

Fortunately, various psychological interventions have been developed to assist patients with cancer in dealing with diagnosis and treatment. A review of the literature over the past 20 years suggests that available interventions for the management of cancer-related stress include one or more of the following basic components: education, stress management (including behavioral training), coping skills training, and emotional support.

Education

The overall goal of education for the cancer patient is to replace the sense of helplessness and inadequacy due to uncertainty and lack of knowledge with a sense of mastery and control. Educational interventions not only offer diagnosis- and treatment-specific information to patients, but may also enhance adjustment and improve coping skills. Those interventions currently in use provide the cancer patient with information about his or her particular diagnosis and treatment side-effects, as well as how to manage the side-effects. Cancer prevention information and how to recognize the warning signs of new or recurrent cancers may also be taught. In addition, general education relevant to a healthy lifestyle is often provided. Such topics as good nutrition, exercise, immune functioning, and effective communication are but a few examples.

Stress Management

Because stress is associated with physical disease and illnesses (e.g., headache, gastrointestinal problems, heart disease, hypertension, and susceptibility to infections), it is logical to assume that making and sustaining even minor lifestyle changes that reduce

stress can improve the odds for better health and a better quality of life. Research has shown that a generalized decrease in stress allows patients to cope more effectively with the overall emotional impact of cancer, enhances compliance with medical regimens, and assists in the management of treatment side-effects (e.g., pain, nausea and vomiting, and insomnia). In such an intervention, patients learn how to recognize personal sources of stress and then how to eliminate or modify the source of stress through problem solving or a change in cognitive appraisal or attitude. Patients are also taught how to change their physiological response to the stressor by replacing it with a relaxation response. Such responses are learned through behavioral training and include the techniques of progressive muscle relaxation (PMR), deep breathing, meditation, biofeedback, hypnosis, self-hypnosis, guided imagery, yoga, and Tai chi.

Coping Skills Training

The goal of coping skills training is to teach the cancer patient how to deal with his or her diagnosis, treatment, and life in general using effective cognitive and behavioral methods. An example of such an intervention is one in which patients are introduced to three types of coping: active-behavioral, active-cognitive (both of which are usually effective coping methods), and avoidance (which is generally ineffective in the long term). Active-behavioral methods may include exercising, improving diet, using physical relaxation techniques, expressing feelings, going out more socially or treating oneself to an outing, and finding out more about the disease. Active-cognitive methods may include thinking about the illness 1 day at a time, praying and trusting one's belief in God or a higher power, accepting the reality of the diagnosis but not equating it with a poor prognosis, forming a mental plan of action, thinking more about the meaning of life and what is really important, using some kind of mental relaxation technique, and trying to maintain and use positive thinking. Avoidance techniques include self-blame; social isolation; procrastination; and tension reduction through eating, sleeping, drinking alcohol, smoking, or taking drugs more than usual. Following instruction, participants are often asked to apply these coping strategies to common theoretical problems and situations they have encountered (e.g., diagnosis, treatment issues, body image, communication, interpersonal relationships, anxiety/depression, and death). After discovering effective ways to deal with these theoretical problems through active-behavioral and active-cognitive means, patients are asked to apply the techniques to their own real-life situations.

Emotional Support

Interventions for patients with cancer can be delivered in individual, patient group, or family settings. Whatever the format, psychosocial interventions can be an invaluable source of support for patients with cancer. A safe environment provides patients with the much-needed opportunity to express the feelings they have about their disease and its treatment.

In group therapy, patients also have the opportunity to learn healthier ways of dealing with their illness from others in similar situations. The sense of community experienced through group participation provides patients with a new and important social network at a time when isolation is common and more emotional support is needed.

Individual psychotherapy is a long-established method that is also used to ease the distress and disruption that accompanies the diagnosis of cancer. Support, compassion, and empathy from a therapist form the cornerstone of such interventions.

Psychotherapeutic interventions have also been used to enhance family support. Improved communication; role flexibility; and adjustment to changing medical, financial, social, and vocational realities are targeted. An atmosphere of openness and shared problem solving in the family is associated with affective improvements in patients with cancer.

Combination Therapies

Research has shown that education, behavioral training, coping skills training, and support have been generally helpful when used separately but are even more powerful and enduring when used in combination, especially in a group format. The explanation for this success may be that patients are given multiple tools with which to cope. It may also be that different individuals come to the cancer experience with different coping skills and resources already available. In this case, their existing coping skills are reinforced and supplemented in those areas in which they are lacking. In addition, the group format provides tremendous emotional support and is more cost

effective than an individually delivered intervention. A summary of the types of interventions that have been proven helpful for patients with cancer is provided in [Table 1](#).

Intervention Outcomes

Over the years, the outcomes of psychosocial interventions for patients with cancer have yielded improvements in both psychological and biomedical functioning.

Psychological Outcomes

Educational interventions for patients with cancer have resulted in increases in knowledge about the disease and treatment compliance and decreases in anxiety and depression. Such interventions have also been shown to enhance coping skills. Behavioral training interventions have resulted in decreased anxiety, whereas supportive interventions, delivered in individual, group, and family formats, have been shown to reduce emotional upset, confusion, social isolation, and anxiety and depression and to improve quality of life, communication, coping skills, sleep, adjustment, body image, and self-concept.

Physical Outcomes

Both educational and behavioral training interventions have been shown to be effective in assisting the patient with cancer-related pain. Behavioral training has also been found to be quite helpful in reducing the frequency, severity, and duration of side-effects from chemotherapy (i.e., anticipatory nausea and vomiting). In fact, some studies demonstrate an increase in food intake as a result of such lowered physiological arousal.

Finally, a growing body of literature suggests that psychosocial interventions for patients with cancer may result in improvements in immune functioning and increases in recurrence-free intervals and survival rates. Behavioral training has been associated with decreased cortisol levels, elevations in immune system

Table 1 Effective therapies for the patient with cancer

<i>Education</i>	<i>Stress management</i>	<i>Coping skills training/ mechanisms</i>	<i>Support</i>
Diagnosis: test results, symptoms	Manage stress source	Active-behavioral (e.g., problem-solving)	Emotional support: expression, validation
Treatment: options, side-effects, dealing with effects	Manage stress reaction	Active-cognitive (e.g., cognitive restructuring)	Group therapy: disease specific, social network
Disease: recurrence, prevention	Eliminate/modify stressor	Avoidance (e.g., drinking or smoking more)	Individual therapy: patient specific
Health: nutrition, exercise	Relaxation response: PMR, meditation, deep breathing, biofeedback, hypnosis, guided imagery, yoga, tai chi		Family support: communication, role flexibility, adjustment

Table 2 A sampling of intervention outcomes

Category	Type of intervention	Results	Source
Psychological outcomes	Education	Decreased depression	Pruitt et al. (1993)
	Behavior training	Decreased anxiety	Gruber et al. (1993)
	Education, support	Increased effectiveness in dealing with psychosocial problems	Gorden et al. (1980)
Physical outcomes	Behavior training	Decreased anticipatory nausea	Arakawa (1997)
		Increased immune parameters	Lekander et al. (1997)
Combined outcomes	Education	Decreased pain intensity	de Wit et al. (1997)
	Support	Decreased emotional distress, reduced pain, increased survival rates	Spiegel et al. (1981, 1989)
	Education, stress management, coping skills training, support	Decreased emotional distress, enhanced coping, increased immune parameters, increased recurrence-free intervals and survival rates	Fawzy et al. (1993)

functioning, and improvements in immune parameters (e.g., natural killer cell activity and lymphocyte number and responsiveness). Supportive interventions and combination therapies involving follow-up assessments over many years have also demonstrated reductions in cancer recurrence and death rates for patients participating in such interventions. A sampling of intervention outcomes in patients with cancer is provided in [Table 2](#).

Conclusion

The emotional and physical stressors encountered by cancer patients are numerous and unique. Based on a review of the literature, cancer patients may benefit both biomedically and psychologically from a variety of psychosocial intervention programs. A structured, psychosocial intervention consisting of education, stress management, coping skills training, and group support offers the greatest potential for positively affecting emotional and physical functioning in patients with cancer. It should be noted that such psychosocial interventions need to be used as an integral part of competent, comprehensive medical care and not as an independent treatment modality for cancer.

See Also the Following Articles

Avoidance; Cancer; Coping Skills; Social Support; Stress Management, CAM Approach.

Further Reading

- Anderson, B. L. (1992). Psychosocial interventions for cancer patients to enhance quality of life. *Journal of Consulting and Clinical Psychology* 60, 552–568.
- Arakawa (1997). *Cancer Nursing* 20, 342.
- Compass, B. E., Haaga, D. A., Keefe, F. J., et al. (1998). Sampling of empirically supported psychological treatments from health psychology: smoking, chronic pain, cancer, and bulimia nervosa. *Journal of Consulting and Clinical Psychology* 66, 89–112.
- de Wit, et al. (1997). *Pain* 73, 55.
- Fawzy, et al. (1993). *Archives of General Psychiatry* 50, 681.
- Fawzy, F. I., Fawzy, N. W., Arndt, L. A., et al. (1995). Critical review of psychosocial interventions in cancer care. *Archives of General Psychiatry* 52, 100–113.
- Gorden, et al. (1980). *Journal of Consulting and Clinical Psychology* 48, 743.
- Gruber, et al. (1993). *Biofeedback & Self Regulation* 18, 1.
- Iacovino, V. and Reesor, K. (1997). Literature on interventions to address cancer patients' psychosocial needs: what does it tell us? *Journal of Psychosocial Oncology* 15, 47–71.
- Lekander, et al. (1997). *Psychotherapy and Psychosomatics* 66, 185.
- Pruitt, et al. (1993). *Psycho-oncology* 2, 55.
- Spiegel, et al. (1981). *Archives of General Psychiatry* 38, 527.
- Spiegel, et al. (1989). *Lancet* 2, 888.
- Trijsburg, R. W., van Knippenberg, F. C. E. and Rijpma, S. E. (1992). Effects of psychological treatment on cancer patients: a critical review. *Psychosomatic Medicine* 54, 489–517.

Cannon, Walter, B. See: Homeostasis.

Captivity, Adaptation to

R H Rahe

University of Washington School of Medicine, Seattle, WA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R H Rahe, volume 1, pp 381–384, © 2000, Elsevier Inc.

Introduction

Stage I: Startle/Disorientation

Stage II: Disbelief

Stage III: Hypervigilance

Stage IV: Resistance/Compliance

Stage V: Depression

Stage VI: Eventual Acceptance

Comment

Introduction

Few life experiences are as difficult and challenging as being held hostage or as a political prisoner. In reviews of the literature and from my own clinical work dealing with former hostages and prisoners-of-war, six stages of adaptation to captivity became apparent. The time course over which these stages occur progressively lengthen from the first to the final stage. More specifically, the first stage lasts only seconds to minutes, whereas the next stage lasts minutes to hours. The third stage generally lasts a few hours to a few days, whereas the fourth stage lasts from several days to a few weeks. The fifth stage is protracted, lasting several weeks to a few months. Finally, the sixth stage can last from several months to a year or two. In the presentation that follows, a general description of each stage is given and mention is made of key captor and captive behaviors, along with useful coping techniques.

Stage I: Startle/Disorientation

The first stage lasts from seconds to minutes. The development of this stage is influenced greatly by the setting in which a captive is taken. Military personnel who are at high risk for capture as prisoners-of-war have often gone through intensive training on what they might encounter. Hostage takings are generally unanticipated and often life-threatening. A common theme in both groups is the abrupt transition from life as usual to sudden, often brutal, subjugation. Such a transition is impossible to assimilate quickly. Captives frequently experience extreme disorientation, seek possible means of escape, and have

anxieties for their lives and the lives of others during these first seconds to minutes.

Captors at this stage are generally highly excitable and fervently driven by political and/or ideological causes. Captives may well be at a loss initially about how to deal with them. To challenge the captors' actions, even by a scornful laugh, may lead to impulsive retaliation and possibly to execution.

Successful coping with capture is constrained by the extreme brevity of this stage. Captives should regain their orientation and control of anxiety as quickly as possible. Some captives cope with the initial shock in a dissociated state, feeling more like observers than participants. Others can be in a state of dazed shock, unable to take any immediate actions. The captives' ability to bring their minds into focus on the details of their surroundings, such as counting the number of captors or studying their behaviors, may greatly help to limit these acute responses to captivity stress.

Stage II: Disbelief

The second stage lasts from minutes to hours. Captives from Western countries, accustomed to protections under the law and an emphasis on human rights, may well respond with utter disbelief when they are suddenly thrust into a situation in which all personal freedoms are forcibly withdrawn. They may think, "This can't be happening," or "Surely we'll be rescued shortly." At some deeper psychological level, captives know that such expectations can be or are wrong, but these thoughts often persist during this stage.

The captors' behavior during this time can be tyrannical, with beatings and physical restraint. Dehumanization includes stripping away the captives' clothes and personal possessions, binding them, blindfolding them, and taking photographs of the subjugated captives for later propaganda use. Captives may even be tortured to obtain early confessions. The captors' plans for the captives are frequently unstable at this stage, with the group moving several times from one location to another being the typical scenario. The captives generally have little chance to communicate with one another, and their attempts to do so are likely to bring swift and painful retaliation.

Many prisoners cope with this stage by convincing themselves that conditions will improve shortly. Hopes for rescue may be sustaining, and sometimes rescue does occur. However, it is probably best for captives to try to escape their circumstances by turning their attention inward. Some captives engage in thoughts of retaliation, but it is more helpful for them to focus on

images of loved ones, valued friends, prayer, and other sustaining elements of life and freedom.

Stage III: Hypervigilance

The third stage lasts from hours to days. Although efforts by the captives to orient themselves to their new environments may start at the very beginning of captivity, the emergence of the hypervigilant stage releases unexpected talents of observation. Assessing mileage while being driven blindfolded, remembering labyrinthine pathways through city streets, and accurately keeping track of passing time are examples of such orientation skills that frequently emerge.

Captor behavior is usually unpredictable during the first days of captivity. The leaders of the captors are often confused about their ultimate goals and objectives. Generally, some form of interrogation begins, with an emphasis on intelligence gathering. Guards tend to be highly attentive to possible escape attempts during these early days of captivity.

Early in their incarceration, captives turn their hypervigilance to orienting tasks. Even when confined in a cell without daylight, nonvisual cues (changing of the guards, when meals are served, etc.) are gleaned to construct a 24-h cycle. The captives examine their surroundings in minute detail, including the behavior of their guards. They also structure their environment as much as possible – sandals placed here, the latrine bucket placed there, and so on.

A frequent mistake made by captives during this stage is to misinterpret bits of their captors' conversations that they overhear, particularly concerning the release of the captives. The captives' wishful thinking about an early rescue leads to the generation of false rumors and disappointment becomes inevitable.

Stage IV: Resistance/Compliance

The fourth stage lasts from days to weeks. This stage frequently begins when the captors begin to coerce the captives into cooperating with them; captors may demand that captives sign confessions of their crimes or make enforced public appearances to help dramatize the captors' achievements. Interrogations at this point change from intelligence gathering to exploitation. The degree to which a captive capitulates to these tactics depends on the severity of the torture applied combined with the captive's ability to resist coercion – which may include torture. It is clear from the literature that anyone can be forced to cooperate with his or her captors if sufficient physical and psychological torture is brought to bear. Nevertheless, people with strong character, and particularly those who have been trained to resist such demands, do much better than others without these assets.

The captors' behavior at this stage is so similar for all prisoner-of-war and civilian hostage situations that it is as if they all had read the same manual. The original architects of many of these behaviors were the secret police in Czarist Russia, with modifications later introduced by the state police in Nazi Germany. Asian refinements were added a decade later by North Korean military. These behaviors include intimidating arrests, imprisonment for indeterminate lengths of time, physical and nutritional deprivation, disturbances of body rhythms, physical and sensory isolation, deprivation of daylight, stressful (often brutal) interrogations, unpredictable responses from guards and inquisitors, frequent threats of death, and attempts to reeducate the captives.

Isolation appears to be the most stressful captor tactic, dreaded by many more captives than physical torture. The settings used are generally cramped, filthy, and pest-ridden. The captives are kept uncomfortably hot or cold, with poor lighting and ventilation. Communication with fellow captives, and even with guards, is usually prohibited. Finally, extremely few allowances are provided for personal hygiene. Periods of isolation may be alternated with friendly interrogations, especially if the captive becomes eager for human contact. Interrogations are almost always unpredictable as to timing (day or night), duration, and approach of the interrogator(s). The overall aim is to soften the captives' resistance and make them susceptible to the will of their captors.

The captives' behavior is highly variable. Some attempt to resist all coercive attempts until they are thoroughly beaten and broken, even to the point of death. Others prove to be unable to withstand even moderate deprivations without making major concessions. Most captives, however, learn to resist torture, isolation, and other deprivations until severe physical and/or psychological damage becomes a likely result of continued resistance. Then, they comply sufficiently for captors to suspend their tortures, giving the captive critical time to rest and recover. Providing clever admissions during interrogations, especially information that takes a long time to validate, can provide precious time for captives to replenish their physical and mental reserves.

One extremely important coping maneuver is physical fitness. Hours of physical training can be carried out even in extremely confined settings. Flexibility exercises minimize the risk of joint dislocations and bone fractures from ensuing torture sessions. Running in place promotes endurance conditioning, with attendant feelings of stress reduction and increased ability to sleep. Perhaps the most important coping activity is communication with fellow captives. The ingenuity of captives in passing voice, hand, tap, or other signals

to their cohorts is truly remarkable. By such methods, communication can be carried out even with people confined to isolation. To feel part of a supportive network of fellow sufferers provides immeasurable sustenance to a captive. Also of extreme importance is faith, prayer, and thoughts of loved ones.

Stage V: Depression

The fifth stage lasts from weeks to months. When captives begin to confront their several recent losses, particularly if they dwell on these losses, they frequently enter into depression. In reality, captives have lost not only their freedom and most everything else that they have valued in life, but they may have even also lost their future. Further, they are deprived of information about the outside world and are often misled by captors as to the degree of interest in their welfare shown by their own government and countrymen. Perhaps most cruel is the indeterminate length of stay in captivity. Prisoners-of-war sometimes reflect with envy on civilian prisoners who have finite prison sentences to serve.

Captives with depression may wish to spend most of their day in bed. Loss of appetite far exceeds the poor quality of the food provided in causing large weight losses. Communication between captives may be avoided even though opportunities for such interactions present themselves. The captives may also feel guilt over their performance during interrogations. Thoughts of suicide may become prevalent, but actual suicide attempts are rare.

The captives' coping with depression is aided enormously by a strong support group. Repeated documentation exists of the virtually life-saving influences of support people in the lives of depressed captives. One excellent coping skill is for captives to use their intellect constructively to combat the potentially devastating boredom. Even a single book can be read and reread, with new meanings found on each reading. Creative work, such as writing, can be extremely helpful. Even without these tools, captives can use their intelligence to elaborate future life plans that they want to achieve. One captive, deprived of paper and pencil, composed books of children stories in his mind and after his release wrote and published the stories. Humor has an immense coping value. Captives getting the best of their guards on occasion not only provides humorous remembrances that can be savored later but gives the captives a moment of control in what otherwise is a totally uncontrolled life situation.

Although nearly all captives experience depressive symptoms during this stage, for some the symptoms are mild and the duration is fleeting. Others become significantly depressed and, in addition, physically ill.

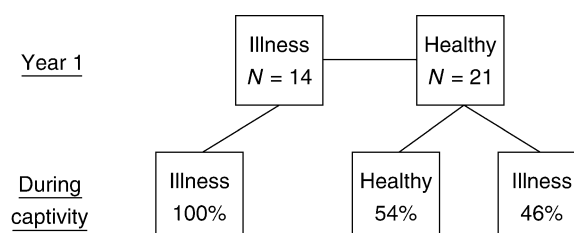


Figure 1 Former American hostages held in Iran ($N = 35$) found to be healthy 1 year after release were frequently healthy throughout captivity. None of the former hostages who reported being ill 1 year after release were healthy during captivity.

The most common physical symptoms are gastrointestinal, dermatological, cardiovascular (hypertension), neurological (headache), orthopedic (injuries), and respiratory. The number of symptoms that captives experience, provided their captivity experience was roughly the same as that of others, often reflects their susceptibility to stress. Figure 1 presents data on illness symptoms reported during and 1 year after captivity for 35 Americans held hostage in Iran. Of the 14 former captives who experienced significant mental and physical symptoms during the first year following their release from captivity, suggesting stress susceptibility, 100% of them also reported having had significant symptoms during captivity. Conversely, of the 21 former captives who reported good health during the first year following their release from captivity, suggesting a low susceptibility to stress, only 46% reported having had significant illness symptoms during captivity.

The captors' behavior at this point is generally limited to custodial care. Unless there is a continuing need for propaganda statements, interrogations are greatly reduced in number. The guards may be changed because the routine generates little excitement for the original captors; escape may become possible at this time, due to the laxity of the new guards. The captors may occasionally reassert their position of control by staging prisoner movements, issuing new policies, and so forth.

Stage VI: Eventual Acceptance

The sixth stage lasts from months to years. Until this stage, captives have actively resisted the idea of a prolonged incarceration. It is now clear that a release, or a rescue, is not likely to occur any time soon. Thus, this final stage of adaptation starts with the conviction that captives must make a more productive use of their time than waiting for rescue if they are to successfully tolerate their ordeal.

Much as a sailor steers a disabled ship by simply concentrating on one wave then the next, captives

find that it's best to live their lives one day at a time. Although they may dwell on the past, not much thought is given to a future more distant than tomorrow. Small things begin to take on huge importance. A frequent preoccupation of captives is what will be served, generally with rice, for the next meal. Meal times, exercise periods, bedtime, and so forth are carried out with compulsive regularity. Social niceties may disappear, and personal appearance may slide. Swearing and crude language can become the norm. Interpersonal difficulties often arise more from intolerance about how a roommate chews his food rather than from philosophical or political differences.

The passage of time becomes distorted. As characterized in Thomas Mann's famous novel, *The Magic Mountain*, although a single day may appear to progress very slowly, weeks and months seem to fly by. Captives have been amazed to find, on release, that they greatly underestimated the time they spent on various activities. The involvement in creative tasks makes time pass swiftly. A German prisoner-of-war, held along with Americans in Vietnam, composed a German-English-French dictionary. He used cloth from his prison pajamas for a book cover, charcoal and water for his ink, paper obtained from his captors, and foil from fellow prisoners' cigarette packets to insulate his book cover. When he was fully absorbed on this project he found that "There just weren't enough hours in the day."

The acquisition of new knowledge is assiduously pursued during long-term captivity. Some captives learn a second language from their cellmates. Others build homes in their imagination – board by board, nail by nail. One prisoner-of-war held in North Vietnam reviewed his entire schooling and found he could recall the names of every classmate, where each one had sat in class, the names of every one of his teachers, and even the subject matter they taught. Reading and writing are important coping activities. Even if they are never delivered, writing letters to family and loved ones is an enormous source of support for captives.

Group activities lend a vital cohesive force to the captive group. These activities range from church services and group games to theater and fitness activities. Even forced labor can become a sustaining activity, as is so aptly illustrated in Boulle's famous novel, *The Bridge over the River Kwai*.

The captors' behavior remains custodial in the main during this period. With the exception of re-education sessions given in some prisoner-of-war settings, captors are primarily concerned with keeping their victims alive and presentable as bargaining chips for future negotiations.

Comment

The emphasis in this report on the stages of adaptation to captivity is on the captives' stress. Very little has been studied about the stresses experienced by the immediate families of the captives. Certainly family members experience their own stages of adaptation, some of which are probably similar to those of captives. However, family members not only have to deal with loss of the captive, but they also must learn to deal with a nearly constant intrusion into their lives by the public and press. The families also need support groups. The success of such groups was demonstrated by the families of Vietnam prisoners-of-war and families of the Americans captured in Iran. These families developed support groups for enhanced communication, emotional solidarity, and political influence, and these associations proved to be extremely effective.

Further Reading

- Boulle, P. (1954). *The bridge over the River Kwai*. New York: Vanguard Press.
- Guarino, L. (1991). *A P.O.W.'s story: 2801 days in Hanoi*. New York: Ivy Books.
- Gunderson, E. K. E. and Rahe, R. H. (1974). *Life stress and illness*. Springfield, IL: Charles C Thomas.
- Hinkle, L. E., Jr. and Wolff, H. G. (1956). Communist interrogation and indoctrination of "enemies of the state." *Archives of Neurology and Psychiatry* 76, 115–174.
- Hubbell, J. G. (1976). *P.O.W.* New York: Reader's Digest Press.
- Mann, T. (1924). *The Magic Mountain*. Berlin: S. Fischer Verlag.
- Rahe, R. H. (1990). Life change, stress responsivity, and captivity research. *Psychosomatic Medicine* 52, 373–396.
- Rahe, R. H. (1995). Stress and psychiatry. In: Kaplan, H. I. & Sadock, B. J. (eds.) *Comprehensive textbook of psychiatry* (vol. 6). Baltimore, MD: William and Wilkins.
- Rahe, R. H. and Arthur, R. J. (1978). Life change and illness studies: past history and future directions. *Journal of Human Stress* 4, 3–15.
- Rahe, R. H., Karson, S., Howard, N. S., et al. (1990). Psychological and physiological assessments on American hostages freed from captivity in Iran. *Psychosomatic Medicine* 52, 1–16.
- Schwinn, M. and Diehl, B. (1976). *We came to help*. New York: Harcourt, Brace, Jovanovich.
- Stockdale, R. and Stockdale, S. (1984). *In love & war*. New York: Harper & Row.
- Wolf, S. and Ripley, H. S. (1954). Reactions among allied prisoners of war subjected to three years of imprisonment and torture by the Japanese. *American Journal of Psychiatry* 104, 180–193.

Captivity, Recovery from

R H Rahe

University of Washington School of Medicine,
Seattle, WA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
R H Rahe, volume 1, pp 385–388, © 2000, Elsevier Inc.

Introduction

Stage I – Brief Euphoria

Stage II – Hyperarousal

Stage III – Compliance/Resistance

Stage IV – Denial

Stage V – Restitution

Stage VI – Gradual Readjustment

Comment

Introduction

There are six stages of adaptation to captivity (see **Captivity, Adaptation to**); the time course over which these stages occur progressively lengthen from the first through the final stage, the initial stage lasting seconds to minutes, the next stage lasting minutes to hours, the third stage lasting the first hours to the first days, the fourth stage lasting the first days to the first weeks, the fifth stage lasting the first weeks to the first months, and the sixth stage lasting from the first months to the first years. The literature, along with my own experience, also suggests that recovery from captivity can be similarly summarized into these six stages.

Stage I – Brief Euphoria

The first stage lasts from seconds to minutes. Despite preconceived notions that the general public may have about how jubilant captives will be on release, a true rejoicing rarely materializes. Captives initially may feel euphoric, but this feeling is fleeting and any planned celebrations may fall flat. When ex-captives are returned to the United States by plane, the aircrew may have champagne aboard and decorations in place only to find that the celebration consists of a brief cheer on take-off followed a few seconds later by quiet conversations and individual contemplation.

Released captives' behavior is, in large part, guided by a deep suspicion that this release may amount to nothing more than another false hope. Emotional restraint and careful inspections of events are typical.

Ex-captives generally realize, however, that the public wishes a show of enthusiasm on their return, and they try not to disappoint it. This display, however, often tends to be hollow and fragile.

Stage II – Hyperarousal

The second stage lasts from minutes to hours. Hyperarousal at the second stage is appreciably different from the hypervigilance described during adaptation to captivity. In the hypervigilance stage, the captives' alertness is high, their observational powers are keen, and they are aware of even subtle cues regarding the behavior and motivations of their captors. In the hyperarousal state, returnees tend to be overstimulated and mentally slowed. The commotion around them, along with the great assortment of new stimuli, can render them punch drunk.

During this early stage of recovery, it is critical to protect former captives from overzealous politicians, well-wishers of all sorts, and particularly the press. Sensitive politicians will make their presence brief and depart shortly along with their entourage, well-wishers can be a significant problem in terms of crowd control, but most difficult to manage are members of the press. During the transition to freedom, former captives often have been severely sleep deprived. Yet, because returnees typically do not wish to offend any of the greeting party, they may talk at length with people who approach them.

If the period of captivity has been a long one, family members may have developed close friendships with representatives of the press. These family members may require repeated explanations as to the need to shield ex-captives from news media exploitation. In their vulnerable state, ex-captives may agree to interviews in which statements are made that they live to regret.

A treatment imperative, then, is to protect returnees from further exhaustion. Because physical examinations and psychological debriefings are often needed, the best place for these assessments may be a hospital with limited access to the public. In the case of the returned hostages from Iran, an Air Force hospital in Germany proved to be ideal for these purposes. The debriefings themselves had a positive effect for many ex-captives through the venting of strong emotions and the opportunity to obtain feedback on how they performed during captivity. The press and even family members were kept at the perimeter of the hospital. Although telephone

contact was encouraged between returnees and their family and friends, actual visits were discouraged. These restrictions allowed necessary medical examinations to be conducted in a restful and recuperative setting.

Stage III – Compliance/Resistance

The third stage lasts from hours to days. Initially returnees are likely to comply with most requests. They are, of course, accustomed to following orders from their guards and captors without much thought given to their own preferences. A danger, therefore, exists that health-care personnel may request too much. For example, ex-captives may complete hours of psychological testing or other paramedical activity not directly related to their immediate care without demurral. As the ex-captives begin to regain feelings of individual power, they start to resist such activities. Who is going to reprimand a noncompliant man or woman who has just returned from months to years of captivity stress?

The transition from compliance to resistance became obvious to physicians caring for the ex-hostage group released from Iran around the third hospital day. Initially, the former captives were extremely candid and cooperative. Most were grateful for the clinical interactions. Two to three days later, however, many started to become somewhat cavalier about keeping their appointments, began to skip ward routines, and stayed out past bedtime hours. Some limits has to be set. Early rehabilitation programs for Korean War POW, conducted as the men returned home aboard ship, were similarly disrupted.

Treatment plans should allow for this emerging independence. For example, some time from the hospital routine should be allowed for shopping, a meal out, and so forth. The hospital setting should not cause returnees to be treated as in-patients; ex-captives should wear leisure clothing instead of hospital pajamas. Doctors, nurses, and other health-care providers must watch their own countertransference reactions. Offering too much too soon and/or attempts to personally help the former captive make up for past deprivations can lead to unprofessional behaviors.

Allowing ample communications among returnees is very important. All returnees from Iran were placed in the same wing of the hospital. Between their physical, psychiatric, and dental examinations, returnees mingled in common areas, passed on information of importance to the group, and shared personal experiences. Many former hostages did not possess a full knowledge of the experiences of other captives until these exchanges.

Stage IV – Denial

The fourth stage lasts from days to weeks. After medical examinations and debriefings and a few days of rest and recuperation, released captives should begin to meet the press. This is a time that many make personal statements regarding their captivity experiences. However, one question the press asks repeatedly is: "What kinds of problems are you now having due to the stresses of your captivity?" Returnees might answer this question quite candidly on their first day or two after release, but as they begin the denial stage their answer will be: "The doctors may think that something bad may happen, but I know that I will be perfectly okay." There is a readiness of the public to believe such answers because such statements reinforce a hero image of the former captives. Such statements, however, frequently omit significant physical and/or psychological difficulties and should not be taken at face value.

It is important for health-care providers not to be taken in by such denial but also not to vigorously challenge it. Gentle reminders to continue with periodic follow-ups should be given and lines of communication should be kept open. Later, the ex-captives' possible entry into treatment, if needed, can be proposed. Often working with family members is also important at this point.

Denial also presents problems such as impulsive actions on returning home. Ex-captives are sometimes prone to buying expensive cars or investing savings accrued during captivity in poorly conceived financial schemes. Many entrepreneurs will try to use the ex-captives' notoriety to support their own enterprises.

Stage V – Restitution

The fifth stage lasts from weeks to months. Efforts toward making restitution are an interesting psychological phenomenon. The area of food is a good example. Due to generally severe and prolonged nutritional deprivation, most hostages return with a lean body mass that, combined with the plentiful exercise during captivity, renders them extremely fit. Some weight gain over the first few months home should be expected. However, moderate to gross obesity was seen in several of the returned hostages from Iran just a few months after their release. When questioned about how they had gained so much weight in such a short time, the typical answer was: "I just felt that when tempted I couldn't deny myself."

A second area of restitution occurs in readjustments among family members. Long-term captivity,

especially with extended periods of social isolation, appears to have the greatest impact. Captives' emotional and intellectual functions may have become dulled from the monotony of prison life. Table manners and bathroom and sleeping habits acquired in captivity may be unacceptably crude in the home situation. Further, family members may have become accustomed to managing their daily lives by themselves and resist direction from the ex-captive. Ex-captives may try to reestablish the family structure and function that existed before their capture and find that they are not wanted. Former captives may also try to provide, over a very short period of time, all the love and attention that they were prevented from giving family members during captivity.

A third problem area is when ex-captives return to work. Employers are frequently tempted to offer long vacations to returnees, but the returnees themselves generally wish to resume work as soon as possible. This is frequently due to a need to reaffirm their previous skills and demonstrate their worth to their organizations. In some cases, particularly those involving people held in long-term captivity, job refresher courses and even retraining are necessary. In the instances of the returned military ex-hostages held along with U.S. Embassy officials in Iran, most were given 3 months of vacation in addition to their first 2 weeks of medical treatment and official receptions. This prolonged vacation period led to problems for several men in terms of how to use this time profitably. The long time away from the military also contributed to the decision by some of the men to resign from the military and switch to other occupations that had been offered to them on their return home.

In addition, family and friends should be advised to limit recitals of the ex-captives' experiences during the first days to weeks after their return. Such early recitals can recreate suffering. In contrast, weeks to months later, giving lectures and speeches on their captivity experience can be therapeutic for releasees. If ex-captives have a talent for writing, the preparation of an article or book about their experiences may help them in working through their captivity experiences.

Stage VI – Gradual Readjustment

The sixth stage lasts from months to years. If the criteria for readjustment from captivity stress simply included whether former captives returned to work or whether their marriages remained intact, the majority would be declared as recovered. If the criterion for recovery is freedom from the psychological scars of captivity, complete recovery may never occur.

Life-long anxieties and depression have been seen in victims of especially severe captivity stress, such as Nazi concentration camp survivors. Hardly a day goes by that they do not suffer distressing psychiatric symptoms emanating from their Holocaust experiences.

Illnesses that develop following ex-captives' return are often an indicative of incomplete recovery. In follow-up studies of American POWs from Korea, as well as from Asian POW camps in World War II (where torture and nutritional deprivation were extremely severe), returned POWs showed significantly increased illness rates for infectious, cardiovascular, degenerative, and psychiatric disorders for the next 25–30 years. Even accidental deaths were higher for returned captives than for controls during this time. In contrast, when U.S. Navy aviators returned from prisoner-of-war captivity during the Vietnam War were similarly studied, the elevated illness rates for physical and psychological illness returned to control-group levels 8 years following their return. Thus, even with the best outcome, these studies point to an enduring effect of severe and prolonged captivity on physical and psychological health.

Some returnees do better than others in terms of illness during their subsequent years of life. How can medical personnel learn to correctly identify people who will probably do well versus those who will probably have multiple problems? The answer is: carry out an illness assessment 1 year after return. When 133 Navy returned prisoners of war from Vietnam were seen after their first year of return, they fell into two groups: The first group, comprising 96 men, reported good health over the previous year; the second group, comprising 37 men, reported significant psychiatric illnesses. The most prevalent psychiatric illnesses were substance abuse and adjustment reaction. As shown in [Figure 1](#), more than two-thirds of men who were healthy at follow-up year 1 stayed healthy during follow-up years 2–5. Of the group (30 men) that reported illnesses during years 2–5, 63% of the illnesses were minor in severity (e.g., mild anxiety disorder). In contrast, of the group that was ill in year 1, more than three-quarters continued to report significant illnesses in years 2–5 ([Figure 2](#)). Further, 59% of these reported illnesses were of major significance (e.g., major depressive disorder).

Although the sample is smaller, this same phenomenon was observed in the Americans held hostage in Iran. As [Figure 3](#) shows, 21 of the 35 former hostages studied reported good health during the first follow-up year. Of this healthy group, two-thirds remained healthy during the years 2–5. Of the group who were ill during years 2–5, 86% reported that these illnesses were minor in severity. In contrast, in the

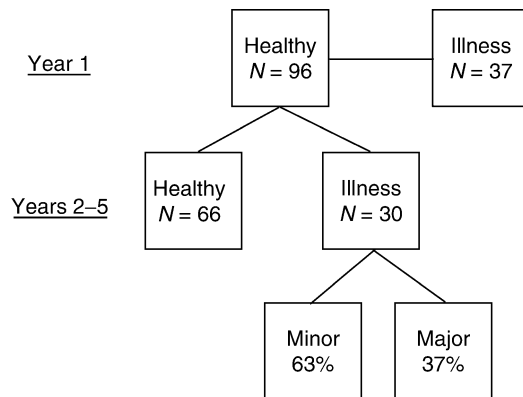


Figure 1 More than two-thirds of former prisoners of war (returned Navy Vietnam POWs; $N = 133$) who were healthy 1 year after release were also healthy 5 years after release. Sixty-three percent of those who were ill at the fifth follow-up year reported illnesses minor in severity.

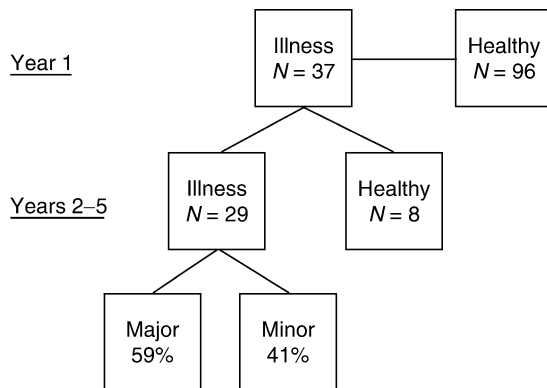


Figure 2 More than three-quarters of former prisoners of war (returned Navy Vietnam POWs; $N = 133$) who were ill 1 year after release were also in ill health 5 years after release. Fifty-nine percent of those who were ill at the fifth follow-up year reported illnesses of major severity.

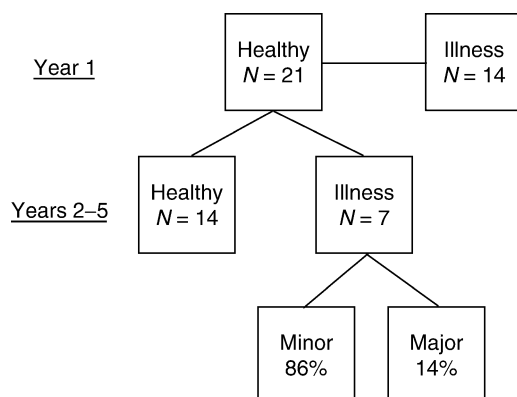


Figure 3 Two-thirds of former American hostages held in Iran ($N = 35$) who were healthy 1 year after release were also healthy 5 years after release. Eighty-six percent of those who were ill at the fifth follow-up year reported illnesses minor in severity.

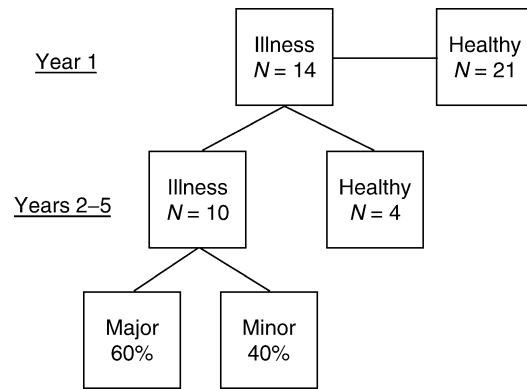


Figure 4 More than 70% of former American hostages held in Iran ($N = 35$) who were ill 1 year after release were also in ill health 5 years after release. Sixty percent of those who were ill at the fifth follow-up year reported illnesses of major severity.

group (14 men) that was initially ill, more than 70% remained ill during years 2–5 (Figure 4). In addition, 60% of the illnesses reported were major in severity.

Comment

The chances to study returned captives from stressful incarcerations, such as prisoners of war or political hostages, are rare. The military and in the U.S. Department of Veterans Affairs have conducted the bulk of such investigations. Follow-up studies of World War II POWs liberated from Japan, POWs returned from the Korean War, and former U.S. Navy POWs held in North Vietnam have spanned 20 years or more. A recent report assessing psychiatric illnesses in the Navy POW group found a 20-year prevalence rate of just 10%. In this day of labeling vast numbers of military stress survivors as having posttraumatic stress disorder (PTSD), the long-term psychological health of this Navy group of severely stressed individuals is truly remarkable.

Part of the answer to the long-term health of these Navy POWs is that they were a group of highly intelligent and talented aviators. Throughout their lives, these individuals had responded to challenges in their lives with success. Many then saw the challenges of surviving being prisoners of war as yet another life challenge that they would overcome. And most of them did so. Such individuals serve as an inspiration to all of us by their indomitable spirit in the face of enormous adversity.

See Also the Following Articles

Cancer Treatment; Captivity, Adaptation to; Combat Reaction, Chronic; Combat, Acute Reactions to; Coping and Stress: A Lens and Filter Model; Defense Services, Stress in.

Further Reading

- Beebe, G. W. (1975). Follow-up study of World War II and Korean War prisoners. II: Morbidity, disability, and maladjustment. *American Journal of Epidemiology* **101**, 400–422.
- Blakey, S. (1978). *Prisoner at war. the survival of Commander Richard A. Stratton*. Garden City, NY: Anchor Press/Doubleday.
- Brill, N. Q. (1946). Neuropsychiatric examination of military personnel recovered from Japanese prison camps. *Bulletin of the US Army Medical Department* **5**, 429–438.
- Clavell, J. (1962). *King rat*. New York: Dell Publishers.
- Engdahl, B., Dikel, T. N., Eberly, R., et al. (1997). Posttraumatic stress disorder in a community group of former prisoners of war: a normative response to severe trauma. *American Journal of Psychiatry* **154**, 1576–1581.
- Hall, R. C. W. and Malone, P. T. (1976). Psychiatric effects of prolonged Asian captivity: a two-year follow-up. *American Journal of Psychiatry* **133**, 786–790.
- Hall, R. C. W. and Simmons, W. C. (1973). The POW wife. A psychiatric appraisal. *Archives of General Psychiatry* **29**, 690–694.
- Matusek, P. (1975). *Internment in concentration camps and its consequences*. New York: Springer-Verlag.
- Neizger, M. D. (1970). Follow-up study of World War II and Korean War prisoners. I: Morbidity, disability, and maladjustment. *American Journal of Epidemiology* **91**, 123–138.
- Nice, D. S., Garland, C. F., Hilton, S. M., et al. (1996). Long-term health outcomes and medical effects of torture among US Navy prisoners of war in Vietnam. *Journal of the American Medical Association* **276**, 375–381.
- Richlin, M., Shale, J. H. and Rahe, R. H. (1980). Five-year medical follow-up of Navy POWs repatriated from Vietnam. *US Navy Medicine* **71**, 1926.
- Segal, H. A. (1954). Initial psychiatric findings of recently repatriated POWs. *American Journal of Psychiatry* **111**, 358–363.
- Segal, J. (1973). Therapeutic considerations in planning the return of American POWs to continental United States. *Military Medicine* **2**, 73–77.
- Spaulding, R. C. (1977). The Pueblo incident: medical problems reported during captivity and physical findings at the time of release. *Military Medicine* **142**, 681–684.

Cardiovascular System and Stress

P Hjelm Dahl

Karolinska Institute and University Hospital, Stockholm, Sweden

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Hjelm Dahl, volume 1, pp 389–403, © 2000, Elsevier Inc.

Background

Historical Perspective

Methodological Comments

Neurohormonal Mechanisms

Baroreflexes and Blood Pressure Regulation

Human Studies of Cardiovascular Responses to Mental Stress

Pathophysiological Aspects

Glossary

Epinephrine (EPI)

The hormonal link of the sympatho-adrenal system; epinephrine is released into the bloodstream from the adrenal glands.

Norepinephrine

The principal neurotransmitter (chemical signaling substance) released from the sympathetic nerves. A neurotransmitter acts locally where it is released. Dream sleep.

Rapid eye movement (REM) sleep Receptors

Molecular structures that carry messages into cells from biological substances (e.g., neurotransmitters and hormones) outside of the cells that they influence. Sympathetic receptors are also termed adrenoceptors.

Renin-angiotensin system

Renin is released from the kidney and regulates the formation of angiotensin II, which is a potent vasoconstrictor and has many other functions. β -Adrenoceptor stimulation increases renin release.

Stroke volume

The volume of blood pumped into the vascular system by one heart beat.

Sympathetic neuroeffector mechanisms

The regulation of sympathetic nerve function (i.e., nerve activity), the release of norepinephrine, and its effects on the cells that it influences via the stimulation of specific receptors (α - or β -adrenoceptors).

<i>Sympathetic system</i>	One of the two classic autonomic systems of nerves that regulate bodily functions. Sympathetic nerves to the heart increase heart rate and contractility via β -adrenoceptor stimulation by norepinephrine, whereas the parasympathetic vagal nerve slows the heart rate via release of its neurotransmitter acetylcholine.
<i>Tachycardia</i>	A rapid heart rate.
<i>Thrombosis</i>	Blood clotting that blocks the flow of blood.
<i>Vasoconstriction</i>	The narrowing of a blood vessel to decrease blood flow.
<i>Vasodilatation</i>	The widening of a blood vessel to increase blood flow.

Background

The cardiovascular system adapts to stress by changes that, from a teleological perspective, are meant to improve the chances of survival when the organism is threatened. The response pattern has been called the defense reaction and serves, in its simplest form, to prepare the organism for fight or flight. Animal experiments by Hilton, Folkow, and others have established a typical pattern of cardiovascular responses to stress, which may be elicited by the electrical stimulation of certain regions of the brain as well as by environmental stimuli in awake animals. The defense reaction involves elevations of cardiac output and blood pressure and a redirection of blood flow from the kidneys and splanchnic organs to skeletal muscle. A similar response pattern is elicited by mental stress in humans. The short-term survival benefits of the changes associated with the defense reaction may not be paralleled by beneficial consequences in the long term, when elicited repeatedly in individuals living in an urban society with everyday stresses.

Research on the importance of psychosocial factors and stress in cardiovascular disease has been rather intense. There is consensus that stress is of importance, but its exact role in cardiovascular pathophysiology has been somewhat illusive for many reasons: what is stress? how do positive and negative stressors differ? how can stress be quantitated? is laboratory stress testing relevant for the everyday situation? and how do personality factors influence cardiovascular responses to stress and for consequences of these responses? Despite this complexity, it is clear that stress (in terms of stimuli that influence the cardiovascular system in a certain direction) is involved in cardiovascular disease and its consequences. This presentation focuses on the effects of mental stress on the cardiovascular system and the mechanisms probably involved, with a few comparisons with the effects of physical exercise and other stressors.

Historical Perspective

From a cardiovascular point of view, it is pertinent to mention the work of Walter B. Cannon, who, in the first decades of the twentieth century, studied the responses of animals to stress. He described the fight-or-flight reaction and a homeostatic theory of how organisms adapt to threatening environmental influences. Cannon recognized the importance of the sympathetic system and proposed that the adrenal glands secreted sympathin, which mediated the responses to stress. However, von Euler could later show that sympathetic nerves use the neurotransmitter norepinephrine, whereas the circulating hormone secreted by the adrenal glands was a related substance, epinephrine. Later, adrenocorticotrophic hormone (ACTH) and glucocorticoids secreted from the adrenal cortex were added to the list of stress hormones. Goldstein reviewed the homeostatic mechanisms in relation to research on stress and distress from Darwin and Bernard to 1995. He noted the lack of a unifying concept of stress and that this reflects the complexity with which the organism may respond to stress due to feedback regulation and the multiplicity of systems involved in the control of homeostasis.

Methodological Comments

In cardiovascular stress research, the conditions for resting measurements are most important because any expression of cardiovascular reactivity to stress depends on the basal value with which the stress value is compared. In fact, the basality of resting measurements (as reflected by, e.g., heart rate and plasma epinephrine levels) may be the most important determinant of study results. The stress test used should, of course, be standardized when comparing stress reactivity in individuals or between groups. Responses should be of sufficient amplitude to allow the discrimination between different conditions and be reasonably reproducible. Many factors, including personality factors, modulate the perception of stress and the degree of arousal caused by the stressor, as well as cardiovascular responses elicited by it (Figure 1).

What Is a Relevant Stressor?

For a cardiologist, stress testing means evaluating whether myocardial ischemia can be provoked by a stimulus that increases myocardial oxygen demand, such as physical exercise. However, mental stress may also provoke ischemia. Any increase in mental activity – even quiet speaking and REM periods (dreaming) during sleep – evokes cardiovascular changes. In fact, REM sleep may even precipitate

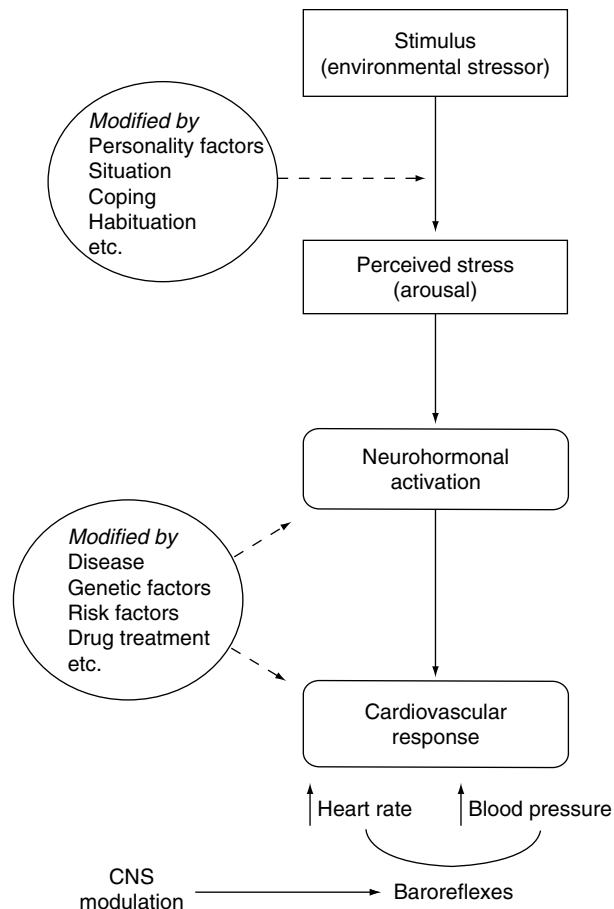


Figure 1 Stimulus–response scheme for cardiovascular responses to psychological stress. The perception of stress is modified by multiple factors, which reinforce or attenuate the arousal evoked by the stimulus. The cardiovascular response is mediated by neurohormonal activation. Several factors may modify the neurohormonal activation, as well as responses to the mediating neurohormones. Finally, the cardiovascular responses may be buffered by baroreflexes aiming at maintaining blood pressure at prestimulus levels, but the efficiency of the baroreflex is attenuated during stress. CNS, central nervous system.

myocardial infarction. Interviews on sensitive personal issues may be most relevant when studying the cardiovascular risks of stress, but they cannot be standardized for comparing individuals or groups. The question is whether laboratory stress tests mimic responses elicited by mental stimuli normally encountered. Herd used an operational definition of relevant stressors whereby “a psychological stressor is one that causes a predictable stress response in a predictable percentage of index subjects” and went on to define the cardiovascular stress response as changes seen in the defense reaction (1991). Various standardized stress tests have been developed that fit this definition.

Among the most common mental stress tests for cardiovascular research are mental arithmetic (MA)

and Stroop’s color word conflict test (CWT). Tests may also be based on auditory feedback or video games. MA is simple to administer, but the test results are influenced by the speed of the test and by the (variable?) harassment of the subject; individuals may also have very different abilities for coping with mathematics. MA produces substantial responses in some studies (e.g., those of Brod and coworkers), but weak responses in other studies. CWT is administered in different ways, and the magnitudes of responses depend on the version used (how color words are presented, how responses are made, and whether there is auditory interference). We have used MA (in various versions, with or without external pacing) and CWT in a modification with auditory interference developed by Frankenhaeuser in the 1960s and have found the latter to be particularly useful.

Adaptation to the same stressor occurs on repeated administration, whereas responses to a novel stressor are either intact or enhanced. Such adaptation involves information processing (coping, etc.) rather than the desensitization of the target organs. Animal experiments suggest that chronic stress can even result in enhanced responses to a novel stressor due to increased synthesis, storage, and release of neurotransmitters.

In clinical research, it may be difficult to find newly diagnosed patients who have not received any treatment. When drug treatment is discontinued, the researcher should minimize carryover effects or rebound phenomena by allowing sufficient time for the treatment effects to disappear. For example, diuretics may have long-lasting effects on sodium and fluid balance, and the withdrawal of a β -blocker may lead to rebound phenomena with increased β -adrenoceptor sensitivity. Such aftereffects may persist longer than is appreciated by many authors and should be taken into account.

Neurohormonal Mechanisms

Before describing the cardiovascular responses to mental stress, I briefly comment on the neurohormonal systems that may be involved and that are discussed in relation to these responses. In general, neurogenic control is suited for fine-tuned influences on individual organs or cells, whereas hormonal influences are diffuse.

The Sympathoadrenal System

The sympathoadrenal system is involved in short-term (phasic) responses to stress. Epinephrine (adrenaline), which is often considered to be *the* stress hormone, is secreted during stress, and epinephrine

levels in plasma or urine reflect the degree of arousal experienced.

The sympathetic nerves may operate in a differentiated fashion, with varying levels of activity and varying responses to stress in different organs. Muscle sympathetic nerve activity during stress may even differ in different parts of the body. Thus, the common term sympathetic tone, which indicates synchronized activity throughout the body, is not appropriate. If possible, sympathetic activity should be studied locally, in the organ of interest.

Norepinephrine is the principal neurotransmitter of sympathetic nerves, although the cotransmitter neuropeptide Y may contribute to vasoconstrictor responses when sympathetic nerves fire at high frequencies. Norepinephrine stimulates α - and β -adrenoceptors. Catecholamines are rapidly cleared from plasma, with half-lives of 1–2 min, but the effect of a sympathetic nerve impulse is much shorter than that, due to local inactivating mechanisms such as the reuptake of norepinephrine into the nerve. Norepinephrine may also function as a circulating hormone when the plasma levels of norepinephrine are markedly elevated, such as during physical exercise. Resting levels of norepinephrine in plasma are $\approx 1 \text{ nmol l}^{-1}$, and threshold levels for hormonal effects are 3–4 nmol l^{-1} . Sympathetic nerve activity may be monitored in humans by:

- Direct recordings of impulse activity in nerves supplying skeletal muscle and skin but not in organs of perhaps greater interest in the context of mental stress.
- Analyses of heart rate variability (for cardiac sympathetic activity).
- Measurements of norepinephrine levels in plasma and norepinephrine release from various organs. Measurements of norepinephrine release and spill-over into plasma can be refined by the use of an isotope dilution method involving infusions of radiolabeled norepinephrine, which was developed by Esler and coworkers.

These techniques all have their technical difficulties and limitations. They complement one another because the release of norepinephrine from the sympathetic nerves is influenced by prejunctional regulation (i.e., variable release per nerve impulse) and the sensitivity of the receptors involved may vary when measuring functional responses, such as heart rate.

Epinephrine may, as a circulating hormone, stimulate α - and β -adrenoceptors in all organs. Under certain circumstances, epinephrine can be taken up into and re-released from sympathetic nerves in such quantities that a role as a co-transmitter is possible, but this is normally not important. Thus, epinephrine

is the classic hormonal link of the sympathoadrenal system.

There is functional differentiation of sympathetic neuroeffector mechanisms; β_1 -adrenoceptors appear to be innervated receptors, and β_2 -adrenoceptors appear to be noninnervated or humoral receptors. This leads to epinephrine being functionally β_2 -selective at physiological concentrations in blood plasma (even though it is also an effective β_1 -agonist in *in vitro* systems) and provides a possibility of studying the importance of epinephrine for physiological responses to stress by examining the β -adrenoceptor subtype involved in the response. Similarly, α_1 -adrenoceptors appear to be innervated and α_2 -adrenoceptors non-innervated in the cardiovascular system, which has therapeutic implications and allows the separation of neurogenic and humoral influences.

Cholinergic Mechanisms

The heart is under dual autonomic control, with excitatory sympathetic nerves (stimulating β_1 -adrenoceptors) and inhibitory vagal nerves (releasing acetylcholine and stimulating muscarinic receptors). Acetylcholine is very rapidly inactivated after its release, allowing beat-to-beat control of the heart rate. The intrinsic rate of the sinus node is ≈ 90 –100 beats/min. Thus, the resting heart rate (usually 60–80 beats/min) is controlled to a large extent by vagal activity. Increased physical fitness reduces the resting heart rate by increasing vagal activity. Stress may thus increase heart rate by activating sympathetic nerves and/or by inhibiting the parasympathetic vagus nerve.

Animal research on vasodilator responses in skeletal muscle has demonstrated sympathetic vasodilator nerves that release acetylcholine. Work by Hilton suggested that such vasodilator nerves did not exist in primates, but early studies in humans suggested that it indeed was possible to elicit cholinergic vasodilation in the forearm with certain stressors. However, the literature on cholinergic vasodilation in human skeletal muscle is divided, and it cannot be definitely stated that this mechanism is of importance or under which conditions it may contribute.

Other Mediators and Mechanisms

The renin–angiotensin system may be activated by stress. Renin release from the kidney is under the influence of several factors, one of which is sympathetic nerve activity and β_1 -adrenoceptor stimulation. Salt deprivation increases renin secretion, and salt loading decreases it. Renin catalyzes the formation of angiotensin II, which is a potent vasoconstrictor peptide with several other biological actions, including the enhancement of aldosterone secretion. Thus, the system is salt-conserving. The relative importance

of angiotensin II and norepinephrine as mediators of vasoconstriction during stress may vary. High sodium intake tends to enhance cardiovascular responses to mental stress, but this is most likely independent of the renin–angiotensin system, which should be downregulated under such conditions.

There are several neuropeptides (e.g., neuropeptide Y and calcitonin gene-related peptide), and other peptides (e.g., atrial natriuretic factor and endothelin) that may be involved in cardiovascular responses to stress. Their roles during stress have not been sufficiently elucidated, in part due to the lack of good antagonists for human use.

Platelets store serotonin (5-hydroxytryptamine), which is released when the cells are activated, and activated platelets synthesize thromboxane A₂. Both of these substances may stimulate platelets (positive feedback) and cause vasoconstriction, especially when the vascular endothelium is not functioning normally. Stress may activate platelets.

Flow-mediated vasodilatation has been recognized for a long time. Furchgott showed the importance of endothelial cells for vasodilator responses and proposed the existence of an endothelium-derived relaxing factor (EDRF). Nitric oxide (NO) was shown by others to be an important component, if not the only EDRF. NO released from the vascular endothelium serves both as a vasodilator and an inhibitor of platelet adhesion and aggregation, and the inhibition of NO synthesis or defective NO production by the endothelial cells may increase vascular tone and the risk of thrombosis. NO appears to be a rather ubiquitous substance in the body, which may be involved in neurogenic mechanisms and influence cardiac function as well. However, the importance of this for stress responses of the cardiovascular system is unknown. Flow-mediated vasodilatation (i.e., NO) is probably important for vasodilator responses to stress in some organs.

Baroreflexes and Blood Pressure Regulation

The cardiovascular system is regulated in a complex manner, with compensatory adjustments when changes are elicited by various stimuli. Normally this fine-tuned system strives to maintain a certain blood pressure (which may be either normal or high, if the individual has hypertension) via arterial baroreflexes, which serve to dampen the change. If the blood pressure rises, there is a reflexogenic inhibition of the heart and reduced vasoconstrictor nerve activity; conversely, if the blood pressure decreases there is a reflexogenic stimulation of the heart with increases in heart rate and cardiac output and a compensatory

increase in vasoconstrictor nerve activity. If cardiac output is lowered (e.g., by drug treatment), peripheral blood flow will decrease to maintain the blood pressure. In the long term, baroreflexes are reset to the new blood pressure level if, for example, a patient with hypertension is treated with a blood pressure lowering drug.

There are also receptors that sense venous pressure and participate in volume regulation. If venous pressure decreases, reflexogenic and hormonal changes occur that serve to conserve salt and water (i.e., renal responses) and to maintain blood pressure by increasing venous tone and reducing peripheral blood flow. The kidney is intimately involved in the regulation of blood pressure according to the Guyton's classic theory of cardiovascular homeostasis. However, there are also indications that blood pressure is a regulated variable and that neurogenic mechanisms are important for both short- and long-term levels of blood pressure.

During stress, the influence of baroreflex regulation of the circulation diminishes, and centrally elicited pressor mechanisms may operate under less influence of reflexogenic modulation than if blood pressure is increased by other maneuvers (Figure 1). As discussed by Julius, there seems to be some blood pressure-sensing mechanism that regulates the blood pressure during stress because interference with one pressor mechanism often leads to another one taking over and the pressor response to the stimulus remaining intact.

Human Studies of Cardiovascular Responses to Mental Stress

Figure 2 illustrates schematically how mental stress elevates blood pressure in humans, based on results from a number of studies. Pressor responses to mental stress generally depend on the increase in cardiac output because systemic vascular resistance decreases. In an early series of investigations, Brod and coworkers studied the cardiovascular responses to mental stress, mainly forced MA (using procedures creating substantial hemodynamic responses), in normotensive and hypertensive subjects. An example is given in Figure 3, which shows that arousal caused by the suggestion of exercise elicited responses that were qualitatively similar to those from mental arithmetic. A few examples of responses to CWT, from our work, are mentioned next.

Cardiac Responses

The increase in cardiac output is caused mainly by an increase in heart rate, but an increase in stroke volume also contributes (Figure 2). Tachycardia may

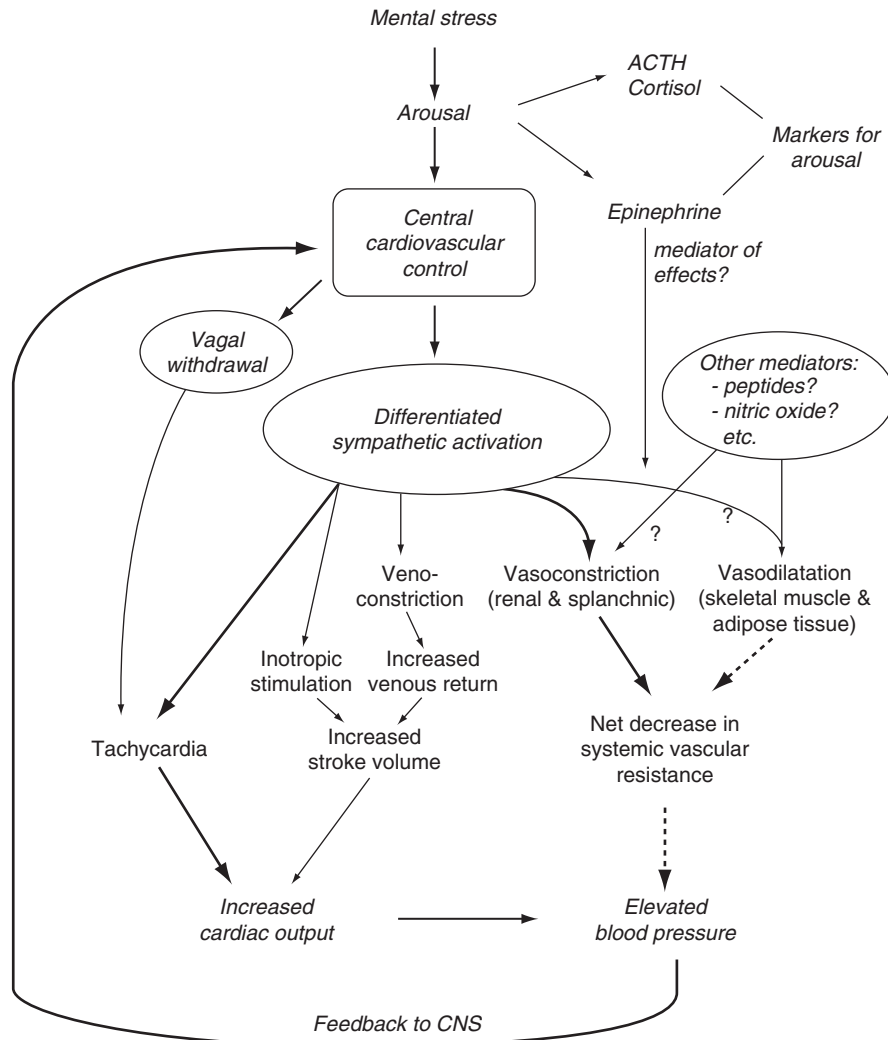


Figure 2 Schematic representation of the cardiovascular response pattern of the defense reaction. The pressor response is mainly cardiac output-dependent, and the increase in cardiac output depends mainly on an increase in heart rate. The vascular response is a net reduction of systemic vascular resistance, but responses of individual organs vary considerably. ACTH, cortisol, and epinephrine are markers of the arousal evoked by the stressor. Established and possible mediators of the cardiovascular responses are indicated. ACTH, adrenocorticotropic hormone; CNS, central nervous system. Modified from Hjelmahl, P. (1993), Plasma catecholamines – analytical challenges and physiological limitations, *Baillière's Clinical Endocrinology and Metabolism* 7, 307–353.

be elicited by both increased sympathetic nerve activity and vagal withdrawal. Increases in stroke volume are related to increased contractility (which can be demonstrated by, e.g., echocardiography), as well as an increase in venous return, which supports the filling of the heart in diastole. Thus, work by Brod and coworkers has demonstrated that stress causes venoconstriction, which results in the centralization of the blood volume and supports cardiac filling. Studies with pacing alone showed that an isolated increase in heart rate does not efficiently elevate cardiac output – an increase in contractility and an adequate distension of the heart in diastole contribute to the stress response.

The effects of β -blockade by metoprolol (β_1 -selective) or propranolol (nonselective) on some invasive hemodynamic variables are illustrated in [Figure 4](#). Healthy volunteers were subjected to CWT (the Frankenhaeuser version), which was found to elicit clear-cut and rather reproducible (see control group) responses. The two β -blockers inhibited the heart rate response similarly, indicating that the sympathetic component was β_1 -mediated (neurogenic). The vagal component in the tachycardia elicited by mental stress is reflected in the persisting, although blunted, tachycardia seen after β -blockade in sufficient doses to eliminate the increase in stroke volume during stress ([Figure 4](#)). The stroke volume response

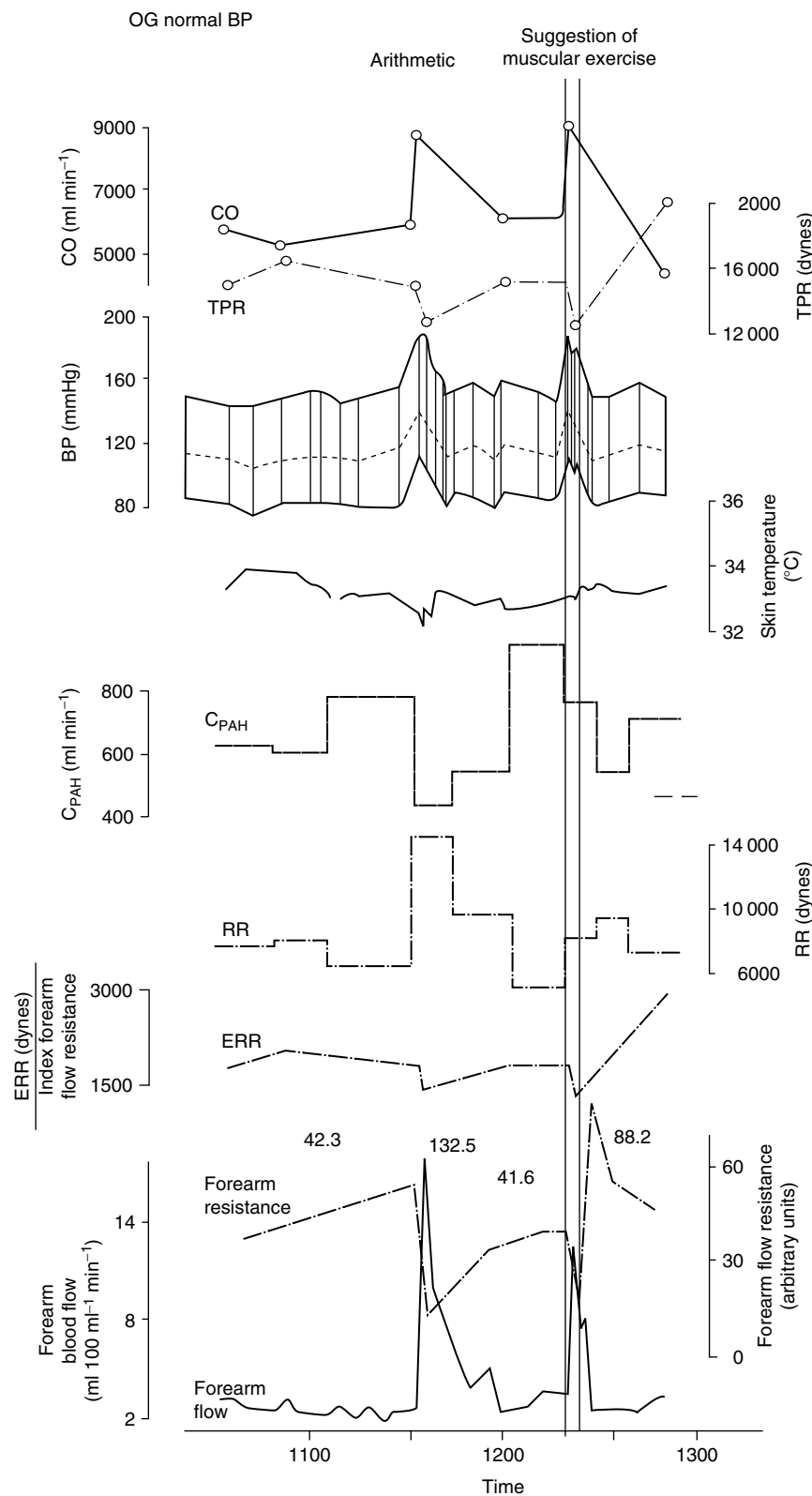


Figure 3 Comparison of general and regional hemodynamic responses to mental arithmetic stress and suggestion of heavy muscular exercise in a healthy subject from the studies of Brod and coworkers. Mental stress and the anticipation of physical stress elicited similar cardiovascular responses, both of which resemble the defense reaction. Illustrated are cardiac output (CO), systemic vascular resistance (TPR), blood pressure (BP), skin temperature, renal paraaminohippurate clearance (C_{PAH}), renal vascular resistance (RR), extrarenal vascular resistance (ERR), forearm vascular resistance, and forearm blood flow. From Brod, J. (1963), Haemodynamic basis of acute pressor reactions and hypertension, *British Heart Journal* **25**, 227–245, with permission from the BMJ Publishing Group.

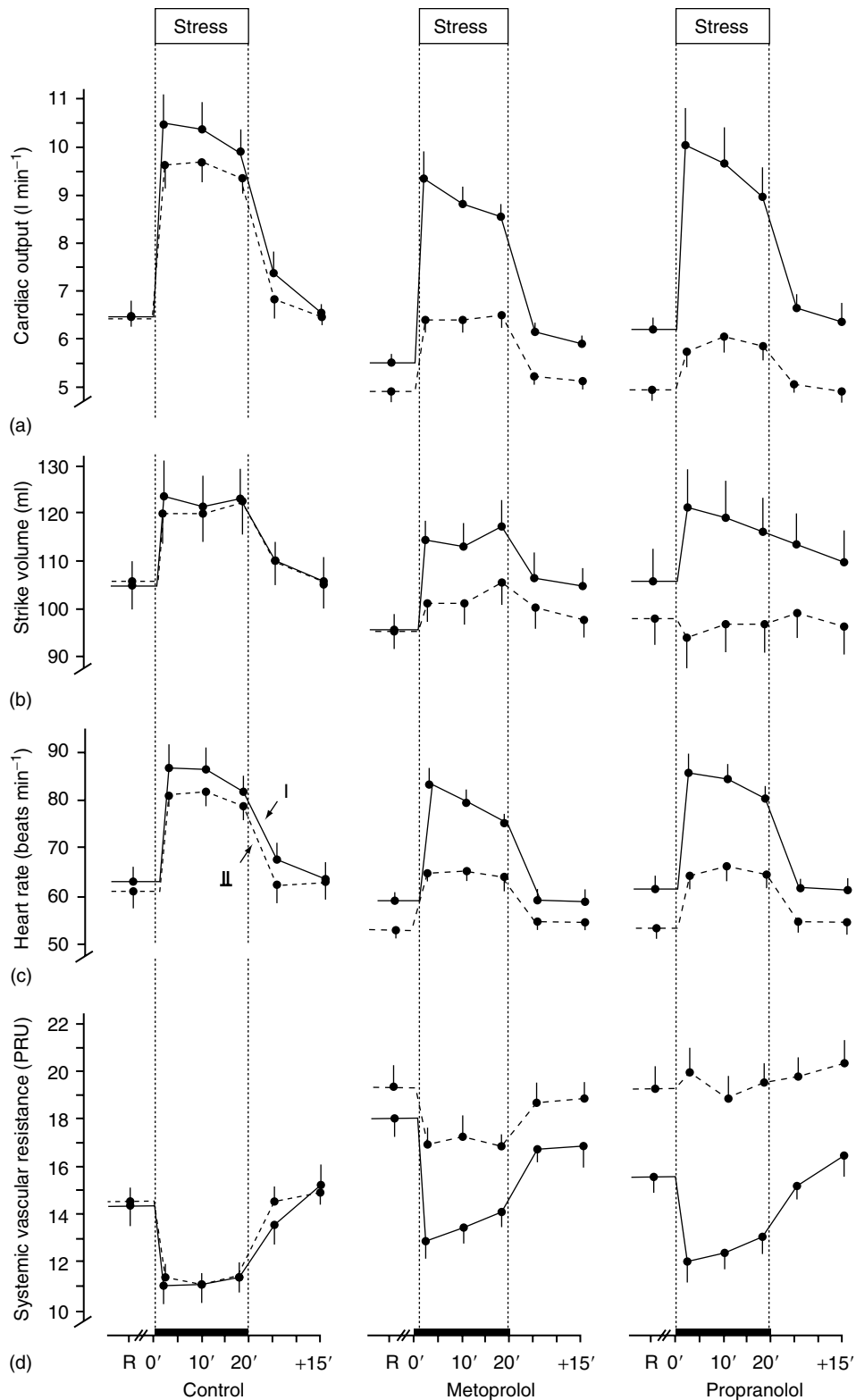


Figure 4 Cardiac and systemic vascular responses to mental stress induced by Stroop's color word conflict test (CWT; the Frankenhaeuser version) in healthy male volunteers. a, Cardiac output (measured by thermodilution); b, stroke volume; c, heart rate; d, systemic vascular resistance. Thirty subjects were divided into three groups, in which CWT (stress) was repeated without treatment (control) or after intravenous injection of the β_1 -blocker metoprolol or the nonselective β -blocker propranolol. The mean arterial pressure response to stress (≈ 20 mmHg) was little affected by the treatments given. PRU, peripheral resistance units. From Freyschuss et al. (1988), *American Journal of Physiology* 255, H443–H451, with permission from the American Physiological Society.

was attenuated by metoprolol and abolished by propranolol, suggesting contributions of both neurogenic (norepinephrine) and hormonal (epinephrine) effects. However, stroke volume is also sensitive to afterload, and the effects of β -blockade on systemic vascular resistance may have influenced the stroke volume response to stress.

The coronary circulation is to a large extent regulated by cardiac metabolism, and it is difficult to separate β -adrenoceptor-mediated coronary vasodilation from that induced by increased cardiac metabolism and flow-mediated vasodilation. However, α -adrenergic coronary vasoconstriction may occur when the endothelium is dysfunctional or when cardiac metabolism is inhibited.

Peripheral Vascular Responses

The decrease in systemic vascular resistance during mental stress is related to vasodilation in skeletal muscle and adipose tissue, whereas there is vasoconstriction in the kidneys, splanchnic organs, and skin (Figure 2). The experiment of Brod and coworkers (Figure 3) illustrates marked vasodilation in the forearm simultaneously with reduced renal clearance of para-aminohippurate (i.e., renal vasoconstriction) and a decrease in skin temperature (i.e., vasoconstriction in skin); other experiments have shown splanchnic vasoconstriction.

Limb blood flow The decrease in systemic vascular resistance in the experiments illustrated in Figure 4 was attenuated by β_1 -blockade and abolished by propranolol. These effects of β -blockade may reflect blockade of vascular β -adrenoceptors (which are of the β_2 subtype) and/or peripheral vascular adaptation to the effects of β -blockade on the cardiac response to stress. The increase in blood pressure during mental stress was little influenced by β -blockade (data not shown), which is in agreement with the results of several other studies. If, as discussed by Julius, blood pressure is the primarily regulated variable, the attenuation of the increase in cardiac output results in compensatory changes in vascular responses in order to achieve the same level of blood pressure during stress as would be obtained without such attenuation. Vasodilator responses are then secondary to, and governed by, cardiac responses. Further support for this idea was obtained in studies of systemic NO synthesis inhibition, which increased vascular resistance and blood pressure at rest but did not influence the blood pressure level achieved during stress.

These experiments included measurements of calf blood flow (by plethysmography) and of adipose tissue and skeletal muscle blood flow (by isotope

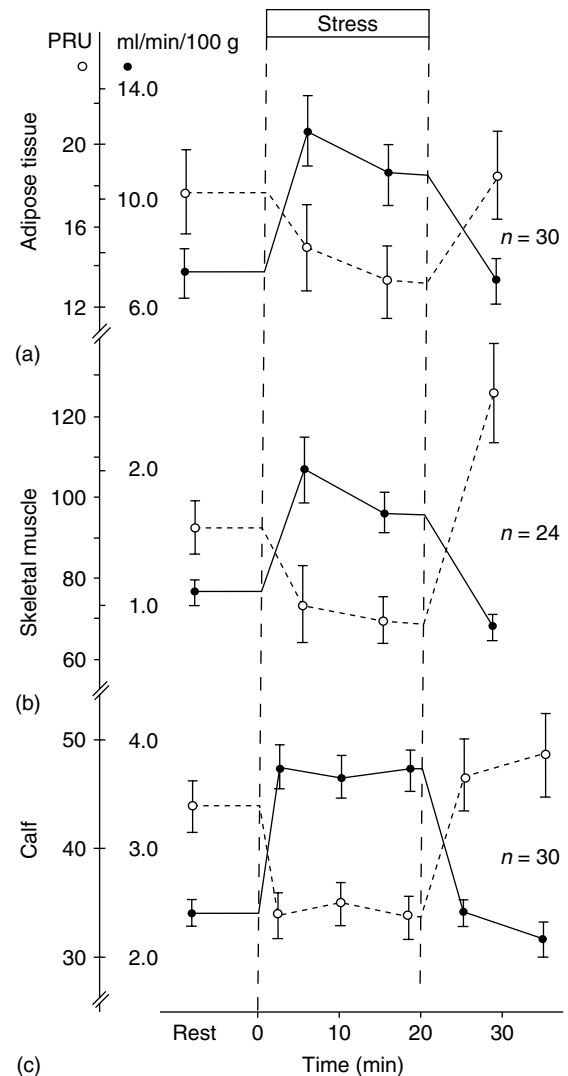


Figure 5 Peripheral vascular responses to mental stress (Stroop's color word conflict test, CWT) in the experiments illustrated in Figure 4. a, Adipose tissue; b, skeletal muscle; c, calf. CWT elicited similar vasodilation in adipose tissue of the abdomen (similar responses were found in adipose tissue on the thigh), skeletal muscle (gastrocnemius muscle), and the calf. Tissue blood flows were determined by the local clearance of ^{133}Xe ; and calf blood flow was measured by venous occlusion plethysmography. PRU, peripheral resistance units. From Linde et al. (1989), *American Journal of Physiology* 256, E12–E18, with permission from the American Physiological Society.

clearance methods). Figure 5 shows similar vasodilation in adipose tissue, skeletal muscle (the gastrocnemius muscle) and the entire calf in response to CWT. β -Blockade inhibited these local vasodilator responses similarly, as previously shown for systemic vascular resistance. In other studies, even greater vasodilator responses than those seen in the leg were found in the human forearm with the same stress test (Figure 6). Of interest are findings by Rusch and coworkers that blood flow responses of the arm and leg

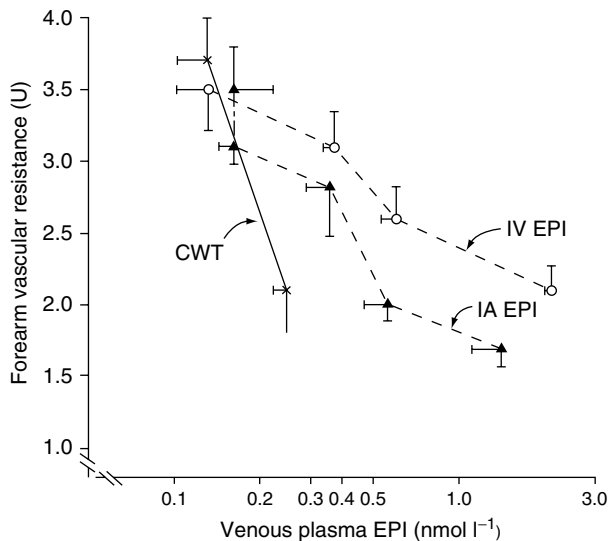


Figure 6 Forearm vascular responses to Stroop's color word conflict test (CWT) and to intravenous (IV) and intra-arterial (IA) infusions of epinephrine (EPI), in relation to epinephrine concentrations before and during interventions. CWT elicited a marked forearm vasodilator response with a small increase in venous plasma epinephrine. Responses to IA epinephrine were greater than those to IV epinephrine, suggesting reflexogenic buffering of the forearm vasodilator response to systemic epinephrine. Interpolation of vasodilator responses to the epinephrine concentration achieved by the same individual during CWT showed that epinephrine may have contributed 10–30% of the vasodilator response to stress. From Lindqvist et al. (1996), *American Journal of Physiology* 270, E393–E399, with permission from the American Physiological Society.

to different stimuli (including MA) differed and that vasodilatation was more prominent in the arm than in the leg. Others have found that differences in the baroreceptor mediated the control of limb blood flow in the leg and the forearm. Recordings of muscle sympathetic nerve activity by Anderson and co-workers showed that MA increased muscle sympathetic activity in the leg but not in the arm, when they were simultaneously recorded (Figure 7). Thus, the sympathetic response to mental stress is so differentiated that differences in muscle sympathetic nerve activity can be seen even between the upper and lower limbs. Increases in muscle sympathetic nerve activity may serve to buffer the vasodilator response to stress.

NO release and flow-mediated vasodilatation may be involved in the blood flow response of skeletal muscle to mental stress. Other possible mediators are acetylcholine (although we found no neurogenic component in the sustained forearm vasodilator response to CWT), vasodilating peptides, and epinephrine. Of interest is that the forearm vasodilator response to CWT (a 50% decrease of local vascular resistance) was not altered by systemic NO synthesis

inhibition, even though previous studies had shown attenuation of stress-induced forearm vasodilatation with local inhibition of NO synthesis. Thus, systemic (centrally mediated?) influences on muscle blood flow may override locally acting influences by, for example, NO released from the endothelium.

Thus, it is well established from work by several groups and with several different stress tests that limb blood flow increases due to vasodilatation in the skeletal muscle and adipose tissue, whereas there is vasoconstriction in the skin. Interestingly, skeletal muscle vasodilatation appears to differ in different regions of the body. The mechanisms involved are not entirely clear.

Renal blood flow The kidneys, which receive $\approx 20\%$ of the cardiac output, are important in the regulation of blood pressure – both in the short term, and in the long term – and appear to be a target for stress effects. Several groups have shown renal vasoconstriction, renin release, and/or altered renal function in response to mental stress. Examples are shown in Figure 3 for the work of Brod and co-workers and in Table 1 for our work. Of particular interest is that renal vascular responses to mental stress may be enhanced in hypertension. Thus, Hollenberg and coworkers found that very mild stress induced by a nonverbal IQ test (which did not increase blood pressure) elicited clear-cut renal vasoconstriction among hypertensives but not among normotensives without a family history of hypertension; those with a family history of hypertension had an intermediate response. Renin release was similarly differentiated among individuals in the Hollenberg study.

Tidgren and coworkers performed a series of investigations of renal vascular responses to different stimuli, based on a renal venous thermodilution technique that allowed rapid measurements of renal blood flow. Table 1 lists some of these results. It may be seen that mental stress (CWT) increased renal vascular resistance by 50% and increased the renal overflows of renin and norepinephrine markedly. Isometric handgrip exercise was less efficient, especially regarding renin release, despite a similar increase in mean arterial pressure (approximately 20 mmHg with both stimuli). The renal vasoconstrictor response to mental stress was similar to that elicited by dynamic exercise of moderate intensity (60% of maximum effort). Heart rate and arterial plasma epinephrine increased more with CWT than with isometric handgrip, whereas dynamic exercise increased heart rate still more, with a smaller increase in epinephrine at a similar level of renal vasoconstriction. Thus, the patterns of responses differ among these three stimuli. The renal vasoconstrictor

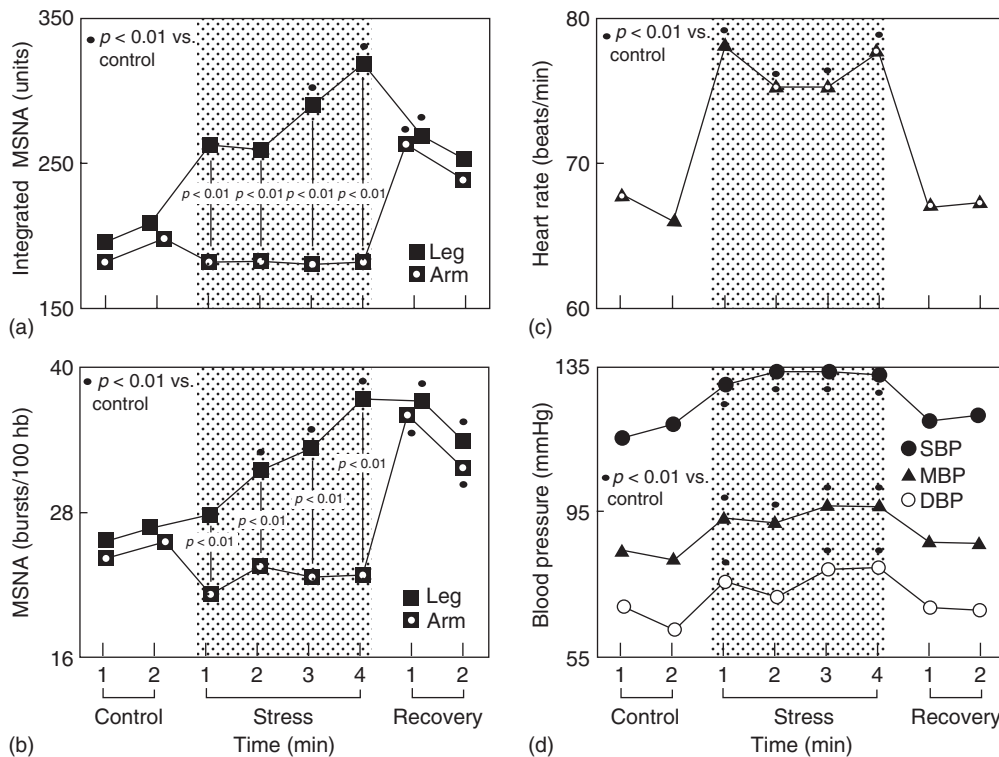


Figure 7 Simultaneously recorded activity in the arm and leg in healthy individuals during mental arithmetic. a, Integrated muscle sympathetic nerve activity (MSNA); b, MSNA (bursts/100 heart beats); c, heart rate; d, blood pressure. MSNA increased in the leg but not in the arm in response to mental stress, indicating that skeletal muscle blood flow is under different neurogenic modulation in different regions of the body during stress.

DBP, diastolic blood pressure; MBP, mean blood pressure; SBP, systolic blood pressure;

From Anderson, Wallin and Mark (1987), *Hypertension* 9(supplement. 3), 114–119, with permission from Lippincott, Williams & Wilkins.

and renin responses to mental stress were correlated to the renal norepinephrine overflow response (i.e., sympathetic nerve activity). The possible importance of epinephrine is discussed next.

Animal studies by di Bona, Koepke, Thorén and others have shown that renal vasoconstriction, as well as decreases in renal sodium excretion and increases in renin release, elicited by mental stress are sympathetically mediated and that these responses are enhanced in spontaneously hypertensive animals. Thus, the sympathetic nerves are important for renal responses to stress. Work in humans is less extensive, but the results discussed here tend to support the conclusion that the kidney is a target organ for stress and that sympathetic nerve activity is involved.

Is Epinephrine an Important Stress Hormone?

When given intravenously, epinephrine increases cardiac output mainly by increasing stroke volume. This is perceived as tachycardia (or the heart is pumping harder), even though the heart rate is relatively little influenced unless the dosage of epinephrine is high. The increase in stroke volume is related to an increase in cardiac contractility as well as a reduction

of afterload due to the systemic vasodilatation elicited by epinephrine. During stress, however, the increase in cardiac output is mainly related to an increase in heart rate, due to increased cardiac sympathetic nerve activity and vagal withdrawal. Even though the human heart is endowed with epinephrine sensitive β_2 -adrenoceptors, the cardiac response to stress is thus mainly neurogenic.

Epinephrine is a potent vasodilator when infused intravenously or intra-arterially. The vasodilator response is elicited by β_2 -adrenoceptor stimulation, and the α agonistic effects of epinephrine become apparent only after blockade of vascular β_2 -adrenoceptors or at supraphysiological concentrations of epinephrine. In the absence of β -blockade, intravenously infused epinephrine does not change mean arterial pressure within the physiological range, but concentration-dependent decreases in diastolic blood pressure (reflecting systemic vasodilatation) and increases in systolic blood pressure (reflecting increased stroke volume) are elicited. The importance of epinephrine for vascular responses to stress may be deduced from comparisons of the plasma concentrations achieved during stress and those required for physiological

Table 1 A comparison of responses to mental stress, isometric handgrip exercise, and dynamic exercise in healthy male volunteers^a

Condition	Renal vascular resistance (PRU)	Renal renin release (units min ⁻¹)	Arterial (nmol l ⁻¹)	Norepinephrine Renal venous (nmol l ⁻¹)	Renal overflow (pmol min ⁻¹)	Epinephrine arterial (nmol l ⁻¹)	Heart rate (beats min ⁻¹)	Source ^b
Rest	0.15 ± 0.01	75 ^c	0.78 ± 0.06	1.09 ± 0.10	252 ± 29	0.26 ± 0.04	59 ± 2	1
Mental stress	0.22 ± 0.02***	247***	1.34 ± 0.11***	2.93 ± 0.54***	708 ± 79***	0.68 ± 0.11***	81 ± 5***	
Rest	0.16 ± 0.02	54 ± 17	0.98 ± 0.13	1.24 ± 0.14	228 ± 34	0.24 ± 0.03	56 ± 2	2
Isometric handgrip	0.20 ± 0.03*	≈60 (n.s.) ^c	1.49 ± 0.16***	2.36 ± 0.29***	508 ± 73***	0.38 ± 0.04***	68 ± 3***	
Rest	0.14 ± 0.02	69 ± 28	0.87 ± 0.08	1.05 ± 0.11	201 ± 33	0.32 ± 0.07	67 ± 3	3
Exercise (60%) ^d	≈0.22**	≈200**	≈3**	≈5**	≈1000**	≈0.5**	≈140***	
(maximum effort)	0.33 ± 0.03***	825 ± 302***	11.4 ± 2.45***	19.8 ± 5.69***	3039 ± 806***	2.12 ± 0.54***	175 ± 6***	

^aMental stress was induced by Stroop's color word conflict test (CWT; Frankenhaeser Version), isometric handgrip exercise used maximum effort, and dynamic exercise used graded supine exercise (levels 2 and 3). Renal blood flow was measured by a renal venous thermodilution technique and renal vascular resistance was calculated on the basis of intra-arterially determined blood pressures. For transformation of catecholamine values, 1 nmol l⁻¹ = 169 pg ml⁻¹ (norepinephrine) or 183 pg ml⁻¹ (epinephrine). Mean values ± SEM (or median) for resting and peak values are shown (*n* = 8–12). * indicates *p* < 0.05; ** indicates *p* < 0.01; *** indicates *p* < 0.001; n.s., not significant; PRU, peripheral resistance units.

^bData from:

1. Tidgren and Hjemdahl (1989), *American Journal of Physiology* **257**, F682–F689.
2. Tidgren and Hjemdahl (1988), *Acta Physiologica Scandinavica* **134**, 23–34.
3. Tidgren et al. (1991), *Journal of Applied Physiology* **70**, 2279–2286.

^cMedian values reported, due to skewed distribution of renin release.

^dEstimated from graphs.

responses, as well as by effects of β -blockade. **Figure 6** illustrates intraindividual comparisons of responses to CWT and infused epinephrine with regard to the forearm circulation; similar comparisons were made in our kidney studies.

Intravenously infused epinephrine may elicit reflexogenic increases in muscle sympathetic nerve activity. This is seen in **Figure 6**, which illustrates that epinephrine is more efficient when infused locally (intra-arterially) than intravenously. When vasodilator responses to CWT and epinephrine were compared, we found that epinephrine contributed to only a small extent (10–30%, on average). Data on local infusions of propranolol support this contention. Thus, epinephrine does not explain the vasodilation in human skeletal muscle during mental stress.

In the kidney study of mental stress (**Table 1**), there was no vasoconstrictor (or vasodilator) response to intravenously infused epinephrine, even at plasma concentrations as high as 6 nmol l^{-1} . Renal norepinephrine overflow was little influenced, but there was a concentration-dependent increase in renin release from the kidney. Epinephrine might thus contribute slightly ($\approx 20\%$) to renin release but not to renal vasoconstriction during mental stress.

In summary, the cardiovascular responses to mental stress in humans cannot be explained by epinephrine, but this hormonal link of the sympathoadrenal system may contribute to certain responses. The conception that epinephrine is the stress hormone is not correct from a functional point of view. However, the epinephrine response to stress does reflect the stress perceived and the arousal of the individual, and it is a good marker for the stimulus intensity of a stressor.

Pathophysiological Aspects

Stress is considered to be involved in the pathogenesis of several cardiovascular disease states, and stress may provoke acute cardiovascular complications, such as myocardial infarction and sudden death in individuals with symptomatic or asymptomatic atherosclerosis. However, for reasons already discussed, it is difficult to define the exact role of stress in cardiovascular disease.

Animal studies have shown that repeated mild activation of the defense reaction (by stress or by electrical stimulation of certain brain regions) may contribute to the development of hypertension. In addition, animals and humans with a genetic predisposition for hypertension often show enhanced cardiovascular and neurogenic reactivity to stress. Animal studies have also shown that stress may accelerate the development of atherosclerosis, as illustrated by

elegant experiments by Clarkson, Kaplan, and Manuck in cynomolgus monkeys showing increased atherosclerosis and by Henry in mice. In humans, it has been more difficult to link stress to atherosclerosis, but there are indications that stress, anxiety, and personality factors are involved. Furthermore, biobehavioral modifications may alter the course of coronary artery disease. Metabolic disturbances, such as diabetes and the metabolic syndrome (a cluster of risk factors with disturbed glucose regulation and often hypertension), are increasingly being linked to atherosclerotic disease. It is of interest that stress may accelerate these processes via the activation of the hypothalamic-pituitary-adrenal axis and release of cortisol. Thus, stress may indeed be involved in the pathogenesis of atherosclerosis, in agreement with popular views on the subject.

Various everyday stresses may trigger serious cardiovascular events, such as myocardial infarction, stroke, and sudden death (due to arrhythmia or myocardial infarction). Work by Mittleman, Muller, Willich, and others has shown diurnal variation in the incidence of myocardial infarction and that physical or mental stress precedes myocardial infarction or sudden death more frequently than would be expected by chance.

Acute myocardial infarction involves thrombosis in the coronary arteries, which may be triggered by many factors, including stress. The hemodynamic changes associated with mental or physical stress increase the risk of atherosclerotic plaque rupture, which is a potent stimulus for local thrombus formation. The mechanisms involved in thrombosis are complex and involve platelets and the coagulation system, as well as factors that oppose coagulation and dissolve blood clots (fibrinolysis). Stress-induced platelet activation may lead to the release of vasoactive substances, which promote ischemia and thrombus formation. Platelet activation may be enhanced by endothelial dysfunction (reduced release of NO and prostacyclin), increased shear stress in an atherosclerotic vessel, and other platelet-activating effects of stress that operate also among healthy individuals. Endothelial dysfunction also promotes vasoconstrictor responses to stress in the coronary arteries and may contribute to imbalances between blood flow and metabolic demands (i.e., myocardial ischemia). Mental stress has indeed been shown to elicit abnormal coronary vasoconstriction and myocardial ischemia among patients with coronary artery disease. The ischemic myocardium is vulnerable to life-threatening arrhythmias, especially when sympathetic activity is increased and vagal activity is reduced, as is the case during stress. Thus, stress elicits cardiovascular and neurohormonal changes that are beneficial to the

healthy subject but detrimental to patients with coronary artery disease. In addition, stress produces prothrombotic alterations that may further increase the risk of myocardial infarction.

Further Reading

- Anderson, E. A., Wallin, B. G. and Mark, A. L. (1987). *Hypertension* 9 (supplement 3), 114–119.
- Bennet, T. and Gardiner, S. M. (eds.) (1996). *Nervous control of blood vessels*. Amsterdam: Harwood Academic.
- Brod, J. (1963). Haemodynamic basis of acute pressor reactions and hypertension. *British Heart Journal* 25, 227–245.
- Byrne, D. G. and Rosenman, R. H. (eds.) (1990). *Anxiety and the heart*. New York: Hemisphere.
- Folkow, B. (1982). Physiological aspects of primary hypertension. *Physiological Reviews* 62, 347–504.
- Fuster, V., Badimon, L., Badimon, J. J., et al. (1992). The pathogenesis of coronary artery disease and the acute coronary syndromes. *New England Journal of Medicine* 326, 242–250, 310–318.
- Goldstein, D. S. (1995). Clinical assessment of sympathetic responses to stress. *Annals of the New York Academy of Sciences* 771, 570–593.
- Goldstein, D. S. (1995). Stress as a scientific idea – a homeostatic theory of stress and distress. *Homeostasis* 36, 177–215.
- Hemingway, H. and Marmot, M. (1998). Psychosocial factors in the primary and secondary prevention of coronary heart disease: a systematic review. In: Yusuf, S., Cairns, J. A., Camm, A. J., Fallen, E. L. & Gersh, B. J. (eds.) *Evidence based cardiology*, pp. 269–285. London: BMJ Books.
- Freyschuss, U., Hjemdahl, P., Juhlin-Dannefelt, A. and Linde, B. (1988). Cardiovascular and sympatho-adrenal responses to mental stress – influence of beta-blockade. *American Journal of Physiology* 255, H1443–H1451.
- Herd, J. A. (1991). Cardiovascular responses to stress. *Physiological Reviews* 71, 305–330.
- Hjemdahl, P. (1993). Plasma catecholamines – analytical challenges and physiological limitations. *Baillière's Clinical Endocrinology and Metabolism* 7, 307–353.
- Hjemdahl, P. (2002). Stress and the metabolic syndrome – an interesting but enigmatic association. *Circulation* 106, 2534–2536.
- Hjemdahl, P., Larsson, P. T. and Wallén, H. (1991). Effects of stress and β -blockade on platelet function. *Circulation* 84, VI-44–VI-61.
- Julius, S. (1988). The blood pressure seeking properties of the central nervous system. *Journal of Hypertension* 6, 177–185.
- Linde, B., Freyschuss, U., Hjemdahl, P. and Juhlin-Dannefelt, A. (1989). Adipose tissue and skeletal muscle responses to mental stress. *American Journal of Physiology* 256, E12–E18.
- Lindqvist, M., Melcher, A. and Hjemdahl, P. (2004). Hemodynamic and sympatho-adrenal responses to mental stress during nitric oxide synthesis blockade. *American Journal of Physiology: Heart Circulatory Physiology* 287, H2309–H2315.
- Lindqvist, M., Kahan, T., Melcher, A., Bie, P. and Hjemdahl, P. (1996). Forearm vasodilator mechanisms during mental stress – possible roles of epinephrine and ANP. *American Journal of Physiology* 270, E393–E399.
- McCarty, R., Horwatt, K. and Konarska, M. (1988). Chronic stress and sympathetic-adrenal medullary responsiveness. *Social Science & Medicine* 26, 333–341.
- Rosmond, R. (2005). Role of stress in the pathogenesis of the metabolic syndrome. *Psychoneuroendocrinology* 30, 1–10.
- Shepherd, J. T. and Weiss, S. M. (eds.) (1987). *Special issue on behavioral medicine and cardiovascular disease*. *Circulation*, 76 (Part 2).
- Septoe, A. and Vögele, C. (1991). Methodology of mental stress testing in cardiovascular research. *Circulation* 83, II-14–II-24.
- Strike, P. C. and Septoe, A. (2003). Systematic review of mental stress-induced myocardial ischemia. *European Heart Journal* 24, 690–703.
- Tidgren, B. and Hjemdahl, P. (1988). Reflex activation of renal nerves in humans: effects on norepinephrine, dopamine and renin overflow to renal venous plasma. *Acta Physiologica Scandinavica* 134, 23–34.
- Tidgren, B. and Hjemdahl, P. (1989). Renal responses to mental stress and epinephrine in man. *American Journal of Physiology* 257, F682–F689.
- Tidgren, B., Hjemdahl, P., Theodorsson, E. and Nussberger, J. (1991). Renal neurohormonal and vascular responses to dynamic exercise in man. *Journal of Applied Physiology* 70, 2279–2286.
- Vane, J. R., Änggård, E. E. and Botting, R. M. (1990). Regulatory functions of the vascular endothelium. *New England Journal of Medicine* 323, 27–36.
- Verrier, R. L. and Mittleman, M. A. (1996). Life-threatening cardiovascular consequences of anger in patients with coronary heart disease. *Cardiology Clinics* 14, 289–307.
- Willerson, J. T., Golino, P., Eidt, J., et al. (1989). Specific platelet mediators and unstable coronary artery lesions: experimental evidence and potential clinical implications. *Circulation* 80, 198–205.
- Willich, S. N. and Muller, J. E. (eds.) (1996). *Triggering of acute coronary syndromes – implications for prevention*. Dordrecht: Kluwer Academic.

Cardiovascular Disease, Stress and

G P Chrousos and G Kaltsas

University of Athens, Athens, Greece

© 2007 Elsevier Inc. All rights reserved.

Prevalence of Atherosclerosis – Cardiovascular Disease
Established and Evolving Risk Factors of Atherosclerosis
Alterations of the Function of the Stress System Leading
to the Development of Atherosclerosis and
Cardiovascular Disease

Glossary

Catecholamines Norepinephrine (norepinephrine), epinephrine (adrenaline), and dopamine are the mediators through which the autonomic nervous system responds to stressors and controls a wide variety of system functions, such as arousal, arterial blood pressure, glucose levels, etc.

Circadian rhythm Many functions of living creatures are subject to periodic or cyclic changes mainly influenced by the nervous system. These changes are intrinsic to the organism independent of the environment and are driven by a biologic “clock”. When they have a period of approximately 24 hours these rhythms are called circadian (around a day).

Cytokines Substances secreted from a number of different immune, including activated macrophages and lymphocytes, and nonimmune cells. According to their subtype they may exert a pro- or anti-inflammatory effect and activate the stress response.

Cushing’s syndrome The presence of symptoms and signs associated with prolonged exposure to inappropriately elevated levels of free plasma glucocorticoids.

Glucocorticoids The final endocrine effectors of the hypothalamic-pituitary-adrenal (HPA) axis. They are steroid pleiotropic hormones that exert their effects through ubiquitously distributed intracellular receptors. Glucocorticoids are crucial for the maintenance of resting and stress-related homeostasis, regulating and assuring the integrity of cardiovascular, metabolic, behavioral and immune homeostasis.

Metabolic syndrome A clustering of factors closely associated with the development of atherosclerotic cardiovascular disease. The most recent definition includes central obesity,

hypertriglyceridemia, low high-density lipoprotein (HDL) cholesterol levels, hypertension, and increased fasting glucose levels or diabetes mellitus type 2.

State of threatened or perceived threatened homeostasis. Stress can be of either physical, social and/or psychological origin. During stress, a coordinated adaptive response of the organism is activated that is crucial in achieving and maintaining homeostasis. If this response is excessive and prolonged or defective and inadequate, the organism is in dyshomeostasis, or allostasis (allo = different) or cacostasis (caco = bad). The main components of the adaptive response are the corticotropin-releasing hormone (CRH) and the locus-caeruleus-norepinephrine (LC-NE)/arousal and autonomic nervous systems and their peripheral effectors, HPA axis and the systemic and adrenomedullary limbs of the autonomic system.

Sympathetic nervous system The SNS along with the parasympathetic system consist the autonomic nervous system. The SNS is under direct control from the CNS, allowing rapid onset of actions of short duration either excitatory or inhibitory) as a result of the abbreviated half-lives of catecholamines.

Prevalence of Atherosclerosis – Cardiovascular Disease

Atherosclerotic lesions (atheroma) consist of asymmetric focal thickening of the innermost layer of the artery, the intima. Pathologic changes caused by atherosclerosis in the cardiovascular system, such as coronary heart disease, aneurisms of the aorta and brain, or peripheral limb ischemia, currently represent the most profound threat to the quality and expectancy of life in developed countries. It has been estimated that cardiovascular diseases (CVD) cause 38% of all deaths in North America and are the most common cause of death in European men under 65 years of age and the second most common cause in women. Ischemia of the myocardial and cerebral circulation – leading to myocardial infarction and stroke, respectively, develops when the atheromatous process prevents blood flow through the involved arteries either as a result of plaque rupture or endothelial erosion. Common risk factors for developing atherosclerosis are hyperlipidemia, hypertension, and

disorders of carbohydrate homeostasis, such as glucose intolerance and diabetes mellitus type 2 (DM2). The spectrum of these traditional risk factors has lately been incorporated under the entity of the metabolic syndrome, which also includes visceral obesity, insulin resistance, and/or blood hypercoagulation. Patients with the metabolic syndrome are at increased risk for developing atherosclerosis and, therefore, constitute a group for early intervention. Over the past two decades, there has been a striking increase in the number of people with the metabolic syndrome; this increase is associated with the global epidemic of obesity and DM2. There is an urgent need for strategies to prevent the emerging global epidemic of metabolic syndrome along with the elevated risk not only of DM2 but also of CVD.

Established and Evolving Risk Factors of Atherosclerosis

When traditional risk factors are taken into consideration, not all cases of CVD can be attributed to their presence. Studies, such as the Framingham Heart Study, have shown that approximately 50% of coronary heart disease cases can be explained by the three major cardiovascular risk factors – elevated serum cholesterol, hypertension, and smoking. In addition, although rigorous treatment of hypercholesterolemia and hypertension was expected to substantially decrease CVD, these diseases still remain the main cause of death globally and are expected to increase further owing to a rapidly increasing prevalence of obesity in developing countries and Eastern Europe and the continuing rise in the incidence of obesity and diabetes in the western world. To explain these observations, additional pathogenetic mechanisms have been thought to contribute, among which is the effect that chronic stress may exert in the development of the metabolic syndrome and the concurrent presence of systemic smoldering inflammation.

Common effector pathways appear to mediate the pathogenesis of metabolic syndrome, atherosclerosis and CVD. These include activation of the hypothalamic-pituitary-adrenal (HPA) axis, the systemic/adrenomedullary sympathetic nervous systems (LC-NE-SNS) and the inflammatory/immune response. These systems, either alone or in combination, can initiate and/or promote mechanisms related to the development of atherosclerosis, such as mobilization of low-density lipoprotein (LDL), which is manufactured and secreted into the circulation by the liver, as a carrier of cholesterol to peripheral tissues. Macrophages, a major component of the immune system, contribute to the vascular accumulation of cholesterol

by absorbing the oxidized forms of LDL and adhering to inflamed endothelium to ultimately form stenosis-causing plaques. Macrophages and the cytokines they produce can also be activated by the stress system, whereas cytokines are also a major link between obesity and endothelial dysfunction. Furthermore, alterations of cortisol secretion, effect, and dynamics have been thought to be related either directly and/or indirectly, through inhibition of systems that exert a protective effect, in the development of atherosclerotic lesions.

Alterations of the Function of the Stress System Leading to the Development of Atherosclerosis and Cardiovascular Disease

The Stress System

During stress, adaptive behavioral and peripheral responses are activated that optimize the ability of the organism to adjust homeostasis. Stress is associated with increased activity of the HPA axis and the LC-NE-SNS, which result in systemic elevations of their endocrine end-effectors, the glucocorticoids (cortisol) and the catecholamines (CAs). These compounds, together with other products of the stress system, such as for instance the cytokine interleukin (IL)-6, influence a variety of adaptive responses. The activation of the HPA axis is regulated by glucocorticoid negative-feedback control through glucocorticoid receptors (GR) in the central nervous system (CNS). Normally, there is a diurnal secretory pattern of cortisol with elevations in the early morning and decreases in the afternoon and evening hours. This basal activity is modified by stressful stimuli, which can be either physical (extreme environmental alteration, trauma, and toxins) and/or events that are perceived as unpleasant, demanding, scary, or threatening. The latter can include recalls of previously experienced or unpleasant or pleasant events which either positively or negatively influence the regulation of the HPA axis. Thus, stress leads to increased integrated cortisol and CA secretion, which can be of short or long duration depending on the severity and duration of the stressor. Following repeated exposure to stressors there is great interindividual variation in the extent of cortisol response habituation. Although studies have been conflicting with some showing activation and others attenuation of the HPA axis function in chronic stress, in both situations abnormalities of cortisol secretory dynamics are evident defining a chronic hyperactivation or hypoactivation of the stress system.

Overt or Subtle Hypercortisolism and Cardiovascular Disease

Excessive and sustained cortisol secretion or chronic administration of pharmacological doses of glucocorticoids (endogenous or exogenous Cushing's syndrome (CS), respectively) have been associated with the entire spectrum of the metabolic syndrome, including visceral obesity, insulin resistance/carbohydrate intolerance/DM2, dyslipidemia, dyscoagulation, and hypertension, along with their morbid sequelae of atherosclerosis and CVD. The causal association between chronic psychosocial stress, hypercortisolism, and metabolic syndrome-related manifestations has been shown experimentally only in cynomolgus monkeys. Major melancholic depression, a state of chronic and sustained hyperactivation of the stress system is also associated with hypercortisolism and such patients have a markedly decreased life expectancy due to increased mortality from CVD (2–3 risk of gender- and age-matched controls). Thus, sustained hypercortisolism, as encountered in CS and melancholic depression, is associated with increased prevalence of atherosclerosis, hypertension, and all the clinical and biochemical features of the metabolic syndrome and an overall increased in mortality from CVD.

The conspicuous similarities between CS and the metabolic syndrome and the development of atherosclerosis in both conditions suggest that hypercortisolism may be involved in the pathogenesis of the latter. There is some evidence that hypercortisolism may be directly related to the pathogenesis of the metabolic syndrome, as in such patients, alterations of the secretory pattern and turnover of glucocorticoids has been noted. Such patients may also have depressive (a severe form of chronic psychological stress) symptomatology and be at increased risk for developing a severe form of metabolic syndrome, depending on genetic, developmental and/or environmental factors. Indeed, some patients with depressive symptomatology and/or chronic active alcoholism associated with visceral obesity and insulin resistance and accompanied by increased cortisol secretion may occasionally be difficult to differentiate from those with frank endogenous CS. The term pseudo-Cushing syndrome has been employed to describe such patients. As depression is highly prevalent, affecting up to 10–15% of the population, a considerable number of patients with pseudo-Cushing are among us.

Increases in body mass index (BMI), blood pressure, and dyslipidemia in nonCushingoid middle-aged men have been correlated with the degree of stress perception and cortisol secretion. It is probable that similar parallel alterations of the other peripheral component of the stress system, the SNS, are also present.

CAs are important modulators of corticotropin-releasing hormone (CRH) and adrenocorticotrophic hormone (ACTH) secretion, particularly during acute and chronic stress and their effects are mediated by both the $\alpha 1$ and $\alpha 2$ subtypes of adrenoceptors. Activation of the $\alpha 2$ subtype may represent a negative-feedback mechanism that prevents excessive glucocorticoid responses to internal and external stressors. Indeed, following $\alpha 2$ inhibition, the ACTH response after combined stimulation with CRH and antidiuretic hormone (AVP) was significantly higher in obese (visceral obesity) compared to normal weight individuals. This finding suggests that obese individuals with abdominal fat distribution may escape the physiologic $\alpha 2$ -adrenoceptor control of their stress system, thus favoring inappropriate HPA axis and SNS excitation (Table 1).

Analyzing further the dynamics of cortisol secretion, it is evident that a nonstressed HPA axis is characterized by an increased variance of circadian cortisol secretion and an appropriate suppression of the morning cortisol levels in response to low-dose dexamethasone administration. In contrast, a chronically stressed HPA axis is characterized by a decreased diurnal variance, mostly due to evening elevation, mild morning decreases of cortisol levels, and inadequate suppression of morning cortisol by overnight dexamethasone. These findings suggest the presence of chronic hypersecretion of CRH in chronically stressed individuals and a reset of their HPA axis leading to an increase of total time-integrated cortisol secretion. In the presence of a properly functioning glucocorticoid negative-feedback system, these alterations of cortisol secretion are minimized resulting in amelioration of the manifestations of chronic hypercortisolism (Tables 1 and 2).

The ability of the glucocorticoid negative-feedback system to limit the production of cortisol during stress can be impaired by early life stress and by exposure of the organism to chronic physical/emotional stress. The sensitivity of target tissues to glucocorticoids, of course, also plays a major role in both influencing the negative-feedback system and the actions of glucocorticoids on their other target tissues. Studies examining the presence of polymorphisms of the glucocorticoid receptor gene described the presence of a correlation between particular polymorphisms and the development of hypertension and/or visceral obesity, whereas others demonstrated that increased dermal tissue glucocorticoid sensitivity was associated with an increased blood pressure, insulin resistance, and hyperglycemia. Interestingly, the same subjects also exhibited an enhanced secretion of cortisol and impaired conversion of cortisol to inactive

Table 1 Somatic consequences of chronic stress system activation/target tissue effects with respect to the development of CVD

HPA axis – cortisol	Locus caeruleus/ norepinephrine CAs, IL-6
<i>CNS effects</i>	
Potential of amygdala/fear – mesocorticolimbic dopaminergic system dysfunction (↑)	↑CAs
Leptin actions ↓	
GH/IGF-1 ↓	
LH/Testosterone – estradiol ↓	
TSH/thyroxine ↓	↓IL-6
<i>Hypertension</i>	
Vasoconstriction (↑) (CAs, angiotensin II, arginine-vasopressin/endothelin)	↑CAs
Vasodilatation (↓) (kallikrein, prostacyclin, NO-synthase, inflammatory cytokines)	
Salt retention ↑, Renin substrate↑	↑Renin, CAs
Visceral fat syndrome (↑)	
Insulin resistance (↑)	
Glucoseogenesis (↑), peripheral glucose disposal (↓), insulin concentration (↑), carbohydrate intolerance (↑), LDL-cholesterol and small dense LDL, FFA, TG (↑), coagulation process (↑),	↑CAs, ↑IL-6

↑, stimulation; ↓, inhibition; CA, catecholamine; CNS, central nervous system, FFA, free fatty acid; HPA axis, hypothalamo-pituitary-adrenal axis; GH, growth hormone; IGF, insulin growth factor; IL, interleukin; LDL, low-density lipoprotein; LH, luteinizing hormone; NO, nitric oxide; TG, triglyceride; TSH, thyroid-stimulating hormone. Adapted from *International Journal of Obesity* 24, The role of stress and the hypothalamic pituitary adrenal axis in the pathogenesis of the metabolic syndrome: neuroendocrine and target tissue related causes. 2000, S50–S55, with permission.

Table 2 Abnormalities of the hypothalamic-pituitary-adrenal (HPA) axis associated with the metabolic syndrome and cardiovascular disease

Function	Type of alteration of the HPA axis
Cortisol metabolic clearance rate	↑
Cortisol production in peripheral tissues (adipose) – ↑ 11β-HSD1 activity	↑
Daily urinary free cortisol excretion	↑
Number of secretory pulses of ACTH	↑
Amplitude of secretory pulses of ACTH	↓
ACTH and cortisol response to ACTH	↑
ACTH and cortisol response to CRH and AVP	↑
ACTH and cortisol response to acute stress	↑
Cortisol response to meals	↑
HPA sensitivity to increased noradrenergic tone	↑
Cortisol suppression by dexamethasone	↓

↑, increased; ↓, decreased; ACTH, adrenocorticotrophic hormone; AVP, antidiuretic hormone; CRH, corticotropin-releasing hormone; HPA, hypothalamic-pituitary adrenal; 11β-HSD1, 11β-hydroxysteroid dehydrogenase type 1. Adapted from *International Journal of Obesity* 24, Activity of the hypothalamo-pituitary-adrenal axis in different obesity phenotypes. 2000, S47–S49, with permission.

metabolites implying a direct or an indirect effect of stress on tissue active hormone availability.

Recent studies have shown alterations of 11β-hydroxysteroid dehydrogenase (11β-HSD) type 2, which converts cortisol to inactive cortisone, particularly in the kidney. However, the metabolic clearance rate of cortisol is mostly influenced by its regeneration from inactive cortisone in the liver, fat, and skeletal muscle by 11β-oxoreductase type 1. Consistent with animal studies, 11β-HSD2 appears to be a constitutive rather than a hormonally regulated enzyme, in contrast to 11β-HSD1 which is downregulated by insulin and estrogens and appears to be dysregulated

in obesity. The activity of 11β-HSD1 in the omental fat of obese people is significantly higher than in lean individuals and the activity of the enzyme is further increased by exposure to cortisol and insulin. The inappropriate constant exposure of fat, particularly visceral fat, to cortisol may be important in determining its differentiation and mass increase leading to visceral obesity and the manifestations of the metabolic syndrome. In addition, it may modify the secretion of other adipose tissue hormones, such as leptin, IL-6, and/or adiponectin, and thus modify the output of adipose cells and influence the information they convey to satiety centers regarding fuel homeostasis.

High circulating cortisol levels or sustained alteration of the normal diurnal pattern of cortisol secretion also suppress the growth, thyroid and reproductive axes and lead to elevation of Ca^{2+} -stimulated cytokine production (and particularly IL-6) contributing further to the exacerbation of the metabolic syndrome by inducing lipolysis and utilization of protein for glyconeogenesis. Activation of the stress system is associated with suppression of growth hormone secretion and resistance of somatomedin C and other growth factor effects on their target organs. The reproductive axis is also inhibited at several levels by various components of the HPA axis. Either directly or via β -endorphin, CRH suppresses luteinizing hormone releasing hormone (LHRH) neurons, and thus gonadotropin secretion from the pituitary gland, whereas glucocorticoids exert inhibitory effects at the levels of LHRH neurons, pituitary gonadotrophs and the gonads and render target tissues of sex steroids resistant to these hormones. These effects and similar ones observed in the thyroid axis, counteract the beneficial effects these hormones may exert in body distribution and fuel homeostasis contributing further to the development of visceral obesity and the metabolic syndrome by antagonizing fat-tissue catabolism and inhibiting muscle and bone anabolism.

Although alterations of the LC-NE-SNS have not been studied in such detail as those of the HPA axis, it is quite likely that SNS hyperactivity, either as a result of a generalized activation of the stress system or as a compensatory mechanism in cases of a faltering HPA axis activity, also participates in the pathogenesis of cardiovascular disorders. A consequence of elevated SNS activity may be through elevated free fatty acid levels which can induce insulin resistance in muscle and amplify the metabolic perturbations of patients with pathological regulation of their HPA axis (Table 1).

Systemic Inflammation and Cardiovascular Disease

Over the last decade, there is mounting evidence that inflammation plays a major role in the development of CVD. Several studies have documented that elevated concentrations of IL-6 or the downstream acute phase reactant C-reactive protein (CRP) predict the development of CVD over many years, suggesting that inflammation makes a contribution to the earlier stages of the disease. Elevated IL-6 levels can predict total and cardiovascular mortality over a 5-year follow-up, the association being independent of prevalent vascular disease, smoking, and traditional risk factors. The predictive value of IL-6 is in fact stronger than that of CRP. This cytokine is not produced only as a result of inflammation but as a component of

noninflammatory stress, while a large proportion of total circulating levels of IL-6 may originate from the adipose tissue. IL-6 participates in the development of CVD through different mechanisms, including induction of endothelial dysfunction, hepatic fibrinogen release, and adhesion molecule expression by endothelial cells and circulating leukocytes. IL-6 also decreases insulin sensitivity of peripheral tissues and increases circulating nonesterified fatty acids, inducing a metabolic syndrome phenotype. Psychological stress can elevate IL-6 levels through Ca^{2+} hypersecretion, possibly from liver Kupfer cells or from the adipose tissue itself. IL-6 is also a potent stimulator of the HPA axis leading to cortisol hypersecretion and the development of the metabolic syndrome through this path as well. It is possible that genetic variations of the IL-6 and other related genes involved in its signaling system modulate the body's response to stress and obesity.

Obesity as a Chronic Inflammatory State

Obesity is the most common and costly nutritional problem in the United States, affecting approximately 33% of the adult population while its incidence among children and adolescents is also rising. Obesity, particularly the visceral type, is the most important potentially preventable factor that substantially increases the risk for developing DM2, CVD, and certain types of cancers. Adipose tissue secretes large amounts of tumor necrosis factor (TNF)- α and IL-6 in a neurally, hormonally, and metabolically regulated fashion. The plasma levels of these cytokines are proportional to the BMI and are further elevated in patients with visceral obesity. The secretion of inflammatory cytokines has a circadian pattern, with elevations in the evening and in the early morning hours. This pattern is maintained in obese subjects, albeit at a higher level, is affected by the quality and quantity of sleep, and correlates with manifestations of the sickness syndrome. Thus, in obesity, hypercytokinemia is frequently associated with sleepiness and fatigue, as well with an activated acute-phase reaction. Obesity, especially the visceral type, can be considered as a chronic inflammatory state, with many of the behavioral, immune, metabolic and cardiovascular manifestations and sequelae of such a state.

Psychosocial Stress and Cardiovascular Disease

There has been considerable information relating alterations in the activity of the stress system and conventional risk factors in developing CVD, the major cause of morbidity and mortality in western societies. A consistent and continuous gradient between socioeconomic status (SES) and the prevalence

of CVD has been found, with people from lower SES having more disease. In some studies the ratio of death rates from the lowest to the highest professional grade may be as high as 3 to 1. Although an inverse relation between SES and cigarette smoking has been noted, this is not obvious when total cholesterol and/or blood pressure are considered; however, an inverse relation between SES and BMI and visceral obesity was noted as well. Not surprisingly, physical inactivity was clearly linked to obesity and lower SES and this combination may contribute further to the SES-disease gradient. In favor of these observations is the finding of the Gothenberg Primary Prevention Study, which demonstrated an odds ratio of CVD morbidity of 2.2, which changed to only 1.9 after adjusting for established risk factors (including blood pressure, smoking, cholesterol, BMI, and exercise habit). The inability of traditional risk factors to explain more than 25% of the SES-CVD gradient suggests that additional factors related to the working conditions might contribute.

The ability to monitor surrogate markers of CVD and also to visualize atherosclerotic plaques directly has helped to evaluate the biologic pathways through which SES may exert its effects. An analysis of psychological factors and blood pressure reactivity (measured as the anticipatory increase in blood pressure before an exercise test) exhibited a clear relation to progression of atherosclerosis. Chronic psychosocial stress directly through the stress system and its hormones and/or indirectly through changes in health-related behaviors may be the main culprit connecting poverty and CVD.

See Also the Following Articles

Inflammation; Neuropeptide Y.

Further Reading

- Banks, J., Marmot, M., Oldfield, Z., et al. (2006). Disease and disadvantage in the United States and in England. *JAMA* **295**, 2037–2045.
- Bjontrop, P. and Rosmond, R. (2000). The metabolic syndrome – a neuroendocrine disorder? *British Journal of Nutrition* **83**, S49–S57.
- Charmandari, E., Tsigos, C. and Chrousos, G. (2005). Endocrinology of the stress response. *Annual Reviews of Physiology* **67**, 259–284.
- Chrousos, G. P. (1998). Stressors, stress and neuroendocrine integration of the adaptive response: 1997 Hans Selye Memorial Lecture. *Annals of the New York Academy of Sciences* **851**, 311–335.
- Chrousos, G. P. (2000). The role of stress and the hypothalamic pituitary adrenal axis in the pathogenesis of the metabolic syndrome: neuroendocrine and target tissue related causes. *International Journal of Obesity* **24**, S50–S55.
- Chrousos, G. P. and Gold, P. W. (1992). The concepts of stress and stress system disorders. *JAMA* **267**, 1244–1252.
- Eckel, R. H., Grundy, S. M. and Zimmet, P. Z. (2005). The metabolic syndrome. *Lancet* **365**, 1415–1428.
- Fang, J., Madhavan, S. and Alderman, M. H. (1996). The association between birth-place and mortality from cardiovascular causes among black and white residents of New York City. *New England Journal of Medicine* **335**, 1545–1551.
- Habib, K. E., Gold, P. W. and Chrousos, G. P. (2001). Neuroendocrinology of stress. *Endocrinology Metabolism Clinics North America* **30**, 695–727.
- Hanson, G. K. (2005). Inflammation, atherosclerosis and coronary artery disease. *New England Journal of Medicine* **352**, 1685–1695.
- Lynch, J. W., Kaplan, G. A., Cohen, R. D., et al. (1996). Do cardiovascular risk factors explain the relation between socioeconomic status, risk of all-cause mortality, cardiovascular mortality, and acute myocardial infarction? *American Journal of Epidemiology* **144**, 934–942.
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* **338**, 171–179.
- Pasquali, R. and Vicennati, V. (2000). Activity of the hypothalamo-pituitary-adrenal axis in different obesity phenotypes. *International Journal of Obesity* **24**, S47–S49.
- Rosenbaum, M., Leibel, R. and Hirsch, J. (1997). Obesity. *New England Journal of Medicine* **396**, 396–407.
- Rosmond, R., Lapidus, L. and Bjontrop, P. (1996). The influence of occupational and social factors on obesity and body fat distribution in middle-aged men. *International Journal of Obesity Related Metabolic Disorders* **20**, 599–607.
- Yudkin, J. S., Kumari, M., Humphries, S., et al. (2000). Inflammation, obesity, stress and coronary heart disease: is interleukin-6 the link? *Atherosclerosis* **148**, 209–214.

Caregivers, Stress and

S H Zarit

Pennsylvania State University, University Park, PA, USA

K Bottigi and J E Gaugler

University of Minnesota, Minneapolis, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S H Zarit and J E Gaugler, volume 1, pp 404–407, © 2000, Elsevier Inc.

Caregiving Burden

The Stress Process

Longitudinal Effects of Caregiving

Psychosocial Interventions for Caregivers

Glossary

<i>Burden</i>	The emotional, psychological, physical, and financial challenges assumed by caregivers.
<i>Caregiver</i>	Person who provides informal (i.e., unpaid) assistance provided to a disabled individual, such as an adult child providing help to an elderly parent with Alzheimer's disease or an associated dementia.
<i>Objective burden</i>	The disruptions and degree of changes caregivers experience in various life domains.
<i>Stress process model</i>	A multidimensional framework that examines the proliferation or spread of stress from care-related dimensions to other life domains to global indicators of well-being.
<i>Subjective appraisals/burden</i>	Emotional reactions to stressors, such as fatigue or feelings of being trapped.
<i>Wear and tear hypothesis</i>	Posits that the longer a caregiver remains in her or his role, the more likely it is that negative outcomes will occur.

With increasing longevity and improvements in medical treatment for a range of chronic diseases, more people are surviving to advanced ages. Although most older people are healthy and independent, a significant minority suffer from a variety of physical, cognitive, and psychiatric disabilities that require regular ongoing care and supervision. Despite the visibility of nursing homes and other institutions, families provide the largest part of informal (i.e., unpaid) long-term care assistance to disabled older adults in the United States, and a major emphasis of public care policy is to keep older

adults in community settings and under the care of families or other informal providers in order to reduce the need for nursing home care.

Families have always cared for their elders, but changes in contemporary society have placed increasing strains on their resources to provide assistance. Elders are living longer after the onset of disabilities, and many suffer from severe chronic disabilities that necessitate extensive care. With women increasingly employed outside the home, the role of caregiver falls either on a spouse, who may have his or her own age-related limitations, or on a daughter or daughter-in-law, who must balance the multiple roles of worker, homemaker, caregiver, and, in some cases, mother to her own young children. It is not surprising, then, that much of the emphasis in research on caregiving has focused on the stressors and emotional strains of informal care. Early research focused on correlates of burden, or the emotional, psychological, physical, and financial challenges assumed by caregivers, as well as caregivers' subjective appraisals of how task performance affects their lives. However, the past two decades have seen the development of sophisticated conceptual models to capture the longitudinal dynamics of family care, as well as the development of psychosocial interventions to alleviate negative outcomes in informal long-term care systems.

Caregiving Burden

Early research on family caregivers to older adults suffering from chronic illnesses such as Alzheimer's disease often attempted to correlate contextual variables (e.g., duration of care, relationship of caregiver to care recipient) and care demands (e.g., severity of behavior problems, intensity of cognitive impairment, extent of activities of daily life [ADL] dependencies) with unidimensional constructs representing burden. Burden was often operationalized as the caregiver's subjective experience of care demands. Subsequent research suggested that burden was a multidimensional construct, with objective and subjective components. Objective indices of burden reflect the disruptions and degree of changes caregivers experience in various life domains, while subjective burden refers to caregivers' emotional reactions to care demands. Studies that incorporate this conceptualization of burden have found that different factors are associated with objective and subjective burden, underscoring the theoretical usefulness in approaching the demands of caregiving as multifaceted.

Other research has moved beyond the caregiver's subjective experience to consider the physiological effects of stress. The chronic stress of caregiving has been found to increase a caregiver's risk for health problems. This increased risk of illness is attributable to an overproduction of stress hormones and/or a decline in the use of preventative health measures by the caregiver, such as having annual physical exams, exercising, and having a healthy diet. Research examining caregiving at other points in the life span suggests that the burdens associated with informal care provision to a chronically ill care recipient can have negative implications at the cellular level, including shortened telomere length, reduced telomerase activity, and increased oxidative stress.

The Stress Process

With the greater concern given to the multidimensionality of caregiving stress and the various factors that account for such outcomes, conceptual models were developed to capture the process of informal long-term care. Perhaps the most noteworthy model is the stress process model (SPM), which is largely grounded in sociological perspectives of stress. Specifically, this approach suggests that the occurrence of some environmental demand that is potentially harmful (primary objective stressor) has an immediate impact on the caregiver's life (primary subjective stressor), for example, leading to feelings of overload or to the loss of valued aspects of the relationship with the other person. This immediate impact may proliferate into other areas of life (secondary role strains), for instance, interfering with employment or creating conflict in the family, and thereby increase the likelihood of poor adaptation (e.g., negative mental or physical health outcomes). Conversely, resources such as social support and coping skills can contain the primary stressors, limiting their impact on other roles and relationships and on well-being. Although this model was developed specifically for caregivers of Alzheimer's patients, its grounding in general sociological theory makes it potentially applicable to caregivers of older adults with other kinds of disabilities and impairments.

Longitudinal Effects of Caregiving

Much of the literature on caregiving is cross-sectional, that is, a one-time snapshot of the caregiver's situation. Many of these studies have attempted to explore the wear and tear hypothesis, which posits that the longer a caregiver remains in her or his role, the more likely it is that negative outcomes, such as caregiver

distress or even care recipient institutionalization, will occur. Longitudinal studies, however, make it possible to examine the dynamic interplay of stressors and adaptation. These studies support more complex views of the longitudinal progression of caregiving than that suggested by the wear and tear hypothesis. Several panel studies of dementia caregiving have found an adaptation effect, in which caregivers who remain in the role report decreases in negative emotional and psychological outcomes, such as burden and depression. These studies have supported counter hypotheses that emphasize the ability of families to tolerate once stressful situations or to adapt to and successfully manage chronic caregiving demands and stressors.

Other analyses have examined various transition points that can occur in informal long-term care, such as onset, institutionalization, and bereavement. While few analyses consider onset of informal care responsibilities, recent research suggests that individuals who experienced less abrupt entries into their caregiving roles are more likely to delay nursing home placement as well as indicate decreases in emotional distress and depression over time. Such findings imply that how people assume care responsibilities or manage problems early in their careers has important longitudinal implications, particularly in chronic diseases.

Given the potential public and private costs of nursing home placement, a number of studies have examined the influence of caregiving indicators (e.g., stress, sociodemographic context) on institutionalization. Several indicators have been identified as potential triggers of nursing home placement. Caregiving stressors (such as feelings of overload or feeling trapped by caregiving responsibilities), depression, or impaired subjective health appear likely to expedite placement and, in some instances, appear more likely to trigger institutionalization than functional indicators of the older care recipient. Other promising work in this area has emphasized that institutionalization is not an endpoint in caregiving. Caregivers remain involved in care and various dimensions of emotional distress and psychological well-being are relatively stable after nursing home placement. Similar studies have emphasized the complex progression of stress and depression of informal caregivers prior to and following bereavement.

Psychosocial Interventions for Caregivers

A major emphasis in the field has been the development of programs, services, and other interventions

designed to relieve the stress of caregivers. Increasingly, these interventions have used stress theory to identify modifiable features of the stress process that can lead to improved outcomes. The most successful interventions with caregivers have been multidimensional, addressing several dimensions of the stress process. Another common component of these interventions has been inclusion of the wider family network in treatment, usually through one or two family meetings. Help from families may be particularly useful, since it is flexible and personal. Family level interventions can also address conflict and misunderstandings about the nature of the elder's disease or how care is being provided. Outcomes of these interventions have included reduced subjective burden, lower negative appraisals of the elder's behavior, decreased depression, and delayed institutionalization. Some of these effects have been found to persist or even increase for 3 years or more beyond the intervention period. Interventions designed to reduce depressive behaviors for elders with comorbid depression and dementia have also been found to have corresponding benefits for caregivers' own feelings of depression and subjective burden.

Another way of addressing the stress on family caregivers is by initiating formal, paid help. A variety of services are available to elders and their families, including care management, in-home respite, and adult day services. The impact of these programs on caregivers' stress and well-being has been mixed. It is often the case, however, that caregivers included in these studies have received only low levels of assistance. In the one study that controlled for amount of exposure to the intervention, caregivers who used adult day services on a regular basis for 3 months or more experienced decreased overload, emotional strain, and depressive symptoms compared to a control group. This issue of therapeutic exposure is particularly relevant for caregiving research. Caregivers who are faced with multiple, chronic, and severe stressors are not likely to gain much benefit from brief or unfocused interventions.

See Also the Following Articles

Allotaxis and Allostatic Load; Alzheimer's Disease; Familial Patterns of Stress.

Further Reading

Epel, E. S., Blackburn, E. H., Lin, J., et al. (2004). Accelerated telomere shortening in response to life stress.

- Proceedings of the National Academy of Sciences of the United States of America* 101, 17312–17315.
- Gaugler, J. E., Kane, R. L., Kane, R. A., Clay, T. and Newcomer, R. (2003). Predicting institutionalization of cognitively impaired older people: utilizing dynamic predictors of change. *The Gerontologist* 43, 219–229.
- Gaugler, J. E., Zarit, S. H. and Pearlin, L. I. (2003). The onset of dementia caregiving and its longitudinal implications. *Psychology and Aging* 18, 171–180.
- Mittelman, M. S., Roth, D. L., Coon, D. W. and Haley, W. E. (2004). Sustained benefit of supportive intervention for depressive symptoms in caregivers of patients with Alzheimer's disease. *American Journal of Psychiatry* 161, 850–856.
- Montgomery, R. J. V., Gonyea, J. G. and Hooyman, N. R. (1985). Caregiving and the experience of subjective and objective burden. *Family Relations* 34, 19–26.
- Pearlin, L. I., Mullan, J. T., Semple, S. J. and Skaff, M. M. (1990). Caregiving and the stress process: an overview of concepts and their measures. *The Gerontologist* 30, 583–594.
- Pinquart, M. and Sorensen, S. (2003). Associations of stressors and uplifts of caregiving with caregiver burden and depressive mood: A meta-analysis. *Journal of Gerontology* 58B(2), P112–P128.
- Roth, D., Haley, W., Owen, J., Clay, O. and Goode, K. (2001). Latent growth models of the longitudinal effects of dementia caregiving: a comparison of African-American and White family caregivers. *Psychology and Aging* 16, 427–436.
- Schulz, R., Mendelsohn, A. B., Haley, W. E., et al. (2003). End-of-life care and the effects of bereavement on family caregivers of persons with dementia. *New England Journal of Medicine* 349, 1936–1942.
- Schulz, R., Belle, S. H., Czaja, S. J., et al. (2004). Long-term care placement of dementia patients and caregiver health and well-being. *Journal of the American Medical Association* 292, 961–967.
- Sorensen, S., Pinquart, M., Habi, X. and Duberstein, P. (2002). How effective are interventions with caregivers? An updated meta-analysis. *The Gerontologist* 42, 356–372.
- Vitaliano, P. P., Zhang, J. and Scanlan, J. M. (2003). Is caregiving hazardous to one's physical health? A meta-analysis. *Psychological Bulletin* 129, 946–972.
- Zarit, S. H. and Leitsch, S. A. (2001). Developing and evaluating community based intervention programs for Alzheimer's patients and their caregivers. *Aging and Mental Health* 5(Supplement), S84–S98.
- Zarit, S. H., Reeve, K. E. and Bach-Peterson, J. (1980). Relatives of the impaired elderly: correlates of feelings of burden. *The Gerontologist* 20(6), 649–655.
- Zarit, S. H., Stephens, M. A. P., Townsend, A. and Greene, R. (1998). Stress reduction for family caregivers: Effects of day care use. *Journal of Gerontology: Social Sciences* 53B, S267–S277.

Caspases See: Apoptosis.

Catecholamines

U Lundberg

Stockholm University, Stockholm, Sweden

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by U Lundberg, volume 1, pp 408–413, © 2000, Elsevier Inc.

Role of the Catecholamines
Summary and Conclusions

Glossary

<i>Dopamine</i>	A transmitter in the central nervous system, with at least five receptors (D1–D5). Dopamine is also a precursor to the formation of epinephrine and norepinephrine and influenced by stress, but generally not considered a stress hormone.
<i>Epinephrine</i>	A stress hormone secreted from the adrenal medulla in response to stimulation from the sympathetic nervous system. It activates both α - and β -adrenoceptors, with effects on heart rate and dilatation of blood vessels in the muscles and constriction of blood vessels in the skin.
<i>Lipolysis</i>	Refers to the release of lipids into the blood stream.
<i>Norepinephrine</i>	A stress hormone secreted from the adrenal medulla that also serves as a neural transmitter in the sympathetic and central nervous system. Most of the circulating levels of norepinephrine are produced by sympathetic nerve endings. Norepinephrine mainly affects α -adrenoceptors causing constriction of blood vessels and increases in blood pressure.
<i>Sympathetic adrenal medullary (SAM) system</i>	The sympathetic adrenal medullary system, where the catecholamines epinephrine and norepinephrine are secreted.

Catecholamines have been of central interest and importance in stress research since the first demonstrations of the role of sympathetic arousal in response

to stress exposure early in the twentieth century. Numerous animal experiments have illustrated the active defense reaction and the emergency function of the adrenal medulla, which increases the organism's chances of survival by fight-or-flight. During the past decades, a considerable number of studies in humans, in laboratory as well as in natural settings, has confirmed and extended the conclusions from animal studies. The aim of this article is to summarize research and conclusions relevant to the role of catecholamines in stress and health, including their assessment and methodological considerations, as well as gender differences in catecholamine responses to stress.

Role of the Catecholamines

Secretion of the catecholamines is regulated by mental influence (from the cortex) on the hypothalamus in the central nervous system, which activates the sympathetic adrenal medullary (SAM) system. Walter B. Cannon first demonstrated the important role of sympathetic activation in emotional responses to stress exposure. Norepinephrine was identified later by Ulf von Euler (1948). On the basis of animal experiments, Cannon described the fight-or-flight response or the emergency function of the adrenal medulla. The SAM system is activated when the individual is challenged in its control of the environment and this defense reaction prepares the body for battle. Catecholamines do not cross the blood–brain barrier but exert their effects peripherally.

The cardiovascular and neuroendocrine functions activated by catecholamines mobilize energy to the muscles, heart, and brain and, at the same time, reduce blood flow to the skin, internal organs and gastrointestinal system. In response to physical threat this is an efficient means for survival as it increases the organism's capacity for fight-or-flight. In modern society, however, elevated catecholamine levels are probably more often caused by threats of a social or mental rather than physical nature. In general, mental efficiency is also positively associated with elevated catecholamine levels. Mental stress is a very potent stimulator of epinephrine output, whereas

norepinephrine, having an important role in blood pressure homeostasis, is more closely associated with physical activity and demands and body posture.

Numerous studies from laboratory experiments, as well as from natural settings, illustrate the sensitivity of the catecholamines to various physical, psychological, and psychosocial conditions in humans and animals. In experimental studies, catecholamine responses have been examined in response to physical stressors, such as noise, electric shocks, heat, cold, and so on, and to mental and cognitive stressors, such as performance tests, harassment, cognitive conflict, time pressure, monotony, fear, and anticipation of stress. The influence of various real-life conditions on catecholamines have been investigated during daily stress at work, home, school, day-care centers, hospitals, on commuter trains or buses, and during intense stress induced by, for instance, parachute jumping, child birth, and the defense of a doctoral thesis.

Health and Behavior

Catecholamines and their concomitant effects on other physiological functions, such as blood pressure, heart rate, and lipolysis, may serve as objective indicators of the stress that an individual is exposed to. However, these bodily effects are also assumed to create a link between psychosocial stress and increased health risks. Lasting elevated catecholamine levels are considered to contribute to the development of atherosclerosis and predispose to myocardial ischemia. The elevated catecholamine levels also make the blood more prone to clotting, thus reducing the risk of heavy bleeding in case of tissue damage but at the same time, increasing the risk of arterial obstruction and myocardial infarction. The role of the catecholamines in hypertension is also of great importance.

Elevated catecholamine secretion in response to challenge has been demonstrated for the stress-related and coronary-prone type A behavior pattern. Additional support for a role of the catecholamines in the development of coronary heart disease (CHD) is obtained from studies showing that high job strain (high demands and low control) and high catecholamine levels are related to elevated risk of CHD. Catecholamine activation is also known to interact with the other major neuroendocrine stress system, namely, the hypothalamic-pituitary-adrenal (HPA) axis and the secretion of cortisol, and form an integrated stress response, which seems to potentiate several health risks including CHD. Negative emotional stimulation activates via higher brain centers both the secretion of catecholamines and cortisol. High epinephrine levels seem to enhance the HPA response and low levels to diminish it and, the corticotropin releasing hormone

(CRH), which activates the HPA axis, also seems to stimulate sympathetic outflow.

Catecholamines also serve an important role in human performance and behavior. The secretion of the catecholamines is not only related to the individual's physical capacity, but also to factors such as mental effort and cognitive performance. Under normal levels of stress and arousal, a further increase in catecholamine output is associated with improved performance, whereas under conditions of very intense stress, additional increase in catecholamine output may cause deterioration in mental functioning and disorganized behavior. In summary, the level of performance and mental wellbeing can be described as an inverted-U function of the intensity of stress, which means positive correlations with catecholamine secretion at low and moderate levels of stress and negative at very high levels. The optimal level of performance is located at higher stress levels for simple tasks compared with more complex ones, and varies between individuals depending on personality characteristics, mental state, and so on.

Assessment

Blood levels of catecholamines change rapidly (within 1 minute) in response to stress exposure. A small but relatively constant fraction of the circulating levels of epinephrine and norepinephrine in the blood is excreted into the urine. Consequently, assessment can be made from blood (usually plasma) as well as from urine. In general, results from studies based on blood and urine samples are highly consistent, showing significant positive relationships between changes in urinary and plasma catecholamines in response to stress.

Plasma catecholamine levels readily reflect acute stress responses during exposure to short-term stress and have been used extensively in animal research. Urinary catecholamines also respond rapidly to stress exposure but, depending on the time between urine sampling, the urine in the bladder provides an integrated measurement representing the mean stress level for a more extended period of stress exposure. Using voluntary voiding, the normal period between urine sampling is 2–3 h, which means that urinary catecholamines are preferable for the study of longer periods of stress or chronic stress exposure. A normal procedure for studying long-term stress could involve taking urine samples at regular intervals over days or weeks.

Due to the rapid changes in plasma catecholamine levels, frequent blood sampling via an indwelling catheter is necessary in order to obtain representative measures of circulating levels for longer periods of time, which is usually not feasible in, for example, the

study of work stress. Other problems associated with venipuncture are that norepinephrine concentrations vary between sampling sites due to differences in regional blood flow and reuptake by the nerve terminals, and that blood sampling in itself may cause increased levels of epinephrine secretion.

For methodological reasons, human plasma catecholamines have been studied almost exclusively in the laboratory, whereas urinary catecholamines have been used extensively in the study of human stress in natural settings. The advantages of urinary measurements in the study of, for example, work stress, are that they do not interfere with the subject's normal habits and environment and that they cause no harm or pain to the subject. In addition, integrated measurements are provided for longer periods of time – relevant for the study of the psychosocial stress associated with different conditions at work.

The amount of urinary catecholamines excreted during a particular period of time can be determined by the concentration of the sample, multiplied by the total urine volume. Thus, in the study of urinary catecholamines, the total volume of the sample has to be measured after voiding. In addition, the pH level of the urine should be adjusted to about 3.0 (using acid) to prevent the catecholamines from deterioration. The acidified urine sample should then be kept frozen (−18 °C) until analyzed.

In cases where the time between voiding of urine is unknown or the total volume of the sample has not been properly measured, the catecholamine concentration can be related to a reference substance such as creatinine, which is considered to be secreted at a constant rate. Provided that reliable measurements are obtained and the subject is able to empty his or her bladder completely, the results from these two methods are highly correlated. The most common method for catecholamine assessment in blood or urine is high-performance liquid chromatography (HPLC) with electrochemical detection. This and other modern methods (e.g., enzymatic methods) have high sensitivity and specificity. An earlier method used extensively during the period 1960–80 is the fluorophotometric method of Euler and Lishajko, which is reliable for the determination of urinary catecholamines but not sensitive enough for the analysis of plasma, where the concentration of the catecholamines is lower.

Methodological Considerations and Confounders

Catecholamines have a pronounced circadian pattern, which has to be taken into consideration in the study of responses to stress. Under normal night-sleep–day-wake conditions, the catecholamines peak

in the middle of the day and reach their lowest levels during night sleep. The circadian rhythm of epinephrine remains relatively stable even during several nights of sleep deprivation, whereas norepinephrine is more influenced by physical activity. Consistent changes in the sleep–wake pattern, for example in habitual night work or east–west traveling over 12-h time zones, will completely reverse the circadian rhythm of the catecholamines in about 1 week. Variations in shift work and frequent east–west flights will disturb the normal circadian rhythm and produce a flattened curve of catecholamine secretion as well as problems in staying alert and sleep disturbances. Other (nonpsychological) factors influencing catecholamine secretion are the intake of caffeine (coffee), alcohol, nicotine (cigarette smoking), medication (β -blockers, diuretics, etc.), and heavy physical exercise.

In the study of stress and catecholamines, the influence by confounders has to be controlled by requesting the participants to refrain from smoking, drinking coffee, and so on, or by controlling the effects on the catecholamines by asking the participants to consume the same amount of cigarettes and coffee in the control condition as in the experimental situation.

Moderate or small seasonal variations in catecholamines have been reported, as well as variations during the menstrual cycle. Norepinephrine levels have also been found to increase with old age. As related to weight, children have higher levels than adults, but as related to body surface area the excretion rates are relatively stable from the first years of life into adulthood.

Individual variations in baseline levels are pronounced but relatively stable over time. For instance, the highest epinephrine level may be 10 times greater than the lowest. This variation is likely to be related to individual differences in renal clearance and body weight, but psychological differences between individuals with high and low baseline catecholamine levels have also been reported. During mild stress, such as carrying out normal daily work, the epinephrine output increases by 50–100% above the resting level, whereas more severe mental stress and intense cognitive and emotional demands, such as child birth, may cause epinephrine to rise more than 10 times the resting level. There are no normal levels of catecholamine and no threshold limits, but clearly pathological levels can be found in association with, for example, adrenal tumors.

In order to reduce the influence of circadian rhythms and of individual differences in baseline levels, it is generally recommended that the individual's response to stressful conditions be compared with his or her corresponding baseline level. The two

measurements have to be obtained at the same time of day, which means that they must be made on separate days. By studying the relative change from baseline rather than absolute levels, the variation between individuals is considerably reduced and the statistical power is increased without having to increase the sample size.

Reliable baseline levels can be obtained in situations where the individual is alone and able to relax completely. In an experimental study this can be achieved by inviting subjects to come to the laboratory after the experimental session and making sure that they are confident that they will not be exposed to any kind of stressor. In order to help the subject to relax, a quiet environment, light entertainment with music and/or suitable reading material, as well as relaxation training can be used. In studying stress in natural settings, such as the work environment, baseline levels of catecholamines can be obtained by giving the worker a paid day off from work. He or she is then instructed to relax alone at home and avoid any activity that might contribute to stress, such as shopping, cleaning, taking care of children, heavy physical exercise, and so on.

Epinephrine secretion is particularly sensitive to novel stimulation and stress associated with anticipation. Exposure to the laboratory and the experimental equipment to be used may induce as high or higher epinephrine levels than the actual stress tests. This means that catecholamine levels before an experiment usually cannot be used as baselines. The anticipation and novelty effects can be reduced or eliminated by asking the subject to visit the laboratory before the experiment to become familiar with the methods of measurement and the equipment to be used. Similarly, at the workplace, the stress induced by the investigation *per se* can be avoided by performing measurements (urine sampling, etc.) one or several days before the actual data collection, to make the participants familiar with all aspects of the study.

Gender Differences

Up to 1970, almost all human studies were performed on men. However, when gender differences were studied in the early 1970s, men were consistently found to be more responsive to experimental stress by showing greater elevations in epinephrine than women. Although women performed as well or usually even better than the men on the various stress tests used in the experiments, their epinephrine secretion did not increase very much. During more intense stress (examination stress), the epinephrine output of women

was found to increase significantly, but still to a lesser extent than that of men in the same situation.

A possible explanation for these gender differences is that the performance tests used to induce stress were less challenging to women than to men. In keeping with this, it was found that emotional stress associated with accompanying a child to hospital had a more pronounced effect on the mothers' catecholamine levels. The importance of gender roles is also illustrated by the fact that women in male-dominated occupations and lines of education seem to respond to performance stress with epinephrine responses similar to their male colleagues. Recent studies comparing men and women matched for education and occupational level showed similar catecholamine responses to stress in women as in men.

Although, the possible influence of biological factors, such as steroid sex hormones, on catecholamine responses cannot be excluded, it seems as if psychological factors and gender role patterns, in general, are more important for the differences in catecholamine responses between men and women.

In men, a significant positive correlation is usually found between perceived stress and performance on the one hand, and catecholamine responses at work, on the other. The more intense or frequent the stress exposure at work the higher the catecholamine output. In women, however, corresponding relationships are more inconsistent. Stress exposure and catecholamine levels at work seem to spill over into nonwork situations. Women with high catecholamine levels at work tend to have high catecholamine levels after work, and women exposed to elevated work stress (e.g., by working overtime) have elevated catecholamine in the evening or during the weekend at home. A likely explanation for this is that there is a greater interaction between the stress from paid employment and unpaid work at home for women than there is for men. This has to be taken into consideration in the study of women's stress and workload.

Positive and Negative Stress

As described by Henry, activation of the SAM system and high catecholamine output are associated with an active coping style and dominant behavior among animals, whereas passive coping and the defeat reaction are associated with activation of the HPA axis, high cortisol levels, and submissive behavior. A similar but less consistent pattern of findings has been reported in humans. Recently, an association between high epinephrine and cortisol levels and low socioeconomic status has been found.

In human stress, epinephrine levels are significantly elevated by stress caused by over- as well as by understimulation, compared with more optimal environmental conditions. This means that work overload, a very high work pace, too much responsibility, role conflicts, as well as simple, monotonous and repetitive jobs or lack of meaningful activities (e.g., unemployment), may contribute to elevated epinephrine levels. However, as demonstrated by Levi, epinephrine output reflects the intensity of stress rather than its emotional valence. Intensive pleasant stimulation, such as watching a funny movie, may induce elevated epinephrine levels as well as does unpleasant stimulation.

Summary and Conclusions

The two catecholamines, epinephrine and norepinephrine, are very sensitive to behavioral, emotional, and cognitive stimulation and the magnitude of the response seems to be closely related to the intensity of perceived stress regardless of the emotional quality. Epinephrine mainly reflects mental stress, whereas norepinephrine is more sensitive to physical demands and body posture. Much less is known about the role of dopamine under stress, but several studies have shown that stress-induced dopamine effects have an important role in the pathogenesis of certain brain diseases and in drug addiction.

Catecholamines may thus serve as objective indicators of the intensity of stress that an individual is exposed to. However, they also play an important role in the development of several health problems and, thus, create a link between stress and disease. In order to obtain valid and reliable measurements, strict control of possible confounders is necessary at all stages of the process – preparing the subjects, determining baselines, collecting, preparing, and storing the samples, performing the hormone assay, and so on.

Acknowledgments

Financial support has been obtained from the Swedish Research Council and the Bank of Sweden Tercentenary Foundation.

See Also the Following Articles

Adrenal Medulla; Fight-or-Flight Response.

Further Reading

- Cannon, W. B. (1914). The emergency function of the adrenal medulla in pain and the major emotions. *American Journal of Physiology* 33, 356–372.
- Cohen, S., Doyle, W. J. and Baum, A. (2006). Socio-economic status is associated with stress hormones. *Psychosomatic Medicine* 68, 414–420.
- Frankenhaeuser, M. (1983). The sympathetic-adrenal and pituitary-adrenal response to challenge: comparison between the sexes. In: Dembroski, T. M., Schmidt, T. H. & Blümchen, G. (eds.) *Biobehavioral bases of coronary heart disease*, pp. 91–105. Basel, New York: Karger.
- Henry, J. P. (1992). Biological basis of the stress response. *Integrative Physiological and Behavioral Science* b, 66–83.
- Henry, J. P. and Stephens, P. M. (1977). *Stress, health, and the social environment. A sociobiologic approach to medicine*. New York: Springer-Verlag.
- Hjemdahl, P., Larsson, P. T., Bradley, T., et al. (1989). Catecholamine measurements in urine with high-performance liquid chromatography with amperometric detection – comparison with an autoanalyser fluorescence method. *Journal of Chromatography* 494, 53–66.
- Levi, L. (1972). Stress and distress in response to psychosocial stimuli. *Acta Medica Scandinavica Supplementum* 528, 1–166.
- Lundberg, U. (1996). The influence of paid and unpaid work on psychophysiological stress responses of men and women. *Journal of Occupational Health Psychology* 1, 117–130.
- Lundberg, U. (2005). Stress hormones in health and illness: the roles of work and gender. *Psychoneuroendocrinology* 30, 1017–1021.
- Lundberg, U., de Château, P., Winberg, J., et al. (1981). Catecholamine and cortisol excretion patterns in three year old children and their parents. *Journal of Human Stress* 7, 3–11.
- von Euler, U. S. (1948). Identification of the sympathomimetic ergone in adrenergic nerves of cattle (sympathin N) with laevo-norepinephrine. *Acta Physiologica Scandinavica* 16, 63–74.
- von Euler, U. S. and Lishajko, F. (1961). Improved technique for the fluorimetric estimation of catecholamines. *Acta Physiologica Scandinavica* 51, 348–355.

Cell Death *See:* Apoptosis.

Central Nervous System *See:* Brain and Brain Regions.

Central Stress Neurocircuits

S Kollack-Walker, H E W Day and H Akil

University of Michigan, Ann Arbor, MI, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S Kollack-Walker, H E W Day, and H Akil, volume 1, pp 414–422, © 2000, Elsevier Inc.

Overview of Stress: Inputs, Outputs, and Evaluative Processes

Hypothalamic-Pituitary-Adrenocortical (HPA) Axis

Initiation of the Stress Response

Termination of the Stress Response

Concluding Remarks

Glossary

<i>Homeostasis</i>	Maintaining constancy of the internal environment of the body.
<i>Stressor</i>	Stimulus input that initiates a stress response.
<i>Stress response</i>	A nonselective response (endocrine, behavioral, autonomic) that targets multiple tissues and organs in response to, or anticipation of, a challenge to homeostasis.

An understanding of central stress neurocircuits requires knowledge of neuroanatomical pathways that underlie detection of various stimuli leading to a stress state, perception of those inputs as stressful (when appropriate), and mediation of select response outputs. In this article, the basic neurocircuitry involved in regulation of the hypothalamic-pituitary-adrenocortical (HPA) axis is reviewed. The HPA axis mediates secretion of glucocorticoids in response to a variety of stimuli ranging from immune responses to hemorrhaging to social defeat. Thus, it is important to understand how such diverse stimuli can lead

to activation of the HPA axis. The stress-induced increases in glucocorticoids act to maintain homeostasis as well as to mobilize energy stores for fight-or-flight reactions. Although these actions can and do occur together, it is possible to isolate homeostatic regulation from the more global defense response. These phenomena can be distinct in terms of arousal, emotional salience, conscious perception, and behavioral responsiveness. As a consequence, it is also important to understand how the brain responds to different types of stressful events.

Overview of Stress: Inputs, Outputs, and Evaluative Processes

Stress has been described as a multidimensional concept consisting of three major components: the stimulus input, an evaluative process, and the response output. The stimulus input, or stressor, refers to an event that elicits a nonselective response (endocrine, behavioral, autonomic) that targets multiple tissues and organs in response to, or anticipation of, a challenge to homeostasis. In many cases, the link between stimulus inputs and response outputs requires perception of a given event as stressful. This evaluative process can include not only processing of stimulus-specific information and coding of the stressor's intensity and intermittency, but also comparison of the current situation to previous experiences. In addition, this evaluative process reflects the organism's ability to cope with the stressor (e.g., degree of controllability, real or perceived). The response output, or stress response, is the body's adaptations designed to re-establish physiological or psychological equilibrium. Depending on the particular stimulus input, a specific array of responses may be initiated ([Figure 1](#)).

This concept of stress has been modified to reflect the notion that some stimulus inputs can lead to a stress response without requiring evaluation. Consequently,

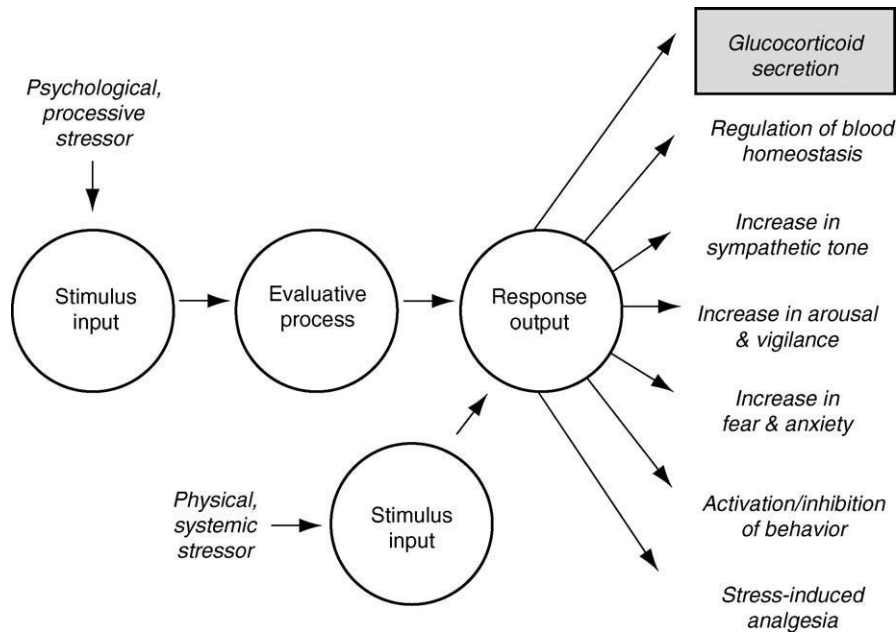


Figure 1 Schematic diagram illustrating the basic relationship between stimulus inputs and response outputs.

various authors in the field have classified stressors in distinct ways: physical versus psychological, or systemic versus neurogenic/processive. The key element associated with such distinctions is the need for a stimulus input to be perceived as stressful. For example, the experience of novelty as a mild stressor requires perception of a situation as different with the potential to encounter threatening stimuli. A novel environment would be classified as processive, involving higher order processing within the central nervous system in order to identify the present situation as unique in comparison to previous experiences. In contrast, other stimulus inputs are thought to elicit a stress response without prior evaluation. For example, hemorrhaging activates various reflexes designed to return blood volume to normal values. In this case, it is not necessary to perceive blood loss as stressful – the reaction is automatic. Hemorrhaging would be considered systemic, involving perturbations in homeostasis or a threat to an individual's internal environment.

Although stressors can be distinguished in terms of the type of stimulus input and the need for evaluative processing, a key component of the response output involves activation of the hypothalamic-pituitary-adrenocortical (HPA) axis. The HPA axis mediates release of glucocorticoids from the adrenal cortex, and glucocorticoids play an important role in energy metabolism and the maintenance of homeostasis. In this article, the basic elements of the HPA axis, including a description of specific brain pathways implicated in HPA axis regulation, are reviewed.

In this discussion, how the brain responds to two distinct stressors – an immune challenge, as an example of a systemic stressor, and social defeat, as an example of a processive and psychological stressor – is considered. Last, the article considers how a stress response is terminated, a process linked to distinct neurocircuits and stress-induced increases in glucocorticoids.

Hypothalamic-Pituitary-Adrenocortical (HPA) Axis

The HPA axis is involved in control of secretion of glucocorticoids from the adrenal cortex. Neurons within the medial parvocellular subdivision of the paraventricular nucleus of the hypothalamus (mpPVN) synthesize corticotropin-releasing hormone (CRH) and project to the median eminence, releasing CRH into hypophyseal portal blood vessels to reach specific cells within the anterior pituitary called corticotrophs. The binding of CRH to its receptor on the corticotroph's cell membrane stimulates a cascade of intracellular events leading to the release of adrenocorticotrophic hormone (ACTH) into the general circulation. CRH is the major ACTH secretagogue, although a number of additional molecules have been implicated in facilitating ACTH release (e.g., vasopressin). ACTH acts upon specific receptors in the adrenal cortex to stimulate the synthesis and release of glucocorticoids into the general circulation. Depending upon the species, the major glucocorticoid secreted from the adrenal cortex may be cortisol (human), corticosterone (rat), or both (hamster).

Initiation of the Stress Response

A variety of approaches have been taken to identify brain regions involved in regulation of the HPA axis. Our current state of knowledge is based upon studies that have utilized tract tracing to identify afferent inputs into the mpPVN, expression of immediate early genes to identify activated brain regions, lesions to determine the importance of a given set of inputs for stress-induced glucocorticoid secretion, and pharmacological and biochemical analyses to determine the role of particular neurotransmitters and neuropeptides in HPA axis regulation. Several key findings obtained from these data are summarized in [Figures 2 and 3](#).

Is there a common set of pathways within the central nervous system critical for initiating a stress response, or do different stressors activate the HPA axis through distinct neurocircuits? The data collected, thus far, provide evidence in support of both arguments. Indeed, systemic (physical) and processive (neurogenic/psychological) stressors appear to activate the HPA axis largely through distinct pathways, although this does not preclude the possibility that multiple pathways may be involved, especially for

complex stressors that have both physical and psychological components. Within the class of processive stressors, similar activation patterns are generally observed, possibly reflecting, at least in part, the activation of core neurocircuits critical for glucocorticoid secretion.

Homeostatic Regulation: Systemic, Physical Stressors

Stressors that are systemic or physical in nature involve a reaction to a threatening stimulus in an individual's internal milieu leading to the activation of regulatory mechanisms that serve to restore homeostasis. A variety of stimulus inputs is under homeostatic control, including body temperature, immune responses, and changes in blood volume and composition. For many of these stimuli, it is absolutely critical that the levels of these variables be maintained within a certain range for survival. Because of this link to survival, systemic stressors are associated with activation of brainstem and circumventricular afferents that have direct projections to mpPVN neurons. In many cases, excitation of the HPA axis is dependent upon ascending catecholaminergic input.

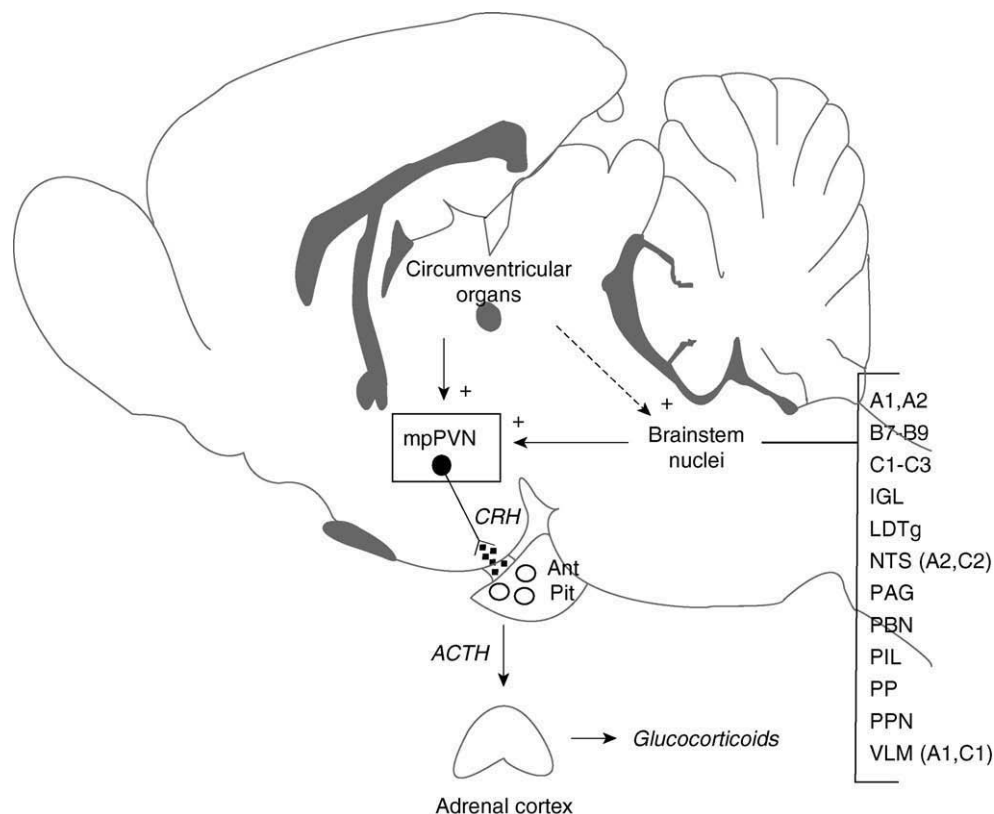


Figure 2 Schematic diagram summarizing the role of brain stem and circumventricular afferents in regulation of the functional activity of mpPVN neurons. This afferent input has been associated with activating the HPA axis in response to systemic and physical stressors.

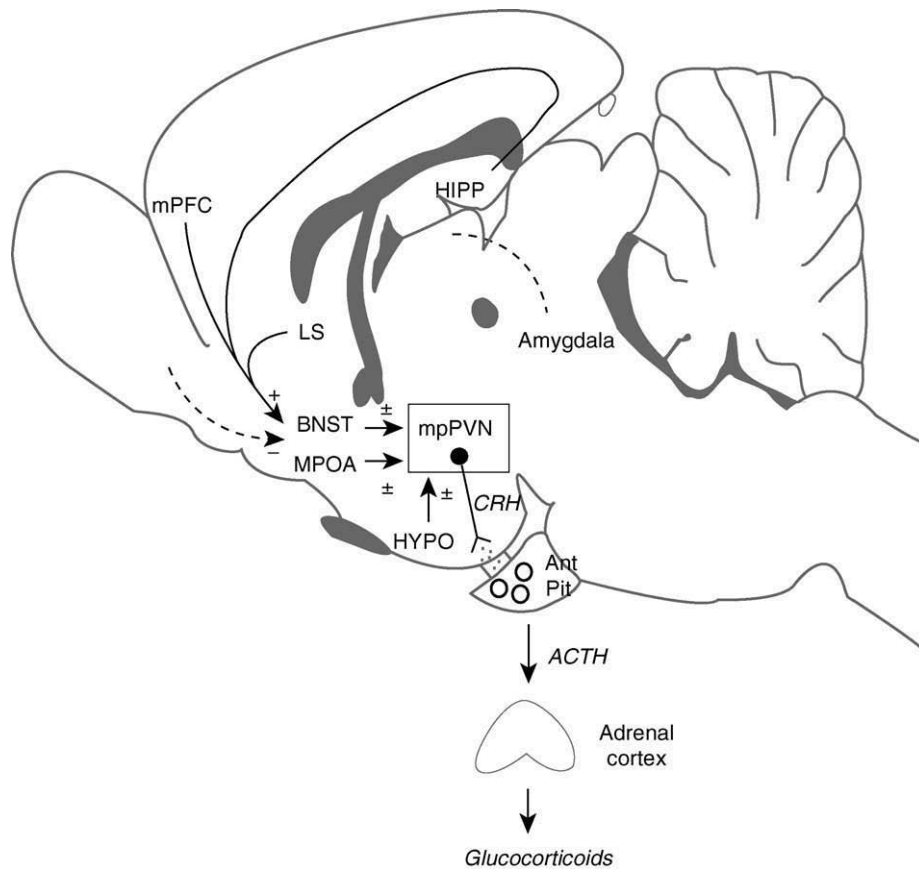


Figure 3 Schematic diagram summarizing the role of BST, MPOA, and hypothalamic and limbic afferents in regulation of the functional activity of mpPVN neurons. The limbic system has been linked to activating the HPA axis in response to processive and psychological stressors.

For example, activation of an inflammatory or immune response increases the production and release of various cytokines (e.g., interleukin-1 β), which leads to excitation of the HPA axis. This rise in glucocorticoids is an adaptive response that serves to inhibit the immune system, limiting its scope of attack and decreasing the likelihood that the immune system will attack self. **Figure 4** illustrates the general pattern of *c-fos* mRNA expression that occurs within the brain of a male rat 30 min following systemic administration of interleukin-1 β (IL-1 β). As evident from the photomicrographs, the pattern of neuronal activation is fairly localized, with increased expression of *c-fos* mRNA observed within bed nucleus of the stria terminalis (BST), central amygdaloid nucleus (CeA), mpPVN, parabrachial nucleus (PBN), nucleus of the solitary tract (NTS), circumventricular organs including area postrema (AP), and barrier-related structures including meninges, cerebral vasculature, and choroid plexus. Additional studies have reported activation of neurons within the medial preoptic area (MPOA), supraoptic nucleus of the hypothalamus (SON), magnocellular and autonomic subdivisions

of PVN, thalamus, and catecholamine cell groups of the ventrolateral medulla (VLM; A1 and C1 cell groups) and NTS (A2 and C2 cell groups).

Increased expression of *c-fos* mRNA within the PVN presumably reflects excitatory control of the pituitary–adrenocortical axis in response to systemic administration of IL-1 β . Catecholaminergic input from the NTS and ventrolateral medulla is critical for activating mpPVN neurons in response to IL-1 β , as lesions that destroy these ascending pathways can drastically reduce *c-fos* expression within mpPVN neurons and other measures of HPA axis activity. In accordance, *c-fos* mRNA levels are increased within neurons of the adrenergic (C1, C2) and noradrenergic (A2) cell groups, and retrograde tract-tracing studies have shown that the majority of these activated neurons project to mpPVN. Systemic administration of IL-1 β can activate these catecholaminergic cell groups via two main pathways. The NTS receives visceral sensory information carried by the vagus and glossopharyngeal nerves, and there is evidence that these sensory afferents can respond to IL-1 β in the periphery. In addition, neurons within the area

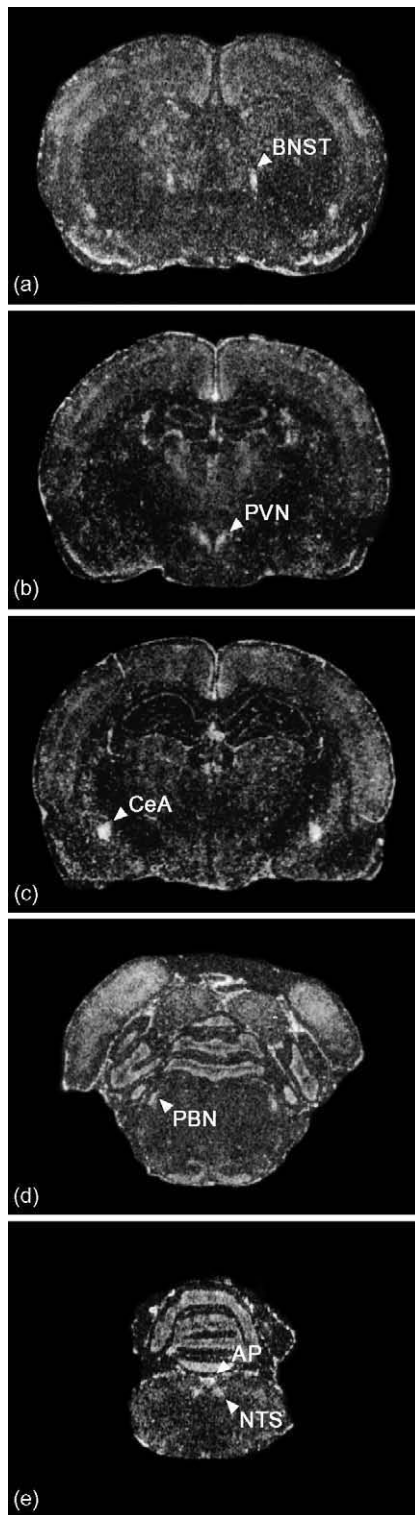


Figure 4 Series of photomicrographs illustrating the location of *c-fos* mRNA within the brain of a male rat following systemic administration of interleukin-1 β . Note the selective pattern of neuronal activation with increases in *c-fos* mRNA present within the bed nucleus of the stria terminalis (BST), paraventricular nucleus of the hypothalamus (PVN), central nucleus of the amygdala (CeA), nucleus of the solitary tract (NTS), and area postrema (AP). PBN, parabrachial nucleus.

postrema can transduce IL-1 β present in the bloodstream into a neural signal that is then transmitted to neurons within the NTS and VLM. In support of this second pathway, *c-fos* mRNA levels are elevated within the area postrema in response to IL-1 β .

The lateral subdivision of the CeA and the dorso-lateral or oval nucleus of the BST (BSTov) also show prominent activation following IL-1 β administration. Both of these brain regions have been implicated in stress responses, particularly autonomic responses, but it is unclear what their role is in the response to IL-1 β administration. They are unlikely to be the primary regions involved in the increase in HPA axis activity for several reasons. First, neurons within the lateral CeA do not project to the PVN, but rather to three main regions – medial CeA, restricted parts of the BST (including the oval nucleus), and the parabrachial nucleus in the pons. Second, while tract-tracing studies show a significant proportion of cells within the NTS and VLM that project to mpPVN and contain Fos protein, a similar finding is not observed for neurons within the CeA or BST. Third, administration of indomethacin, a prostaglandin synthesis inhibitor, blocks the ability of neurons in area postrema to transduce IL-1 β into a neural signal. This effect results in decreased activation of neurons within NTS and VLM, leading to decreased activation of mpPVN neurons and a reduction in ACTH release. However, indomethacin administration does not reduce *c-fos* expression within CeA, suggesting that the mechanisms associated with HPA axis regulation in response to IL-1 β administration are distinct from those controlling activational processes within CeA. That said, there is some evidence to suggest that the ventral BST modulates the HPA axis response to IL-1 β and that the CeA may modulate this BST region.

Fight or Flight Response: Processive, Psychological Stressors

Stressors that are processive and psychological in nature involve a reaction to a real or perceived threatening stimulus in an individual's external environment. Such reactions are associated with increased arousal and vigilance, often eliciting a state of fear (or increased anxiety) and, in many instances, leading to a global response output that has been termed fight or flight. The image conveyed by fight or flight is one of an active state – secretion of glucocorticoids and other neurohormones, activation of the sympathetic division of the autonomic nervous system leading to increases in heart rate, blood pressure, and respiration, and display of behavioral responses designed to effectively cope with the threatening event.

A processive stressor requires the processing of exteroceptive signals from multiple sensory modalities that lead to perception of the stimulus input as stressful. This evaluative process may also reflect psychological variables, including predictability and controllability of the stressor as well as coping strategies. Because of this need for higher level processing, processive stressors are thought to activate limbic afferents that control HPA axis activity via synaptic connections with the BST, MPOA, and hypothalamus.

The effects of social defeat in the male Syrian hamster have been studied as one model of a processive, psychological stressor. Using a resident-intruder paradigm, the pairing of two isolated male hamsters results in fighting, with one male (dominant) behaving aggressively, attacking and chasing the other male (subordinate), which reacts submissively, defending against attack and displaying flight. The subordinate male is stressed as evidenced by elevated levels of glucocorticoids, while the dominant male is not. **Figure 5** illustrates the general pattern of *c-fos* mRNA expression that occurs within the brain of a socially defeated male 30 min after the onset of fighting. This series of photomicrographs reveals a global pattern of activity encompassing vast regions of the neocortex, septum, amygdala, hypothalamus, and brainstem. By analyzing *c-fos* mRNA expression within different control groups and in dominant and subordinate males, it is possible to identify patterns of neuronal activity associated with handling, fighting, and the stress of defeat. Activation observed in numerous brain regions in handled control, dominant, and subordinate males presumably reflects the detection of exteroceptive signals, leading to an increase in arousal and conscious awareness. In contrast, selective activation of brain regions in subordinate males most likely reflects processes associated with the perception of the stimulus input as threatening and stressful as well as mediation of various response outputs leading to the fight-or-flight response. In subordinate males, *c-fos* mRNA increased selectively within the cingulate cortex (medial prefrontal cortex), lateral septum, BST, MPOA, several hypothalamic nuclei (paraventricular, anterior, ventromedial, arcuate, and supramammillary), CeA, amygdalohippocampal area, periaqueductal gray, dorsal raphe, cuneiform nucleus, and locus coeruleus.

The limbic system plays an important role in regulating HPA axis activity in response to processive and psychological stressors. Following defeat, significant increases in neuronal activation were observed within mPFC, lateral septum, and amygdala. The hippocampus also plays an important role in HPA axis activity, although *c-fos* mRNA levels were not selectively increased within this brain region. This negative finding has been reported following other

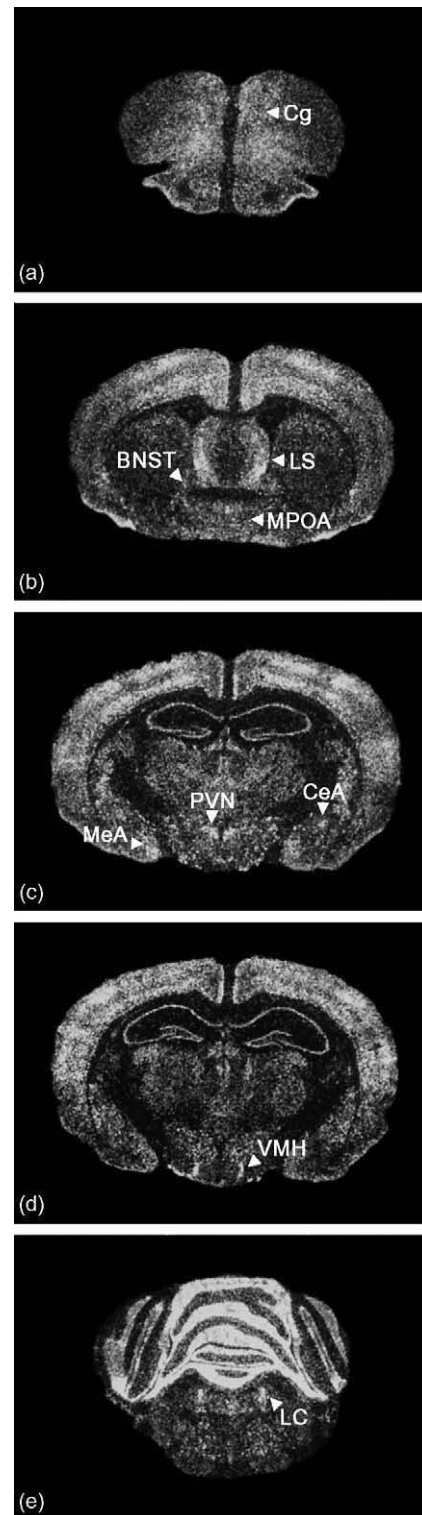


Figure 5 Series of photomicrographs illustrating the location of *c-fos* mRNA within the brain of a male hamster following social defeat. Note the global pattern of brain activity observed throughout the brain. A selective pattern of neuronal activation was observed in the cingulate cortex (Cg), lateral septum (LS), bed nucleus of the stria terminalis (BNST), medial preoptic area (MPOA), paraventricular (PVN) and ventromedial (VMH) nuclei of the hypothalamus (PVN), central nucleus of the amygdala (CeA), and locus coeruleus (LC). MeA, medial nucleus of the amygdala.

stress paradigms. The majority of studies have linked mPFC, lateral septum, and hippocampus to inhibition of glucocorticoid secretion (see next section). In contrast, numerous studies have implicated the amygdala in excitatory regulation of the HPA axis. For example, electrical stimulation of neurons within the medial, cortical, and central nuclei have all been shown to stimulate glucocorticoid secretion. Furthermore, lesions of the medial and central nuclei can decrease ACTH and corticosterone secretion in response to stress. Following defeat, a significant increase in *c-fos* mRNA expression was observed within CeA. This increase is quite unusual for processive stressors, which in general produce quite meager or no increases in *c-fos* mRNA or Fos protein expression in this area. This is in contrast to systemic stressors such as IL-1 β , which cause robust increases of *c-fos* gene products in the CeA. In fact, recent evidence has suggested that many processive stressors, including novelty, loud noise, and restraint, inhibit the neurons in the CeA (and related BSTov). It is possible that brain activity following defeat reflects psychological and physical processes, with increased expression of *c-fos* mRNA within CeA (and dorsolateral BST) synonymous with brain activity observed following systemic or physical stressors (e.g., processing of painful stimuli associated with biting attacks).

Although activated neurons were also detected within the medial and cortical subdivisions of the amygdala, the levels present in subordinate males were not different from those observed in handled control or dominant males. It is at present unclear whether *c-fos* expression within the medial and cortical amygdaloid nuclei following social defeat reflects an increased propensity to stimulate glucocorticoid secretion that is differentially regulated at more central levels (e.g., enhanced in subordinate animals but inhibited in handled control and dominant males). Alternatively, activation within these brain regions may be associated with processes that are not specific to stress, such as coupling sensory cues with behavioral arousal.

While there is some evidence to suggest that the medial and central amygdaloid nuclei can project to the PVN, the input is limited and primarily focused toward autonomic subdivisions of this nucleus. The majority of anatomical studies suggest that the limbic system can influence the HPA axis via synaptic relays within the BST, MPOA, and hypothalamus. Indeed, lesions of neurons within the BST, MPOA, and hypothalamus have been shown to have either excitatory or inhibitory effects on HPA axis activity. Furthermore, tract-tracing studies reveal that many neurons within these three brain areas project directly to

the mpPVN. Following defeat, significant increases in *c-fos* expression were observed throughout the BST, MPOA, and hypothalamus. Additional work will be required to identify specific groups of neurons within these brain areas that are activated during defeat and that project to mpPVN.

The global pattern of brain activation seen following defeat is also observed in response to many other processive or psychological stressors, including conditioned fear, audiogenic stress, predator odor stress, footshock, restraint, immobilization, and swim stress. Following many of these stressors, significant increases in neuronal activation can be seen within various components of the limbic system. Several of these areas have been shown to correlate well with ACTH release following audiogenic stress, including the PVN, ventral lateral septum, medial and ventral BST, and medial preoptic area. It remains to be tested whether some or all of these areas form a core neural circuitry mediating the HPA axis response following processive stress.

Termination of the Stress Response

In addition to understanding how the HPA axis is activated during a stressful event, it is equally important to identify the mechanisms involved in terminating or limiting this response. Specific brain regions and stress-induced increases in glucocorticoids play an important role in terminating the stress response.

Neuronal Regulation

Local neurocircuits utilizing γ -aminobutyric acid (GABA) as a neurotransmitter have been implicated in inhibitory regulation of the HPA axis. Lesions placed within the BST, MPOA, or hypothalamus (including the arcuate, ventromedial, and suprachiasmatic nuclei) can increase the levels of ACTH and corticosterone secreted basally and in response to stress. All of these regions contain substantial populations of GABA-containing neurons, and, most recently, it has been shown that mRNA levels of the GABA biosynthetic enzyme glutamic acid decarboxylase increase within these brain areas in response to stress.

In addition, neurons within the lateral septum, mPFC, and hippocampus have also been linked to inhibitory regulation of the HPA axis. For example, lesions placed within any of these brain regions can prolong glucocorticoid secretion in response to stress. In addition, electrical stimulation of the hippocampus has been shown to inhibit glucocorticoid secretion. These data suggest that one function of these limbic brain regions may be to limit or terminate activity

within the HPA axis. The major neurotransmitter utilized by neurons within the mPFC and hippocampus is the excitatory neurotransmitter glutamate. It is possible that afferents from the mPFC and hippocampus act to excite GABA neurons within the BST, MPOA, and hypothalamus, leading to increased inhibition of the HPA axis.

Hormonal Regulation

Glucocorticoids act to inhibit the HPA axis at multiple levels: pituitary, PVN, and higher brain centers; and in different time domains: fast, intermediate, and slow (genomic). The genomic, and, to some extent, intermediate forms of inhibition appear to involve the classic mechanism of steroid–receptor-mediated regulation of gene transcription. Elevated levels of glucocorticoids bind to glucocorticoid receptors (GRs), activating the receptor and causing it to translocate to the nucleus where the steroid–receptor complex binds to DNA, leading to decreased transcription of CRH and AVP genes in the hypothalamus and decreased transcription of the POMC gene in the pituitary. The mechanisms that underlie fast feedback inhibition are not well understood. This form of inhibition occurs too rapidly to be explained by changes in gene transcription. It is possible that glucocorticoids act at GRs to mediate nongenomic effects, or, alternatively, glucocorticoids may interact with a different class of glucocorticoid-sensitive receptors (e.g., membrane receptors).

Glucocorticoids are also thought to act at higher brain centers to inhibit the HPA axis. For example, lesions of the hippocampus lead to an increase in the expression of CRH and AVP mRNA within mpPVN neurons, even though basal levels of glucocorticoids are elevated, suggesting that glucocorticoid negative feedback has been reduced following hippocampal damage. Lesions of the hippocampus have also been shown to attenuate the ability of exogenous glucocorticoids to inhibit secretion of stress hormones, although not all studies support this contention. Similarly, implants of glucocorticoids into the mPFC or MPOA have been shown to block restraint-induced secretion of stress hormones. These findings suggest that glucocorticoids interact with neuronal elements to increase inhibitory control over the HPA axis. How this effect is mediated at the level of neurons and neurocircuits is presently unknown.

Concluding Remarks

This article was designed to provide a general understanding of how the brain responds to an acute stress

exposure, highlighting important differences in brain activity in response to stressors that are systemic and physical in nature versus those that are processive and psychological. However, the relationship between stress and brain activity may be more complex than simply a reflection of the class of stressor involved. Indeed, how the brain responds to stress can be influenced by a number of associative factors, including the intensity, intermittency, chronicity, and context of the stress exposure as well as the individual's ability to effectively cope with stress. The challenge will be to understand how these more subtle aspects of stress responsiveness are coded within central stress neurocircuits.

See Also the Following Articles

Amygdala; Brain and Brain Regions; Hippocampus, Overview.

Further Reading

- Akil, H. and Morano, M. I. (1996). The biology of stress: from periphery to brain. In: Watson, S. J. (ed.) *Biology of schizophrenia and affective disease*, pp. 15–48. Washington, D.C: American Psychiatric Press.
- Burow, A., Day, H. E. W. and Campeau, S. (2006). A detailed characterization of loud noise stress: intensity analysis of hypothalamo-pituitary-adrenocortical axis and brain activation. *Brain Research* 1062, 63–73.
- Campeau, S., Day, H. E. W., Helmreich, D. L., Kollack-Walker, S. and Watson, S. J. (1998). Principles of psychoneuroendocrinology. *Psychiatric Clinics of North America* 21, 259–276.
- Crane, J. W., Buller, K. M. and Day, T. A. (2003). Evidence that the bed nucleus of the stria terminalis contributes to the modulation of hypophysiotropic corticotropin-releasing factor cell responses to systemic interleukin-1 β . *Journal of Comparative Neurology* 467, 232–242.
- Cullinan, W. E., Herman, J. P., Helmreich, D. L. and Watson, S. J. (1995). A neuroanatomy of stress. In: Friedman, M. J., Charney, D. S. & Deutch, A. Y. (eds.) *Neurobiological and clinical consequences of stress: from normal adaptation to PTSD*, pp. 3–26. New York: Raven.
- Day, H. E. W. and Akil, H. (1996). Differential pattern of c-fos mRNA in rat brain following central and systemic administration of interleukin-1-beta: implications for mechanism of action. *Neuroendocrinology* 63, 207–218.
- Day, H. E. W., Badiani, A., Uslaner, J. M., Oates, M. M., Vittoz, N. M., Robinson, T. E., Watson, S. J. and Akil, H. (2001). Environmental novelty differentially affects c-fos mRNA expression induced by amphetamine or cocaine in subregions of the bed nucleus of the stria terminalis and amygdala. *Journal of Neuroscience* 21, 732–740.
- Day, H. E. W., Nebel, S., Sasse, S. K. and Campeau, S. (2005). Inhibition of the central extended amygdala

- by loud noise and restraint stress. *European Journal of Neuroscience* 21, 441–454.
- Day, T. A. (2005). Defining stress as a prelude to mapping its neurocircuitry: no help from allostasis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 29, 1195–1200.
- Herman, J. P. and Cullinan, W. E. (1997). Neurocircuitry of stress: central control of the hypothalamo-pituitary-adrenocortical axis. *Trends in Neuroscience* 20, 78–84.
- Herman, J. P., Ostrander, M. M., Mueller, N. K. and Figueiredo, H. (2005). Limbic system mechanisms of stress regulation: hypothalamo-pituitary-adrenocortical axis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 29, 1201–1213.
- Kollack-Walker, S., Watson, S. J. and Akil, H. (1997). Social stress in hamsters: defeat activates specific neurocircuits within the brain. *Journal of Neuroscience* 17, 8842–8855.
- Levine, S. and Ursin, H. (1991). What is stress? In: Brown, M. R., Koob, G. F. & Rivier, C. (eds.) *Stress: neurobiology and neuroendocrinology*, pp. 3–21. New York: Marcel Dekker.
- Sawchenko, P. E., Brown, E. R., Chan, R. K. W., Ericsson, A., Li, H.-Y., Roland, B. L. and Kovacs, K. J. (1996). The paraventricular nucleus of the hypothalamus and the functional neuroanatomy of visceromotor responses to stress. *Progress in Brain Research* 107, 201–222.

Cerebral Metabolism, Brain Imaging

K P Ebmeier, C L Donaghey and N J Dougall
University of Edinburgh, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by K P Ebmeier, volume 1, pp 423–428, © 2000, Elsevier Inc.

Imaging Methods

Structural and Functional Brain Changes Associated with Stress-Related States

Conclusions

Glossary

<i>Amygdala</i>	Part of the temporal cortex; associated with anxiety and other emotional function.
<i>Atrophy</i>	Reduction in cell mass.
<i>Cingulate cortex</i>	Evolutionary old limbic cortex overlying the corpus callosum, whose fibers connect the two brain hemispheres.
<i>Cortisol</i>	A glucocorticosteroid.
<i>Dopamine</i>	A neurotransmitter related to motivation, movement, and emotion.
<i>Glucocorticosteroids</i>	Hormones secreted by the adrenal gland, related to stress.
<i>Hippocampus</i>	Evolutionary old part of the temporal lobe, associated with memory and emotional functions; part of the limbic system.
<i>Ligand</i>	Substance binding to a receptor, e.g., for a neurotransmitter.
<i>Orbitofrontal cortex</i>	Part of the frontal cortex nearest to the eyes, connected with limbic structures.
<i>Radiotracer</i>	Radioactive substance injected in small (trace) amounts for diagnostic purposes.

Serotonin

A neurotransmitter among others related to mood and anxiety.

Tomography

Three-dimensional reconstruction of brain maps from radiation given off by the brain (emission tomography) or from shadows of traversing rays (transmission scan).

Imaging Methods

Positron Emission Tomography

Positron emission tomography (PET) uses a scanner to detect the two gamma photons traveling in opposite directions that result from the annihilation of a single positron. The origin of these photons can, therefore, be localized on a straight line between the two detectors registering a coincident signal. Although the energy from positron emission is relatively high, the emitters have a short half-life (minutes), which keeps the administered dose of radiation low. However, a particle accelerator (cyclotron) and a radio chemist are required to generate the radio-emitting isotopes and to synthesize the required compounds *in situ*. Positron emitters in common use are ^{18}F -fluoro-deoxyglucose (FDG), which is taken up into neurons just like the energy provider glucose, but then is not further metabolized so that it reflects the regional pattern of brain activity. Cerebral blood flow markers, such as ^{15}O -labeled water, give an almost identical distribution of activity to FDG. ^{11}C and ^{18}F can be incorporated into many receptor ligands to provide radiolabels representing specific binding to these receptors. Examples are ^{11}C -raclopride, a dopamine $\text{D}_{2,3}$ receptor ligand, ^{11}C -methyl-spiperone, a $\text{D}_{2,3,4}$ ligand, and

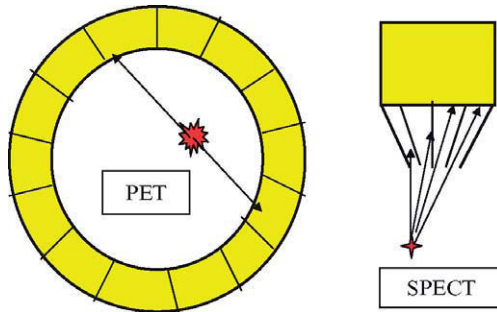


Figure 1 Principles of localizing source of radiation in PET and SPECT (see text).

^{18}F -fluoro-dopa, a marker of dopamine synthesis. The spatial resolution of PET is in the region of 5 mm.

Single Photon Emission Computerized Tomography

Radioisotopes that emit photons in the gamma frequency range with a radioactive half-life in the range of hours can be employed as single photon emission computerized tomography (SPECT) ligands. In contrast to PET, in which theoretically all coincident signals are exploited to localize the source of the radiation, SPECT needs to restrict the field of view of the detectors with collimators in order to localize the origin of the signal (Figure 1). Collimators absorb much of the activity from the field of vision, so that SPECT is less sensitive than PET. Tracers used are $^{99\text{m}}\text{Tc}$ -labeled exametazime and bismuth and ^{123}I -labeled IMP for perfusion imaging and, e.g., ^{123}I -iomazenil for (benzodiazepine) receptor binding. Modern SPECT cameras have a resolution of 5–10 mm.

Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) is based on manipulating the resonance frequencies of nuclei, such as ^1H protons or ^{31}P phosphorus, by systematically changing magnetic field strength across the imaging volume. As magnetic field strength is directly proportional to the magnetic resonance frequency of hydrogen nuclei in H_2O , their spatial position can be frequency encoded. The strength of signal in the frequency domain (k space) then corresponds to the proton density at a particular spatial coordinate. Apart from proton density, relaxation times such as T_1 and T_2^* are extracted from such frequency domain data and provide important information in illness and during the normal functioning of the brain. The resolution for MRI is in the 1–2 mm range.

Functional MRI

T_2^* changes, for example, occur when oxygenated hemoglobin loses its oxygen. In active brain regions, regional blood flow increases beyond the consumption of oxygen, so that a local excess of oxygen-saturated

hemoglobin results. This forms the basis of blood oxygen level-dependent (BOLD) imaging, the most commonly used functional (f) MRI method. fMRI is likely to take over from emission CT as the method of choice for brain activity imaging. It does not involve any radiation exposure and can, therefore, be repeated many times. It also has a greater spatial (2–5 mm) and time resolution (about 100 ms) than PET or SPECT.

Magnetic Resonance Spectroscopy

Local magnetic field strengths vary relative to their molecular, i.e., chemical, context. This results in slight changes of resonance frequency that can be measured with magnetic resonance spectroscopy (MRS). MRS thus allows the chemical imaging of certain biologically important molecules. Nuclei commonly used in MRS are ^1H , which, after suppression of the water signal, yields measures of compounds such as glutamate, lactate, *N*-acetyl-aspartate, choline, and creatinine, and ^{31}P , which gives estimates of the markers of energy metabolism ATP/ADP and of membrane metabolism, phosphomono- or phosphodiesteres, and phosphocreatinine. The volumes studied with MRS lie in the range of several cubic centimeters.

Structural and Functional Brain Changes Associated with Stress-Related States

A number of psychiatric and medical syndromes have a *prima facie* or conceptual association with stress. The obvious connection is with posttraumatic stress disorder (PTSD), but some affective disorders, such as major depressive disorder (MDD) and bipolar affective disorder (BAD), present with aspects of the classical stress response; medically caused hypercortisolism, such as Cushing's syndrome and aging in some patients, may be relevant in this context.

Posttraumatic Stress Disorder

Posttraumatic stress disorder (PTSD), as defined by the *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition, of the American Psychiatric Association, is an anxiety disorder that arises and persists in the aftermath of an emotionally traumatic event. Specifically, patients with PTSD (1) re-experience phenomena that may occur spontaneously or in response to reminders of the traumatic event (i.e., flashbacks), (2) avoid reminders of the event, and (3) exhibit generalized hyperarousal (i.e. exaggerated startle response). One of the core symptoms of PTSD in trauma survivors is a disturbance in memory function. Traumatic events such as acts of terrorism, war, and sexual or physical abuse can trigger PTSD that can persist for months or years following the actual event.

Kolb proposed a neuropsychological hypothesis of PTSD in the late 1980s, which implies that stress-related neuroanatomical changes might underlie symptom formation in PTSD. Numerous empirical studies have provided support for the notion that trauma can trigger neurochemical and neuroanatomical changes leading to emotional dysregulation and exaggerated stress reactivity in susceptible individuals. Such research has implicated the hippocampus, amygdala, anterior cingulate, Broca's area, and medial prefrontal cortex.

Structural imaging studies Adults with PTSD related to combat exposure or childhood physical or sexual abuse have smaller hippocampi compared with healthy volunteers. Findings have been generally consistent. However, results are equivocal as to whether the reduction in hippocampal volume occurs predominantly in the left or the right hippocampus. Suggestions as to why hippocampal volume is reduced in this diagnostic group are that diminished hippocampal size may be either a consequence of exposure to trauma or a risk factor for the development of psychiatric complications such as PTSD.

It has long been known that the hippocampus is involved in memory function and learning; the hippocampus together with the amygdala, orbitofrontal cortex, and anterior cingulate are thought to be part of a circuit involved in the processing of information and the creation of emotional and declarative memories. Core symptoms of PTSD are disturbances of memory function and deficits in attention, so research has aimed to examine the possible relationship between reduced hippocampal volume and memory and attention deficits in this diagnostic group, so far unsuccessfully.

Pitman and colleagues hypothesized in the early 2000s that pre-existing diminished whole brain and hippocampal volume may increase the risk of exposure to traumatic events capable of causing PTSD. Because the hippocampus is involved in declarative memory, those with smaller hippocampi from birth or due to adverse development during childhood may be less able to learn from experience and therefore avoid dangerous situations. Exposure to traumatic events may be increased, with the consequent risk of developing PTSD. Another possibility is that a pre-existing anatomical abnormality may increase the individual's vulnerability to developing PTSD after exposure to a traumatic event. In particular, there is evidence that the hippocampus is involved in the extinction of conditioned responses; individuals with brain abnormalities such as reduced hippocampal volume may, therefore, be poorer extinguishers of conditioned responses and thereby be more likely to

develop PTSD. This vulnerability hypothesis (i.e., primary hippocampal abnormalities result in increased risk of PTSD) needs to be supported by expensive and difficult-to-conduct prospective imaging studies. More economically, examining PTSD-unaffected twins of PTSD victims can tease apart congenital, in lieu of pre-existing, from secondary hippocampal abnormalities. First evidence from this approach is now emerging.

Functional imaging studies The majority of functional brain imaging studies have employed a symptom provocation paradigm, to track the brain equivalents of PTSD and also because measuring brain function at rest makes it difficult to control the variability of possible resting mental states. Symptom provocation paradigms in combination with functional neuroimaging provide an important tool for visualizing neuroanatomical correlates of psychiatric symptoms. Liberzon and colleagues investigated the neurobiological role of limbic brain regions in PTSD, such as amygdala, hippocampal formation, hypothalamus, thalamus, and nearby paralimbic cortex, such as the anterior cingulate cortex, orbitofrontal cortex insula, and temporal poles. Fourteen PTSD patients and two control groups (11 combat controls and 14 healthy volunteers) were scanned with SPECT to measure regional cerebral perfusion in two settings: immediately after exposure to a provocative stimulus and after a nonprovocative control stimulus. During exposure to traumatic stimuli, the researchers noted that the PTSD patients demonstrated the characteristic exaggerated psychological and psychophysiological responses, similar to those documented in previous literature. PTSD patients had increased rCBF in relevant areas of the limbic brain in the vicinity of the left amygdala/nucleus accumbens. No comparable activation was found in either control group. Given the role that these neuroanatomical structures play in the regulation of mental processes in PTSD, the findings from this study support and extend the existing evidence of limbic system and extended amygdala involvement in PTSD symptoms.

Functional MRI studies have employed the masked-faces paradigm to assess automatic amygdala responsivity to general threat-related stimuli in PTSD. Findings from such studies indicate that patients with PTSD exhibit exaggerated automatic responses within the amygdala to general threat-related stimuli, dissociated from medial frontal activation. Furthermore, research in this area has suggested that the magnitude of amygdala response may be able to distinguish subjects with PTSD from those without PTSD, leading to the conclusion that the amygdala plays a fundamental role in the pathophysiology of PTSD.

Although functional neuroimaging studies of patients with PTSD have largely been performed in adult populations, a few have aimed to investigate whether adolescents who developed PTSD after experiencing traumatic events show a differential brain response compared to those who did not. Findings appear to be consistent with those from studies of PTSD in adults.

Magnetic resonance spectroscopy studies Significant biochemical changes affecting the function of an anatomical structure may not be reflected in volumetric changes. Supportive and connective cells and other structural elements may maintain the general shape or volume of a structure despite significant changes in neuron number, size, or viability. One way to assess this possibility is through MRS.

The few MRS studies that have been conducted in PTSD have showed lower levels of N-acetyl-aspartate (NAA) in the hippocampus of patients compared with healthy controls. Research has also examined the relationship between hypothalamic-pituitary-adrenocortical measures and hippocampal NAA in those with and without PTSD. Such research has shown that cortisol levels were positively correlated with hippocampal NAA. This suggests that within the normal range, cortisol may have a trophic effect on hippocampal neurons. This relationship between cortisol and hippocampus may vary depending on the levels of glucocorticoids in particular diagnostic groups. MRS findings so far suggest that hippocampal neuronal viability, density, and function are impaired in PTSD, consistent with the volume loss in hippocampus reported by MRI studies.

Depression

One of the more robust findings in depression research is an impaired regulation of stress hormones. Neuroimaging in MDD has brought about a clearer understanding of the impact of stress mechanisms on the brain. The role of stress in depression is complex and varies with the frequency or duration of stressful life events and the length and number of cycles of depressive episodes. Pooling individuals who have similar lifetime stress burdens and a comparable clinical history for depression in case control imaging studies may be potentially informative, but quantifying lifetime stress burden is challenging and rarely done, as there is no satisfactory, i.e., simple and valid, proxy measure. Interindividual response to stress varies widely, and stressful events are experienced differently from one individual to another.

Structural imaging studies Glucocorticoid steroid receptors are highly concentrated in the hippocampus; the excess of glucocorticoids associated with

repeated stress is neurotoxic in animal models, contributing to local neuronal dysfunction and ultimately cell death. A large number of MRI studies have looked for differences in regional brain volumes, especially hippocampal volume, between depressed subjects and comparison groups, usually healthy volunteers. Less attention has been paid to exploring relationships between hippocampal volume and neuropsychological impairment, hypercortisolism, stress burden of repeated depressive episodes, and estimating volumes of other parts of the brain, for example, the amygdala.

Although brain volume reductions have been reported for frontal cortex, amygdala, and limbic areas including subgenual cingulate cortex, hippocampal reductions, mainly on the left, have been replicated most frequently (Figure 2). Hippocampal volume differences are associated with length and number of depressive episodes and cortisol concentration, but are not associated with age. Evidence from meta-analytic studies suggests that including the amygdala in a combined amygdala-hippocampal structural analysis decreases the ability to detect volumetric differences in either structure. In depressed patients, poor verbal memory performance and later dementia were correlated with hippocampal gray matter reduction, although the evidence for mediation by cortisol is limited.

Functional imaging studies Functional imaging studies (PET) of patients with MDD have generally reported increased cerebral blood flow and metabolism in the amygdala, but reduced flow in dorsolateral prefrontal cortex and anterior cingulate cortex. Furthermore, elevated amygdala activity in MDD was positively correlated with both depression severity and plasma cortisol concentrations.

Although increased amygdala activity could represent an exaggerated response to the stress of imaging itself, amygdala metabolism has been found to be elevated in MDD compared with controls during wakefulness and during rapid eye movement sleep. Treatment with antidepressants reduces metabolism in the anterior cingulate cortex and amygdala to normal levels, with clinical improvement and a reduction in hypercortisolism. Metabolism in the hippocampus of MDD patients has been less well investigated than in the amygdala, but both reductions and increases have been found in the hippocampi of MDD subjects.

Neuroreceptor imaging studies In contrast to structural and functional imaging, there has been relatively little exploration of the role of neurotransmitter systems in the regulation of depression. Both SPECT and PET have been used with radiotracers to investigate

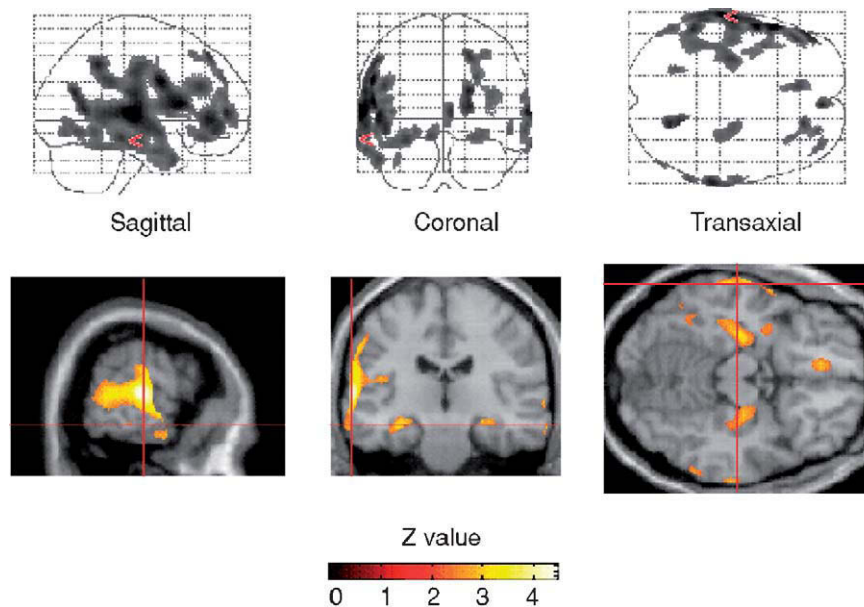


Figure 2 Left (medial) temporal reduction of gray matter density in chronic treatment resistant depression. Data presented in Shal et al. (1998, 2002).

mainly the serotonergic and dopaminergic systems. Serotonin (5-HT) is implicated in the mechanism of depression and clinically modified with selective serotonin reuptake inhibitors (SSRIs). Primarily, the postsynaptic 5-HT_{1A} and 5HT_{2A} receptors have been explored and are among several that are known to mediate serotonergic neurotransmission.

5-HT_{1A} is abundant in the limbic system and, in particular, the hippocampus. In animal studies, stress negatively modulates 5-HT_{1A} receptors through corticosteroid secretion. Decreases in 5-HT_{1A} binding associated with chronic stress can be prevented in animals with antidepressant treatment. Changes observed in the hippocampi of suicide victims with a history of depression have been found to be similar to changes associated with chronic stress in animals, i.e., abnormally decreased hippocampal 5-HT_{1A} levels, suggesting that alterations in hippocampal 5-HT_{1A} levels and in the corticoid receptor balance may be a mechanism by which stress triggers depressive episodes.

In depression, decreased 5-HT_{1A} binding has been found in multiple brain regions, including mesiotemporal cortex, midbrain raphe, as well as hippocampus/parahippocampus, postcentral gyrus, and occipital pole. Differences between mean stressed plasma cortisol concentrations in depression and healthy volunteers are not correlated with reduced 5-HT_{1A} receptor binding. 5-HT_{1A} receptor-binding potentials have been found to be reduced across multiple brain regions in both medicated (SSRIs) and unmedicated depressed subjects compared with healthy volunteers.

Reduced 5-HT_{1A} receptor binding was unchanged by treatment with SSRIs. Results in imaging studies of 5-HT₂ and D₂ receptor binding have been less conclusive in depression and may be associated with subsyndromes.

MRS studies MRS has been used to identify neurochemical effects of treatment and predictors of treatment response with MDD and BDD. Significantly smaller hippocampal regions accompanied by a corresponding significantly elevated choline/creatinine ratio have been found in drug-resistant severe MDD compared with matched healthy controls. However, reductions in choline have also been reported so that there is no current consensus.

Cushing's Disease/Syndrome

Cushing's syndrome (CS) results from the chronic exposure to higher than physiological levels of endogenous cortisol, the naturally occurring glucocorticoid (GC), which often produces neuropsychological and emotional changes in affected patients. Because patients with CS have high circulating levels of endogenous cortisol, they offer an opportunity to study cognitive abnormalities after chronic exposure to elevated levels of cortisol. Structural imaging studies in patients with CS have documented atrophy of the hippocampus. Studies looking at neuronal dysfunction and loss as a result of endogenous levels of glucocorticoids, especially within the hippocampus, may provide a better general understanding of the effects

of high levels of glucocorticoids on hippocampus-mediated cognitive functions and the pathophysiology of CS. Furthermore, patients with CS provide a unique opportunity to study the reversibility of the effects of endogenous cortisol in humans. The potential for reversibility of glucocorticoid-induced alterations of hippocampal structure and function has important theoretical and clinical implications.

In a seminal study, Starkman and colleagues used MRI to investigate the reversibility of cortisol-induced changes in hippocampal volume in 22 patients with Cushing's disease who were examined prior to and following treatment. The results from this study indicate that the hippocampal formation is able to increase in volume following sustained reduction of previously elevated cortisol concentrations. With remission of Cushing's disease, hippocampal volume increased in patients up to 10%. Furthermore, the percent increase in hippocampal volume was significantly correlated with the magnitude of change in urinary free cortisol level, a relationship that was region specific, as there was no significant association between change in cortisol and change in caudate head volume. These results are consistent with observations in animals and demonstrate that hippocampal change in response to sustained elevation of cortisol is reversible, at least in part, once cortisol levels have decreased to normal. However, patients were not imaged prior to the onset of Cushing's disease. It is, therefore, not known whether the increase in hippocampal volume that was observed in this study represents partial or complete reversibility of the decrease of hippocampal volume that occurred as a result of exposure to high levels of cortisol.

Human Aging

Despite evidence of memory decreases during pharmacological glucocorticoid treatment and correlations between memory and cortisol levels in certain disease conditions, such as CS, it remains unclear whether exposure to the endogenous glucocorticoid cortisol levels seen during physical and psychological stress in humans can inhibit memory performance in otherwise healthy adults. Results from research in this area has indicated that exposure to cortisol at doses and plasma concentrations associated with physical and psychological stress in humans can reversibly decrease verbal declarative memory function in otherwise healthy humans. It has been suggested that the results from this area of research are relevant to the interpretation of decreased memory performance during periods of extended stress in humans, in which plasma cortisol elevations are present throughout memory encoding and retrieval processes.

There is further evidence that elderly human subjects with high cortisol levels increasing over the years have impaired hippocampus-dependent memory and show a reduction in hippocampal volume compared with elderly subjects who have moderate cortisol levels that have been decreasing over the years.

One of the main issues that have arisen from research relating to reduced hippocampal volume and impairment on hippocampus-mediated memory function in elderly populations is whether deficits are due to the atrophying effects of increased cortisol levels or some other underlying organic brain pathology such as dementia of the Alzheimer's type. Because age is a risk factor for dementia, it becomes increasingly difficult in older adults to determine the etiology of reduced hippocampal volume and hippocampus-mediated memory.

Conclusions

Technical advances in the field of neuroimaging have made it possible to examine brain activity and brain structure in life. It is possible to measure the concentration of biologically relevant molecules and the number of receptors in a particular neurotransmitter system. The usual way of employing these methods is by selecting a group of patients and a matched control group and comparing them with regard to an objective measure derived from the brain image. There are two possible limitations in our ability to draw valid conclusions from such data. The first is that because of the cost of imaging studies, a relatively small sample from the larger population of patients has to be selected. Whether the sample actually studied is representative for the disorder depends on whether its selection has been subject to any bias. Many sources of bias can occur in the selection for an investigation that needs a significant amount of cooperation from subjects. The ideal, i.e., that subjects are randomly selected from the underlying population, is virtually never achieved, so that conclusions always have to be tentative and need independent replication. The other limitation is inherent in the cross-sectional design of virtually all studies. An association study usually does not allow for any establishment of causality. Hippocampal atrophy may be the result of chronic stress, or it may be the predisposing factor that makes subjects vulnerable to developing the stress-related disorder. To circumvent this difficulty, longitudinal follow-up designs are necessary. These are costly and require an early identification of subjects at risk. They are also likely to require a longer time to come to fruition, so regarding the relevance of the changes described to date, the jury is still out.

See Also the Following Articles

Depression and Manic-Depressive Illness; Major Depressive Disorder; Posttraumatic Stress Disorder – Clinical; Posttraumatic Stress Disorder – Neurobiological basis for; Posttraumatic Stress Disorder, Neurobiology of.

Further Reading

- Anand, A. M. I. T. and Shekhar, A. N. A. N. (2003). Brain imaging studies in mood and anxiety disorders: special emphasis on the amygdala. *Annals of the New York Academy of Sciences* 985(1), 370–388.
- Bremner, J. D., Randall, P., Vermetten, E., et al. (1997). Magnetic resonance imaging-based measurement of hippocampal volume in posttraumatic stress disorder related to childhood physical and sexual abuse—a preliminary report. *Biological Psychiatry* 41(1), 23–32.
- Campbell, S., Marriott, M., Nahmias, C. and MacQueen, G. M. (2004). Lower hippocampal volume in patients suffering from depression: a meta-analysis. *American Journal of Psychiatry* 161(4), 598–607.
- de Kloet, E. R., Vreugdenhil, E., Oitzl, M. S. and Joels, M. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Reviews* 19(3), 269–301.
- Drevets, W. C. (2003). Neuroimaging abnormalities in the amygdala in mood disorders. *Annals of the New York Academy of Sciences* 985(1), 420–444.
- Drevets, W. C., Frank, E., Price, J. C., et al. (1999). Pet imaging of serotonin 1A receptor binding in depression. *Biological Psychiatry* 46(10), 1375–1387.
- Grossman, R., Buchsbaum, M. S. and Yehuda, R. (2002). Neuroimaging studies in post-traumatic stress disorder. *Psychiatric Clinics of North America* 25(2), 317–340.
- Liberzon, I., Taylor, S. F., Amdur, R., et al. (1999). Brain activation in PTSD in response to trauma-related stimuli. *Biological Psychiatry* 45(7), 817–826.
- Lopez, J. F., Chalmers, D. T., Little, K. Y. and Watson, S. J. (1998). Regulation of serotonin1a, glucocorticoid, and mineralocorticoid receptor in rat and human hippocampus: implications for the neurobiology of depression. *Biological Psychiatry* 43(8), 547–573.
- O'Brien, J. T., Lloyd, A., McKeith, I., Gholkar, A. and Ferrier, N. (2004). A longitudinal study of hippocampal volume, cortisol levels, and cognition in older depressed subjects. *American Journal of Psychiatry* 161(11), 2081–2090.
- Sala, M., Perez, J., Soloff, P., et al. (2004). Stress and hippocampal abnormalities in psychiatric disorders. *European Neuropsychopharmacology* 14(5), 393–405.
- Shah, P. J., Ogilvie, A. D., Goodwin, G. M. and Ebmeier, K. P. (1997). Clinical and psychometric correlates of dopamine D2 binding in depression. *Psychological Medicine* 27, 1247–1256.
- Shah, P. J., Ebmeier, K. P., Glabus, M. F. and Goodwin, G. M. (1998). Cortical grey matter reductions associated with treatment-resistant chronic unipolar depression. Controlled magnetic resonance imaging study. *British Journal of Psychiatry* 172, 527–532.
- Shah, P. J., Glabus, M. F., Goodwin, G. M. and Ebmeier, K. P. (2002). Chronic, treatment-resistant depression and right fronto-striatal atrophy. *British Journal of Psychiatry* 180, 434–440.
- Starkman, M. N., Giordani, B., Gebarski, S. S., Berent, S., Schork, M. A. and Schteingart, D. E. (1999). Decrease in cortisol reverses human hippocampal atrophy following treatment of Cushing's disease. *Biological Psychiatry* 46(12), 1595–1602.
- Videbech, P. and Ravnkilde, B. (2004). Hippocampal volume and depression: a meta-analysis of MRI studies. *American Journal of Psychiatry* 161(11), 1957–1966.

Chaperone Proteins and Chaperonopathies

A J L Macario and E Conway de Macario
New York State Department of Health and The
University of Albany (SUNY), Albany, NY, USA

Published by Elsevier Inc.

Pathology and Medicine
Mechanism
To Fold, Refold, or Degrade

Proteins
Definition
Classification
Biology
Phylogenetic Domains
Physiology

Glossary

Chaperone machine A functional chaperoning complex formed by the chaperone Hsp70(DnaK), the co-chaperones Hsp40(DnaJ), and a nucleotide-exchange factor, for example, GrpE in prokaryotes and BAG-1 (BCL2-associated pathanogen, where BCL2

	stands for B cell lymphoma 2) and HspBP1 (Hsp70-binding protein 1) in eukaryotes.
<i>Chaperonin</i>	A molecular chaperone belonging to the 60-kDa family.
<i>Chaperonin system type (group) I</i>	A chaperoning complex formed by the 60-kDa chaperonin and a 10-kDa co-chaperonin (e.g., GroEL and GroES, respectively, in bacteria) that is present in bacteria, in the bacterial-related eukaryotic-cell organelles, and in some archaea.
<i>Chaperonin system type (group) II</i>	A chaperoning complex formed by subunits of the 60-kDa family in archaea (i.e., the thermosome subunits) and by related proteins of various sizes in the eukaryotic cell cytosol.
<i>Molecular chaperone</i>	A molecule that assists polypeptides in the process of folding to achieve a final functional conformation or to regain such conformation after reversible denaturation. The prototype is the stress protein called Hsp70(DnaK).
<i>Phylogenetic domain</i>	The highest order of classification of all living cells (domain, kingdom, phylum, class, order, family, genus, species) that represents the major evolutionary lines of descent, or phylogenetic lineages. Three domains have been identified: Bacteria (or eubacteria), Archaea (archaeobacteria), and Eucarya (eukarya, eukaryotes); the former two are prokaryotes.
<i>Stress (heat shock) proteins</i>	The products of the stress (heat shock) genes. Many of these are molecular chaperones, for example, the components of the chaperone machine.
<i>Stress (heat shock) response</i>	The series of intracellular events caused by cell stressors. One of the most prominent of these events is the induction of the stress (heat shock) genes, resulting in the production of stress (heat shock) proteins.

Proteins

A protein molecule is a chain whose links are amino acids. The number, type, and order of amino acids in the chain are specified by the gene that encodes the protein, and they determine the primary structure. To reach a functional status, a protein must acquire higher levels of complexity. It must fold onto itself to achieve secondary and tertiary levels of structure. In addition, some proteins must associate with others to form a quaternary structure before they can become fully functional.

The functional conformation of a protein, once acquired, may be partially or completely lost in a process called protein denaturation. This can happen when a cell is stressed. There are a number of cell stressors of a physical or chemical nature that cause

protein denaturation. Although the information necessary to achieve a final functional conformation is contained in the primary structure, many proteins need assistance to fold properly or to refold after reversible denaturation. This assistance is provided by molecular chaperones.

Definition

Chaperone proteins, or molecular chaperones, are proteins that assist others to fold properly during or after synthesis, to refold after partial denaturation, and to translocate to the cellular locales at which they reside and function. Chaperones are also involved in the regulation of their own genes and in the presentation of proteins destined for degradation by proteases. Some chaperones are themselves proteases.

Many chaperones are stress (heat shock) proteins, in that they increase in quantity on cell stress (also termed heat shock, although heat is not the only stressor known) so as to counteract the protein-denaturing effects of stressors (e.g., heat, alcohol, ischemia-hypoxia, heavy metals, and hypersalinity). For this reason, stress proteins and molecular chaperones are usually treated together, as a single large group of heterogeneous components, termed Hsp (for heat shock protein). Thus, there are many cell stressors that cause stress or heat shock and many stress or heat shock proteins, but not all of the latter are molecular chaperones. Likewise, not every molecular chaperone is a stress protein.

As may be inferred from their multiplicity of function, molecular chaperones are promiscuous; they interact with a variety of molecules. They are also ubiquitous; they occur in all cells, tissues, and organs, with very few known exceptions. Furthermore, molecular chaperones are present in the various compartments of the eukaryotic cell, nucleus, cytosol, and organelles such as mitochondria and chloroplasts; they are also present in the prokaryotic cell cytoplasm and periplasm.

Classification

Stress proteins, including the chaperones, are grouped into families on the basis of similarities in molecular masses ([Tables 1 and 2](#)). The genes encoding most of these proteins are stress-inducible, but they are also expressed constitutively (i.e., in the absence of stress) at lower levels.

The best studied are the Hsp70 and Hsp60 families, but considerable progress has also been made in the investigation of proteins belonging to other families, for example, Hsp40 and Hsp90, the high-molecular-weight proteases, and the small heat shock proteins (sHsp).

Table 1 Stress (heat shock) proteins and chaperones in prokaryotes

Family		Examples	
Name(s)	Mass (kDa)	Bacteria	Archaea
Heavy; high molecular weight; Hsp100	100 or higher	Clp	No ^a
Hsp90	81–99	HtpG	No
Hsp70; DnaK; chaperones	65–80	DnaK; Hsc66	Hsp70(DnaK) ^b (only some species)
Hsp60	55–64		L
Chaperonins group I		GroEL	GroEL (only some species)
Chaperonins group II		No	TF55; thermosome subunits (up to five)
Hsp40; DnaJ	35–54	DnaJ	Hsp40(DnaJ) (only some species)
Other	≤34		
small Hsp; sHsp		sHsp: Hsc20; Hsp15; alpha-crystallin family; other	sHsp: Hsp16.5; α-crystallin family; other
miscellanea		Hsp33; GroES; GrpE; PPIase; PDlase; lbpA/B; SecB; other	GroES (only some species); GrpE (only some species); PPIase; PDlase; prefoldin, 1 or 2 subunits; Sso7d

^aNo indicates not yet found, or not yet investigated in many species.

^bArchaeal species with Hsp70(DnaK) also have Hsp40(DnaJ) and GrpE; no species has yet been found having only one or two of these three chaperones.

Table 2 Stress (heat shock) proteins and chaperones in eukaryotes^a

Family		Examples			
Name(s)	Mass (kDa)	Cytosol	Mitochondrion	Endoplasmic reticulum ^b	Chloroplast
Heavy; high molecular weight; Hsp100	≥100	Hsp101 (Hsp102); Hsp104; Hsp105 alpha and beta; Hsp110	ClpXP/Hsp100	Hsp170 (Grp170)	
Hsp90	81–99	Hsp90(Hsp82; Hsp83; Hsp90 alpha and beta); Grp94	No	Hsp90 (GRP94; Grp94; endoplasmin; gp96))	Hsp93 (atHsp93-III; atHsp93-V)
Hsp70; chaperones	65–80	Hsp70(Hsp72); Hsc70 (Hsp73); SSA1-4; SSB1-2	mhsp70; SSC1; mt-Hsp70 (Grp75); Hsp78; Ssq1; Ecm10	BiP(Grp78); Erp72; Ssi1p; Kar2p; Lhs1p	Hsp70(DnaK) Hsp70B; Hsc70-2
Hsp60	55–64				
chaperonins group I		No	Cpn60 (mtGroEL; mtHsp60)	No	Cpn60 (Rubisco binding protein)
chaperonins group II		CCT (TRiC; TCP1 complex): up to 9 subunits	No	RBP	No
Hsp40; DnaJ	35–54	Hsp40; Hdj1(MAS5); Hdj2(DjA1; HSDJ); DjA4; Sis1p; Ydj1p	Hsp40; Mdj1p; Mdj2p; DjA3	Hsp40; Hsp47; Sec63p; Erdj3; Scj1p; JEM1p	Hsp40(DnaJ); ANJ1; TLP40; CDJ2
Other	≤34				
small Hsp; sHsp		alpha-crystallin family; Hsp10; Hsp20(Hsp26); Hsp22; Hsp27	Hsp10(Cpn10; mtGroES)	Cpn20	
miscellanea		PPIase; PDlase; prefoldin (Gim)(up to 6 subunits);	GrpE(mtGrpE; Yge1; Mye1p); PPIase; Atp12p; Hep1	Cpn20; PPIase; PDlase	GrpE; Hsp18; Hsp25; Hsp26; Cpn10; Cpn20; atAcid

^aChaperones build multimolecular assemblies to exercise their functions. These assemblies are formed by identical or different chaperone molecules (subunits) and by cofactors. In addition, these chaperone complexes interact with other chaperone complexes (networking) and with the protein-degradation machinery (e.g., the ubiquitin–proteasome system) to integrate the protein quality-control system of the cell. Synonyms are given in parentheses. No indicates not yet found, or not yet investigated in many species.

^bIn addition to the chaperones listed in the endoplasmic reticulum column, this organelle also has Grp58(Erp57) and lectins (calreticulin and calnexin, approximately 45 and 60 kDa, respectively) that assist, together with associated factors, the folding and oligomerization of many glycoproteins.

Biology

Chaperones and stress proteins can be viewed from an evolutionary-phylogenetic standpoint, while keeping in mind that the separation between prokaryotes and eukaryotes is not absolute, inasmuch as the eukaryotic cell contains components of prokaryotic ancestry. For instance, the mitochondrion is the remnant of an ancient alpha-proteobacterium that entered into a symbiotic association with a primitive eukaryotic cell. Similarly, the chloroplast descends from a primitive symbiotic cyanobacterium. Even the eukaryotic nucleus contains, in its DNA, genes that were derived from bacteria. In this regard, genes encoding chaperones of the Hsp60 and Hsp70 families are of particular interest. These chaperones are present in the mitochondria, chloroplasts, endoplasmic reticulum (ER), and other organelles (see [Table 2](#)), but their genes reside in the nucleus. These genes are transcribed in the nucleus, but the chaperones are synthesized in the cytosol and are finally translocated to the destination organelle. In fact, it has been the comparative analyses of Hsp70 and Hsp60 amino acid sequences that have convincingly shown the origin of the genes that encode these proteins and have provided strong support for the idea that the eukaryotic cell is an assemblage of diverse parts, several of which are prokaryotic in origin.

Phylogenetic Domains

Since the late 1970s, the idea that all living cells can be sorted into three main lines of evolutionary descent, or phylogenetic domains, has gained acceptance, although many points remain controversial and the boundaries between domains at the beginning of evolution are less well-defined than was initially believed. Nevertheless, it is helpful to think of stress proteins and chaperones within the context of the three domains, Bacteria (eubacteria), Archaea (archaeobacteria), and Eucarya (eukaryotes) – the former two being prokaryotes. Why is this helpful? Because archaea (i.e., the members of the domain Archaea) are indeed different from bacteria and eucarya or eukaryotes (i.e., the members of the domains Bacteria and Eucarya, respectively) in terms of the Hsp70 and Hsp60 chaperones. For example, bacteria possess the Hsp70 (also named DnaK), Hsp40 (also named DnaJ), and GrpE stress proteins that constitute the molecular chaperone machine (KJE) and the Hsp60–Hsp10 molecules that constitute the GroEL/S or chaperonin group I system. In contrast, many archaea do not have the chaperone KJE or GroEL/S, and eukaryotes have GroEL/S homologs only inside their organelles. It is noteworthy that many archaea do not have KJE because this

machine is present in all bacteria and eukaryotes without any known exception. Also remarkable is the presence of GroEL/S in some archaea because until recently it was thought that this chaperoning complex was exclusive to the bacteria and bacteria-related organelles.

The structures and functions of the molecular chaperone machine and of the chaperonin system type I are relatively well characterized in the bacterium *Escherichia coli*. There is also considerable information on the eukaryotic Hsp70 complex and on the chaperonin system type II. Intriguingly, some archaea either do not have an identifiable chaperone machine or have one of bacterial type with the Hsp70(DnaK), Hsp40(DnaJ), and GrpE components, but they all have a type II chaperonin system. The fact that archaea have a eukaryotic type of chaperonin system is remarkable because archaea, like bacteria, are prokaryotes.

Also interestingly, eukaryotes have the bacterial type of Hsp60 chaperonin system in their organelles, bearing witness to the bacterial origins of the latter. Thus, one of the most exciting fields of biology at present is the analysis of chaperonins and chaperones in organisms representing the three phylogenetic domains and domain branches to understand the evolution of these molecules and also to elucidate the relationships among species, and between organelles and their prokaryotic ancestors.

Another point of great interest, from both the evolutionary and the physiological standpoints, is that a number of archaeal species do not have Hsp70 (DnaK) or the other components of the chaperone machine, Hsp40(DnaJ) and GrpE. This peculiarity is remarkable because it represents the only exception known to the rule that every organism has an Hsp70 system (in fact, Hsp70 is considered one of the most highly conserved proteins in nature). The absence of the gene triad that encodes the components of the chaperone machine is noteworthy also because it raises the issue of how these organisms can achieve protein biogenesis and survive stress without the molecular chaperone machine.

Physiology

Prokaryotes, with few exceptions, have only a single *hsp70(dnaK)* gene, which is active under normal physiological conditions and is activated further by stressors. In the latter case, an increase in Hsp70 (DnaK) over the basal or constitutive levels is observed; it may be due to increased transcription and/or translation, and/or decreased mRNA and/or Hsp70(DnaK) degradation. In contrast, eukaryotes have more than one *hsp70(dnaK)* gene. Essentially, one gene is active physiologically and responds to

physiological signals (it is said to be constitutively induced), whereas the other gene is induced by stressors. The former gene is designated cognate *hsp70*, abbreviated *hsc70*, and the protein is termed Hsc70. In fact, eukaryotic cells may have more than one representative of the two types of the *hsp70* gene, and both types may respond to physiological stimuli, such as stimuli derived from cell-cycle and cell-differentiation events. Therefore, *hsp70(dnaK)* is also interesting from the point of view of cell physiology during development and histogenesis and is an important player in pathology and disease.

Pathology and Medicine

Molecular chaperone genes and proteins, as well as chaperonins and other stress genes and proteins, are being studied at the present time to determine their roles in the maintenance of health and in the development of disease. It is conceivable that the failure of the chaperoning processes due to a defective chaperone (chaperonopathy) will lead to the accumulation of misfolded proteins that do not function or to misplaced proteins that do not reach their destination in the cell to function properly. In both instances, the accumulation of nonfunctional proteins will encumber the cell and disturb its normal physiology. This is why protein degradation is also an important cellular process, part of which depends on molecular chaperones.

Chaperonopathies may be caused by genetic defects (e.g., chaperone-gene mutation) or by post-translational modifications of the chaperone molecule, such as those observed in the elderly as a manifestation of the aging process, which is characterized by a progressive accumulation of biochemical damage in many proteins.

Molecular chaperones may be exploited as tools for preventing protein denaturation or for restoring partially denatured proteins (e.g., during surgery) when ischaemia-hypoxia (a strong cell stressor) is unavoidable. The use of chaperones for prevention and treatment of pathological conditions is emerging as a new medical field: chaperonotherapy.

Mechanism

The mechanism of action of chaperones at the molecular level has not been fully elucidated for any organism or any particular family of chaperones, but there is considerable information on the Hsp70(DnaK) and Hsp60 (chaperonin) systems in *E. coli*. How do chaperones interact to help a polypeptide to acquire its final three-dimensional (native) configuration, based on its primary structure? It is known that chaperones

neither provide steric information nor form part of the final product. The Hsp70(DnaK) system in the cytosol is believed to bind nascent polypeptides so as to prevent their aggregation, either among one another or with other polypeptides in the immediate vicinity. In this manner, the Hsp70(DnaK) machine maintains the newly made polypeptide in a folding-competent status. In contrast, the Hsp60 (chaperonin) system GroEL/S acts posttranslationally and assists in the folding of polypeptides ranging in size from 25 to 55 kDa. Hence, without chaperones, most polypeptides would aggregate and the few that would proceed in the right direction toward achieving a physiologically active conformation would do so very slowly and inefficiently. Within this context, it is easy to realize that chaperones and chaperonins, and other stress proteins such as the sHsp, play a critical role in cell survival during stress, when the tendency toward protein denaturation and aggregation is strong.

The binding of polypeptides to, and their release from, Hsp70(DnaK) is ATP-dependent and involves the other members of the Hsp70(DnaK) machine, Hsp40(DnaJ), and a nucleotide-exchange factor (e.g., GrpE in prokaryotes). Hsp40 binds the polypeptide in need of assistance and presents it to Hsp70-ATP complex. As this complex binds the Hsp40-tagged polypeptide, Hsp40 catalyzes ATP hydrolysis, which increases the strength of the binding between Hsp70-ATP and the polypeptide. GrpE then comes into action and catalyzes the exchange of ATP for ADP on Hsp70. When Hsp70 is once again complexed with ATP, its affinity for the polypeptide diminishes and the latter is released, either because it has folded correctly or because it has to be processed further by the Hsp60 system GroEL/S.

The GroEL/S complex in *E. coli* is a cylindrical barrel. GroEL (the Hsp60 component) forms two heptameric rings with a central cavity. GroES (the Hsp10 subunit) forms a single heptameric ring and serves as a lid for the two-ring GroEL barrel. The polypeptide-folding mechanism is believed to be as follows. The unfolded polypeptide reaches the entrance to the GroEL barrel either unassisted or ushered by the chaperone machine, enters into the cavity, and binds to hydrophobic patches in the wall of the barrel. The GroES ring, together with ATP, then binds to the GroEL cylinder, thereby closing it. Thus, a hydrophilic folding chamber (also called cage) is created, and the polypeptide is in an environment in which it becomes soluble and can fold correctly. ATP binding to the other GroEL ring releases GroES, opening the barrel, and this allows the folded polypeptide to escape into the cytosol. The cycle is repeated if the folding process was incomplete.

The details just described have been inferred from experiments *in vitro* with *E. coli*. It is not known to what extent the *in vitro* findings reproduce the *in vivo* mechanisms. Also, the mechanism of Hsp70–Hsp60 chaperoning in locales other than the cytosol, for example, in the mitochondria, differs from the mechanism described earlier. Similarly, the mechanisms of action of the other families of chaperones are also different in many details.

We may conclude, from studies pertaining to the *E. coli* cytoplasm, that, whereas the Hsp70(DnaK) chaperone machine intervenes early in protein synthesis (during translation) to prevent aggregation of the majority of the nascent polypeptides, the Hsp60 chaperonin system acts later. The GroEL/S complex intervenes posttranslationally and only to assist that fraction of the new polypeptides that reach it either unassisted or ushered by the chaperone machine. The rest of the newly made polypeptides in *E. coli* do not use the GroEL/S chaperonin system, and they probably fold in the cytosol assisted only by the chaperone machine.

To Fold, Refold, or Degrade

The protein content and quality of a cell vary according to the cell's needs imposed by physiological (e.g., cell cycle or differentiation stage) and environmental changes (e.g., nutrient limitation, pH drop, or temperature or osmolarity upshift). The diverse array of proteins and the quantities of each of these must be adjusted precisely to the requirements of every moment of the cell's life. This delicate balance is achieved by various complementary, coordinated mechanisms that effectively either increase or decrease the concentration of any given protein in its native functional conformation in any given cell locale. Also, when proteins are damaged by stressors, they must be restored to normality (native as opposed to denatured or unfolded, or misfolded, status) or they must be eliminated if the damage is too severe and beyond repair.

Stress proteins, particularly those that are molecular chaperones, play a key role not only in protein biogenesis and in restoring partially damaged proteins to a functional pool, but they are also involved in eliminating molecules that cannot be recovered. For the latter purpose, stress proteins act in conjunction with proteases or they use their own proteolytic power (many stress proteins are proteases, and some are both chaperones and proteases). Because of this dual role – assistance in the generation and recovery of polypeptides, as well as in their destruction – stress proteins are essential tools by which the cell maintains its protein balance.

Proteins for degradation are either normal or abnormal. Many normal proteins have a short life span; they are naturally unstable, as required by their transient tasks inside the cell. Examples are transcription regulatory factors, termed activators, that need to be eliminated to downregulate the expression of certain genes. The same is true for some repressors, which have to be removed when the gene that they regulate must be upregulated or induced. Many abnormal (denatured) proteins are generated by stress, as noted earlier. In addition, abnormal proteins may appear in the cell in the absence of stress, as a result of genetic mutations or of transcriptional, translational, and/or processing errors.

Whatever the case, unfit or unnecessary proteins must be either eliminated or restored to a functional status; for each one, the cell must make the decision as to whether to try to fold, or refold, or to degrade. In the last alternative, the cell will not only be freed from potentially detrimental molecules or molecular debris, but it will at the same time acquire building blocks to synthesize new molecules. Proteins destined for degradation must be recognized by proteases; for this purpose, they have tags that are seen by the proteolytic enzymes. Some molecular chaperones provide the tag and present the protein for degradation to the protease.

In conclusion, molecular chaperones not only assist in the generation of protein molecules, but they also contribute to the elimination of proteins when they are potentially dangerous or no longer necessary (dispensable) to the cell.

See Also the Following Articles

Chaperonopathies; Heat Resistance; Heat Shock Genes, Human; Heat Shock Response, Overview; Proteases in Prokaryotes and Eukaryotic Cell Organelles; Proteases in the Eukaryotic Cell Cytosol; Viral Virulence and Stress.

Further Reading

- Deuerling, E. and Bukau, B. (2004). Chaperone-assisted folding of newly synthesized proteins in the cytosol. *Critical Reviews in Biochemistry and Molecular Biology* 39, 261–277.
- Ferrari, D. M. and Soeling, H-D. (1999). The protein disulphide-isomerase family: unraveling a string of folds. *Biochemical Journal* 339, 1–10.
- Hartl, F. U. and Martin, J. (1995). Molecular chaperones in cellular protein folding. *Current Opinion in Structural Biology* 5, 92–102.
- Jackson-Constan, D., Akita, A. and Keegstra, K. (2001). Molecular chaperones involved in chloroplast protein import. *Biochimica et Biophysica Acta: Molecular Cell Research* 1541, 102–103.

- Kovacheva, S., Bedard, J., Patel, R., et al. (2005). *In vivo* studies on the roles of Tic110, Tic40, and Hsp93 during chloroplast protein import. *The Plant Journal* **41**, 412–428.
- Lee, A. S. (2001). The glucose-regulated proteins: stress induction and clinical applications. *Trends in Biochemical Sciences* **26**, 504–510.
- Macario, A. J. L. and Conway de Macario, E. (2001). The molecular chaperone system and other anti-stress mechanisms in Archaea. *Frontiers in Bioscience* **6**, d262–283 [online].
- Macario, A. J. L. and Conway de Macario, E. (2004). The pathology of anti-stress mechanisms: a new frontier. *Stress* **7**, 243–249.
- Macario, A. J. L., Malz, M. and Conway de Macario, E. (2004). Evolution of assisted protein folding: the distribution of the main chaperoning systems within the phylogenetic domain Archaea. *Frontiers in Bioscience* **9**, 1318–1332 [online].
- Martin, J. (2004). Chaperonin function – effects of crowding and confinement. *Journal of Molecular Recognition* **17**, 465–472.
- Martin, J., Gruber, M. and Lupas, A. N. (2004). Coiled coils meet the chaperone world. *Trends in Biochemical Sciences* **29**, 455–458.
- Maruyama, T., Suzuki, R. and Furutani, M. (2004). Archaeal peptidyl prolyl *cis*-transisomerases (PPIases): update 2004. *Frontiers in Bioscience* **9**, 1680–1700. [online].
- Morano, K. A., Liu, P. C. C. and Thiele, D. J. (1998). Protein chaperones and the heat shock response in *Saccharomyces cerevisiae*. *Current Opinion in Microbiology* **1**, 197–203.
- Riggs, D. L., Cox, M. B., Cheung-Flynn, J., et al. (2004). Functional specificity of co-chaperone interactions with Hsp90 client proteins. *Critical Review in Biochemistry and Molecular Biology* **39**, 279–295.
- Scharf, K-D., Siddique, M. and Vierling, E. (2001). The expanding family of *Arabidopsis thaliana* small heat stress proteins and a new family of proteins containing alpha-crystallin domains (Acd proteins). *Cell Stress & Chaperones* **6**, 225–237.
- Schlichter, T. and Soll, J. (1997). Chloroplastic isoforms of DnaJ and GrpE in pea. *Plant Molecular Biology* **33**, 181–185.
- Spiess, C., Meyer, A. S., Reissmann, S., et al. (2004). Mechanism of the eukaryotic chaperonin: Protein folding in the chamber of secrets. *Trends in Cell Biology* **14**, 598–604.
- Townsend, P. A., Stephanou, A., Packham, G. and Latchman, D. S. (2005). BAG-1: A multifunctional pro-survival molecule. *International Journal of Biochemistry and Cell Biology* **37**, 251–259.
- Wimmer, B., Lottspeich, F., van der Klei, I., et al. (1997). The glyoxysomal and plastid molecular chaperones (70-kDa heat shock protein) of watermelon cotyledons are encoded by a single gene. *Proceedings of the National Academy of Sciences USA* **94**, 13624–13629.
- Zhao, Q., Wang, J., Levichkin, I. V., et al. (2002). A mitochondrial specific stress response in mammalian cells. *EMBO Journal* **21**, 4411–4419.

Chaperonopathies

A J L Macario and E Conway de Macario
New York State Department of Health and The
University of Albany (SUNY), Albany, NY, USA

Published by Elsevier Inc.

Definitions and Classifications
Molecular Pathology
Impact of Defective Chaperones
Examples of Chaperonopathies

Glossary

Acquired chaperonopathy
A chaperonopathy that is due to a structural defect in a molecular chaperone acquired during the life of the individual and that is not passed on to the offspring, i.e., is not hereditary.

Chaperone complex

A team of two or more chaperones, cochaperones, and cofactors assembled to carry out the chaperoning process.

Chaperone networking

The interaction of a chaperone complex with another, distinct chaperone complex, or with a protein-degrading complex.

Chaperonopathy
Client polypeptide or substrate

A pathological condition in which a molecular chaperone is abnormal.

A polypeptide recognized by a chaperone or chaperone complex as needing assistance for folding or refolding, or as destined for degradation.

Dysregulatory chaperonopathy
Genetic chaperonopathy

Increase or decrease in the levels of a molecular chaperone caused by up- or down-regulation, respectively, of its parent gene.
A chaperonopathy characterized by a structural defect in a molecular chaperone due to a genetic (hereditary) cause, such as a gene mutation.

<i>Genetic polymorphism</i>	The occurrence, in a population of organisms of the same species, of more than one allele or genetic marker at the same locus. A polymorphism is distinguished by the fact that the least frequent allele or marker occurs more often than if it were due only to mutation.
<i>Primary chaperonopathy</i>	A condition in which the molecular chaperone pathology is not the consequence of another pathological disorder.
<i>Proteinopathy</i>	A pathological condition characterized by the occurrence of an abnormal protein molecule, usually with a tendency to misfold, aggregate, and precipitate.
<i>Secondary chaperonopathy</i>	A condition in which a molecular chaperone pathology is associated with another disorder and is the consequence of it, rather than its cause.

Definitions and Classifications

A chaperonopathy can be defined as a pathological entity characterized by a defect in a chaperone or co-chaperone. In structural chaperonopathies, the structure of the chaperone molecule is altered because its

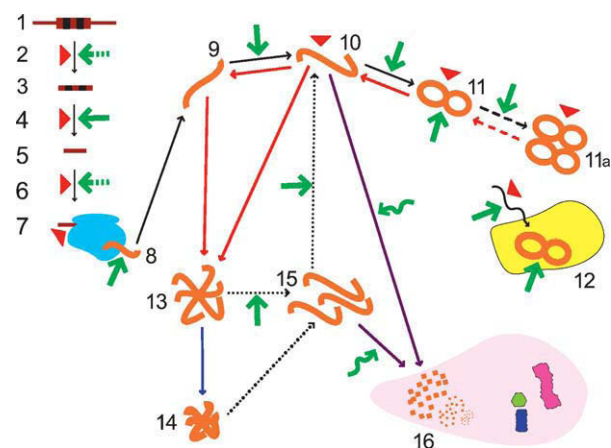


Figure 1 Schematic (not to scale) of the life of a protein, from parent gene to fully folded molecule to degradation, in a eukaryotic cell. Crucial steps that require chaperone assistance, and sites of stress impact whose adverse effects might enhance senescence and/or disease development, are noted. 1, Protein-coding gene (brown rectangle) with two introns (black). 2, Transcription. 3, Primary transcript (pre-mRNA), still with the two introns. 4, Pre-mRNA processing, including intron removal and exon splicing. 5, Mature mRNA (brown rectangle). 6, mRNA transport from nucleus to cytosol via nuclear pore complex. 7, mRNA translation on the ribosome (light blue). 8, Nascent polypeptide (orange, thick wavy line) emerging from the ribosome, when folding begins, assisted by ribosome-associate chaperones. 9, Newly made polypeptide released by the ribosome on its way to posttranslational folding. 10, The same polypeptide farther along the folding pathway. 11, New protein, correctly folded into its tertiary structure (native, functional conformation); some proteins achieve a functional status by forming oligomers (quaternary structure), 11a. 12, Some

parent gene is mutated, or because of another mechanism acting beyond the DNA, namely, during or after gene transcription. An example of the latter post-DNA mechanisms resulting in a structural defect in the chaperone molecule is an aberrant posttranslational modification.

Other chaperonopathies do not involve a structural modification of the chaperone molecule itself, but rather are characterized by quantitative (increase or decrease) and spatial (distribution in cells and tissues) changes. Levels of chaperones may be elevated or

proteins are translocated into organelles (e.g., mitochondrion; yellow) through the organellar membranes. 13, Aggregate, formed by unfolded or misfolded polypeptides. 14, Precipitate of aggregated polypeptides (protein deposit or inclusion body). 15, Unfolded or partially folded polypeptides freed from the aggregates by the action of aggregate-dissolving mechanisms involving chaperones. 16, Protein degradation products, depicted inside a shaded area that symbolizes a dynamic graveyard, in which degradation proceeds from oligopeptides of decreasing lengths down to individual amino acids, which will ultimately be recycled as building blocks for new proteins, in step 8. Some of the known proteolytic machineries found in prokaryotic and eukaryotic cells are depicted: the proteasome-ubiquitin system (magenta), tricorn protease (light green), and other proteases (dark blue). Also, the proteasome (immunoproteasome) generates short antigenic peptides by cleaving large protein antigens for use by antigen-presenting cells. Black solid arrows: physiological route from gene to functional protein. Black wavy arrow: translocation pathway from cytosol into an organelle, for example a mitochondrion. Green solid arrows: critical steps that require chaperone assistance. Green broken arrows: critical steps for which chaperone assistance is probably needed but not yet fully documented. Red arrowheads: molecule (or structure, function, pathway) adversely affected by stress. Red arrows: route from unfolded, partially folded, or misfolded proteins toward aggregation; this route is followed by abnormal proteins even in the absence of stress, and by normal and abnormal proteins that have been denatured by stress. Blue solid arrow: route from protein aggregation toward precipitation of aggregates (inclusion body formation). Violet arrows: routes to protein degradation followed by normal proteins in physiological cellular processes (e.g., in gene regulation, via activator or repressor degradation) and by abnormal proteins that cannot be folded or rescued (refolded) by chaperones (there should also be a similar arrow from 9 to 16, which was omitted for the sake of clarity). Green wavy arrows: participation of chaperones in protein degradation pathways. Black dotted arrows: recovery route from aggregate to free polypeptides to refolding; possibly this route is very actively utilized in proteinopathies, as the chaperoning systems attempt to rid the organism of pathologic aggregates, and to avoid protein deposition. A similar course of events may occur during senescence, with the caveat that the required chaperoning systems probably become less and less efficient with age. Black discontinuous (squares alternating with solid circles) arrow from 14 to 15: possible but improbable recovery route from protein deposit toward the release of free, re-foldable polypeptides. Note that the need for chaperone assistance and the stress-sensitive sites are widespread. Modified from Macario, A. J. L. and Conway de Macario, E. (2002). Sick chaperones and ageing: a perspective. *Ageing Research Reviews* 1, 295–311, with permission.

Table 1 Genetic chaperonopathies and human chaperone gene polymorphisms^a

<i>Chaperone Hsp gene/protein affected</i>	<i>Disease/syndrome</i>
sHsp Hsp27	Williams Charcot-Marie-Tooth Distal hereditary motor neuropathy
Alpha crystallins	Childhood cataracts
Alpha-B-crystallin	Desmin-related myopathy
Hsp22	Distal motor neuropathy
Chaperone cofactors for microtubule biogenesis	
Cofactor C (RP2)	X-linked retinitis pigmentosa
Cofactor E	Sanjad-Sakati and Kenny-Caffey Progressive motor neuronopathy
Chaperonin of group I, Hsp60	
Mitochondrial Hsp60 (Cpn60)	Hereditary spastic paraplegia
Chaperonin of group II, CCT subunits	
Alpha subunit	McKusick-Kaufman and Bardet-Biedl
Delta subunit	Hereditary sensory neuropathy
Peptidyl-prolyl <i>cis-trans</i> isomerase (FKBP)	Leber congenital amaurosis
<i>Genetic polymorphism</i>	<i>Abnormality/disease</i>
<i>hsp70-1</i> promoter region, allele (A)-110	Does not favor longevity in women
<i>hsp70-Hom</i> , T247C, threonine instead of methionine at position 493	Does not favor longevity
<i>hsp70-1</i> , promoter region, allele (C)-110	Associates with Parkinson's disease in Taiwanese

^aFor explanations and references, see Macario, A. J. L., Grippo, T. M. and Conway de Macario, E. (2005). Genetic disorders involving molecular-chaperone genes: a perspective. *Genetics in Medicine* 7, 3–12; and Macario, A. J. L. and Conway de Macario, E. (2005). Sick chaperones, cellular stress, and disease. *New England Journal of Medicine* 353, 1489–1501.

Abbreviations: CCT, chaperonin containing TCP-1 (see footnote to Table 3); FKBP, FK506-binding protein (FK506 is a compound that affects the immune system so that immune responses are suppressed).

Reproduced from Macario, A. J. L. and Conway de Macario, E. (2005). Sick chaperones, cellular stress, and disease. *New England Journal of Medicine* 353, 1489–1501, © Copyright 2005 Massachusetts Medical Society. All rights reserved.

decreased in cells in certain diseases compared with levels in the same cells from healthy individuals. Likewise, the distribution pattern of a chaperone inside a cell (e.g., within the cytosol, or within organelles), or in various cell types and tissues, may differ from that typical of healthy counterparts.

Quantitative alterations may be caused by up- or downregulation of the chaperone gene or by other posttranscriptional mechanisms such as increased (or decreased) stability of mRNA and/or protein or translation rate.

In the case of chaperones whose parent genes are stress inducible, i.e., genes that respond to stressors and are active in the stress response, quantitative changes may be due to up- or downregulation of stress gene regulators such as the HSF (heat shock factor) genes in eukaryotes. If a chaperone gene is up- or downregulated by its own regulatory mechanisms, the resulting pathologic entity will be a dysregulatory chaperonopathy.

From another standpoint, chaperonopathies may be characterized as primary or secondary. The former are due to defects inherent in the chaperone molecule, defects that have an impact on cells and tissues, and thus becomes a central part of pathogenesis. Examples of primary chaperonopathies are those due to inherited defects in the chaperone-encoding genes.

Secondary chaperonopathies are the consequence of a disease rather than the reverse; the defective chaperone is not a primary factor in pathogenesis, but it may contribute to it. Among these are chaperonopathies associated with the process of aging and chaperonopathies caused by stressors that damage the chaperone molecule. Likewise, some quantitative chaperonopathies may be classified as secondary.

Another cause of chaperone failure may reside not in the chaperone itself, but in the client polypeptide. At least two causes have been recognized. (1) The substrate is genetically defective and lacks the segment of sequence that would be recognized by the chaperone machine to proceed with folding; consequently, the client polypeptide remains unfolded and malfunctioning. (2) The substrate does interact with the chaperone, but because the substrate has certain biochemical-biophysical characteristics, this interaction inhibits rather than initiates the chaperoning process that leads to folding.

Similarly, chaperonopathies may be classified as genetic or acquired. The former are due to a genetic defect, e.g., mutation in the chaperone-encoding gene itself or in the gene encoding a chaperone gene regulator (e.g., the HSF gene). Polymorphisms in these genes may also account for some chaperonopathies.

A genetic chaperonopathy can be described at any of several levels: (1) a locus, gene, or exon that is altered, causing the absence of a chaperone molecule, or its structural abnormality, leading to malfunction; (2) a gene product, from primary transcript to protein (pre-mRNA, or mRNA, or protein, Figure 1) can be structurally and functionally abnormal; (3) a chaperone complex or chaperone network that is structurally

Table 2 Examples of chaperonopathies associated with aging^a

<i>Hsp/chaperone/HSF</i>	<i>Experimental-clinical observation</i>	<i>Conclusion</i>
Alpha-A-crystallin	Low levels and PTM and truncation in aged retina	Senescence-associated quantitative and qualitative changes
Alpha crystallins	Decreased solubility and PTM in the eye lens	Senescence-associated qualitative changes
Hsp70 and Hsc70	Aged cultured hepatocytes, melanocytes, and skin fibroblasts, and fibroblasts from old humans: low response to heat shock	Senescence-associated quantitative and functional changes
Hsc70	Low levels in retinas from aged humans and monkeys	Same as above
Hsp27, Hsp70, Hsc70,	High basal levels in aged, cultured, RMHS-treated fibroblasts from humans	Quantitative and functional changes due to aging and chronic stress
Hsp90	Decreased with RMHS and aging in human skin fibroblasts	Same as above
Hsp90	Low levels and chaperone capacity in the cytosol of aged hepatocytes	Functional changes associated with aging
HSF-1	Low levels in aged fibroblasts in culture and from the elderly	Senescence-associated quantitative and functional changes

^aFor source of data, explanations, and references, see Macario, A. J. L. and Conway de Macario, E. (2005). Sick chaperones, cellular stress, and disease. *New England Journal of Medicine* **353**, 1489–1501.

Abbreviations: PTM, posttranslational modification; RMHS, repeated mild heat shock; HSF, heat shock factor (a regulator of transcription of heat shock genes in eukaryotes).

Table 3 Examples of disease-associated chaperonopathies^a

<i>Hsp/chaperone</i>	<i>Experimental-clinical observation</i>	<i>Disease</i>
Alpha-B-crystallin	High in IBM muscle fibers	IBM
Alpha-B-crystallin	High in glial inclusions	Tauopathies (CBD, PSP, FTDP-17)
Hsp25	Decreased in motoneurons	ALS (mouse model)
Alpha-B-crystallin, Hdj1, Hdj2, Hsp70, alpha and beta SGT proteins	Low in brain	Huntington's
Hsp72	High in plaques and tangles, and in brain cortex	Alzheimer's
Grp78	High in neurons that are otherwise normal	Alzheimer's
Hsp60, Hsp70RY, Hsc70, alpha-B-crystallin, Grp75, Grp94	Altered distribution pattern in several regions of the brain	Alzheimer's

^aFor source of data, explanations, and references, see Macario, A. J. L. and Conway de Macario, E. (2005). Sick chaperones, cellular stress, and disease. *New England Journal of Medicine* **353**, 1489–1501.

Abbreviations: IBM, inclusion body myositis; CBD, corticobasal degeneration; PSP, progressive supranuclear palsy; FTDP-17, fronto-temporal dementia with parkinsonism linked to chromosome 17; ALS, amyotrophic lateral sclerosis; SGT, small glutamine-rich tetra-trycptide repeat containing (protein).

and functionally deficient due to a defect in one of the components, e.g., a chaperone or cochaperone molecule; (4) a cell, tissue, or organ abnormality that is the consequence of the chaperone deficiency; and (5) developmental and clinical manifestations, i.e., the phenotype.

Acquired chaperonopathies are those that are not heritable; they feature defects, quantitative or structural, in the chaperone molecule due to posttranscriptional aberrations. Acquired chaperonopathies can be described at various levels, as for the genetic chaperonopathies, except that in the acquired conditions the DNA (locus, gene, exon) does not carry hereditary defects.

The chaperones discussed in this article are proteins; hence, the chaperonopathies can be considered a subset of the proteinopathies. In these conditions, a

protein is structurally and functionally altered, and it usually has a tendency to misfold, aggregate, and precipitate. Abnormal chaperones can also follow this fate.

Molecular Pathology

Structural defects, genetic or acquired, in a chaperone molecule may affect one or more of the chaperone's functional domains and may have an impact on one or more of its functions.

A chaperone molecule must be able to recognize the client polypeptide in need of assistance for folding or refolding, and it must also be able to interact with the other members of the chaperone complex (the actual chaperoning machine). Furthermore, a chaperone must be able to interact with other chaperoning

machines and with protein-degrading machines (to build the chaperone networks involved in protein quality control). Interaction with protein-degrading machines is necessary when the client polypeptide cannot be folded properly and has to be eliminated. Chaperone complex formation and networking involve distinct segments (molecular domains) of the chaperone molecule. If one of these segments is altered because of a mutation or a posttranslational modification, the particular interaction in which the affected segment is involved will not occur, will be ineffective, or will proceed at a slower than normal rate. This perturbation will affect the functioning of the entire molecule, and also the structure and function of the whole chaperone complex and network of which the affected molecule is part. As the chaperone complex is affected, so will all of those cellular processes that require the complex; pathology will ensue.

Impact of Defective Chaperones

Figure 1 shows the life history of a typical protein molecule, from the production of the message (transcription of its parent gene) encoding the amino acid sequence to its final functional form. The figure also shows the points of impact of cell stressors along the way from transcription to the final native conformation, including migration and translocation to the cell locale in which the mature protein molecule resides and functions, and the consequences of misfolding, aggregation, and deficient structure, which lead to degradation.

Deficient chaperones can cause pathological disorders of various severities in any cell, tissue, or organism, whether animal or plant. They can also affect prokaryotic cells, rendering them vulnerable to environmental changes and other agents.

Examples of Chaperonopathies

Examples of genetic, structural chaperonopathies and of polymorphisms of human chaperone genes are listed in **Table 1**. Chaperonopathies associated with senescence and with disease are listed in **Tables 2** and **3**, respectively.

See Also the Following Articles

Chaperone Proteins and Chaperonopathies; Heat Resistance; Heat Shock Genes, Human; Heat Shock Response, Overview; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles; Viral Virulence and Stress.

Further Reading

- Bloemendal, H., de Jong, W., Jaenicke, R., et al. (2004). Aging and vision: structure, stability and function of lens crystallins. *Progress in Biophysics and Molecular Biology* 86, 407–485.
- Chapple, J. P., Grayson, C., Hardcastle, A. J., et al. (2001). Unfolding retinal dystrophies: a role for molecular chaperones? *Trends in Molecular Medicine* 7, 414–421.
- Cloos, P. A. C. and Christgau, S. (2004). Post-translational modifications of proteins: Implications for aging, antigen recognition, and autoimmunity. *Biogerontology* 5, 139–158.
- Csermely, P. (2001). Chaperone overload is a possible contributor to ‘civilization disease.’ *Trends in Genetics* 17, 701–704.
- Jackson, P. K. (2004). Post-translational control and ubiquitination. In: Epstein, C. J., Erickson, R. P. & Wynshaw-Boris, A. (eds.) *Inborn errors of development*, pp. 804–814. Oxford, UK: Oxford University Press.
- Macario, A. J. L. and Conway de Macario, E. (2001). Molecular chaperones and age-related degenerative disorders. *Advances in Cell Ageing Gerontology* 7, 131–162.
- Macario, A. J. L. and Conway de Macario, E. (2002). Sick chaperones and ageing: a perspective. *Ageing Research Reviews* 1, 295–311.
- Macario, A. J. L. and Conway de Macario, E. (2004). The pathology of cellular anti-stress mechanisms: a new frontier. *Stress* 7, 243–249.
- Macario, A. J. L. and Conway de Macario, E. (2005). Sick chaperones, cellular stress, and disease. *New England Journal of Medicine* 353, 1489–1501.
- Macario, A. J. L., Grippo, T. M. and Conway de Macario, E. (2005). Genetic disorders involving molecular-chaperone genes: a perspective. *Genetics in Medicine* 7, 3–12.
- Nardai, G., Csermely, P. and Soti, C. (2002). Chaperone function and chaperone overload in the aged: a preliminary analysis. *Experimental Gerontology* 37, 1257–1262.
- Slavotinek, A. M. and Biesecker, L. G. (2001). Unfolding the role of chaperones and chaperonins in human disease. *Trends in Genetics* 17, 528–535.
- Soti, C. and Csermely, P. (2003). Aging and molecular chaperones. *Experimental Gerontology* 38, 1037–1040.
- Verbeke, P., Fonager, J., Clark, B. F. C. and Rattan, S. I. S. (2001). Heat shock response and aging: mechanisms and applications. *Cell Biology International* 25, 845–857.

Chemical Warfare

J Berberich

Agentase, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

Use of Chemical Weapons

Classification of Chemical Weapons

Vesicants

Nerve Agents

Miscellaneous Lesser Chemical Warfare Agents

Glossary

<i>Acetylcholinesterase</i>	An enzyme found in the blood and at nerve endings that catalyzes the hydrolysis of acetylcholine.
<i>Blister agent</i>	A chemical compound that burns the eyes, lungs, and skin and typically forms blisters on the skin.
<i>Blood agent</i>	A chemical, typically containing a cyanide group, that prevents the normal utilization of oxygen by tissue.
<i>Chemical warfare</i>	Intentional use of chemicals, typically poisons, contaminants, or irritants, in order to kill, injure, or incapacitate.
<i>Choking agent</i>	A chemical agent that prevents normal breathing and causes a victim to suffocate.
<i>Lachrymation</i>	Excessive secretion of tears.
<i>Miosis</i>	Contraction of the pupil of the eye in response to increased light or to certain chemicals or pathological conditions.
<i>Nerve agent</i>	A chemical agent that inhibits the enzyme acetylcholinesterase and prevents the transmission of nerve impulses.
<i>Toxic industrial chemical</i>	Chemicals from industrial processes that may be hazardous to individuals.
<i>Vesicant</i>	A chemical compound that burns the eyes, lungs, and skin and typically forms blisters on the skin.

Use of Chemical Weapons

World War I was the beginning of large-scale chemical warfare. The Germans first used chlorine and phosgene gas released from gas cylinders and artillery shells; 85% of all deaths resulting from exposure to chemical warfare agents during World War I were caused by phosgene. The first use of sulfur mustard was also by the Germans during World War I. Sulfur mustard was such an effective chemical warfare agent

that it became known as the king of the battle gases. The Germans began designing a second generation of chemical weapons, the nerve agents, before World War II. Although they were never used, the Germans maintained large stockpiles of the nerve agents tabun and sarin. After World War II, the British synthesized the very potent nerve agent VX.

Even after World War II, chemical warfare agents were still in use. During the 1960s, mustard agents, riot control agents, and perhaps nerve agents were used by Egypt during the Yemen War. Iraq used sulfur mustard, tabun, and soman during the war with Iran, and it is believed that they may have used cyanide against the Kurds. There is some speculation that Iraq used chemical weapons during the first Persian Gulf War (this being the cause of Gulf War Syndrome). Many veterans of this war complain of a number of symptoms, including recurrent headaches, joint stiffness, nausea, anxiety, and depression. No conclusion has been drawn on the cause of Gulf War Syndrome, but many conclude that multiple chemical exposure, contact with pesticides, negative reactions to vaccines or side effects of chemicals given to protect against nerve agent exposure may have caused some symptoms.

Terrorist groups are now the biggest threat for using chemical weapons. Sarin was used by the Aum Shinrikyo cult in Japan in the mid-1990s. It was first used in Matsumoto in 1994 and killed seven and injured over 600. The second occurrence was in the Tokyo subway in 1995, which killed 12 people and injured more than 5000. This cult had also used VX for the assassination of important political leaders.

Chemical warfare presents a number of unique physical and psychological stresses to an individual. Military personnel not only risk the physical stress and damage caused by contact with chemical warfare agents, but also cope with the severe psychological stress that comes with the fear of risk of exposure. Wartime already presents significant fear and confusion for a soldier, and the threat of chemical warfare exacerbates this. The first responders who must respond to chemical warfare incidences also experience significant stress caused by the need to wear cumbersome protective gear, the potential accidental exposure to the chemical warfare agent, and the obligation to help the injured, dying, or dead.

For those living with the dread of war or terrorism, the stress, fear and worries can lead to severe anxiety and have a serious impact on normal life. People's mood, cognition, and behavior may be disrupted as

a result of the fear, uncertainty, and panic that may accompany a chemical warfare incident, a training drill, or a hoax. In fact, the psychological stress itself may mimic many of the physiological effects caused by exposure to chemical weapons.

Classification of Chemical Weapons

A variety of chemicals exist that may be used as chemical warfare agents. Those considered to be the biggest threat for the modern soldier are the nerve gases and mustard gases. These agents are actually liquids at ambient temperatures and are generally aerosolized in order to be efficiently dispersed. The low volatility of these agents allows them to persist in the environment for a long period of time. It is the persistence of these agents that poses a great risk for first responders and health-care workers. Other toxic chemicals such as cyanide, chlorine, and phosgene are true gases at ambient conditions and rapidly dissipate from the environment, making them less of a long-term hazard. However, some of these gases are heavier than air and may remain in low-lying areas for extended periods.

Chemical warfare agents may also include riot control agents, such as tear gas and pepper sprays, and even herbicides. Agent Orange was a herbicide used during the Vietnam War to remove foliage and expose hidden enemy soldiers. With the growing fears of terrorism, two of the largest classes of chemicals of concern are the toxic industrial chemicals (TICs) and toxic industrial materials (TIMs). If a terrorist organization were to release any of these chemicals into the environment, the impact would be devastating.

Vesicants

The vesicants are a broad class of chemical agents that produce vesicles or blisters on the skin and the mucous membranes. These chemicals include sulfur mustard, the nitrogen mustards, Lewisite, and phosgene oxime.

Sulfur mustard was developed by Germany before World War I and was noted to smell like mustard and taste like garlic. Pure sulfur mustard is known as HD; it is a strong alkylating agent, known for its ability to react with a number of biomolecules, including proteins, DNA, and RNA. The organs most commonly affected by mustards are the skin, eyes, and those in the respiratory system. Sulfur mustard does not have an effect at first contact, and a dose-dependent latent period exists before effects begin to occur. Nausea and burning of the eyes are often the first

symptoms of mustard exposure. Blistering and itching of the skin, chest pains, coughing, lesions, or blindness may begin to occur within a few hours or days after exposure. There is a poor understanding of exactly how mustard produces blisters. Most symptoms begin to heal within few weeks of exposure; however, neurosis, depression, and continuous eye problems have been noted as long-term effects. Of the soldiers exposed to mustard gas during World War I, only 3–4% died of its direct effects. Death typically occurs due to secondary infections, shock, or respiratory insufficiency. There also exists some data that suggest that chronic exposure to mustard may increase the likelihood of developing cancer of the mouth, throat, and respiratory tract.

The nitrogen mustards, referred to as HN1, HN2, and HN3, are also potent vesicants. Although the nitrogen mustards are structurally similar to the sulfur mustards and have common chemical reactions, they have never found use as weapons. They produce damage to the eyes, respiratory tract, and skin similar to sulfur mustard; however, onset may take place more rapidly. Nitrogen mustard (HN2) was manufactured and used for a number of years in chemotherapy.

Lewisite was developed during World War I to be less persistent than mustard since waging an attack on mustard-containing areas was difficult. Lewisite was never used in war. The toxic effects of Lewisite are similar to those of sulfur mustard; however, the onset of effects occurs within seconds of exposure. Like other vesicants, Lewisite is a potent inhibitor of enzymes, inhibiting via the interaction of the Lewisite arsenic group with enzyme cysteine residues. Like sulfur mustard, the exact mechanism of toxicity is poorly understood.

Nerve Agents

The nerve agents are a class of organophosphorous esters, similar to pesticides, that are rapidly lethal via any route of exposure. There are two major classes of nerve agents: V and G. The G series agents include tabun (GA), sarin (GB), soman (GD), and GF. The V series agents include VX and Russian VX. The relative toxicity of the nerve agents is $VX > GD \sim GF > GB > GA$. The nerve agents are about 1000 times more potent than the most potent commercially available pesticides. VX is especially potent since it has a low volatility, making it a serious contact threat, and it is stable in the environment. The G series agents are more volatile and do not pose as much of a contact hazard.

The nerve agents act by irreversibly binding to esterase enzymes, especially acetylcholine esterase, which is present in the blood and at the neuromuscular junctions in the peripheral and central nervous systems. The binding of nerve agents to acetylcholine esterase causes inhibition of the enzyme and leads to an accumulation of acetylcholine. The clinical effects of nerve agents are due to the build-up of acetylcholine within the nervous system, which, depending on dose and route of exposure, may lead to a variety of symptoms, including miosis, nausea, tightness in the chest, vomiting, increased sweating and salivation, lachrymation, headaches, weakness, twitching, muscular and abdominal cramps, diarrhea, blurring of vision, anxiety, and tremors. Death can occur through respiratory arrest or cardiac arrhythmia. Severe poisoning by nerve agents may cause long-term physiological, psychological, and neurological effects such as cerebral impairment and cardiac malfunctioning as well as posttraumatic stress disorder and differences in intellectual and motor skills. Exposure to small concentrations of nerve agents may cause depression, forgetfulness, inability to concentrate, bad dreams, insomnia, and irritability.

Miscellaneous Lesser Chemical Warfare Agents

The cyanides, often referred to as blood agents, were used as chemical weapons during World War I by the French and Austrians. Cyanides have also reportedly been used during the Iran-Iraq War. The cyanide gases are difficult to deliver since they rapidly disperse into the environment due to their high volatility. There are three main types of cyanide: hydrogen cyanide (AC), cyanogen chloride (CK), and cyanogen bromide. The cyanogen halides (cyanogen bromide and cyanogen chloride) react in the body to produce hydrogen cyanide and hence have a very similar toxic effect. The target of cyanide toxicity is cellular respiration in the mitochondria. The cyanide ion reversibly inhibits cytochrome oxidase, which is the terminal oxidase in the mitochondrial respiratory chain. Inhibition leads to tissue asphyxia, resulting in central respiratory arrest and death. In addition, cyanogen chloride and cyanogen bromide act as irritants and may cause ocular and respiratory tract irritation. Recovery from cyanide poisoning, if it is not fatal, is typically rapid.

Phosgene and chlorine comprise a class of chemical gases called pulmonary agents (also referred to as choking agents) since they cause severe lung irritation that leads to fluid accumulation in the lungs (pulmonary edema), which can choke the victim. Phosgene

injury is caused by hydrochloric acid that is formed in the lungs and by diamide formation that cross-links cellular components. Chlorine (often referred to as CL) and phosgene (designated CG) were used during World War I. Phosgene accounted for 85% of all World War I deaths caused by chemical weapons. The risk of use of chlorine and phosgene remains high since they are used extensively in industry, making them readily available for use by terrorist groups.

Riot control agents, most commonly referred to as tear gases, are nonlethal compounds used to temporarily disable individuals by irritating exposed mucous membranes and skin. Most tear gases are effective within seconds to minutes, but their duration of effect is brief once an individual has been removed from exposure. Common riot control agents include chloroacetophenone (CN), the active ingredient in Chemical Mace, and *o*-chlorobenzilidene malonitrile (CS). These chemicals are ideal for use as riot control agents since the dose required to be effective is significantly lower than the lethal dose. These agents are typically aerosolized or released as smoke, and they cause burning irritation in the eyes, lachrymation (tear production), chest tightness, cough, sneezing, and a burning sensation of the skin. Oleoresin capsicum (OC), commonly referred to as pepper spray, is also considered a riot control agent.

Incapacitating agents are anticholinergic compounds that affect the central nervous system. These include the compound 3-quinuclidinyl benzilate (BZ) and also lysergic acid diethylamide (LSD). Those exposed to incapacitating agents suffer from drowsiness, poor coordination, slowed thought processes, and delirium. Hallucination may also be experienced. These effects can last from a few hours to a number of days. The Bosnian Serbs were alleged to have used BZ against civilians.

Biological toxins are sometimes referred to as chemical weapons since they are not living organisms. Many biological toxins mimic a number of chemical weapons. For example, the effects of capsaicin (from the oil of hot peppers) mimic those of artificially manufactured lachrymators. Other toxins include botulinum toxin, ricin, and mycotoxins. Botulinum toxin may cause paralysis and death and is about 15 000 times more potent than VX. Ricin is a plant toxin that is easily produced from castor beans and inhibits protein synthesis in the ribosome. Botulinum toxin and ricin have been used to perform assassinations. Tricothecene mycotoxins are produced by molds. Exposure to mycotoxins resembles exposure to vesicants. Soviet-supported forces have been alleged to use tricothecene mycotoxins in aerosolized droplets

against Afghanistan, Laos, and Kampuchea. They may have also been used during the Iran-Iraq war.

Although not commonly referred to as a chemical warfare agent (since it was not specifically used to cause direct harm to individuals), the herbicide Agent Orange was used extensively in Vietnam to remove foliage and expose enemy soldiers. Agent Orange is a mixture containing the phenoxyacetates 2,4-D and 2,4,5-T; some preparations contained tetrachlorodibenzodioxin (TCDD). TCDD is a dioxin that may be a carcinogen and a teratogen. Following the Vietnam War, veterans began to report a number of health problems, including skin rashes, cancer, psychological symptoms, birth defects, and handicaps in their children. Agent Orange may affect the nervous system causing psychiatric illnesses and problems with the nerves responsible for movement and sensation.

Finally, there is a growing fear that terrorist organizations will move away from the more traditional chemical warfare agents and begin using so-called agents of opportunity. These include a number of common and not-so-common industrial chemicals, TICs and TIMs. Examples include chlorine, ammonia, formaldehyde, fluorine, and sulfur dioxide. These are chemicals that are commonly used in industrial processes, relatively available, easy to disperse, and significantly toxic. Release from a chemical plant could produce a large plume, threatening nearby communities. Also, a truck carrying these chemicals could easily be hijacked and its contents released in a highly populated area. It is this ease of availability

that potentially makes TICs and TIMs more ideal than traditional chemical warfare agents for use as a terrorist's weapon.

See Also the Following Articles

Gulf War Syndrome, Psychological and Chemical Stressors; Persian Gulf War, Stress Effects of; Terrorism.

Further Reading

- Brown, M. A. and Brix, K. A. (1998). Review of health consequences from high-, intermediate- and low-level exposure to organophosphorous nerve agents. *Journal of Applied Toxicology* 18, 393–408.
- Gunderson, C. H., Lehmann, C. R., Sidell, F. R. and Jabbari, B. (1992). Nerve agents: a review. *Neurology* 42, 946–950.
- Marrs, T. C. (1996). *Chemical warfare agents: toxicology and treatment*. Wiley: Chichester.
- Maynard, R. L. (1999). Toxicology of chemical warfare agents. In: Ballantyne, B., Marrs, T. C. & Syversen, T. (eds.) *General and applied toxicology* (2nd edn., pp. 2079–2109). London: MacMillan.
- National Research Council (1999). *Chemical and biological terrorism: research and development to improve civilian medical response*. Washington, D.C: National Academy Press.
- Reutter, S. (1999). Hazards of chemical weapons release during war: new perspectives. *Environmental Health Perspectives* 107, 985–990.
- Suchard, J. R. (2002). Chemical and biological weapons. In: Goldfrank, L. R., Flomenbaum, N. E., Lewin, N. A., Howland, M. A., Hoffman, R. S. & Nelson, L. S. (eds.) *Goldfrank's toxicologic emergencies* (7th edn., pp. 1527–1551). New York: McGraw-Hill.

Chernobyl, Stress Effects of

A Tønnessen and L Weisaeth

National Competence Centre on Traumatic Stress,
Oslo, Norway

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
L Weisaeth, volume 1, pp 435–437, © 2000, Elsevier Inc.

The Chernobyl Accident: Health Consequences

Short-Term Stress Effects

Long-Term Stress Effects

Short- and Long-Term Stress Effects in Far-Off Areas

Glossary

*Acute radiation
syndrome
(ARS)*

ARS is the collective term for the hematopoietic, gastrointestinal, and central nervous system/cardiovascular systems response to whole-body acute or subacute exposure to penetrating ionizing radiation (e.g. x-ray or gamma radiation) above a certain threshold (roughly 1–2 Gy).

The study of the causes of disease.

Etiology

Pathogenesis

The origin and development of disease.

Shock trauma

A severe psychological experience, usually involving a threat to life, for which the individual was totally unprepared.

<i>Thyroid gland</i>	An endocrine gland, which is located in the base of the neck and produces thyroxine.
<i>United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR)</i>	UNSCEAR mandate in the United Nations system is to assess and report levels and effects of exposure to ionizing radiation and the Committee was established by the General Assembly of the United Nations in 1955. Governments all over the world rely on the Committee's estimates as the scientific basis for evaluating radiation risk and for establishing protective measures.

The Chernobyl Accident: Health Consequences

The nuclear reactor accident in Chernobyl, Ukraine, on April 26, 1986, created acute, short-term, and long-term effects both locally and farther away. The Chernobyl nuclear power plant (NPP) is located in the Kiev region of Ukraine, 15 km from the southern border of the Gomel region of Belarus and 150 km from the western part of the Bryansk region in Russia.

In contrast to the Three Mile Island accident in 1979, in which the safety containment held the release, at Chernobyl there were extremely large releases of radioactivity to the environment because the design of the reactor did not provide for the additional safety containment. Furthermore, it is unlikely that containment could have withstood the kind of pressure that the Chernobyl accident created. The graphite used for moderating the RBMK reactor kept a fire and the releases going for approximately 10 days, and it led to substantial cross-border radioactive fallout. The radioactive fallout resulted in three main spots of contamination within the former Soviet Union; they have been termed the Central, the Bryansk-Belarus, and the Kaluga-Tula-Orel spot.

The far-field fallout was chiefly dependent on the meteorological conditions; these affected the direction taken by the atmospheric releases and whether precipitation occurred while the release passed a particular area. The radioactive clouds first spread to Scandinavia and then turned toward the Balkans, affecting various other parts of Europe. It is important to note that the washout from the Chernobyl fallout plume resulted in very heterogeneous fallout, especially compared to previous experiences with atomic bomb test fallout. Although, the levels of bomb fallout had also been associated with washout and precipitation levels, they were not associated with such extreme local variations as observed with the Chernobyl fallout. The previous experiences with more homogenous bomb fallout made quite a few

experts believe that the early measurements from a few areas were valid indications for larger areas. However, with such a patchy fallout pattern, this was not the case. For instance, in Norway when the central Oslo area was not affected with fallout, experts made conclusions about the rest of the country that later turned out to be premature. UNSCEAR has estimated that 57% of the collective effective dose was received in Europe, outside the former USSR.

To avoid deterministic health effects from exposure to ionizing radiation, the city of Pripjat (49 000 inhabitants) located 3 km from the Chernobyl NPP was evacuated by Soviet authorities on Sunday, April 27. However, until the evacuation was ordered, the inhabitants of the town were given no official information about the accident and, for instance, the amusement park was open on Saturday evening. In fact, the official news report in the mass media of the Chernobyl accident started on Monday, April 28 as a local incident at the Swedish NPP Forsmark; as an employee entered the NPP, he triggered the alarm system because of Chernobyl fallout on his foot. The Forsmark NPP dispatched emergency measurement crews, which found increased radiation levels everywhere; hence, the Forsmark NPP was evacuated. On Monday, Swedish and Finnish authorities exchanged measurement data and pressed the Soviet authorities in the evening to communicate to the world that there had been an accident.

Because of the information blockage by Soviet authorities, the trust – and lack of trust – in the available information was a central issue. Also, within the same power structure held by the Soviet state, gigantic clean-up and intervention measures were conducted, some of which might have a quite limited dose-reduction effect compared to their rather large negative psychosocial effects.

On May 2 and 3, an additional 11 000 inhabitants in 14 villages in the 10-km zone were evacuated, and 42 000 inhabitants of 83 villages within the 30-km zone were evacuated between May 4 and 7. Furthermore from June to September, an additional 57 villages in Belarus, one village in Ukraine, and four villages in Russia (Bryansk region) were resettled, resulting in a total of 116 000 evacuated individuals. There is an ongoing discussion about whether the health gains from reducing the exposure to ionizing radiation by the ordered evacuations of all these people outweighed the psychosocial and stress impacts of being evacuated.

Extensive research on the health effects of the two major civilian nuclear reactor accidents (Three Mile Island and Chernobyl) has pointed to one particular finding that is considered to be of major importance, namely the impact of the psychological

fallout. Psychological impacts are important not only after the event of a nuclear accident itself but also with regard to the threat of a nuclear accident, the threat of nuclear material gone astray, or, with the war on terror, the threat of a terrorist event involving radioactive material.

Before we return to these psychological effects, a description of the more immediate and somatic effects is given.

On-Site Consequences

During the night of April 25–26, 1986, 176 staff members were working the night shift on Units 1–4 of the Chernobyl NPP while 268 staff members were working night shift on constructing two new additional reactors some 1.5 km to the southeast of the other reactors. A plant employee working immediately above the exploding reactor was killed instantly. His body could not be recovered. Another employee was crushed by debris and badly burned, and he died from his injuries within hours.

Those participating in the heroic immediate firefighting efforts were the most heavily affected by exposure to ionizing radiation immediately after the accident. More than 100 firefighters from the plant and from the nearby city of Pripyat participated in the vital task of limiting the raging fires and saving the other reactors also from exploding; around 5:00 a.m., all fires except the graphite fire in the core were extinguished. The heroic efforts of the firefighters, emergency workers, and plant personnel were vital in limiting the consequences. These efforts caused these workers to experience many injuries and exposure to much ionizing radiation, but saved reactor Unit 3 of the NPP (without their efforts a hydrogen fire or explosion and fire in the oil of the turbines might have had even larger consequences).

An explosion in the core of the reactor had not been thought to be possible. In harmony with human behavior in other crisis situations, hours passed until the operating crew of Unit 4 accepted and realized the consequences of the reports from the rescue workers that the core had been destroyed. The denial of these observations led to a delay in their realizing that highly dangerous graphite and other debris from the core were causing the fires on the roofs (including the roof of Unit 3). Radiation levels were so high in the damaged part of the plant and outside that the monitoring equipment could not measure it – the readings went completely off scale. The firemen and the plant personnel had no personal dosimeters and were very severely exposed. Within 1 h, the first of many cases of acute radiation syndrome (ARS)

appeared, during the first 12 h 132 emergency workers were hospitalized in Pripyat, suspected to be suffering from ARS. On Saturday, a specialized emergency team was established and examined some 350 people within the first 3 days, identifying 299 suspected cases of ARS.

Those with indications of ARS were sent mainly for specialized treatment to Moscow and to hospitals in Kiev. During the subsequent days, some 200 people were admitted for examination. The specialized treatment center in Moscow admitted the more severe cases, and after screening, it treated 115 patients for ARS. These patients were categorized into four categories according to the severity of the ARS. All the 20 patients with fourth-degree ARS (bone marrow dose range 6.1–16 Gy) died; 14 of the 21 patients with third-degree ARS (bone marrow dose range 4.2–6.4 Gy) survived; 42 out of the 43 with second-degree ARS survived; and all 31 with first-degree ARS survived.

After extensive effort, the Moscow center managed to establish that the two most significant dose factors were that of a general relatively uniform whole-body gamma irradiation and extensive beta irradiation of extensive body surfaces. The beta-radiation burns caused extensive skin lesions that proved to be a significant problem, and 16 out of the 28 deaths have skin injuries listed along with other causes of death. All in all, there were a total of 32 deaths during the initial period.

There is an ongoing controversy regarding the number of deaths among the 600 000 liquidators, with many heterogeneous subgroups including the early-involved firefighters and the soldiers who removed debris and built the sarcophagus over the wrecked reactor.

Far-Field Effects

One of the somatic health impacts of Chernobyl that is currently demonstrated is the large increase in the prevalence of thyroid diseases, including cancer, in children in the affected areas in the former Soviet Union, mainly in Belarus and Ukraine. In Belarus, Ukraine, and Russia, 5 million people were exposed to low levels of ionizing radiation over an extended period of time. It has been important, but difficult, to make a clear distinction between the immediate and deferred direct effects arising from exposure to radiation and the possible indirect psychosomatic effects. Complex psychosomatic or somatopsychological processes may be at work. People in the affected areas naturally feared that they had been exposed to radiation that might cause future health damage.

The political strategy of the Soviet experts and authorities in the management of the postaccident

situation involved, first, the withholding of information from the population by creating a total information blockade and, second, the psychiatrization (blaming the victim) of the health consequences of the accident. For example, the term radiophobia was widely used starting in 1986 to explain the population's reactions, despite the objective evidence of there being valid reasons for a certain degree of fear.

Short-Term Stress Effects

The explosion in the NPP contained stressful stimuli typical of shock trauma. This is likely to produce acute stress reactions and later posttraumatic stress disorder (PTSD, a psychiatric disorder characterized by intrusive recollections and the reliving of the trauma, along with symptoms of avoidance and hyperarousal. Unlike the stimuli carrying a high risk of PTSD, which were time-limited dangers, additional stressors inherent in the radioactive contamination were ongoing, future oriented, somatically based, and not confined to a single past event that could be processed by the senses.

The Chernobyl NPP accident started a sequence of events that has continued to unfold over several years, thereby creating a situation of chronic stress. For the people exposed, the Chernobyl nuclear accident had no clearly defined low point from which things would gradually get better. Exposures to a high-risk environment, even when the feared consequences, such as cancer, may never materialize, elicit stress reactions that are influenced by individual perceptions of the danger. The fear was aggravated by the lack of adequate information immediately after the accident.

Human contamination of the biosphere is a comparatively new type of crisis. Rather than merely an ecological and medical emergency, this was also a social and political crisis. In the affected regions, covering thousands of square kilometers, a culture of uncertainty developed. The environmental contamination was not possible to see, hear, smell, taste, or touch. In a silent disaster, it is impossible for humans to determine if and when they are being exposed because they are totally dependent on experts for information. Because of the secretiveness of the Soviet authorities, news about the accident was released several days after the accident. Only then did the populations in nearby towns become aware of the radiation to which they had been exposed and the consequent threat of long-term illness. These people were struck by fear and the distrust of political authorities. When people were evacuated from their homes, they suffered considerable stress because they were kept ignorant about what was going on. They underwent disruption in

community infrastructure and social interaction and faced uncertainty about future housing and employment. Many evacuees who moved to new settlements were particularly depressed in their new homes because of financial difficulties, fear of isolation, and concern for the health of their children. The tense situation caused considerable stress, which, combined with the constant fear of health damage from the radioactive fallout, led to a rising number of health disorders being reported to local outpatient clinics. People received late and reluctant information, leading to widespread rumors. Only 30 months after the accident did Soviet authorities admit the contamination of some villages. This obvious disinformation resulted in widespread indifference, denial of danger, and public distrust in the government. People suffered from various symptoms frequently called chronic radiation sickness, the exact definition of which varied; fatigue, loss of memory, loss of appetite, and psychosomatic symptoms were among the most frequently documented symptoms in the clinics.

The findings of increased somatic complaints, with large discrepancies between the findings of thorough clinical examinations and the self-reports is reminiscent of Lifton's psychosomatic bind, found among Hiroshima atomic bomb survivors. Thus, substantial proportions of the populations in the former Soviet Union, severely affected with Chernobyl fallout, live with an intense focus turned inward, listening for all kinds of physical health complaints that might possibly be the first signals confirming their expectations of becoming ill consequent to the impact of Chernobyl.

Similarly, in recent years, several authors have suggested that psychological factors may be involved in the etiology and pathogenesis of the health problems occurring after toxic exposure. A number of terms have been proposed to describe this syndrome: chronic environmental stress syndrome, informed of radioactive contamination syndrome, and toxic stress syndrome.

Long-Term Stress Effects

The breakup of the Soviet Union in 1989 increased the level of stress in the affected populations dramatically. Added to the ecological crisis, the chronic political, economic, military, social, and psychological crisis introduced a cluster of stressors for ordinary men and women. Insecurity, unpredictability, and shortages in goods and other life necessities caused severe hardships. Within this total situation of need, the relative importance of the Chernobyl problem was reduced. The numerous health problems reported by the local population of all ages are, however,

attributed to the radioactivity. Even the fall of the Soviet Union has been attributed to the Chernobyl accident. However, problems in everyday life are so demanding that many people cannot muster the necessary energy to maintain adequate precautions concerning radioactive contamination. Thus, attribution and indifference appear to be common stress reactions. Studies by Johan Havenaar and colleagues, in particular, in the affected regions have shown that the psychological impact of exposure to noxious agent has induced a wide range of health effects, including symptoms of anxiety; depression; non-specific somatic complaints, especially fatigue and headaches; loss of concentration; subclinical physiological changes; increased illness awareness; lowered threshold for illness roles; and increased help-seeking behaviors.

Evelyn Bromet and colleagues have shown how mothers with small children are a vulnerable group both after the Three Mile Island and the Chernobyl accident, and seemingly their immediate coping is important also for the long-term impacts.

Short- and Long-Term Stress Effects in Far-Off Areas

Paradoxically, people in remote areas – as far away as Norway and Sweden – were alarmed at a fairly early stage due to nuclear monitoring and Western news media. These countries received heavy fallout due to the unfortunate wind direction and rain. Reactions varied among public awareness, alarm, and shock. Studies have shown that an information crisis developed in the first weeks following the Chernobyl disaster. The information crisis resulted from a combination of factors: (1) the shortcomings of previous public education on the matter; (2) the shortage of reliable data on the location and concentration of radioactive material, including the premature, too optimistic expert conclusions erroneously taking the central-location measurements to be indicative of the whole country; (3) the diffuse threat, ambiguous risk, and theoretical complexity of the subject; (4) the seeming contradiction between the health authorities' assurances that, on the one hand, the radiation level was not dangerously high and, on the other hand, some precautions had to be taken; (5) the experts' obvious difficulties in conveying simple information on complicated matters and statistical risks; and (6) the media's focusing on conflicts, particularly between experts.

In Scandinavia, it was found that, although the majority of the population was psychologically affected by the crisis in that all but few had some reactions to it, only a small proportion (1–3%) seemed to

have developed posttraumatic stress reactions of such a severity that they needed or sought help. Data analysis showed that the degree of exposure, sex, age, educational level, general threat perception, and previous mental health were associated both with information and with the reaction variables. Women were significantly more alarmed than men.

The Southern Saamis was the group that was worst affected in Scandinavia because radiocesium from Chernobyl fallout went into the Lichen (reindeer) Saami food chain. The Southern Saami reported that repeated whole-body measurements aided them in turning to a more active coping style in which they believe there indeed is something they might do – the application of countermeasures to influence possible adverse health effects. This is in stark contrast to some of the more fatalistic coping patterns found among the populations in many contaminated places in the territories of the former Soviet Union.

Providing the public with immediate, simple, and reliable information is of crucial importance in the management of such silent disasters. In fact, in such disasters, information can itself be a threat. It is not realistic for the relevant authorities to start producing information materials or to start educating the public during the height of a crisis; thus, it is important that preparatory efforts be made. A key issue for the public is what people themselves might do to influence the possible health impacts on themselves and their family in case of an accident, and this perspective should be center stage for authorities charged with the challenging task of providing information.

Further Reading

- Baum, A. (1986). Toxins, technology, disaster. In: Vanden Bios, G. R. & Bryants, B. K. (eds.) *Cataclysm, crisis and catastrophes: psychology in actions*, pp. 9–53. Washington, DC: American Psychological Association.
- Bromet, E. J. (1998). Psychological effects of radiation catastrophes. In: Peterson, L. E. & Abrahamson, S. (eds.) *Effects of ionizing radiation: atomic bomb survivors and their children (1945–1995)*. Washington, DC: Joseph Henry Press.
- Collins, D. L. and de Carvalho, A. B. (1993). Chronic stress from the Goiania ¹³⁷Cs radiation accident. *Behavioral Medicine* 18(4), 149–157.
- Green, B. L., Lindy, J. D. and Grace, M. (1994). Psychological effects of toxic contamination. In: Ursano, R. J., McGaughey, B. G. & Fulerton, C. S. (eds.) *Individual and community response to trauma and disaster: the structure of human chaos*, pp. 154–176. Cambridge, UK: Cambridge University Press.
- Guskova, A. K., Barbanova, A. V., Baranov, A. Y., et al. (1988). Acute radiation effects in victims of the Chernobyl nuclear power plant accident. In: United Nations

- Scientific Committee on the Effects of Atomic Radiation (ed.) *UNSCEAR (1988) sources and effects of ionizing radiation*. 1988 Report to the General Assembly, with annexes, pp. 613–631. New York: United Nations.
- Havenaar, J. M. (1996). After Chernobyl: psychological factors affecting health after a nuclear disaster. PhD thesis, Universiteit Utrecht.
- Lee, T. R. (1996). Environmental stress reactions following the Chernobyl accident. In: International Atomic Energy Association (ed.) *One decade after Chernobyl: summing up the consequences of the accident*, pp. 283–310. Vienna: IAEA.
- Shigematsu, I. (ed.) (1991). *The international Chernobyl project technical report: assessment of radiological consequences and evaluation of protective measures*. Report by an international advisory committee. Vienna: IAEA.
- Tønnessen, A. (2002). Psychological reactions to nuclear threats – information, coping and the uncertainties of outcome at the individual level. PhD thesis, University of Oslo.
- Vyner, H. M. (1988). *Invisible trauma: the psychological effects of invisible environmental contaminants*. Lexington MA: Lexington Books.
- Weisæth, L. (1991). Psychosocial reactions in Norway to nuclear fallout from the Chernobyl disaster. In: Couch, S. R. & Kroll-Smith, J. S. (eds.) *Communities at risk: collective responses to technological hazards*, pp. 53–80. New York: Peter Lang.
- World Health Organisation (1996). *Health consequences of the Chernobyl accident: results of the IPHECA pilot projects and related national programmes, scientific report*. Geneva: WHO.

Child Abuse

C C Swenson and L Saldana

Medical University of South Carolina, Charleston,
SC, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
C C Swenson and C E Ezzel, volume 1, pp 438–441, © 2000,
Elsevier Inc.

Child Physical Abuse
Child Neglect

Glossary

<i>Factitious disorder by proxy</i>	Disorder that occurs when adults deliberately produce or feign physical or psychological symptoms in a child under their care with the presumed motivation a psychological need to assume the sick role.
<i>Failure to thrive</i>	In infancy this is indicated by a growth delay with postural signs (poor muscle tone, persistence of infantile postures) and behavioral signs (unresponsive, minimal smiling, few vocalizations) and may be a result of child neglect.
<i>Multisystemic therapy for child abuse and neglect (MST-CAN)</i>	An ecological treatment model that takes into account all of the etiological factors related to a problem behavior and treats those factors through changes in the family and social ecology. The delivery of treatment is generally in the home and at times convenient for

families, with 24 hours a day, 7 days a week on-call coverage for crises. This treatment is an adaptation of standard MST that has established efficacy with deep-end delinquents. MST-CAN, developed at the Medical University of South Carolina, has been shown to be effective with physically abused children and their families.

Parent-child interaction therapy (PCIT)

A short-term early intervention approach in which young children and their parents are provided with direct coaching in interactions during play or specific tasks. PCIT is appropriate for young children demonstrating externalizing and internalizing behavior problems. PCIT has recently been established as effective with physically abused children and their families in a clinical trial at the University of Oklahoma Health Sciences Center.

Reinforcement-based intensive outpatient therapy (RBT)

A treatment model developed at Johns Hopkins University and applied to mothers with heroin or cocaine addiction. The model provides intensive treatment and focuses on skills needed to become abstinent, monitoring of substance use, abstinence-contingent housing, prosocial recreational activities, social skills, and employment counseling.

Shaken baby syndrome

An injury to the brain of an infant that results from vigorous shaking. Shaken baby syndrome is a form of physical abuse and often results in death to the child.

Child Physical Abuse

Child physical abuse is defined as physical injury to a child under age 18 by an adult. Acts of physical abuse may vary from hitting that leaves bruising to actions that involve welts, cuts, bites, broken bones, burns, poisoning, and internal injuries to soft tissue and organs. In special cases, termed factitious disorder by proxy, adults inflict injury or induce illness in a child under their care, present the child for medical care, and feign no knowledge of the etiology.

Epidemiology

According to reports from state agencies, nearly 3 million children are abused and neglected annually. Among those reported for suspicions of abuse, roughly 25% are investigated for physical abuse. In a congressionally mandated prevalence study of child abuse and neglect called the Third National Incidence Study, a broader sample of cases that extended beyond child protection reports was included. Of the 2 815 600 abused or neglected youth, 22% were physically abused.

Etiology

The complexity of child abuse and neglect is highlighted in the literature, showing that multiple risk factors relate to whether physical abuse will occur. These risk factors include (1) parent factors such as history of abuse, substance abuse, deficits in knowledge of child development, depression, poor impulse control, negative perception of the child, and neuropsychological dysfunction; (2) child factors such as age, aggression, noncompliance, and delayed development; (3) family factors such as high conflict, spouse or partner abuse, and single marital status; (4) social network factors such as social isolation, low use of community resources, and limited involvement in community activities; and (5) community factors such as economic disadvantage, instability, poor organization, and high resource need coupled with low resources available for child care. Although each of the systems noted above (i.e. family, community) is part of every child's life, the particular risk factors within each system will differ from family to family. For example, in one family, substance abuse, child aggression, and isolation from social networks may be the key factors driving the abuse. In another family, partner violence, child developmental delay, living in an unstable neighborhood, and low parenting skills may be the driving risk factors.

Medical and Mental Health Effects

Physical abuse injuries range from mild (e.g., bruises) to severe (e.g., broken bones, gunshot wound) and

account for 10% of injuries among children who present for emergency room care. Physical injuries that might be mild for an older child can be fatal for an infant. For example, when babies are shaken, the resulting damage to the brain can leave children severely developmentally delayed but may also result in death. This act against infants is termed shaken baby syndrome. Some children do not need medical care following physical abuse, but an assessment of potential injury is always a safe bet. Mental health effects of child physical abuse vary by the child, experience, and context of the abuse. These effects generally occur in five domains: (1) aggression and behavioral dysfunction toward peers and adults; (2) poor interpersonal relations and social competence such as social skills deficits, poor social problem solving, and peer rejection; (3) trauma-related emotional symptoms such as depression, posttraumatic stress disorder (PTSD), anxiety, and suicidality; (4) developmental deficits in relationship skills including anxious attachment in very young children, low frustration tolerance, and difficulty managing conflict; and, (5) cognitive and neuropsychological impairment such as receptive and expressive language difficulties and low performance in math and reading. In addition, physical abuse places children at risk of long-term mental health difficulties in adulthood such as substance abuse, violent crime, depression, and relationship problems.

Treatment Approaches

In a number of practice settings, physically abused children participate in individual treatment and parents go to parenting groups. The literature on physical abuse indicates that behavioral parent training has empirical support with this population for reducing parental aggression toward children. However, the problem of child abuse and neglect is so complex that individual child treatment or parent training alone will not resolve the risk factors that have led to the abuse in the first place. Since 1996, advances have been made in research regarding the treatment of physically abused children and their families. Through randomized clinical trials, three treatments are showing promise: parent-child interaction therapy (PCIT), trauma-focused cognitive-behavioral therapy (CBT), and multisystemic therapy for child abuse and neglect (MST-CAN). All three are treatments that address parent and child risk factors. This characteristic is important in that past research has focused on treatment of the parent or child only. Given the multidetermined etiology that relates to physical abuse, the importance of continued evaluation of comprehensive treatments and dissemination of effective treatments is paramount. Parental substance abuse has become a major factor in child

abuse and neglect reports and out-of-home placement. Treatments that successfully address substance abuse and child abuse are sorely needed.

Child Neglect

Despite inconsistencies in state definitions, generally neglect is defined as physical (refusal or delay in seeking health care; inadequate supervision, hygiene, nutrition, and safety), emotional (failure to provide adequate attention, warmth, and affection and/or exposure to substance abuse or violence), and/or educational (failure to promote education or school attendance and inattention to special education needs) omissions in parenting. Recent evidence suggests that these neglectful behaviors fall into three distinct categories: physical, psychological, and environmental neglect. Child neglect has been grossly underresearched in the United States, leading to the phrase neglect of neglect. Throughout the literature, child abuse and neglect often are grouped together despite the clear distinction between acts of commission and those of omission. Whereas many people think of child abuse when considering child maltreatment, child neglect is in fact more prevalent, costly, and detrimental.

Epidemiology

Of the 896 000 indicated child maltreatment cases in the United States in 2002, over 60% of those children were exposed to neglect. The number of youths who experience child neglect has more than doubled since the mid-1980s, an increase more than eight times greater than the rise in children's population. Moreover, prevalence studies suggest that approximately 45% of child maltreatment fatalities are the result of neglect, with 1 400 deaths recorded in 2002. These figures likely provide a gross underestimate of the incidence of child neglect in our country, as acts of omission are far more difficult to identify than those of commission. Studies from the general population suggest that as many as 2 million children are endangered by neglect.

Etiology

As with child physical abuse, the factors related to neglectful parenting practices are multidetermined. Evidence suggests that risk factors include (1) parent factors such as mental illness, low cognitive functioning, substance abuse, and a history of child maltreatment; (2) family factors such as domestic violence, social isolation, high levels of family stress, and poverty; (3) child factors such as child physical disability or chronic illness; and (4) community factors such

as impoverished neighborhoods, acceptance among community members of low monitoring or low school attendance, and poor community cohesion.

Although the etiology of neglectful parenting is multifaceted, of particular note is the influence of substance abuse. There appears to be an association between the staggering increase in neglectful parenting practices and the increase in identified substance abuse among child welfare cases. Children whose parents abuse substances are more than four times as likely to be neglected as those who are not abusing substances, and parental substance abuse is more often a factor in reports of child neglect than other forms of child maltreatment. Parental substance abuse, which is identified in up to 79% of child protection cases, not only is one of the strongest risk factors for abuse and neglect, but also is the deciding factor in the majority of cases in which children are taken into custody and placed out of the home.

Medical and Mental Health Effects of Child Neglect

The medical and mental health effects of neglect on the child are dependent on the nature of the neglectful parenting practices. Potential consequences are immense and can cause deleterious physiological and emotional stress on the child and family. Compared to other forms of maltreatment, child neglect has been associated with the most profound developmental and cognitive delays and has the greatest risk of enduring effects. Children who are neglected often evidence language and academic deficits, inadequate social skills, poor growth development including nonorganic failure to thrive, medical problems, and poor attachments and interpersonal relationships. Moreover, those who are neglected by substance-abusing mothers are at increased risk for conduct problems including delinquency, substance abuse, and maltreatment of their own children. The most severe forms of neglect, however, result in child fatality.

Treatment Approaches

There is a dearth of controlled clinical trials with sufficient power to examine outcomes with regard to neglect. Despite these deficits in the literature, there is evidence to suggest that an ecological approach to treatment is beneficial with neglectful families. Across studies, results suggest that neglectful parents benefit from skills training including improved problem solving, reducing safety hazards, increasing hygienic and home cleaning practices, and improving child nutrition and intellectual stimulation. Further, studies have demonstrated positive parental response to protocols targeting improving adaptive attachment styles, affective relationships, and recognition of children's

emotional needs. Despite these advances in treatment approaches with neglectful parents, there is a strong need for evidenced-based practice in this area. In particular, an urgent need exists for treatment models to address co-occurring substance abuse and child abuse and neglect. At present, the majority of treatments address the two problems independently. Work has begun in Connecticut on a project known as Building Stronger Families. This project involves the integration of MST-CAN and reinforcement-based therapy. The latter has established efficacy with serious substance abuse.

See Also the Following Articles

Child Sexual Abuse; Childhood Stress; Posttraumatic Stress Disorder in Children; Child Physical Abuse.

Further Reading

Chaffin, M., Silovsky, J. F., Funderburk, B., et al. (2004). Parent-child interaction therapy with physically abusive parents: efficacy for reducing future abuse reports. *Journal of Consulting and Clinical Psychology* 72, 500–510.

Donohue, B. (2004). Coexisting child neglect and drug abuse in young mothers. *Behavior Modification* 28, 206–233.

Dubowitz, H. (1999). *Neglected children: research, practice, and policy*. Thousand Oaks, CA: Sage.

Gaudin, J. (1993). *Child neglect: a guide for intervention*. Washington, D.C.: U.S. Department of Health and Human Services.

Gruber, K., Chutuape, M. A. and Stitzer, M. L. (2000). Reinforcement-based intensive outpatient treatment for inner city opiate abusers: a short-term evaluation. *Drug and Alcohol Dependence* 57, 211–223.

Kolko, D. J. and Swenson, C. C. (2002). *Assessing and treating physically abused children and their families: a cognitive behavioral approach*. Thousand Oaks, CA: Sage Publications.

National Clearinghouse on Child Abuse and Neglect (2004). *Child maltreatment 2002: summary of key findings*. Washington, D.C.: U.S. Department of Health and Human Services.

Sedlak, A. J. and Broadhurst, D. D. (1996). *Executive summary of the Third National Incidence Study of Child Abuse and Neglect*. Washington, D.C.: U.S. Department of Health and Human Services.

Swenson, C. C., Saldana, L., Joyner, C. D. and Henggeler, S. W. (in press). *Ecological treatment for parent to child violence. Interventions for children exposed to violence*. New Brunswick, NJ: Johnson & Johnson.

Child Maltreatment and Genetic Factors See: Child Abuse; Child Sexual Abuse; Childhood Stress.

Child Physical Abuse

C C Swenson

Medical University of South Carolina, Charleston, SC, USA

© 2007 Elsevier Inc. All rights reserved.

Prevalence

Etiology

Mental Health Effects

Treatment Approaches

Glossary

Child physical abuse Involves physical behavior by an adult toward a child that results in injury to the child. Signs of physical abuse may

Multisystemic therapy for child abuse and neglect (MST-CAN)

include bruises, welts, cuts, broken bones, skull fractures, burns, poisoning, internal injuries of soft tissue and organs, and injuries to the bone and tissue joints of a child under the age of 18 years

An adaptation of multisystemic therapy for delinquency that involves treatment of the entire family through interventions targeting multiple risk factors. Treatment is delivered in the home and at times convenient for families, with 24 hours a day, 7 days a week on-call coverage for crises. This model, developed at the Medical University of South Carolina, is applied to families who have come under the guidance of child protection due to physical abuse.

<i>Munchausen by proxy syndrome (MBPS)</i>	A form of child abuse in which a parent fabricates or produces illness in a child and/or creates physical signs that persistently result in unnecessary medical treatment.
<i>Parent-child interaction therapy (PCIT)</i>	A short-term intervention currently applied to young children and their parents. The method involves direct coaching of parents in interactions during play or specific tasks. Through a recent clinical trial conducted at the University of Oklahoma Health Sciences Center, PCIT was established as effective for treating young physically abused children.
<i>Shaken baby syndrome</i>	Applies to infants who undergo vigorous shaking that leads to acceleration-deceleration injuries to the brain. These infants generally present at less than 1 year of age with seizures, vomiting, lethargy or bradycardia, hypotension, respiratory irregularities, coma, or death.

Child physical abuse is a major public health problem that generally involves acts of physical violence by a parent or carer toward a child. Because physically abusive acts range greatly in intensity and duration, the subsequent mental health consequences vary widely. Treatment must involve the child, parent, and entire family.

Prevalence

Data regarding the prevalence of child physical abuse (CPA) in the United States are gathered from child protection agencies and through national surveys. These data indicate that annually, nearly 3 million children are abused and neglected. Among those reported for suspicions of abuse, roughly 25% are investigated for physical abuse. In an evaluation of reporting practices of professionals, the National Center on Child Abuse and Neglect (NCCAN) conducted a series of National Incidence Studies of Child Abuse and Neglect. They found that approximately 2.3/1000 children are physically abused. The 1995 Gallup poll estimated the rate of CPA to be 49/1000 children, more than five times the NCCAN report.

Etiology

The etiological factors related to CPA are multiple and complex. They range across many different systems (e.g., parent, child, family) and vary by individual families. Current research has not established what factors cause physical abuse. At best, the understanding of etiology involves factors that

correlate with physical abuse, also called risk factors. These risk factors are specific to the child, parent, family, and community.

The children most at risk of physical abuse are those who (1) are younger, (2) have developmental delays, (3) have other special needs such as chronic medical conditions, and (4) exhibit noncompliant behavior. A number of parent factors increase the risk of physical abuse. First is childhood abuse history. Roughly one-third of adults who were physically abused as children go on to abuse their own children. Second, certain cognitive factors increase the parent's risk of harm to the child. Parents who view their child in a negative light, distort beliefs about the reasons why children act as they do (e.g., he cries to make me upset), expect the child to be responsible for the parent's welfare, and have unrealistic expectations beyond age and skill capability are at greater risk of abusing their child. Parents who physically abuse are more likely to show emotional difficulties such as difficulty regulating emotions, including irritability, sadness, anxiety, explosiveness, hostility, anger, and perceptions that life is more stressful. They are more likely to show deficits in positive parenting. Psychiatric difficulties such as depression and posttraumatic stress disorder may make general problem solving difficult and hamper an individual's capacity to regulate emotions and parent in a nonphysical way. Parental substance abuse, in particular, is a factor that places parents at significant risk of abusing their children and is a major factor in the decision to remove children from their families and place them in state custody. Family factors that increase the risk of physical abuse include a volatile home environment in which there is domestic violence, limited psychosocial resources, and general family stressors such as frequent moves or family dissolution. Existing research has identified economic disadvantage, instability and isolation, and neighborhood burden as factors that increase the risk of child maltreatment. The latter factor, neighborhood burden, refers to greater childcare need with fewer resources.

Mental Health Effects

Mental health difficulties related to CPA include those that occur in the short term and those that persist. It should be noted that some physically abused children will not experience mental health problems. Others will experience problems that dissipate with time or with intervention, and yet others will experience problems that endure into adulthood and result in patterns of behavior and coping that disrupt the management of behavior, emotions, and relationships.

Risk and Buffering Factors

In part, the emergence and persistence of symptomatology relate to factors specific to the event, cultural and community context, family, and individual characteristics. Such factors may potentiate a negative or a buffering effect. In relationship to the event, exposure to multiple incidents of maltreatment increases vulnerability for the development of mental health problems, such as substance abuse, severe suicidal behavior, and emotional problems. In fact, as the number and type of traumatic events accumulate, the outcomes (i.e., internalizing and externalizing disorders) appear to worsen. In addition, chronicity of physical punishment during childhood also has been associated with delinquent behavior, symptoms of depression, suicidal thoughts, and alcohol abuse later in life.

Factors specific to the child, such as gender (e.g., girls are at risk of internalizing disorders; boys are at risk of externalizing) or developmental disabilities may place him or her at risk of psychiatric symptomatology. In families, greater maternal stress and poor conflict management are associated with severity of functional impairment. The combination of low family cohesion and high family conflict are common risk factors for aggression and substance abuse problems. Social and family support and placing responsibility on the adult who abused rather than the child relate to more positive mental health outcomes.

Short-Term Effects

Multiple effects have been associated with the occurrence of CPA. Behavioral and emotional difficulties include externalizing problems, such as aggression, and internalizing problems, such as anxiety or depression. Social and relationship difficulties include poor peer and adult relations that are exacerbated by difficulty managing conflict and low confidence in regulating negative mood. Cognitive and neuropsychological impairment is typically shown through skills deficits in expressive and receptive language, reading comprehension, and auditory attention. More severe forms of physical abuse result in risk of death. These include shaken baby syndrome and Munchausen's by proxy. In the latter, parents have been shown to conduct behaviors such as smothering to create a medical problem that then wins them attention.

Long-Term Effects

The potential long-term impact of CPA falls into three categories: aggression, substance abuse, and emotional problems. Aggression is expressed toward peers, family members, and dating partners and may

include violent crimes or physical abuse of one's own child. CPA is associated with substance use, more lifetime treatment, and a more severe course of substance abuse later in adulthood. For adults, a history of CPA has been found to relate to a host of emotional difficulties, including somatization, anxiety, depression, hostility, paranoid ideation, psychosis, dissociation, and self-injurious and suicidal behavior.

Treatment Approaches

Treatments for the effects of CPA are those that are directed toward child, parent, or family symptomatology and comprehensive approaches that address all persons involved or impacted. Most studies targeting child symptoms have been with preschoolers who are in a day treatment program or receiving instruction in peer relations and have shown success in increasing social skills and peer relations. One cognitive-behavioral group therapy study with school-age physically abused children found reductions on self-report measures of posttraumatic stress disorder, general anxiety, depression, and anger symptoms but no improvement on social competence and externalizing behaviors, potentially highlighting the need to include carers in treatment. In a small number of studies, parent training has been shown to decrease abusive behaviors in parents. More comprehensive models, such as cognitive-behavioral treatment with parent and child, parent-child interaction therapy, and multisystemic therapy for child abuse and neglect, are showing promise for reducing child symptoms, adult risk factors related to CPA, and re-abuse. Importantly, recent comprehensive pilot programs such as the Building Stronger Families program being implemented by the Connecticut Department of Children and Families are targeting risk factors within the family and ecology with an intensive focus on parental substance abuse.

See Also the Following Articles

Childhood Stress; Cognitive Behavioral Therapy; Domestic Violence; Posttraumatic Therapy; Posttraumatic Stress Disorder in Children; Dissociation; Depression and Manic-Depressive Illness; Multiple Personality Disorder; Anxiety.

Further Reading

- Borrego, J., Urquiza, A. J., Rasmussen, R. A. and Zebell, N. (1999). Parent-child interaction therapy with a family at high risk for physical abuse. *Child Maltreatment* 4, 331-342.
- Chaffin, M., Kelleher, K. and Hollenberg, J. (1996). Onset of physical abuse and neglect: psychiatric, substance

- abuse, and social risk factors from prospective community data. *Child Abuse & Neglect* 20, 191–203.
- Chaffin, M., Silovsky, J. F., Funderburk, B., et al. (2004). Parent-child interaction therapy with physically abusive parents: efficacy for reducing future abuse reports. *Journal of Consulting and Clinical Psychology* 72, 500–510.
- Chapman, D., Whitfield, C., Felitti, V., Dube, S., Edwards, V. and Anda, R. (2004). Adverse childhood experiences and the risk of depressive disorders in adulthood. *Journal of Affective Disorders* 82, 217–225.
- Chasnoff, I. J. and Lowder, L. A. (1999). Parental alcohol and drug use and risk for child maltreatment: a timely approach to intervention. In: Dubowitz, H. (ed.) *Neglected children: research, practice, and policy*, pp. 132–155. Thousand Oaks, CA: Sage.
- Finkelhor, D. and Dziuba-Leatherman, J. (1994). Children as victims of violence: a national survey. *Pediatrics* 94, 413–420.
- Hien, D. and Honeyman, T. (2000). A closer look at the drug abuse-maternal aggression link. *Journal of Interpersonal Violence* 15, 503–522.
- Kolko, D. J. (1996). Individual cognitive behavioral treatment and family therapy for physically abused children and their offending parents: a comparison of clinical outcomes. *Child Maltreatment* 1, 322–342.
- Kolko, D. J. and Swenson, C. C. (2002). *Assessing and treating physically abused children and their families: a cognitive behavioral approach*. Thousand Oaks, CA: Sage.
- Meyers, J. E. B., Berliner, L., Briere, J., Hendrix, C. T., Jenny, C. and Reid, T. A. (2002). *The APSAC handbook on child maltreatment* (2nd edn., pp. 205–232). Thousand Oaks, CA: Sage.
- Swenson, C. C. and Brown, E. J. (1999). Cognitive-behavioral group treatment for physically abused children. *Cognitive and Behavioral Practice* 6, 212–220.
- Swenson, C. C. and Chaffin, M. (2006). Beyond psychotherapy: treating abused children by changing their social ecology. *Aggression and Violent Behavior* 11, 120–137.
- Swenson, C. C., Randall, J., Henggeler, S. W. and Ward, D. (2000). Outcomes and costs of an interagency partnership to serve maltreated children in state custody. *Children's Services: Social Policy, Research, and Practice* 3, 191–209.
- Swenson, C. C., Saldana, L., Joyner, C. D. and Henggeler, S. W. (2006). Ecological treatment for parent to child violence. In: Lieberman, A. & DeMartino, R. (eds.) *Interventions for children exposed to violence*, pp. 155–185. New Brunswick, NJ: Johnson & Johnson Pediatric Institute.
- Swenson, C. C., Brown, E. J. and Lutzker, J. R. (in press). Issues of maltreatment and abuse. In: Freeman A. & Reinecke, M. (eds.) *Personality disorders in childhood and adolescence*.

Child Sexual Abuse

J A Cohen

Drexel University College of Medicine, Philadelphia, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J A Cohen, volume 1, pp 450–453, © 2000, Elsevier Inc.

Definition and Epidemiology

Symptomatology

Assessment

Treatment

Definition and Epidemiology

Child sexual abuse is not a clinical disorder or diagnosis in itself. It is, rather, a variety of events or experiences to which there may be a wide range of behavioral and emotional responses. In this sense, sexual abuse is best conceptualized as a life stressor rather than as a distinct clinical entity. The legal definition of child sexual abuse varies from state to

state, and there is little consensus regarding what acts constitute sexual abuse among mental health-care providers. A common operational definition of sexual abuse is sexual exploitation involving physical contact between a child and another person. Exploitation implies an inequality of power between the child and the abuser, on the basis of age, physical size, and/or the nature of the emotional relationship. Physical contact includes anal, genital, oral, or breast contact. The definition obviously encompasses a number of different behaviors; sexual abuse is therefore not a unitary phenomenon.

It is difficult to know the exact incidence and prevalence of sexual abuse in our society because most sexual abuse is not reported immediately, and sometimes it is never reported. The Third National Incidence Study on Child Abuse and Neglect estimated that more than 300 000 children were sexually abused in the United States in 1993. However, this study includes only reported cases of sexual abuse. This is probably an underestimation of the frequency of sexual abuse, as there is evidence that only about one-fifth of true abuse cases are actually reported.

Although sexual abuse occurs across religious, racial, and socioeconomic lines, several factors may put children at higher risk for sexual abuse. There is some controversy over the relationship between social class and child sexual abuse. Some researchers have found that reported cases of sexual abuse have come predominantly from families in the lower socioeconomic levels of society, but other studies have not supported this relationship. It is likely that this apparent relationship is due to a reporting bias because economically disadvantaged families are more likely to be scrutinized by social service agencies.

Symptomatology

There is also a great deal of variability in the behavioral, emotional, and social consequences of child sexual abuse. As a group, sexually abused children have been found to manifest significantly more behavioral and emotional difficulties than do nonabused children. Some researchers have found significant levels of self-reported depression in these children, while others have not. A significant number of these children and adolescents exhibit posttraumatic stress disorder (PTSD) and other anxiety symptoms. Recent, well designed studies which have controlled for factors such as genetics and family environment variables have demonstrated that child sexual abuse significantly increases risk for adverse psychological outcomes including depression, suicide attempts, substance abuse, anxiety, and subsequent sexual assault (Brent et al., 2002; Nelson et al., 2002).

Certain clinical issues appear to be common to most sexually abused children. These include feeling different or damaged compared to other children; feeling responsible for the abuse; feeling angry, fearful, or sad; taking on inappropriate behaviors and roles (including sexual ones); and losing trust in others. Briere's research on the Trauma Symptom Checklist for Children (TSCC) provided empirical support for these difficulties. The factor analysis of the TSCC demonstrated six discrete symptom fields: difficulties with sexual concerns, dissociation, anger, posttraumatic stress, depression, and anxiety symptoms.

Finkelhor discussed four traumagenic dynamics of sexual abuse, which included traumatic sexualization, stigmatization, powerlessness, and betrayal. The Children's Attributions and Perceptions Scale (CAPS) measures issues parallel to these traumagenic dynamics, in addition to measuring attributions about the sexual abuse. This study demonstrated that feeling different from peers (stigmatization), personal attribution for negative events, impaired trust (betrayal), and decreased perceived credibility (one

aspect of powerlessness) were significant issues for sexually abused children compared to a normal control group. Wolfe et al. developed an instrument, The Children's Impact of Traumatic Events Scale (CITES) to measure abuse-specific thoughts and attributions. Factor analyses revealed distinct issues including attributions, betrayal, guilt, helplessness, intrusive thoughts, sexualization, and stigmatization. Thus, there is strong empirical evidence that sexually abused children have clinical issues and difficulties specific to sexual abuse, which can be measured by newly developed standardized instruments.

Friedrich demonstrated that sexually abused children exhibit significantly higher rates of sexually inappropriate behavior than nonabused populations. These may arise as a result of learning inappropriate behaviors through the abusive experience or from the fact that children with preexisting sexually inappropriate behavior are more likely to be targeted as victims of sexual abuse. However, there may be other, non-abuse-related reasons for children to exhibit sexually inappropriate behaviors. These behaviors are therefore not diagnostic of the sexual abuse. It is also important to note that the majority of sexually abused children do not exhibit sexually inappropriate behaviors.

It should be clear from this discussion that there is no such entity as a child sexual abuse syndrome. That is, there is no characteristic presentation that clinically distinguishes sexually abused children from non-sexually abused children. Rather, sexual abuse may result in a variety of clinical presentations. Sexual abuse therefore cannot be diagnosed by the clinical features present at evaluation. Such a determination depends on careful history taking and interviews with the child and significant others.

Assessment

Because child sexual abuse is an event (or events) rather than a clinical syndrome, and psychological responses to this experience differ considerably, one would expect psychiatric diagnoses in sexually abused children to also vary considerably. This has been documented in the literature. McLeer et al. evaluated psychiatric diagnoses in sexually abused and non-sexually abused children referred for outpatient pediatric assessment, using the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS). The only significant difference in diagnoses between the two groups was the higher prevalence of PTSD in the sexually abused (42.3%) compared to the non-sexually abused (8.7%) children. In other diagnostic categories, there were no significant differences.

The process of determining accurate *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (DSM-IV) psychiatric diagnoses in sexually abused children is similar to that used in nonabused populations. The issue of determining whether or not a particular child has experienced sexual abuse is a complex medical, psychiatric, and legal process, which is beyond the scope of the present article. In addition to determining DSM-IV diagnosis, there are several special areas of concern related to the sexual abuse experience that should also be addressed when assessing sexually abused children. These include possible legal involvement, child protective services interventions, medical issues related to the abuse, and other abuse-related issues of the child and family. Factors such as the type of sexual abuse experienced, how long it was going on, whether the parent(s) or other siblings also experienced sexual abuse, and the identity of the perpetrator may have direct effects on the child's and/or family's emotional reactions, attitudes, and concerns about the abuse. Some families may be most upset because the abuse involved actual intercourse, whereas to other families this may be less important than the fact that the abuse was perpetrated by a trusted family member. This information is helpful in determining appropriate therapeutic interventions.

The child's and family's perceptions of why the sexual abuse occurred are also important. Does the family blame the perpetrator, the child, the parents, or some combination? The child's symptoms may be more closely related to these attributions than to the specific details of the abuse itself. Specifically, children who blame themselves rather than the perpetrator may be more likely to feel guilt, shame, or poor self-esteem; children who blame the abuse on the unpredictability of the world may have more difficulty with fear and anxiety symptoms. Similarly, parents who blame the abuse in part on the child's behavior or appearance may be less likely to be supportive of the child, or even less likely to view the experience as abuse.

The parent's emotional reaction to their child's sexual abuse also has an effect on child symptomatology. Low levels of maternal support and maternal depression correlate with greater behavioral and emotional symptoms in sexually abused children. Higher levels of parental distress related to the child's sexual abuse have also been found to predict higher levels of behavioral problems in the child.

Because of the complex legal and child protective issues involved in cases of child sexual abuse, it is essential for the clinician to clearly define his or her role in the case. There are many differences between treatment and forensic roles, and these should be kept separate and distinct as much as possible. It

is regarded as a conflict of interest to evaluate a child for credibility of sexual abuse allegations or to evaluate a child and parents for custody determination following sexual abuse, and then to provide ongoing treatment for that child or family.

Treatment

Several randomized controlled treatment studies for sexually abused children have been published since 1990. At this time, trauma-focused cognitive-behavioral therapy (TF-CBT) has the strongest empirical support for treating this population. TF-CBT is a hybrid treatment model incorporating cognitive-behavioral principles with trauma-sensitive interventions along with family and developmental theory. TF-CBT includes the following treatment components, summarized by the acronym "PRACTICE:" Parenting skills (psychoeducation); Relaxation skills; Affective modulation skills; Cognitive coping skills; Trauma narrative and cognitive processing of the sexual abuse or other traumatic experiences; In vivo desensitization to trauma reminders; Conjoint child-parent treatment sessions; and Enhancing safety and future developmental trajectory.

Deblinger et al. evaluated the effectiveness of TF-CBT in decreasing PTSD and other psychological difficulties in 100 sexually abused children. Subjects were randomly assigned to one of four treatment conditions: child treatment only, mother treatment only, mother and child treatment, or usual community care. Children who received the TF-CBT treatment exhibited significantly fewer PTSD symptoms than children who did not receive this treatment. Children whose mothers received TF-CBT exhibited fewer externalizing symptoms and depressive symptoms compared to children whose parents did not receive TF-CBT.

Cohen and Mannarino compared TF-CBT to non-directive supportive therapy (NST) provided during individual sessions to sexually abused preschoolers and their primary caretakers. Children receiving TF-CBT showed more symptomatic improvement than those receiving NST. This differential treatment response was particularly strong with regard to children's sexually inappropriate behavior. Follow-up assessments indicated that these treatment group differences were sustained over the course of 12 months. This study also evaluated the impact of several child and family factors on treatment outcome. The strongest predictor of outcome was the parent's emotional reaction to the sexual abuse. Specifically, higher levels of parental distress at pretreatment strongly predicted higher levels of the child's post-treatment symptomatology.

In a parallel study for sexually abused children aged 8–14 years, TF-CBT was superior to NST in improving depressive symptoms and social competency. At 12 month follow-up, TF-CBT completers experienced a significantly greater improvement in PTSD and dissociation than noncompleters. Intent-to-treat analyses indicated significant group $\times$ time effects in favor of the TF-CBT group on measures of depression, anxiety, and sexual problems. In these older children, a higher level of parental support was a strong predictor of positive treatment outcome.

In a large multisite study, Cohen et al. randomly assigned 229 8- to 14-year-old sexually abused children who had experienced multiple traumas. Over 90% of this cohort met full criteria for PTSD diagnosis at pretreatment. Children were randomly “assigned” to TF-CBT or to child-centered therapy (CCT) along with their nonoffending caretaking parent. This study demonstrated that even for these multiply traumatized children, TF-CBT was superior to CCT with regard to improvement in PTSD, depression, anxiety, shame, behavior problems, and abuse-related attributions. Parents receiving TF-CBT also experienced significantly greater improvement in their own depressive symptoms and positive parenting than those receiving CCT. Thus, there is evidence that trauma-focused CBT interventions are efficacious in decreasing symptomatology in multiply traumatized sexually abused children. All of the above studies also lend support to the importance of including nonoffending parents in treatment. More research is needed to further assess the efficacy of CBT and other potentially effective interventions and to clarify what components of these treatments may be most effective for which subgroups of children. Such TF-CBT treatment interventions may be provided either individually or in a group setting.

Other CBT-related approaches have been developed and are currently being tested for adolescents experiencing chronic PTSD related to sexual abuse. For example, the SPARCS (Structured Psychotherapy for Adolescents Recovering from Chronic Stress) and STAIR (Skills Training in Affective and Interpersonal Regulation) models show promise in this regard.

Other approaches to treating sexually abused children have received less empirical evaluation. One study has evaluated the efficacy of psychodynamic therapy for this population. Trowell and colleagues randomly assigned sexually abused children to 18 sessions of psychoeducation groups or 30 sessions of individual psychoanalytic therapy. The psychoanalytic group experienced greater improvement in PTSD symptoms, but the design of this study makes it difficult to determine whether this differential response was due to treatment type (psychoanalytic vs.

psychoeducation), modality (individual vs. group), or dosage (30 sessions vs. 18). There are probably many efficacious therapeutic techniques for sexually abused children. Follow-up studies have indicated that many sexually abused children recover fully from this stressful experience and do not suffer ongoing negative sequelae. Hopefully, future research will more clearly elucidate which treatment approaches are most likely to result in this positive outcome.

See Also the Following Articles

Child Abuse; Child Physical Abuse; Posttraumatic Stress Disorder in Children; Nightmares.

Further Reading

- Brent, D. A., Oquendo, M., Birmaher, B., et al. (2002). Familial pathways to early-onset suicide attempts: Risk for suicidal behavior in offspring of mood-disordered suicide attempters. *Archives of General Psychiatry* 59, 801–807.
- Briere, J. (1995). *The Trauma Symptom Checklist for Children*. Odessa, FL: Psychological Assessment Resources.
- Cohen, J. A. and Mannarino, A. P. (1993). A treatment model for sexually abused preschool children. *Journal of Interpersonal Violence* 8(1), 115–131.
- Cohen, J. A. and Mannarino, A. P. (1996). A treatment outcome study for sexually abused preschool children: initial findings. *Journal of the American Academy of Child and Adolescent Psychiatry* 35(1), 42–60.
- Cohen, J. A. and Mannarino, A. P. (1998). Interventions for sexually abused children: initial treatment outcome findings. *Child Maltreatment* 3(1), 17–26.
- Cohen, J. A., Deblinger, E., Mannarino, A. P. and Steer, R. (2004). A multisite, randomized controlled trial for sexually abused children with PTSD symptoms. *Journal of the American Academy of Child and Adolescent Psychiatry* 43, 393–402.
- Cohen, J. A., Mannarino, A. P. and Knudsen, K. (2005). Treating sexually abused children: one year follow-up of a randomized controlled trial. *Child Abuse and Neglect*.
- Cohen, J. A., Mannarino, A. P. and Deblinger, E. (2006). *Treating trauma and traumatic grief in children and adolescents*. New York: Guilford Press.
- Deblinger, E. and Heflin, A. H. (1996). *Treating sexually abused children and their nonoffending parents: a cognitive behavioral approach*. Sage: Newbury Park, CA.
- Deblinger, E., Lippmann, J. and Steer, R. (1996). Sexually abused children suffering posttraumatic stress symptoms: initial treatment outcome findings. *Child Maltreatment* 1(4), 104–114.
- Finkelhor, D. (1987). The trauma of child sexual abuse: two models. *Journal of Interpersonal Violence* 2(4), 348–366.
- Finkelhor, D. and Barron, L. (1986). Risk factors for sexual abuse. *Journal of Interpersonal Violence* 1(1), 43–71.
- Friedrich, W. N. (1998). *The Child Sexual Behavior Inventory*. Odessa, FL: Psychological Assessment Resources.

- Mannarino, A. P. and Cohen, J. A. (1996). Family related variables and psychological symptom formation in sexually abused girls. *Journal of Child Sex Abuse* 5(1), 105–119.
- McLeer, S. V., Callaghan, M., Henry, D. and Wallen, J. (1994). Psychiatric disorders in sexually abused children. *Journal of the American Academy of Child and Adolescent Psychiatry* 33(3), 313–319.
- Nelson, E. C., Heath, A. C., Madden, P. A. F., et al. (2002). Association between self-reported childhood sexual

abuse and adverse psychosocial outcomes. *Archives of General Psychiatry* 59, 139–145.

Trowell, J., Kolvin, I., Weeramanthri, T., Sadowski, H., Berelowitz, M., Glasser, D. and Leitch, I. (2002). Psychotherapy for sexually abused girls: psychopathological outcome findings and patterns of change. *British Journal of Psychiatry* 160, 234–247.

Wolfe, V. V., Gentile, C. and Wolfe, D. A. (1989). The impact of child sexual abuse on children: a PTSD formulation. *Behavioral Therapy* 20, 215–228.

Childbirth and Stress

S Ayers and E Ford

University of Sussex, Brighton, UK

© 2007 Elsevier Inc. All rights reserved.

Introduction

Physical Outcome

Maternal Satisfaction

Psychopathology

Conclusion

Glossary

<i>Baby blues</i>	A short period of mood swings and emotional reactivity in the 2 weeks following childbirth.
<i>Epidural anesthesia</i>	Analgesia injected into the space around the spinal cord to block pain in the abdomen.
<i>Episiotomy</i>	A cut made to the woman's perineum to help the baby be delivered.
<i>Intrapartum Oxytocin</i>	During labor. A hormone used to induce labor and stimulate contractions of the uterus.
<i>Parity</i>	Number of live children born to a woman.
<i>Polymorphic (cycloid) psychosis</i>	Psychosis with rapidly changing and variable psychotic symptoms and emotional states as a cardinal feature.
<i>Postnatal depression</i>	Depressive disorder in women in the year following childbirth.
<i>Postpartum</i>	After childbirth.
<i>Prophylaxis</i>	Measures taken to prevent a disease.
<i>Puerperal</i>	Pertaining to childbirth.

Introduction

Childbirth is an intense event that involves extreme physical stress and carries emotional, cognitive,

social, and cultural significance. Physically, women have to cope with acute changes as the uterus contracts, the cervix dilates, and the baby and placenta are delivered. General responses to stress such as release of endorphins, epinephrine, and catecholamines have been shown in laboring women, as well as increases in cardiac output and blood pressure. Emotionally, labor and the birth of the baby usually involve intense positive and negative emotions. Cognitive demands are also placed on the woman as she copes with the events of birth. Interpersonal dynamics between the woman and her birth partner and medical attendants can be supportive or can increase stress levels if birth attendants are perceived as unhelpful or dismissive. Finally, birth and motherhood are associated with many cultural expectations and perceived norms. In Western societies, for example, birth is highly medicalized, yet often emphasis is placed on women trying to achieve a so-called natural birth with minimal intervention or pain relief.

Childbirth is therefore an intense acute stressor that many women go through. It also enables the study of the interaction between pre-event and event factors in perceived stress and the effect of this on physical and psychological outcomes. This is consistent with the diathesis stress approach to health outcome, in which pre-existing vulnerability (physical, psychological, and social) interacts with event characteristics to determine outcomes. However, the classification of childbirth as an acute stressor is rather artificial, and the impact of childbirth must be placed in the context of previous adjustment to pregnancy and postnatal adjustment to motherhood and a new baby, both of which are also potentially stressful and difficult.

Research into the role of psychosocial factors in birth outcome has generally established that

psychosocial factors are important in physical outcomes, such as morbidity and mortality, and psychological outcomes, such as maternal satisfaction and psychopathology. The postnatal period is associated with a number of psychological problems that are seen as specific to childbirth, namely, the baby blues, postnatal depression, and puerperal psychosis. However, it is still debated whether childbirth has specific effects that distinguish postnatal psychopathology from psychopathology occurring at other times. More recently, childbirth has also been linked with posttraumatic stress disorder.

The experience of birth and the outcome of birth are therefore influenced by multiple psychosocial factors. The importance of these factors varies for different outcomes and between individuals. For example, a woman with a history of psychological problems may have a high vulnerability to postnatal psychopathology, so other factors during and after birth may not be as important as they are for a woman with no previous history of psychological problems. This is one of the challenges of research in this area, and it is therefore an area characterized by conflicting findings. It is thus important to consider methodological quality of studies and weight of evidence when drawing conclusions. This article provides a brief overview of some of the current research and understanding of psychosocial factors that are important in three broad areas of birth outcome: physical outcome, maternal satisfaction, and psychopathology; then it looks generally at what conclusions can be drawn regarding childbirth and stress.

Physical Outcome

The physical outcome of birth includes complications during birth, type of labor and delivery, and complications with the infant. As in other areas of health, low socioeconomic status is associated with increased complications of birth (e.g., pre-term labor) and maternal and infant morbidity (e.g., low birth weight) and mortality. It is hypothesized that the effect of low socioeconomic status is due to material deprivation but is mediated by psychosocial factors such as level of education, access to information, a lack of social support, and negative life events. These in turn may lead to stress, anxiety, lowered self-esteem, and depression, with cognitive problems such as dysfunctional attitudes, which result in a set of maladaptive coping strategies such as learned helplessness and risk taking.

It is well established that the physical outcome of birth can be affected by psychosocial factors during pregnancy and birth such as support or anxiety. In the literature, continuous support during labor has

received the most attention. A series of experimental studies has been carried out in countries where it is usual for women to give birth without a birth companion. These studies compare women allocated a doula during birth to women who have no birth companion. A doula is a layperson (usually female) trained in supporting laboring women, but not medically qualified. A review of 15 doula studies involving over 12 000 women showed that women with the continuous support of a birth companion used less analgesia during birth, were less likely to have an operative birth, and were less likely to report dissatisfaction with their birth experiences. In general, continuous support during birth was associated with greater benefits when the provider was not a member of the hospital staff, when the support began early in labor, and in settings in which epidural analgesia was not routinely available.

A number of explanations have been put forward for why support in labor has such a powerful effect on birth outcome, and these typically focus on biological or psychological mechanisms. From a biological perspective it has been hypothesized that the complicated neuroendocrinal control of labor and birth can be disrupted by stress in various ways, reducing the efficacy of the birthing process. For example, it has been posited that labor pain leads to a strong stress response, resulting in increased autonomic activity, uncoordinated uterine contractions, and abnormal fetal heart rate, which all impede the progress of labor. Whether this effect is a direct result of the pain or is mediated by stress or anxiety is not established.

Psychological explanations suggest that support in labor improves outcomes by increasing women's feelings of control and competence, decreasing the stress response, and thereby reducing the necessity for medical interventions. Emotional support may also reduce fear and anxiety and their associated effects.

Maternal Satisfaction

Research looking at maternal satisfaction has largely focused on women's satisfaction with the birth generally and has not looked at specific elements such as satisfaction with coping, birth outcome, and health-care provision. The main factors during birth that have been examined in relation to maternal satisfaction are support during birth, pain relief, perceived control, and levels of intervention during birth. For example, as previously discussed, continuous support during labor is associated with increased maternal satisfaction, as is continuity of care. It was found that women assigned to a one-to-one midwifery care model, who were cared for by the same caregivers throughout their pregnancy and birth, reported

significantly more satisfaction with their experience than women assigned to obstetrician-led care who met different caregivers at each visit, without any difference in clinical outcome.

Most research on satisfaction with the birth experience has focused on pain relief, in the belief that if the pain is controlled or alleviated the experience of giving birth will be more positive. This appears to be the case, with studies showing that women who receive more effective pain relief rate their experience as more satisfying. However, most studies have shown that satisfaction is high overall, regardless of pain relief received, suggesting that many other factors are also important. A woman's expectation of the pain may play a role, with research showing that women who expect more pain are likely to be more satisfied afterward. Furthermore, if women expect to be able to give birth without pain relief, they report a decline in satisfaction if they did have to use analgesia, notably epidural anesthesia.

Other factors associated with maternal satisfaction are personal control, having expectations met, and social class. Extensive medical interventions (electronic fetal monitors, episiotomies, perineal shaves, and enemas) are negatively correlated with maternal satisfaction. However, active management of labor (e.g., early rupturing of membranes, vaginal assessments, and early use of oxytocin for slow progress) does not affect women's satisfaction overall compared to routine management.

A systematic review concluded that expectations, support, the quality of the caregiver-patient relationship, and involvement in decision making (control) are the most important factors in determining maternal satisfaction. Secondary factors are pain, socioeconomic status, preparation for birth, interventions, and continuity of care.

Psychopathology

Baby Blues

The baby blues are a period of emotional lability and overreactivity that appears in the first 2 weeks following birth. The blues do not seem to be part of a continuum with postpartum depression or puerperal psychosis, and sadness and depression are not typical features. Instead, women seem to exhibit mood fluctuations such as elation, hostility, tearfulness, and general lability of affect. Some women who experience severe blues will go on to develop postnatal depression. The blues affect between 26 and 85% of women, the variability in figures being due to the fact that the baby blues are not classified as a psychiatric disorder and therefore not uniformly defined.

Several competing theories have been proposed to account for the blues, most notably a hormonal/biological account, suggesting that the large fluctuations in hormones following delivery are responsible. The changes in concentrations of hormones such as estrogen, progesterone, and cortisol are hypothesized to cause a neurotransmitter imbalance, which is expressed by widely fluctuating moods. It has been suggested that this could be evolutionarily adaptive, as this excessive emotionality separates the mother from her usual concerns and makes her available solely for the baby, facilitating mother-baby bonding. Historically, it may also have been responsible for promoting some of the rituals that surround the birth of a baby, such as extensive periods of confinement and good support for the mother. However, there is no consistent evidence linking changes in hormone levels with fluctuating postnatal mood. Other theories suggest that psychosocial factors are important. While demographic factors such as age and race are not predictive of the baby blues, in the West the blues have been associated with economic insecurity and lack of social support. It has also been proposed that it may be a grief reaction in women who have been disappointed with pregnancy and birth events and outcome, and has been shown to be correlated to pessimism in late pregnancy.

The baby blues appear to be time limited and require no specific treatment other than support. However, the blues should be distinguished from depression and sadness in the immediate postpartum period, which is predictive of postnatal depression later.

Postnatal Depression

Postnatal depression (PND) is the most well known of postpartum psychiatric disorders. Between 10 and 20% of women are believed to suffer from it during the first year after birth. However, the disease still remains underdiagnosed and undertreated. Recently, controlled studies have shown there is no significant difference between postnatal women and appropriate female controls in the incidence of depression. Therefore, postnatal women are no more at risk of developing depression than at other times in their lives. However, it is argued that PND has a greater psychosocial impact than depression at other times in a woman's life, as it can cause considerable strain in relationships and family life, affect the mother-infant relationship, and possibly also affect the infant's development. Currently self-report scales are used to screen for depression in postnatal women. These screening tools can be a useful measure for identifying women at risk, although patients who score above

threshold on screening questionnaires are heterogeneous and may be suffering from anxiety, obsessive, or posttraumatic stress disorders.

Like the baby blues, postnatal depression has also been postulated to be a result of a hormone imbalance in the period following birth. A large drop in levels of estrogen and progesterone occurs after birth, suggesting that postpartum mood changes can be explained by the effect of the withdrawal of these hormones. This has been supported by studies in which women were administered these hormones in the postpartum period, resulting in alleviation of the depression. A further hypothesis is that hormonal changes trigger pathological mechanisms at the neurotransmitter level, resulting in depression. However, many studies measuring levels of hormones in women after birth have found no differences in hormone concentrations between women with and without depression, suggesting that the link between hormone levels and depression is weak if not nonexistent.

It has been argued that because prevalence of PND is no greater than depression in women in general, PND should be seen as primarily a psychosocial disorder. Studies have shown associations with previous depression and other psychiatric illness, stressful life events, and disturbed relationships. Other predictors include high parity, stress in pregnancy, social isolation, material deprivation, and low social support. A history of depression, and depression during pregnancy, appear to be some of the strongest predictors of PND. Furthermore, genetic factors have been shown to explain 25–38% of the variance. The psychosocial model would suggest that good social support following childbirth would protect women from depression, and studies seem to bear this out. If, as suggested, prevalence of depression postnatally is not greater than prevalence in women of childbearing age in general, and if PND is related to similar factors to depression in nonpostnatal women, it may be hypothesized that PND is not a separate illness from depression in general.

Puerperal Psychosis

Puerperal psychosis (PP) is a rare disorder occurring in about one woman in a thousand following childbirth. It generally takes the form of mania, severe depression (with delusions, confusion, or stupor), or acute polymorphic (cycloid) psychosis. It does not overlap completely with bipolar mood disorder, although there is evidence of a link between the two, especially as the majority of patients presenting with a severe postpartum psychiatric disorder appear to have a cycloid psychosis. It has been proposed that PP may be one of a spectrum of bipolar disorders. In susceptible women, childbirth, abortion, and even

menstruation can trigger bipolar episodes. However, other researchers propose that PP has a pattern of symptoms unlike any other psychiatric disorder; for example, patients exhibit a high rate of disorientation, confusion, and depersonalization. Hallucinations appear in about half of patients with PP, suggesting an overlap with schizophrenia, and mania is not as pronounced in postpartum patients as in those with a diagnosis of bipolar disorder. Thus, some argue that PP is a separate disorder from other psychoses.

Puerperal psychosis appears to be highly heritable, with 10–25% of relatives of women with PP having some kind of severe psychiatric disorder. It also tends to recur, with a chance of about 1 in 4 in each pregnancy. It can be treated safely with newer neuroleptic drugs. In certain circumstances electroconvulsive therapy can be useful, and lithium has been explored as a prophylaxis.

Postnatal Posttraumatic Stress Disorder

Posttraumatic stress disorder (PTSD) is characterized by three types of symptoms: re-experiencing (flashbacks, nightmares), avoidance (of reminders of the stressor, emotional numbing), and hyperarousal (insomnia, difficulties concentrating). There is converging evidence that approximately one-third of women appraise their birth experience as traumatic and 1–2% of women will develop PTSD after childbirth. Case studies suggest that the disorder might result in an impaired mother–baby relationship, sexual dysfunction, and fear of subsequent childbirth (tokophobia), although this has not yet been confirmed by quantitative research.

Studies of factors associated with postnatal PTSD suggest an interaction between prenatal factors, birth experience, and postnatal factors. The type of delivery appears to be important, with women who undergo operative deliveries more at risk of postnatal PTSD. However, this is not a consistent finding. Indeed, one study noted that the majority of women with PTSD had an obstetrically normal vaginal delivery. Thus, in common with PTSD following other events, there is unlikely to be a direct relationship between the severity of birth and PTSD. Other labor and birth characteristics variously associated with rating birth as traumatic are medical intervention, pain in first stage, a long labor, feelings of powerlessness, and receiving inadequate information. PTSD has been associated with feelings of fear for self and baby during labor and birth. It has also been found that low control, higher trait anxiety, and greater fear were associated with symptoms of PTSD.

Several prenatal vulnerability factors have been investigated, and those that emerge as important are a history of trauma or psychological problems and

possibly trait anxiety. Dysfunctional attitudes are predictive of postpartum depressive symptoms, and it has been hypothesized that there are cognitive vulnerabilities to PTSD as well, although this has not yet been examined in relation to postnatal PTSD. Postnatally, social and professional support in the period after birth may be a protective factor, and aftercare trauma services might be useful. Several studies have analyzed the effectiveness of debriefing services with a midwife or doctor following birth in preventing PTSD, but results have been inconsistent.

Conclusion

The association between different psychosocial factors and the outcomes discussed in this article demonstrates the complexity of this field. In summary, physical outcomes are associated with socioeconomic status and continuous support in labor. Maternal satisfaction is associated with expectations, support in labor, the quality of caregiving in labor, and involvement in decision making. Postnatal depression is associated with previous history of depression or other psychiatric illness, additional stressors, and low social support. Puerperal psychosis is associated with a personal history or family history of psychiatric illness. Postnatal PTSD is associated with a history of trauma and psychiatric illness, type of delivery, and life threat.

Models of the interaction between prenatal, birth, and postnatal factors therefore tend to be complex psychosocial models. However, it could be argued that models specific to childbirth are not necessary, as, in many cases, the determinants of postnatal psychopathology are similar to psychopathology at other times. For example, the finding that postnatal depression is associated with previous depression, additional stressors, and low social support is similar to research findings regarding vulnerability to depression generally. Equally, the finding that there is no dose-response relationship between the obstetric severity of birth and postnatal PTSD is comparable to research into PTSD following other events. This etiological similarity between postnatal psychopathology and psychopathology at other times implies that stress and coping processes similar to those in other types of life events act to determine the outcome of childbirth. Therefore, childbirth can be used as a research paradigm with which to study stress processes and outcomes that generalize to other acute life events. As such, it may also help us better understand the complexity of stress processes.

See Also the Following Article

Depression and Manic-Depressive Illness.

Further Reading

- Ayers, S. (2004). Delivery as a traumatic event: prevalence, risk factors, screening and treatment. *Clinical Obstetrics and Gynecology* 47(3), 552-567.
- Brockington, I. (2004). Postpartum psychiatric disorders. *The Lancet* 363, 303-310.
- Brownridge, P. (1995). The nature and consequences of childbirth pain. *European Journal of Obstetrics and Gynaecology and Reproductive Biology* 59(supplement), S9-S15.
- Clement, S. and Page, L. (eds.) (1998). *Psychological perspectives on pregnancy and childbirth*. Edinburgh, UK: Churchill Livingstone.
- Elliott, S. A. (2006). Postnatal depression. In: Ayers, A., Baum, A., Newman, S., McManus, C., Wallston, K., Weinman, J. & West, R. (eds.) *Cambridge handbook of psychology, health and medicine* (2nd edn.). Cambridge, UK: Cambridge University Press.
- Goodman, P., Mackay, M., et al. (2004). Factors related to childbirth satisfaction. *Journal of Advanced Nursing* 46(2), 212-219.
- Hodnett, E., Gates, S., et al. (2003). Continuous support for women during childbirth. *The Cochrane Database of Systematic Reviews* 3, CD003766.
- Hodnett, E. D. (2002). Pain and women's satisfaction with the experience of childbirth: a systematic review. *American Journal of Obstetrics and Gynecology* 186(5, part 2), S160-S172.
- Kyman, W. (1991). Maternal satisfaction with the birth experience. *Journal of Social Behaviour and Personality* 6(1), 57-70.
- Quine, L. and Steadman, L. (2006). Pregnancy and childbirth. In: Ayers, A., Baum, A., Newman, S., McManus, C., Wallston, K., Weinman, J. & West, R. (eds.) *Cambridge handbook of psychology, health and medicine* (2nd edn.). Cambridge, UK: Cambridge University Press.
- Robertson, E., Grace, S., et al. (2004). Antenatal risk factors for postpartum depression: a synthesis of recent literature. *General Hospital Psychiatry* 26(4), 289-295.
- Rutter, D. R. and Quine, L. (1990). Inequalities in pregnancy outcome: A review of psychosocial and behavioural mediators. *Social Science and Medicine* 30(5), 553-568.
- Slade, P., MacPherson, S., et al. (1993). Expectations, experiences and satisfaction with labour. *British Journal of Clinical Psychology* 32, 469-483.
- Stanton, A. L., Lobel, M., Sears, S. and DeLuca, R. S. (2002). Psychosocial aspects of selected issues in women's reproductive health: current status and future directions. *Journal of Consulting and Clinical Psychology* 70(3), 751-770.

Childhood Stress

S Sandberg

University College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S Sandberg, volume 1, pp 442–449, © 2000, Elsevier Inc.

Definitions of Stress

Measurement of Stress

Effects of Stress

Glossary

<i>Contextual threat</i>	The level of threat caused or implied by the life event to an average child of the same age, sex, and biographical characteristics as the child in question.
<i>Life event</i>	Any event or circumstance occurring in the life of a child that may have the potential to alter his or her present state of mental or physical health.
<i>Limbic kindling</i>	A psychophysiological process in which the limbic system (a border between the new and the old brain, forming a circuit primarily concerned with emotions, memory, olfaction, and neuroendocrine control) receives electrical stimulation resulting in increased responsiveness to low levels of stimulation over time. In this way, repeated traumatization, as in child abuse, or a trauma followed by intrusive memories could kindle limbic nuclei, leading to behavioral changes.
<i>Long-term psychosocial experience</i>	Ongoing process in which the delineation of a beginning and end stage is not always possible. The experience may be adverse (e.g., parents' marital disharmony or financial hardship), desirable or hedonically positive (e.g., a rewarding hobby), or socially required for normal development or for protection against negative experiences (e.g., supportive relationship with a parent or having a close friend).
<i>Modifier</i>	A variable, aspect, or characteristic that exerts modifier effects on risk. If the risk is intensified, then the variable carries a vulnerability effect; if the risk is diminished, then the variable carries a protective effect. A vulnerability/protective factor is a moderator and is demonstrated statistically with a significant interaction effect.

Negative life event

A life event exerting a particular undesirable impact on the child.

Positive life event

A life event that is hedonically positive, enhances self-esteem, or has a beneficial effect on the child's life circumstances.

Risk factor

A variable that negatively influences outcome regardless of the presence or absence of adversity.

Stress, stressor, and stress reaction

A form of stimulus, an inferred inner state, and an observable response reaction, respectively.

Definitions of Stress

Stress, Stressor, and Stress Reaction

Child stress can be defined as any intrusion (stressor) into children's normal physical or psychosocial life experiences that acutely or chronically unbalances physiological or psychological equilibrium, threatens security or safety, or distorts physical or psychological growth and development and the psychophysiological consequences (stress reaction) of such an intrusion or distortion.

Harmful vs. Beneficial Stress

The traditional view of stress as a significant discrepancy between the demands of a situation and the organism's capacity to respond, resulting in detrimental consequences, implies that all stress is harmful. However, as far as children are concerned, the opposite is sometimes true. Mild stressors are often benign and sometimes even growth promoting.

Additive and Synergistic Stress

Apart from their individual effects, stressors can also act additively and synergistically. For example, deprivation and malnutrition can each compound fetal alcohol effects. Even the same cause can stress the child through two or more different mechanisms. Thus, maternal alcoholism can stress the fetal brain chemically and then psychologically stress the newborn infant through maternal unavailability, neglect, or abuse. Poverty can stress the child through malnutrition, inadequate environmental stimulation, and poor parenting due to worry, preoccupation, and discouragement.

Measurement of Stress

Chronic Stress

There exist a variety of questionnaire and interview instruments for the assessment of chronic stress in

children, usually stemming from enduring adverse circumstances such as environmental or socioeconomic deprivation, parental unemployment or illness, and family discord or criminality. Also, the two widely used diagnostic classification systems, the *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (DSM-IV) and the *International Classification of Diseases* (ICD-10), contain a separate axis for systematic measurement of stressors of likely relevance for children's health.

Life Events

Questionnaire measures The first-generation measures of stress in childhood were based on the questionnaire format to provide overall scores of degree of life change. The assumption was that change per se, not necessarily the unpleasant nature of the experience, was stressful for children. In parallel with these overall approaches, numerous studies of specific life events such as family breakup, bereavement, and disasters such as floods, earthquakes, or hijacking were carried out.

Interview measures

The meaning of life experiences The realization that it was necessary to take into account the social context of life events in order to assess their meaning and hence their stressful quality constituted perhaps the most significant conceptual advance. The clear implication of this advance was that interviews rather than questionnaire approaches were needed for adequate assessment of life events.

Undesirability, cognitive appraisal, and independence A major reappraisal of psychosocial stress research involved the emphasis on the need to differentiate between desirable and undesirable life changes and to highlight the role of cognitive appraisal of the events. The distinction between life events that could have been brought about by the person's own actions and those that are independent of behavior constitutes another conceptual advance. This distinction is especially important when considering stressful life events as causal agents for psychiatric disorders.

Contextual threat The concept of contextual threat has represented a decisive step in the assessment of children's life events. In determining the stressful quality of life events, it is necessary to take into account their social context, i.e., their meaning to an individual. For example, it is unreasonable to treat moving as a unitary event. The effects would be very different for a child who moved within the same neighborhood and maintained the same school and

friends compared to one whose new home was in another town, or even another country, resulting in loss of friends, surroundings, and school. To understand this difference, interviews rather than questionnaire approaches are needed.

The assessment is usually made separately for short-term (based on immediate impact), and long-term (the likely impact occurring some 10 days later) threat. In addition, it has become customary to conceptualize the overall rating of contextual threat as being made up of one or more components that vary somewhat depending on the instrument used. These components commonly include the following: loss of an attachment figure (e.g., losing a parent or a close friend), loss of a valued idea (a major disappointment, letdown, or humiliation), risk of loss of attachment figure (the threat of losing someone close), physical jeopardy (being in physical danger due to accident, etc.), trauma as witness (being witness to an unpleasant or frightening incident), and psychological challenge (the event involving the subject in a new role or with new responsibilities).

Some life events encountered by children rate high on contextual threat because they cause a major alteration in the life circumstances (e.g., the death of a parent). In other events, the threat is primarily cognitive. This means that the event may not result in any actual change in the external environment, but instead drastically changes the child's perception of an aspect of him- or herself, or of other people or things in a way that presents a threat to the child's self-esteem (e.g., severe humiliation) or reduces his or her perceived sense of security (e.g., a parent threatening to abandon the child). Life events may also involve a combination of real life change and cognitively mediated threat (e.g., the parents' marital separation).

Effects of Stress

Developmental Interferences

Physical development The ill effects of compromised prenatal, perinatal, and early postnatal care on the child's development and well-being have long been recognized. For example, in the multiple associations between prematurity, illness, and outcome, environmental factors are known to play a central role. Likewise, children suffering from the nonorganic failure to thrive syndrome frequently exhibit a complex entanglement of caloric and protein intake, touch, and the release of growth hormone.

Neurodevelopment

Neuroanatomy Stress can also influence the development of the brain. To develop properly,

undifferentiated neuronal systems are dependent on environmental cues such as neurochemical and hormonal factors. These processes in turn are dependent on the child's sensory experiences. A narrow window appears to exist during which specific experiences are required for the optimal development of the brain. After birth, brain development mainly consists of an ongoing process of wiring and rewiring. Early experiences can, for example, cause the final number of synapses to increase or decrease. Infants confined to their cots and with little opportunity to explore or experiment, and who are rarely spoken to, may fail to develop neural pathways that facilitate later learning.

Neurotransmitters and hormones play a major role in neuronal migration, differentiation, and synaptic proliferation. Massive increases in catecholamine activity caused by severe or prolonged stressors are expected to have a significant impact on brain development. The hippocampus is a likely target for the effects of early stress because neurogenesis in this area continues into postnatal life. The amygdaloid nuclei are critically involved in the formation of emotional memory and in stress-induced memory enhancement. Repeated trauma may lead to limbic kindling. Thus, during critical periods in early childhood, experiences can organize brain systems; nurture may become nature. This would also suggest that it is more appropriate to speak of early childhood malleability than resilience.

Primate research has presented strong evidence that chronic unpredictable stress during pregnancy may have long-lasting effects on brain activity and behavior of the offspring. Furthermore, prenatal stress appears to be capable of affecting the neurobiology of the neonate so that its responses to common events and stimuli are markedly altered. These effects on brain mechanisms are more likely to be diffuse, rather than specific. As a consequence, the neonate may, for example, be more fussy and difficult, affecting how the caregiver is able to cope. Thus, prenatal stress appears to set the infant's neurobiology into a state in which vital attachments are hard to achieve and adverse experiences are more likely. The adverse experiences can in turn result in further neurobiological deflection of development.

Prenatal effects on fetal neurobiological development may also be easily confused with genetic predisposition. It therefore remains a challenge to determine how prenatal environment contributes to persisting and traitlike organization of neurobiological systems. This, in turn, would markedly further our understanding of how genetic and postnatal environmental factors play a role in the expression of developmental psychopathology, for example.

Neurophysiology When a child is exposed to a strong stressor, a series of interactive, complicated neurophysiological reactions take place in the brain, the autonomic system, the hypothalamus-pituitary-adrenocortical axis, and the immunosystem, with concomitant release of stress hormones. Neurophysiological activation during acute stress is usually rapid and reversible. As such, it has an adaptive purpose. Following severe prolonged or repeated stress, however, these changes may no longer be reversible, and become maladaptive. One explanation for this is that the organisms have not yet evolved mechanisms to deal adaptively with chronic stress. Stress-induced sensitization may also occur, with the systems underlying the stress response becoming more sensitive to future adverse experiences. Children growing up in a persistently threatening environment develop stress response systems in midbrain and brain stem that are overreactive and hypersensitive. This development may be highly adaptive, but only as long as the child remains in a violent, chaotic environment. Profound cognitive disturbances can accompany this process.

Some young children, especially girls, however, react in the opposite way by freezing rather than by becoming visibly anxious or hyperactive. Flooding of adrenocortical hormones during periods of acute stress may also have negative consequences on memory and the anatomical substrates of memory.

Trauma response Traumatic events impact structures in the central nervous system – though the exact mechanism of this remains unknown at present. Structures underlying cognitive, affective, sensory, integrative, regulatory, neuroendocrine, and motor functions are coordinated in the life threat response. This may lead neuronal networks, excited during a life-threatening event, to become identified for later potentiation, forming the basis for permanent registration in memory.

In younger children whose brains are maturing in use-dependent interaction with their environment, aspects of such memory states may become lifelong traits. It has been suggested, for example, that children using a dissociative adaptive defense in an acute response to trauma would later demonstrate primarily dissociative or somatic symptoms. On the other hand, children who primarily use hyperarousal adaptation to an acute stressor would be more likely to develop chronic hyperarousal symptoms, such as startle response, anxiety, motor hyperactivity, sleep disturbance, or tachycardia.

Psychological development

Cognitive development Persistent environmental stress and deprivation can significantly impair the

development of the child's cognitive abilities, including attention skills, memory, and language as well as intellectual capabilities. Conversely, a child's level of cognitive development is an important determinant of how he or she will deal with trauma. Ability to label and express feelings, to receive reality-based feedback, and to incorporate the experience into a narrative are relevant developmental achievements.

Emotional development Because a child's most pressing emotional needs change over time, a particular trauma may have more influence during one stage of development than during another. The cognitive interpretation of the same trauma may also be different depending on the child's developmental level.

Effects on Physical Health

Numerous studies in the fields of pediatrics and psychosomatic medicine have shown that children with increased psychosocial stress are significantly more likely to experience illness and hospitalization as well as use health services more frequently than other children. Stress as a precipitating and/or provoking factor has been implicated for asthma, appendicitis, rheumatoid arthritis, and leukemia, for example. High levels of psychosocial stress have also been shown to predict greater morbidity in children who already have a chronic disease. For example, recent studies on children with asthma have demonstrated time-related associations between exposure to stress and acute worsening of symptoms.

The role of stress in viral infections has been the focus of research in more recent years. Well-controlled prospective and experimental studies have shown that adverse life events and other stressful experiences significantly increase a person's susceptibility to acute and recurrent respiratory infections. One likely explanation for this association is that stress compromises the body's immune responses, with individual differences in susceptibility possibly being accounted for by differences in psychobiological reactivity. It is acknowledged, for example, that the immune system and the central nervous system form a bidirectional communication network. In this network, pro-inflammatory cytokines play a critical role.

Effects on Mental Health

A range of research strategies have demonstrated beyond all reasonable doubt that psychosocial factors exercise causal effects on children's psychosocial functioning. They also have contributed in a major way to the multifactorial causation of psychiatric disorders in childhood. In this respect, the main risk

stems from chronic psychosocial adversity such as unhappy family life, parental psychopathology (including substance abuse and criminality), and parental stress caused by poverty or unemployment, or being a victim of bullying or other victimization. Acutely stressful life events, especially those that carry long-term threat to the psychological security of the child, add to the risk caused by chronic stressors, and in some instances constitute sufficient risk in their own right. Negative life events that are associated with chronic psychosocial adversity, rather than those that arise out of the blue, have also been shown to provoke an onset of psychiatric disorder, particularly depression, in children. Their main effect, however, appears to be to determine the timing of a new episode of disorder in a child already at risk due to chronic stress or made vulnerable through lack of emotional support.

Reasons for Individual Differences

Personal vulnerability Compared with adults, far less work has been carried out to examine the importance of personal vulnerability to stressful experiences in children. The role of close, supportive relationships, however, has received support. Thus, difficulties in the child's own friendships as well as their mother's lack of confiding relationships have been shown to predict depression and other emotional disorders. Likewise, the overall risk of psychiatric disorder in the presence of severely negative life events is significantly enhanced in children lacking a confiding relationship with a parent or a close adult relative.

Exposure to stressful experiences The very large individual differences in children experiencing stress and adversity are acknowledged. A variety of explanations are likely to apply. First, some stressful life events represent chance or acts of fate, including natural and man-made disasters such as war, floods, earthquakes, and shipping accidents. Second, stressful life events can also be a function of structural factors in the society, making negative experiences much more likely in some segments of the population than in others. In children, the effect of low social class on stressful life events appears to be an indirect one, via chronic psychosocial adversities such as family ill health and interparental conflicts, which in turn increase the likelihood of acute negative events. Likewise, family poverty is associated with multiple risk factors for children. Many of these risks are mediated via parental responses to such pressures. Being poor is linked with chronic psychological distress in parents, thus impeding effective parenting through harsh discipline, inconsistency, lack of monitoring, family

conflict, and multiple changes in family configuration. Third, many of the stressful life events occurring to children are a function of their parents' behavior in one way or another. This is the case with divorce, abuse, and a parent being arrested, for example. Such stressful, family-dependent life events tend to be rooted in parental psychopathology, substance abuse, or criminality.

Genetic factors are also implicated as bases of individual differences in environmental risk exposure – a phenomenon known as gene–environment correlation. The parents who pass their genes to their children are usually the same parents who provide the rearing experience. For example, antisocial adults are much more likely than other people to provide negative rearing environments to their children. Antisocial behavior in one or both parents in turn predicts a much increased rate of marital breakdown and negative parent–child interactions and, hence, life events that are stressful to children. Gene–environment correlation can also lead to environmental risk that impinges differently on the children.

Susceptibility There are large individual differences in children's responses to stressful life experiences. A major reason for this may lie in the crucial importance of the personal meaning of the particular experience involved. In addition, previous experiences are likely to color a person's reaction to a new experience. Stressful life events can be strengthening as well as weakening; they can steel as well as sensitize. Other factors determining individual differences are the child's age, sex, temperament, locus of control, and family support.

Several other explanations may account for the individual differences. One of these is gene–environment interaction, a concept introduced more recently. In this case, the vulnerability to stress is viewed as being genetically determined rather than acquired through experience.

Mechanisms

Additivity of stressful experiences There is no doubt that a single major life event such bereavement or kidnapping can be quite sufficient on its own to provoke psychiatric disorder. However, it also appears that there is a cumulative effect when several major life events occur together during a short period of time. Likewise, the multiplicative effect of the co-occurrence of several chronic adversities has been demonstrated in children.

Potentiation of stress Chronic adverse circumstances such as unhappy family life; parental psychopathology, substance abuse, or criminality; economic

hardship; or stressors in the school environment all in themselves increase the risk of child psychiatric disorder. Their presence also makes it more likely that acute stressful events will occur, often stemming directly from the ongoing adversity. Discord leading to divorce or criminality resulting in jail sentence, from the child's point of view, both involve separation from a parent. Apart from increasing the likelihood of stressful life events to occur, chronic adversity is also apt to potentiate their adverse effects.

Vulnerability and protective mechanisms There also exist complex vulnerability and protective mechanisms, i.e., those that increase and those that decrease stress effects. These include interactions, not just main effects. Such mechanisms need to be considered in relation to the child's individual characteristics as well as to their experiential properties.

A moderator is a variable that influences the strength or the direction of a relationship between a predictor variable and a criterion variable. An example is a finding that in the context of a child's chronic physical illness, family stress is negatively associated with child psychological adjustment. When the nature of the association is further examined, it may turn out that the effect is stronger (or weaker) in the presence of other contextual variables. For example, the strength or the direction of the relationship between stress and adjustment may depend on the type of coping used by the family. That is, a significant association may emerge only when a parent (or child) copes in a maladaptive manner.

A protective factor either ameliorates negative outcomes or promotes adaptive functioning. The protective factor serves its protective role only in the context of adversity; a protective factor does not operate in conditions of low adversity. Conversely, a vulnerability factor is a moderator that increases the chances for negative (maladaptive) outcomes in the presence of adversity. Similar to a protective factor, a vulnerability factor operates only in the context of adversity.

Mediating mechanisms There are a number of mechanisms through which psychosocial stressors lead to child disorder. Some acute life events bring about lasting adverse alterations in life circumstances. Examples include bereavement and parental divorce. Thus, for a child, the ill effects arising from the loss of a parent are highly dependent on whether the loss is associated with family discord or leads to serious impairment in parenting. Loss that does not have these adverse sequelae may be painful at the time, but usually remains of limited significance as a causal factor for persisting disorder.

Stress may also be associated with loss that is only in the person's mind. The mechanism of cognitive threat is a likely possibility. Young children, for example, appraise their situation and contemplate their life circumstances in ways that differ from those typical of adults.

Recurrent hospital admission is a known stressor in young children, with the key mediating mechanism being the disturbed parent-child relationship engendered by the child's behavior upon the return home. A comparable situation is the birth of a sibling. Here the disturbance of the prior parent-child relationship is more important, suggesting that the main stress for the child may reside in the parental reaction rather than in the child's response to the event as such.

Chain events With children, it is also important to consider mediating mechanisms in long-term consequences, i.e., how adverse early experiences influence later psychosocial functioning. The sequelae frequently depend on chain events that involve a number of different links reflecting separate psychological processes. These include learned maladaptive emotional responses, self-perpetuating behavior patterns or habits, lowered self-esteem influencing the ability to deal with future experiences, poor coping skills, internal working models of insecure relationships, or behavior that serves to shape later environments or increase the individual's exposure to stressful life events.

Positive experiences as protectors Compared with the volume of research that has been carried out on the effects of negative experiences on children's health, studies examining the role of positive life experiences are lacking. Furthermore, the meager evidence available so far fails to support the notion that hedonically positive events and life situations would have an ameliorating effect on psychiatric risk. The situation, however, may be different with regard to somatic manifestations of illness. Positive life events have been shown to offset the increased risk for new asthma exacerbations associated with severely negative life events.

Beneficial stress There are both beneficial and harmful stress experiences. The distinction is likely to lie in the child's temperamental qualities and in how the particular past experience has been coped with. This has implications for prevention. It is possible that successful prevention only partly lies in the avoidance of stressors, but largely lies in helping children to overcome negative experiences. The overcoming of stress can be strengthening, ensuring that

children encounter stressors at times and in ways that make it more likely that they will come out on top with a sense of accomplishment rather than with feelings of fear and humiliation.

Turning points Turning points in development is another issue likely to affect long-term sequelae of childhood stress. A turning point refers to an experience either internal (as with puberty) or external (as with dropping out of school) that carries with it the potential for enduring change because the experience either closes doors or opens up opportunities, or because it is associated with a lasting alteration in life circumstances. Most turning points concern very ordinary life experiences, but, in each case, the particular individual circumstances are crucial. Therefore, life events need to be conceived of not just in terms of their threat features, but also in terms of their effects on the person's social group and life opportunities.

Specificity of stressors vs. specificity of outcome It is conceivable that specific associations between particular stressors and particular outcomes are mediated by psychological, biological, or social processes and the specific nature of the stressor. Little evidence of this, however, exists so far. This may be because until recently very few studies testing full specificity designs (i.e., specific stressors linked to specific outcomes, via specific mediators, in the context of specific moderators) have been carried out in children. Across the various stressors such as exposure to poverty, illness, marital conflict, divorce, violence, or abuse, the only consistent evidence appears to be in relation to sexual abuse. Sexual abuse has been found to be associated specifically with internalizing problems, posttraumatic stress disorder, and sexual acting out.

The lack of specificity may have several explanations, including significant rates of co-occurrence and comorbidity of psychopathology in childhood, the lack of a well-established taxonomy of stressors, and the fact that specific characteristics and environmental contexts of the child moderate the associations between stressor and outcome in unique ways. Particular stressors may also be linked with particular outcomes only in the presence of particular moderating and mediating processes. Furthermore, stressors categorized by type may include a variety of stressors that are different in nature. It is known, for example, that poverty is associated with a range of additional stressful experiences, which include conflict (e.g., exposure to violence, physical abuse, divorce/marital conflict) and stressors involving loss (e.g., deaths, separations, disasters).

See Also the Following Articles

Child Abuse; Child Sexual Abuse; Child Physical Abuse; Posttraumatic Stress Disorder in Children.

Further Reading

- Arnold, L. E. (1990). *Childhood stress*. New York: Wiley.
- Friedman, R. J. and Chase-Lansdale, P. L. (2002). Chronic adversities. In: Rutter, M. & Taylor, E. (eds.) *Child and adolescent psychiatry* (4th edn., pp. 261–276). Oxford: Blackwell Science.
- Goodyer, I. (1990). *Life experiences, development and psychopathology*. Chichester, UK: Wiley.
- Grossman, A. W., Churchill, J. D., McKinney, B. C., et al. (2003). Experience effects on brain development: possible contributions to psychopathology. *Journal of Child Psychology and Psychiatry* **44**, 33–63.
- Maier, S. F. (2003). Bi-directional immune-brain communication: implications for understanding stress, pain, and cognition. *Brain, Behavior and Immunity* **17**, 69–85.

- McMahon, S. D., Grant, K. E., Compas, B. E., Thurm, A. E. and Ey, S. (2003). Stress and psychopathology in children and adolescents: is there evidence of specificity? *Journal of Child Psychology and Psychiatry* **44**, 107–133.
- Perry, B. D. (1996). *Maltreated children. brain development and the next generation*. New York: Norton.
- Sandberg, S. and Rutter, M. (2002). The role of acute life stresses. In: Rutter, M. & Taylor, E. (eds.) *Child and adolescent psychiatry* (4th edn., pp. 287–298). Oxford: Blackwell Science.
- Sandberg, S., Jarvenpaa, S., Penttinen, A., Paton, J. Y. and McCann, D. C. (2004). Asthma exacerbations in children immediately following stressful life events: a Cox's hierarchical regression. *Thorax* **59**, 1046–1051.
- Sandberg, S., Paton, J. Y., Ahola, S., McCann, D. C., McGuinness, D., et al. (2000). The role of acute and chronic stress in asthma attacks in children. *Lancet* **356**, 982–987.
- Shea, A., Walsh, C., MacMillan, H. and Steiner, M. (2004). Child maltreatment and HPA axis dysregulation: relationship to major depressive disorder and post traumatic disorder in females. *Psychoneuroendocrinology* **30**, 162–178.

Cholesterol and Lipoproteins

C M Stoney

National Center for Complementary and Alternative Medicine, National Institutes of Health, Bethesda, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by C M Stoney and M Finney, volume 1, pp 454–459, © 2000, Elsevier Inc.

Structure and Function of Cholesterol and Lipoproteins
Environmental Influences on Lipid Concentrations and Lipid Variability
Chronic Stress Mechanisms of Lipid Variability
Acute Stress Mechanisms of Lipid Variability
Interventions to Reduce Stress-Related Increases in Lipids
Clinical Implications of Variations in Lipids

Glossary

- Apolipoprotein* Proteins located on the surface of lipoproteins that transport proteins and bind to specialized enzymes.
- Cholesterol* A fat-soluble molecule in the body used to form cholic acid in the liver, steroid hormones in the gonads and adrenals, and used in the formation of cell membranes.
- Hemoconcentration* The generally transient and reversible filtration of fluid from the intravascular space.

Lipids

Chemical compounds in the body that are miscible with each other and which provide energy and a variety of intracellular functions. They include triglycerides, phospholipids, and cholesterol.

Lipoprotein

Large complexes of molecules that circulate and transport triglycerides, phospholipids, cholesterol, and proteins in the blood.

Lipid reactivity

Response of one or more of the lipids circulating in the blood to an external stimulus.

Structure and Function of Cholesterol and Lipoproteins

Lipoproteins are large, complex structures circulating in the blood that contain triglycerides, phospholipids, cholesterol, and proteins. Triglycerides are synthesized in the liver and adipose tissue, and are utilized as energy sources. Phospholipids are transported in lipoproteins, are formed in all cells of the body, and are particularly important in the formation of cellular membranes. Cholesterol, both exogenous and endogenous forms, also circulates in the lipoproteins and is used to form cholic acid in the liver. As with phospholipids, the most important function of cholesterol is the formation and permeability of

cell membranes. High circulating concentrations of low-density lipoprotein-cholesterol (LDLC) and low circulating concentrations of high-density lipoprotein-cholesterol (HDL) are associated with an increased risk of coronary heart disease (CHD).

Until recently, there has been a widely held understanding that blood cholesterol concentrations were relatively stable, showing only moderate elevations with aging, and transient but predictable elevations in triglycerides and related lipids postprandially, i.e. period of time after a meal. It is well documented that there are large inter- and intra-individual variations and fluctuations in plasma cholesterol, triglycerides, and LDLC concentrations. The variability within individuals is in the order of 10–20% over the course of a day or a week, and the degree of variability remains high even under reasonably standardized conditions.

Investigations of the factors responsible for such fluctuations are clinically relevant, because in order to make reasonable clinical decisions regarding assessment and treatment of hypercholesterolemia (condition describing a high concentration of cholesterol in the blood), an understanding of normal, biologically based variations in cholesterol concentrations are essential. As importantly, however, is the understanding of the causes of this variability in order to better understand the process of cholesterol regulation.

Environmental Influences on Lipid Concentrations and Lipid Variability

A variety of stimuli alter the blood concentrations of lipids. For example, numerous investigations have demonstrated that psychological stress reliably influences blood concentrations of cholesterol and lipoproteins. A separate and smaller literature has shown that individuals differ in their lipid responses to dietary challenge. Importantly, exaggerated lipid responsivity to diet is associated with an increased risk of CHD. To the extent that exaggerated lipid responses to environmental stimuli are atherogenic, investigation of the role of stress on lipid concentrations is of potential clinical importance. The task among behavioral scientists dedicated to investigating the phenomenon of stress-associated lipid changes is to delineate the potential mechanisms responsible for the effects, which will ultimately lead to either a confirmation or refutation of the notion that stress-associated lipid changes are atherogenic.

Several models have been proposed as potential pathways by which stress alters lipid concentrations in the blood; one such model is presented in [Figure 1](#). Although the pathways are not mutually exclusive

and are likely to be different under varying conditions of psychological stress, models such as these suggest ways of considering the current literature and point to obvious avenues for future research endeavors.

Chronic Stress Mechanisms of Lipid Variability

During chronically or pervasively stressful situations, cholesterol concentrations are elevated. A number of naturalistic studies have demonstrated that the stress associated with a major disaster or major stressor, such as the experience of earthquake, loss of job and income, and bereavement, can result in substantial increases in total cholesterol, LDLC, and triglycerides. Among better controlled laboratory studies of chronic stress, generally similar associations are noted between chronically stressful situations and elevated atherogenic lipids. Although there is not complete agreement, the preponderance of studies supports the notion that stress changes, or in some way is associated with, lipid concentrations in the blood. A meta-analysis of the bulk of the literature examining chronic or pervasive stressor effects on lipids demonstrates overall support for the notion that chronic stress is associated with elevated total cholesterol and LDLC concentrations ([Table 1](#)). Several prospective studies have further substantiated the notion that even when cholesterol concentrations are not elevated during periods of increased workload, there are still positive and significant associations between increased perceived stress and total cholesterol.

Behavioral Influences

Although it has been tempting to suggest that lipid changes are a result of stress-associated behavioral changes, such as diet and exercise modifications, which in turn increase lipid concentrations, data have not confirmed this as an adequate explanation. For example, a very early series of single case studies examined individuals while residing in a metabolic ward for periods of time ranging from 1 to 5 months. Diet, physical activity, exercise, and medication use were not only monitored, but actually closely regulated, and total cholesterol concentrations were measured frequently. Simultaneously, diaries were maintained periodically by both patients and the nursing staff, reporting on mood and perceived stress experiences. Total cholesterol levels were quite variable over time, despite the tightly controlled conditions and behaviors of the patients while in the ward. Moreover, total cholesterol fluctuations upwards were more prominent during times of self-reported perceived stress.

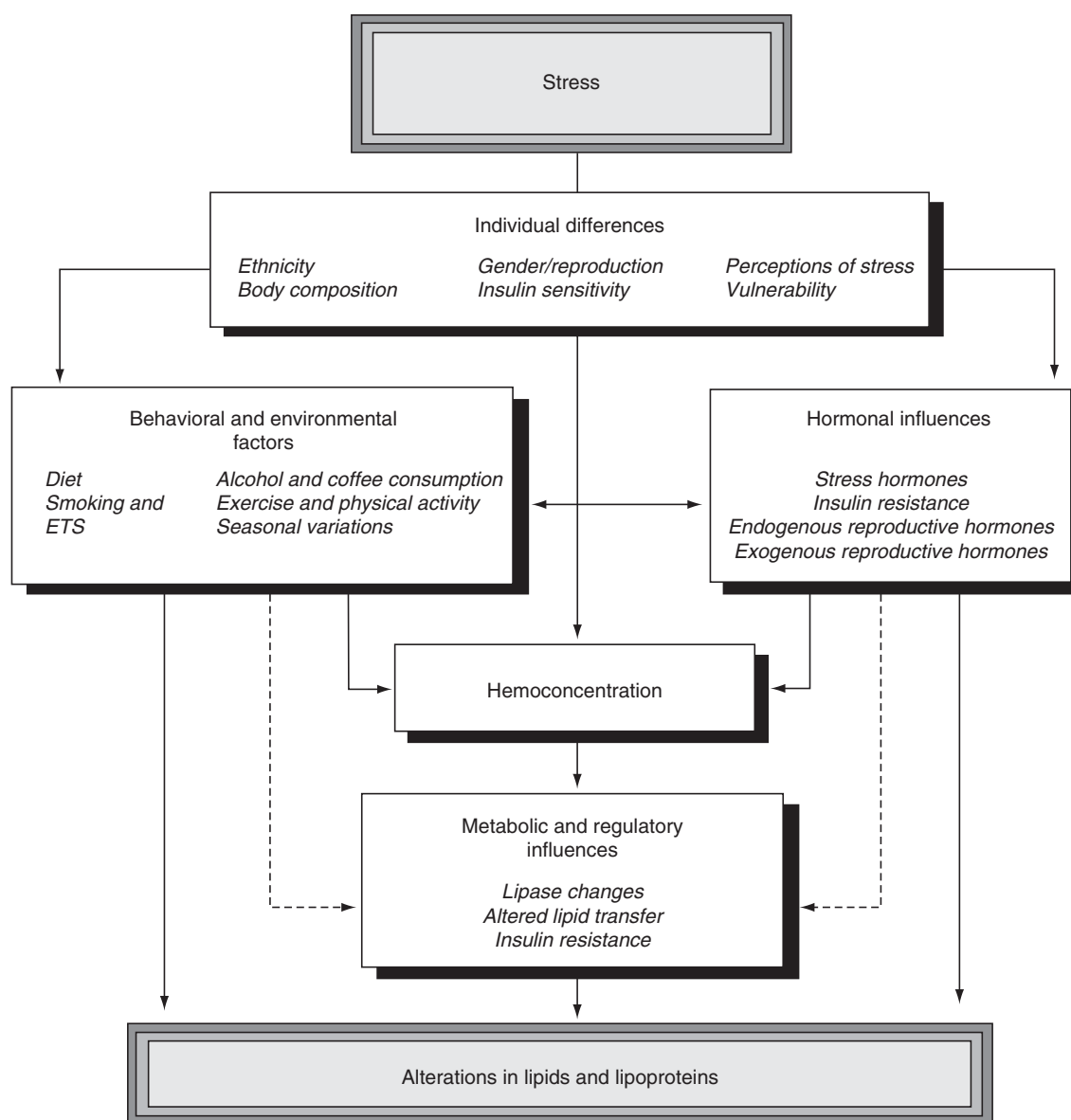


Figure 1 Potential pathways by which stress alters lipid concentrations in the blood (reprinted from Stoney & West (1997) *Lipids, personality and stress: Mechanisms and Modulators*. In: Hillbrand & Revben (eds) *Lipids, Health, and Behavior*, pp. 47–66. Washington, DC: American Psychological Association, with permission of the American Psychological Association). ETS, environmental tobacco smoke. Solid lines indicate that there is evidence for these pathways. Hashed lines indicate incomplete evidence for these hypothesized pathways.

Table 1 Meta-analysis on the effects of chronic stressors on lipid concentrations

Lipid variable	Number of studies	Unweighted effect size estimates (d)	P value
Total cholesterol	14	0.406	0.0001
LDLC	9	0.358	0.0001
Triglycerides	3	−0.014	0.05
Free fatty acids	5	0.071	0.05

LDLC, low-density lipoprotein-cholesterol.

More contemporary studies of much larger groups of men and women similarly found that the chronic stress-associated lipid elevations were maintained even after rigidly controlling variations in diet, exercise, sleep, medication use, and physical activity. For example, healthy men and women undergoing several weeks of predictable but potent work stress had elevated blood total cholesterol, LDLC, and triglycerides relative to lower work stress periods, independent of changes in diet, exercise, physical activity, medication use, and sleep patterns.

That said, in certain circumstances and for certain vulnerable individuals, stress is associated with changes in behavioral risk factors that can have important effects on cholesterol concentrations. For example, both chronically and acutely stressful situations cause some people to eat more, while others eat less. Among those who are so-called 'stress eaters', the increased eating may be, in part, a consequence of higher stress-related cortisol elevations. Importantly, at least one study has reported that those who report increasing caloric intake during stress have higher total cholesterol/HDLc during stress. However, even among this subset of vulnerable individuals, the increased lipid concentrations are small. Thus, although there may be a portion of individuals for whom lipid alterations during chronic stress are an indirect consequence of stress-induced changes in eating patterns, this mechanism likely does not explain the larger total cholesterol elevations that occur during stress among the majority of individuals.

Hormonal Influences

One way that stress might influence lipid concentrations is through stress-induced hormonal changes that affect lipid metabolism. For example, stress-induced sympathetic nervous system activity leads to concomitant increases in catecholamines (epinephrine and norepinephrine), glucocorticoids (cortisol), and free fatty acids. These increases have resultant effects on metabolism of the lipoproteins. For example, catecholamines can directly stimulate adipose tissue to release free fatty acids into circulation through the process of lipolysis. In particular, epinephrine induces the release of free fatty acids; during epinephrine infusion and mental stress, free fatty acid concentrations increase concomitantly. Epinephrine-induced increase in free fatty acids may be the result of increased blood flow through adipose tissue or stimulation of adipose- β 2 adrenoreceptors. In either case, the accumulation of circulating free fatty acids can trigger the production of triglyceride-rich, very low-density lipoprotein (VLDL) catabolism, which in turn will eventually result in increased concentrations

of circulating LDLc. Further, epinephrine will decrease VLDL secretion and increase LDLc uptake by stimulating α 1 adrenoreceptors.

Like epinephrine, norepinephrine influences lipoprotein metabolism. Increased norepinephrine activity stimulates adipose β adrenoreceptors. This may result in diminished lipoprotein lipase activity, a subsequent decline in triglyceride clearance, lower concentrations of HDL, and increasing levels of VLDL, intermediate-density lipoproteins (IDL) and LDL. Norepinephrine also decreases hepatic triglyceride lipase activity, which can increase plasma concentrations of both HDLC and triglyceride-risk lipoproteins.

Cortisol also has profound influences on the mobilization of lipids and lipid metabolism through activation of the hypothalamic-pituitary-adrenocortical (HPA) axis. Cortisol sensitizes adipose tissue to the lipolytic effects of catecholamines and increases concentrations of free fatty acids. Moreover, cortisol and free fatty acids stimulate the secretion of VLDL, increase the synthesis of hepatic triglycerides, inhibit insulin secretion, and increase insulin insensitivity in the tissues. Insulin regulates triglyceride synthesis and hepatic production of VLDL. Thus, the decreased activity of insulin during stress inhibits the regulation of triglycerides and the production of VLDL. Further, insulin resistance and increased levels of cortisol suppress hepatic LDL receptors and catabolism, thereby delaying LDL clearance. Cortisol also stimulates apolipoprotein B (apo B) secretion from cultured rat hepatocytes.

This pathway of stress-related, epinephrine-induced free fatty acid release from adipose tissue, enhanced hepatic cholesterol secretion, cortisol-associated lipolysis, and increased circulating LDLc concentrations has some empirical support. Chronic epinephrine infusions into cynomolgus monkeys produce significant total cholesterol concentration increases, and the same pattern has been noted among humans. Additional data regarding the effects of cortisol are also available. For example, several types of stressors simultaneously elevate LDLc, total cholesterol, apo B, insulin, and plasma cortisol. In some studies, the changes in cortisol positively predict changes in lipids. For example, a recent study reported that individuals with high cortisol stress recovery (reflecting higher cortisol stress responses) also had the highest LDLc and total cholesterol/HDLc ratios. Thus, the sympathetic-adrenal-medullary axis and the HPA axis are likely to play a key role in modulating lipid and lipoprotein metabolism during times of pervasive psychological stress. There are continuing efforts to delineate these pathways more definitively.

Acute Stress Mechanisms of Lipid Variability

The hypothesis that chronic stress affects lipids due to stress-induced hormonal changes has considerable appeal, but it is unlikely to account for the abundant additional data demonstrating rapid and transient increases in plasma cholesterol concentrations during acute psychological stress. For example, many laboratory-based investigations have shown that very short-term stressors, such as those that might occur over a period of several minutes to 1–2 h, can induce rapid and transient increases in blood concentrations of total cholesterol, LDLC, apo B, and triglycerides. A meta-analysis of these investigations has supported the notion that short-term stress reliably elevates total cholesterol, LDLC, triglycerides, and free fatty acids (Table 2). These elevations are clearly not a function of hormonally induced increased synthesis of lipids because the enhanced rate of lipoprotein entry into the circulation would simply take longer than the time period represented in these studies. Kinetic studies document that the manufacture of VLDL from fatty acids and subsequent conversion of LDL requires several hours to days to accomplish. Investigations of acute stress effects have also demonstrated little concordance between catecholamine and cortisol responses and acute changes in blood lipid concentrations. The effects of acute psychological stress on lipid concentrations are therefore likely to operate through a different mechanism or set of mechanisms.

Hemoconcentration Influences

Psychological stress causes a generally transient and rapid decline in the volume of blood plasma, with a resultant mild hemoconcentration of blood cells. This phenomenon is similar to the acute decreases in intravascular plasma volume that occur with postural changes, with a resultant posture-induced hemoconcentration of serum lipids. These plasma volume shifts can occur by increased red blood cell mass, alterations in kidney functioning, and/or through hemodynamically induced fluid shifts. Whatever the exact mechanism of the fluid volume shifts, it is clear that the resultant hemoconcentration leads

to higher lipid concentrations in the blood. Thus, the increased lipid concentrations in the blood during acute stress are likely to be partially, but not fully, accounted for by hemoconcentration. Generally speaking, during mild stressors that induce no hemodynamic or hormonal (e.g., catecholamine and blood pressure) changes, a significant portion of the generally small lipid increases might reasonably be accounted for by plasma volume shifts. However, during more substantial psychological stressors that are accompanied by alterations in catecholamines, blood pressure, and cortisol, increases in cholesterol and LDLC are generally more substantial. In some, but not all investigations, the lipid (generally total cholesterol and triglyceride) and lipoprotein (LDLC) elevations remain significant even after correction for transvascular plasma volume shifts. For those investigations that do not show significant lipid elevations after correction for plasma volume shifts, it is possible that the failure to retain significance is related to decreased power in combination with modest effect sizes. The fact that some lipoproteins are consistently elevated during stress (e.g., LDLC) and others are not (e.g., HDLC) suggests that transvascular shifts cannot account for all of the stress-related changes observed, because otherwise all lipid particles would show the same pattern of response.

Metabolic Influences

This state of affairs has led to the speculation that there may be a metabolic explanation for the acute stress-induced elevation in blood lipid concentrations. For example, stress-associated elevations in lipids and lipoproteins may be due to metabolic effects of sympathetic activation. Pharmacologic blockade of sympathetic activation with a nonselective adrenoceptor antagonist among men completing several acutely challenging tasks blocked stress-associated increases in free fatty acids and triglycerides, but not total cholesterol or LDLC. These data add support to the notion that different lipids are likely influenced by environmental and psychological factors through different mechanisms.

Stoney et al. have reasoned that if stress does have a direct influence on certain blood lipid levels, that

Table 2 Meta-analysis on the effects of acute stressors on lipid concentrations

Lipid variable	Number of studies	Unweighted effect size estimate (<i>d</i>)	<i>P</i> value
Total cholesterol	25	0.306	0.0001
LDLC	14	0.291	0.0002
Triglycerides	24	0.341	0.0001
Free fatty acids	37	0.717	0.0001

LDLC, low-density lipoprotein-cholesterol.

this may be reflected in the efficiency with which individuals metabolize a high-fat meal during stress. In a test of this hypothesis, a standard intravenous load of intralipid (a triglyceride-rich product) was administered to a group of middle-aged men and women under a condition of psychological stress, and then again during a condition of no psychological stress. In this test, which is similar conceptually to a glucose tolerance test, a triglyceride determination was made and then a standard bolus of intralipid was administered. Serial triglyceride determinations were then made over time, and the disappearance rate constant (K_2) was calculated. Thus, K_2 defines the clearance rate of the exogenous fat solution, with a higher K_2 value indicating a more efficient clearance of triglycerides from the circulation. When administering this test to a group of healthy individuals, Stoney et al. found that K_2 values declined significantly during stress, in relation to a nonstress period, showing that the increases in lipids during acute stress were due to altered clearance rates. They also found that individuals who were slow to clear a triglyceride load were more hostile than those who were more efficient at clearing the triglyceride load imposed. These data are the most compelling to date to indicate that blood lipid concentrations are altered during psychological stress because of altered triglyceride metabolism and that this may be related to an increased risk of CHD.

Interventions to Reduce Stress-Related Increases in Lipids

Outside of altering health behaviors or broad lifestyle changes, very few clinical interventions have been tested as potential means to reduce the stress-associated elevations in lipids. Stress reduction techniques such as relaxation, meditation, and self-monitoring likely have effects on both behavioral risk factors and lipid levels. For example, individuals with highly demanding jobs and who spent 3 weeks away from their job responsibilities and in a spa had more substantial decreases in LDLC and LDL/HDL ratio than did those who also spent time in a spa but who had less demanding occupations. However, no studies have explicitly tested the mechanisms by which reductions in perceived stress related to reductions in lipid and lipoprotein concentrations.

Clinical Implications of Variations in Lipids

There is little doubt about the presence of a direct relationship between blood cholesterol concentrations and the development and initiation of cardiovascular disease. For example, chronically elevated LDLC is associated with an elevated incidence of

coronary artery disease (CAD), and lowering LDLC concentrations produces a regression of coronary atherosclerosis and coronary artery events in some individuals. Elevations in triglycerides and low levels of HDLC are also correlated positively with both CAD and CHD although it is not clear to what degree lowering triglyceride or raising HDLC levels also results in a diminished risk for cardiovascular disease.

Although there are no prospective investigations of the clinical relevance of stress-associated elevations in lipids, data from several cross-sectional investigations are informative. A number of investigations have demonstrated that healthy individuals at an elevated risk for the future development of cardiovascular disease have more profound short-term elevations in blood lipid concentrations during stress, relative to a nonstress period. For example, men, postmenopausal women, and healthy young offspring of parents with documented myocardial infarction all have greater elevations in total cholesterol and LDLC during acute psychological stress, relative to women, age-matched premenopausal women, and healthy offspring of parents with no cardiovascular disease, respectively. Additionally, psychosocial factors that have been shown to be related independently to CHD risk, such as hostility and anxiety, are also associated with total cholesterol and LDLC concentrations. Future cross-sectional and prospective investigations examining the extent to which stress-induced elevations in lipid concentrations are atherogenic are warranted.

Further Reading

- Brennan, F. X., Cobb, C. L., Silbert, L. H., et al. (1997). Peripheral β adrenoreceptors and stress-induced hypercholesterolemia in rats. *Physiology and Behavior* 60, 1307–1310.
- Brindley, D. N., McCann, B. S., Niaura, R. S., et al. (1993). Stress and lipoprotein metabolism: modulators and mechanisms. *Metabolism* 42(Suppl 1), 3–15.
- Stoney, C. M. and West, S. (1997). Lipids, personality, and stress: Mechanisms and modulators. In: Hillbrand, M. & Spitz, R. T. (eds.) *Lipids, Health, and Behavior*, pp. 47–66. Washington, D.C.: American Psychological Association.
- Stoney, C. M., Niaura, R., Bausserman, L., et al. (1999). Lipid reactivity to stress: I. Comparison of chronic and acute stress responses in middle-aged airline pilots. *Health Psychology* 18, 241–250.
- Stoney, C. M., West, S. G., Hughes, J. W., et al. (2002). Acute psychological stress reduces plasma triglyceride clearance. *Psychophysiology* 39, 80–85.
- Strass-Blasche, G., Ekmekcioglu, C. and Marktl, W. (2003). Serum lipid responses to a respite from occupational and domestic demands in subjects with varying levels of stress. *Journal of Psychosomatic Research* 55, 521–524.

Chronic Fatigue Syndrome

A J Cleare and S Wessely

King's College London, Institute of Psychiatry,
London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by A J Cleare and S Wessely, volume 1, pp 460–466,
© 2000, Elsevier Inc.

Historical Aspects
Symptomatology
Epidemiology
Etiology
Treatment
Prognosis
Synthesis

Glossary

*Myalgic
encephalo-
myelitis (ME)*

Term invented to describe an outbreak of unexplained illness that affected the staff of the Royal Free Hospital in 1955. It has become linked to chronic fatigue syndrome in the last 20 or more years. Because encephalomyelitis is a specific pathological process not found in these patients, the term is avoided in the scientific literature.

Neurasthenia

A term introduced to medicine in 1869 to describe an illness characterized by chronic physical and mental exhaustion after minimal effort and accompanied by numerous somatic symptoms, a precursor for modern chronic fatigue syndrome. Originally thought to be an organic disease of the central nervous system caused by depletion of the energy supply to the brain resulting from such insults as overwork or infection, it was later seen as a psychological condition linked to depression and anxiety. Now an archaic term, although retained in the latest *International Classification of Diseases* (ICD-10).

*Postviral
fatigue
syndrome
(PVFS)*

A term used to describe chronic fatigue syndrome apparently triggered by a viral infection. Because of the frequency of viral infections and the lack of any compelling evidence implicating viral infection in the persistence of most cases of CFS, the term is to be avoided unless there is clear-cut evidence of a relevant infective trigger.

Somatization

A condition in which the patient presents with a physical symptom that is attributed to a physical disease but is more likely to be associated with depression or anxiety.

Chronic fatigue syndrome (CFS) is an operationally defined syndrome characterized by a minimum of 6 months of severe physical and mental fatigue and fatigability, made worse by exertion. Other symptoms such as muscle pain, sleep disorder, and mood disturbance are common. [Table 1](#) lists the current consensus criteria, although the Centers for Disease Control (CDC) identified ambiguities in these criteria, and more recommendations for their resolution, in 2003 (Reeves et al, 2003). Other common causes of chronic fatigue must be considered before the diagnosis can be made.

Historical Aspects

The term chronic fatigue syndrome (CFS) is a relatively recent arrival on the medical scene, but it is not a new illness. The condition known to the Victorians as neurasthenia is clearly its forerunner, not just in terms of symptoms, but presumed etiologies, treatments, and much else. As neurasthenia gradually shifted from a physical to a psychological etiology, its popularity declined, and the diagnosis had effectively vanished by World War II, except in parts of Asia. Chronic fatigue syndrome appeared in internal medicine in the 1980s in response to a variety of events, including renewed claims to have found a physical, infective cause for the symptoms, a series of well-publicized epidemic outbreaks, and the general change in doctor–patient relationships during this period. The term CFS and its variants are now widely used in the English-speaking and Scandinavian world, but not elsewhere.

Symptomatology

The hallmark of CFS is severe fatigability, of both mental and physical functions, after minimal exertion. It is accompanied by other common symptoms, chiefly muscle pain, sleep disorder, and mood disturbance. Some sufferers complain of symptoms affecting most bodily functions. CFS should not be diagnosed without some complaints of physical fatigue and fatigability (reduced functional capacity, symptoms after minor exertion, postexertional myalgia, and others)

Table 1 Consensus operational criteria for chronic fatigue syndrome^a**Patients must have the following:****1. Chronic fatigue that**

Is of new onset (not lifelong)

Has lasted at least 6 months

Has resulted in substantial functional impairment (occupational, social, or personal activities)

Is not substantially alleviated by rest

2. At least four of the following symptoms for at least 6 months

Memory or concentration impairment

Sore throat

Tender lymph nodes

Muscle or multijoint pain

Headaches

Unrefreshing sleep

Postexertional malaise lasting > 24 h

3. No medical illness explaining the symptoms**4. No diagnosis of**

Psychosis (schizophrenia, mania or psychotic depression)

Eating disorder (anorexia nervosa, bulimia nervosa)

Dementia

Alcohol or substance abuse

Severe obesity (body mass index > 45)

^aData from Fukuda et al. (1994). *Annals of Internal Medicine* 121, 953–959.

and mental fatigue and fatigability (poor concentration, symptoms made worse by mental effort, subjective memory disturbances, etc.). Because there are many causes of chronic fatigue and fatigability, it is essential that an appropriate medical evaluation is undertaken before diagnosing CFS, including physical examination, mental state examination, and simple investigations.

Epidemiology

Chronic fatigue is one of the commonest somatic symptoms – between 20 and 30% of most populations studied report symptoms such as exhaustion and feeling tired all the time. Some of these are short-term complaints, but around 10% of those attending general practice will complain of chronic fatigue sufficient to interfere with their lifestyle and lasting at least 6 months. Operationally defined CFS is less common, ranging from 0.1 to 3% depending on the definition used, with the best estimates suggesting that approximately 1% of primary care attendees fulfill the diagnostic criteria. Even when those with comorbid psychiatric disorders are excluded, a reasonable estimate of prevalence of pure CFS is approximately 0.5%. Few of these subjects will use a term such as CFS, myalgic encephalomyelitis (ME), or its variants to describe the illness. Population surveys have found similar prevalence of CFS in people

of different socioeconomic status. Women are at higher risk than men (relative risk 1.3–1.7, depending on the diagnostic criteria used), while studies of ethnicity suggest that White individuals have a lower risk of CFS compared with Latinos, African Americans, and Native Americans. Different characteristics have been reported in clinic populations (e.g., an excess of White professionals), almost certainly because of biases in referral patterns.

An operational definition should not be taken as endorsing the concept of CFS as a discrete disorder – in fact, epidemiological research suggests that it is not, and is best viewed as the extreme end of a normally distributed spectrum of fatigability, much in the same way as high blood pressure is not a discrete disorder distinguishable from moderate blood pressure.

While it is unlikely that CFS is a single illness, efforts to define valid subgroups by symptom profile or etiological factors have not yet been particularly fruitful. Dividing those with an abrupt onset to symptoms from those with a more insidious origin seems to have some merit, the former having a more favorable prognosis. Differences also exist between the minority of cases with long illness histories, severe disability, and multiple symptoms, who show overlap with the concept of somatization disorder, and the larger group with less disability, fewer symptoms, and shorter illness durations.

Another area of confusion is the relationship between CFS and other medically unexplained syndromes. Especially in specialist practice, there is considerable overlap between CFS and the condition recognized by rheumatologists as fibromyalgia. It is likely that the boundaries between CFS and other unexplained and/or controversial syndromes such as irritable bowel syndrome and multiple chemical sensitivity are similarly fuzzy. Another area of overlap is between chronic fatigue and chronic pain. The similarities are many: the increased rates of affective and somatic symptoms, the role of avoidance behavior in sustaining symptoms, the role of physical factors as precipitants, the need for rehabilitation rather than cure, and the role of social and economic factors in perpetuating disability and causing conflict both with and between doctors.

Etiology

Biological Factors

Genetics Studies suggest that there is a genetic propensity to suffer from chronic fatigue, with the concordance between monozygotic twins more than twice that of dizygotic twins. Multivariate modeling suggests that part of this vulnerability is conferred

through the genetic tendency to suffer from anxiety or depression, but that part is independent of psychiatric disorder.

Muscle Although muscle pain and complaints of muscle fatigability are common, there is little objective evidence that symptoms have a primary neuromuscular origin. Most studies of neurophysiological function have been normal, particularly once the secondary effects of inactivity are taken into account. There may be a small group whose symptoms are related to primary muscle dysfunction, but given the importance of mental fatigue and fatigability, central rather than peripheral explanations are more likely to be relevant for most patients.

Viruses While it is common for patients to date the onset of their symptoms to common viral infections, a controlled prospective study found that subjects who consulted their physician with viral symptoms showed the same rate of subsequent chronic fatigue as those consulting for other reasons. There is, however, consistent evidence that certain infective agents causing severe illness, such as the Epstein–Barr virus (EBV), viral hepatitis, Q fever, or viral meningitis, are associated with prolonged fatigue reactions (Table 2). Several factors increase the risk of developing prolonged fatigue after such infections. These include pre-existing fatigue, older age, female gender, previous psychiatric disorder, and the response to symptoms (including a prolonged convalescence from illness and oversolicitous social support). The response of the doctor may be important, too: provision of a sick note and an unclear diagnosis have both been associated with the development of CFS.

Neurochemistry Early studies in normal subjects showed that fatigue after exercise was associated with a rise in plasma levels of tryptophan and, theoretically, increased brain serotonin (5-HT) production. In CFS, increased cerebrospinal fluid (CSF) levels of 5-HIAA relative to controls suggested increased turnover of

5-HT in the CNS (brain and spinal cord). Many studies have used direct challenges of brain serotonergic receptors, measuring neuroendocrine responses as indices of the functional status of these receptors. These have generally revealed enhanced responses, suggesting increased serotonergic neurotransmission. These findings are important when set in the context of similar studies in major depression that reveal the opposite change: reduced serotonergic function. The importance of these changes is not yet understood. It is possible that they are of physiological significance, since 5-HT is known to be important in the physiological control of sleep, appetite, and mood. In parallel with the opposite 5-HT changes, these features are disturbed in opposite directions in classic endogenous depression (insomnia, anorexia, agitation) and CFS (hypersomnia, hyperphagia, and retardation). Given what we know about brain physiology, it is also conceivable that these changes could be secondary to low cortisol levels or the effects of inactivity, sleep disturbance, or disruption of circadian rhythms.

Recent advances in neurochemical imaging have allowed direct visualization of cerebral serotonin receptors in CFS. Early studies in CFS report a reduction in 5-HT_{1A} receptor binding throughout the brain, most markedly in the hippocampi bilaterally (Figure 1), and reduced serotonin transporter binding in the rostral anterior cingulate gyrus. It has been suggested that these findings represent receptor down-regulation in the face of chronically raised synaptic serotonin levels, but other explanations are also possible, including a longstanding effect of prior depressive illness.

Although there have been individual studies of other neurotransmitter dysfunction in CFS, further study is needed before drawing any conclusions.

Neuroendocrine Dysfunction Abnormal function of the hypothalamic-pituitary-adrenocortical (HPA) axis (Figure 2) has been demonstrated in several centers. Most, but not all, of the better-quality studies using serial samples of blood or saliva, or 24-h urinary collections, have found reduced basal cortisol output, again the reverse of that seen in major depression. Studies of the circadian rhythm suggest a blunted early morning rise of corticotropin (ACTH). This is paralleled by the findings of an impaired ACTH or cortisol response to a variety of challenges to the HPA axis, including CRH, arginine vasopressin (AVP), Synacthen (1–24 adrenocorticotropin), insulin, naloxone, exercise, social stress, and awakening. There is also evidence of heightened negative feedback and glucocorticoid receptor function. However, there is no evidence for a unique or uniform dysfunction of the HPA axis. Given the many factors that may

Table 2 Fatigue following severe infections

	<i>Rate of persistent fatigue 6 months after diagnosis (%)</i>
Viral meningitis	25
Epstein–Barr virus glandular fever	12–16
Viral hepatitis	20
Q fever	42 ^a
Common viruses	1
Baseline risk	1

^aThere were also very high rates (26%) in controls in this study.

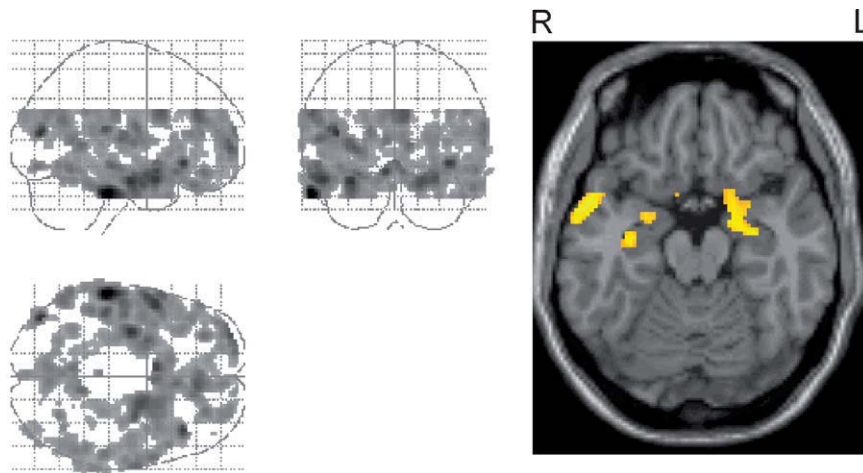


Figure 1 Cerebral 5-HT_{1A} receptor binding in CFS, measured using positron emission tomography and [¹¹C]WAY-100635. On the left, the statistical parametric mapping analysis (using a threshold of $P < 0.01$) is shown in three dimensions, revealing a widespread reduction in 5-HT_{1A} receptor binding in CFS versus controls throughout the brain. On the right, results using a stricter threshold ($P < 0.001$) are projected onto a standardized T1-weighted magnetic resonance image, showing bilateral reductions in 5-HT_{1A} receptor binding centered on the hippocampus. From Cleare, A. J., Messa, C., Rabiner, E. and Grasby, P. (2005). Brain 5-HT_{1A} receptor binding in chronic fatigue syndrome measured using positron emission tomography and [¹¹C]WAY-100635. *Biological Psychiatry* **57**, 239–246.

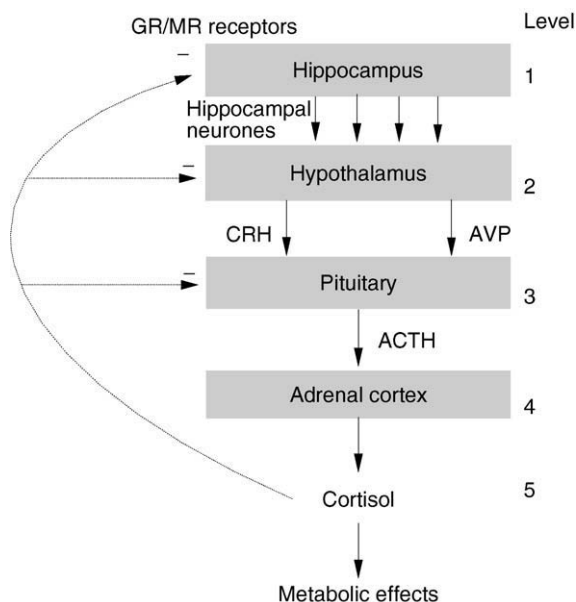


Figure 2 Components of the hypothalamic-pituitary-adrenocortical (HPA) axis and directions of positive and negative feedback loops. Abbreviations: CRH, corticotropin-releasing hormone; ACTH, corticotropin; AVP, arginine vasopressin; MR, mineralocorticoid receptors; GR, glucocorticoid receptors. See Table 4.

impinge on the HPA axis in CFS, such as inactivity, sleep disturbance, psychiatric comorbidity, medication, and ongoing stress, it seems likely that HPA axis disturbance in CFS is heterogeneous and of multifactorial etiology.

One issue with all of the above studies is that they were largely undertaken on samples of patients that had been ill for several years, and the relevance of these

HPA axis changes to the onset of illness is unknown. Prospective studies of groups at high risk of developing chronic fatigue (post-EBV infection and postsurgery) have found that the development of fatigue 6 months later was not associated with any HPA axis changes.

Thus, at present it appears that HPA axis changes are not important during the early stages of the genesis of fatiguing illnesses. However, the low cortisol levels that may develop at a later stage of the illness do appear to be of clinical relevance, since their reversal using supplementation with a low dose of hydrocortisone can improve fatigue in some patients. Also, there is now evidence that HPA axis changes can be reversed by modifying behavioral features of the illness such as inactivity, deconditioning, and sleep disturbance, using cognitive behavioral therapy.

Consistent evidence for other disruption of other neuroendocrine systems is lacking.

Neuroimaging The literature to date on CFS has been conflicting. Though some early studies did suggest that there was reduced blood flow to many areas of the brain, most notably the brain stem, more methodologically rigorous studies, particularly those in which comorbid depression was controlled for, suggested that there was no decreased blood flow and some subcortical areas of increased flow. Furthermore, a study that compared 22 pairs of monozygotic twins discordant for CFS found no alterations in regional cerebral blood flow.

Structurally, some studies have found an increased number of white matter hyperintensities, something nonspecific and also found in mood disorders. Some

findings of lateral ventricular enlargement, global loss of gray matter, and reduced hippocampal neuronal density have been reported, but there are as yet no replicated findings.

Finally, functional MRI studies are beginning to identify the neural correlates of the symptom of fatigue; these may enhance our understanding of the brain mechanisms underlying the sense of effort.

Immunology Initial views of CFS emphasized immunological changes as an inherent part of the syndrome. Subsequent studies have given conflicting results – normal, increased, or decreased cell-mediated and humoral immune markers. The most frequent finding is a mild immune activation, but it is unclear whether this is secondary to the decreased cortisol tone (see previous discussion) or frequent mood disorder found in CFS. There is also some evidence of a shift toward a Th2-type immune response. However, there has been little evidence of a relationship between immune dysfunction and clinical parameters or outcome, and no evidence at all of any clinical immunological problems. Likewise, although abnormal cytokine release in response to antigen challenge is a tempting model for CFS, this would have to be centrally mediated, and has not been proven.

Autonomic nervous system Several studies have investigated the possible role of the autonomic nervous system in CFS. A number of studies have found an increased rate of neurally mediated hypotension, while others have found evidence for sympathetic overactivity and/or parasympathetic underactivity. Whether these are causal factors in CFS remains under debate since both anxiety and prolonged inactivity are associated with similar changes.

Psychological Factors

Life events The literature regarding the effects of life events on CFS is still conflicting. It is clear that, as in other conditions, adverse life events in CFS are strongly associated with the development of psychological disorders. There is also some evidence that negative life events may be related to the development of, and chronicity of, fatigue, with the best evidence suggesting an excess of negative life events in the 12 months prior to the onset of CFS. Positive life events may be protective.

Several specific stresses have been studied: around one-third of patients suffer from significant fatigue following major surgical operations, while the overtraining syndrome (renamed the underperformance syndrome [UPS]) afflicts athletes who engage in periods of excessive exercise and bears some similarities with the symptoms of CFS. Others have studied

fatigue syndromes occurring during cancer treatment or α -interferon administration. These conditions share with CFS the interaction of psychological and physiological factors in producing the final clinical picture.

Psychiatric disorder The relationship between psychiatric disorder and CFS remains controversial, but there is no doubt that such a relationship exists. Exact rates of psychiatric disorder vary from study to study, but usually between one-half and two-thirds of clinical samples fulfill criteria for psychiatric disorder as well as CFS. The greater the number of symptoms and the greater the disability, the greater the chance of psychiatric disorder. Importantly, a series of controlled studies have shown that these rates are too high to be explained simply as a reaction to physical disability (Table 3), though clearly this is part of the explanation. Other possible explanations include misdiagnosis, a neurobiological origin common to CFS and psychiatric disorder, and the overlap between the criteria used to diagnosis CFS and the common psychiatric disorders. No study has ever reported complete congruence between CFS and psychiatric disorder. Furthermore, factor analytic studies

Table 3 Studies comparing prevalence of psychiatric disorder in CFS with prevalence in medically ill controls

<i>Control group</i>	<i>Prevalence of psychiatric disorder (%)</i>		<i>Relative risk of psychiatric disorder in CFS</i>
	<i>CFS</i>	<i>Controls</i>	
Neuromuscular diseases	72	36	2.0
Myopathies	45	6	7.5
Rheumatoid arthritis	41	13	3.3
Multiple sclerosis	23	8	2.9
Multiple sclerosis	45	16	2.8

Table 4 Sites of possible HPA axis abnormalities in CFS shown in Figure 2

<i>Level (Figure 2)</i>	<i>Finding</i>
1	Hippocampal atrophy
1	Impaired rate-sensitive negative feedback
1	Reduced suprahypothalamic drive to HPA axis
2, 3	Enhanced negative feedback
2	Impaired CRH and/or AVP release from hypothalamus
3	Impaired pituitary response to CRH
4	Blunted adrenal cortex response to ACTH
4	Adrenal gland atrophy
1–4	Impaired HPA response to a variety of stressors
5	Low basal cortisol

of postviral cohorts support the existence of fatigue-like symptoms at least partially separate from anxiety and depression symptoms.

The commonest psychiatric disorder reported in studies of CFS is depression (perhaps because of the considerable overlap of operational criteria), which has led to a relative neglect of anxiety disorders. However, in some respects CFS shares more similarities with anxiety than with depression, particularly in terms of neuroendocrine dysfunction, the role of fear and avoidance behavior, and treatment response.

The role of personality remains unclear, with the literature split as to whether certain personality traits like perfectionism predispose an individual to CFS.

Neuropsychological function CFS patients frequently complain of considerable difficulties with memory, attention, and concentration. However, the results of neuropsychological testing in CFS are inconsistent. In general, the objective findings do not reflect the severity of the subjective complaints encountered. Some changes may be explained by depression, anxiety, and/or sleep disturbance. There is evidence of deficits in higher and more complex cerebral function, such as selective attention. The selection of comparison groups is important; in a twin study, although CFS patients showed reduced performance on timed cognitive tasks, so did their nonfatigued twins. Nevertheless, it is plausible that sufferers find everyday tasks that require increased cognitive resources to be difficult, and that although they are able to perform reasonably well on such tasks, it is at the price of increased mental effort. Physical exercise also worsens effortful cognitive processing in CFS, suggesting that a disorder of the sense of effort lies at the heart of CFS.

Cognitive-behavioral factors The cognitive-behavioral model distinguishes factors that may have precipitated chronic fatigue from secondary factors thought to maintain fatigue and disability. The latter include the following:

1. Effects of inactivity: prolonged rest is an effective strategy in the short term but counterproductive in the medium and long term. With the passage of time, more symptoms and greater fatigue occur at progressively lower levels of exertion. Inactivity therefore sustains symptoms and increases sensitivity to them.

2. Inconsistent activity: many sufferers are not persistently inactive; instead, excessive or prolonged rest is followed by a burst of activity that is out of proportion to the preceding level of inactivity. This pattern may also be reinforced by the sense of frustration often encountered in sufferers, and perhaps also by

pre-existing personality and lifestyle factors. This sudden increase in activity culminates in exhaustion, for which the inevitable response is further rest. This stop-start pattern means that while extremes of disability are often avoided, sufferers are unable to build up a sustained level of recovery. This pattern may underlie the characteristic complaint of CFS sufferers of delayed fatigue and myalgia, well-recognized physiological phenomena that occur between 24 and 48 h after any exertion in excess of a person's current (and not previous) fitness.

3. Illness beliefs and fears about symptoms: these can influence disability, mood, and behavior in any illness. In CFS, unhelpful and inaccurate illness beliefs, reinforced by ill-informed media coverage, include the fear that any activity that causes an increase in fatigue is damaging or impossible; that doing too much causes permanent muscle damage; that CFS is due to a persistent virus or progressive immune disorder; that CFS is irreversible or untreatable; and that rest is the only remedy. Research suggests that the tendency of CFS patients under specialist care to ascribe their illness to a predominantly biological cause is not the issue – what predicts the outcome is how the meaning given to symptoms influences the choice of strategies to deal with such symptoms.

4. Symptom focusing: many patients have come to rely on day-to-day monitoring of their own symptoms as the best method of determining activity levels hour by hour. Concern about the meaning and significance of symptoms, interpreted as warning signals or relapses, is heightened by the unpredictable nature of CFS. Increased concern leads to heightened awareness, selective attention, and body watching, which can then intensify both the experience and perceived frequency of symptoms. A vicious circle is set up, thereby confirming illness beliefs and reinforcing disability.

Social Factors

Social influences on CFS operate in two arenas. The first is on the micro level. Social support influences the outcome of most conditions, and CFS is unlikely to prove an exception. One recent study suggested that certain forms of support can be helpful, but others can be detrimental. Oversolicitous support from a spouse or partner may, as in chronic pain, have a paradoxically adverse effect on outcome.

CFS also exists in the macro arena. It has become a cause célèbre, one likely to excite passion from many quarters. This reflects the difficulties that patients and doctors have in dealing with illnesses of uncertain etiology that cannot be easily pigeonholed. It also reflects the changing nature of the doctor-patient relationship and the increasing reluctance of many

people to accept what they see as inadequate or dismissive attitudes from doctors.

Interaction of Predisposing, Precipitating, and Perpetuating Factors

An alternative method of dissecting the various components of pathophysiology in CFS is to identify areas where factors may act as risk factors, triggers, and perpetuating factors. [Figure 3](#) shows the preceding factors separated in such a schema.

Treatment

Pharmacotherapy

Antidepressants Two placebo-controlled studies of fluoxetine in CFS showed divergent results, one indicating no benefit, even in depressed CFS patients, the other showing mild benefit to both fatigue and depression. Two studies of monoamine oxidase inhibitors (phenelzine and moclobemide) also showed mild benefit, primarily on symptoms rather than disability. Low dose tricyclics are helpful if pain or insomnia predominate, but there are no randomized controlled trials specifically in CFS.

Corticosteroids Because of the finding of reduced cortisol levels, two studies investigated a replacement strategy using hydrocortisone (cortisol). A full replacement dose (30 mg daily) produced mild benefit but also endogenous adrenal suppression. Smaller doses (5–10 mg daily) produced large reductions in fatigue in about 28% compared to 9% in the placebo group, without adrenal suppression. Potential side effects probably preclude such strategies as routine treatments but suggest that low cortisol levels may be a significant perpetuating factor in fatigue chronicity. Three randomized controlled trials used the mineralocorticoid fludrocortisone (one in combination with hydrocortisone) in order to treat a presumed autoimmune disorder and found no benefits.

Other Claims have been made for several other treatments, including magnesium supplements, oil of evening primrose oil, nicotinamide adenine dinucleotide and immunotherapy. None have yet passed the tests of replication or safety.

Graded Exercise Therapy

Given the perpetuating effects of inactivity and chronic fatigue, graded exercise is an obvious treatment for CFS. Research has shown that, providing account is taken of the level of deconditioning, CFS patients can perform graded exercise without adverse effects. Clinical trials have confirmed the considerable

benefits. However, improvement is not related to measures of fitness, suggesting that the benefits of exercise are more behavioral than physiological and may reflect increased confidence and decreased avoidance behavior. Nevertheless, many patients are reluctant to engage in such programs, usually because of fears of possible adverse consequences.

Cognitive-Behavioral Therapy

Cognitive-behavioral therapy (CBT) is a promising line of treatment. It encourages a collaborative approach to treatment, emphasizes the importance of engagement, and has lower dropout rates than those encountered in simple exercise programs. Treatment depends upon challenging some of the unhelpful illness beliefs encountered in some specialist clinics, overcoming behavioral avoidance by encouraging consistent, modest activity followed by graded increases, and looking at cognitive factors associated with predisposition and relapse (previous discussion). Controlled trials confirm the benefits of treatment in specialist centers: in one study, two-thirds of patients had a good outcome 6 months after the end of treatment, compared to one in five receiving relaxation as a control for nonspecific effects of treatment. One problem, however, has been the shortage of skilled therapists. There is now evidence that less specialized therapists (but not, regrettably, general practitioners) can also be trained to deliver effective CBT for CFS. Good outcome is associated with altered beliefs about the relative merits of rest versus exercise. Benefits are largely maintained at 5-year follow-up. CBT may also reverse some of the HPA axis changes seen in CFS.

Prognosis

The prognosis of patients attending specialist clinics and not receiving rehabilitation is poor. A systematic review of the long-term prognosis of specialist clinic attenders, undertaken prior to the introduction of effective rehabilitation strategies, found that children had better outcomes than adults: 54–94% of children showed definite improvement in symptoms, whereas 20–50% of adults showed some improvement in the medium term and only 6% returned to premorbid levels of functioning. Factors associated with a poor prognosis are a longer duration of illness, fatigue severity, comorbid depression and anxiety, and holding an exclusively physical attribution for CFS. This latter factor is not a problem *per se*, but is probably associated with continuing behavioral avoidance and firm beliefs about the adverse effects of activity. Given that most patients do not reach a specialist clinic until they have been ill for some years, it may

Predisposing factors (risk factors) in fatigue generation:

Female gender
Past psychological illness
Somatic attributional style
Illness in close family members
Childhood experiences
History of fatigue
History of other medically unexplained symptoms
Genetic

Precipitating factors (triggers) in fatigue generation:

Serious infective illness (EBV, Q fever, viral meningitis, viral hepatitis)
Life events
Operative stress
Overtraining
Depression/anxiety
Cancer
Physical trauma?
Head injury?

ACUTE OR
SUBACUTE FATIGUE

Perpetuating factors:

Response to fatigue: Prolonged bed rest or time off work
Fixed somatic attribution of cause
Demoralization, depression, anxiety

Effects of inactivity: Fatigue
Physical deconditioning

Response of doctor: Issuing sick note
Imprecise diagnosis

Behavioral factors: Use of avoidance as coping strategy

Cognitive factors: Loss of control
Belief that exercise will cause damage or worsen fatigue

Biological factors: Reduced cortisol levels
Hypothalamic disturbance
Circadian rhythm disruption or desynchronization
Altered central 5-HT receptor sensitivity
Autonomic disturbances

Sleep: Insomnia, hypersomnia, or poor sleep efficiency

Social factors: Reinforcement by others
Work-related problems
Lack of social support

CHRONIC FATIGUE OR
CHRONIC FATIGUE SYNDROME

Figure 3 Various potential factors in the generation of acute, subacute, and chronic fatigue. The model separates those that are thought to be risk factors or triggers and those that may be responsible for fatigue perpetuation in CFS. Many of the latter may be amenable to treatment. Factors interact substantially and may be more important in some patients than in others.

also simply be that chronicity predicts chronicity. There is insufficient knowledge about the prognosis of the larger problem of CFS in the community and primary care, but it is likely to be more optimistic. Adequate treatment improves prognosis.

Synthesis

One theme that emerges from the literature of all the fatigue syndromes is the possibility of a general disorder of perception, perhaps of both symptoms and disability. At the heart of this misperception lies the sense of effort. CFS patients clearly experience increased effort in everyday physical and mental tasks, reflected in sense of painful muscle exertion and painful cognitive processing. This increased effort is the not the result of increased neuromuscular or metabolic demands (a Victorian concept), nor does it result in any substantial decline in actual muscle or cognitive performance. The result is a mismatch between patients' evaluation of their physical and mental functioning and the external evidence of any consistent deficits. The basis of this disorder of effort must remain speculative, particularly since the perception of effort is a complex topic. It is possible that it is because the sufferer needs to devote more attention, or even energy, to processes that the rest of us find automatic, be it muscular exertion or mental concentration. Why that should be remains unclear.

From this fundamental problem come other problems identified in CFS, such as increased symptom monitoring, decreased tolerance, and increased anxiety. These are not unique to CFS and have also been described, for example, in fibromyalgia and irritable bowel syndrome. Thus, some centrally mediated disorder of perception of information may underlie the experience of fatigue syndromes and explain the widespread discrepancies between the intensity of symptoms and disability and objective testing of a number of different parameters, both physiological and neuropsychological.

CFS continues to present to clinicians and is associated with high levels of disability and distress. Yet it is an elusive and awkward issue for modern scientific medicine, particularly in the social context in which medicine is now practiced. As outlined previously, the fullest understanding at the present time involves using a multidimensional approach in which due attention is given to both physical and psychological factors and to their interaction. Unitary models of the condition have little scientific basis at present; as with other multifactorial conditions, it is important to consider the roles of predisposing, precipitating, and perpetuating factors. For now, the most effective interventions are aimed at possible perpetuating

factors. As more is understood in the future, it is possible that there will be greater potential for targeting risk factors, the prevention of CFS after identified triggers, and the modification of the biological processes underlying fatigue.

See Also the Following Articles

Cognitive Behavioral Therapy; Fibromyalgia.

Further Reading

- Cleare, A. J. (2003). The neuroendocrinology of chronic fatigue syndrome. *Endocrine Reviews* 24, 236–252.
- Joint Working Group of the Royal Colleges of Physicians Psychiatrists General Practitioners. (1996). *Chronic fatigue syndrome: report of a committee of the Royal Colleges of Physicians, Psychiatrists and General Practitioners*. London: Royal Colleges of Physicians.
- Komaroff, A. and Fagioli, L. (1996). Medical assessment of fatigue and chronic fatigue syndrome. In: Demitrack, M. & Abbey, S. (eds.) *Chronic fatigue syndrome: an integrative approach to evaluation and treatment*, pp. 154–184. New York: Guildford Press.
- Lyall, M., Peakman, M. and Wessely, S. (2003). A systematic review and critical evaluation of the immunology of chronic fatigue syndrome. *Journal of Psychosomatic Research* 55, 79–90.
- Moss-Morris, R., Petrie, K., Large, R. and Kydd, R. (1996). Neuropsychological deficits in chronic fatigue syndrome: artifact or reality? *Journal of Neurology Neurosurgery and Psychiatry* 60, 474–477.
- Price, J. R. and Couper, J. (2001). Cognitive behaviour therapy for chronic fatigue syndrome in adults (Cochrane Review). In: Oakley-Browne, M., Churchill, R., Gill, D., Trivedi, M. & Wessely, S. (eds.) *Depression, anxiety and neurosis module of the Cochrane Database of Systematic Reviews*, Issue 4. Oxford: Update Software.
- Reeves, W., Lloyd, A., Vernon, S., et al. (2003). Identification of ambiguities in the 1994 Chronic Fatigue Syndrome Research Case Definition and recommendations for resolution. *BioMed Central Health Services Research* 3, 25.
- Reid, S., Chalder, T., Cleare, A., Hotopf, M. and Wessely, S. (2005). Chronic fatigue syndrome. *Clinical Evidence* 14, 1366–1378.
- Wessely, S. (1994). The history of chronic fatigue syndrome. In: Straus, S. (ed.) *Chronic fatigue syndrome*, pp. 41–82. New York: Mark Dekker.
- Wessely, S. (1995). The epidemiology of chronic fatigue syndrome. *Epidemiologic Reviews* 17, 139–151.
- Wessely, S., Hotopf, M. and Sharpe, M. (1998). *Chronic fatigue and its syndromes*. Oxford: Oxford University Press.
- Wessely, S., Nimnuan, C. and Sharpe, M. (1999). Functional somatic syndromes: one or many? *Lancet* 354, 936–939.

- White, P. (1997). The relationship between infection and fatigue. *Journal of Psychosomatic Research* 43, 346–350.
- Whiting, P., Bagnall, A. M., Sowden, A. J., Cornell, J. E., Mulrow, C. D. and Ramirez, G. (2001). Interventions

- for the treatment and management of chronic fatigue syndrome: a systematic review. *Journal of the American Medical Association* 286, 1360–1368.
- Wright, J. and Beverley, D. (1998). Chronic fatigue syndrome. *Archives of Disease in Childhood* 79, 368–374.

Chronic Social Stress: GR Sensitivity in Leukocytes

A Weizman and B Rotberg

Sackler Faculty of Medicine, Tel-Aviv University,
Tel-Aviv, Israel

© 2007 Elsevier Inc. All rights reserved.

Stress and the Immune System
Glucocorticoid Receptors Resistance Model
Clinical Implications

Glossary

<i>Anti-depressants</i>	Drugs used to treat depression. Their main mode of action is presumed to be related to elevation of synaptic monoamine levels.
<i>Cytokines</i>	Proteins that are secreted from immunocompetent cells. They act as ligands to transmembrane receptors on immune cells and other tissues to modulate the immune response. Examples are interleukins or tumor necrosis factor.
<i>Glucocorticoid receptors (GRs)</i>	Intracellular receptors with high affinity to glucocorticoids. They belong to the superfamily of nuclear steroid receptors such as thyroid hormone, estrogen, and testosterone. After the glucocorticoid binds to its receptor, the newly formed glucocorticoid–receptor complex translocates into the cell nucleus, where it stimulates or inhibits transcription of genes. Among other things, it modulates transcription of cytokines.
<i>Leukocytes</i>	White blood cells, consisting of granulocytes, lymphocytes, and monocytes. They are immunocompetent cells that act as a defense against infection and progression of malignant cells.
<i>Receptor sensitivity</i>	The responsiveness of a cellular system to a ligand. It reflects the binding of the ligand to its receptor and the effectiveness of the signal transduction.

Stress and the Immune System

Stress is an unavoidable part of human life. Be it infection, physical trauma, or social stress, the body handles the challenges of stress in similar ways, by activating the hypothalamic-pituitary-adrenocortical (HPA) axis and the autonomic sympathetic nervous system. This activation leads to increased levels of circulating glucocorticoids and adrenaline. Leukocytes have glucocorticoid and adrenergic receptors and respond to their signal. The leukocytes change positions (marginate to blood vessel walls or redistribute to lymph nodes and to the skin) and secrete pro-inflammatory cytokines such as interleukin-1 (IL-1), IL-6, and tumor necrosis factor- α (TNF- α). If an immunological challenge is not met, this process ends after 3 to 5 days. Glucocorticoids that are involved in activating this pro-inflammatory process also serve as the signal to stop it. They exert this effect via two pathways: negative feedback at the HPA axis and direct immunosuppressive effect on leukocytes.

Social stress is frequently a chronic experience, taking place day after day for weeks, months, or even years. Examples are work overload, unemployment, living in a violent neighborhood, and caring for a loved one with a chronic disease. Our understanding of the influence of chronic social stress on the immune system has changed over the years. Initially, an immunosuppressive model was suggested. Individuals exposed to chronic social stress were reported to be more prone to infectious or neoplastic diseases. The explanation was that excess glucocorticoid levels cause an immunosuppressive state, similar to that found in subjects taking high doses of steroids. A limitation of that model was that it did not explain the observation that people under chronic stress are also more susceptible to allergic, autoimmune, and cardiovascular diseases, that is, disorders related to overactive immunity. That led to a model that emphasizes dysregulation of the immune system, in which

some functions are dampened while others are enhanced. A proposed explanation involves different effects of chronic stress on two subtypes of lymphocytes, T helper 1 (TH1) and T helper 2 (TH2). TH1 secretes cytokines (IL-2 and interferon- γ) that promote function of natural killer cells and cytotoxic T cells. Those cells secrete free radicals to destroy and phagocyte pathogens. TH2 secretes different cytokines (IL-4 and IL-10) that stimulate B cells to produce antibodies. Chronic stress suppresses TH1 but not TH2, thus causing suppression of cellular immunity but activation of humoral immunity. Excess humoral activity might lead to autoimmune disorders and allergy. As mentioned earlier, glucocorticoids that are involved in initiating this complex process serve also as the signal to stop it. With chronic social stress, it was demonstrated that leukocytes are less sensitive to glucocorticoid signaling to stop inflammation. The explanation for this is the glucocorticoid receptors (GRs) resistance model.

Glucocorticoid Receptors Resistance Model

Empirical Evidence

Several studies have demonstrated that the leukocytes of people exposed to chronic social stress have decreased sensitivity to the immunosuppressive effects of glucocorticoids. These studies used *in vitro* assays to compare the concentration of synthetic glucocorticoids needed to suppress the production of pro-inflammatory cytokines by leukocytes of subjects exposed to chronic social stress with controls. Some of the studies also assessed the steroid concentration needed to suppress leukocyte proliferation. Parents of children with cancer, exhausted industrial employees, and elderly caregivers of dementia patients were found to have increased resistance of leukocytes to the anti-inflammatory effects of glucocorticoids as compared to their respective control groups. Their leukocytes secreted high levels of pro-inflammatory cytokines despite incubation with high doses of glucocorticoids. Higher levels of glucocorticoids were needed to suppress leukocyte proliferation in stressed subjects as compared to unstressed subjects.

Pathophysiology

The mechanism that leads to glucocorticoid resistance in leukocytes is not fully understood. One explanation may be that chronic social stress leads to elevated levels of circulating glucocorticoids and to downregulation of leukocyte GR density. Another explanation, supported by several studies, is that

pro-inflammatory cytokines lead to decreased function of GRs. One pathway is by inhibiting GR translocation to the nucleus, and another one is induction of nonfunctioning isoforms of GR (GR-beta). Thus, a vicious cycle is created and a pro-inflammatory state is maintained even after the stress ends.

Clinical Implications

Stress is a fundamental part of human life. Chronic social stress is recognized as a common phenomenon with detrimental consequences. It has been linked to increased probability for infections, cancer, autoimmune disorders, allergy, cardiovascular disease, and drug abuse. Not everyone is vulnerable to the same degree to such challenge. Some of the differences are constitutional, like age, sex, or genetic make-up. Stress occurs when we perceive the demands that we are facing to be beyond our resources; thus, enhancing our resources may protect us. For example, social support was found to act as a buffer to increased GR sensitivity in parents of children with cancer. Improving our coping skills may help us handle stressful challenges. A promising direction of research on buffers to the detrimental outcomes of stress evolved from studies on major depression. Many patients suffering from major depression have been found to have dysregulation of the HPA axis, as do subjects confronting chronic social stress. This dysregulation is manifested through increased secretion of cortisol and decreased sensitivity of GR so that the net effect is probably too little glucocorticoid signaling. About half of the patients with major depression fail to suppress secretion of cortisol following an external dose of a synthetic steroid, namely, dexamethasone. This nonsuppression in dexamethasone suppression test (DST) is a predictor of poor prognosis, pointing to a high probability of relapse. The mechanism leading to nonsuppression in DST is probably related to decreased brain GR sensitivity. Nonsuppression in DST was correlated to elevated levels of pro-inflammatory cytokines, thus probably reflecting decreased sensitivity of GRs in leukocytes as well. High levels of pro-inflammatory cytokines such as IL-1 β , IL-6, and TNF- α were linked to increased risk of cardiovascular diseases, diabetes, osteoporosis, and brain cortical atrophy, observed in patients suffering from major depression. Treatment with antidepressants (AD) was reported to enhance glucocorticoid signaling by increasing GR number and function. The mechanisms involved in this response include induction of GR transcription as well as promoting translocation of GRs from the cytoplasm to the nucleus. Accumulating data indicate that the cyclic AMP (cAMP)/protein

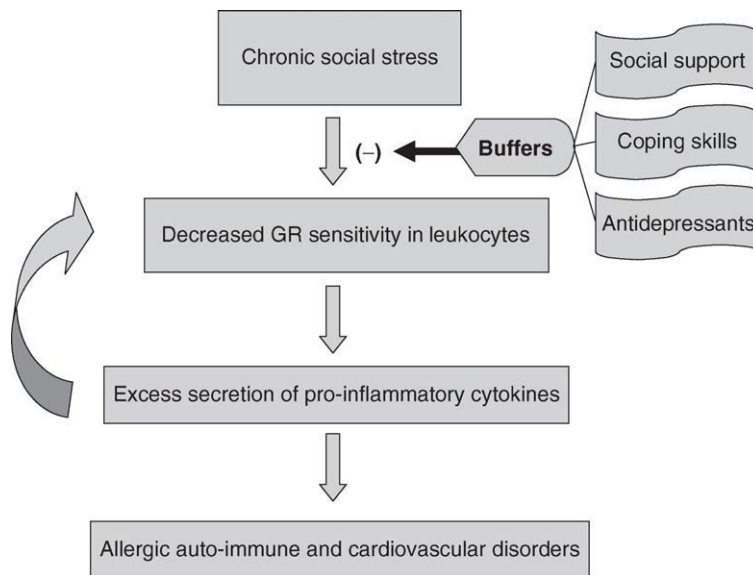


Figure 1 Buffers to stress's detrimental outcomes.

kinase A (PKA) pathway is a key factor in the regulation of GR sensitivity. Phosphorylation of GR by PKA leads to higher sensitivity of the receptor, and hypoactivity of that pathway may be linked to GR resistance in major depression. Thus, ADs may enhance GR function by increasing phosphorylation of GR by the cAMP/PKA cascade.

Several studies involving glucocorticoid agonists as well as glucocorticoid antagonists were shown to have some efficacy for patients with major depression. These advancements in our understanding might lead the way to better prevention and treatment of people facing chronic social stress (Figure 1).

See Also the Following Articles

Antidepressant Actions on Glucocorticoid Receptors; Social Stress, Animal Models of; Posttraumatic Stress Disorder – Clinical; Cytokines, Chronic Stress, and Fatigue.

Further Reading

Bamberger, C. M., Schulte, H. M. and Chrousos, G. P. (1996). Molecular determinants of glucocorticoid receptor function and tissue sensitivity to glucocorticoids. *Endocrine Reviews* 17, 245–261.

Bauer, M. E., Vedhara, K., Perks, P., Wilcock, G. K., Lightman, S. L. and Shanks, N. (2000). Chronic stress in caregivers of dementia patients is associated with

reduced lymphocyte sensitivity to glucocorticoids. *Journal of Neuroimmunology* 103, 84–92.

Carmine, M. P. (2006). The glucocorticoid receptor: part of the solution or part of the problem? *Journal of Psychopharmacology* 20(Supplement 4), 79–84.

McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* 338, 171–179.

Miller, G. E., Cohen, S. and Ritchey, A. K. (2002). Chronic psychological stress and the regulation of pro-inflammatory cytokines: a glucocorticoid-resistance model. *Health Psychology* 21, 531–541.

Raison, C. L. and Miller, A. H. (2003). When not enough is too much: the role of insufficient glucocorticoid signaling in the pathophysiology of stress-related disorders. *American Journal of Psychiatry* 160, 1554–1565.

Segerstrom, S. C. and Miller, G. E. (2004). Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry. *Psychology Bulletin* 130, 601–630.

Webster, J. C., Oakley, R. H., Jewell, C. M. and Cidlowski, J. A. (2001). Proinflammatory cytokines regulate human glucocorticoid receptor gene expression and lead to the accumulation of the dominant negative beta isoform: a mechanism for the generation of glucocorticoid resistance. *Proceedings of the National Academy of Sciences USA* 98, 6865–6870.

Wirtz, P. H., von Kanel, R., Schnorpfeil, P., Ehlert, U., Frey, K. and Fischer, J. E. (2003). Reduced glucocorticoid sensitivity of monocyte interleukin-6 production in male industrial employees who are vitally exhausted. *Psychosomatic Medicine* 65, 672–678.

Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress

M P Antoch and R V Kondratov

Lerner Research Institute, Cleveland Clinic Foundation,
Cleveland, OH, USA

© 2007 Elsevier Inc. All rights reserved.

Anticancer Therapy and Genotoxic Stress

The Circadian Clock System: A Universal Timing
Mechanism

Circadian Clock System Modulates *in Vivo* Response to
Genotoxic Treatments

Glossary

Chronotherapy Timed delivery of therapeutic drugs at empirically established periods of higher efficacy and lower toxicity to the organism.

Circadian rhythms The 24-h periodicities in various biological processes (behavioral, physiological, and biochemical) that have evolved in response to the 24-h periodicity of the Earth's rotation. Controlled by an internal clock mechanism, these rhythms can persist in the absence of environmental cues (for example, in constant darkness), displaying oscillations with periodicities close to 24 h. They can also respond to external stimuli (light, in particular), which coordinate the phases of the internal clock and the light-dark cycle.

Genotoxic stress Condition induced in cells by exposure to different types of external and internal potentially harmful agents, which results in DNA damage and is recognized by the organism's stress response systems.

Molecular circadian oscillator Cell-autonomous oscillator that is formed by the products of core clock genes, which exhibit the 24-h periodicity in their expression/activity.

Peripheral oscillators Circadian oscillators that operate in peripheral (nonneural) tissues and generate rhythmic output in physiology.

Suprachiasmatic nucleus (SCN) The anatomical site of the mammalian circadian master clock. The SCN is composed of ~20 000 neurons residing in the anterior of the hypothalamus. The individual SCN neurons are capable of generating the 24-h periodicities in electrical and metabolic activity and are responsible for synchronization of peripheral circadian oscillators.

Anticancer Therapy and Genotoxic Stress

In the course of their lifetime, all cells, tissues, and entire organisms may be exposed to a variety of harmful agents that cause DNA damage (genotoxic stress). The endogenous sources of DNA-damaging agents result from normal cellular metabolism and include free radicals and peroxides. The most common sources of exogenous genotoxic agents are ultraviolet light, ionizing radiation, oxidative stress, and various chemical mutagens. DNA damage may also result from routine errors in DNA replication and recombination. To prevent detrimental effects on the genome, cells have evolved a response system that induces cell cycle arrest to allow sufficient time to repair the incurred damage, activates the appropriate DNA repair pathway, or, in the case of irreparable damage, induces apoptosis. Thus, response to genotoxic stress is a rather complex process that includes induction of multiple genes associated with cell cycle control, replication, DNA repair, apoptosis, signal transduction, etc. As a result, the cell's fate after genotoxic stress is determined by the superposition of multiple stress response pathways and depends on cell type, cell status, the level of damage, and environmental conditions.

Understanding the molecular mechanisms of cellular response to DNA-damaging agents as well as developing the tools for manipulating stress response have important medical applications, especially in the field of clinical oncology. Thus, two major types of existing anticancer treatments – chemotherapy and radiation – rely on induction of various types of damage to biological macromolecules in tumor cells. Because many of them specifically target DNA, these treatments induce genotoxic stress, which consequently induces the cellular stress response system. Though designed to selectively kill tumor cells, both chemotherapy and radiation are accompanied by nonspecific damage to normal cells and tissues, resulting in undesirable side effects.

One of the approaches that were developed to minimize these side effects was based on the observation that *in vivo* response to various therapies can display significant variation depending on the time of treatment. It has been suggested that these variations may result from differential daily response of both tumor and normal cells, and the final outcome reflects superposition of these effects. While the data on daily variations in therapeutic response of tumors are still

sporadic and sometimes inconsistent, daily variations in drug and radiation responses of normal tissues have been studied in more detail. Testing the effect of timed administration of more than 30 anticancer drugs in mice and rats clearly demonstrated that the drug's toxicity varied significantly throughout the day. In addition, several clinical studies indicate that chronomodulated drug administration results in a significant decrease in complications associated with certain types of cancer treatments while maintaining the same therapeutic effect. These variations in the organism's response to anticancer therapy may depend on different parameters: daily variations in the pharmacokinetics of a particular medicine, blood and tissue concentration of various growth factors, cytokines, and hormones, as well as daily variations in the activity of the intracellular genotoxic stress response system. The chronotherapeutic approach takes advantage of natural daily variability in the physiological response of an organism, which, along with many other parameters, is controlled by the circadian clock system.

The Circadian Clock System: A Universal Timing Mechanism

In mammals, various physiological parameters (such as the sleep–wake cycle, endocrine function, blood pressure, energy metabolism, and behavior) show a nearly 24-h rhythm that is controlled by an internal circadian clock system. This system is capable of generating periodicities even in the absence of environmental cues and at the same time responds to various external stimuli to synchronize the phases of the internal clocks with the environment.

The mammalian circadian clock is organized as a hierarchical network of numerous oscillators, with the master clock residing in the neurons of the hypothalamic suprachiasmatic nucleus (SCN). The master clock in the SCN synchronizes the phases of multiple peripheral clocks throughout the body with the natural 24-h light–dark cycle. Importantly, these peripheral clocks can respond to other synchronizing stimuli (for example, the liver's clock can be coordinated with the food intake schedule, thus functioning independently of the SCN). The ubiquitous nature of the clock mechanism and its ability to respond to a variety of external stimuli underscore the important role the circadian clock system plays in controlling different physiological parameters, including the response to genotoxic stress.

Insights on the mechanistic aspects of global circadian regulation were made possible by recent progress in understanding the functional organization of the molecular clock. It is now well established that, at

the molecular level, the mammalian circadian clock in all types of cells is formed by interdependent transcription/translation feedback loops (Figure 1). Two basic helix-loop-helix (bHLH) PAS-domain transcription factors, CLOCK and BMAL1, form the positive elements of the central oscillatory loop. The CLOCK/BMAL1 heterodimer binds to E box elements in the promoters of target genes and drives rhythmic transcription of three *Period* (*Per1*, *Per2*, and *Per3*) and two *Cryptochrome* (*Cry1* and *Cry2*) genes. When PER and CRY proteins are translated, they are translocated to the nucleus, where they act as potent inhibitors of CLOCK/BMAL1-mediated transcription. CLOCK/BMAL1 also controls transcription of *Rev-erb α* , which feeds back to periodically repress expression of *Bmal1*, thus regulating a positive limb of the feedback loop. Positive and negative elements of these regulatory loops function to coordinate transcriptional events at appropriate times to provide 24-h periodicities. In addition, several post-translational modifications (phosphorylation and sumoylation) play an important role in regulating clock proteins' stability, cellular localization, and functional activity.

The basal transcriptional nature of the core clock mechanism suggests that its control of the output signaling, which translates into daily variations in physiological and metabolic changes, is also transcription based. Indeed, large-scale transcriptional profiling studies have demonstrated that a large number of genes display 24-h periodicities in their mRNA abundance, suggesting that they function under direct (through the E box elements in their promoter regions) or indirect (through the intermediate targets)

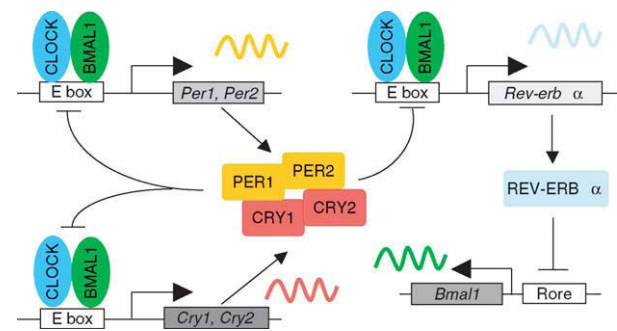


Figure 1 Simplified schematic representation of the mammalian circadian autoregulatory feedback loop. Positive elements of the loop (CLOCK and BMAL1) dimerize to activate rhythmic transcription of *Per* and *Cry* genes through E box enhancer elements. The nuclear-localized CRY and PER proteins interact with CLOCK and BMAL1 to negatively regulate CLOCK/BMAL1-mediated transcription. *Rev-erb α* transcription is regulated by the same components that control *Per* and *Cry* transcription, and the resulting circadian accumulation of REV-ERB α leads to periodic repression of *Bmal1* transcription.

control of the circadian clock system. Interestingly, circadian transcriptional control demonstrates remarkable tissue specificity, suggesting that additional profiling of drug- and radiation-sensitive tissues may identify critical stress response genes that function under circadian control.

Circadian Clock System Modulates *in Vivo* Response to Genotoxic Treatments

The establishment of the functional link between the circadian and stress response systems was made possible by the availability of various circadian mutant mouse models. All circadian mutant mice display a common phenotype manifested by behavioral arrhythmicity and disrupted rhythmic gene expression. However, at the molecular level they represent two extreme states of circadian clock function in regard to CLOCK/BMAL1 transcriptional activity. The *Clock* mutant and *Bmal1*-deficient mice are characterized

by constant low levels of CLOCK/BMAL1 transcriptional activity, while mice with a deficiency of both CRYs contain CLOCK/BMAL1 in its maximally active state. The comparison of the sensitivity of these three types of mutant mice to the anticancer drug cyclophosphamide has shown a direct correlation of survival with the CLOCK/BMAL1 functional status. Thus, when compared to wild-type animals, mice with a deficiency in the positive components of the circadian transcriptional loop (CLOCK and BMAL1) are more sensitive to drug-induced genotoxic stress, while animals with a deficiency in negative elements (CRY1 and CRY2) are more resistant. At the same time, wild-type animals demonstrate daily variations in sensitivity to cyclophosphamide-induced genotoxic stress, which also correlates with normal daily variation in CLOCK/BMAL1 activity (Figure 2). The fact that these arrhythmic mutants display opposite phenotypes suggests that sensitivity to genotoxic stress does not depend on persistence of

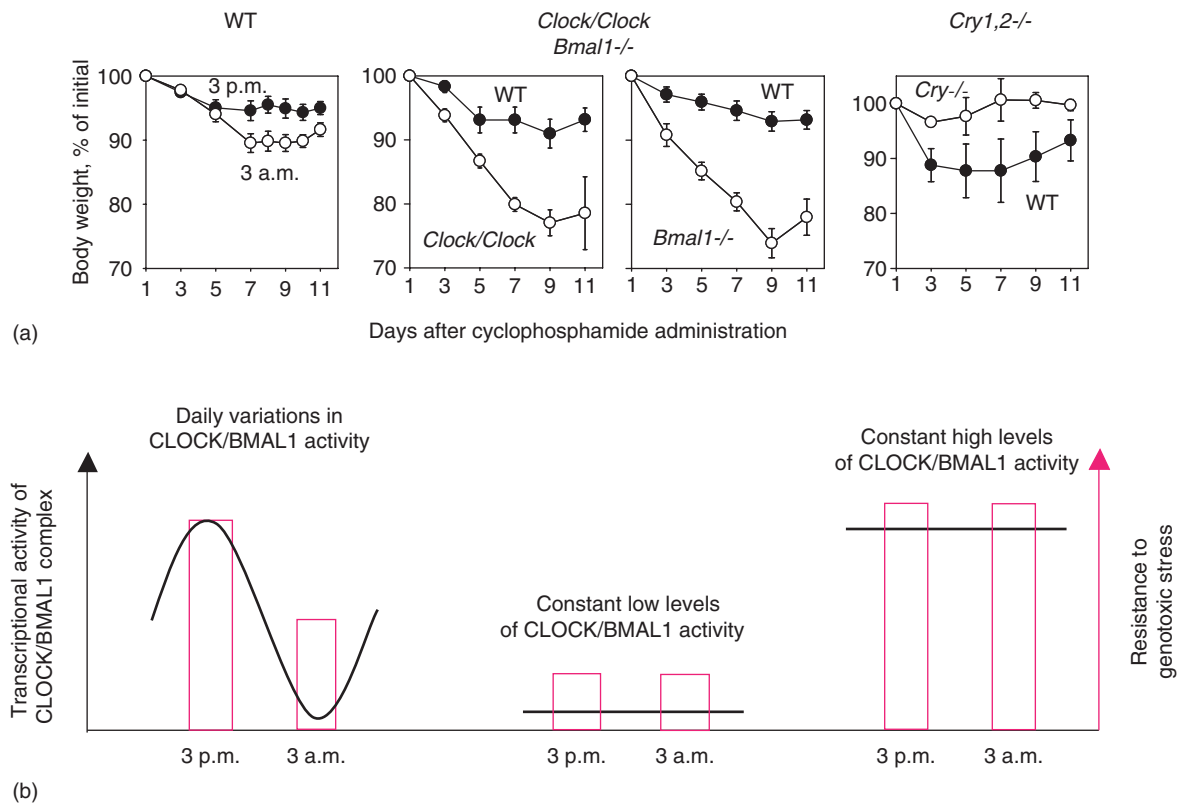


Figure 2 Animals' *in vivo* sensitivity to genotoxic stress correlates with the functional activity of CLOCK/BMAL1 transcriptional complex. (a) Body weight loss of the wild-type and circadian mutant mice after cyclophosphamide treatment. Wild-type animals injected at 3 a.m. demonstrate greater sensitivity to the drug-induced toxicity compared to their littermates injected at 3 p.m. *Clock/Clock* and *Bmal1*^{-/-} animals demonstrate greater sensitivity to cyclophosphamide at all times tested. *Cry1*^{-/-}*Cry2*^{-/-} mice are more resistant to the treatment. (b) Schematic representation of the correlation between daily or locus-specific variations in CLOCK/BMAL1 functional activity (black line) and resistance to the treatment (red bars) in animals of various circadian genotypes.

rhythmicity per se, but rather is guided by the activity of the major circadian transactivation complex.

Mutants with a deficiency of the PERIOD protein represent another example of the functional link between the circadian and genotoxic stress response systems. These mice display higher sensitivity *in vivo* to long-term effects of low doses of ionizing radiation, while the primary thymocytes from these animals are more resistant to high doses of radiation in short-term experiments. Taken together, these data suggest that circadian variations in an organism's sensitivity to various types of genotoxic treatments are determined by the rhythms in expression and activity of the core circadian proteins.

How can rhythms in the activity of circadian proteins modulate response to genotoxic stress? It is well known that different types of cellular stress lead to induction of multiple transcription factors (i.e., p53, NF- κ B or AP1 induction by radiation, induction of NF- κ B during the immune response, or HIF1 α in response to hypoxia). They in turn activate downstream transcriptional targets, encoding factors important for cell death or survival. The balance of these factors will determine the cellular and organism response to genotoxic treatments. Since the circadian clock controls the output at the transcriptional level, some of the responsive genes may be direct targets of CLOCK/BMAL1. Although not yet confirmed, this view is supported by the fact that the CLOCK/BMAL1 complex directly controls such important regulators of transcription as DBP, REV-ERB α , and ROR α . In addition, the CLOCK/BMAL1 complex may display dual transcriptional activity, acting both as transcriptional activator and transcriptional repressor, depending on its interaction with CRYs. Importantly, while functioning as a repressor, it could interfere with the ability of other transcription factors to activate the same promoter. Thus, those genes that are controlled by both the circadian and stress response systems will be induced in response to genotoxic treatments only at those times of day when CLOCK/BMAL1 is functioning as an activator and

will be attenuated at the times when it functions as a repressor. As a result, the balance between pro-death and pro-survival pathways will vary depending on the status of the circadian transcriptional complex. Importantly, these recent findings identify CLOCK/BMAL1 as a potential target for a rational search for pharmacological modulators of the organism's response to genotoxic stress that may potentially be used in clinical practice to minimize the side effects of cancer treatment.

See Also the Following Articles

Cancer Treatment; Circadian Rhythms, Effects of Prenatal Stress in Rodents; Circadian Rhythms, Genetics of.

Further Reading

- Antoch, M., Kondratov, R. and Takahashi, J. (2005). Circadian clock genes as modulators of sensitivity to genotoxic stress. *Cell Cycle* 4, 901–907.
- Fu, L., Pelicano, H., Liu, J., Huang, P. and Lee, C. (2002). The circadian gene Period2 plays an important role in tumor suppression and DNA damage response in vivo. *Cell* 111, 41–50.
- Gorbacheva, V., Kondratov, R., Zhang, R., et al. (2005). Circadian sensitivity to the chemotherapeutic agent cyclophosphamide depends on the functional status of the CLOCK/BMAL1 transactivation complex. *Proceedings of the National Academy of Sciences USA* 102, 3407–3412.
- Hrushevsky, W. J. M. (1994). *Circadian cancer therapy*. Boca Raton, FL: CRC Press, Inc.
- Levi, F. (2006). Chronotherapeutics: the relevance of timing in cancer therapy. *Cancer Causes Control* 17, 611–621.
- Matsuo, T., Yamaguchi, S., Mitsui, S., Emi, A., Shimoda, F. and Okamura, H. (2003). Control mechanism of the circadian clock for timing of cell division in vivo. *Science* 302, 255–259.
- Panda, S., Antoch, M. P., Miller, B. H., et al. Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell* 109, 307–320.
- Sehgal, A. (2003). *Molecular biology of circadian rhythms*. Hoboken, NJ: John Wiley & Sons, Inc.

Circadian Rhythm Effects on Cardiovascular and Other Stress-Related Events

R Manfredini, B Boari, R Salmi, A M Malagoni and F Manfredini

Vascular Diseases Center, University of Ferrara, Italy

© 2007 Elsevier Inc. All rights reserved.

Chronobiology and Circadian Rhythms
Circadian Rhythm of Cardiovascular Events
Circadian Rhythm of Underlying Risk and Triggering Factors
Stress and Circadian Rhythms
Circadian Rhythms and Possible Implications for Treatment of Cardiovascular Diseases

Glossary

<i>Aortic aneurysm</i>	Dilatation or ballooning of the wall of thoracic or abdominal aorta.
<i>Chronobiology</i>	The study of biological systems as affected by time.
<i>Circadian rhythm</i>	Biological rhythm with regular recurrence in cycles of about 24 h.
<i>Coagulation</i>	Process of interaction of blood coagulation factors to form a fibrin clot.
<i>Hypertension</i>	Persistently high systemic arterial blood pressure.
<i>Myocardial infarction</i>	Necrosis of the myocardium, as a result of interruption of blood supply.
<i>Stroke</i>	Sudden and severe cerebral attack, ischemic or hemorrhagic.

Chronobiology and Circadian Rhythms

Chronobiology is a branch of biomedical sciences devoted to the study of biological rhythms. Biological rhythms exist at any level of living organisms and, according to their cycle length, may be divided into three main types: (1) circadian rhythms (from the Latin *circa diem*, characterized by a period of approximately 24 h), (2) ultradian rhythms (period shorter than 24 h, e.g., hours, minutes or even seconds), and (3) infradian rhythms (period longer than 24 h, e.g., days, weeks, months). Circadian rhythms are the most commonly and widely studied biological rhythms.

The cardiovascular system is organized according to a specific temporal order that is oscillatory in nature, and most cardiovascular functions exhibit circadian changes. Such predictable-in-time differences in the physiological status of the cardiovascular system

give rise to rhythmic variations in the susceptibility of human beings to morbid and mortal events. On the other hand, the pathological mechanisms of cardiovascular disease themselves exhibit temporal changes in both their manifestation and severity, leading to predictable-in-time differences in their ability to precipitate and graduate the overt expression of disease. A growing body of evidence suggests that the occurrence of cardiovascular events is not evenly distributed in time, but shows peculiar temporal patterns that vary with time of the day, the day of the week, and the month of the year. These patterns coincide with the temporal variation in the pathophysiological mechanisms that trigger cardiovascular events and the physiological changes in the body rhythms. These contribute to the definition of the new concept of chronorisk, in which several factors, not harmful if taken alone, are capable of triggering unfavorable events when presenting all together within the same temporal window.

Circadian Rhythm of Cardiovascular Events

Myocardial Ischemia and Infarction

Coronary heart disease (CHD) is the number one cause of death worldwide, leading to more than 7 million deaths annually. CHD occurs when the coronary arteries are unable to supply adequate oxygenated blood to myocardial tissue, leading to myocardial ischemia and possible acute myocardial infarction (AMI). A reversible injury to myocardial tissue leads to ischemia, while AMI represents cell death after prolonged ischemia. The development of AMI is based on two major components. The first is vascular injury and/or dysfunction with loss of a non-thrombogenic endothelial surface, most frequently due to atherosclerotic plaques. In response to intraluminal stress, for example, the circadian periodic or incidental rise in blood pressure, plaques may rupture and release thrombogenic factors into the bloodstream. The second is activation of the blood-clotting mechanism beyond the capacity of the simultaneously activated fibrinolytic mechanisms to remove the fibrin formed and prevent the formation of a clot, thus leading to occlusion of the vessel.

As for circadian variation in onset, both stable coronary disease and ischemic attacks with ST segment

elevation show a peak in the morning hours, particularly in the first few hours after waking and commencing activity. Myocardial infarction shows an evident circadian rhythm as well, characterized by a higher frequency of onset during morning hours. It has been estimated that the incidence rate of AMI onset is 40% higher in the morning period than throughout the rest of the day, and nearly 28% of morning infarctions and 22% of sudden cardiac deaths (accounting for approximately 9% and 7% of all AMIs and sudden deaths, respectively) are attributable to the morning excess. Compared with all other times of the day, morning shows a relative risk of AMI of 1.4, rising to a threefold increase in the first 3 h after waking, when almost 25% of all AMIs occur. A series of studies has indicated that this circadian pattern in the incidence of AMI was similar in different subgroups of patients and was independent of gender, age, and clinical characteristics, for example, lower or higher ejection fraction, presence of one-, two-, or three-vessel disease, type of infarction (Q-wave or non-Q-wave), and site of infarction (anterior or inferior). Diabetes mellitus, however, represents a peculiar entity. Diabetic subjects, who are more at risk of AMI compared with nondiabetics, show attenuation of morning peak or, in some cases, the absence of a significant circadian variation of AMI onset. It is thus possible that diabetic subjects are exposed to a steady and high risk of myocardial infarction during the 24 h, maybe due to a marked diminution of parasympathetic activation during sleep.

Drugs may also suppress diurnal variation. Specifically, an absence or attenuation of the morning peak of AMI has been observed in patients taking β -blockers. Additionally, in those taking aspirin, the primary morning peak of infarction was attenuated: there was an additional 25% protective effect associated with aspirin specifically between 4 and 10 a.m., the hours of awakening from sleep and assuming the upright posture.

Stroke

Stroke is the end result of a heterogeneous group of vascular diseases that are manifested by ischemia or hemorrhage of one or more vessels of the brain. Worldwide, stroke is the second leading cause of death and the leading cause of disability and long-term care costs. Although the frequency of and mortality from stroke have declined in the United States and many other industrialized countries during the past 30 years, in the United States stroke continues to affect ~700 000 persons and to cause ~165 000

deaths per year, and a similar or even higher incidence of stroke is reported in other countries. A number of studies have indicated that the incidence of all subtypes of stroke (e.g., lacunar, thrombotic, embolic, and hemorrhagic) shows a significant circadian rhythm, characterized by a higher frequency during morning hours. A comprehensive meta-analysis of 31 separate studies in which the clock time of events was verifiable confirmed the prominent 24-h pattern of cerebrovascular events. Thus, for example, the incidence of transient ischemic attack (TIA) and hemorrhagic and ischemic stroke reaches a peak between 6 and 12 a.m. Approximately 55% of all ischemic strokes, 34% of all hemorrhagic strokes, and 50% of all TIAs occurred during this temporal window.

Rupture/Dissection of Aortic Aneurysm

Acute aortic dissection and rupture of aortic aneurysms are serious, life-threatening cardiovascular emergencies that carry high fatality rates. In fact, many patients with either an acute aortic dissection or rupture of an aortic aneurysm die before hospitalization, and data from the large population enrolled in the International Registry of Acute Aortic Dissection (IRAD) show a 24% mortality rate in patients reaching hospitals even in centers with advanced technology and considerable expertise. Also, several studies have found a circadian variation for aortic aneurysm, characterized by an increased risk of both aortic rupture as well as aortic dissection during early hours of morning. Analysis of the IRAD registry showed that this morning peak (around 8–9 a.m.) was independent of gender, type A or B dissection, age $<$ or $\geq$ 70 years, and presence or absence of hypertension.

Circadian Rhythm of Underlying Risk and Triggering Factors

Several variables indicative of cardiovascular and cerebrovascular status exhibit prominent circadian rhythms and may contribute to the circadian pattern of cardiovascular events ([Figure 1](#)).

Vascular Tone and Arterial Blood Pressure

Sympathetic activity is circadian rhythmic, reaching peak level in the morning when vascular receptors exhibit greatest sensitivity, and this can play a role in the sharp morning rise in arterial blood pressure (BP) and heart rate (HR). The morning rise in BP is significantly amplified by biological responses to events that occur upon awakening from nighttime sleep, that is, assumption of upright posture, commencement of physical activity, and exposure to

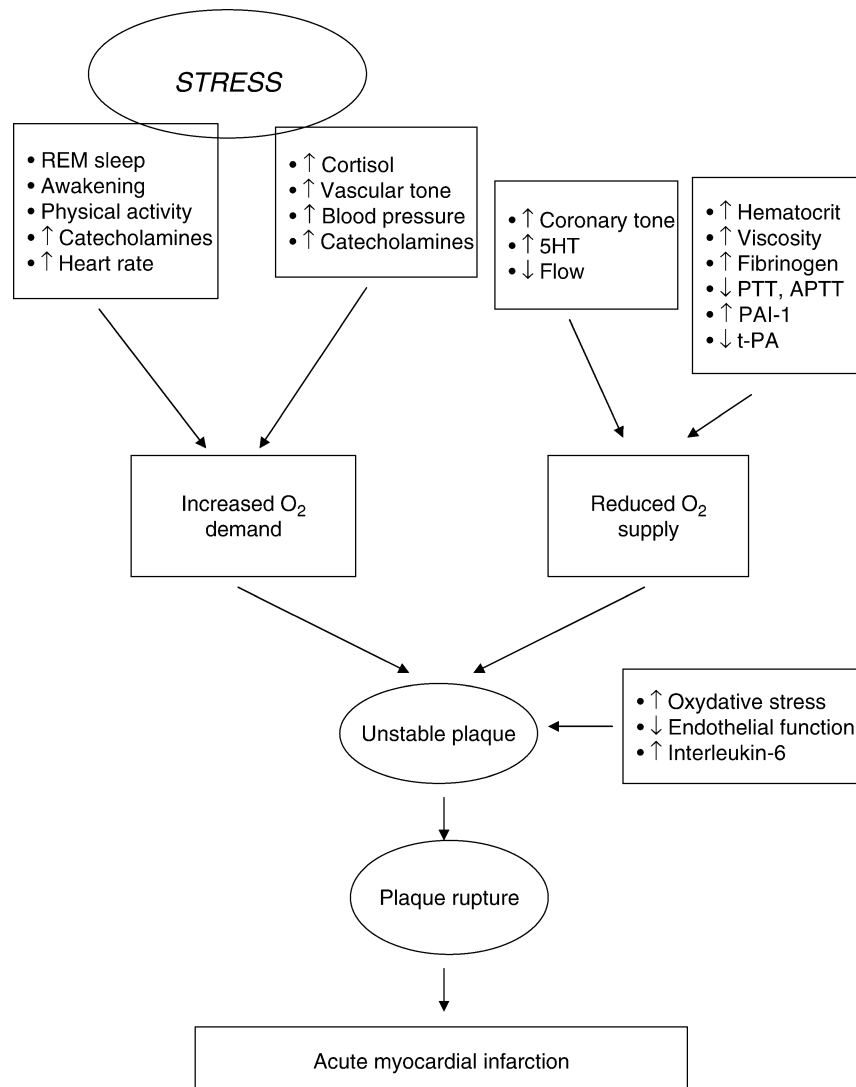


Figure 1 Schematic outline of the way in which several concurring risk factors might contribute to trigger the onset of acute myocardial infarction during morning hours. Abbreviations: REM, rapid eye movement; 5HT, 5-hydroxytryptamine; PTT, partial thromboplastin time; APTT, activated partial thromboplastin time; PAI-1, plasminogen activator inhibitor-1; t-PA, tissue plasminogen activator.

emotional stress, all of which further elevate catecholamine secretion and sympathetic tone. The circadian variation in BP, characterized by a double peak in the early morning and in the early evening, plays an important role in the onset of cardiovascular and cerebrovascular events. Moreover, night BP levels are important as well. In normotension and in uncomplicated essential hypertension, in fact, systolic BP (SBP) and diastolic BP (DBP) normally decline during nocturnal sleep by 10–20% relative to the respective daytime mean levels. Some persons exhibit a so-called nondipping pattern, that is, a less than 10% decline (or even rise) in SBP and/or DBP during nighttime sleep relative to their daytime levels. The cumulative effect over time of abnormally elevated nighttime BP load on target organs, for example,

heart, brain, and kidney, may lead to deleterious effects.

Blood Coagulation and Fibrinolysis

Circadian rhythms in blood coagulation and fibrinolysis are also key components of the multifactor pathogenesis of cardiovascular and cerebrovascular events and their 24-h patterns. The *in vitro* aggregation response of platelets to stimulants such as ADP and epinephrine is circadian rhythmic, and each of these rhythms peaks during the late night or morning hours and declines to the trough in the afternoon. Activation of platelet aggregation is brought about by the increased secretion of catecholamines, which is initiated by the activation of the sympathetic nervous

system during nocturnal episodes of rapid eye movement (REM) sleep. Since REM episodes are more frequent during final hours of nighttime rest in normal sleepers, the rise in catecholamine concentration commences at that time, and increases further in the morning due to baroreflex with the change from supine to upright posture upon awakening and the commencement of physical activity and exposure to emotional stress. Moreover, circadian rhythms have been found in several factors of the coagulation cascade, on the intrinsic as well as the extrinsic and common pathways of blood coagulation. Circadian rhythms in factors II (thrombin), VII (tissue thromboplastin), VIII, and IX, with peak activity in the morning hours, were observed in clinically healthy human subjects.

As a result of these rhythms, the global screening tests of different steps in the coagulation cascade, such as prothrombin time (PT) and activated partial thromboplastin time (aPTT), show the lowest values (indicative of blood hypercoagulability) in the morning. The extent of the differences in PT, aPTT, and thrombin time (TT) was found to be about 10% of the respective means.

Activation of the coagulation cascade is followed in healthy subjects by an increase in fibrinolytic activity, which serves to limit the coagulation process and avoid vascular occlusion. Fibrinolysis also shows an evident circadian rhythmic variation. The circadian rhythm in fibrinolysis, although dependent on multiple influences, is to a large extent determined by the interaction between tissue plasminogen activator (t-PA) and its potent inhibitor, tissue plasminogen activator inhibitor-1 (PAI-1). Both t-PA antigen and PAI-1 antigen are circadian rhythmic, showing peak concentrations in the morning, but the circadian rhythm of t-PA activity is almost 12 h out of phase with the circadian rhythm in t-PA antigen concentration. Thus, PAI-1, the activity of which parallels its concentration, suppresses t-PA activity during the late night and early morning hours when thromboembolic events are most frequent. The circadian rhythm in PAI-1 is of large amplitude, showing excursions as great as 250%, thereby contributing in diurnally active subjects to the relative hypercoagulability of the blood in the morning and relative hypocoagulability of the blood in the evening.

Stress and Circadian Rhythms

At least two major physiological systems play a crucial role in the adaptation of the organism to environmental challenges: the circadian system and the stress response system, which are tightly coupled.

Activation of the hypothalamic-pituitary-adrenocortical (HPA) axis in response to stressful stimuli is

a crucial mechanism that allows an organism to maintain homeostasis and adapt to environmental conditions. It is known that HPA has a circadian activity in basal activity, and the rhythmicity in secretory activity is controlled by the suprachiasmatic nucleus (SCN) of the hypothalamus. However, the SCN is not the only biological clock. In fact, cloning of the clock genes in the late 1990s revealed that clock genes are expressed and oscillate with a circadian rhythm in each organ or cell, suggesting that each of them has its own internal clock. Thus, this clock system has been denominated as a peripheral clock in comparison with the central clock in SCN. It seems that the SCN-based central clock is capable of synchronizing peripheral clock systems.

Stress-induced secretion of corticotropin-releasing factor (CRF) from the paraventricular nucleus (PVN) of the hypothalamus triggers the synthesis and release of ACTH and, ultimately, glucocorticoids. Thus, the hypothalamic PVN receives convergent information from both stress and the circadian clock, and it has been demonstrated that, at least in mice, the mammalian orthologs *Per1*, *Per2*, and *Per3* of the *Drosophila melanogaster* clock gene *Period* are expressed in different areas of the central nervous system (e.g., PVN) and peripheral tissue.

Stress is capable of exerting major effects upon the circulatory system of coronary patients, particularly in the morning. Stress acts on the cardiovascular system in multiple ways, and a direct link has been demonstrated in both laboratory and ambulatory studies between emotional stress and changes in HR, systolic and diastolic BP, left ventricular stroke work, cardiac output, blood viscosity, and coagulation. These and other stress-related changes can trigger angina in coronary syndrome patients and are perceived to act in a primary or secondary way in the triggering cerebrovascular events through the sudden increase in BP and blood coagulation. Acute coronary syndromes may also be triggered by emotional stress, and studies demonstrated that emotional stress preceded the onset of AMI in 11% of men and 16% of women. Anger is important as well, and it has been reported that there is a twofold increased relative risk of acute coronary syndromes in the 2 h after an emotional burst of anger. Mental stress has also been shown to inhibit flow-mediated brachial artery vasodilatation, a validated indicator of endothelial function, and the hemodynamic response to mental stimuli, consisting of the combination of a reduction in blood supply and an increase in cardiac demand, might explain the role of mental stress in the onset of myocardial ischemia. Indeed, the effect of stress on human cardiac, vascular, and hemostatic status appears to be circadian rhythmic, with the effects

being greater in the morning. Experimental studies on rats showed the molecular mechanisms of stress-induced heart attacks, with a rapid activation of the p44/p42 mitogen-activated protein kinase, followed by transient upregulation of the immediate early genes of the smooth muscle cells of the coronary arteries, endothelial cells, and myocardium. It is interesting to observe that the 24-h pattern of acute cardiovascular and cerebrovascular events is tightly coupled in time, suggesting that they share numerous risk factors. An exaggerated cardiovascular response to behavioral challenges has been suggested as a potential factor that enhances the risk of hypertension and CHD. Thus, individuals who exhibit large increases in BP during acute psychological stress are at risk for atherosclerosis, probably secondary to endothelial damage by BP surge, via increased flow turbulence and subsequent release of inflammatory and pro-atherogenic cytokines.

Emotional stress can also play a role in the triggering of stroke. In patients with acute ischemic stroke, the finding that salivary cortisol levels are positively correlated with 24-h and nighttime BP suggests that stress may induce endocrine changes that result in BP increase. In asymptomatic older adults, stress-induced BP reactivity is enhanced in those diagnosed with silent cerebrovascular disease, independently of the resting BP levels, suggesting that mental stress at the start of the day can intensify the morning BP surge and contribute to the higher frequency of strokes at this time.

Circadian Rhythms and Possible Implications for Treatment of Cardiovascular Diseases

In conclusion, it seems that the temporal rhythmic organization of the vascular functions is such that the defense mechanisms against disease are incapable of providing the same degree of protection at all times of the day. A point of interest is the possible influence on chronobiologic rhythms and outcome of cardiovascular events. At present, only a limited amount of data is available on the prognostic implications of the concept of chronorisk. For example, the clock time of onset seems to affect the clinical course and outcome of patients with AMI. Moreover, the concept of chronorisk implies that the delivery of protective or preventive medications should be synchronized in time in proportion to need as determined by established rhythmic patterns in cardiovascular risk, during the 24 h and in a manner that will avert or minimize undesired side effects. A chronotherapeutic approach to cardiovascular diseases has been suggested. For example, a growing interest is given to

morning BP surge as an important risk factor. It is known that some drugs have different effects when administered in the morning compared to in the evening. In fact, it should not be assumed that a long-acting, once-daily antihypertensive drug administered in the morning will have the same antihypertensive effect as when administered in the evening. Thus, the effects of time of drug administration on the pharmacodynamics of antihypertensive drugs are an important consideration for antihypertensive therapeutics. Improved control of the morning BP surge by means of a tailored antihypertensive therapy might provide maximal benefit during the particularly vulnerable time periods.

See Also the Following Articles

Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress; Circadian Rhythms, Genetics of; Heart Disease/Attack; Cardiovascular Disease, Stress and; Ultradian Rhythms.

Further Reading

- Andreotti, F., Davies, G. J., Hackett, D. R., et al. (1988). Major circadian fluctuations in fibrinolytic factors and possible relevance to time of onset of myocardial infarction, sudden cardiac death and stroke. *American Journal of Cardiology* **62**, 635–637.
- Elliott, W. J. (1998). Circadian variation in the time of stroke onset: a meta-analysis. *Stroke* **29**, 992–996.
- Haus, E., Cusulos, M., Sackett-Lundeen, L. and Swoyer, J. (1990). Circadian variations in blood coagulation parameters, alpha-antitrypsin antigen and platelet aggregation. *Chronobiology International* **7**, 203–216.
- Lemmer, B. (2005). Chronopharmacology and controlled drug release. *Expert Opinion on Drug Delivery* **2**, 667–681.
- Manfredini, R., Gallerani, M., Portaluppi, F. and Fersini, C. (1996). Relationships of the circadian rhythms of thrombotic, ischemic, hemorrhagic, and arrhythmic events to blood pressure rhythms. *Annals of the New York Academy of Science* **783**, 141–158.
- Manfredini, R., Gallerani, M., Portaluppi, F., Salmi, R. and Fersini, C. (1997). Chronobiological patterns of onset of acute cerebrovascular diseases. *Thrombosis Research* **88**, 451–465.
- Manfredini, R., Boari, B., Gallerani, M., et al. (2004). Chronobiology of rupture and dissection of aortic aneurysms. *Journal of Vascular Surgery* **40**, 382–388.
- Manfredini, R., Boari, B., Smolensky, M. H., et al. (2005). Circadian variation in stroke onset: identical temporal pattern in ischemic and hemorrhagic events. *Chronobiology International* **22**, 417–453.
- Manfredini, R., Boari, B., Smolensky, M. H., et al. (2005). Seasonal variation in onset of myocardial infarction – a 7-year single-center study in Italy. *Chronobiology International* **22**, 1121–1135.

- Muller, J. E. (1999). Circadian variation and triggering of acute coronary events. *American Heart Journal* 137 (4 Pt 2), S1–S8.
- Portaluppi, F., Manfredini, R. and Fersini, C. (1999). From a static to a dynamic concept of risk: the circadian

- epidemiology of cardiovascular events. *Chronobiology International* 16, 33–49.
- Stone, P. H. (2004). Triggering myocardial infarction. *New England Journal of Medicine* 351, 1716–1718.

Circadian Rhythms, Effects of Prenatal Stress in Rodents

S Maccari

University of Lille 1, Villeneuve d'Ascq, France, and
Università La Sapienza, Rome, Italy

O Van Reeth

Université Libre de Bruxelles, Brussels, Belgium, and
Università Cattolica del Sacro Cuore, Rome, Italy

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
S Maccari and O Van Reeth, volume 1, pp 467–470,

© 2000, Elsevier Inc.

Prenatal Stress, an Early Environmental Change during
a Critical Period

Circadian Rhythms and Adaptation to External
Environment

Influence of Prenatal Stress on Circadian Rhythmicity,
Sleep, and Behavior

Conclusion

Glossary

Circadian rhythms Biochemical, physiological and behavioral processes occurring in a cyclic fashion with a period of about 24 h (*circa die* = about a day).

Free-running Term used to characterize the expression of circadian rhythms in the absence of any 24 hr synchronizing information from the external environment.

Sleep-wake cycle The various vigilance states were characterized as follows. Wake (W): low voltage fast electroencephalogram (EEG) activity, high electromyogram (EMG) activity and numerous eye movements; Light slow wave sleep (SWS1): high voltage slow cortical waves interrupted by low voltage fast waves and reduced EMG activity; Deep slow wave sleep (SWS2): continuous high

amplitude slow wave activity in EEG, very low EMG activity and no electrooculogram (EOG) activity; Paradoxical sleep (PS): low voltage fast cortical waves with a regular theta rhythm, absence of muscular tone and presence of rapid eye movements.

Phase shift

A discrete displacement (advance or delay) of an oscillation (e.g. circadian activity rhythm) along the time axis, in response to a synchronizing stimulus.

Prenatal stress procedure by restraining pregnant females

Pregnant females were individually restrained three times a day (at 09:00, 12:00 and 17:00 h) for 45 min in transparent plastic cylinders (7 cm diameter, 19 cm long) and exposed to bright light. This manipulation was performed daily during the last week of pregnancy until delivery. Control pregnant females were left undisturbed in their home cages.

Prenatal stress procedure by gestational hypoxia

Pregnant females were exposed to hypoxia from the 5th to 20th day of gestation, by housing them in Plexiglas chambers with a normobaric hypoxic atmosphere consisting of 10% O₂ – 90% N₂. Control pregnant females were kept in normoxia (21% O₂).

Prenatal Stress, an Early Environmental Change during a Critical Period

The prenatal environment exerts a profound influence on the development of an organism, inducing changes that extend from early to later life. In animal models of prenatal stress (PNS), it has been shown that adult prenatally stressed rats exhibit abnormalities under stressful conditions, such as increased anxiety-like behaviors, hyperemotionality-like behaviors, or depression-like behaviors. Among the biological

substrates possibly mediating these behavioral disorders, the hypothalamic-pituitary-adrenal (HPA) axis, involved in an animal's ability to cope with stress, appears to be a good candidate. Indeed, PNS induces an increase in stress-induced plasma adrenocorticotrophic hormone (ACTH) levels, a prolongation of stress-induced corticosterone secretion, and a decreased binding capacity of hippocampal corticosteroid receptors. These behavioral and physiological changes seem to depend on the gender of the animals, with substantially more marked effects of PNS in females than in males. Among the neurotransmitters involved in those hormonal and behavioral responses, acetylcholine plays a critical role. Indeed, PNS has long-term effects on the development of forebrain cholinergic systems and induces an increase in hippocampal acetylcholine release after acute stress or corticotropin releasing hormone (CRH) microinjections. Those long-term neuroendocrinological effects are mediated, at least in part, by a stress-induced maternal corticosterone increase during pregnancy.

Taken together, these studies support the idea that PNS disrupts the organism's capabilities to cope with stress and that prenatally stressed rats have biological traits similar to those observed in depressed patients. Indeed, like depressed patients, prenatally stressed rats do escape from the feedback inhibition responsible for returning corticosterone secretion to basal levels after stress. Finally, like depressed patients, in whom cholinergic hyperactivity is described, prenatally stressed rats exhibit cholinergic hypersensitivity after a CRH challenge.

Circadian Rhythms and Adaptation to External Environment

Most biochemical, physiological, or behavioral processes within the organism fluctuate on a regular basis throughout the 24-h day. These daily rhythms arise from an internal time-keeping system, the circadian clock, located in the hypothalamic suprachiasmatic nuclei (SCN). In the absence of any environmental input, these rhythms persist with a period of approximately 24 h and are therefore referred to as circadian rhythms.

The circadian time-keeping system is essential for the optimal functioning and adaptation of the organism to the dramatic changes in the environment imposed by Earth's rotation on its axis. Indeed, the circadian system maintains temporal synchronization not only between the organism and daily changes in the physical environment but also between diverse internal physiological processes. For many years, the circadian system was thought to be sensitive only to changes in the light-dark (LD) cycle. However, it

has been shown that circadian function and/or sleep patterns can also be reset by neurochemical or behavioral stimuli. Among those stimuli, steroids can have marked effects on the functioning of the circadian system, and chronic stress in adult rats can induce changes in sleep patterns and circadian rhythms.

Perinatal events seem to have complex influences on the long-term development of circadian function. For instance, it is known that the abnormal development of visual system in anophthalmic mice has clear effects on the integrity of the SCN and on the development of behavioral rhythms. Restricted access to their natural mother has been shown to shift the endogenous corticosterone rhythm of rat pups. Despite the fact that postnatal events have complex influences on the long-term development of the circadian clock, very little attention has been given to the possible effects of prenatal stress on circadian clock function. Previous studies reported that sustained hypoxia in adults affects the functioning of the internal clock located in the SCN. We hypothesize that gestational hypoxia, found in many pathological conditions in pregnant women and naturally associated with life at high altitude, can cause long-lasting consequences for the synchronization of the circadian clock to light.

Influence of Prenatal Stress on Circadian Rhythmicity, Sleep, and Behavior

We recently showed a phase advance in the timing of the rise of corticosterone in prenatally stressed rats compared to control rats. More precisely, PNS induced higher levels of total and free corticosterone secretion at the end of the light period in both males and females and induced hypercorticism over the entire diurnal cycle in females. The effects of prenatal stress on corticosterone secretion could be mediated, at least in part, by a reduction in hippocampal corticosteroid receptors at specific times of day. Results also show that prepartal stress can alter the pattern of corticosterone secretion in pregnant females. We also showed significant phase advances in the circadian rhythm of locomotor activity relative to the entraining LD cycle in prenatally stressed rats. When subjected to an abrupt shift in the LD cycle, male and female prenatally stressed rats resynchronized their activity rhythm to the new LD cycle slower than control rats. In temporal isolation in constant darkness, the free-running period of locomotor activity was significantly shorter in prenatally stressed rats compared to control rats.

In view of those changes in circadian rhythmicity, we investigated the effects of PNS on the sleep-wake cycle in adult rats. Male prenatally stressed rats

exhibited a significant increase in the amount of paradoxical sleep (PS) over the 24-h recording session, positively correlated to plasma corticosterone levels. Other changes include an increased sleep fragmentation and light slow wave sleep (SWS1) time and a slight decrease in the percentage of deep slow wave sleep (SWS2) relative to total sleep time. The present results demonstrate that exposure to PNS can produce long-term and selective changes in both the structure and continuity of sleep. Although there were preliminary reports of abnormal sleep-like behaviors in prenatally stressed monkeys and prenatally stressed humans, our data provide the first polygraphic demonstration of the long-term effects of PNS on the sleep–wake cycle when the animals reach adulthood.

Using gestational hypoxia as an alternative prenatal stress procedure, we demonstrated that, once they become adults, rats born to hypoxic mothers had significant alterations in their circadian rhythm of locomotor activity (recorded in freely accessible running wheels). Under a regular LD cycle, they showed a phase advance of their rhythm of activity and were less active than control rats. After an abrupt 6-h phase delay in the LD cycle, rats from the prenatal hypoxic group (PNH) took significantly more time to resynchronize to the new LD cycle than the controls. Under constant darkness, PNH and control rats had a similar period of activity, but the response of PNH rats to a light pulse in the early subjective night was less marked than of control rats. When submitted to acute restraint stress, PNH rats had a prolonged secretion of corticosterone compared to controls. These results indicate that prenatal hypoxia is a factor that has long-lasting consequences for the functional output of the biological clock and the hormonal response to stress.

Conclusion

Taken together, our results indicate that PNS or PNH induces an increased stress response and abnormal circadian and sleep functions in adult rats, suggesting an underlying dysfunction of the circadian clock. Our results parallel, to some extent, circadian and sleep changes found in depressed patients. Indeed, abnormalities in a variety of overt circadian rhythms, including the cortisol rhythm, have been documented in depressed patients. Also, alteration in the sleep–wake cycle is a hallmark of human depression, with changes including a shortened PS latency, an increase in the amount and frequency of PS during the first part of the night, increased sleep fragmentation, and a decrease in the amount of slow wave sleep.

Added to our previous findings of higher anxiety-like behaviors and hyperemotionality-like behaviors

in prenatally stressed rats, the dysfunction of their HPA axis and circadian timing system and the observation of long-term changes in their sleep structure support the validity of PNS models as a possible animal model of depression. In contrast to other stress-related animal models of depression, the persistence of all induced abnormalities after stressor removal in PNS or PNH rats can be seen as particularly advantageous for the design and testing of new therapeutic strategies in mood and sleep disorders.

Further Reading

- Barbazanges, M., Piazza, P. V., Le Moal, M., et al. (1996). Maternal glucocorticoid secretion mediates long-term effects of prenatal stress. *Journal of Neuroscience* 16, 7783–7790.
- Day, J. C., Koehl, M., Deroche, V., et al. (1998). Prenatal stress enhances stress- and corticotropin-releasing factor-induced stimulation of hippocampal acetylcholine release in adult rats. *Journal of Neuroscience* 18, 1886–1892.
- de Kloet, E. R., Sibug, R. M., Helmerhorst, F. M., et al. (2005). Stress, genes and the mechanism of programming the brain for later life. *Neuroscience and Biobehavioral Reviews* 29(2), 271–281.
- Dugovic, C., Maccari, S., Weibel, L., et al. (1999). High corticosterone levels in prenatally-stressed rats predict persistent paradoxical sleep alterations. *Journal of Neuroscience* 19(19), 8656–8664.
- Koehl, M., Barbazanges, A., Le Moal, M., et al. (1997). Prenatal stress induces a phase advance of circadian corticosterone rhythm in adult rats which is prevented by postnatal stress. *Brain Research* 759(2), 317–320.
- Koehl, M., Darnaudery, M., Dulluc, J., et al. (1999). Prenatal stress alters circadian activity of hypothalamo-pituitary-adrenal axis and hippocampal corticosteroid receptors in adult rats of both gender. *Journal of Neurobiology* 40(3), 302–315.
- Maccari, S., Darnaudery, M., Morley-Fletcher, S., et al. (2003). Prenatal stress and long-term consequences: implications of glucocorticoid hormones. *Neuroscience and Biobehavioral Reviews* 27, 119–127.
- Maccari, S., Piazza, P. V., Kabbaj, M., et al. (1995). Adoption reverses the long-term impairment in glucocorticoid feedback induced by prenatal stress. *Journal of Neuroscience* 15, 110–116.
- Morley-Fletcher, S., Darnaudery, M., Koehl, M., et al. (2003). Prenatal stress in rats predicts immobility behaviour in the forced swim test: effects of a chronic treatment with tianeptine. *Brain Research* 989, 246–251.
- Morley-Fletcher, S., Darnaudery, M., Mocaer, E., et al. (2004). Chronic treatment with imipramine reverses immobility behaviour, hippocampal corticosteroid receptors and cortical 5-HT(1A) receptor mRNA in prenatally stressed rats. *Neuropharmacology* 47, 841–847.
- Poncet, L., Pequignot, J. M., Cottet-Emard, J. M., et al. (1999). Altered daily rhythms of brain and pituitary indolamines and neuropeptides in long-term hypoxic rats. *American Journal of Physiology* 277, R66–R75.

- Rosenwasser, A. and Wirz-Justice, A. (1997). Circadian rhythms and depression: clinical and experimental models. In: Redfern, P. H. & Lemmer, B. (eds.) *Physiology and pharmacology of biological rhythms* (vol. 125), pp. 457–486. Berlin: Springer-Verlag.
- Turek, F. W. and Van Reeth, O. (1996). Circadian rhythms. In: Fregly, M. J. & Blatteis, C. M. (eds.) *Handbook of physiology. Sec. 4: Adaptation to the environment*, pp. 1329–1359. New York: Oxford University Press.
- Vallée, M., Mayo, W., Dellu, F., et al. (1997). Prenatal stress induces high anxiety and postnatal handling induces low anxiety in adult offspring: correlation with stress-induced corticosterone secretion. *Journal of Neuroscience* 17(7), 2626–2636.
- Vincent, J., Mamet, J., Lee, F., et al. (2002). Prenatal hypoxia impairs circadian synchronization and response to the biological clock to light in adult rats. *Journal of Physiology* (London) 543, 387–395.

Circadian Rhythms, Genetics of

F W Turek and M H Vitaterna

Northwestern University, Evanston, IL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by F W Turek and M H Vitaterna, volume 1, pp 471–477, © 2000, Elsevier Inc.

The Molecular Circadian Clock

The Molecular Clock Is Not Confined to the SCN

Disruption of Molecular Clock Leads to Tissue and Organismal Dysfunction

Glossary

<i>Circadian clock</i>	An internal timing mechanism that regulates the expression of circadian rhythms. In mammals, a master circadian clock is located in the hypothalamic suprachiasmatic nucleus (SCN), although recent findings indicate that most if not all tissues of the body contain circadian clocks that are directly or indirectly regulated by the clock in the SCN.
<i>Circadian clock genes and clock-controlled genes</i>	While only a small number of genes and their protein products appear to be involved in the transcriptional/translational feedback loop that makes up the molecular core of the circadian clock, this molecular clock in turn appears to regulate the expression of hundreds of downstream clock-controlled genes.
<i>Circadian rhythms</i>	Biochemical, physiological, and behavioral events that reoccur in a cyclic fashion with a period of about 24 h.
<i>Entrainment</i>	Process by which an internal circadian clock is entrained or synchronized to the period of an external stimulus that

usually has a period of about 24 h. In nature, the light–dark cycle is the primary external environmental signal that entrains the circadian clock to the 24-h period of the day that is due to the rotation of the earth on its axis.

Free-running

Term used to describe circadian rhythms persisting in the absence of any 24-h synchronizing information from the external environment. The expression of free-running circadian rhythms with a period close to (e.g., 23–24 h) but rarely exactly 24 h demonstrates the endogenous nature of the circadian clock underlying the expression of circadian rhythms.

Period

The length of a cycle of the rhythm. The period of a free-running rhythm is considered a fundamental property of the circadian clock.

Phase

The timing of the rhythm. During entrainment, the timing of a phase reference point (such as activity onset) relative to the external time cue (such as lights on) should be stable. This relative timing is referred to as the phase angle of entrainment. The rhythm-resetting response to a perturbation such as a brief light exposure is measured as a phase shift.

The Molecular Circadian Clock

Early studies in invertebrates demonstrated that the period of the circadian clock was clearly under genetic control since 24-h rhythms developed in organisms never experiencing any 24-h changes in the physical environment. Furthermore, rearing mice on non-24-h cycles as extreme as 20 or 28 h has only small (i.e., on the order of a few minutes) and transient effects on

the free-running period of the circadian clock when animals are subsequently housed under conditions of constant darkness (i.e., the free-running period remains within a few minutes of 24 h). Until recently, evidence that genetic differences define the characteristics of the circadian clock system in mammals has been determined primarily by comparing properties of circadian rhythms between different inbred strains of the same species.

With the discovery of the first mammalian circadian clock gene (named *Clock*) in 1997, a new era opened for discovering how the mammalian circadian clock ticks. The *Clock* gene was initially identified through a behavioral phenotypic screen of mice that were the offspring of mutagenized animals. Vitaterna et al. identified a mutant animal that demonstrated a lengthened circadian period under free-running conditions (constant darkness). Descendants of this mutant with two copies of the mutant gene initially showed a free-running period of about 27–28 h in constant darkness, with subsequent breakdown in circadian rhythmicity. Next, the *Clock* gene was cloned and found to encode a member of the bHLH-PAS (basic helix-loop-helix, Per-Arnt-Sim domain) family of transcription factors. The role of the putative *Clock* gene was confirmed by creating transgenic mice in which the altered phenotype of *Clock* mutants was rescued by insertion of the wild-type *Clock* gene.

Other key clock genes in mammals have since been characterized. Three murine homologs of the *Drosophila period* gene, *mPer1*, *mPer2*, and *mPer3* have been identified. While double-knockout animals lacking *mPer1* and *mPer2* show behavioral arrhythmicity when placed in constant conditions, there appears to be partial redundancy of function between *mPer* family members. Moreover, *mPer3* does not appear to have an essential function in the core clock mechanism.

A second bHLH-PAS transcription factor, BMAL1, has been identified and shown to be a core circadian clock gene. *In vitro*, CLOCK-BMAL1 dimers activate transcription of *mPer*, and the *Clock* mutation described previously is associated with reduced expression of *mPer1*, *mPer2*, and *mPer3* in the suprachiasmatic nucleus (SCN). These data support a role for CLOCK (and the CLOCK-BMAL1 dimer) in activation of transcription of all three *mPer* homologs.

Two cryptochrome genes, *Cry1* and *Cry2*, also appear to be components of the core circadian system. *Cry2* null mutants exhibit lengthened circadian periods, while *Cry1* null mutants exhibit shortened circadian periods. In *Cry* double-knockout mice, animals are completely arrhythmic immediately after the introduction of constant darkness. CRY1 and CRY2

are both synthesized in the SCN, where at least *Cry1* exhibits rhythmic expression. These proteins inhibit CLOCK-BMAL1-driven transcription, most probably due to binding of CRY1 and CRY2 to clock components. Importantly, CRY1 and CRY2 are able to translocate *per* gene products to the nucleus.

Posttranslational modification of clock proteins seems to be important in the modulation of circadian rhythm. Casein kinase I ϵ (CKI ϵ), which phosphorylates CRY1, CRY2, PER1, PER2, and BMAL1, is defective in the *tau* mutant in Syrian hamsters, in which the free-running period in constant darkness is 22 h and 20 h in heterozygous and homozygous animals, respectively. In this mutant, CKI ϵ has a lower rate of phosphorylation. Recent data also suggest that a second kinase, CKI δ , is important for the mammalian clock.

The current concept of the core clock is a system of interacting transcriptional/translational feedback loops (Figure 1). The two transcription factors, CLOCK and BMAL1, form dimers that bind to DNA and promote rhythmic transcription of three *Per* genes, two *Cry* genes, and the *Rev-Erb α* gene (step 1). The proteins encoded by *mPer1* and *mPer2* bind to the proteins encoded by *Cry1* and *Cry2*, forming complexes with casein kinases (step 2). These translocate into the nucleus, where CRY and PER proteins act as negative regulators, directly inhibiting CLOCK/BMAL1 activity and indirectly the transcription of *mPers*, *Crys*, and *Rev-Erb α* (step 3). Positive feedback is mediated by rhythmic modulation of *Bmal1* transcription. Thus, the REV-ERB α protein, whose synthesis is increased by binding of CLOCK/BMAL1 dimers to DNA, inhibits the transcription of *Bmal1* (step 4).

Recent studies indicating that other genes involved in the posttranslational modification of circadian clock proteins can alter the circadian clock suggest that many other genes may be found that can modify the core genes/protein products of the molecular clock and alter its periodicity.

The Molecular Clock Is Not Confined to the SCN

With the establishment of the bilaterally paired SCN of the hypothalamus as the site of the circadian clock in mammals, considerable attention was focused on how the SCN can generate circadian rhythms at the neural as well as molecular levels. However, once the core components of the molecular circadian clock were identified, a surprising discovery was made: the clock genes not only are expressed in the SCN, but also are rhythmically expressed in the SCN and peripheral tissues throughout the body, including the

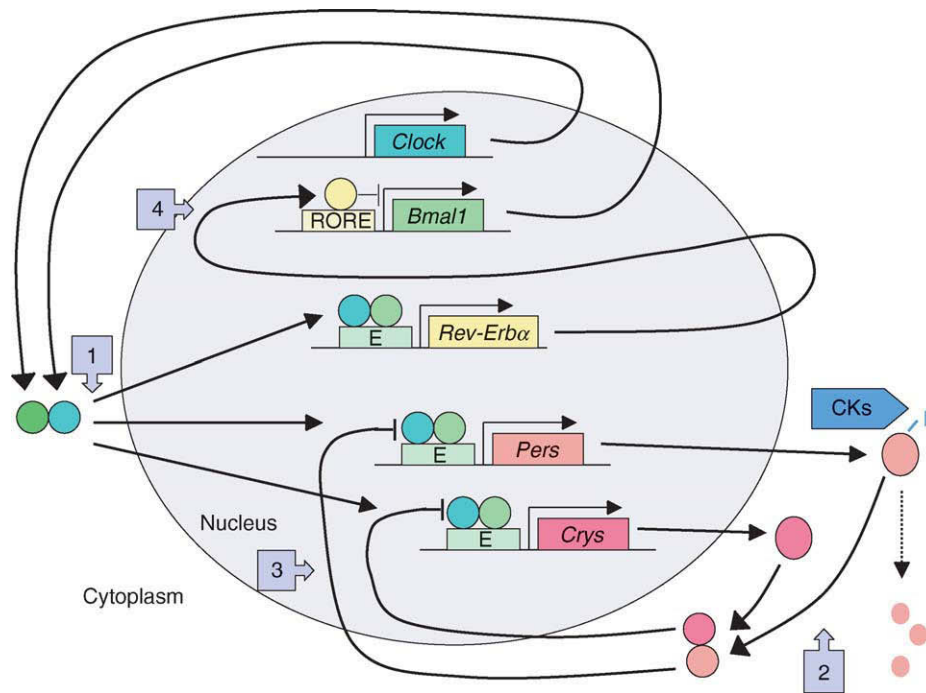


Figure 1 See text for explanation of figure.

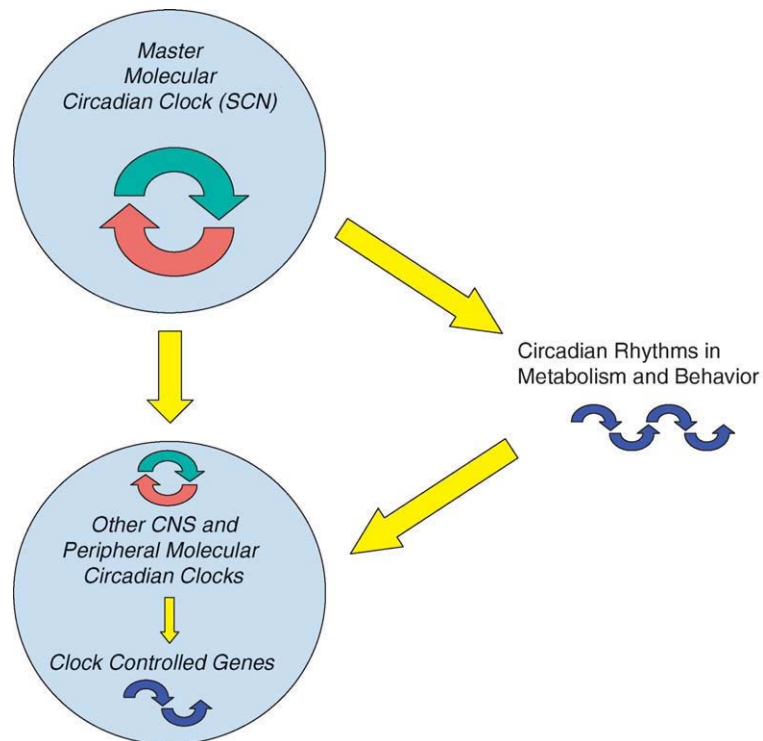


Figure 2 See text for explanation of figure.

liver, heart, lungs, and muscle. Furthermore, these circadian oscillations in the peripheral tissues can persist for many cycles *in vitro*, revealing the endogenous nature of tissue-specific circadian rhythmicity. Most importantly, not only are the clock genes oscillating

in various tissues, but also they appear to be regulating the transcription of hundreds of downstream clock-controlled genes. In a simplistic form, our understanding of the circadian clock system is as follows (Figure 2):

1. The circadian clock genes and their protein products compose a transcriptional/translational molecular feedback loop (with posttranslational components) that can generate precise circadian timing in the master circadian clock in the SCN.

2. The SCN communicates circadian information to other central and peripheral tissues via neural and/or neurohormonal transmission vectors.

3. Non-SCN circadian clocks appear to consist of a molecular core, as does the SCN. Thus, the molecular clock in the SCN (which is itself entrained by the light–dark cycle) somehow entrains a similar molecular clock in other CNS and peripheral tissues. This entrainment may be direct or indirect via the SCN control of metabolic and behavioral rhythms (e.g., feeding, sleep–wake cycle, body temperature).

This simple three-step process does not rule out the possibility that downstream (from SCN) circadian oscillators will not persist on their own (e.g., in SCN-lesioned animals). Nor does it exclude the possibility that they can be entrained by environmental factors (e.g., when feeding occurs out of phase with the normal SCN-timed control of feeding). However, under normal entrained conditions, it appears that the SCN serves as the conductor of an orchestra in regulating the timing, directly or indirectly, of hundreds of downstream oscillators, which in turn are regulating the timing and expression of hundreds, if not thousands, of clock-controlled genes. Such intricate and diverse timing raises profound questions about the role of such timing in health and disease at many levels of organization.

Disruption of Molecular Clock Leads to Tissue and Organismal Dysfunction

A foundation of the field of biological rhythms is that the circadian clock system has evolved to allow organisms to maintain two different kinds of synchronization. One type of synchronization, referred to as external synchronization, enables the organisms to maintain synchrony with the external 24-h environment. Thus, the animal, or plant, is doing the right thing at the right time of the day or night such that optimum survival and reproductive success are ensured. A second, and less visible but no less important, type of synchrony imposed by the circadian clock is referred to as internal synchronization. In internal synchrony, the vast array of internal cellular, metabolic, physiological, endocrine, and behavioral rhythmic processes are maintained in temporal relationships to allow for life processes to be carried out at optimal efficiency. Until recently, this concept of internal synchrony was more often applied to the functioning of the organism as a whole. That is, in the case of

mammals, the SCN was seen as the master circadian clock regulating the timing of the various organ systems and behavior of the animal to ensure optimal internal synchrony for survival and reproductive fitness. In the case of humans engaged in shift work or experiencing jet lag, the internal synchrony would be perturbed (i.e., internal desynchronization), leading to less than optimal organ/behavioral performance. However, with the discovery that most if not all tissues and organ systems of the body have their own circadian molecular core machinery, the concept of internal synchrony takes on new meaning. One can now talk about internal desynchronization at the local cellular or tissue level. Such a paradigm shift in thinking about the meaning of internal synchronization has profound implications for health and disease.

The fact that the expression of hundreds of genes is under local (i.e., cellular) circadian control argues that the phase relationship between the rhythmic expression of these genes has functional significance at the cellular and/or tissue level. Just as a breakdown of internal synchrony at the macro-level (i.e., organ system/whole organism) is disruptive to health and well-being, one can expect that a similar breakdown at the micro-level (cellular/tissue) could lead to dysfunction at the local level. Thus, we may have a perfectly functioning circadian clock system at the macro-level, but still have a dysfunctional system at a local tissue level. Age-related diseases are often associated with tissue (CNS or peripheral) dysfunction due to age-related breakdown (e.g., due to oxidative stress, DNA repair dysfunction, accumulated protein/DNA damage) in cellular processes. However, the newly discovered circadian molecular clock machinery and its overall importance for cellular/tissue timing raises the intriguing possibility that many age-related diseases may have at their core a breakdown in the timing or internal temporal organization. Many diseases are due to malfunction of a particular tissue, organ, or organ system, whether due to external or internal factors, and death is often associated with whatever vital system breaks down first. An intriguing offshoot of this concept is that disease may be due to temporal disorganization at the cellular or tissue level, whether due to external or internal factors, and death may be due to the abnormal timing at the molecular/genetic level in specific tissue/organ systems whose normal function is critical for the survival of the organism. The recent finding that a mutation of a canonical circadian clock gene, *Clock*, can lead to a major disruption at tissue and behavioral levels in energy balance, as well as obesity and the metabolic syndrome, may be an example whereby temporal disorganization leads to disease conditions at the macro and micro organizational levels.

See Also the Following Articles

Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress; Seasonal Changes in Stress Responses; Sleep Loss, Jet Lag, and Shift Work; Sleep, Sleep Disorders, and Stress; Seasonal Rhythms.

Further Reading

- Akashi, M., Tsuchiya, Y., Yoshino, T. and Nishida, E. (2002). Control of intracellular dynamics of mammalian period proteins by casein kinase I epsilon (CKIepsilon) and CKIdelta in cultured cells. *Molecular and Cellular Biology* 22(6), 1693–1703.
- Albrecht, U., Sun, Z. S., Eichele, G. and Lee, C. C. (1997). A differential response of two putative mammalian circadian regulators, mper1 and mper2, to light. *Cell* 91, 1055–1064.
- Antoch, M. P., Song, E., Chang, A., et al. (1997). Functional identification of the mouse circadian Clock gene by transgenic BAC rescue. *Cell* 89, 655–667.
- Bunger, M. K., Wilsbacher, L. D., Moran, S. M., et al. (2000). Mop3 is an essential component of the master circadian pacemaker in mammals. *Cell* 103(7), 1009–1017.
- Cardone, L., Hirayama, J., Giordano, F., et al. (2005). Circadian clock control by SUMOylation of BMAL1. *Science* 309(5739), 1390–1394.
- Eide, E. J., Vielhaber, E. L., Hinz, W. A. and Virshup, D. M. (2002). The circadian regulatory proteins BMAL1 and cryptochromes are substrates of casein kinase Iepsilon. *Journal of Biological Chemistry* 277(19), 17248–17254.
- Gekakis, N., Staknis, D., Nguyen, H. B., et al. (1998). Role of the CLOCK protein in the mammalian circadian mechanism. *Science* 280, 1564–1569.
- Griffin, E. A., Jr., Staknis, D. and Weitz, C. J. (1999). Light-independent role of CRY1 and CRY2 in the mammalian circadian clock. *Science* 286(5440), 768–771.
- Iitaka, C., Miyazaki, K., Akaike, T. and Ishida, N. (2005). A role for glycogen synthase kinase-3beta in the mammalian circadian clock. *Journal of Biological Chemistry* 280(33), 29397–29402.
- Jin, X., Shearman, L., Weaver, D., et al. (1999). A molecular mechanism regulating output from the suprachiasmatic circadian clock. *Cell* 96, 57–68.
- King, D. P., Zhao, Y., Sangoram, A. M., Wilsbacher, L. D., et al. (1997). Positional cloning of the mouse circadian Clock gene. *Cell* 89, 641–653.
- Kume, K., Zylka, M. J., Sriram, S., et al. (1999). mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop. *Cell* 98(2), 193–205.
- Lowrey, P. L. and Takahashi, J. S. (2000). Genetics of the mammalian circadian system: photic entrainment, circadian pacemaker mechanisms and posttranslational regulation. *Annual Review of Genetics* 34, 533–562.
- Lowrey, P. L., Shimomura, K., Antoch, M. P., et al. (2000). Positional syntenic cloning and functional characterization of the mammalian circadian mutation tau. *Science* 288(5465), 483–492.
- Miyamoto, Y. and Sancar, A. (1998). Vitamin B2-based blue-light photoreceptors in the retinohypothalamic tract as the photoactive pigments for setting the circadian clock in mammals. *Proceedings of the National Academy of Sciences of the USA* 95(11), 6097–6102.
- Panda, S., Antoch, M. P., Miller, B. H., et al. (2002). Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell* 109(3), 307–320.
- Reppert, S. M. and Weaver, D. R. (2002). Coordination of circadian timing in mammals. *Nature* 418(6901), 935–941.
- Shearman, L. P., Zylka, M. J., Weaver, D. R., Kolakowski, L. F. J. and Reppert, S. M. (1997). Two period homologs: circadian expression and photic regulation in the suprachiasmatic nuclei. *Neuron* 19, 1261–1269.
- Shearman, L. P., Jin, X., Lee, C., Reppert, S. M. and Weaver, D. R. (2000). Targeted disruption of the mPer3 gene: subtle effects on circadian clock function. *Molecular and Cellular Biology* 20(17), 6269–6275.
- Sun, Z. S., Albrecht, U., Zhuchenko, O., et al. (1997). RIGUI, a putative mammalian ortholog of the Drosophila period gene. *Cell* 90, 1003–1011.
- Takumi, T., Taguchi, K., Miyake, S., et al. (1998). A light-independent oscillatory gene mPer3 in mouse SCN and OVL. *EMBO Journal* 17, 4753–4759.
- Tei, H., Okamura, H., Shigeyoshi, Y., et al. (1997). Circadian oscillation of a mammalian homologue of the Drosophila period gene. *Nature* 389, 512–516.
- Turek, F. W., Joshi, C., Kohsaka, A., et al. (2005). Obesity and metabolic syndrome in circadian Clock mutant mice. *Science* 308(5724), 1043–1045.
- Van der Horst, G. T. J., Muijtens, M., Kobayashi, K., et al. (1999). Mammalian Cry1 and Cry2 are essential for maintenance of circadian rhythms. *Nature* 398, 627–630.
- Vitaterna, M. H., King, D. P., Chang, A.-M., et al. (1994). Mutagenesis and mapping of a mouse gene, Clock, essential for circadian behavior. *Science* 264, 719–725.
- Vitaterna, M. H., Selby, C. P., Todo, T., et al. (1999). Differential regulation of mammalian period genes and circadian rhythmicity by cryptochromes 1 and 2. *Proceedings of the National Academy of Sciences of the USA* 96(21), 12114–12119.
- Vitaterna, M. H., Pinto, L. H. and Turek, F. W. (2005). Molecular genetic basis for mammalian circadian rhythms. In: Kryger, M. H., Roth, T. & Dement, W. C. (eds.) *Principles and practice of sleep medicine*, pp. 363–374. New York: W.B. Saunders.
- Yoo, S. H., Yamazaki, S., Lowrey, P. L., et al. (2004). PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues. *Proceedings of the National Academy of Sciences of the USA* 101(15), 5339–5346.
- Zylka, M. J., Shearman, L. P., Weaver, D. R. and Reppert, S. M. (1998). Three period homologs in mammals: differential light responses in the suprachiasmatic circadian clock and oscillating transcripts outside of brain. *Neuron* 20, 1103–1110.

Cocaine See: Drug Use and Abuse.

Cognition and Stress

M W Eysenck

University of London, Egham, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M N O Davies and G Underwood, volume 1, pp 478–483,

© 2000, Elsevier Inc.

Appraisal Processes

Interpretive Bias

Attentional Bias

Assessment of the Cognitive Approach

Glossary

<i>Appraisal</i>	An evaluation of whether the situation is threatening and the availability of coping strategies.
<i>Attentional bias</i>	A tendency to attend preferentially to threat-related rather than neutral stimuli.
<i>Coping strategies</i>	A wide range of different ways of handling stressful situations, some focusing primarily on behavior and others on cognition or emotion.
<i>Interpretive bias</i>	A tendency to interpret ambiguous situations in a threat-related fashion.
<i>Transactional models</i>	A theoretical approach based on the notion that stress is determined by the interaction of the situation and the individual.

There is a consensus that stress responses (whether psychological, behavioral, or physiological) depend at least in part on various cognitive processes. More specifically, it is assumed that it is important to distinguish between the actual situation (defined in some relatively objective fashion) and the situation as it is perceived by the individual. There are systematic individual differences in the way any given threatening situation is perceived, and many of these differences can be understood by assuming that individuals differ in the cognitive processes determining how the situation is perceived or interpreted.

The emphasis on cognitive processes forms part of a reconceptualization of stress. The traditional view was that stress is an almost inevitable consequence of an individual's being exposed to a stressor (e.g., loud noise or complex situation). The contemporary

view (expressed with particular clarity by Lazarus and Folkman) is that stress should be considered in terms of transactional models. More specifically, whether any given stimulus or situation produces stress depends on complex interactions between the stimulus or situation and the individual's appraisal and coping processes. Cognitive appraisal is of major importance within this transactional approach because stress is experienced only when the individual perceives that environmental demands exceed his or her available resources.

Appraisal Processes

Most theoretical approaches concerned with the role of cognition in influencing stress responses emphasize the crucial role of appraisal processes. A seminal influence in this area is the contribution of Lazarus, who argued that cognitive appraisal is of central importance in determining individuals' reactions to threatening stimuli and situations. More specifically, he argued that cognitive appraisal can be subdivided into three forms of appraisal: (1) primary appraisal, in which the immediate situation is perceived as being positive, stressful, or irrelevant to well-being; (2) secondary appraisal, in which the individual takes into account the resources that he or she has available for coping successfully with the current situation; and (3) re-appraisal, in which the preceding primary and secondary appraisals are modified, if necessary, on the basis of ongoing monitoring of the current situation and the coping strategies being used by the individual.

There is a considerable amount of evidence demonstrating that manipulations of cognitive appraisal have significant (and predictable) effects on self-report measures of emotion and on psychophysiological responses. In addition, individual differences in spontaneous cognitive appraisals are predictive of stress responses. For example, Gaab and colleagues exposed participants to a simulated job interview and a mental arithmetic task in front of an audience. Their key finding was that anticipatory cognitive appraisal accounted for up to 35% of the variance in the salivary cortisol response.

Coping Strategies

Of crucial importance during secondary appraisal is a consideration of various coping strategies, which are

ways of dealing with a stressful situation. Examples include problem-focused coping, emotion-focused coping, and avoidance-based coping. Numerous coping strategies have been proposed. Indeed, in a 2003 review, Skinner and colleagues produced a list of 400 different coping strategies based on previous research. They produced a coherent structure for coping based on various systems put forward by earlier theorists. As a consequence, Skinner and colleagues placed coping strategies into a limited number of loosely organized core families of coping.

Interpretive Bias

The notion that cognitive appraisals are important in determining an individual's stress level has been applied extensively in the clinical field. In essence, it is assumed by cognitive therapists that unduly negative and pessimistic cognitive appraisals of environmental or internal threat-related stimuli and situations produce high levels of anxiety and stress in patients with anxiety disorders. The term interpretive bias is used to describe such pessimistic cognitive appraisals, and there is much evidence that anxious patients have interpretive biases for those stimuli and situations most closely associated with their particular disorder. For example, patients suffering from social phobia experience extremely high stress levels when anticipating or being involved in demanding social situations, and patients with panic disorder produce exaggerated responses when physiologically aroused.

The issue of causality is important. The finding that patients with anxious disorders have various interpretive biases could be due either to the possession of interpretive biases triggering or maintaining clinical anxiety or to being clinically anxious, causing the development of interpretive bias. Relevant evidence was discussed by Yiend and Mackintosh. Healthy individuals who were trained to select threat-related meanings of homographs responded to stressful videos with a greater increase in anxious mood than did those not receiving such training. Such findings indicate that interpretive bias can increase anxious and stressed responses to stressful situations.

Attentional Bias

Another way in which cognitive processes are related to stress and anxiety is via attentional bias, which is a tendency to focus attention on threat-related rather than neutral stimuli when both types of stimuli are present concurrently. Healthy individuals who are highly stressed or anxious and patients with anxious disorders exhibit attentional bias to a much greater extent than other people.

The association of attentional bias with stress and anxiety does not in and of itself demonstrate that attentional bias has causal efficacy in producing stressed and/or anxious states. However, this causal issue has been investigated in a number of recent studies. Two experimental approaches are of particular importance. First, attentional bias can be assessed in a group of individuals prior to receiving very stressful information, and the extent to which prior attentional bias predicts subsequent stress responses can be measured. This approach was adopted by MacLeod and Hagan, who studied women subsequently receiving a diagnosis of cervical pathology. The extent of attentional bias to subliminal stimuli predicted emotional reactivity after receiving the diagnosis.

Second, individuals can receive extensive training designed to increase or decrease attentional bias. For example, MacLeod and colleagues required participants to respond as rapidly as possible to probe stimuli that replaced one of two concurrently presented stimuli (one threat-related and the other neutral). Attentional bias was produced by consistently having the probe stimuli replace threat-related stimuli, and opposite attentional bias (i.e., avoidance of threat-related stimuli) was produced by consistently having the probe stimuli replace neutral stimuli. Participants who had been trained to produce an attentional bias showed enhanced negative emotional responses to a stressful task compared to those trained to produce an opposite attentional bias.

Assessment of the Cognitive Approach

Cognitive appraisal, interpretive bias, and attentional bias can take many forms and occur in a wide range of situations. For example, some appraisal processes seem to operate relatively automatically and below the level of conscious awareness, whereas other appraisal processes are conscious and deliberate. Another distinction is between appraisal processes that revolve primarily around the individual and his or her ability to cope with a stressful situation and appraisal processes that are strongly influenced by the prevailing social context. Most research in this area has focused on conscious appraisal processes based on the individual's assessment of his or her individually based coping strategies. As a consequence, there should be more emphasis in future research on socially based appraisal processes occurring below conscious awareness.

It is important to ascertain the limits of contemporary cognitive approaches to stress. Four potential limitations were discussed by McNally. First, sometimes appraisal-based approaches produce findings that appear almost tautological. For example, claiming that an individual feels stressed in a given situation

because he/she finds that situation threatening may seem almost self-evident. However, such a claim represents progress in the sense that it identifies cognitive processes as being intimately involved in stress responses, and it provides an impetus to discover more about the processes responsible for such an appraisal.

Second, much of what is known about appraisal processes and interpretive biases is based on self-report data. Self-report data are subject to various distortions and inaccuracies. In addition, individuals in stressful situations typically have only limited access to information about the precise factors causing them to produce certain cognitive appraisals or interpretive biases. It is undeniable that self-report data are limited in various respects. However, much research has included various measures not dependent on self-report; these include behavioral measures of interpretive and attentional biases and psychophysiological measures of stress responses.

Third, there are complex issues relating to causality. In spite of the evidence that manipulations designed to alter cognitive appraisals, interpretive bias, or attentional bias have predicted effects on stress and other responses, it is probable that causality can also operate in the opposite direction, with stress influencing cognitive processes.

Further Reading

- Eysenck, M. W. (1997). *Anxiety and cognition: a unified theory*. Hove, UK: Psychology Press.
- Gaab, J., Rohleder, N., Nater, U. M., et al. (2005). Psychological determinants of the cortisol stress response: the role of anticipatory cognitive appraisal. *Psychoneuroendocrinology* 30, 599–610.

- Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford: Oxford University Press.
- Lazarus, R. S. and Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.
- Macleod, C. (1993). Cognition in clinical psychology: measures, methods or models? *Behaviour Change* 10, 169–195.
- MacLeod, C., Campbell, L., Rutherford, E., et al. (2004). The causal status of anxiety-linked attentional and interpretive bias. In: Yiend, J. (ed.) *Cognition, emotion and psychopathology: theoretical, empirical and clinical directions*, pp. 172–189. Cambridge, UK: Cambridge University Press.
- Macleod, C. and Hagan, R. (1992). Individual differences in the selective processing of threatening information, and emotional responses to a stressful life event. *Behaviour Research and Therapy* 30, 151–161.
- MacLeod, C., Rutherford, E., Campbell, L., et al. (2002). Selective attention and emotional vulnerability: assessing the causal basis of their association through the experimental manipulation of attentional bias. *Journal of Abnormal Psychology* 111, 107–123.
- McNally, R. J. (2001). On the scientific status of cognitive appraisal models of anxiety disorder. *Behaviour Research and Therapy* 39, 513–521.
- Parkinson, B. and Manstead, A. S. R. (1992). Appraisal as a cause of emotion. In: Clark, M. S. (ed.) *Review of personality and social psychology* (vol. 13), pp. 122–149. New York: Sage.
- Skinner, E. A., Edge, K., Altman, J., et al. (2003). Searching for the structure of coping: a review and critique of category systems for classifying ways of coping. *Psychological Bulletin* 129, 216–269.
- Yiend, J. and Mackintosh, B. (2004). The experimental modification of processing biases. In: Yiend, J. (ed.) *Cognition, emotion and psychopathology: theoretical, empirical and clinical directions*, pp. 190–210. Cambridge, UK: Cambridge University Press.

Cognitive-Behavioral Therapy

G A Fava

University of Bologna, Bologna, Italy and State University of New York, Buffalo, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G A Fava, volume 1, pp 484–486, © 2000, Elsevier Inc.

Definition of Cognitive-Behavioral Therapy

Specific Techniques

Applications of Cognitive-Behavioral Therapy in the Stress Response

Glossary

Cognitive therapy
Exposure

A brief psychotherapeutic approach aimed at changing thinking.

A brief psychotherapeutic approach based on exposure to feared situations or stimuli.

Problem-solving training
Rational-emotive behavior therapy

A brief psychotherapeutic approach for improving problematic situations and/or interpersonal conflicts.

A brief psychotherapeutic treatment that relies on persuasion and reason.

Stress-management approaches

A set of diverse techniques for reducing psychophysiological activation.

Definition of Cognitive-Behavioral Therapy

The term cognitive-behavioral therapy subsumes a number of diverse therapeutic procedures, such as cognitive therapy, rational-emotive psychotherapy, problem-solving interventions, and exposure. In a restrictive definition, it indicates psychotherapeutic approaches that explicitly seek to produce change in cognition as a means of influencing other phenomena of interest, such as affect or behavior. This definition excludes specific behavioral interventions, such as exposure, that may work through cognitive mediation but that are not primarily intended to produce change by influencing thinking. A broader definition stems from the growing awareness that most psychotherapeutic approaches combine cognitive and behavioral elements, although in different ways and to different extents.

As a result, both cognitive and behavioral approaches may legitimately belong to the broader rubric of cognitive-behavioral therapy. In recent years there has been an increasing awareness of the common ingredients that pertain to the psychotherapeutic process (Table 1) and cognitive-behavioral approaches. The main characteristics that cognitive-behavioral therapies share are the following:

1. The therapy tends to be short (fewer than 20 sessions).
2. The therapy is structured and follows a pre-arranged schema of development.
3. The attention is centered on current problems.
4. Patient and therapist agree on the goals of treatment, which are clearly defined.
5. The patient is encouraged to perform self-observation of thinking and/or behavior, which may lead to their characterization (functional analysis).

Table 1 Nonspecific therapeutic ingredients that are shared by most forms of psychotherapy

<i>Ingredient</i>	<i>Characteristics</i>
Attention	The therapist's full availability for specific times
Disclosure	The patient's opportunity to ventilate thoughts and feelings
High arousal	An emotionally charged, confiding relationship with a helping person
Interpretation	A plausible explanation of the disturbed symptoms
Ritual	A ritual or procedure that requires the active participation of both patient and therapist and that is believed by both to be the means of restoring the patient's health

6. Homework assignments (whether consisting of self-observation or performing specific tasks) are given and reviewed by the therapist.
7. The therapy aims at developing the patient's control over his or her own problems or behaviors.

In randomized controlled trials, behavior therapy (exposure) alone and cognitive therapy alone resulted in similar improvements in obsessive-compulsive disorder, agoraphobia with panic attacks, posttraumatic stress disorder, and hypochondriasis. The fact that similar results can be achieved by different therapeutic approaches is leading to a weakening of classic paradigms and theories underlying each specific psychotherapeutic approach and to the necessity of paradigm shifts in our understanding of clinical changes. According to an evidence-based viewpoint, the term cognitive-behavioral therapy has a broad connotation and may indicate any brief psychological therapy that uses the ingredients listed above. The relative weight of each (or additional) ingredient may vary from one technique to the next, as defined by treatment protocols. In clinical practice, it may even vary from one case to another, or according to the stages of application of psychotherapy. Different ingredients may be used simultaneously from the beginning of treatment, or in a sequential order. For instance, in the cognitive-behavioral prevention of recurrent depression, behavioral techniques may characterize the first phase of treatment, whereas cognitive restructuring is applied only at a later point in time.

Specific Techniques

The therapeutic techniques that are used in cognitive-behavioral therapy may range from modification of emotional and relaxation responses (e.g., stress management techniques) to cognitive restructuring (e.g., cognitive therapy). The most important approaches (in terms of clinical acceptance and effectiveness or future implications) are the following:

1. Exposure
2. Response prevention
3. Operant conditioning
4. Relaxation
5. Stress-inoculation training
6. Problem-solving training
7. Rational-emotive behavior therapy
8. Cognitive therapy
9. Well-being therapy

Exposure (mainly graded, planned, homework exposure to distressing stimuli) has emerged as the most effective behavioral technique in anxiety disorders.

In the setting of disturbances such as obsessive-compulsive disorder, exposure to avoided cues may be associated with response prevention (staying in contact with those cues during the ensuing discomfort). Exposure appears to be a key factor in therapeutic change and may be implemented by diverse protocols, whose effectiveness, however, depends on the degree of *in vivo* exposure by the patient as a homework assignment. Operant conditioning derives from Skinner's work on the importance of contingencies in behavior, that is, whether a given response is followed by a positive or negative reinforcer. More than a specific strategy, it is a technique that may be incorporated in more complex strategies, such as dietary programs. Techniques based on relaxation methods achieved wide popularity in the 1970s and have been documented to be effective when emphasis is placed on enhanced muscular control, as in many cases of functional medical disorders. Relaxation may be seen as part of stress-management approaches that are aimed at reducing arousal and that combine various therapeutic ingredients, such as stress-inoculation training. This technique combines cognitive restructuring with behavioral self-management techniques. Patients are encouraged to apply these skills to a series of increasingly difficult tasks in order to cope with problematic life events. Problem-solving training emphasizes social problem solving as the primary coping strategy. This transactional approach uses a set of sequential procedures for dealing with problematic situations and interpersonal conflicts. In Ellis's rational-emotive behavior therapy, patients are taught to examine rationality as the primary mechanism of change. Beck's cognitive therapy is today probably the most widely endorsed cognitive intervention. Patients are encouraged to identify their beliefs, to treat them as hypotheses to be tested, and are taught alternative explanations. Both rational-emotive therapy and cognitive therapy incorporate behavioral components in their approach. Well-being therapy uses cognitive restructuring and behavioral interventions for enhancing psychological well-being. It is aimed at changing beliefs, attitudes, and behaviors that are detrimental to well-being.

Applications of Cognitive-Behavioral Therapy in the Stress Response

Cognitive-behavioral intervention may have two distinct yet ostensibly related roles in the modulation of stress response. One is concerned with a reduction in the psychophysiological activation that is associated (e.g., stress management approaches). The other is

related to the transactional model of stress, which emphasizes cognitive appraisal (how an event is interpreted by the individual) and coping (e.g., cognitive therapy). It may include supporting successful behavior change strategies to decrease health-damaging behaviors, increasing levels of social support, identifying physical symptoms that are expressions of emotional distress, and helping patients develop alternative coping strategies.

There is substantial evidence, based on randomized controlled studies, that cognitive-behavioral treatment may improve the outcome and quality of life of several psychiatric and medical disorders. Evidence-based applications of cognitive-behavioral therapy in psychiatric disorders include alcoholism and drug dependence, anorexia nervosa, bulimia nervosa, generalized anxiety disorder, hypochondriasis, mood disorders, obsessive-compulsive disorder, panic disorders and agoraphobia, personality disorders, posttraumatic stress disorder, psychosis and schizophrenia, simple phobia, and social phobia. Evidence-based applications of cognitive-behavioral therapy in medical settings include arthritis, asthma, cancer, chronic fatigue syndrome, chronic pain, coronary heart disease, diabetes, epilepsy, HIV infection and AIDS, hypertension, inflammatory bowel disease, irritable bowel syndrome, obesity, preparation for medical procedures, smoking, and tension headache. Further, it may improve health-promoting behaviors, such as changing harmful lifestyle patterns and habits and modifying illness attitudes and behaviors. In view of the role of psychological well-being in human health and disease, it is conceivable, but yet to be tested, that well-being therapy may offer another route to the improvement of the stress response. The absence of wellness may in fact create conditions of vulnerability to life's adversities, and an increase in psychological well-being may increase an individual's coping skills.

There is increasing evidence that cognitive-behavioral therapy exerts its effects by modulating the functioning of specific sites in limbic and cortical regions. It may thus work at the molecular level to alter stress-related gene expression and protein synthesis or to influence mechanisms implicated in learning and memory acquisition in stress-related neuronal structures.

See Also the Following Articles

Aerobic Exercise and Stress Reduction; Behavior Therapy; Cognition and Stress; Family Therapy; Group Therapy; Posttraumatic Therapy; Psychotherapy.

Further Reading

- Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.
- Ellis, A. T. (1962). *Reason and emotion in psychotherapy*. New York: Lyle Stuart.
- Emmelkamp, P. M. G., Bouman, T. K. and Scholing, A. (1992). *Anxiety disorders*. Chichester: Wiley.
- Fava, G. A. and Freyberger, H. (eds.) (1998). *Handbook of psychosomatic medicine*. Madison, CT: International Universities Press.
- Fava, G. A. and Ruini, C. (2003). Development and characteristics of a well-being enhancing psychotherapeutic strategy: well-being therapy. *Journal of Behavior Therapy and Experimental Psychiatry* **34**, 45–63.
- Fava, G. A., Ruini, C., Rafanelli, C., Finos, L., Conti, S. and Grandi, S. (2004). Six year outcome of cognitive behavior therapy for prevention of recurrent depression. *American Journal of Psychiatry* **161**, 1872–1876.
- Frank, J. D. and Frank, B. (1991). *Persuasion and healing* (3rd edn.). Baltimore, MD: The John Hopkins University Press.
- Lambert, M. J. (ed.) (2004). *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th edn.). New York: Wiley.
- Marks, I. (1987). *Fears, phobias, and rituals*. New York: Oxford University Press.
- Marks, I. (2002). Fear reduction by psychotherapies. *British Journal of Psychiatry* **176**, 507–511.
- Meichenbaum, D. (1985). *Stress inoculation training*. New York: Pergamon Press.
- Sobel, D. S. (2000). The cost-effectiveness of mind-body medicine interventions. *Progress in Brain Research* **122**, 393–412.

Colgi Complex See: Protein Synthesis.

Combat Reaction, Chronic

R H Rahe

University of Washington School of Medicine, Seattle, WA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R H Rahe, volume 1, pp 491–494, © 2000, Elsevier Inc.

Introduction**Problems with a Diagnosis of PTSD****Introduction**

Chronic stress reaction is appreciably different from the acute reaction described in a separate article. In the American and British militaries, dating from the American Civil War, a chronic stress disorder was generally what was referred to when medical personnel used terms such as soldier's heart, neurasthenia, battle neurosis, shell shock, traumatic neurosis, and battle fatigue. Of interest is that any of the above terms were literally ordered out of existence in the United States Army by General Omar N. Bradley in 1943. Colonel Long wrote: "Of the possible diagnostic

terms discussed, this term [combat exhaustion] was chosen because it was thought to convey the least implication of neuropsychiatric disturbance." Since the early 1980s this disorder has been frequently referred to as posttraumatic stress disorder (PTSD).

As in acute combat reactions, causes of the chronic condition are multiple. Military men developing a chronic reaction have typically been under combat stress for many weeks, months, or perhaps even years. Their physiological state appears to be characterized by hypoarousal. Instead of fight-or-flight physiology, these men tend to show signs and symptoms of conservation/withdrawal. The affected soldier frequently isolates him/herself from peers and shows signs and symptoms of depression. Once established, this chronic reaction is extremely difficult to reverse. In animal studies of defeat situations, a similar syndrome is seen. Signs and symptoms, predispositions, precipitants, treatment, and prevention of chronic combat reaction are outlined in subsequent sections of this article.

Signs and Symptoms

Early signs include impairment of previously displayed skills, severely lowered frustration tolerance, excessive griping, and even mild-to-moderate

paranoia. As these characteristics are frequently manifested in all combat troops, it can be a subtle distinction to determine when they exceed the normal range and signal concern. In sum, when these traits cause a human to significantly isolate him/herself from his/her peers and his/her military performance is substandard, it is time to evaluate the individual. If excessive alcohol or drug use, and/or signs and symptoms of depression are also part of the picture, the diagnosis should be strongly suspected.

If urine and blood samples could be collected in the field, catecholamines should be found to be at normal, or even low, levels in men with chronic combat reactions. In contrast, cortisol should be found to be significantly and persistently elevated throughout the circadian cycle. In a variety of chronic stress reactions, the baseline for cortisol secretion appears to be reset to a higher level. A further biochemical correlate of chronic stress is serum cholesterol. This high-density lipid is frequently seen in low serum concentrations during acute stress situations but shifting to high concentrations over a prolonged stress period (months to years).

In its most severe form, the soldier with a chronic combat reaction is so slowed in his/her thinking and motor movements that he/she appears to resemble a robot. His/her vacant-eyed look seems to focus far into the distance and has been labeled the 1000 yard stare. At this point the person is totally unable to function.

Predispositions and Precipitants

Many of the susceptibility factors making individuals more prone to chronic stress reactions are the same factors lending susceptibility to acute reactions to stress. These factors included childhood traumas, female gender, family and personal history for psychiatric

illness, minimal social skills, low levels of education, and antisocial traits. These traits are frequently accompanied by recent, pre-deployment, life stress.

Whereas intensity of battle is a major precipitant for an acute combat reaction, duration of combat is a large force shaping a chronic combat reaction. As in the acute disorder, high casualty rates add to the likelihood of this reaction. Frequently, a direct relationship has been seen between wounded-in-action rates and disabling chronic psychological reactions to combat.

A critical environmental precipitant is lack of mobility and/or progress in battle. This condition was widely prevalent during World War I, when trench warfare resulted in men spending up to 2–3 years in the same labyrinth of trenches and tunnels. In *Good-bye to All That*, British author Robert Graves wrote about his years in the trenches, where the war never seemed to progress and the only way out was to be wounded or killed.

An interpersonal precipitant is loss of confidence in leadership. If a mistake in tactics is made, for example a strike hits friendly forces, the soldier may then believe that his/her lack of confidence in command is justified.

On a personal level, when a soldier loses his/her sense of purpose, and especially when he/she perceives that there is no end in sight to long-lasting stress, the risk for this condition increases markedly. Preexisting neurosis and character pathology are personal factors which increase the liability of a soldier to develop a chronic stress reaction. As in acute battle stress, a severely dysfunctional early family life, possible abuse, lack of socialization experiences while growing up, and low educational attainment have been shown to predispose these individual to chronic battle stress reactions (Figure 1).

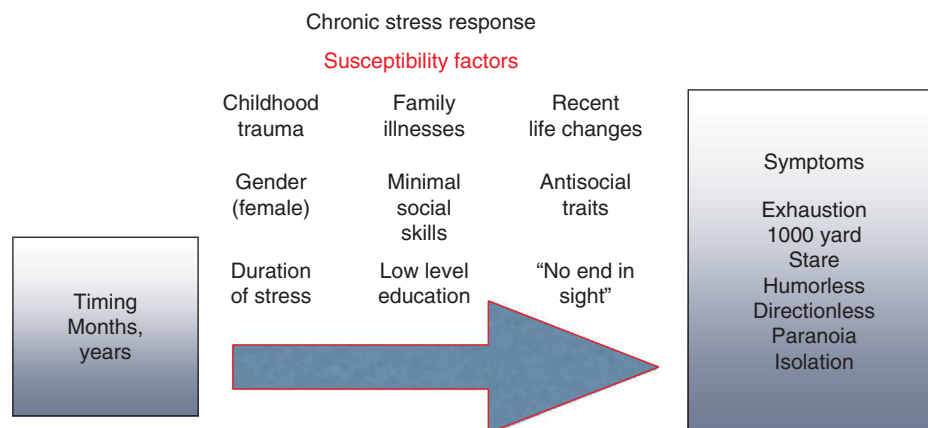


Figure 1 A presentation of selected early developmental and personal factors influencing susceptibility to symptom response following chronic stress.

Several resistance factors making reactions to stress less likely to occur are the opposite of susceptibility factors. These factors are a healthy childhood and a healthy family medical history, high school graduation, plentiful socialization when growing up, and an ability to find and give social support. Added to these factors are good leadership qualities, optimism tempered by reality, and an ability to pace oneself throughout a long campaign (Figure 2).

The progression of chronic stress reactions to recovery is approximately 40%, compared to a 60% likelihood of developing a subsequent development of a disorder. These subsequent disorders include several possibilities – not just chronic PTSD (Figure 3).

Treatment

The first line of treatment for PTSD is buddy and/or group support. This will be most effective in the early stages of the disorder. Even then, because the affected individual is constantly complaining and mistrustful,

it takes a strong commitment on the part of the group to bring him/her back into the fold. Such a group commitment has been seen to function extremely well in prisoner-of-war settings. Here, if buddy or group efforts failed to halt the withdrawal of a fellow prisoner, that person often died. The cause of death appeared to be chiefly determined by the person's loss of hope and giving up.

When specialized care is required for this condition, it means more time (from a few days to 3–4 weeks), more professionals (psychiatrists, psychologists, and possibly exercise therapists), and more resources (medications, a site for prolonged and specialized care) than is the case for acute combat reactions. The treatment site should be located near the front lines of battle but should be secure enough that it does not have to be moved frequently. The Israeli Defense Forces experimented with such a camp just inside their northern border during the Lebanon incursion. Here the men received medications (chiefly hypnotic and anxiolytic

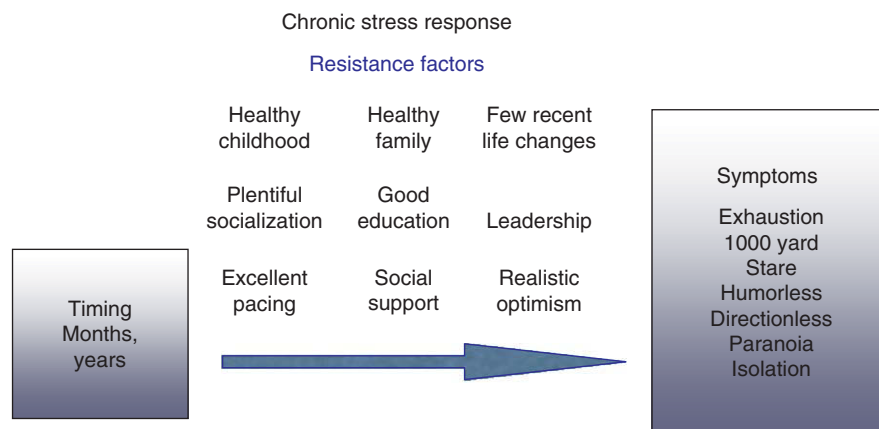


Figure 2 A presentation of selected early developmental and personal factors influencing resistance to symptom response following chronic stress.

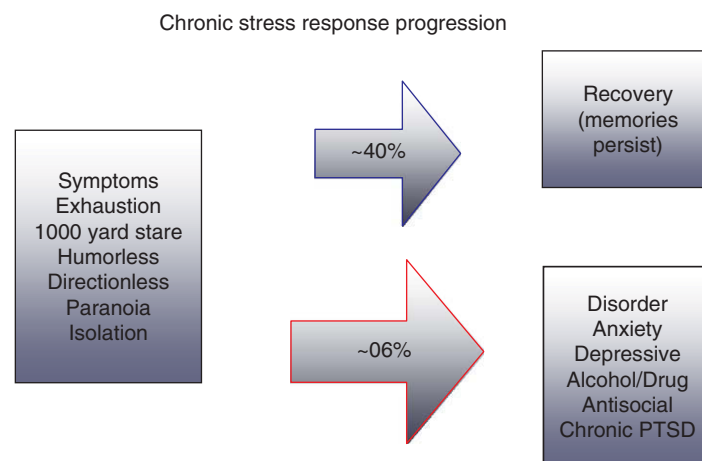


Figure 3 The progression of a chronic stress response to either recovery or to a psychiatric disorder with a more prolonged and uncertain course.

drugs), rest and replenishment, daily group therapy, and daily exercise sessions. This “walk ‘em and talk ‘em” approach resulted in return to duty rates between 40% and 70%.

Therapy may also include attempts to modify prebattle life stresses. If, for example, the individual is worried about the health of his wife or children, some communication (preferably by telephone) might be allowed. Personal visits, especially periods of home leave, are frequently a big mistake! This is because, once the person becomes enmeshed in his family setting his resolve to return to the front begins to weaken and his return to duty becomes increasingly unlikely.

In the long term, treatment of chronic combat reactions should include an appraisal of the person’s need for renewal. Often, job redefinition and/or retraining become appropriate. In animal studies of chronic stress, when conservation/withdrawal is seen, the animal becomes particularly motivated to learn new, and more adaptive, behaviors.

Today, treatment is routinely offered to soldiers returning from Afghanistan and Iraq – unlike the long delay in treatment that was the case for many servicemen from the Vietnam War. Dr. Michael Dinneen, a United States Navy psychiatrist, studied return to function of servicemen following their exposure to chronic stress. He emphasized that in assessing outcomes it is critical to know initial levels of functioning prior to the stress experience. If prior functioning was low, then recovery following stress meant a return to the low pre-existing level. Treatment following stress may well help these people make that recovery, as untreated individuals frequently develop PTSD (Figure 4). For individuals with moderate functioning prior to stress, early treatment not only helped them recover, but also often enhanced their pre-stress functioning (Figure 5). Lastly, for individuals with good

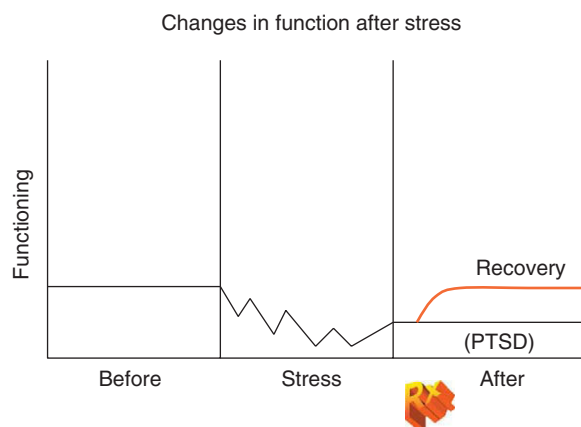


Figure 4 An illustration depicting a person’s functional outcome following stress dependent upon their prior functioning and if they received early treatment.

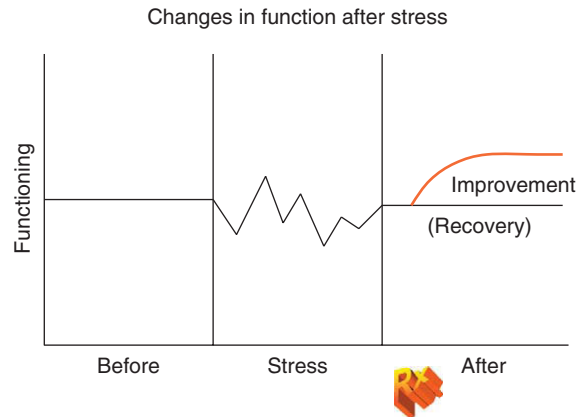


Figure 5 An illustration depicting a person’s functional outcome following stress dependent upon their prior functioning and if they received early treatment. Note the possibility of improved function with treatment.

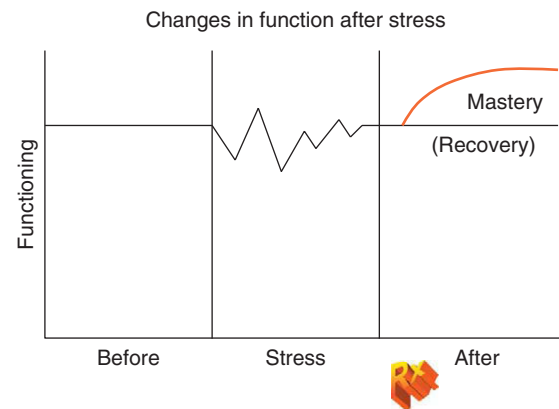


Figure 6 An illustration depicting a person’s functional outcome following stress dependent upon their prior functioning and if they received early treatment. Note the possibility of mastering the experience with treatment.

prior functioning, early post-stress treatment often facilitated an improvement in their post-stress recovery to a “mastering” level (Figure 6).

Prevention

Most line commanders will strive to control environmental stresses such as fatigue, heat, cold, hunger, and sleep deprivation in their men. However, the maintenance of communication and morale in a situation of long and deadly conflict is no easy task. A tremendously important support mechanism for American prisoners of war in Vietnam was maintenance of communication. Via the chain of command, each man was given an important role to carry out in their ongoing struggle to survive with dignity.

Another important preventive measure is having predictable intervals set aside for rest and recovery. This may not be possible in some long-term battle

situations, but when it is, the effects are extremely salutary. Knowing that one would be rotated home 1 year to the day from arrival in Vietnam greatly assisted many soldiers in getting through their combat tours. It should be noted, however, that this policy also had undesirable effects, in that it inhibited group bonding and restricted a man's identity with his unit.

One interpersonal preventive measure frequently underemployed for combat troops is to allow for closure after a period of prolonged stress. Closure was inadvertently done well in the days of troop ship transport for soldiers returning from Europe and Asia following World War II. Aboard ship the men had ample opportunities to discuss their war experiences, compare their reactions with others, and to talk out several distressing thoughts and emotions. In a related manner, bringing men home as a unit markedly helps to prevent this disorder. Talking among trusted buddies on the return home is extremely supportive.

In terms of personal preventive measures, one of the most disturbing early stresses reported by prisoners of war was their awareness of all the things they had left undone prior to being taken captive. A soldier has enough to think about during the early stages of a long-term stress without additional concerns regarding the lack of a will, the lack of sufficient insurance, and other problems left for the family. This problem surfaced again in America's recent military engagements in Afghanistan and Iraq. Especially National Guard units, where their tours of duty were extended from initial order for 6 months up to 15 months of combat duty, were affected. These soldiers often discovered their jobs were no longer waiting for them, their families could no longer afford their quarters, children were ill, and so forth, and yet they did not know when they might go home.

Almost all persons experiencing chronic stress find physical fitness, especially of an endurance nature, to be of tremendous assistance. Endurance exercises not only help to pass time during periods of lull but also aid in maintaining feelings of vigor and well being. Of nearly equal importance is that exercise greatly facilitates sleep. Finally, a fitness program may even counter the tendency toward hypoactivity seen in chronic combat reactions.

Perhaps the single most important personal preventive strategy for long-term stress is developing a sense of pace. Prisoners of war and hostages held for 1 year or more testify to the critical importance of the acquired skill of taking one day at a time. Much like a mountain climber who slowly but relentlessly proceeds up the slope, a soldier must learn to pace him/herself so that sustained function is possible over many days, months, and even years. Also like the

mountain climber, some energy should be held in reserve. Occasionally, one has to draw upon this reserve in order to meet setbacks and/or unexpected demands. Personal traits of acceptance, humility, and humor can substantially facilitate this pacing.

A word should be said about the incredible importance of humor in prolonged stress situations. The ability to laugh, not only with others but also at oneself, is vital. Former prisoners of war have claimed that single instances of a humorous circumstance made them feel better for weeks to months later.

Personal beliefs, such as conviction of cause, faith, and basic optimism, are tremendously sustaining. Even if these personal characteristics are not present in individuals at the beginning of a prolonged period of stress, they can be developed – especially through supportive group interactions.

Problems with a Diagnosis of PTSD

Recovery from a chronic stress reaction occurs far more often than what it is generally believed to be the case. This misunderstanding is in large part due to today's habit of diagnosing both acute and chronic stress reactions as PTSD. Early following a trauma, victims can be plagued by intrusive memories and dreams of the trauma, tend to avoid reminders of their experiences and develop feelings of emotional numbing. If the victim subscribes to high numbers of these symptoms, with moderate-to-high intensities, psychological raters routinely diagnose PTSD. These evaluators are apparently unaware that persons with acute and chronic stress reactions describe very similar symptom pictures. A very recent investigation found that persons suffering from nontraumatic life events gave a higher PTSD symptom profile than did other subjects that had experienced traumatic events. Therefore, the symptom profile described above is not specific to PTSD.

Another problem with liberal use of this diagnosis is the common belief that no-one ever recovers from PTSD. One large element supporting this belief is that military veterans with this diagnosis often receive pensions for this diagnosis as long as they remain symptomatic. It literally pays the man or woman to remain disabled. Yet studies of severely traumatized military veterans, such as former prisoners of war, have found that over 90% of these men not only recovered from their torture experiences but also went on to live long and productive lives. A fuller description of how these former prisoners of war achieved such remarkable recoveries is presented in another article entitled: *Captivity, Recovery from*.

A third problem resulting from the trauma criterion for PTSD in the *Diagnostic and Statistical Manual of Mental Disorders, 4th edition* (DSM IV) was that it

was expanded from personally experiencing a major trauma to now including those who only witnessed it. With this definition, at least 200 000 000 Americans fulfilled this criterion after watching television footage of the 11 September 2001 terrorist attacks in New York and Washington, DC.

After years of just focusing on the trauma and the symptom picture, recent studies have introduced the importance of genetics, family medical history, past medical history, developmental factors, and resistance to stress – as noted above. Trauma by itself is a necessary, but not by itself sufficient, explanation of the etiology of this disorder.

Lastly, distortion and fabrication have entered the picture. A recent study of persons applying for a Veterans Administration pension for PTSD found that 52% of the applicants had either unclear evidence in their medical records of combat exposure, or no evidence at all of being in combat. Another 5% of applicants' military records showed that they had never served in Vietnam. A final 2% of applicants were found to have never been in the military. These six out of 10 applicants with little to no evidence of combat stress scored higher on their PTSD symptom profiles than did the four out of 10 applicants with clear evidence in their military files of combat exposure. A final complication was found for those who did serve in combat. Some of these men reported traumas more devastating than what actually occurred when their statements were compared to their military records.

These problems with compensation for Vietnam combat veterans should not cloud the fact that many individuals did, in fact, experience war or civilian traumas of extreme severity. Further, several of these people went on to develop long-lasting physical and/or psychological disabilities. It was also true that Vietnam veterans returning home did not receive the treatment for chronic stress outlined above. In particular, they came home quickly by jet aircraft, alone from their unit, and were frequently welcomed with spit and abuse from antiwar citizens that blamed them for the war. To their credit, the Veterans Association today offers immediate access to both physical and psychological evaluation and treatment for veterans involved in today's conflicts. The military also is alert to these needs for returning soldiers remaining

on active duty. As mentioned above, early medical and psychological treatment, with a goal of recovery, can prevent some of the prolonged disability seen in Vietnam era veterans.

Further Reading

- Brill, N. Q. and Beebe, G. W. (1955). *A follow-up study of war neuroses*. Washington, DC: VA Medical Monograph.
- Burkett, B. G. and Whitley, G. (1998). *Stolen valor*. Dallas, TX: Verity Press.
- Freuh, B. C., Elhai, J. D., Grubaugh, A. L., et al. (2005). Documented combat exposure of US veterans seeking treatment for combat-related post-traumatic stress disorder. *British Journal of Psychology* 186, 467–472.
- Kushner, F. H. (1973). Doctor's report from a V.C. prison. *Medical World News*. April.
- Mullins, W. S. (ed.) (1973). *Neuropsychiatry in World War II* (Vol. II). Washington, DC: Office of the Surgeon General (Army).
- McNally, R. J. (2002). Progress and controversy in the study of posttraumatic stress disorder. *Annual Review of Psychology* 54, 229–252.
- Nice, S., Garland, C. F., Hilton, S. M., et al. (1996). Long-term outcomes and medical effects of torture among US Navy prisoners of war in Vietnam. *JAMA* 276, 375–381.
- Rahe, R. H., Rubin, R. T. and Arthur, R. J. (1974). The three investigator study: serum uric acid, cholesterol, and cortisol variability during the stresses of everyday life. *Psychosomatic Medicine* 36, 358–368.
- Rahe, R. H. and Genender, E. (1983). Adaptation to and recovery from captivity. *Military Medicine* 148, 577–585.
- Sachar, E. J. (1980). Hormonal changes in stress and mental illness. In: Krieger, D. T., Hughes, J. D. & Sunderland, M. A. (eds.) *Neuroendocrinology*. Sinquer Boston, MA: Associates Inc.
- Saskia, S. L., Mol, A. A., Job, F. M., et al. (2005). Symptoms of post-traumatic stress disorder after non-traumatic events: evidence from an open population study. *British Journal of Psychology* 186, 494–499.
- Schwinn, M. and Diehl, B. (1973). *We Came to Help*. New York: Harcourt, Brace, and Jonanovich.
- Segal, J. (1973). Therapeutic considerations in planning the return of American POWs to the continental United States. *Military Medicine* 138, 73–77.
- Yahuda, R. and McFarlane, M. B. B. S. (1995). Conflict between current knowledge about posttraumatic stress disorder and its original conceptual bases. *American Journal of Psychology* 152, 1705–1713.

Combat Stress Reaction

M Dobson

New South Wales Health Department,
North Sydney, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1,
pp 495–500, © 2000, Elsevier Inc.

Introduction

Combat Stress Reaction and Combat-Related
Posttraumatic Stress Disorder
Risk Factors for Combat Stress Reaction
Treatment of Combat Stress Reaction
Prevention of Combat Stress Reaction
Summary

Glossary

<i>Combat stress reaction (CSR)</i>	The acute psychological breakdown experienced by combat soldiers as a result of their exposure to the stress of individual, unit, and battle factors in the war zone.
<i>Expectation</i>	A treatment principle that states that psychiatric casualties should be treated with the expectation that soldiers will return to their unit.
<i>Immediacy</i>	A treatment principle that states that psychiatric casualties should be treated as soon as possible after symptom onset.
<i>Posttraumatic stress disorder (PTSD)</i>	A chronic pathological reaction following exposure to a traumatic event characterized by the intrusive reexperiencing of the traumatic event, the avoidance of stimuli associated with the traumatic event, a reduced responsiveness to the external world, and a range of autonomic and cognitive symptoms.
<i>Proximity</i>	A treatment principle that states that psychiatric casualties should be treated on or near the battlefield.
<i>Risk factor</i>	A variable that epidemiological evidence suggests is associated with an increased probability of a specific health outcome; the nature of this association is not necessarily causal.

CSR is an acute psychological breakdown experienced by combatants in a war zone. It is characterized by labile polymorphic manifestations ranging from overwhelming anxiety, irritability, and restlessness to psychological withdrawal, confusion, and psychomotor

retardation. Other characteristics include apathy, anxiety, depression, paranoia, conversion symptoms, signs of sympathetic arousal, and psychosomatic reactions such as diarrhea, abdominal pains, nausea, and vomiting. The defining criterion for CSR is that the soldier ceases to function effectively as part of his or her combat unit by acting in a manner that endangers his or her life and/or the lives of his or her fellow soldiers.

Introduction

CSR is a significant concern for the military because of its potential impact on the fighting effectiveness of combat units. It has been estimated that the ratio of psychiatric casualties to the number of soldiers killed and the number wounded in action during a battle of average intensity is 1 : 1 : 4. Belenky reported that the ratio of psychiatric casualties to wounded was 30 : 100 in the 1973 Arab-Israeli War and 23 : 100 in the 1982 Lebanon War. By way of comparison, the ratio of psychiatric to wounded casualties for the American army in World War II was 36 : 100. All these estimates attest to the significant impact of psychiatric casualties on the fighting capacity of a military unit.

For a number of reasons, the availability of scientific evidence on CSR is limited. In part, this is due to the difficulty of conducting prospective epidemiological investigations on the acute stress responses of combat soldiers during wartime. It is also a reflection of the elusive nature of CSR. That is, the polymorphic and labile nature of CSR means that clear diagnostic criteria cannot be developed. Another reason is that the introduction of the PTSD diagnosis in the *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn.; DSM-III) marshaled an extensive research effort into examining the long-term effects of combat exposure, particularly among U.S. veterans of the Vietnam War. Although the political and social factors driving this research effort were understandable, the voluminous research literature available on PTSD serves only to highlight the absence of empirical data on CSR. The only exception to this is a series of controlled studies on Israeli CSR casualties by Zahava Solomon and her colleagues. The aim of this article is to provide an overview of CSR by reviewing the literature on (1) the relationship between CSR and combat-related PTSD, (2) the identification factors for CSR, (3) the effectiveness of frontline treatment for CSR, and (4) the use of prevention programs to reduce CSR symptoms.

Combat Stress Reaction and Combat-Related Posttraumatic Stress Disorder

An adequate conceptualization of CSR must define the relationship between this acute reaction and the range of other stress responses that may be experienced by combat soldiers. Combat exposure can elicit a range of stress responses that can be divided into three categories: (1) a transient, nonpathological stress response, (2) CSR, and (3) PTSD. This division recognizes the need to distinguish between a soldier's adaptive response to combat stress and an acute or chronic reaction. An adaptive response is the manifestation of a transient stress response that is commensurate with the stress of daily life on or near the battlefield. CSR differs from a transient stress response in that it is an acute stress reaction that incapacitates soldiers to the degree that they can no longer function in combat. In some cases, CSR can be transient, with soldiers able to return to military duty after a few days rest from combat. In other cases, it may be the precursor to the development of combat-related PTSD. This form of PTSD is a chronic psychopathological stress reaction that occurs in soldiers after they have been exposed to a combat-related trauma. It should be noted that CSR is not a necessary precursor to the onset of combat-related PTSD. However, it has been suggested that CSR creates a latent vulnerability that can lead to the subsequent development of combat-related PTSD long after the war has ended.

CSR and combat-related PTSD are similar in that they share a number of common symptoms such as increased physiological arousal, startle reactions, restlessness, and psychological withdrawal. However, CSR differs from PTSD in that it may not necessarily develop following exposure to a specific combat event. CSR also differs in its symptoms of confusion, nausea, vomiting, and paranoid reactions. The critical pathological difference between these stress responses is that the symptoms associated with combat-related PTSD will endure. In PTSD, the symptom duration is expected to be at least a month, with the trauma being persistently reexperienced throughout the course of the disorder.

The relationship between CSR and combat-related PTSD has been extensively investigated in a series of studies examining the postwar adaptation of Israeli combat veterans. Solomon, Weisenberg, Schwarzwald, and Mikulincer assessed the prevalence and severity of combat-related PTSD in 382 Israeli CSR casualties 1 year after they had received treatment. Control group subjects were selected from those combat veterans who had fought in the same units as the CSR casualties but had not exhibited any symptoms or presented for

treatment. Matching with CSR casualties also occurred on the following variables: age, education, military rank, and assignment. The results showed that 59% of the CSR group developed PTSD, compared to 16% in the control group. There was also a significant difference in the number of PTSD symptoms reported. The CSR group patients reported a higher number of PTSD symptoms than their control group counterparts who also had PTSD. These findings suggest that those CSR casualties who develop PTSD experience a more intense form of the disorder than those combat veterans who develop PTSD without a history of CSR.

In a subsequent study, Solomon and Mikulincer found that 1, 2, and 3 years after the war, symptoms of combat-related PTSD were more prevalent among CSR casualties than among their matched controls. Similar trends were observed for social functioning, perceived self-efficacy in battle, and somatic complaints. The only reduction in the symptom levels observed during the study occurred when the intensity of intrusion and avoidance symptoms dropped substantially in the third year. However, the authors noted that this reduction in symptom levels should not be seen as evidence that the CSR veterans were getting better. These findings suggest that CSR casualties are more vulnerable to the subsequent development of combat-related PTSD.

Risk Factors for Combat Stress Reaction

The identification of risk factors for CSR is central to an understanding of the etiology of combat-related psychopathology. As discussed earlier, the epidemiological literature on Israeli combat veterans has shown that a higher percentage of CSR casualties subsequently developed PTSD than did controls with no CSR history. This suggests that an understanding of the risk factors for CSR may provide further insight into the factors associated with the development of chronic combat-related stress reactions such as PTSD.

The intensity of combat exposure has been identified as a significant risk factor for CSR. Belenky reported the findings from an Israeli Defence Force study that correlated objective measures of combat stress such as battle type and availability of tactical and logistic support with the incidence of physical and psychiatric casualties. This study found that battalions exposed to higher levels of combat stress had both a higher casualty rate and a higher ratio of psychiatric to physical casualties.

In addition to the intensity of combat exposure, there are a variety of other factors in the combat environment that may affect CSR onset. In a review of the clinical presentation of 158 U.S. army stress casualties during Operation Desert Storm, McDuff

and Johnson identified a number of environmental stressors associated with stress reactions. Although this study does not purport to be a systematic investigation of CSR risk factors, it demonstrated that a soldier's exposure to fatigue, cold, sleep deprivation, poor unit leadership, lowered morale, and threats to personal safety led to the onset or worsening of combat-related stress symptoms.

The identification of modifiable risk factors for CSR will play an important role in the development of effective prevention programs for combat soldiers. Despite this, there are few published studies that focus solely on CSR risk factors. A search of the recent research literature identified two controlled studies that investigated the association between a range of variables and CSR onset. Solomon, Noy, and Bar-On examined the association between demographic, military, and personality variables and CSR. One of the strengths of this study is that the researchers were allowed access to official military records in order to collect data on these variables. Age, education, combat suitability, military rank, and type of military service were identified as risk factors for CSR. These risk factors suggest that the onset of CSR can be accounted for by a combination of individual and situational factors.

This study found that 80% of the sample of psychiatric casualties were reservists. The psychiatric casualty rate for reservists was in marked contrast to the rate for career soldiers, which was the lowest of any of the groups studied. It appeared that a soldier's reservist status was a marker for a number of interrelated risk factors, all of which increased vulnerability to combat stress. Because the reservists were the oldest group of soldiers studied, the authors suggested that their age may have reduced their capacity to deal with the physical aspects of frontline service. Furthermore, reservists were more likely to have been exposed to combat in previous wars, which in turn may have reduced their capacity to cope with any subsequent exposure to combat stress. Being older, reservists were more likely to have established family and career commitments. These factors increased the stress on reservists when they had to adapt from civilian to military life. Also, in contrast to the career soldier, reservists had to make this change in a relatively short period of time. Solomon and Flum investigated the association between life events and CSR. They hypothesized that a combatant's exposures to stressful life events before the war may increase their vulnerability to CSR. However, the findings of this study did not provide support for this hypothesis. There was no relationship between the number, magnitude, or type of life events experienced by individuals prior to combat and CSR onset. However,

prewar life events were implicated in the subsequent onset of PTSD.

Treatment of Combat Stress Reaction

Frontline treatment interventions for psychiatric casualties are based on the principles of proximity, immediacy, and expectancy. These principles indicate that casualties should be treated in a setting close to the combat zone (proximity) as soon as possible after the onset of the CSR (immediacy) with the expectation that the soldiers will return to the frontline following their recovery (expectation). In the frontline setting, the main aim of treatment is to restore a casualties' depleted physiological condition by satisfying their needs for sleep, food, and drink. This usually lasts for a few days and takes place in an area of relative safety. During this time, casualties are provided with the opportunity to discuss their traumatic experiences.

One of the advantages of providing treatment on or near the frontline is that it allows casualties the opportunity to maintain their role as soldiers in a military unit. This enables access to their combat peers and opportunities for ventilation, which, in turn, reinforces the soldiers' sense of commitment to their combat unit. Another benefit of treatment in this setting is that it acts as a form of desensitization by exposing the soldiers to conditions similar to those that precipitated the CSR. Also, the immediate provision of treatment in the frontline places the expectation on the soldiers that they will be able to return to their combat duties.

Other principles of combat psychiatry that have been cited in the literature include simplicity and centrality. Simplicity refers to the use of uncomplicated treatments such as rest, food, and hot showers to provide the CSR casualty with a respite from the stresses of combat. Centrality refers to the process of assessing all CSR casualties at a central screening area so that an experienced team of mental health professionals can minimize the likelihood of inappropriate evacuations to the rear.

Although the principles of frontline treatment have been in use since World War I, it is only recently that the question of their effectiveness has been subjected to empirical investigation. Belenky highlighted the successful use of frontline treatment by the Israeli army during the Lebanon War, with 60% of CSR casualties returning to combat duties within 72 h. This return rate confirms the earlier findings of Solomon and Benbenishty, who investigated the relative contribution of proximity, immediacy, and expectancy in CSR recovery. Their study followed up 82% of the Israeli army CSR casualties of the Lebanon War to compare the effectiveness of frontline treatment

with treatment in a civilian setting at the rear (for security reasons, the precise number of subjects involved in this study could not be provided by the authors). Data on the immediacy and expectancy principles were collected through the use of self-report questionnaires. The outcome measures in this study were return to military unit and PTSD onset.

The results demonstrated a strong association between each of the frontline treatment principles and CSR casualties' unit return rate. A high percentage of CSR casualties who were treated at the front (43%) or near the border with Lebanon (59%) returned to their units. In contrast, a significantly lower percentage of CSR casualties returned to their units after being airlifted to the rear (21%) or treated in a civilian facility at the rear (28%). A high percentage of those soldiers who received treatment immediately after the onset of a CSR returned to their units (44%), whereas those who waited until after the war had ended before seeking treatment showed the lowest return rate (24%). Of the three treatment principles, expectancy displayed the strongest association with the unit return rate. CSR casualties who perceived that their therapist expected them to return at all costs had a higher unit return rate (53%) than casualties who were unclear about their therapist's expectations (26%).

The application of frontline treatment principles also had a significant impact on PTSD. Although the relationship between proximity and PTSD did not reach statistical significance, there was a strong trend suggesting that treatment proximity was a factor in the subsequent development of PTSD. When treatment occurred at the front or near the border, the percentages of CSR casualties who subsequently developed PTSD were 52 and 48%, respectively; when treatment occurred outside the war zone, this percentage increased significantly (66%). There was also a significant relationship between treatment immediacy and PTSD onset. In this study, the highest rate of PTSD (74%) was observed among those CSR casualties who had waited until after the war to seek treatment. Similar to the unit return rate findings, Solomon and Benbenishty reported a strong relationship between expectancy and PTSD. Casualties who perceived that their therapist expected that they would return to their unit at all costs exhibited the lowest rate of PTSD (55%).

Prevention of Combat Stress Reaction

Because combat is an unpredictable and chaotic experience, prevention programs for combat soldiers can only be formulated in general terms. The development of specific prevention programs targeting

CSR must address the complex interaction among individual, unit, and battle factors that determine a soldier's vulnerability to combat stress. These programs need to recognize that the occupation of combat soldier routinely exposes individuals to multiple stressful and/or traumatic events and that, over time, this reduces the soldiers' capacity to cope with combat stress. This is further complicated by the soldiers' perception of combat and their role in the war zone. Soldiers' subjective interpretations of a battlefield event as stressful or traumatic will be shaped by their acceptance of the violent nature of their occupation as well as by their previous experiences with such an event.

Prevention programs for CSR should also acknowledge that combat soldiers can be exposed to high levels of prolonged combat stress without ever being exposed to a specific traumatic event. This means that prevention programs that target specific mental health outcomes such as PTSD should not be seen as providing a complete alternative to the complex task of preventing CSR. As military personnel, combat soldiers also have to deal with all the normal occupational stresses that are experienced on a daily basis by the members of any large hierarchical organization in the civilian setting. It has generally been expected that, in the combat setting, the individuals most likely to experience this form of prolonged occupational stress, such as combat soldiers or medical personnel, will by virtue of their selection, training, and background be able to cope with the effects of their exposure.

The U.S. army's approach to the management of combat stress during Operation Desert Storm highlighted the importance of identifying and managing individual, unit, and battle factors in the onset of CSR. Examples of individual factors include combat experience, sleep, food, fitness, and stability of personal life. Also, soldiers' perceptions of their preparedness for combat, the loss of a family member at home, and/or a major threat to family during their service can affect their capacity to deal with combat stress. Unit factors include length of rest periods, the tactical situation, adequacy of supplies, communications, terrain, group cohesiveness, leadership, and morale. As noted previously, the intensity of combat exposure is a known risk factor for CSR. Specific examples of combat stressors include exposure to human suffering, risk of capture by the enemy, fear of torture, threat of personal injury, and death.

It is argued that prevention programs must target the range and type of combat stressors that soldiers are likely to encounter. Stress management techniques can be used to educate soldiers about the inevitability of experiencing fear during combat. The value of this experience should be emphasized, and soldiers should

be assisted in managing any anticipatory anxieties that they might have about maintaining control of their emotions in combat. Other stress management techniques such as problem solving, conflict resolution, and relaxation can also be used to strengthen soldiers' abilities to cope with the stressors that they are likely to encounter in the course of fulfilling their military duties as a combat soldier. These techniques can also be used by combatants who have the added responsibility of leadership to manage their own stress levels.

Military training provides an opportunity for the prevention of combat stress through the use of realistic training exercises. Using live ammunition, these exercises are designed to simulate the life-threatening challenges inherent in the combat experience. The successful completion of such exercises should increase soldiers' confidence in their ability to deal with the future stress of combat. As Oei, Hennessy, and Lim noted in their review of the literature on military training, the use of well-rehearsed drills during training exercises should lead to the development of automatic reactions in soldiers. When these are activated during combat, they reduce the level of cognitive demands on individuals, which in turn minimizes the effect of combat stress.

For military personnel who are in a support role, which may or may not involve active participation in combat, a general program of education about the likely occupational and interpersonal stresses that are frequently encountered in this setting would be beneficial. Based on the experiences of military personnel during the Vietnam War, this might include discussing (1) the problems associated with fighting an ambiguous enemy, (2) the definition of individuals' roles in the war zone, (3) a realistic appraisal of their contribution to the war effort, and (4) reasonable expectations of leadership.

Summary

CSR is an acute psychological breakdown that incapacitates soldiers to the degree that they are no longer able to function in combat. It is characterized by labile and polymorphic manifestations ranging from overwhelming anxiety and confusion to psychological withdrawal and psychomotor retardation. The variety of symptoms that are manifested during CSR has meant that clear diagnostic criteria have not been developed. Risk factors for CSR include the intensity of combat exposure, age, education, combat suitability, military rank, and type of military service. Treatment for CSR is based on the principles of proximity, immediacy, and expectancy. These treatment principles emphasize the need to treat CSR casualties

in a setting near the combat zone as soon as possible after symptom onset and with the expectation that the soldier will be able to return to combat after a brief period of rest and recuperation. The task of preventing CSR provides a complex challenge to military psychiatry. It requires an understanding of the interplay between a diverse range of individual, unit, and battle variables. Based on the evidence available to date, the use of realistic exercises during military training and the implementation of stress management techniques can play a role in minimizing the impact of combat stress.

See Also the Following Articles

Combat, Acute Reactions to; Combat Reaction, Chronic; Korean Conflict, Stress Effects of; Persian Gulf War, Stress Effects of.

Further Reading

- American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd edn.). Washington, DC: American Psychiatric Association.
- Belenky, G. (1987). Psychiatric casualties: the Israeli experience. *Psychiatric Annals* 17(8), 528–531.
- Camp, N. (1993). The Vietnam war and the ethics of combat psychiatry. *American Journal of Psychiatry* 150, 1000–1010.
- Dobson, M. and Marshall, R. (1997). Surviving the war zone experience: preventing psychiatric casualties. *Military Medicine* 162, 283–287.
- Grinker, R. and Spiegel, J. (1945). *Men under stress*. New York: Blakiston.
- Jones, F. and Hales, R. (1987). Military combat psychiatry: a historical review. *Psychiatric Annals* 17(8), 525–527.
- Kormos, H. (1978). The nature of combat stress. In: Figley, C. (ed.) *Stress disorders among Vietnam veterans*, pp. 3–22. New York: Brunner/Mazel.
- McDuff, D. and Johnson, J. (1992). Classification and characteristics of army stress casualties during operation desert storm. *Hospital and Community Psychiatry* 43(8), 812–815.
- Marshall, R. and Dobson, M. (1995). The treatment of trauma reactions in war veterans: the seductive elegance of the PTSD diagnosis. In: Vialle, W. (ed.) *Why psychology?: selected papers from the 29th annual Australian Psychological Society conference*.
- Miller, L. (1995). Tough guys: psychotherapeutic strategies with law enforcement and emergency services personnel. *Psychotherapy* 32, 592–600.
- Oei, T., Lim, B. and Hennessy, B. (1990). Psychological dysfunction in battle: combat stress reactions and post-traumatic stress disorder. *Clinical Psychology Review* 10, 355–388.
- Solomon, Z. (1993). *Combat stress reaction: the enduring toll of war*. New York: Plenum.
- Solomon, Z. and Benbenishty, R. (1986). The role of proximity, immediacy, and expectancy in frontline treatment

- of combat stress reaction among Israelis in the Lebanon war. *American Journal of Psychiatry* 143(5), 613–617.
- Solomon, Z. and Flum, H. (1988). Life events, combat stress reaction and post-traumatic stress disorder. *Social Science & Medicine* 26(3), 319–325.
- Solomon, Z., Laror, N. and McFarlane, A. (1996). Acute posttraumatic reactions in soldiers and civilians. In: Van der Kolk, B., McFarlane, A. & Weisaeth, L. (eds.) *Traumatic stress: the effects of overwhelming experience on mind, body and society*, pp. 102–114. New York: The Guilford Press.
- Solomon, Z. and Mikulincer, M. (1987). Combat stress reactions, posttraumatic stress disorder, and social adjustment: a study of Israeli veterans. *Journal of Nervous and Mental Disease* 175(5), 277–285.
- Solomon, Z., Noy, S. and Bar-On, R. (1986). Risk factors in combat stress reaction – a study of Israeli soldiers in the 1982 Lebanon War. *Israeli Journal of Psychiatry & Related Sciences* 23(1), 3–8.
- Solomon, Z., Weisenberg, M., Schwarzwald, J., et al. (1987). Posttraumatic stress disorder among frontline soldiers with combat stress reaction: the 1982 Israeli experience. *American Journal of Psychiatry* 144(4), 448–454.
- Weisaeth, L. (1996). PTSD: the stressor-response relationship. *Baillières Clinical Psychiatry* 2(2), 191–216.

Combat, Acute Reactions to

R H Rahe

University of Washington School of Medicine,
Seattle, WA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
R H Rahe, volume 1, pp 487–490, © 2000, Elsevier Inc.

Introduction

Acute Reactions to Combat Stress

Predispositions and Precipitants

Treatment

Prevention

Introduction

Disabling psychological reactions to stress, particularly to combat stress, have been witnessed frequently but codified differently in the various wars occurring over the past 150 years. Diagnoses coined for combat stress have varied from soldier's heart in the Civil War to shell shock in World War I to war neurosis and battle fatigue in World War II to combat stress in the Korean War and Vietnam conflict. In most instances, what was defined was either a precipitous or gradual breakdown in a serviceman's ability to perform his combat functions, accompanied by a variety of acutely distressing and/or chronically disabling mind–body symptoms.

Drawing from the military psychiatry literature, as well as from studies of human responses to stress encountered in a variety of civilian settings, a case can be made for distinguishing acute combat reactions from chronic ones. In so doing, clinical

presentations, predispositions, precipitating factors, treatment approaches, and expectations about return to duty become more specific.

Acute Reactions to Combat Stress

The Israeli Defense Forces in the 1980s and 1990s used the term battle shock to denote the acute and disabling psychological reactions to combat that appear within minutes to hours after stress exposure. In civilian settings, we find very similar conditions of acute stress reactions to trauma. These reactions are characterized by an abrupt onset of physiological hyperarousal following an exposure to an intensely stressful event(s) occurring during the previous hours, days, or weeks. The signs and symptoms of this disorder and its treatment are considered here.

Unsettling but manageable manifestations of this reaction are common to all humans facing trauma. It is, therefore, the degree of response rather than symptoms themselves that leads to a diagnosis. An early sign of vulnerability to military stress is a marked reluctance of the soldier to leave a secure setting. These people often are the last in line when heading toward danger, casting frequent glances back toward safety. They may unnecessarily check and recheck their equipment – these adventitious movements representing a displacement of anxiety. They may also show considerable difficulty understanding instructions and carrying out even simple tasks.

Widely opened eyes, dilated pupils, and tremulousness and rapid shallow breathing signal strong sympathetic nervous system activation. Sympathetic nervous system hyperarousal is the major physiological component of an acute stress reaction. When

blood or urinary measures for catecholamines are gathered from soldiers with this disorder, epinephrine and norepinephrine are both found to be significantly elevated. It is also likely that the soldier's serum cortisol is also elevated throughout this reaction.

As an acute combat reaction worsens, the affected soldier may not take cover during an assault, or he may remain hidden in a bunker and unable to fire his weapon or to care for a buddy in trouble. In its most severe form, affected individuals may show an overflow of undirected motor activity, which may mimic tics or seizures. Conversely, hyperarousal can also lead to a freezing of motor function, which may resemble paralysis.

An acute combat stress reaction can develop, and resolve, within a matter of minutes. Because recovery from this entity can be very rapid, the soldier may show a dramatic transition from gross panic at one instant to rational thought and behavior a few minutes later. Therefore, a cardinal rule is not to move an afflicted soldier further than the battle aid station closest to the front lines and assume there will be a rapid recovery.

Predispositions and Precipitants

It should be emphasized that no single factor by itself, even if the stress is tremendous, totally accounts for the development of an acute stress reaction. Rather,

it is the sum of separate forces acting in concert over a brief period of time. These forces usually include developmental factors, environmental conditions, and interpersonal stresses.

Developmental factors influencing susceptibility to acute stress response include childhood family traumas, especially abuse and neglect; poor socialization while growing up; low educational attainment; and little success in managing major life stresses (Figure 1). These individuals often join the military to escape from their troubled lives. The last thing they expect, and what their early lives has left them unprepared to handle, is to go into combat.

In contrast, resistance factors to acute stress reaction include enriched childhood and developmental experiences, prior stress experiences that the person handled successfully, high intelligence and educational achievement, plentiful social support, and prior training on how to handle stress (Figure 2). A positive attitude and an ability to resist rash behaviors are traits that enhance resistance to stress.

Environmental stresses, such as fatigue, hunger, sleep loss, cold, and heat, are common experiences of men in battle. They take on additional significance when other precipitants of acute battle stress co-exist, for example, when a battle is being waged in sustained heat (such as a desert war) and soldiers are dehydrated. The high magnitude of the stress situation, especially with a high casualty rate, makes stress

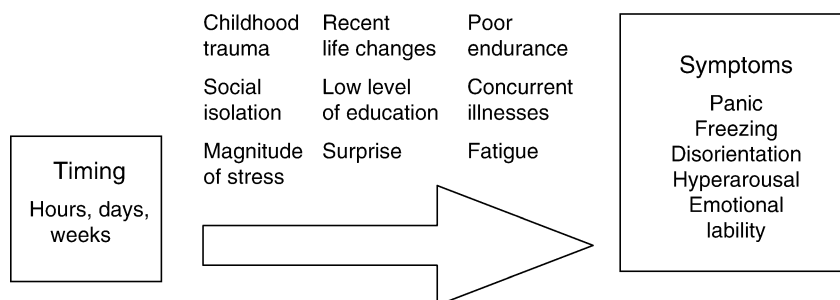


Figure 1 Early developmental and personal susceptibility factors for acute stress response.

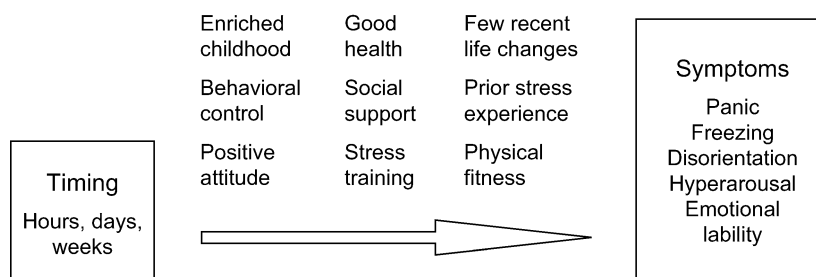


Figure 2 Early developmental and personal resistance factors for acute stress response.

reactions more likely to occur. A surprise event leading to disorientation, concurrent illness, and poor physical endurance add even further to a susceptibility to acute stress response.

Interpersonal factors also contribute to susceptibility to stress, including recent life stresses outside of combat, such as family upheaval at home, financial problems, and difficulties getting along with peers. Reserve status, rather than a regular commission, and membership in a support unit (usually armor) are frequently an interpersonal precipitant of this disorder. Soldiers, who are members of elite units, in which the esprit de corps and morale are generally quite high, tend to perform far better in combat than men in less cohesive units. A further interpersonal stress occurs when reserves are brought forward to replace the dead and injured. These reserves may be assigned to tanks that are still bloody from the previous campaign and find themselves communicating, by radio, with people they have never met. A replacement soldier is seen as a new guy by the others and is a reminder that one of their buddies was killed or severely wounded. It may take weeks to months for a new soldier in a unit to lose this new guy label.

Factors of older age and inexperience are two additional interpersonal precipitants of this stress response. An infantryman in his late twenties or early thirties has a much higher probability of developing an acute combat reaction than one in his late teens or early twenties. Soldiers are especially vulnerable to acute stress reactions when they witness death for the first time. One of the major functions of military training is to prepare soldiers for the rigors of battle so that they can function almost reflexively in that environment. Yet it is impossible to train men for the experience of seeing their buddies killed.

Treatment

The treatment of an acute combat reaction should be immediate. It is far easier to reverse this reaction in its early stages than when the condition lingers. This means that the location where treatment is provided is usually at, or very near, the front lines of battle. Significantly, the person who carries out treatment is usually not a doctor; he is more likely to be a buddy or a medic/corpsman.

The most important first step in treatment is to help the affected individual gain control over his hyperaroused physiology. Rapid, shallow breathing will quickly cause hypocapnia with resultant light-headedness and feelings of panic. Hypothermia, hyperthermia, and tachycardia can produce similar effects. The

first rule in such a situation is to stop! Once the person begins to cool down (often literally), rational thinking can return. Having the affected person count slowly or make respiratory expirations last longer than inspirations is an easy remedy. When rational thinking returns, reorientation can take place and priorities can be reestablished.

An acutely stressed individual often has a need to talk about his recent trauma. A buddy or medic should listen to the story once, at most twice. Further recitations will have very little anxiety-reducing effect. Following one or two recitals, a soldier should be evaluated to determine if he is ready to be sent back to his unit (with a buddy) or needs rest and replenishment over the next 24 to 48 hours. These men can be treated at a battle aid station or in a separate section of a field hospital. Medications, if required, are generally hypnotics and anti-anxiety preparations. Because of this condition's ready reversibility, return to duty rates can be as high as 90%.

It has become a rule not to mix soldiers with disabling psychological reactions to combat with others who are physically wounded. When that occurs, men with psychological reactions quickly identify with the wounded and their psychological disabilities increase. A special effort must be made to keep these two classes of patients apart; otherwise, in the normal logistical flow of patients rearward, those with acute combat reactions will be separated from their unit. A special problem arises when a soldier experiences both an acute combat reaction and a physical wound. He is usually treated for the wound and not for his psychological condition.

In civilian settings, people with acute stress disorders should ideally be treated with psychiatric first aid. Groups of disordered individuals should be gathered in a large room such as a cafeteria. They should be able to sit around small tables and undergo a brief interview concerning their stress experience (again, no more than two recitals), sign a roster of attendance with home address and telephone, and be given some paper and a pen to write out their recent stress experience. Putting their stresses down in writing helps them to regain their concentration and reduces inclinations to repeatedly talk about their stories to others. Snacks should be available; foods high in fat, such as donuts, are most comforting. Doctors, nurses, and chaplains should act like they are social workers; that is, they should periodically circle the room asking victims if they have eaten recently, direct them to Red Cross personnel to inquire about family and friends, and to provide counseling as needed. When ready to leave the room, transportation to their homes should be facilitated if possible. It is

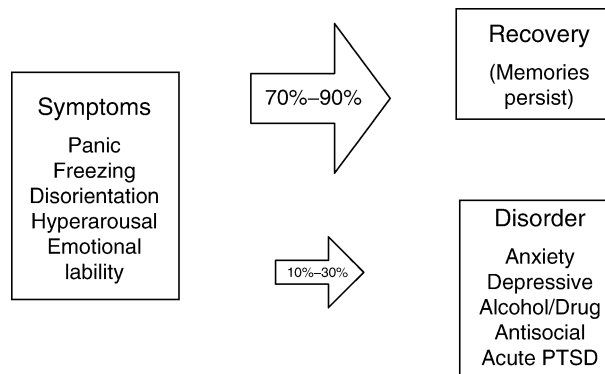


Figure 3 Acute stress response progression to rapid recovery or a psychiatric disorder with a more uncertain outcome.

counterproductive for psychiatrists and psychologists to set up an office and expect victims to come in for a lengthy session.

The good news about acute reactions to stress, including combat stress, is that they are short-lived and recovery from distressing symptoms is generally seen in 70–90% of affected people (Figure 3). A lasting disorder can occur in 10–30% of victims. Recovery from acute stress does not imply that disturbing memories of the event go away. Instead, several memories may persist during the person's lifetime. They do change from frequent awareness to distant recollections over time. However, during and shortly following a new and severe acute stress situation, these memories can make a renewed, and usually temporary, appearance.

Prevention

The environmental, interpersonal, and personal aspects of prevention should be considered as much as the precipitating event of this disorder in a multifactorial framework. It takes efforts in several areas to appreciably reduce the likelihood of acute stress reactions.

It is military doctrine to keep soldiers properly fed, warm, and rested whenever possible. When these conditions cannot be met, as is frequently the case in battle, command should realize that the chances for acute stress reactions are thereby increased. Although resupplies may not be under a commander's control, communications with his men regarding the progress of the campaign, the motivations and armament of the enemy, and so forth are certainly within his influence. Also, battle plans should be phrased in clear, concise, easy-to-remember language.

A most effective preventive force is realistic stress training. Seldom, however, is such training carried out. Perhaps the best example of preventive training that is currently performed is survival, evasion,

resistance, and escape training conducted by the U.S. Navy and U.S. Air Force. These 1–2 week programs are designed for aviators, who by the nature of their duties are at high risk for capture and prisoner-of-war status. However, realistic training for combat troops could also be provided. An important element of such training is allowing soldiers to actually experience panic symptoms under controlled conditions. The treatment steps of stopping and controlling hyperphysiology (allowing return of rational thinking and action) are then implemented. Also, mental rehearsal prior to stress training of the steps a soldier should follow in a panic situation should be part of this training.

Physical fitness leading to increased endurance leads to increased parasympathetic tone. Parasympathetic activation may well act as a brake on the sympathetic nervous system's hyperarousal during panic-producing situations. A soldier's personal commitment to the battle and attitude under stress are greatly shaped by his training. In civilian settings, a positive attitude, with realistic optimism can stave off a stress reaction. It was John Wooden, the eminently successful basketball coach at UCLA, who encouraged his players to "Be quick, but don't hurry." Such an attitude imparts a perception of personal control over fast-moving events.

Further Reading

- Brill, N. Q. and Beebe, G. W. (1955). *A follow-up study of war neuroses*. Washington, DC: U.S. Government Printing Office.
- Belenky, G. L., Tyner, C. F. and Sodetz, F. J. (1983). *Israeli battle shock casualties: 1973 and 1982*. Report NP-83-4. Walter Reed Army Institute of Research, Division of Neuropsychiatry, Washington, DC.
- Glass, A. J. (1958). Observations upon the epidemiology of mental illness in troops during warfare. In: *Symposium on preventive and social psychiatry*, pp. 185–197. 15–17 April. Washington, DC: U.S. Government Printing Office.
- Grinker, A. A. and Spiegel, J. P. (1945). *Men under stress*. Philadelphia: Blakiston.
- Keller, S. and Seraganian, P. (1984). Physical fitness and autonomic reactivity to psychosocial stress. *Journal of Psychosomatic Research* 28, 279–287.
- Medical and surgical history of the war of the rebellion* (vol. 1). (1870). Washington, DC: U.S. Government Printing Office.
- Rahe, R. H. (1988). Acute versus chronic psychological reactions to combat. *Military Medicine* 153, 365–372.
- Shalev, A. Y., Schreiber, S. and Galai, T. (1993). Early psychiatric responses to traumatic injury. *Journal of Traumatic Stress* 6, 441–450.
- Trimble, M. R. (1981). *Post-traumatic neurosis*. New York: John Wiley & Sons.

- Turner, M. A., Kiernan, M. D., McKechnie, A. G., et al. (2005). Acute military psychiatric casualties from the war in Iraq. *British Journal of Psychiatry* 186, 476–479.
- Yehuda, R., McFarlane, A. C. and Shalev, A. Y. (1998). Predicting the development of posttraumatic stress

- disorder from the acute responses to a traumatic event. *Biological Psychiatry* 44, 1305–1313.
- Yitzhake, T., Solomon, Z. and Kotler, M. (1991). The clinical picture of acute combat stress reaction among Israeli soldiers in the 1982 Lebanon War. *Military Medicine* 156, 193–197.

Common Cold and Stress

A Smith

Centre for Occupational and Health Psychology,
Cardiff University, Cardiff, UK

© 2007 Elsevier Inc. All rights reserved.

Introduction

Early Research on Naturally Occurring Colds

Virus Challenge Studies

Duration and Type of Stressor

Underlying Mechanisms

Conclusions

Glossary

<i>Illness</i>	Presence of symptoms or signs.
<i>Infection</i>	Presence of virus in the body.
<i>Negative life events</i>	Major events such as bereavement, divorce, and loss of job.
<i>Secretory immunoglobulin A (sIgA)</i>	The most abundant immunoglobulin; it provides a barrier to invasion, and there is evidence that there is an inverse relationship between levels of sIgA and vulnerability to infection

Introduction

Research on stress and the common cold is a relatively new topic. The area was not reviewed prior to the 1990s, but since then it has become an important topic because it exemplifies many of the general features of the associations between stress and health, provides a good example of interactions between the brain, behavior, and the immune system, and has implications for life-threatening infectious diseases such as AIDS. Initial approaches to understanding the common cold used biomedical models in which the person was seen as a passive victim of exposure to the virus. However, the possible importance of psychological factors began to be considered, for two main reasons. First, when exposed to cold-producing

viruses, only a proportion of people develop the illness. Second, the severity and duration of symptoms vary widely among those who become ill. These two facts led to consideration of factors such as stress in the etiology and pathogenesis of the common cold.

A variety of measures of infection and illness have been examined in the research. Studies of infection have used virus isolation and antibody change as the outcomes. The severity and duration of the illness have been assessed using symptom checklists (e.g., runny nose, blocked nose, sore throat) and signs (e.g., weight of nasal secretion). Use of medication and the seeking of medical care have also been used as outcome measures. Different types of biological mechanisms by which stress may play a role in susceptibility to colds have also been proposed. Stress may alter the biological susceptibility (predispose some individuals). Second, it may trigger a process that increases virus replication (an initiating effect). Finally, it may contribute to an ongoing pathogenic process that leads to maintenance of the disease. As well as biological mechanisms, it has been suggested that stress may be associated with colds because it increases the perception of sensations in the nose, leads to these sensations being labeled as symptoms, and biases reporting the presence of illness. In addition, stress may increase use of medical care.

Before reviewing the evidence on stress and susceptibility to the common cold, it is necessary to consider some conceptual and methodological issues. Many studies of stress and the common cold are largely looking at correlations, and in some cases disease causes stress rather than vice versa. This problem is reduced in prospective studies, especially viral challenge research. Second, factors other than stress influence susceptibility: exposure; prior exposure (immunity); age, gender, social status; nutritional status; pregnancy; and biological rhythms (circadian, menstrual, and seasonal). Some of these may be correlated with both stress and infection and could, therefore, provide alternative explanations for postulated associations

between stress and colds. Each factor may also make an independent contribution to susceptibility to colds, and the more of these variables that are controlled, the greater the chance of showing effects of stress in the context of multiple environmental, social, and biological predictors.

Early Research on Naturally Occurring Colds

Cohen and Williamson reviewed the literature on stress and infectious disease. Some early research, using retrospective reports of stress, showed that the frequency and/or severity of upper respiratory tract infections (URTIs) are related to negative life events, upsets and worries, and personal failures. Other research has failed to demonstrate such effects, and some studies suggest that stress is just one of many psychosocial factors that have to be considered. Prospective studies, with measures of stress at baseline, have also provided some support for stress influencing susceptibility to URTIs. Diary methodologies have also been used. For example, Stone et al. found that undesirable daily events increased 3 to 4 days prior to an URTI episode, and events rated as desirable decreased 4 to 5 days prior to onset. One problem with such studies is that it is often unclear whether the effects are specific to URTIs or whether the results show that stress influences general illness behaviors. Other studies have included objective verification of the illness (either clinical diagnosis or virological assays), and, again, some of the research using these methods shows that stress influences susceptibility but other studies have produced negative results.

Virus Challenge Studies

A major problem with studies of naturally occurring colds is that they do not control for exposure to the virus. This has been overcome by conducting viral challenge studies. Early research using this approach produced conflicting results, which probably reflected the small sample sizes, lack of control of important predictors of infection, inadequate conceptualization of stress, and confounding effects of medication. These problems were overcome in a study by Cohen et al., carried out at the MRC Common Cold Unit, Salisbury, UK. The results of this study showed that susceptibility to colds was related to stress in a dose-response manner. This effect was still apparent when analyses controlled for possible confounding factors. It also appeared to be a general effect that was not

virus specific. Use of measures of infection then allowed analyses to be conducted to determine whether stress influenced susceptibility to infection or reporting of symptoms following infection. The results showed that perceived stress and negative mood were associated with increased risk of infection but that negative life events increased the likelihood of reporting illness after infection. It has often been suggested that stress influences health by increasing negative health-related behaviors (e.g., smoking, alcohol consumption). These factors were controlled for in this study, and the effect of stress could not be attributed to these (although both smoking and alcohol did influence susceptibility to colds).

Duration and Type of Stressor

Further studies of stress and susceptibility to colds considering the impact of different types of stressor have been carried out in the United States. The first major finding was that it is chronic stress that is important in increasing susceptibility to colds. It has also been shown that work stress (e.g., job insecurity) and interpersonal stress lead to the greatest risk of getting a cold. Other research has examined other psychosocial characteristics that may underlie or modify the effects of stress on susceptibility to colds. Social integration was found to be an important factor, with high social support being associated with a reduced risk of infection. Further analyses addressed the question of whether the reported effects might be attributable to personality characteristics, not to stress. Personality (extraversion) was found to be an important risk factor, but this effect was independent of those of stress.

Underlying Mechanisms

The review of the above studies confirms that stress is an important factor in susceptibility to the common cold, and the analyses have eliminated many alternative explanations such as health practices and other psychological characteristics. Further analyses have been carried out to determine whether the effects of stress are attributable to endocrine responses. High levels of both epinephrine and norepinephrine were associated with an increased risk of colds, but these effects did not account for the stress-colds association. Immunological mechanisms were also investigated, and it was found that the effect of stress did not reflect natural killer cell activity or antibody production. It is, of course, possible that other aspects of the immune system underlie the stress-colds effect (e.g., cytokines), but these require further investigation.

Another possibility is that stress influences local changes in the nose. There have been a large number of studies of stress and secretory IgA (sIgA). This is the most abundant immunoglobulin; it provides a barrier to invasion, and there is evidence that there is an inverse relationship between levels of sIgA and vulnerability to infection. Chronic stress suppresses levels of sIgA, which may be part of the explanation for the association between stress and susceptibility to colds. An alternative speculation is that stress has a direct effect on the nasal mucosa, which may lead to viruses becoming established and replicating at a faster rate. Further research is needed to address the effects of stress on the anatomy and physiology of the nose.

Conclusions

In conclusion, there is now strong evidence that stress influences susceptibility to the common cold. This has been demonstrated in studies that have controlled exposure to the virus and other confounding factors. Chronic stress is required to demonstrate these effects, and research has shown that work stress and interpersonal stress pose the greatest risk. Many possible mechanisms have been put forward to account for the association between stress and colds, and a large number of these can be discounted as the sole underlying mechanism. It is possible that stress influences many mechanisms and that it is the combined effects of these that is important. Alternatively, the effect of stress on susceptibility to colds may reflect specific changes to the upper respiratory tract rather than more central mechanisms.

See also the Following Articles

Cytokines; Infection; Workplace Stress.

Further Reading

Belfer, M. L., Shader, R. I., Mascio, A. D., Harmatz, J. S. and Nahum, J. P. (1968). Stress and bronchitis. *British Medical Journal* 3, 805–806.

- Cohen, S. and Williamson, G. M. (1991). Stress and infectious disease in humans. *Psychological Bulletin* 109(1), 5–24.
- Cohen, S., Tyrrell, D. A. J. and Smith, A. P. (1991). Psychological stress and susceptibility to the common cold. *New England Journal of Medicine* 325, 606–612.
- Cohen, S., Tyrrell, D. A. J., Russell, M., Jarvis, M. J. and Smith, A. P. (1993). Smoking, alcohol consumption and susceptibility to the common cold. *American Journal of Public Health* 83, 1277–1283.
- Cohen, S., Tyrrell, D. A. J. and Smith, A. P. (1993). Negative life events, perceived stress, negative affect and susceptibility to the common cold. *Journal of Personality and Social Psychology* 64, 131–140.
- Cohen, S., Gwaltney, J. M., Doyle, W. J. and Newsom, J. T. (1995). State and trait negative affect as predictors of objective and subjective symptoms of respiratory viral infections. *Journal of Personality and Social Psychology* 68(1), 159–169.
- Cohen, S., Doyle, W. J., Skoner, D. P., Rabin, B. S. and Gwaltney, M. D. (1997). Social ties and susceptibility to the common cold. *Journal of the American Medical Association* 277(24), 1940–1944.
- Cohen, S., Frank, E., Doyle, W. J., Skoner, D. P., Rabin, B. S. and Gwaltney, J. M. (1998). Types of stressors that increase susceptibility to the common cold in healthy adults. *Health Psychology* 17(3), 214–223.
- Jacobs, M. A., Spilken, A. Z. and Norman, M. M. (1969). Relationship of life change, maladaptive aggression, and upper respiratory infection in male college students. *Psychosomatic Medicine* 31, 31.
- Jemmott III, J. B. and Locke, S. E. (1984). Psychological factors, immunologic mediation, and human susceptibility to infectious diseases: how much do we know? *Psychological Bulletin* 95, 78–108.
- McClelland, D. C., Alexander, C. and Marks, E. (1982). The need for power, stress, immune function, and illness among male prisoners. *Journal of Abnormal Psychology* 9, 61–70.
- Sarason, I. G., Sarason, B. R., Potter, E. H. and Antoni, M. H. (1985). Life events, social support, and illness. *Psychosomatic Medicine* 47, 156–163.
- Stone, A. A., Reed, B. R. and Neale, J. M. (1987). Changes in daily event frequency precede episodes of physical symptoms. *Journal of Human Stress* 13, 70–74.

Community Studies

C J Holahan

University of Texas at Austin, Austin, TX, USA

R H Moos

Dept. of Veterans Affairs Health Care System and
Stanford University Medical Center, Palo Alto, CA, USA

L M Groesz

University of Texas at Austin, Austin, TX, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
C J Holahan, R H Moos, and J D Ragan, volume 1, pp 501–506,
© 2000, Elsevier Inc.

Conceptualizing Life Stressors

Life Stressors and Adaptation

Stress Moderators

Integrative Models

Glossary

<i>Chronic stressors</i>	Stressors that endure over time, such as ongoing problems in a marriage and long-term difficulties with a troublesome supervisor at work.
<i>Coping strategies</i>	Cognitive and behavioral efforts to reduce or adapt to stressful conditions and associated emotional distress; most approaches distinguish between strategies oriented toward approaching and avoiding the problem.
<i>Life change events</i>	Stressors that occur at a discrete point in time, such as the death of a family member or losing one's job.
<i>Life crises and traumas</i>	Qualitatively more severe stressors, such as natural disasters or being the victim of a violent crime, such as rape.
<i>Personal resources</i>	Relatively stable personality and cognitive characteristics that operate as stress moderators, such as dispositional optimism and self-efficacy.
<i>Social resources</i>	Social factors that operate as stress moderators, such as emotional support and assistance from family members and friends.
<i>Stress moderators</i>	Personal and social resources and coping strategies that function either as protective or vulnerability factors when stressors occur.
<i>Stress resistance</i>	Individuals' adaptive capacity for resilience and constructive action in the face of challenge, emphasizing protective factors that can moderate the effects of stressors.

Conceptualizing Life Stressors

Community studies of life stressors investigate (1) life change events, (2) chronic stressors, and (3) major life crises and traumas (see [Table 1](#)). It should be noted that, although life change events and traumas have an acute onset, they often initiate chronic stressors, such as persistent economic and family problems. In fact, these enduring secondary stressors may account for many of the adverse effects of acute stressors.

Life Change Events

Initially, community studies of life stressors focused on acute life change events, which occur at a discrete point in time, such as the death of a close friend or losing one's job. Life change events generally are indexed by self-report questionnaires that may use objectively weighted life change units to reflect the relative level of adaptive demand associated with a particular life event. These weighted life change units are summed to reflect cumulative life change during a given period of time, typically the previous 12 months. Investigators also use in-depth interviews that, although time consuming, are able to assess the personal meaning of life changes. Although initial studies assumed that all life change, whether positive or negative, involved adaptive demands that were detrimental, later research revealed that the adverse health effects of life change are specific to negative life events. Later studies also showed that weighted life change units contributed little beyond an unweighted sum of events experienced in predicting illness. Based on these findings, most contemporary research on life change

Table 1 Major categories of life stressors with representative stressors^a

<i>Life change events</i>	<i>Chronic stressors</i>	<i>Major crises and traumas</i>
Divorce	Family role strain	Victim of natural disaster
Death of a close friend	Work role strains	Life-threatening illness
Legal problem	Ongoing caregiving demands	Personal war exposure
Fired at work	Ongoing financial problems	Victim of violent crime

^aAdapted from "Stress," Holahan, Ragan, & Moos, 2004, in Spielberger (Ed.), *Encyclopedia of Applied Psychology*, San Diego, CA: Academic Press. Copyright 2004 by Elsevier Science (USA).

events uses the sum of negative events experienced during a given period of time.

Chronic Stressors

Early community studies of life stressors were limited by their narrow focus on acute stressors. Eventually, recognizing that many life stressors are ongoing, investigators began to examine chronic stressors, which endure over time, such as social role strains and everyday hassles. Researchers studying social role strains often measure chronic family or work difficulties. Moreover, there has been a growing research interest in role strains associated with the interface between work and family life. Some stress studies have focused on additional burdens that often are associated with family life, such as caregiving for a chronically ill family member. For example, increasing interest is being devoted to caregivers of older family members suffering from Alzheimer's disease. Investigators have also studied minor life stressors that are experienced chronically, such as everyday hassles with troublesome neighbors or traffic problems. Increasingly, studies in this area also examine the health consequences of ongoing macrostressors, such as discrimination, national or regional economic recessions, and the ongoing threat of terrorism. As with acute stressors, the unweighted sum of these types of ongoing stressors provides an index of cumulative chronic stressors during a given period of time, typically the previous 12 months.

Major Life Crises and Traumas

In addition to indexes of cumulative stressors, community studies also focus on specific major life crises and traumas, which are qualitatively more severe than life events or role strains. For example, researchers have studied natural disasters, such as Hurricane Andrew in Florida; technological disasters, such as the Three Mile Island nuclear accident in Pennsylvania; and acts of terrorism, such as the destruction of the World Trade Center in New York. Current community studies also focus on how people cope with serious medical illnesses, such as heart disease, cancer, and diabetes. Interest in traumatic stressors is evidenced by a growing body of research on posttraumatic stress disorder (PTSD), a debilitating anxiety-based disorder caused by personal experience of a life-threatening stressor. Initially, PTSD research focused on stressors such as exposure to war as a civilian or soldier or being the victim of a violent crime such as rape. Over time, the definition of trauma has been broadened to encompass witnessing events that

place others at risk of death or serious injury or witnessing individuals being badly injured or killed by such events.

Life Stressors and Adaptation

Initially community studies focused exclusively on the adverse health effects of life stressors. Eventually, however, the research focus broadened to encompass adaptive processes of stress resistance and stress-related growth.

Stressors and Illness

Findings with diverse community samples have shown that negative life change events are associated with psychological stress reactions that involve depression and anxiety. Interpersonal problems and losses are especially likely to be associated with depressive reactions and threatening events with anxiety. Even though chronic stressors may be less severe than acute life events, their effects may last longer and be more pervasive. Thus, chronic stressors generally are more strongly linked to psychological distress than are acute life change events. More severe trauma exposure produces a recognized pattern of PTSD symptoms, including psychologically reexperiencing the trauma through flashbacks and nightmares, emotional numbing, and heightened arousal.

Life stressors may elicit physiological stress reactions (e.g., circulatory, digestive, and musculoskeletal problems) that are mediated through the activation of the sympathetic nervous system. In addition, life stressors are associated with susceptibility to less serious infectious diseases (e.g., colds, influenza, and herpes) and possibly to the onset and exacerbation of autoimmune diseases (e.g., rheumatoid arthritis). Underlying these latter findings is an apparent correlation between life stressors and both cellular (e.g., reduced lymphocyte proliferation) and humoral (e.g., decreased antibody production) indices of immune suppression.

Stress Resistance

Despite the consistency of these stressor effects, they typically account for only a small portion of the variance in illness. Moreover, individuals showed highly variable reactions to stressors; many people remain healthy despite being exposed to stressful circumstances. These findings engendered a new approach to conceptualizing the stress and coping process, described as stress resistance. From an initial emphasis on people's deficits and vulnerabilities, contemporary

community studies of life stressors have evolved to place increasing emphasis on protective factors that can counteract the potential adverse effects of life stressors.

A guiding assumption of stress resistance research is that personal and social resources and adaptive coping strategies can help individuals to manage stressful circumstances effectively and to remain healthy when stressors occur. Stress resistance research fundamentally changed the underlying assumptions that guided early investigators' conceptualization of the stress process. From initially placing an emphasis on people's deficits and vulnerabilities, stress research progressed to placing increasing emphasis on individuals' adaptive strengths and capacity for resilience and constructive action in the face of challenge.

Stress-Related Growth

Stress researchers have also come to recognize that life stressors can be constructive confrontations that challenge an individual and provide an opportunity for personal growth. Individuals may emerge from a stressful experience or life crisis with greater self-confidence, new coping skills, closer relationships with family and friends, and a richer appreciation of life. Stress-related growth can occur even when stressors involve profound adaptive demands, such as serious financial setbacks and the deaths of friends or family members. Of course, beneficial outcomes to life crises may emerge only after an initial stage of emotional distress and disorganization. Yet many people are able to grow stronger in the face of adversity.

Approximately half of women who experience a marital breakup show long-term improvements in psychological functioning. These women may become more assertive, develop more realistic views of themselves, and experience increased self-esteem with successful new careers. Similarly, survivors of serious illness often show more concern for and sense of community with others, a change in focus from the constant pressure of work to family relationships, and heightened awareness of religious and humanitarian values. Even in the context of stressors as profound as being a prisoner of war, a diagnosis of terminal illness, and the death of a close family member, some individuals are able to grow stronger, richer, and more caring.

Stress Moderators

Contemporary interest in adaptive outcomes in the stress process has encouraged efforts to identify factors that moderate the health effects of life stressors. Investigators have focused on the stress-moderating

roles of personal and social resources, coping strategies, and personal agendas.

Personal and Social Resources

A variety of relatively stable personality characteristics operate as protective factors that can help individuals remain mentally and physically healthy when stressors occur. Personality characteristics that relate broadly to personal control, such as dispositional optimism and self-efficacy, are especially important. Individuals who are more optimistic about the eventual outcome of a stressful experience and those who are more confident in their ability to manage adaptive demands manage stressful episodes more successfully than individuals who lack these personal strengths. In contrast, negative personality characteristics, such as neuroticism and low self-esteem, can operate as vulnerability factors that increase health risk under high stressors.

Social resources also play a key role in protecting individuals from the potential adverse effects of life stressors. Social resources include guidance and assistance from one's family members, friends, and broader social network. Emotional support, such as the nurturance of a caring family member or attentive listening of a trusting confidant, is especially important as a protective factor during stressful times. Of course, social ties are complex and may entail criticism or competition as well as support. In fact, studies that consider both the positive and negative aspects of relationships show that negative aspects of relationships are at least as strongly related to poorer adjustment as positive aspects of relationships are related to better adjustment.

Box 1 The *Life Stressors and Social Resources Inventory* (LISRES) provides a comprehensive picture of the interrelated stressors and social resources in a person's life. The inventory includes nine indices of life stressors and seven indices of social resources. **Figure 1** shows an illustrative LISRES profile for a 66-year-old woman with rheumatoid arthritis (scores are standardized, with a mean of 50 and a standard deviation of 10). Although she reported well above-average physical health stressors, she had below-average stressors in six of the other seven domains. Moreover, she also had above-average social resources in her work setting and in her relationships with her spouse, extended family, and friends. At an 18-month follow-up, this favorable balance of moderate stressors and high resources enabled her to manage quite well; she reported high self-confidence and below-average depression.

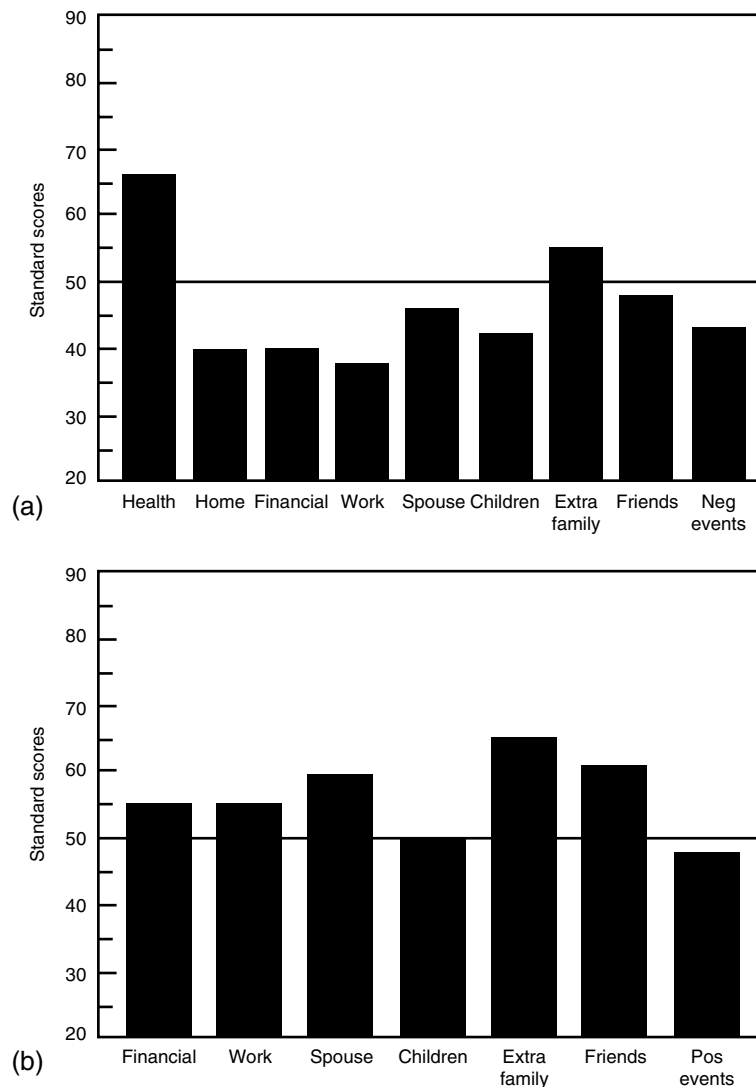


Figure 1 (a) Life stressors and (b) social resources profile for a 66-year-old woman with rheumatoid arthritis. Reproduced by special permission of the Publisher, Psychological Assessment Resources, Inc., 16204 N. Florida Avenue, Lutz, Florida 33549, from the LISRES-A by Rudolf H. Moos, Ph.D., Copyright 1994 by PAR, Inc. Further reproduction is prohibited without permission of PAR, Inc.

Coping Strategies

A large number of community studies has examined the role of coping strategies in stress resistance. Coping encompasses cognitive and behavioral efforts to reduce or adapt to stressful conditions and associated emotional distress. Most conceptualizations distinguish between coping strategies that are oriented toward approaching and confronting the problem and strategies that are oriented toward reducing tension by avoiding dealing directly with the problem (see Table 2). People who rely more on approach coping strategies, such as problem solving and information seeking, tend to adapt better to life stressors. In contrast, avoidance coping strategies, such as denial and wishful thinking, generally are associated with more psychological distress.

Although we can draw overall conclusions about the relative efficacy of approach and avoidance coping strategies, such generalizations oversimplify the process of adaptation. Adaptation is best when coping efforts match situational demands. Individuals who are flexible in their choice of coping should show better adaptation than people who have a more restricted or rigid coping repertoire. Avoidance coping may be adaptive in the short term with time-limited stressors, such as pain, blood donation, and uncomfortable medical diagnostic procedures. However, avoidance generally is maladaptive as a long-term coping strategy. The manageability of a stressor also shapes coping effectiveness; approach coping processes are most effective in situations that are controllable.

Table 2 Four basic categories of coping strategies along with eight associated coping subtypes^a

<i>Basic coping categories</i>	<i>Coping subtypes</i>
Cognitive approach	Logical analysis ("Did you think of different ways to deal with the problem?") Positive reappraisal ("Did you think about how you were much better off than other people with similar problems?")
Behavioral approach	Seeking guidance and support ("Did you talk with a friend about the problem?") Taking problem-solving action ("Did you make a plan of action and follow it?")
Cognitive avoidance	Cognitive avoidance ("Did you try to forget the whole thing?") Resigned acceptance ("Did you lose hope that things would ever be the same?")
Behavioral avoidance	Seeking alternative rewards ("Did you get involved in new activities?") Emotional discharge ("Did you yell or shout to let off steam?")

^aSample coping items are shown in parentheses. Adapted and reproduced by special permission of the Publisher, Psychological Assessment Resources, Inc., 16204 North Florida Avenue, Lutz, Florida 33549, from the Coping Responses Inventory by Rudolf Moos, Ph.D., Copyright 1993 by PAR, Inc. Further reproduction is prohibited without permission from PAR, Inc.

Personal Agendas

An individual's personal agendas in the domain in which a stressor occurs shape the meaning and impact of the stressor. People function in multiple environments that often reflect different personal needs and commitments. Individuals appraise life stressors in the light of these personal agendas, and such appraisals can either enhance or reduce the stressfulness of a situation. For example, because chronic stressors provide a context within which new acute events are interpreted, they may alter the threshold at which events have negative impacts on health. Job loss during a period of bereavement may adversely affect health in part because it amplifies a more pervasive sense of personal loss.

Similarly, when an event threatens or disrupts a domain in which a person has central commitments, that event is more likely to have an adverse outcome. For example, women who experience a new severe event in a life domain in which they have marked difficulties, or in a domain that matches an area of marked personal commitment, are more likely to develop depression than are women who experience a new severe event in some other domain. At the same time, strong commitments can sustain coping efforts in the face of profound obstacles, such as coping with bereavement or a life-threatening illness.

Integrative Models

Most recently, community studies have been examining integrative frameworks reflecting reciprocal links among stressors, resources, coping, and adaptation. Integrative models have addressed how maladaptive behavior generates life stressors, the role of stressors in eroding protective resources, and how protective factors operate together in a mutually supportive way.

Stress Generation

In a revision of traditional conceptualizations of the stress process, the stress generation model describes how depressed individuals, through their depression and related behaviors, generate life stressors, particularly interpersonal stressors. These stressors, in turn, are linked to increases in subsequent depressive symptoms. The stress generation model has been applied across diverse age groups to predict depressive outcomes in both clinical and nonclinical samples. An association between depression and subsequent stressors is evident for both acute events and chronic strains. In a similar way, avoidance coping is prospectively linked to both acute and chronic stressors, independent of depression. Cognitive avoidance may permit incipient stressors, such as financial problems, to worsen. Behavioral avoidance may actively promote new stressors, such as when emotional discharge aggravates strains in family relationships.

Resource Depletion

Another set of community studies has examined a support deterioration model that describes how life stressors can deplete psychosocial resources. Research on natural disasters, combat, and caregiving has shown that stressors can deplete both personal and social resources. Resource depletion may be tangible and intrinsically tied to a stressor; loss of income or loss of a loved one entail real reductions in resources. However, resource depletion also has a cognitive component; perceived support appears to be especially vulnerable to deterioration after crises. Support diminishes after the immediate crisis period even though crisis victims face enormous ongoing demands. Moreover, in community-level disasters, individuals who could potentially provide support are often crisis victims themselves. Resource loss may help to explain the association between life stressors and a decline in psychological functioning. For example, disaster-related stressors deplete perceived support; in turn, this loss in perceived support mediates the connection between initial stressors and later depression.

Coping Mediation

Resources models of coping examine how personal and social resources relate to better functioning under high stressors by fostering adaptive coping efforts. At one level, these models identify a key mechanism through which personal and social resources operate; at another level, they demonstrate that adaptive coping requires a strong underlying psychosocial foundation. Among community adults facing high stressors, adaptive personality characteristics and family support operate prospectively as coping resources; in turn, coping mediates between initial resources and later health status. Similarly, among women coping with breast cancer and students adjusting to law school, more adaptive coping efforts operate as a mediator through which optimism relates to better psychological adjustment. People with higher levels of self-efficacy, for instance, tend to approach challenging situations in an active and persistent style, whereas those with lower levels of self-efficacy are less active or tend to avoid such situations. Social resources can enhance coping efforts by bolstering feelings of self-confidence, as well as by providing informational guidance that aids in assessing threat and in planning coping strategies.

Acknowledgments

Preparation of this manuscript was supported in part by NIAAA grant AA12718 and by the Department of Veterans Affairs Health Services Research and Development Service. Opinions expressed herein are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs.

See Also the Following Articles

Disasters and Mass Violence, Public, Effects of; Life Events Scale.

Further Reading

Dohrenwend, B. P. (ed.) (1998). *Adversity, stress, and psychopathology*. New York: Oxford University Press.

- Folkman, S. and Moskowitz, J. T. (2004). Coping: pitfalls and promise. *Annual Review of Psychology* 55, 745–774.
- Hammen, C. (1999). The emergence of an interpersonal approach to depression. In: Joiner, T. & Coyne, J. (eds.) *The interactional nature of depression: advances in interpersonal approaches*, pp. 21–35. Washington, DC: American Psychological Association.
- Kaniasty, K. and Norris, F. H. (2001). Social support dynamics in adjustment to disasters. In: Sarason, B. R. & Duck, S. (eds.) *Personal relationships: implications for clinical and community psychology*, pp. 201–224. New York: John Wiley.
- Kemeny, M. E. (2003). The psychobiology of stress. *Current Directions in Psychological Science* 12, 124–129.
- Kessler, R. C. (1997). The effects of stressful life events on depression. *Annual Review of Psychology* 48, 191–214.
- McNally, R. J. (2003). Progress and controversy in the study of posttraumatic stress disorder. *Annual Review of Psychology* 54, 229–252.
- Moos, R. H. and Moos, B. S. (1994). *The life stressors and social resources inventory*. Odessa, FL: Psychological Assessment Resources.
- Moos, R. H. (2002). The mystery of human context and coping: an unraveling of cues. *American Journal of Community Psychology* 30, 67–88.
- Segerstrom, S. C. and Miller, G. E. (2004). Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry. *Psychological Bulletin* 130, 601–630.
- Snyder, C. R. (ed.) (1999). *Coping: the psychology of what works*. New York: Oxford University Press.
- Taylor, R. D. and Wang, M. C. (eds.) (2000). *Resilience across contexts: family, work, culture, and community*. Mahwah, NJ: Lawrence Erlbaum.
- Taylor, S. E., Repetti, R. L. and Seeman, T. (1997). Health psychology: what is an unhealthy environment and how does it get under the skin? *Annual Review of Psychology* 48, 411–447.
- Tedeschi, R. G., Park, C. L. and Calhoun, L. G. (eds.) (1998). *Posttraumatic growth: positive changes in the aftermath of crises*. Mahwah, NJ: Lawrence Erlbaum.
- Yehuda, R. and McEwen, B. S. (2004). Protective and damaging effects of the biobehavioral stress response: cognitive, systemic and clinical aspects: ISPNE XXXIV meeting summary. *Psychoneuroendocrinology* 29, 1212–1222.

Comorbid Disorders and Stress

S Sundram

Mental Health Research Institute of Victoria,
Northern Psychiatry Research Centre, and
University of Melbourne, Melbourne, Australia

Avril Pereira

Mental Health Research Institute of Victoria,
Melbourne, Australia

© 2007 Elsevier Inc. All rights reserved.

Introduction

Prevalence of Comorbidity

Substances

Etiological Theories

Comorbidity and Stress

Major Depression, Substance Use Disorders, and Stress

Conclusion

Glossary

<i>Comorbidity</i>	The concurrence within the one individual of a substance use disorder and another psychiatric disorder.
<i>Dopamine</i>	A catecholamine neurotransmitter central to reward processes and psychiatric disorders.
<i>Mesolimbic dopamine pathway</i>	The trajectory within the brain of dopamine-releasing neurons from the ventral tegmental area to the nucleus accumbens, amygdala, hippocampus, and other related regions.
<i>Nucleus accumbens</i>	A brain region rich in dopamine neurons and believed to integrate reward responses and to be affected in substance use disorders.
<i>Substance-use disorder (SUD)</i>	The repeated uncontrolled use of a substance resulting in deleterious health and social outcomes; there are two forms, dependence and the less severe abuse.

Introduction

Comorbidity simply refers to the concurrent presence of two or more discrete disorders in the one individual. In psychiatry, it refers to at least one disorder being a recognizable *Diagnostic and Statistical Manual for Mental Disorders* (4th edn.; DSM-IV) or *International Classification of Diseases* (10th edn.; ICD-10) diagnosis and the second or subsequent disorders being either a physical medical illness or another psychiatric disorder. Commonly, however, the second disorder is substance-related, and it is in

this context that comorbidity (also referred to as dual diagnosis) is discussed here.

Substance-related disorder refers to the consumption of a substance by an individual resulting in deleterious health, social, or psychological outcomes. This definition allows the exclusion of substance use within culturally determined or religious domains and the recognition that some substance use may not necessarily be harmful depending on the substance and volume of use. Nevertheless, this definition encompasses a broad and heterogeneous group of disorders that can be applied to any particular substance. These disorders can be classified into disorders induced by the substance and disorders of the regulation of substance use. The first category includes intoxication, withdrawal, and those disorders that mimic other psychiatric disorders: deliriums of intoxication and withdrawal; dementia; amnestic, psychotic, mood, anxiety, and sleep disorders; and sexual dysfunction. The second category incorporates dependence and abuse. Dependence, which is considered more severe than abuse, comprises three or more of the following: tolerance (needing an increased amount of the substance for the equivalent effect or a diminished effect with same amount), withdrawal (physiological response to the reduction of substance use), prolonged or excessive ingestion, unfulfilled desire to reduce use, excessive time spent on substance use to the detriment of other activities, and use in spite of deleterious health effects. Abuse emphasizes the persistent use of the substance in the face of work, social, or forensic problems when the criteria for dependence are not met.

Commonly used substances include alcohol, caffeine, nicotine, cannabis, amphetamines, opioids, cocaine, hallucinogens, phencyclidine and ketamine, inhalants, and benzodiazepines; most commonly the use of more than one substance (polysubstance use) is observed. The pattern of substance use generally follows social criteria so that legal substances, such as alcohol, nicotine, and caffeine are most widely used, followed by decriminalized substances, and then by the most readily available illicit substances. For example, alcohol-use disorders are relatively less common in religious and social groups that prohibit its use.

When we combine these lists of disorders and substances, a large number of permutations for substance-related disorders is possible. All substances must induce intoxication; however, not every substance causes every other disorder. This is dependent on its individual pharmacodynamic and pharmacokinetic properties and, thus, it is critical to emphasize the differences in

the individual behavioral and neurobiological effects of each substance and their interaction with other psychiatric disorders. For example the sedating and myorelaxant properties of benzodiazepines contrast with the stimulating effects of amphetamine.

Given that comorbidity refers to coexisting but discrete disorders, the relationship between the disordered regulation of substance use and other psychiatric disorders and stress is of primary relevance here; SUD is used here to refer to disorders of substance-use regulation, namely dependence and abuse. This article outlines the observed relationships between SUD and other psychiatric disorders. It discusses the possible nature of these relationships with reference to the role of stress and details this in reference to major depression because of the well-documented effect of stress on relapses of both SUD and major depression.

Prevalence of Comorbidity

The lifetime prevalence of SUD in the general community has been estimated to be 16.7% (excluding nicotine and caffeine); the 12-month prevalence figures for any mental disorder is 26.2%. Thus, even if no relationship existed between the two clusters, comorbidity would be predicted to be 4% by chance alone. The reality is, however, that approximately 76% of men and 65% of women with SUD also have an additional psychiatric diagnosis, most commonly that of another SUD. Other, non-SUD psychiatric diagnoses also occur at rates greater than in the general community, and the reciprocal relationship is also true, suggesting that the two clusters of disorders are not independent. This high rate of comorbidity needs to be considered within the context of the various methods of ascertainment used (retrospective or prospective, structured interview, self-report, or caregiver report), the diagnostic systems used (DSM-III, DSM-III-R, DSM-IV, or ICD-10), and the sample populations (urban or rural; and psychiatric inpatient, outpatient, general community, drug treatment program, or correctional facility). This limits the extent to which studies may be compared across regions or over time. Given the diversity of substances and co-occurring disorders, the prevalence of comorbidity inevitably differs by diagnostic classification. Notwithstanding this, the incidence of comorbidity appears to have increased over the past 2 decades. In 1980 the United States National Center for Health Statistics reported that 12% of psychiatric inpatients treated that year had a SUD in addition to a non-substance-use psychiatric disorder. A decade later, the Epidemiological Catchment Area (ECA) study recorded that 48% of individuals with schizophrenia had a lifetime diagnosis of SUD. Subsequent estimates

of the prevalence of such comorbidity range from 35 to 53% in community and inpatient settings, respectively. It is noteworthy that epidemiological surveys that estimate the lifetime and 12-month prevalence of psychiatric disorders in the general population such as the 1990–1992 National Comorbidity Survey (NCS) and the 2001–2003 NCS-2 and NCS-Replication have enabled the analysis of comorbidity rates unbiased by sample selections of affected individuals. These surveys, however, are also limited in that they underrepresent several population segments such as the homeless or those in institutions. The examination of the so-called low-prevalence psychotic disorders is also excluded. It is unsurprising, therefore, that prevalence estimates of comorbidity from community population surveys are generally lower than seen in clinical settings.

Substances

Unfortunately, unraveling comorbidity is complicated by polysubstance use, particularly of alcohol with other drugs. According to the NCS, 78% of alcohol-dependent men and 86% of women met the criteria for a lifetime diagnosis of another psychiatric disorder, including drug dependence. Furthermore, although demographic characteristics such as male gender, younger age, less education, history of incarceration, and symptoms of conduct disorder and antisocial personality disorder predict substance misuse, these are not consistently associated with the types of substances misused. For example, age is related to lifetime prevalence of cannabis and cocaine but not alcohol-use disorder. In addition, attempts to find demographic or clinical correlates related to the order of onset of substance use and psychiatric disorders have not detected consistent differences to demonstrate a causal relationship. A key point to note is that no psychiatric diagnosis is exclusively related to a specific type of SUD.

Alcohol

Consistent across several studies, alcohol is the most commonly abused substance (when nicotine and caffeine are not measured), followed by cannabis. Among individuals with psychotic illnesses, approximately 39% of men and 17% of women had a lifetime diagnosis of alcohol abuse or dependence compared with a 12-month prevalence in the general population of 9% and 4%, respectively (National Survey of Mental Health and Well-being). These data concur with the ECA study, which found that 47% of those with schizophrenia spectrum disorders met the criteria for at least one comorbid SUD (34% for alcohol, 28% for any

drug use disorder). A high rate of alcohol comorbidity with affective illness is also described with adverse outcomes of mood destabilization and induction of depressive or manic episodes. In their 1990 review of studies on hospitalized alcoholics, Merikangas and Gelernter found that depressive symptomatology ranged from 16 to 59% and that this met a diagnosis of secondary major depression in 8–53% of cases. Although there is a consensus that a high comorbidity of alcohol abuse/dependence and depression occurs in clinical samples, there is far less agreement concerning the nature or direction (primary or secondary) of this relationship. This distinction is necessary in determining whether alcohol misuse and depression are different symptomatic expressions of a common dysfunctional neurobiological substrate or whether one condition induces the other. In contrast, alcohol abuse is also frequently comorbid with bipolar disorder and posttraumatic stress disorder, but in these cases it is usually temporally secondary.

Nicotine

In the case of nicotine, investigation of a subset of NCS participants found that psychiatric disorders predicted an increased risk for the first onset of daily smoking and subsequent nicotine dependence. Most notable are people with schizophrenia, who have smoking rates of between 70 and 90% (compared to approximately 25% for the general population), resulting in excess morbidity and mortality in this group. As well, nicotine dependence is twice as prevalent in patients with mood, anxiety, or personality disorders than in the general population. The reasons for this increased rate of use include self-medication to ameliorate the symptoms of the disorder or the side-effects of treatment.

Conversely, it has also been proposed that daily smoking is a causal factor in panic disorder and agoraphobia, conditions that for some individuals may be relieved by smoking cessation. Of interest also are data on twins centered on the familial aggregation of smoking and depression, indicating common or shared genes as a source of this association. Defining the interaction between genes and environment in such comorbidity is important in understanding why increased risk for substance use occurs.

Illicit Drugs

The risk relationship between illicit substance use and other psychiatric disorder appears to be reciprocal, with one predicting the increased likelihood of the other. Moreover, approximately one in five people report the use of more than one illicit drug. The substances most frequently misused include cannabis,

amphetamines, opioids (heroin), cocaine, hallucinogens, ecstasy or MDMA (N-methyl-3,4-methylenedioxy-methamphetamine), and benzodiazepines (nonprescribed use). As may be anticipated, individuals with multiple dependencies experience the highest rates of psychiatric comorbidity. Most psychiatric illnesses including psychotic, mood, anxiety, or alcohol-use disorders are often comorbid with illicit SUDs. In such psychiatric cohorts, gender differences in drug use are commonly substance-type-specific. For example, men are twice more likely than women to meet the criteria for cannabis abuse or dependence; however, in cocaine and opioid abusers, the numbers are generally similar between the sexes. It is also prudent to highlight that differences in the availability of illicit drugs in the study location may explain the data gathered. Thus, whereas cocaine is the illicit drug most commonly used by incarcerated women in the United States, cannabis, opioids, and benzodiazepines were most frequently used in an Australian cohort.

Etiological Theories

Each class of abused drugs is chemically diverse and produces effects through actions at different specific cellular and molecular targets within the brain; nevertheless, the pathways underlying reward and dependence are thought to be similar. The common mode of action involves all drugs of abuse releasing dopamine, either directly or indirectly, within the shell of the nucleus accumbens. This dopamine is released from neurons that originate in the ventral tegmental area (VTA) and terminate in the nucleus accumbens, amygdala, the bed nucleus of the stria terminalis, and other regions and are termed the mesolimbic dopamine projection. This is thought to be the path for all natural rewards such as food and sex and also for other nondrug rewards such as gambling. However, repeated drug use that results in drug dependence is more complex than simple changes in extracellular dopamine or receptor sensitivity. Drug dependence comprises a number of aspects including the need for the drug, the motivation to use it, the reward or pleasurable aspect of the use, and the relief from withdrawal symptoms. The brain regions and neurochemical pathways underpinning these distinct behavioral processes differ and are not fully elucidated, although they are considered to be similar between drugs of abuse. How these pathways are involved with the manifestation of other psychiatric disorders and how the pathologies of both sets of disorders may interact and be modulated by stress are poorly understood. Hence, although there is convincing epidemiological evidence linking psychiatric disorder

and SUDs and stress, direct biological evidence explaining this association is limited. Several competing theories have been proposed to explain this high co-occurrence.

Common Factor Models

Common factor models posit that comorbidity is the result of shared vulnerability to both disorders. Specific factors can independently increase the risk of developing both disorders and thus the increased comorbidity is explained. For example, common genetic predisposition may underlie both the index and the comorbid condition. Other possible common factors include social class of rearing, family disruption, cognitive functioning, and exposure to an environmental toxin (e.g., lead).

Genetics There is clear evidence that genetic factors play a role in both most major psychiatric disorders, including schizophrenia, bipolar disorder, and major depression, and in SUDs. The question is whether genetics contributes to comorbidity. Studies have shown that comorbid patients are more likely to have relatives with a SUD than patients without a dual diagnosis; however, this does not imply linked genetic factors. An alternative question is whether genetic vulnerability to one disorder increases the risk of exposure to the other disorder. Data addressing this point indicate that the genetic risk of schizophrenia is not associated with an increased risk of SUD in relatives, or vice versa. Such findings suggest that shared genetic factors do not explain the increased rates of comorbid substance use in schizophrenia. This conclusion is also supported by twin studies on the heritability of schizophrenia and alcoholism, in which the rates of the disorders between monozygotic and dizygotic twins were compared and no correlation was found. A 2003 large multivariate modeling study using same-sex twins demonstrated that common genetic factors for SUDs loaded with adult antisocial behavior and conduct disorder but not with internalizing disorders such as major depression and anxiety disorders. These data contribute to dispelling notions of a common shared genetic vulnerability but not of disorder-specific interactions. Consequently, although there may be some genetic contribution to comorbidity, the problem remains of identifying which factors are inherited and how they interact with other personal traits and facets of the environment to elevate risk.

Secondary Substance Disorder Models

Several models postulate that psychiatric illness increases the vulnerability of a patient to developing a SUD. These models are broadly characterized into

two types: the self-medication model and the supersensitivity model.

Self-Medication Model Self-medication denotes that individuals use specific substances to attenuate adverse states caused either by their psychiatric disorder or by its treatment. The underlying assumption is that substances are used in a selective manner based on their distinct pharmacological effects. Examples of symptom alleviation include the use of nicotine by people with schizophrenia to normalize deficits in sensory processing (attention) or the use of stimulant drugs (amphetamines and cocaine) by depressed patients to counter anergic symptoms (lack of energy and amotivation). An example of substance use mitigating the side-effects of psychotropic medications is nicotine relieving the extrapyramidal motor and dysphoric side-effects induced by antipsychotic medication; it is proposed that it releases dopamine to ameliorate the dopamine D2 receptor blockade induced by the medication. As intuitively appealing as the self-medication hypothesis is, the evidence is not supportive. First, much substance use, especially cigarette smoking, precedes the first symptoms of the psychiatric disorder, let alone the first treatment. Second, the factors most consistently associated with substance use in psychiatric patients, such as availability, cost, compliance with the peer group, facilitation of social interaction, intoxication, and relaxation, are the same as that observed in the general community. Similarly, patients seldom report self-medication as a reason for substance use. Finally, the choice of substance does not appear to be related to the psychiatric diagnosis or type or severity of symptoms.

The Supersensitivity Model This model contends that a combination of genetic and early environmental (pre- or postnatal) events interact with an environmental stressor to precipitate the onset or relapse of a psychiatric disorder in a vulnerable individual, where vulnerability is defined as a compromised biological sensitivity to stress. Therefore, whereas psychotropic medications can be protective or decrease vulnerability, substance abuse may increase it. This sensitivity may then make psychiatric patients more likely to experience negative consequences from using even small amounts of the substance. The assertion is that these negative outcomes of substance use, rather than the use itself, is what differentiates comorbid patients from the general population. Hence, given a similar volume of use, a greater proportion of people with psychiatric disorders will present with problematic use. Various observations provide support for this model, including the lesser amounts of substance used by comorbid patients, induction of clinical

symptoms by low-dose drug (amphetamine) challenge tests, and negative effects such as relapse following the administration of small quantities of the substance. Similarly, a longitudinal examination of the course of drinking alcohol in a psychiatric cohort found that fewer than 5% were able to sustain symptom-free drinking over time without negative consequences, compared to approximately 50% within the general population.

Secondary Psychiatric Disorder Model

This model posits that the SUD induces the psychiatric disorder, based on observations of the psychotogenic effects of drugs such as stimulants (amphetamines and cocaine), hallucinogens (lysergic acid diethylamide), and cannabis and the depressive effects of alcohol. However, the complex interaction of neurobiological and psychosocial factors that underlie each disorder often makes the assignment of causality in this direction difficult.

Comorbidity between schizophrenia and cannabis use illustrates this point. Cannabis may induce a transient psychotic state characterized by hallucinations and agitation that may be difficult to differentiate from a primary psychotic disorder. Further, cannabis use commonly precedes the onset of schizophrenia. Trying to determine the issue of causality, however, is hard given that the available evidence is largely based on epidemiological studies. A 2002 meta-analysis concluded that cannabis use accounted for approximately 8% of the attributable risk for developing schizophrenia. This may account for the anecdotal clinical observations of cannabis causing schizophrenia and the conflicting prevalence data that frequently show no relation between the two disorders.

With respect to alcohol use and depression, the notion is that alcohol increases the risk and severity of depression, although no causal association has been established. This is complicated by the general observation that alcohol abuse and dependence develops after the onset of bipolar disorder, again emphasizing the need to carefully differentiate particular disorders and their comorbidity.

Bidirectional Models

Bidirectional models suggest interactional effects between psychiatric and SUDs that account for the high co-occurrence rates. For example, a SUD could trigger a psychiatric illness in a vulnerable individual, which is maintained by continued substance use due to cognitive factors, social conditioning, beliefs, and motivations. Despite evidence that substance misuse worsens the course of psychiatric illness, these postulates, although highly plausible, remain largely theoretical and untested.

Comorbidity and Stress

An extension of causal explanations for the co-occurrence of psychiatric and SUDs incorporates stress as central to the genesis and perpetuation of the comorbid state. It should be noted that the role of stress has already been well established independently in both SUDs and other psychiatric disorders. This is based on increasing evidence that these disorders are associated with maladaptive neurobiological changes in brain stress circuits that principally comprise the hypothalamic-pituitary-adrenal (HPA) axis, corticotropin releasing hormone (CRH), and noradrenergic systems. The primary stress mediators of this endocrine system include the glucocorticoids (adrenal steroids such as cortisol) and catecholamines (dopamine, norepinephrine, and epinephrine), which in the short term are vital for the adaptation and maintenance of homeostasis but which can act detrimentally with repeated long-term release. As previously mentioned, chronic drug use is linked with neuroadaptations in the brain reward pathways, predominantly involving mesolimbic dopamine projections. One possible mechanism of interaction could be that stress activation of the HPA axis causes the release of glucocorticoids that interact with mesolimbic dopamine release and stimulate neuronal pathways implicated in substance misuse. Neuroadaptations in stress and reward circuits may, accordingly, underlie the adverse states often associated with psychiatric disorders and SUDs. It is feasible therefore that common mechanisms influencing stress, motivation, and behavior may contribute to comorbidity.

Data indicate that stress plays a role in the development and relapse of comorbid disorders. For instance, high levels of exposure to adversity or stressful events have been causally related to the onset of depressive and anxiety disorders. Similarly, stressors may facilitate substance misuse by increasing the reinforcing efficacy of substances that ameliorate the anxiety and dysphoric symptoms induced by stress. Further support is obtained from animal studies in which stress can induce drug self-administration, increase the sensitivity to the reinforcing effects of drugs, and precipitate the reinstatement or relapse of substance use even after prolonged periods of withdrawal. Hence, from a theoretical perspective, stress has been proposed as a factor that links psychiatric disorders and SUDs. This model postulates that a chronic distress state due to repeated stress exposure underlies the severity of psychiatric symptoms and substance use. This adverse state is then associated with maladaptive responses such as drug use in order to maintain homeostasis. Maladaptive responses may, in turn, alter HPA or CRH circuits and various neurotransmitter systems to induce substance

craving and poor behavioral or coping strategies, causing neuroadaptations in the brain stress and reward pathways. Increased drug use to levels characteristic of dependence leads to further changes in these circuits. Such alterations thereby promote the susceptibility to recurrent stress exposure in comorbid individuals. Supporting this paradigm is the notion that various genetic and environmental vulnerability factors (such as family history and social influences, individual differences in responsiveness to substances, specific personality traits, early life trauma, and poor cognitive functioning) aid development of distress states.

Major Depression, Substance Use Disorders, and Stress

A number of the issues discussed here may be usefully illustrated using the relationship among major depression (MD), SUDs, and stress. As previously noted, the high rate of MD in people with a primary SUD and vice versa underscores the increased risk of an individual with one disorder developing the other. Furthermore, there is considerable evidence relating stressors to the onset, perpetuation, and relapse to either MD or SUD. Hence, it may be that life stressors predispose to each disorder independently, that the stress in predisposing the individual to one disorder renders him or her vulnerable to the development of the other, or that underlying factors such as genes or early environment predispose the individual to experiencing stressors, MD, and SUDs. In parallel with understanding these interactions among MD, SUDs, and life stressors from a psychosocial epidemiological perspective, the involvement of CRH and the HPA axis in both MD and SUD suggests shared neural substrates.

Unraveling these associations is only just beginning; however, there is some clinical, biochemical, and neuroimaging evidence to this end. Clinically, MD and substance withdrawal share similarities, including dysphoria or lowered mood, irritability, anxiety, sleep and appetite changes, and impaired concentration. This is especially so for alcohol withdrawal, in which a depressive phase indistinguishable from MD is present initially in up to 80% of patients but commonly remits without antidepressant treatment.

In functional imaging studies using positron emission tomography or functional magnetic resonance imaging of MD reduced frontal metabolism, anterior cingulate hypoactivity and amygdala hyperactivity indicate altered frontolimbic functioning. Similarly, in people with SUDs, decreased frontal and anterior cingulate activity has been described, in particular in distressed cocaine-dependent subjects and methamphetamine abusers with mood and anxiety symptoms.

This group of methamphetamine abusers also demonstrated amygdala hyperactivity, a finding noted in abstaining cocaine-dependent people with cravings induced by image cues. Although these data are not definitive, they indicate that MD and SUDs share common neural regions of dysregulation that are associated with negative affect and stress-related drug seeking.

In acute withdrawal from nicotine and cocaine, both the distress and depressive symptoms positively correlate with changes in CRH and HPA axis responses. Blunted ACTH and cortisol responses to CRH have also been described in people with SUDs, both with and without depressive symptoms. Consistent with this, individuals abstaining from nicotine, alcohol, or multiple substances demonstrated blunted cortisol responses to psychological stressors; D-fenfluramine induced a similar response in abstaining heroin users. These attenuated responses may indicate overactivity of the HPA axis and are commonly seen in MD. The role of this axis is supported by the observations in animal studies that the stress of repeated foot shocks precipitated a relapse to drug use that was prevented by CRH antagonists. In addition, the extrahypothalamic role of CRH in the bed nucleus of the stria terminalis and amygdala is important in mediating the effects of external stressors on relapses to drug use in animal models. Thus, overall increased CRH-related function has been implicated for substance withdrawal, at least for psychostimulants, opioids, ethanol, nicotine, and benzodiazepines. This same change is described in MD with increased CRH levels in the cerebrospinal fluid of depressed subjects and a decrease in CRH levels with antidepressant or electroconvulsive treatment. Hence, it may be that stress as reflected through the perturbation of CRH has a common effect on MD and substance withdrawal.

The most commonly noted neurochemical system changes in MD are the monoamines, especially serotonin and norepinephrine (NE), and CRH and the HPA axis. Some evidence supports a role for dopamine dysfunction; however, therapeutic strategies augmenting dopamine neurotransmission are of equivocal benefit. In contrast, for SUDs dopamine release in the nucleus accumbens is the central mechanism of reward and motivation for use and as previously noted CRH-HPA axis dysregulation appears during withdrawal. Therefore the neuroimaging findings of a shared frontolimbic circuit dysfunction in MD and SUDs probably involve other linking neural systems. Although speculative, one possibility is through CRH-HPA axis and NE, where either peripheral glucocorticoids cross the blood-brain barrier or NE from the locus coeruleus bind to its receptors in the nucleus

accumbens and ventral tegmental area, enhancing DA release within the nucleus accumbens similar to drugs of abuse. Thus, a well-described finding in MD of altered CRH-HPA axis and NE activity can influence the central pathway involved in SUDs. This probably oversimplified schema suggests that a primary MD increases vulnerability to a subsequent SUD and that substance use may be compensatory in reestablishing homeostasis in a deficient reward-motivation pathway. Although it is problematic to extrapolate from clinical observations to neural mechanisms, this proposal is consistent with a much greater decrease in cocaine use in depressed cocaine users compared to nondepressed users when both were treated with antidepressants. Similar but more modest effects have been noted for both alcohol and nicotine users. Also, MD is known to sensitize smokers to stress, which increases the motivation for smoking. Moreover, for abstaining smokers with a psychiatric disorder, depression often accompanies nicotine withdrawal, with relief from these symptoms prompting a relapse. Smokers therefore anticipate that nicotine will provide partial relief from stress and depression, as it does from the symptoms of withdrawal. In support of this premise is the finding that stress and nicotine both induce synaptogenesis (synapse formation) in the VTA. Furthermore, increase in dopamine release caused by nicotine is suggested to improve anhedonia (loss of pleasure), often associated with depressive illness.

Conclusion

This article describes the high rate of comorbidity between SUDs and other psychiatric disorders, briefly outlines competing hypotheses to explain this observation, proposes how stress may contribute to this, and illustrates this with reference to major depression. However, despite the strong epidemiological evidence of a high rate of comorbidity between SUDs and other psychiatric disorders that is exacerbated by stress, the divergent pathologies of major psychiatric disorders and the limited understanding of the neuroadaptations involved in SUDs suggest

that this interaction differs between the disorders. Moreover, although there is limited evidence to support the self-medication hypotheses, alternative explanatory hypotheses remain plausible. Thus, untangling the interaction of stress-related factors in the neurobiological findings in these disorders and their effects on the brain changes from substance dependence will be essential in determining the correct hypothesis of comorbidity. The result of this endeavor could have profound effects on the prevention and treatment of SUD and other psychiatric disorder comorbidity.

See Also the Following Articles

Alcohol, Alcoholism and Stress: A Psychobiological Perspective; Catecholamines; Dopamine, Central; Hypothalamic-Pituitary-Adrenal; Major Depressive Disorder; Schizophrenia.

Further Reading

- Brady, K. T. and Sinha, R. (2005). Co-occurring mental and substance use disorders: the neurobiological effects of chronic stress. *American Journal of Psychiatry* **162**, 1483–1493.
- Kessler, R. C., Chiu, W. T., Demler, O., et al. (2005). Prevalence, severity and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry* **62**, 617–627.
- Khantzian, E. J. (1985). The self-medication hypothesis of addictive disorders: Focus on heroin and cocaine dependence. *American Journal of Psychiatry* **142**, 1259–1264.
- Lowinson, J. H., Ruiz, P., Millman, R. B. and Langrod, J. G. (eds.) (2004). *Substance abuse: a comprehensive textbook* (4th edn.). Philadelphia: Lippincott, Williams & Wilkins.
- Mueser, K. T., Drake, R. E. and Wallach, M. A. (1998). Dual diagnosis: a review of etiological theories. *Addictive Behaviours* **23**, 717–734.
- Nestler, E. J. (2005). Is there a common molecular pathway for addiction? *Nature Neuroscience* **8**, 1445–1449.
- Swendsen, J. D. and Merikangas, K. R. (2000). The comorbidity of depression and substance use disorders. *Clinical Psychology Review* **20**, 173–189.

Comparative Anatomy and Physiology

A G Watts

University of Southern California, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
A G Watts, volume 1, pp 507–511, © 2000, Elsevier Inc.

Stress Responses Defined

Stress Responses in Invertebrates

Stress Responses in Vertebrates

Impact of Stress on the Regulation of Energy Metabolism

Glossary

<i>Cyclostome</i>	A member of the order Cyclostomata (class Agnatha): hagfish and lampreys. These primitive vertebrates have relatively rudimentary endocrine and nervous systems that are thought to represent evolutionarily early examples of these functions.
<i>Cytokine</i>	Regulatory proteins secreted by the white blood cells and a variety of other cells in the body. Cytokines can regulate cells of the immune system and modulate immune responses, but act in a relatively local manner.
<i>Elasmobranch</i>	Cartilaginous fish (e.g., sharks, skates, and rays).
<i>Fenestrated capillary</i>	A specialized capillary that has transcellular circular openings in the endothelial walls allowing the passage of large molecules in or out of the capillary. They are particularly prevalent in regions where secretory products are released into the vasculature.
<i>Mesoderm</i>	One of the three principal layers of the mammalian embryo from which the vascular tissues are derived.
<i>Tetrapod</i>	A four-limbed vertebrate.

Stress Responses Defined

Using the simplest and most wide-ranging definition, a stressor can be thought of as any alteration in the animal's environment (internal or external) requiring a response that, if not invoked, could develop into a threat to survival. In some instances, the disturbance might only require a simple avoidance reflex that moves the animal away from an external threat; in others it might invoke an internally directed motor response nullifying a homeostatic disturbance.

Considered in this broader manner, stress responses are fundamental characteristics displayed by all living organisms. Protozoans display an avoidance reaction to a detrimental change in the local environment, for example, increased CO₂ concentrations in the water; more complex multicellular organisms such as coelenterates and platyhelminths have sensory-motor networks that allow them to make the appropriate distinction between a threat and a meal. All animals use some sort of avoidance (generally reflexive) reaction that can be considered a stress response to an aversive presence in the environment. However, more complex invertebrates and all vertebrates have also evolved sets of physiological responses that provide the organism with the ability to adapt their metabolic and immune systems to deal with the consequences of stress. In these animals, physiological and neural systems involving learning and memory also provide organisms with anticipatory features allowing them to generate conditioned or habituated responses as well as the capability of dealing with the more extended consequences of stress.

Stress Responses in Invertebrates

Stress responses in invertebrates can be divided into two broad categories: (1) those concerned with abrogating the stressor and (2) adaptive changes invoked to deal with the metabolic consequences of the stressor. In invertebrates, the nervous system generally deals with the first, whereas cellular interactions of a chemical nature (endocrine and immune) generally deal with the second.

Neural

For avoidance responses, sessile invertebrates generally confront a threat by retracting vulnerable body parts into a protective shell or the environmental surroundings, whereas more mobile species have the ability to move away from the threat as fast as possible. In most metazoans, the effector of these responses is the nervous system. Simple invertebrates such as coelenterates have avoidance reactions that may involve little more than invoking the stimulus-response properties of their nerve nets. In more complex invertebrates, these avoidance reactions are neurophysiologically quite complex and involve the regulation of specific neural networks by sensory inputs (e.g., air currents). These networks coordinate the very rapid muscle contraction required for fleeing (e.g., jumping, flying, turning, or running). These responses

appear to involve populations of giant interneurons found in the abdominal ganglia that may, because of their fast conduction velocities, bring about the very rapid stimulus–response interactions required for these behaviors. However, the detailed role of these neurons and the wider functioning of this circuitry remains unclear. Although all of these avoidance responses have the characteristics of a reflex, studies from species such as *Aplysia* have shown that contributions from learning and memory mechanisms cannot be overlooked.

Endocrine

Arthropods use hormones to regulate their metabolism in a way that is beneficial during periods of stress. In a manner somewhat analogous to the secretions of the vertebrate chromaffin cell–adrenal medulla system, some insects (e.g., locusts and cockroaches) produce adipokinetic hormones (AKH) in the gland cells of the corpora cardiaca. These hormones increase the heart rate and the carbohydrate and protein metabolism. They also produce hyperglycemia through the breakdown of glycogen to trehalose in the fat bodies and muscles and its facilitated transport into the hemolymph. Starvation, increased activity, and stress all elevate the secretion rate of these hormones.

Complex invertebrates such as annelids, insects, and mollusks possess quite sophisticated immune functions that use interleukin-like signaling molecules and cytotoxic activities derived from natural killer cells. These systems are activated by stressors. Some invertebrates also synthesize corticotropin releasing hormone (CRH) and proopiomelanocortin-(POMC-) derived peptides that are related to their vertebrate homologs. In molluscs, CRH and adrenocorticotrophic hormone (ACTH) are capable of releasing biogenic amines from hemocytes (cells that appear to have both endocrine and immune functions), and at least some of the cells that are capable of phagocytosis also exhibit ACTH immunoreactivity. ACTH-like molecules have even been identified in some protozoans, suggesting their great phylogenetic antiquity. Indeed, the existence of those endocrine signals capable of regulating carbohydrate metabolism and immune responses – properties required to generate adaptive responses to stressors – appear to have developed very early in evolution.

Stress Responses in Vertebrates

Not surprisingly, stress responses in vertebrates have more components, variety, and complexity than those found in invertebrates. Although some vertebrate stress responses are similar in principle to those exhibited by invertebrates (e.g., avoidance, endocrine

secretions that increase energy substrate availability, and adaptations that restrict reproduction), because of their more complex central nervous systems, vertebrates possess a greater flexibility and adaptative capability.

Comparative Anatomy

The general organization of those parts of the neural and endocrine systems involved with mediating stress responses are remarkably well conserved across all vertebrate classes.

The neurohypophysis The neurohypophysis comprises the median eminence of the tuber cinereum, the infundibular stem, and the neural lobe of the pituitary gland. The median eminence and the neural lobe both contain terminals of neuroendocrine motor neurons. However, neurons that regulate ACTH secretion terminate in the median eminence. With appropriate increases in firing rate, they release releasing factors (CRH and, in some mammals, vasopressin) into the hypophyseal portal blood. The neurons that release oxytocin and vasopressin/vasotocin into the general circulation terminate on fenestrated capillaries in the neural lobe. The neural lobe appears to be better developed in tetrapods than in fish, and this may be related to the larger amount of stored hormone required for the increased hydrostatic stresses of terrestrial existence. With a few exceptions (e.g., in cyclostomes, the neurohypophysis is quite poorly developed), the generalized organization of the neurohypophysis is retained across the vertebrate spectrum.

Corticotropes The anterior pituitary of all vertebrates contains cells that synthesize ACTH. In evolutionarily older species (e.g., reptiles and elasmobranchs) there is less intermingling of cell types within the anterior pituitary than is found in more recently evolved species. POMC is cleaved to produce ACTH in the anterior pituitary corticotropes in all vertebrate classes, in which it acts to increase glucocorticoid secretion. Negative glucocorticoid feedback signals also appears to operate in all vertebrate classes examined.

Catecholaminergic and corticosteroidogenic tissues The analog of the adrenal gland is critical to the stress response of all vertebrates. In mammals, it is called the adrenal (or in older literature, the suprarenal) gland because of its position close to the kidneys. In mammals, the adrenal gland has a characteristic cortical and medullary structure – the medulla is neural in origin and synthesizes catecholamines for release into the bloodstream on sympathetic stimulation; the

cortex is mesodermal in origin and is the site of glucocorticoid and mineralocorticoid synthesis. Although cells performing these functions are found in all vertebrates, the way they are arranged and their anatomical association vary from class to class. In amphibians, reptiles, and fish, this tissue is found between the renal tissue and is sometimes called the interrenal gland.

All vertebrates possess tissue that can secrete catecholamines into the bloodstream during stress. Catecholamines (usually adrenaline and noradrenaline) universally increase cardiovascular muscle tone, increase heart rate, and can rapidly mobilize fat and carbohydrate stores. There has been a trend during evolution toward a close relationship between catecholamine-synthesizing and steroidogenic tissue. Thus, in the cyclostomes and some elasmobranchs, chromaffin cells are widely scattered along large veins and the dorsal body wall, but steroidogenic tissue has been more difficult to identify. In many vertebrate classes, catecholaminergic tissue is interspersed with steroidogenic tissue. In mammals, however, chromaffin cells are aggregated in the adrenal medulla and are separate from the steroid-synthesizing cells in the adrenal cortex, although a close functional relationship does exist.

Sympathetic branch of the autonomic nervous system Sympathetic motor activation plays an important role in the stress responses in virtually all vertebrates. However, it is poorly developed in cyclostomes, in which there are no sympathetic chains and relatively little sympathetic innervation of the viscera. In teleost fish, catecholamine secretion is regulated by preganglionic sympathetic activation. However, other regulators such as glucocorticoids can also play an important role. It is possible that the chromaffin cell system first evolved to function as the principal regulator of the cardiovascular and energy metabolism under all conditions, not just stress. But as the sympathetic innervation of the cardiovascular system became more sophisticated and was able to regulate visceral organ function under basal conditions, the actions of the chromaffin system during stress became the more important function.

Juxtaglomerular apparatus Renin and the cells that contain juxtaglomerular granules are found in all vertebrates other than cyclostomes. In more recently evolved vertebrates, these are aggregated in a specific juxtaglomerular apparatus (juxtaglomerular cells and the macula densa), whereas in others only juxtaglomerular cells are seen. In teleost and more primitive fish species, these are often quite distal from the glomerular apparatus.

Comparative Physiology

Vertebrates invoke sets of stress responses that target three critical physiological systems – those involved with fluid balance, energy metabolism, and tissue damage and immune function. Each set has specific neural and endocrine components.

Impact of stress on the regulation of fluid balance

Protecting the fluid compartments of the body is a prime physiological concern, and the physiological systems that contribute to this function are usually activated during stress. Two sets of hormones have evolved to regulate fluid balance; the renin–angiotensin system protects the extracellular fluid compartment, whereas the posterior pituitary lobe hormones regulate the intracellular fluid compartment. In addition, autonomic reflexes are available to regulate cardiovascular and hemodynamic function through the control of vascular tone. Both systems are stimulated by stress.

Renin–angiotensin system The renin–angiotensin system targets vascular smooth muscle, mineralocorticoid-synthesizing tissue in the adrenal cortex, and some circumventricular organs in the brain. Renin is present in all jawed vertebrates, but is absent in the cyclostomes. Its original function may have been to increase blood pressure through vasoconstriction, with its dipsogenic and other effects (e.g., increased release of mineralocorticoids) appearing later.

Vasotocin/vasopressin Maintaining body fluid composition between narrow limits is a requirement for survival for all organisms. This becomes absolute in animals that live in environments of varying temperature or water availability and in animals that move from one major environment to another. Any variation that challenges fluid balance constitutes a stressor that needs to be regulated. The vertebrate brain has evolved a very simple reflex loop to help with this regulation, and such a function has remained reasonably unchanged for the 500 million years since the evolution of the cyclostomes. It involves direct projections from the osmoreceptors to the neuroendocrine neurons that synthesize vasopressin (in mammals) or vasotocin (in birds, reptiles, and most amphibians and fish, including cyclostomes). In this manner, one of the simplest and phylogenetically oldest vertebrate neuroendocrine stress response is probably the reflex release into the circulation of vasopressin/vasotocin that occurs as plasma osmolality increases. The original function of vasotocin (which is still its primary action in some amphibians) may have been a V1 receptor-mediated vasoconstrictive effect that reduces blood flow to the kidney and thereby fluid loss. (In this

respect, vasotocin is a more potent vasoconstrictor than its evolutionary descendant, vasopressin). However, during evolution a V2-mediated action on kidney tubules became the predominant action.

Impact of Stress on the Regulation of Energy Metabolism

Glucocorticoids

The stress-induced release of glucocorticoids is found widely throughout vertebrate classes, in which they act to promote gluconeogenesis. Stress-induced hyperglycemia reportedly occurs in all classes of vertebrates that have been investigated, from cyclostomes to mammals, although the contribution of glucocorticoids to this effect is far from clear in more simple vertebrates. It is worth considering that, in many respects, the hypothalamic-pituitary-adrenal (HPA) motor response to stress is a more proactive than a reactive event. Although ACTH and corticosterone secretion are obviously triggered by the stressor, they are not the motor responses that act to remove the immediate consequences of the stress, as occurs with, for example, adrenaline's actions to generate hypoglycemia. The actions of elevated plasma corticosterone levels (mediated in part by interactions with leptin, insulin, and the thyroid hormones) anticipate the possibility that the stressor will lead to a debilitating sequence of events, particularly the catabolic effects of negative energy balance. In those vertebrate classes that have been studied, the principal function of corticosteroids is on the intermediate metabolism, particularly in stimulating proteolysis and then gluconeogenesis.

Catecholamines of chromaffin cell origin

Increased circulating catecholamines lead to hyperglycemia in all vertebrates examined, although noradrenaline and adrenaline show differences in efficacy across different classes. Also, the tissue that is targeted varies; in mammals both muscle and liver are responsive to adrenaline, whereas only liver glycogen is mobilized in cyclostomes. Stress will increase plasma catecholamines in teleost fish and elasmobranchs.

Inhibition of reproduction

A variety of stressors inhibits reproductive function in most vertebrates. The simplest of these is the negative

energy balance resulting from food restriction or starvation. Here a mechanism (possibly involving elevated neuropeptide Y synthesis in the hypothalamic arcuate nucleus) leads to the inhibition of reproductive function, ensuring that limited energy supplies are not used to raise offspring in adverse conditions. Related stressors such as social crowding may also restrict reproductive capacity for the same reason, although the neural mechanism responsible for this effect is unclear.

Impact of stress on the regulation of immune function

Immune activation and the consequent increases in circulating interleukins that occur during some stressors provide a marked stimulus to the vertebrate HPA axis. In mammals, this elevates CRH release into the portal vasculature and, hence, ACTH release. Increased glucocorticoid release in response to cytokines has been noted in all the mammals, reptiles, birds, and amphibians that have been investigated. However, it should be noted that in nonmammalian classes, data are very scanty. In turn, corticosteroids have a suppressive effect on immune function. During acute stress, this operates as a dampening effect to restrain any adverse effect of immune stimulation, but in chronic stress this can have a deleterious effect leading to increased infections and mortality. These types of effect have been seen in a wide variety of vertebrates.

Further Reading

- Bentley, P. J. (1998). *Comparative vertebrate endocrinology*. Cambridge, UK: Cambridge University Press.
- Burrows, M. (1996). *The neurobiology of an insect brain*. Oxford: Oxford University Press.
- Dores, R. M. and Lecaude, S. (2005). Trends in the evolution of the proopiomelanocortin gene. *General Comparative Endocrinology* **142**, 81–93.
- Flik, G., Klaren, P. H., Van den Burg, E. H., et al. (2006). CRF and stress in fish. *General Comparative Endocrinology* **146**, 36–44.
- Highnam, K. C. and Hill, L. (1977). *The comparative endocrinology of the invertebrates* (2nd edn.). Baltimore, MD: University Park Press.
- Ottaviani, E. and Franceschi, C. (1996). The neuroimmunology of stress from invertebrates to man. *Progress in Neurobiology* **48**, 421–440.
- Wendelaar Bonga, S. E. (1997). The stress response in fish. *Physiological Reviews* **77**, 592–625.

Concentration Camp Survivors

J D Kinzie

Oregon Health & Science University, Portland, OR, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 515–518, © 2000, Elsevier Inc., with revisions made by the Editor.

History of Concentration Camp Syndrome
 Nazi Concentration Camp Survivors
 Prisoners of War Concentration Camp Survivors
 Cambodian Pol Pot Concentration Camp Survivors
 Concentration Camp Survivors and the Global War On
 Terrorism
 Summary

Glossary

<i>Concentration camp</i>	A camp where prisoners of war, enemies, aliens, or political prisoners are confined. The term implies that physical mistreatment, torture, indiscriminate murder, and genocide can occur.
<i>Depression</i>	A specific psychiatric disorder consisting of lowered mood and symptoms, such as decreased interest in most activity, loss of appetite and weight, poor sleep, fatigue, feelings of worthlessness, poor concentration, and recurrent thoughts of death.
<i>Holocaust</i>	The systematic destruction of over six million European Jews by the Nazis before and during World War II.
<i>Posttraumatic stress disorder (PTSD)</i>	A specific psychiatric disorder following psychological trauma with symptoms of reexperiencing, such as nightmares and flashback; avoidance behavior, such as attempts to avoid thoughts and memories of the events; and hyperarousal symptoms, such as irritability, startle response, and difficulty sleeping.

History of Concentration Camp Syndrome

There is probably no more stressful environment than that of a concentration camp, where people are detained for their combatant or alien status, their political views, or even their ethnicity. The purposes of such camps are often to control, punish, or kill its occupants. Words cannot describe the worse conditions, but crowded living, forced labor, starvation diets, lack of medical treatment, torture, and random or systematic murder are usual. Understandably, the physical and mental effects of the survivors are severe.

A constant and persistent group of symptoms were found among survivors of Nazi concentration camps, which then became known as the concentration camp syndrome. This consists of marked anxiety, motor restlessness, hyperapprehensiveness (startle reactions), sleep disturbance, nightmares, difficulty in concentration, obsessive rumination about the traumatic events, depression, and survivor guilt. Some symptoms seem to represent a form of brain damage, and early researchers found that beatings, malnutrition, and disease were the cause of this disorder; however, it eventually became apparent that the severe psychological trauma itself was the cause of this disorder, and it could occur in previously emotionally healthy persons. Several psychiatrists describing the survivors reported that none have escaped the effects; although the effects may be a subtle personality change, such as apathy, seclusiveness, suspiciousness, hostility, and mistrust. The concentration camp syndrome never became a formal American Psychiatric Association diagnosis, but it had much similarity with diagnosis of posttraumatic stress disorder (PTSD), which was recognized in 1980.

Nazi Concentration Camp Survivors

Background

Six million people (mostly Jews) were killed in Nazi concentration camps. Some camps (Treblinka, Auschwitz-Birkenau, Dachau, Chelmno, Sobibor, Belzek, and Majdanek) were for exterminations. Others had less systematic murder, but were characterized by starvation, forced labor, torture, endemic disease, beatings, inadequate clothing, death marches, and a lack of proper medical facilities. While living with constant threats to their lives, camp prisoners were subjected to medical experimentation, gross humiliation, and separation from family and were forced to witness severe cruelty to others. They were defenseless without privacy or any means of protection. Further details may be obtained from *A Teacher's guide to the Holocaust*. The experience lasted 4 years before liberation came at the end of World War II. Much information up to now has been accumulated on the effects of the survivors.

Physical Effects

In a study of physical sequelae of the Holocaust in Norway, survivors had a higher mortality, more frequent hospitalizations, and higher incidents of sick periods. Physical symptoms, such as headaches, persistent dizziness, and gastrointestinal complaints, which can persist 50 years later, were common.

However, most survivors have reported symptoms with a psychogenic component. Few differences have been found for such diagnostic disorders as arthritis, cancer, and Parkinson's disease.

Psychiatric Effects

There have been many community-based, large-scale studies on the Holocaust survivors, some 40 years after the events. Unlike the previous studies on patients seeking clinical help, these have consistently shown more distress in symptoms and lower mood in survivors than in controls.

In one study, 43% of survivors had symptoms compared to 28% of controls. However, there was considerable overlap and some survivors had fewer symptoms than controls. In one study, 40 years after the Holocaust, 46.8% of survivors met the diagnosis of PTSD. Although the trauma experiences of the Holocaust victims were universally severe, some evidence that the more severe trauma, i.e., surviving a death camp setting or being incarcerated longer, had more symptoms. Greater psychological problems were seen in those who experienced the Holocaust as children and showed more PTSD symptoms. Female victims report more distress than males. The Nazi concentration camp victims tended to remain vulnerable to reactivation of symptoms during stress, whether this redevelopment of PTSD symptoms occurred in actual trauma, such as scud missile attacks during the Gulf War, or a perceived threat, such as a sense of increased anti-Semitism. The Nazi concentration camps were so horrible and the evil committed was so beyond comprehension that clinical terms may not do justice to the existential and theological despair of mankind's inhumanity.

Prisoners of War Concentration Camp Survivors

Background

Wars and upheaval in this century have led to thousands experiencing captivity by the enemy often with malnutrition, overcrowding, torture, poor diet, inadequate shelter, and terrorizing executions. The long-term effects have been reported in survivors in the United States, Australia, Israel, Canada, The Netherlands, United Kingdom, and, more recently, Bosnia Herzegovina.

Physical Effects

American prisoners of war (POWs) from Japan had excess mortality from accidents and tuberculosis. POWs from the Korean War had increased cognitive

deficits that were related to the amount of captivity and weight loss.

Psychiatric Effects

Follow-up studies of the POWs demonstrated a uniformly high rate of persistent and debilitating psychiatric disorders. High rates of psychiatric hospitalizations for neurosis were found in POWs from World War II of Germany and Japan. In another study, high rates of anxiety reaction, alcoholism, and even schizophrenia have been found with greater occurrence in the POWs held by Japan, which inflicted more cruelty on the prisoners. More than 40–50 years after World War II and the Korean War, over half had at sometime a diagnosis of PTSD. The severely traumatized group of POWs held by Japan had a PTSD lifetime prevalence of 84% and a current diagnosis of 58%. PTSD also greatly increases the risk of other disorders, including depression and alcoholism. From a series of studies, we know that the POW concentration camp experience has severe and persistent effects, with up to 50% having PTSD, depression, and alcohol abuse. Many studies found that few survivors actively sought medical treatment and were suffering quite privately.

Cambodian Pol Pot Concentration Camp Survivors

Background

One of the most recent and best-studied concentration camp experiences was that afflicted on the Cambodian people by the radical Marxist Pol Pot regime between 1975 and 1979. During this brutal time, millions of people were rounded up and placed in isolated camps with forced labor, no medical facilities, starvation diets, indiscriminate beatings, and murders, often with family members watching helplessly. One-quarter to one-third of Cambodia's seven million people died during this time. Many survivors went to Thailand as refugees and were later given asylum in Western countries.

Physical Effects

Even after they have improved physically from acute effects, many Cambodian refugees continue to report somatic complaints; 15–20% reported much health impairment. A very high percentage of Cambodian clinic patients (59% of males and 49% of females) have documented hypertension. Incidence of type 2 diabetes is about the same as the population in the United States (about 13.5%) but its prevalence is higher in young people than in the United States. It is difficult to know if these effects are due to refugee

status or dietary in addition to the probable long-term effects of trauma.

Psychiatric Effects

The first studies on Cambodian concentration camp survivors found that the symptoms of nightmares, startle reaction, intrusive memories of events, a strong tendency to avoid thinking about this, and marked concentration problems persisted. In addition to these symptoms of PTSD, most have major depressive symptoms. In a community study (nonpatients), 50% of Cambodians were found to have PTSD and 50% also had depression. The depression symptoms increased over time, unlike the PTSD symptoms, which tend to wax and wane, but continue to affect about 50% of Cambodians even 12 years later. An important finding was the marked sensitivity to stress. Cambodians who witnessed accidents, underwent surgery, or even saw violence on television, such as viewing the destruction of the twin towers on 11 September 2001, had acute PTSD symptoms involving the trauma in Cambodia. Interestingly, drug and alcohol abuse was a rare condition among Cambodians, unlike the POWs.

Concentration Camp Survivors and the Global War On Terrorism

As a result of the global war on terrorism, detention centers for terrorist suspects or unlawful enemy combatants have been set up in Afghanistan, Iraq, and most published, Guantanamo Bay. These camps administered by United States military in no way match the horror of The Nazi or Khmer Rouge; however, evidence indicated that most detainees have been in legal limbo without an open and fair trial. As of June 2005, the United States was holding about 520 foreign terrorism suspects at Guantanamo Bay. Released prisoners, usually held for many years without charge, have alleged ongoing torture, sexual degradation, and religious persecution. As of August 2003, 29 inmates had attempted suicide (the more recent number is 34) and one-fifth of all prisoners were taking antidepressants. There are no studies on the effects of the experience of the detainees, but if other concentration camps can be a model, many will suffer long-term psychological changes and develop PTSD, depression, and other health problems. The very existence of such camps represents a profound moral and ethical challenge for democratic societies.

Summary

The concentration camp syndrome, as first described following the Nazi Holocaust, represented a

catastrophic stress and resulted in a common symptom pattern present in most survivors. These symptoms included marked anxiety, restlessness, startle reaction, sleep disturbances, nightmares, poor concentration, and excessive rumination about the traumatic events. The concentration camp syndrome was never recognized as an official American Psychiatric Association diagnosis but its symptoms are the core of what is now called PTSD. The symptoms have been found in other concentration camp survivors, such as POWs and the more recent Cambodian concentration camp survivors. The response is the result of extreme stress and indicates that such stress can produce a psychiatric disorder that is severe and persistent in many people.

An associated aspect of the disorder is the tendency to have reactivation on symptoms when further stress occurs. A real or symbolic threat can reproduce symptoms or behavior (such as extreme fear or nightmares) just as if the person were in the concentration camp. This speaks to the powerful memories that such extreme stress can cause, even though these memories may be repressed for a time.

Treatment of the concentration camp syndrome originally was not satisfactory, probably because the early methods used psychoanalysis (or similar techniques), which seemed too stressful for the patients. Current techniques using medication to reduce nightmares, sleep disorders, and hyperarousal, and supportive psychotherapy seem to help the patient integrate the experience and reconnect emotionally to others.

An emerging issue is that children of concentration camp victims also seem to suffer indirect effects and may be more vulnerable to PTSD than others. Even under the most extreme conditions, only about 50% of concentration camp victims suffer the full PTSD (concentration camp syndrome) or other physical or psychiatric disabilities. Reasons for this vulnerability of some and resiliency of others are unknown, but genetic factors are thought to play a part. The relationship between such massive psychological trauma and the development of psychiatric disorders is an important area of research.

The global war on terrorism has resulted in multiple concentration camps for detainees. Evidence is beginning to accumulate that torture occurs at such places and their may be severe long-term mental and physical for the survivors.

See Also the Following Articles

Holocaust, Stress Effects of; Korean Conflict, Stress Effects of; Prisoners of War; Survivor Guilt.

Further Reading

- Camp X-ray (and Guantanamo Bay) Available online at: http://en.wikipedia.org/wiki/Camp_X-ray (and Guantanamo Bay). Last accessed 1 October 2006.
- Danieli, Y. (1998). *International handbook of multigenerational legacies of trauma*. New York: Plenum Press.
- Eitinger, L. (1980). The concentration camp syndrome and its late sequela. In: Demsdale, J. (ed.) *Survivors, victims, and perpetrators: essays on the Nazi Holocaust*. New York: Hemisphere.
- Guantanamo Bay. Available online at: http://en.wikipedia.org/wiki/Guantanamo_Bay. Last accessed 1 October 2006.
- Florida Center for Instructional Technology. A teacher's guide to the Holocaust. Available online at: <http://fcit.coedu.usf.edu/holocaust/timeline/camps.htm>. Last accessed 1 October 2006.
- Hunter, E. J. (1988). The psychological effects of being prisoners of war. In: Wilson, J. P., Harel, Z. & Kahana, B. (eds.) *Human adaptation to extreme stress from the Holocaust to Vietnam*. New York: Plenum Press.
- Kinzie, J. D. (1988). The psychiatric effects of massive trauma on Cambodian refugees. In: Wilson, J. P., Harel, Z. & Kahana, B. (eds.) *Human adaptation to extreme stress from the Holocaust to Vietnam*. New York: Plenum Press.
- Krystal, H. (ed.) (1968). *Massive psychic trauma*. New York: International Universities Press.
- Marsella, A., Bornemann, T. and Ekblad, S. (eds.) (1994). *Amidst peril and pain*. Washington, DC: American Psychiatric Association.
- Matussek, P. (1975). *Internment in concentration camps and its consequences*. New York: Springer-Verlag.
- Stenger, C. A. (2003). *American prisoners of war in WWI, WWII, Korea, Vietnam, Persian Gulf, and Somalia*. Washington, DC: Department of Veterans Affairs Advisory Committee on Former Prisoners of War.

Congenital Adrenal Hyperplasia (CAH)

A Solomon and P-M G Bouloux

Royal Free Hospital and Medical School, London, UK

© 2007 Published by Elsevier Inc.

Classification
Molecular Genetics
Biochemical Abnormalities
Clinical Features
Diagnostic Investigations
Treatment of Congenital Adrenal Hyperplasia
Conclusion

Glossary

<i>Allele</i>	One of two or more copies of a particular gene located on a chromosome.
<i>Androgens</i>	Sex steroids that stimulate the development of male sex organs and male secondary sexual characteristics.
<i>Azoospermia</i>	The complete absence of sperm in seminal fluid.
<i>Compound heterozygote</i>	A pair of alleles that are both mutated with abnormalities distinct from one another.
<i>Homozygote</i>	A pair of genes that are identical.

<i>Hyperplasia</i>	The increase in the size and number of normal cells in an organ or tissue.
<i>Hypogonadism</i>	A condition that results from the failure of sex steroid production.
<i>Phenotype</i>	The expression of the characteristics of a person as determined by his or her genes and influenced by environmental factors.
<i>Virilization</i>	The clinical consequences of excessive androgen production in females; characterized by temporal balding, a male body form, muscle bulk, deepening of the voice, enlargement of the clitoris, and hirsutism.

Congenital adrenal hyperplasia (CAH; previously known as adrenogenital syndrome) represents a diverse range of inherited disorders resulting from altered activity in the enzymatic steps required in steroid hormone synthesis. In each, cortisol production is reduced, leading to an increase of pituitary adrenocorticotrophic hormone (ACTH; corticotropin) secretion by negative feedback, which causes adrenal hyperplasia and the overproduction of either cortisol precursors or androgens. The incidence of the classic disorder is 1 in 5000–20 000 live births in most populations. This rate increases in certain specific geographic locations, including the Yupic population of

Alaska, the island of La Réunion, Brazil, and the Philippines. There is wide racial variation and phenotypic heterogeneity. Among Caucasians a 1/60 carrier frequency is found; in the Ashkenazi Jewish population, the carrier rate for nonclassic CAH is up to 19%.

Classification

CAH can be classified according to the specific type of enzyme deficiency involved. Classic CAH refers to the more severe form of enzyme deficiency with a prenatal, neonatal, or early childhood presentation. Nonclassic CAH represents a milder variant of the disorder, associated with less severe enzyme deficiency, a later presentation, and usually mild hyperandrogenism, often mimicking polycystic ovarian syndrome (PCOS). When presenting in adulthood, symptoms may be unmasked by various factors including cytochrome p450-inducing drugs such as anticonvulsants. CAH can also be classified as salt-wasting, non-salt-wasting, virilizing, or feminizing (see [Table 1](#)).

Molecular Genetics

The gene encoding the 21-OH enzyme is found on chromosome 6 (6p.21.3), and within this locus are both active and inactive (pseudo-)genes known as CYP21A2 and CYP21A, respectively. Most mutations involve the recombination of active and inactive genes, creating compound heterozygotes (i.e., different mutations on the two alleles) – with phenotype correlating with the less severely mutant allele – explaining the phenotypic variability of the disorder. Alongside common mutations, novel, highly variable

mutated alleles of CYP21A2 are increasingly being recognized, especially in the nonclassic phenotype. An unusual mechanism involving inactivating mutations of the P450 oxidoreductase enzyme causing simultaneous combined P450P17 and P450C21 deficiency has now been described in several families and provides significant insight into the pathophysiology of the disorder. Interestingly, *in vivo* this mechanism appears to confer a combination of androgen excess prenatally and androgen deficiency postnatally.

Biochemical Abnormalities

The family of disorders causing CAH are due to deficient activity (either the reduction in the amount of enzyme or impaired enzymatic function) in one of the five steps involved in cortisol biosynthesis ([Figure 1](#)). 21-Hydroxylase (21-OH) deficiency is the most frequent form. This enzyme converts progesterone to deoxycorticosterone and 17-hydroxyprogesterone (17-OHP) to 11-deoxycortisol in the endoplasmic reticulum (ER). Less commonly, deficient activity of 11 β -hydroxylase (11 β -OH) in the mitochondria or 3 β -hydroxysteroid dehydrogenase (3 β -HSD) and 17 α -hydroxylase (17 α -OH) in the ER can result in the syndrome. Last, the steroidogenic acute regulatory protein (StAR) is essential for transporting cholesterol to the mitochondria for steroid biosynthesis in the adrenal and gonad, and an abnormality in the gene encoding this protein leads to the very rare but devastating congenital lipid adrenal hyperplasia.

[Figure 2](#) illustrates the mechanism whereby deficient 21-OH enzyme activity leads to a decrease in cortisol production and an increase in adrenal androgen

Table 1 Classification and presentation of congenital adrenal hyperplasia at diagnosis^a

Deficiency	Salt-wasting	Virilization	Feminization (undervirilization)	Precocious puberty	Delayed puberty	Hypertension	Infertility, menstrual disturbance, hirsutism, and acne
21-OH (>90%)	Yes (not in late onset nonclassic cases)	Yes (severity varies)	No	Yes (possible presenting complaint)	No	No	Yes
11 β -OH (~5%)	No (almost all cases)	Yes (severity varies)	No	Yes	No	Yes (most cases)	Yes
3 β -HSD	Yes	No (mild or none at birth)	Yes (in males with classic cases)	Yes	No	No	Yes
17 α -OH	No	No	Yes	No	Yes	Yes	No
StAR	Yes	No	Yes	No	Yes	No	No

^a3 β -HSD, 3 β -hydroxysteroid dehydrogenase; 11 β -OH, 11 β -hydroxylase; 17 α -OH, 17 α -hydroxylase; 21-OH, 21-hydroxylase; StAR, steroidogenic acute regulatory protein.

and metabolic acidosis. Sodium conservation improves with age in some patients, although it is unclear why this should occur.

Virilization and precocious puberty In females, clitoral enlargement in early childhood may occur due to overproduction of adrenal androgens. In males, the penis may enlarge but the testes remain small. In both, pubic hair growth may occur early and excess linear growth with concomitant epiphyseal closure is common. Circulating androgens may prematurely activate the hypothalamic-pituitary-gonadal axis, and the child may have true precocious puberty. Short stature due to premature epiphyseal closure invariably occurs unless the condition is treated early.

Delayed puberty Delayed puberty is a rare manifestation of CAH and can occur in poorly controlled 21-OH deficiency with the suppression of gonadotropins by excess adrenal androgens.

Primary amenorrhea Primary amenorrhea is an uncommon presentation that may occur in less severe variants of 17 α -OH deficiency or 3 β -HSD deficiency when phenotypic females with poorly developed secondary sexual characteristics (breast development and pubic hair) fail to menstruate due to deficient production of ovarian estrogen.

Impaired fertility A degree of impaired fertility found in some patients is multifactorial. Live birth rates resulting from patients with classic CAH continue to improve – ovulation may be spontaneous or assisted by clomiphene and/or gonadotropins.

Nonclassic Congenital Adrenal Hyperplasia

Nonclassic forms of CAH have their onset of symptoms in late childhood or in adulthood.

Precocious puberty Less severe forms of CAH with no evidence of salt-wasting can present for the first time in later childhood. Accelerated linear growth and bone age can result in reduced ultimate adult height unless the correct diagnosis is made and treatment given.

Hypertension In those types of CAH in which excess mineralocorticoids are produced, hypertension may result, which may become apparent only in later life.

Acne and hirsutism in the female Nonclassic types of CAH with excess androgen production can be

indistinguishable from those women with polycystic ovaries, which are frequently present on ultrasound. Approximately 1–3% of hyperandrogenism in adult females is due to late-onset CAH.

Impaired fertility and menstrual disorders Fertility can be impaired in both males and females with CAH. In females, this is due to excess progesterone and androgens suppressing the normal cyclicity of the hypothalamic-pituitary-gonadal axis plus various factors affecting fertilization and implantation. In males, azoospermia may be reversed with glucocorticoid replacement and various forms of assisted conception.

17 α -Hydroxylase Deficiency

The specific condition of 17 α -OH deficiency deserves highlighting due to the resultant deficiency of sex steroid production. It exhibits marked phenotypic heterogeneity, with a disproportionately higher prevalence found in Brazil. There may be a severely impaired development of external genitalia in both sexes; genotypic males are usually raised as females due to a lack of *in utero* and postnatal male phenotypic development. In addition, the mechanism of this metabolic defect generates excess mineralocorticoid with its associated clinical features.

Diagnostic Investigations

Once the clinical suspicion is raised, the diagnosis of CAH can be made by measuring the level of steroid precursors in the basal state (see [Table 2](#)) or after ACTH stimulation (see [Figure 3](#)). 17-OHP can also be measured in saliva; examining the urinary steroid profile for abnormal concentrations or ratios of metabolic products of the various steroids can be diagnostic (see [Table 2](#)). The adrenal glands can be imaged by ultrasound, computed tomography (CT) scan, or magnetic resonance imaging (MRI) and may show evidence of bilateral hyperplasia in contrast to the focal lesion, which may be seen with an adrenal tumor. Prenatal diagnosis can be made by the measurement of 17-OHP or androstenedione in amniotic fluid. Human leukocyte antigen (HLA) typing and linkage analysis of cultured fetal cells from chorionic villus sampling in the first trimester or amniocentesis later in gestation can reliably predict genotype for 21-OH deficiency. Automated sequencing and polymerase chain reaction using DNA extracted from uncultured or cultured fetal cells can be employed to screen for specific mutations in the genes responsible for CAH.

Table 2 Basal biochemical profile in congenital adrenal hyperplasia^a

Deficiency	Serum							Urine		
	17-OHP	A ₄	DHEAS	T	S	P	PRA	17KS	Pregnantriol	THS
21-OH	↑↑	↓	-,↑	↑	-,↓	-	-,↑↑	↑↑	↑↑	↓
11β-OH	-,↑	↓	↑	↑	↑↑	-	-,↓↓	↑↑	↑	↑↑
3β-HSD	-,↑	2,↑	↑	↓	-,↓	↓	↑	↑↑	2,↓	2,↓
17α-OH	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
StAR	↓	↓	↓	↓	↓	↓	↑	↓	↓	↓

^a-, ↓, ↑, ; 3β-HSD, 3β-hydroxysteroid dehydrogenase; 11β-OH, 11β-hydroxylase; 17α-OH, 17α-hydroxylase; 17KS, 17-ketosteroid; 17-OHP, 17-hydroxyprogesterone; 21-OH, 21-hydroxylase; A₄, androstenedione; DHEAS, dehydroepiandrosterone sulfate; P, PRA, S, 11-deoxycortisol; StAR, steroidogenic acute regulatory protein; T, testosterone; THS.

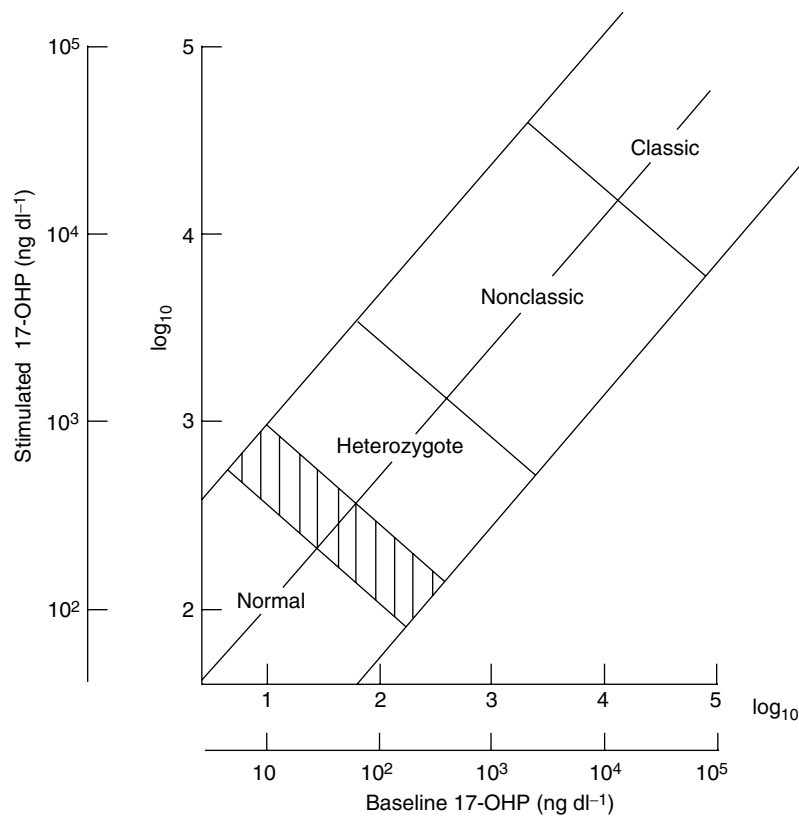


Figure 3 Nomogram for genotyping 21-OH deficiency based on baseline plasma 17-OHP and values obtained 60 min following stimulation with 250 μg ACTH (Synacthen). Note there is considerable overlap between the normal unaffected individual and the heterozygote. 17-OHP, 17-hydroxyprogesterone. From M. C. Young and I. A. Hughes, used with permission.

Treatment of Congenital Adrenal Hyperplasia

The aims of treatment are usually to correct cortisol and mineralocorticoid deficiency and to suppress adrenal androgen production, thereby preventing virilization. If successful, this should allow normal maturation, growth, development, and fertility in adult life. The patient should be monitored for the potential adverse effects of overtreatment with glucocorticoids.

The daily dose of glucocorticoid therapy should correspond to the normal adrenal production rate of approximately 7 mg/m²/day usually given in split dosage. The decision on dosage schedule and preparation of replacement glucocorticoid is still controversial. Larger dosing at night with longer-acting, more potent glucocorticoids such as dexamethasone decreases the exaggerated morning ACTH (and also adrenal androgen) peak of the circadian rhythm but theoretically can suppress nocturnal growth hormone

release. The dosage of glucocorticoid should be increased two- to threefold in the event of a febrile illness, and for more severe stresses higher doses or even parenteral therapy may be required. The theoretical concern in relation to the bone density effects of glucocorticoids does not appear to be borne out by the evidence available from carefully treated patients. Glucocorticoids are not used for replacement in the treatment of nonclassic CAH but are used to suppress adrenal hyperandrogenism by inhibiting the morning ACTH peak by delivery of small doses of dexamethasone (e.g., 0.5 mg) or prednisone (e.g., 5 mg) at night. In the nonclassic and classic variants, an antiandrogen can be used in combination with glucocorticoids to treat symptoms of hyperandrogenism.

The adjustment of the dose of glucocorticoid depends on clinical assessment combined with the results of monitoring in the cases of 21-OH deficiency, 17-OHP, adrenal androgens, and/or urinary 17-ketosteroid production. These are, however, not ideal biomarkers, and efforts continue to find more specific ways of monitoring the disease.

Mineralocorticoid replacement therapy with 9 α -fludrocortisone should be used in all those patients with salt-wasting forms of CAH. The use of mineralocorticoid therapy in nonclassic late-onset variants of the disorder with raised plasma renin activity, which reflects salt loss, often allows a reduction in the dosage of glucocorticoid and subsequently less risk of iatrogenic side effects. Renin levels can be used in the monitoring of fludrocortisone dosage. In refractory cases, especially when subfertility, obesity, and hirsutism may be prominent features, adrenalectomy is sometimes performed, followed by adrenal steroid replacement.

Prenatal treatment of affected females *in utero* with dexamethasone to prevent virilization of the fetus has been used with varying success. To be effective, treatment must be started at less than 10 weeks (preferably 5–7 weeks) gestation and continued until delivery in affected females but discontinued in males and unaffected females. Mendelian inheritance dictates that without diagnostic confirmation, treatment will only be appropriate in one out of eight fetuses. Maternal side effects such as weight gain, hypertension, diabetes, and edema are common but fortunately more severe side effects are rare. This field lacks consensus, and registries have been created in certain locations, such as Sweden, where the gold standard is that treatment is provided only within the framework of a prospective trial.

The timing of corrective surgery and correct gender assignment is the subject of much debate, with earlier single-procedure operations becoming more popular. There is still controversy over the ideal mechanisms

for assigning gender, and multidisciplinary teams have been found to be the optimal strategy for handling the delicate issues involved. In severe forms of 21-OH and 11 β -OH deficiency, virilization may be so severe that genotypic females may need to be raised as males and ovariectomy will be required with prostheses inserted into the scrotum at puberty. There is evidence that exposure to androgens *in utero* can alter gender and sexual identity in affected females, in whom there is an above-average increase in tomboyish behavior and lesbianism. These sequelae demonstrate the need for early and adequate treatment of the condition and long-term follow-up to ensure well-balanced individuals, progress through puberty, and the formation of stable relationships. Sexual function in adulthood is an area of increasing discussion, and the psychosocial impact of this condition has only recently been fully explored.

Conclusion

CAH is a heterogeneous condition that causes multiple imbalances in metabolic pathways of steroid hormone synthesis. Knowledge of the genetic basis for the condition has allowed increasingly specific methods of screening and treatment to be developed.

See Also the Following Articles

Adrenocorticotrophic Hormone (ACTH); Androgen Action.

Further Reading

- Arlt, W. and Stewart, P. M. (2005). Adrenal corticosteroid biosynthesis, metabolism, and action. *Endocrinology & Metabolism Clinics of North America* 34(2), viii, 293–313.
- Arlt, W., Walker, E. A., Draper, N., et al. (2004). Congenital adrenal hyperplasia caused by mutant P450 oxidoreductase and human androgen synthesis: analytical study. *Lancet* 363(9427), 2128–2135.
- Friaes, A., Rego, A. T., Aragues, J. M., et al. (2006). CYP21A2 mutations in Portuguese patients with congenital adrenal hyperplasia: identification of two novel mutations and characterization of four different partial gene conversions. *Molecular Genetics and Metabolism*.
- Gonzalez, A. A., Reyes, M. L., Carvajal, C. A., et al. (2004). Congenital lipoid adrenal hyperplasia caused by a novel splicing mutation in the gene for the steroidogenic acute regulatory protein. *Journal of Clinical Endocrinology and Metabolism* 89(2), 946–951.
- King, J. A., Wisniewski, A. B., Bankowski, B. J., et al. (2006). Long-term corticosteroid replacement and bone mineral density in adult women with classical congenital adrenal hyperplasia. *Journal of Clinical Endocrinology and Metabolism* 91(3), 865–869.

- Lajic, S., Nordenstrom, A., Ritzen, E. M., et al. (2004). Prenatal treatment of congenital adrenal hyperplasia. *European Journal of Endocrinology* 151(supplement 3), U63–U69.
- Loidi, L., Quinteiro, C., Parajes, S., et al. (2006). High variability in CYP21A2 mutated alleles in Spanish 21-hydroxylase deficiency patients, six novel mutations and a founder effect. *Clinical Endocrinology* (Oxford) 64(3), 330–336.
- Merke, D. P. and Bornstein, S. R. (2005). Congenital adrenal hyperplasia. *Lancet* 365(9477), 2125–2136.
- Ogilvie, C. M., Crouch, N. S., Rumsby, G., et al. (2006). Congenital adrenal hyperplasia in adults: a review of medical, surgical and psychological issues. *Clinical Endocrinology* (Oxford) 64(1), 2–11.
- Rumsby, G., Avey, C. J., Conway, G. S., et al. (1998). Genotype-phenotype analysis in late onset 21-hydroxylase deficiency in comparison to the classical forms. *Clinical Endocrinology* (Oxford) 48(6), 707–711.
- Seckl, J. R. and Miller, W. L. (1997). How safe is long-term prenatal glucocorticoid treatment? *Journal of the American Medical Association* 277, 1077–1079.
- Speiser, P. W. and White, P. C. (2003). Congenital adrenal hyperplasia. *New England Journal of Medicine* 349(8), 776–788.
- Wudy, S. A., Hartmann, M. F., Draper, N., et al. (2004). A male twin infant with skull deformity and elevated neonatal 17-hydroxyprogesterone: a prismatic case of P450 oxidoreductase deficiency. *Endocrine Research* 30(4), 957–964.

Conservation of Resources Theory

S E Hobfoll

Department of Psychology, Kent State University, Kent, USA

J S Ford

Memorial Sloan-Kettering Cancer Center, New York, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S E Hobfoll and J S Kay, volume 1, pp. 519–525, © 2000, Elsevier Inc.

Stress

Occurs in circumstances that represent a threat of loss or actual loss of the resources required to sustain the individual-nested-in-family-nested-in-social-organization. Further, because people will invest what they value to gain further, stress occurs when individuals do not receive reasonable gain following resource investment, this itself being an instance of loss.

Traumatic stressors

Severe, typically infrequent and unexpected events that usually include serious threat to life and well-being.

Conservation of Resources Theory

Resources

Resource Categories

Principles of Conservation of Resources Theory

Corollaries of Conservation of Resources Theory

Conclusion

Glossary

Conservation of resources (COR) theory

States that stress has central environmental, social, and cultural bases in terms of the demands on people to acquire and protect the circumstances that insure their well-being and distance themselves from threat to well-being.

Resources

Those objects, personal characteristics, conditions, or energies that are valued by the individual or that serve as the means for attainment of other objects, personal characteristics, conditions, or energies.

High-risk situations such as war, terrorism, natural disasters, and crime and victimization as well as the onset of a serious illness put people's coping resources to a test and can result in psychological stress and distress. Because these situations are widespread, it is important that we examine the effects that high-risk situations have on people and why some can go through them relatively unscathed, whereas others are almost completely debilitated. As many as 2 million American households experience injuries and physical damage each year from fire, floods, hurricanes, tornadoes, and earthquakes. Over the past two decades, natural disasters and other calamities have killed about 3 million people worldwide and adversely affected the lives of at least 800 million more people. Between 1974 and 1980, there were 37 major catastrophes in the United States alone. In addition, few families are spared the pain of having a family member who has a serious illness or one who passes away.

Although in the United States we think of war as something that does not affect us directly, global conflict is widespread, and communications have forced us to share in the suffering many go through in wars elsewhere. The events of September 11th, 2001 in the form of the attack on the World Trade Center in New York and the Pentagon in Virginia made terrorism all too real for Americans as well. This article explores conservation of resources theory as applied to high-risk situations as one theoretical backdrop for explaining the differences in people's coping successes and their subsequent psychological reactions.

Conservation of Resources Theory

Conservation of resources (COR) theory examines and describes the nature of psychological stress and its likely consequences. Traditionally, stress theories have concentrated on people's individual appraisals of stressful situations as the determining factor of how much distress they will experience. COR theory states that stress is neither first nor foremost a product of individuals' appraisal of events, but that it has central environmental, social, and cultural bases in terms of the demands on people to acquire and protect the circumstances that ensure their well-being and distance themselves from threats to well-being. COR theory posits that stress emanates from difficulty achieving the common goals toward which members of a culture strive. In this regard, stress is largely culturally determined because most of the major demands placed on people have a shared social context and that culture is a social phenomenon. Through personal experience and learning, people come to recognize what they need in order to affirm the acquisition and ownership of what is important directly, indirectly, and symbolically for success within their culture and for sheer survival. These things that individuals value are called resources.

Resources

Resources are those objects, personal characteristics, conditions, or energies that are valued by the individual or that serve as the means for attainment of other objects, personal characteristics, conditions, or energies. COR theory's basic tenet is that people strive to retain, protect, and build resources and that what is threatening to them is the potential or actual loss of these valued resources. In addition, people also endeavor to foster that which they value. Therefore, people work to acquire resources that they do not have, preserve those resources they have, protect resources when they are threatened, and cultivate

resources by positioning themselves so that their resources can be put to best use.

Stress occurs in circumstances that represent a threat of loss or actual loss of the resources required to sustain the individual-nested-in-family-nested-in-social-organization. Further, because people will invest what they value to gain further, stress is predicted to occur when individuals do not receive reasonable gain following resource investment, this itself being an instance of loss.

It follows that psychological stress is a reaction to the environment in which there is (1) the threat of a net loss of resources, (2) the net loss of resources, or (3) a lack of resource gain following the investment of resources. Individual-nested-in-family-nested-in-social-organization implies that these levels are enmeshed because there is no organization or family without individuals and individuals must rely on social attachments for well-being, self-esteem, and survival.

Resource Categories

COR theory identifies resources whose loss is likely to result in stress. Since there is a common basis of human survival, most of these resources are valued across cultures, whereas others are more culturally or familially determined. Three different methods have been considered to categorize these resources. The first method categorizes resources into internal and external types. Several researchers have used this distinction. Internal resources are those that are possessed by the self or are within the domain of the self. This includes resources such as optimism, self-esteem, and a sense of mastery. External resources are those resources that are not provided by the individual, but are external to it. These include resources like social support, employment, and economic status.

Another method of classification of resources is a structurally based system that divides resources that have meaningful differences. The resulting four resource categories are (1) object resources, (2) personal resources, (3) condition resources, and (4) energy resources. Object resources (e.g., shelter, transportation) are valued because of some aspect of their physical nature or because of their acquiring secondary status values based on their rarity and expense. Personal characteristics are resources to the extent that they generally aid stress resistance. These include both skills and personal traits. Some examples of these are social competence, self-esteem, and a sense of mastery. Conditions are resources to the extent that they aid in obtaining other valued conditions or are themselves goals that people value. Some examples of condition resources are marriage and

tenure. These resources are important because they lay the structural groundwork for access to other resources. Energies include resources such as time, knowledge, and money. They acquire value from their ability to be exchanged for resources in the previous three categories. Because of the common basis for human survival, many resources are valued across cultures. However, the order of importance of resources will vary along with differing cultural values, and some resources may be present in some cultures and not others.

The third method of classification of resources is based on the proximity of the resource to survival. Primary resources are those that directly correlate to survival. These include ample food, shelter, safety, and clothing. Secondary resources are those that contribute indirectly to primary resources. These include social support, marriage, and optimism. Tertiary resources include those things that are symbolically related to primary or secondary resources. These include money as well as workplace and social conditions that allow availability or access to secondary resources and resources that signify social status.

Principles of Conservation of Resources Theory

Principle 1

There are two major principles that follow from the basic proposition of COR theory. The first principle is that *resource loss is disproportionately more salient than resource gain*. Therefore, given an equal amount of loss and gain, loss will have a much greater impact. According to COR theory, it is the loss and the threat of loss of resources that principally define stress. This principle has also been developed in work by Tversky and Kahneman in their prospect theory. Their theory states that the gradient of loss is steeper than the gradient for gain, which results in a bias in favor of loss. In experiments they have shown that when problems are framed or worded in terms of loss, greater risk will be taken to try to preserve resources than if the same situation is framed in terms of potential gain. Studies that separate loss from gain events also confirm that only loss events are related to psychological distress and illness. Therefore, it appears that losing resources will be much more notable and will have a greater impact than the equivalent gain.

If loss is more salient than gain, resource loss should have a greater impact on psychological distress than gain. This was examined empirically. First, groups of students, community members, and psychologists nominated resources that they valued.

Next, different groups of people added resources that they felt were important that did not already appear on the list and deleted resources they felt were not valued. This process was repeated with about 50 different groups until no new resources were added that had not been deleted by more than one group prior and until no new deletions were judged to be necessary, ending with a final list of 74 resources.

Next, 255 students and 74 community members indicated whether they had lost or gained each of the 74 resources recently as well as during the past year. They repeated this process twice, separated by 3 weeks. They were also administered well-known anxiety and depression scales. Results indicated that neither recent resource gain nor resource loss during the past year had any direct impact on psychological distress for either group. Recent resource loss and resource loss during the past year, in contrast, had major negative effects on psychological distress. People were deeply negatively effected when they lost resources, but were hardly impacted when they experienced resource gain.

Researchers have applied COR theory to study the impact of both Hurricane Hugo, which affected South Carolina in 1990, and the Sierra Madre earthquake that hit Los Angeles county in 1991. They examined how resource loss influenced the mobilization of resources and also how loss effected mental health. It was found that the greater the resource loss, the more coping individuals engaged in and the more psychologically distressed they became. They also found that the influence of resource loss was both of greater magnitude and independent of the influence of positive coping responses. Supporting COR theory, resource loss was more important in predicting psychological distress than were personal characteristics and coping behavior.

The impact of resource loss on coping, physical health, and psychological well-being was examined further after Hurricane Andrew struck South Florida in 1992. Comparing the impact of resource loss to other factors, it was found that resource loss had the single most profound influence. The greater the resources people lost, the greater was their psychological distress and the worse their immune resistance. It was also found that individuals' sense of optimism counterbalanced much of the negative influence of resource loss. Therefore, those people who had resources, including optimism, were more likely to successfully cope with Hurricane Andrew than did those individuals who lacked resources.

Traumatic stress entails the rapid loss of resources and the resources lost are usually the ones held with highest value by individuals. The rapidness of

resource loss is related to the fact that traumatic stressors (1) attack people's most basic resources, (2) typically occur unexpectedly, (3) make excessive demands on remaining resources, (4) are outside of the realm for which resource utilization strategies have been developed, and (5) leave a powerful mental image that is easily evoked by cues associated with the event. The excessive demands placed on an individual by a war or natural disaster are such that no amount of resources could prevent severe initial reaction to stress. At best, it is expected that a healthy individual could make a reasonably rapid recovery that could be a matter of weeks, months, or years, depending on the trauma. For many traumatic events, it would be expected that negative sequelae be lifelong, but hopefully limited to certain life domains so that most normal functioning continues.

Chronic stressful conditions also have been found to chip away at resources and strong resource reservoirs. For example, social support has been found to diminish for mothers of chronically ill children and for women with breast cancer. In addition, under chronic stress conditions, resources that remain have been found to have increasingly limited effectiveness. This may be the case because loss begets further loss and can lead to loss spirals. Under a chronic stressful condition, not only do people experience more and more loss, they also may deplete resources until they have few resource reserves to contribute to the stress resistance process. For this reason, COR theory has also become a major explanatory system for understanding the stress of burnout in the workplace. In burnout, resources are slowly depleted to produce exhaustion, alienation, and psychological distress.

Although loss is more salient than gain in COR theory, resource gain is an important facet of stress, even if it is secondary. Resource gain is important because it is woven with loss. Loss can be prevented, counterbalanced, or forestalled through resource gain. If initial losses occur, previously stored resources can be applied to minimize the impact of loss. Resource gain also increases in meaning in the face of loss. This occurs because people take stock of their resources when loss occurs. People enact gain cycles in the wake of loss, in part to offset current resource loss, but also because they become more aware of future losses and look to prevent them. Consistent with the idea that resource gain is important in the face of loss, researchers examining workers who were laid off from their jobs found that those people with financial savings did not experience the harmful effects of unemployment. However, they found that prior to the worker's job loss, their savings

had little positive impact, supporting the idea that resource gain or surpluses (in this case a savings account) increases in significance during a loss.

Principle 2

The second principle that follows from COR theory is that *people must invest resources in order to protect against resource loss, recover from loss, and to gain resources*. COR theory states that resources are central to the experience of stress. Stress occurs when resources are lost or threatened, and people use resources to prevent or offset loss and to make other resource gain. This investment of resources occurs by several mechanisms. The first mechanism of resource investment deals with the total investment of a resource. The second kind of resource investment involves risking the resource without total investment. Resource investment may also occur directly or through substitution. Resource investment may counterbalance loss, protect against threat of loss, or contribute toward resource gain.

When mastery, optimism, social support, finances, or status are put into place to offset a major stress event, the importance of principle 2 is underscored. First, the resource must be ample enough to be used. Low self-esteem will have limited value, just as little money can do little good to offset financial losses. The second principle of COR theory also suggests a more subtle point. Specifically, because people typically have little experience with major stressors, they may not know how to employ their resources in these special circumstances. Just as combat troops must learn to apply their skills under increasing pressure, people will learn and adjust to even catastrophic circumstances. However, the cost of resource investment will accelerate as people will at first misuse resources, hold back on resource use to preserve resource integrity, and will need to allow for a quick depletion of resources in some instances.

Corollaries of Conservation of Resources Theory

Corollary 1

In addition to the two major principles presented previously, there are four corollaries of COR theory that outline rules that allow for definitive predictions as to how resources function over time. The first corollary states that those with greater resources are less vulnerable to resource loss and more capable of orchestrating resource gain. Conversely, those with fewer resources are more vulnerable to resource loss and less capable of achieving resource gain. Moreover, those who lack resources are more likely to

experience extreme consequences, as without adequate resource reserves they are less likely to have resources to invest in the wake of initial losses.

Resources may be used either individually or in combination with other resources. In addition, stress often makes multiple demands on people that require different combinations of resources. However, those people with greater resources will be less negatively affected by initial resource loss and will be more likely to create gain cycles because they can invest resources that are not required for everyday functioning. Those people with fewer resources will be more deeply impacted by a major crisis or by chronic demands and will have few resource reserves to assemble. Therefore, initial setbacks will be devastating and result in immediate and rapid loss spirals.

Corollary 2

The second corollary of COR theory states that *those who lack resources are not only more vulnerable to resource loss, but that initial loss begets future loss*. People rely on resources to apply to losses, and therefore with each loss there are fewer resources that can be utilized or invested in gains that might influence the occurrence of stress. With a depleted resources pool, future challenges are increasingly less likely to be met and a downward spiral increases in momentum. This results in depleted resource reserves that are less capable of mobilizing to defend against future challenges. This further suggests that loss cycles will have initially higher velocity for resource-poor individuals or groups, as they are from the outset in a resource-challenged state characterized by their resources being already stretched in protection of the self, family, or social system. Therefore, this corollary predicts that loss cycles will have progressing momentum and strength.

When exposed to very high risk situations, people may lose not only their tertiary and secondary resources, such as transportation and a sense of mastery, but also may lose primary resources, such as food and shelter that are necessary for survival. In this way, the psychological stress that people experience in a high-risk situation is not only taxing because of its severity, but also because of the vast array of resources that it may effect and deplete. Hence, in studies of the impact of September 11th, psychosocial and material resource losses were the most major predictors of posttraumatic distress.

When individuals are not equipped to gain resources, especially if they are in a loss spiral, they are also likely to be particularly vulnerable. For example, in many disasters such as war, the country may be in such turmoil that it is very difficult for

people to gain resources such as employment or money. Individuals may also not be equipped to gain resources because they might have limited access to opportunities to protect themselves or to gain access to the resources available to others in society. Furthermore, resources are not distributed equally in society, and those people who lack resources are most vulnerable to additional losses. Those with weak resources have been found to have the least success under high stress. Loss spirals develop because people lack the resources to offset the cascading of loss. If resources are used to prevent loss of other resources, such loss would be predicted to lead to further decreases in the likelihood of possessing necessary resource reserves. In this manner, people who have weak or few resources prior to the onset of a high-risk situation, like a natural disaster, will psychologically do worse and will experience more loss spirals than those who had stronger or greater resources prior to the disaster.

Environmental circumstances, like high-risk situations such as war, natural disasters, chronic illness, and rape, often threaten or result in multiple depletions of people's resources. These losses are important because resources have instrumental value to people and they have symbolic value in that they help to define for people who they are. Hence, rapid resource loss impacts people's ability to mobilize resistance resources as challenges to the social fabric that binds communities and sustains people as a safety net. Further, ongoing major losses ebb away at people's sense of identity that may result in helplessness and other extreme behaviors evidenced in panic, mob action, and even barbarism.

Corollary 3

The third corollary of the COR theory states that *those who possess resources are both more capable of gain and that initial resource gain begets further gain*. If initial resource gains are made, then greater resources become available for investment. When initial gains are achieved, additional resources become available for investment and individuals and social systems are less vulnerable to loss and loss spirals. In addition, individuals do not necessarily need to rely on these resource surpluses for reserves, so they can benefit further by employing them for gains' sake. However, because resource loss is more potent than resource gain, gain cycles will have less momentum or speed and less impact than loss cycles and people will be motivated to sustain resistance reserves for future misfortune. Seen in this light, in the case of major and traumatic stress, gain cycles will often only occur well after the events' initial impact. Even then,

gain cycles will often begin from a depleted base and might better be framed as rebuilding cycles.

Corollary 4

The fourth and final corollary affirms *that those who lack resources are likely to adopt a defensive posture to guard their resources*. For those with few resources, the cost of resource investment surpasses demands and makes the individual or organization vulnerable. A defensive posture keeps a maximum of resources in reserve in case the person needs to offset a future loss.

Those people who lack resources are predicted to take a defensive posture in order to guard their resources. Those with weak resources have been found to use their resources best when challenged by everyday stressors, but to have the least success under high-stress conditions. In studying victims of Hurricane Andrew, researchers found that resource loss resulted in a marked increase in use of denial coping. This, in turn, placed victims at increased risk for PTSD. To expend resources from a depleted system may be too risky a venture. Instead, a shutdown response may best preserve limited resource pools until the storm (literally in this case) blows over.

Conclusion

In conclusion, loss is particularly salient because it is disproportionately weighted in the human experience and it is harder to prevent loss than to obtain resource gain. Therefore resource loss is more powerful and more potent than resource gain. COR theory states that when loss occurs it is more depleting of resources than gain is resource generating. It is clear that high-risk situations deplete resources as well as resource reserves and can result in loss spirals. COR theory defines psychological stress as a reaction to the environment in which there is the threat of a net loss of resources, the net loss of resources, or a lack of resource gain following the investment of resources. In high-risk situations such as war, natural disasters, or sudden illness, many if not all of these threats can occur. Therefore, COR theory would explain people's distress following a high-risk situation not as an individualized response based solely on characteristics of the individual, but as one that occurs because of the threats to resources as shared within a community.

Although loss is more salient than gain, this is not to imply that resource gain is not important. Resource gain is important because it can prevent, counterbalance, or forestall loss. However, in a

high-risk situation, many factors may prevent resource gain and therefore limit the role of gain in forestalling loss cycles. Because of the long tail of resource loss that follows major stressful circumstances, follow-up to protect the individual and social system are needed on a long-term basis.

See Also the Following Article

Disasters and Mass Violence, Public, Effects of.

Further Reading

- Bolger, N., Vinokur, A., Foster, M. and Ng, R. (1996). Close relationships and adjustment to a life crisis: the case of breast cancer. *Journal of Personality and Social Psychology* 70, 283–294.
- Carver, C. S. (1993). *Coping with Hurricane Andrew*. Paper presented at the 15th International Conference of the Stress and Anxiety Research Society, Madrid, Spain.
- Freedy, J. R., Saladin, M. E., Kilpatrick, D. G., Resnick, H. S. and Saunders, B. E. (1994). Understanding acute psychological distress following natural disaster. *Journal of Traumatic Stress* 7, 257–273.
- Freedy, J. R., Shaw, D. L., Jarrell, M. P. and Masters, C. R. (1992). Towards an understanding of the psychological impact of natural disasters—an application of the conservation resources stress model. *Journal of Traumatic Stress* 5, 441–454.
- Galea, S., Ahern, J., Resnick, H., Kilpatrick, D., Bucuvalas, M., Gold, J., et al. (2002). Psychological sequelae of the September 11 terrorist attacks in New York City. *New England Journal of Medicine* 346, 982–987.
- Hobfoll, S. E. (1988). *The ecology of stress*. New York: Hemisphere.
- Hobfoll, S. E. (1998). *Stress, community and culture*. New York: Plenum.
- Kessler, R. C., Turner, J. B., Blake, J. and House, J. S. (1988). Effects of unemployment on health in a community survey: main, modifying and mediating effects. *Journal of Social Issues* 44, 69–85.
- Lazarus, R. S. and Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer-Verlag.
- Lee, R. T. and Ashforth, B. E. (1996). A meta-analytic examination of the correlates of the three dimensions of job burnout. *Journal of Applied Psychology* 81, 123–133.
- Schönplflug, W. (1985). Goal-directed behavior as a source of stress: psychological origins and consequences of inefficiency. In: Frese, M. & Sabini, J. (eds.) *The concept of action in psychology*, pp. 172–188. Hillsdale, NJ: Erlbaum.
- Tversky, A. and Kahneman, D. (1974). Judgement under uncertainty: heuristics and biases. *Science* 185, 1124–1131.

Control and Stress

A Steptoe

University College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A Steptoe, volume 1, pp 526–531, © 2000, Elsevier Inc.

Behavioral Control and Physiological Stress Responses

Perceived Control

Cognitive Control and Coping

Control and Well-Being

Conclusions

Glossary

<i>Behavioral control (instrumental or objective control)</i>	The situation when the organism has at its disposal a behavioral response that can prevent, reduce, or terminate stressful stimulation.
<i>Cognitive control</i>	A form of psychological coping related to perceived control over one's reactions to aversive events.
<i>Locus of control</i>	A set of generalized control beliefs concerning the extent to which individuals feel that aspects of their lives are determined by their own activities and abilities (internal locus of control) or are governed by external factors.
<i>Perceived control</i>	The belief that one can exert control over stressful events, based on the appraisal of the situation and appropriate perceptions of competence (self-efficacy expectancies) and expectations concerning the outcome of making a response.

Stressors come in many guises, for example, confrontation with a predator, death of a close relative, conflict at work, or living in crowded noisy conditions. When we try to understand what aspects of situations make them more or less threatening, the concept of control emerges as a crucial feature. Stressors differ in the extent to which they can be controlled, and behavioral control is typically associated with a diminution of stress responses. Sometimes the perception of control is sufficient to reduce stress reactions, even though control itself is never actually exercised. For example, studies of people who are anxious about dentistry have shown that telling people they can give a signal that will stop the procedure reduces pain and anxiety, even though they never use the signal. Humans

and other animals also seek out information about impending threatening events and ways of coping with the situation. This phenomenon is known as cognitive control and is an important weapon in the armory of psychological coping. Finally, there are control beliefs or the sense of control that people have in specific areas or over their lives and destiny more generally. A sense of control is generally adaptive for health, but the need for control may under some circumstances be maladaptive and lead to an inflexibility in coping with difficult problems. Since control has so many facets, there is a danger that the construct is misused and disguises careless reasoning. In this article, the different aspects of control and stress are described, along with evidence related to physiological responses, well-being, and health. The ways in which control can be enhanced to improve quality of life, ameliorate pain, and promote rehabilitation are also discussed.

Behavioral Control and Physiological Stress Responses

Behavioral control can be defined as having at one's disposal a behavioral response that can prevent, reduce, or terminate stressful stimulation. Natural stressors vary across the entire spectrum of controllability. Some events, such as the unexpected death of a family member, are outside personal control. But in many situations there is an element of behavioral control. People who regularly find themselves stuck in traffic jams on their morning journey to work may feel helpless, but actually have some control; they might choose to use some other form of transport, to start their journey at a different time, to select a different route, to stop and take a break until the traffic clears, and so on.

Study of the impact of behavioral control over stress is complicated by the fact that many of the most aversive life events are uncontrollable. Comparisons between controllable and uncontrollable events may therefore be confounded with the aversiveness of the experience. Experimental studies provide the best opportunity to examine the effects of behavioral control, since aversiveness can be held constant. The yoked design championed by Jay Weiss and later by Martin Seligman and others has served to illustrate the effects of behavioral control most clearly. In this paradigm, a pair of rodents is exposed to identical intermittent aversive stimuli (electric shocks or loud noise). One animal is able to make a behavioral response that either terminates the stimuli or delays

Table 1 Adverse effects of uncontrollable stress in comparison with exposure to matched controllable aversive stimulation

Decreased food/water consumption
Greater weight loss
Higher plasma corticosterone
Increased gastric lesions
Reduced production of specific antibodies
Reduced lymphocyte reactivity
Decreased cytotoxic activity of natural killer cells
Decreased tumor rejection
Increased susceptibility to malignancy

their onset, while the second animal has no response available. The amount of aversive stimulation experienced by both depends on the effectiveness of the performance of the one in the behavioral control condition, and any differences in physiological responses or pathology will result from the availability of behavioral control.

Table 1 summarizes some of the many physiological and pathological advantages conferred by control in this design. It can be seen that behavioral control is associated with amelioration of physiological stress responses and pathological outcomes. These range from reduced levels of neuroendocrine response to slower proliferation of experimentally implanted malignancies. The regions of the brain responsible for these effects have been identified. Uncontrollable stressors appear to sensitize serotonergic neurons in the dorsal raphe nucleus, while this activation is blocked by projects from the central medial prefrontal cortex when stressors are controllable. But not all the biological responses to uncontrollable stress are detrimental. For example, uncontrollable conditions tend to lead to greater stress-induced analgesia than matched controllable conditions, and this mechanism may permit organisms to endure stressful stimulation at a reduced level of physical discomfort.

Research in humans has gone some way toward duplicating these effects, with differences in the magnitude of acute cardiovascular and endocrine responses. However, an important caveat concerns the effort or response cost associated with maintaining behavioral control.

Control and Effort

Effects of the type shown in **Table 1** have generally emerged in studies in which behavioral control is easy to exert. For example, electric shock may be avoided by a lever press on a simple schedule of reinforcement. However, control may require great effort to maintain and can be associated with a degree of uncertainty as to whether it has been successful.

Under these circumstances, physiological activation may be greater than that elicited in uncontrollable situations. This was illustrated in the classic human experimental studies carried out by Paul Obrist. Obrist randomized human volunteers to three conditions in a shock avoidance reaction time study. All participants had to respond rapidly to aversive stimuli and were informed that successful performance would lead to avoidance of electric shock. Three criteria were employed: an easy condition in which even slow reaction times would be successful, a hard condition in which participants had to maintain or improve performance over time, and an impossible condition in which the reaction time criterion was too fast for most people. The impossible condition is comparable to uncontrollable stress, while the other two conditions represent easy and effortful behavioral control, respectively. The number of shocks administered was held constant across conditions. It was found that the greatest physiological activation was present not in the impossible or uncontrollable condition, but in the hard condition. This physiological reaction pattern was sustained by selective activation of β -adrenergically mediated sympathetic nervous system responses. Later studies have established that the enhanced physiological reactions associated with control under these circumstances are related to the effort expended, with greater effort being correlated with heightened norepinephrine responses.

Control and Fear

Lack of control over aversive stimuli potentiates fear responses. Using the yoked control design, it has been found that rats subsequently placed in a different environment and given two brief shocks so as to generate conditioned fear showed greater freezing responses if they had previously had no control over shock compared with those that had had control. These effects are long-lasting, but can be blocked by benzodiazepines. Fear responses associated with uncontrollable stress are again related to release of serotonin in the dorsal raphe nucleus and amygdala. Experiments using *in vivo* microdialysis have shown that uncontrollable stress leads to increased levels of extracellular serotonin in both these sites, while animals exposed to controllable stress show no such elevation.

Research on humans also suggests that control is relevant to fear responses. Fear of pain during medical and dental procedures is strongly associated with perceptions of uncontrollability. Fear of social threats (for example, when the individual is threatened with physical or sexual assault) is greater when people believe the situation is uncontrollable.

Control and Work

Control is important to understanding the physiological and health consequences of stressful events in human life. One area in which control has emerged as a key variable is in the study of work stress. Jobs vary on many dimensions of control. Some work is self-paced, while other tasks are externally paced by machines or other people. In some jobs, workers have choice about their posture, when they can take rest breaks, and how the work should be done. Flexibility in work hours, participation in decision making, and autonomy at work are all factors that vary across jobs. In psychophysiological studies, it has been found that cardiovascular and neuroendocrine activation is greater during externally paced than self-paced work, even when the actual pace of work is held constant. The same patterns are present in the real work environment. Figure 1 illustrates the pattern of ambulatory blood pressure measured every 20 min over the working day in middle-aged men and women divided into low and high job control on a standard questionnaire measure. Ambulatory

pressure was greater throughout the day in people working in low-control jobs after controlling statistically for covariates.

Several prospective epidemiological observational studies have evaluated the demand/control model of work stress. They have shown that low control at work, either alone or in combination with high demands, predicts future coronary heart disease in previously healthy adults independently of standard disease risk factors. It has also been found that control over work pace influences health during pregnancy, and that women with lower control jobs may be at greater risk for preeclampsia and low back pain, after statistical adjustment for age, parity, education, smoking, and type of work.

Conversely, greater autonomy and participation at work promote higher levels of self-reported job satisfaction, commitment to work, better performance, and reduced levels of emotional distress, staff turnover, and absenteeism. However, these studies of work illustrate a complication in human research in control, since it is clear that control cannot be seen solely as an objective characteristic of the environment, but depends on subjective perceptions.

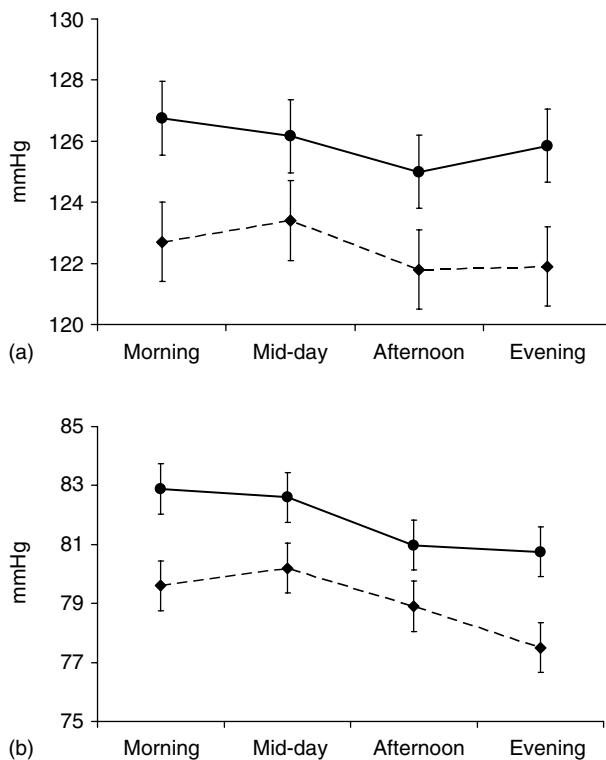


Figure 1 Mean systolic blood pressure (a) and diastolic pressure (b) in the morning, mid-day, afternoon, and evening periods in high (dashed line) and low (solid line) job control groups, adjusted for gender, employment grade, age, body mass index, smoking status, and physical activity. Error bars are standard errors of the mean. From Steptoe and Willemsen (2004). *Journal of Hypertension* 22, 915–920.

Perceived Control

Perceived control is the sense that one has control over aversive events. Objective differences in behavioral control are presumably mediated by perceived control. Thus, the benefits of having control over a particular aspect of work (such as being able to make a personal telephone call if necessary) will be apparent only if the worker realizes that this option is available. Perceived control is the result of a subjective judgment based on the individual's appraisal of the situation. It has two elements. The first is a perception that the situation is potentially controllable, and the second is that the individual can take the correct action. There may be circumstances in which events are perceived as controllable, but no action is taken since people lack (or believe they lack) appropriate behavioral competence. Psychologists such as Albert Bandura have postulated a distinction between outcome expectations and self-efficacy expectations to characterize these two aspects.

It is generally found that distress is associated with a perception that events are uncontrollable, and this phenomenon may operate independently of objective behavioral control. Thus, two individuals may experience the same event, such as the breakdown of an important personal relationship. One may perceive that the event was outside his or her control, due to the callousness of his or her partner, the circumstances in which they lived, or the predatory behavior

of a third party. Another person may perceive him- or herself to have been responsible; if the person had not been preoccupied with work or had responded better to the needs of his or her partner, then the relationship might not have floundered. One student who has failed an academic examination may attribute this to lack of preparation and effective revision, while another may blame arbitrary examiners or bad luck in the selection of questions.

Human beings seek to understand their experiences in life by developing explanations for aversive events and often construe events retrospectively as having been controllable (for example, "If I had not stayed later than I planned at that party, and had not decided to walk down that street, I would not have been assaulted"). Such cognitive attributions may be maladaptive in leading to self-blame for occurrences over which one has little control, but also may help to impose sense and order on a world that may otherwise appear cruel and arbitrary.

Sense of Control

Perception of control is not restricted to individual aversive experiences, but also relates to domains of life or even to life in general. The term locus of control is used to describe these generalized control beliefs. This construct was first introduced by Rotter, who argued that people varied in the extent to which they believed that important outcomes were determined by their own internal abilities and activities or by external factors. The idea of locus of control has since been elaborated upon in two important ways. First, locus of control may be domain specific, with different levels of belief in personal control over different aspects of life (health, work, personal relationships, etc.). Thus, people may feel a strong sense of control over their work, but believe that their living environment (traffic, difficult neighbors, etc.) is out of their own control. Second, it may be useful to understand locus of control as consisting of three dimensions: belief in internal control, chance, and powerful others. For example in the domain of health, people can be categorized according to the extent to which they believe that their health is determined by their own actions (internal control), that whether they remain healthy is a matter of fate or luck (belief in chance), and that their health is influenced by the competence of doctors and other health professionals (belief in powerful others). A dimensional perspective implies that people may simultaneously hold strong beliefs about more than one set of influences. They may believe, for instance, that they have quite a lot of control over their health, while at the same time believing that doctors are also very influential.

Sense of control is an individual belief pattern that also has developmental and social determinants. Animal studies indicate that monkeys provided with experiences of controllable (contingent) events early in life are subsequently less reactive to stressful events in adulthood. The experience of a major uncontrollable event in childhood, such as the death of one's mother, is associated with increased likelihood of depression following negative life events in adult life. Differences in sense of control may also relate to social class, with economically more deprived people facing greater external obstacles and fewer opportunities to influence events than the better off. One study assessed sense of mastery (e.g., agreement with the statement "Whether I am able to get what I want is in my own hands") and perceptions of constraints (e.g., "There is little I can do to change many of the important things in my life") in three large population surveys. It was found that sense of mastery was greater in higher than in lower income sectors of the sample, while perceived constraints were less. The financially better off also reported greater life satisfaction and less depression and rated their health as better. However, not all the lower income respondents had low mastery and high constraint ratings. Some individuals in the low-income sectors had high mastery ratings, and they reported greater life satisfaction, less depression, and health levels comparable with those of the high-income group.

Sense of control has cultural and political determinants as well. Surveys in Eastern Europe during the first decade after the collapse of communism identified a profound sense of lack of control over life, as people witnessed the destruction of familiar institutions and ways of life and experienced greater difficulty in fulfilling daily needs. Those years were characterized by a rapid deterioration in health and well-being in many sectors of the population of Russia and several Central-Eastern European countries.

Cognitive Control and Coping

Another closely related phenomenon is cognitive control, which can be defined as the belief that one can control one's reactions to events. It is a common characteristic to seek out information about events even when they cannot be controlled behaviorally. The person who has an abnormal result on a medical test and is scheduled for further examination may try to find out about the test and its consequences from acquaintances, the Internet, or other sources of information. The purpose of this information seeking is not to control the outcome, but to feel in control by reducing uncertainty and increasing the predictability

of the situation and to manage reactions to the result by anticipatory coping. This type of control, sometimes known as secondary control, is closely allied with the repertoire of other coping responses that people mobilize in an effort to regulate their responses to aversive experiences. It has important uses in medical settings, where the procedures for psychological preparation for surgery and stressful medical procedures have the effect of increasing cognitive control by providing information and helping patients mobilize a range of coping responses. In the self-regulation model of health developed by Leventhal, control is a key dimension in the cognitive representation of health threats, and variations in controllability influence effective adaptation.

People vary in the extent to which they exert effort to establish control over the situations with which they are confronted, and some individuals appear to have a high need for control. This may be maladaptive in some circumstances. Studies of work stress indicate that when high need for control is blocked by low extrinsic rewards (such as poor pay or promotion prospects), the individual may be placed at increased risk for coronary heart disease. People with a high need for control may find it particularly difficult to come to terms with experiences over which they have no control, such as the injury of a family member in a traffic accident.

Control and Well-Being

Control and Aging

Growing old is associated with a diminution of control. When people retire from work, their economic power is reduced, they have less authority than before, and their social networks shrink. Impaired physical health may increase people's sense of vulnerability, while they are exposed more often to uncontrollable events such as the death of contemporaries. A number of studies have documented the links between aging and the loss of control, though there are exceptions. Successful aging, or the maintenance of well-being and effective functioning despite chronological progression, is coupled with maintenance of beliefs in personal control. People may adapt by lowering their valuation of the importance of areas of life over which they lose control and by investing in social relationships more selectively than before. The impact of increasing sense of control and autonomy on the well-being of elderly population has also been investigated. Randomized trials of increased involvement in decision making in institutionalized elderly

groups have demonstrated improvements in mood and activity levels associated with enhanced control.

Control and Pain

The effects on pain of having control over the source of painful stimulation have been studied extensively. These experiments suggest that having behavioral control leads to a reduction in the subjective and physiological impact of the painful stimulus and may also increase pain tolerance. However, the benefits of control depend on the confidence that people have in their response efficacy, so perceptions are again important. It is interesting that the favorable effects of control over acute pain are observed despite the fact that stress-induced analgesia is elicited when there is a lack of control. It is likely that control allows behavioral and cognitive coping methods to be deployed that in turn reduce distress. Brain imaging studies have shown that when pain is perceived to be controllable, there is an attenuation of activation in the anterior cingulate, insular, and secondary somatosensory cortex, areas of the brain that are associated with pain processing. Interestingly, placebo analgesia results in alterations in brain activity in pain-sensitive brain regions, indicating that the control associated with beliefs leads to changes in the experience of pain.

Chronic clinical pain is often accompanied by feelings of helplessness, lack of control, and depression. Some authorities have argued that there may be a progression from feelings of personal helplessness in the early phases of a chronic pain problem to general helplessness (the belief that no one is able to control the pain) in later stages. The impact of chronic pain on well-being is that of a uncontrollable stressor, since by definition a pain that becomes chronic has not been alleviated and so has not been controlled. The generalization of distress about pain to disability and depression mirrors the phenomenon of learned helplessness.

The links between control and pain can be harnessed for patient care. As noted earlier, providing patients with control over sources of potential pain (for example, during dental examinations) can increase tolerance. Patient-controlled analgesia is used in a variety of medical settings with both adults and children and has benefits in terms of both emotional well-being and analgesic use. Several studies have found that when patients can control the level of analgesia, they use less medication than when it is provided by the physician. Rehabilitation programs for patients with chronic pain have a strong element

of increasing perceived control by providing patients with greater and more effective coping options.

Control and Illness

Most serious illnesses bring with them strong feelings of loss of control. When people have heart attacks or are diagnosed with cancer, they feel that they are no longer in control of their bodies, and their future plans are put in jeopardy so that they lose a sense of control over their destiny. Loss of control is often accompanied by other facets of helplessness, including hopelessness, distress, and passivity. Similar patterns occur in chronic health problems. For example, high levels of perceived control over asthma are associated with better self-management skills, reduced risk of hospitalization, and less restricted activity, independent of the severity of asthma itself. Some studies have found that recovery from serious illness such as stroke and major fractures is positively correlated with high internal control beliefs.

Many procedures for improving the emotional state of patients have the effect of enhancing perceived control. Thus, the methods used to prepare women for childbirth increase perceived control by providing a repertoire of coping response. Stress management procedures during cardiac rehabilitation enhanced perceived control and patient self-efficacy by increasing confidence in the resumption of normal activities. A particular area in which the issue of control has been contentious surrounds patient choice. To what extent should patients be involved in medical decision making? A number of studies indicate that although patients desire information about issues surrounding medical decisions, they sometimes prefer to delegate responsibility for the decision itself to the doctor. This should not be too surprising, since it is the typical pattern found in any society with specialist roles. While we might want to know all the facts, we would expect the car mechanic to decide how to fix an engine or the chef to use judgment and skill to prepare the complicated dish we have ordered. Patients often wish to be involved in decision making, particularly when there is a genuine choice of more than one effective alternative, but seldom wish to take authority completely away from the physician.

Conclusions

Control has emerged over recent years as a powerful explanatory variable in stress research. It is relevant

to a wide variety of settings, from basic experimental research on neuroendocrinology and neurochemistry to the experience of people in different socioeconomic sectors of society. There is a danger that such an attractive integrating notion will be misused and employed as a global explanation for completely unrelated phenomena. However, if care is taken not to confuse the different meanings of control, the construct has considerable value in understanding stress.

See Also the Following Articles

Childbirth and Stress; Coping and Stress: A Lens and Filter Model; Demand-Control Model; Life Events Scale; Surgery and Stress; Workplace Stress.

Further Reading

- Bandura, A. (1997). *Self-efficacy: the exercise of control*. New York: Freeman.
- Cockerham, W. C. (1999). *Health and social change in Russia and Eastern Europe*. New York: Routledge.
- Guadagnoli, F. and Ward, P. (1998). Patient participation in decision-making. *Social Science and Medicine* 47, 329–339.
- Katouzian, H. D., Cameron, L. D. and Leventhal, H. (eds.) (2002). *The self-regulation of health and illness behaviour*. Taylor & Francis.
- Maier, S. F. and Watkins, L. R. (2005). Stressor controllability and learned helplessness: the roles of the dorsal raphe nucleus, serotonin, and corticotropin-releasing factor. *Neuroscience and Biobehavioral Reviews* 29, 829–841.
- Marmot, M. G., Bosma, H., Hemingway, H., Brunner, E. and Stansfeld, S. (1997). Contribution of job control and other risk factors to social variations in coronary heart disease incidence. *Lancet* 350, 235–239.
- Rodin, J. (1986). Aging and health: effects of the sense of control. *Science* 233, 1271–1276.
- Skinner, E. A. (1996). A guide to constructs of control. *Journal of Personal and Social Psychology* 71, 549–570.
- Steenland, K., Fine, L., Belkic, K., et al. (2000). Research findings linking workplace factors to CVD outcomes. *Occupational Medicine* 15, 7–68.
- Stepptoe, A. and Appels, A. (eds.) (1989). *Stress, personal control and health*. Chichester, UK: Wiley.
- Stepptoe, A. and Willemsen, G. (2004). The influence of low job control on ambulatory blood pressure and perceived stress over the working day in men and women from the Whitehall II cohort. *Journal of Hypertension* 22, 915–920.
- Wager, T. D., Rilling, J. K., Smith, E. E., et al. (2004). Placebo-induced changes in fMRI in the anticipation and experience of pain. *Science* 303, 1162–1167.

Coping and Stress: A Lens and Filter Model

R H Rahe

University of Washington School of Medicine,
Seattle, WA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
R H Rahe, volume 1, pp 541–546, © 2000, Elsevier Inc.

Introduction

A Life Stress and Illness Model

Discussion

Introduction

Stressful life events can include challenges of high magnitude. However, many life change events occurring in people's lives are of low to moderate intensity and are part of their normal life experience. When do these events of lesser magnitude reach a level at which they can lead to illness? In addition, how is it that some individuals tolerate brutally traumatic life events with only temporary discomfort while others exposed to the same events react with severe and long-lasting disability? To help answer these questions, I have found that the optical lens and photographic filter model presented here can be helpful.

A Life Stress and Illness Model

An early activity that helped in the creation of the optical lens and photographic filter model was a scaling experiment to determine the relative significances of 42 typical life change events. This allowed for quantitative estimates to be made of a person's recent life change events during an arbitrary period of time. Recent life change magnitudes found for large samples of individuals were seen to correlate positively and significantly with their near-future (the following 6 months to 1 year) illness experiences. The original list of events was enlarged in 1975 and rescaled in 1978 and 1995. When comparing the average life change magnitude values over 30 years (1965–1995) the average increase in magnitude was 45%. Thus, life change stress appeared to increase at a rate of 1.5% per year over this 30-year interval.

A model was created using optical lenses and photographic filters to display individuals' exposures to stressful life change events, their perceptions of the significances of these events, their psychological and physiological reactions to these perceptions, their

coping with these reactions, and the possible subsequent illness behaviors leading to diagnosed illnesses. The full six-step model is shown in [Figure 1](#).

Step 1: Perception

In [Figures 1 and 2](#), individuals' recent life change exposure is represented by a series of light rays (lines). The thickness of these light rays indicates the differing life change magnitudes of the events. Thick dark rays represent highly significant events, such as marriage, divorce, and death of a parent. Thinner rays signify less meaningful events, such as a residential move or a purchase of an average-priced automobile.

The influence of perception in altering people's evaluations of their recent life change events is shown in step 1 using a polarizing filter. Just as a polarizing filter diminishes the intensities of some light rays and augments others, the perception of the significance of the recent events may be diminished or augmented according to people's past experiences with similar events and their current social supports. Early life influences, such as intelligence, educational, and social class, can also lead to alterations in their perceptions of life change events. For example, a poorly educated, financially challenged person may perceive the stress of losing his or her job as substantially more of a life change than a highly educated, financially secure individual who has the time and resources to find other employment.

Step 2: Psychological Defenses

Once challenged by significant recent life change events, people quickly employ a variety of psychological defense mechanisms, such as denial, displacement, repression, and reaction formation in an unconscious, almost reflexive, manner. These defenses tend to diffract away the potential impact of these perceived life events on people's psychophysiology. A negative optical lens shows this diffraction for step 2 in the model ([Figure 3](#)).

Some individuals employ their psychological defenses so strongly, and so persistently, that they show nearly complete insulation from the expected psychophysiological responses (in step 3). For example, it has been reported that some patients with myocardial infarction who have been hospitalized in coronary care units are in denial to such a degree that they convince themselves they never had a heart attack. Such strongly defended people have even been shown to develop fewer cardiac arrhythmias and to experience better survival rates than

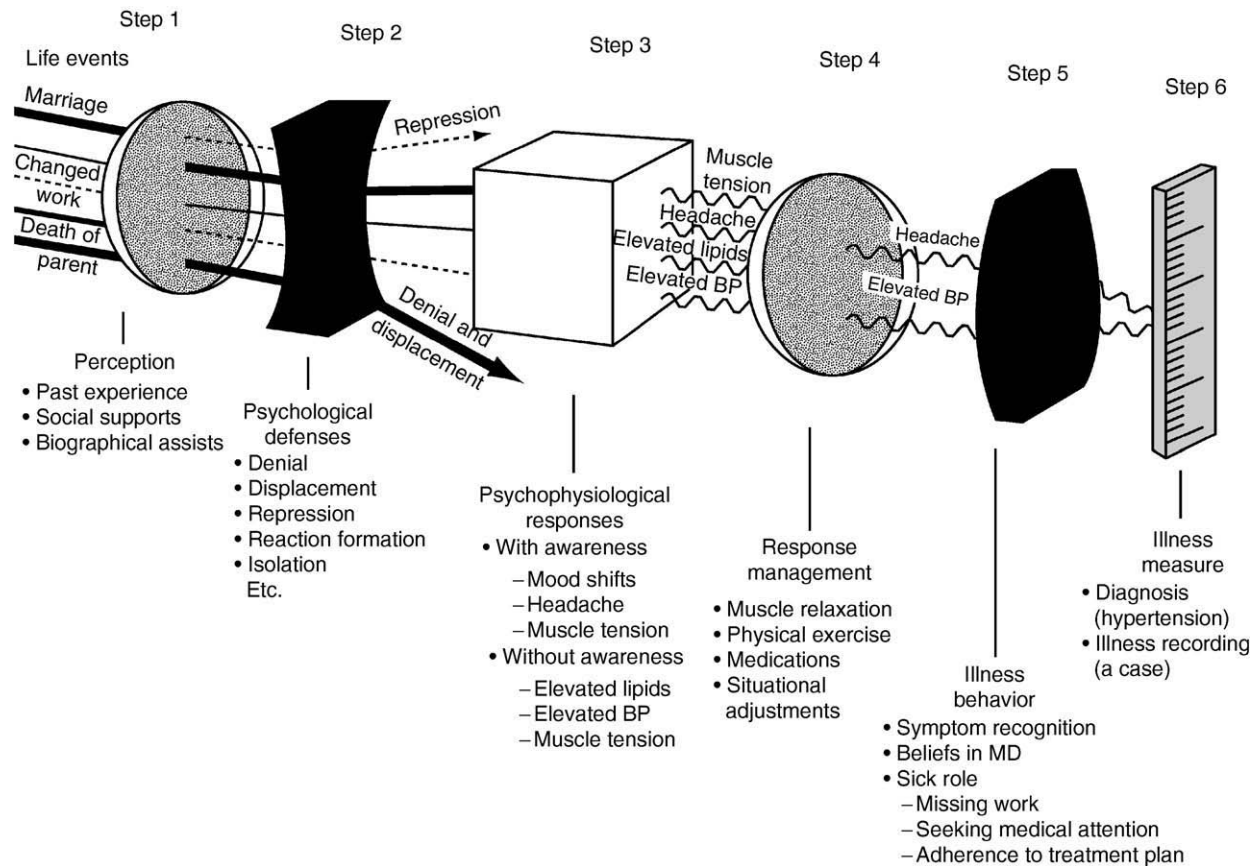


Figure 1 The lens and filter model.

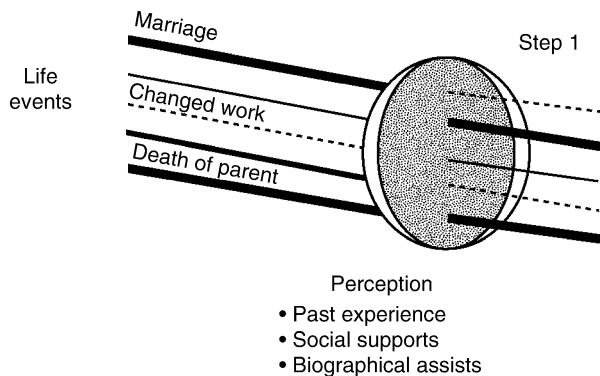


Figure 2 Step 1.

do more realistically oriented patients in the same coronary care unit.

Although psychological defenses can be extremely helpful in the short term, they often prove to be maladaptive in the long run. To use the coronary patients example once again, their use of one or more psychological defenses may help them survive the first few days after infarction, but these patients

may not eventually learn to deal with the realities of this illness. They may not stop smoking, reduce their body weight, begin a physical fitness program, or learn to control their typical unhelpful psychological responses to life stress.

Step 3: Psychophysiological Responses

The box in Step 3 represents the body's internal milieu where various psycho-physiological processes take place. These responses can be divided into two categories: responses of which people are aware, such as sweating, pain, or muscle tension; and responses that occur outside people's awareness, such as elevated serum cortisol or a rise in blood pressure (BP; see Figure 3).

An example of incompletely defended life stress events might be found in people who have cancer. Such people may experience the psychological responses to recent life stress of helplessness and hopelessness. These psychological responses have been shown to lead to reduced immunological competence, facilitating an exacerbation of this disease.

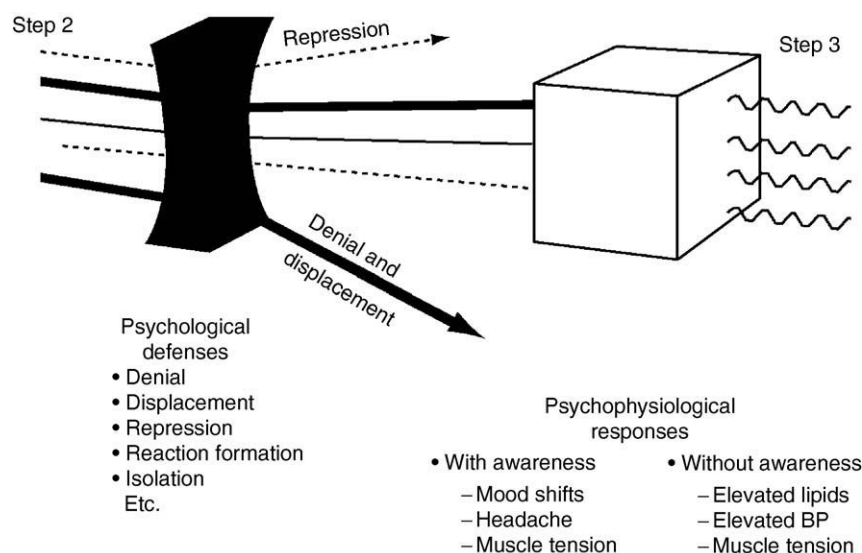


Figure 3 Steps 2 and 3.

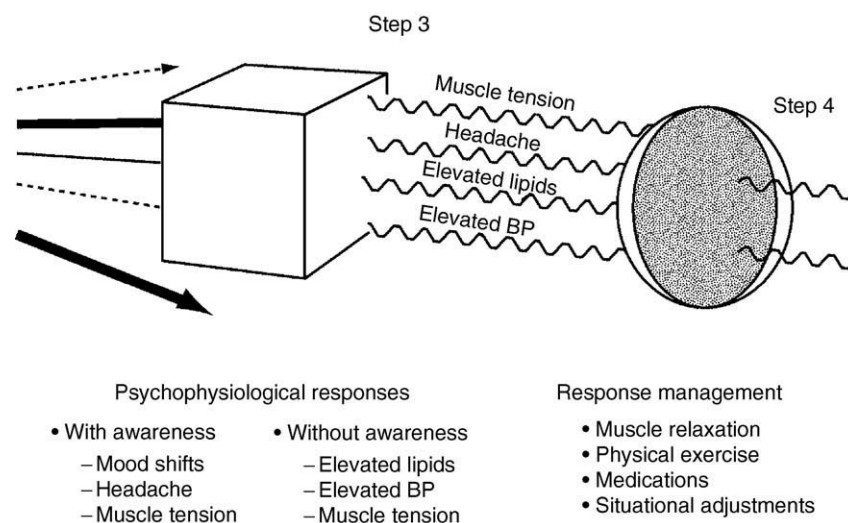


Figure 4 Steps 3 and 4.

Step 4: Response Management

Once individuals become aware of their psychophysiological responses to recent life stress, and especially if these responses are seen as worrisome, these responses are labeled symptoms. People may then decide to employ one or several response management (coping) techniques. For example, individuals may try muscle relaxation, meditation, exercise, or start taking various medications. A color filter is used here in the model to indicate that successful coping can absorb some of the psychophysiological reactions (symptoms) in much the same manner that a color

filter absorbs light rays of certain frequencies (**Figure 4**). Coping in this model is defined as consciously regulated practices and behaviors designed to reduce and, it is hoped, to eliminate symptoms.

Coping, or response management, is a potent resource for dealing with environmental challenges. Because defenses are of short-term effectiveness, coping can achieve more lasting results. To use the cancer patients example once again, after experiencing feelings of helplessness and hopelessness, they may institute behaviors such as seeking medical care, using social supports, and developing their spirituality.

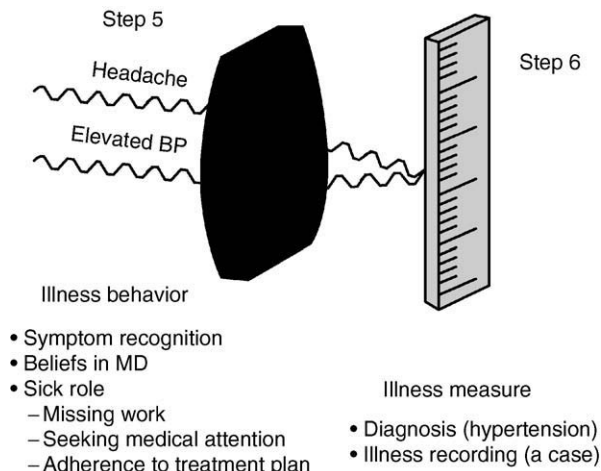


Figure 5 Steps 5 and 6.

These coping activities may help to reduce, or even eliminate, their depressive symptoms. In addition, this coping may also counteract some of the psychophysiological activations that promoted the expression of their cancer in the first place.

Step 5: Illness Behavior

When people's abilities to cope with their environmental challenges are insufficient and their attendant psychophysiological arousal cannot be absorbed, body symptoms are likely to be sustained, leading to tissue breakdown, organ system dysfunction, and disease. A positive lens is pictured in [Figure 5](#) to represent illness behavior. As symptoms of headache and elevated BP continue in the figure, the affected individuals may then decide to seek medical care. It is clear from epidemiological studies that at least as many people with sustained symptoms secondary to poorly managed life stress fail to seek medical care as those who do. Thus, doctors see only a portion of people experiencing mind and body effects of unresolved life stress leading to bodily symptoms and disease.

Step 6: Illness Measure

The final steps in the model ([Figure 5](#)) represents people with long-lasting symptoms progressing to the point of illness who then seek medical consultation and receive one or more diagnoses. The patient's illnesses and their treatments are usually documented subsequently in a medical record.

Many stress and illness studies use medical records as a criterion for defining a case. Thus, people exposed to a significant recent life stress, such as surviving a flood or an earthquake, would only be part

of such a stress and illness study if they followed the entire path displayed in [Figure 1](#) and became a case. People who defended well against and/or coped successfully with the trauma and those who became symptomatic but never reported for medical care would be missed.

Discussion

According to the lens and filter model, people who manage to remain healthy following severe life stress do so by successfully managing steps 1–4. Those who do not handle these four steps effectively are likely to be the ones who develop lasting symptomatology and disease (steps 5 and 6).

Perception, psychological defenses, and coping skills seem to develop early in life. It is extremely beneficial to be endowed with an abundance of biological (genetic) and biographical (acquired) assets. Returned prisoners of war and political hostages who generally stayed healthy during and following captivity often appear to have been born and bred to handle stress well. That is, they had a high number of biological and biographical assets, such as high intelligence, good parenting, ample social skills, and perseverance in adversity, to assist them in managing their perceptions of stress, appropriate use of defenses, and successful coping. Conversely, others who became clinically ill during and following captivity tended to display more negative perceptions of stress, exaggerated and/or prolonged defenses, and restricted coping behaviors – both early in their lives and during and following capture.

Recent application of this model has been found in military men and women returning home from duty in Afghanistan and Iraq. The effective approach to recovery from combat stress is to not tell their war experiences over and over (see **Combat, Acute Reactions to; Combat Reaction, Chronic**). Rather, they should begin to understand the power of their biological and biographical assets and liabilities on their recent stress experiences, see how current life stresses (which they can do something about) add to their current symptoms, and evaluate the effectiveness of their current coping capabilities. Often, an answer to their physical and psychological discomforts can be found in the expansion of their coping repertoires.

See Also the Following Articles

Acute Stress Response: Experimental; Acute Trauma Response; Combat, Acute Reactions to; Combat Reaction, Chronic.

Further Reading

- Cassileth, B. R., Lusk, E. J. and Hutter, R. (1984). Concordance of depression and anxiety in patients with cancer. *Psychological Reports* 54, 588–590.
- Croog, S. H., Shapiro, D. S. and Levine, S. (1971). Denial among male heart patients. *Psychosomatic Medicine* 33, 385–397.
- Fawzy, F. L., Kemeny, M. E. and Fawzy, N. W. (1990). A structured psychiatric intervention for cancer patients. II: Changes over time in immunological measures. *Archives of General Psychiatry* 47, 729–735.
- Hackett, T. P. (1978). The use of groups in the rehabilitation of the post-coronary patient. *Advances in Cardiology* 24, 127–135.
- Holmes, T. H. and Rahe, R. H. (1967). The social readjustment rating scale. *Journal of Psychosomatic Research* 11, 213–218.
- Rahe, R. H. (1969). Life crisis and health change. In: May, P. R. A. & Wittenbom, R. (eds.) *Psychotropic drug response: advances in prediction*, pp. 92–125. Springfield, IL: Charles C Thomas.
- Rahe, R. H. (1988). Anxiety and physical illness. *Journal of Clinical Psychiatry* 49(10, supplement), 26–29.
- Rahe, R. H. (1990). Life change, stress responsivity, and captivity research. *Psychosomatic Medicine* 52, 373–396.
- Rahe, R. H. and Arthur, R. J. (1978). Life change and illness studies: past history and future directions. *Journal of Human Stress* 4, 3–13.
- Rahe, R. H., Karson, S., Howard, N. S., et al. (1990). Psychological and physiological assessments on American hostages freed from captivity in Iran. *Psychosomatic Medicine* 52, 1–16.
- Rahe, R. H., Ward, H. W. and Hayes, V. (1979). Brief group therapy in myocardial infarction rehabilitation: three- to four-year follow-up of a controlled trial. *Psychosomatic Medicine* 41, 229–242.

Coping Skills

A DeLongis and E Puterman

University of British Columbia, Vancouver, BC, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A DeLongis and M Preece, volume 1, pp 532–540, © 2000, Elsevier Inc.

Stress and Coping

Chronic Illness

Structured Coping Programs: Learning How to Cope Effectively

Summary

Glossary

<i>Cognitive reappraisal</i>	A form of coping that focuses on changing one's attitudes and beliefs toward a stressful situation.
<i>Coping</i>	Cognitive and behavioral efforts to manage stress.
<i>Emotion-focused coping</i>	Attempts to manage one's emotions during stressful periods.
<i>Problem-focused coping</i>	Attempts to directly change a stressful situation.
<i>Relationship-focused coping</i>	Attempts to manage one's social relationships during stressful periods.

Coping describes cognitive and behavioral responses to a stressful situation. These responses are determined by the interaction between the characteristics of the person and characteristics of the situation. When these responses are particularly effective in reducing strain, due to the expertise with which they are selected and employed, we could say these responses are particularly skillful. When an individual experiences a stressor that threatens to overwhelm current personal resources, such as a chronic illness, an intervention that emphasizes coping skills training can be helpful. However, there is no one coping skill that, once learned, becomes a magic bullet, effective with all kinds of stress. An important part of being a skilled copier is being able to match a stressful situation with the most effective strategy that is appropriate for the given context. Given this, flexibility in coping appears to be a key aspect of effective stress management.

Stress and Coping

Everyone experiences stress at times during his or her life. However, what is extremely stressful to one individual may not be at all stressful to another. How a situation is evaluated, or appraised, is an important determinant of the degree of stress experienced by the individual. Coping is loosely defined as the things we

think and actions we take to ameliorate the negative aspects of a stressful situation. How we manage stress is critical because overwhelming stress can cause psychological distress as well as have both short- and long-term negative consequences for physical health. The ability to cope successfully with a stressful situation depends on a number of factors. Primary among these are the resources one brings to the stressful situation. These resources include one's personality, age, financial assets, education, previous experiences, social support, and physical and mental health. Both resources and subjective appraisals are important determinants of the degree of stress experienced by the individual. If a situation is not appraised as taxing one's resources or ability to cope, it may not be experienced as stress at all. Features of the situation, such as the desirability, controllability, and severity of the stressor, are important in shaping coping responses. Certainly there are many types of stressors that are a tremendous strain for almost any individual.

We review here ways of coping with stress. Reviewing the entire literature on effective and maladaptive coping across a multitude of stressors is beyond the scope of this article. Here we focus primarily on chronic illness, which has been the center of a great deal of research. At some point in their lives, all individuals must cope with chronic illness, either personally or through a significant other. Finally, we review evidence-based coping skills interventions.

Ways of Coping

Problem-focused coping Problem solving is a strategy that quickly comes to mind as an effective way of attempting to remedy a negative situation. Defining the problem, generating alternative solutions, comparing these alternatives in terms of their likely costs and benefits, selecting a likely solution, coming up with a plan, and then acting on it are steps that most of us take when we determine a situation to be one that we can do something about. These strategies primarily involve changing aspects of the situation. Other problem-focused coping strategies may be geared toward changing ourselves, such as learning new skills and procedures, thereby increasing one's coping resources.

Many problem-focused coping strategies that can be enumerated as applicable in specific situations do not generalize broadly across all types of stressors. For example, problem-focused strategies used to cope with chronic pain may not be useful for coping with a pressing deadline at work. Even in situations in which problem-focused strategies could be used to great effect, if the individual perceives the level of threat to be very high, it may be impossible for such

strategies to be employed effectively. Excessive threat has been found to have negative effects on cognitive functioning and on the capacity for information processing. A common example of this often occurs in a physician's office, when patients are given bad news about their health. At such times, patients may have difficulty comprehending any additional information about treatment and prognosis that the physician may wish to convey. This phenomenon is due to a reduction in cognitive functioning brought about by the patient's experience of feelings of threat or harm.

Another important determinant of problem-focused coping is the individual's perception of the situation as one that is amenable to change. In a study of middle-age people coping with stressful encounters, Lazarus and colleagues found that people adapted better to stress when they used more problem-focused coping for an encounter that the person felt could be manipulated. However, if the encounter was judged as one requiring acceptance, more emphasis was put on emotion-focused coping.

Emotion-focused coping The primary goal of emotion-focused coping is to lessen emotional distress. Some of the ways this may be achieved are through avoidance, distancing, and wishful thinking. Another approach may be to minimize the threat of the situation, for example, via cognitive reappraisal. Cognitive reappraisals involve changing the meaning of the situation without changing it objectively. For example, a stressed person may focus on how much worse things could be, perhaps by comparing him- or herself to others less fortunate. Noting how much better one's situation is than the situation of others is called "engaging in downward social comparisons." There are also emotion-focused strategies directed toward diverting attention from the problem, such as engaging in physical exercise, meditating, having a drink, expressing one's anger, or seeking emotional support. Similar to problem-focused coping, a coping response that is helpful in one situation may be harmful in another. For example, denying the severity of chronic heart disease may be beneficial right after diagnosis to alleviate anxiety and anger, but it is problematic during rehabilitation.

Relationship-focused coping Recently, attention has also been directed toward the ways that people attempt to maintain their important social relationships during periods of stress. People in enduring relationships cannot concern themselves only with the situation they are facing, but must also consider the implications for their partner and for the relationship.

This focus on relationships during stressful episodes may be particularly important when a loved one has a chronic illness or disability. For example, O'Brien and DeLongis pointed out the usefulness of empathic coping by caregivers of Alzheimer's patients. Other types of relationship-focused coping involve providing support and attempting to compromise when there is a disagreement. Not all attempts to manage relationships during stressful periods have positive effects. Coyne has found that emotional overinvolvement can have serious negative consequences on the health of family members.

Although some people are more skilled at coping with stress than are others, it is impossible to identify a set of coping strategies that can be called "good" ways of coping. The context in which the stressful situation occurs, the type of problem, the other people involved, and the personality characteristics of the individual are only some of the factors that affect how to best deal with the situation. Further, two people might use identical strategies with different degrees of success, depending on how skillfully the strategy is implemented. However, researchers have pinpointed some general ways of coping that appear to be either adaptive or maladaptive in particular types of situations. A major type of stressor is chronic illness, but differing types of illnesses require different coping skills.

Chronic Illness

Due to medical and technological advances, acute illnesses have given way to increases in chronic illnesses as the primary cause of death in advanced countries. Given the improvements in treatment of chronic illnesses, people who adhere to medical advice can often live relatively normal lives for a long time after diagnosis. It is therefore important to understand how successful adherence and coping with illness can aid in adaptation.

Adherence

Ajzen and Fishbein identified two factors that predict someone's intention to adhere to medical recommendations for a chronic illness, in particular, and to any health behavior, in general. First, a person's attitudes toward the behavior, determined by the extent the person values the outcome of the behavior, predict the person's intentions. Second, this intention is predicted by the perception that important people in the person's life value the behavior and the extent he or she cares about what these people think. Although these factors predict intention well, they fare less well in predicting actual adherence to medical regimens and recommendations for adaptation to chronic

illness. People may intend to perform a health behavior but might not if they are not able or ready.

Adhering to health recommendations is often strikingly low, particularly for lifestyle-change recommendations associated with chronic illness. Rates of adherence are influenced by the experience of side effects, treatment complexity (e.g., one pill a day at dinner versus a handful at different times per day, as in some cases of HIV treatment), the quality and quantity of support received from both the health professionals and important others, and cultural norms. Furthermore, some people avoid recommendations from health professionals and do not adhere to a prescribed regimen, and still others may deny that they are at risk as a way of coping with the stress associated with being ill. These methods of coping may have negative consequences for illness progression. Other people may actively cope by strictly sticking to the medical regimen.

Coping with Chronic Illnesses

When confronted with illness, people's physical and psychological worlds often become disrupted. People with chronic illness face potential mental and physical health problems and disruptions in social functioning. Those with chronic illnesses are challenged by having to alter their self-perception, to integrate the concept of being chronically ill into their lives, and to find meaning in their illness and lives. People who find positive meaning in their illnesses and accept the physical, mental, and social changes that occur with a chronic illness adjust better to illnesses than those who do not. In the following sections, we review four major health problems that impact individuals and thus receive increased attention and research: cancer, chronic heart disease, HIV disease, and chronic pain.

Coping with cancer The question of whether psychological factors, such as coping strategies, affect cancer development has intrigued researchers for years. Some research has suggested that a "fighting spirit," or coping by fighting angrily against the diagnosis, may be associated with living longer with cancer, whereas passively accepting ("stoic acceptance") of the disease and feeling hopeless may predict poorer courses. Under some circumstances, people who minimize or deny the illness's impact on their lives and continue on with daily activities also have longer survival rates. In contrast, repressing negative emotions may be linked with speedier cancer progression. Although the jury is out on whether social support and coping effect initiation, disease course, and mortality rates among cancer patients, research clearly

indicates that how people psychologically cope with the illness and whether support is received affect the psychological well-being, adjustment, and quality of life of cancer patients. The benefits of psychological interventions are apparent in the emotional lives of cancer patients, even though claims of extending patient's lives may be premature.

Coping with Chronic Heart Disease

Reacting to stress with anxiety and anger has been linked to the development of chronic heart disease (CHD), especially in men. In addition, during the first stage of recovery following a heart attack, coping responses that involve denial and that may serve to reduce anxiety and anger toward the illness appear to be more beneficial than are more active coping strategies. During the first stage of the disease, supportive interventions are often most useful. However, after a patient's discharge from a coronary care unit, more active coping strategies are important. For example, interventions that focus on stress management, as well as educational interventions that emphasize lifestyle changes and compliance with medical advice are important in increasing both health and well-being. Good nutrition and exercise are critical factors in preventing CHD as well as in recovery from a heart attack, and many preventive interventions are focused on changing these health-related behaviors.

Coping with human immunodeficiency virus disease The course of human immunodeficiency virus (HIV) disease often begins with an acute, flulike illness, but then remits into an asymptomatic stage that could last between 2 and 10 years. People with HIV may then cycle among (1) chronic infections, (2) symptoms that lead to a diagnosis of acquired immune deficiency syndrome (AIDS), and (3) back into an asymptomatic stage, depending on the effectiveness of the antiretroviral medications and other factors. The risk of continuously cycling through these stages and the fears associated with them may cause severe distress for those affected by HIV.

When people learn of a positive HIV antibody test result, they usually react with increased psychological distress. Blaming oneself for the illness, denying the reality that one has contracted and is living with HIV, and holding onto false hope have been associated with poor psychological adjustment to HIV disease and lower levels of CD4 cells, markers of immune functioning. On the other hand, those who find meaning and who create positive experiences in their lives tend to fare better psychologically. Finally, active coping strategies that promote taking care of one's

health, getting information about the illness, seeking support, and adhering to medication have all been associated with better psychological and physical outcomes for people with HIV.

Psychological interventions for people with HIV disease range from encouraging people to seek out information regarding their illness to helping them weigh the pros and cons of disclosure and drug adherence. Therapists can also help patients plan their drug-adherence schedules and teach them coping skills to deal with potential side effects. As well, psychological interventions may promote psychological adaptation to the illness and help people find meaning in their experience with HIV.

Coping with chronic pain Pain is considered chronic when the symptoms have lasted for a period longer than 6 months. Common types of chronic pain are lower back pain; joint pain as a result of rheumatoid arthritis or osteoarthritis; headaches; pain resulting from chronic illnesses, such as HIV disease and cancer; and various myofascial pain syndromes, usually associated with diffuse musculoskeletal pain. Research has demonstrated that psychosocial factors, such as social support, coping, and patients' beliefs, predict approximately half of the variation of long-term disability resulting from chronic pain.

Most coping efforts are directed toward two main goals: (1) prevention or reduction of pain and (2) maximization of quality of life despite pain. A host of factors, such as the nature of the pain, the expectation of relief, and the specific situation, influences the effectiveness of a particular coping strategy. However, there are consistent findings that point to a relationship between particular classes of coping strategies and the degree of impairment.

Catastrophizing is a cognitive process that includes negative self-statements and excessively negative beliefs about the future. Pain patients may express such things to themselves or to others, saying, "This pain will never go away" or "I can't cope with this pain anymore." Catastrophizing and avoiding activities related to illness-specific exercises and social-leisure activities have been found to be associated with poor outcomes. The diversity of these outcomes is remarkable: increased pain and disability, greater use of the health care system and medication, limited activities, and depressive symptoms in patients with chronic pain.

Probably the most adaptive method of coping with chronic pain involves behaviors or cognitions that divert the person's attention from the pain and toward some more positive experience. Distracting oneself with other activities such as socializing,

reading, and painting reduces emphasis on the experience of pain and has beneficial effects on mood as well. Activities requiring a focus of attention tend to reduce the activation of the sympathetic nervous system, leading to further reductions in pain intensity. Guided relaxation, or transformational imagery techniques, in addition to reducing pain, can also serve to increase feelings of self-efficacy.

Structured Coping Programs: Learning How to Cope Effectively

Coping interventions often focus on learning to appraise stressful contexts more adaptively and then to cope more effectively with the demands of these contexts. In other words, reducing the experience of stressors as threatening and adjusting the sense of control people have over the stressors are key to adaptive coping. We review here three evidence-based coping skills interventions designed for patients with chronic illnesses: coping effectiveness training (CET), cognitive-behavioral therapy (CBT), and mindfulness-based therapies.

Coping Effectiveness Training Program

Folkman, Chesney, and colleagues developed a group-based coping intervention at the Center for AIDS Prevention Studies at the University of California at San Francisco. This intervention follows closely Lazarus's model of stress and coping. The intervention focuses on helping individuals to change their appraisals, improve their coping skills, and obtain beneficial social support when needed. Although this intervention has been tested with HIV-positive men, it is designed to apply to a broad range of stressful contexts and has been applied to various chronic problems, such as spinal cord injury.

First, participants learn to distinguish between global and specific stressful situations. It is much easier to feel in control of a specific stressor (e.g., fears associated with receiving HIV antibody test results) than a vague, complex, global condition (e.g., coping with AIDS more generally). Second, participants learn to distinguish between those aspects of specific stressors that are changeable (e.g., finding ways to spend time while waiting for lab results) and those that are unchangeable (e.g., the lab results per se).

Emotion- and problem-focused coping strategies are then evaluated as possible methods of coping with the unchangeable and changeable situations, respectively. Strategies developed and practiced include cognitively reframing the situation as less stressful, selectively attending to positive aspects and distancing oneself from distressing aspects of the situation,

asking for information, and using humor to diminish the negative impact of the situation. They also consider maladaptive forms of emotion-focused coping, such as smoking, drug use, and overeating, and the resulting negative impact associated with these strategies. Various decision-making strategies are taught, including ways to evaluate the probable outcomes of different approaches to a problem. Participants are also trained in communication skills, such as listening, communicating, and negotiating.

CET also has a component to help people obtain the social support they need from others. Support is considered to be informational, emotional, or tangible, and participants are taught to identify those members of their social network who are most able to provide the specific type of support required and how to best enlist their aid.

Cognitive-Behavioral Therapy

Many people who have a chronic illness or psychological problem see themselves as incapable of changing their situation. CBT directly challenges people to alter the way they think about themselves, the way they behave, and the way they deal with problems and their illnesses. The theoretical underpinning behind these therapies is that cognitions, behaviors, and emotions are intricately interconnected and that altering one affects and alters the other two.

CBT is a structured, evidence-based intervention that combines two effective forms of therapy: behavioral and cognitive therapy. Treatment usually begins with a behavioral component, breaking long-term goals or problems into small steps that can easily be taken by an individual. For example, a person with debilitating chronic back pain who wants to be able to work again may be encouraged to break down the goal into a shorter-term goal, such as being able to go for a 10-min walk again. In such a way, the person is encouraged to drop the overwhelmingly large long-term goal for simpler smaller goals that are manageable and attainable. Asking a patient to engage in activities, such as exercise and social activities, has the added benefit of getting the person active again and challenges the person's thinking that he or she cannot do anything. Accordingly, the person's self-efficacy can be slowly and incrementally heightened.

The cognitive component of CBT aids patients in coming to understand how their patterns of thinking might cause psychological problems. Maladaptive ways of experiencing oneself and situations can cause and enhance the psychological problems associated with mental or physical illness. Therapy and homework are designed to heighten the awareness of behaviors, feelings, and thoughts associated with specific

stressors, situations, or conditions. Then, the person is encouraged to evaluate the pros and cons of his or her attitudes and beliefs regarding the stressor and to offer positive interpretations or cognitive reappraisals of situations with the intention of challenging the negative thought pattern. The patient, therefore, learns to actively seek new ways of perceiving or appraising situations that are less damaging to their psychological and physical well-being.

Mindfulness-Based Therapies

Over the past 2 decades, we have seen growing interest in incorporating mindfulness-based approaches into therapies for such conditions as chronic pain, stress, depression relapses, and CHD, to name a few. The goals of these therapies are to teach people how to manage everyday stress and pain and take responsibility for their well-being. Mindfulness involves moment-by-moment awareness of what a person is experiencing, such as paying close attention to breathing, noises, sensations in the body, inner feelings and thoughts, and our reactions to specific situations. Mindful awareness involves meditation exercises and is often coupled with stretching to help attain a mindful state. The person observes the moment-by-moment sensations, cognitions, and emotions without judging whether they are bad or good, important or trivial, or sick or healthy (nonjudgmental observation). The person is encouraged to recognize and accept his or her limits and body, as opposed to pushing and pulling at his or her physical and psychological limits. The goal is for the person to experience a new relationship with the mental or physical condition instead of doing the same old catastrophizing, avoiding, or other maladaptive coping strategy.

Mindfulness-based approaches have been applied to various physical and emotional problems and have received empirical support. Questions have been raised, however, regarding the methodology employed in studies designed to evaluate the therapies. Others consider mindfulness-based therapies as an alternative to more traditional approaches that involve learning skills that directly manipulate the stressor, negative thoughts, or feelings.

Summary

There are a number of factors that determine whether an individual can cope effectively with a particular stressor. An important first step is to appraise the situation in a positive yet realistic manner. Further, breaking up a global stressor into those aspects that can be changed and those that must be accepted can facilitate the identification of potential coping

strategies. Practicing new coping strategies, considering the possible consequences of several different approaches, and obtaining support from others can result in a reduction in the negative consequences of stress. For those situations that threaten to overwhelm an individual's current resources, coping interventions that provide support, information, and skills training can be effective.

Acknowledgments

Anita DeLongis's research is supported by the Social Science and Humanities Research Council of Canada. Eli Puterman's research is supported by the Social Science and Humanities Research Council of Canada and the Michael Smith Foundation for Health Research in British Columbia.

See Also the Following Articles

Behavior Therapy; Cancer; Cardiovascular System and Stress; Caregivers, Stress and; Health Behavior and Stress; HIV Infection/AIDS; Pain; Problem-Solving Skills Training; Psychological Stressors, Overview; Psychosocial Factors and Stress; Relaxation Techniques.

Further Reading

- Ajzen, I. and Fishbein, M. (2000). The prediction of behavior from attitudinal and normative variables. In: Higgins, E. T. & Kruglanski, A. W. (eds.) *Motivational science: social and personality perspectives*, pp. 177–190. New York: Psychology Press.
- Baer, R. A. (2003). Mindfulness training as a clinical intervention: a conceptual and empirical review. *Clinical Psychology: Science and Practice* 10, 125–143.
- Baum, A., Revenson, T. A. and Singer, J. E. (eds.) (2001). *Handbook of health psychology*. Mahwah, NJ: Lawrence Erlbaum.
- Chesney, M., Chambers, A., Taylor, D. B., et al. (2003). Coping effectiveness training for men living with HIV: results from a randomized clinical trial testing a group-based intervention. *Psychosomatic Medicine* 65, 1038–1046.
- Craig, K. D. (2005). Emotions and psychobiology. In: McMahon, S. & Koltzenburg, M. (eds.) *Melzack and Wall's Textbook of pain* (5th edn.). pp. 231–239. Edinburgh: Elsevier/Churchill Livingstone.
- DeLongis, A. and Holtzman, S. (2005). Coping in context: the role of personality and social relationships. *Special issue on daily process methodology. Journal of Personality* 73, 1–24.
- Folkman, S., Lazarus, R. S., Dunkel-Schetter, C., et al. (2000). The dynamics of a stressful encounter. In: Higgins, E. T. & Kruglanski, A. W. (eds.) *Motivational science: social and personality perspectives*, pp. 111–127. Philadelphia: Psychology Press.

Garssen, B. (2004). Psychological factors and cancer development: evidence after 30 years of research. *Clinical Psychology Reviews* 24, 315–338.

Revenson, T. A., Kayser, K. and Bodenmann, G. (eds.) (2005). *Couples coping with stress: emerging perspectives on dyadic coping*. Washington, DC: American Psychological Association.

Teasdale, J. D. (2004). Mindfulness-based cognitive therapy. In: Yiend, J. (ed.) *Cognition, emotion and psychopathology: theoretical, empirical and clinical directions*, pp. 270–289. New York: Cambridge University Press.

Turk, D. C. and Okifuji, A. (2002). Psychological factors in chronic pain: evolution and revolution. *Journal of Consulting and Clinical Psychology* 70, 678–690.

Coronary Heart Disease See: Atherosclerosis; Heart Disease/Attack.

Corticosteroid Receptor Genes: Functional Dissection in Mice

F Tronche

CNRS, Collège de France, Paris, France

© 2007 Elsevier Inc. All rights reserved.

Classic Inactivation of Glucocorticoid Receptor and Mineralocorticoid Receptor Genes

Tissue-Specific Corticosteroid Receptor Genes Inactivation
Modifying Glucocorticoid Receptor and Mineralocorticoid Receptor Gene Dosage in an Organism or Cell Type

Glossary

Additional transgenesis

Adding an extra gene to the mouse genome; often used for the random insertion of a synthetic gene into the mouse genome by injection into a monocellular zygote followed by the selection of a mouse colony carrying the desired gene.

Corticosteroid receptor genes

Genes encoding the glucocorticoid and mineralocorticoid receptors. These receptors are the intracellular receptors for corticosteroid hormones and function as transcription factors in the cell nucleus, controlling the expression of target genes. These genes are similar to those of the receptors for other steroid hormones: androgens, progesterone, and estrogens. All these genes are derived from a common ancestor, and they form the steroid receptor gene subfamily of the nuclear receptor gene family.

Gene targeting

A technique for the modification of a given chromosomal DNA sequence by

Molecular genetics

homologous recombination in cultured mouse embryonic stem cells. Live animals carrying the mutation can be derived from this modified cell line.

The field of biology concerned with studies of the structure and function of genes at the molecular level. The development of reverse genetic tools, facilitating the introduction of the desired modification into the mouse genome has made it possible to investigate the consequences of targeted mutations at the level of a mammal organism.

Null allele

Often used to define a gene allele that cannot lead to the synthesis of gene product, synonym of (–) or knock-out. A major difficulty of gene targeting in mice is to ensure that the gene is indeed invalidated and that no wild-type or modified protein is produced.

Tissue-specific gene inactivation

The deletion of a given gene from the mouse chromosome only in certain cell types, the gene remaining intact throughout the rest of the body. Such deletions are achieved by introducing two 34-bp DNA sequences – the *loxP* sites – on either side of the DNA segment to be deleted. These sites render the modified allele sensitive to Cre recombinase, which, if present in the desired cell type, binds to *loxP* sites and excises the intervening DNA segment from the chromosome. The *loxP* site and Cre recombinase are derived from the P1 bacteriophage.

The adaptation of organisms is based on their ability to respond to environmental variation by adapting their gene transcription profile to meet actual physiological needs. This is well illustrated by the adrenal release of glucocorticoids (GCs) in response to stress. These compounds may act on various tissues, coordinating immune, metabolic, endocrine, and behavioral responses by activating receptors, which, in turn, directly modulate gene transcription. This system is beneficial in healthy organisms, but dysfunctions in this system constitute a severe threat to the health and well-being of the organism. In humans, the dysregulation of the stress response is associated with the development of physiological and behavioral disorders, such as depression, anxiety, memory deficits, and drug dependence.

GCs bind to two related transcription factors, the ubiquitously expressed low-affinity glucocorticoid receptor (GR) and the high-affinity mineralocorticoid receptor (MR), which acts as a bona fide glucocorticoid receptor in a few cell types, such as hippocampal neurons, but is mostly expressed in cells involved in salt-balance control, in which it serves as a mineralocorticoid (aldosterone) receptor. In these cells, the MR is indeed protected from GCs by 11 β -hydroxysteroid dehydrogenase (11 β -HSD2), an enzyme that converts GCs into inactive forms.

The GR and MR genes belong to the nuclear receptor gene family, the members of which have several characteristics in common. The GR and MR proteins contain two highly structured domains: the DNA-binding domain (DBD) and the ligand-binding domain (LBD). In addition, two domains involved in controlling transcription (the activation function, AF, domains) are located in the N-terminal part of the protein and within the LBD. In the absence of ligand, the receptor exists in an inactive state as part of a large cytoplasmic complex. On hormone binding, the activated GR monomer translocates into the nucleus, where it modulates the transcription of target genes by remodeling chromatin and interacting with the basal transcriptional machinery, coactivators, and other transcription factors. The GR may activate or repress gene expression, by binding directly to its cognate response element in the promoter region of target genes and via DNA-binding-independent mechanisms involving other transcription factors (Figure 1). However, due to the wide-ranging effects of GCs, the precise mechanisms underlying the molecular actions of GR and MR and the nature of their cellular targets throughout the body remain largely to be defined with precision.

To address these issues, various animal models have been developed, based on the use of molecular genetic tools established for the precise modification

of the mouse genome. These tools include classic gene inactivation or mutation by homologous recombination in embryonic stem (ES) cells and tissue-specific gene inactivation based on the Cre-*loxP* system, both of which affect the endogenous gene. Other approaches modify gene dosage by adding a transgene to the genome for the constitutive or inducible (tetracycline system) expression of GR or MR. Critical analysis of the phenotypes resulting from these modifications has both confirmed the expected roles of GR and MR and uncovered entirely unexpected functions.

Classic Inactivation of Glucocorticoid Receptor and Mineralocorticoid Receptor Genes

Glucocorticoid Receptor Is Absolutely Required for Survival

Homologous recombination in ES cells has been used to generate various mutations inactivating the GR gene in mice. The resulting phenotypes were somewhat different. The first attempt involved insertion of a DNA cassette within the second exon to prevent the processing of GR mRNA. This approach generated a hypomorphic allele because alternative splicing of the GR gene can lead to the formation of a truncated mRNA encoding an N-terminally truncated GR protein that contains the DBDs and LBDs (allele GR⁻ or GR^{hypo}; Figure 2). A second mutation was then designed to ensure complete gene inactivation. This mutation consisted of a deletion of the third exon (GR^{null} allele), which encodes an essential part of the DBD, and it introduced a frameshift in the open reading frame following the fusion of the second and the fourth exon. This mutation thus ensures that no GR protein isoform containing the DBDs and LBDs can be formed.

GR^{null/null} animals develop to term, but die at birth because they are unable to breathe due to collapsing of the lungs. They present defects in the maturation of several tissues, including a very thin stratum corneum (the outermost layer of the epidermis) and a defect in pancreas organization. GC levels are tightly controlled by an endocrine cascade (see **Hypothalamic-Pituitary-Adrenal**), which is in turn regulated by a negative feedback loop. The absence of GR leads to a profound imbalance in the hypothalamic-pituitary-adrenal (HPA) axis. Corticotropin releasing hormone (CRH), adrenocorticotrophic hormone (ACTH), and GC levels increase. Consistent with the increase in GC levels, the region responsible for GC synthesis – the adrenal cortex – is markedly enlarged and displays tissue disorganization. Surprisingly, adrenal

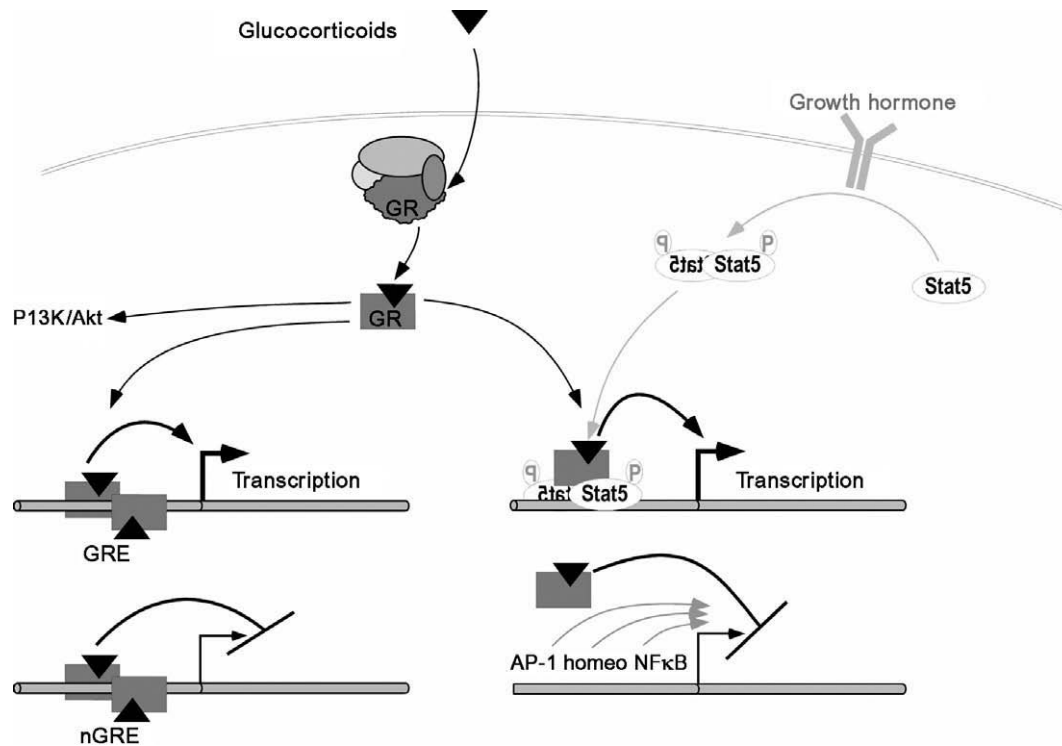


Figure 1 The glucocorticoid receptor acts by various molecular mechanisms. Due to their lipophilic nature, glucocorticoids can easily cross the cell membrane. Hormone binding to GR induces a change in the conformation of the receptor, leading to its release from an inactivating multiprotein complex and its translocation into the nucleus. The receptor then binds to glucocorticoid response elements (GRE) in DNA, as a dimer, or to negative GRE (nGRE), thereby activating or repressing the transcription of target genes. It can also modulate the activation of gene transcription by other factors, acting as a coactivator of Stat5 – a transcription factor induced by phosphorylation on GH binding to its receptor – or repressing transcription factors such as AP-1, NF- κ B, or some homeoproteins. GR may also act by other modes, interfering with intracellular signaling cascades by stimulating Akt phosphorylation, for example.

chromaffin cell development is not impaired. Differentiated chromaffin cells are scattered throughout the gland, but are no fewer in number in mutant neonates than in wild-type neonates. This unexpected result demonstrates that GR is not essential for the development of this cell type *in vivo*, whereas the contrary is suggested by results obtained in cell culture. However, GR^{null/null} chromaffin cells do not synthesize epinephrine due to a lack of expression of the gene-encoding phenylethanolamine *N*-methyltransferase (PNMT), the enzyme responsible for converting norepinephrine into epinephrine.

Hematopoietic stem cells from GR^{null/null} animals have been used to demonstrate the role of GR activation in stress erythropoiesis. In the absence of GR, cultured erythroblasts proliferate poorly and rapidly differentiate into erythrocytes. *In vivo*, this has important consequences for survival and adaptation. Chimeric adult animals subjected to irradiation and grafted with mutant stem cells such that they carry only GR^{null/null} erythroblasts cannot rapidly reconstitute a normal number of red blood cells after erythrolisis or

bleeding. In hypoxic conditions, similar to conditions at an altitude of 5000 m, they are unable to adapt rapidly by increasing their number of red blood cells.

GR^{hypo/hypo} animals (originally named GR^{-/-}) display a milder phenotype. Unlike GR^{null/null} animals, all of which die at birth, 20% of GR^{hypo/hypo} mice pass this stage, grow normally, and reach adulthood. Several of the phenotypes of GR^{null/null} animals, such as the almost total absence of the stratum corneum and the pancreatic defect, are not observed in GR^{hypo/hypo} mice. The differences between the two mutations presumably result from the presence of residual GR protein isoforms in GR^{hypo/hypo} animals. A third mutation involving the deletion of the alternative exon 1c and exon 2 has been generated (GR^{ko}). No skin or pancreas phenotypes have been reported for GR^{ko/ko} animals. The inactivation strategy, consisting of the deletion of the second exon, is similar to that of the GR^{hypo} allele and may potentially allow the formation of proteins with a DBD. Further investigations are required to understand the differences between GR^{ko/ko} and GR^{null/null} animals.

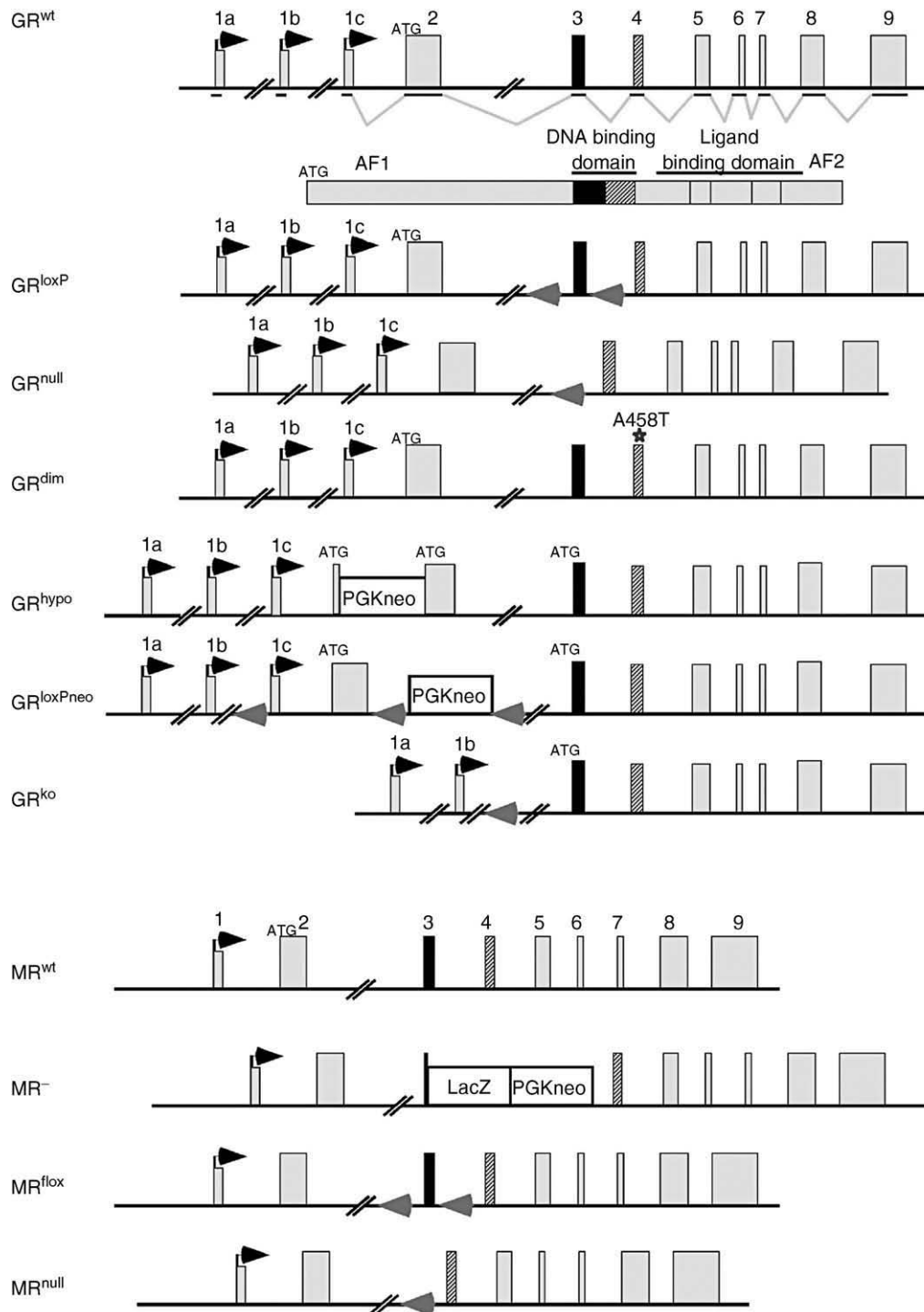


Figure 2 Schematic representation of the various GR and MR alleles generated in mice by gene targeting. Promoters are represented by black arrows, exons by boxes and *loxP* sites by gray arrows. The position of the *dim* point mutation is indicated by a star.

Inactivation of the Mineralocorticoid Receptor Gene Leads to Type I Pseudohypoaldosteronism

Mineralocorticoid hormone (aldosterone) is a key element in homeostatic sodium balance control, activating MR in the polarized cells of sodium-transporting

epithelia (distal part of the nephron, colon, salivary, and sweat glands). MR is also expressed in neurons, cardiomyocytes, and adipocytes, in which it may be activated by both mineralocorticoids and GCs. The MR gene has been inactivated by replacing the third

exon with the *lacZ* gene in embryonic stem cells. This exon encodes the first zinc finger of the DBD. $MR^{-/-}$ animals develop normally but die approximately 10 days after birth due to a massive loss of renal sodium and water. Mutant animals present major dehydration, with a high hematocrit and hyperkalemia. The renin–angiotensin–aldosterone system (RAAS), which controls salt balance, is strongly induced. These symptoms correspond to exacerbated human type I pseudohypoaldosteronism, due to mutations in the MR (heterozygous) or ENaC genes. In $MR^{-/-}$ animals, the molecular mechanism underlying the major loss of renal sodium remains unclear. Amiloride-sensitive sodium reabsorption rates have been shown to be low, but the levels of ENaC mRNA encoding the amiloride-sensitive epithelial sodium channels remain unchanged, as do the sodium/potassium ATPase mRNA levels.

$MR^{-/-}$ mice may be kept alive by salt supplementation. Animals may survive if they are given daily subcutaneous injections of NaCl solution until weaning and are supplied with oral NaCl thereafter. Adult animals display high levels of renal sodium excretion, hyperkalemia, and persistent activation of the RAAS. These survivors have been used to study other MR functions. MR is strongly expressed in brain structures, such as the hippocampus. In rescued adult $MR^{-/-}$ animals, the dentate gyrus region of the hippocampus is thinner than in control littermates due to lower levels of neurogenesis, resulting in a lower density of granule cells. The absence of MR may be the primary cause of this defect, but higher levels of GR signaling due to higher circulating GC levels may also be involved.

Tissue-Specific Corticosteroid Receptor Genes Inactivation

The early death of mice lacking the GR or the MR gene precludes the study of the functions of these genes in adult physiology or behavior. In addition, the disruption throughout the animal of genes expressed in many different cell types and encoding proteins with highly diverse functions, results in a phenotype reflecting the combined absence of these functions. This phenotype may therefore be difficult to interpret. The ability to inactivate a gene in a single cell type is therefore a major breakthrough and has clear advantages over the disruption of the gene throughout the body. The Cre–*loxP* system has been used to inactivate the GR and MR genes in a tissue-specific manner. This system is based on the use of CRE recombinase, which binds to a 34-bp DNA sequence (*loxP*) and excises any DNA segment flanked by two such sites. Therefore, if a *loxP* site is present on either side of a mouse gene, that gene will be selectively

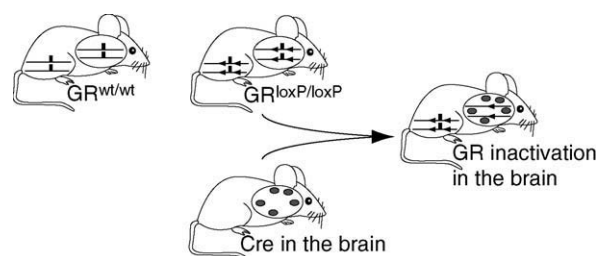


Figure 3 Principle of tissue-specific gene inactivation using the Cre–*loxP* system. When crossed with transgenic mice expressing the Cre recombinase in the brain (gray circle), the DNA fragment flanked with *loxP* sites in the same orientation (gray arrow) is deleted in this organ. It remains intact in all other cells of the organism.

deleted in cell types expressing the Cre recombinase but will remain intact in the rest of the body (Figure 3). Conditional alleles have been generated by flanking the third exons of the GR and MR genes with two *loxP* sites, through homologous recombination in ES cells. This DNA modification is very small given the overall size of the GR and MR genes (34 bp versus >100 kb), particularly because the neomycin resistance gene, used for ES cell selection, is removed from these targeted alleles. The insertion of *loxP* sites has no effect on the endogenous expression of the genes, but it makes them sensitive to the presence of Cre recombinase. Mice carrying the GR^{loxP} or MR^{loxP} alleles were crossed with transgenic animals expressing Cre recombinase in various tissues, leading to the conversion, in these tissues, of the GR^{loxP} or MR^{loxP} alleles into null alleles. Analyses of mutant mice lacking GR or MR in particular tissues confirmed some known or suspected functions of the receptor and revealed a number of previously unsuspected functions. Using this approach, we were able to dissect *in vivo* some of the roles of GR because we were able to isolate its function in a given cell type from its effects in the rest of the body.

Tissue-Specific Glucocorticoid Receptor Gene Inactivation

Liver-Specific Glucocorticoid Receptor Inactivation Results in Body-Growth Deficiency The GR gene was inactivated in hepatocytes, by expressing the Cre recombinase under the control of transcriptional regulatory elements from both albumin and α -feto-protein genes, leading to early and robust recombination (AlfpCre transgene). These animals were crossed with mice carrying the GR^{loxP} allele. The resulting mutant mice ($GR^{loxP/loxP}$; Tg:AlfpCre) are referred to as $GR^{AlfpCre}$ mice. Surprisingly, mice lacking GR in their hepatocytes are much smaller than control mice. This difference is first observed at 3–4 weeks of age and is more pronounced in males, suggesting a

dysfunction in growth hormone (GH)-dependent postnatal body growth. The defect is not systemic. GC and GH levels are normal, but the GH signaling pathway is not functional in hepatocytes. The binding of GH to its receptor activates the phosphorylation of signal transducer and activator of transcription 5 (Stat5), which in turn stimulates the transcription of GH-regulated genes. Some of these genes encode secreted factors mediating peripheral body growth. In the absence of GR, the expression of GH-regulated target genes controlled by Stat5, such as insulin-like growth factor (IGF-I), Spi2.1, acid labile subunit (ALS), and major urinary protein (MUP), is strongly reduced. This may be accounted for by physical interaction between GR and Stat5, as shown by coimmunoprecipitation and chromatin immunoprecipitation (ChIP) experiments and by GR acting as a coactivator for Stat5-dependent transcription on GH stimulation. This hypothesis is supported by the similar growth phenotype observed in animals lacking Stat5 throughout the body (Stat5^{-/-} animals) and by analyses of gene expression at the genomic scale, which have identified a subset of coregulated genes, including all known GH-regulated genes in the liver. Significantly, in addition to the role for hepatic GR in the control of body growth, this result demonstrates the essential role of the liver for body growth.

GCs were named based on the function of these molecules in the hormonal control of glycemia. They globally promote glycemia by acting on the fat tissue, muscles, pancreas, and liver, where they stimulate glucose synthesis. Glucose homeostasis was studied in GR^{AlfpCre} animals to isolate the effects of GCs on hepatocytes from those on other cell types. Basal glycemia was normal in GR^{AlfpCre} mice. However, in challenging situations, such as prolonged fasting, the induction of gluconeogenesis is impaired due to deficiencies in the induction of enzymes participating in this pathway, such as phosphoenol pyruvate carboxykinase (PEPCK), TAT, and glucose-6-phosphate (G6P). This leads to hypoglycemia in mutant animals, whereas control littermates maintain normal glucose levels. GR^{AlfpCre} mice have also been studied in pathological situations. The contribution of hepatic GR to the hyperglycemia observed in diabetes was assessed by inducing type I diabetes mellitus, by streptozotocin injection, in GR^{AlfpCre} and control animals. Hyperglycemia was less severe in the absence of hepatocytic GR, probably due to the observed defective induction of PEPCK, reducing gluconeogenesis.

GR^{AlfpCre} mice were used by U. Schibler and colleagues to confirm *in vivo* an unexpected role for GCs – their contribution to resetting the circadian clocks in peripheral tissues. In mammals, the circadian system consists of a master pacemaker in the suprachias-

matic nucleus of the hypothalamus and slave clocks in most peripheral cell types. The mechanisms underlying the resetting of these peripheral clocks remain unclear, but GCs seem to play a role. The injection of dexamethasone, a synthetic GC, induces the expression of circadian genes and shifts the circadian phase of peripheral organs. This effect involves GR in peripheral cells, as shown in GR^{AlfpCre} animals, in which a shift is observed after dexamethasone injection in all organs except the liver.

GR gene inactivation in the liver has also been reported in animals carrying a GR allele in which exon 1c and exon 2 are flanked by *loxP* sites and a neomycin resistance selection cassette. On recombination, this allele generates the GR^{ko} allele. Liver-specific mutants were obtained using a different Cre-expressing transgenic mouse line (AlbCre). Defective cell proliferation was observed after partial hepatectomy, but the body growth rate remained normal. This difference may be due to the use of the AlbCre transgene, which gives later recombination than AlfpCre for other *loxP* targets. It may also be due to differences in the GR mutation because removal of exons 1c and 2 is still potentially compatible with the formation of GR proteins with a DBD, as for the GR^{hypo} allele (as already discussed).

The Glucocorticoid Receptor Gene and Pancreas Programming There is considerable evidence to suggest that elevated maternal GC levels may influence embryonic development, with long-term consequences for the metabolism or behavior of the child into adulthood. Maternal food restriction results in a low birth weight and impaired glucose homeostasis in adulthood. This phenomenon may be related to the association between low birth weight and the appearance of cardiovascular diseases, obesity, insulin resistance, and diabetes at adulthood in humans. It has been suggested that maternal GCs may affect embryonic pancreas development because the consequences of maternal food restriction, including a reduction of the β -cell mass in the pancreas, can be blocked by adrenalectomy. In addition, the absence of GR in developing pancreatic cells leads to an increase in β -cell mass that persists into adulthood. This model was developed using a recombinase transgene expressed under the control of the pdx promoter (GR^{PdxCre} animals). Further studies on these animals should make it possible to analyze the role of pancreatic GR in adult animal metabolism.

The Glucocorticoid Receptor Gene and Mammary Glands GCs are known to influence mammary gland function *in vivo* and to stimulate milk-protein gene expression in cell culture. Mice deprived of GR

in their mammary epithelial cells were obtained using a Cre recombinase transgene expressed under the control of the whey acidic protein (WAP) gene promoter (GR^{WAPiCre}). They produce and secrete milk and nurse their litters until weaning, suggesting that GCs do not act on the mammary gland via GR or that they act on other cells than those secreting WAP. The absence of GR reduces cell proliferation, retarding lobulo-alveolar development.

The Glucocorticoid Receptor Gene and T Cells GCs act as potent anti-inflammatory and immunosuppressive therapeutic agents. The endogenous production of these hormones is involved in the optimization of the immune repertoire during thymocyte maturation and in downregulating the immune response in mature T cells. The role of GR in T cells from animals carrying the GR^{ko} allele was studied by grafting hematopoietic stem cells from fetal GR^{ko/ko} mouse liver into irradiated animals and using the lckCre transgenic line to invalidate the GR^{loxPneo} allele in thymocytes (TGRKO). Analyses of these animal models demonstrated that GR is not required for T-cell development.

In vivo, GR^{ko/ko} thymocytes are resistant to apoptosis induced by stimulation of the T-cell receptor (TCR). This is not a cell-autonomous property of GR^{ko/ko} thymocytes because these cells display normal sensitivity *in vitro*, demonstrating that the evoked GC release on TCR stimulation is involved in this response. TCR stimulation in TGRKO animals is associated with a high mortality rate, probably due to the uncontrolled induction of cyclooxygenase-2 (COX-2) because the selective inhibition of this enzyme prevents the death of mutant mice.

Inactivation of the Glucocorticoid Receptor Gene in the Brain GCs affect many aspects of brain physiology, in addition to repressing the HPA axis by controlling the CRH and vasopressin expression in the neurons of the hypothalamic paraventricular nuclei. In response to stress, these hormones modulate the activity of several neurotransmitter systems and affect behavior. In healthy animals, this response helps the organism to adapt to its environment, but dysfunctions in this mechanism may be associated with the development of behavioral disorders.

GCs can affect neuronal activity and neurotransmitter release, modify cellular shape by remodeling dendritic formations, compromise neuronal survival, and reduce the rate of renewal of granule cells in the dentate gyrus by downregulating neurogenesis.

The GR gene was inactivated in the brain by crossing mice carrying the GR^{loxP} allele with transgenic animals expressing Cre recombinase in neuronal

precursor cells under the control of the rat nestin promoter and a neuron-specific enhancer (the NesCre transgenic line). GR^{NesCre} mice, which have no GR in their neurons and glial cells, survive to adulthood. Because the GR-mediated negative feedback on the HPA axis is disrupted at the hypothalamic level in mutant mice, GC levels are 15 times higher than in control animals. The negative feedback loop continues to function in the pituitary gland (in which the GR gene is not inactivated), but this is clearly not sufficient to control hormone levels. The level of CRH production is very high in the hypothalamus of mutant animals, whereas vasopressin levels are unaffected. The increase in CRH levels induces ACTH production in the pituitary gland, and both mRNA and intracellular protein levels are high. Surprisingly, circulating ACTH levels are significantly reduced. This demonstrates that high GC levels are not sufficient to repress ACTH expression completely but are sufficient to repress ACTH secretion. The mechanisms responsible for higher levels of GC secretion in the presence of lower circulating ACTH circulating levels remain to be elucidated. A similar divergence between ACTH and GC levels is observed in other animal models and in some pathological situations, such as major depression. The high GC levels in GR^{NesCre} mice may have an effect on peripheral organs expressing GR, leading to symptoms reminiscent to those observed in patients with Cushing's syndrome (due to high GC levels), such as a reduced size, low bone density, and changes in fat distribution. GR^{NesCre} animals present a decreased food intake and a reduced metabolic efficiency. In GR^{NesCre} animals, GC levels cycle with a circadian rhythm and are further increased in response to stress.

In the brain of mutant mice, stress-released GCs are not perceived by GR that is absent. We therefore studied behaviors influenced by stress responses in GR^{NesCre} mice to distinguish between those involving GR activation in the brain and those involving other mechanisms. Mutant mice display lower levels of anxiety in the elevated zero maze and the dark-light box, two tasks based on a behavioral conflict between exploring and avoiding an aversive compartment. Despair or depression-like behavior, measured in the forced swim test, was also reduced in the absence of GR. CRH is often considered to be the anxiogenic mediator of the HPA axis. Interestingly, the fact that CRH expression is strongly induced in the hypothalamus of mutant animals, but unaffected in the amygdala, suggests that changes in CRH are not responsible for changes in the level of anxiety of GR^{NesCre} mice. This indicates that central GR may be directly involved in anxiety modulation. A more restricted deletion of the GR^{loxPneo} allele, using a

recombinase expressed in most forebrain neurons (FBGRKO mice), also affects emotional behavior. Despair behavior is more marked, but anxiety levels are reduced, in these mutants in the dark/light box and elevated plus-maze behavioral tasks. This difference between GR^{NesCre} and FBGRKO mice may reflect the restriction of the GR mutation to a neuronal subpopulation in FBGRKO mice. This mutation leads to an increase in GC levels that, in turn, may act on GR-positive neurons. More work is required to define the nature of the neuronal cells in which GR activation modulates particular types of behavior.

The absence of brain GR in GR^{NesCre} mice also has a weak effect on learning and memory retention in a water-maze task designed to test spatial memory. Learning is similar between mutant and control littermates, but retention is slightly affected in mutant mice. Further training sessions, however, overcame this memory impairment, indicating that GR is not essential for the formation of explicit memory.

Addiction is favored by stress responses. Physical and psychological stressors facilitate drug addiction in rodents. The acquisition of drug self-administration, its intensity, and the reinstatement of drug taking even after prolonged periods of withdrawal are facilitated or increased by stress. Even in the absence of stress, GR gene inactivation in GR^{NesCre} mice profoundly reduces the motivation for cocaine-induced self-administration, a test considered to be the best experimental model of drug abuse. In addition, cocaine-induced behavioral sensitization, a process implicated in the development of addiction and defining the increase in the behavioral effects of cocaine observed over repeated drug injections, is abolished in these mice. The molecular mechanisms underlying these effects are unknown, but the absence of GR leads to strong downregulation of the mitogen-activated protein kinase (MAPK) pathway responsible for impaired expression of the immediate early gene *Zif268* (also named *Egr-1*, *NGFI-A*, or *Krox24*). This is particularly evident in hippocampus GR^{NesCre} animals. The MAPK pathway and *Zif268* expression in the hippocampus of control animals are induced by immobilization – an acute stress for mice – but this induction is defective in GR^{NesCre} animals.

Tissue-Specific Mineralocorticoid Receptor Gene Inactivation

In the brain, MR is abundantly expressed in the hippocampus, where it serves as a high-affinity GR. Conditional MR gene inactivation was recently achieved by flanking the third exon with two *loxP* sites (MR^{fllox} allele) and using a transgenic line expressing Cre recombinase throughout the forebrain,

under the control of the CamKII α gene on a bacterial artificial chromosome. The absence of MR does not affect the circadian rhythm or stress response of the HPA axis. MR^{CamKCre} mice have a learning deficit in spatial navigation, as shown in a water-maze task, but display normal retention once the task is learned. MR^{CamKCre} mice also display working memory impairment, as assessed in a radial maze. In contrast to what has been observed in GR mutant animals, anxiety is not affected by the mutation. Significantly, studies on MR^{CamKCre} mice recently demonstrated that MR is responsible for certain rapid effects of GCs on hippocampal signaling that do not involve gene expression.

Function-Selective Glucocorticoid Receptor Mutation

GR acts via different molecular mechanisms involving specific protein–protein interactions, with itself or with other factors, to control gene expression (Figure 1). If these different interactions engage different protein interfaces, it should be possible to define mutations that selectively abolish a given mechanism without affecting the others. This is the case for a mutation in the D-loop region of the DBD that results in the replacement of alanine 458 by a threonine residue (GR^{dim}). This mutation affects the cooperative binding of GR homodimers to glucocorticoid-response elements (GREs), resulting in reduced occupancy of these binding sites. Interactions with other known proteins are not affected. Following its introduction into the mouse genome by homologous recombination in embryonic stem cells, this mutation was indeed shown to reduce the expression of genes known to be controlled via binding to GREs, such as the TAT gene in the liver. In contrast, the GR^{dim} protein is still able to repress the activating enhancer binding protein (AP-1)-dependent expression of the collagenase-3 (MMP-13) and stromelysin-1 (MMP-3) genes and pro-inflammatory NF κ B activity and to activate the expression of Stat5-dependent genes in the liver. Surprisingly, all homozygous mutant mice survived to adulthood. The fact that GR^{null/null} mice die shortly after birth, whereas GR^{dim/dim} mice do not, indicates that the interaction of GR with other proteins is more important for survival than its ability to bind DNA cooperatively on GRE target sites. In adult mice, various phenotypes have implicated defects in GR binding to GREs. CRH synthesis and ACTH release are not impaired in these mice. In contrast, transcription of the *POMC* gene (encoding ACTH) is upregulated in corticotropic cells and therefore requires the binding of GR dimers to DNA. In the hippocampus, dimerization and cooperative DNA

binding seem to be involved in neuronal activity. GR activation in GR^{dim/dim} mice does not lead to increases in calcium currents or in the response to serotonin as in wild-type mice, indicating that dimerization is required for these responses.

Modifying Glucocorticoid Receptor and Mineralocorticoid Receptor Gene Dosage in an Organism or Cell Type

We have so far discussed studies aiming to define the function of the GR and MR genes by means of gene mutation abolishing expression or generating modified protein. The function of corticosteroid receptors has also been studied using approaches that increase or decrease the levels of the corresponding gene products. GR gene dosage has been increased by introducing two additional copies of the receptor gene into the mouse genome using yeast artificial chromosomes. As the entire GR gene, including all the regulatory elements responsible for gene transcription, is present in the 290-kb yeast artificial chromosomes (YAC) used, the expression of the GR transgene should mimic the endogenous expression profile. These animals have a downregulated HPA axis under basal conditions and after acute stress, and they display increased resistance to lipopolysaccharide-induced endotoxic shock. A behavioral analysis of this animal model showed lower levels of despair behavior, with a decrease in GR dosage (GR^{null/wt} animals) giving the opposite response.

Mice over- or underexpressing GR in a given tissue have also been established by additional transgenesis. The first mice with impaired GR function were generated by overexpressing an antisense RNA directed against the 3' prime untranslated region of the GR mRNA. This mouse model has been of great value for studying GR function, but it is problematic for several reasons. The main problem is that it is difficult to determine the actual levels of GR protein in a given cell because transgene expression varies as a function of cell type. In addition, the antisense RNA is expressed ectopically in organs outside the nervous system, probably due to positional effects relating to the use of the human neurofilament promoter. On average, GR levels were 50% lower than normal in the brain, resulting in changes to HPA axis regulation and stress-related behavior and in changes in energy balance and lipid metabolism. GR has been overexpressed in various tissues, including the pancreas, brain, and T cells, using tissue-specific promoters or the tetracycline inducible system.

MR function has also been investigated by means of conditional downregulation and overexpression

strategies based on the tetracycline system. This approach revealed unexpected roles for MR in cardiac tissue. The heart-specific expression of an antisense mRNA directed against MR causes severe heart failure and cardiac fibrosis in the absence of hypertension or hyperaldosteronism, leading to massive water retention. In adult mice, this phenotype was fully reversed within a few days following the abolition of mRNA expression. Mice with conditional heart-specific MR overexpression in the heart also present a cardiac phenotype with marked arrhythmia and high early mortality rates, as observed in arrhythmia-related sudden death in humans.

The different animal models presented here are invaluable tools for the study of corticosteroid receptor gene function. They allowed us to define more precisely the known or expected corticosteroid receptor function, such as, the GR's roles in the HPA axis negative feedback loop and in glucose homeostasis and its functions in immunosuppression and anti-inflammation and the MR's role in the control of the body salt balance. These animal models also clarified the role of GR in the modulation of stress-related behaviors and revealed unexpected roles such as the requirement of GR for GH signaling in hepatocytes. The results are still fragmented, but further analysis of the existing models and of new models targeting other tissues should allow a more integrated understanding of corticosteroid receptor gene functions at the level of the organism. They should also allow us to distinguish in these functions the ones relevant for hormonal control and stress response from the ones that are not associated with the binding of corticosteroids. It is indeed now clear that, in some situations, steroid receptors may be activated by other means than ligand binding, such as phosphorylation. The generation of precisely modified animal models is fastidious and will undoubtedly benefit from complementary approaches, such as the use of short interfering RNA (siRNA) or small molecules that will facilitate combinatorial studies targeting several genes simultaneously.

Acknowledgments

I thank Monique Lazar and Christoph Kellendonk for their critical reading of this review and all the members of the Molecular Genetics, Neurophysiology and Behavior laboratory for our stimulating discussions. I also thank Günther Schütz for his support and for having introduced me to this research field. The work in the CNRS UMR7148 unit is supported by the CNRS, the NRJ foundation, the MILDT, the FRM, the French Ministry of Research (ACI and ANR programs), and the European Union (EAR-NEST consortium).

See Also the Following Articles

Corticosteroid Receptors; Corticosteroids and Stress; Hypothalamic-Pituitary-Adrenal.

Further Reading

- Balsalobre, A., Brown, S. A., Marcacci, L., et al. (2000). Resetting of circadian time in peripheral tissues by glucocorticoid signaling. *Science* **289**, 2344–2347.
- Bauer, A., Tronche, F., Wessely, O., et al. (1999). The glucocorticoid receptor is required for stress erythropoiesis. *Genes & Development* **13**, 2996–3002.
- Beggah, A. T., Escoubet, B., Puttini, S., et al. (2002). Reversible cardiac fibrosis and heart failure induced by conditional expression of an antisense mRNA of the mineralocorticoid receptor in cardiomyocytes. *Proceedings of the National Academy of Sciences USA* **99**, 7160–7165.
- Berger, S., Bleich, M., Schmid, W., et al. (1998). Mineralocorticoid receptor knockout mice: pathophysiology of Na⁺ metabolism. *Proceedings of the National Academy of Sciences USA* **95**, 9424–9429.
- Berger, S., Wolfer, D. P., Selbach, O., et al. (2006). Loss of the limbic mineralocorticoid receptor impairs behavioral plasticity. *Proceedings of the National Academy of Sciences USA* **103**, 195–200.
- Boyle, M. P., Brewer, J. A., Funatsu, M., et al. (2005). Acquired deficit of forebrain glucocorticoid receptor produces depression-like changes in adrenal axis regulation and behavior. *Proceedings of the National Academy of Sciences USA* **102**, 473–478.
- Boyle, M. P., Kolber, B. J., Vogt, S. K., et al. (2006). Forebrain glucocorticoid receptors modulate anxiety-associated locomotor activation and adrenal responsiveness. *Journal of Neuroscience* **26**, 1971–1978.
- Brewer, J. A., Kanagawa, O., Sleckman, B. P., et al. (2002). Thymocyte apoptosis induced by T cell activation is mediated by glucocorticoids in vivo. *Journal of Immunology* **169**, 1837–1843.
- Brewer, J. A., Khor, B., Vogt, S. K., et al. (2003). T-cell glucocorticoid receptor is required to suppress COX-2-mediated lethal immune activation. *Nature Medicine* **9**, 1318–1322.
- Cole, T. J., Blendy, J. A., Monaghan, A. P., et al. (1995). Targeted disruption of the glucocorticoid receptor gene blocks adrenergic chromaffin cell development and severely retards lung maturation. *Genes & Development* **9**, 1608–1621.
- Cole, T. J., Myles, K., Purton, J. F., et al. (2001). GRKO mice express an aberrant dexamethasone-binding glucocorticoid receptor, but are profoundly glucocorticoid resistant. *Molecular and Cellular Endocrinology* **173**, 193–202.
- Delaunay, F., Khan, A., Cintra, A., et al. (1997). Pancreatic beta cells are important targets for the diabetogenic effects of glucocorticoids. *Journal of Clinical Investigations* **100**, 2094–2098.
- Deroche-Gamonet, V., Sillaber, I., Aouizerate, B., et al. (2003). The glucocorticoid receptor as a potential target to reduce cocaine abuse. *Journal of Neuroscience* **23**, 4785–4790.
- Finotto, S., Kriegstein, K., Schober, A., et al. (1999). Analysis of mice carrying targeted mutations of the glucocorticoid receptor gene argues against an essential role of glucocorticoid signalling for generating adrenal chromaffin cells. *Development* **126**(13), 2935–2944.
- Froger, N., Palazzo, E., Boni, C., et al. (2004). Neurochemical and behavioral alterations in glucocorticoid receptor-impaired transgenic mice after chronic mild stress. *Journal of Neuroscience* **24**, 2787–2796.
- Gass, P., Kretz, O., Wolfer, D. P., et al. (2000). Genetic disruption of mineralocorticoid receptor leads to impaired neurogenesis and granule cell degeneration in the hippocampus of adult mice. *EMBO Report* **1**, 447–451.
- Gesina, E., Tronche, F., Herrera, P., et al. (2004). Dissecting the role of glucocorticoids on pancreas development. *Diabetes* **53**, 2322–2329.
- Gold, P. W., Loriaux, D. L., Roy, A., et al. (1986). Responses to corticotropin-releasing hormone in the hypercortisolism of depression and Cushing's disease: pathophysiologic and diagnostic implications. *New England Journal of Medicine* **314**, 1329–1335.
- Hubert, C., Gasc, J. M., Berger, S., et al. (1999). Effects of mineralocorticoid receptor gene disruption on the components of the renin-angiotensin system in 8-day-old mice. *Molecular Endocrinology* **13**, 297–306.
- Karst, H., Berger, S., Turiault, M., et al. (2005). Mineralocorticoid receptors are indispensable for nongenomic modulation of hippocampal glutamate transmission by corticosterone. *Proceedings of the National Academy of Sciences USA* **102**, 19204–19207.
- Karst, H., Karten, Y. J., Reichardt, H. M., et al. (2000). Corticosteroid actions in hippocampus require DNA binding of glucocorticoid receptor homodimers. *Nature Neuroscience* **3**, 977–978.
- Kellendonk, C., Eiden, S., Kretz, O., et al. (2002). Inactivation of the GR in the nervous system affects energy accumulation. *Endocrinology* **143**, 2333–2340.
- Kellendonk, C., Opherk, C., Anlag, K., et al. (2000). Hepatocyte-specific expression of Cre recombinase. *Genesis* **26**, 151–153.
- Mittelstadt, P. R. and Ashwell, J. D. (2003). Disruption of glucocorticoid receptor exon 2 yields a ligand-responsive C-terminal fragment that regulates gene expression. *Molecular Endocrinology* **17**, 1534–1542.
- Opherk, C., Tronche, F., Kellendonk, C., et al. (2004). Inactivation of the glucocorticoid receptor in hepatocytes leads to fasting hypoglycemia and ameliorates hyperglycemia in streptozotocin-induced diabetes mellitus. *Molecular Endocrinology* **18**, 1346–1353.
- Ouvrard-Pascaud, A., Sainte-Marie, Y., Benitah, J. P., et al. (2005). Conditional mineralocorticoid receptor expression in the heart leads to life-threatening arrhythmias. *Circulation* **111**, 3025–3033.
- Pazirandeh, A., Jondal, M. and Okret, S. (2005). Conditional expression of a glucocorticoid receptor transgene in thymocytes reveals a role for thymic-derived glucocorticoids in thymopoiesis in vivo. *Endocrinology* **146**, 2501–2507.

- Pepin, M., Pothier, F. and Barden, N. (1992). Impaired type II glucocorticoid-receptor function in mice bearing antisense RNA transgene. *Nature* **355**, 725–728.
- Reichardt, H. M., Kaestner, K. H., Tuckermann, J., et al. (1998). DNA binding of the glucocorticoid receptor is not essential for survival. *Cell* **93**, 531–541.
- Reichardt, H. M., Tuckermann, J. P., Gottlicher, M., et al. (2001). Repression of inflammatory responses in the absence of DNA binding by the glucocorticoid receptor. *EMBO Journal* **20**, 7168–7173.
- Reichardt, H. M., Umland, T., Bauer, A., et al. (2000). Mice with an increased glucocorticoid receptor gene dosage show enhanced resistance to stress and endotoxin shock. *Molecular and Cellular Biology* **20**, 9009–9017.
- Revest, J. M., DiBlasi, F., Kitchener, P., et al. (2005). The MAPK pathway and Egr-1 mediate stress-related behavioral effects of glucocorticoids. *Nature Neuroscience* **8**, 664–672.
- Ridder, S., Chourbaji, S., Hellweg, R., et al. (2005). Mice with genetically altered glucocorticoid receptor expression show altered sensitivity for stress-induced depressive reactions. *Journal of Neuroscience* **25**, 6243–6250.
- Shteyer, E., Liao, Y., Muglia, L. J., et al. (2004). Disruption of hepatic adipogenesis is associated with impaired liver regeneration in mice. *Hepatology* **40**, 1322–1332.
- Tronche, F., Casanova, E., Turiault, M., et al. (2002). When reverse genetics meets physiology: the use of site-specific recombinases in mice. *FEBS Letters* **529**, 116–121.
- Tronche, F., Kellendonk, C., Kretz, O., et al. (1999). Disruption of the glucocorticoid receptor gene in the nervous system results in reduced anxiety. *Nature Genetics* **23**, 99–103.
- Tronche, F., Kellendonk, C., Reichardt, H. M., et al. (1998). Genetic dissection of glucocorticoid receptor function in mice. *Current Opinion in Genetics and Development* **8**, 532–538.
- Tronche, F., Opherk, C., Moriggl, R., et al. (2004). Glucocorticoid receptor function in hepatocytes is essential to promote postnatal body growth. *Genes & Development* **18**, 492–497.
- Tuckermann, J. P., Reichardt, H. M., Arribas, R., et al. (1999). The DNA binding-independent function of the glucocorticoid receptor mediates repression of AP-1-dependent genes in skin. *Journal of Cell Biology* **147**, 1365–1370.
- Wei, Q., Lu, X. Y., Liu, L., et al. (2004). Glucocorticoid receptor overexpression in forebrain: a mouse model of increased emotional lability. *Proceedings of the National Academy of Sciences USA* **101**, 11851–11856.
- Wintermantel, T. M., Bock, D., Fleig, V., et al. (2005). The epithelial glucocorticoid receptor is required for the normal timing of cell proliferation during mammary lobuloalveolar development but is dispensable for milk production. *Molecular Endocrinology* **19**, 340–349.

Corticosteroid Receptors

O C Meijer and E R de Kloet

Leiden University, Leiden, Netherlands

B S McEwen

Rockefeller University, New York, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by O C Meijer, E R de Kloet, and B S McEwen, volume 1, pp 557–569, © 2000, Elsevier Inc.

Corticosteroid Receptor Properties
Cellular Effects of Corticosteroid Hormone Activation
Behavior and the Stress Response
Corticosteroid Receptors in Pathology

Glossary

Adrenalectomy The surgical removal of the adrenal glands from the body.

Corticosteroid hormones Hormones synthesized and secreted from the adrenal gland in response to stress

and also during the diurnal rhythm. The main endogenous glucocorticoids are corticosterone (in rodents) and cortisol (in humans).

Corticosteroid receptors

Intracellular receptor molecules for corticosteroid hormones. On hormone binding, they function as transcription factors in the nucleus of the cell to change mRNA and protein synthesis of target genes. There are two types: mineralocorticoid receptors and glucocorticoid receptors.

Hypothalamic-pituitary-adrenal (HPA) axis

The brain-controlled activation pathway for the secretion of corticosteroid hormones after stress through a hormonal cascade, involving the release of corticotropin releasing hormone from the paraventricular nucleus of the hypothalamus in the brain and adrenocorticotrophic hormone from the pituitary gland.

Hippocampus

A cortical brain structure involved in declarative and spatial memory and the cognitive aspects of emotions, with high

expression of both types of corticosteroid receptors.

Stress A state caused by a real or perceived challenge to homeostasis, which is accompanied by a hormonal stress response.

The effects of cortisol and corticosterone (CORT), the main glucocorticoid hormones in primates and rodents, respectively, are mediated by two types of intracellular receptor molecules: glucocorticoid receptors (GRs, or type II receptors) and mineralocorticoid receptors (MRs, or type I receptors). In peripheral tissues, CORT acts almost exclusively via GRs to affect energy stores, bone metabolism, and inflammatory and immune responses. MRs in epithelial cell types, as present in kidney and colon, are protected from endogenous CORT by the enzyme 11 β -OH-steroid dehydrogenase (11 β -HSD) so that aldosterone may bind to regulate the sodium balance in the body. In nonepithelial targets, such as the brain, both MRs and GRs function as receptors for CORT to act in either synergy or antagonism. GRs and MRs are members of the steroid hormone receptor superfamily and, accordingly, act as transcription factors to change the expression levels of target genes. These genomic effects have a relatively slow onset and long duration. However, CORT may also acutely affect the activity of neuronal cells via non-genomic mechanisms involving MR and possibly GR. However, no such receptors have been identified in mammals to date. MRs and GRs have different pharmacological properties and distributions in the brain. Corticosteroid receptors in the brain and pituitary play a crucial role in the regulation of the stress response and in mediating the effects of CORT on mood and cognition.

Corticosteroid Receptor Properties

Activation of Corticosteroid Receptors

The lipophilic CORT diffuses readily through the plasma membrane of cells to bind to its receptors. Corticosteroid receptors are part of a cytoplasmic multiprotein complex that consists of one receptor

molecule and several heat shock proteins (HSPs). The binding of CORT leads to a rapid chain of events that consists of the dissociation of the HSP, multiple phosphorylation steps, and increased affinity of the ligand-activated receptor for nuclear domains: the Ligand-receptor complex translocates to the cell nucleus to exert its action on gene expression and protein synthesis (Tables 1 and 2).

An important determinant of corticosteroid receptor activation is access of the ligand to the receptor, which depends on several factors in addition to the total plasma concentration of CORT. First, the circulating hormone is bound to corticosteroid-binding protein (CBG)/transcortin and with a much lower affinity to serum albumin. Of the average CORT concentration circulating during a 24-h period in the rat, less than 5% is not bound to CBG, (i.e., is free and biologically available).

Second, the enzyme 11 β -HSD type 2 converts cortisol and CORT to their inactive 11-dehydro metabolites in mineralocorticoid target tissues. Conversely, the 11 β -HSD type 1 isoform, which is often colocalized with GRs, can catalyze the reverse reaction and generate cortisol from cortisone within a target cell. This reaction takes place in the liver and fat cells, and it may also be relevant for certain areas in the brain.

A third determinant of access, which particularly pertains to synthetic glucocorticoids, is the *mdr1a* P-glycoprotein. This protein is expressed in the apical membranes of endothelial cells of the blood-brain barrier. *mdr1a* P-glycoprotein functions as an energy-dependent pump that limits access to the brain of xenobiotic agents, including synthetic steroids. Accordingly, the brain (but not the pituitary) is resistant to the penetration of moderate amounts of dexamethasone.

Corticosteroid Receptor Diversity in Pituitary and Brain

Although the GRs and MRs act as high-affinity receptors for CORT in the brain, these receptor types differ significantly with respect to their pharmacology, distribution, and physiological effects.

Table 1 Characteristics of the two corticosteroid receptor types

	Affinity	Agonists	Antagonists	Distribution
Mineralocorticoid receptor (type 1)	High affinity for cortisol and corticosterone ($K_d \approx 0.5$ nM)	Aldosterone (epithelial cell types as in kidney), cortisol (brain)	Spironolactone, RU 28328	Limited distribution in periphery (kidney, colon, lymphocytes) and brain (limbic brain)
Glucocorticoid receptor (type 2)	Lower affinity for cortisol and corticosterone ($K_d \approx 5.0$ nM)	Cortisol, dexamethasone, RU 28362	RU 38486	Ubiquitous expression in periphery and brain

Table 2 Milestones in brain corticosteroid receptor research

Year	Milestone
1968	First report on selective retention of [³ H]corticosterone in hippocampus
1975	Low retention of [³ H]dexamethasone in hippocampus; high in pituitary
1980	Selective MR agonists discriminate between MRs and GRs
1982	<i>In vitro</i> hippocampus and kidney cytosol contain MRs and GRs
1983	GR acts as transcription factor through DNA binding
1985	MR and not GR retains [³ H]corticosterone in hippocampus
1985	Cloning of GR
1985	First brain map of immunoreactive GR neurons
1987	Cloning of MR
1988	11 β -HSD confers aldosterone specificity to MR
1990	Transrepression of AP-1-induced transcription via GR
1991	First brain map of immunoreactive MR neurons
1993	Differential effects of MR and GR on transrepression
1994	<i>In vitro</i> demonstration of MR/GR heterodimers
1994	Mechanistic differentiation between transactivation and transrepression GR
1995	mdr1a Pgp in blood–brain barrier hampers brain entry of dexamethasone
1995	GR knockout mice
1996	Colocalized immunoreactive MR/GR clusters in hippocampal neuronal nuclei
1998	Knock-in mice that express a dimerization-impaired mutant of GR
2005	Non-genomic effects in hippocampus via MR

Structure The human GR gene is localized on chromosome 5, whereas the MR gene is localized on chromosome 4, indicating an early duplication of their common ancestor during evolution. Despite this early divergence, their genomic organization is very similar, notwithstanding differences at the 5' and 3' ends. The overall structure of GR and MR proteins follows that of the other members of the steroid hormone receptor family. The molecules are generally viewed as consisting of three domains: a highly conserved DNA-binding domain (95% homology between the MR and the GR), flanked by the C-terminal ligand-binding domain (60%), and a large N-terminal domain that contains a transactivation function and is highly variable between the MR and the GR (15%) and among species.

Pharmacology Although GRs and MRs are high-affinity receptors for CORT, they differ in their affinities for several ligands. In the case of cortisol and CORT, an MR has an approximately 10-fold higher affinity than a GR. The important consequence of this is that GRs and MRs are occupied differentially during the day and during stress responses. MRs

are already occupied extensively under basal trough conditions, whereas the saturation of GRs requires hormone levels that occur after stress or after the circadian peak. Because of the high, almost tonic, occupation of MRs, it has been hypothesized that the amount of bioactive receptor protein is an important level of regulation for MRs, whereas the GR signal depends primarily on ligand concentration. Synthetic glucocorticoids have higher affinity for GRs than for MRs, but their access to the brain GRs is hampered by the mdr1a gene-encoded P-glycoprotein.

Distribution GRs are expressed ubiquitously in many different tissues and cell types. In tissues such as the liver, lung, and adrenal medulla, GRs are crucial for appropriate development. MRs have a much more limited distribution, and as receptors for CORT (as opposed to mineralocorticoids), they have been characterized in the brain and some lymphoid tissues. The pituitary, outside the blood–brain barrier, contains GRs as well as MRs, although no specific function for MRs has been described for any of its cell types.

Throughout the brain, immunocytochemical and *in situ* hybridization procedures have shown a widespread distribution of GRs in neurons and glial cells. Particularly high GR concentrations are found in the limbic system (hippocampus, with relatively low concentrations in the CA3 region; septum; and amygdala), in the parvocellular neurons of the paraventricular nucleus (PVN) of the hypothalamus, and in the supraoptic nucleus. In the PVN, the biosynthesis and release of parvocellular vasopressin, corticotropin releasing hormone (CRH), and other neuropeptides are under glucocorticoid control. GRs are also present in relatively high concentrations in the ascending monoaminergic neurons of the brain stem. Moderate GR levels are also found in many thalamic nuclei and in patchlike distribution in the striatal areas, as well as throughout the cortical hemispheres (see [Figure 1](#)).

In the brain, MRs have a more restricted topography than GRs. High MR densities have been found in the neurons of the hippocampal formation, lateral septum, medial and central amygdala, olfactory nucleus, layer II of the cortex, and in brain-stem sensory and motor neurons. This distribution of MRs is essentially the same as that discovered in 1968 by McEwen with the cell-nuclear retention of radioligand after the administration of tracer doses of [³H]-corticosterone to adrenalectomized rats. Aldosterone-preferring MRs involved in salt homeostasis are localized in the anterior hypothalamus and circumventricular organs, such as the choroid plexus.

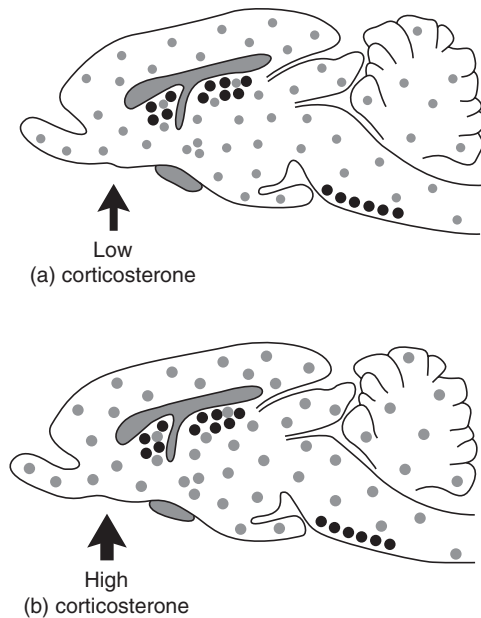


Figure 1 Distribution of MRs (black circles) and GRs (grey circles) in the rat brain. a, Basal corticosterone secretion; b, Peak corticosterone secretion. MR expression is more restricted and especially high in the hippocampus; the MRs are substantially occupied by corticosterone under basal conditions (i.e., at the circadian trough under nonstressed conditions). The GRs show a widespread distribution and are only partially occupied under basal conditions; they become more fully activated at the circadian peak of corticosterone secretion or after stress. Reprinted from M. Joëls and E. R. De Kloet (1992), Control of neuronal excitability by corticosteroid hormones. *Trends in Neuroscience* 15, 25–30. Used with permission.

The subcellular localization of MRs and GRs was studied in hippocampal neurons by dual labeling immunocytochemistry and confocal microscopy. It was observed that MRs and GRs are distributed non-homogeneously over the nucleus. Both receptors are concentrated in approximately 1000 clusters scattered throughout the nucleoplasm. Many clusters contain exclusively either MRs or GRs, although a significant number of domains were found to contain both receptor types. The latter clusters are candidate sites where the two receptors could interact to establish a coordinated regulation of gene expression. This implies that GR and MR homodimers, as well as the possible MR–GR heterodimers, are associated with distinct nuclear domains.

Corticosteroid Receptor Variants Alternative splicing of the 3' end of the human GR pre-mRNA creates an hGR β variant in addition to the common hGR α variant. However, it should be noted that rodent GR pre-mRNA lacks this splice site and GR β cannot be detected in rodent tissue. The translation of GR α and GR β mRNA produces two proteins that are almost

identical but that differ in their C-terminals, which implies that GR α binds cortisol, whereas GR β does not. However, *in vitro* experiments suggest that GR β is capable of binding to glucocorticoid-responsive elements (GREs) and can form homodimers as well as heterodimers with GR α . Accordingly, GR β has been shown to block GR-mediated transactivation *in vitro*, whereas it does not seem to interfere with transrepression. In the human brain, GR β mRNA is found in the hypothalamus and the hippocampus, but the relative abundance is subject to debate and may be only 1% relative to GR α . Thus, validation of the possible role of GR β as a dominant-negative factor awaits further studies. Different forms of GR may also arise from the use of alternative translation start sites for its mRNA, adding yet another level of complexity to the receptor diversity.

The expression of the human MR gene may result in the formation of at least four transcripts, which are derived from two different promoters. The two main transcripts, MR α and MR β , differ only in the 5' untranslated exon 1 and thus are translated into the same 985-amino-acid MR protein. In rats, an MR transcript derived from exon 1 is detectable in very low amounts. Other intriguing splice variants have been described, including transcripts that code for truncated proteins. The use of an alternative splice site between exon 3 and 4 creates a 12-bp insertion, which, after translation, corresponds to four additional amino acids in the amino acid sequence bridging the two zinc finger domains of the DNA-binding domain. The functional analysis and relevance of these MR variants await further studies.

Variations in MRs and GRs in the form of single nucleotide polymorphism are quite frequent in the normal human populations. Although the biology of these variants is poorly understood, some of these variant receptors can be associated with the metabolic profile, stress responsiveness, and even life span.

Receptor variants may play an important role in disease states. Thus, individuals with familial glucocorticoid resistance may express a GR variant and, as a consequence, have deficient GR function and impaired glucocorticoid feedback resulting in hypercorticism. However, they lack the symptoms of Cushing's syndrome because adrenocorticotrophic hormone (ACTH) and cortisol levels are also at a higher set point and GRs are, in general, less sensitive to glucocorticoid activation. The symptomatology of familial glucocorticoid resistance is usually related to the overproduction of adrenal mineralocorticoids and androgens in response to ACTH. Although these severe inherited deficits in GRs, resulting in asymptomatic hypercorticism, are rare disorders

in humans, there are also pronounced species differences in guinea pigs, prairie voles, and New World monkeys. These animals display a normal circadian rhythm in HPA axis activity, but ACTH and cortisol are circulating at a much higher level. Apparently, the elevated set point in HPA axis regulation is an adequate adaptation.

Regulation of Transcription through Corticosteroid Receptors

As members of the steroid hormone receptor superfamily, ligand-activated GRs and MRs exert their effects by acting as transcription factors to up- or downregulate protein and peptide levels in the cell. MRs and GRs may stimulate or repress transcription by several mechanisms (see Figure 2).

Transactivation via Glucocorticoid-Responsive Elements GRs and MRs contain a nearly identical DNA-binding domain that recognizes specific DNA elements in the regulatory regions of genes: glucocorticoid response elements (GREs; consensus sequence: GGTACAnnnTGtT/cCT). The steroid receptors bind as homodimers and perhaps also as heterodimers to GREs to stimulate transcription. In general, GRs are more potent activators of transcription than MRs, at least in *in vitro* conditions. Heterodimers of MRs and GRs have been shown in cell systems to have, at times, characteristics that are different from either type of homodimer. The strong synergizing effect of GR-activated transcription on multiple GREs is not observed with MR activation, probably due to the

limited homology of the N-terminal sequences. Tyrosine amino transferase and phenylethanolamine-N-methyltransferase (PNMT) are examples of genes that are regulated via GREs.

Repression via Negative Glucocorticoid-Responsive Elements GRs (and possibly MRs) can also repress gene transcription by binding to DNA. DNA elements that are involved are known as negative GREs (nGREs). The sequence of nGREs can be highly variable and differs from the consensus sequence for positively acting GREs. The mechanism by which transcription is repressed also differs between cases. One mechanism involves the binding of GRs to the nGRE to occlude adjacent or overlapping binding sites on the DNA for positively acting transcription factors. An nGRE has been described for the human POMC gene.

Repression via Protein-Protein Interactions The mode of action generally referred to as transrepression involves the repression by GRs of gene transcription activated by other transcription factors, such as activating protein (AP)-1, nuclear factor (NF)- κ B, and cAMP response element binding protein (CREB). GRs interfere with these other factors via protein-protein interactions. This may happen independently of the binding of GRs to the DNA, and dimerization of GRs is not required. The target genes involved lacked GREs.

Corticosteroid receptors may also activate genes in synergy with other, nonreceptor transcription factors.

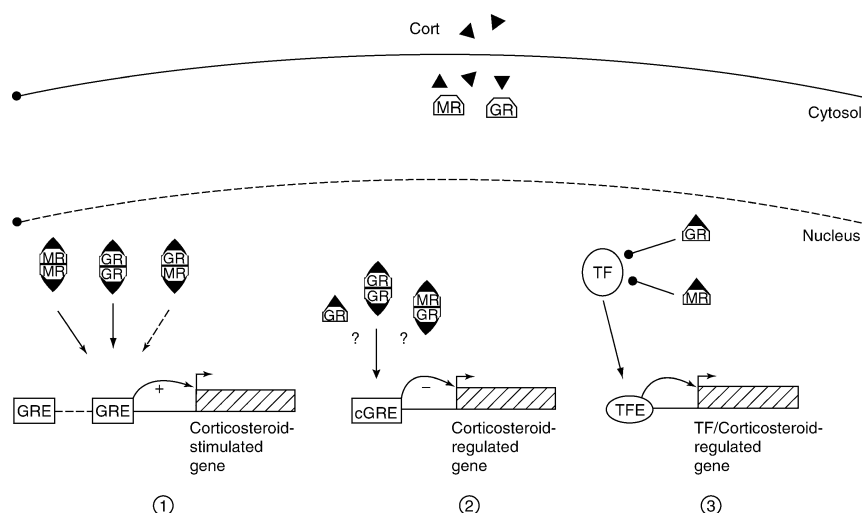


Figure 2 Three molecular mechanisms of regulation of gene expression. (1) GRs and MRs bind as homo- and heterodimers to GRE sequences to stimulate transcription. (2) GRs (and possibly MRs) bind to negative GRE sequences that diverge from consensus GRE. Through mechanisms such as occlusion of adjacent sites they repress transcription. (3) GRs and MRs interfere through protein-protein interactions with other transcription factors (TFs) at times independent of binding of the steroid receptors to the DNA. This results in antagonism of the effects of the TF (e.g., AP-1 or NF- κ B).

On elements known as composite GREs (cGREs), GRs bind to the DNA in close proximity to another transcription factor and may either repress or enhance the effect of this other factor. In the case of the proliferin cGRE, the GR enhances the effect of the adjacently bound transcription factor AP-1 greatly if the latter consists of c-jun–c-jun homodimers. In contrast, GR binding leads to repression of the action of c-jun–c-fos heterodimers. Interactions with other transcription factors depend by definition on the presence of these factors, which can be regulated by extracellular signals other than steroid hormones. This demonstrates that the effects of CORT vary widely and are context-dependent.

Dissociation between MR- and GR-mediated events may arise through these cross-talk mechanisms because it has been shown that GRs suppresses AP-1 activity under conditions in which an MR is ineffective. It is probable that the unique N-terminal regions of the GR and the MR account for the different properties with respect to transrepression.

Cellular Effects of Corticosteroid Hormone Activation

The activation of intracellular corticosteroid receptors in many cell types may induce a variety of cellular responses, influencing diverse processes such as cellular structure, energy metabolism, or signal transduction. Changes in cellular structure usually develop over the course of several days, whereas effects on energy metabolism or signal transduction may already become apparent within an hour. In peripheral tissues, the activation of GRs leads to pronounced cellular effects, for example, apoptosis of thymocytes (via antagonism of NF- κ B activity) and induction of gluconeogenesis in hepatocytes (through induction of critical enzymes). Here, focus lies on the effects of the MRs and GRs in the brain because these are important factors in the regulation of the HPA axis and of adaptation to stress in general. Although the delayed onset of the observed effects – as opposed to the rapidly induced effects by neurosteroids – favors gene-mediated effects via MRs and GRs, the molecular mechanism has, in many cases, not been resolved.

Differential Effects at the Level of a Single Structure: Mineralocorticoid and Glucocorticoid Receptors in the Hippocampus

The principal cell types in the hippocampus, a brain structure involved in declarative memory processes and (possibly) mood, express both MRs and GRs. The effects of differential occupation on firing characteristics have been particularly well studied for

CA1 pyramidal cells. A conspicuous feature of CORT actions on cellular activity in the hippocampus is the apparent lack of effect when neurons are studied under basal conditions. Only when neurons are shifted from their basal condition (e.g., by the action of neurotransmitters) do CORT effects become visible. The general picture emerging is that conditions of predominant MR activation (i.e., at the circadian trough at rest) are associated with the maintenance of neuronal excitability, so excitatory inputs to the hippocampal CA1 area result in considerable excitatory hippocampal output. In contrast, additional GR activation (e.g., following acute stress) generally depresses the transiently activated CA1 hippocampal output. A similar effect is seen after adrenalectomy (ADX), indicating a U-shaped dose–response dependency of these cellular responses to corticosteroids. Most of the CORT effects on hippocampal physiology have been shown to depend on *de novo* protein synthesis and even on DNA binding of the GR, indicating that the effects on transcription of as yet unidentified target genes are responsible for the observed phenomena.

The U shape of CORT on neuronal excitability, which is determined by differential MR and GR activations, is reflected in a number of parameters. An intrinsic cellular property is the voltage-gated Ca current; this current is small under conditions of predominant MR occupation and is elevated when GRs become activated in addition. Basal and stimulus-induced intracellular Ca levels were also found to be increased by substantial GR activation, which may be explained by steroid effects on Ca-buffering or extrusion mechanisms. In accordance with these steroid effects, the Ca-dependent accommodation and after-hyperpolarization (AHP) amplitude were found to be small with predominant MR activation.

MR and GR activations also modulate the responses to several neurotransmitters. With respect to amino acid-mediated synaptic transmission in the hippocampus, responses to both excitatory (glutamatergic) and inhibitory (GABAergic) inputs are maintained at a stable level when MRs are predominantly activated (i.e., with low CORT levels). When GRs become occupied, in addition, as a consequence of rising CORT levels, excitatory transmission (and thus CA1 hippocampal output) is reduced. At very high steroid levels, inhibitory networks are also impaired. These GR-mediated effects may involve the impairment of energy metabolism. The high efficiency of amino acid transmission with predominant MR activation is also reflected in long-term plastic changes involving this hippocampal network. Plasticity-related phenomena are most pronounced with moderate CORT levels so that most of the MRs and only part of the GRs are activated; when CORT levels are either

reduced or enhanced, synaptic potentiation/depression is far less effective. Thus, the varying concentrations of CORT as a consequence of stress will have consequences for the efficacy of synaptic potentiation in the hippocampus.

In addition to amino acid input, hippocampal neurons also receive considerable input mediated by biogenic amines such as acetylcholine, norepinephrine, and serotonin. Norepinephrine acts via a β -adrenergic receptor to increase excitability of the CA1 hippocampal area. High CORT levels, occupying GRs, reduce the norepinephrine effect, again leading to a reduction in responsiveness of the hippocampal neurons.

Serotonin (5-hydroxytryptamine, 5-HT) evokes many different effects in the CA1 hippocampal neurons, of which 5-HT-1A receptor-mediated hyperpolarization of the membrane is most prominent. This hyperpolarization is small in amplitude with predominant MR occupation; additional GR activation increases 5-HT responses. This is seen both when CORT is applied exogenously and after acute stress. Although CORT through MR activation also downregulates 5-HT-1A receptor mRNA expression and binding capacity, acute functional impairment is probably not brought about through reduced binding capacity.

Effects on Neurotransmitter Systems

CORT influences many cell types in the brain, and it is worthwhile to analyze its effects at the level of neurotransmitter systems, as opposed to a single cell type or nucleus. Major targets for coordinate regulation through corticosteroid receptors are the serotonin, dopamine, and norepinephrine systems. Modulation may take place at the level of both the presynaptic cell and the postsynaptic cell. As for most aspects of CORT action, the effects are permissive in nature – the genomic effects of steroid receptor activation are revealed in the context of activated systems.

Serotonin: Coordinate Regulation via Mineralocorticoid and Glucocorticoid Receptors A concept of coordinate regulation via MRs and GRs has emerged for the serotonin system, particularly in relation to its projection on the hippocampus. Under basal conditions, when corticosteroid levels are low, there is predominant activation of MRs. As described earlier, the response of hippocampal 5-HT-1A receptors under these conditions is low. In addition, MRs in certain hippocampal areas suppress the expression of the 5-HT-1A receptor. Thus, a situation of predominant MR occupation is associated with low responses to serotonin in the hippocampus. Serotonin release in the hippocampus is increased during and after stress

conditions that are associated with increased levels of CORT, that is, increased activation of GRs in serotonergic and hippocampal neurons. When the stress-induced GR occupation is prevented, no increased release of serotonin is observed in animals. CORT acts via GRs, presumably on the activity of the rate-limiting enzyme for serotonin synthesis (tryptophan hydroxylase), and via the attenuation of negative feedback inhibition by serotonin in the serotonergic neurons in the midbrain. At the postsynaptic level in the hippocampus, acute GR activation leads to increased responses to 5-HT-1A receptor activation. Thus, whereas MR-mediated actions of CORT inhibit hippocampal responses, GR occupation is required for the stress-induced increased activity of serotonin cells and for the increased responsiveness of 5-HT-1A receptor-mediated responses.

When CORT is chronically elevated to levels in the GR-occupying range, hyporesponses to serotonin are observed in hippocampal neurons and endocrine challenge tests. It is unclear which molecular changes underlie the transition from stimulatory to inhibitory effects of high levels of CORT on serotonergic transmission.

Catecholamines Dopaminergic and catecholaminergic cells also depend on GR occupation for development and function. As in the case of stress-induced serotonin activity, increased activity of the mesolimbic dopamine system in the adult rat depends on GR activation. Dopamine (DA) release in the nucleus accumbens and neuronal and behavioral responses to dopamine agonists such as apomorphine are stimulated by CORT via GR. Sensitization to repeated treatment with drugs that increase dopaminergic transmission, such as cocaine, depend on GR activation. GR-mediated activation and sensitization of DA systems are important for understanding the relationship between stress and drug addiction in the context of individual response characteristics and have implications for psychopathology. Increased DA is involved in the symptoms of psychosis and depression, which are often associated with elevated basal levels of circulating CORT and escape from dexamethasone suppression.

In the periphery, there are clear effects of CORT on epinephrine – the expression of the PNMT gene in the adrenal medulla depends on GR activation. CORT also potentiates epinephrine responses at the target level. In the central nervous system, GRs are abundantly present in the majority of the neurons in midbrain norepinephrine cell groups. Although studies with ADX rats show a permissive function of corticosteroids for norepinephrine cells during ontogeny and stress (e.g., in relation to memory formation),

the effect of GR activation in these cells and their postsynaptic targets are not as clear as for other transmitters.

Postsynaptically, CORT region-dependently regulates adrenergic receptor expression; that is, it induces the inhibitory $\alpha 2$ receptor in the hypothalamic PVN. In the hippocampus, the electrophysiological effects of norepinephrine are attenuated via a GR-dependent mechanism. Stress levels of CORT reduce the cyclic AMP generation induced by norepinephrine acting via β -adrenergic receptors in the hippocampus.

Neuropeptide Systems In addition to the classical neurotransmitter systems, many peptidergic systems are controlled via corticosteroid receptors, both at the level of peptide synthesis and at the level of postsynaptic receptors. These peptides are involved in many aspects of brain function, including hypothalamic control of the HPA axis and regulation of metabolism and feeding. The best-studied examples are POMC-derived peptides, CRH and arginine vasopressin (AVP) because of their critical role in the activation of the HPA axis. Both CRH and AVP mRNA levels are suppressed by CORT in the parvocellular division of the PVN, presumably via GR activation. However, this type of regulation is cell type-dependent because in the amygdala CRH mRNA is upregulated by corticosteroids. In the hippocampus, corticosteroid receptors mediate the regulation by CORT of neuropeptide Y and dynorphin mRNA levels; in the caudate-putamen and nucleus accumbens, corticosteroid receptors mediate regulation by CORT of preproenkephalin, neurokinin A, and preprodynorphin mRNA levels.

Cellular Viability

Structural changes of hippocampal neurons were observed both with chronic absence and chronic overexposure to CORT, indicating that the steroid-dependent expression of genes is of crucial importance for hippocampal integrity.

Chronic Absence of Corticosteroids The removal of the adrenal glands results within 3 days in apoptotic-like degeneration of mature granule cells in the rat dentatus gyrus but not in other hippocampal fields. The degeneration and associated loss of synaptic function can be prevented by treatment with MR ligands. In contrast to most brain regions, the dentate gyrus shows neurogenesis even during adulthood, which can be blocked by chronic stress and adrenal steroids. Lack of MR activation as a consequence of ADX results in increased cell birth and cell death. It has been argued that MR activation maintains

the balance between cell birth and cell death in the dentate gyrus by enhancing the excitatory input to this region. However, newborn neurons in the dentate gyrus do not express corticosteroid receptors, suggesting that the steroid effects on neurogenesis are indirect and possibly involve the modulation of excitatory transmission and neurotrophins.

Chronic Exposure to High Corticosteroid Levels

Ample evidence now shows that chronic elevation of CORT levels leads to atrophy in parts of the hippocampus, as well as to increased vulnerability to excitotoxic insults. CA3 pyramidal neurons seem to be particularly vulnerable, although effects in the CA1 and dentate subfield have also been reported. Three weeks of restraint stress in rats causes the regression of apical dendrites of hippocampal CA3 pyramidal neurons. This effect is mimicked by 3 weeks of treatment with a high dose of CORT and does not occur when CORT secretion is blocked by cyanoketone treatment. The regression occurring after 3 weeks of high levels of CORT is reversible, and available data support the view that an imbalance of excitatory over inhibitory signals, leading indirectly to an enhanced Ca influx in the CA3 pyramidal neurons, contributes to the observed degeneration. Lack of neurotrophic capacity may also contribute to enhanced neuronal vulnerability. Indeed, chronic stress or the administration of CORT affects brain-derived neurotrophic factor (BDNF), nerve growth factor (NGF), and basic fibroblast growth factor (bFGF), and tumor growth factor (TGF) expression in the hippocampus.

A critical duration of overexposure to steroids is essential for the atrophy of CA3 neurons to develop; temporary high levels of the hormone decrease responsiveness to excitatory amino acids and reduce excitability. It is only in conditions of chronically enhanced corticosteroid exposure that local depolarization is sustained and the breakdown of inhibitory networks occurs. Hippocampal neurons are subsequently subjected to a substantial GR-dependent rise in intracellular Ca levels. Following this line of reasoning, the beneficial effects exerted by phasic activation of GRs are turned into damaging actions when GRs are activated chronically.

Behavior and the Stress Response

The differential effects on neuronal functioning that CORT exerts via MRs and GRs, and the differential occupation of these receptors as a function of the behavioral state of the animal, suggest fundamentally different roles for CORT via MR- and GR-mediated processes. MRs are occupied by basal levels of CORT and thus mediate tonic hormonal influences. This

might function to prepare the animal for upcoming challenges and determine the magnitude and/or nature of initial responses. This proactive mode of control includes the effects on the processing of sensory information, interpretation of environmental stimuli, and selection of the appropriate behavioral strategy. GRs, however, are occupied after plasma CORT levels have become elevated after challenges/stress. Their role may be to generally mediate reactive processes that serve to restore homeostasis, to facilitate behavioral adaptation of the animal, and to prepare the animal for a next exposure to a stressor.

Corticosteroid Receptors and Hypothalamic-Pituitary-Adrenal Axis Control

Both MRs and GRs are involved in the feedback control of HPA axis activity, but the temporal and neuroanatomical aspects of feedback differ substantially for each receptor type. Apart from primary feedback sites, the hippocampus has received a lot of attention, particularly because it contains the highest concentrations of both MRs and GRs in the brain.

Pituitary corticotrophs and PVN (being part of the HPA axis) contain the primary feedback sites for stress levels of CORT. The effects of CORT on pituitary, PVN, and brain-stem catecholaminergic nuclei involve the activation of GRs. Limbic inputs impinging on the PVN (via hypothalamic GABAergic neurons) express high levels of the MR in addition to the GR, suggesting the dual regulation of these inputs by CORT. Moreover, the effect of these inputs can be either inhibitory on HPA activity (from hippocampus) or excitatory (e.g., from amygdala).

Pharmacological experiments in rats have shown that MR occupation suppresses HPA axis activity under basal conditions around the circadian trough; in ADX animals, hormone levels can be normalized with a median inhibitory concentration (IC_{50}) of approximately 0.5 nM in terms of circulating free CORT, which is in the range of the MR K_d value. In intact animals, MR antagonists lead to an increase in CORT levels at the time of the circadian trough. Accordingly, MR activation serves to control HPA axis activity in a proactive manner.

At the circadian peak, much higher levels of exogenous corticosteroids are required to normalize ADX-induced increases in HPA axis activity, and half-maximal suppression is achieved by a free concentration of approximately 5 nM, close to the K_d of GRs. However, the exclusive activation of GRs is insufficient to suppress the circadian peak, and MR activation appears to be indispensable. One episodic rise in CORT levels by injection or ingestion via the normal evening drink is sufficient to occupy both receptor types and to maintain ACTH levels with

small amplitude changes over the 24-h period. These GR-mediated effects observed after exogenous glucocorticoids thus are also involved in maintenance of HPA activity.

Elevated CORT facilitates the termination of the HPA response to stress, and various temporal (fast and slow feedback) domains have been distinguished. These GR-mediated effects triggered in response to stress represent the reactive mode of feedback operation.

Hippocampal Mineralocorticoid and Glucocorticoid Receptors Lesioning and electrical stimulation studies suggest an overall inhibitory influence of the hippocampus on HPA activity. The effects of intracerebroventricular MR antagonists in rats (see earlier discussion on the elevation of basal trough levels of CORT and enhanced adrenocortical responses to novelty) are likely mediated by hippocampal MRs. Consistent with the MR specificity of the response, a CORT implant in the dorsal hippocampus has been shown to suppress ADX-induced elevations in ACTH levels, whereas dexamethasone implants were ineffective. Finally, the systemic administration of spironolactone can increase basal HPA activity in humans, although this response has not been noted in all studies. Thus, hippocampal MRs appear to mediate the effect of CORT in maintaining the tone of basal HPA activity.

GR activation in the hippocampus, however, may not be involved in negative feedback but, rather, serves to suppress hippocampal output, leading to a disinhibition of the HPA axis. In fact, there is some experimental evidence that the intrahippocampal (in contrast to intracerebroventricular) administration of GR agonists does lead to an activation of the HPA axis. The nature of HPA regulation via hippocampal MRs and GRs is also consistent with the cellular actions they evoke in the hippocampus; predominant MR activation (comparable with local antiglucocorticoid application) maintains hippocampal excitability and, through transsynaptic inhibitory projections to the PVN, basal HPA activity. Conversely, with rising CORT concentration, GR activation suppresses the hippocampal output, resulting in a disinhibition of PVN neurons. The functions mediated by both receptor types are linked. A deficiency in an MR is predicted to allow a CORT response more readily, thus leading to more pronounced GR-mediated effects. This illustrates the importance of balance in MR- and GR-mediated effects involved in HPA regulation.

Indirect Effects Observations on HPA regulation have often been made without consideration of the

steroid effects on higher brain functions involved in arousal and processing of information. Brain areas projecting to the PVN have profound and long-lasting consequences for feedback regulation. Many systems with inputs to the PVN generally express GRs and contain numerous colocalized neuropeptides, which also regulate PVN activity in their own right.

The activation of a particular afferent neuronal network innervating the PVN area is stressor-specific and depends on the nature of the stimulus. If it constitutes a direct threat to survival through physical stressors (e.g., respiratory distress, hemorrhage, inflammation, infection, and trauma), the ascending aminergic pathways promptly activate the autonomic and neuroendocrine centers in the hypothalamus. If sensory stimuli are subject to appraisal and interpretation, processing in higher brain regions is required, which may subsequently lead to the modulation of GABAergic tone and a change in synthesis of CRH, AVP, and other neuropeptides of the PVN secretagogue cocktail. The activation of brain stem and limbic circuitry is not separated but is, in fact, mutually interactive because stress-induced CORT enters the brain readily and feeds back on all components of the neural stress circuitry, but in a context-dependent manner.

Control over Learning and Memory via Mineralocorticoid and Glucocorticoid Receptors

The activation of central corticosteroid receptors does not necessarily cause a behavioral change but, rather, influences information processing, thereby affecting the likelihood that a particular stimulus elicits an appropriate behavioral response. For example, the GR-mediated effects on the dopaminergic reward system may render animals more susceptible to addiction but by themselves cause no such state. Through coordinate MR- and GR-mediated actions in higher brain areas (e.g., the neocortical regions and limbic areas such as the hippocampus, septum, and amygdala), CORT affects learning and memory processes. Therefore, when studying the modulation of behavioral responses by CORT, the time and duration of the hormone action, as well as its context, need to be considered.

The hippocampal formation plays a key role in animals' reactivity to novelty and provides an essential contribution to learning and memory. Stress hormones, including CORT, are secreted during learning and are necessary for the establishment of an enduring memory. In this context, the role of CORT on the acquisition, consolidation, and retrieval of information has been studied for more than 4 decades. Many different tests revealed that the administration of exogenous CORT in the appropriate

temporal context (i.e., in close relation to training) potentiated memory in a dose-related fashion. From studies involving the use of specific corticosteroid receptor antagonists, antisense oligonucleotides, and transgenic animals, it is clear that these effects of corticosteroids require GR activation. These effects via GR in the context of stress-induced elevated hormone levels represent a reactive mode of action.

Studies with specific antagonists have shown that MRs modulate ongoing behavioral activity, as demonstrated by the effects of MR blockade during, but not after, the training session. These effects mediated by the MR regulate sensory integration underlying the evaluation of environmental information and response selection and thus have the capacity to subsequently affect memory storage for spatial and avoidance behavior. MR-mediated effects on behavioral response selection and reactivity may again be interpreted as a proactive mode of action, aimed to maintain homeostasis.

Thus, hippocampal MRs mediate the effects of CORT on the appraisal of information and response selection, whereas GR function does not so much modify these aspects of sensory integration as promote the processes underlying the consolidation of acquired information. However, exposure to high CORT levels in a different context (e.g., as a consequence of an additional stressor) may interfere with the consolidation processes because it may signal different, and more relevant, stimuli to the animal. The MR- and GR-mediated effects on information processing facilitate behavioral adaptation. This also promotes the inhibitory control exerted by higher brain circuits over HPA activity.

Corticosteroid Receptors in Pathology

The inappropriate activation of corticosteroid receptors through aberrant levels of hormone leads to well-known pathologies: Cushing's disease, immunosuppressed states (glucocorticoid excess), Addison's disease (glucocorticoid deficiency), and pseudomineralocorticoid excess states that are associated with high blood pressure. In the brain, the situation is much more complex. Depending on the region, both MRs and GRs can mediate the effects of CORT, and the balance in activation of these two receptor types may be critical in determining the effects of CORT on neuronal systems. For example, animal studies show that the balance of MR- and GR-mediated activities in hippocampal circuits determines excitability, stress responsiveness, and behavioral reactivity. Also, the effects of corticosteroid receptor activation often depend on the cellular context as a result of cross-talk with other signaling cascades. Genetically determined

differences in any of the signaling pathways involved can lead to a predisposition to adverse effects of CORT.

Significant changes in CORT signaling in the brain can result from chronic stress exposure and the inefficient operation of the HPA axis, resulting in over- or underexposure to circulating CORT, and from genetically determined or acquired changes in relative abundance of MRs and GRs (and associated factors). All these changes may lead to disturbances in mood and affect and in neurodegenerative phenomena via direct actions on relevant brain structures, as well as via the modulation of the HPA axis activity. However, the specific contribution of brain MRs and GRs to these cognitive and affective deficits has not been established.

Disturbed CORT signaling in brain appears to be causally involved in enhancing vulnerability to depression. In this context, it is extremely interesting that the tricyclic antidepressants not only restore 5-HT and noradrenergic transmission in brain but also increase the expression of brain corticosteroid receptors, in parallel with the normalization of HPA tone. Recent data suggest that the blockade of central GRs can be used as an effective treatment against psychotic major depression. In contrast to depression, posttraumatic stress disorder, chronic fatigue syndrome, and fibromyalgia are associated with a hypoactivity of the HPA axis.

Evidence from animal experiments and in humans shows that a dysregulated HPA axis, corticosteroid receptor deficits, and elevated CORT levels during stress are associated with age and may lead to an enhanced vulnerability to neurodegenerative disorders, particularly in the hippocampus. GR-mediated impairment in the energy metabolism of hippocampal pyramidal cells is thought to play a role here. The glucocorticoid cascade hypothesis has received much attention over the last years – hypercorticism leads, over time, to damage to the hippocampus, which in turn leads to the disinhibition of the HPA axis and further elevation of CORT levels. A central assumption of this hypothesis is that hippocampal GRs are thought to promote escape from negative feedback via downregulation. However, there is little evidence for a direct negative feedback function of hippocampal GRs; the role of the hippocampus appears to be indirect and involved more with the shutoff of the neural activity that drives HPA output, operating in concert with the output from the amygdala and pathways via the bed nucleus of the stria terminalis. Thus, the mechanisms that lead to the progressive hyperactivity and dysregulation of the HPA axis in some individuals as they age are evidently more

complex than envisioned in the glucocorticoid cascade hypothesis, and their elucidation will require a more extensive understanding of the role of other brain structures than the hippocampus and the hypothalamus in HPA regulation.

See Also the Following Articles

11 β -Hydroxysteroid Dehydrogenases; Catecholamines; Corticosteroids and Stress; Hypothalamic-Pituitary-Adrenal; Metabolic Syndrome and Stress; Mineralocorticoid Receptor Polymorphisms.

Further Reading

- Ahima, R., Krozowski, Z. and Harlan, R. (1991). Type 1 corticosteroid receptor-like immunoreactivity in the rat CNS: distribution and regulation by corticosteroids. *Journal of Comparative Neurology* **313**, 522–538.
- Arriza, J. L., Weinberger, C., Cerelli, G., et al. (1987). Cloning of human mineralocorticoid receptor cDNA: structural and functional kinship with the glucocorticoid receptor. *Science* **237**, 268–275.
- Belanoff, J., Rothschild, A., Cassidy, F., et al. (2002). An open label trial of C-1073 (mifepristone) for psychotic major depression. *Biological Psychiatry* **52**, 386–392.
- Chandler, V. L., Maler, B. A. and Yamamoto, K. R. (1983). DNA sequences bound specifically by glucocorticoid receptor in vitro render a heterologous promoter hormone responsive in vivo. *Cell* **33**, 489–499.
- Cole, T. J., Blendy, J. A., Monaghan, A. P., et al. (1995). Targeted disruption of the glucocorticoid receptor gene blocks adrenergic chromaffin cell development and severely retards lung maturation. *Genes & Development* **9**, 1608–1621.
- De Kloet, E. R., Vreugdenhil, E., Oitzl, M. S., et al. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Review* **19**, 269–301.
- De Kloet, E. R., Wallach, G. and McEwen, B. S. (1975). Differences in corticosterone and dexamethasone binding to rat brain and pituitary. *Endocrinology* **96**, 598–609.
- Edwards, C. R. W., Stewart, P. M., Burt, D., et al. (1988). Localization of 11 β -hydroxysteroid dehydrogenase: tissue specific protector of the mineralocorticoid receptor. *Lancet* **2**, 986–989.
- Funder, J. W., Pearce, P. T., Smith, R., et al. (1988). Mineralocorticoid action: target tissue specificity is enzyme, not receptor, mediated. *Science* **242**, 583–586.
- Fuxe, K., Wikström, A. C., Okret, S., et al. (1985). Mapping of the glucocorticoid receptor immunoreactive neurons in the rat tel- and diencephalon using a monoclonal antibody against rat liver glucocorticoid receptors. *Endocrinology* **117**, 1803–1812.
- Heck, S., Kullmann, M., Gast, A., et al. (1994). A distinct modulating domain in glucocorticoid receptor monomers in the repression of activity of the transcription factor AP-1. *EMBO Journal* **13**, 4087–4095.

- Herman, J. P. and Cullinan, W. E. (1997). Neurocircuitry of stress: central control of the hypothalamo-pituitary-adrenocortical axis. *Trends in Neuroscience* **20**, 78–84.
- Hollenberg, S. M., Weinberger, C., Ong, E. S., et al. (1985). Primary structure and expression of a functional glucocorticoid receptor cDNA. *Nature* **318**, 635–641.
- Joëls, M. and De Kloet, E. R. (1992). Control of neuronal excitability by corticosteroid hormones. *Trends in Neuroscience* **15**, 25–30.
- Jonat, C., Rahmsdorf, H. J., Park, K. K., et al. (1990). Antitumor promotion and antiinflammation: down-modulation of AP-1 (fos/jun) activity by glucocorticoid hormone. *Cell* **62**, 1189–1204.
- Karst, H., Karten, Y. J., Reichardt, H. M., et al. (2000). Corticosteroid actions in hippocampus require DNA binding of glucocorticoid receptor homodimers. *Nature Neuroscience* **3**, 977–978.
- Karst, H., Berger, S., Turiault, M., et al. (2005). Mineralocorticoid receptors are indispensable for nongenomic modulation of hippocampal glutamate transmission by corticosterone. *Proceedings of the National Academy of Sciences of the USA* **102**, 19204–19207.
- Krozowski, Z. S. and Funder, J. W. (1983). Renal mineralocorticoid receptors and hippocampal corticosterone-binding species have identical intrinsic steroid specificity. *Proceedings of the National Academy of Sciences USA* **80**, 6056–6060.
- Liu, W., Wang, J., Sauter, N. K., et al. (1995). Steroid receptor heterodimerization demonstrated *in vitro* and *in vivo*. *Proceedings of the National Academy of Sciences USA* **92**, 12480–12484.
- McEwen, B. S. (1999). Stress and hippocampal plasticity. *Annual Review of Neuroscience* **22**, 105–122.
- McEwen, B. S., Weiss, J. M. and Schwartz, L. S. (1968). Selective retention of corticosterone by limbic structures in rat brain. *Nature* **220**, 911–912.
- Moguilewsky, M. and Raynaud, J. P. (1980). Evidence for a specific mineralocorticoid receptor in rat pituitary and brain. *Journal of Steroid Biochemistry* **12**, 309–314.
- Pearce, D. and Yamamoto, K. R. (1993). Mineralocorticoid and glucocorticoid receptor activities distinguished by non-receptor factors at a composite response element. *Science* **259**, 1161–1165.
- Reichardt, H. M., Kaestner, K. H., Tuckermann, J., et al. (1998). DNA binding of the glucocorticoid receptor is not essential for survival. *Cell* **93**, 531–541.
- Reul, J. M. H. M. and De Kloet, E. R. (1985). Two receptor systems for corticosterone in rat brain: microdistribution and differential occupation. *Endocrinology* **117**, 2505–2511.
- Schinkel, A. H., Wagenaar, E., Van Deemter, L., et al. (1996). Absence of the *mdr1a* P-glycoprotein in mice affects tissue distribution and pharmacokinetics of dexamethasone, digoxin, and cyclosporin A. *Journal of Clinical Investigations* **96**, 1698–1705.
- Trapp, T. and Holsboer, F. (1996). Heterodimerization between mineralocorticoid and glucocorticoid receptors increases the functional diversity of corticosteroid action. *Trends in Pharmacology* **17**, 145–149.
- Van Steensel, B., Van Binnendijk, E. P., Hornsby, C. D., et al. (1996). Partial colocalization of glucocorticoid and mineralocorticoid receptors in compartments in nuclei of rat hippocampal neurons. *Journal of Cell Science* **109**, 787–792.
- Veldhuis, H. D., Van Koppen, C., Van Ittersum, M., et al. (1982). Specificity of adrenal steroid receptor system in the rat hippocampus. *Endocrinology* **110**, 2044–2051.
- Yudt, M. R. and Cidlowski, J. A. (2002). The glucocorticoid receptor: coding a diversity of proteins and responses through a single gene. *Molecular Endocrinology* **16**, 1719–1726.

Corticosteroid-Binding Globulin (Transcortin)

B E P Murphy

McGill University, Montreal, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by B E P Murphy, volume 1, pp 547–556, © 2000, Elsevier Inc.

Binding of Corticosteroid-Binding Globulin to Cell Membranes

Measurement of Corticosteroid-Binding Globulin

Use of Corticosteroid-Binding Globulin to Measure Steroids by Competitive Protein Binding

Glossary

Affinity

The degree of association between two entities; expressed as the ratio of complexed to free reactants at equilibrium (the affinity constant), which is the reciprocal of the dissociation constant.

Discovery and Definition of Corticosteroid-Binding Globulin

Biological Role of Corticosteroid-Binding Globulin

Physical Characteristics of Corticosteroid-Binding Globulin

Physiological Variation

<i>Binding capacity</i>	An expression of the number of binding sites available for a ligand.
<i>Corticosteroids</i>	Steroids made by the adrenal cortex.
<i>Ligand</i>	An entity bound by noncovalent forces to another (usually larger, usually protein) entity.
<i>Steroids</i>	Molecules containing the perhydrocyclopentanophenanthrene nucleus, including sterols, bile acids, sex hormones, adrenocortical hormones, cardiac glycosides, and sapogenins.

Discovery and Definition of Corticosteroid-Binding Globulin

The presence of a human plasma protein binding corticosteroids with high affinity ($5-8 \times 10^7 \text{ M}^{-1}$) and low capacity (10^{-9} – 10^{-8} M) was discovered in 1956 independently by I. E. Bush and by W. H. Daughaday. In 1958, A. A. Sandberg and W. R. Slaunwhite discovered a similar protein, which they called transcortin and which turned out to be identical with corticosteroid-binding globulin (CBG). Similar proteins have been found, many by U.S. Seal and R. P. Doe, in almost all species of vertebrates. A CBG-like protein has even been described in the eukaryotic fungus *Candida albicans*, with a K_a of $1.6 \times 10^8 \text{ M}^{-1}$ at 0°C , which might represent a primitive form of either the mammalian glucocorticoid receptor or of CBG.

CBGs differ from species to species in their relative affinities for various steroids. Most bind cortisol and corticosterone the most strongly; many bind progesterone very strongly; they bind aldosterone weakly. The guinea pig possesses a second corticosteroid-binding protein – progesterone-binding globulin (PBG) – in addition to CBG; PBG rises more than 100-fold during pregnancy. Although CBG binds testosterone and estradiol weakly, many species have, in addition to CBG, a sex hormone-binding globulin (SHBG), which binds these and related steroids more strongly. CBG is very low or undetectable in the plasma of the spiny anteater (an egg-laying mammal) and of the New World squirrel monkey.

Biological Role of Corticosteroid-Binding Globulin

Serum proteins such as albumin and CBG increase the solubility in biological fluids of steroids that are otherwise poorly soluble in aqueous solutions. However, the limited capacity and high affinity of CBG for glucocorticoids and progesterone has led to the accepted model of CBG action in which CBG provides a plasma store of tightly bound steroid that is less able to diffuse into cells, whereas free or unbound steroid

passes freely across cell membranes to initiate hormone action. Because the binding to CBG is rapidly reversible, steroid may quickly be transferred from the bound to the free state. Thus CBG is a major determinant regulating the concentration of the hormones that it binds. On exposure to serine protease from inflammatory granulocytes, CBG is cleaved and releases all or most of its bound steroid; thus, it seems likely that this process may permit the enhanced delivery of cortisol to sites of inflammation. The binding of CBG to cell membranes and entry into cells may also be important in its action.

Changes of CBG in fetal life occur as a result of normal development, possibly promoting the delivery of glucocorticoids to sites of rapid growth and tissue remodeling.

In the human, the unbound fraction of cortisol is usually approximately 10% of total cortisol; in Cushing's syndrome, in which CBG levels are decreased and total cortisol levels increase, this fraction rises to approximately 25% of the total, whereas in adrenal insufficiency it falls to approximately 6% of total cortisol.

Physical Characteristics of Corticosteroid-Binding Globulin

Structure

The molecular weight of CBG is 50–60 kDa, and there is one cortisol-binding site per molecule. The degree of sequence similarity among species is 60–70%. CBG is a monomeric glycoprotein (with the possible exception of New World primate CBGs, which may circulate as dimers) consisting of approximately 380 amino acids with a molecular weight of approximately 42 000 (Figure 1). Relatively large differences in the apparent molecular size among species is probably due to variations in carbohydrate composition. The human CBG sequence is thought to contain six consensus sites for N-glycosylation. Only two of these sites have been retained throughout evolution, and these are located in highly conserved regions and may therefore be functionally important. The gene for CBG is localized to chromosome 14. The deduced primary structure of CBG suggests a greater than 30% homology with the SERPIN (serine protease inhibitor) superfamily of proteins. This family includes α_1 -proteinase inhibitor (A1-P1), α_1 -antichymotrypsin (ACT), and thyroxine-binding globulin. A1-P1 and ACT are located close to the region containing the human CBG gene. The PBG of guinea pig has a molecular weight of approximately 80 000 kDa and has one binding site per molecule. Its properties have been reviewed by Westphal.

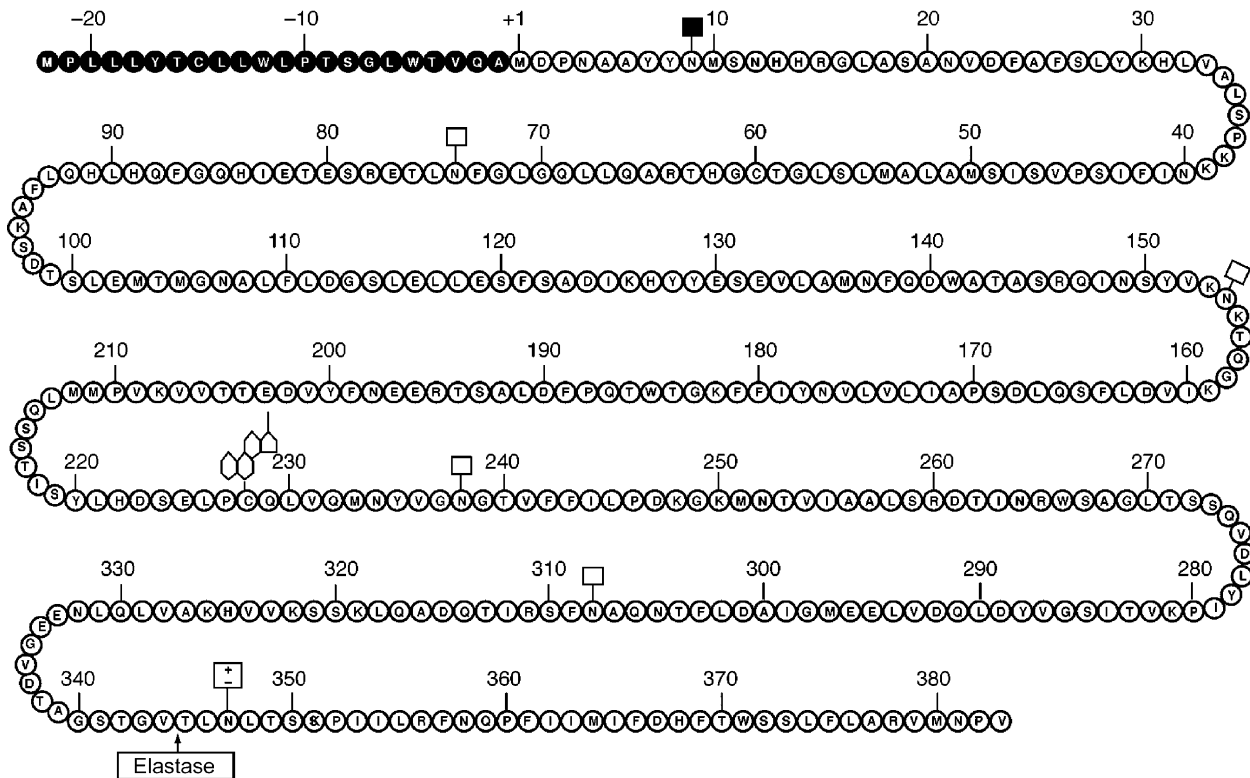


Figure 1 Primary structure of the human CBG precursor. The N-terminus of the mature polypeptide is inferred from published data, and the proposed signal peptide comprises 22 amino acids, shown in shaded circles. Potential sites for N-glycosylation are shown as squares, with a filled square indicating the site that is known to be used, and $\pm$ representing partial utilization. The elastase cleavage site is marked by an arrow, and the cysteine residue, which is conserved between species, is indicated. From Hammond, G. L. (1990), *Molecular properties of corticosteroid binding globulin and the sex-steroid binding proteins*, *Endocrine Reviews* 11, 65–79. Copyright 1990, The Endocrine Society.

Biosynthesis

The liver is the main site of CBG synthesis, and CBG is synthesized by hepatocytes in culture. CBG mRNA has been detected in the extrahepatic tissues (the spleen, ovary, testis, and lung) of various species but only in low amounts, and their significance is uncertain. CBG production is regulated independently in mother and fetus. Guinea pig PBG is made in only very small amounts by the liver, and the large rise in pregnancy is due mainly to synthesis by the placenta.

Degradation of Corticosteroid-Binding Globulin

The half-life of human CBG is approximately 5 days under usual circumstances but is decreased with severe stress. The serum clearance in early postnatal life in the rat is approximately twice as rapid (50% in ~ 7 h) as in the adult animal (~ 14 h).

Temperature Dependence

CBG binding in all species is temperature dependent, being maximal at approximately 4°C . At temperatures above 60°C , it is rapidly destroyed (in ~ 20 min).

Stability is increased at higher cortisol concentrations. Between 37 and 40°C there is a rapidly increasing rate of dissociation, which may be important physiologically in febrile states, resulting in a more rapid release of cortisol to the tissues. Similarly, local elevation of temperature as a consequence of inflammation may contribute to more available cortisol at these sites. Binding activity is usually retained after a single freezing but is lost with repeated thawing and freezing. Many CBGs in serum are stable for long periods at 4°C ; human CBG is stable for months or years provided bacterial contamination is prevented (e.g., by the addition of 0.1% sodium azide). Guinea pig PBG is more resistant to heating than most CBGs.

Affinity

The affinity constant for the binding of cortisol to human CBG is approximately $1 \times 10^9 \text{ M}^{-1}$ at 4°C and $5 \times 10^7 \text{ M}^{-1}$ at 37°C . The CBGs of most other species are in the range of 10^7 – 10^9 M^{-1} at 4°C and are also lower at 37°C . A relatively high association constant ($0.9 \times 10^8 \text{ M}^{-1}$) has been observed for ovine plasma at 37°C .

pH

Binding activity is also dependent on pH; for human CBG, it is maximal at approximately pH 8, but irreversible denaturation takes place below pH 5. In the guinea pig, PBG is more resistant to low pH than is CBG.

Electrophoresis

Electrophoretically, CBG behaves as an α_1 -globulin.

Relationship to Albumin Binding

Steroids are also bound by albumin in all species. The free fraction is usually 6–14% of the total, with the CBG-bound fraction being 67–87% and the albumin-bound fraction 7–19%. In the squirrel monkey, approximately one-half is free and one-half is bound to albumin; this is the only vertebrate species in which CBG appears to be virtually absent.

Effect of Reducing Agents

The binding of cortisol and progesterone to purified human CBG is reversibly decreased by adding strong reducing agents such as dithiothreitol and β -mercaptoethanol but not by the milder sodium ascorbate. Only the affinity is decreased; the number of binding sites remains the same.

Binding Specificity

Mammals Although the specificity and affinity of the steroid-binding site vary among species, maximal binding usually coincides with the biologically more important glucocorticoid (either cortisol or corticosterone) in a given species. The ratio of cortisol to

corticosterone varies widely among species, and the relative binding affinities for the two steroids also differ widely. In the green iguana there is high binding of cortisol (3056 nmol l⁻¹) but no detectable binding of corticosterone, whereas in the porcupine fish the binding to corticosterone (1250 nmol l⁻¹) is much higher than that of cortisol (67 nmol l⁻¹). For human plasma, Dunn et al. have calculated the binding distribution of endogenous steroids in normal men and in nonpregnant and pregnant women (Table 1).

In many species, progesterone, 11-desoxycortisol and desoxycortisol are also bound to CBG (Table 2). Alterations in ring A of strongly binding steroids cause a drastic drop in binding affinity; this is less marked in the CBG of avian species and in PBG.

Progesterone-binding globulin of guinea pig Whereas in human and many other species, CBG binds progesterone strongly, the guinea pig is the only species known to possess PBG, a separate CBG-like protein, which binds progesterone almost exclusively.

Marsupials The affinity of CBG in marsupials is similar to those of mammals, but the binding site concentrations are generally lower.

Synthetic steroids Some synthetic steroids, such as prednisolone, which is a glucocorticoid, and danazol (WII 131), which is not, also bind to CBG and compete with cortisol for binding sites. Dexamethasone, another synthetic glucocorticoid, is not bound to human CBG, but it does bind to chicken and dog CBG. Bovine serum is unusual in that, in addition to

Table 1 CBG-bound endogenous steroids in human plasma^a

Steroid	Men		Women					
	nmol/l ^b	% ^c	Late pregnancy nmol/l	%	Follicular nmol/l	%	Luteal nmol/l	%
Cortisol	400	90	740	95	400	90	400	90
Cortisone	72	38	110	5	54	39	54	38
Corticosterone	12	78	48	90	7.0	78	13	78
4-Androstene-3,17-dione	4.1	1.4	13	2.3	5.4	1.4	5.4	1.3
5-Pregnen-3 β -ol-20-one	2.2	0.4	15	0.4	2.2	0.2	4.4	0.2
17-Hydroxyprogesterone	1.8	41	14	66	1.8	42	5.9	41
Androsterone	1.5	0.5	1.5	1.4	1.5	0.5	1.5	0.5
Testosterone	1.3	3.4	4.7	4.7	1.3	2.3	1.3	2.2
Etiocholanolone	1.2	0.4	1.2	1.3	1.2	0.5	1.2	0.4
Aldosterone	0.4	21	5.8	41	0.2	22	0.5	21

^aData from Dunn, J. F., Nisula, B. C. and Rodbard, D. (1981), Transport of steroid hormones: binding of 21 endogenous steroids to both testosterone-binding globulin and corticosteroid-binding globulin in human plasma, *Journal of Clinical Endocrinology and Metabolism* **53**, 58–68.

^bTotal concentration.

^cPercentage bound to CBG.

Table 2 Steroid competition for sites on CBG in various species and on PBG of pregnant guinea pig^a

Species	Plasma concentration (%)	Steroid added ^b							
		F	B	S	P	H	E	T	A
Human	2.5	100	82	89	20	73	16	4	3
Rhesus monkey	0.5	20	100	17	43	7	4	6	2
Dog	1.25	100	54	85	9	44	44	5	5
Cat	1.25	100	80	104	20	50	80	<1	<1
Beef	12.5	100	57	60	29	27	68	15	3
Rabbit	1.25	100	40	56	16	19	64	14	2
Rat	0.5	30	100	17	27	2	9	9	7
Mouse	0.5	12	100	15	6	15	0	0	0
Chicken	2.5	100	83	125	89	37	270	10	9
Horse	1.0	100	30	28	2	22	2	0.5	<1
Guinea pig ^c	0.025	<0.5	<0.5	0.9	100	0.6	<0.5	8	<0.5

^aThe tracer used was ³H-corticosterone for all but pregnant guinea pig serum, in which ³H-progesterone was used. The relative binding activity of the steroid added was determined at 50% displacement of the steroid indicated at 100%.

A, aldosterone; B, corticosterone; E, cortisone; F, cortisol; H, 17-hydroxyprogesterone; P, progesterone; S, 11-desoxycortisol; T, testosterone.

^bDetermined at 4–10°C, using Florisil as adsorbent (which gives slightly different values from equilibrium dialysis).

^cPBG; serum from guinea pig in late pregnancy was used.

a CBG, which binds cortisol strongly, it contains a protein that binds triamcinolone acetonide, a synthetic glucocorticoid, approximately 50 times more strongly than cortisol.

Physiological Variation

In the human, the CBG binding capacity of men and women is similar (600–700 nM) and changes little from ages 1–60 years. Plasma levels of CBG vary from 10 to 500 nM in most vertebrate species. They are exceptionally high in the green iguana (3000 nM) and squirrel (2500 nM), and they are low in New World monkeys – only 1–10% of those in Old World primates and humans. The affinity is also lower, resulting in a much higher fraction of unbound plasma cortisol, possibly to compensate for the target organ resistance to glucocorticoids, which characterizes this group. In the squirrel monkey, the only mammal known that appears to lack CBG, half the cortisol is albumin-bound and half is free. In rat and guinea pig, female CBG levels are approximately twice those of males, a difference that is abolished by neonatal gonadectomy.

The concentration of CBG in the lymph is slightly lower than that in the blood, whereas concentrations in the synovial fluid and pleural effusion are higher and those in the spinal fluid and amniotic fluid are very low. CBG passes into milk – in human milk the concentration has been reported to be 1–15% that of plasma and in rat milk approximately 25% that of plasma. A CBG-like protein that binds progesterone strongly has been reported in human breast cyst

fluids. Because there is little or no CBG or albumin in saliva, salivary cortisol is often used as an index of free cortisol; however, the salivary gland contains 11-hydroxysteroid dehydrogenase type 2, which converts approximately half the cortisol to cortisone, so salivary values are approximately one-half those of free cortisol in plasma.

Pregnancy

In human pregnancy, CBG levels rise to levels more than twice those of nonpregnant women. The affinity is unchanged. The increase is due to the increased levels of estrogens in pregnancy. Progesterone binding, both free and total, increases steadily until the end of pregnancy; the free fraction rises from 6% of total at week 24 to 13% at week 40.

Elevated levels of CBG, up to 30-fold, have been observed by Seal and Doe to occur during pregnancy in mammals with a hemochorial type of placentation (rodents, cats, dogs, moles, shrews, manatees, and most primates) but not in those with other types such as endotheliochorial, choriovitelline, epitheliochorial, or syndesmochorial; the biological significance of this difference is not known. The increase in CBG is minimal in the rat (see Figure 2). In the guinea pig, CBG rises three- to fourfold, whereas PBG rises more than 100-fold. PBG differs from CBG in that it is not stimulated by estrogen.

Fetus and Neonate

Whereas fetal corticosteroid levels rise in many species during pregnancy, CBG levels vary considerably with species; in many, the levels fall rapidly

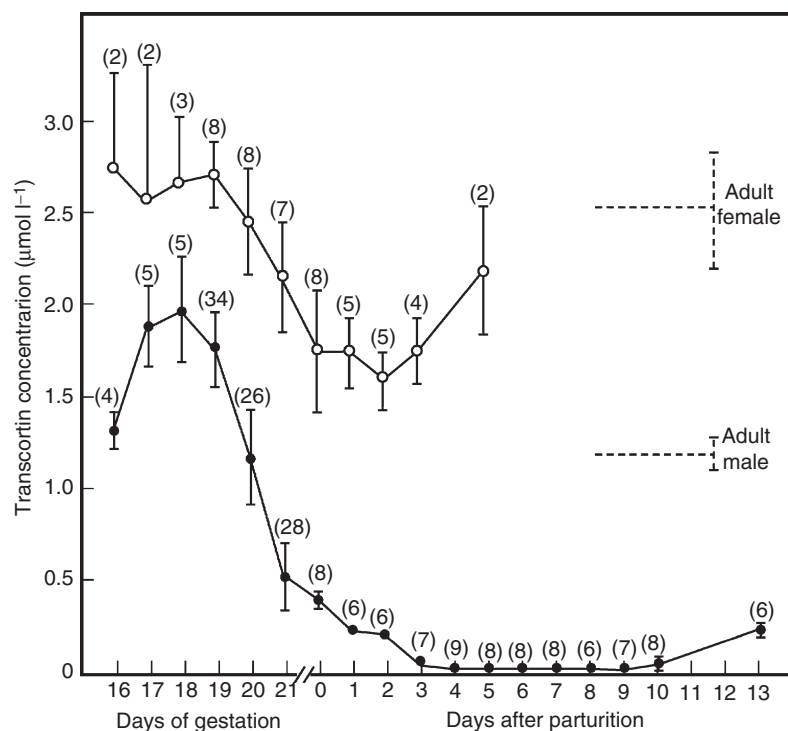


Figure 2 Serum CBG concentrations in fetal and neonatal rats (●) and in pregnant and lactating rats (○). The concentrations are expressed as means $\pm$ SD, and the number of animals or pools (days 16–18 for the fetuses) is given in parentheses. The normal serum concentration of four female and four male rats on day 154 of life is indicated on the right. From Van Baelen, H., Vandoren, G. and DeMoor, P. (1977), Concentration of transcortin in the pregnant rat and its foetuses, *Journal of Endocrinology* **75**, 427–431. Reproduced by permission of the Society for Endocrinology.

postpartum. In most species, early postpartum fetal CBG levels are lower than adult values.

Compared to the adult (700 nmol l^{-1}), CBG levels are lower in the human fetus and neonate, approximately 250 nmol l^{-1} in late gestation and immediately postpartum; rise to approximately 500 nmol l^{-1} by 30 days postpartum; and rise to adult levels by several months of age.

The level of CBG in the fetal rat rises from approximately 1.3 mol l^{-1} at 16 days gestation to a maximum of approximately $2 \mu\text{mol l}^{-1}$ at 28 days, after which it falls rapidly and subsequently rises slowly to adult values by 6 weeks (Figure 2). Thus CBG production is apparently independently regulated in mother and fetus. The decline in fetal rat CBG before birth may promote an increase in free corticosterone levels necessary for lung maturation. Despite the low neonatal values of CBG in rats, hippocampal glucocorticoid receptor occupancy/translocation (i.e., hormone-bound receptor) was noted by Viau et al. to be similar at all ages, both under basal conditions and following stress.

The sheep differs from this pattern; fetal CBG levels increase markedly over the last 2 weeks before term, from 50 nmol l^{-1} cortisol bound to 250 nmol l^{-1}

at term, whereas unbound and albumin-bound cortisol increase from 3 to 58 nmol l^{-1} and the total cortisol rises 13- to 15-fold. During this time, no changes occur in the maternal plasma. After birth, the CBG binding capacity of the neonate falls rapidly to about 90 nmol l^{-1} by 4 days and to 17 nmol l^{-1} by 14 days; total and free corticoids also decrease, but more slowly. The increase of CBG in fetal sheep plasma before term has been shown to depend on an intact fetal pituitary. The affinity of fetal CBG is similar to that of adult CBG. Possibly the rise in CBG contributes to the onset of parturition because the administration of glucocorticoids induces parturition in this species.

In the baboon, fetal serum CBG capacity at 100–132 days gestation is similar to that of the mother and decreases at term by half. Fetal serum cortisol is low ($\sim 100 \text{ nmol l}^{-1}$) before 168 days gestation and reaches approximately 1400 nmol l^{-1} after term delivery. Affinity constants are the same throughout.

Hereditary Deficiency

Hereditary deficiency has been shown in a few human families, in which the concentration was reduced to approximately one-half the normal. However,

familial low levels of CBG are rare and no convincing case of completely absent CBG in the human has been documented, suggesting that absence of CBG may be a lethal mutation, although cases with low CBG have no clinical apparent abnormalities. CBG appears to be identical in all human races. A variant of CBG, containing a slightly altered carbohydrate, occurs in pregnant women and accounts for 4–14% of total plasma CBG at term; its function is not yet clear.

Changes with Diet

Slight changes occur with diet; a high carbohydrate diet may result in a 20% fall in CBG.

Possible Fatal Drop in Corticosteroid-Binding Globulin

A remarkable change in CBG occurs in the male *Antechinus stuartii*, a shrewlike marsupial. In this species, all the males die within 3 weeks of the beginning of their first and only mating period in late winter. Over this period there is a precipitous fall in CBG levels, whereas the plasma corticosteroid concentrations remain high. No such changes occur in the females. Because these changes in the males result in a large increase in free cortisol, it is possible that the males die of hypercorticism.

Response of Corticosteroid-Binding Globulin to Stress

In humans, patients with septic shock or multitrauma have markedly low CBG levels in the early phase, with levels gradually increasing to normal by 7 days; in contrast, the free cortisol index shows high levels in the early phase, decreasing toward normal during the second phase. These observations suggest that CBG plays an active role by regulating availability to target tissues.

Acute severe stress in rats (e.g., immobilization for 1 h or tail shocks over 2 h) causes a fall in CBG, which is apparent by 6 h, maximal at 24 h, and persists for 24–48 h. Despite the low values of CBG in the neonatal rat, in response to stress (restraint for 1 h or separation from the mother) the adrenocorticotrophic hormone (ACTH) and the free corticosterone levels increase, and this increase persists for 48 h. Such increases are associated with an increased glucocorticoid receptor occupancy/translocation similar to that of stressed adults.

In rats in which an inflammatory process was induced by a subcutaneous injection of turpentine, CBG levels had decreased by two-thirds, whereas the corticosteroid levels were very high. Although the high corticosteroid levels might be expected to decrease CBG levels, the effect was still present after

adrenalectomy, suggesting that changes in glucocorticoid concentration are not necessary for this change. At the same time as the CBG levels were low, the levels of haptoglobin, an acute-phase serum protein rose 6- to 10-fold. CBG levels in pregnant rats treated similarly showed a threefold decrease of CBG, whereas their 19-day-old embryos were not affected.

Alexander and Irvine found that social stress in horses caused a decrease in CBG binding capacity (horses new to a herd were confined in a small yard with dominant resident horses for 3–4 h daily for 3–14 days). This decrease in CBG, seen after 3–4 days and increasing over the 2-week period, was accompanied by an increase in free cortisol, so that the total cortisol concentrations remained unaltered.

In comparing the homology of CBG cDNA, it was found that there is no sequence homology with any steroid hormone receptor. However, the homology with the SERPIN superfamily of proteins exceeds 30%. The proteins in the SERPIN superfamily share a typical tertiary structure responsible for the marked inhibition of serine proteases. In contact with neutrophil elastase, released by neutrophils in large amounts at sites of inflammation, CBG is cleaved, resulting in a 10-fold fall in its affinity for cortisol and thereby releasing its cortisol. Such a mechanism may permit the free cortisol to remain in high concentration at sites of inflammation while maintaining basal free cortisol concentrations at other sites.

Response of Corticosteroid-Binding Globulin to Other Pathological Factors in Humans

Increases in CBG also occur with the administration of oral contraceptive agents or other estrogen-containing preparations that increase synthesis. Tamoxifen and clomiphene also raise CBG levels. On the other hand, androgens such as testosterone decrease them; other anabolic steroids such as methandrostenedione and oxymetholone increase them slightly. CBG levels are reduced in patients with liver cirrhosis, nephrotic syndrome, and, in some cases, hypertension and hyperthyroidism, whereas elevated levels have been reported in some cases of acute leukemia, lymphoma, chronic active hepatitis, lung cancer, untreated adrenal insufficiency, adrenocortical hyperplasia, and disorders of male and female puberty. A slightly enhanced CBG capacity has been observed in juvenile-onset diabetics. A decreased CBG concentration, corrected by vitamin B₁₂ administration, has been observed in some cases of pernicious anemia.

Glucocorticoids decrease CBG levels, so increased CBG levels are found after adrenalectomy, whereas decreased levels are found in Cushing's syndrome (whether endogenously or exogenously produced).

A rapid fall in CBG occurs in septic shock, reaching a nadir at approximately 24 h due to a marked increase in rate of removal because the half-life of CBG is approximately 5 days; this may occur by proteolysis at sites of inflammation. In cases of shock due to severe acute infection by *Candida albicans*, the CBG activity may be virtually absent.

Binding of Corticosteroid-Binding Globulin to Cell Membranes

A number of reports suggest that CBG binds to various cell membranes, including prostate, epididymis, testis, mammary carcinoma cells, uterus, pituitary, kidney, spleen, muscle, lung, and placenta. Although many of these await confirmation, it appears that CBG does bind to some cell membranes with a K_D of approximately 0.9 μM at physiological temperatures. This binding requires calcium and/or magnesium ions and is negligible at 4 or 23°C. The physiological significance of this binding is still unknown. In rat white adipose tissue, del Mar Grasa et al. found that CBG was seen in a thin layer surrounding the cells; they suggested that CBG in this location may act as a protective barrier limiting the access of glucocorticoids to adipocytes.

Measurement of Corticosteroid-Binding Globulin

CBG is measured either by its binding capacity for glucocorticoids (described by Westphal), by radioimmunoassay, or more recently by enzyme-linked immunosorbent assay.

Use of Corticosteroid-Binding Globulin to Measure Steroids by Competitive Protein Binding

Human CBG was the first protein used for the determination of glucocorticoids by B. E. P. Murphy in 1963 using the principle of competitive protein binding (displacement analysis and radiotransinassay), which is identical to that of radioimmunoassay. In this type of analysis, the specificity of the assay depends on the binding properties of the binding protein and the conditions under which the assay is carried out. The endogenous steroid or ligand is quantitated according to its competition with a tracer ligand, usually a radioisotopic form of the endogenous ligand, for binding sites on the protein, and the displacement of the tracer steroid is estimated against a series of standards of the ligand being measured. Thus, for example, the binding of radiolabeled cortisol will

be gradually lowered by increasing amounts of endogenous cortisol, giving a standard curve.

Although no CBG is entirely specific for cortisol alone, the amounts of competing steroids in human plasma are usually small. In human pregnancy, progesterone is an important competitor, but it can be easily removed by extraction into hexane, which does not extract cortisol; the extracted progesterone can then itself be measured using CBG and a series of progesterone standards. Even better, progesterone may be measured using PBG of pregnant guinea pig serum, which has high specificity for progesterone and very little for cortisol or corticosterone. Because no protein is ever specific for a single ligand (and monoclonal antibodies are no exception), care must be taken to ensure that specificity is achieved under the conditions used. Because of the high affinity of CBG for cortisol, the sensitivity of such a method is very great, so that picogram amounts are readily measured. If chromatographic techniques are employed, other competing steroids can also be measured. Because different species of CBGs provide proteins with varying specificity and affinity, they can be used for the assay of many different steroids ([Table 1](#)). Maximum sensitivity is achieved by using a high dilution of the protein and a low volume for equilibration (usually 0.1 ml). Precision and accuracy are similar to those of radioimmunoassay.

See Also the Following Articles

Adrenal Cortex; Circadian Rhythms, Genetics of; Corticosteroids and Stress; Cushing's Syndrome, Medical Aspects; Cushing's Syndrome, Neuropsychiatric Aspects; Glucocorticoid Negative Feedback; Glucocorticoids, Overview; Glucocorticoids, Role in Stress.

Further Reading

- Alexander, S. L. and Irvine, C. H. G. (1998). The effect of social stress on adrenal axis activity in horses: the importance of monitoring corticosteroid-binding globulin capacity. *Journal of Endocrinology* 157, 425–432.
- Ballard, P. L., Kitterman, J. A. and Bland, R. D. (1982). Ontogeny and regulation of corticosteroid-binding globulin capacity in plasma of fetal and newborn lambs. *Endocrinology* 110, 59–366.
- Beishuizen, A., Thijs, L. G. and Vermes, I. (2001). Patterns of corticosteroid-binding globulin and the free cortisol index during septic shock and multitrauma. *Intensive Care Medicine* 27, 1584–1591.
- Breuner, C. W. and Orchinik, M. (2002). Beyond carrier proteins: plasma binding proteins as mediators of corticosteroid action in vertebrates. *Journal of Endocrinology* 175, 99–112.

- Dunn, J. F., Nisula, B. C. and Rodbard, D. (1981). Transport of steroid hormones: binding of 21 endogenous steroids to both testosterone-binding globulin and corticosteroid-binding globulin in human plasma. *Journal of Clinical Endocrinology and Metabolism* **53**, 58–68.
- Hammond, G. L. (1990). Molecular properties of corticosteroid binding globulin and the sex-steroid binding proteins. *Endocrine Reviews* **11**, 65–79.
- Murphy, B. E. P. (1967). Some studies of the protein-binding of steroids and their application to the routine micro and ultramicro measurement of various steroids in body fluids by competitive protein-binding radioassay. *Journal of Clinical Endocrinology and Metabolism* **27**, 973–990.
- Rosner, W. (1991). Plasma steroid-binding proteins. *Endocrinology & Metabolism Clinics of North America* **20**, 697–720.
- Seal, U. S. and Doe, R. P. (1967). The role of corticosteroid-binding globulin in mammalian pregnancy. In: Martin, L., (ed.) *Proceedings of Second International Congress on Hormonal Steroids*, pp. 697–706. Amsterdam, New York: Excerpta Medica Foundation.
- Van Baelen, H., Vandoren, G. and DeMoor, P. (1977). Concentration of transcortin in the pregnant rat and its fetuses. *Journal of Endocrinology* **75**, 427–431.
- Viau, V., Sharma, S. and Meaney, M. J. (1996). Changes in plasma adrenocorticotropin, corticosterone, corticosteroid-binding globulin, and hippocampal glucocorticoid receptor occupancy/translocation in rat pups in response to stress. *Journal of Neuroendocrinology* **8**, 1–8.
- Westphal, U. (1971). *Steroid-protein interactions*. New York: Springer-Verlag.
- Westphal, U. (1986). *Steroid-protein interactions II*. New York: Springer-Verlag.

Corticosteroids and Stress

A Munck

Dartmouth Medical School, Lebanon, NH, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A Munck, volume 1, pp 570–577, © 2000, Elsevier Inc.

Glucocorticoids, Mineralocorticoids, and Stress

Basal and Stress-Induced Glucocorticoid Levels: Functions in Stress?

Corticosteroid Receptors

Permissive and Suppressive Glucocorticoid Actions

Metabolic Actions

Cardiovascular Actions

Actions on Electrolyte and Fluid Volume Balance

Anti-inflammatory and Immunosuppressive Actions

Actions on the Central Nervous System

Effects via Cytokines and Other Mediators

Unified View of the Protective Roles of Glucocorticoids in Stress

Glossary

11 β -Hydroxy-steroid dehydrogenase type 2 (11 β -HSD2) An enzyme in mineralocorticoid target cells (e.g., the kidney and some brain cells) that inactivates glucocorticoids, thus allowing aldosterone access to the mineralocorticoid receptors; may also

Adrenocorticotrophic hormone (ACTH) Cytokines

Glucocorticoid receptors (GRs)

Glucocorticoids (GCs)

Hypothalamic-pituitary-adrenal (HPA) axis

regulate glucocorticoid concentrations in some glucocorticoid target cells.

A peptide that controls the synthesis and secretion of glucocorticoids from the adrenal cortex.

Peptide hormones, produced particularly by activated cells of the immune system, that transmit signals between cells and organs via specific cell membrane receptors.

Protein molecules with a relatively low-affinity ($K_d \sim 40$ nM) specific binding site for glucocorticoids that are present in the cytoplasm and nucleus of almost all nucleated cells; bind to chromatin and regulate gene expression when activated by ligand binding.

Cortisol and corticosterone, the steroidal stress hormones produced by the adrenal cortex under the control of ACTH; they act on almost all nucleated cells, influencing innumerable metabolic and physiological functions. The term also refers to synthetic analogs such as dexamethasone.

The physiological link between the central nervous system and adrenocortical production of GCs. In response to signals including stress, decreased feedback from GCs, and certain cytokines, the hypothalamus produces the peptides

Mineralocorticoid receptors (MRs)

corticotropin releasing hormone and vasopressin, which regulate the pituitary output of ACTH.

Protein molecules with high-affinity ($K_d \sim 2$ nM) specific binding site for aldosterone, cortisol, and corticosterone; present in the cytoplasm and nucleus of the kidney, brain, and some other cells. It binds to chromatin and regulates gene expression when activated by ligand binding.

Mineralocorticoids

Aldosterone, the steroid hormone produced in the adrenal cortex under control of angiotensin II, which stimulates mainly kidney cells to reabsorb sodium and water; also used for the much weaker natural steroid deoxycorticosterone and for synthetic analogs.

Glucocorticoids, Mineralocorticoids, and Stress

Although Thomas Addison's discovery of the essential role of the adrenal gland for survival dates back to the mid-nineteenth century, it was not until the 1930s that Hans Selye promulgated the physiological concept of stress and implicated the hormones of the adrenal cortex in protecting the organism against stress. By then Walter Cannon had established the role of the adrenal medulla and the sympathetic system in fight-or-flight responses, but the adrenal medulla, in contrast to the cortex, is not essential for life. An intimate relation between the adrenal cortex and stress became evident from observations that stress from all kinds of challenges to homeostasis – cold, infection, inflammation, trauma, hypoglycemia, psychological shock, emotional distress, pain, heavy exercise, hemorrhage, and others – stimulates the secretion of adrenocortical hormones and causes hyperplasia of the adrenal cortex. Furthermore, Addisonian patients, adrenalectomized animals, and other adrenally insufficient organisms were the least capable of resisting the sometimes fatal onslaught of stress, but were protected by the administration of adrenocortical extracts.

At first the adrenal cortex was thought to regulate a multitude of effects on metabolism and salt and water balance, along with stress, through a single hormone. Eventually, two principal types of steroid hormones were identified, the GCs and the mineralocorticoids. Each was found to be essential for survival, the mineralocorticoids by controlling electrolyte and water balance, mainly via actions on the kidney, and the GCs by protecting against stress through mechanisms that were to remain obscure for decades. This article

discusses the mechanisms by which GCs protect against stress from a modern perspective.

Basal and Stress-Induced Glucocorticoid Levels: Functions in Stress?

As shown schematically in [Figure 1](#), basal levels of cortisol in humans and primates (and of corticosterone in rats and mice) vary in a characteristic circadian pattern, rising during sleep to a maximum at the start of the waking period and then subsiding during the active period. Superimposed on this periodicity are erratic spikes in cortisol levels caused by brief episodes of secretion. Stress at any time causes a rapid increase in GC secretion to peak GC levels that may be several times higher than those attained under basal conditions.

Here, we discuss the questions: What role do basal GC levels play in protecting the organism against stress? Is there some significance to the pattern of circadian variation? What role do stress-induced GC levels play? Tentative answers to these questions are already intimated in [Figure 1](#). Basal levels of GCs exert mainly permissive actions that prepare or prime some of the organism's homeostatic defense mechanisms for action, reaching a peak of preparedness as the diurnal activities begin. Stress-induced levels probably function mainly to suppress such defense mechanisms in order to prevent them from overshooting and damaging the organism. To answer the questions posed more completely, it is necessary first to survey some of the basic mechanisms and physiological consequences of GC actions.

Corticosteroid Receptors

The intracellular receptors for the GCs and mineralocorticoids, the GRs and MRs, respectively, were

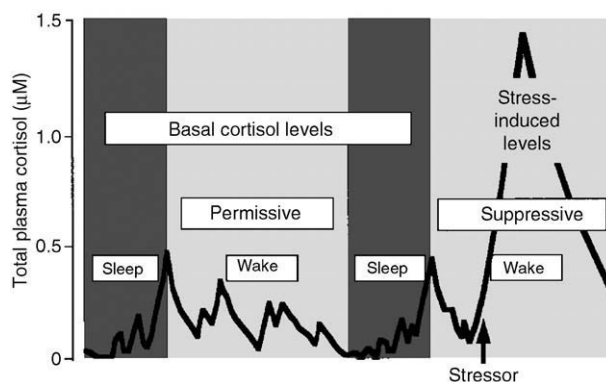


Figure 1 Schematic depiction of diurnal variation of basal cortisol levels in humans and a stress-induced surge. Permissive and suppressive refer to the roles proposed for each of these glucocorticoid regimens in protecting against stress.

identified during the 1960s and cloned in the 1980s. GRs were characterized by their presence in GC target cells and their ability to specifically bind radioactively labeled cortisol and related GCs. MRs were characterized in mineralocorticoid target tissues by their ability to bind labeled aldosterone. Cortisol and corticosterone, the natural GCs, bind to GRs with dissociation constants in the range of normal GC concentrations in the blood. Similarly, aldosterone binds to MRs with a dissociation constant in the range of normal aldosterone concentrations in the blood, which are 100–1000 times lower than GC concentrations.

In the late 1980s, cortisol and corticosterone were found to bind to MRs with almost the same dissociation constant as aldosterone, that is, with a much higher affinity than to GRs. That posed a major quandary – because GCs would normally be expected to flood mineralocorticoid target cells and saturate MR binding sites, how does aldosterone bind to MRs? The quandary was resolved with the identification in mineralocorticoid target cells of the enzyme 11β -HSD2, which protects the MRs by rapidly inactivating GCs to their 11-keto forms, allowing aldosterone access to MRs. MRs in certain cells are unprotected by 11β -HSD2 and respond to GCs at concentrations far lower than those required to activate the GRs. Some cells in male and female reproductive tracts have 11β -HSD2 that may protect them from excessive levels of GCs.

The role of MRs as mediators of GC actions has been studied most extensively in the central nervous system. Both GRs and MRs are found in the brain and spinal cord, MRs being particularly abundant in the dentate gyrus and pyramidal cells of the hippocampus as well as other regions of the limbic system. There appears to be no 11β -HSD2 in the hippocampus, so hippocampal MRs are bound and activated mainly by GCs. Only the MRs in the anterior hypothalamus and circumventricular organs are protected by 11β -HSD2 and are therefore controlled by aldosterone. GRs are more widely dispersed in the central nervous system than MRs and are present in both neurons and glial cells.

GRs and MRs, when activated by hormone binding, initiate hormone actions by translocating to the nucleus and activating or repressing particular genes. Most GC and mineralocorticoid actions are thought to be produced through such genomic mechanisms, which are relatively slow, usually taking hours to become manifest. There is mounting evidence, however, that certain rapid effects of these and other steroid hormones, particularly in the nervous system, may be mediated by nongenomic effects transmitted through membrane receptors.

Permissive and Suppressive Glucocorticoid Actions

In the 1950s, Dwight Ingle found that adrenalectomized animals can respond normally to moderate stress if administered GCs at basal levels and concluded that, although stress-induced levels are required to protect against severe stress, they may not be necessary for moderate stress. He proposed that basal GC levels have permissive effects that are necessary for some homeostatic functions to respond to stress, distinguishing those effects from the regulatory effects of high GC levels espoused by Selye. Selye had introduced a unified theory of stress and proposed the controversial hypothesis that, by increasing levels of adrenocortical hormones, stress causes such diseases of adaptation as rheumatoid arthritis and allergy.

Permissive actions of the GCs on many defense mechanisms have since been found. They are generally produced by basal levels of GCs and often enhance the actions of such mediators of stress-induced responses as catecholamines and other stress hormones. Permissive actions thus occur during the unstressed periods indicated in [Figure 1](#) and can be regarded as preparing for a stressful episode. They can account for the substantially improved resistance to moderate stress of organisms lacking adrenocortical function, including Addisonian patients, when given maintenance levels of GCs. Among several mechanisms for permissive actions is the induction by the GCs of receptors for mediators of stress responses. Some other ways GCs are known to increase the sensitivity of signaling pathways are by strengthening coupling of a receptor to G-proteins and by increasing cAMP synthesis.

The suppressive actions of GCs are better known than the permissive actions. They have been extensively investigated because of their close link to the widespread clinical applications of GCs as immunosuppressants, as anti-inflammatory agents, and for treatment of lymphoproliferative diseases. For many years after they were described in 1949, the anti-inflammatory and immunosuppressive actions were regarded as the pharmacological effects of great therapeutic value but no physiological significance. Usually they are manifested at relatively high, stress-induced GC levels. They are exerted through many mechanisms; one mechanism that is fairly general is the suppression of cytokines and other mediators of stress responses.

Metabolic Actions

Early observations that GCs raise and maintain blood glucose levels gave the hormones their name. They also stimulate hepatic glycogen deposition. Long,

Katzin, and Fry, in a classic 1940 paper, showed that GCs stimulate gluconeogenesis from amino acids derived from protein catabolism and decrease glucose oxidation. Ingle later found that GCs decrease glucose use and cause insulin resistance. Much work since then has elucidated some of the molecular mechanisms of these and other actions. In the case of gluconeogenesis, there is increased activity of phosphoenolpyruvate carboxykinase and other rate-limiting enzymes. The inhibition of glucose transport and insulin resistance are due to decreased translocation of glucose transporters to the cell membrane and decreased synthesis of insulin receptor substrate-1 (IRS-1). Liver glycogen is increased partly in response to the activation of glycogen synthase. Selye suggested that GCs protect against stress by providing an increased need for glucose. This idea did not survive experimental tests, which showed that glucose alone cannot substitute for GCs.

GCs are permissive for gluconeogenesis and glycolysis stimulated by catecholamines and glucagon, reactions that respond to stress in a fight-or-flight situation as well as to insulin-induced hypoglycemia. GC induction of glycogen synthesis can be regarded as a preparative action that anticipates future stressful episodes.

How GC actions on carbohydrate metabolism protect the organism may be viewed in two distinct ways. The ethological view, which considers how evolutionary pressures have molded GC actions for the survival of, for example, a confrontation between predator and prey, emphasizes the permissive role of GCs in acutely raising blood glucose levels to fuel muscular activity. The metabolic view, on the other hand, sees GCs, through their anti-insulin and permissive actions, as protecting the organism from insulin-induced hypoglycemia, as can occur after a meal. GCs share this role with other counterregulatory hormones, catecholamines and glucagon.

Cardiovascular Actions

It has long been known that Addisonian patients, adrenalectomized animals, and other organisms deficient in GCs or GC activity are relatively insensitive to the actions of vasoactive agents such as norepinephrine, vasopressin, and angiotensin II. GCs at basal levels permissively enhance that sensitivity in both vascular and cardiac tissues. These actions may be due partly to the enhancement of α_1 B and β_2 catecholamine receptors and angiotensin II type 1 receptors in smooth muscle cells, as well as increased sensitivity of the G-protein-coupling receptor and adenylyl cyclase. GCs also increase cardiac output, which may be related to induction in cardiocytes of Na,K-ATPase.

Despite evidence that GCs suppress catecholamine release, which counters the permissive actions, permissive enhancement clearly predominates in the cardiovascular stress response.

Actions on Electrolyte and Fluid Volume Balance

GC deficiency is associated with decreased ability to excrete water, which at one time was used as a test for GC insufficiency. It appears to be due to a decreased rate in glomerular filtration and an increased synthesis of vasopressin (VP), both of which are reversed by administered GCs. The underlying mechanisms are not known, but they may be related to the glucocorticoid-induced synthesis and secretion of atrial natriuretic peptide (ANP) and increased levels of ANP receptors on endothelial cells. The inhibition of VP synthesis is part of the negative feedback mechanism by which GCs regulate their own concentration and appears to be due to the decreased synthesis of VP.

Some of these actions are central to the role of GCs in stress. In rats, the stress of hemorrhage has been found to elicit a powerful vasoconstrictor response from VP, catecholamines, and renin. Without GCs, the VP response can overshoot, leading to ischemia and death in adrenalectomized animals. Stress-induced GCs protect the organism by limiting the increase in VP levels and hence the ischemia.

Anti-inflammatory and Immunosuppressive Actions

High levels of GCs suppress many functions. The inflammatory and immune systems are among those in which suppressive actions are most pronounced. Thymus enlargement in the absence of the adrenal cortex was recognized before 1900. Selye showed that stress causes thymus atrophy, a GC action now known to be due to the induction of lymphocyte apoptosis.

The anti-inflammatory actions of GCs, which account for most clinical applications of these hormones, were reported in 1949 by Hench, Kendall, Slocumb, and Polley, who found that treatment with GCs dramatically improved the symptoms of patients with rheumatoid arthritis. These observations had two unexpected consequences: (1) they contradicted Selye's hypothesis of diseases of adaptation, which rapidly lost support, and (2) GC physiologists were unable to reconcile the finding that GCs suppressed a defense mechanism, in this case inflammation, with their longstanding belief that stress-induced GC levels helped resist stress by enhancing activity of defense mechanisms. Despite an alternative view proposed by Marius Tausk (in a 1951 review that unfortunately had very

limited circulation), physiologists categorized the anti-inflammatory actions as the pharmacological consequences of high unphysiological levels of GCs. The closely related immunosuppressive actions suffered the same fate. As a result, the spectacular development of new applications and synthetic analogs for GCs, the miracle drugs of the 1950s, proceeded with little or no involvement by GC physiologists.

The view advanced in germinal form by Marius Tausk in the context of GC therapy was revived independently in 1984 in the context of GC physiology. It held that the anti-inflammatory and immunosuppressive actions of GCs, far from being pharmacological, are prototypes for the physiological role of stress-induced GCs in stress – that role is to prevent stress-induced defense reactions from overshooting and harming the organism.

A great deal of evidence now supports the view that anti-inflammatory and immunosuppressive actions are physiological and are exerted by endogenous GCs. In rats, for example, adrenalectomy or administration of the glucocorticoid antagonist RU486 increases inflammatory responses, showing that endogenous GCs can control inflammation. Autoimmune reactions are also under the control of endogenous GCs. Striking experiments have shown that Lewis rats, which are extremely sensitive to arthritis after injection with streptococcal cell wall polysaccharide (SCW) and to experimental allergic encephalomyelitis (EAE) after injection with myelin basic protein, have a defect in biosynthesis of corticotropin releasing hormone (CRH), the hypothalamic hormone that, along with VP stimulates ACTH, is released by the pituitary. That defect reduces the ACTH, and hence the GC, response to an antigenic challenge. As expected, treatment with GCs protects Lewis rats from SCW.

GCs may also influence immune responses through permissive actions. For example, mitogenic responses of rat peripheral or splenic T cells in culture are reduced if the rats have been adrenalectomized. The responses can be restored by treating the rats with low physiological doses of corticosterone, as well as by brief exposure of the cells to low concentrations of corticosterone or aldosterone. These observations led to the suggestion that the permissive effects in this case were mediated by MRs, and possibly involve induction of interleukin (IL)-2 receptors. The acute phase response is a case in which GCs both potentiate the induction of acute phase proteins by cytokines such as IL-1 and IL-6 and suppress the production of the cytokines.

The permissive actions of GCs at stress-induced levels were revealed by experiments in which normal humans were given an endotoxin to stimulate increases in plasma tumor necrosis factor (TNF)- α and IL-6. At

various times before or after the endotoxin, cortisol was infused at rates that gave stress-induced levels. When given within 6 h of the endotoxin, cortisol suppressed the cytokine response, but, when given 12, 32, 72, or 144 h before endotoxin, cortisol markedly enhanced the response.

Actions on the Central Nervous System

Electrophysiological experiments have been conducted with hippocampal tissue isolated from adrenalectomized rats and treated with GCs at various concentrations. Low concentrations of corticosterone, which mainly activate the MRs, were found to decrease after hyperpolarization of the neuronal membranes and enhance neuronal excitability, whereas high concentrations of corticosterone, or specific glucocorticoid agonists that activate the GRs but not MRs, suppressed excitability. Thus, GCs at low concentrations, corresponding to basal levels, permissively maintain neuronal excitability via the MRs, whereas GCs at high concentrations, corresponding to stress-induced levels, suppress stimulated neuronal activity via the GRs.

Effects via Cytokines and Other Mediators

As already indicated, suppressive actions of GCs are often due to the inhibition of production or activity of cytokines, the hormones of the immune system, and of other mediators. Such mediators constitute a communications network between cells of defense systems that respond to stress-induced challenges. Stress-induced mediators known to be suppressed by GCs are listed in [Table 1](#). Neurotransmitters, such as the catecholamines, respond to sudden fight-or-flight challenges; inflammatory agents respond to tissue damage; cytokines and chemokines respond to infection and inflammation; and so on. Mediators are essential for propagating stress responses. By blocking them, GCs can dampen the stress responses, preventing them from overshooting and damaging the organism. At the same time, GCs also protect the organism from toxicity of most of these mediators when they are present in excess.

Many well-known GC effects on cell traffic may be initiated through the inhibition of chemokines (also known as chemotactic cytokines), which regulate the traffic and homing of leukocytes by binding to cell-surface receptors and which are actively secreted during inflammation. Some mediators are suppressed by GCs under some conditions and stimulated under others. There are also mediators, such as macrophage colony stimulating factor (M-CSF), that appear not to be affected by GCs at all.

Table 1 Effects of glucocorticoids on mediators and mediator receptors^a

	Mediators suppressed	Mediators suppressed/ stimulated ^b	Mediator receptors induced
Cytokines	IL-1, IL-2, IL-3, IL-5, IL-6, IL-11, IL-12, IL-16, IFN- γ , TNF- α , GM-CSF	IL-4, IL-10, TGF- β	IL-1, IL-2, IL-4, IL-6, IFN- γ , GM-CSF
Chemokines	IL-8, MCP-1, RANTES, MIP-1 α , CINC/gro, LIX, SCF		
Inflammatory agents	5-HT, bradykinin, histamine, eicosanoids, Cox-2, collagenase, elastase, MMP-9, plasminogen activator		5-HT
Hormones and neurotransmitters	ACTH, CRH, VP, oxytocin, β -endorphin, LH, insulin, norepinephrine, nitric oxide, substance P	Epinephrine	CRH, VP, insulin, norepinephrine
Cell adhesion molecules	ICAM-1, LFA-1, e-selectin ELAM-1		

^a5-HT, serotonin; ACTH, adrenocorticotrophic hormone; CINC/gro, cytokine-induced neutrophil chemoattractant/growth-induced gene product; Cox, cyclo-oxygenase; CRH, corticotropin releasing hormone; ELAM, endothelial-leukocyte adhesion molecule; GM-CSF, granulocyte macrophage colony stimulating factor; ICAM, intercellular adhesion molecule; IFN, interferon; IL, interleukin; LFA, lymphocyte function-associated antigen; LIX, lipopolysaccharide-induced CXC chemokine; MCP, monocyte chemoattractant protein; MIP, macrophage inflammatory protein; MMP, matrix metalloproteinase; RANTES, regulated on activation, normal T-cell-expressed and secreted; SCF, stem-cell factor; TGF, tumor growth factor; TNF, tumor necrosis factor; VP, vasopressin.

^bMediators suppressed by GCs under some conditions and stimulated under others.

Table 1 lists several mediators that have receptors that are induced by GCs. The possibility that the induction of mediator receptors accounts for some permissive actions has already been noted. Some cytokines, such as IL-1 and IL-6, not only communicate between cells of the immune and inflammatory systems, but also, as first proposed by Besedovsky and Sorkin, may signal the HPA system that the immune system is activated, thereby constituting a negative feedback loop through which GCs are summoned to control activated immune reactions.

Unified View of the Protective Roles of Glucocorticoids in Stress

As described so far, GCs exert actions that are predominantly permissive on the cardiovascular system; that are mainly suppressive on other systems, such as inflammation; and that are both permissive and suppressive on still other systems, such as the neural and immune systems. How are these opposing influences reconciled when both are present, with one enhancing and the other suppressing the defense response? This question can be analyzed with a simple but fairly realistic model in which suppressive actions are assumed to be due to the reduction in the concentration of a mediator and permissive actions are due to the induction of the mediator's receptors, as illustrated in **Figure 2**.

In **Figure 2**, a lymphokine receptor (R) is permissively induced by cortisol over a concentration range corresponding roughly to that in **Figure 1** (keeping in mind that at low cortisol concentrations less than 10% of total cortisol is free), following a

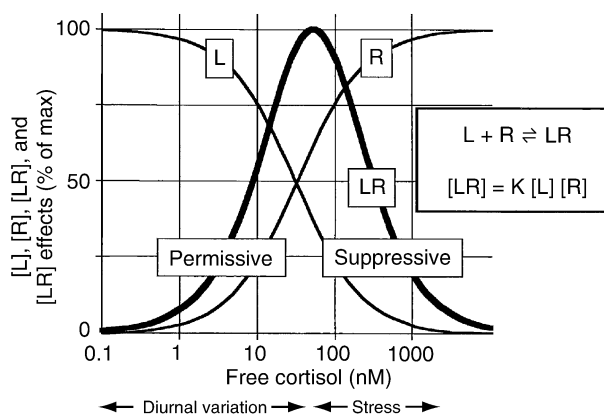


Figure 2 Model of a cortisol-regulated lymphokine system. Cortisol is assumed to permissively increase the concentration of lymphokine receptors (R) and to suppress the concentration of the lymphokine (L), according to dose-response curves that correspond to binding of cortisol to GRs with a dissociation constant of 30 nM. The concentration and activity of the lymphokine-receptor complex (LR) is determined by the equations in the inset, yielding the bell-shaped curve with a rising permissive branch at basal cortisol concentrations and a falling suppressive branch at stress-induced concentrations. Brackets denote concentration. Below the cortisol concentration scale the ranges covered during basal diurnal variation and stress are indicated roughly.

dose-response curve chosen to parallel the binding of cortisol to GRs. Lymphokine (L) is suppressed according to an identical dose-response curve of opposite sign. Through these actions, cortisol regulates the concentration of the lymphokine-receptor complex (LR), as determined by the equations in the inset. The bell-shaped curve for LR and LR activity is, then, a composite dose-response curve that represents

schematically the outcome of permissive and suppressive GC actions on the hypothetical defense mechanism that the lymphokine regulates (the simple curves for R and L represent the outcome for defense mechanisms subject to pure permissive and suppressive actions, respectively).

This model assumes that both permissive and suppressive actions are mediated by GRs. If permissive actions were mediated by MRs, the dose-response curve for induction of R and the ascending limb of the bell-shaped curve would be shifted to the left, as would the peak of the bell-shaped curve; similar qualitative conclusions would apply, however.

It should be evident that any permissive and suppressive actions with roughly comparable dose-response curves, whether transmitted via a lymphokine or other mechanism, will yield similar bell-shaped curves. Such curves correspond well with the physiological roles of GCs we have sketched earlier. Basal GC concentrations fluctuate during normal diurnal variation from the left end of the cortisol scale in [Figure 2](#) to approximately where the bell-shaped curve peaks and can be viewed as permissively priming some of the organism's defense mechanisms in preparation for a period of activity. Suppressive actions are present at all cortisol concentrations, but they predominate only at high stress-induced levels toward the right end of the cortisol scale, where they are required to prevent defense mechanisms from overshooting. In the absence of GCs, an organism may succumb to stress either because its defense mechanisms are not adequately primed (perhaps what happens in Addisonian crisis if the cardiovascular system cannot respond) or because the defense mechanisms overshoot (as occurred in the examples cited earlier after hemorrhage or an inflammatory challenge), depending on the nature of the stressor.

Thus, in summary, GCs can be thought of as protecting the organism from stress through two separate but related mechanisms. First, they are essential before a stressful episode to permissively prime certain defense mechanisms, a task generally accomplished by basal GC levels. Second, they are essential after a stressor to prevent some defense mechanisms from

overshooting, a task generally requiring stress-induced GC levels.

Further Reading

- Besedovsky, H. O. and del Rey, A. (1996). Immuno-neuro-endocrine interactions: facts and hypotheses. *Endocrine Review* 17, 64–102.
- de Kloet, R. E. (2003). Hormones, brain and stress. *Endocrine Regulation* 37, 51–68.
- Ingle, D. J. (1954). Permissibility of hormone action: a review. *Acta Endocrinologica* 17, 172–186.
- Makara, G. B. and Haller, J. (2001). Non-genomic effects of glucocorticoids in the neural system: evidence, mechanisms and implications. *Progress in Neurobiology* 65, 367–390.
- Munck, A., Guyre, P. M. and Holbrook, N. J. (1984). Physiological functions of glucocorticoids in stress and their relation to pharmacological actions. *Endocrine Review* 5, 25–44.
- Reichardt, H. M., Umland, T., Bauer, A., et al. (2000). Mice with an increased glucocorticoid receptor gene dosage show enhanced resistance to stress and endotoxin shock. *Molecular and Cellular Biology* 20, 9009–9017.
- Sapolsky, R. M., Romero, L. M. and Munck, A. (2000). How do glucocorticoids influence stress responses? integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrine Review* 21, 55–89.
- Selye, H. (1946). The general adaptation syndrome and the diseases of adaptation. *Journal of Clinical Endocrinology and Metabolism* 6, 117–230.
- Sternberg, E. M., Young, S. W., III, Bernardini, R., et al. (1989). A central nervous system defect in biosynthesis of corticotropin-releasing hormone is associated with susceptibility to streptococcal cell wall-induced arthritis in Lewis rats. *Proceedings of the National Academy of Sciences USA* 86, 4771–4775.
- Tausk, M. (1951). Hat die Nebenniere tatsächlich eine Verteidigungsfunktion? *Das Hormon (Organon, Holland)* 3, 1–24.
- Webster, J. I., Tonelli, L. and Sternberg, E. M. (2002). Neuroendocrine regulation of immunity. *Annual Review of Immunology* 20, 125–163.
- Wieggers, G. J., Labeur, M. S., Stec, I. E. M., et al. (1995). Glucocorticoids accelerate anti-T cell receptor-induced T cell growth. *Journal of Immunology* 155, 1893–1902.

Corticotropin Releasing Hormone (CRH)

A T Lim

Mental Health Research Institute of Victoria, Victoria,
Australia

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1,
pp 578–581, © 2000, Elsevier Inc.

Structure and Gene

Sites of Production and Pathways

Biochemical Actions and Physiological Functions

Glossary

*Adenylyl
cyclase-cyclic
adenosine
monophos-
phate (cAMP)
system*

An important component of a common signal transduction system that generates cAMP from adenosine triphosphate (ATP) following the activation of the membrane-bound enzymes adenylyl cyclases. The latter is triggered by subunits of G-proteins that are activated by receptors belonging to the superfamily of seven membrane-spanning receptors. The augmented level of intracellular cAMP in turn activates protein kinase A, which phosphorylates proteins to accomplish important cell functions.

G-proteins

A superfamily of proteins that act as intermediates between seven membrane-spanning receptors and membrane-bound enzymes to generate biochemical signals. Following activation by the receptors, the subunits of G-proteins augment or suppress the activity of adenylyl cyclases, phospholipase C, and ion channels through direct bindings to these proteins to modulate their functions.

*Hypophy-
siotropic
hormones*

Small proteins or peptides produced in cell bodies of hypothalamic neurons and transported along axons to the external layer of median eminence. The peptides are considered neurohormones because they are released into portal blood to modulate the functions of various types of pituitary cells. These neurohormones include corticotropin releasing hormone, growth hormone releasing hormone, luteinizing hormone releasing hormone, thyrotropin releasing hormone, and somatostatin.

*In situ
hybridization*

A method commonly used to identify sites of protein or peptide production at cellular level by demonstrating the

presence of corresponding mRNA signals. Specific mRNA signals on fixed tissue sections or cultured cells are detected by cross-binding with short sequence of radioactive or labeled nucleotides complementary to the equivalent part on the mRNA of interest.

Portal blood

The pool of blood that flows from the median eminence of the hypothalamus toward the pituitary gland. It carries hypophysiotropic neurohormones released from the median eminence to act on receptors in various cell types of the anterior pituitary.

*Radioimmuno-
assay*

A sensitive technique particularly useful for measuring small quantities of peptides in picogram concentrations. It is performed by first raising antibodies against peptides or antigens of interest. Displacement of the bound radioactive antigens or peptides from the antibodies by standard amounts of unlabeled antigen or synthetic peptides is established. By comparison to the standard displacement curve, the unknown quantity of the peptide is estimated by calculating the extent of its displacement of the labeled peptide from the antibodies.

The term corticotropin releasing hormone (CRH) was first coined about 40 years ago to describe a hypothalamic derivative that explicitly induces the release of pituitary adrenocorticotrophic hormone (ACTH). In 1981, the 41-amino-acid sequence of CRH was characterized by Vale and his colleagues at the Salk Institute. The neuropeptide has satisfied all the requirements as a hypophysiotropic hormone that specifically activates pituitary corticotrophs to augment the production of proopiomelanocortin (POMC), the precursor of ACTH, and the release of ACTH.

Structure and Gene

CRH(human, rat): Ser-Glu-Glu-Pro-Pro-Ile-Ser-Leu-Asp-Leu-Thr-Phe-His-Leu-Leu-Arg-Glu-Val-Leu-Glu-Met-Ala-Arg-Ala-Glu-Gln-Leu-Ala-Gln-Gln-Ala-His-Ser-Asn-Arg-Lys-Leu-Met-Glu-Ile-Ile-NH₂ (molecular weight, MW, 4758.1).

The sequence of CRH is highly preserved across species. In the case of humans and rats, they are identical and share greater than 90% homology when compared to that of monkeys and sheep.

The C-terminus of CRH-41 is critical for biological activity and the minimal sequence that retains full CRH activity in the rat species is 15–41. The shorter peptide has a much higher median effective dose (ED_{50}) than CRH-41 because of its lower affinity to the receptors. Whereas the C-terminal truncated peptide CRH(9–41) in α -helix conformation acts as an antagonist to type I CRH receptors, human CRH(6–33) and CRH(9–33) do not possess biological activities but bind strongly to CRH-binding proteins. The amino acid sequence of CRH shares a structure similarity of greater than 40% with urocortin, amphibian sauvagine, and fish urotensin I, all of which are biologically active in releasing ACTH from pituitary corticotrophs.

The human and rat CRH genes share great similarity; they both contain two exons separated by an intervening sequence. The first exon contains approximately 10 bp of the 5' untranslated region of the mRNA. The second exon contains 15 bp of the 5' untranslated region of the mRNA, the protein coding region, and the complete 3' untranslated region of the mRNA. The two genes are very homologous, suggesting that it has been highly conserved through evolution. The DNA sequence encoding the CRH protein shows a 94% homology, and the homology decreases to 80% for the rest of the CRH precursor. Similarly, a greater than 90% homology between humans and rats is also found in 5' flanking DNA sequences representing DNA regulatory elements for CRH gene regulation. In this region, there are consensus sequences putative for interactions with transcription proteins mediating the biochemical effects of glucocorticoids and derivatives from protein kinase A and protein kinase C activities.

Sites of Production and Pathways

In the brain, CRH is produced predominantly in the parvocellular subdivision of paraventricular nuclei (PVN) of the hypothalamus. These CRH-producing cells are hypophysiotropic and are involved in the regulation of pituitary corticotrophs because approximately 90% of the cells send their axons to the external layer of the median eminence (ME). From this pathway, CRH-41 is released into the portal blood to act on pituitary corticotrophs. Another CRH-41 pathway reaches the ME from magnocellular neurons of the PVN and the supraoptic and the accessory magnocellular nuclei of the hypothalamus. This pathway contains low levels of CRH-41, perhaps 10% of the amount found in the parvocellular pathway, and the neuropeptide comes predominantly from oxytocinergic magnocellular neurons and the neurites project mainly to the neural lobe of the pituitary.

A substantial proportion of CRH-producing neurons in the PVN are capable of expressing arginine vasopressin (AVP). Prominent expression of the AVP mRNA signal, as determined by *in situ* hybridization of AVP peptides by immunohistochemistry, has been demonstrated in CRH neurons after the removal of circulating glucocorticoids by adrenalectomy. In adrenalectomized rats, there is a concurrent increase in the production of the two neurohormones in CRH neurons and in their release into portal blood, as confirmed by radioimmunoassay. CRH and AVP act synergistically and induce a marked augmentation of pituitary POMC synthesis and ACTH release from pituitary corticotrophs. Although the physiological significance of this biological process remains unclear, it may act as a safety measure to enable pituitary ACTH to rapidly increase circulating glucocorticoids when the adrenal steroids reach a low level.

CRH-containing neurons are also present in the hippocampus, amygdala, and cortex. Although CRH axons from these sites do not project to the ME, they show synaptic formations in regions that are rich in CRH type I receptors and may be involved in modulating anxiety, cognition, and memory. This is in contrast to the distribution of CRH type II receptors, which correlates better with that of the axonal terminals of urocortin-producing neurons. Unlike CRH, hypothalamic or central urocortin neurons do not take part in the regulation of pituitary corticotrophs.

CRH-41 and its receptors are also found in cells intrinsic to spinal cord, peripheral lymphoid organs, gonads, placental endometrium, and adrenal glands. In these tissues, the peptides are likely to play a paracrine role by modulating the functions of adjacent cells or blood vessels in the same tissues. In some of these organs, CRH is known to induce vascular relaxation and involves in inflammatory responses. It is of interest to note that, whereas the CRH mRNA species of the rat hypothalamus is approximately 1400 bases, those found in the adrenal and gonads of the same species are larger by approximately 200 and 500 bases. These additional lengths of base pairs are found mainly at the nontranslated sites of the mRNA.

Biochemical Actions and Physiological Functions

Hypothalamic CRH is the principal hypophysiotropic activator of pituitary corticotrophs for the release of ACTH. Circulating ACTH, in turn, augments the synthesis and the release of corticosteroids from adrenal glands. In addition to exerting biochemical actions on the brain and peripheral tissues involved in stress response, one of the important roles of corticosteroids is to produce a negative feedback to dampen

the activity of hypothalamic CRH neurons and pituitary corticotrophs. These hierarchical levels of interactions and the negative feedback loop of circulating adrenoglucocorticoids form the basis of the stress circuit commonly known as the hypothalamic-pituitary-adrenal (HPA) axis.

The biochemical effect of CRH on corticotrophs is mediated through the activation of type I CRH receptors (CRH-RI) in corticotrophs of the anterior pituitary and intermediate lobe. The receptor is a 415-amino-acid protein comprising seven membrane-spanning domains. It belongs to the superfamily of G-protein-coupled receptors and is also activated by urocotin, urotensin, and sauvagine.

CRH binds to the receptors with high specificity and affinity with a K_d in the nanomolar range. Following its activation, the receptor couples GTP-binding proteins to effector enzymes, adenylyl cyclases that produce intracellular cAMP. The latter acts on protein kinase A to augment POMC mRNA expression and the release of ACTH. Consequently, derivatives of cAMP and agents that increase the intracellular levels of this second messenger (e.g., forskolin) mimic the effects of CRH-41 in all respects, whereas the inhibitor of cAMP-dependent kinase blocks the stimulating effects of CRH-41.

In addition to acting on the adenylyl cyclase-cAMP system, CRH-41 also increases the level of cytosolic free calcium through facilitating the influx of calcium from the extracellular medium. In addition, CRH-41 activates pathways related to arachidonic acid metabolism and produces metabolites that may stimulate or inhibit ACTH secretion.

Interaction with Arginine Vasopressin

Whereas AVP also stimulates the release of ACTH, the action of AVP is mediated through V3 receptors in the pituitary corticotrophs. The V3 receptors belong to the G-protein-coupling receptor superfamily, but the receptor acts by triggering membrane-bound phospholipase C to increase the turnover of inositol phosphate and diacylglycerol. This biological process is markedly different from that of CRH-41 because it results in enhancing protein kinase C activity. Hence, the stimulating effect of AVP synergizes with the action of CRH-41 in amplifying ACTH release by two- to threefold through the interaction of protein kinase A- and protein kinase C-dependent pathways.

In addition to AVP, there are other peptides and neurotransmitters that are also capable of stimulating the release of ACTH in a less efficient manner than CRH. The effects of these compounds on ACTH release in general are additive to CRH-41. These include oxytocin, angiotensin II and cholecystokinin,

noradrenaline, and adrenaline. On the other hand, a truncated form of atrial natriuretic peptide (ANP (5–28)) of hypothalamic origin, which is released into the portal blood, inhibits the baseline CRH-stimulated expression of POMC mRNA and the release of ACTH.

Effects of Corticotropin Releasing Hormone in the Brain

In addition to modulating pituitary corticotrophs, CRH-41 also acts on the central nervous system through two receptors: CRH-RI and CRH-RII. The CRH-RII has 431 amino acids and shares a 68% similarity with its counterpart. The CRH-RI is found in forebrain with a widespread signal throughout all areas of the neo-, olfactory, and hippocampal cortices. It is also present in subcortical limbic structures, in the septal region, and amygdala; low levels are seen in ventral thalamic and medial hypothalamic nuclei. In brain stem, high levels of receptors are found, including in the cerebellar and deep nuclei and precerebellar nuclei. The CRH-RI has been implicated in mediating behavioral changes associated with anxiety and the development of a functional HPA axis. CRH-41 has also been implicated as playing an important role in augmenting the central sympathetic system, involving in the modulation of the central cardiovascular system. On the other hand, CRH-41 suppresses growth and the reproductive system by inhibiting the release of growth hormone releasing hormone and luteinizing hormone releasing hormone at the hypothalamic level.

See Also the Following Articles

Adrenocorticotrophic Hormone (ACTH); Corticotropin Releasing Factor-Binding Protein; Corticotropin-Releasing Factor Receptors; Vasopressin.

Further Reading

- Antoni, F. A. (1986). Hypothalamic control of adrenocorticotropin secretion: advances since the discovery of 41-residue corticotropin-releasing factor. *Endocrine Review* 7, 351–378.
- Chen, R., Lewis, K. A., Perrin, M. H., et al. (1993). Expression cloning of a human corticotropin-releasing-factor receptor. *Proceedings of the National Academy of Sciences USA* 90(19), 8967–8971.
- Potter, E., Sutton, S., Donaldson, C., et al. (1994). Distribution of corticotropin-releasing factor receptor mRNA expression in the rat brain and pituitary. *Proceedings of the National Academy of Sciences USA* 91(19), 8777–8781.
- Smith, G. W., Aubry, J. M., Dellu, F., et al. (1998). Corticotropin releasing factor receptor 1-deficient mice display

decreased anxiety, impaired stress response, and aberrant neuroendocrine development. *Neuron* 20(6), 1093–1102.

Swanson, L., Sawchenko, P. E., Lind, R. W., et al. (1987). The CRH motoneuron: differential peptide regulation in neurons with possible synaptic, paracrine and endocrine outputs. In: Ganong, W. F., Dallman, M. F. & Roberts, J. L. (eds.) *Special issue on the hypothalamic-pituitary-adrenal axis revisited. Annals of the New York Academy of Sciences* 512, 12–23.

Tan, T., Yang, Z., Huang, W., et al. (1994). ANF(1–28) is a potent suppressor of Pro-opiomelanocortin (POMC)

mRNA but a weak inhibitor of EP-LI release from AtT-20 cells. *Journal of Endocrinology* 143, R1–R4.

Turnbull, A. V. and Rivier, C. (1997). Corticotropin-releasing factor (CRF) and endocrine responses to stress: CRF receptors, binding protein, and related peptides. *Proceedings of the Society for Experimental Biology and Medicine* 215(1), 1–10.

Vale, W., Spiess, J., Rivier, C. and Rivier, J. (1981). Characterization of a 41-residue ovine hypothalamic peptide that stimulates secretion of corticotropin and beta-endorphin. *Science* 213, 1394–1397.

Corticotropin Releasing Factor Receptor Deficiency in Mice

N J Justice and K-F Lee

The Salk Institute, La Jolla, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Corticotropin Releasing Factor Receptors and Ligands
 CRFR1-Deficient Mice Have a Defective Hypothalamic-Pituitary-Adrenal Axis Response to Stress
 CRFR2-Deficient Mice Are Hypersensitive to Stress
 UCN1-Deficient Mice Have Increased Anxiety and Hearing Impairment
 CRFR1/CRFR2-Double-Deficient Animals Have a Blunted Hypothalamic-Pituitary-Adrenal Axis Stress Response
 Conclusion

Glossary

<i>Corticotropes</i>	Endocrine cells in the anterior pituitary that release adrenocorticotrophic hormone into the bloodstream.
<i>Corticotropin releasing factor (CRF) family</i>	A family of peptide ligands that bind CRF receptors; they were identified based on homology to CRF.
<i>CRF receptor-deficient mice</i>	The mice resulting from genetic manipulation of DNA to remove the genes encoding CRF receptors.
<i>CRF receptors</i>	Two seven-transmembrane G-protein-coupled receptors that are bound and activated by CRF family peptide ligands.
<i>Hypothalamic-pituitary-adrenal (HPA) axis</i>	The neuroendocrine circuit that coordinates the release of corticosteroids into the bloodstream in response to stress.

Paraventricular nucleus of the hypothalamus (PVN)

A nucleus in the hypothalamus containing neurons that release factors into portal circulation, which then stimulate the release of hormones from the anterior pituitary into the bloodstream.

Stress response

The cumulative physiological and behavioral response to real or perceived threats to homeostasis.

Corticotropin releasing factor (CRF) is released from the paraventricular nucleus of the hypothalamus (PVN) into the portal circulation in response to stress. In the anterior pituitary, CRF receptors on pituitary corticotropes respond to CRF by releasing adrenocorticotrophic hormone (ACTH), which circulates in the bloodstream and, in turn, causes the release of glucocorticoids from the adrenal cortex. These corticosteroids are the primary effectors of the hypothalamic-pituitary-adrenal (HPA) axis stress response. CRF and its family of peptide ligands are also released at many other sites in the brain and in the periphery, where they effect cellular and physiological changes in addition to the activation of the HPA axis. The nature of these changes is determined by where CRF and related ligands bind and activate CRF receptors. Loss-of-function studies in mice deficient for CRF receptors have confirmed the importance of these receptors in bringing about corticosteroid release and have provided evidence for additional roles of CRF receptors in modulating the amplitude, duration, and behavioral outcome of the stress response. Here, we review what these receptor-deficient mice have taught us about how CRF receptors function to mediate and modulate stress.

Corticotropin Releasing Factor Receptors and Ligands

Two receptors for CRF have been identified. CRF type 1 receptor (CRFR1) is a high-affinity receptor for CRF that mediates the response of anterior pituitary corticotropes to CRF. CRF type 2 receptor (CRFR2), which shares 70% amino acid identity with CRFR1, has a lower affinity for CRF than the other CRF family peptide ligands. Urocortin (UCN1), a peptide ligand with homology to CRF, binds both CRFR1 and CRFR2 with high affinity, whereas urocortin 2 (UCN2) and urocortin 3 (UCN3) bind only CRFR2 with high affinity. Each peptide ligand has a unique distribution in the brain and peripheral tissues and overlaps in distinct areas with CRFR1 and CRFR2. For example, CRFR1 is expressed in the anterior pituitary, where it responds to CRF release from the PVN. In contrast, CRFR2 is expressed in the lateral septum, where it most likely responds to UCN1 from the Edinger–Westphal nucleus and UCN3 from the paraventricular region of the hypothalamus. The discrete localization of each unique receptor–ligand pair suggests a distinct role for each receptor in the stress response. By selectively removing these genes from mice, we have begun to understand these roles.

CRFR1-Deficient Mice Have a Defective Hypothalamic–Pituitary–Adrenal Axis Response to Stress

CRFR1-deficient mice display abnormally low levels of circulating corticosteroids. This can be explained by a lack of responsiveness of anterior pituitary cells to CRF. The inability to stimulate ACTH release, which then activates corticosteroid release, not only blunts the stress response but also prevents the proper development of the adrenal cortex. The ACTH-responsive zona fasciculata of the adrenal gland fails to develop appropriately and undergoes atrophy in CRFR1-deficient animals. The supplementation of ACTH in postnatal animals (postnatal day 10 to postnatal day 21) can rescue adrenal gland development, indicating that an early postnatal CRFR1-dependent release of ACTH is necessary for proper adrenal gland development. The combination of adrenal gland atrophy and reduced ACTH release from the pituitary results in a profoundly reduced corticosteroid release in response to stress. Consistent with the notion that CRF and CRFR1 function as a ligand–receptor pair, this CRFR1-deficient phenotype mirrors the phenotype observed in CRF-deficient mice.

In addition to having a defective HPA axis, CRFR1-deficient animals display reduced anxiety-like behavior. In the elevated plus-maze and in the

dark–light emergence task, two well-defined behavioral paradigms that measure anxiety-related behavior, CRFR1-deficient animals display decreased anxiety. This could be due to a blunted HPA response in mutant animals or a function of CRFR1 at other sites in the nervous system. To distinguish between these possibilities, investigators provided exogenous corticosteroids, which failed to rescue the anxiety-like behavioral phenotype. These data suggest that CRFR1 might modulate stress-induced behavior through actions in the brain, independent of HPA axis stimulation. Indeed, data from a forebrain-specific CRFR1-deficient animal support this interpretation. In the forebrain-specific loss of function, the pituitary retains expression of CRFR1. These animals have normal resting ACTH and corticosteroid levels as well as an intact HPA response. However, they display a reduction in anxiety-related behavior similar to full animal CRFR1-deficiency. Together, these findings strongly suggest that CRFR1 is involved in generating anxiety-like behavior through action in the forebrain, independent of its role in activating the HPA axis in the pituitary.

CRFR2-Deficient Mice Are Hypersensitive to Stress

In contrast to CRFR1-deficient mice, CRFR2-deficient mice display a hyperactive HPA axis and increased anxiety-like behavior. ACTH levels in CRFR2-deficient mice rise to approximately twice that of wild-type littermates following 2 min of restraint stress. Corticosteroid levels also remain elevated for a longer period of time following stress in the CRFR2 mutants. Consistent with overactive HPA axis dynamics, CRFR2-deficient animals appear more anxious than their wild-type littermates. In the elevated plus-maze and open field test, CRFR2-deficient mice display increased anxiety-like behavior. This might be due to a lack of modulation of the HPA axis, which is normally provided by CRFR2. Alternatively, it could be due to independent functions of CRFR2 apart from the regulation of the HPA axis.

CRFR2 is the most widely expressed CRF receptor in peripheral tissues. It was discovered based on its high levels of expression in the heart and skeletal muscle. Recently, additional functions for CRFR2 have been suggested by peripheral phenotypes observed in the CRFR2-deficient animal. For example, CRFR2-deficient animals have elevated blood pressure and fail to respond to UCN1 with an increase in heart contractility. In addition, an increase in temperature of the interscapular brown adipose tissue of CRFR2-deficient animals has been

reported, suggesting metabolic abnormalities. Consistently, these animals consume more food and gain less body fat when fed a high fat diet. These findings highlight the importance of CRF receptors in mediating peripheral changes in response to stress as well as central nervous system and neuroendocrine stress responses.

UCN1-Deficient Mice Have Increased Anxiety and Hearing Impairment

In order to gain more insight into the CRFR2-deficient mouse phenotype, a UCN1-deficient mouse has been generated. Interestingly, UCN1-deficient animals display increased anxiety-like behavior, similar to the CRFR2-deficient animal. Indeed, when the investigators examined CRFR2 expression in the lateral septum, where UCN1 projections meet CRFR2-expressing neurons, they saw greatly reduced levels of CRFR2 expression. This finding is consistent with a role for UCN1/CRFR2 in the lateral septum to control anxiety-related behavior. This control, however, must be independent of the modulation of the HPA by CRFR2 because the UCN1-deficient animal displays no changes in the dynamics of HPA function. This suggests that the role played by CRFR2 in regulating the HPA is either independent of UCN1 or that other ligands for CRFR2 function redundantly in the absence of UCN1.

An additional phenotype observed in the UCN1 knockout mouse is hearing impairment. UCN1-deficient mice showed an approximately twofold increase

in the threshold for reaction to an auditory stimulus compared with wild-type mice. Although a role for UCN1 in development of the auditory system was not expected, it is consistent with CRFR1 and CRFR2 expression in the organ of Corti. Moreover, UCN1 is expressed on fibers of the lateral superior olive that project to the inner hair cells of the organ of Corti. Curiously, the change in auditory threshold indicates a problem with the outer hair cells. The investigators hypothesize that UCN1 from the inner hair cells is involved in a paracrine signal to the outer hair cells, which is required during development of the organ.

CRFR1/CRFR2-Double-Deficient Animals Have a Blunted Hypothalamic-Pituitary-Adrenal Axis Stress Response

The CRFR1-deficient mouse has an impaired HPA axis and displays decreased anxiety-like behavior, whereas the CRFR2-deficient mouse has a hypersensitive HPA axis and increased anxiety-like behavior. How do mice deficient for both CRFR1 and CRFR2 respond to stress? CRFR1-CRFR2-deficient mice have a decreased HPA response and decreased anxiety-related behavior, displaying a very similar phenotype to CRFR1-deficient animals. This suggests that the deletion of CRFR1 has a dominant impact on behavioral phenotypes associated with loss of CRF receptors, whereas CRFR2 most likely modulates behavior through central nervous system actions. The central functions of CRFR1 and CRFR2 can affect the HPA axis by changing neural activity in the PVN.

Table 1 CRF receptor- and ligand-deficient mouse phenotypes

Genotype	HPA function	Anxiety-related behavior	Peripheral tissue development/function	Source ^a
CRFR1-deficient	Inactive	Decreased anxiety	Abnormal adrenal cortex development	6, 7
CRFR2-deficient	Hyperactive	Increased anxiety	Altered cardiac function, altered metabolism	1, 2, 3, 4
CRF-deficient	Inactive	Not reported	Abnormal adrenal cortex development	5
UCN1-deficient	Normal	Increased anxiety	Hearing impairment	8

^aData from:

¹Bale, T. L., Contarino, A., Smith, G. W., et al. (2000). Mice deficient for corticotropin-releasing hormone receptor-2 display anxiety-like behaviour and are hypersensitive to stress. *Nature Genetics* **24**, 410–414.

²Carlin, K. M., Vale, W. W. and Bale, T. L. (2006). Vital functions of corticotropin-releasing factor (CRF) pathways in maintenance and regulation of energy homeostasis. *Proceedings of the National Academy of Sciences USA*, **103**, 3462–3467.

³Coste, S. C., Kesterson, R. A., Heldwein, K. A., et al. (2000). Abnormal adaptations to stress and impaired cardiovascular function in mice lacking corticotropin-releasing hormone receptor-2. *Nature Genetics* **24**, 403–409.

⁴Kishimoto, T., Radulovic, J., Radulovic, M., et al. (2000). Deletion of *chrh2* reveals an anxiolytic role for corticotropin-releasing hormone receptor-2. *Nature Genetics* **24**, 415–419.

⁵Muglia, L., Jacobson, L., Dikkes, P., et al. (1995). Corticotropin-releasing hormone deficiency reveals major fetal but not adult glucocorticoid need. *Nature* **373**, 427–432.

⁶Smith, G. W., Aubry, J. M., Dellu, F., et al. (1998). Corticotropin releasing factor receptor 1-deficient mice display decreased anxiety, impaired stress response, and aberrant neuroendocrine development. *Neuron* **20**, 1093–1102.

⁷Timpl, P., Spanagel, R., Sillaber, I., et al. (1998). Impaired stress response and reduced anxiety in mice lacking a functional corticotropin-releasing hormone receptor 1. *Nature Genetics* **19**, 162–166.

⁸Vetter, D. E., Li, C., Zhao, L., et al. (2002). Urocortin-deficient mice show hearing impairment and increased anxiety-like behavior. *Nature Genetics* **31**, 363–369.

CRFR1 activates the PVN, whereas CRFR2 might decrease activity in the PVN to modulate HPA dynamics. The modulations made by central actions of CRFR1 and CRFR2, however, cannot override the changes in behavior that result from a defective HPA due to CRFR1 deficiency. Research is in progress to further understand how the complex actions of these receptor systems in the brain coordinate with the HPA axis to influence behavior.

Conclusion

CRFR1- and CRFR2-deficient mice have improved our understanding of the prominent roles these receptors play in the stress response. CRFR1 functions to activate the HPA axis and generate anxious behavior. Within the HPA axis, CRFR2 functions to control overactivity. In addition, CRFR1 and CRFR2 act within the central nervous system to influence behavior. CRFR2 also acts in the periphery to alter metabolism and cardiac function. UCN1 has an unexpected role in the development of the auditory system (Table 1). The diverse effects of CRF receptor deficiency on HPA axis, central nervous system, and peripheral tissue functions demonstrate the important contributions of CRF and its family of peptide ligands to adaptive responses to stress.

See Also the Following Articles

Corticotropin Releasing Factor (CRF); Corticotropin-Releasing Factor Receptors; Urocortins; CRH Anxiety Circuitry in Brain; Genetic Polymorphisms in Stress Response; Corticotropin-releasing factor (CRF) Family of Neuropeptides – Role in Inflammation.

Further Reading

- Bale, T. L., Contarino, A., Smith, G. W., et al. (2000). Mice deficient for corticotropin-releasing hormone receptor-2 display anxiety-like behaviour and are hypersensitive to stress. *Nature Genetics* **24**, 410–414.
- Bale, T. L., Picetti, R., Contarino, A., et al. (2002). Mice deficient for both corticotropin-releasing factor receptor 1 (CRFR1) and CRFR2 have an impaired stress response

- and display sexually dichotomous anxiety-like behavior. *Journal of Neuroscience* **22**, 193–199.
- Bale, T. L. and Vale, W. W. (2004). CRF and CRF receptors: role in stress responsivity and other behaviors. *Annual Review of Pharmacology and Toxicology* **44**, 525–557.
- Carlin, K. M., Vale, W. W. and Bale, T. L. (2006). Vital functions of corticotropin-releasing factor (CRF) pathways in maintenance and regulation of energy homeostasis. *Proceedings of the National Academy of Sciences USA* **103**, 3462–3467.
- Coste, S. C., Kesterson, R. A., Heldwein, K. A., et al. (2000). Abnormal adaptations to stress and impaired cardiovascular function in mice lacking corticotropin-releasing hormone receptor-2. *Nature Genetics* **24**, 403–409.
- Kishimoto, T., Radulovic, J., Radulovic, M., et al. (2000). Deletion of *Crhr2* reveals an anxiolytic role for corticotropin-releasing hormone receptor-2. *Nature Genetics* **24**, 415–419.
- Muglia, L., Jacobson, L., Dikkes, P., et al. (1995). Corticotropin-releasing hormone deficiency reveals major fetal but not adult glucocorticoid need. *Nature* **373**, 427–432.
- Muller, M. B., Zimmermann, S., Sillaber, I., et al. (2003). Limbic corticotropin-releasing hormone receptor 1 mediates anxiety-related behavior and hormonal adaptation to stress. *Nature Neuroscience* **6**, 1100–1107.
- Owens, M. J. and Nemeroff, C. B. (1991). Physiology and pharmacology of corticotropin-releasing factor. *Pharmacological Reviews* **43**, 425–473.
- Rivier, C., Brownstein, M., Spiess, J., et al. (1982). In vivo corticotropin-releasing factor-induced secretion of adrenocorticotropin, beta-endorphin, and corticosterone. *Endocrinology* **110**, 272–278.
- Smith, G. W., Aubry, J. M., Dellu, F., et al. (1998). Corticotropin releasing factor receptor 1-deficient mice display decreased anxiety, impaired stress response, and aberrant neuroendocrine development. *Neuron* **20**, 1093–1102.
- Timpl, P., Spanagel, R., Sillaber, I., et al. (1998). Impaired stress response and reduced anxiety in mice lacking a functional corticotropin-releasing hormone receptor 1. *Nature Genetics* **19**, 162–166.
- Vale, W., Spiess, J., Rivier, C., et al. (1981). Characterization of a 41-residue ovine hypothalamic peptide that stimulates secretion of corticotropin and beta-endorphin. *Science* **213**, 1394–1397.
- Vetter, D. E., Li, C., Zhao, L., et al. (2002). Urocortin-deficient mice show hearing impairment and increased anxiety-like behavior. *Nature Genetics* **31**, 363–369.

Corticotropin Releasing Factor-Binding Protein

P J Lowry, C F Kemp and R J Woods

University of Reading, Reading, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 582–585, © 2000, Elsevier Inc.

Introduction

Distribution

Properties

Regulation of Gene Expression

Role of Corticotropin Releasing Hormone-Binding Protein in the Stress Response

Glossary

Corticotropin releasing factor (CRF)

A hypothalamic peptide comprising 41 amino acid residues that is secreted in response to stressful stimuli into the hypothalamic-pituitary portal system causing release of adrenocorticotrophic hormone (ACTH) from the anterior pituitary gland.

Corticotropin releasing factor-binding protein (CRF-BP)

A 37-kDa protein that binds CRF and is found in the circulation of higher primates as well as, in both membrane-bound and free forms, in the brain of all vertebrates hitherto examined.

Introduction

Under experimental conditions, the corticotropin releasing hormone-binding protein (CRF-BP) is able to bind CRF and render it biologically inactive. Because CRF-BP expression overlaps with CRF receptor expression in several areas of the brain, including the pituitary gland, it has been postulated that its function is to modulate the neuroendocrine and transmitter activities of CRF. The presence of CRF-BP in the periphery of higher primates as well indicates that this protein is involved in modulating the action of CRF in the periphery, where CRF has known proinflammatory effects (this is in contrast to the centrally derived anti-inflammatory action of CRF via the release of cortisol). Thus, the CRF-BP may play a pivotal role in the crossover of endocrine, immune, and higher functions in both acute and chronic stress responses.

Distribution

Central Corticotropin Releasing Factor-Binding Protein

From histochemistry and *in situ* hybridization, membrane-bound CRF-BP is observed to be expressed in the rat brain predominantly in the cerebral cortex, including archipaleo- and neocortical fields (see [Figure 1](#)). It is present also in subcortical limbic structures (amygdala and bed of the stria terminalis) and sensory relays of the auditory, olfactory, vestibular, and trigeminal systems; several raphe nuclei; and some cell groups of the brain stem reticular core. Hypothalamic expression occurs only in the ventral premammillary and dorsomedial nuclei. It is expressed in both neurons and glial astrocytes as well as in pituitary corticotrophs. Its wide distribution does not consistently match the more limited distribution of CRF, and the expression of the CRF 1 receptor (CRF-R1) and CRF 2a receptor (CRF-R2a) in the brain does not coincide completely with either CRF or CRF-BP mRNA.

Peripheral Corticotropin Releasing Factor-Binding Protein

It is suspected that, in addition to the brain form, only higher primates have circulating CRF-BP. Humans express CRF-BP in the liver, and it circulates in the plasma of both sexes at concentrations, measured by immunoradiometric assay, of between 50 and 100 $\mu\text{g liter}^{-1}$. The liver is the probable source of CRF-BP in plasma because concentrations are much reduced in patients with chronic liver disease. By contrast, kidney failure is associated with elevated plasma levels, indicating that CRF-BP (not bound to ligand) is normally cleared by the renal route. Plasma levels are remarkably constant from day to day. In molar terms, plasma CRF-BP vastly exceeds CRF and therefore never approaches saturation by ligand except in the last few weeks of pregnancy. The placenta of primates also expresses CRF, and it has been suggested that the binding protein serves in late pregnancy to help prevent the stimulation of the maternal pituitary by releasing large amounts of placental CRF into the circulation. Not only is hormonal activity abolished after binding to CRF-BP, but it is in the bound form that most CRF may be removed from the maternal circulation. The CRF-BP is also measurable in fetal blood and amniotic fluid and may have a further role in the modulation of CRF activity associated with the onset of labor. CRF stimulates the release of prostaglandins from

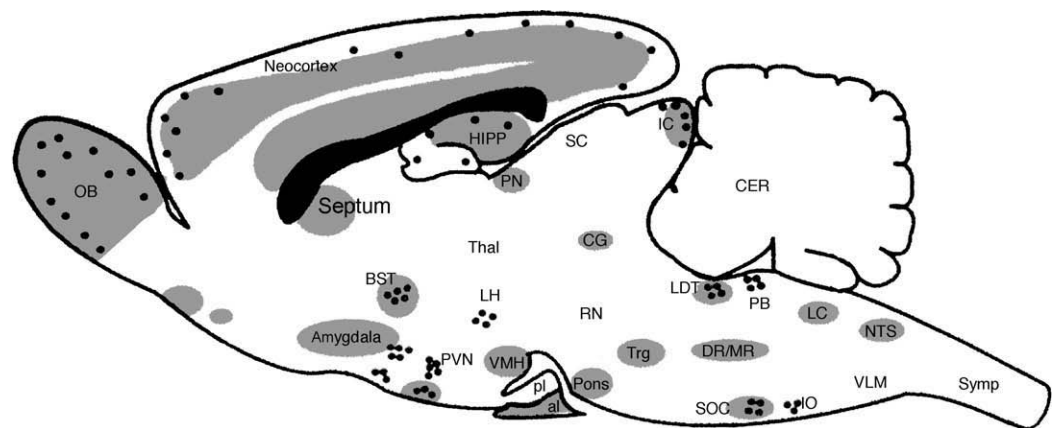


Figure 1 The distribution of the corticotropin releasing factor-binding protein in the rat brain. al, ; BST, bed nucleus of the stria terminalis; CER, cerebellum; CG, central grey; DR/MR, dorsal/medial raphe nucleus; HIPP, hippocampus; IC, inferior colliculus; IO, inferior olive; LC, locus cereolus; LH, lateral hypothalamus; LTD, lateral dorsal tegmental nucleus; NTS, nucleus of the solitary tract; OB, olfactory bulb; PB, parabrachial nucleus; pl, ; PN, perfect nucleus; Pons, pontine nuclei; PVN, paraventricular nucleus; RN, red nucleus; SC, superior colliculus; SOG, superior olivary complex; Symp, sympathetic related areas; Thal, thalamus; Trg, trigeminal; VLM, ventrolateral medulla; VMH, ventromedial hypothalamic nucleus.

fetal membranes and potentiates their induction of myometrial contractility. These activities are presumed to be modulated by CRF-BP in the fetal and maternal circulation.

In contrast to late pregnancy, when the plasma levels of CRF-BP are lower than normal, in both rheumatoid arthritis and septicemia the concentrations are significantly elevated (by up to 50%). Whereas the lowering of CRF-BP levels in pregnancy seems to be triggered by the secretion of placental CRF, the raised levels in these active immune states appear to be caused by increased secretion by the liver. Presumably this is due to the activation of the acute-phase response elements found in the 5' flanking region of the CRF-BP gene. It is likely that the bioactivity of circulating CRF is always attended by some degree of modification imposed by the presence of its binding protein.

Properties

Ligand Binding

The affinity of human CRF-BP for human CRF is high ($K_d = 0.1$ nM). That for ovine CRF, which differs at 7 of the 41 amino acid residues, is 200-fold lower; however, CRFs from both species have equal affinities for the pituitary CRF-R1. Sheep CRF-BP has an 86% homology in its amino acid sequence with human and rat CRF-BP, yet it binds human CRF with a 10-fold higher affinity than ovine (CRF $K_d = 0.15$ nM, as opposed to 1.75 nM). Furthermore, a truncated version of human CRF, CRF (6–33), binds CRF-BP with reduced affinity but does not bind to CRF-R1, CRF-R2a, or CRF-R2b. These considerations indicate that the structural determinants required for ligand binding by the

Table 1 Order of ligand affinities for CRF-BP and the CRF receptors

CRF-BP	Urotensin > urocortin ≥ hCRF > sauvagine > oCRF
CRF-R1	Urocortin > hCRF = oCRF = urotensin = sauvagine
CRF-R2a	Urocortin > sauvagine > urotensin > hCRF > oCRF
CRF-R2b	Urocortin > sauvagine = urotensin > hCRF = oCRF

binding proteins and CRF-R1 are very different. The recent discovery of urocortin, a new member of the CRF family of peptides, has shown that CRF receptor-binding protein interactions may be further complicated by the participations of additional ligands. Discrepancies in the central distribution of the CRF receptors, CRF-BP, and CRF terminals is evidence that novel ligands for CRF-BP remain to be discovered. The relative affinities for CRF and CRF-like ligands are shown in Table 1.

With the exception of the synthetic peptide antagonist α -helical CRF(9–41), CRF-like ligands interacting with CRF-BP *in vitro* induce dimerization. The formation of dimers has not been demonstrated to occur *in vivo*, but CRF administered to humans intravenously is cleared rapidly from the circulation. The possibility therefore exists that CRF-BP not only inactivates CRF but, by dimerizing, targets the ligand for elimination.

Physicochemical Properties

CRF-BP has a molecular mass of 37 kDa, and the mature polypeptide consists of 298 amino acid residues. Its structure is composed predominantly of a β -sheet with little evidence of α -helical folding.

CRF-BP is subject to two posttranslational modifications. These have been investigated using human

recombinant CRF-BP. It has a consensus sequence for N-glycosylation of asparagine residue 204. Enzymic removal of the glycosyl moiety does not affect the binding of CRF. A potential phosphorylation site exists at serine residue 233, although phosphorylation has been found by experiment to occur instead at serine 206. On storage, affinity-purified recombinant CRF-BP also undergoes cleavage between amino acid residues 234 and 235, even in the presence of protease inhibitors, yielding fragments of 28 and 9.6 kDa molecular mass. Some degree of protein unfolding has been observed, by tryptophan excitation fluorescence, to occur prior to this *in vitro* cleavage. It is not known whether it is accompanied by dephosphorylation. Protein fragments of similar size have been observed in the synovial fluid of patients suffering from rheumatoid arthritis, although it is not clear whether this is a consequence of the inflammation or if it is involved in the pathogenesis of this condition. Under nonpathological conditions, insulinlike growth factor (IGF)-binding proteins also undergo dephosphorylation and cleavage, both of which are regulated by and affect ligand binding. In the case of IGF-binding protein-3, cleavage is regulated by IGF-1. The cleavage of IGF-binding proteins reduces their affinity for ligand, whereas, after cleavage of CRF-BP, the binding of human CRF to the 28-kDa N-terminal fragment is unimpaired.

Regulation of Gene Expression

The characterization of the upstream region of the CRF-BP gene has revealed acute-phase response elements, one of which (located at -305) is known to bind the transcription factor nuclear factor (NF)- κ B. This factor, which regulates immunoglobulin and interleukin expression, is thought to be important in activating liver-specific genes, such as angiotensin, in response to acute stress, inflammation, and injury. Another element (located at -676) binds interferon (INF)-1, which is known to regulate the interferon gene.

More recent evidence has revealed multiple potential transcription-factor binding sites in the 5' flanking region of both the rat and human CRF-BP gene. These include a cAMP response element binding protein (CREB)/activating transcription factor (ATF) site, a liver factor (LF)-A1 site, an oxytocin (OTX) site, an activating enhancer binding protein (AP)-2 site, two stimulating protein (Sp)1 sites, and two AP-1 sites, as well as those mentioned previously. The CREB/ATF and AP-2 sites suggest that cAMP may be important for the regulation of the CRF-BP gene, and this is corroborated by the fact that primary rat neuronal or astrocyte cultures, stimulated with forskolin

or isobutylmethylxanthine, have been shown to increase levels of CRF-BP mRNA. It has also been demonstrated using the G_s-protein-coupled CRF receptor in the AtT-20 pituitary cell line that CRF-BP gene expression is positively regulated by CRF, suggesting a mechanism for refining the action of CRF on the stress axis at the level of gene expression.

Role of Corticotropin Releasing Hormone-Binding Protein in the Stress Response

The stress axis and the immune system are interrelated, and stress influences the course and pathology of the immune response to infectious disease and to autoimmune stimuli. Intravenous administration of CRF, a key component of the stress axis, elicits visceral and neuroendocrine changes that mimic the body's reaction to stress. Peripheral CRF receptors have been identified on cells of the testes, ovaries, adrenal medulla, heart, uterus, skeletal muscle, lung, kidney, and vasculature. Furthermore, CRF receptors have been reported to exist on circulating lymphocytes and macrophages as well as on mast cells. There is evidence for the hypothesis that peripheral proinflammatory effects of CRF in humans may be modulated by circulating CRF-BP, whereas the central anti-inflammatory response to stress may be modulated by membrane-bound CRF-BP in the brain. It has been shown that CRF is displaced from CRF-BP when human brain tissue is treated with an analog of CRF that displaces bound CRF. *In vivo*, the displacement of CRF from CRF-BP by the central administration of a truncated form of CRF (CRF(6-33)) to rats had an anxiolytic effect. It improved learning and memory presumably by making more free CRF available to interact with CRF receptors. These observations demonstrate that CRF-BP is certainly capable of modulating the actions of CRF.

See Also the Following Articles

Corticotropin Releasing Factor (CRF); Corticotropin-Releasing Factor Receptors.

Further Reading

- Behan, D. P., De Souza, E. B., Lowry, P. J., et al. (1995). corticotropin releasing factor (CRF) binding protein: a novel regulator of CRF and related peptides. *Frontiers in Neuroendocrinology* 16, 362-382.
- Behan, D. P., Maciejewski, D., Chalmers, D., et al. (1995). Corticotropin-releasing factor binding protein (CRF-BP) is expressed in neuronal and astrocytic cells. *Brain Research* 689, 259-264.

- Chalmers, D. T., Lovenberg, T. W., Grigoriadis, D. E., et al. (1996). Corticotropin-releasing factor receptors: from molecular biology to drug design. *TIPS* 17, 166–172.
- Kemp, C. F., Woods, R. J. and Lowry, P. J. (1998). The corticotropin releasing factor-binding protein: an act of several parts. *Peptides* 19(6), 1119–1128.
- Linton, E. A., Wolfe, C. D. A., Behan, D. P., et al. (1988). A specific carrier substance for human corticotropin releasing factor in late gestational maternal plasma that could mask its ACTH releasing activity. *Clinical Endocrinology* 28, 315–324.
- Turnbull, A. V. and Rivier, C. (1997). Corticotropin-releasing factor (CRF) and endocrine responses to stress: CRF receptors, binding protein and related peptides. *Proceedings of the Society for Experimental Biology and Medicine* 215, 1–10.

Corticotropin See: Adrenocorticotrophic Hormone (ACTH).

Corticotropin-Releasing Factor Circuitry in the Brain – Relevance for Affective Disorders and Anxiety

D A Gutman and C B Nemeroff

Emory University School of Medicine, Atlanta, GA, USA

© 2007 Elsevier Inc. All rights reserved.

Glucocorticoids in Affective Disorders: Early Evidence
Corticotropin-Releasing Factor
Corticotropin-Releasing Factor Receptors
Extrahypothalamic Corticotropin-Releasing Factor
Circuits and Depression
Corticotropin-Releasing Factor in Depression
Corticotropin-Releasing Factor and Anxiety Disorders
Small-Molecule Corticotropin-Releasing Factor
Antagonists
Conclusions and Future Directions

Glucocorticoids in Affective Disorders: Early Evidence

Evidence linking the hypothalamic-pituitary-adrenal (HPA) axis abnormalities with psychiatric symptoms dates back over 100 years. The occurrence of depression, anxiety, and in extreme cases psychosis, in both Cushing and Addison diseases, which are associated with excessive or markedly reduced levels of circulating glucocorticoids, respectively, served as an initial impetus for researchers to scrutinize HPA-axis function in psychiatric disorders. Some of the first direct evidence that posited a role between glucocorticoids

and mood stemmed from studies that were carried out in the 1950s.

The concatenation of several lines of evidence revealed abnormalities in glucocorticoid (i.e., cortisol) function in depressed patients. These include elevated plasma cortisol concentrations, increased 24-h urinary free cortisol concentrations, and increased levels of cortisol metabolites in urine. Although not observed in every patient with major depression, elevated cortisol secretion in depression is among the most reproducible findings in all of biologic psychiatry.

Structural changes in the components of the HPA axis have also been documented in depressed patients. These include pituitary gland enlargement as demonstrated by magnetic resonance imaging (MRI) and moreover, the adrenal glands have also shown to be enlarged in depression as assessed by both imaging studies, as well as postmortem analysis of the adrenals, presumably due to the effects of adrenocorticotrophic hormone (ACTH) hypersecretion.

Corticotropin-Releasing Factor

Although Saffran and Schally identified a crude extract that promoted the release of ACTH from the pituitary in 1955, the ultimate regulator of ACTH and cortisol release, corticotropin-releasing factor (CRF), was not isolated and chemically characterized until 1981. Working with extracts derived from 500 000 sheep hypothalami, Vale and colleagues at the Salk Institute elucidated the structure of CRF. This

discovery led to the availability of synthetic CRF, which allowed for a comprehensive assessment of the HPA axis. Based on findings from numerous studies it is clear that CRF coordinates the endocrine, immune, autonomic, and behavioral responses of mammals to stress.

The hypothalamus serves as a regulatory point for a number of key functions including body temperature, hunger, thirst, and circadian cycles. Within the hypothalamus, CRF is synthesized primarily in the parvocellular neurons of the paraventricular nucleus (PVN). These PVN CRF neurons receive input from a variety of brain regions, including the amygdala, the bed nucleus of the stria terminalis (BNST), and the brain stem. CRF-containing neurons within the hypothalamus then project to the median eminence. The median eminence serves as a bridge between the hypothalamus and the anterior lobe of the pituitary gland. Within the median eminence, the release and release-inhibiting hypothalamic hypophysiotropic hormones are released from nerve terminals and enter the portal blood system which supplies the anterior pituitary gland.

In response to a variety of stressors, this neural circuit becomes activated, thereby releasing CRF into the hypothalamo-hypophysial portal system, where it activates CRF receptors on corticotrophs in the adenohypophysis to promote the synthesis of pro-opiomelanocortin and the release of its major posttranslation products, ACTH and β -endorphin. ACTH, released from the anterior pituitary, stimulates the production and release of cortisol from the adrenal cortex. These same hypothalamic CRF neurons also project to the spinal cord and brainstem nuclei, including the locus coeruleus (LC), the site of the noradrenergic neurons that project to the forebrain. Magnocellular neurons (literally large cells) in the PVN also produce vasopressin and oxytocin. Vasopressin has been regarded to potentiate the action of CRF on ACTH secretion.

Corticotropin-Releasing Factor Receptors

Two CRF receptor subtypes, CRF₁ and CRF₂, with distinct anatomic localization and receptor pharmacology, have been identified in rats and humans. Both receptors are members of the large G-protein receptor superfamily. The CRF₁ receptor is predominantly expressed in the pituitary, cerebellum, and neocortex in rats. Considerable evidence from laboratory animal studies has shown that CRF₁ receptors may specifically mediate some of the anxiogenic-like behaviors observed after administration of CRF.

In agreement with these findings, mice that have had their CRF₁ receptor knocked out were found to exhibit an impaired stress response. The CRF₁ receptor knockout mice were less anxious than their wild-type litter mates when tested in the elevated plus maze, a commonly used animal model of anxiety. This paradigm consists of an elevated maze in which two of the arms contain walls, whereas the other two arms are open. Rodents that spend more time on the open arms, where they are potentially exposed to predators or more likely could fall, are generally interpreted to be less anxious than animals that spend the preponderance of time in the closed arms. In addition, data in these transgenic mice showed a significant reduction in stress-induced release of ACTH and corticosterone (the primary glucocorticoid in rodents).

CRF₂ receptor knockout mice have also been generated. Deletion of the CRF₂ receptor gene during development has provided conflicting results, with increased anxiety observed in some but not all paradigms tested. These studies suggest that CRF₂ receptor blockade may lead to states of increased anxiety, though it is likely that both the environment and the genetic background on which the knockouts were bred contribute to the behavioral phenotype of these animals.

Research using selective CRF₂ receptor agonists and antagonists have also provided inconsistent results. Several studies have used the selective CRF₂ receptor antagonist antisauvagine-30 (ASV-30), which has been reported to be between 100- and 1000-fold selective for the CRF₂ receptor, depending upon whether the radiolabeled ligand is sauvagine or ASV-30, respectively. Direct administration of ASV-30 into the lateral septum was shown to reduce anxious behavior induced by immobilization stress in the elevated plus maze paradigm or by previous association with foot shock in mice. These behavioral data were corroborated in rats, where direct injection of ASV-30 into the cerebral ventricles reduced anxious behavior in the plus maze, defensive withdrawal, and a conditioned anxiety paradigm.

Compounds which selectively activate (i.e., agonists) the CRF₂ receptor have also been discovered. The peptides urocortin 2 and urocortin 3 are structurally and ancestrally related to CRF but show between 100- and 1000-fold selectivity at the CRF₂ receptor versus the CRF₁ receptor. Urocortin 3 has been shown to mildly suppress locomotion and has an anxiolytic-like profile in mice. However, another study from the same group of researchers demonstrated that urocortin 2 was inactive in the plus maze after acute administration but did increase exploratory behavior in the plus maze 4 h later.

Thus, compounds reported to be both selective agonists and antagonists at the CRF₂ receptor have been shown to possess anxiolytic effects; clearly the precise role of this receptor in modulating stress-induced behaviors remains obscure.

Extrahypothalamic Corticotropin-Releasing Factor Circuits and Depression

Although initially investigated for its role as one of the key modulators of the HPA axis, further research has revealed that CRF controls not only the neuroendocrine, but also the autonomic (e.g., heart rate, blood pressure, etc.), immune, and behavioral responses to stress in mammals. Results from both clinical studies and a rich body of literature conducted primarily in rodents and nonhuman primates have highlighted the importance of extrahypothalamic CRF neurons. In rodents, primates, and humans, CRF and its receptors are heterogeneously localized with high concentrations in a variety of regions, including the amygdala, thalamus, hippocampus, and prefrontal cortex, among others. These brain regions serve as key regulators of the mammalian stress response and emotion.

The presence of CRF receptors in both the dorsal raphe and LC, the origin of the major serotonergic- and noradrenergic-containing perikarya, respectively, also deserves comment because most available antidepressants, including the tricyclic antidepressants and selective serotonin-releasing inhibitors (SSRIs), are believed to act via modulation of the serotonergic and/or noradrenergic systems. The neuroanatomic proximity of CRF and monoaminergic systems provides evidence for an interaction between CRF systems and certain antidepressants, thereby suggesting a mechanism by which antidepressants may affect the CRF system.

Corticotropin-Releasing Factor in Depression

Among the affective disorders, the role of CRF in major depression has been the most thoroughly investigated. Shortly after the isolation and characterization of CRF, a standardized intravenous CRF stimulation test was developed to assess HPA-axis activity. In this paradigm, CRF is administered intravenously (usually at a dose of 1 µg kg⁻¹ or a fixed dose of 100 µg), and the ACTH and cortisol responses are measured at 30-min intervals over a 2–3-h period. Numerous studies have now documented a blunted ACTH and β-endorphin response to intravenously administered ovine or human CRF in depressed patients compared with nondepressed people; the cortisol response in depressed patients and nondepressed

control subjects did not consistently differ. This was one piece of evidence that contributed to the initial hypothesis that CRF may be hypersecreted in a subset of depressed patients.

More specifically, the attenuated ACTH response to CRF has been hypothesized to be due to chronic hypersecretion of CRF from nerve terminals in the median eminence, which results in downregulation of CRF receptors in the anterior pituitary, and/or to chronic hypercortisolemia and its associated negative feedback. CRF receptor downregulation results in reduced responsivity of the anterior pituitary to CRF, as has been repeatedly demonstrated in laboratory animals.

In support of the involvement of extrahypothalamic CRF systems in the pathophysiology of depression, numerous studies have demonstrated elevated CRF concentrations in the cerebrospinal fluid (CSF) of depressed patients, though some discrepant results have also been reported. It is likely that this increased CSF CRF is due to hypersecretion not only from hypothalamic CRF nerve terminals, but from extrahypothalamic sites as well.

Elevated CSF CRF concentrations have also been detected in depressed suicide victims. A reduction in concentrations of CRF in CSF has been reported in healthy volunteers treated with the tricyclic antidepressant desipramine and in depressed patients following treatment with fluoxetine or amitriptyline, providing further evidence of an interconnection between antidepressants, monoamine neurons, and CRF systems. Similar effects have also been reported after electroconvulsive therapy (ECT) in depressed patients.

The alterations in CSF CRF concentrations appear to represent a state, rather than a trait, marker of depression (i.e., a marker of current depression rather than a marker of vulnerability to depression). Furthermore, high and/or increasing CSF CRF concentrations despite symptomatic improvement of major depression during antidepressant treatment may be the harbinger of early relapse.

Consistent with altered concentrations of CRF found in clinical studies of depression, CRF binding site density and messenger ribonucleic acid (mRNA) expression have shown alterations in both preclinical and clinical studies, presumably in response to changes in CRF availability. Our group has previously reported a marked (23%) reduction in the number of CRF binding sites in the frontal cortex of suicide victims compared with controls; we have now replicated this finding in a second study. Raadsheer and colleagues demonstrated an increase in CRF mRNA expression in the PVN of depressed patients compared with controls.

Increased CRF mRNA and decreased CRF₁ mRNA have also been detected in the brains of suicide victims in subregions of the frontal cortex. Although conducted in different laboratories and on different tissue, and keeping in mind the relative difficulty in obtaining and analyzing human tissue, the general pattern of increased CRF concentrations and/or CRF mRNA and the relative decrease in CRF binding sites is consistent with the well-documented phenomenon of receptor up- and downregulation and the CRF hypersecretion hypothesis of depression.

While the exact mechanism contributing to CRF hyperactivity remains obscure, studies from our group and others have documented long-term persistent increases in HPA-axis activity and extrahypothalamic CRF neuronal activity after exposure to early untoward life events, e.g., neglect and child abuse, respectively, in both laboratory animals (rats and nonhuman primates), and patients. Early life stress apparently permanently sensitizes the HPA axis and extrahypothalamic CRF neurons and leads to a greater risk of depression developing later in life.

To measure HPA-axis responsivity to stress in humans, the Trier Social Stress Test (TSST) was developed. This laboratory paradigm involves a simulated 10-min public speech followed by a difficult mental arithmetic task. The TSST has been validated as a potent activator of the HPA axis in humans. Our group reported increased HPA-axis responsivity, presumably due to hypersecretion of CRF, after exposure to the TSST in both depressed and nondepressed women who were exposed to severe physical and emotional trauma as children. These data strongly suggest that CRF systems are particularly sensitive to the effects of early adverse life events.

Corticotropin-Releasing Factor and Anxiety Disorders

Involvement of CRF in anxiety disorders has been well documented in both animal and human studies. Best studied among this diverse group of disorders is posttraumatic stress disorder (PTSD), which has been most commonly studied in Vietnam combat veterans, though it is now widely recognized to occur after other life-threatening situations including victims of rape, natural disasters, and physical and sexual abuse as children. Combat veterans with PTSD exhibit significantly elevated CSF CRF concentrations, as well as alterations in the ACTH response to CRF challenge. In contrast to studies from depressed patients, low serum cortisol and urinary free cortisol levels have been repeatedly, yet unexpectedly, detected in PTSD, especially after dexamethasone administration. One

possible mechanism that has been proposed by Yehuda and colleagues to explain these findings is heightened negative glucocorticoid feedback within the HPA axis in chronic PTSD patients.

Elevated CSF CRF concentrations have not been detected in patients with panic disorder, though a blunted ACTH response after CRF administration has been observed. Increased or normal concentrations of CSF CRF have been observed in patients with obsessive-compulsive disorder (OCD), though significant decreases in CSF CRF concentrations have been noted following a therapeutic response to clomipramine. Patients with generalized anxiety disorder (GAD), however, exhibit similar CSF CRF concentrations in comparison to normal controls.

Not surprisingly, increased concentrations of CSF CRF occur in alcohol withdrawal, a condition of sympathetic nervous system arousal and increased anxiety. In contrast, CSF CRF concentrations are reduced or are normal in abstinent chronic alcoholics with normal plasma cortisol concentrations. Although human data are limited in other addictive states, there is also a burgeoning laboratory animal literature which suggests that CRF may play a role in both the withdrawal and/or stress-induced drug relapse of a number of addictive compounds including nicotine, alcohol, cocaine, and benzodiazepines. To understand the exact role of CRF in anxiety will require considerable additional study.

Small-Molecule Corticotropin-Releasing Factor Antagonists

Although space constraints do not permit an extensive review of the entire preclinical literature, several additional points are worth noting. Findings from numerous studies have shown that when CRF is directly injected into the central nervous system (CNS) of laboratory animals it produces effects reminiscent of the cardinal symptoms of depression, including decreased libido, reduced appetite and weight loss, sleep disturbances, and neophobia. Certainly, by the late 1980s, a number of research groups, including our own, had hypothesized that a lipophilic, small-molecule CRF receptor antagonist that readily penetrates the blood-brain barrier after oral administration would represent a novel class of antidepressant and/or anxiolytic agents.

CRF₁ receptor antagonists possess activity in both animal models of anxiety and depression. CRF receptor antagonists have been tested in many different paradigms, including the elevated plus maze, foot shock, restraint stress, and defensive withdrawal. Pretreatment with CRF receptor antagonists decreases

measures of anxiety induced by stressors. There is also some evidence that CRF receptor antagonists may reduce the effects of drug withdrawal and stress-induced relapse to drug seeking in rats. Based on this premise, newly developed CRF₁ receptor antagonists represent a novel putative class of antidepressants. Such compounds show activity in nearly every preclinical screening test for antidepressant and anxiolytic drugs.

Despite the rich preclinical and clinical literature supporting a potential role for CRF₁ receptor antagonists, there has only been one published study investigating the effects of a CRF₁ receptor antagonist in humans. A small open-label study examining the effectiveness of R121919, a CRF₁ receptor antagonist, in major depression was completed more than 5 years ago. This study of 20 patients showed that R121919 (5–40 mg day⁻¹ or 40–80 mg day⁻¹ for 30 days) was well tolerated and did not significantly affect plasma ACTH or cortisol concentrations at baseline or following CRF challenge. It is important that any potential CRF antagonist not lead to adrenal insufficiency, which can, of course, have grave medical consequences. Hamilton Depression Rating Scale and Hamilton Anxiety Scale severity scores were both significantly reduced following 30-day treatment with this drug. Although this small, pilot study does not provide unequivocal proof, it does provide further evidence that a selective CRF-receptor antagonist may possess antidepressant and antianxiety properties in humans. Although this drug is no longer in clinical development because of hepatotoxicity, several novel CRF₁ antagonists are currently under investigation.

Conclusions and Future Directions

Since the discovery of CRF more than 25 years ago, evidence has accumulated that it plays a preeminent role in the physiology of the stress response. Moreover the evidence of its involvement in pathophysiologic states such as major depression and anxiety disorders is compelling. The recent introduction of small-molecule CRF receptor antagonists as a putative novel class of antidepressant and anxiolytic drugs remains very promising. These compounds block the

actions of exogenous and endogenous CRF in a variety of *in vivo* models, supporting a putative role for these agents in the treatment of stress and/or anxiety and affective disorders. The promising clinical results in patients with depression in the completed open trial of R121919 is of great interest and the results of controlled studies are eagerly awaited.

Acknowledgments

This work was supported in part by MH-58922 (Sylvio O. Conte Center for the Neuroscience of Mental Disorders) and MH-42088.

See Also the Following Articles

Anxiolytics; Central Stress Neurocircuits; Corticotropin Releasing Factor (CRF); Corticotropin-Releasing Factor Receptors; Dex-CRH Test.

Further Reading

- Bremner, J. D., Licinio, J., Darnell, A., et al. (1997). Elevated CSF corticotropin-releasing factor concentrations in posttraumatic stress disorder. *American Journal of Psychiatry* 154, 624–629.
- Gutman, D. A., Owens, M. J. and Nemeroff, C. B. (2005). Corticotropin-releasing factor receptor and glucocorticoid receptor antagonists: new approaches to antidepressant treatment. In: denBoer, J. A., George, M. S. & terHorst, G. J. (eds.) *Current and future developments in psychopharmacology*, pp. 133–158. Amsterdam: Benecke NL.
- Heim, C., Newport, D. J., Miller, A. H., et al. (2000). Long term neuroendocrine effects of childhood maltreatment. *JAMA* 284, 2321.
- Sanchez, M. M., Young, L. J., Plotsky, P. M., et al. (1999). Autoradiographic and *in situ* hybridization localization of corticotropin-releasing factor 1 and 2 receptors in nonhuman primate brain. *Journal of Comparative Neurology* 408, 365–774.
- Steckler, T. and Holsboer, F. (1999). Corticotropin-releasing hormone receptor subtypes and emotion. *Biological Psychiatry* 46, 1480–1508.
- VanPett, K., Viau, V., Bittencourt, J. C., et al. (2000). Distribution of mRNAs encoding CRF receptors in brain and pituitary of rat and mouse. *Journal of Comparative Neurology* 428, 191–212.

Corticotropin-Releasing Factor (CRF) Family of Neuropeptides – Role in Inflammation

A Gravanis and A N Margioris

University of Crete, Heraklion, Greece

© 2007 Elsevier Inc. All rights reserved.

The Corticotropin Releasing Factor Family of
Neuropeptides and Their Receptors
The Inflammatory Response
Corticotropin Releasing Factor Peptides and Immune Cells
Corticotropin Releasing Factor Peptides Modulate
Inflammation at a Local Paracrine Level
Corticotropin Releasing Factor Antagonists as New
Therapeutic Agents of Inflammatory Diseases

Glossary

<i>Corticotropin releasing factor (CRF)</i>	A hypothalamic neuropeptide produced in response to stress; it regulates the production of cortisol from the adrenal glands via the hypothalamic-pituitary-adrenal axis. CRF is also produced in the brain and peripheral tissues, including the immune cells.
<i>Inflammation</i>	The first line of immune defense against a broad spectrum of molecules; the clinical sign characteristics are swelling, redness, fever, and pain. It can be acute or chronic, local or generalized. Exposure to chemicals, microorganisms or physical damage initiates the inflammatory response characterized by a cellular and noncellular (exudative) component.
<i>Urocortins</i>	Neuropeptides with high homology/chemical resemblance to CRF. They are produced in the brain and peripheral tissues, including the immune cells. There are three urocortin molecules: urocortin 1, urocortin 2 (or stresscopin-related peptide), and urocortin 3 (or stresscopin).

The Corticotropin Releasing Factor Family of Neuropeptides and Their Receptors

Corticotropin Releasing Factor

CRF, a 41-amino-acid hypothalamic peptide, is the primary hypothalamic hormone involved in the mammalian response to stress. Stress is defined as any threat to homeostasis. CRF coordinates the homeostatic

mechanisms necessary for an organism to cope with internal or external threats. The stress-regulating CRF is produced in the parvocellular neurons of the paraventricular hypothalamic nucleus, which receive innervation from higher central nervous system (CNS) areas and from the periphery. CRF is released into the hypophyseal portal circulation and travels to anterior pituitary lobe, where it binds to specific CRF sites, the CRF receptors type 1, stimulating the production of proopiomelanocortin (POMC) by corticotroph cells. POMC is then cleaved into adrenocorticotrophic hormone (ACTH), a systemic hormone that stimulates cortisol production from the adrenal gland; cortisol is the principal stress hormone in humans. At the same time, several parvocellular CRF neurons project into locus coeruleus, the command center of sympathetic system, the second major adaptation mechanism to stress. CRF enhances central sympathetic outflow. In addition to its synthesis in the hypothalamus, CRF is also produced throughout the brain, where it works as a peptide neurotransmitter. In general, in the brain CRF stimulates the stress axes and mediates anxiogenic effects. CRF is also present in peripheral tissues including the heart, gastrointestinal (GI) tract, skin, gonads, adrenals, and several components of the immune system including the lymphocytes, splenocytes, neutrophils, mast cells, and monocytes/macrophages.

The Urocortins

There are three urocortin molecules. Urocortin 1 is a 40-amino-acid peptide first isolated from rat brain by molecular cloning in 1995. It exhibits 45% homology with CRF and 63% homology with urotensin, a neuropeptide found in teleost fish. Urocortin 1 strongly binds and stimulates the CRF1 and CRF2 receptors. Like CRF, it can stimulate ACTH secretion from anterior pituitary corticotrophs *in vitro* and *in vivo*. Urocortin 1 is more potent than CRF in decreasing feeding in both meal-deprived and free-feeding rats. Urocortin 1 suppresses feeding via CRF2 receptors.

Urocortin 2, also known as stresscopin-related peptide, and urocortin 3 (stresscopin) are 38-amino-acid peptides and share high homology. They are selective ligands for the CRF2 receptor and thus they share several common effects with urocortin 1 in

appetite suppression, delay of gastric emptying, and cardiovascular effects. They are expressed in several areas of the brain, including the hypothalamus, amygdala, and brain stem, and in the periphery in the GI tract, skin, adrenals, and cardiovascular system.

The Corticotropin Releasing Factor Receptors

The biological effects of CRF peptides are mediated by CRF receptor type 1 (CRF1) and CRF2. CRF2 exists in at least three isoforms, resulting from alternative splicing, the CRF α , CRF β , and CRF γ . The CRF receptors belong to the class B subdivision of G-protein-coupled membrane receptor superfamily. CRF1 and the CRF2 α and CRF2 γ receptors are expressed throughout the brain, whereas CRF2 β is mainly expressed in the periphery, in the lung, skeletal muscles, gonads, cardiac myocytes, skin, GI tract, adrenals, and many immune cells. In the CNS, the CRF2 receptors mediate a central anxiolytic response, opposing the general anxiogenic effect of CRF, which is mediated by the CRF1 receptor. Mice overexpressing CRF show anxiogenic-like responses compared to wild-type mice, whereas mice lacking the CRF1 receptor show an anxiolytic-like behavior. It thus appears that the activation of the CRF1 receptors prepare the organism to respond to stress at a somatic level (activation of the hypothalamus-pituitary-adrenal axis and the sympathetic nervous system) and behaviorally (i.e., the anxiogenic-like response). CRF2 receptors adapt the organism to stress and ameliorate the stress response. Finally, a third type of receptor, the CRF-binding protein (CRFBP), is a pseudoreceptor whose primary function is to compete with CRF receptors for the CRF ligands, thus making it unavailable to the bioactive CRF receptors. CRFBP is mainly expressed in the CNS and peripheral nervous system.

The Inflammatory Response

A Brief Description of Inflammation

Inflammation is the first line of immune defense against a broad spectrum of molecules. This biologically ancient and nonspecific process is part of the innate immunity system. It does not possess a mechanism for memorizing the offending event and, thus, does not confer a lasting immunity against it. If innate immunity is overwhelmed, then a biologically more recent immune mechanism takes control – acquired immunity, a complex process that memorizes the event and confers lasting and specific immunity.

Basic Concepts in the Physiology of the Inflammatory Response

Inflammation can be acute or chronic, local or generalized. Exposure to chemicals, microorganisms or physical damage initiates the inflammatory response characterized by a cellular and a noncellular (exudative) component. The first cells to gather at an inflammation site are the neutrophils (polymorphonuclear cells). Neutrophils live for 1–2 days, and they are gradually replaced by monocytes, which mature to macrophages. All these cells cross the vascular wall and enter the inflammation site. There, they secrete pro- and anti-inflammatory substances, including cytokines, chemokines, and adhesion molecules; clean the area from microbes and cellular debris by phagocytosis; chemoattract other immune cells; and finally act as antigen-presenting cells. The pro-inflammatory cytokines interleukin (IL-)1 and IL-6 and tumor necrosis factor (TNF-) α activate endothelial cells to produce specific membrane receptors (vascular adhesion molecule 1, VCAM-1; intercellular adhesion molecule 1, ICAM-1; selectins), which bind immune cells. The exudative component consists of fluid moving from the intravascular space, due to the increased permeability of the vascular wall. It consists of fibrins and several types of immunoglobulins. These phenomena at an inflammation site produce the clinical characteristics of inflammation: swelling, redness, fever, and pain (tumor, rubor, calor, and dolor, in Latin). If the inflammatory response is large, it can cause systemic symptoms including fever, chills, fatigue, loss of appetite, and mental confusion. If the acute inflammatory response cannot isolate and destroy the offending agent, T lymphocytes take over, shifting the response to chronic inflammation by recruiting B lymphocytes, macrophages, and fibroblasts. Note that arteriosclerosis is now considered the most significant (in terms of threat to health) consequence of a chronic low-level inflammatory response.

Corticotropin Releasing Factor, Urocortins, and Inflammation

CRF regulates the production of cortisol via the hypothalamic-pituitary-adrenal axis. Cortisol, like all glucocorticoids, affects the immune system at multiple levels. Most of its effects are suppressive. Indeed, glucocorticoids are used extensively in suppressing the immune system in autoimmune diseases. During the last 20 years, it became gradually apparent that CRF and the urocortins affect the immune system independently of cortisol in a direct and paracrine

manner via the CRF receptors present, as previously stated, in most immune cells. CRF and the urocortins reach the immune cells and the sites of inflammation either via axonal transport through the autonomic nerves or by being produced ad hoc by the epithelial, the endothelial, and the immune cells themselves. Indeed, it has been shown that human mast cells, lymphocytes, monocytes/macrophages, murine neutrophils, and splenic T lymphocytes produce the CRF family of peptides as well as their receptors. In addition, it is now known that the levels of these peptides and their receptors increase at inflammation sites at a level parallel to the degree of the inflammatory response. In particular, high levels of CRF and urocortins have been found in experimentally induced inflammation, for example, in the inflamed synovia of patients with rheumatoid arthritis, in the inflamed colonic mucosa of patients with ulcerative colitis, and in the inflamed uvea. The concentration of CRF peptides also increases in human endometrium at the sites of egg nidation and implantation, both well-described inflammatory phenomena.

Corticotropin Releasing Factor Peptides and Immune Cells

Mast Cells

It appears that mast cells represent one of the principal immune cell targets of CRF peptides. Indeed, the exposure of mast cells to CRF results in their immediate degranulation of pro-inflammatory mediators, an effect blocked by CRF1 antagonists (molecules that bind to the CRF1 receptor, inhibiting its function). Human mast cells synthesize CRF, urocortins, and CRF1 and CRF2 receptors. The activation of immunoglobulin E receptor induces the synthesis and secretion of both CRF and urocortins. It has recently been suggested that in the skin, mast cells, neurons, and keratinocytes interact via several substances including the CRF peptides. This cutaneous cellular network may be involved in the exacerbation of atopic dermatitis, psoriasis, and other skin ailments following stressful situations (neurogenic inflammation).

Macrophages

Macrophages are among the initiator cells during the inflammatory response and represent the main source of pro-inflammatory cytokines. The activation of macrophages occurs through antigenic signals, including bacterial lipopolysaccharide (LPS), which binds to toll-like receptor (TLR)-4; the activation

of the latter results in an induction of cytokine production and secretion. CRF augments LPS-induced pro-inflammatory cytokine production from macrophages. This effect of the CRF family of peptides appears to be due to their effect on the expression of TLR-4 receptors on macrophages. This effect occurs at the transcriptional level. Indeed, CRF peptides activate TLR-4 in macrophages transfected with a construct containing the proximal region of the TLR-4 promoter linked to luciferase gene. This effect is abolished on mutation of the proximal to the initiation codon PU.1 binding site or on mutation of the activating enhancer binding protein (AP-1) binding element. In addition, CRF peptides block the well-known inhibitory effect of LPS on TLR-4 expression.

Corticotropin Releasing Factor Peptides Modulate Inflammation at a Local Paracrine Level

Multiple published reports suggest that CRF and the urocortins modulate the inflammatory response at a local, ad hoc level. Indeed, several published reports have shown that the CRF peptides affect the inflammatory response at several steps, including neutrophil and macrophage recruitment and migration, phagocytosis, production of cytokines and chemokines, production of adhesion molecules, capillary permeability, angiogenesis, and antigen presentation. Furthermore, the blockade of CRF receptors in several experimental models of inflammation or in stress-induced colitis suppresses immune cell migration, exudate production, the levels of cytokines, ionic and macromolecular permeability of vascular wall, and several other phenomena associated with inflammation.

Corticotropin Releasing Factor Peptides in the Gastrointestinal Tract

As previously stated, CRF, the urocortins, and their receptors are expressed throughout the GI tract. CRF is mainly produced by the normal enterochromaffin cells in the lower GI tract, whereas the urocortins are produced mostly in the upper GI tract by epithelial cells. The activation of the CRF1 receptor stimulates colonic motility, whereas the activation of the CRF2 receptor inhibits gastric emptying.

Gastrointestinal inflammation As previously stated, CRF and urocortins play a major role in the inflammatory phenomena taking place in the GI tract. In patients with ulcerative colitis, CRF intensifies mucosal inflammation in a paracrine manner. Thus, in the lower GI tract, CRF appears to be a major

pro-inflammatory agent. In the stomach, the story appears to be a little different. Urocortin appears to be the major player in the stomach, where it acts as an anti-inflammatory agent. We have studied the role of stomach urocortin in humans with *Helicobacter pylori* (HP) gastritis. The concentration of urocortin is higher in gastric biopsies from patients with active HP gastritis than in normal controls. Following the apparent eradication of HP infection, by clinical and microscopic criteria, urocortin levels increase dramatically compared to pretreatment values. The nonresponders to HP eradication treatment do not appear to be able to elevate their levels of urocortin 1 in their gastric epithelium. In summary, it appears that CRF plays a pro-inflammatory role in inflammatory diseases of the bowel, whereas urocortins play an anti-inflammatory role in the stomach. These differences between upper and lower GI tracts may explain in part the lack of severe chronic inflammatory diseases in the stomach compared to the prevalence of chronic inflammatory bowel syndromes.

Corticotropin Releasing Factor Peptides and Endometrial Decidualization

CRF, the urocortins, and their receptors are expressed in the human endometrium, the maternal reception area for the fertilized egg. During the early stages of egg apposition, endometrial stroma is differentiated into decidua, a thick and sticky area that attracts and traps the fertilized egg. The decidualization of endometrial stroma has characteristics of a local aseptic inflammatory reaction, playing a crucial role in the implantation of the fertilized egg and the continuation of pregnancy. It now appears that the CRF peptides play a major role in the decidualization of endometrial stroma. Indeed, they potentiate the decidualizing effect of progesterone, the principal gonadal steroid hormone of pregnancy. Interestingly, progesterone appears to induce the expression of endometrial CRF, thus forming a local self-augmenting loop with the CRF peptides. In addition to progesterone, several other locally produced pro-inflammatory immune factors modulate decidualization, including prostaglandins and interleukins. Endometrial CRF peptides and the interleukins interact with progesterone to fine-tune and control the phenomenon of decidualization. The sequence of events appears to be as follows. Progesterone, in addition to its strong direct decidualizing effect, induces the production of endometrial CRF peptides, which potentiate the effect of progesterone on stromal decidualization. They also affect local immune modulators inhibiting, among others, the enhancing effect of prostaglandin E2 (PGE2), inducing the inhibitory effect of IL-1 and

stimulating the inducing effect of IL-6. This complex interaction results in the optimal decidualization of human endometrium, making it ready to receive the fertilized egg.

Endometrial implantation of the fertilized egg Apart from the endometrium, the implanting blastocyst (i.e., the fertilized egg composed of approximately 16–32 cells) also secretes several pro-inflammatory mediators, including IL-1 and PGE2. Blastocyst-deriving IL-1 plays an essential role in the implantation process, inducing the production of endometrial CRF peptides at the very site of implantation, rendering thus the local endometrial surface extremely adhesive for its attachment. This process leads to the formation of the egg nidus. The pivotal role of the CRF peptides on egg implantation is supported by several lines of evidence showing a significantly higher concentration of CRF peptides at the early implantation site compared to the interimplantation uterine areas. In addition, IL-1 receptor antagonists block egg implantation and early pregnancy in mice. *In vivo* experiments in mice have shown that intraperitoneal injections of CRF antibodies at day 2 of pregnancy decrease the number of fetuses by at least 60%. These data are further supported by experiments in rats using CRF1 receptor antagonists, which cause a dramatic reduction of the number of implantation sites. Thus, blocking of CRF has an antinidation effect at an early stage of pregnancy. Recent experimental findings support this hypothesis. It should be noted that the implanting egg should be viewed as a semi-xenograft containing paternal antigens, which may activate the maternal immune system, leading to embryo resorption. CRF, by stimulating the expression of the pro-apoptotic FasL protein in the decidual and trophoblastic cells, induces apoptosis (a genetically programmed form of cell death) of the surrounding maternal T lymphocytes, thus rescuing the implanted egg from the maternal immune defense. Expression of Fas ligand by fetal extravillous trophoblast cells induces apoptosis of activated T lymphocytes expressing the Fas receptor on their membranes. Inadequate CRF-mediated self-induction of FasL in extravillous trophoblasts may cause recurrent spontaneous abortions. Indeed, bridging of the local immune privilege may be detrimental to the developing fetus. These findings may also lead to new insights into the pathophysiology of preeclampsia.

Corticotropin Releasing Factor Peptides in the Vascular Endothelium

Multiple reports support the hypothesis that the urocortins, acting via the CRF2 receptors, suppress

locally produced inflammatory responses, the main cause of arteriosclerosis. Indeed, the CRF2 receptors are highly expressed throughout the human cardiovascular system. Interestingly, urocortins and the CRF2 receptor are also highly expressed by the endothelial cells of coronary arteries. Statins, the anti-cholesterol medications, induce the expression of vascular endothelium urocortins and their receptors. The infusion of urocortins affects human vasculature in several ways. They cause a potent vasodilation in an endothelium-independent manner. They protect cardiomyocytes during inflammation of the coronary vasculature via the suppression of angiotensin II-induced reactive-oxygen-species production from vascular endothelial cells. Urocortins also inhibit the apoptosis of mesenteric arterial smooth muscle cells via L-type calcium channels in spontaneously hypertensive rats.

Corticotropin Releasing Factor Peptides in the Heart

Urocortin 1, urocortin 3, and the CRF2 receptor are expressed in human heart. Urocortins, via the CRF2 receptor, have multiple and potent cardioprotective effects. They have an anti-apoptotic effect on cardiomyocytes. More specifically, following cardioplegic arrest and subsequent reperfusion of hearts, cardiomyocytes die from apoptosis through mitochondrial injury. Urocortins protect cardiomyocytes from mitochondrial-mediated apoptosis. Urocortins also protect the ischemia-reperfusion injury of the heart muscle by inhibiting free radical activity. Finally, the infusion of urocortins ameliorates the hemodynamic, hormonal, and renal changes that occur during heart failure.

CRF cardiomyocyte apoptosis Apoptosis contributes to myocyte cell loss in a variety of cardiac pathologies, including cardiac failure and pathologies related to ischemia/reperfusion injury. The apoptotic process involves pro- and anti-apoptotic proteins. Indeed, apoptosis occurs when the equilibrium between pro- and anti-apoptotic mediators is disturbed. Cardiac urocortin expression is enhanced by ischemia/reperfusion injury *in vitro*, and the addition of exogenous urocortins reduces cell death caused by ischemia/reperfusion. In the isolated perfused heart, urocortins improve hemodynamic recovery and partially prevent the reduction in high-energy phosphates following ischemia/reperfusion. It now appears that cardiac urocortins may play a central role in cardiac myocyte protection against iatrogenic ischemia/reperfusion injury associated with bypass surgery. Indeed, it is now obvious that synthetic CRF2 receptor agonists may be extremely important

in cardioplegic therapies, ischemia/reperfusion, and heart failure.

Corticotropin Releasing Factor Peptides in the Skin

A brain–skin connection with local neuroimmunoenocrine circuitry underlies the pathogenesis of allergic and inflammatory skin diseases triggered or aggravated by stress. CRF and the urocortins, along with other neuropeptides, represent part of the brain–skin connection. CRF peptides are produced in the skin, and their production is regulated by ultraviolet radiation, glucocorticoids, and the phase of hair cycle. The skin also expresses the CRF1 and CRF2. CRF1 is expressed in the epidermal and dermal compartments, whereas CRF2 is predominantly expressed in dermal structures. CRF affects several physiological parameters of epidermal melanocytes via the CRF1 receptor. Most prominently, CRF acts on epidermal melanocytes as a survival, anti-apoptotic factor under the stress of starvation as well as an inhibitor of growth factor-induced cell proliferation.

Skin disorders CRF has been implicated in the pathogenesis of skin disorders exacerbated by stress. Indeed, CRF induces skin vascular permeability through neurotensin acting on mast cells. High levels of the CRF1 receptor transcript have been found in the inflammatory lesions of contact dermatitis and chronic urticaria. It should be noted that the local application of CRF on skin reduces the inflammatory and hyperalgesic processes following a vesicant-induced skin injury. CRF has been implicated in the development of acne, seborrhea, androgenetic alopecia, skin aging, xerosis, and other skin disorders associated with alterations in lipid formation of sebaceous origin. CRF has been also found to play a role in the development of psoriasis.

Corticotropin Releasing Factor Peptides in Muscle Inflammation

Fibromyalgia CRF peptides may play a role in muscle physiology and in the inflammatory phenomena taking place in this tissue. Fibromyalgia (FMS) is a debilitating disorder characterized by chronic diffuse muscle pain, fatigue, sleep disturbance, depression, and skin sensitivity. There are no genetic or biochemical markers, and patients often present with other comorbid diseases, such as migraines, interstitial cystitis, and irritable bowel syndrome. Diagnosis includes the presence of 11/18 trigger points, but many patients with early symptoms might not fit this definition. Pathogenesis is still unknown, but there has been evidence of increased CRF and

substance P (SP) in the cerebrospinal fluid (CSF) of FMS patients, as well as increased SP, IL-6, and IL-8 in their serum. Increased numbers of activated mast cells were also noted in skin biopsies. The hypothesis was put forward that FMS is a neuroimmunoendocrine disorder in which increased release of CRF and SP from neurons in specific muscle sites triggers local mast cells to release pro-inflammatory and neurosensitizing molecules. There is no curative treatment, although low doses of tricyclic antidepressants and the serotonin-3 receptor antagonist tropisetron are helpful. Recently, new formulations containing the natural anti-inflammatory and mast cell-inhibitory flavonoid quercetin may hold promise as therapeutic agents.

Muscle thermogenesis CRF may also induce skeletal muscle thermogenesis (i.e., loss of energy as heat), protecting patients against excessive intramyocellular lipid storage and hence against skeletal muscle lipotoxicity, local inflammation, and, most important, the all-important development of insulin resistance. Finally, the administration of synthetic CRF2 receptor agonists may hold promise as a treatment for the prevention of skeletal muscle atrophy.

Corticotropin Releasing Factor Antagonists as New Therapeutic Agents of Inflammatory Diseases

The CRF system has been implicated in the pathophysiology of a variety of diseases, including depression, epilepsy, addiction, and Alzheimer's. The CRF system has been also shown to play an important role in GI disorders. During the last decade, an important effort was made to develop nonpeptide micromolecular antagonists of CRF receptors with potential therapeutic applications in the treatment of several of these diseases. These, low-molecular-weight, non-peptide CRF antagonists belong to four main classes of molecules: monocyclic, bicyclic, and tricyclic compounds and a miscellaneous category that includes compounds that are unlike the traditional small-molecule CRF antagonists.

In addition to the promising therapeutic role that nonpeptide CRF1 antagonists may have for the treatment of CNS-related diseases, these molecules may be useful in the treatment of localized chronic inflammation in several systems. For instance, these compounds open new therapeutic options in the control of lower-GI inflammatory conditions associated to CRF, such as chronic inflammatory bowel syndromes and irritable bowel syndrome. On the other hand, urocortin or synthetic micromolecular CRF2 agonists may be useful in the treatment of upper-GI

inflammatory diseases. Similarly, in reproduction, because CRF1 blockade appears to prevent implantation by reducing the inflammatory reaction of endometrium to the invading blastocyst, CRF1 antagonists may play a role as a new class of nonsteroidal inhibitors of implantation. Given the promising future of CRF antagonists in the therapy of depression and anxiety disorders, their ability to cause hypofertility and early miscarriages should seriously be borne in mind in prescribing these compounds.

Synthetic CRF2 receptor agonists may be extremely important in cardioplegic therapies, ischemia/reperfusion, and heart failure, protecting against iatrogenic ischemia/reperfusion injury associated with bypass surgery. Finally, the administration of synthetic CRF2 receptor agonists may hold promise as a treatment preventing skeletal muscle atrophy.

See also the Following Articles

Anxiety; Cardiovascular System and Stress; Corticotropin Releasing Factor (CRF); Depression and Manic-Depressive Illness; Depression Models; Stress Management and Cardiovascular Disease; Ulceration, Gastric; Urocortins; Inflammation; Cardiovascular Disease, Stress and; Stress, NPY and Cardiovascular Diseases.

Further Reading

- Bale, T. L., Giordano, F. J. and Vale, W. W. (2003). A new role for corticotropin-releasing factor receptor-2: suppression of vascularization. *Trends in Cardiovascular Medicine* **13**, 68–71.
- Bale, T. L. and Vale, W. W. (2004). CRF and CRF receptors: role in stress response and other behaviors. *Annual Review of Pharmacology and Toxicology* **44**, 525–557.
- Coste, S. C., Quintos, R. F. and Stenzel-Poore, M. P. (2002). Corticotropin-releasing hormone-related peptides and receptors: emergent regulators of cardiovascular adaptations to stress. *Trends in Cardiovascular Medicine* **12**, 176–182.
- Gravanis, A. and Margioris, A. N. (2005). The corticotropin-releasing factor (CRF) family of neuropeptides in inflammation: potential therapeutic applications. *Current Medicinal Chemistry* **12**, 1503–1512.
- Karalis, K., Muglia, L. J., Bae, D., et al. (1997). CRF and the immune system. *Journal of Neuroimmunology* **72**, 131–136.
- Keller, P. A., Elfick, L., Garner, J., et al. (2000). Corticotropin releasing hormone: therapeutic implications and medicinal chemistry developments. *Bioorganic & Medicinal Chemistry* **8**, 1213–1223.
- Makrigiannakis, A., Zoumakis, E., Kalantaridou, S., et al. (2003). Corticotropin-releasing hormone (CRF) and immunotolerance of the fetus. *Biochemical Pharmacology* **65**, 917–921.

Martinez, V., Wang, L., Million, M., et al. (2004). Urocortins and the regulation of gastrointestinal motor function and visceral pain. *Peptides* 10, 1733–1744.

McCarthy, J. R., Heinrichs, S. C. and Grigoriadis, D. E. (1999). Recent advances with the CRF1 receptor: design of small molecule inhibitors, receptor subtypes and clinical indications. *Current Pharmaceutical Design* 5, 289–315.

Slominski, A., Wortsman, J., Pisarchik, A., et al. (2001). Cutaneous expression of corticotropin-releasing hormone (CRF), urocortin, and CRF receptors. *FASEB Journal* 15, 1678–1693.

Theoharides, T. C., Donelan, J. M., Papadopoulou, N., et al. (2004). Mast cells as targets of corticotropin-releasing factor and related peptides. *Trends in Pharmacological Sciences* 25, 563–568.

Corticotropin-Releasing Factor Receptors

D E Grigoriadis

Neurocrine Biosciences, Inc., San Diego, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D E Grigoriadis, volume 1, pp 586–593, © 2000, Elsevier Inc.

Introduction

Corticotropin Releasing Factor Receptor Family

Pharmacological Characteristics

Distribution of Corticotropin Releasing Factor Receptor Subtypes

Summary and Conclusion

Glossary

Adenylate cyclase

One of the signaling molecules for G-protein-coupled receptors. On activation by a specific G-protein, this transmembrane protein either increases or decreases the rate of conversion of ATP to cAMP.

Corticotropin releasing factor (CRF)

A 41-amino-acid peptide primarily secreted by the hypothalamus into the hypophyseal portal vasculature to act on the pituitary and cause the release of adrenocorticotrophic hormone.

G-protein

A heterotrimeric protein that has high affinity for GTP and interacts with the cytoplasmic domains of a receptor, transducing the signal from the activated (ligand-bound) receptor to the second-messenger signaling protein inside the cell.

G-protein-coupled receptor

A protein that typically spans a cell membrane seven times with extracellular and intracellular loops coordinating the binding of a ligand outside the membrane with a signal transduction mechanism inside the membrane.

Hypothalamic-pituitary-adrenal (HPA) axis

The hormonal pathway that is thought to mediate the primary stress response. CRF secreted from the hypothalamus acts on the anterior pituitary to release adrenocorticotrophic hormone (ACTH), which in turn acts on the adrenals to produce and release glucocorticoids. Glucocorticoids feed back to the pituitary to decrease ACTH release and also act on the hypothalamus to decrease CRF release in a negative feedback loop.

Introduction

More than half a century ago, Geoffrey Harris first proposed the concept that the hypothalamus plays a primary role in the regulation of the pituitary-adrenocortical axis and extended the notion put forth by Walter B. Cannon that an organism's ability to maintain homeostasis in the face of external stressors was coordinated by a specific portion of the central nervous system (CNS) operating in an automatic fashion. Subsequently, during the 1950s, Guillemin and Rosenberg and Saffran and Schally independently observed the presence of a factor in extracts of the hypothalamus that could stimulate the release of adrenocorticotrophic hormone (ACTH, corticotropin) from anterior pituitary cells *in vitro*. This extract was termed corticotropin releasing factor (CRF). Although CRF was the first hypothalamic hypophysiotropic factor to be recognized, its chemical identity remained unknown largely due to the presence in hypothalamic extracts of other weaker secretagogues of ACTH secretion such as vasopressin, catecholamines, and angiotensin II. These agents, along with their synergistic effects with CRF on ACTH secretion and in combination with the relative lack of specificity of the *in vitro* bioassays, hindered the purification of this peptide. The development of radioimmunoassays for ACTH and quantitative

in vitro methods for assaying hypophysiotropic substances along with the use of ion exchange and high-performance liquid chromatographic techniques led to the successful purification of CRF from sheep hypothalamic extracts in 1981 by Vale and his colleagues at the Salk Institute. They were the first to report the isolation, characterization, synthesis, and *in vitro* and *in vivo* biological activities of a 41-amino-acid hypothalamic peptide that fulfilled all the criteria for a corticotropin releasing factor. Thus, from the first basic understanding of homeostasis in the early 1940s, this neuropeptide has gained widespread acknowledgement of its role in mediating the HPA axis and regulating an organism's response to physical, emotional, and environmental stress.

Since the elucidation of the structure of CRF, there have been countless studies that have demonstrated the importance of this peptide and its role as a neurotransmitter and neuromodulator in the CNS apart from its role in the direct stimulation of ACTH from the pituitary through hypothalamic release. These reports have detailed the anatomical distribution of both CRF-containing cell bodies and fibers in the CNS and have identified the peptide in regions of the brain shown to clearly regulate autonomic function and mediate the behavioral responses to stress. For example, the neocortex contains high densities of CRF-containing neurons that are localized primarily to prefrontal, cingulate, and insular areas and that appear to regulate behavioral actions of the peptide. This extrahypothalamic distribution of CRF suggests that this peptide has numerous effects in the brain, and its widespread but discrete localization in various brain regions, both rat and human, place it in a position to regulate mood and stress-related behaviors. In fact, the exogenous administration of CRF into discrete brain regions has demonstrated direct interactions between the CRF and the norepinephrine systems (such as in the locus coeruleus) or the serotonergic systems (such as in the dorsal and median raphe nuclei) and corresponds to the effects seen under various stress paradigms in which these systems are activated. This further implicates this particular neuropeptide in mediating the central mechanisms through which various stressors can alter behavior.

The relatively recent cloning of multiple CRF receptor subtypes and the identification of a specific binding protein for the peptide (known as CRF-BP) has precipitated a new era in CRF research. In addition, the identification of a new related family of endogenous peptides, the urocortins, has expanded the notion that this system is under multiple regulatory factors and plays a much more complex role in the CNS than previously thought. This article focuses primarily on recent studies that elucidate the molecular

biological, pharmacological, anatomical, and functional characteristics of mammalian CRF receptors.

Corticotropin Releasing Factor Receptor Family

Molecular Biology/Receptor Structure

The CRF receptors belong to the superfamily of G-protein-coupled receptors and, as such, have been shown to contain seven putative transmembrane domains and function through the coupling of a stimulatory guanine-nucleotide binding protein. These receptors also fall within the recently described, and still-growing, family of gut-brain neuropeptide receptors, which includes receptors for calcitonin, vasoactive intestinal peptide (VIP), parathyroid hormone, secretin, pituitary adenylate cyclase-activating peptide, glucagon and growth hormone-releasing factor. These receptors all share considerable sequence homology, and all stimulate adenylate cyclase in response to their respective agonist activation. Almost all of the initial characterization of the CRF receptors was performed using various native tissues from a variety of species. As such, there was only one form of the receptor that could be experimentally discerned. As will be described later in the chapter, this form was subsequently identified as the CRF₁ receptor. It has been very difficult to pharmacologically study the CRF₂ receptor in native tissues due to its discrete localization and relatively low abundance. Cloning of these receptor subtypes however, has enabled the pharmacological and biochemical characterization of these proteins and allowed the discovery of potential therapeutics through a variety of chemical-screening strategies.

CRF₁ Receptors

Vale and colleagues first cloned the CRF₁ receptor in 1993 from a human Cushing's anterior pituitary corticotrophic adenoma using the technique of expression cloning. The protein was characterized as a 415-amino-acid protein with seven putative transmembrane domains and five potential extracellular N-linked glycosylation sites. In addition, the protein contained putative sites for protein kinase C phosphorylation in the first and second intracellular loops and in the C-terminal tail, as well as casein kinase II and protein kinase A phosphorylation sites in the third intracellular loop. The human gene was subsequently localized on chromosome 17 at position 17q12–22. Shortly after the elucidation of this human form, several groups identified the CRF₁ form of the receptor from a variety of other species. All species of CRF₁ mRNAs thus far identified encode proteins of 415

amino acids that are 98% identical to one another. The human CRF₁ gene contains at least two introns and is found in a number of alternative splice forms, none of which to date has been found to have any physiological significance. This is in contrast to the mouse, in which the gene has been identified as containing at least 12 introns. The various kinase and phosphorylation sites predicted on these proteins, although not fully characterized, may serve as regulatory elements in the control of receptor expression and/or function. The potential N-linked glycosylation sites on the N-terminal extracellular domain are characteristic of most G-protein-coupled receptors and confirm the glycosylation profiles determined by earlier chemical-affinity cross-linking studies. Indeed, the molecular weight predicted from the deglycosylated forms of the CRF₁ in the earlier biochemical studies was virtually identical to that obtained from the cloned amino acid sequence. Furthermore, mRNA distribution of the CRF₁ in the rat or human brain correlates extremely well with the previously identified binding sites for radiolabeled CRF using autoradiographic techniques in frozen brain sections. This correlation was expected (retrospectively) because the peptide CRF itself has a low affinity for the CRF₂ and does therefore not label this subtype in brain sections to any detectable level. There have been reports in the literature of splice variants of the CRF₁ mRNA; however, these have not yet been shown to produce physiologically functional proteins. This does not preclude the notion that splice variants of the CRF₁ may exist; however, their full characterization will have to await functional elucidation in a native tissue.

CRF₂ Receptors

Shortly after the cloning and characterization of the CRF₁, another subtype of this family for this hormone was identified through homology. This receptor constituted the CRF₂ subtype of the family and initially represented two splice forms termed CRF_{2(a)} and CRF_{2(b)}. The CRF_{2(a)}, originally described by Lovenberg et al., is a 411-amino-acid protein with approximately 71% identity to the CRF₁. The CRF_{2(b)}, which has been cloned from rat, mouse, and human, contains 431 amino acids and differs from the CRF_{2(b)} isoform in that the first 34 amino acids in the N-terminal extracellular domain are replaced by a unique sequence, 54 amino acids in length. The CRF_{2c} has most recently been identified and, thus far, is found only in the human brain. This splice variant uses yet a different 5' alternative exon for its N-terminus and replaces the first 34-amino-acid sequence of the CRF_{2(a)} with a unique 20-amino-acid sequence.

Between the CRF₁ and CRF_{2s}, there exist very large regions of amino acid identity, particularly between transmembrane domain five and transmembrane domain six. This similarity strongly argues for conservation of biochemical function because it is this region that is thought to be the primary site of G-protein coupling and signal transduction. All three CRF₂ Receptors contain five potential N-glycosylation sites, which are analogous to those found on the CRF₁. The genomic structure of the human CRF₂ gene is similar to that of the mouse CRF₁, already described, and has 12 introns, the last 10 of which interrupt the coding region in identical positions. These gene sequences, however, diverge significantly at the 5' end. The chromosomal mapping of the human CRF₂ gene has been localized to chromosome 7 p21-p15.

Endogenous Ligands

The peptide structure of CRF, once the amino acid sequence was determined, was found to bear strong similarities to two nonmammalian peptides, sauvagine from the frog and urotensin I from the fish. All of these peptides have the ability to potently release ACTH from cultured rat pituitary cells, which identifies them as high-affinity agonists for CRF Receptors. Although the mammalian CRF peptides from a variety of species all had high affinity for the CRF₁, they had lower affinity for the CRF₂, creating some doubt as to the likelihood that these peptides were the physiological activators for the CRF₂. It was interesting and quite unexpected, however, that, in addition to their high affinity for the CRF₁, the nonmammalian peptides sauvagine and urotensin I also had a high affinity for the CRF₂. In 1995, using antibodies for urotensin I, the first of a unique and novel family of mammalian peptides related to CRF was discovered in the rat and termed urocortin (UCN)1. UCN1 had a high affinity for both CRF Receptor subtypes, also bound with high affinity to the CRF-BP, and was localized in areas corresponding to the distribution of the CRF₂ itself. The human homolog of UCN was later identified and cloned from a human brain library and defined as the endogenous ligand for the CRF₂. Since then, two other mammalian peptides have been identified from the mouse and human, termed UCN2 (also referred to as stresscopin-related peptide) and UCN3 (also referred to as stresscopin). Unlike UCN1, UCN2 and UCN3 are much more selective for the CRF₂ subtype and have little or no affinity for the CRF-BP. These peptides have enabled a greater understanding of this increasingly complex system in the regulation of central stress responses.

To date, there are no known endogenous physiological antagonists for the CRF Receptor system;

however, various truncations of the agonist peptides, with deletions of the first 8–11 N-terminal amino acids, have resulted in potent peptide antagonists. Of these, α -helical CRF(9–41), d-Phe CRF(12–41), and astressin are among the most potent and well characterized, having a high affinity for both the CRF₁ and CRF₂ receptors and being able to functionally block receptor activity both *in vitro* and *in vivo*. Recently, chemical modifications have resulted in two additional peptides that now have 400- to 500-fold selectivity for the CRF₂ over the CRF₁. Antisauvagine-30, an N-terminally truncated form of sauvagine, and the more recent cyclized astressin-2B have been used to dissect out specific roles of the CRF₂ subtype in a variety of *in vivo* studies. The pharmacological profiles of these peptide agonists and antagonists have elucidated some of the discrete biochemical requirements of the two receptor subtypes and have greatly enhanced the understanding of the physiology of the CRF system and the important role this neurohormone plays in mediating the stress response. Furthermore, there is now increasing evidence that this system is fully integrated in diseases such as anxiety and depression and may provide a target for novel therapeutics in the treatment of these disorders.

Pharmacological Characteristics

The rat and human forms of the CRF Receptor family have been transfected into a variety of mammalian cell lines that do not normally express these receptors, including COS-7 (monkey kidney), CHO (Chinese hamster ovary), and Ltk⁻ (mouse fibroblast) cells. Although rat tissue was commonly available for the study of the CRF₁, the expression in stable mammalian cell lines of the human forms of the CRF Receptor subtypes enabled, for the first time, direct biochemical and pharmacological study without the requirement for scarce human brain tissue.

The CRF Receptor subtypes exhibit a clear and distinct pharmacological profile with respect to the endogenous and related peptides known to activate this system. In addition, the CRF_{2(a)} and CRF_{2(b)} display very subtle differences in the pharmacological profiles with respect to one another, which suggests that some specific conformation of the protein can be attributed to the N-terminal extracellular domains and is critical for the high-affinity binding of peptides. As a direct result of the distinct pharmacological difference between the CRF₁ and CRF₂ subtypes, it became necessary to radiolabel a nonmammalian CRF-like ligand ([¹²⁵I]sauvagine) in order to completely characterize the CRF₂ subtype. Previous studies examining the detailed binding characteristics of the CRF peptides with respect to the peptide structure have revealed

that all the CRF peptides – be they agonist or antagonist – must be amidated at their C-terminus for potent activity. The deamidation of any of these peptides results in a virtually complete loss of activity at either receptor subtype, suggesting that at least some similarity between the two subtypes must exist.

Ligand Binding Profile of CRF₁ Receptors

CRF₁ receptors expressed in the mammalian cell lines described demonstrate reversible, saturable, high-affinity binding to CRF and its related peptides with the pharmacological and functional characteristics comparable to those found in the variety of animal or human tissues previously described in the literature. Human and rat CRF₁ receptors in stably transfected Ltk⁻ cells demonstrate binding to a single homogeneous population of receptors with apparent affinities (K_D) of 130 and 168 pM and receptor densities (B_{max}) of 97 and 588 fmol/mg protein, respectively. The pharmacological rank orders of potency in the stably transfected cell lines with either the human or rat CRF₁ receptors were identical to the established profile for the CRF receptor in the rat frontal cortex, where UCN1, urotensin I, sauvagine > ovine CRF (oCRF), rat/human CRF (r/hCRF), bovine CRF > astressin > D-Phe r/hCRF(12–41) > α -helical oCRF(9–41) $\gg$ UCN2, UCN3, astressin-2B, r/hCRF(6–33), r/hCRF(9–33), r/hCRF(1–41)OH, VIP, arginine vasopressin (AVP). These data demonstrate that, although there are subtle differences between the rat and human forms of the receptor in terms of amino acid sequence, the differences are not significant enough to alter the ability of these peptides to bind or interact with this subtype.

Ligand Binding Profile of CRF₂ Receptors

The CRF₂ isoforms have also been expressed in stable mammalian cell lines and demonstrated to bind CRF-related peptides and function through the stimulation of cAMP production. In general, although some minor differences exist in the rank order profile of the CRF₂ splice variants to one another (mainly related to fragments of CRF peptides), as a subtype they are quite distinct from either the CRF₁ subtype or the CRF-BP. When expressed in stable cell lines, these receptors exhibit the following pharmacological rank order profile: UCN1, UCN2, sauvagine, urotensin I, astressin-2B > UCN3 > astressin > r/hCRF > D-Phe CRF(12–41), α -helical CRF(9–41) $\gg$ oCRF $\gg$ r/hCRF(6–33), oCRF(6–33), r/hCRF(1–41)OH, GHRH, AVP, and VIP. The data thus far on the CRF_{2(a)} suggests a rank order of potencies similar to the other two splice variants with UCN, sauvagine > urotensin I > r/hCRF. Without exception, it is

clear that the mammalian CRF peptides oCRF and r/hCRF have weaker affinity for the CRF₂ subtype, whereas UCN has equal affinity for both the CRF₁ and CRF₂ receptors and UCN2 has selectivity for the CRF₂ receptors. Interestingly, the relatively weaker affinity of UCN3 for the CRF₂ receptors may indeed suggest that this peptide is the endogenous ligand for a different and thus far undiscovered CRF receptor subtype.

Guanine Nucleotide Characteristics of Corticotropin Releasing Factor Receptors

Characteristic of receptor systems that are coupled to adenylate cyclase as a second-messenger system is the exclusive link to a guanine nucleotide-binding protein. These transducing proteins mediate the ligand-binding event at the receptor to an activation of an intracellular pathway, in this case beginning with the conversion of ATP to cAMP. *In vitro*, guanine nucleotides have been shown to decrease the specific binding of agonists for a great number of seven-transmembrane neurohormone or neurotransmitter receptors coupled to G-proteins, presumably by uncoupling the G-protein from the receptor and changing its conformation such that it now has low affinity for its agonist. Both CRF receptor subtypes transfected in stable mammalian cell lines were used to determine the guanine nucleotide sensitivity of the cloned receptors. Guanine nucleotides inhibited the binding of [¹²⁵I] oCRF (for CRF₁ receptors) or [¹²⁵I]sauvagine (for either CRF₁ receptors or CRF₂ receptors) by 50–70% (where 100% inhibition was defined by 1 μ M unlabeled r/hCRF). The nonhydrolyzable forms of guanine nucleotides, Gpp(NH)p and GTP- γ -S, were more potent than GTP itself. The order of potencies were GTP- γ -S > Gpp(NH)p > GTP with median effective dose (ED₅₀) values of ~20–45, 200–800 and 2500–3000 nM, respectively. These effects on the binding in both human and rat CRF₁- or CRF₂-stable cell lines were specific for the guanine nucleotides because the adenosine nucleotide ATP was ineffective in altering the binding at equimolar concentrations (ED₅₀ > 10 000 nM). Recently, using a variety of peptide ligands in conjunction with the guanine nucleotide GTP- γ -S, three distinct receptor states have been elucidated for the CRF₁. These states represent RG, the G-protein-coupled form of the receptor sensitive to guanine nucleotides; R, the G-protein-uncoupled form of the receptor; and R0, a high-affinity uncoupled state of the CRF₁ that is GTP- γ -S insensitive. Although the nature of this state requires much further investigation for the CRF system, this conformation has been described for other GPCRs, including the β_2 -adrenergic, M2 muscarinic, formyl peptide, and natural killer (NK)1 cell receptors where

they have been identified to be in a complex with arrestin. Significantly, these three states for the CRF₁ have also been confirmed to exist in native rat tissues such as the cerebellum and may offer new approaches for the discovery of molecules targeting conformationally restricted binding sites.

Second-Messenger Characteristics of Corticotropin Releasing Factor Receptors

In addition to the adenylate cyclase system, other signal transduction mechanisms have been postulated to be involved in the actions of CRF. For example, CRF has been shown to increase protein carboxyl-methylation and phospholipid methylation in AtT-20 cells. Preliminary evidence suggests that CRF may regulate cellular responses through products of arachidonic acid metabolism. Furthermore, although the evidence in anterior pituitary cells suggests that CRF does not directly regulate phosphatidylinositol turnover or protein kinase C activity, the stimulation of protein kinase C, either directly or by specific ligands (vasopressin or angiotensin II), enhances CRF-stimulated adenylate cyclase activity, ACTH release, and inhibits phosphodiesterase activity. Thus, the effects of CRF in native tissues such as the anterior pituitary cells, neurons, and other cell types expressing CRF receptors are likely to involve complex interactions among several intracellular second-messenger systems. All this evidence has thus far been gathered from native tissues that exhibit binding profiles consistent with the CRF receptor. These studies have most probably been examining second-messenger profiles of the CRF₁ in various tissues, primarily due to the ubiquitous nature of this subtype compared to the CRF₂ in the rat. Clearly, the most common and well-studied second-messenger system for the CRF receptor is the coupling to the adenylate cyclase system. In fact, there is no evidence to date that the heterologous expression of hCRF receptors transfected into mammalian cells function in any manner other than through the stimulation of the production of cAMP, unless forced to couple through overexpressed, genetically manipulated cell systems.

In the pituitary gland, CRF initiates a cascade of enzymatic reactions, beginning with the receptor-mediated stimulation of adenylate cyclase, which ultimately regulates proopiomelanocortin (POMC) peptide secretion and possibly synthesis. POMC-derived peptide secretion mediated by the activation of adenylate cyclase is dose-related and exhibits appropriate pharmacology consistent with the activation of the CRF₁. Similarly, in the brain and also in spleen, the pharmacological rank order profile of CRF-related peptides for stimulation of adenylate cyclase is analogous to the profile seen in the pituitary

and in keeping with the affinities of these compounds for receptor binding. In addition, the putative CRF receptor antagonist α -helical oCRF(9–41) inhibits the CRF-stimulated adenylate cyclase observed in brain and spleen homogenates.

To further characterize the stable CRF receptor transfectants, the ability of CRF-related and unrelated peptides to stimulate CRF₁- or CRF₂-mediated cAMP production from the cells has been examined. The rank order of the peptides in stimulating cAMP production in these cells was in keeping with their rank order in inhibiting [¹²⁵I]oCRF and [¹²⁵I]sauvagine binding to the CRF₁ and CRF₂, respectively, and at least for the CRF₁, this rank order was virtually identical to the second-messenger pharmacological profile previously reported in the brain, pituitary, and spleen.

In addition to the stimulation studies, CRF-stimulated cAMP production could be competitively inhibited by the putative peptide antagonists D-Phe CRF(12–41) and α -helical CRF(9–41) in both CRF₁- and CRF₂-expressing cell lines, demonstrating that the expressed CRF receptors in these cell lines were indeed functional. Interestingly, both D-Phe CRF(12–41) and α -helical CRF(9–41) inhibited the stimulated cAMP response with the same median effective concentration (EC₅₀) in both receptor subtype preparations. This suggests that, although there is some fundamental difference in the pharmacological specificity of these two receptor subtypes when examining agonists, they must still possess some similarities in their structure of the binding site, at least as far as peptide antagonists are concerned. These results clearly suggest that not only subtype-specific and subtype-selective compounds can be identified for these two receptor subtypes but that mixed antagonists are possible and could bring to light possible subtle functions of the two receptor subtypes.

Distribution of Corticotropin Releasing Factor Receptor Subtypes

Receptor autoradiographic studies and more recent *in situ* hybridization and immunocytochemical studies have localized CRF binding sites and CRF Receptor mRNA, respectively, in slide-mounted sections of many tissues from a variety of species, identifying these proteins in anatomically and physiologically relevant areas. For example, in the pituitary gland, CRF₁ expression is detectable in both the anterior and intermediate lobes, with particularly high expression in clusters within the anterior lobe. In fact, there exists a good correlation between the actions of CRF on corticotropes of the anterior pituitary and the distribution of [¹²⁵I]oCRF or [¹²⁵I]sauvagine binding sites and CRF₁ mRNA. Within the anterior

lobe, CRF₂ expression is detectable in only a few scattered cells. Thus, in terms of HPA axis activity, CRF₂ receptors may mediate CRF effects at the level of the hypothalamus, whereas CRF₁ receptors are responsible for CRF-induced changes in ACTH release in pituitary corticotropes.

Although there is a general correspondence of the distribution of [¹²⁵I]oCRF or [¹²⁵I]sauvagine binding sites and CRF₁ or CRF₂ mRNA in the brain with the functional actions of CRF, there are several brain areas with some apparent discrepancies. For example, CRF has potent electrophysiological and behavioral effects in the locus coeruleus, but the distribution of binding sites and CRF₁ or CRF₂ mRNA is not noteworthy in this brain-stem nucleus. Physiologically, it is tempting to speculate that other subtypes that are as yet unidentified may be responsible for the actions of the neurotransmitter in those regions where CRF seems to have profound actions but no appreciable labeling of the receptor is observed. The localization and distribution of CRF₂ Receptors in the mammalian brain, coupled with the discovery of the urocortins, has accounted for some of these discrepancies. However, species differences in the localization of receptor subtypes and peptides among rodent, non-human primate, and human tissues have made it somewhat difficult to extrapolate behavioral studies in rodent models to human effects. It is important to note that, because most of the autoradiographic studies in the literature thus far in rodents have used [¹²⁵I]oCRF or various forms of [¹²⁵I]rat/human CRF, the primary labeling, localization, and distribution are almost exclusively attributable to the CRF₁ subtype.

Within the rat brain, CRF₁ expression was very high in neocortical, cerebellar, and sensory relay structures, whereas CRF₂ expression was generally confined to subcortical structures, including the lateral septal region, the bed nucleus of the stria terminalis, the amygdaloid area, and the olfactory bulb. The contrast in expression patterns between CRF Receptor subtypes was particularly evident within the septal region; CRF₂ mRNA expression was very high in the lateral septal nuclei but very low in the medial septum, a region where CRF₁ mRNA abundance was most evident. The lateral septum, by virtue of widespread reciprocal connections throughout the brain, is implicated in a variety of physiological processes. These range from higher cognitive functions, such as learning and memory, to autonomic regulation, including food and water intake. In addition, the septum plays a central role in classic limbic circuitry and is thus important in a variety of emotional conditions, including fear and aggression. The lateral septum thus acts as both an integrator of limbic circuitry and an interface between the telencephalic and diencephalic

areas. The high-level of CRF₂ expression in this area suggests a role for CRF₂ Receptors in modulating limbic circuitry at the level of the lateral septum.

The heterogeneous distribution of CRF₁ and CRF₂ mRNA suggests distinctive functional roles for each receptor in CRF-related systems. CRF₂ mRNA was evident throughout the rostrocaudal extent of the hypothalamus, particularly within the paraventricular nucleus (PVN), whereas CRF₁ expression was limited. The distribution of cells expressing CRF₂ mRNA within the PVN coincides with the cellular distribution of CRF mRNA, suggesting a possible autoreceptor role for CRF₂ Receptors in this nucleus. Because the CRF neurons of the PVN play a classical hypophysiotropic role in controlling ACTH release from the anterior pituitary, it is tempting to speculate that CRF₂ Receptors may act to selectively regulate hypophysiotropic-and/or autonomic-related CRF neurons. As such, this receptor subtype may be central to the control of the mammalian HPA stress system. Further, more detailed studies using double-labeling of the peptides and their receptors are required to determine the discrete role of this subtype. Whatever the case may be, the high level of CRF₂ expression within the hypothalamic structures is certainly suggestive of a role for CRF₂ Receptors in modulating autonomic CRF neurons. With regard to HPA axis activity, we may speculate that the PVN CRF₂ Receptors could also act to modulate CRF neurosecretory neurons in response to stress, in essence providing a short-loop feedback role. Ono and colleagues have proposed such a role for CRF in a positive ultrashort-loop feedback effect on stress-induced ACTH release. In this regard, the comparatively lower affinity of CRF for the CRF₂ subtype may be of physiological design to allow activation of CRF₂ Receptors only under conditions of chronic stress or prolonged CRF release.

A more detailed examination of the precise localization of the CRF₂ splice variants, CRF_{2(a)} and CRF_{2(b)}, indicates discrete anatomical distributions in the rat. The CRF_{2(a)} form is primarily expressed within the brain, whereas the CRF_{2(b)} variant is localized in both the CNS and periphery. Within the brain, it appears that the CRF_{2(a)} form represents the predominant neuronal CRF₂ variant, whereas the CRF_{2(b)} splice form is localized on nonneuronal elements, such as the choroid plexus of the ventricular system and cerebral arterioles. In the periphery, CRF_{2(b)} mRNA is expressed at high levels in both the cardiac and skeletal muscle, with lower levels evident in both the lung and intestine. The CRF_{2(c)} isoform has yet to be identified in the rodent; however, reverse transcription polymerase chain reaction (RT-PCR) analysis of human brain mRNA demonstrated expression in the

septum, amygdala and hippocampus, and frontal cortex. The full characterization of the CRF_{2(c)} subtype has not yet been elucidated in terms of native characteristics and function largely due to its nonexistence in rodent tissues. This type of species selectivity will make it difficult to completely examine the role of the CRF₂ subtype unless suitable animal models can be found.

Finally, in addition to the expression in the CNS, CRF receptors have been localized to a number of peripheral tissues. The gastrointestinal (GI) tract is an organ system that has a discrete localization pattern of CRF₁ and CRF₂ and is sensitive to a variety of stressors thought to act via the CRF system, both centrally and peripherally. GI responses are differentially affected by stress and these can be reproduced by the central administration of CRF. For example, stress decreases gastric emptying (upper GI tract) while increasing colonic motility and transit (lower GI tract). These effects seem to be independent of the HPA response and most likely reflect central autonomic function mediated by central CRF₂ and CRF₁, respectively. The fact that both these receptor subtypes are also found throughout the stomach and bowel, including the colon, suggests a much more complex interaction of this stress response system in the gut and that the precise mechanisms of regulation remain to be elucidated.

Summary and Conclusion

CRF is the primary regulator of the endocrine, autonomic, and behavioral response to stress via mechanisms in both the brain and the periphery. The endocrine stress response begins with the activation of the HPA axis, resulting in the release of CRF into the hypophysial portal vasculature. CRF then mediates its actions through cell surface receptors coupled to the activation of adenylyl cyclase. Centrally, CRF is found throughout the brain and plays a major role in mediating emotional, psychological, or behavioral stress. As described here, CRF receptors can be classified into two distinct subtypes: the CRF₁ (411 amino acids) and CRF₂, with the latter comprising three different isoforms, CRF_{2(a)}, CRF_{2(b)}, and CRF_{2(c)} (encoded by 415, 431, and 397 amino acid proteins, respectively). These receptors differ with respect to their structures, tissue distributions, and pharmacological specificities for CRF-related ligands. UCN1, a mammalian CRF-related peptide with close sequence homology to fish urotensin, interacts with a high affinity with all CRF receptors, whereas UCN2 and UCN3 are more selective for the CRF₂ subtype. The areas of distribution of these CRF receptors in the brain are correlated well with the

immunocytochemical distribution of the CRF and urocortin pathways as well as the pharmacological sites of action of CRF and the urocortins. The species differences that exist with respect to the distinct subtypes of the CRF₂ precludes our complete correlation of animal studies with humans; however, the identification of the multiple subtypes has led to an enhanced understanding of this complex system. It is becoming increasingly clear that the multiple receptor subtypes must mediate different functional and pathophysiological actions of this family of neuropeptides, and future studies aimed at identifying additional subtypes or peptides will serve to fully elucidate the role of this key neurohormone in the pathophysiology of stress and stress-related disorders.

Further Reading

- Chen, R., Lewis, K. A., Perrin, M. H., et al. (1993). Expression cloning of a human corticotropin-releasing-factor receptor. *Proceedings of the National Academy of Sciences USA* 90(19), 8967–8971.
- De Souza, E. B. (1987). Corticotropin-releasing factor receptors in the rat central nervous system: characterization and regional distribution. *Journal of Neuroscience* 7(1), 88–100.
- Grigoriadis, D. E. (2005). The corticotropin releasing factor receptor: a novel target for the treatment of depression and anxiety-related disorders. *Expert Opinion on Therapeutic Targets* 9(4), 651–684.
- Grigoriadis, D. E., Haddach, M., Ling, N., et al. (2001). The CRF receptor: structure, function and potential for therapeutic intervention. *Current Medicinal Chemistry: Central Nervous System Agents* 1, 63–97.
- Hauger, R. L., Grigoriadis, D. E., Dallman, M. F., et al. (2003). International Union of Pharmacology. 36: Current status of the nomenclature for receptors for corticotropin-releasing factor and their ligands. *Pharmacological Reviews* 55(1), 21–26.
- Heinrichs, S. C. and Tache, Y. (2001). Therapeutic potential of CRF receptor antagonists: a gut-brain perspective. *Expert Opinion on Investigational Drugs* 10(4), 647–659.
- Hoare, S. R. J. (2005). Mechanisms of peptide and nonpeptide ligand binding to class B G-protein coupled receptors. *Drug Discovery Today* 10(6), 417–427.
- Kehne, J. and De Lombaert, S. (2002). Non-peptidic CRF1 receptor antagonists for the treatment of anxiety, depression and stress disorders. *Current Drug Targets: CNS & Neurological Disorders* 1(5), 467–493.
- Keller, P. A., Elfick, L., Garner, J., et al. (2000). Corticotropin releasing hormone: therapeutic implications and medicinal chemistry developments. *Bioorganic & Medicinal Chemistry* 8(6), 1213–1223.
- Kostich, W. A., Chen, A., Sperle, K., et al. (1998). Molecular identification and analysis of a novel human corticotropin-releasing factor (CRF) receptor: the CRF2gamma receptor. *Molecular Endocrinology* 12(8), 1077–1085.
- Liaw, C. W., Lovenberg, T. W., Barry, G., et al. (1996). Cloning and characterization of the human corticotropin-releasing factor-2 receptor complementary deoxyribonucleic acid. *Endocrinology* 137(1), 72–77.
- Monnikes, H., Tebbe, J. J., Hildebrandt, M., et al. (2001). Role of stress in functional gastrointestinal disorders: evidence for stress-induced alterations in gastrointestinal motility and sensitivity. *Digestive Diseases* 19(3), 201–211.
- Perrin, M. H. and Vale, W. W. (1999). Corticotropin releasing factor receptors and their ligand family. *Annals of the New York Academy of Sciences* 885, 312–328.
- Sanchez, M. M., Young, L. J., Plotsky, P. M., et al. (1999). Autoradiographic and in situ hybridization localization of corticotropin-releasing factor 1 and 2 receptors in nonhuman primate brain. *Journal of Comparative Neurology* 408(3), 365–377.
- Skelton, K. H., Owens, M. J. and Nemeroff, C. B. (2000). The neurobiology of urocortin. *Regulatory Peptides* 93 (1–3), 85–92.
- Tache, Y., Martinez, V., Wang, L., et al. (2004). CRF1 receptor signaling pathways are involved in stress-related alterations of colonic function and viscerosensitivity: implications for irritable bowel syndrome. *British Journal of Pharmacology* 141(8), 1321–1330.
- Vale, W., Spiess, J., Rivier, C., et al. (1981). Characterization of a 41-residue ovine hypothalamic peptide that stimulates secretion of corticotropin and beta-endorphin. *Science* 213(4514), 1394–1397.

Cortisol Awakening Response

A Steptoe

University College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

Definition

Origins of the Cortisol Awakening Response

Measurement of the Cortisol Awakening Response

Factors Affecting the Cortisol Awakening Response

Significance of Cortisol Awakening Response for Stress Research

Relationship with Health Outcomes

Conclusion

Glossary

Area under the curve A method of estimating aggregate cortisol output from a series of timed samples. It is typically calculated by assuming that the secretion rate is uniform across the time period. For example, if the salivary cortisol level is 18.0 nmol/liter at T_1 , 22.0 at T_2 (14 min later), and 26.0 at T_3 (20 min after T_2), the area under the curve is calculated as $((18.0 + 22.0) / 2) \cdot 14 + ((22.0 + 26.0) / 2) \cdot 20 = 760.0$.

Carotid intima-media thickness The depth of the intima-medial layer of the wall of the carotid arteries, assessed noninvasively using duplex ultrasonography. This is an index of subclinical atherosclerosis and correlates with coronary artery disease.

Cortisol awakening response (CAR) The increase in cortisol during the first hour after waking up, typically peaking at 20–40 min after waking.

Definition

The CAR is the change in cortisol concentration that occurs during the first hour after waking from sleep. The CAR has been studied in detail only over the past decade, following the introduction of salivary cortisol sampling. This has allowed people to collect samples noninvasively under normal life conditions at home instead of under the contrived conditions of the sleep or endocrinological laboratory. The CAR is now a topic of intense investigation.

Cortisol levels are low in the night but rise in the early hours before waking. After waking up, most people show a further rise, the concentration peaking 20–40 min later. This is followed by a progressive

reduction of cortisol during the day. There are several reasons why the CAR has become an important research topic. First, there is evidence that the CAR is under somewhat independent control from cortisol output during the remainder of the day. Several studies have shown little association between the CAR and cortisol levels during the rest of the day or the slope of cortisol concentration decline into the evening. Second, twin studies have documented a genetic influence on the CAR that is distinct from the heritability of daytime cortisol levels. Third, the CAR is associated with stress and health in potentially significant ways, suggesting that it is a useful marker of hypothalamic-pituitary-adrenocortical (HPA) function.

A number of different measures of the CAR have been devised, depending on how many saliva samples are obtained over the first hour after waking. Serial sampling suggests that that peak CAR is observed around 30 min after waking, so the simplest method of assessment involves measures at two time points: immediately after waking and 30 min later; the CAR is then defined as a simple difference score. More elaborate assessments involve measurements of several time points, such as immediately after waking and 15, 30, 45, and 60 min later; under these circumstances, the area under the curve can be computed to estimate total cortisol output. Difference scores and area under the curve measures are consistently intercorrelated.

Origins of the Cortisol Awakening Response

The secretion of cortisol is pulsatile, and it has been estimated that each secretory burst induces an increase in salivary free cortisol of approximately 2.5 nmol liter⁻¹. The CAR is a result of one to four secretory bursts over the period following waking. It is consistent across days under stable environmental conditions, but changes in magnitude when conditions vary.

The CAR is sensitive to light in the environment and can be enhanced by augmented light exposure early in the day. It has been proposed that this effect is mediated by the suprachiasmatic nucleus of hypothalamus, which is light sensitive and is involved in circadian regulation. Direct neural connections between the suprachiasmatic nucleus and the adrenal cortex have been identified and may modulate the sensitivity of the adrenal gland to adrenocorticotrophic hormone.

The functional role of the CAR has not been established. It may be important metabolically in mobilizing the organism during the transition from sleep to waking. However, no association between the CAR and blood glucose levels has been observed. A role in immunological regulation has also been suggested, but direct evidence is lacking at present. The relevance of psychological factors is apparent from some studies, and it is further endorsed by the observation that the CAR is abolished in individuals with global amnesia, even though the remainder of the circadian rhythm of cortisol output is retained. This hints at the possibility that the CAR may depend either on recent life experiences or on anticipation of the activities and events of the forthcoming day.

Measurement of the Cortisol Awakening Response

Participants in CAR studies are instructed to take saliva samples immediately after waking and then at later defined intervals. Saliva is collected using cotton rolls or by directly spitting into test tubes, and participants are instructed not to brush their teeth, eat, drink, or smoke throughout the sampling period because these factors can influence cortisol values. Because the CAR is dynamic, accurate timing is critical, and failure to adhere to the sampling schedule may confound studies and produce misleading results. If the waking sample is delayed, the CAR may already have commenced, leading to an apparent diminution in the magnitude of the rise of the next 20–30 min. Similarly, if later samples are mistimed, then the magnitude of the CAR will not be assessed correctly.

The significance of timing errors has been demonstrated using saliva collection devices that are time-stamped. One study showed that more than 50% of community volunteers failed to take the 30-min post-waking sample at the correct time. Noncompliant participants produced smaller CARs. The impact of delays in collecting the waking sample is illustrated in [Figure 1](#). This figure compares salivary cortisol profiles during the first hour of the day in elderly (65- to 80-year-old) volunteers who reported no delay in taking the first sample after waking and those who admitted delaying for 10 min or more. The no delay group showed a typical CAR with increases from waking to 30 min, whereas a decrease during same period was found in the delay group.

Several methods of improving the accuracy and reliability of CAR measurement have been proposed. Electronically time-tagged saliva collection methods provide data about the intervals between measurements, but information about the timing of waking up is still provided by the participants. Asking

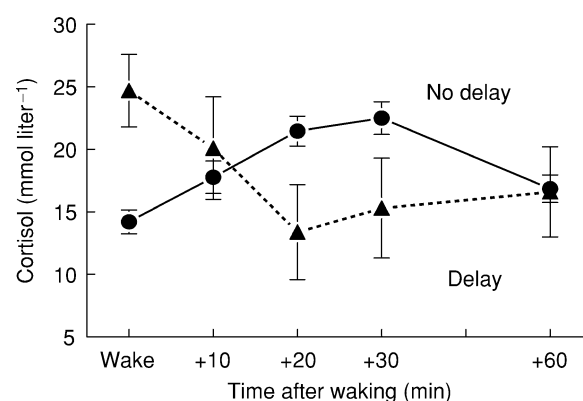


Figure 1 Mean salivary cortisol levels immediately after waking and at 10, 20, 30 and 60 minutes after waking in elderly volunteers who delayed (dashed line) and did not delay (solid line) saliva sampling after waking in the morning. Error bars are standard errors of the mean. From Wright, C. E., & Steptoe, A. (2005). Subjective socioeconomic position, gender and cortisol responses to waking in an elderly population. *Psychoneuroendocrinology*, 30, 582–590. Used with permission.

participants to record when they wake and when they take each sample provides a check and limits can be set on acceptable delays, but this method again depends on self-reports. The measurement of waking time using actigraphy is being explored. Another approach is to check for consistency across days and to exclude individuals who show widely discrepant CAR patterns on different days.

Factors Affecting the Cortisol Awakening Response

A wide range of factors potentially affect the CAR. In this section, other factors other than stress that affect the magnitude of the CAR are briefly described.

Gender and Age

Several investigations have shown that the CAR is greater in women than men. Results have not been completely consistent and may depend on whether measures are taken on a work or leisure day. The greater CAR in women may also be affected by the stressful demands they often encounter early in the day. Women are frequently responsible for preparing not only themselves for the day ahead but also children and for doing household chores as well. Most studies of the CAR have shown little association with age in adult populations.

Time of Waking

There has been controversy about the association between time of waking and the CAR. Several studies have shown that the CAR is greater in people who

wake up early than in those who wake later, but no association with time of waking has also been reported in a number of large-scale investigations. The issue is complicated by two factors. First, people who say they wake up late in the morning may in fact have woken earlier and have subsequently dozed. Their CAR may therefore have partly taken place prior to the reported waking time. Second, waking early may be a reflection of greater stress, worry, or work demands, which themselves influence the CAR. In one study, the greater CAR recorded in people waking before starting an early work shift was no longer significant once concomitant stress had been taken into account.

Sleep Duration and Quality

The majority of studies to date have not shown any association between sleep duration or time of going to bed and the CAR. Experiments in which healthy participants were deliberately disturbed during the night showed no change in the CAR when they later woke in the morning. The CAR is also unaffected by whether the person wakes spontaneously or with an alarm clock. On the other hand, a negative relationship between cortisol levels on waking and sleep quality has been described, and cortisol levels on waking are lower in individuals diagnosed with primary insomnia.

Smoking

Data relating smoking with the CAR are mixed, but it has recently been shown that the CAR is greater in smokers than in nonsmokers, even when they have not smoked during the early morning period.

Significance of Cortisol Awakening Response for Stress Research

The relationship between the CAR and stress-related processes is intriguing. The early studies of the CAR indicated greater responses in people under chronic stress due to work or other demands, suggesting that the CAR might be a convenient stress measure. The CAR is also greater on work than on leisure days, indicating that the stressfulness of the forthcoming day may be significant. This is illustrated in [Figure 2](#), which shows CAR results from a sample of working men and women measured on work and leisure days. It is evident that cortisol levels on waking did not differ between the two days. However, the CAR was substantially greater on work days than on weekend days. In addition, this study found a larger CAR in women than men on work days but not on weekend days.

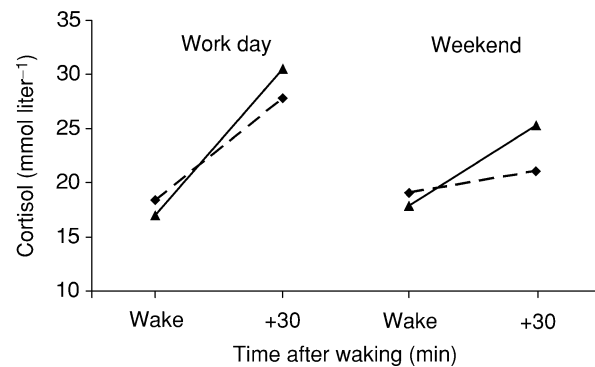


Figure 2 Mean salivary cortisol on waking (Wake) and 30 minutes later (+30) on a work day and weekend day in men (solid lines) and women (dashed lines). Error bars are standard errors of the mean. From Kunz-Ebrecht, S. R., Kirschbaum, C., Marmot, M., & Steptoe, A. (2004). Differences in cortisol awakening response on work days and weekends in women and men from the Whitehall II cohort. *Psychoneuroendocrinology*, 29, 516–528. Used with permission.

Subsequent studies have identified several conditions in which the CAR is apparently reduced, indicating that stress-related processes may both accentuate and inhibit the CAR. [Table 1](#) summarizes the evidence to date. An enhanced CAR has consistently been found to relate to work stress and to measures such as job strain, work overload, and overcommitment to work. Both clinical depression and depressive symptoms within the nonpsychiatric population are characterized by a heightened CAR. By contrast, a diminished CAR has been recorded in people suffering from posttraumatic stress disorder and chronic fatigue and in those who have experienced early life adversity. Other types of chronic strain, such as caring for a demented relative, have been associated with a heightened CAR, while the effects of early life adversity have been variable. Research on socioeconomic status has shown inconsistent results, with elevated CAR in some studies and reduced CAR in others. The literature on burnout is also mixed, with most studies showing clinically defined burnout problems to have no effect on the CAR.

Nearly all of these effects are cross-sectional, so the causal significance of associations is unclear. However, one recent study assessed changes in financial strain over a 3-year period and demonstrated that a reduction in financial strain was accompanied by a decrease in the CAR, indicating a parallel between changes in chronic stress and the magnitude of the CAR.

The explanation for these variations among studies is not clear. They may reflect genuine differences between the patterning of the CAR in relation to different types of stressful experience. But it should be noted that several of these studies did not control rigorously for the timing of sampling. It is conceivable

Table 1 Psychosocial factors and the cortisol awakening response

<i>Factors associated with increased CAR</i>	<i>Factors associated with reduced CAR</i>	<i>Inconsistent effects</i>
Work vs. leisure days Chronic stress, work overload Job strain, overcommitment to work Depression and depressive symptoms Financial strain Loneliness Neuroticism Caregiving	Posttraumatic stress disorder Chronic fatigue	Socioeconomic status Burnout Early life adversity

that in some conditions of chronic stress, people are less reliable in the timing of measurements, leading to an apparent diminution in the CAR.

Relationship with Health Outcomes

The CAR has not been studied extensively in relation to health to date. One study divided respondents into people who did and did not report any health problem or chronic disease and showed a smaller CAR in the less healthy group. There have been a number of other interesting observations suggesting that associations may exist with physical health risk. For example, the CAR has been positively associated with waist/hip ratio in men, suggesting that it may relate to HPA dysfunction in abdominal adiposity. Another study assessed relationships with the rate of healing of a wound imposed experimentally on healthy volunteers. Wound healing was slower in people reporting high levels of subjective stress and in those with a larger CAR, indicating that stress may impair wound healing in part through disturbances in cortisol regulation. Another indication of the health relevance of the CAR is the observation that in women a greater CAR was positively associated with progression in carotid intima-media thickness and so correlated with the development of atherosclerosis. Other studies have not measured the CAR directly but, rather, assessed the cortisol level in the early morning. The values recorded in these studies may reflect the peak CAR or some time point soon afterward. Notably, the early morning cortisol level is positively associated with risk of future clinical depression in prospective studies of adolescents and adults and with cardiovascular disease risk factors.

Conclusion

The CAR has only begun to be studied in recent years in stress research, but it is already providing valuable information about the relationship among HPA function, psychosocial factors, and health risk.

See Also the Following Articles

Corticosteroids and Stress; Demand–Control Model; Effort-Reward Imbalance Model; Posttraumatic Stress Disorder, Neurobiology of; Salivary Cortisol; Sleep, Sleep Disorders, and Stress.

Further Reading

- Backhaus, J., Junghanns, K. and Hohagen, F. (2004). Sleep disturbances are correlated with decreased morning awakening salivary cortisol. *Psychoneuroendocrinology* 29, 1184–1191.
- Clow, A., Thorn, L., Evans, P., et al. (2004). The awakening cortisol response: methodological issues and significance. *Stress* 7, 29–37.
- Goodyer, I. M., Herbert, J., Tamplin, A., et al. (2000). Recent life events, cortisol, dehydroepiandrosterone and the onset of major depression in high-risk adolescents. *British Journal of Psychiatry* 177, 499–504.
- Ebrecht, M., Hextall, J., Kirtley, L. G., et al. (2004). Perceived stress and cortisol levels predict speed of wound healing in healthy male adults. *Psychoneuroendocrinology* 29, 798–809.
- Eller, N. H., Netterstrom, B. and Allerup, P. (2005). Progression in intima media thickness – the significance of hormonal biomarkers of chronic stress. *Psychoneuroendocrinology* 30, 715–723.
- Harris, T. O., Borsanyi, S., Messari, S., et al. (2000). Morning cortisol as a risk factor for subsequent major depressive disorder in adult women. *British Journal of Psychiatry* 177, 505–510.
- Kudielka, B. M., Broderick, J. E. and Kirschbaum, C. (2003). Compliance with saliva sampling protocols: electronic monitoring reveals invalid cortisol daytime profiles in noncompliant subjects. *Psychosomatic Medicine* 65, 313–319.
- Kunz-Ebrecht, S. R., Kirschbaum, C., Marmot, M., et al. (2004). Differences in cortisol awakening response on work days and weekends in women and men from the Whitehall II cohort. *Psychoneuroendocrinology* 29, 516–528.
- Pruessner, M., Hellhammer, D. H., Pruessner, J. C., et al. (2003). Self-reported depressive symptoms and stress levels in healthy young men: associations with the

- cortisol response to awakening. *Psychosomatic Medicine* 65, 92–99.
- Step toe, A., Kunz-Ebrecht, S. R., Brydon, L., et al. (2004). Central adiposity and cortisol responses to waking in middle-aged men and women. *International Journal of Obesity and Related Metabolic Disorders* 28, 1168–1173.
- Step toe, A., Brydon, L. and Kunz-Ebrecht, S. (2005). Changes in financial strain over three years, ambulatory blood pressure, and cortisol responses to awakening. *Psychosomatic Medicine* 67, 281–287.
- Wolf, O. T., Fujiwara, E., Luwinski, G., et al. (2005). No morning cortisol response in patients with severe global amnesia. *Psychoneuroendocrinology* 30, 101–105.
- Wright, C. E. and Steptoe, A. (2005). Subjective socioeconomic position, gender and cortisol responses to waking in an elderly population. *Psychoneuroendocrinology* 30, 582–590.
- Wüst, S., Wolf, J., Hellhammer, D. H., et al. (2000). The cortisol awakening response – normal values and confounds. *Noise and Health* 7, 77–85.

Cortisol See: Corticosteroids and Stress.

C-Reactive Protein

W J Kop and A A Weinstein

Uniformed Services University of the Health Sciences,
Bethesda, MD, USA

© 2007 Elsevier Inc. All rights reserved.

C-Reactive Protein: Function and Measurement
Relationship between C-Reactive Protein and
Cardiovascular Disease
Stress-Related Psychosocial Cardiovascular Risk Factors
and C-Reactive Protein
Conclusion and Future Directions

Glossary

Acute coronary syndrome A general term indicating a group of clinical symptoms compatible with acute myocardial ischemia, including myocardial infarction and unstable angina pectoris.

Acute-phase response Complex series of biological reactions in response to infection, physical trauma, malignancy, or other physical challenges, involving blood clotting and inflammatory parameters. The acute-phase response can be initiated short-term (within min) and may last several days.

Antibody Protein produced by the immune system that recognizes and helps fight infections and other foreign substances in the body.

Atherosclerotic plaque

Atherosclerosis is the hardening (loss of elasticity) and narrowing of medium or large arteries. The plaque is also called an atheroma, which involves the abnormal inflammation-dependent accumulation of white blood cells, lipids, and other cells within the arterial walls.

C-reactive protein (CRP)

A marker of inflammation and the acute-phase response that can be measured in blood samples.

Cytokine

A wide range of small proteins released by cells that have specific effects on cell interactions and cell function. Cytokines include the interleukins (ILs; e.g., IL-6), lymphokines, and cell signal molecules (e.g., tumor necrosis factor and interferons). Cytokines may stimulate (pro-inflammatory) or inhibit (anti-inflammatory, e.g., IL-4) the inflammation response. Cytokines that are bound to antibodies have a stronger effect on the immune system than unbound cytokines. One of the white blood cell types involved in the immune response. The two major types of lymphocytes are T cells and B cells.

Lymphocyte

Monocyte

One of the white blood cell types that ingests microorganisms, other cells, or foreign substances. When a monocyte enters tissue, it develops into a macrophage (which surrounds, kills, and removes microorganisms).

<i>Myocardial ischemia</i>	Insufficient blood flow to part of the heart muscle, reflecting an imbalance between blood supply and cardiac demand.
<i>Pathogen</i>	An organism that causes disease in another organism, such as viruses and bacteria.
<i>Phagocytosis</i>	The process of ingesting and destroying microorganisms or other foreign substances by phagocytes. Phagocytes are cells that can ingest and destroy these foreign substances.
<i>Vital exhaustion</i>	A state characterized by unusual fatigue and loss of energy, increased irritability, and feelings of demoralization. Often precedes cardiac events. In the original conceptualization of the exhaustion construct, the term vital was included to reflect the far-reaching consequences of this condition on daily life function (similar to vital depression). We will refer to this construct as “exhaustion” in the remainder of this chapter.

C-Reactive Protein: Function and Measurement

C-reactive protein (CRP) is a member of the pentraxin protein family (molecular weight 120 kD) and a well-established inflammation marker. In the early 1930s, Tillett and Francis discovered CRP in the serum of patients with acute inflammation as a protein that reacted with the C polysaccharide of pneumococcus. Subsequent research has indicated that elevated CRP is a characteristic indicator of the acute-phase response in inflammatory disorders. The acute-phase response involves a shift in the plasma proteins released by the liver (including a decrease in albumin and an increase in CRP). The three acute-phase reactants that have been most extensively investigated are CRP, fibrinogen, and serum amyloid A (SAA), and these acute-phase proteins are well-documented inflammatory markers and mediators of atherosclerosis.

CRP is released from hepatocytes and its release is primarily stimulated by the cytokine interleukin (IL)-6. CRP mimics the function of an antibody, and is capable of complement activation. More specifically, CRP interacts with phagocytic cells at an injury site, binding to the phosphorylcholine of cell-wall polysaccharides from a wide range of bacteria and fungi. When CRP binds to a bacterium, the pathogen will become vulnerable for phagocytosis, and the immunological complement cascade will be activated. Unlike antibodies, CRP and other acute-phase reactants have no structural diversity and are not specifically released or targeted. In addition to the usual hepatic storage and release, CRP can also be stored

in the hepatic endoplasmic reticulum, and this storage diminishes during inflammation thus leading to more efficient CRP secretion. CRP may also be released from other sources, including atherosclerotic plaques, vascular tissue, lymphocytes, and monocytes. CRP is a primary marker of the acute-phase reaction because of its rapid (24–28 h) and marked response to a wide variety of inflammatory conditions. Thus, CRP is part of the innate (nonadaptive) host response to infection and other pathogens.

CRP can be measured by a variety of methods including radioimmunoassay, immunonephelometry, immunoturbidimetry, immunoluminometry, and enzyme-linked immunosorbent assay (ELISA). Methods that enable precise measurements at low CRP concentrations have been defined as high sensitivity (hs) CRP assays. High sensitivity CRP is necessary to enable assessments in apparently healthy individuals using a lower limit of approximately 0.3 mg l^{-1} , with an assay imprecision of less than 10% at a CRP concentration of less than 1.0 mg l^{-1} . Based on the biological variability of CRP, the proposed allowable analytical error has been proposed to lie between 15% and 32%. Additional standardization is still necessary for some of the available assays, but a few well-established commercial hsCRP assays are available. CRP has a long half-life (approximately 19 h), and levels are relatively stable over time, with a 5-year correlation of approximately 0.4.

Diseases with a clinically significant inflammatory component are commonly associated with CRP levels exceeding 10 mg l^{-1} . However, the elevated CRP-related risk for cardiovascular disease is documented to occur at levels between 3.0 mg l^{-1} and 10.0 mg l^{-1} . Such low-grade elevations occur in approximately one-third of the adult population, primarily in the following settings: (1) diseases with tissue dysfunction or minor damage that are not primarily inflammatory (e.g., hypertension and atrial fibrillation); (2) conditions related to metabolic or hypothalamic pituitary dysregulation (e.g., obesity, insulin resistance, sleep apnea, exhaustion, and depression); and (3) adverse health behaviors (cigarette smoking, low physical activity, and high-fat diet).

Twin studies suggest a substantial genetic component (monozygotic concordance for CRP = 0.40). Multiple genetic polymorphisms for elevated CRP levels have been documented, but none are associated with an altered CRP amino acid sequence. A wide range of relatively minor environmental challenges (e.g., secondary smoke inhalation, pollutants, and estrogen-containing substances) and inflammatory stimuli related to minor injuries (e.g., dental problems and minor respiratory infection) are known to elicit low-grade CRP responses. Depending on the nature

of the research question and study sample, these background factors need to be taken into account by statistical methods or research design.

Relationship between C-Reactive Protein and Cardiovascular Disease

Epidemiological studies have demonstrated that CRP and other acute-phase reactants predict incident and recurrent cardiovascular events. Based on population studies, clinical cut-off points for CRP have been set as: low ($<1 \text{ mg l}^{-1}$), moderate ($1\text{--}3 \text{ mg l}^{-1}$), and high ($>3 \text{ mg l}^{-1}$). CRP levels of more than 10 mg l^{-1} should lead to referral for noncardiac diagnosis because of the high likelihood of an active inflammatory disorder. Three processes are important to consider in the relationship between CRP and cardiovascular disease: (1) the role of inflammation in the onset and gradual progression of atherosclerosis; (2) immune system involvement in acute atherosclerotic plaque rupture; and (3) secondary increases in inflammatory parameters following transient myocardial ischemia and/or myocardial infarction, or damage to the vasculature.

CRP may reflect early stages of atherosclerosis characterized by cell adhesion and engulfment of lipids and T cells. CRP may promote these gradual immune system-related processes. Elevated CRP levels have long-term predictive value for adverse cardiovascular outcomes, although some evidence suggests that the predictive value is most pronounced in the first 2 years of follow-up. The predictive value of CRP is statistically independent of other inflammatory predictors of cardiovascular disease (IL-6 and tumor necrosis factor (TNF)- α), and standard cardiovascular risk factors such as cholesterol.

In advanced coronary atherosclerosis, CRP may also promote transient increases in the risk of plaque rupture by promoting plaque instability and activation. Plaque rupture is associated with activation of the blood clotting process, which is linked to inflammatory parameters. The resulting sudden obstruction of coronary blood flow may lead to sustained myocardial ischemia and infarction.

In addition to promoting gradual atherosclerosis and plaque instability, elevated CRP levels may also reflect secondary inflammatory responses in patients with cardiovascular disease. Myocardial tissue damage or transient ischemia, infection, and mechanical cell injury are all known to activate numerous biological pathways, including the cytokine and coagulation systems. CRP levels tend to peak at 48–54 h postmyocardial infarction, and the increase is proportional to the magnitude of myocardial damage as documented by troponin levels.

Recent meta-analyses have confirmed the significant association between CRP and adverse cardiovascular risk, although the magnitude of risk is lower (odds ratio 1.58, 95% confidence interval 1.48 to 1.68) than originally reported. In addition, recent Mendelian randomization studies indicate that although genetic polymorphisms give rise to elevated CRP levels, these genetic factors are not associated with increased cardiovascular risk. These findings suggest that CRP may be a marker of the underlying atherosclerotic process and/or a measure that partly reflects other risk indicators. Additional studies are needed to further establish the epidemiological and experimental evidence for the role of CRP and other inflammatory markers in adverse cardiovascular disease progression. At present, CRP is not used to directly guide clinical practice in cardiovascular medicine.

Stress-Related Psychosocial Cardiovascular Risk Factors and C-Reactive Protein

Because of the well-documented predictive value of CRP for adverse cardiovascular health outcomes, we will primarily focus on stress-related psychological factors that are known to be associated with atherosclerotic disease progression. Chronic and acute psychological factors are associated with increased risk of coronary artery disease and its clinical manifestations as acute coronary syndromes, including myocardial infarction. These psychological constructs are characterized by increased levels of perceived stress and/or exposure to potentially stressful environmental challenges. The pathophysiological mechanisms accounting for the relationship between psychological factors and coronary disease progression are likely to involve inflammatory processes. This section selectively reviews the interplay between psychological factors and CRP. A detailed review on the psychoneuroimmunological processes in cardiovascular disease has been presented by Kop and Cohen (2007: 921–943).

Prior investigations have shown that psychological risk factors for coronary disease can be classified into three broad categories, based on their duration and temporal proximity to coronary syndromes: (1) chronic factors, such as negative personality traits (e.g., hostility) and low socioeconomic status (SES); (2) episodic factors with a duration of several months up to 2 years, among which are depression and exhaustion; and (3) acute factors, including mental stress and outbursts of anger. Accumulating evidence indicates that the stage of coronary disease is a major determinant of the inflammatory mechanisms

involved in coronary disease progression and acute coronary syndromes.

Chronic Psychosocial Risk Factors and C-Reactive Protein

The chronic psychological factors discussed here are low SES and hostility, because of their well-documented relationship with coronary disease and the emerging evidence that these measures are also related to CRP. Both factors are associated with elevated levels of distress as a result of the purported increased exposure to challenging environmental circumstances and reduced resources to cope with such challenges.

SES can be defined based on a combination of education, income, job status, as well as other dimensions. Low SES is associated with elevated levels of CRP. For example, Owen et al. (2003: 286–295) documented elevated CRP levels among lower versus higher SES participants (1.18 ± 0.75 versus 0.75 ± 0.8 mg l⁻¹, $p = 0.002$) independent of sex, age, body mass, waist-to-hip ratio, smoking status, alcohol consumption, and season of the year. This finding has been replicated in larger epidemiological studies. The main potentially confounding factor in the association between SES and CRP is overweight. Nonetheless, most studies support a relationship between SES and CRP, even when statistically adjusting for body mass and other potentially confounding factors. However, because low SES is associated with a wide range of health behaviors, environmental exposures to pollutants and infectious agents, and psychological contextual factors, further studies are needed to clarify the role of SES in inflammatory factors such as CRP.

Hostility is a psychological trait characterized by a mixture of anger and disgust, and is associated with emotions such as resentment, indignation, and contempt. Hostility is predictive of long-term cardiovascular events, and this personality trait has been described as the toxic component of the type A behavior pattern. Research suggests modest (r values between 0.2 and 0.4) correlations between hostility and CRP levels. Measures of aggressive behavior are stronger predictors of CRP than general hostile attitudes. The mechanisms by which trait hostility, and particularly aggression, result in elevated inflammatory factors may involve stress-activated neurohormonal and autonomic nervous system pathways. Noradrenergic stimulation may increase gene expression for various inflammatory cytokines including IL-6, which may result in CRP release. Frequent responses to aversive interpersonal interactions may result in heightened noradrenergic drive in hostile individuals and thus result in increased production of pro-inflammatory cytokines and CRP.

Episodic Factors and C-Reactive Protein: Exhaustion and Depression

Prolonged emotional challenges may result in general distress, exhaustion, and depression. These stress-related conditions wax and wane with the intensity of the environmental challenges and the success of an individual's coping strategies. The predictive value of these conditions for cardiovascular events is strongest in the first 2 years after assessment. Because of the transient nature of these conditions and their stronger predictive value for adverse events within 2 years following assessment, these factors are referred to as episodic psychological risk factors.

Episodes of prolonged distress have been associated with markers of immunosuppression. The psychoneuroimmunology literature generally categorizes the immunosuppressive correlates of distress-related conditions along with the long-term immunological consequences of depressive disorders and other conditions characterized by negative affect (e.g., bereavement, separation, and daily hassles). Based on the Cardiovascular Health Study, our group has documented significantly elevated CRP levels among exhausted men (6.82 ± 2.10 mg l⁻¹) versus nonexhausted men (3.05 ± 0.16 mg l⁻¹; $p = 0.007$) aged over 65 years. The association for women was less strong, but still statistically significant when adjusting for co-variables including age, gender, race, diabetes mellitus, smoking status, and systolic blood pressure. Elevated CRP levels in depression, exhaustion, and other markers of prolonged distress have been consistently reported in the literature, with only a few exceptions (3/30 reviewed articles).

The mechanisms accounting for the associations between exhaustion and depression with CRP include biological pathways involving the central and autonomic nervous systems, and adverse health behaviors (e.g., overweight status and smoking). Prolonged distress commonly results in norepinephrine released from the locus coeruleus as well as hypothalamic corticotropin-releasing hormone (CRH), which are the main effectors of what Selye has described as the general adaptation response. The most common finding in depression, especially melancholia, involves activation of the CRH system and hence elevated cortisol levels. The clinical nature of depression (i.e., melancholic depression versus atypical depression and exhaustion) may influence the association between depressive symptoms and immune system parameters. CRH is elevated in typical depression, and acts as the main regulatory hormone in the acute-stress response, resulting in the release of pro-inflammatory cytokines, as well as a wide range of other immune system responses. These effects may be less strong or even reversed in atypical depression and exhaustion.

The autonomic nervous system also plays an important role in the acute-phase reaction, and parasympathetic outflow inhibits macrophage activation via the cholinergic anti-inflammatory pathway. Exhaustion and depression are associated with decreased parasympathetic nervous system activity, which may thus contribute to elevations in pro-inflammatory cytokines and elevated CRP.

The neuroimmunological responses to prolonged distress generally occur in the context of health behavior-mediated physical changes, particularly being overweight, affecting multiple metabolic processes. Thus, an overall imbalance of normal homeostatic function may be characteristic of episodic risk factors, resulting in prolonged CRP elevations.

Acute Psychological Factors and C-Reactive Protein

Acute psychological factors relevant to cardiovascular disease involve responses to acute mental or emotional challenges, such as outbursts of anger. These acute psychological factors can act as triggers of myocardial infarction by inducing cardiac ischemia and promoting plaque rupture and thrombus formation in advanced stages of coronary artery disease. There is increasing interest in the role of acute stress-induced inflammatory responses in relation to cardiovascular disease risk.

Some evidence suggests that acute mental challenge tasks induce elevations in CRP in individuals with inflammation-related conditions. However, in healthy individuals the results have been largely negative. One study examined CRP 45 min and 2 h postmental stress in healthy individuals, without significant CRP elevations. The optimal timing of CRP reactivity is an issue of current debate. Studies have demonstrated an immediate response in patients with cardiac problems and individuals with rheumatic arthritis (<5 min poststress).

A few methodological issues are relevant to the acute CRP response. Acute psychological distress elicits increases in arterial blood pressure and changes in blood rheology, including decreases in plasma volume and increases in blood viscosity. CRP is a relatively small molecule, but when plasma volume decreases in response to mental challenge, CRP does not readily pass into the interstitial space because of its molecular structure. Therefore, studies examining the acute CRP response need to adjust for changes in plasma volume. CRP levels at baseline are also highly correlated with levels obtained during mental stress (r values of >0.85), which results in statistical significance of very small absolute changes from baseline. Of note is that IL-6 responses tend to be unrelated to CRP responses, which suggests that mechanisms

other than hepatic release may play a role in the acute CRP response. These may include endothelial cells and other tissues, and the local release could possibly be mediated by circulating neurohormones. However, acute CRP responses may not be as important when compared with other acute responses of the immune system relevant to cardiovascular disease (e.g., IL-6, shift in lymphocyte distribution towards increased CD8+ and decreased CD4+). The timing of IL-6 should theoretically precede hepatic CRP release, and current data do not support such a pathway. Storage of biologically active signals can cause an immediate physiological effect, but the translation from messenger ribonucleic acid (mRNA) to protein takes several hours. Acute increases in circulating CRP are therefore not likely to reflect production of newly synthesized protein by tissues such as vascular endothelial cells or adherent leukocytes.

Conclusion and Future Directions

CRP is a critical component of the innate immune system, providing early defense against infectious agents and other pathogens. Substantial individual differences have been observed in CRP levels in the range below clinical inflammation (i.e., CRP levels <10 mg l^{-1}). Psychological factors are associated with elevated CRP levels, and distress-related inflammatory processes may therefore partially account for the relationship between psychosocial factors and adverse cardiovascular health outcomes. Evidence for elevated CRP levels is strongest for episodic risk factors such as depression and exhaustion. These conditions reflect the consequences of prolonged psychological distress, and elevated CRP levels in these conditions have a plausible biobehavioral explanation. Evidence for chronic psychosocial factors is also consistent, with the strongest support for low SES as predictor of elevated CRP. Acute CRP responses have been documented in patients with rheumatic arthritis and coronary disease, but the acute effects are small and probably do not reflect an increase in CRP secretion but rather a redistribution of CRP.

Further studies are needed to document the relative importance of catecholamines and hypothalamic pituitary adrenal axis hormones in elevated CRP levels. These studies need to be conducted in the context of important behavioral concomitants of inflammation, particularly assessments of adipose tissue because metabolic imbalance may be a common explanatory factor. In addition, prolonged stress may lead to increased allostatic load which may further promote the cascade of distress and elevated inflammatory markers.

Many of the psychological factors mentioned above are interrelated, which partially results from

the currently available measurement tools but may also reflect a general stress-related psychological factor. Such a general prolonged distress factor could account for increased levels of pro-inflammatory markers. More research is needed to identify potentially common environmental and/or genetic factors for vulnerability for stress-related psychological traits and high levels of inflammatory parameters.

Minor CRP elevations are associated with a diverse range of conditions. Current research suggests that the stimulus for CRP production involves a mild degree of tissue stress or injury, suggesting the presence of challenged cells, rather than a subsequent inflammatory response to these cellular challenges. Nonetheless, CRP has potentially pragmatic utility for cardiovascular-risk stratification, and further improvements in the assessment of inflammatory processes will contribute to a better understanding of the biopsychosocial factors involved in cardiovascular and other inflammation-related disorders.

Acknowledgments

Supported in part by a grant from the NIH (HL066149).

See Also the Following Articles

Cardiovascular System and Stress; Cytokines, Stress, and Depression; Heart Disease/Attack; Depression, Immunological Aspects; Inflammation.

Further Reading

- Casas, J. P., Shah, T., Cooper, J., et al. (2006). Insight into the nature of the CRP-coronary event association using Mendelian randomization. *International Journal of Epidemiology* 35, 922–931.
- Cesari, M., Penninx, B. W., Newman, A. B., et al. (2003). Inflammatory markers and onset of cardiovascular events: results from the Health ABC study. *Circulation* 108, 2317–2322.
- Danesh, J., Wheeler, J. G., Hirschfield, G. M., et al. (2004). C-reactive protein and other circulating markers of

- inflammation in the prediction of coronary heart disease. *New England Journal of Medicine* 350, 1387–1397.
- Kop, W. J. and Cohen, N. (2007). Psychoneuroimmunological pathways involved in acute coronary syndromes. In: Ader, R. (ed.) *Psychoneuroimmunology*, pp. 921–943. Amsterdam, Boston: Academic Press.
- Kop, W. J., Gottdiener, J. S., Tangen, C. M., et al. (2002). Inflammation and coagulation factors in persons >65 years of age with symptoms of depression but without evidence of myocardial ischemia. *American Journal of Cardiology* 89, 419–424.
- Kushner, I., Rzewnicki, D. and Samols, D. (2006). What does minor elevation of C-reactive protein signify? *American Journal of Medicine* 119, e17–28.
- Miller, G. E., Freedland, K. E., Carney, R. M., et al. (2003). Pathways linking depression, adiposity, and inflammatory markers in healthy young adults. *Brain Behavior and Immunity* 17, 276–285.
- Owen, N., Poulton, T., Hay, F. C., et al. (2003). Socio-economic status, C-reactive protein, immune factors, and responses to acute mental stress. *Brain Behavior and Immunity* 17, 286–295.
- Ridker, P. M., Rifai, N., Rose, L., et al. (2002). Comparison of C-reactive protein and low-density lipoprotein cholesterol levels in the prediction of first cardiovascular events. *New England Journal of Medicine* 347, 1557–1565.
- Roberts, W. L. (2004). CDC/AHA workshop on markers of inflammation and cardiovascular disease: application to clinical and public health practice: laboratory tests available to assess inflammation—performance and standardization: a background paper. *Circulation* 110, e572–e576.
- Rozanski, A., Blumenthal, J. A. and Kaplan, J. (1999). Impact of psychological factors on the pathogenesis of cardiovascular disease and implications for therapy. *Circulation* 99, 2192–2217.
- Segerstrom, S. C. and Miller, G. E. (2004). Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry. *Psychological Bulletin* 130, 601–630.
- Steptoe, A., Willemsen, G., Owen, N., et al. (2001). Acute mental stress elicits delayed increases in circulating inflammatory cytokine levels. *Clinical Science* 101, 185–192.
- Suarez, E. C. (2004). C-reactive protein is associated with psychological risk factors of cardiovascular disease in apparently healthy adults. *Psychosomatic Medicine* 66, 684–691.

Crime Victims

I Robbins

St. George's Hospital, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 594–597, © 2000, Elsevier Inc.

Victimization

Psychological and Physical Health Consequences

Responses to Specific Crimes

Support and Treatment Services

Glossary

Acute stress disorder (ASD)

A disorder that may occur in the immediate aftermath of severe trauma. The diagnostic features include dissociation in the form of numbing, reduction in awareness, derealization, depersonalization, and dissociative amnesia. There is anxiety and increased arousal, which often accompanies reexperiencing of the trauma. As a consequence there is avoidance of reminders of the trauma and impairment in social or occupational functioning. The diagnosis of ASD requires three dissociative symptoms but only one symptom from the intrusion, avoidance, and arousal categories.

Posttraumatic stress disorder (PTSD)

Occurs following exposure to a traumatic event in which the person witnessed or was confronted with an event that involved actual or threatened death or serious injury to themselves or others and where their response involved intense fear, helplessness, or horror. It is characterized by persistent reexperiencing in the form of recurrent, intrusive thoughts or images and/or distressing dreams or flashbacks. There may be intense distress or arousal when confronted with reminders of the event. As a consequence there may be persistent avoidance of things associated with the trauma and numbing of general responsiveness. There may also be sleep disturbance, irritability or outbursts of anger, difficulty in concentration, hypervigilance, and exaggerated startle response. The symptoms result in significant impairment of social, occupational, or other important areas of functioning and need to last longer than 1 month for a diagnosis to be made.

Victimization

It is difficult to estimate the extent of criminal victimization in part because most victims of crime do not report the event to the police, resulting in an underrepresentation within the criminal statistics compiled from police data. Most crime differentially targets and damages victims who are poor, marginalized, and disempowered within society; individuals are usually targeted because of what they represent rather than because of who they are. Examples of this can include racially motivated attacks or sexual assaults. Within the United Kingdom the recent British Crime Survey of 1998 pointed to an 83% increase in crime since 1981, with the largest increase being in violent crime. Women are more likely to be at risk for sexual or domestic violence, whereas men are more likely to report physical violence from strangers.

Being the target of a crime may result in the individual feeling that he or she is a victim. Ochberg reported that victims feel diminished, pushed down, exploited, and invaded. There is a feeling of stigmatization and of being isolated by the experience, with a shattering of basic assumptions about the predictable orderly nature of the world where bad things are perceived as only happening to people who deserve it. Individuals lose their sense of autonomy, and their belief in being able to control their own lives is shattered, with a consequent increase in feelings of vulnerability. This is often associated with a belief that others do not understand unless they have experienced being a victim themselves. This feeling may be reinforced by the critical response they experience from friends or family when their recovery is not sufficiently rapid.

There may also be a considerable amount of self-blame, which may take one of two forms: behavioral or characterological. Behavioral self-blame is concerned with aspects of behavior that, if changed, could reduce the possibility of reoccurrence, whereas characterological self-blame implies that the victimization is attributable to the sort of person that the victim is. Clearly, behavioral self-blame, since it implies the possibility of increased control over events, is more healthy than characterological blame.

Psychological and Physical Health Consequences

Victims of violence may experience a sense of detachment or depersonalization at the time of the attack.

While this may be a protective mechanism in the immediate aftermath of an attack, it may well hinder subsequent recovery, and recent evidence suggests that it may predict the subsequent development of posttraumatic stress disorder (PTSD). While dissociation is not a feature of PTSD, it is one of the major features of the diagnosis of acute stress disorder (ASD). Recent research has found that a diagnosis of ASD 1 month after experiencing a violent crime predicted the development of PTSD in 83% of cases at 6-month follow-up.

Many of the psychological consequences fit within the PTSD framework, with as many as 27% of all female crime victims meeting the criteria for diagnosis. Assaultive violence is more damaging than other types of crime, with higher rates of PTSD relating to increased perception of threat to life and extent of physical injury. In addition to PTSD, depression, anxiety, and substance abuse are common consequences of criminal victimization.

As well as the direct effects of the crime in terms of physical damage, victims are more likely to have poorer physical health and report increased drug and cigarette and alcohol consumption, health-care neglect, risky sexual behavior, and eating disorders.

Responses to Specific Crimes

Rape and Sexual Assault

Definitions as to what constitutes rape vary across countries. Within the United Kingdom the definition has recently been extended to include nonconsensual anal intercourse as well as vaginal intercourse. This change allows sexual attacks on men to be treated for the first time in law as rape, although it has to be acknowledged that the majority of victims of rape are women. Most sexual offences are unreported, and the rate of successful prosecution is low. Rape trauma syndrome was described by Burgess and Holstrom in 1974, but is now regarded as a variant of PTSD. Being the victim of a completed rape, being injured, and the extent of the perception of threat to life are predictive of increased rates of subsequent mental health problems in the longer term, as are prior victimization, previous psychological problems, and the lack of available social support. A third of women who report rape develop long-term psychological and social problems. A similar pattern was described in male victims of sexual assault by Mezey and King.

Stalking

Stalking has as yet been relatively poorly researched. Pathe and Mullen described severe social disruption and psychological distress with high levels of anxiety,

persistent intrusive recollections and flashbacks, and suicidal thoughts. There are profound economic and social consequences of stalking, as victims often feel compelled to leave their employment or change their address. Over a third of victims in the Pathe and Mullen study met the criteria for PTSD.

Domestic Violence

Domestic violence, like sexual violence, is primarily but not exclusively directed against women. It is defined as an act carried out with intent to physically injure another person, usually an intimate partner. Recent surveys in the United Kingdom have suggested a lifetime prevalence of domestic violence of 1 in 4 women and an annual prevalence of 1 in 9 women. Battered woman syndrome, first described in the 1970s, contains within it many of the features of PTSD and describes the emotional, cognitive, and behavioral consequences. It is associated with apparent learned helplessness, whereby the victims of domestic violence appear to be unable to extricate themselves from abusive relationships. Rather than being helplessness, this unwillingness to leave an abusive partner may in fact be a rational appraisal of the danger involved in freeing themselves from an abusive partner. The degree of risk is seen in terms of the possibility that the domestic violence may end up causing severe physical injury and can progress to homicide. The consequences of domestic violence include depression, anxiety, suicidal behavior, substance abuse, and somatization. It is a frequent cause of divorce and homelessness and may be associated with child abuse.

Workplace Violence

Workplace assaults have increased in frequency and severity in recent years and are associated with increased job stress, reduced job satisfaction, and the likelihood of carrying weapons to work. Males are most likely to be involved in fatal workplace assaults, while women are more likely to be involved in nonfatal assaults. Health-care workers are particularly affected, with the rates for health- and social-care workers being 10 times those in non-health-care industries. Around a quarter of nurses and doctors report physical assaults in the course of their work.

Murder

Murder, unlike death by natural causes, disproportionately affects the young, leaving relatives, particularly parents, feeling as if their future has been taken from them. Relatives of murder victims often feel stigmatized and isolated, with a sense of shame and betrayal, which results in their being unable to communicate their distress or make emotional

contact with others. The impact of traumatic bereavement may include physical symptoms, cognitive impairment, depression, and phobic avoidance as well as impaired work and social functioning. There is often a feeling of being let down by the criminal justice system, which compounds the sense of loss.

Robbery and Burglary

The effects of burglary may include PTSD but are generally less severe and long lasting. They may, however, include depression and anxiety and a sense of violation, which may be the most distressing and difficult aspect to resolve. Robbery, unlike burglary, involves direct contact with the perpetrator as well as a degree of threat to life and is therefore more likely to result in PTSD. Predictors of good recovery include a lower perception of threat to life, having had a preexisting view of the world as meaningful and orderly, and a rapid reduction of symptoms within the first month, whereas a depressive and avoidant coping style, fear of future violence, and increased somatic symptoms over time are associated with a poorer outcome.

Mass Shootings and Terrorist Crimes

While terrorist crimes are relatively infrequent, their social impact is much more widespread, inducing a climate of fear and uncertainty. For individuals caught up in terrorist attacks, the degree of threat to life and the extent of physical injury sustained during the attack are the best predictors of psychological problems, particularly PTSD, both in the immediate and the longer term. In the context of Northern Ireland a number of studies have found significant numbers of those with direct experience of terrorist incidents to be suffering from PTSD. This is in contrast to research on the impact of terrorist violence on the general population, which has been unable to detect a relationship between terrorist violence and resultant mental health problems in the population at large.

Studies that have looked at the impact of shootings tend by their very nature to be small scale but have found significant levels of distress and high rates of PTSD and other psychiatric disorders, with 33% or more being diagnosed as suffering from ASD in the immediate aftermath and this diagnosis being predictive of PTSD symptoms at follow-up several months later. Similarly, being held hostage has been related to high levels of subsequent distress both in victims and in their families. Victims of hostage taking may experience strong attachment and paradoxical gratitude toward the captors, with positive emotions including compassion and romantic love occurring. This may

be expressed as profound gratitude for being allowed to live and has become referred to as Stockholm syndrome.

Support and Treatment Services

There are few culturally accepted rituals used to support victims of crime. This means that a major strategy in treatment and support services involves normalizing the process. In the first instance this may be best achieved by victim support schemes, which offer practical assistance such as accompanying people to court hearings, guiding them through the process of applying for compensation, and dealing with the complexities of the criminal justice system. They also offer emotional support and the opportunity to ventilate emotions following a crime. Referrals to schemes are often automatically made by the police but may also be made from mental health professionals or may be requested by victims themselves. Despite the high prevalence of PTSD in victims of crime who participate in the criminal justice system, there is still ample evidence that they do not have adequate access to mental health services. This occurs despite awareness that victims of serious crimes and the families of murder victims may develop psychiatric illnesses that require referral to mental health services for specialist treatment.

There has been an assumption that early intervention is more successful. There are few controlled studies of the impact of early intervention. Those that do exist seem to be evenly divided between three categories, i.e., bringing about improvement, having no impact, and resulting in deterioration, so as yet it is not possible to assume that early intervention is either effective or at least harmless. Research on victims of serious crimes such as rape suggests that in the immediate aftermath the majority of victims would meet the diagnostic criteria for PTSD but that by 4–5 months this decreases substantially to less than half. If recovery does not occur during this time then subsequent improvements may be slow, resulting in chronic problems. The fact that the majority of victims may recover spontaneously should not diminish the need for provision of adequate services for the significant minority who do not improve.

The effects of criminal victimization may be severe and incapacitating and may have long-term economic and social consequences. Victims of crime are rarely vocal on their own behalf, and as a consequence their needs may be unrecognized, both by the population at large and by health professionals in particular.

See Also the Following Articles

Sexual Assault; Terrorism; Violence.

Further Reading

- Brewin, C. J., Andrews, B., Rose, S. and Kirk, M. (1999). Acute stress disorder and posttraumatic stress disorder in victims of violent crime. *American Journal of Psychiatry* 156, 360–366.
- Burgess, A. W. and Holmstrom, L. L. (1974). Rape trauma syndrome. *American Journal of Psychiatry* 131, 981–986.
- Davis, R. C., Taylor, B. and Lurigio, A. J. (1996). Adjusting to criminal victimisation: the correlates of post crime distress. *Violence & Victims* 11(1), 21–38.
- Eisele, G. R., Watkins, J. P. and Mathews, K. O. (1998). Workplace violence at government sites. *American Journal of Industrial Medicine* 33(5), 485–492.
- Figley, C. R. (1985). *Trauma and its wake* (vol. 1). New York: Brunner Mazel.
- Hilberman, E. (1976). *The rape victim*. New York: Basic Books.
- Kilpatrick, D. G., Saunders, B. E., Amick-McMullan, A., Best, C. L., Veronen, L. J. and Resnick, H. (1989). Victim and crime factors associated with the development of crime-related post traumatic stress disorder. *Behavioral Therapy* 20, 199–214.
- Kilpatrick, D. G., Saunders, B. E., Veronen, L. J., Best, C. L. and Von, J. M. (1987). Criminal victimisation: lifetime prevalence, reporting to police and psychological impact. *Crime and Delinquency* 33(4), 479–489.
- Lees, S. (1996). *Carnal knowledge: rape on trial*. London: Hamish Hamilton.
- Maguire, M. (1982). *Burglary in a dwelling*. London: Heinemann.
- Mezey, G. C. and King, M. B. (1989). The effects of sexual assault on men: a survey of 22 victims. *Psychological Medicine* 19, 205–209.
- Mirlees-Black, C., Budd, T., Partridge, S. and Mayhew, P. (1998). *The 1998 British Crime Survey*. England and Wales: HMSO.
- Ochberg, F. M. (1988). *Post traumatic therapy and victims of violence*. New York: Brunner Mazel.
- Pathe, M. and Mullen, P. (1997). The impact of stalkers on their victims. *British Journal of Psychiatry* 170, 12–17.
- Resnick, H. S., Kilpatrick, D. G., Dansky, B. S., Saunders, B. E. and Best, C. L. (1993). Prevalence of civilian trauma and post traumatic stress in a representative national sample of women. *Journal of Consulting and Clinical Psychology* 61(6), 984–991.
- Van der Kolk, B. A., McFarlane, A. C. and Veisaeth, L. (eds.) (1996). *Traumatic stress: the effects of overwhelming experience of mind, body and society*. London: The Guildford Press.
- Walker, L. E. (1979). *The battered woman*. New York: Harper and Row.

Crisis Intervention

D Hamaoka, D Benedek, T Grieger and R J Ursano
Uniformed Services University of the Health Sciences,
Bethesda, MD, USA

© 2007 Elsevier Inc. All rights reserved.

Impact of Crises
Planning for Crisis
Interventions During and After the Crisis
Critical Incident Needs Assessment Teams
Conclusion

Glossary

Crisis The critical turning point of a situation or event; can be an individual crisis or a population-level/community crisis.

Psychological first aid A group of evidence-informed interventions that can be helpful in the immediate aftermath of a crisis or traumatic event.

Resiliency The dynamic process of healthy response and coping in the face of adversity.

Risk communication Scientifically based method for communicating effectively under high-stress conditions.

Terrorism Intentional acts of human malevolence with the primary goal of causing terror; implemented by those who wish to coerce societies by inducing fear, shock, horror, and revulsion, often with ideological, religious, and political agendas.

A crisis is defined as a critical turning point of a situation or event. Crises can occur at the individual or community levels and affect one person or an entire population. Typically crises arise for after severely stressful events. For example, stressful events include traumatic events (i.e., car accidents) affecting individuals and large-scale disasters (i.e., hurricane, earthquake, or terrorism) that have profound effects on communities and nations. Crises are of variable

duration; they can be short-lived or persist for months and years. Over time, crises attenuate, resolve, or worsen, depending on the nature of the inciting event and its management. During severe crises, affected individuals often experience a period of stress, uncertainty, and anxiety and are concerned about their own safety or the safety of others. Each crisis often is followed by a series of crises (e.g., secondary crises such as loss of job, illness, or dislocation) that further stress the individual, group, and community.

Impact of Crises

Crises evoke a variety of reactions. Resilience and recovery are the rule, and most individuals do not develop chronic problems. For some, however, there are adverse psychological and behavioral responses. In crises precipitated by traumatic events such as disaster or terrorism, many people may experience sleeping difficulties; feel worried, sad, and anxious; increase alcohol and tobacco use; and change their regular behavior (e.g., alter their usual means of travel). Challenges to their faith and spiritual beliefs may also occur (Figure 1).

Many acute negative behavioral and emotional responses remit over time and do not require formal treatment. This tendency toward recovery is often credited to resiliency, a dynamic process of health recovery and coping in the face of adversity. Optimism, intelligence, humor, creativity, and active coping are related to resilience and positive outcomes after crises. Through active coping, individuals accept the impact of traumatic events and implement attainable, concrete measures to improve things.

Although many people experience distress after a crisis, some experience more persistent psychological

sequelae, such as anxiety, insomnia, increased smoking, increased alcohol consumption, and bereavement. This group may benefit from supportive psychological interventions, including psychological first aid and brief pharmacological interventions for sleep or anxiety. A still smaller group will develop psychiatric illness, including anxiety disorders (acute stress disorder, ASD; and posttraumatic stress disorder, PTSD), major depressive disorder, and substance use disorders. Such individuals require more formal (and perhaps more prolonged) interventions, including psychiatric treatment.

Individual responses to a crisis, which can include a traumatic event, a severe life stressor, or disaster, depend on a number of factors. Responses are influenced by proximity and involvement with the precipitating event. For a traumatic event, this means the severity of the trauma (e.g., degree of life threat). Responses are affected by psychological factors as well as interpersonal, family, and community stressors. In addition, research suggests social contexts, past experiences, future expectations, and genetic makeup interact with the characteristics of traumatic event to produce an individual's psychological response. Other identified risk factors that appear to increase the severity and/or duration of negative response include gender (e.g., women are more likely than men to develop acute PTSD), low level of social support, previous psychiatric illness, previous history of trauma, and ongoing negative life events after the trauma.

Groups and communities are also affected by crises, including the closing of a major community employer, the death of a beloved leader, and disasters. A community's response to a disaster often runs a predictable, but at times paradoxical, course. Communities temporarily coalesce immediately after a traumatic event. This is also known as the honeymoon period. During this time, individual heroics, a sense of working for a common cause, altruism, and "we will survive" attitude pervade the community. As time evolves, however, this optimism can change to disillusionment and often exposes the social fault lines of conflict, suspicion, and differential resources that are present along racial, ethnic, economic, and religious divides.

Community (as well as individual) responses to disasters are more pronounced when the trauma is intentional or the disaster humanmade. A relevant, modern-day example is terrorism. Here, acts of human malevolence are implemented with one primary goal – to cause terror. Terrorist acts are implemented by those who wish to coerce societies by inducing fear, shock, horror, and revulsion, often with ideological, religious, and political agendas.

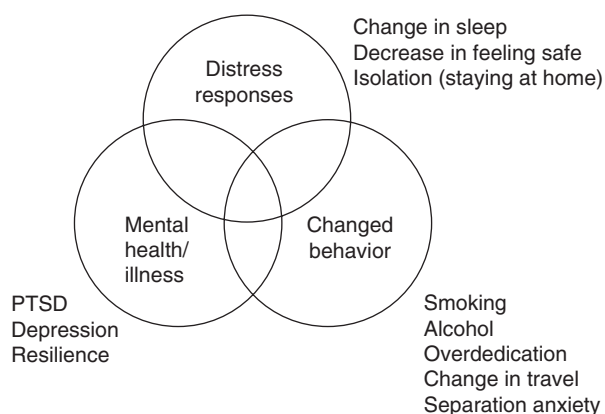


Figure 1 Psychological consequences of disasters, terrorism, and crises. PTSD, posttraumatic stress disorder. Adapted from Institute of Medicine (2003), *Preparing for the psychological consequences of terrorism: a public health strategy*, Washington, DC: National Academies Press.

Although acts of terrorism may lead to death, injury, property damage, and the evacuation and displacement of communities, the main aim is to challenge a society's sense of well-being, cohesion, and security. The severity and length of the crisis may increase if chemical, biological, nuclear, radioactive, or high-yield explosives (CBNRE) agents are used because these agents are particularly effective in causing terror. The infectiousness of biological agents, the persistence of chemical weapons, and the delayed effects of radioactive agents, in particular, perpetuate fear and induce terror.

A variety of interventions are used to attenuate the course of and speed the recovery from a crisis. Although all such interventions have been developed to be helpful for those affected, it should be noted that some interventions are more helpful than others, and there are some interventions that may be harmful. The effectiveness of any crisis intervention depends on a number of factors, including timing and availability. Some interventions, for example, must take place immediately after the onset of a crisis (such as ensuring safety and basic needs). Others interventions take place weeks to months later, as the focus, concerns, and emphasis shift over the evolution of the crisis. Beyond life-saving actions, subsequent interventions can help mitigate the risk of or the degree of problems for those who survive. The principles of crisis intervention for large-scale disasters are discussed in the following sections.

Planning for Crisis

Large-scale community crises, such as a disaster, illustrate the range of principles important to crisis intervention. Although the term crisis intervention may seem to refer to actions taken only after disasters occur, steps that can be taken prior to the occurrence of such events should also be part of public health planning for crisis intervention. For example, an inventory of available resources (including personnel, material, and monetary) can be developed before a crisis. Important government and community leaders and spokespeople can be identified, at-risk populations accounted for (Table 1), and gaps in the current support and response systems can be examined. Educating leaders in the principles of risk communication may provide them with skills to assist in calming people, dispelling rumors, and maintaining the leaders' credibility after a disaster. Similarly, pre-event education of the populace about personal preparedness is also an important part of crisis intervention. Preparedness plans include forming a communication plan with loved ones, establishing and mapping out

Table 1 Populations at risk for psychological problems after disaster

Previous exposure to trauma, particularly childhood
Direct exposure to the trauma/event (to include physically injured)
Those with premorbid psychiatric illness
Those experiencing acute losses
First responders (police, firefighters, emergency medical technicians)
Female gender
Those with minimal social support
Body handlers
Children
Elderly
Physically disabled
Those with negative life experiences after the trauma

evacuation routes, having extra medical supplies, developing care plans for pets, and having extra necessities on hand in the event of an evacuation (e.g., at least a half tank of gas in the car at all times, extra batteries, plenty of bottled water, and nonperishable food items).

Interventions During and After the Crisis

In the immediate aftermath of a large-scale crisis such as a disaster, the most important interventions focus on medical emergencies and life-threatening emergencies. Although it is not known (and it is difficult to predict) how the public will react to a large-scale disaster such as the Madrid or London bombings or the anthrax attacks in the United States, previous responses to disasters have shown generally effective and collective action. Depending on the cause of the traumatic event, there may be a large number of people seeking medical care. These numbers will very likely increase and overwhelm medical capacity if CBRNE agents are involved. Biological, chemical, and radioactive agents (due to their invisible, odorless nature) tend to induce a great number of people who are not at actual risk of exposure to believe they might have been exposed.

Crisis intervention, from a mental health standpoint, includes working with medical personnel to perform initial triage. Those who are not reassured easily may benefit from a brief stay in an area that has been set aside, preferably close to the emergency room. In this way, acutely distressed patients have the opportunity to reconstitute while still being monitored. In dealing with this population, terms such as the worried well and psychological casualties should be avoided. These terms are pejorative and convey a message that it's all in their head. Recognizing that these individuals are experiencing distress and require

caring responses can produce calm. The creation of a voluntary registry for individuals who are seen in the emergency room is not only a good public health intervention but can be therapeutic as well.

In addition to attending to the immediate medical needs, early crisis intervention addresses the basic needs of the survivors and includes safety from further harm as well as providing food, water, and shelter. As these needs are met, psychological first aid (PFA) can be employed. PFA is an evidenced-informed intervention that can be helpful in the immediate aftermath (hours to days) of an event (Table 2). The principles of PFA include establishing a sense of safety, facilitating social connectedness, fostering optimism, decreasing arousal, and restoring a sense of self-efficacy (e.g., the ability to take positive action). PFA can be thought of as flexible, supportive, and unlikely to cause harm; its main objectives are to limit distress, emphasize healthy behavior and activities, and minimize negative health behaviors. Although education and training in the PFA principles are required, the application of PFA can be accomplished by laypeople and does not require specific mental health expertise.

Brief simple conversations and informal on-site talks with survivors and responders can be of great assistance. This early crisis intervention avoids mental health labeling, offers support, education, and problem-solving techniques. Later interventions include cognitive-behavioral therapy (CBT). CBT has demonstrated efficacy in the prevention of PTSD in those with acute stress disorder after trauma exposure. One well-publicized disaster intervention known as psychological debriefing has not been shown to prevent PTSD and may be harmful in certain settings. Supportive and educational groups should be conducted by experienced and well-trained personnel; should be accompanied by clear objectives, evaluation, and referral procedures; and should never be mandatory.

Crisis intervention also includes good risk communication, a scientifically based method for

communicating effectively under high-stress conditions. The development of an effective risk communication strategy (as part of preevent intervention) is of vital importance in enabling leaders to inform and direct diverse populations. Individuals in the community look to their leaders for information, inspiration, a sense of control, optimism, and help during their period of grief. A major goal of the leaders should be to enlist the public as a partner. Information must be delivered frequently by credible and consistent sources. Messages should avoid speculation, never mix facts with reassurance, recommend specific steps people may take to protect themselves, and inform people when the next messages will be delivered. Good risk communication helps reduce negative psychological responses, encourage responsible safety behaviors, build trust, and minimize rumors and misinformation.

Critical Incident Needs Assessment Teams

Community crisis intervention that incorporates these principles in population health strategies must allocate and target resources at the individual, group, and community levels. Deploying critical incident needs assessment teams (CINATs) can be a helpful initial response to community disasters and crises. Such teams initiate planning, obtain on-site assessment, begin leadership consultation, and provide initial on-site guidance and support. CINATs thus are mental health-public health disaster response teams. CINATs are multidisciplinary and combine a public health approach and mental health knowledge to identify and respond to crisis. Teams initially quantify and identify needs in order to appropriately direct intervention and outreach resources.

CINATs recognize that interventions and responses are integrative, depend on a collaborative effort, and must use the community's inherent resiliency to help promote recovery. Teams can work effectively in responding to crises at workplaces (e.g., a shooting at a school or an airport after a crash). They target the individual and group levels that share a common task, mission, culture, structure, and/or physical proximity.

To be effective, CINATs require familiarity with the community or group they are deploying to assist. They may already be familiar with the given community or have received education and training about the community structure, culture, and leadership. These multidisciplinary groups include psychiatrists, psychologists, social workers, and mental health technicians. Additional individuals with particular areas of expertise, such as clergy and individuals responsible

Table 2 Principles of psychological first aid

Establish safety; identify safe areas and behaviors
Maximize individuals' ability to care for self and family and provide measures that allow individuals and families to be successful in their efforts
Teach calming skills and maintenance of natural body rhythms (e.g., nutrition, sleep, rest, exercise)
Maximize and facilitate connectedness to family and other social supports to the extent possible
Foster hope and optimism while not denying risk

Table 3 Grief leadership actions after disasters^a

Performs public announcements, appearances, and briefings
Presents calm demeanor
Organizes memorial services
Attends funerals, grieves
Endorses the various assistance programs
Attempts to describe loss in positive terms (acknowledging sacrifice and contributions)
Presents future goals and objectives

^aAdapted from Wright K. M. and Bartone P. T. (1994), Community responses to disaster: the Gander plane crash. In: Ursano, R. J., McCaughey, & Fullerton C. S. (eds.) *Individual and community responses to trauma and disaster*, pp. 267–284, Cambridge, UK: Cambridge University Press.

for security, communications, management, employee assistance, and human resources may be added.

In the aftermath of a community crisis, these pre-identified teams deploy to evaluate mental and behavioral health needs, identify high-risk groups, assist leadership function, and identify needed resources. CINATs avoid pathologizing appropriate responses to trauma and loss, and they identify those at greatest risk for subsequent problems. CINATs are a first population-level intervention. Subsequent care is arranged and distributed based on this epidemiological assessment. These teams also provide early support, education, PFA, and teaching about grief leadership (Table 3). CINATs may identify the need for additional supports including family support centers complete with legal assistance, casualty affairs assistance, Red Cross, and adult and child mental health counseling.

Conclusion

Crises affect both individuals and communities. Disasters are a severe form of crisis. Interventions can foster resiliency and mitigate adverse responses and health risk behaviors for individual and community crises. Interventions include PFA for individuals, CINAT for community assessment and early intervention, and traditional health care for those more severely affected. With appropriate planning, proper implementation, considerate timing, and coordinated execution, an effective and efficient response can foster resiliency, limit impairment, and speed recovery.

Crisis intervention is not one size fits all approach and requires thoughtful consideration and planning. Crisis interventions must always be acceptable to the survivors and their culture. Even the most well-thought-out plans cannot account for every possibility. Leaders and helpers require flexibility in their

approach and must meet disaster victims where they are – both literally and figuratively. Finally, those providing crisis intervention require support. They often do so with altruistic and noble intentions, and it is imperative that they take care of themselves as well as those they attempt to help.

Further Reading

- Barbera, J., Macintyre, A., Gostin, L., et al. (2001). Large-scale quarantine following biological terrorism in the United States: scientific examination, logistic and legal limits, and possible consequences. *Journal of the American Medical Association* 286(21), 2711–2717.
- Benedek, D. M., Ursano, R. J., Fullerton, C. S., et al. (in press). Responding to workplace terrorism: applying military models of behavioral health and public health response. *Journal of Workplace Behavioral Health*.
- Bisson, J. I., McFarlane, A. C. and Rose, A. (2000). Psychological debriefing. In: Foa, E. B., Keane, T. M. & Friedman, M. J. (eds.) *Effective treatments for PTSD: practice guidelines from the International Society of Traumatic Stress Studies*, pp. 317–319. New York: Guilford Press.
- Bonanno, G. A. (2004). Loss, trauma, and human resilience: have we underestimated the human capacity to thrive after extremely aversive events? *American Psychologist* 59(1), 20–28.
- Bryant, R. A., Harvey, A. G., Dang, S. T., et al. (1998). Treatment of acute stress disorder: a comparison of cognitive-behavioral therapy and supportive counseling. *Journal of Consulting and Clinical Psychology* 66(5), 862–866.
- Department of Veterans Affairs and Department of Defense (2004). Medical response to weapons of mass destruction. Public Law 107–287, Section 3. Washington DC: Government Printing Office.
- Galea, S., Ahern, J., Resnick, H., et al. (2002). Psychological sequelae of the September 11 terrorist attacks in New York City. *New England Journal of Medicine* 346(13), 982–987.
- Glass, T. A. and Schoch-Spana, M. (2002). Bioterrorism and the people: how to vaccinate a city against panic. *Clinical Infectious Diseases* 34(2), 217–223.
- Holloway, H. C., Norwood, A. E., Fullerton, et al. (1997). The threat of biological weapons: prophylaxis and mitigation of psychological and social consequences. *Journal of the American Medical Association* 278(5), 425–427.
- Institute of Medicine (2003). *Preparing for the psychological consequences of terrorism: a public health strategy*. Washington, DC: The National Academies Press.
- North, C. S., Nixon, S. J., Shariat, S., et al. (1999). Psychiatric disorders among survivors of the Oklahoma City bombing. *Journal of the American Medical Association* 282(8), 755–762.
- Schlenger, W. E., Caddell, J. M., Ebert, L., et al. (2002). Psychological reactions to terrorist attacks: findings from the National Study of Americans' Reactions to

- September 11. *Journal of the American Medical Association* 288(5), 581–588.
- Tusaie, K. and Dyer, J. (2004). Resilience: a historical review of the construct. *Holistic Nursing Practice* 18(1), 3–8, 9–10.
- Wright, K. M. and Bartone, P. T. (1994). Community responses to disaster: the Gander plane crash. In: Ursano, R. J., McCaughey & Fullerton, C. S. (eds.) *Individual and community responses to trauma and disaster*, pp. 267–284. Cambridge, UK: Cambridge University Press.
- Ursano, R. J., Kao, T. C. and Fullerton, C. S. (1992). Post-traumatic stress disorder and meaning: structuring human chaos. *Journal of Nervous and Mental Disease* 180(12), 756–759.

Critical Thermal Limits

J Roth

University of Giessen, Giessen, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J Roth, volume 1, pp 598–600, © 2000, Elsevier Inc.

Critical Thermal Limits of Ambient Temperature
Critical Thermal Limits of Body Temperature

Glossary

<i>Homeotherm</i>	The pattern of temperature regulation in a species in which the cyclic variation in core temperature is maintained within arbitrarily defined limits ($\pm 2^\circ\text{C}$) despite much larger variations in ambient temperature.
<i>Thermal balance</i>	Equality between heat gain and heat loss to maintain body temperature at a constant level.
<i>Thermoeffector</i>	An organ and its function that affects heat balance in a controlled manner as a part of the process of temperature regulation.

Critical thermal limits can be defined for homeotherms (1) as the highest and lowest ambient temperatures at which the rate of evaporative heat loss or the rate of metabolic heat production must be increased to maintain thermal balance, (2) as the highest and lowest ambient temperatures at which the capacity for temperature regulation is exceeded and core temperature changes (the development of hyper- or hypothermia), or (3) as the lethal highest and lowest body core temperatures.

Critical Thermal Limits of Ambient Temperature

Critical Thermal Limits for Activation of Thermoeffectors

Within a certain range of ambient temperature, heat production is at its minimum. This temperature range is called thermoneutral zone or metabolically indifferent zone. Within the thermoneutral zone, there are no regulatory changes in metabolic heat production or evaporative heat loss. The absolute highest or lowest limits, as well as the width of the thermoneutral zone, depend on body size, heat insulation, the heat transfer coefficient of the surrounding medium (e.g., water, wind), and the status of acclimation. For example, the metabolism of the arctic white fox does not start to rise until the ambient temperature decreases below -40°C , whereas a nude human starts to increase heat production at external air temperatures of 27°C and lower. The zones of thermal neutrality for a number of representative homeothermic species are shown in [Figure 1](#).

Critical Thermal Limits for the Thermoregulatory Capacity

The so-defined highest critical thermal limit depends on the capacity of the thermoeffector (i.e., the capacity to dissipate heat by evaporation of water) and on body size. At a low relative humidity of approximately 15%, the average external heat tolerance limits are 60°C for humans, 56°C for dogs and cats, 42°C for rabbits, 39°C for rats, and 37°C for mice. As long as a sufficient amount of water is supplied, these species can tolerate the listed ambient temperatures for longer periods. The limit of cold tolerance primarily depends on the effectiveness of the heat insulation that covers the body. Thus, it is not surprising that naked humans,

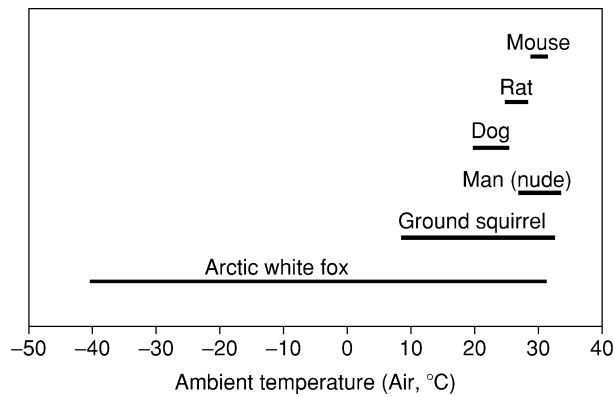


Figure 1 Thermoneutral zones (metabolically indifferent zones) of some homeothermic species. The left end of each bar indicates the ambient temperature at which the minimal resting metabolic heat production starts to increase. The right end of each bar indicates the ambient temperature at which mechanisms of heat dissipation (sweating, panting) are activated. Modified from Precht, H., Christophersen, J., Hensel, H. and Larcher, W. (1973). *Temperature and life*. Berlin: Springer-Verlag, Used with permission.

without an insulating hair coat, possess only moderate cold resistance. Unclothed humans can maintain a constant body core temperature for approximately 1 h at an ambient temperature of -1°C , whereas rats can withstand an external temperature of -25°C and the arctic white fox a temperature of -80°C for 1 h without any decrease in core temperature. In contrast to the quality of insulation, the extent of possible changes in the metabolic rate is a less important factor for the external limit of cold tolerance. In homeothermic animals whose insulation is poorer, the cold-induced increase of metabolic heat production is stronger than in well-insulated animals and starts at higher ambient temperatures.

Critical Thermal Limits of Body Temperature

This kind of thermal range refers only to the highest and lowest body core temperatures that a given species can survive. With regard to the lethal body temperature limits, the time factor (i.e., the duration of extreme body temperatures) is critically important.

Upper Limit of Body Temperature

The highest lethal body temperatures of homeotherms fall within quite a narrow range, which is only a few degrees above the normal body core temperature. The upper lethal body temperatures of some species are summarized in Table 1. It should be noted that a number of organs or even parts of the body can be warmed to temperatures higher than the general lethal body temperature. Due to this fact, experimentally

Table 1 Highest lethal body temperature in some representative homeothermic species^a

Species	Lethal temperature ($^{\circ}\text{C}$)	Remarks
Mouse	43.3	
Rat	42.5	50% lethal
Guinea pig	42.8	50% lethal
Rabbit	43.4	50% lethal
Cat	43.4	50% lethal
Dog	41.7	50% lethal
Human	45.0	Short-term survival
	43.0	Usually fatal

^aData compiled from Precht, H., Christophersen, J., Hensel, H. and Larcher, W. (1973). *Temperature and life*. Berlin: Springer-Verlag, Used with permission.

Table 2 Lowest lethal temperatures in some representative homeothermic species^a

Species	Lethal temperature ($^{\circ}\text{C}$)	Remarks
Rat	13	Cardiac arrest
Rat	-3	Supercooled, artificial hypothermia, revival
Guinea pig	17-21	
Cat	14-16	
Dog	15-18	
	0-2	Extracorporeal circulation revival
Human	24-26	Average accidental hypothermia
	18	Lowest accidental hypothermia, revival
	9	Artificial hypothermia, 45 min cardiac arrest, revival

^aSome cases of revival after artificial hypothermia and rewarming are included. Data compiled from Precht, H., Christophersen, J., Hensel, H. and Larcher, W. (1973). *Temperature and life*. Berlin: Springer-Verlag, Used with permission.

induced local hyperthermia is used in oncotherapy with the aim of destroying tumor cells.

Lower Limit of Body Temperature

The lowest lethal body temperatures for homeotherms are usually between 15 and 20°C . By the use of special techniques for cooling and rewarming, it is possible to lower the body temperature of a subject (human or animal) far below the usually lethal levels and to revive him or her unharmed. The lower lethal body temperatures of some species, including a few cases of revival under extreme conditions, are summarized in Table 2.

See Also the Following Articles

Heat Resistance; Hyperthermia; Hypothermia; Pressures, Effects of Extreme High and Low.

Further Reading

Cossins, A. R. and Bowler, K. (1987). *Temperature biology of animals*. London: Chapman and Hall.

Fregly, M. J. and Blatteis, C. M. (1996). *Handbook of physiology* (vol. 1). New York: Oxford University Press.

Precht, H., Christophersen, J., Hensel, H. and Larcher, W. (1973). *Temperature and life*. Berlin: Springer-Verlag.

Sharma, H. S. and Westman, J. (1998). *Progress in brain research: brain function in hot environments* (vol. 115). Amsterdam: Elsevier.

Crowding Stress

L Kovács

Babes-Bolyai University, Cluj-Napoca, Romania

P Csermely

Semmelweis University, Budapest, Hungary

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Csermely, volume 1, pp 601–603, © 2000, Elsevier Inc.

Small and big phenotypes

Small and big phenotypes are well separated in most organisms. Smalls are optimized for low resources (crowded conditions), while bigs are optimized for high resources. These phenotypes are epigenetically inheritable, and their conversion often requires three generations.

Introduction

Crowding Stress: Psychosocial Effects

Physiological Changes in Crowding Stress

Possible Molecular Mechanisms of Crowding Stress

Crowding of Flies and Worms

Cell Crowding

Molecular Crowding

Conclusions

Glossary

Amyloidosis

A severe pathological change of various organs and tissues during which aggregated amyloid fibers develop and induce the destruction of affected cells.

Channeling

Interaction of enzymes catalyzing consecutive enzyme reactions in which the product of the first reaction becomes the substrate of the second enzyme by a direct molecular transfer largely avoiding free diffusion.

Hypothalamic-pituitary-adrenocortical (HPA) axis

A major mechanism of the stress response, involving three major constituents: the corticotropin-releasing hormone (CRH), corticotropin (ACTH), and glucocorticoids.

Molecular crowding

A term to denote a dense population of molecules (usually macromolecules) where aggregation, diffusion, hydration, and other properties of the individual molecules are significantly altered.

Introduction

Studies on crowding stress consider an exceptionally high number of variables. Consequences of crowding stress may differ greatly, depending on whether population density is raised by an increased number of species living in the same area or by reducing their living space. If crowding is increased to such an extent that it leads to confinement, malnutrition, or an increased incidence of infections, other complications develop. Crowding stress may be acute (transient), i.e., the effects manifest after a few days, or chronic, i.e., changes occur after prolonged overcrowding lasting for weeks, months, or even years. Stress conditioning (or stress tolerance) can be observed in crowding stress: repeated stress exposure significantly diminishes the acute stress-induced effects occurring later. While mice or rats are the most commonly used species in crowding stress experiments, studies have been performed with almost all types of domesticated animals, various birds, fishes, and even humans. Though the conclusions of these studies can be directly compared only within the same species, some general trends can be observed. This article focuses on these general aspects of crowding stress.

Crowding Stress: Psychosocial Effects

Crowding as a chronic source of stress constitutes a major threat to psychological well-being. Crowding leads to anxiety and social instability. Dense populations are characterized by considerably increased

aggressive behavior. Crowded monkeys (even well fed), including females and young, have brutal fights, wounding and killing each other. Crowding stress adversely affects gonadal functions, and if it occurs during pregnancy it may inhibit reproductive activity of even the second generation through masculinization of female pups. Chronic crowding leads to deficits in learning tasks and has been used in animal models to induce depression. In human populations, crowding stress evokes prominent psychosocial reactions: it is proposed to be an important factor in the development of increased urban insanity/schizophrenia. Moreover, substance abuse (alcohol, amphetamine, morphine, etc.) and addictive behavior are prompted by a stressful social environment, e.g., crowding stress.

Physiological Changes in Crowding Stress

Recently detailed studies were performed on the effect of crowding on birds. With increased brood size, nestlings of zebra finch *Taeniopygia guttata* grow less, and have decreased testosterone levels and a lower T cell response. These birds are significantly lighter and have shorter wing and tarsus length in adulthood. Females allocate less testosterone in the yolk of their eggs in crowded conditions. This hormone has a positive effect on the growth and muscular development of the embryos. Consequently, newborn birds already start with a growth deficit.

Crowding-related growth deficits can be observed and explained in a wider context. The average height of U.S. men becomes smaller by 1.75 inches (4.5 cm) as the population density increases from 55 persons per square mile to 60 000 persons per square mile. Obviously this change is affected by a large number of variables including the availability of health services, car use/abuse, pollution rates, and stress levels. However, small and big phenotypes seem to be well separated in several species including humans. Smalls are optimized for survival, while bigs were preferentially developed for proliferation. Smalls will develop and succeed under low resources (crowded conditions), while bigs prevail under ample resources (non-crowded conditions). Smalls and bigs properties are coded at the epigenetic level and are not readily interchangeable. As many as three generations may be needed for a phenotype switch from small to big or vice versa. A question for further exciting studies is how acute and prolonged crowding stress affect the switch between these phenotypes.

Crowding stress (especially if chronic) suppresses immune functions. Disturbed immune regulation leads to increased autoantibody levels and may be one of the factors behind the increased occurrence of childhood asthma. Various infections and increased

susceptibility to poisoning are more likely to occur under crowded conditions. A widely established example indicates that household overcrowding is related to an increased prevalence of ulcer-inducing *Helicobacter pylori* infections. *H. pylori* infections and stress-induced gastric lesions significantly contribute to the development of ulcers and stomach cancer. Due to digestive problems and occasional appetite loss, chronic stress induces weight loss. In several organs, such as in kidneys and adrenals, chronic crowding stress induces intensive amyloidosis. Chronic overcrowding in many cases leads to hypertension in the resting state or to relative hypertension after exercise.

Possible Molecular Mechanisms of Crowding Stress

Crowding stress activates the hypothalamic-pituitary-adrenocortical axis (HPA axis) and enhances basal level or reactivity of plasma corticosterone secretion. This stress-related stimulation is triggered by the corticotropin-releasing hormone (CRH) system. HPA stimulation by other HPA-related biochemical factors, such as vasopressin, carbachol, and nicotine, is significantly diminished under crowded conditions. Moreover, crowding considerably impairs the HPA axis response to cholinergic and adrenergic stimulations. The mechanism of crowding-induced inhibition is best known in the case of nicotine: social stress affects signal transmission from membrane nicotinic receptors of different subtypes through ion channels into the cell. Crowding stress seems to induce an adaptive response to the non-CRH-induced HPA response to avoid the overstimulation of this important regulatory mechanism.

Repeated, short stresses induced by restraint or crowding attenuate the acute restraint stress-induced stimulatory action of the HPA axis. This indicates the occurrence of stress tolerance (stress conditioning) in the HPA axis response to acute stress. As a possible mechanism a short hypersecretion of corticosterone may induce a prolonged feedback inhibition of the HPA axis activity.

HPA stimulation is possibly the cause of decrease in appetite and consequent weight loss. HPA stimulation leads to compromised immune function and suppression of gonadal functions. The latter effect has an important role in regulation of population size, decreasing the chance of fertilization. Chronic HPA stimulation may lead to osteoporosis, chronic gastrointestinal pain, and retarded growth. Thus, prolonged activation of the HPA axis may explain many of the psychological effects of overcrowding, such as gastrointestinal problems, weight loss, sensitivity to infections, and decreased reproductive activity.

Additionally, crowding stress induces lipid peroxidation and impairs cellular signaling mechanisms, especially in elderly subjects. Impaired signaling may significantly contribute to immune suppression and decreased adaptive mechanisms.

Crowding of Flies and Worms

Signaling mechanisms can be studied more easily in simple organisms. Larval crowding in the fruit fly, *Drosophila melanogaster*, induces HSP (heat shock protein) expression and leads to increased adult longevity and adult thermal stress resistance. Flies that had been exposed to larval crowding exhibit greater starvation resistance and lipid content than populations that did not experience larval crowding. Crowding suffered in larval stage suspends the usual buffering of phenotypic variation: adults of crowded *D. melanogaster* larvae display an increased variability of thorax and wing length, as well as sternopleural and abdominal bristle number.

Food limitation and overcrowding also induce arrested development of the worm, *Caenorhabditis elegans*, leading to the formation of the so-called dauer larva. Daf-7, a homolog of the human transforming growth factor- β (TGF- β), prevents dauer larva commitment. Several other members of the dauer larva regulating Daf family are receptor serine-threonine kinases similar to the human TGF- β receptor. Mutations of another signaling pathway of *C. elegans* may quadruple the adult lifetime of the worm in addition to disturbing its dauer larva development. Thus, disturbances in signaling due to crowding stress may have profound consequences in the longevity of (simpler) organisms.

Cell Crowding

Experimenters often use cell cultures, where the cell density may be much smaller than under physiological conditions. During their proliferation, cells increase their density (in adherent cell lines the culture approaches confluency), and cell crowding may gradually develop. Cell crowding significantly alters the efficiency of autocrine and paracrine hormonal regulation and profoundly changes the influence of neighboring cells as well as the extracellular matrix on the individual cells. Cell crowding usually diminishes cell proliferation. However, tumor cells may escape from this control by expressing various molecules such as integrin $\alpha V\beta 6$ or lytic enzymes against components of the extracellular matrix. The extent of cell crowding should be always considered when interpreting the physiological relevance of experimental results with cell cultures.

Molecular Crowding

If the total volume of a macromolecular species occupies a significant fraction of the total volume of the solution, we refer to such a medium as crowded. Under experimental conditions molecular crowding is induced by polyethylene glycol or by dextrane. An intracellular environment, where the total amount of macromolecules usually occupies more than one-third of the total volume, is a typical example of molecular crowding.

Molecular crowding exerts profound quantitative effects on macromolecular interactions in living systems and induces an increased association of macromolecules. Crowding reduces diffusion rates. As a compensatory mechanism, it enhances channeling between enzymes catalyzing consecutive enzyme reactions as well as improves signaling efficiency in organized signaling cascades. By extending the range of intracellular conditions, where macromolecular interactions occur, crowding acts as a metabolic buffer.

Another potential outcome of molecular crowding is the effect on the properties of cellular water. The crowded environment in the cell results in a significant decrease of the proportion of cellular water being in contact with macromolecules such as proteins and DNA. Macromolecules begin to compete for water molecules, their hydration becomes compromised, and, consequently, osmotic stress occurs. The large amount of macromolecules and their immobilized hydrate shell constitute a large excluded volume. Thus, macromolecular crowding affects all those biochemical process in which a change of excluded volume occurs. Such a process is the collapse of newly synthesized polypeptide chain into compact functional proteins, the unfolding of proteins induced by stress, and the association of proteins into nonfunctional aggregates such as plaques in human amyloid diseases.

In conclusion, molecular crowding has a profound effect on the chemistry of life via its influence on association of macromolecules and on the cellular water properties.

Conclusions

Crowding may occur at various levels, from molecules through cells to organisms. Elements of crowded populations have an increased chance for extensive interactions, which increases community formation but may also lead to a high number of unspecific interactions, against which the elements have not previously developed an adaptive response. The unusual effects behave as perturbations of the elements and may lead to their destabilization – thus

crowding stress occurs. Crowding stress profoundly affects the behavior of the element at all levels studied, be it a molecule, a cell, or a simple or higher organism up to humans. Crowding often leads to the reduction of the element number either by diminishing the birth of new elements or by destabilizing, segregating, and/or destroying previously existing elements. This article gave a number of examples of these changes.

See Also the Following Articles

Hypothalamic-Pituitary-Adrenal; Prison; Psychosocial Factors and Stress; Social Stress, Animal Models of.

Further Reading

- Bar, J., Cohen-Noyman, E., Geiger, B. and Oren, M. (2004). Attenuation of the p53 response to DNA damage by high cell density. *Oncogene* 23, 2128–2137.
- Bateson, P., Barker, D., Clutton-Brock, T., et al. (2004). Developmental plasticity and human health. *Nature* 430, 419–421.
- Bugajski, J., Boricz, J., Gold, R. and Bugajski, A. J. (1995). Crowding stress impairs the pituitary-adrenocortical responsiveness to the vasopressin but not corticotropin-releasing hormone stimulation. *Brain Research* 618, 223–228.
- Csermely, P., Péntes, I. and Tóth, S. (1995). Chronic overcrowding decreases cytoplasmic free calcium levels in the T lymphocytes of aged CBA/Ca mice. *Experientia* 51, 976–979.
- Ellis, R. J. (2001). Macromolecular crowding: obvious but underappreciated. *Trends in Biochemical Sciences* 26, 597–604.
- Galpin, O. P., Whitaker, C. J. and Dubiel, A. J. (1992). *Helicobacter pylori* infection and overcrowding in childhood. *Lancet* 339, 619.
- Gil, D., Heim, C., Bulmer, E., et al. (2004). Negative effects of early developmental stress on yolk testosterone levels in a passerine bird. *Journal of Experimental Biology* 207, 2215–2220.
- Goeckner, D. J., Greenough, W. T. and Mead, W. R. (1973). Deficits in learning tasks following chronic overcrowding in rats. *Journal of Personal and Social Psychology* 28, 256–261.
- Haller, J., Baranyi, J., Bakos, N. and Halász, J. (2004). Social instability in female rats: effects on anxiety and buspirone efficacy. *Psychopharmacology* 174, 197–202.
- Imasheva, A. G. and Bubli, O. A. (2003). Quantitative variation of four morphological traits in *Drosophila melanogaster* under larval crowding. *Hereditas* 138, 193–199.
- Nagaraja, H. S. and Jeganathan, P. S. (2002). Voluntary alcohol drinking and caloric intake in rats exposed to crowding. *Indian Journal of Medical Research* 116, 111–116.
- Rohwer, J. M., Postma, P. W., Kholodenko, B. N. and Westerhoff, H. V. (1998). Implications of macromolecular crowding for signal transduction and metabolite channelling. *Proceedings of the National Academy of Sciences USA* 95, 10547–10552.
- Sørensen, J. G. and Loeschke, V. (2001). Larval crowding in *Drosophila melanogaster* induces Hsp70 expression, and leads to increased adult longevity and adult thermal stress resistance. *Journal of Insect Physiology* 47, 1301–1307.
- Xigeng, Z., Yonghui, L., Xiaojing, L., et al. (2004). Social crowding sensitizes high-responding rats to psychomotor-stimulant effects of morphine. *Pharmacology, Biochemistry, and Behavior* 79, 213–218.

Cultural Factors in Stress

J W Berry

Queen's University, Kingston, ONT, Canada

B Ataca

Bogazici University, Istanbul, Turkey

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J W Berry and B Ataca, volume 1, pp 604–610, © 2000, Elsevier Inc.

Culture as Adaptation

Cultural Stress

Acculturative Stress

Glossary

- Acculturation** A process of cultural and psychological change that results from contact between two cultural groups.
- Adaptation** A process of change that seeks to improve the fit between cultural groups and/or individuals and their habitat; it may result in outcomes that range from well-adapted to maladapted.
- Culture** A shared way of life of a group of people, including their symbolic, social, and material products; cultures are transmitted to new members over generations.

<i>Ecology</i>	The study of relationships between human organisms and their physical habitats, including their biological, cultural, and psychological adaptations to these habitats.
<i>Strategies</i>	Various ways employed by cultures and individuals to adapt to their habitats, including adjustment, reaction, and withdrawal (e.g., coping strategies and acculturation strategies).
<i>Stress</i>	The cultural and psychological consequences that occur when changes exceed the capacity of groups and individuals to adapt.

Culture as Adaptation

Among the many approaches to understanding culture is the view that human groups develop a way of dealing with recurrent problems in their ecosystems; these solutions are widely shared among members of a society and transmitted to their offspring. That is, cultures are adaptive to context, and, to the extent that these adaptations are successful, their fit is enhanced and stress is reduced: "Culture is man's most important instrument of adaptation" (Cohen, 1968: 1). The empirical foundations for this conception of culture were laid by Forde and Kroeber, who demonstrated that in Africa and North America, culture areas generally mapped onto ecological areas; in both continents, broadly shared features of culture were associated with ecological features in a particular group's habitat. This conception of culture has permitted work in both anthropology and cross-cultural psychology to align with stress, coping, and adaptation frameworks that are widely used in psychology.

When ecological conditions are relatively stable and there is sufficient carrying capacity (i.e., support from the habitat), cultural adaptations are also stable. However, when there is ecosystem disturbance due to internal or external forces (such as natural disaster or invasion) or when there is a chronic shortfall in the ability of the habitat to sustain the cultural group, then new adaptations are required if the group is to survive. In these terms, ecosystem changes constitute stressors, attempts to find innovative ways to manage day-to-day existence constitute coping, and the solutions achieved (successful or not) constitute adaptation.

Concern with these two sources of influence (internal and external) has given rise to an ecocultural model that attempts to understand cultural and psychological phenomena as the result of adaptation to ecological factors and to those that arise from contact with other cultures (the process of acculturation). Both during the process and when there is limited

adaptive success, two kinds of stress may result: cultural stress and acculturative stress.

For both kinds of stress, it is useful to consider some fundamental strategies of coping with these internal and external stressors. There are at least three ways to achieve adaptation: adjustment, reaction, and withdrawal. In the case of adjustment, behavioral changes are in a direction that reduces the conflict (that is, increases the congruence) between the environment and the behavior by changing the behavior to bring it into harmony with the environment. In general, this variety is the one most often intended by the term adaptation and may indeed be the most common form of adaptation. In the case of reaction, behavioral changes are in a direction that retaliates against the environment; these may lead to environmental changes that, in effect, increase the congruence between the two, but not by way of cultural or behavioral adjustment. In the case of withdrawal, behavior is in a direction that reduces the pressures from the environment; in a sense, it is a removal from the adaptive arena. These three varieties of adaptation are similar to the distinctions in the psychological literature made between moving with or toward, moving against, and moving away from a stimulus.

Cultural Stress

As noted previously, cultural stress occurs when extant or novel situations within the culture place demands on the group and its individual members that exceed their capacity to respond. Rather than maintaining or improving their fit, the responses of the group and individuals fail, and adaptation is unsuccessful. From the point of view of ecological anthropology, examples of successful adaptations are the most common; this is because those that are not successful fail to survive and are unavailable for contemporary observation. However, archaeology provides numerous examples of cultures that disappeared as a result of failure to adapt. Despite this imbalance in cases, there are examples of societies and individuals under cultural stress due to the extreme circumstances that they face at present. Two of these are the Ik of northern Uganda and the Mossi and Fulani of Burkina Faso.

The classic portrayal of the Ik was presented by Turnbull. This cultural group had lived for centuries as hunting and gathering nomads in a semidesert region at the intersection of Uganda, Sudan, and Kenya. Their main territory was a valley that had been taken over by the government to create a national park, from which they were excluded. With their old economic subsistence base taken away and agriculture impossible, they attempted to survive by

raiding the cattle herds of neighboring groups. As the carrying capacity of their territory was reduced (exacerbated by failure of the rains), the social fabric of the Ik deteriorated. According to Turnbull, they “abandoned useless appendages . . . those basic qualities such as family, cooperative sociality, belief, love, hope and so forth, for the very good reason that in their context, these militated against survival” (1972: 289).

This portrait is one of extreme ecological change, one that was rather quickly followed by social, cultural, and psychological change, all of them maladaptive. In this extreme case, the stressors were so severe that the cultural resources were incapable of providing any coping strategies other than intense striving for individual survival. Severe hunger generated “loss of any community of interest, familial or economic, social or spiritual” (Turnbull, 1972: 157) and eventually resulted in a maladaptive “everyone for himself” strategy that in the end failed.

A second example of cultural stress, one that also resulted from ecological change, is that of two groups in the Sahel region of West Africa. Ongoing environmental degradation (deterioration of the soil, loss of nutrients, reduction in wildlife) has led to loss of food resources and income and constituted a major set of stressors in the lives of two societies in this region. The Fulani (pastoralists) and the Mossi (agriculturalists) had differentially adapted to this semiarid ecosystem, one grazing their cattle over a large territory, the other enclosing small land areas for cultivation. In this study, coping with environmental change was directly assessed (by questionnaire), as was their locus of control and two psychological outcomes (feelings of marginality and of personal stress). With respect to coping, two factors emerged, the first representing a combination of problem solving and support seeking and the second mainly avoidance strategies; this split resembles the Lazarus and Folkman distinction between problem-focused and emotion-focused coping. For the locus of control measure, three factors were obtained: the first represented individual effort, the second nonpersonal control, and the third was uninterpretable. Both the marginality and stress scales produced single factors.

Structural analyses of the impact of ecological and cultural factors on the two psychological outcomes were carried out. The carrying capacity of the ecosystem was introduced as the latent variable, with four variables (environmental degradation, land use, cattle, and modernity) used as independent variables. The ecosystem provided a higher carrying capacity for the pastoralists than for the agriculturalists, and consequently the former were less marginalized and stressed. According to Van Haaften and Van de Vijver

(1996: 426), “this finding is in agreement with the common observation that nomadic people (the pastoralists) are less susceptible to environmental stressors than are sedentary people (the agriculturalists). Unlike the latter, the former can move away from environmental stressors.”

With these two examples, it is possible to identify links between ecological, cultural, and psychological changes in human populations. Evidence from these anthropological and psychological sources shows clearly that when cultural groups experience stressors, collective and individual coping sometimes leads to successful adaptation but sometimes does not, resulting in stress.

Acculturative Stress

When ecosystem and cultural changes are introduced from outside, the process of acculturation is initiated. In essence, acculturation refers to both the cultural and psychological changes that follow from contact between two or more cultural groups. At the cultural group level, these changes can occur in the physical, political, economic, or social domains (e.g., urbanization, loss of autonomy and livelihood, and the reorganization or even the destruction of social relationships). At the individual level, changes in the psychology of the individual take place.

It had been previously thought that acculturation inevitably brings social and psychological problems. However, such a negative and broad generalization no longer appears to be valid. Variability in psychological acculturation exists and is associated with three differing views about the degree of difficulty that is thought to exist during acculturation: behavioral shifts, acculturative stress, and psychopathology. The first is one in which changes in an individual's behavioral repertoire take place rather easily and are usually nonproblematic. This process encompasses three subprocesses: culture shedding, culture learning, and culture conflict. The first two involve the selective, accidental, or deliberate loss of behaviors and their replacement by behaviors that allow the individual a better fit with the larger society. Most often this process has been termed adjustment, since virtually all the adaptive changes take place in the acculturating individual, with few changes occurring among members of the larger society. These adjustments are typically made with minimal difficulty, in keeping with the appraisal of the acculturation experiences as nonproblematic. However, some degree of conflict may occur, in which incompatible behaviors create difficulties for the individual.

When greater levels of conflict are experienced, and the experiences are judged to be problematic but

controllable and surmountable, then acculturative stress is the appropriate conceptualization. Drawing on the broader stress and adaptation paradigms, this approach advocates the study of the process of how individuals deal with acculturative problems on first encountering them, and over time. In this sense, acculturative stress is a stress reaction in response to life events that are rooted in intercultural contact. Within this orientation, depression (due to cultural loss) and anxiety (due to uncertainty about how to live) were the problems most frequently found in a series of studies in Canada. More recently, general and acculturation-related sources of immigrants' stress have been differentiated in the literature. It was found that acculturation-specific hassles, above and beyond general hassles, had a negative effect on psychological distress. In-group hassles of Vietnamese immigrant students, i.e., stressors that result from interactions with Vietnamese peers and family, and out-group hassles of Iranian immigrants, i.e., stressors that result from interactions or lack of interactions (perceived or real) with out-group members in Canada, were associated with depression.

When major difficulties are experienced, the psychopathology perspective is most appropriate. According to this perspective, changes in the cultural context exceed the individual's capacity to cope because of the magnitude, speed, or some other aspect of the change, leading to serious psychological disturbances such as clinical depression and incapacitating anxiety.

Long-term adaptation to acculturation is variable, ranging from a well to poorly-adapted situation in which individuals can manage their new lives very well to one in which they are unable to carry on in the new society. Short-term changes during acculturation are sometimes negative and often disruptive in character. For most acculturating individuals, however, after a period of time, some long-term positive adaptation to the new cultural context usually takes place.

Variations in adaptation result in part from the acculturation strategy adopted by individuals. These strategies derive from individuals orienting themselves to two fundamental issues faced during acculturation: (1) To what extent do individuals seek to maintain their heritage culture and identity? (2) To what extent do they seek to have contact with and participate in the larger society? These two issues can be responded to along two attitudinal dimensions, with generally positive or negative (yes or no) responses at opposite ends. Orientations to these issues intersect to define four acculturation strategies. From the point of view of nondominant groups, when individuals do not wish to maintain their cultural identity and seek daily interaction with other

cultures, the assimilation strategy is defined. In contrast, when individuals place a value on holding on to their original culture and at the same time wish to avoid interaction with others, the separation alternative is defined. When there is an interest in maintaining one's original culture while in daily interactions with other groups, integration is the option; here, there is some degree of cultural integrity maintained while at the same time seeking, as a member of a cultural group, to participate as an integral part of the larger social network. Finally, when there is little possibility or interest in cultural maintenance (often for reasons of enforced cultural loss) and little interest in having relations with others (often for reasons of exclusion or discrimination), marginalization is defined. Acculturation strategies have been shown to have substantial relationships with positive adaptation: integration is usually the most successful, marginalization is the least, and assimilation and separation strategies are intermediate.

Adaptation is also multifaceted. It can be primarily internal or psychological (e.g., sense of well-being, self-esteem) or sociocultural, linking the individual to others in the new society (e.g., competence in the activities of daily intercultural living). Other forms of adaptation, including marital adaptation and economic adaptation, have also been proposed in the acculturation literature.

Each of these four domains of adaptation to acculturation may entail acculturative stress. In all four, the source of the stressors lies in the contact between groups and the resulting process of cultural change. It is important to note that acculturative stress is not a unique form of stress, but has unique antecedents (residing in the process of acculturation). Moreover, many of the outcomes are closely linked to the unique features of the acculturation process (such as experiencing cultural identity problems, acquiring culturally appropriate social skills, rearranging marital relations to take new cultural expectations into account, and experiencing difficulties due to employment discrimination in the new society). Thus, acculturative stress can be said to be a special kind of stress because of both its special set of stressors and its special set of outcomes. Following are examples of these four facets of acculturative stress.

The initial distinction between psychological and sociocultural adaptation was proposed and validated by Ward and colleagues. As noted earlier, psychological adaptation mostly involves one's psychological well-being and satisfaction in a new cultural context, whereas sociocultural adaptation refers to one's ability to acquire culturally appropriate knowledge and skills and to interact with the new culture and manage daily life. Conceptually, these two forms of

adaptation reflect two distinct theoretical approaches to acculturation. Psychological adaptation is better interpreted in terms of a stress and coping model; sociocultural adaptation is better viewed from a social learning perspective. Stress and coping models are based on the notion that both positive and negative life changes are intrinsically stressful in that they require adaptive reactions. Acculturation has been viewed as entailing such stress-inducing life changes, which increase susceptibility to physical and mental illness. The second approach draws on a combination of social skills and culture learning models. Individuals experiencing culture change are socially unskilled in the new cultural setting. Social learning models emphasize the importance of the acquisition of social skills and knowledge appropriate to the new culture.

While conceptually distinct, psychological and sociocultural adaptation are empirically related to some extent (correlations between the two measures are in the 0.4 to 0.5 range). However, they are also empirically distinct in the sense that they usually have different experiential antecedents. Research has shown that psychological adaptation, defined in terms of well-being or mood states (e.g., depression, anxiety, stress), is predicted by personality, life changes, and social support variables. Locus of control, life changes, and personal relationship satisfaction accounted for a substantial portion of variance in psychological well-being in student and adult sojourners. In contrast, assessed in terms of social difficulty, sociocultural adaptation was predicted by variables that are related more strongly to cognitive factors and social skills acquisition, such as cultural knowledge, cultural distance, cultural identity, language ability, length of residence in the new culture, and amount of contact with hosts. Extending this framework to immigrant couples, Ataca and Berry also found that these two dimensions of adaptation were associated with different variables. Psychological adaptation of married Turkish immigrant couples in Canada was associated with the personality variable of hardiness, social support, acculturation attitudes, and perceived discrimination, while sociocultural adaptation was mostly related to variables that are instrumental in acquiring social skills in the new culture, namely, English language proficiency and contact with Euro-Canadians.

Ataca and Berry also introduced the concept of marital adaptation, which relates to the accommodation of spouses in the process of acculturation when each is faced with the new culture and forms of behavior and different ways of acculturating. Their findings supported the distinctiveness of

marital adaptation from psychological and sociocultural adaptation. Better marital adaptation was mostly associated with variables specific to marital life, namely, fewer marital stressors and greater marital support; yet, marital variables also displayed close relations with psychological adaptation. This is in line with the literature on the relationship between marital variables and psychological distress among immigrants. Marital stressors were found to have an impact on both the marital distress and the depressive and psychosomatic symptoms of Indo-Canadian women. Naidoo found that South Asian women in Canada who had supportive husbands were less stressed. While the marital life of Turkish immigrants was related to psychological well-being, it was unrelated to sociocultural adaptation. The social skill learning necessary to function in the new cultural context was neither impeded nor facilitated by the conjugal relationship.

A fourth aspect of adaptation, that of economic adaptation, refers to the process by which individuals cope with and reestablish sustainable work relationships in the new economic system (including problems of status loss, under- or unemployment, and status mobility). Immigrants typically suffer high rates of under or unemployment. Lack of knowledge of language, lack of training, and discrimination force many immigrants into low-skill jobs. Conversely, high levels of education, occupational skills, and professional training received in the home country are often not recognized by authorities in the new country, resulting in underemployment. Under- or unemployment, low incomes, and the concomitant loss in one's socioeconomic status constitute sources of stress and are related to psychological symptoms. Aycan and Berry found that a greater decline in present status in Canada compared to the departure status from Turkey was associated with high acculturative stress among Turkish immigrants. Those who experienced greater status loss were also less satisfied and less likely to describe themselves as accomplished in economic life. The longer the immigrants were unemployed, the more likely they suffered from acculturative stress, negative self-concept, alienation from the society, and adaptation difficulties.

Studying the adaptation of both professional and nonprofessional Turkish immigrants in Toronto, Ataca and Berry found that those of low socioeconomic status were more stressed yet more satisfied with their lives than those of high socioeconomic status. This finding again points to the consequence of status loss that professional immigrants experience. Immigrants of lower status make comparisons with what their economic condition used to be like

and feel satisfied, while those of higher status make comparisons with their cohorts in Turkey and feel deprived. In this context, it was reported that low socioeconomic status and family income were also highly associated with high feelings of disturbance among South Asian women in Canada. Socioeconomic status is again an important predictor of stress in the studies with African American and Mexican American females and Korean American women.

Apart from personal factors, contextual factors are also important in the adaptation of immigrant and acculturating groups. Of utmost importance in the society of settlement are the attitudes of the dominant society. Perceptions and personal experiences of intolerance, hostility, and discrimination have been documented to serve as stressors and to have a negative impact on psychological adaptation.

In conclusion, stress does not take place in a vacuum: contextual factors, especially cultural ones that have been outlined here, clearly play a role. The importance of ecological and acculturative factors has been emphasized in order to provide a macro-level view of how stress and coping are shaped. Within these, micro-level experiences have been identified to illustrate the close links between culture and stress.

See Also the Following Articles

Cultural Transition; Economic Factors and Stress; Health and Socioeconomic Status; Indigenous Societies; Social Status and Stress.

Further Reading

- Aldwin, C. M. (1994). *Stress, coping, and development*. New York: Guilford.
- Ataca, B. and Berry (2002). Psychological, sociocultural, and marital adaptation of Turkish immigrants in Canada. *International Journal of Psychology* 37, 13–26.
- Aycan, Z. and Berry, J. W. (1996). Impact of employment related experiences on immigrants' psychological well-being and adaptation to Canada. *Canadian Journal of Behavioural Science* 28, 240–251.
- Berry, J. W. (1970). Marginality, stress and ethnic identification in an acculturated Aboriginal community. *Journal of Cross-Cultural Psychology* 1, 239–252.
- Berry, J. W. (1990). Psychology of acculturation. In: Berman, J. (ed.) *Cross-cultural perspectives*, (vol. 37), pp. 201–234. Lincoln, NE: University of Nebraska Press.
- Berry, J. W. (1997). Immigration, acculturation, and adaptation. *Applied Psychology: An International Review* 46, 5–68.
- Berry, J. W. (2001). A psychology of immigration. *Journal of Social Issues* 57, 615–631.
- Berry, J. W. (2003). Conceptual approaches to acculturation. In: Chun, K., Balls-Organista, P. & Marin, G. (eds.) *Acculturation: advances in theory, measurement, and applied research*, pp. 17–37. Washington, D.C: American Psychological Association.
- Berry, J. W. (2004). An ecocultural perspective on the development of competence. In: Sternberg, R. J. & Grigorenko, E. (eds.) *Culture and competence*, pp. 3–22. Washington, D.C: American Psychological Association.
- Berry, J. W. (2005). Acculturative stress. In: Wong, P. & Wong, L. C. J. (eds.) *Handbook of multicultural perspectives on stress and coping*, pp. 283–294. New York: Springer.
- Berry, J. W. and Kim, U. (1988). Acculturation and mental health. In: Dasen, P. R., Berry, J. W. & Sartorius, N. (eds.) *Health and cross cultural psychology: towards applications*, pp. 207–238. Newbury Park, CA: Sage.
- Berry, J. W. and Sam, D. L. (2006). Cultural and ethnic factors in health. In: Ayers, S., et al. (eds.) *Cambridge handbook of psychology, health and medicine* (2nd edn.). Cambridge, UK: Cambridge University Press.
- Berry, J. W., Kim, U., Minde, T. and Mok, D. (1987). Comparative studies of acculturative stress. *International Migration Review* 21, 491–511.
- Cohen, Y. (1968). Culture as adaptation. In: Cohen, Y. (ed.) *Man in adaptation*, pp. 40–60. Chicago, IL: Aldine.
- Dion, K. L., Dion, K. K. and Pak, A. (1992). Personality-based hardiness as a buffer for discrimination-related stress in members of Toronto's Chinese community. *Canadian Journal of Behavioral Science* 24, 517–536.
- Dona, G. and Berry, J. W. (1994). Acculturation attitudes and acculturative stress of Central American refugees in Canada. *International Journal of Psychology* 29, 57–70.
- Dyal, J. A., Rybensky, L. and Somers, M. (1988). Marital and acculturative strain among Indo-Canadian and Euro-Canadian women. In: Berry, J. W. & Annis, R. C. (eds.) *Ethnic psychology: research and practice with immigrants, refugees, native peoples, ethnic groups, and sojourners*, pp. 80–95. Lisse, Netherlands: Swets and Zeitlinger.
- Forde, D. (1934). *Habitat, economy and society*. London: Methuen.
- Kroeber, A. (1939). *Cultural and natural areas of North America*. Berkeley, CA: University of California Press.
- Lay, C. H. and Nguyen, T. (1998). The role of acculturation-related and acculturation non specific daily hassles: Vietnamese-Canadian students and psychological distress. *Canadian Journal of Behavioral Science* 30, 172–181.
- Lazarus, R. S. and Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer-Verlag.
- Liebkind, K. (1996). Acculturation and stress: Vietnamese refugees in Finland. *Journal of Cross Cultural Psychology* 27, 161–180.
- Liebkind, K. and Jasinskaja-Lahti, I. (2000). The influence of experiences of discrimination on psychological stress: a comparison of seven immigrant groups. *Journal of Community and Applied Social Psychology* 10, 1–16.
- Naidoo, J. C. (1985). A cultural perspective on the adjustment of South Asian women in Canada. In: Lagunes, I. R. & Poortinga, Y. H. (eds.) *From a different perspective:*

- studies of behavior across cultures*, pp. 76–92. Lisse, Netherlands: Swets and Zeitlinger.
- Noh, S., Beiser, M., Kaspar, V., Hou, F. and Rummens, J. (1999). Perceived racial discrimination, depression and coping. *Journal of Health and Social Behavior* 40, 193–207.
- Safdar, A. F. and Lay, C. H. (2003). The relations of immigrant-specific and immigrant nonspecific daily hassles to distress controlling for psychological adjustment and cultural competence. *Journal of Applied Social Psychology* 33, 299–320.
- Sam, D. L. and Berry, J. W. (eds.) (2006). *Cambridge handbook of acculturation psychology*. Cambridge, UK: Cambridge University Press.
- Turnbull, C. (1972). *The mountain people*. New York: Simon and Schuster.
- Van Haaften, E. H. and Van de Vijver, F. J. R. (1996). Psychological consequences of environmental degradation. *Journal of Health Psychology* 1, 411–429.
- Ward, C. (1996). Acculturation. In: Landis, D. & Bhagat, R. (eds.) *Handbook of intercultural training* (2nd edn., pp. 124–147). Thousand Oaks, CA: Sage.
- Ward, C. and Kennedy, A. (1993). Psychological and socio-cultural adjustment during cross cultural transitions: A comparison of secondary school students overseas and at home. *International Journal of Psychology* 28, 129–147.
- Ward, C., Bochner, S. and Furnham, A. (2001). *The psychology of culture shock* (2nd edn.). East Sussex, UK: Routledge.

Cultural Transition

M S Kopp

Semmelweis University, Budapest, Hungary

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M S Kopp, volume 1, pp 611–614, © 2000, Elsevier Inc.

Cultural Transition in the Central Eastern European (CEE)
Countries as Field Experience
Changing Attitudes and Values and Cultural Transition

Glossary

- Chronic stress** According to the original concept of the general adaptation theory of Janos (Hans) Selye, the three phases of stress are alarm reaction, resistance phase, and, the physiologically most harmful phase, exhaustion, that is, chronic stress. Chronic stress theory incorporates the learned helplessness model, the psychosocial and psychiatric models of depression, the control theory of stress and health, and the concept of vital exhaustion.
- Cognitive appraisal** The interpretation of an event or situation with respect to one's attitudes, values, and well-being.
- Coping** Constantly changing cognitive and behavioral efforts to manage specific external and/or internal demands that are

*Depressive
symptomatology*

Learned helplessness

Social capital

Social status

appraised as taxing or exceeding the resources of the person.

Symptoms of mood disturbances characterized by negative thoughts (for example, feelings of helplessness, inadequacy, loss of control, and low self-esteem), decreased motivation and interest in life, and such physical symptoms as sleep disturbances and fatigue. In civilized countries the prevalence of depressive symptomatology is around 20%.

A condition of loss of control created by being subjected to unavoidable trauma (such as shock and relative deprivation). Being unable to avoid or escape (flight or fight) an aversive situation produces a feeling of helplessness that generalizes to subsequent situations. Learned helplessness is the best animal model of human depressive symptomatology.

Refers to features of social organization, such as trust, norms, and networks, that can improve the efficiency of society by facilitating coordinated actions. Indicators are levels of trust, perceived reciprocity, and density of membership in civic associations.

An individual's worth relative to other group members. This evaluation of worth must be at least tacitly understood and agreed upon by interacting individuals.

Sociocultural identification

Introjection of socially meaningful behavioral patterns (usually social roles), through significant persons, via nonverbal communication and contextual and situational cues, acquiring thereby also motivation to actualize the given behavior patterns in fantasy and in action. Learning by model, model following behavior, model learning, and imitation are behavioral descriptions of the same phenomenon. Identification is a psychoanalytic term, meaningful from a developmental and intrapsychic point of view.

Cultural Transition in the Central Eastern European (CEE) Countries as Field Experience

Cultural Transition and Health in CEE Countries

While in modern societies living conditions have improved in many respects, other aspects of life have been fundamentally damaged. Important factors of personality development, for instance, the mother-child relationship, social models of the extended family, and the order in which values are passed on, are being challenged. The accelerated pace of life and unpredictable environmental changes add to these factors. Therefore, it is understandable that in all developed countries the number of adults and children suffering from the symptoms of anxiety and depression has increased. Compared to other countries, more dramatic changes have been experienced in the suddenly changing CEE and Eastern European countries. For example, in Hungary the prevalence of severe depressive symptomatology increased from 2.9 to 7.1% between 1988 and 1995.

A unique field experience of cultural and socioeconomic transition has taken place in the CEE countries during the past several decades. Until the end of the 1970s, the mortality rates in the CEE countries, including Hungary, were better than those of Britain and Austria. Subsequently, whereas in Western Europe the life expectancy rose continuously, in the CEE countries, such as in Hungary, Poland, the Czech and Slovak Republics, and Romania, this tendency reversed. By 1990 the Hungarian mortality rate was the highest in Europe. This deterioration cannot be ascribed to deficiencies in health care, because during these years there was a significant decrease in infant mortality rate and improvements in other dimensions of health care. Until 1990 it was not due to a declining standard of living: between 1970 and 1988 the GDP rose by more than 200% in Hungary, and in 1988 the economic status of the lowest socioeconomic strata was even better than in 1970.

The changes in Hungary and other CEE countries since 1970 have been due to a gradually intensifying social and economic polarization. Whereas a large part of the population lived at nearly the same low level in 1970, by the end of the 1980s a large fraction of society had achieved a higher socioeconomic level, having obtained one or more cars, their own property, and substantially higher income. Gaps within society had consequently widened.

Multivariate analysis showed that it was not the relatively worse socioeconomic situation itself that resulted in the higher morbidity rates, but the subjective appraisal of the situation and the consequent chronic stress – not a difficult social situation itself, but the subjective experience of relative disadvantage was the most significant health risk factor. When, for example, an individual does not have a car and therefore feels disadvantaged and that he or she cannot properly provide for his or her family, this is the state of mind that intensifies deterioration of both mental and physical health.

Self-Destructive Cycles between Cultural and Socioeconomic Transition: Chronic Stress, Depressive Symptomatology, and Health

Returning to the original concept of the general adaptation theory of János Selye, the three phases of stress are alarm reaction, resistance phase, and, the physiologically most harmful phase, exhaustion, that is, chronic stress. Chronic stress theory could incorporate the learned helplessness model, the psychosocial and psychiatric models of depression, the control theory of stress and health, and the concept of vital exhaustion. Furthermore, such a unified stress model could best explain the morbidity and mortality crisis in the middle-aged male population in Central and Eastern Europe in the last decades. A self-destructive circle develops from the chronic stress of the enduring, relatively disadvantageous socioeconomic situation and depressive symptoms. This circle resulting in depressive symptoms plays a significant role in the increase of morbidity and mortality rates in the lower socioeconomic groups of the population. It is not the bad socioeconomic situation itself, but rather the subjective evaluation and cognitive appraisal of the relative disadvantage that seems to be the most significant risk factor of health deterioration, since at equal living standards, Hungarian and other health statistics were better than those of other European countries until the 1970s.

Amid rapid socioeconomic and cultural changes, those left behind can continually blame themselves or their environment, see their future as hopeless, and experience permanent loss of control and helplessness. Negative judgment of one's own situation

and feelings of powerlessness and loss of control are the main background factors in the development of a chronic stress situation that might result in depressive symptomatology. This view of life at a time when society is rapidly polarizing, especially when the only goal of most of society is individual advancement, becomes very common.

Paralleling the increase in socioeconomic gaps within society there was an enormous decrease in perceived social support and social capital and increase in the sense of loss of control, not only in individuals, but in masses of people who felt rejected. The sudden transition of society continuously created situations in which the psychological and physiological balance could be maintained only with great difficulty; accordingly, there was a need for a change in attitudes and ways of coping. Only people with flexible coping resources adapted successfully.

Health Consequences of Chronic Stress and Depressive Symptomatology

There are many possible explanations for chronic stress and the consequent depressive symptomatology having an important role in higher morbidity and mortality rates during periods of cultural and socioeconomic transition.

The depressive condition affects perceived state of health and can lead to disability even without organic illness. Depression has a very close relationship with self-destructive behaviors, such as smoking and alcohol abuse, and suicidal behavior is especially common among depressive people. Those suffering from permanent mood disorder and depression are more vulnerable to various diseases and are less able to improve their social situation, so that they easily fall into a sustained vicious cycle. In recent decades, depression, vital exhaustion, and hopelessness have been identified as important independent risk factors for coronary disease. Learned helplessness or learned hopelessness, which can be regarded as the most appropriate model for depression, is associated with decreased immunological activity and affects tumor growth and vulnerability to various infections.

Mediating Role of Social Cohesion and Social Capital

In recent years, studies have suggested that cultural and social identity, sociocultural identification, social capital, and social cohesion are among the most important health protection factors in modern societies. Where they exist, wealthier people are prepared to make sacrifices for the community, and the disadvantaged do not feel that they have been completely left

to themselves in a hostile world. Trust in each other, esteem for shared values, the acceptance and internalization of cultural and social identity result in a high level of social cohesion in the society, which is the foundation not only of health but also of economic wealth and prosperity.

Marmot, Kawachi, and Wilkinson, who drew attention to the fundamental risk of inequality within society, also considered the phenomena of social cohesion and social capital to be the most important factors behind differences between countries and regions. For a lower-level British civil servant, for example, it was not his or her relative poverty itself that was the cause of his or her bad health, because during German bombing the death rates of the London population greatly decreased despite difficult living circumstances. The weakening of social cohesion and social capital is a significant factor in the emergence of large social differences and the associated deterioration of health – not only in CEE countries, but also in the developed countries.

Changing Attitudes and Values and Cultural Transition

Psychological Definition of Freedom and Democracy

An important question is whether the crisis experienced in human communities nowadays is a necessary part of civilization. The latest results of behavioral science point out that, beyond the concomitant phenomenon of lifestyle, we can see the root cause of this crisis in the fact that in modern society new possibilities for arousing anxiety have taken place.

During previous centuries the basic optimization principles of life were survival, subsistence, and maintaining the family, which dominated people's behavior. As a result of technological development, focus on these principles are no longer fundamentally necessary in civilized countries. However, the previously fixed order and communal forms of passing on values have also ceased. A young person's self used to be formed on the basis of feedback from the family, the extended family, and contemporaries in the community of the village or small town in which he or she lived. Nowadays, this process takes places within a nuclear or broken family and through mass communication, which seemingly provides feelings of community, but in reality the person rarely receives the continuity of patterns and values with which they can identify themselves.

While there are undeniable benefits in the move away from more tight-knit, rule-bound traditional

societies, under the conditions of continuous cultural and socioeconomic transition, the loss of balance in the human–environment system becomes more apparent. A solitary, anxious human being, deprived of relationships, values, goals, and self-esteem, can be used for the necessary functions of society and can be arbitrarily changed and manipulated. So we must recognize that technological developments have established the preliminary conditions for creating anxiety.

In order to maintain or increase their power, the person or group who possesses the most information can afford to either deprive others of information or give it as a reward, thereby creating anxiety in the environment. Arousing anxiety is not only a means, but also the essence of a dictatorship. The most effective tactic employed by a dictator is to create anxiety. A dictator wants to assure total decisional freedom for him- or herself. This is characteristic of all dependent relationships, in which the least dependent has the right to deprive the other person of information and thus keep the other in a situation over which he or she have no control. This not only occurs in totalitarian societies, but also can happen in families, schools, or jobs. Thus, we encounter the abuse of power and different ways of arousing anxiety every day, but we still do not give these situations enough attention, we do not recognize them in their embryonic forms, and we do not understand how this slow-action poison kills.

We can also arrive at the psychological definition of freedom and democracy through understanding the essence of anxiety. In a psychological sense, freedom means possessing the necessary, essential information needed to evaluate, solve, and control our situation, and being able to contribute actively to shaping our situation. Democracy means these rights – both the right and the responsibility of information and action.

Misuse of Anxiety in Cultural Transition

How does this apparently purely psychological and medical phenomenon, anxiety, become a basic concept that can shape society? We can ask whether anxiety has an adaptive function: is there any need for anxiety? Imagine a completely anxiety-free person: he or she does not recognize prospective danger and hence does not avoid dangerous situations and cannot be controlled by social sanctions. Some criminals and people with antisocial personality disorder belong to this type. In such cases, those in their environment suffer while they themselves follow their momentary drives.

Social co-existence requires a certain degree of anxiety. One learns as a child in which situations to

expect punishment – a mother's frowning look or a bad mark in school. The avoidance of anxiety is a very important driving force from the first moments of life. If the rules of punishment are foreseeable and are commensurate with the mistakes made, then anxiety can have an adaptive function. But if the rules are opaque or do not exist, then the powerlessness and defenselessness can become a weapon in the hands of those who have power. If parents punish from a whim rather than for a reason, then the child will not understand and will not be able to understand when he is threatened by danger. If a teacher or university lecturer asks questions from material that the student could not learn, he or she hammers insufficiency and failure into the student. If leaders, managers, politicians, those who possess more information do not make the rules clear, then they assure their arbitrary authority by causing anxiety in deprived masses. The twentieth century has created the possibility of causing such total anxiety, spread by the mass communication media just as Hitler's or Stalin's radio speeches were transmitted by the loudspeakers of a whole empire.

It is obvious from the preceding that, since creating anxiety can involve issues of money, benefit, and power, there are enormous advantages in knowing how to deal with these powers effectively. Films, newspapers, magazines, and books infiltrate the community and human relationships and can belittle or deny values and mold people into a mass that can be led. This type of mass manipulation can occur even if the manipulator's aim differs from what actually takes place.

The ideology of the consumer society, to control human behavior through the liberation of instincts and search for pleasure, degrades human beings to the level of animals since an animal's decisions are influenced by its instincts to preserve its physiological balance. To achieve dominance by any means is a natural need in such a system.

On every level of society there is, day by day, a life-and-death struggle between two behaviors: either accepting and respecting other people, establishing agreement and a community of free people and bringing the value of democracy into being or arousing anxiety by withholding information from others. Not the slogans but the behavior classifies the participants.

The question is whether humans will be able to adapt themselves to the circumstances formed by history, or whether their lack of adaptive ability will lead the human race into extreme peril. This adaptation has to be realized on the level of both the individual and society.

See Also the Following Articles

Anxiety; Cultural Factors in Stress; Social Capital; Social Networks and Social Isolation; Social Status and Stress; Social Support.

Further Reading

Kawachi, I. and Berkman, L. F. (2000). Social cohesion, social capital, and health. In: Berkman, L. F. & Kawachi, I. (eds.) *Social epidemiology*. New York: Oxford University.
Kopp, M. S. and Réthelyi, J. (2004). Where psychology meets physiology: chronic stress and premature

mortality-the Central-Eastern European health paradox. *Brain Research Bulletin* 62, 351–367.
Marmot, M. (2004). *The status syndrome: how social standing affects our health and longevity*. New York: Times Books, Henry Holt and Company.
Skrabski, Á., Kopp, M. S. and Kawachi, I. (2004). Social capital and collective efficacy in Hungary: cross sectional associations with middle aged female and male mortality rates. *Journal of Epidemiology and Community Health* 58, 340–345.
Wilkinson, R. G. (1994). The epidemiological transition: from material scarcity to social disadvantage? *Daedalus* 123(4), 61–77.

Cushing's Syndrome, Medical Aspects

S R Bornstein and M Gruber

Technical University of Dresden, Dresden, Germany

H S Willenberg

University Hospital Duesseldorf, Duesseldorf, Germany

C A Stratakis and G P Chrousos

National Institute of Child and Health and Human Development, National Institutes of Health, Bethesda, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S R Bornstein, C A Stratakis, and G P Chrousos, volume 1, pp 615–620, © 2000, Elsevier Inc.

Transsphenoidal surgery (TSS)

Surgical procedure of choice to remove pituitary tumors. This procedure utilizes the nasal approach, avoiding the necessity of opening the skull.

Epidemiology

The overall incidence of spontaneous Cushing's syndrome (CS) is approximately two to four new cases per million of population per year, with a female to male preponderance. Approximately 10% of these cases occur in children and adolescents. Adrenocorticotrophic hormone (ACTH)-dependent CS accounts for about 85% of endogenous cases in adults, adolescents, and children older than 7 years of age (**Table 1**). Cushing's disease is responsible for 80% of these cases; the remaining 20% is caused by ectopic ACTH or, very rarely, ectopic corticotropin-releasing hormone (CRH) secretion. Benign cortisol-secreting adenomas or adrenocortical carcinomas are rare and account for 10–15% of endogenous cases in children older than 5. In children younger than 5, adrenocortical carcinomas and primary pigmented nodular adrenal disease (PPNAD) are the more frequent causes of endogenous CS. Primary pigmented nodular adrenal disease is a form of adrenal tumors that occurs frequently in association with Carney complex, a multiple endocrine neoplasia and lentiginosis syndrome. With the increasing utilization of glucocorticoids for a wide range of nonendocrine diseases, the exogenous, iatrogenic CS has become more frequent and is more prevalent than the endogenous forms.

Epidemiology

Etiology

Pathophysiology

Clinical Picture

Laboratory Findings

Radiographic Findings

Treatment

Glossary

Adrenocorticotrophic hormone (ACTH)
Corticotropin-releasing hormone (CRH)
Cushing's syndrome

Pituitary hormone stimulating release of adrenal glucocorticoids.

The main regulator of the endocrine stress axis. Stimulates pituitary ACTH.

Caused by the state of hypercortisolemia.

Table 1 Classification of Endogenous Cushing's Syndrome and Rate of Occurrence in Children Older than 7 Years^a

<i>Classification</i>	<i>Percentage</i>
ACTH-dependent	85
Pituitary (Cushing's disease) (includes MEN-I)	80
Ectopic ACTH (includes MEN-I)	20
Ectopic CRH	Rare
ACTH-independent	15
Adrenal adenoma (includes MEN-I)	30
Adrenal carcinoma	70
Primary pigmented adrenocortical disease (PPNAD)	0.5–1
McCune-Albright's syndrome	Rare
"Transitional states"	Rare

^aModified from Magiakou and Chrousos (1994).

Etiology

Endogenous CS results from increased secretion of cortisol by the adrenal cortex and is due to ACTH hypersecretion (ACTH-dependent CS) or the autonomous hyperfunction of the adrenocortical cells (ACTH-independent CS). Some forms of ACTH-independent CS have recently been discovered to be due to aberrant expression of neuroendocrine or cytokine receptors (such as GIP, vasopressin, β -receptor, or interleukin-1), putting the adrenal tissue under a tropic stimulus, which is not under the control of the regular negative feedback. Cushing's disease, the ACTH-dependent form of endogenous CS, is due to an ACTH-secreting, benign corticotroph tumor of the anterior pituitary gland. Ectopic ACTH-secreting tumors occur mostly in patients with bronchial carcinoma. In addition, other tumors may present with ectopic ACTH production (thymus, liver, renal carcinoma, etc.). Exogenous CS results from chronic administration of glucocorticoids (for treatment of neoplastic, autoimmune, and other diseases) or ACTH (for the treatment of certain seizure disorders).

Pathophysiology

The main stimulus for the release of cortisol from the adrenal cortex is ACTH, which is produced by the corticotrophs of the anterior pituitary, which, in turn, are under the regulatory influence of hypothalamic CRH and vasopressin (Figure 1). The ambient plasma free cortisol levels regulate ACTH secretion in a negative feedback fashion. Normally, less than 10% of plasma cortisol is in the free form, with the majority being bound to cortisol-binding globulin (or transcortin).

In the physiological state, ACTH and cortisol have a circadian pattern of secretion, established early in infancy and in accordance with other human

circadian biological activities. The peak of ACTH and cortisol secretion occurs in the morning (between 7:00 and 8:00 a.m.) and is lowest in the late evening hours (around midnight). This physiological circadian rhythm is blunted in CS.

The adrenals have an amazing ability to adapt to states of hyper- or hypofunction. They use more than 10 times the cardiac output and blood volume that would correspond to their organ weight. Therefore, they are unique in demonstrating signs of hormonal excess and stress. Chronic stimulation of adrenal tissue with ACTH in Cushing's disease or ectopic ACTH-secreting tumors will cause hypertrophy of the adrenals, leading in particular to an increase in size of the zona fasciculata. The adrenal cortex undergoes hypervascularization and the cells increase in mitochondria, smooth endoplasmic reticulum, and filopodia, indicating increased activation of steroidosynthesis.

Glucocorticoids are vital for survival because they participate greatly in the stress response of humans. Glucocorticoids affect all tissues and organ systems, and the glucocorticoid receptor is ubiquitously present. Glucocorticoids have profound catabolic activities: they inhibit growth and reproduction, promote proteinolysis, and suppress the immune and inflammatory reaction. These effects are behind some of the cardinal clinical manifestations of CS: delayed growth and bone maturation, hypogonadism, dermatopathy, loss of muscular tissue, and frequent fungal or saprophytic infections.

Clinical Picture

Table 2 summarizes the clinical features of CS in children and adolescents. An early and common sign in almost all patients is obesity, which is generalized or truncal and is characterized by moon facies (facial rounding) and plethora. Growth retardation or complete arrest is present in 90% of the children. Other clinical manifestations include sleep disturbances, muscle weakness and fatigue, hirsutism, and typical purple skin striae. Hypertension, carbohydrate intolerance or diabetes, amenorrhea, advancement or arrest of pubertal development, and easy bruising or spontaneous fractures of ribs and vertebrae may also be encountered. Although all patients may exhibit some of these features at the time of diagnosis, few, if any, will have all of them. Photographs of the patient, taken over a period of years, are helpful in the clinical evaluation.

The clinical presentation can identify the cause of CS in children and adolescents. Thus, severe and rapidly progressing CS points toward an adrenal neoplasm or, although rare in children, ectopic ACTH syndrome. Patients with adrenocortical carcinomas

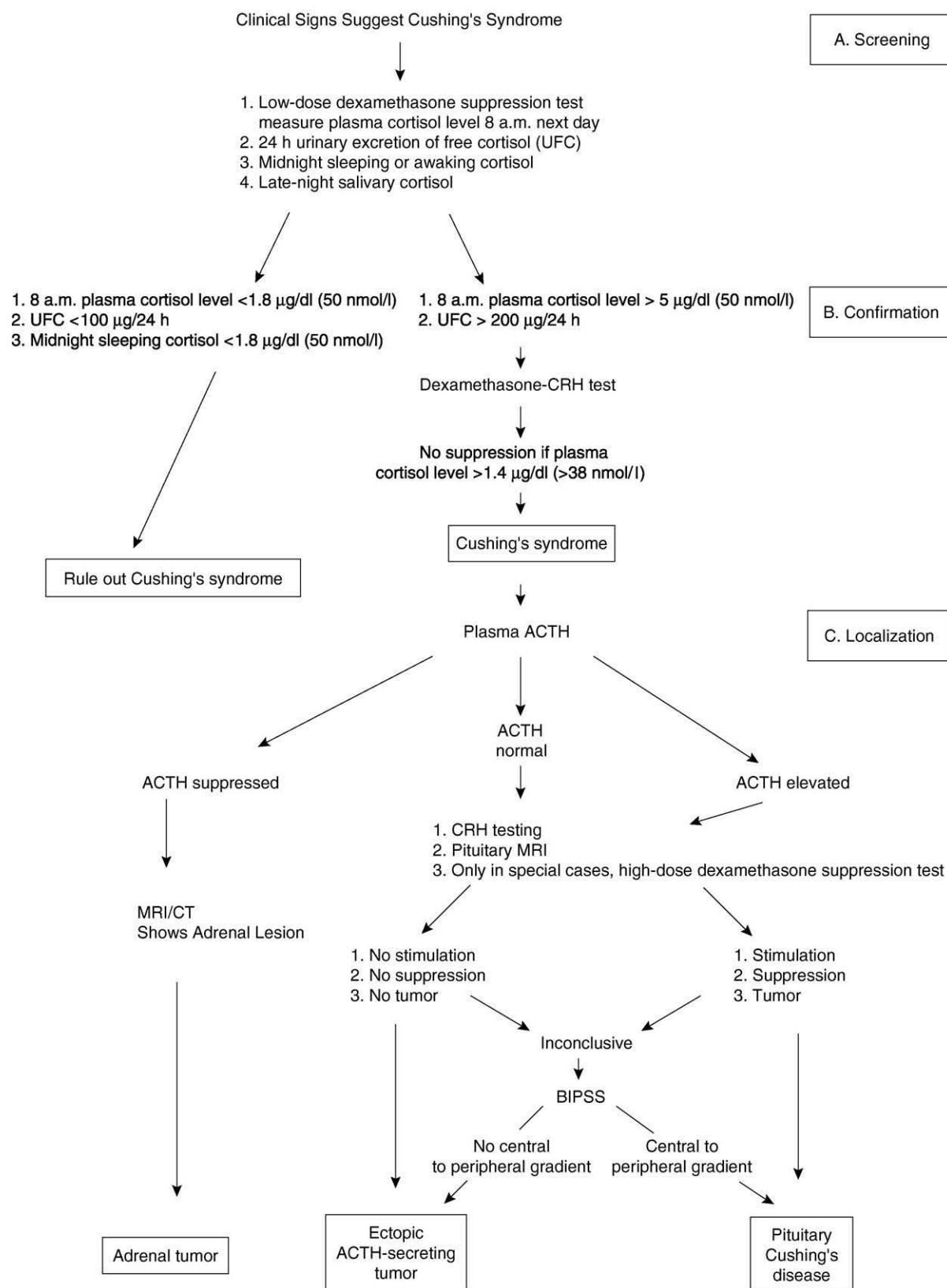


Figure 1 Diagnostic algorithm for Cushing's syndrome. Modified from Bornstein, S. R., Stratakis, C. A., and Chrousos, G. P. (1999). Adrenocortical tumors: recent advances in basic concepts and clinical management. *Annals of Internal Medicine* **130**, 711–759.

Table 2 Clinical presentation of CS in pediatric patients^a

Symptoms/signs	Frequency (percentage)
Weight gain	90
Growth retardation	83
Menstrual irregularities	81
Hirsutism	81
Obesity (Body Mass Index .85 percentile)	73
Violaceous skin striae	63
Acne	52
Hypertension	51
Fatigue-weakness	45
Precocious puberty	41
Bruising	27
Mental changes	18
"Delayed" bone age	14
Hyperpigmentation	13
Muscle weakness	13
Acanthosis nigricans	10
Accelerated bone age	10
Sleep disturbances	7
Pubertal delay	7
Hypercalcemia	6
Alkalosis	6
Hypokalemia	2
Slipped femoral capital epiphysis	2

^aModified from Magiakou *et al.* (1994).

may be asymptomatic or may have abdominal pain or fullness, symptoms and signs of CS (30%), virilization (20%), combined symptoms of CS and virilization (30%), feminization (10%), or hyperaldosteronism (5–10%). CS is periodic or intermittent in a significant percentage (~10%) of children and adolescents.

Laboratory Findings

The first step in diagnosing noniatrogenic CS is the biochemical documentation of endogenous hypercortisolism, which can usually be accomplished by outpatient tests. The first test is measurement of 24-h urinary free cortisol (UFC) excretion corrected for body surface area and/or the determination of 24-h urinary 17-hydroxycorticosteroid (17OHCS) excretion (corrected per gram of excreted creatinine). UFC excretion is a good first-line test for documentation of CS, with, assuming correct collection, very few false negative results. The second test is the single-dose dexamethasone suppression test. The overnight 1-mg (in young children, 15 µg/kg body weight) dexamethasone suppression test is a useful screening procedure for hypercortisolism. It is simple, with a low incidence of false normal suppression (less than 3%). The same test, however, has

a high incidence of false positive results (approximately 20–30%). A plasma cortisol level >1.8 µg/dl (50 nmol/l) suggests endogenous hypercortisolism and needs to be followed by a 24-h urine collection for determination of UFC or 17OHCS levels.

CS is generally excluded if the aforementioned tests are normal, although one should remember that periodic and intermittent cortisol hypersecretion, which may occur in patients with CS of any etiology, may confuse the picture.

In the past few years, a simple test reached importance in the first-line diagnosis of CS, especially for outpatient assessment. The measurement of late-night salivary cortisol is easy, and the cortisol in this test is stable under room conditions. Because of the variety of the diagnostic ranges in this test, the cut-off point of the local laboratory has to be known. The sensitivity and the specificity of this test are between 95 and 98%.

Another important and highly sensitive test is the measurement of the midnight sleeping or awaking cortisol levels, because of the circadian rhythm of cortisol secretion in healthy persons. A sleeping midnight cortisol level of less than 1.8 µg/dl (50 nmol/l) excludes CS. Once CS has been diagnosed, the differential diagnosis is investigated via dynamic testing of the function of the hypothalamic-pituitary-adrenal (HPA) axis as well as imaging studies. It is essential that dynamic endocrine testing is performed while the patient is hypercortisolemic, and, in order to avoid mistakes, this state should always be documented at the time of testing. All adrenal blocking agents should be discontinued for at least 6 weeks prior to testing.

The major tests in the differential diagnosis of CS and their interpretations include the following (see also [Figure 1](#)):

1. Low-dose dexamethasone suppression test and dexamethasone-CRH test: The low-dose dexamethasone suppression test (0.5 mg every 6 h over 2 days) and the dexamethasone-CRH test are reliable procedures for differentiating Cushing's disease from subclinical CS. In the combined dexamethasone-CRH test, after application of 0.5 mg dexamethasone every 6 h over 2 days (start on first day at 12.00 a.m.; the finish dose is given at 6.00 a.m.), a CRH-stimulating test will follow at 8.00 a.m. If the plasma cortisol is lower than 1.4 µg/dl (38 nmol/l) after 15 min, the CS is excluded. In subclinical CS, the cortisol is depressed at this time.
2. CRH stimulation test. The ovine (o) CRH test is of greater value than the dexamethasone suppression test in differentiating Cushing's disease and ectopic ACTH syndrome.

3. Overnight 8-mg dexamethasone suppression test. This test is currently not recommended for routine use. Because of its sensitivity of about 80% of cortisol suppression in CS, it yields no further information than the low-dose test.
4. Petrosal-sinus sampling with CRH test. This procedure can help to differentiate between a pituitary and a non-pituitary source of corticotropin secretion: if there is a ratio of two between basal central and basal peripheral sampling of corticotropin, or after stimulation with CRH a ratio of three, the diagnosis of Cushing's disease is clear.

Radiographic Findings

Imaging techniques can help clarify the etiology of hypercortisolism. These include computed tomographic (CT) scanning and magnetic resonance imaging (MRI) of the pituitary gland and CT scan, MRI, and ultrasound imaging of the adrenal glands. CT and MRI scans of the chest and abdomen are also used when tumors secreting ectopic ACTH are suspected.

Treatment

The preferred treatment for CS depends upon the specific cause of the hypercortisolism, which must be firmly and unequivocally established. The optimal treatment is correction of hypercortisolism without permanent dependence on hormone replacement.

Cushing's Disease

Transsphenoidal surgery (TSS) Transsphenoidal surgery is the treatment of choice for most cases of CS caused by pituitary microadenomas. It is indicated if the presence of a pituitary microadenoma can be demonstrated preoperatively by imaging techniques or bilateral inferior petrosal sinus blood sampling (BIPSS). In most specialized centers, the success rate of first TSS exceeds 90%. If BIPSS has lateralized the microadenoma and the surgeon cannot identify it at surgery, 75–80% of patients can be cured by ipsilateral hemihypophysectomy. Usually, successful TSS leads to the cure of hypercortisolism without using permanent glucocorticoid replacement. Approximately 5% of patients suffer recurrences; this is most common in patients with pituitary macroadenomas. The success rate of repeat TSS is considerably lower (approximately 50–60%). Complications include transient or permanent, partial or complete anterior pituitary insufficiency, including, in order of frequency, hypothyroidism, growth hormone deficiency, hypogonadism, and adrenal insufficiency. Transient diabetes

insipidus and inappropriate antidiuretic secretion syndrome are also common. Success after TSS is defined as a decrease in serum cortisol or UFC to an undetectable level in the immediate postoperative period. After a successful TSS, a period of adrenal insufficiency occurs in most patients cured, during which glucocorticoids must be replaced at a dose of 12–15 mg/m²/day. This abnormality of the HPA axis can last as long as a year or longer. It is rare, though possible, that it is permanent. The replacement dose of hydrocortisone is maintained for 3 months, and adrenocortical function is evaluated at that time with a rapid Cortrosyn test (250 mg ACTH 1–24 iv bolus, with plasma cortisol measured at 0, 30, and 60 min). If the test is normal (cortisol >18 and >20 µg/dl (500 nmol/l and 550 nmol/l) at 30 and 60 min, respectively), glucocorticoid replacement is discontinued. If the response is subnormal, the patient is reevaluated at 3-month intervals. About 70–80% of the patients will have a normal test at 6 months postoperation. Patients should be given extra glucocorticoids during stress (two times replacement dose for minor stress such as febrile illness or dental surgery, and 8–10 times replacement dose for major stress, such as major trauma or surgery).

Pituitary X-radiation with or without mitotane This is an alternative treatment that may be used after failure of TSS, when there is presence of cavernous sinus wall invasion by the tumor, or in patients judged unsuitable for surgery. The total recommended dosage of pituitary irradiation is 4500 to 5000 rad. High-voltage, conventional X-radiation is administered in 180- to 200-rad fractions over a period of 6 weeks. Success rate is ~40–50%, without mitotane and ~80% with mitotane (see later).

Pharmacotherapy Drug therapy alone is rarely used to treat Cushing's disease, except temporarily, prior to definitive treatment. Mitotane (Lysodren) is the only available pharmacological agent that both inhibits biosynthesis of corticosteroids and destroys adrenocortical cells secreting cortisol, thus producing a long-lasting effect. Therapy with mitotane (up to 3 g/day) is usually combined with radiotherapy. Aminoglutethimide, up to 1 g/day orally in four divided doses, or metyrapone, up to 1 g/day in four divided doses, may also be used alone or in combination with mitotane. During treatment, UFC excretion should be monitored and the regimen should be titrated to maintain this parameter in the normal range. If adrenal insufficiency is suspected, oral hydrocortisone should be added.

Ketoconazole (up to 1400 mg/day or 20 mg/kg/day) or trilostane have also been used alone or in combination with mitotane to control some of the symptoms and metabolic abnormalities associated

with Cushing's disease. Combinations are recommended because they usually prevent breakthroughs that occur when the drugs are used alone. In addition, one can use moderate doses with less side effects from each agent.

There are also new compounds such as fluconazole and glitazones, and for ectopic CS there are somatostatin analogs such as octreotide, lanreotide, and cabergolin.

Bilateral adrenalectomy The indications for this procedure have been dramatically changed by the success and low morbidity of TSS. Bilateral adrenalectomy is a possible treatment for patients who have failed selective pituitary adenomectomy or hemihypophysectomy. The major disadvantages of this surgery are that the individual is committed to lifelong daily gluco- and mineralocorticoid replacements, it does not eliminate the cause underlying the hypersecretion of ACTH, and relapses, though rare, may occur as a result of growth of adrenal rest tissue or an adrenal remnant. Furthermore, perioperative mortality is approximately three times higher than that of TSS, although it can be minimized by careful perioperative preparation. Also, Nelson syndrome (i.e., large pituitary macroadenomas secreting great amounts of ACTH and lipotropin resulting in skin hyperpigmentation) may occur in approximately 10–15% of patients with Cushing's disease treated with bilateral adrenalectomy. These tumors can impair vision and should be treated surgically and by radiotherapy.

Ectopic ACTH-Dependent CS

The preferred treatment for ectopic ACTH secretion is surgery aimed at completely excising the tumor, if resectable and of known location. If there is evidence of adjacent lymph node invasion, local radiation can be recommended after surgery for some tumors, such as carcinoids. If surgical resection is impossible or if the tumor is hidden, blockade of steroidogenesis is indicated; if necessary, bilateral adrenalectomy is considered.

See Also the Following Articles

Adrenal Cortex; Cushing's Syndrome, Neuropsychiatric Aspects; Pituitary Regulation, Role of.

Further Reading

- Arnaldi, G., et al. (2003). Diagnosis and complications of Cushing's syndrome: a consensus statement. *Journal of Clinical Endocrinology and Metabolism* 88(12), 5555-5560.
- Aron, D. C., et al. (1997). Effectiveness versus efficacy. The limited value in clinical practice of high dose dexamethasone suppression testing in the differential diagnosis of ACTH-dependent Cushing's syndrome. *Journal of Clinical Endocrinology and Metabolism* 82(6), 1780–1785.
- Findling, J. W., et al. (2005). Screening and diagnosis of Cushing's syndrome. *Endocrinology and Metabolism Clinics of North America* 34(2), 384–502.
- Ilias, I., et al. (2005). Cushing's syndrome due to ectopic corticotropin secretion: twenty years' experience at the National Institutes of Health. *Journal of Clinical Endocrinology and Metabolism* 90(8), 4955–4962.
- Katabami, T., et al. (2005). Discrepancies in results of low- and high-dose dexamethasone suppression tests for diagnosing preclinical Cushing's syndrome. *Endocrinology Journal* 52(4), 463–469.
- Monson, J. P., et al. (2000). The epidemiology of endocrine tumours. *Endocrine-Related Cancer* 7(1), 29–365.
- Newell-Price, J., et al. (1995). A single sleeping midnight cortisol has 100% sensitivity for the diagnosis of Cushing's syndrome. *Clinical Endocrinology (Oxford)* 43(5), 545–550.
- Newell-Price, J., et al. (1998). The diagnosis and differential diagnosis of Cushing's syndrome and pseudo Cushing's states. *Endocrine Reviews* 19(5), 647–672.
- Newell-Price, J., et al. (2006). Cushing's syndrome. *The Lancet* 367, 1605–1617.
- Niemann, L. K., et al. (2002). Medical therapy of Cushing's disease. *Pituitary* 5(2), 77–82.
- Reimondo, G., et al. (2005). Evaluation of the effectiveness of midnight serum cortisol in the diagnostic procedures for Cushing's syndrome. *European Journal of Endocrinology* 153(6), 803–809.
- Riedl, M., et al. (2006). Long term control of hypercortisolism with fluconazole: case report and in vitro studies. *European Journal of Endocrinology* 154(4), 519–524.
- Terzolo, M., et al. (2004). Subclinical Cushing's syndrome. *Pituitary* 7(4), 217–223.
- Viardot, A., et al. (2005). Reproducibility of nighttime salivary cortisol and its use in the diagnosis of hypercortisolism compared with urinary free cortisol and overnight dexamethasone suppression test. *Journal of Clinical Endocrinology and Metabolism* 90(10), 5730–5736.
- Yaneva, M., et al. (2004). Midnight salivary cortisol for the initial diagnosis of Cushing's syndrome of various causes. *Journal of Clinical Endocrinology and Metabolism* 89(7), 3345–3351.

Cushing's Syndrome, Neuropsychiatric Aspects

M N Starkman

University of Michigan, Ann Arbor, MI, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M Starkman, volume 1, pp 621–625, © 2000, Elsevier Inc.

Neuropsychiatric Symptoms in Patients with Spontaneous Untreated Cushing's Syndrome

Improvement in Neuropsychiatric Symptoms after Treatment

Brain-Imaging Studies in Cushing's Syndrome

Mechanisms for the Behavioral Effects of Cortisol

Glossary

<i>Hippocampus</i>	A brain structure that is part of the limbic system and key to learning and memory.
<i>Libido</i>	Desire for sexual activity.
<i>Mental status evaluation</i>	Standard bedside clinical tests used to examine mental functions.
<i>Neuropsychological studies</i>	Tests of various aspects of cognition, including verbal functions, visuospatial functions, learning, and memory.
<i>Sleep electroencephalogram (EEG)</i>	An electroencephalogram taken during sleep to measure brain electrical activity and study the structure of sleep.

Cushing's disease, the major type of spontaneous Cushing's syndrome, is the classic endocrine disease characterized by hypersecretion of cortisol. Intriguingly, emotional disturbances were already recognized as a feature of the disease in Harvey Cushing's original description in 1932. This association was then confirmed by retrospective chart reviews. More recently, studies of patients with spontaneous Cushing's syndrome examined prior to treatment have characterized the disturbance and shown that the majority manifest clinical features similar to those seen in patients with a primary psychiatric depressive disorder. Patients with Cushing's disease and syndrome exhibit a consistent constellation of symptoms, including abnormalities in mood (irritability and depression), vegetative functions (decreased libido and increased insomnia), and cognitive functions (decreased concentration, learning, and memory).

Neuropsychiatric Symptoms in Patients with Spontaneous Untreated Cushing's Syndrome

Mood and Affect

Irritability, a very frequent symptom seen in close to 90% of patients, is usually the earliest behavioral symptom to appear. It begins close to the onset of the earliest physical symptom (weight gain) and prior to the appearance of other physical manifestations of Cushing's syndrome. Patients describe themselves as feeling overly sensitive, unable to ignore minor irritations, and impatient with or pressured by others. In addition, overreactivity and easy development of anger occur. Patients feel that they are often on the verge of an emotional explosion and that the intensity of anger experienced is also increased.

Depressed mood is reported by 60–80% of patients. There is a range in the intensity of depressed mood. Some patients describe short spells of sadness; others experience feelings of hopelessness and giving up. Suicide attempts are infrequent but may occur. Hypersensitivity and oversentimentality lead to crying spells. For some, crying also occurs as the behavioral response to anger and frustration and feeling unable to respond effectively. Patients also experience spontaneous onset of depressed mood or crying in the absence of any preceding upsetting thought or event.

The time course of the mood disturbances is noteworthy. Most patients report that their mood disturbances are intermittent rather than sustained. Sometimes they wake up depressed and remain depressed throughout the day or the next day as well. Alternatively, the onset of depressed mood and/or crying might occur during the day, often suddenly. The duration of each depressive episode is usually 1 to 2 days and is rarely longer than 3 days at a time. A frequent weekly total is 3 days per week. There is no regular cyclicity, however, so patients cannot predict when a depressive day will occur. Although there are intervals when they might not experience pleasure, these patients do not experience the unrelenting, unremitting inability to experience pleasure that is characteristic of patients with severe psychiatric depressive illness. There are intervals when patients retain the capacity for pleasure, finding enjoyment in hobbies and interpersonal relationships. At times,

they may find it difficult to initiate such activities, although once others mobilize them, they are able to enjoy them. Some patients do experience decreased interest in their environment.

Social withdrawal is less common, but can occur. Sporadic withdrawal might occur because of the patient's need to remove him- or herself from a situation of overstimulation that elicits the fear of impending emotional dyscontrol. Guilt is infrequent; if present, it is not excessive, self-accusatory, or irrational, but is related to remorse about angry outbursts and inability to function well at work and in the family. Hopelessness is infrequent, but, if present, is attributed to the increasing physical and emotional disability that has puzzled the patient's physicians and been resistant to treatment interventions until the diagnosis of Cushing's syndrome is finally made.

A minority of patients experience episodes of elation-hyperactivity early in the course of the illness. These episodes of elation are described as a high. During these episodes, patients are more ambitious than usual and might attempt to do more than their ability and training make reasonable. Increased motor activity is present, with restlessness and rapidly performed activities. Patients report embarrassment that their speech is both loud and rapid. As the illness progresses and new physical signs of Cushing's syndrome begin to appear, this type of episode becomes rare or disappears entirely.

A percentage of patients report generalized anxiety. New-onset panic disorder has also been observed in patients with Cushing's syndrome. In addition, even patients who do not experience psychic anxiety describe episodic symptoms of autonomic activation such as shaking, palpitations, and sweating.

Biological Drives

Abnormalities in four areas of basic biological vegetative drives are present:

1. Fatigue. This is reported by 100% of patients.
2. Libido. A decrease in libido is very frequent and reported by close to 70% of patients. In fact, this is one of the earliest manifestations of Cushing's disease, beginning when the patient is experiencing the first onset of weight gain.
3. Appetite and eating behavior. More than 50% of patients have an alteration in their appetite; in 34% appetite is increased; in 20% it is decreased.
4. Sleep and dreams. Difficulty with sleep, particularly middle insomnia and late insomnia (early morning awakening), is found in more than 80% of patients. Difficulty with early insomnia (not falling

asleep at bedtime) is not as frequent. One-third of patients report an alteration in the frequency or quality of their dreams, which are increased in their frequency and intensity and become bizarre and very vivid. Some patients report they have lost the ability to wake themselves out of a nightmare. Sleep EEG studies indicate that there are many similarities between Cushing's disease patients and patients with major depressive disorder. For example, both groups show significantly less total sleep time, lower sleep efficiency, and shortened rapid eye movement (REM) latency compared to normal subjects.

Cognition

Cognitive symptoms are a prominent part of the clinical picture. Most patients report difficulty with concentration, inattention, distractibility, and shortened attention span. Some patients report episodes of scattered thinking, while others note slow thinking. Thought blocking may occur in more severe instances. Patients may find themselves using incorrect words while speaking or misspelling simple words. Perceptual distortions are very rare.

Impairment of learning and memory is among the most frequent symptoms and is reported by 80% of patients. Patients report problems with registration of new information. They commonly repeat themselves in ongoing conversations. They easily forget items such as appointments made, names of people, and location of objects.

Most patients have no disorientation or overt clouding of consciousness, and their waking electroencephalographs (EEGs) are not characteristic of delirium. On mental status bedside clinical evaluation, difficulties with tests such as mental subtraction and recall of recent U.S. presidents are seen in close to 50% of the patients.

Detailed neuropsychological studies reveal that individuals vary in the severity of degree of cognitive dysfunction, with minimal or moderate to severe decrements in a variety of subtests seen in close to two-thirds of patients. Verbal functions and learning are particularly affected. Once learned, the percent retention of verbal material is not subnormal. The overall pattern of decrements in cognitive function suggests the involvement of both frontal lobe and hippocampus.

Specific Features of the Neuropsychiatric Symptoms

The psychiatric symptoms that develop in Cushing's syndrome are not simply a nonspecific response to

severe physical illness. Irritability and decreased libido occur early, often before patients are aware that they have any physical problems other than weight gain. Later, when depressed mood appears, it is not simply the demoralization seen in the medically ill, but episodic sadness and crying, sometimes occurring in the absence of depressive thought content. The incidence of depressive disorders is greater than in comparison groups with other types of pituitary tumors or hyperthyroidism. Although they have difficulty with concentration and memory, patients with Cushing's syndrome are not delirious.

The overwhelming majority of patients with spontaneous Cushing's syndrome do not have very severe neuropsychiatric disturbances of a psychotic or confusional nature, as patients receiving corticosteroids for treatment of certain physical disorders sometimes may. This can be understood by considering that patients with spontaneous Cushing's syndrome differ from those receiving high-dose steroids in several respects. The level of circulating cortisol in the former is not as high as the equivalent amount of steroid administered to the latter. Patients with Cushing's syndrome also differ from those receiving exogenous steroids in that they are exposed to sustained elevated cortisol levels for months to years and are less subject to sudden acute shifts and rapid rates of change of steroid levels that occur during short-term treatment with high-dose corticosteroids.

Although similar in many respects to the major depressive disorders, the depression seen in Cushing's syndrome does have certain distinguishing clinical characteristics. Irritability is a prominent and consistent feature, as are symptoms of autonomic activation such as shaking, palpitations, and sweating. Depressed affect is often intermittent, with episodes lasting 1 to 3 days and recurring very frequently at irregular intervals. Patients usually feel their best, not their worst, in the morning. Psychomotor retardation, although present in many patients, is usually not so pronounced as to be clinically obvious and is usually apparent only in retrospect after improvement with treatment. The majority of these patients are not withdrawn, monosyllabic, unsponaneous, or hopeless. Their guilt, when present, is not irrational or self-accusatory and is primarily related to their realistic inability to function effectively. Significant cognitive impairment, including disorders of concentration, learning, and memory, is a very consistent and prominent clinical feature.

Improvement in Neuropsychiatric Symptoms after Treatment

Significant improvements in depression rating scores occur after treatment produces decreases in cortisol.

In treated patients, improvements in depressed mood begin with a decrease in the frequency of days when the patient feels depressed. In addition, each episode lasts a shorter period of time, perhaps only a few hours instead of 1 to 2 days. The patients also describe a change in the quality of the depressive mood, so that they no longer experience it being as deep. They no longer feel depressed without some external precipitating reason. A mood change comes on gradually and abates gradually, rather than appearing suddenly as before. Crying becomes less frequent, is less easily elicited by environmental upsets, occurs only with some identifiable external precipitant, and is of shorter duration.

The time course of improvement in depressed mood compared to improvement in other neurovegetative symptoms is of interest. In patients with Cushing's disease who manifest depressed mood at initial evaluation and are subsequently studied during the first 12 months after treatment, improvement in symptoms other than depressed mood occur prior to improvement in depressed mood. Depressed mood is less likely than irritability and sleep, for example, to be among the first cluster of symptoms to improve. Interestingly, this lag is similar to that seen in psychiatric patients with depressive episodes treated with antidepressants, in whom improvements in sleep and psychomotor activity often occur prior to improvement in depressed mood.

Cognitive testing after treatment indicates improvement in many, but not all, affected functions. After cortisol levels decline to normal concentrations, verbal learning and verbal functions such as verbal fluency improve.

Antidepressant medication is more effective after treatment has normalized cortisol levels. When selective serotonin reuptake inhibitors are administered prior to treatment, patients report only partial effectiveness, so that only mood or only vegetative symptoms improve. After treatment, antidepressants can be helpful when depressed mood lingers or is exacerbated by the sharp and rapid decline in steroid levels.

The long-term impact of Cushing's disease on subjective well-being after cure has been studied by using validated health-related questionnaires that assess quality of life. While there is partial resolution of physical and psychosocial decrements after treatment, at long-term follow-up many report some decrease in quality of life.

Brain-Imaging Studies in Cushing's Syndrome

Brain-imaging technology has provided further information about the effects of cortisol on brain structure and function. The hippocampus is a brain

structure key to learning and memory. During active Cushing's syndrome, greater elevations in cortisol are associated with smaller hippocampal formation volume. Reduced hippocampal formation volume is associated with lower scores for verbal learning and recall. After treatment reduces cortisol concentrations to normal, the hippocampal formation increases in volume more consistently and to a greater degree than other brain structures examined for comparison. Greater decreases from the initial pretreatment cortisol levels are associated with greater increases in hippocampal formation volume. Studies indicate that patients with Cushing's disease show a reduction in the metabolism of glucose in several brain regions. Studies using functional brain imaging in Cushing's syndrome indicate altered activation in the verbal learning (encoding) systems, including the hippocampus. In addition, accuracy of emotion processing is decreased, and altered activation patterns occur in the brain regions subserving it.

Mechanisms for the Behavioral Effects of Cortisol

The neuropsychiatric abnormalities seen in untreated Cushing's syndrome are related, at least in part, to the effect of elevated levels of cortisol. Cortisol has pleiotropic effects in the central nervous system. Possible mechanisms for the effect of cortisol on mood, cognition, and vegetative functions include those already shown to be actions of glucocorticoids in animals. These include a direct effect on receptors of central nervous system cells, the synthesis or function of neurotransmitters, effects on glucose metabolism and electrolytes, increasing sensitivity of the brain to other neuroactive substances such as excitatory amino acids, effects on nerve growth factors, and neurogenesis.

While this article has focused on a single adrenal steroid, cortisol, levels of other adrenal glucocorticoids as well as sex steroids produced by the adrenal, such as testosterone and dehydroepiandrosterone (DHEA), are also altered in patients with Cushing's syndrome. These substances likely modify the effect of cortisol on psychopathology, and their role remains to be elucidated.

See Also the Following Articles

Adrenal Cortex; Adrenal Medulla; Hypocortisolism and Stress; Major Depressive Disorder; Pituitary Regulation, Role of.

Further Reading

Brunetti, A., Fulham, M. J., Aloj, L., et al. (1998). Decreased brain glucose utilization in patients with

Cushing's disease. *Journal of Nuclear Medicine* 39, 786–790.

Cohen, S. I. (1980). Cushing's syndrome: a psychiatric study of 29 patients. *British Journal of Psychiatry* 136, 120–124.

Jeffcoate, W. J., Silverstone, J. T., Edwards, C. R. and Besser, G. M. (1979). Psychiatric manifestations of Cushing's syndrome: response to lowering of plasma cortisol. *Quarterly Journal of Medicine* 48, 465–472.

Kelly, W. F., Checkley, S. A., Bender, D. A. and Mashiter, K. (1983). Cushing's syndrome and depression – a prospective study of 26 patients. *British Journal of Psychiatry* 142, 16–19.

Lindsay, J. R., Nansel, T., Baid, S., Gumowski, J. and Nieman, L. (2006). Long-term impaired quality of life in Cushing's syndrome despite initial improvement after surgical remission. *Journal of Clinical Endocrinology and Metabolism* 91, 447–453.

Loosen, P. T., Chambliss, B., DeBold, C. R., Shelton, R. and Orth, D. N. (1992). Psychiatric phenomenology in Cushing's Disease. *Pharmacopsychiatry* 25, 192–198.

Mauri, M., Sinforiani, E., Bono, G., et al. (1993). Memory impairment in Cushing's disease. *Acta Neurologica Scandinavica* 87, 52–55.

Shibley, J. E., Schteingart, D. E., Tandon, R. and Starkman, M. N. (1992). Sleep architecture and sleep apnea in patients with Cushing's disease. *Sleep* 15, 514–518.

Sonino, N., Fava, G. A., Belluardo, P., Girelli, M. E. and Boscaro, M. (1993). Course of depression in Cushing's syndrome: response to treatment and comparison with Grave's disease. *Hormone Research* 39, 202–206.

Starkman, M. N. (1987). Commentary on M. Majewska: actions of steroids on neuron: role in personality, mood, stress and disease. *Integrative Psychiatry* 5, 258–273.

Starkman, M. N. and Schteingart, D. E. (1981). Neuropsychiatric manifestations of patients with Cushing's syndrome: relationship to cortisol and adrenocorticotrophic hormone levels. *Archives of Internal Medicine* 141, 215–219.

Starkman, M. N., Schteingart, D. E. and Schork, M. A. (1981). Depressed mood and other psychiatric manifestations of Cushing's syndrome: relationship to hormone levels. *Psychosomatic Medicine* 43, 3–18.

Starkman, M. N., Schteingart, D. E. and Schork, M. A. (1986). Cushing's syndrome after treatment: changes in cortisol and ACTH levels, and amelioration of the depressive syndrome. *Psychiatric Research* 17, 177–188.

Starkman, M. N., Gebarski, S. S., Berent, S. and Schteingart, D. E. (1992). Hippocampal formation volume, memory dysfunction, and cortisol levels in patients with Cushing's syndrome. *Biological Psychiatry* 32, 756–765.

Starkman, M. N., Giordani, B., Gebarski, S., Berent, S., Schork, M. A. and Schteingart, D. E. (1999). Decrease in cortisol reverses human hippocampal atrophy following treatment of Cushing's disease. *Biological Psychiatry* 46, 1595–1602.

Starkman, M. N., Giordani, B., Berent, S., Schork, M. A. and Schteingart, D. E. (2001). Elevated cortisol levels in Cushing's disease are associated with cognitive decrements. *Psychosomatic Medicine* 63, 985–993.

- Starkman, M. N., Giordani, B., Gebarski, S. S. and Schteingart, D. E. (2003). Improvement in learning associated with increase in hippocampal formation volume. *Biological Psychiatry* 53, 233–238.
- Tucker, R. P., Weinstein, H. E., Schteingart, D. E. and Starkman, M. N. (1978). EEG changes and serum

- cortisol levels in Cushing's syndrome. *Clinical Electroencephalography* 9, 32–37.
- Whelan, T. B., Schteingart, D. E., Starkman, M. N. and Smith, A. (1980). Neuropsychological deficits in Cushing's syndrome. *Journal of Nervous Mental Diseases* 168, 753–757.

Cytokines

G D Marshall

University of Mississippi Medical Center, Jackson, MS, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G D Marshall, Jr., and J L Rossio, volume 1, pp 626–632, © 2000, Elsevier Inc.

Relationships between Stress and Immunity
Impact of Stress and Behavior on Cytokine Expression
Impact of Cytokines on Stress and Behavior
Summary

Glossary

<i>Cytokine</i>	Small molecular weight glycopeptide produced by a variety of cell types that serve to modulate immune and inflammatory responses.
<i>Immunoendocrinology</i>	The study of the impact of soluble products of the immune system on endocrine and neuroendocrine networks as part of the bidirectional communication between the immune and endocrine systems.
<i>Psychoneuroimmunology</i>	The study of the relationships between neurocognitive behaviors, neuroendocrine response, and immune changes occurring in response to stress.
<i>Regulatory T cells</i>	CD4+ (and some CD8+) cells that reside in both the central and peripheral immune systems whose main function is to regulate the intensity and duration of immune responses by inhibiting both cytokine production and effector function.
<i>Stress</i>	Perturbation of homeostatic state by environmental factors demanding an adaptive response from the host.
<i>T helper cell</i>	Lymphocyte that plays a central role in immune modulation by selectively activating various immune and nonimmune cells through the production of specific cytokine combinations.

Relationships between Stress and Immunity

Stress Defined

Stress is best thought of as a psychophysiological process, usually experienced as a negative emotional state. Stressors, defined as events with the potential for posing threat, harm, or challenge to the host, are judged in the context of personal and environmental factors, and if appraised as threatening or challenging, initiate specific physiologic responses directed at reducing or eliminating the stress (fight-or-flight). The field of psychoneuroimmunology seeks to establish the link between behavior, neuroendocrine functions, and immune responses.

Impact of Stress on Health

A common clinical observation is the often adverse relationship between stress and human disease. Various sources have estimated that up to 75% of all visits to physicians' offices are stress-related. This appears to be particularly true in relationship to immune-based dysfunction such as increased susceptibility to infections, autoimmune, allergic, and asthmatic diseases. Stress is also suspected to play a role in morbidity and mortality in other immune-based diseases such as cancer, HIV disease, inflammatory bowel diseases, and even aging. Stress may also cause persistent increases in sympathetic nervous system activity, including increased blood pressure, heart rate, and catecholamine secretion, as well as platelet aggregation which may explain, at least in part, the known association between stress, immune alterations, and cardiovascular disease. Emerging evidence suggests that sleep alterations can modulate the stress–health relationship. Acute stressors and adverse naturalistic events are associated with sleep disruptions as measured by self-report and polysomnography. Poor sleep, in turn, is associated with subsequent decrements in mental health including symptom reporting, incident cases of mood and anxiety disorders, and

immune function. Sleep disruptions have also been associated with adverse physical health outcomes, including increased morbidity and mortality compared to populations with adequate patterns and duration of sleep.

Role of Cytokines in the Normal Immune Response

In the normal host, presentation of antigen for a specific protective immune response elicits a complex series of events that results in a mixed cellular and humoral protective response, the intensity and nature of which depends upon the specific inciting antigen. Generally speaking, extracellular parasites (i.e., bacteria) incite primarily a humoral response while intracellular parasites (i.e., virus, fungi, mycobacteria, etc.) elicit a cell-mediated response. Antiviral immunity is particularly complex because both arms are necessary for host resistance – cellular immunity to eliminate the virus-infected host cells and humoral immunity that produces antiviral neutralizing antibodies to prevent reinfection. The host immune response has a variety of mechanisms to direct the immune response into the humoral versus cellular direction, i.e., nature of antigen-presenting cells, major histocompatibility complex (MHC) restriction, and availability of specific T and B cell components. However, the central control of the cellular versus humoral response to an antigenic challenge appears to be via production of specific cytokine milieu.

All cytokines have certain properties in common. They are all small molecular weight peptides or glycopeptides. Many are produced by multiple cell types such as lymphocytes, monocytes/macrophages, mast cells, eosinophils, and even endothelial cells lining blood vessels. Each individual cytokine can have multiple functions depending on the cell that produces it and the target cell(s) on which it acts (called pleiotropism). Also, several different cytokines can have the same biologic function (called redundancy). Physiologically it appears that most cytokines exert their most important effects in a paracrine and/or autocrine rather than endocrine fashion. Their major functions appear to involve host defense or maintenance and repair of the blood elements.

Cytokines are categorized by their major specific function(s) and there are five major categories. Interferons (IFNs) are so named because they interfere with virus replication. There are three major types based upon the source of the IFN. IFN alpha (IFN α) is produced by the buffy coat layer from white blood cells, IFN beta (IFN β) by fibroblasts, and IFN gamma (IFN γ) by activated T cells, natural killer (NK) cells, and dendritic cells (DC). IFN γ is an important immunoregulatory molecule in

determining intracellular versus extracellular host defense responses.

The colony stimulating factors (CSFs) are so named because they support the growth and differentiation of various elements of the bone marrow. Many are named by the specific element they support such as granulocyte colony stimulating factor (G-CSF), macrophage colony stimulating factor (M-CSF), and granulocyte macrophage colony stimulating factor (GM-CSF). Other CSFs include interleukin (IL)-3, which can stimulate a variety of hematopoietic precursors and is being evaluated as a therapy in aplastic anemia and bone marrow transplantation; and c-Kit ligand (stem cell factor) which has been demonstrated as a cytokine necessary to cause the differentiation of bone marrow stem cells into their various precursor elements for eventual differentiation into erythrocytes, white blood cells, and megakaryocytes (platelets).

The tumor necrosis factors (TNFs) are so called because injecting them into animals causes a hemorrhagic necrosis of the animals' tumors. TNF- α is produced by activated macrophages and TNF- β is produced by activated T cells (both T helper (Th) and cytotoxic T lymphocytes (CTL)). These molecules appear to be involved in the pathogenesis of septic shock and much research is aimed at trying to inhibit their activity in septic patients.

The newest group described is the chemokines, small molecular weight proteins that promote chemotaxis and activation of cells to tissue sites experiencing inflammatory reactions. Currently more than 50 human chemokines and 18 receptors have been identified. They appear to be intimately involved in generating selective types of inflammatory cell infiltrates both because of the specific chemokine milieu produced as well as selective expression of particular chemokine receptors on specific inflammatory cell subpopulations. They are categorized into subfamilies (CXC, CC, CX3C, and so on) based on their chemical composition of N-terminal cysteinyl receptors. Examples include RANTES (regulated upon activation, normal T cell expressed and secreted), macrophage inhibitory peptide (MIPs), and eotaxin.

The largest group is the ILs, so named because their fundamental function appears to be communication between (inter-) various populations of white blood cells (leukocytes – leikin). They are produced by a variety of cell types such as monocytes/macrophages, T cells, B cells, and even nonleukocytes such as epithelial cells.

A central source of these cytokines comes from CD4+ helper T cell subpopulations, often referred to as Th1 and Th2 cells. Human Th1 cells secrete a specific cytokine profile including IFN γ and TNF- β .

These cytokines are important helper factors in the generation of cellular immune responses including antigen-specific CTL and NK cells. Additionally, IFN γ in particular has an antagonistic activity against Th2 cytokines. IL-12, produced primarily by activated macrophages and other antigen presenting cells (APC), plays a central role in inducing IFN γ production. In contrast, Th2 cells secrete IL-4, IL-5, IL-9, IL-10, and IL-13, which are involved in isotype switching of B cells as well as proliferation and differentiation into antibody-secreting plasma cells. In particular, IL-4 and IL-13 are involved in the isotype switch from immunoglobulin (Ig)M to immunoglobulin (Ig)E, the antibody responsible for classical allergic disease. IL-4 and IL-10 are also regulatory cytokines, antagonizing the activities of Th1 cytokines. In humans, Th cells can have a third cytokine secretory pattern. Th0 cells are capable of secreting all of the above mentioned cytokines and are often thought of as precursors that differentiate into the Th1 or Th2 pathway with antigen challenge. Thus it can be said that the nature, intensity, and duration of a specific immune response depends upon the delicate balance between Th1 and Th2 numbers and/or activities. Since it is now appreciated that many other cell types besides Th cells produce IFN γ , IL-4, IL-5, IL-9, and IL-10, many scientists now refer to these as type 1 and type 2 cytokines, supporting cellular and humoral immune responses, respectively.

Impact of Stress on Neuroendocrine and Immune Function

Extensive research efforts have also examined various psychosocial variables in order to predict patterns of adjustment to stressful situations. These variables include optimism, social support, and coping responses, all of which have been found to buffer the psychological, cardiovascular, endocrine, and immune effects of stress and predict overall adjustment to stressful life events. Further, social support and coping strategies can moderate these effects. There is also research indicating that coping style mediates the effects of social support on distress, such that patients who receive social support may use fewer avoidance coping strategies and therefore have lower overall distress. Similarly, in addition to the direct effects optimism has on adjustment to stressful situations, some researchers have found that the benefits of an optimistic disposition are mediated through its effects on both social support and coping responses. These data lend support to the notion that many of these immune imbalances can be reversed with various forms of stress management.

Impact of Stress and Behavior on Cytokine Expression

It has been suggested that acute versus chronic stress may exert distinct effects on host immunity. Acute stress may initially produce an adaptive immune response that is actually protective of the normal host. NK cell activity, as well as certain potentially antiviral cytokines, increases in normal individuals who undergo acute laboratory stressors. Interestingly, chronic stress exposure appears to result in the opposite effect with decreases in NK function and antiviral cytokines. It has also been reported that acute psychological and/or physiological stresses can be associated with adverse clinical outcomes. It is plausible that chronic stress disrupts the ability of some individuals to generate a protective homeostatic immune response when faced with acute stressful events.

Inflammatory Cytokines

When the host is subjected to infection, trauma, or injury, a series of cellular and molecular events occur to defend against potential invading organisms and initiate tissue repair. Collectively, these events are called inflammation and are of central importance in the maintenance of homeostasis. The normal components of inflammation involve the production of a variety of cytokines that function to recruit and activate the cellular components of inflammation and tissue repair which include neutrophils, mast cells, macrophages, and fibroblasts. This cellular cascade can be initiated by local microbial invasion through production of substances such as endotoxin (lipopolysaccharide – LPS) or the host neuroendocrine system with production of molecules such as substance P. Acute inflammation is characterized by the recruitment of polymorphonuclear (PMN) cells, basophils/mast cells, and the production of fibrinogen to aid in hemostasis. If the inflammatory response continues and becomes chronic, the PMN cells are replaced by mononuclear cells such as lymphocytes and monocytes/macrophages.

Stimulation of neurons can cause localized release of substance P which results in tissue edema from peripheral vasodilatation. Substance P can also stimulate tissue mast cells to become activated and release preformed histamine and other chemical mediators of inflammation including cytokines such as IL-4 and GM-CSF. Augmenting the response, histamine can bind to H1 and H2 receptors located on sensory nerve bodies and axons, further stimulating the release of substance P and thus potentiating the inflammatory reaction.

The major inflammatory cytokines are often considered to be TNF- α , IL-1 β , and IL-6. This is accompanied by the production of hepatic proteins collectively known as acute phase reactants (APR) that function to limit the activity and spread of invading microorganisms. Studies have shown the increase in serum inflammatory cytokine levels with certain types of acute stress likely due to mononuclear phagocyte activation. Epinephrine has a similar effect on increasing inflammatory cytokine production *in vitro* and is known to be increased *in situ* during acute stressful events. Acute stress also causes the increased production of substance P which appears to promote upregulation of many cytokines including the primary inflammatory TNF- α and IL-1 β . The substance P effect may be direct through stimulation of mononuclear cells or indirect through the induction of increased adrenal catecholamine secretion with stress.

Chronic stress is usually associated with diminished inflammatory response and poor wound healing. The inflammatory cytokine-APR networks can be inhibited by corticosteroids. Chronic stress, through the activation of the hypothalamic-pituitary-adrenocortical (HPA) axis, can increase systemic cortisol levels, thus acting as an inhibitor of inflammation. The increase in inhibitory cytokines, especially IL-10, which accompanies chronic stress may also contribute to the anti-inflammatory activity of stress.

Finally, in certain inflammatory diseases, there appears to be a paradoxical response to chronic stress that results in exacerbation of inflammatory pathways. While the mechanisms have not been established, it appears to involve more the systemic imbalance between pro- and anti-inflammatory cytokines rather than the absolute levels of any specific cytokine. Further, therapeutic interventions for stress reduction have demonstrated a salutary clinical effect in subpopulations of patients identified as having high stress levels. Work is in progress to determine whether this correlates with changes in any cytokine levels.

Immunoregulatory Cytokines

Recent work has demonstrated that the examination of only a single component of an immune response can lead to inaccurate conclusions regarding immunosuppression versus dysregulation. The reporting of a Th1/Th2 cytokine network subsequently expanded to include other cytokine-producing cells helps to explain the observations that an inappropriate humoral response to an antigen that requires a cellular response (and vice versa) is as harmful to the host as is generalized immunosuppression. Cytokines

can function in an immunoregulatory fashion, both by upregulating and downregulating the overall immune response as well as directing it toward a cellular, humoral, or mixed response. This is in response to antigen-presenting cells directing a host response to be primarily cellular or humoral against an intracellular versus extracellular parasite invasion.

Stress has been reported to have distinct effects on immunoregulatory cytokine production. A fundamental cytokine for expansion of an ongoing immune response is IL-2. Corticosteroids decrease IL-2 production by T cells. Investigations using rodent models have shown that IL-2 levels and production capability are decreased with various forms of stress. The continuance of diminished IL-2 is observed in adrenalectomized animals and indicates the complex nature of this phenomenon. Similar observations have been made for IFN γ . In rodents, IL-2 and IFN γ are produced by Th1 cells while IL-4, IL-5, and IL-10 are produced by Th2 cells. Several studies in animal and humans have shown the decrease in Th1 (type 1) cytokines with a concomitant rise in Th2 (type 2) cytokines in stress models such as chronic restraint, medical student exams, major depressive disorders, and chronic pain.

In contrast, acute stress in normal animals and humans has been reported to cause at least a transient increase in NK cell activity accompanied by increased IFN γ and decreased IL-10 production. This may represent an innate host immune response for protection against intracellular parasites similar to the demargination of granulocytes for host defense against extracellular parasites in the peripheral blood observed with acute stress. Some studies report that the changes in type 1/type 2 cytokine production are different with acute stress superimposed on a chronically stressed host. This could, at least in part, explain the observations of increased susceptibility to and durations of infections, particularly viral, observed in chronically stressed individuals.

Effector Cytokine Networks

Aside from inflammatory and immunoregulatory activities, cytokines also have direct effector functions. These functions include thermoregulation, antiviral, cytotoxic, and cell trafficking. IL-1, IL-6, and TNF- α have all been described as endogenous pyrogens and function to alter the core temperature as a host defense against invading microorganisms. Antipyretic drugs decrease the sensitivity of the hypothalamus to the effects of these cytokines. IFN γ can have a direct antiviral IFN effect. It also activates macrophages to increase antigen processing of viral particles and CTL and NK cells to kill virally infected

target cells. CTL make TNF- β (lymphotoxin) in response to specific antigen challenge that can directly kill the target cell. In order for a host immune response to be protective, the elements of the response must reach and remain at the site of the invading microorganism. This occurs primarily with alterations in leukocyte trafficking through the production of chemokines and adhesion molecules. Specific combinations are critical to insure the proper type, sequence, and duration of inflammatory cell influx.

The effects of stress on these effector networks have not been extensively investigated. Body temperature in situations of acute stress has been reported to increase in both humans and animal models. No significant long-term changes have been reported with chronic stress. As IFN γ levels decrease with chronic stress, it is not surprising that increased incidence, duration, and susceptibility to viral infections occur in these affected populations. This may be the result of decreased CTL and NK cell activity seen with chronic stress. Adhesion molecule expression on lymphocytes and target tissue cells as well as chemokine secretion are both inhibited by glucocorticoids and thus could be expected to diminish in stress environments where HPA axis activity is high.

Impact of Cytokines on Stress and Behavior

It has been appreciated for some time that cytokines can have a direct effect on the central nervous system (CNS) and thus psychological components of the host. This can occur directly through effects on certain parts of the brain that control behavior directly (i.e., hippocampus) or indirectly through the neuroendocrine network involving the hypophyseal-pituitary-adrenal axis. The end result is that cytokine production can affect behavior in both positive and negative ways for host homeostasis.

Effects of Cytokines on the Central Nervous System

This area is only recently emerging as a field of study. Much impetus for study has been as a result of the increased use of recombinant cytokines for the treatment of a variety of diseases. The CNS side effects observed with these agents has led to speculation as to the components of the CNS most affected. Although there is evidence to suggest that recombinant cytokines do not directly cross the blood-brain barrier (BBB), the relationships are none-the-less clear. The patterns with the cytokines are remarkably similar: an acute toxicity (flu-like illness) followed by chronic phase (neuras-thenia) with continued cytokine administration. The

nature and severity of the chronic symptoms appear to be dose related but occur virtually with all regimens given for more than 2 months. More severe symptoms such as psychomotor and cognitive disorders as well as delirium and coma may develop at any point in the therapeutic regimen and appear to be related to underlying CNS dysfunction (aging, history of brain injury, and concomitant CNS disease).

Natural and recombinant IFNs (α , β , and γ), T cell growth factors (IL-2, IL-4, and IL-12), proinflammatory cytokines (IL-1, IL-6, and TNF- α), and hematopoietins (GM-CSF, IL-3, and stem cell factor) have all been utilized to treat patients with a variety of illnesses. Their biochemistry and receptor bindings are largely unique yet all share similar CNS toxicities suggesting mechanisms based upon common albeit as yet undescribed physiologic activities.

Another area of research that suggests a direct effect of cytokines on CNS activity is pain perception in areas of inflammation. Pain is a protective host response generated in answer to tissue injury. It is mediated by afferent sensory nerves in the midst of tissue injury producing enhanced sensitivity to further noxious (hyperesthesia) as well as innocuous (allodynia) stimuli. The peripheral sensory nerve releases substance P into the injured area which in turn can modulate the production and secretion of inflammatory cytokines (such as TNF- α and IL-1) from resident inflammatory cells which promote the pain response. The effect appears to be indirect through production of nerve growth factor, prostaglandins, etc., since no cytokine receptors have been demonstrated on peripheral nerve cells.

Pain control in peripheral as well as CNS tissue involves endogenous opiate receptors. Corticotropin-releasing hormone (CRH) is part of the HPA axis upregulated in stress responses. It also upregulates the secretion of endogenous opiate precursors in both peripheral and central nervous tissues. IL-1 can potentiate the CRH response. Of note, both IL-1 and CRH receptors are present on lymphocytes and are upregulated during inflammation. Binding of either IL-1 or CRH causes these lymphocytes to secrete β -endorphins. Although IL-1 can mediate hyperesthesia in normal tissue, it appears that IL-1, through upregulation of CRH receptors, can mediate analgesia through inducing the local production of β -endorphin by lymphocytes.

Effects of Cytokines on Neuroendocrine Networks

The effect of HPA as well as sympathetic adrenal medullary (SAM) hormones on immune function has been extensively studied. Much work has been done to characterize the effects of various cytokines

on neuroendocrine networks to further define the bidirectional communication between neuroendocrine and immune systems. Research to date strongly suggests that the mechanism for this bidirectional crosstalk involve shared ligands and receptors. While stress is known to effect immunity primarily through its effect on neuroendocrine networks, stimulation of the immune system by invading microorganisms produce a cytokine milieu that can have a direct effect on the neuroendocrine network through CNS receptor binding.

There are multiple neuroendocrine hormones produced by components of the immune system including adrenocorticotrophic hormone (ACTH), CRH, thyroid-stimulating hormone (TSH), luteinizing hormone (LH), follicle-stimulating hormone (FSH), prolactin (PRL), growth hormone (GH), and somatostatin (SOM). This explains the observation of various immune activations (including cytokine production) noted *in vitro* after incubation of immune cells with these hormones. The exact physiologic role of these immune cell-derived neuropeptides is still not established.

However, when a microorganism invades the host, immune activation leads to cytokine production that finds its way to the BBB via the peripheral circulation. The cytokines do not pass but bind to neuronal cells in the hypothalamus via their specific cytokine receptors. Although there are many different possible cytokine combinations that may bind CNS tissue, common neuroendocrine effects include upregulation of ACTH and downregulation of TSH release. IL-1, IL-2, IL-6, and TNF- α all activate the adrenal axis through the stimulation CRH and/or ACTH directly. IL-1 and TNF- α inhibit the gonadal axis through downregulation of luteinizing hormone-releasing hormone (LHRH). TNF- α and IFN γ can both inhibit the thyroid axis through downregulating thyrotropin-releasing factor (TRF).

Effects of Cytokines on Behavior

Neurocognitive defects with various cytokines have been described in many of the studies using recombinant cytokine therapy. Further, somnolence, psychoses, depression, and personality changes have all been reported during and after cytokine infusions. Yet it is the so-called sickness behavior (malaise, social withdrawal, somnolence, and hyperesthesia) associated with cytokine release after infection, trauma, or even stress that best demonstrates the effects of cytokines on behavior.

Another condition implicating the relationship between cytokines and behavior is depression. Major depressive disorders commonly occur after

stressful episodes and are associated with increased HPA hormone levels. Interestingly, cardinal clinical symptoms of depression (somnolence, anorexia, and diminished libido) can all be induced by cytokines. IL-1, TNF- α , and IFN γ are all soporific when administered systemically. Anorexia is a major clinical sign in both depression and infections. TNF- α , IL-1, and IL6 have all been reported to suppress appetite in both animal models and humans receiving recombinant cytokine therapies. Libido requires an intact hypophysial-pituitary activity to stimulate production of sex hormones. These can also be regulated by the inflammatory cytokines noted above. Of note, several studies report increased IL-1 and IL-6 levels with decreased IL-2 and IFN γ in plasmas from depressed patients.

Summary

The relationships between stress, distress, immunity, and health are now generally accepted but still not fully defined. Cytokines play a central role in immune responses and are quite sensitive to the effects of both acute and chronic stress through the neuroendocrine network. This can be both a physiological protective system as well as pathologically imbalanced or suppressed. Many factors influence the balance between cytokine and neuroendocrine networks and impact host defense mechanisms against infection, hypersensitivity diseases, and cancer. The stressors that influence host defense can be psychological, physiological, or a combination. Cytokines can, in turn, influence neuroendocrine networks, CNS function, and even behavior directly. A better understanding of the cytokine–neuroendocrine crosstalk and the influence of various stressors on that crosstalk offers opportunity for future interventional strategies for stress-related illnesses.

Further Reading

- Agarwal, S. K. and Marshall, G. D. (1998). Glucocorticoid-induced type 1/type 2 cytokine alterations in humans: a model for stress-related immune dysfunction. *Journal of Interferon and Cytokine Research* **18**, 1059–1068.
- Calcagni, E. and Elenkov, I. (2006). Stress system activity, innate and T helper cytokines, and susceptibility to immune-related diseases. *Annals of the New York Academy of Sciences* **1069**, 62–76.
- Cohen, H., Ziv, Y., Cardon, M., et al. (2006). Maladaptation to mental stress mitigated by the adaptive immune system via depletion of naturally occurring regulatory CD4+CD25+ cells. *Journal of Neurobiology* **66**, 552–563.
- Glaser, R., Kiecolt-Glaser, J. K., Marucha, P. T., et al. Stress-related changes in proinflammatory cytokine

- production in wounds. *Archives of General Psychiatry* 56, 450–456.
- Hori, T., Katafuchi, T., Take, S., et al. The autonomic nervous system as a communication channel between the brain and the immune system. *Neuroimmunomodulation* 2, 203–215.
- Ippoliti, F., Santis, W. D., Volterrani, A., et al. (2006). Psychological stress affects response to sublingual immunotherapy in asthmatic children allergic to house dust mite. *Pediatric Allergy and Immunology* 17, 337–345.
- Johnson, J. D., Campisi, J., Sharkey, C. M., et al. (2005). Catecholamines mediate stress-induced increases in peripheral and central inflammatory cytokines. *Neuroscience* 135, 1295–1307.
- Kiecolt-Glaser, J. K., Glaser, R., Gravenstein, S., et al. (1996). Chronic stress alters the immune response to influenza virus vaccine in older adults. *Proceedings of the National Academy of Sciences of the United States of America* 93, 3043–3047.
- Marshall, G. D. Jr., Agarwal, S. K., Lloyd, C., et al. (1998). Cytokine dysregulation associated with exam stress in healthy medical students. *Brain, Behavior, and Immunity* 12, 297–307.
- Marsland, A. L., Cohen, S., Rabin, B. S., et al. (2006). Trait positive affect and antibody response to hepatitis B vaccination. *Brain, Behavior, and Immunity* 20, 261–269.
- Marx, C., Ehrhart-Bornstein, M., Scherbaum, W. A., et al. (1998). Regulation of adrenocortical function by cytokines – relevance for immune-endocrine interaction. *Hormone and Metabolic Research* 30, 416–420.
- Miller, A. H. (1998). Neuroendocrine and immune system interactions in stress and depression. *Psychiatric Clinics of North America* 21, 443–463.
- Moynihan, J. A., Kruszewska, B., Brenner, G. J., et al. (1998). Neural, endocrine, and immune system interactions. Relevance for health and disease. *Advances in Experimental Medicine and Biology* 438, 541–549.
- Myint, A. M., Leonard, B. E., Steinbusch, H. W., et al. (2005). Th1, Th2, and Th3 cytokine alterations in major depression. *Journal of Affective Disorders* 88, 167–173.
- Pucak, M. L. and Kaplin, A. I. (2005). Unkind cytokines: current evidence for the potential role of cytokines in immune-mediated depression. *International Review of Psychiatry* 17, 477–483.
- Redwine, L., Mills, P. J., Sada, M., et al. (2004). Differential immune cell chemotaxis responses to acute psychological stress in Alzheimer caregivers compared to non-caregiver controls. *Psychosomatic Medicine* 66, 770–775.
- Viswanathan, K. and Dhabhar, F. S. (2005). Stress-induced enhancement of leukocyte trafficking into sites of surgery or immune activation. *Proceedings of the National Academy of Sciences of the United States of America* 102, 5808–5813.
- Viveros-Paredes, J. M., Puebla-Perez, A. M., Gutierrez-Coronado, O., et al. (2006). Dysregulation of the Th1/Th2 cytokine profile is associated with immunosuppression induced by hypothalamic-pituitary-adrenal axis activation in mice. *International Immunopharmacology* 6, 774–781.
- Wilson, C. J., Finch, C. E. and Cohen, H. J. (2002). Cytokines and cognition – the case for a head-to-toe inflammatory paradigm. *Journal of the American Geriatrics Society* 50, 2041–2056.
- Yang, E. V. and Glaser, R. (2002). Stress-associated immunomodulation and its implications for responses to vaccination. *Expert Review of Vaccines* 1, 453–459.

Cytokines, Chronic Stress, and Fatigue

S Jain and P J Mills

University of California San Diego,
San Diego, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Cytokines: Description and Classification
Cytokines: Central Nervous System and Endocrine Interactions
Cytokine Measurement
Cytokines and Chronic Stress
Cytokines and Fatigue in Clinical Disorders
Chronic Stress, Fatigue, and Cytokines: Relevance for Disease Processes
Conclusion

Glossary

Cancer-related fatigue A persistent feeling of tiredness, weakness, or lack of energy related to cancer and/or its treatment; it is more severe than general fatigue and interferes with regular functioning and quality of life.

Chronic fatigue syndrome (CFS) A complex syndrome of unknown etiology marked by persistent fatigue that lasts for over 6 months and interferes with daily functioning.

C-reactive protein (CRP) An acute-phase protein that is produced by the liver during inflammation. CRP increases in response to the cytokine interleukin-6. Elevated CRP levels are currently considered the strongest prognostic indicator of increased risk for stroke as well as for myocardial infarction and

	death from coronary events. Recent studies indicate that elevated CRP is associated with cognitive decline.
<i>Cytokines</i>	A diverse group of potent, low-molecular-weight proteins and glycoproteins that mediate important physiological processes within the immune system as well as the nervous and endocrine systems. Major types of cytokines include the interleukins (ILs), tumor necrosis factors (TNF), and interferons (IFNs).
<i>Fatigue</i>	A persistent feeling of tiredness, weakness, or lack of energy. There are many components of fatigue, including physiological (such as flulike symptoms and anemia), psychological (including depression), cognitive (including memory and attention problems), and chronological (such as circadian rhythms disorders and sleep disruption).
<i>Multiple sclerosis (MS)</i>	An autoimmune disorder characterized by the degeneration of myelin in the central nervous system, leading to muscle weakness, cognitive disorientation, and fatigue.
<i>Soluble cytokine receptors</i>	Biologically active receptors that circulate in the extracellular matrix. Soluble cytokine receptors may function as antagonists (i.e., inhibiting the effects of a cytokine by binding to it and blocking it from attaching to its cell-bound receptor) or agonists (i.e., promoting the effects of a cytokine by binding to it and forming a complex that binds to a different receptor subunit that initiates signal transduction).
<i>Vascular endothelial growth factor (VEGF)</i>	An important cytokine that is produced by the endothelial cells. It is an angiogenic cytokine, meaning it supports the formation of new blood vessels, including those for tumors.
<i>Vital exhaustion</i>	A syndrome marked by persistent fatigue, often stemming from an inability to cope with one or more chronic stressors, and implicated as an independent risk factor for negative cardiovascular events.

Cytokines: Description and Classification

Description

Cytokines are potent immunotransmitters that play a pivotal role in immune system response and communication with other physiological systems, both to maintain homeostasis and to respond appropriately to infection and injury. The functions of cytokines are diverse: They assist in the development and

proliferation of immune cell subsets, promotion of inflammatory as well as noninflammatory processes, and alteration of neurochemical and neuroendocrine processes that affect overall physiology and behavior. Cytokines may be thought of as similar to neurotransmitters and hormones in that they are mediators of specific physiological responses, rely on receptor-ligand interactions, and have self (autocrine), local (paracrine), and distal (endocrine) effects.

In addition to their multiple effects within the immune system, cytokines also have extensive bidirectional communication with the central nervous system (CNS) and hypothalamic-pituitary-adrenal (HPA) axis. Probably due to this communication, cytokines are correlated with a number of psychological states, including chronic stress and fatigue.

Classification

There are major families of cytokines. The interleukins (ILs) are a large class of cytokines that promote cell-to-cell interactions and the stimulation of humoral or cell-mediated immune responses. The tumor necrosis factors (TNFs) include a host of cytokines characterized by several molecules, such as TNF- α and TNF- β and soluble TNF receptor-I (sTNFRI) and TNF receptor-II (sTNFRII). The activation of the TNF family promotes a variety of cell functions related to inflammation as well as immune organ development and maintenance, including cell proliferation and adhesion, cell differentiation, apoptosis, and cell survival. The interferons (IFNs) play an important role immunosurveillance, antiviral, and antitumor effects. IFN- α and IFN- β inhibit virus replication in infected cells, and IFN- γ stimulates major histocompatibility complex (MHC) presentation on antigen-presenting cells, aiding in the recognition and lysing of foreign cells. In addition, IFNs initiate cascades of cytokine responses, which result in a further immune activation.

Soluble receptors of cytokines are formed either by the cleavage of portions of transmembrane protein complexes, thereby becoming part of the extracellular matrix, or by translation from alternatively spliced mRNAs. They can serve as both antagonists or agonists. Examples of soluble cytokine antagonists include IL-1 receptor antagonist (IL-1Ra) and TNFRI receptor antagonist. Examples of soluble cytokine agonists include IL-6 receptor (IL-6R). In certain situations, some cytokine receptors may function as agonists or antagonists, depending on their isoforms (e.g., TNFRII).

Because of their notable variability in structure and function, there have been many attempts to classify cytokines. A classification system that has

proven useful for stress and behavioral medicine researchers is the classification of cytokines as either pro-inflammatory or anti-inflammatory. Pro-inflammatory cytokines, which include IL-1, IL-2, IL-6, TNF- α , and IFN- γ , promote a variety of cell functions that stimulate and enhance inflammation through various methods, including promoting the differentiation of cytotoxic T cells, enhancing increased vascular permeability and cellular adhesion and migration to tissues, and stimulating the release of acute-phase proteins from the liver. These inflammatory immune responses are often described as T-helper-1 (Th1) responses, referring to the T-helper cell subset that generally produces cytokines that initiate inflammatory processes.

Anti-inflammatory cytokines, which include IL-3, IL-4, IL-5, IL-10, and IL-13, are sometimes described as immunosuppressors due to their ability to inhibit the Th1-mediated inflammatory response (often via direct antagonism of Th1-secreted inflammatory cytokines). However, these cytokines themselves promote certain increases in immune response, most notably increased overall production of antibodies and increased eosinophil and mast cell production. Often called the T-helper-2 (Th2) response, these cascades of cytokine-induced immune activation support allergic reactions. It is important to note that some cytokines support both pro- and anti-inflammatory effects depending on the situation (e.g., IL-6 and IL-8), thus rendering the nomenclature of cytokines as either pro-inflammatory or anti-inflammatory less than perfect. IL-6 is also classified as a myokine because it is produced by contracting skeletal muscle and plays an important role in the anti-inflammatory effects of acute exercise.

Cytokines: Central Nervous System and Endocrine Interactions

Cytokines and the Central Nervous System

It is well understood that cytokines are secreted by certain classes of brain cells, including microglial cells and astrocytes. The endogenous expression of cytokines and their receptors have been found in the hypothalamus, basal ganglia, cerebellum, circumventricular sites, and brain-stem nuclei. Included in the considerably large list of brain-active cytokines are IFN- α and IFN- γ ; TNF- α and TNF- β ; and IL-1, -2, -3, -4, -5, -6, -8, -10, and -12. Studies involving the systematic administration of cytokines in some of the brain regions previously mentioned indicate that cytokines promote the release of neurotransmitters, including norepinephrine, dopamine, and serotonin. Thus, in addition to their immunoprotective effects

(such as the regulation of infiltrating leukocytes during times of infection) within the brain, cytokines may promote neurochemical cascades that directly affect mood and behavior. In addition, elevated pro-inflammatory cytokines may also be associated with neurodegenerative disorders; for example, increased IL-6, TNF- α , and IL-1 β have been found in patients with Parkinson's disease.

Cytokines also affect the CNS via peripheral mechanisms. Although cytokines are too large to effectively cross the blood-brain barrier, there are several posited indirect mechanisms of action. One hypothesis is that cytokines enter the brain via passive transport in areas where the blood-brain barrier is not present (e.g., circumventricular sites). Another is that cytokines might bind to cerebral vascular endothelium, facilitating the release of active second messengers such as nitric oxide. Yet another hypothesis is that cytokines might be transported across the blood-brain-barrier via carrier-mediated transport. Finally, it has been posited that cytokines might affect the CNS indirectly via the stimulation of peripheral afferent nerve terminals, via the vagus.

Cytokines and the Hypothalamic-Pituitary-Adrenal Axis

The HPA is part of the neuroendocrine system that is responsible for the cortical as well as adrenal release of hormones in response to stress. Considerable progress has been made in understanding the complex interactions between cytokines and the HPA. Briefly, it is now well understood that complex and dynamic interactive communication exists between cytokines and the HPA and that the regulation of cytokine release, as well as HPA responses to immune insults, are governed in part by positive and negative feedback loops between the two systems. In particular, pro-inflammatory cytokines have been shown to stimulate HPA stress responses, whereas Th2 cytokines can inhibit this activation. For example, pro-inflammatory cytokines appear to activate corticotropin releasing hormone (CRH) and arginine vasopressin neurons in the parvocellular paraventricular nucleus within the hypothalamus. This activation results in a downstream HPA cascade in which CRH is released from the hypothalamus, promoting the release of adrenocorticotrophic hormone (ACTH) from the anterior pituitary gland and resulting in release of the glucocorticoids corticosterone and cortisol from the adrenal cortex. There are several postulated mechanisms of action for how this cascade is initiated by pro-inflammatory cytokines, some of which involve mediating effects of cytokines on the HPA via afferent vagal fiber activity.

In addition to their effects on the anterior pituitary and adrenal cortex via CRH release from the hypothalamus, pro-inflammatory cytokines may also affect the anterior pituitary and adrenal cortex directly, resulting in similar end-organ effects (release of corticosterone from the adrenal cortex). For example, IL-6 is synthesized and released within the human adrenal gland itself, promoting glucocorticoid release. The multitude of sites of action allows pro-inflammatory cytokines several pathways of promoting a similar end-organ response so that even if higher-level actions of cytokines on hypothalamic or anterior pituitary structures are inhibited (e.g., via antagonism by a Th2 cytokine such as IL-10), some level of glucocorticoid release into the circulation is preserved. In turn, glucocorticoid actions on cytokines help to maintain homeostasis via negative feedback loops. For example, cortisol inhibits cellular synthesis and release of pro-inflammatory cytokines, thus acting to preserve homeostasis in the system. However, the effects of glucocorticoids on maintaining this homeostasis are dampened in cases of chronic stress, possibly due to the ability of the pro-inflammatory cytokines to promote receptor desensitization, downregulation, or prevalence of negative isoforms of the glucocorticoid receptor, causing decreased glucocorticoid sensitivity or glucocorticoid resistance. Thus, cytokines are intimately intertwined with HPA responses, providing a potent influence on stress responses and appearing to play a very active role in HPA modulations during chronic stress and fatigue.

Cytokine Measurement

When we examine cytokines in biological fluids, there are two broad categories for methods of measurement. Immunoassays measure the prevalence of cytokines or their soluble receptors by using either radioisotope-tagged antibodies (i.e., radioimmunoassays, RIAs), or enzyme-linked antibodies (i.e., enzyme-linked immunosorbent assays, ELISAs) that are specific for certain peptides that are part of the cytokine structure. The strengths of immunoassays in general are their ease of use and relatively low cost, combined with a relatively high specificity due to the use of monoclonal antibodies. However, these methods are not immune to methodological problems, and they do not provide information on functionality of the cytokine. Bioassays measure cytokine functionality as indexed by specific biological responses such as chemotaxis (movement through a chemical diffusion gradient), proliferation (increase in numbers of the particular cell line), cytotoxicity (ability of cells to kill pathogens), expression of cell surface molecules, or subsequent release of specific proteins. Bioassays

thus give the researcher information not simply about soluble cytokine levels, but also about some aspect of their functionality. Although they are very sensitive tests, they are generally less specific, less reliable, and more time-consuming than immunoassays.

Other, newer methods also exist for examining whole cells' capacities to produce and release cytokines. Two such notable methods are flow cytometry and enzyme-linked immunospot (ELISPOT), which measure the abilities of single cells to produce or release cytokines, respectively. Most studies discussed in this article relied on ELISA methods for measuring basal (resting) levels of circulating cytokines in peripheral blood, although some studies examined cytokines after stimulating immune cells with a mitogen known to enhance cytokine release (e.g., lipopolysaccharide, LPS).

Cytokines and Chronic Stress

A few studies have examined the effects of chronic stressors on cytokines in humans. Most studies have been conducted with Alzheimer caregivers and the diagnosis of vital exhaustion. Interestingly, chronic stress appears to be associated with immunosuppression in Alzheimer caregivers but associated with pro-inflammatory activity in vital exhaustion.

Chronic Stress Studies with Alzheimer's Disease Caregivers

An early study examining chronic stress and wound healing in Alzheimer caregivers compared to matched controls showed a decrease in IL-1 mRNA secretion in response to LPS stimulation of peripheral blood leukocytes in Alzheimer's caregivers. Significantly, this may have contributed to slower wound healing in this group. Another study reported lower IL-1 β and IL-2 responses to virus-specific stimulation for Alzheimer caregivers than for controls. Alzheimer caregivers have also shown increased intracellular IL-10 (but no change in INF- γ or IL-2) levels in T-helper and T-cytotoxic cells compared to controls, with the difference between these groups being significantly greater for younger subjects. A recent longitudinal study, in which IL-6 levels for Alzheimer's caregivers and controls were tracked over 6 years, showed an almost fourfold rate of increase of IL-6 levels in Alzheimer's caregivers compared to matched controls. This result was consistent even for caregivers whose spouses had died during the 6-year period. Another study examining IL-6 associations with aging and chronic stress in women indicated that Alzheimer's caregivers showed significantly higher levels of IL-6 compared to age-matched women who were experiencing moderate forms of

stress (i.e., moving), as well as compared to older and younger control subjects. These findings suggest that chronic stress associated with Alzheimer's caregiving in the elderly promotes increases in Th2 cytokines and the inhibition of Th1 cytokines. The findings also suggest pathways by which caregiving in the elderly leads to significantly increased risk for deleterious health outcomes even well after the death of the spouse being cared for.

Chronic Stress Studies with Vital Exhaustion

Vital exhaustion has been implicated as an independent risk factor for negative cardiovascular events, such as first stroke and myocardial infarction. Vital exhaustion in women has been linked to elevated circulating TNF- α levels. A study comparing vitally exhausted middle-age males with controls indicated that vital exhaustion was associated with increased levels of IL-6, IL-1Ra, and IL-10 levels, as well as increased procoagulant activity. Similarly, a recent study examining vital versus nonvital exhaustion in male industrial workers showed decreased glucocorticoid sensitivity for highly exhausted men, both by dexamethasone inhibition of LPS-stimulated IL-6 release and by the median inhibitory concentration (IC₅₀; a measure of glucocorticoid sensitivity that is independent of absolute cytokine release). In addition, vitally exhausted workers showed significantly increased resting levels of CRP. Together, the findings suggest that chronic stress-related fatigue in otherwise healthy women and men may be associated with a shift toward a pro-inflammatory profile. These findings are important because they point to some of the linkages between stress-related fatigue in vital exhaustion and the risk of coronary events. Longitudinal studies in this area are needed to elucidate the direction of relationship between elevations in pro-inflammatory cytokines and vital exhaustion.

Cytokines and Fatigue in Clinical Disorders

As already indicated, cytokines have been shown to be associated with fatigue in otherwise healthy individuals. However, alterations in cytokines have also been examined in the context of several clinical disorders marked by persistent fatigue, such as cancer, chronic fatigue syndrome, and multiple sclerosis.

Cytokines and Fatigue in Cancer

Fatigue is one of the most frequent complaints of cancer patients, with studies showing 40–75% of patients reporting feeling tired and weak, and with rates up to 95% during chemotherapy and/or

radiotherapy treatment. Much of the fatigue that cancer patients experience is probably attributable to cytokines that are elevated either by the cellular damage from the cancer itself or by its treatment. Although high levels of certain endogenous cytokines (e.g., TNF- α) are associated with tumor genesis and growth, paradoxically, the exogenous administration of pro-inflammatory cytokines attenuate cancer growth and thus are often used in cancer treatment. This exogenous administration of cytokines for the treatment of cancer often elicits sickness behavior, a syndrome characterized by flulike symptoms including fatigue and depression. In addition, treatments such as chemotherapy and radiation therapy are associated with elevations in pro-inflammatory cytokines (such as IL-1Ra, IL-6, and TNF- α), which, in turn, are associated with elevations in fatigue. We have shown that circulating levels of VEGF are elevated in response to chemotherapy for breast cancer and that these elevated levels are associated with the significantly increased feelings of fatigue and poorer quality of life that result from chemotherapy. Elevations in cytokines may also persist along with fatigue well after treatment. For example, significantly higher serum levels of IL-1Ra and sTNFR_{II} have been found among breast cancer survivors who report a high level of fatigue compared to low-fatigue breast cancer patients, independent of depression.

In addition, sleep loss, a common correlate of fatigue, is associated with elevations in pro-inflammatory cytokines. For example, pro-inflammatory cytokines such as IL-6 and TNF- α are fatigue-inducing cytokines that are elevated during the day in disorders of excessive daytime sleepiness and in healthy subjects following sleep deprivation. Thus, evidence for the interaction between cytokines and cancer-related fatigue suggests that higher levels of endogenous or exogenous pro-inflammatory cytokines predict higher fatigue. However, higher levels of fatigue pretreatment may predict poorer outcomes related to treatment as well. For example, our group has found that breast cancer patients with high fatigue prior to chemotherapy experience poorer sleep in response to chemotherapy (both subjectively and objectively) compared to those with lower fatigue prior to chemotherapy. Whether these alterations in sleep patterns for high-fatigue patients are associated with alterations in cytokine responses to chemotherapy remains to be elucidated.

Cytokines and Fatigue in Multiple Sclerosis

The immune profile of multiple sclerosis (MS) appears to be characterized by self-reactive T cells, as well as an relative elevation in pro-inflammatory

or Th1 cytokines such as IL-2, TNF- α , and IFN- γ compared to anti-inflammatory cytokines such as IL-4 and IL-10. Fatigue is a common complaint in MS, with over one-third of patients reporting significant fatigue. Although the specific etiology of fatigue in these patients remains to be completely understood, a few studies have reported significant associations with fatigue and IL-1 and TNF- α in MS patients. One study has also reported positive associations with CRP and fatigue in MS patients. Several studies have also related fatigue to HPA axis dysregulation. Most studies have reported HPA axis hyperactivity in MS, with a handful of studies reporting significant associations with HPA axis hyperactivity and fatigue. Many researchers in this field suggest that a potential mechanism of fatigue in MS patients stems from elevations in pro-inflammatory cytokines (such as IL-1 and IL-6), which impair in glucocorticoid receptor signaling, driving HPA axis hyperreactivity, which leads to feelings of fatigue. However, this theory remains to be tested.

Cytokines and Fatigue in Chronic Fatigue Syndrome

The immune profile in CFS is complex and mechanisms that are linked to the particular experience of fatigue in these patients are yet to be understood. Common symptoms in CFS include decreased natural killer (NK) cell cytotoxic activity and lymphocyte proliferation, and increased allergic and autoimmune activities. Such a profile is consistent with a relative elevation in Th2 versus Th1 cellular activity; however, findings in the literature cannot yet confirm this contention. For example, findings on basal cytokine levels have been mixed; whereas several studies report elevations in IL-1 α and IL-1 β for CFS patients versus controls, a similar number of studies have reported no differences. Similar equivocal results have been reported for levels in TNF- α , TNF- β , IL-6, and IL-2 for CFS patients versus controls, and findings on HPA dysregulation are also mixed. Heterogeneity in findings is probably due to study differences in methodology as well as the ranges of patient symptoms and comorbidities studied.

With respect to fatigue ratings and cytokines in CFS, very little work has systematically looked at the relations between fatigue ratings and circulating cytokine levels. However, a recent study reports significant associations with fatigue ratings and LPS-stimulated levels of TNF- α and IL-6 compared to control subjects, who showed significant associations with fatigue and stimulated levels of IL-6 but not of TNF- α .

Cytokines and Fatigue in Clinical Disorders: Summary

Studies examining the relation of cytokines and fatigue in clinical populations are burgeoning, with more studies needed to confirm the preliminary results. However, a common thread in studies with cancer, MS, and CFS appears to be a consistent association with pro-inflammatory cytokines with reports of fatigue in these patients. However, the immunological profile in each of these disorders is quite complex, and future studies will need to be more specific in their assessment of disease stage/severity as well as other potential covariates (including depression) before these preliminary results can be confirmed.

In addition, given the preliminary findings on HPA dysregulation and fatigue, as well as the close interaction of cytokines with the HPA axis, future studies should examine in detail the interactions with potential elevations in pro-inflammatory cytokines and glucocorticoid sensitivity. The use of structural equation modeling methods combined with a careful assessment of data should shed light on proposed theories of cytokine-induced dysregulation of HPA axis activity and its relation to fatigue.

Chronic Stress, Fatigue, and Cytokines: Relevance for Disease Processes

There is a possibility that the chronic prevalence of fatigue in otherwise healthy individuals may also increase their risk for cardiovascular disorders via a cytokine-mediated response. We already mentioned that vital exhaustion is characterized by a generalized increased inflammatory profile. Not surprisingly, individuals who experience vital exhaustion also report a general lack of energy and some symptoms of depression, such as hopelessness. A relatively recent large-scale study indicated that vital exhaustion was independently associated with a twofold risk of mortality from coronary heart disease in older men. In addition, a recent study reports that vital exhaustion is a significant independent risk indicator for first stroke in middle-age men and women, implying that this state of persistent and excessive fatigue (and its negative effects) is not necessarily a consequence of existing cardiovascular disorders. Given that elevations in IL-6 as well CRP are associated with greater risk of peripheral vascular disease, myocardial infarction, and stroke and that elevations in these inflammatory markers are found in patients with vital exhaustion, it is not unreasonable to suspect that chronic elevations in inflammation associated with vital exhaustion lead over time to an increased risk

for deleterious cardiovascular events. Longitudinal studies with younger, chronically stressed men and women will help to elucidate the potential effects of persistent stress and fatigue on glucocorticoid resistance, cytokine balance, and subsequent immune response over time.

Obviously, alterations in cytokines that are associated with chronic stress and fatigue affect more than cardiovascular disease progression. Increases in pro-inflammatory cytokines associated with chronic stress and fatigue may directly exacerbate disease progression for disorders such as rheumatoid arthritis, MS, and diabetes, which are already characterized by increased pro-inflammatory activity. What remains uncertain is whether persistent fatigue in clinical populations may affect tolerance for other therapies. For example, it will be important to better understand whether patients with higher fatigue (and higher inflammation) respond more poorly to inflammation-inducing treatments such as INF treatment, chemotherapy, and radiation, both in terms of poorer physiological tolerance or response to treatment and in terms of the psychological ramifications of treatment, such as increased depression. To date, there are very few studies that have examined such potential differential psychological and physiological responses with respect to fatigue. These studies will be very important because they may identify a potential need to ameliorate fatigue in some patients either before or during treatment to enhance psychological and physiological responses to treatment.

Conclusion

Cytokines have emerged as not only important mediators of immune function but also as immunotransmitters that have widely varying and far-reaching effects on other physiological systems, including the CNS and HPA. Imbalances in cytokines and their effects on psychosocial and physiological functioning are reflected in numerous populations of relevance to chronic stress and fatigue, including (but not limited to) cancer, cardiovascular disorders, MS, and CFS.

In addition, because of their far-reaching effects, it is becoming clearer that cytokines play an important role in the psychophysiological states of chronic stress and fatigue. It is our hope that continued rigorous research in these areas will further elucidate the linkages between cytokines and these psychological states, as well as identifying effective behavioral and/or pharmacological interventions aimed at relieving the associated symptomatology.

See Also the Following Articles

Chronic Fatigue Syndrome; Cytokines; Cytokines, Stress, and Depression; Fatigue and Stress; Multiple Sclerosis.

Further Reading

- Banks, R. E. (2000). Measurement of cytokines in clinical samples using immunoassays: problems and pitfalls. *Critical Reviews in Clinical Laboratory Science* 37(2), 131–182.
- Gottschalk, M., Kumpfel, T., Flachenecker, P., et al. (2005). Fatigue and regulation of the hypothalamo-pituitary-adrenal axis in multiple sclerosis. *Archives of Neurology* 62(2), 277–280.
- John, C. D. and Buckingham, J. C. (2003). Cytokines: regulation of the hypothalamo-pituitary-adrenocortical axis. *Current Opinions in Pharmacology* 3(1), 78–84.
- Kurzrock, R. (2001). The role of cytokines in cancer-related fatigue. *Cancer* 92(6S), 1684–1688.
- Lipman, A. J. and Lawrence, D. P. (2004). The management of fatigue in cancer patients. *Oncology* 18(12), 1527–1535.
- Mills, P. J., Parker, B., Dimsdale, J. E., et al. (2005). The relationship between fatigue and quality of life and inflammation during anthracycline-based chemotherapy in breast cancer. *Biological Psychology* 69, 85–96.
- Patarca, R. (2001). Cytokines and chronic fatigue syndrome. *Annals of the New York Academy of Sciences* 933, 185–200.
- Vollmer-Conna, U., Fazou, C., Cameron, B., et al. (2004). Production of pro-inflammatory cytokines correlates with the symptoms of acute sickness behaviour in humans. *Psychology and Medicine* 34(7), 1289–1297.

Cytokines, Stress, and Depression

B E Leonard

University of Maastricht, Maastricht, The Netherlands
and National University of Ireland, Galway, Ireland

C Song

University of Prince Edward Island, Charlottetown, PE,
Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
C Song, volume 1, pp 633–638, © 2000, Elsevier Inc.

Introduction

Psychological Stress and Activation of the Endocrine
System

Stress and the Role of Corticotropin-Releasing Hormone
Receptors

Depression and the Hypothalamic-Pituitary-Adrenocortical
Axis

The Chronic Effects of Stress: Depression and
Neurotoxicity

Changes in Pro-Inflammatory Cytokines in Depression
Conclusion

Glossary

Corticotropin-releasing hormone (CRH)

Brain peptide that mediates the neural control of adrenocorticotrophic hormone (ACTH) release from the anterior pituitary gland. CRH has also been implicated as a central neurotransmitter involved in anxiety and depression.

Cytokines

These are biologically active peptides that elicit responses from the endocrine, immune, and neurotransmitter systems by acting on specific cytokine receptors situated on the cell membrane. Cytokines have a wide range of actions that include cellular differentiation, growth, stimulation of antibody synthesis, and inflammatory changes. Regarding the psychopathology of depression, **pro-inflammatory cytokines** (e.g., interleukin (IL)-1, IL-6, tumor necrosis factor (TNF)- α , and interferon (IFN)) have been shown to increase in the blood and brain where they contribute to the symptoms of depression (such as anhedonia [loss of feeling of pleasure], anorexia, sleep disturbance, and anxiety). Pro-inflammatory cytokines are also thought to increase neuronal apoptosis that could contribute to neurodegeneration. The rise in the pro-inflammatory cytokines

in depression is associated with a reduction in the **anti-inflammatory cytokines** (such as IL-4 and IL-10) that normally counteract the effects of the pro-inflammatory cytokines.

Term used by a founder of psychoendocrinology, Hans Seyle, to describe the long-term effects of stressful stimuli that result in hypertrophy of the endocrine and immune organs. Seyle showed that if the stress persists, this eventually leads to complete exhaustion of the animal and death.

Tendency towards the establishment of a stable internal environment. This is usually achieved by biological mechanisms that are activated by a negative feedback mechanism. Thus the rise in the concentration of a hormone, immune regulator, or neurotransmitter triggers an inhibitory mechanism that slows the further release of active molecule.

Presence of an excess of adrenocortical hormones (usually cortisol in humans or corticosterone in nonprimates) frequently found in major depression.

Term used to describe the different types of molecules that accumulate in tissues as part of the inflammatory process. In depression, the rise in pro-inflammatory cytokines triggers a cascade of changes that result in the rise in prostaglandin E2 and nitric oxide. Together with the glucocorticoids and pro-inflammatory cytokines, these inflammatory mediators are thought to contribute to the increase in neuronal apoptosis and possible neurodegenerative changes associated with chronic depression.

This is the mechanism whereby the increase in, for example, circulating glucocorticoid hormones activates the central glucocorticoid receptors that then inhibit the further release of corticotropin-releasing hormone. This reduces the release of adrenocorticotrophic hormone from the anterior pituitary gland thereby reducing the further synthesis and release of glucocorticoids from the adrenal gland.

Term used to describe the proliferation of new neurons in the brain. Experimental evidence shows that stress inhibits the development of new pyramidal cells in the hippocampus due to the reduction in availability of neurotrophic factors such as brain-derived neurotrophic factor

General adaptation syndrome

Homeostasis

Hyper-cortisolemia

Inflammatory mediators

Negative feedback mechanism

Neurogenesis

(BDNF). Chronic antidepressant treatment has been shown to reverse the effect of stress on the neurogenesis of cells in the hippocampus.

Introduction

The causal relationship between negative emotions and psychological distress on physical health has been confirmed in recent years by epidemiological studies in which an increase in the prevalence of coronary heart disease, osteoporosis, diabetes, and dementia have been associated with chronic stressful life events. Research has largely focused on changes in neurotransmitter function, as well as the endocrine and immune systems, in explaining the basis of the causal relationship between stress and ill-health as these systems provide the main pathways of communication between the brain and the peripheral organs. The anatomical links between the brain, endocrine system, and peripheral nervous system is illustrated in Figure 1.

Psychological Stress and Activation of the Endocrine System

It is well established that, in the face of any threat, either real or perceived, an organism mounts a series of co-ordinated and specific hormonal, autonomic, immune, and behavioral responses that allow it to fight, or escape from, the threat. Normally, the characteristics and intensity of the response to the stress matches that posed by the threat and the duration of

the response is normally no longer than is necessary to achieve a successful outcome. If the response to the stress is inadequate or excessive then it could give rise to pathological changes.

The scientific basis of adaptation to stress was suggested by Walter Cannon who, in the early part of the twentieth century, introduced the concept of homeostasis. This involved the initiation of the fight or flight response whereby changes in the sympatho-adrenomedullary system co-ordinated the physiological aspects underlying homeostasis. This concept was developed further by Hans Seyle (1936) who stressed the importance of the hypothalamic-pituitary-adrenocortical (HPA) axis in the stress response. Seyle described the general alarm reaction as an early response to a noxious stimulus that was characterized by a nonspecific activation of the endocrine and the sympatho-medullary system. Continued exposure to the same stressful stimulus was proposed to lead to long-term changes in the endocrine and immune systems. These were characterized by hypertrophy of the pituitary and adrenal glands. Seyle termed this chronic effect of stress the general adaptation syndrome. An organism could recover from the changes once the stressful stimulus was terminated by exhaustion, but death could occur if the stress continued.

The concept of stress being physically and functionally damaging was initially thought to be due to the hypersecretion of glucocorticoids, a concept that was supported by the observation that various diseases of the adrenals and pituitary were associated with psychopathological changes that resolved when the HPA axis was normalized. However, it is now

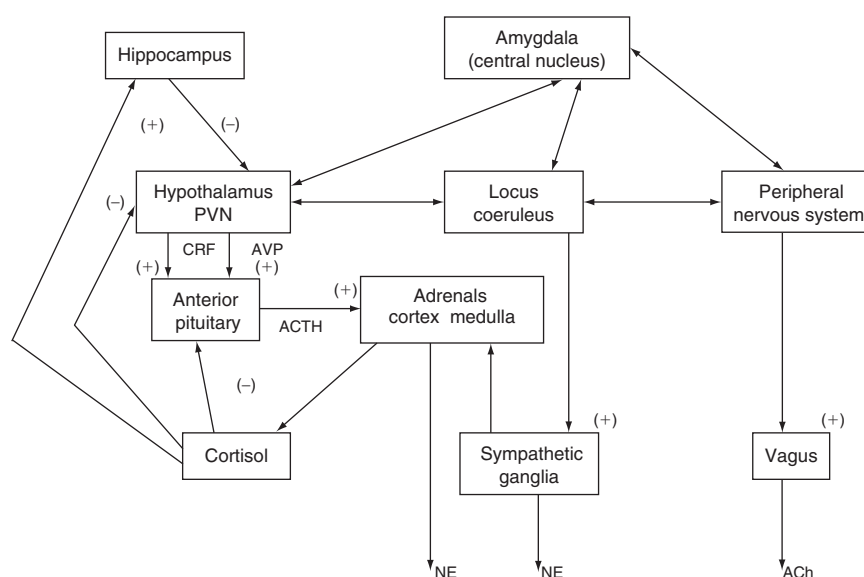


Figure 1 The relationship between stress and neuroendocrine system. (+), positive feedback; (-), negative feedback; ACh, acetylcholine; ACTH, adrenocorticotrophic hormone; AVP, arginine vasopressin; NE, norepinephrine; PVN, paraventricular nucleus.

accepted that the stress response is often beneficial in protecting the organism from harmful situations and that adaptation, a learned response, is important in adjusting to future adverse situations. Nevertheless, problems arise when stress becomes repetitive and sustained, and adaptation becomes difficult. This forms the basis of stress-induced ill-health that can result in hypertension, hypercortisolism, abnormal immune function, and ultimately mental ill-health, particularly depression.

Stress and the Role of Corticotropin-Releasing Hormone Receptors

Depression may be interpreted as the result of the sustained hyperactivity of the stress axis that results from the overactivity of the corticotropin-releasing hormone (CRH) pathway, an effect that is correlated with the hyperactivity of the central and peripheral sympatho-adrenal system and the secretion of the glucocorticoids. In addition to the psychological effects of maladaptation to stress, physical damage may also occur as indicated by a loss of bone mass, hippocampal atrophy, and neurodegeneration.

CRH produces its effects by activating CRH-1 and CRH-2 receptors. It is now evident that the fight or flight response involves activation of CRH-1 receptors whereas the much slower response system that promotes recovery and adaption to stress, is activated by urocortin acting on CRH-2 receptors.

The fight or flight response, the fast response, occurs when CRH activates CRH-1 receptors on the paraventricular nucleus (PVN) of the hypothalamus, together with noradrenergic neurons of the locus coeruleus. CRH-1 receptors also occur on the corticotrophs in the anterior pituitary which, on activation, results in the release of the adrenocorticotrophic hormone (ACTH). This is followed by the activation of the adrenal cortex by ACTH that leads to the synthesis and release of glucocorticoids.

Epinephrine and norepinephrine are released from the adrenal medulla in response to the stress-induced increase in the activity of the locus coeruleus. The fast response is particularly sensitive to the negative feedback activation of the mineralocorticoid receptors, under conditions of normal stress, or glucocorticoid receptors, under conditions of high stress.

It is now apparent that the slow stress response system is driven by the peptides urocortins 2 and 3, two peptides that are distinct from the CRH family of peptides. The urocortins have been identified as selective, high-affinity ligands for CRH-2 receptors. They are synthesized in the hypothalamus and amygdala and their terminals are located in the hypothalamus and brain stem where the CRH-2 receptors

are mainly located. The CRH-2 receptors are thought to facilitate coping and behavioral adaptation. Further, the central administration of these peptides evokes anxiolytic-like responses whereas CRH causes anxiogenic effects in rodents.

The stress response to adverse stimuli is driven not only by CRH but also by arginine vasopressin (AVP), a peptide that is synthesized in a different region of the PVN of the hypothalamus to CRH. Stress increases the proportion of cells that synthesizes AVP. The secretion of CRH is usually accompanied by AVP but whereas AVP usually has a low efficacy for stimulating the release of ACTH compared to CRH, AVP potentiates the effects of CRH thereby enhancing ACTH release. Chronic stress desensitizes the CRH-1 receptors in the pituitary gland but, under these circumstances, the AVP receptors are sensitized thereby sustaining the release of ACTH and the activation of the adrenal cortex.

In addition to its role in stimulating the HPA axis and the central sympathetic system, CRH neurons are activated by different neurotransmitters from several brain regions. Excitatory inputs to the CRH pathways include serotonin, norepinephrine and peptides such as neuropeptide-Y, while the inhibitory inputs include γ -aminobutyric acid (GABA) and the opioids.

Different types of stressful stimuli activate the CRH pathway in different ways. Thus painful stimuli (infections, inflammation, toxic agents, hypoglycemia, and hypo- and hypertension) activate monosynaptic pathways to the CRH neurons. By contrast, stressful, psychological stimuli activate neurons of the limbic-cortical region that subserve nociception, emotional, and cognitive functions. It now appears that these emotional circuits in the central amygdala, frontal cortex, and hippocampus project to a GABAergic network that surrounds the PVN. The effect of this GABAergic system on the function of the hypothalamus varies and can be either disinhibitory to CRH release, when the input is from the amygdala, or inhibitory when the input is from the hippocampus. This pathway functions to process cognitive and emotional aspects of a stressful situation to be assessed.

Depression and the Hypothalamic-Pituitary-Adrenocortical Axis

Some 50 years ago, Board and colleagues (1956) first reported that cortisol was elevated in the plasma of depressed patients but it was Sachar and coworkers (1970) who showed that the hypersecretion of cortisol was accompanied by an altered circadian pattern. They also showed that the frequency, magnitude, and duration of the secretory bursts of cortisol

Table 1 Changes reported to occur in the hypothalamic-pituitary-adrenal axis of patients with major depression

<i>Short-term stress: acute exposure</i>	<i>Pathophysiological effects: chronic stress</i>
Energy mobilization	Myopathy, fatigue, steroid-induced diabetes
Control of stress reactions	Hypertension
Increased vascular tone	Osteoporosis
Suppression of growth	Amenorrhea, infertility
Suppression of fertility	Impaired immune function
Anti-inflammatory effects	Affective disorders
	Neurotoxicity

secretion were elevated. More recently, it has been shown that a high morning salivary cortisol concentration is a possible risk factor for depression. The possible link between the hypersecretion of cortisol and the symptoms of depression is also supported by the observation that, therapeutically, glucocorticoids can precipitate depression even in those who were not apparently vulnerable to the condition. Conversely, low doses of glucocorticoids are administered to rodents have rewarding effects, changes that have been attributed to the stimulation of the mesencephalic dopaminergic system. It is possible that this system forms part of the coping strategy to stress whereas chronic, high glucocorticoid concentrations are causally responsible for provoking both anxiety and depression. [Table 1](#) summarizes the major changes that have been reported in the HPA axis of patients with major depression.

The Chronic Effects of Stress: Depression and Neurotoxicity

Several studies have reported that there is a reduction in the volumes and metabolic activity of the hippocampus, amygdala, and pre-frontal cortex in depressed patients. Such changes could be associated with the cause of depression, or could be a consequence of depression. Whatever the relationship, it is apparent that the shrinkage of the hippocampus could account for the inability of the patient to readily adapt to stressful situations. Whether the decrease in the hippocampal volume is attributable to hypercortisolemia, as is frequently stated, is not certain as patients with posttraumatic stress disorder also have a reduced hippocampal volume but frequently suffer from hypocortisolemia.

The connection between the shrinkage of some of the limbic regions and neuronal atrophy has also been the subject of debate. For example, some investigators have reported no signs of increasing neuronal death in the hippocampus of depressed patients, and

while there is evidence that the gray matter and glia numbers are reduced in the prefrontal cortex of depressed patients, there does not appear to be an equivalent loss of neurons. By contrast, animal studies clearly indicate that chronic stress has a profound effect on the structure of hippocampal neurons, changes that are effected through the actions of the glucocorticoids. For example, the dendrites of the pyramidal cells in the CA3 region of the hippocampus retract after an acute stress and become irreversibly damaged if the stress becomes chronic. Other experimental studies have shown that the CA1 neurons of the hippocampus are particularly vulnerable to neurotoxic damage caused by glucocorticoids. Glucocorticoids are known to block glucose uptake by the neurons thereby initiating the excitotoxic cascade of changes caused by the excessive activation of *N*-methyl-D-aspartate (NMDA) receptors. Furthermore, glucocorticoids block the reuptake of glutamate into glia and thereby enhance the concentration of the excitatory amino acid in the vicinity of the neurons. In addition for these changes that could account for glucocorticoid induced neurotoxicity, there is also evidence that stress, and glucocorticoids, suppress neurogenesis. Neurogenesis is enhanced by neurotrophic factors such as brain-derived neurotrophic factor (BDNF) and by the activation of 5-HT_{1A} receptors.

In rodent studies, stress has been shown to increase the turnover of serotonin in the prefrontal cortex, nucleus accumbens, amygdala, and lateral hypothalamus. These effects of stress are mediated through the glucocorticoids, which have major effects on the expression of both the 5-HT_{1A} and 5-HT_{2A} receptors. It is known that the 5-HT_{1A} receptors are under the tonic inhibition by the adrenal steroids, by activation of the type 1 glucocorticoid receptors. Conversely, 5-HT_{2A} receptor expression is increased following chronic stress, or by exogenous glucocorticoid administration, and decreased in response to adrenalectomy. Thus the 5-HT_{1A} and 5-HT_{2A} receptors appear to play an antagonistic role in stress which may help to explain why elevated basal cortisol concentrations are attenuated in depressed patients by the 5-HT_{1A} agonist ipsapirone. These results suggest that the dysfunctional HPA axis in depression results in a glucocorticoid subsensitivity of the postsynaptic 5-HT_{1A} receptors. The link between the serotonergic system, stress, and the effect of stress has been the subject of a recent review.

The results of these experimental studies support the hypothesis that stress and glucocorticoids reduce neurotrophic factor synthesis in the hippocampus, a process that contributes to hippocampal damage. There is evidence that the cellular actions of BDNF

and serotonin are interlinked and a reduction in BDNF reduces the density of 5-HT_{1A} receptors and increases that of 5-HT_{2A} receptors in the frontal cortex. This results in increased aggressiveness in BDNF-deficient rodents, a similar situation to that occurring following the drug-induced depletion of serotonin. These changes probably contribute to the reduction in neurogenesis thereby resulting in a reduction in the size of the hippocampus that is reported to occur in patients with chronic depression. Conversely, chronic antidepressant treatment has been shown to increase neurogenesis; this correlates with the attenuation of the symptoms of depression.

It is still uncertain from the results of these studies whether the shrinking of the hippocampus in patients with chronic depression is the cause or the result of stress. If a smaller than normal hippocampus is genetically determined, as seems possible from the results of primate studies, then it seems possible that this could contribute to an increased vulnerability to the effects of stress. For example, this could result in a decreased ability of the patient to cope with stress. It is also possible that while there may not be gross changes in the cytoskeleton, there could be changes in the synaptic plasticity that would not be evident in the imaging studies on patients with depression. Thus, while it is presently uncertain that chronic stress results in increased neuronal apoptosis, partly because of the limited sensitivity of the imaging methods available, experimental studies clearly indicate that persistent high concentrations of glucocorticoids cause atrophy of the hippocampus, prefrontal cortex, amygdala, and other brain regions that are considered to be anatomically associated with depression. The changes in the monoamine neurotransmitters in the brain, which form the basis of the amine hypothesis of depression, are therefore postulated to occur as a consequence of the primary changes in the HPA axis and immune system.

If the vulnerability of the adverse effects of stress is a causative factor in the etiology of depression, it would be anticipated that drugs activating the central glucocorticoid receptors would have antidepressant activity. There is some clinical evidence to support this view. Thus DeBattista and colleagues have reported that the intravenous administration of a high dose of cortisol (15 mg) produces an improvement in the symptoms of depression while in recent controlled studies the glucocorticoid antagonist mifepristone has been shown to have a significant antidepressant effect in patients with psychotic depression. This suggests that the glucocorticoid antagonists block the action of endogenous cortisol on glucocorticoid receptors in afferent pathways. Presumably a very high bolus dose of cortisol acts by desensitizing

these receptors. From the results of the experimental and clinical studies, it may be concluded that the overactivity of the CRH system plays a crucial role in the psychophysiology of depression. This could result in long-term structural and functional changes in neurons in the limbic system that may contribute to increased apoptosis. This raises the important question whether the neurodegenerative changes that occur in depression could contribute to the onset of dementia in vulnerable individuals. Hypercortisolemia is a common feature of chronic depression and some types of dementia while qualitatively similar changes have been reported to occur in patients with depression, and also in those suffering from dementia.

Changes in Pro-Inflammatory Cytokines in Depression

The evidence implicating a role for the pro-inflammatory cytokines in the etiology of depression is provided by studies on the changes in the concentrations of interleukin (IL)-1, IL-6, and tumor necrosis factor (TNF)- α in the plasma of depressed patients. Additional evidence is provided by the effects of interferon (IFN)- α on nondepressed patients being treated for hepatitis or malignancy. The results of such studies imply that the pro-inflammatory cytokines are causative factors in the symptoms of major depression. These symptoms include depressed mood, anxiety, cognitive impairment, lack of motivation, loss of libido, sleep disturbance, and deficits in short-term memory. These symptoms usually disappear once the plasma cytokine concentrations normalize. The changes appear to be a consequence of the neurotransmitter and endocrine changes induced by the cytokines rather than the pathological condition for which the treatment has been given. It is perhaps not surprising therefore to find that the symptoms of depression frequently occur in patients who are recovering from a chronic infection, those with multiple sclerosis, allergies, and rheumatoid arthritis. In all these situations, pro-inflammatory cytokines are known to be overexpressed. The initial studies linked depression with an abnormality of the immune system, such as impaired mitogen-stimulated lymphocyte proliferation and reduced natural killer cell activity in untreated depressed patients. These immune changes largely returned to normal once the patient has recovered from the depressive episode.

Some research into the immune changes that occur in depression has concentrated on pro- and anti-inflammatory cytokines, soluble cytokine receptors, and acute-phase proteins. For example, positive acute-phase proteins have been shown to increase,

while the negative acute-phase proteins decrease, in depression. These changes are known to be a consequence of the stimulant action of IL-6 on the liver. In addition, Song et al. (1994) showed that complement proteins (C3, C4) and immunoglobulin M are increased in depressed patients. Such changes are evidence that immune activation involves both the inflammatory cytokines and the B cells that are activated by the pro-inflammatory cytokines. Further evidence of immune activation in depressed patients is provided by the studies showing that the plasma concentrations of IL-1, IL-6, IFN, soluble IL-6, and soluble IL-2 receptors, and the soluble IL-1 receptor antagonist, are raised. These changes are correlated with a rise in acute-phase proteins. Effective antidepressant treatments attenuate these immune changes. In addition to the changes in the pro-inflammatory cytokines, there is also evidence of an increased number of T helper, T memory, activated T cells, and B cells that act as a source of plasma cytokines. From these changes, it would appear that in depression there is an imbalance between the inflammatory and anti-inflammatory arms of the immune system, the cytokines from the inflammatory Th-1 pathway (such as IFN- γ) becoming predominant over those of the anti-inflammatory Th-2 pathway (such as IL-10). A recent study has shown that the Th-3 cytokine, transforming growth factor (TGF)- β 1, whose function is to re-establish the balance between the Th1 and Th2 pathways, is increased in depressed patients following effective antidepressant treatment. Although TGF is reported to be a regulatory cytokine that maintains the balance between the pro- and anti-inflammatory cytokines, precisely how the changes are established is presently unclear. It is of interest that many clinical and experimental studies show that the chronic administration of antidepressants normalize the immune changes thereby suggesting that antidepressants have an anti-inflammatory action. This suggests that antidepressants may modulate monoaminergic function indirectly as a consequence of their actions on the disturbed immune system and the HPA axis.

Conclusion

Chronic stress, by initiating changes in the HPA axis and the immune system, acts as a trigger for depression. There is experimental and clinical evidence that the rise in the concentrations of pro-inflammatory cytokines and cortisol, which occurs following chronic stress and in depression, contribute to the neurodegenerative changes in the hippocampus, prefrontal cortex, and amygdala. The structural changes in the brain could contribute to the chronicity of

depression and be a prelude to dementia. Effective antidepressant treatment attenuates the endocrine and immune changes thereby suggesting that antidepressants have central anti-inflammatory actions that may contribute to their therapeutic efficacy.

See Also the Following Article

Depression, Immunological Aspects.

Further Reading

- Board, E., Persky, H. and Hamburg, D. A. (1956). Psychological stress and endocrine function. *Psychosomatic Medicine* 18, 324–333.
- Brambilla, F. (2000). Psychoneuroendocrinology: research on the pituitary-adrenal-cortical system. *Psychosomatic Medicine* 30, 576–607.
- Bremner, J. D., Narayan, M., Andersoin, E. R., et al. (2000). Hippocampal volume reduction in major depression. *American Journal of Psychiatry* 157, 115–118.
- Chrousos, G. B. (1998). Stressors, stress and neuroendocrine integration of the adaptive response. *Annals of the New York Academy of Sciences* 750, 311–335.
- Chrousos, G. B. and Gold, P. W. (1998). A healthy body and a healthy mind – and vice versa; the damaging power of uncontrollable stress. *Journal of Clinical Endocrinology and Metabolism* 83, 1842–1845.
- Connor, T. and Leonard, B. E. (1998). Depression, stress and immunological activation: the role of cytokines in depressive disorders. *Life Sciences* 62, 583–606.
- DeBattista, C., Posener, J. A., Kalehzan, B. M. and Schatzberg, A. E. (2000). Acute antidepressant effects of intravenous hydrocortisone and CRH in depressed patients: a double-blind, placebo controlled study. *American Journal of Psychiatry* 150, 656–657.
- Duman, R. S., Malberg, J., Nakgawa, S. D., et al. (2000). Neural plasticity in mood disorders. *Biological Psychiatry* 48, 732–739.
- Fuchs, K. C. F. and Gould, E. (2000). Mini-review: *in vivo* neurogenesis in the adult brain: regulation and functional implications. *European Journal of Neuroscience* 12, 2211–2214.
- Leonard, B. E. (2005). The HPA and immune axes in stress: the involvement of the serotonergic system. *European Child and Adolescent Psychiatry* 20(suppl.3), S302–S306.
- Maes, M., Smith, R. and Scharpe, S. (1995). The monocyte T-lymphocyte hypothesis of major depression. *Psychoneuroendocrinology* 20, 111–116.
- McEwan, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* 338, 171–179.
- Myint, A-M., Kim, Y-K., Verkirck, R., et al. (2005). Th1, Th2, and Th3 cytokine alterations in major depression. *Journal of Affective Disorders* 88, 169–173.
- Reul, J. M. H. and Holsboer, F. (2002). CRF receptors 1 and 2 in anxiety and depression. *Current Opinion in Pharmacology* 2, 23–33.

- Sachar, E., Hellman L., Fukushima, D. A. and Gallagher, T.F. (1970). Cortisol production in depressive illness. *Archives of the General and Psychiatrists* 23, 289–298.
- Sapolsky, R. M. (1996). Why stress is bad for your brain. *Science* 273, 749–750.
- Selye, H. (1936). Syndrome produced by diverse noxious agents. *Nature* 138, 32.
- Smith, R. (1991). The macrophage theory of depression. *Medical Hypotheses* 35, 298–306.
- Song, C. and Leonard, B. E. (1995). The effect of olfactory bulbectomy in the rat, alone and in combination with antidepressants and endogenous factors, on immune function. *Human Psychopharmacology* 10, 7–18.
- Song, C., Earley, B. and Leonard, B. E. (1995). Behavioral, neurochemical and immunological response to CRF administration. *Annals of the New York Academy of Sciences* 771, 55–72.
- Watanabe, Y., Sakai, R. R. and McEwan, B. S. (1993). Stress and antidepressant effects on hippocampal and cortical 5HT1A and 5HT2A receptors and transport sites for serotonin. *Brain Research* 615, 87–94.

Cytotoxic Lymphocytes

M A Fletcher and N G Klimas

Miller School of Medicine, University of Miami and the Veterans Administration Medical Center, Miami, FL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M A Fletcher, N G Klimas, and R Patarca, volume 1, pp 639–644, © 2000, Elsevier Inc.

Types of Cytotoxic Effector Cells

Mechanisms of Target Cell Killing

Effects of Stress Hormones on Cytotoxic Lymphocytes

Effects of Stress on Cytolytic Lymphocytes

Stress, Cytotoxic Lymphocytes, and Human Diseases

Glossary

<i>Cytokines</i> (<i>lymphokines</i>)	Peptides or glycopeptides that act in a manner analogous to hormones to regulate the immune system.
<i>Cytolytic</i> (<i>cytotoxic</i>) <i>lymphocytes</i>	White blood cells that have the capacity, by means of direct cell-to-cell contact, to cause the lysis and subsequent death of target cells.
<i>Large granular lymphocytes</i> (<i>LGLs</i>)	Lymphocytes that are larger and contain more intracellular granules than most lymphocytes. These lymphocytes, called natural killer (NK) cells, can form conjugates with and lyse target cells, such as tumor cells or virally infected cells, without the mediation of antigen recognition. Association of an antibody against the target cell with such LGLs through Fc receptor interaction with the Fc portion of the antibody can result in a type of cytotoxicity called

antibody-dependent cellular cytotoxicity (ADCC). Exposure of the LGL to activating cytokines, such as interleukin-2, results in a more potent cytotoxic cell called the lymphokine-activated killer cell (LAK).

T lymphocytes

White blood cells that have on their surface membranes an antigen recognition structure, the T cell receptor. Most cytolytic T lymphocytes are CD8 T cells that can interact with the major histocompatibility complex (MHC) peptides on the surface of target cells.

Types of Cytotoxic Effector Cells

Lymphocyte-mediated cytotoxicity is a form of cellular immunity whose existence was first documented in experiments on transplant rejection. In 1960, Andre Govaerts showed that lymphocytes from dogs that had received and rejected a kidney transplant destroyed cultured kidney cells from the donor dog *in vitro*. Govaerts' experiments provided the first proof for the existence of cellular immunity. In 1967, K. Theodor Brunner, Albert A. Nordin, and Jean-Charles Cerottini developed the radioactive chromium (^{51}Cr)-release assay as a quantitative tool to assess lymphocyte-associated cytolysis. Radioactive chromium binds to intracellular proteins and is used to label target cells. Release of intracellular proteins occurs relatively quickly upon lysis of target cells, and the amount of released ^{51}Cr is proportional to the amount of lysed cells. Use of this assay provided proof for the existence of cytotoxic T lymphocytes (CTLs) that, without antibodies or complement, specifically kill the target cells toward which they had been sensitized through immunization or transplantation.

Moreover, the cytolytic activity measured with this assay reflected changes in the frequency, rather than in the lytic capacity, of individual effector CTLs.

The ^{51}Cr -release assay also allowed the discovery of other types of lymphocyte-mediated cytolytic activities. The second model of cytolysis, defined by Peter Perlmann and Goran Holm in 1968, required that the target cell first react with antibodies to activate the cytolytic mechanism. This antibody-dependent cellular cytotoxicity (ADCC) was nonspecific, was present in nonimmunized individuals, and lacked memory, i.e., the cytolytic activity did not increase with immunization. Cells that mediate ADCC were termed killer (K) cells and were shown to inhibit tumor growth upon activation with cytokines or biological response modifiers (lymphokine-activated killer [LAK] cells).

A third type of lymphocyte-mediated cytolysis was discovered by several immunologists, including Ronald B. Herberman, Rolf Kiessling, and George Klein. They were looking for tumor-specific cellular immunity, but found that lymphocytes from both cancer patients and normal controls could lyse tumor cells. These lymphocytes were called natural killer (NK) cells since their activity against tumor-transformed and virus-infected cells occurred naturally without immunization. Most NK cells are large granular lymphocytes (LGLs) and are constitutively cytotoxic.

Cytolytic mechanisms involving nonlymphocytic cells, such as monocytes through their expression of tumor necrosis factor (TNF)- α or nitric oxide, are not discussed in this article but can be assessed using the same approaches.

Mechanisms of Target Cell Killing

Lymphocyte-mediated cytolysis can be conceptually divided into three distinct stages: conjugate formation (binding of effector to target), triggering (signal transduction), and lethal hit (granule exocytosis). Studies of lytic (killers), nonlytic conjugate-forming (binders), and non-conjugate-forming (free) cells suggest that the ability to perform the different stages of cytotoxicity may be acquired at different points in cell maturation, with lytic function preceding responsiveness to target cell-specific triggering. Once triggered, the cytolytic process may propagate as effector cells detach from lysed targets and recycle to initiate new lytic interactions.

Direct membrane-to-membrane contact between cytolytic and target cells (conjugate formation) is a prerequisite for lysis to occur. Cell membrane receptors and adhesion molecules mediate conjugation,

and transmission electron microscopy reveals extensive cellular membrane interdigitation at the contact region, forming an immunological synapse. In the case of CTLs, target cell adhesion results from T cell receptor–major histocompatibility complex (MHC)–peptide interactions and adhesion strengthening steps, although several other molecules mediate binding of the different types of lymphocytolytic effectors to target cells. In the case of NK cells, effector function involves receptors and ligands on the surface of the effector cells and their targets, but does not result from antigen recognition. Following the effector cell–target cell direct contact, the intracellular cytotoxic granules move to the immunological synapse. Perforin from these intracellular vesicles facilitates the release of the granzymes and the passage of these serine protease through target cell membranes by a process involving endocytosis, and perhaps pore formation. The most abundant granzymes are granzyme A and B. They induce cell death through alternate and nonoverlapping pathways. Granzymes initiate cell death mechanisms that operate through the activation of apoptotic cysteine proteases (caspases), but can also cause cell death in the absence of activated caspases. The second pathway involves the engagement of the so-called death receptors such as Fas/CD95 on target cells by their ligands (FasLs) on cytolytic cells, resulting in classical caspase-dependent DNA fragmentation and apoptosis. Granule exocytosis also requires extracellular calcium for the regulated secretion of lytic granule components, binding of the secreted perforin to the membrane of the target cell, and the polymerization of perforin to polyperforin. The antigen-specific degranulation that occurs in CTLs leads to the nascent appearance on the lymphocyte surface of a granule-associated protein, CD107. Detection of this marker using flow cytometry forms the basis of a new type of assay for functional assessment of CTLs.

Although perforin lyses a variety of target cells nonspecifically, cytolytic cells are generally more resistant to the lytic effect of perforin. The protective mechanism is insufficiently characterized at present. However, in some cases a phenomenon known as fratricide can be elicited by T cells presenting antigens to other CTLs, resulting in perforin-mediated cytolysis. These results may be important in the regulation of CD8 $^{+}$ CTLs in immune responses.

Effects of Stress Hormones on Cytotoxic Lymphocytes

Elevations in cortisol, norepinephrine (NE), and epinephrine (E) may be accompanied by decrements in

immune function. Adrenal cortical hormones may directly impair or modify several components of cellular immunity, including the cytolytic activity of T lymphocytes and NK cells. Corticosteroids inhibit cellular responses to antigens and impair NK cell activity. Corticosteroids may communicate with lymphocytes via transcriptional cytoplasmic receptors. Hence, the literature supports the concept that elevated levels of cortisol are associated with impaired immune system functioning with accompanying depression of cytokine production. Sympathetic noradrenergic fibers innervate both the vasculature and parenchymal regions of several lymphoid organs. Elevations in peripheral catecholamines may depress immune functioning. This interaction is likely mediated by β -adrenergic receptors on lymphocytes. Neuropeptide Y (NPY) also is stored in sympathetic nerve terminals and is released along with catecholamines during stress-induced activation. It is known to depress NK cell cytotoxicity *in vitro*. In contrast, pituitary and adrenal peptide stress hormones such as met-enkephalin, β -endorphin, and substance P stimulate T cell and NK cell responses. Lymphocytes have receptors for NPY and for many other neurotransmitter/neurohormones, including serotonin, cholinergic agonists, and β -adrenergic agonists.

Effects of Stress on Cytolytic Lymphocytes

Biological, psychological, and social factors that are involved in physiological stress responses and capable of immunomodulatory effects are discussed in the following sections.

Biophysical Stimuli

Several biophysical stimuli, including tobacco, ethanol, and recreational drug usage, are associated with immunomodulation. Smoking tobacco is reported to be associated with significant decreases in cytotoxic T cells. Ethanol use is correlated with depressed cell-mediated immunity and diminished NK cytotoxicity. Intravenous drug usage is associated with depressed cellular immune functioning. The specific mechanisms by which these substances affect immune functioning have not been fully elucidated, but some of these substances are associated with alterations in those physiological stress response systems noted previously. For instance, nicotine in cigarette smoke is associated with catecholamine discharge, and ethanol consumption is also linked with catecholamine elevations due to blockage of reuptake.

Psychological Factors

Acute psychological stress influences cellular immune functions. Data support the hypothesis that altered trafficking of lymphocytes with cytolytic potential occurs as a component of the fight-or-flight response to an acute stressor. Depending on the time frame of measurements, early increases in circulating CD8 T cells and NK cells can be seen, along with an increase in the ability of NK cells to kill target cells. Later, however, the per cell lytic activity decreases. For example, in a field study of first-time tandem parachutists, immediately after jumping from an airplane, elevation in noradrenalin levels was observed that correlated with elevated levels of peripheral blood CD8 T cells and NK cells and with an increase in functional capacity of NK cells. Within 1 h, however, the cytotoxic cell numbers and function decreased significantly below starting values. Similar results are observed in laboratory studies using techniques such as mental arithmetic, speech stressors, or discussions of marital discord. Such data support the hypothesis that altered trafficking of lymphocytes with cytolytic potential occurs as a component of the fight-or-flight response to an acute stressor. The resultant *in vivo* exposure of immune cells to neuropeptides, hormones, and cytokines leads to altered functional responses in CD8 T cells and NK cells.

Clinical depression or depressed affect is linked with impaired cellular immune functioning, although the literature regarding this relationship is controversial. Depression is associated with neuroendocrine abnormalities, including hypercortisolemia and elevated NPY. The impact of depression on immunity may be stronger in older subjects and in those with more severe depression. Severity of depressive symptomatology was associated with decreased NK activity. Depressive symptoms were predictive of decrements in NK activity among women when measured before and after the death of their husbands. Sleep deprivation, a condition often associated with depression, leads to reduced NK cytotoxicity.

Intrusive thoughts to an adverse situation or event are related to changes in immunological parameters. In the months following Hurricane Andrew, there was a lower level of NK cytotoxicity in those individuals of a community sample who had intrusive cognitions regarding the storm. There are several reports of reduction in immune function in patients with major depressive disorders and in persons undergoing severe life stress. There is an association between increased sympathetic nervous system activity and reduced NK cell activity in depression and Alzheimer caregiver stress. Plasma concentrations of NPY in

depressed patients were elevated compared to controls. In both depressed patients and in Alzheimer spousal caregivers, circulating concentrations of NPY, but not catecholamines, were inversely correlated with NK activity.

Social Variables

Social stressors are associated with elevations in stress hormone levels and impaired immune functioning in animals and humans. Social stressors most commonly associated with immunomodulation of cytotoxic activity in studies of humans include loneliness, marital disruption, job loss, family illness, and bereavement. Urinary excretions of NE and E are elevated in bereaved subjects and in subjects threatened with a loss as compared to normals. Animals subjected to uncontrollable stressors had immune system decrements such as decreased NK cell cytotoxicity. The experience of sustained stressful periods (e.g., studying for medical school examinations) by normal individuals is associated with poorer control of latent viral infections, impaired lymphocyte production of cytokine, diminished T cell killing of virally infected cells, and decreases in NK cell number and lytic activity. It is suggested that perception and utilization of social support buffers the effects of self-preoccupying helplessness in the face of elevated life stress and that this provides the opportunity for task-oriented thinking and active coping. Thus, disengagement from active, adaptive coping, possibly accompanying each of the previously noted stressors, may lead to self-preoccupying helplessness and consequent physiological arousal. Such a sequence of events could explain the catecholamine and glucocorticoid elevations and impaired cellular immune functioning repeatedly associated with these conditions. In a series of behavioral immunology studies in which social support, mood, and NK activity were evaluated among breast cancer patients, a lack of social support predicted poorer NK activity. A longitudinal study of homosexual men noted that NK cytotoxicity was decreased in men who had lost friends and/or partners to AIDS.

Stress, Cytotoxic Lymphocytes, and Human Diseases

Cytolytic cells are critical for immune surveillance against fungal, bacterial, and viral infections. They also play a vital role in cellular resistance to malignancy and tumor metastasis. Granule-mediated cytotoxicity is the major mechanism by which lymphocytes kill viruses, intracellular bacteria, and tumors. Acute psychological stress is often followed by increased susceptibility to infections.

A strong correlation exists between levels of psychological stress and the development of colds in rhinovirus-innucleated volunteers as well as between stress and upper respiratory infections, necrotizing ulcerative gingivitis infections, and acute Epstein-Barr virus (EBV) infection (mononucleosis). Because the cellular immune system has a major role in the prevention of viral infections and in the suppression of the activation of latent viral infections, any factor that changes the function of the immune system may alter disease progression. Abnormalities of the stress response have been hypothesized as a potential trigger or mediator of chronic fatigue syndrome (CFS). In CFS, the activity of NK cells is reduced relative to controls. In CFS patients, intracellular levels of perforin are reduced relative to controls, in both NK cells and in CD8 T cells. Stress may be a precipitating factor, which promotes neuroendocrine and immune dysfunction in vulnerable individuals, resulting in an increased risk of developing chronic multisymptom illness, such as CFS as well as mood/anxiety disorders.

Antibody titers to EBV are consistently elevated in stressed subjects. Reactivation of latent herpesvirus infections, including EBV, cytomegalovirus, human herpes viruses type 6 (HHV-6), and herpes simplex types 1 and 2 (HSV-1 and -2), is an important consequence of the loss of cellular immune function. Elevated antibody titers to EBV have been found to be associated with an avoidant style of cognitive processing in college students (suggesting indirectly a decrement in cellular immune function and poorer control of latent herpesviruses).

In HIV infection, the antiviral role of CD8 T cells and NK cells is thought to be of great importance. Mathematical modeling suggests that, on average, an HIV-specific CD8 T cell can kill 0.7 to 3 target cells per day. Some HIV-infected individuals, called long-term survivors, survive for longer periods than others; a subset of these remain asymptomatic for extended periods and are denoted long-term nonprogressors. Among a myriad of factors that may contribute to disease progression in HIV/AIDS, the stress/endocrine/immune connection should not be neglected. In a pilot study of long-term nonprogressors, it was noted that men who had very low CD4 counts had normal NK cytotoxicity. This suggests that natural NK cell immunity may act as a compensatory protective mechanism in the face of virtual absence of T helper cells.

See Also the Following Article

Natural Killer (NK) Cells.

Further Reading

- Betts, M. R., Brenchley, J. M., Price, D. A., et al. (2003). Sensitive and viable identification of antigen-specific CD8⁺ T cells by a flow cytometric assay for degranulation. *Journal of Immunological Methods* **281**, 65–78.
- Esterling, B., Antoni, M., Fletcher, M. A., Margulies, S. and Schneiderman, N. (1994). Emotional disclosure through writing or speaking modulates latent Epstein-Barr virus reactivation. *Journal of Consulting and Clinical Psychology* **62**, 130–140.
- Fletcher, M. A., Maher, K. and Klimas, N. G. (2002). Natural killer cell function in chronic fatigue syndrome. *Clinical and Applied Immunological Reviews* **2**, 129–139.
- Glaser, R. and Kiecolt-Glaser, J. K. (1998). Stress-associated immune modulation: relevance to viral infections and chronic fatigue syndrome. *American Journal of Medicine* **105**(3A), 35S–42S.
- Ironson, G., Wyings, C., Schneiderman, N., et al. (1997). Post-traumatic stress symptoms, intrusive thoughts, loss and immune function after Hurricane Andrew. *Psychosomatic Medicine* **59**, 128–141.
- Ironson, G., Balbin, E., Solomon, G., Schneiderman, N., Fahey, J. and Fletcher, M. A. (2001). Preservation of natural killer cell cytotoxicity and number in healthy AIDS patients with low CD4 counts. *AIDS* **15**, 2065–2073.
- Irwin, M., Brown, M., Patterson, T., Hauger, R., Maseovich, A. and Grant, I. (1991). Neuropeptide Y and natural killer cell activity: findings in depression and Alzheimer caregiver stress. *FASEB Journal* **5**, 3100–3107.
- Kiecolt-Glaser, J., Glaser, R., Strain, E., et al. (1986). Modulation of cellular immunity in medical students. *Journal of Behavioral Medicine* **9**, 311–320.
- Klimas, N. G., Morgan, R., Blaney, N., et al. (1990). Alcohol and immune function in HIV-1 seropositive, HTLVII/III seronegative and positive men on methadone. *Progress in Clinical Biological Research* **325**, 103–111.
- Klimas, N. G., Walling, J., Patarca, R., et al. (1994). Clinical and immunologic in AIDS patients following adoptive therapy with activated CD8⁺ T cells and interleukin-2 infusion. *AIDS* **8**, 1073–1081.
- Maher, K., Klimas, G. and Fletcher, M. A. (2005). Chronic fatigue syndrome is associated with diminished intracellular perforin. *Clinical Experimental Immunology* **142**, 505–511.
- Miller, L., Goldstein, G. and Murphy, M. (1982). Reversible alterations in immunoregulatory T cells in smoking. *Chest* **82**, 526–529.
- Schedlowski, M., Jacobs, R., Stratmann, G., et al. (1993). Changes of natural killer cells during acute psychological stress. *Journal of Clinical Immunology* **13**, 119–126.
- Solomon, G. F., Benton, D., Harker, J. O., Bonivida, B. and Fletcher, M. A. (1994). Prolonged asymptomatic states in HIV-seropositive persons with fewer than 50 CD4⁺ cells per mm³. *Annals of the New York Academy of Sciences* **741**, 185–190.
- Wick, W. D., Yang, O. O., Corey, L. and Self, S. G. (2005). How many human immunodeficiency virus type 1-infected target cells can a cytotoxic T-lymphocyte kill? *Journal of Virology* **79**, 13579–13586.

D

Death Anxiety

R Kastenbaum

Arizona State University, Tempe, AZ, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R Kastenbaum, volume 1, pp 645–651, © 2000, Elsevier Inc.

Empirical and Clinical Studies of Death Anxiety
Theories of Death Anxiety

Glossary

<i>Existentialism</i>	A philosophical position that emphasizes the individual as alone in the world and responsible for his or her own life and values.
<i>Hospice</i>	A programmatic approach to providing comfort to terminally ill people and their families.
<i>Ontological confrontation</i>	An experience that forces one to acknowledge the reality and salience of death.
<i>Near-death experience (NDE)</i>	An altered state of consciousness usually occurring after traumatic injury that is often recalled as a comforting and revelatory episode.
<i>Palliative care</i>	Medical and nursing procedures designed to relieve pain and suffering rather than extend life by all possible means.
<i>Thanatophobia</i>	Fear of death and of whatever has become associated with death.

Empirical and Clinical Studies of Death Anxiety

The experience of anxiety is much the same whatever its source: a sense of danger and foreboding accompanied by physical manifestations such as changes in heart rate and respiration. Precisely how death

anxiety differs from other forms of anxiety – if at all – is a question that has received little attention. Instead, most academic studies have relied on self-report questionnaires. Respondents usually respond by agreeing or disagreeing with direct questions such as “I fear dying a painful death.” Fortunately, there have also been multilevel and interview studies that enhance our understanding.

Self-Reported Death Anxiety in Everyday Life

Most studies have been conducted with healthy adults, with attention given to their level of anxiety and gender, age, and other demographic variables.

The general population: How anxious about death? Most adults in the United States report themselves as having moderately low levels of death anxiety. This consistent finding often has been interpreted to mean that we labor to keep our anxieties hidden both from ourselves and others. Supporting this interpretation are studies that have detected physiological indications of stress at the same time that subjects report feeling no death-related anxiety at all. It is therefore argued that we are all anxious about death on the gut level, though are reluctant to admit so. An alternative interpretation is that we usually have a low walking-around level of death anxiety that only surges up when a threat is detected. There is a readiness for emergency response, but from moment to moment life is not an emergency. This issue is revisited in the section on theoretical approaches.

Is there a gender difference in death anxiety? Women in the United States usually report higher levels of death anxiety than men. The few studies conducted in other nations find the same pattern. We can choose among several interpretations:

- Women are too anxious about death.
- Men are not anxious enough about death.

- There is no actual difference in level of death anxiety between the sexes, but women are more willing to acknowledge and express their concerns.

In weighing these alternative interpretations, it is useful to recognize the exceptional contribution made by women in providing care and comfort to dying people and their families. Much of the leadership and manpower in the international hospice movement has come from women, who also are far more likely than men to enroll in death education courses. Women have generally been more responsive to the needs of people with life-threatening conditions and more cognizant and accepting of their own feelings about death. Men generally seem more guarded and uncomfortable when circumstances remind them of their vulnerability. These gender differences have been related to childhood socialization patterns in which the expression of feelings, especially those of vulnerability, is encouraged in girls but discouraged in boys.

Women may be too anxious about death if the criterion is subjective comfort, but men may be not anxious enough if the criterion is the readiness to respond to the mortal jeopardy of self and others. Research also suggests gender differences in the type of situations that arouse the most anxiety. Men express more fear of extinction and the way their post-death self will be remembered. Women express more fear of the dying process and of being unable to make adequate arrangements for the well-being of their families. These differences should not be overestimated. There are marked individual differences in personality within each gender.

Is there an age difference in death anxiety? Do we become more anxious as we grow older because of the decreasing distance from death? Or do we become less anxious because we develop a more mature outlook on life and death?

Age differences, when found, indicate less fear of death among elderly people. Concerns are more likely to center around the challenges and stresses of daily life, such as protecting relationships and self-esteem. However, it should be kept in mind that each generation passes through unique historical circumstances with distinctive configurations of social attitudes, economic up- and downturns, technological innovations, and threats to life. The fact that most elders have been reporting low or moderate levels of death anxiety probably represents the interaction of their distinctive life histories and current health and socioeconomic status, as well as their personal accommodations to the aging process. Qualitative studies have identified specific concerns that are most salient for

people at different points in their life course. Young people seem most apprehensive about the possible loss of loved ones, death as punishment, and the finality of death. Midlife adults express most concern about premature death and fear of pain in dying. Older adults do not fear death as much as they do the prospect of becoming helpless and dependent on others during their final phase of life.

There is little support for the hypothesis that people become more anxious as the distance from death decreases with advanced age. Self-actualization and existential theories propose that acceptance of death is a mark of maturity. This is an attractive proposition, but not an established fact. It might be overreaching to insist that philosophical acceptance of death is the only way a mature person can come to terms with mortality. Furthermore, other explanations might account for the relatively low death anxiety found among elderly adults. Some people experiencing social isolation, financial concern, and age-related physical problems express a readiness to have their lives come to an end – an attitude that is also reflected in the high completed suicide rates for elderly white men. Low death anxiety might then be related to dissatisfaction with a life that is no longer considered worth living.

What are the demographics of death anxiety? People in favorable socioeconomic circumstances tend to report lower levels of death anxiety. Education, affluence, and status serve as buffers against death anxiety (though not impenetrable buffers). Growing up in an intact family and having a secure interpersonal environment also offer some protection against high death anxiety. Married people generally live longer and are less likely to commit suicide. We might therefore expect that married people would experience less anxiety than those who are single or divorced, but, surprisingly, this proposition has not been systematically examined.

Religious beliefs and practices have a complex relationship with death anxiety. Anthropologists and historians have argued that religion originated as a response to the fear of death. A particular religion operating within a particular cultural context might be effective in controlling death anxiety with comforting rites of passage and images of blessings in the afterlife. The opposite effect is also possible, however. Presenting death in frightening images is an effective way of keeping anxiety bubbling and therefore motivating people to seek protection by obedience to the establishment. In today's mass and heterogeneous societies, religion offers a broad array of beliefs and practices, some of which seem likely to reduce and others to increase death anxiety. Some people of faith

might have less fear of death because they expect a joyful afterlife, while others are beset by anxiety about damnation. Surveys continue to indicate that most people in the United States do believe in an afterlife. Studies of terminally ill people suggest that there is more anxiety about suffering and dependency during the dying process than fear of death as such. Religious faith and belief in an afterlife seem to help many people face the prospect of death, but they are still vulnerable to anxiety about the dying process and the effects of their death on others. It has become clear, however, that many people feel more secure during their final illnesses when they know that they will be given full benefit of the religious and social customs that are meaningful within their group.

What are the personality and lifestyle correlates of death anxiety? People who are afflicted with many other fears are likely to have relatively high death anxiety as well. People with humorless authoritarian personalities are the least likely to sign the organ donation forms attached to their drivers' licenses, suggesting a higher level of death-related anxiety. A dominating need to control situations and to win at all costs has been related to a feeling of vulnerability: catastrophic anxiety and death must be warded off by exercising power over the course of events. Similarly, some people who seem to live outside of themselves through causes and cults do so to achieve a sense of participatory immortality and thereby escape the sting of death anxiety. By contrast, people with deep interpersonal attachments who have a strong sense of purpose in life tend to have a lower level of death anxiety.

Death Anxiety in Particular Situations

Our focus now shifts from characteristics of the individual to the types of situations in which individuals may find themselves. We draw now upon a wide range of observations beyond self-report questionnaires.

Transitional situations We often experience heightened death anxiety in transitional situations. Starting or losing a job and moving to a new community are common examples of transitions that have no evident relationship with death but that can engage feelings of uncertainty and threat. Separation, divorce, and other significant changes in relationships can also arouse a sense of vulnerability, which, on the emotional map, is not far from fear of mortality. Even detection of the first gray hair can trigger the recognition that one is no longer young, thereby arousing the lurking fears of aging and death. Determined attempts to look young often express a convergence

of fears: "I will be less attractive, less valued, less useful – and I will become the sort of person who is most likely to die – the old-timer."

The community at large can also experience heightened stress during periods of transition and uncertainty. Apprehension about unemployment, inflation, inadequate health-care benefits, and difficulties in coping with bureaucracy and technology can all contribute to a sense of helplessness, which, in turn, increases permeability to awareness of mortality. A society that has lost confidence in its purpose and competence may contribute to heightened death anxiety throughout all ages, ranks, and echelons.

Touched by death Exposure to death does not always make us anxious. Sometimes we seal off such episodes quickly before they can penetrate our awareness. At the extreme, we may deny the significance of the event, as when the victim of a heart attack insists that he is just fine and does not need medical attention. More often, though, we use subtle defensive strategies to avoid the impact of an exposure to death. For example, at a funeral we may not permit ourselves to acknowledge a possible resonance between the neighbor who died of pulmonary disease and our own habit of smoking. "It's too bad about the neighbor, but I don't smoke the same brand of cigarettes." Many of us are gifted in the ability to shed exposures to death in order to preserve the comforting routines of everyday life. Nevertheless, some exposures to death do get through to us. Usually these exposures take one of two forms: a personal brush with death or the death of another person.

It is not uncommon for the stress reaction to develop after a life-threatening emergency. A competent driver or pilot, for example, is likely to respond to an emergency with quick and skillful actions. Only when the danger has passed will the emotional reaction seize its opportunity, perhaps taking the form of a transient posttraumatic stress episode. Anxiety aroused due to a close call with death usually dissipates quickly, although there may be heightened sensitivity to future risk situations. What has become known as a near-death experience (NDE) is a different case. These episodes almost invariably involve risk to life that the individual could not deal with through personal initiative and action (e.g., the passenger in a car hit in a broadside collision). The NDE is an episode split off from the person's usual life and marked by unusual, dream-like events. Instead of intense anxiety, however, many reports depict a remarkable feeling of serenity and liberation. Some people believe that they not only have been touched by death but also were actually in death for a short period of time. There are varying interpretations of

the actual state of being during a NDE, but a great many reports that anxiety about death completely dissipated after the experience.

By contrast, the death of another person is often mentioned as the wake-up call that leads to anxious realization of one's own mortality. Most often it is the death of a parent that produces an increased sense of personal vulnerability. "Death is coming after me, now!" as a 50-ish woman exclaimed after the death of her mother. The death of a long-lived cultural icon can also lead to feelings of vulnerability and abandonment. In addition, we are likely to experience an upsurge of death-related anxiety when a person seen as much like ourselves passes away. Our behavior toward others when they are terminally ill and during the funeral process can be markedly influenced by anxieties aroused by our feeling that "this could be happening to me."

Facing death: Life-threatening illness The prospect of our own death may flash before us in a quickly developing episode as already noted. It is a different matter, however, when the threat is persistent, as in the case of life-threatening illness. Coping with a threat to life varies with both the individual and the medical condition (e.g., a cardiac disorder that might end one's life at any time or advanced pulmonary disease in which every breath is a struggle as the disease continually progresses). Nevertheless, there are also some common issues that often intensify anxiety. These include (a) uncertainty about diagnosis and prognosis, (b) learning that one has a life-threatening condition, (c) progression of symptoms and dysfunctions that suggest that treatment efforts are failing, (d) interpersonal cues that others also recognize the end is approaching, along with concern about losing one's most valued relationships, (e) concern about practical end-of-life matters, (f) concern about the overall meaning of one's life, and (g) anxiety about experiencing helplessness and pain in the end phase of the dying process. Patients often are aware of the anxiety that exists in others around them: it has been found that physicians often experience intense discomfort in the terminal-care situation and therefore try to limit their contacts with the patients. It is not unusual for death anxiety to be reinforced and multiplied throughout the entire terminal-care interpersonal network.

There is some evidence to suggest that episodes of intense anxiety can trigger sudden death in people with coronary heart disease. These findings support earlier observations that a person literally can be scared to death. However, a moderate level of anxiety might be more protective than a denial reaction. It has

often been suggested, but not yet firmly established, that high levels of anxiety can also increase the risk of death in other life-threatening conditions.

Anxiety peaks most often at two points. The first jolt usually occurs when the person discovers that his or her illness is terminal. Suicidal ideation can develop during this chaotic period of panic and depression. Trustworthy and sensitive communication on the part of family and health-care providers can prevent a plunge into the depths of despair or an anxious leap into self-destruction. The second anxiety peak, again accompanied by depression, arises as a result of continued physical deterioration, fatigue, and apprehension regarding progressive loss of function. "I'm just so tired of dying" is how more than one person has expressed it. Here anxiety is focused less on death than on the fear of being abandoned and isolated while continuing to suffer for no good reason that the person can see. Euthanasia is not sought by all people in the end-phase of their terminal illnesses, however. Many continue to experience life as meaningful and to be protected from disabling anxiety by positive belief systems and loving support from the people most important to them.

Anxiety and depression associated with terminal illness have often been the result of social isolation and inadequate control of pain and other symptoms. The pioneering contributions of Elisabeth Kubler-Ross and Dame Cicely Saunders have done much to improve this situation. Kubler-Ross, a psychiatrist, encouraged professional and family caregivers to overcome their own anxieties about interacting with terminally ill people. Saunders, a nurse and physician, launched and guided the international hospice movement that provides a broad spectrum of symptom relief with a combination of compassionate care and palliative expertise. As their own lives neared the end, Kubler-Ross and Saunders both received the benefits of the hospice care they had helped make possible for people around the world. Recently, the medical establishment has shown signs of agreeing with the hospice premise that pain should be regarded as one of the vital signs to be assessed and effectively treated in all physician-patient interactions. Team approaches to the comfort of terminally ill people and their families with priorities given to the dying person's own lifestyle and values have also contributed to reducing anxiety throughout the final phase of life.

Major mental health problems People who feel that their worlds are exploding or collapsing often express intense anxiety about death. Some of the most striking expressions of death anxiety are found during psychotic episodes, whether functional or triggered

by a drug-altered state of mind. Whatever the specific cause, it is usually the sense of losing total control over one's life that has precipitated the episode. Dying and death become metaphorical expressions of this panic. Severely depressed people may even present themselves as dead, thereby having passed beyond the reach of death anxiety. Helpful interventions include providing a higher degree of stability and simplicity in the sociophysical environment and creating situations in which the individual can begin to experience again a sense of competence and control. Approaches as different as expressive and drug therapies can also be useful.

Theories of Death Anxiety

General theories of stress and anxiety have potential for contributing to the understanding of death-related concerns. Here we focus on theories that were intended specifically for comprehension of death anxiety.

Freud and Becker: Rival Theories

Two theories may be considered classic for their bold positions and widespread influence. These views could not be more opposed in their fundamental propositions, although there are commonalities as well.

Sigmund Freud held that we cannot truly have a fear of death because we cannot believe in our own death. Notice that person who is standing off to the side and viewing your death – it is you, the spectator self who is still very much alive. Personal death is not comprehended by the foundational logic of our unconscious processes, while on the conscious level, we have not actually experienced death. An expressed fear of death (thanatophobia), then, is a phobic symptom that conceals the actual, deeper source of anxiety lodged in childhood experiences and stimulated by current life developments. This influential view relegated death anxiety to a derived, secondary status and thereby not of much interest in its own right. Later in his own pained and stressful life, Freud came to see dying and death as issues very much deserving of attention, but it was his earlier formulation that continued to prevail.

Ernest Becker challenged Freud's interpretation from an existential standpoint. For Becker, all anxiety is an expression of death anxiety. Our core fear of nonbeing attaches itself to an almost infinite variety of situations and symbols. This is where society comes in – it exists largely to protect us from being consumed by death anxiety. Becker charges that our language and customs are designed to disconnect us from ontological encounters with mortality. Even our

technologically resplendent civilization remains basically a theatrical performance piece to which we all contribute our talents and suspension of disbelief. We agree not to notice that we all happen to be mortal. The person who really believes in death – and is therefore really anxious about it – is often the person we call schizophrenic. Being normal means being fairly successful in protecting ourselves against facing up to reality.

Both the Freudian/psychoanalytic and the Becker/existential theories have served as useful guides to observation and reflection. Different as they are, however, both share the premise of universal application and the flaw of unverifiability. There are situations in which one theory or the other seems applicable, but also situations in which neither seem sufficient.

Emerging Theories

Several new theories have emerged, each offering a distinctive perspective and drawing upon a distinctive realm of empirical research. Terror management theory is an empirically oriented offshoot of Becker's position. Its core proposition is that self-esteem and a positive worldview serve as effective barriers to death anxiety. When we feel ourselves imperiled, we can call upon our sense of self-worth and confidence as well as our belief that our society is basically sound and the universe a friendly place in which to dwell. Whatever contributes to growth of self-esteem and confidence and whatever contributes to the moral character and resilience of society can stand between ourselves and mortal terror.

Developmental learning theory places death anxiety within the mainstream of sociobehavioral research. We learn to fear, avoid, or cope with death-related phenomena just as we learn to deal with all other life challenges: through a long process of maturation and social interaction. It is not necessary to assert either that all anxiety or no anxiety is related to death or that we all have some kind of fixed response to death-linked situations and symbols. As a society and as individuals, the shape and intensity of our anxieties and coping strategies are pretty much in our own hands, and are likely to vary from generation to generation.

Edge theory shifts the emphasis from fear of death to the readiness to respond to threats to life. Two frequently held assumptions are set aside. First, the question of whether or not we can truly comprehend death is not as important as it once seemed. The point is that we do recognize threats to survival, whether of ourselves or of others. A vigilance function is with us from birth. Like other living creatures, we are equipped with the ability to detect serious threats

and go on alert. The second assumption set aside is that we try desperately to hide our intense anxieties about death from ourselves and others. According to edge theory, however, we do not carry about a burden of repressed death anxiety. Instead, our moderate everyday approach to life is occasionally challenged by a potentially dangerous situation. We have a momentary sense of being on the edge of nonbeing. In such situations we go on alert and try to identify the source of the danger. If we can then either dismiss or take action to overcome the threat we experience only a sense of heightened activation. It is when we feel on the edge and cannot respond to the threat that are likely to be flooded by death anxiety.

Points of difference aside, current research and theory suggest that the understanding of stress management throughout life requires attention to the ways in which we come to terms with our mortality.

See Also the Following Articles

Anxiety; Control and Stress; Fear; Freud, Sigmund; Pain; Self-Esteem, Stress and Emotion; Suicide, Psychology of.

Further Reading

- Becker, E. (1973). *The denial of death*. New York: Free Press.
- Clark, D. (2003). Saunders, Cicely. In: Kastenbaum, R. (ed.) *Macmillan encyclopedia of death and dying* (vol. 2), pp. 743–745. New York: Macmillan Reference USA.
- Freud, S. (1951). Thoughts for the times on war and death. In: *Collected works* (vol. IV), pp. 288–317. London: Hogarth Press.
- Hayslip, B., Jr. (2003). Death denial. Hiding and camouflaging death. In: Bryant, C. (ed.) *Handbook of death and dying*, pp. 34–42. Thousand Oaks, CA: Sage.
- Kastenbaum, R. (2004). *Death, society, and human experience* (8th edn.). Boston, MA: Allyn & Bacon.
- Kastenbaum, R. (2004). *On our way. The final passage through life and death*. Berkeley, CA: University of California Press.
- Kaufmann, W. (1976). *Existentialism, religion and death*. New York: New American Library.
- Kubler-Ross, E. (1969). *On death and dying*. New York: Macmillan.
- Tomer, A. (2003). Terror management theory. In: Kastenbaum, R. (ed.) *Macmillan encyclopedia of death and dying* (vol. 2), pp. 885–887. New York: Macmillan Reference USA.

Death Guilt See: Survivor Guilt.

Defense Services, Stress in See: Survivor Guilt.

Defensive Behaviors

D C Blanchard, M Yang, M Hebert and R J Blanchard

University of Hawaii, Honolulu, HI, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D C Blanchard, M Hebert and R J Blanchard, volume 1, pp 652–656, © 2000, Elsevier Inc.

Defensive Behaviors and the Stimuli That Elicit Them

Across-Species Generality

Defensive Behaviors as Independent Biobehavioral Systems

Acute versus Chronic Behavioral Defenses

Defensive Behaviors, Stress, and Psychopathology

Glossary

Chronic mild stress (CMS)

A pattern of anhedonia involving reduced positive response toward normally rewarding items, such as a sweetened liquid or intracranial stimulation in reward sites, following intermittent exposure to a variety of mildly stressing stimuli or events. These behavioral outcomes, like those of both learned helplessness and subordination, have been shown to be responsive to antidepressant drugs. An array of evolved behavioral responses (e.g., flight or freezing) to threat stimuli such as predators or attacking conspecifics. Individual defensive behaviors are

Defensive behavior

	modulated by features of both the threat stimulus and the situation in which they are encountered. Many defensive behaviors are relatively consistent across mammalian species and may be a component of stress- or defense-related psychopathologies.
<i>Defensive threat/attack</i>	Display of weapons (e.g., teeth and claws) and strength (e.g., through loud vocalization) in response to close contact by an attacker, followed by bites or blows to particularly vulnerable sites on the attacker's body (typically, face and eyes).
<i>Dominance/subordination</i>	A pattern of relationships between a consistent winner (dominant) and a consistent loser (subordinate) in agonistic encounters. Both dominant and subordinate often show behavioral and physiological signs of stress, but these may be modulated by compensatory mechanisms in the dominant that are not available to the subordinate. Subordination has been used as a model of chronic social stress.
<i>Learned helplessness (LH)</i>	A pattern of reduced learning, particularly in active avoidance situations and of reduced activity in a threat context, following exposure to uncontrollable aversive events. This, like the chronic mild stress paradigm, has been proposed as a model of depression.
<i>Offensive attack</i>	An attack, typically (but not always) toward conspecifics, in the context of a dispute over important resources such as territory, access to mates, or food. An offensive attack is aimed toward species-typical sites on the opponent, often areas in which bites or blows do little serious damage. In contrast to defensive threat/attack, fear reduces offensive attack.
<i>Risk assessment</i>	A pattern of orientation, attention, and exploration with regard to potential threat stimuli, enabling an animal to gather information about their location, identity, and threat status. This information facilitates an optimally adaptive response to a genuine threat or promotes a return to nondefensive behavior if no threat is present.

Defensive behaviors are those activities that occur in response to the host of life-threatening dangers encountered in every natural environment, from predators, from conspecific attack, and from threatening features of the environment. They constitute the behavioral component of the stress response and have evolved on the basis that, to particular types of threat

stimuli and in particular situations, each such behavior has proved to be optimally adaptive in terms of enhancing the extended reproductive success of the individual. These behaviors range from active investigation of threat stimuli to actions facilitating escape from or termination of threat. Because the extended reproductive success of females may, in an important way reflect the survival of offspring in their care, the optimal defensive behaviors for males and females may differ: The magnitude and mechanisms of gender difference in defense are poorly understood, but there is speculation that these may be related to the substantial and consistent differences in vulnerability to particular psychiatric disorders for men and women.

Defensive Behaviors and the Stimuli That Elicit Them

Much recent work on defensive behaviors has focused on response to predators or attacking conspecifics, threat stimuli (especially the former) that elicit a number of defensive behaviors in rodents without the necessity of pain or prior experience. Although the defensive behaviors of higher primates are undoubtedly more complex than those of rodents and may involve some specific learning components, there appears to be an essential continuity of defense patterns across mammals such that research on the behavioral, neural, and neurochemical aspects of the defense systems may be very relevant to analysis of defense-related behaviors, including psychopathologies, in people.

Defensive behaviors are modulated by features of both the threat stimulus and the situation in which it is presented). The major (self-)defensive behaviors are:

1. *Flight* from the threatening stimulus, and avoidance of that stimulus if it is at a distance. These behaviors are facilitated by the presence of an escape route or a place of concealment and are dominant responses to high-level threat when such features are available.

2. *Freezing* when no escape route or place of concealment is available. Freezing involves orientation toward the threat source and immobility in a species-typical posture. Like flight, it increases in intensity as the threat approaches, up to a level of proximity that elicits defensive threat/attack.

3. *Defensive threat* occurs as the threat stimulus nears the subject. It consists of a weapons display (teeth bared, claws unsheathed) and loud sonic vocalizations such as screams.

4. *Defensive attack* occurs at even shorter threat-subject distances. This attack is often oriented toward

particularly vulnerable sites such as the predator/attacker's face and eyes.

5. *Risk assessment* occurs when the threat is potential rather than present (e.g., predator odor or novel situation) or of low intensity and includes sensory sampling (visual/auditory scanning, sniffing) and approach and investigation of the threat source, alternating with rapid retreat. Risk assessment is associated with low-profile postures and rapid movement interspersed with periods of immobility; these reduce the likelihood of detection as the animal approaches and investigates the threat stimulus. These activities enable the gathering of information, facilitating identification and localization of the threat source. If it proves to be a real danger, defenses such as flight or defensive threat/attack ensue. Alternatively, if risk assessment permits a determination that the situation is safe, then the animal can resume its normal activities (food gathering, self-care, or offspring care, etc.).

6. *Alarm vocalizations* occur in many (perhaps most) mammalian species with a colonial lifestyle. Such cries warn other group members of the appearance of a predator. In some primate species, different types of predator (e.g., terrestrial vs. aerial) elicit different alarm cries, permitting the listener to take defensive action appropriate to that predator.

These defensive behaviors all occur in the context of physical threat. Threat to resources elicits a very different behavior, offensive attack or aggression, which has a much weaker association with fearfulness or psychological stress. Both offensive and defensive (i.e., self-defensive) attack toward conspecifics in mammals tend to be oriented toward specific areas on the body of the opponent, and these target sites may provide an additional means of differentiating offensive and defensive attack modes. The former tends to be aimed at sites where bites or blows will produce pain without a high risk of damage to vital organs, whereas the latter is typically targeted toward the face of the threatener. This, for both conspecific and predator attackers, is both the site of particularly vulnerable and important organs, such as the eyes, and of the major offensive weapons of many animals, that is, the teeth. Because conspecific attack is so precisely targeted, an additional group of defenses against conspecific threat involves concealment of these especially likely targets of attack. These target-protection defenses, in combination with attacker behaviors that seek to thwart the defender (target concealment) strategy and reach the preferred attack target, may give fighting among conspecific mammals a dancelike quality in which the actions of one animal are immediately countered by appropriate actions of the other.

All of these defensive behaviors involve a strong component of orientation toward (or, in the case of flight, away from) the threat stimulus. In contrast, analyses of chronic internal pain (a stressful event not involving an external stimulus) suggest that the behavioral defense to this may be quiescence, involving reduced activity and inattention to external stimuli.

Across-Species Generality

These defensive behaviors and their specific associations with particular elements of the threat stimulus and the environment in which it is encountered show a high degree of cross-species generality in mammals. The generality is probably not limited to mammals, but mammals do appear to rely more strongly on this range of behavioral defenses than do other animals. Structural defenses (e.g., armor, venom, and poisonous tissues) and cryptic coloration are emphasized in lower vertebrates and invertebrates, whereas birds show an obvious specialization toward flight.

The behavioral repertory, including defensive behavior, of higher primates is more varied and sophisticated than that of more primitive mammals (e.g., rodents). However, the same defensive behaviors noted for rodents also occur in primates, albeit with greater elaboration of cognitive and communicatory elements. Examples are threat-differentiated alarm vocalizations, social or observational learning of threat stimuli (based on the defensive behaviors of conspecifics to these stimuli), and the formation of conspecific coalitions involving mutual defense and protection. Studies of defensive behaviors in humans are necessarily more indirect. However, choices of responses to scenarios designed to focus on features of the threat stimulus and the situation in which it is encountered strongly suggest that the relationship between stimulus/situational features and particular defenses in humans are broadly similar to those described in nonhuman animals.

Defensive Behaviors as Independent Biobehavioral Systems

Much classic research on defensive behavior had an initial "site" focus, in that it involved attempts to describe the immediate response of an animal to electrical or chemical stimulation of particular brain structures. Although a number of forebrain sites yield behaviors (e.g., vocalizations) that appear to be related to defensiveness, the most thoroughly investigated area has been the periaqueductal gray (PAG) region of the midbrain. This contains a number

of longitudinally coursing columns in which stimulation can produce defensive threat/attack, flight, and freezing or quiescence. Some controversy surrounds the interpretation of stimulation of the ventral PAG column in terms of the last two alternatives. Responses related to defensive threat appear to be more common in the anterior components of these systems, whereas flight and freezing occur with stimulation in more posterior sites. It is particularly interesting that the PAG has been mapped in both cats and rats and that the specific areas where particular behaviors are elicited by stimulation appear to be very consistent.

Typically such studies did not provide a strong threat stimulus, or even an opponent to which the animal could orient and react, making interpretations of behavior more difficult. More recent attempts to determine the neural systems of defense have involved specific stimuli, notably presentation of a cat or cat odor to the subject rat. These studies indicate that cat stimuli elicit a strong c-Fos expression in specific areas of the amygdala, lateral septum, and specific nuclei of the hypothalamus, notably the dorsal pre-mammillary nucleus. Lesions of the last of these strongly reduce responsivity to a cat or cat odor, but do not have a major impact on freezing following foot shock, suggesting considerable stimulus or stimulus modality specificity of control of defense in some forebrain structures. In addition to anatomical differences, there appear to be differences in the neurochemistry of the biobehavioral defensive systems, providing a basis for findings of differential response of these systems to psychoactive drugs.

Acute versus Chronic Behavioral Defenses

In addition to acute exposure paradigms, a number of models of chronic social or predator threat have been devised and used to evaluate behavioral and physiological effects of longer-term stress. In a particularly desirable habitat, with females and a burrow system, male rodents quickly create a strong dominance hierarchy based on victory and defeat. Defeated subordinate male rats spend most of their time within a particular chamber or tunnel, from which they attempt to exclude the victorious male (dominant); defensive threat/attack is a particularly effective defense in a tunnel. From these sites, subordinates also consistently monitor the dominant's location, avoiding the surface area when it is present and fleeing when they are in the open area and the dominant appears. Subordinates spend most of their waking time in defensive behavior and lose weight even when food and water are freely available.

Much of the defensive behavior of a subordinate rat to a dominant is very similar to that seen on initial exposure to a threat stimulus, except that it is more protracted. However, prolonged exposure to threat/stress may produce deleterious behavioral changes such as reduced activity, a shift away from active defense toward more passive defenses, and possibly a shift to a quiescent state in which the animal becomes less, or less effectively, attentive to possible environmental dangers. In learning situations, particularly those involving an active avoidance response, deficits in responding by animals chronically exposed to an uncontrollable stressor are characterized as indicating LH. A potentially similar phenomenon in subordinate rats involves greatly prolonged risk assessment without transition to nondefensive behavior. This may also reflect a learning deficit and it, like the active avoidance deficit of LH rats, may be quite maladaptive. The prolonged (and often inadequately cautious) risk assessment behavior exposes the animal to danger if a threat is present, whereas the failure to resume normal behavior patterns is maladaptive if there is no threat present.

Another behavior associated with chronic stress (in a CMS model) is anhedonia, typically indexed by reduced consumption of sweetened liquids. In fact, the CMS animals may actually show heightened preference for more intensely sweetened liquids.

Defensive Behaviors, Stress, and Psychopathology

Defensive behaviors – both those seen in response to acute threat and the altered patterns expressed after chronic threat/stress – have many parallels to the behavioral components of a range of human psychopathologies. In particular, risk-assessment behaviors appear to be very similar to the increased vigilance and scanning characteristic of generalized anxiety disorder, whereas reduced activity, anhedonia, and deficits in active avoidance learning have parallels in depression.

An additional parallel between the chronic stress models and depression is through alterations in the functioning of the hypothalamic-pituitary-adrenal (HPA) axis and changes in brain neurotransmitter systems. Different physiological systems have often been emphasized in investigations of the various chronic stress models (LH, CMS, and subordination), and so a consensus on their commonality of effects is difficult to obtain. HPA axis changes appear to be most severe for subordinates, many of whom show greatly decreased levels of corticosterone binding globulin, reduced corticosterone (CORT) response to acute stress, and reduced corticotropin-releasing

factor (CRF) in the paraventricular nuclei of the hypothalamus, indicating the involvement of brain mechanisms in this deficit. Subordinates also show sharply reduced levels of testosterone, reflecting both peripheral and central effects of stress, with the former appearing prior to the latter. For the most part, LH is not associated with these hormonal effects, particularly when the criterion of experiencing uncontrollable versus controllable shock is applied. This very stringent criterion, using chronic but controllable shock as a control condition, may obscure some important changes in stress-linked physiological systems. As befits the term chronic mild stress, the CMS model also fails to produce notable hormonal changes. However, all three models have been associated with brain neurochemical changes, particularly involving biogenic amines, and a number of these brain changes, as well as the behaviors with which the models are associated, have been shown to be reversible after treatment with antidepressants. These hormonal and brain effects, in addition to the changes in defense and other behaviors after chronic stress, suggest that animal models of chronic stress may provide some very relevant analogs to depression and further emphasize the potential value of use of these behaviors as pre-clinical models for research on the pharmacology of emotion-linked disorders.

This approach has produced many interesting results. Risk-assessment behaviors show a selective response to anxiolytic drugs such as classic benzodiazepines. Antipanic drugs (e.g., alprazolam, imipramine, fluoxetine, and phenelzine) decrease flight behaviors, whereas the panicogenic agent yohimbine increases flight. As with clinical findings, the effects of these antipanic drugs on flight require chronic administration, whereas the initial injection of imipramine and fluoxetine may actually exacerbate defense. Both CMS and LH models have proved to be responsive to a variety of antidepressant drugs. However, the various behavioral effects of the latter model, such as reduced activity and disrupted learning of an active avoidance response, may show differential patterns of response to antidepressant and anxiolytic drugs. Both acute and chronic social defeat produce behavior changes that are normalized by fluoxetine, a selective serotonin reuptake inhibitor that has antidepressant as well as antipanic effects. In general, these findings suggest that, although acute and chronic stress may

produce a wide pattern of changes in defensive and other behaviors, it is the behavioral response itself, rather than the stress paradigm, that may provide the clearest and most direct relationship to particular stress-linked psychopathologies.

See Also the Following Articles

Anxiety; Depression Models.

Further Reading

- Bandler, R. and Keay, K. A. (1996). Columnar organization in the midbrain periaqueductal gray and the integration of emotional expression. *Progress in Brain Research* **107**, 285–300.
- Blanchard, D. C., Griebel, G. and Blanchard, R. J. (2003). The mouse defense test battery: pharmacological and behavioral assays for anxiety and panic. *European Journal of Psychology* **463**, 97–116.
- Blanchard, D. C., Hynd, A. L., Minke, K. A., et al. (2001). Human defensive behaviors to threat scenarios show parallels to fear- and anxiety-related defense patterns of nonhuman mammals. *Neuroscience and Biobehavioral Reviews* **25**, 761–770.
- Blanchard, D. C., Spencer, R., Weiss, S. M., et al. (1995). The visible burrow system as a model of chronic social stress: behavioral and neuroendocrine correlates. *Psychoneuroendocrinology* **20**, 117–134.
- Canteras, N. S., Ribeiro-Barbosa, E. R. and Comoli, E. (2001). Tracing from the dorsal premammillary nucleus prosencephalic systems involved in the organization of innate fear responses. *Neuroscience and Biobehavioral Reviews* **25**, 661–668.
- Davis, M. (1997). Neurobiology of fear responses: the role of the amygdala. *Journal of Neuropsychiatry and Clinical Neuroscience* **9**, 382–402.
- Dielenberg, R. A., Hunt, G. E. and McGregor, I. S. (2001). “When a rat smells a cat”: the distribution of Fos immunoreactivity in rat brain following exposure to a predatory odor. *Neuroscience* **104**, 1085–1097.
- Gray, J. A. and McNaughton, N. (2000). *The neuropsychology of anxiety: an enquiry into the functions of the septo-hippocampal system* (2nd edn.). Oxford: Oxford University Press.
- Maier, S. F., Ryan, S. M., Barksdale, C. M., et al. (1986). Stressor controllability and the pituitary-adrenal system. *Behavioral Neuroscience* **100**, 669–674.
- Willner, P., Muscat, R. and Papp, M. (1992). Chronic mild stress-induced anhedonia: a realistic animal model of depression. *Neuroscience and Biobehavioral Reviews* **16**, 525–534.

Demand–Control Model

T Theorell

Karolinska Institute, Stockholm, Sweden

© 2007 Elsevier Inc. All rights reserved.

History

The Model and Its Components

The Model in Relation to Heart Disease

Demand–Control and Effort–Reward: Independent Predictions?

Other Health Outcomes

Physiological Mechanisms

Childhood Confounding

Examples of Organizational Interventions in Workplaces Aiming at Decreased Heart Disease Risk

Critique and Place of the Demand–Control–Support Model

Glossary

<i>Decision authority</i>	The component of decision latitude that is the possibility for employees to influence how to do work.
<i>Decision latitude</i>	The possibility for employees to influence decisions regarding work.
<i>Iso-strain</i>	The combination of high psychological demands, poor decision latitude, and poor social support at work.
<i>Job strain</i>	The combination of high psychological demands and poor decision latitude at work.
<i>Skill discretion</i>	The component of decision latitude that is the possibility for employees to develop skills at work (improving possibility to make decisions); also called intellectual discretion.
<i>Work social support</i>	Support from workmates and superiors.

History

The demand–control–support model was introduced in the early 1970s by Karasek in his doctoral thesis and published internationally in 1979. It was further developed in relation to psychophysiological theory and empirically tested in relation to cardiovascular disease by Karasek and Theorell during the 1980s. The support dimension was tested and discussed in more detail for the first time in 1990 by Johnson and Hall. The demand–control theory states that the combination of high psychological demands and low decision latitude (job strain) is particularly dangerous

to health. The demand–control–support model states that the effects of job strain are worsened by poor social support and ameliorated by good social support; on the other hand, high demands with high decision latitude (active work) may be associated with psychosocial growth and improved coping.

The demand–control model is related to the general stress theory of Selye (demands) and to the sociological alienation theory of Gardell. Karasek combined these two theories into a two-dimensional environmental theory. When the two dimensions are combined with the support dimension, a three-dimensional model arises (see [Figure 1](#)). The ideal situation is a job without excessive psychological demands in which the employees have good decision latitude and good support (relaxed with good support). The worst situation is a job situation with high psychological demands in which employees have low decision latitude and poor support (high strain with poor social support or iso-strain). According to the theory, all the other situations – high demands and good decision latitude, active with good and poor social support, and low demands and poor decision latitude, passive with good and poor social support – occupy intermediate positions with regard to illness risk and health.

Frankenhaeuser had earlier described another two-dimensional model focused on the individual's reaction to stressors. When the effort dimension is combined with the joy/uneasiness dimension in the reaction, it becomes evident that a high degree of effort may be combined either with joy (eustress) or with uneasiness (distress). This model has a similarity with the demand–control model, but the demand–control model is an environmental model, whereas the Frankenhaeuser model is an individual-reaction model. This means, for instance, that individual distress reactions are more common in high strain and eustress reactions more common in active work.

Later, the Frankenhaeuser and the demand–control models were incorporated into other models. Examples are the effort–reward imbalance model (see [Effort-Reward Imbalance Model](#)) and the demands–resources model. In the latter model, resources include personal resources as well as external ones such as decision latitude and social support.

The Model and Its Components

Decision latitude (DL) is defined as possibility for the individual employee to exert control over his or

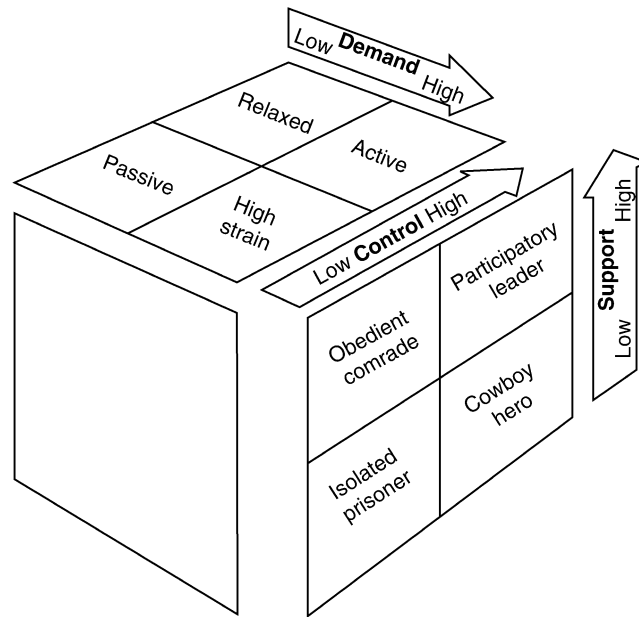


Figure 1 A three-dimensional model of the psychosocial work environment. From Karasek, R. A. and Theorell, T. (1990) *Healthy work*, New York: Basic Books.

her work situation; it does not refer primarily to the individual's ability to exert control because the focus in the demand-control-support model is on the organization's demands on the employee and the possibility that it gives control to the employee to handle these demands. DL has two components: decision authority and skill discretion. Both of these components are related to the exertion of control. Decision authority is the influence that the employee has over how the work is done (control over tasks). Questions regarding decision authority typically deal with influence on tempo and production techniques and whether employees have any say in decision making. Skill discretion is the development of competence and skills (control over knowledge). Typical questions regarding skill discretion deal with the possibility of learning new things, variations, and skill requirements.

Psychological demands in the demand-control model correspond to several kinds of external demands in the work situation. They are added up to produce a total score. It has been pointed out in recent discussions that this assessment practice may create problems. For instance, long working hours and high intensity should perhaps be treated as separate variables because they are patterned differently in different social groups and occupational branches. Similarly, emotional and cognitive demands are different. Social support refers to emotional as well as instrumental support from both superiors and work-mates. The three basic dimensions – psychological demands, decision latitude, and social support – are

assessed using many different questionnaires; the most widely used questionnaire for the assessment of the total model is Karasek and colleagues' Job Content Questionnaire (JCQ).

The Model in Relation to Heart Disease

There is a growing body of research showing that the demand-control-support model is related to heart disease risk. Most of the empirical studies have been performed on demands and decision latitude. That poor social support contributes to cardiovascular mortality, however, was shown in a 1989 prospective study of the Swedish working population that showed that a combined self-rated measure of iso-strain was associated with a 9-year excess risk of cardiovascular mortality in men in the upper tertile of job strain, compared to those in the lowest tertile (relative risk of dying 1.77, with 95% confidence interval 1.28–2.44). Reviews of the demand-control model have been published, for instance, by Belkic and colleagues, who performed a methodological analysis of the sources of error and strengths in published studies. Their conclusion was that there is convincing evidence that job strain is a risk factor for cardiovascular disease.

There have also been prospective studies, however, with negative findings. One possible explanation for the discrepancy in findings is the age factor because the negative studies have mainly been long-term follow-up studies of relatively old employees. When a large proportion of the participants in a study are relatively old, they may retire during the follow-up

period, and it has been shown that the effect of exposure to job strain diminishes after retirement. A prospective European study by Kornitzer and colleagues with a 5-year follow-up of 20,435 middle-age men previously free from coronary heart disease has recently been made. One hundred and eighty of these men had myocardial infarctions during follow-up. When comparisons were made with the relaxed group, there was clearly an excess risk of myocardial infarction in the job strain group (after adjustments odds ratio 1.5, with 95% confidence limits 1.0–2.3) but not in the active or passive groups.

In prospective cohort studies with full control of retirements, clear associations have been shown. For instance, in the 2002 Finnish Valmet study, the excess cardiovascular mortality risk during follow-up was 2.2-fold (95% confidence interval 1.2–4.2) after adjustments for other risk factors.

The findings for women have been less consistent than findings for men. It has been argued that the demand-control model may be less relevant for women than for men. This could be true. However, direct comparisons between men and women are difficult because of the smaller female samples and because in several countries women work part-time more often than men.

Demand-Control and Effort-Reward: Independent Predictions?

Two studies of cardiovascular disease, the 1998 prospective Whitehall II study and the 2004 Stockholm SHEEP study have used the effort-reward imbalance (see **Effort-Reward Imbalance Model**) and the demand-control models together, and the findings indicated that they predict coronary heart disease episodes partly independently of one another. In the Whitehall II study of British civil servants, the decision authority component of the demand-control model and the effort-reward imbalance model both made independent predictions of new episodes of coronary heart disease among previously healthy state employees. In the Stockholm SHEEP study, a large population-based case control study of first myocardial infarctions, self-rated job strain as well as the intrinsic and proxy measures of the extrinsic parts of the effort-reward model were studied as separate variables. The results showed different results for men and women. For men, a combination of job strain and imbalance between extrinsic effort and reward was the best predictor of myocardial infarction status. In women, the intrinsic part of the effort-reward model (overcommitment) had the same role as the extrinsic one did in men – a combination of overcommitment and job strain was the best predictor.

It is not known to what extent self-reported assessments of the psychosocial work environment (on which the majority of research reports have been based) reflect individual characteristics (which may distort the perception of reality) and to what extent they reflect true environmental conditions. Critics argue that subjectivity bias may explain most of the observed associations between the psychosocial work environment and coronary heart disease. The general impression, however, is that objective assessments of the demand-control-support dimensions predict in ways similar to the self-reported more subjective ratings, although the associations are attenuated.

In their study of imputed and self-reported assessments of demand and decision latitude in first myocardial infarction cases and referents (SHEEP), Theorell and colleagues showed that there was no evidence of recall bias (distortion that could cause false associations) for decision latitude self-reports in the myocardial infarction cases. The analysis of demands was more difficult because the relationship between self-reports and imputed scores was much weaker than for decision latitude.

Other Health Outcomes

Mental disease outcomes have also been studied in relation to the demand-control-support model. A prospective design seems to be particularly important when mental disease is studied because a slow onset of a mental disease could influence the work environment and the relationships could be affected by this. There are a few such published studies. Stansfeld and colleagues found in their follow-up of British civil servants that low decision latitude, high demands, and poor social support were all related to an elevated risk of development of new mental illnesses (defined as a bad score on the 30-question General Health Questionnaire). Niedhammer and colleagues reported similar findings on the development of new depressive symptoms during follow-up of a cohort of French workers (GAZEL study). All the three dimensions in the model – high demands, low decision latitude and poor support – were associated with a risk of development of depressive symptoms.

Findings with regard to musculoskeletal disorders have been more mixed in the literature. In most studies, at least one of the three dimensions has a relationship with the risk of developing musculoskeletal disorders. The psychosocial factors interact in complicated ways with physical demands, and there is a debate as to whether or not an adjustment should be made for physical demands when the association between psychosocial job factors and musculoskeletal disorders is studied. In addition, the relationships

seem to be different in different occupational groups, in men versus women, and in different cultural settings and for different diagnoses (low back pain and neck/shoulder pain, for instance).

Other illnesses have also been studied. For instance, Westerberg and Theorell showed that there is an association between low decision latitude and functional gastrointestinal disorder. In men, a follow-up of individuals with such episodes showed that in men (but not in women) good support from colleagues was associated with rapid recovery.

Physiological Mechanisms

The psychosocial work environment could influence illness risks either through influences on lifestyle factors such as cigarette smoking, physical activity, and eating habits or through more direct physiological mechanisms. A possible indirect way of studying the effect of the demand-control-support model on lifestyle factors is to explore the association among the model, cigarette smoking, and serum lipids. Such studies have shown that low decision latitude is related to a high prevalence of smoking and high serum lipids and that this association holds even after adjustment for social class and education.

Most researchers presently agree that all the association between job strain and risk for cardiovascular illness cannot be explained by lifestyle-related risk factors. Therefore, other mechanisms more directly related to biological stress reactions have been explored. Blood pressure variations measured automatically at repeated intervals during the 24-h circadian cycle have been related to job strain in several studies. A high systolic and diastolic blood pressure during work hours has been shown in job strain in several studies, and in some studies this elevation has extended also to leisure hours and to sleep. A high morning cortisol, especially half an hour after awakening, seems to be a characteristic of the job strain day. Recently, coagulation and inflammation parameters have been explored. Several studies have shown a relationship between job strain and/or low decision latitude and plasma fibrinogen, even after adjustment for other risk factors. Interleukins and other markers of inflammation are presently being studied. Longitudinal studies of variations in demands, decision latitude, and social support have been performed. These studies have shown that increasing job strain is associated with increasing systolic blood pressure, decreasing sleep quality, and, in men, lowered serum testosterone. Karasek and co-workers have started exploring the associations between intra-individual variations in decision latitude during the daily round of life and variations in autonomic indices as

they are reflected in 24-h recordings of heart-rate variability.

In summary, it is likely that the working conditions associated with job strain are associated with increased energy mobilization and decreased anabolism. Both of these processes may increase cardiovascular disease risk.

Childhood Confounding

Theoretically, there is a possibility that adverse material childhood circumstances (infections, nutrition, drinking water, etc.) could explain the relationship between psychosocial work conditions and coronary heart disease. The argument would be that subjects with a poor childhood are more likely than others to end up in bad jobs with poor psychosocial environments. Life course research is beginning to address these kinds of questions. For instance, in the Valmet study the prospective relationships between work stress and risk of cardiovascular death were adjusted for a number of childhood factors, and both job strain and effort-reward imbalance remained as significant predictors even after these adjustments.

Examples of Organizational Interventions in Workplaces Aiming at Decreased Heart Disease Risk

If low decision latitude, high demands, and poor social support contribute to health and disease, they represent conditions that could be the focus of interventions. For several reasons, the health effects of such interventions are difficult to evaluate. However, particularly in the Scandinavian setting, evaluations of the health effects of organizational changes aiming at improved worker participation have been made.

The Stockholm group has performed an intervention study aiming at improved psychosocial knowledge in managers. Managers in an insurance company had mandatory psychosocial education once every second week (a half an hour lecture and a 90-min group discussion) for a whole year. The education program comprised all relevant aspects of psychosocial working conditions, such as the role of demand, decision latitude, social support, and effort-reward imbalance. Their employees were examined before the study, after half a year, and after a whole year with regard to psychosocial work conditions and serum cortisol (when they arrived to the office in the morning). Employees in another comparable part of the same organization (with managers not subjected to the psychosocial training) were followed at the same intervals (130 subjects in each group). Whereas cortisol levels remained unchanged in the comparison

group, the employees in the intervention group had a substantial significant decrease in serum cortisol during the follow-up year. Psychosocial questionnaire data from the same groups of employees indicated that the development of decision authority was more favorable in the intervention group than in the control group, although demands and work pace developed in the same way in the two groups. These results indicate that managers could be one target group in psychosocial work-site interventions and that the improvement of decision authority for employees may be a crucial variable. Bond and Bunce published a British study in which three workplaces were randomly allocated to intervention and three similar workplaces were assigned to the control condition. The intervention followed the principles of participation activation research (PAR), which aims at improved employee participation in decisions at the workplace. A 1-year follow-up showed that the employees in the intervention group had lowered amounts of sick leave and improved self-ratings of performance and mental health. Such changes were not seen in the control group, and statistics showed that the improvement in employee participation accounted for nearly all of this.

In a psychosocial intervention program in Sweden, a similar strategy was used, which involved all the employees in the workplaces. Compared to the control group, the intervention group showed improved decision authority and improved lipoprotein patterns (which are associated with decreased cardiovascular risk).

Critique and Place of the Demand-Control-Support Model

There have been several criticisms of the demand-control-support model. We present the most important ones here.

Old Model in a Changing World

Society is changing, and the basic concepts in the demand-control-support model may not be as important as they were in the past. The model is 30 years old and was formulated on the basis of a society with a large industry sector dominated by production lines. It has turned out, however, that the model is applicable to most kinds of jobs in society.

Another argument is that it is possible that the increasing demands in people's work life that have been documented in North American and European work environment surveys during the 1990s and early 2000s will make the model invalid in the end. If demands are extreme in all parts of the work life, there will be no opportunity for employees to use

the decision latitude that they are given in the system. At least in European men, this had not happened during the latter half of the 1990s because the recent study by Kornitzer shows the same order of magnitude of excess myocardial infarction risk that was observed previously. For mental disorders, there are very few prospective studies, but the British Whitehall II follow-up study from the late 1990s and the French GAZEL study both show that the model is relevant for mental disease in both men and women during this period. It is true, however, that the world of work is constantly changing and that there is a need for new questions and new dimensions to be added. It has been pointed out, for instance, that the large changes on the company level that are becoming increasingly frequent belong to a high level of decision latitude that is not captured by the instruments that are presently used (control over). A concept that is related to decision latitude, justice, has recently been introduced in this research field.

A theoretical model has to achieve the right balance between generality and specificity. The model has been sufficiently general to generate theoretical interest but not so general that it is trivial. For practitioners exploring work-site conditions, it will never be enough to use general demand, decision latitude, and support questions. We have to formulate more specific questions for the individual work site as well.

Psychometric Difficulties

Although the three dimensions in the model mostly form distinct entities in factor analyses based on mixed working populations, they may still be psychometrically imprecise. The demand dimension contains emotional and quantitative demands, for instance. These kinds of demands may mean very different things in different groups. Decision latitude has two components that behave slightly differently in different populations. There has been a tendency among researchers during later years to focus on decision authority only rather than on decision latitude (the combination of skill discretion and decision authority). In social support, finally, there is support both from superiors and from workmates, and these components could also have differential impacts on health. Accordingly, there is more work to do in improving the instruments. Today there are many other sets of questionnaires that could be used in the assessment of demand, decision latitude, and support at work.

No True Interaction among the Components

The three dimensions in the model seem to operate relatively independently of one another. There are

very few published studies that have shown a true multiplicative interaction between demands and decision latitude in generating risk for illness. This is true, but the combination of high demands and low decision latitude has been a more precise risk indicator in most of the epidemiological studies than low decision latitude or high demands separately.

No Individual Component

The demand-control-support model has its focus on the environment, and its use has been intended to be a tool in the improvement of work design. Of course, the individual is also important in the pathogenesis of bad health. Assessments of the model components could certainly be supplemented by assessments on the individual level.

The demand-control-support model has been of great importance to the psychosocial work environment field because it has stimulated a large number of research projects and generated increased interest in job stress and illness during the past 15 years. It captures both negative and positive aspects and has great value in practical work at real work sites.

See Also the Following Articles

Workplace Stress; Effort-Reward Imbalance Model.

Further Reading

- Belkic, K. L., Landsbergis, P. A., Schnall, P. L., et al. (2004). Is job strain a major source of cardiovascular disease risk? *Scandinavian Journal of Work, Environment & Health* 30(2), 85–128.
- Bond, F. W. and Bunce, D. (2001). Job control mediates change in work organization intervention for stress reduction. *Journal of Occupational Health Psychology* 6, 290–302.
- Bosma, H., Peter, R., Siegrist, J., et al. (1998). Two alternative job stress models and the risk of coronary heart disease. *American Journal of Public Health* 88(1), 68–74.
- Brunner, E. J., Kivimäki, M., Siegrist, J., et al. (2004). Is the effect of work stress confounded by socio-economic factors in the Valmet study? *Journal of Epidemiology and Community Medicine* 58(12), 1019–1020.
- Collins, S. M., Karasek, R. A. and Costas, K. (2005). Job strain and autonomic indices of cardiovascular disease risk. *American Journal of Industrial Medicine* 48(3), 182–193.
- Demerouti, E., Bakker, A. B., Nachreiner, F., et al. (2001). The job demands-resources model of burnout. *Journal of Applied Psychology* 86(3), 499–512.
- Gardell, B. (1982). Worker participation and autonomy: a multilevel approach to democracy at the workplace. *International Journal of Health Services* 12, 527–558.
- Greiner, B. A., Krause, N., Ragland, D., et al. (2004). Occupational stressors and hypertension: a multi-method study using observer-based job analysis and self-reports in urban transit operators. *Social Science & Medicine* 59(5), 1081–1094.
- Johnson, J. V. and Hall, E. M. (1988). Job strain, workplace social support and cardiovascular disease: a cross-sectional study of a random sample of the Swedish working population. *American Journal of Public Health* 78, 1336–1342.
- Johnson, J., Stewart, W., Hall, E., et al. (1996). Long-term psychosocial work environment and cardiovascular mortality among Swedish men. *American Journal of Public Health* 86, 324–331.
- Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: implications for job redesign. *Administrative Science Quarterly* 24, 285–307.
- Karasek, R. A., Brisson, C., Kawakami, N., et al. (1998). The job content questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics. *Journal of Occupational Health Psychology* 3, 322–355.
- Karasek, R. A. and Theorell, T. (1990). *Healthy work*. New York: Basic Books.
- Kivimäki, M., Leino-Arjas, P., Luukkainen, R., et al. (2002). Work stress and risk of cardiovascular mortality: prospective cohort study of industrial employees. *British Medical Journal* 325(7369), 857.
- Kristensen, T. S., Bjorner, J. B., Christensen, K. B., et al. (2004). The distinctions between work pace and working hours in the measurement of quantitative demands at work. *Work and Stress* 18, 305–322.
- Niedhammer, I., Goldberg, M., Leclerc, A., et al. (1998). Psychosocial factors at work and subsequent depressive symptoms in the GAZEL cohort. *Scandinavian Journal of Work, Environment & Health* 24, 197–205.
- Orth-Gomér, K., Eriksson, I., Moser, V., et al. (1994). Lipid lowering through work stress reduction. *International Journal of Behavioral Medicine* 1(3), 204–214.
- Peter, R., Siegrist, J., Hallqvist, J., et al. (2002). Psychosocial work environment and myocardial infarction: improving risk estimation by combining two complementary job stress models in the SHEEP Study. *Journal of Epidemiology and Community Health* 56, 294–300.
- Schnall, P. L., Belkic, K., Landsbergis, P., et al. (2000). Why the workplace and cardiovascular disease? *Occupational Medicine* 15, 1–16.
- Stansfeld, S. A., Fuhrer, R., Shipley, M. J., et al. (1999). Work characteristics predict psychiatric disorder: prospective results from the Whitehall II Study. *Occupational and Environmental Medicine* 56(5), 302–307.
- Steptoe, A., Cropley, M., Griffith, J., et al. (2000). Job strain and anger expression predict early morning elevations in salivary cortisol. *Psychosomatic Medicine* 62(2), 286–292.
- Theorell, T. (2002). Job stress and fibrinogen (Editorial). *European Heart Journal* 23, 1799–1801.
- Theorell, T., Emdad, R., Arnetz, B., et al. (2001). Employee effects of an educational program for managers

- at an insurance company. *Psychosomatic Medicine* 63, 724–733.
- Theorell, T., Karasek, R. A. and Eneroth, P. (1990). Job strain variations in relation to plasma testosterone fluctuations in working men – a longitudinal study. *Journal of International Medicine* 227, 31–36.
- Theorell, T., Tsutsumi, A., Hallquist, J., et al. (1998). Decision latitude, job strain, and myocardial infarction: a study of working men in Stockholm. *American Journal of Public Health* 88, 382–388.
- Wahlstedt, K. (2001). Postal work – work organizational changes as tools to improve health. *Acta Universitatis Upsaliensis*.
- Westerberg, L. and Theorell, T. (1997). Working conditions and family situation in relation to functional gastrointestinal disorders: the Swedish Dyspepsia Project. *Scandinavian Journal of Primary Health Care* 15(2), 76–81.

Dementia and Stress See: Glucocorticoids – Adverse Effects on the Nervous System; Metals, Oxidative Stress and Brain Biology; Oxidative Stress and Aging.

Denial See: Emotional Inhibition.

Dental Stress

T K Fábián, G Fábián and P Fejérdy

Semmelweis University Budapest, Budapest, Hungary

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by T K Fábián and G Fábián, volume 1, pp 657–659, © 2000, Elsevier Inc.

Introduction

Psychological Coupling of Mouth and Teeth
Dental Fear, Phobic Reactions, and Panic Disorder
Psychosomatic Manifestations in Dentistry
Prevention and Treatment Possibilities

Glossary

Biofeedback Use of sensors and instrumentation capable of detecting, amplifying, and displaying a function-relevant biological signal to achieve self-regulation of a given physiological function (e.g., muscle tonus in relaxation).

Conscious sedation Deep sedation for minor operative procedures in which the administered drug

(benzodiazepine) renders the patient calm, conscious, and cooperative; however, there is often total amnesia of the operation.

Denture intolerance

Appearance of psychogenic symptoms reported by the patient in relation to fixed or removable dentures or to the treatment procedure, with the absence of detectable pathological conditions of the related tissues and the absence of detectable insufficiency of the denture.

Mental hygiene

Interdisciplinary view of maintaining mental health and preventing mental sickness during all kinds of human services (e.g., medicine, dentistry, teaching, pastoral counseling, social work, justice).

Odontophobia

Phobic reactions related to dental treatment, the dentist, or the dental surgery office.

Oral stage

Highly important early phase of psychological development in which breastfeeding plays a prominent role in communication between mother and infant.

Orofacial

Localized in, or related to, the hard or soft tissues of the mouth, tongue, teeth, or face.

<i>Panic disorder</i>	Disorder with the symptom of recurrent intensive feelings of alarm linked with various bodily reactions and fear of death.
<i>Photoacoustic stimulation</i>	The simultaneous stimuli of flashing lights delivered through closed eyes and rhythmic noises.

Introduction

Stress-related problems in dentistry are a collection of various psychological conditions. The most common and widely known phenomenon is dental fear, which may lead to phobic reactions or panic attack in some cases. Another large group of stress-related reactions in dentistry is formed by several orofacial psychosomatic manifestations based on the rich psychological coupling of the mouth and teeth.

Psychological Coupling of Mouth and Teeth

The mouth and teeth are important aspects of an individual's facial aesthetics and sexual characteristics and play an important role in speech and meta-communication (beginning as early as the oral stage of infancy). The mouth, teeth, and tongue take part in sexual contact and act as organs of the senses of touch and taste, leading to a reach representation of these organs in the central nervous system. In addition, they have a strong symbolic meaning related to sexuality, strength, and aggressiveness. Tooth loss, especially edentulousness, symbolizes the loss of sexual characteristics and the living force and evokes a symbolic meaning of growing old, evanescence, and death. Another important psychosomatic coupling is that masticatory muscle tonus seems to be more strongly coupled to psychological stress than other muscles of the body; salivary secretion (including excretion of defense proteins such as secretory immunoglobulin A (sIgA) and the salivary chaperone Hsp70) is also well known to be strongly influenced by psychological stress. These multiple psychological couplings of the mouth, teeth, tongue, masticatory muscles, and saliva are responsible for the reach psychological phenomena appearing in the field of dentistry.

Dental Fear, Phobic Reactions, and Panic Disorder

Problems relating to dental fear primarily occur before the age of 20. Most patients suffering from dental fear experience at least partial remission at times, and a large proportion of these patients are

aware of having been exposed to fear-provoking dental treatment in the past. In some cases, an unconscious psychological trauma may also be a cause of such reactions. Patients with dental fear, including patients having phobic reactions (odontophobia), are usually agreeable to using medications and/or several psychological techniques to reduce fear during dental treatment. Patients with panic disorder and patients fearing injections (needle phobia) are exceptions, however, because they usually recognize their indisposition (usually simple collapse or panic attack) as a result of a supposed life-threatening allergic reaction to injected anesthetics, which develops as a rigid, uncompromising behavior. The differential diagnosis of a real allergy to local anesthetics, panic disorder, or needle phobia is extremely important. True allergy can cause life-threatening anaphylactic shock. In such patients, the use of local anesthetics should be strictly avoided. In the case of panic disorder, some data in the literature suggest that the immune reaction regulated by IgE can be increased, which means a possible higher risk of anaphylactic shock. Because of this, great care should be exercised if using local anesthetics with these patients. For needle-phobic patients, the contraindication for the use of local anesthetics (if any) is only psychological.

Psychosomatic Manifestations in Dentistry

Because of the multiple psychological couplings and importance of the mouth and teeth, many of the psychopathological mechanisms related to sexuality, aggressivity, autoaggressivity, or death anxiety can lead to orofacial manifestations, especially if these important psychological and symbolic functions are damaged by tooth or mouth disorders. Acute psychological stress conditions (e.g., existential trauma, workplace problems, relationship problems with the sexual partner) or chronic conditions (e.g., depression, neuroses, chronic anxiety, death anxiety, schizophrenic or paranoid reactions) may serve as a background. These manifestations more frequently appear in women, usually in the second half of life, although they can appear in children or in young adults as well. The most frequent symptoms are atypical facial pain, burning mouth syndrome, myofascial pain, temporomandibular dysfunction, bruxism and other parafunctions, gagging, psychogenic taste disorders, certain recurrent oral ulcerations or inflammations, some oral allergic reactions, psychogenic occlusal problems, tic, psychogenic salivation problems, and oral discomfort. The symptoms may appear singly or in combination. Such symptoms appear frequently in orthodontic and prosthodontic

treatments (psychogenic denture intolerance), since orthodontic treatments and dentures have a great impact on the aesthetic, sexual, nutritional, phonetic, and, consequently, symbolic function of the teeth. In addition, orthodontic and prosthodontic treatments can be rather uncomfortable, extremely expensive, and time-consuming, which can cause the patient to develop strong emotions toward the dentist (and/or assistant). These factors may induce pressure, aggression, and complication of the dentist–patient–assistant relationship.

Prevention and Treatment Possibilities

Prevention of stress-related problems can be based on introducing a framework of mental hygiene to the dental treatment, including skilled communication with the patient and monitoring of the patient–dentist–assistant interpersonal relationships. Great care should be taken to acquiesce to the wishes of the patient regarding his or her treatment, both conscious and unconscious. Conscious wishes (e.g., to achieve a nice smile, for treatment to be inexpensive and painless) are usually easily detectable. Understanding unconscious wishes (e.g., to look younger, to stop the appearance of aging, to be loved by the dentist) may be more challenging but is similarly important for preventing the manifestation of symptoms.

Several possibilities for treatment can be considered. In the case of dental fear, sedation is recommended. Short-acting benzodiazepines are the most frequently used drug, either intraorally as a pre-medication or intravenously for deep (conscious) sedation during treatment. A mixture of nitrous oxide and oxygen is also frequently used for sedation during treatment; it is often advantageously combined with several hypnosis techniques. Hypnosis or relaxation techniques as single therapies can be effective as well.

In the case of psychosomatic manifestations, the patient usually regards his or her symptoms as a somatic or technical problem and asks for a repetition or correction of previous dental treatment. An early diagnosis is crucial, because repeated somatic dental treatment worsens the prognosis, fortifies the patient's false conception, and may lead to further loss of oral tissue. Special techniques should be used to lead the patient to accept psychotherapy or psychiatric treatment. For this purpose, psychotherapy-oriented discussions should be gradually introduced during placebo and/or palliative dental treatment (alleviating or relieving symptoms without surgery). Relaxation, biofeedback, and hypnotic techniques can be useful as well, most advantageously

as a combination with a form of physiotherapy (e.g., magnetotherapy, transcutaneous electrical nerve stimulation (TENS), photoacoustic stimulation, massage). For older patients, a religious-based treatment (e.g., pastoral therapy, pastoral counseling) may also be needed as an effective means of treating death anxiety-related manifestations.

In many cases, a dentist skilled in the above methods can effect a notable improvement in the patient's symptoms or even complete recovery. In other instances, a referral to a psychotherapist or psychiatrist may be necessary.

See Also the Following Articles

Anxiety; Fear; Hypnosis; Pain; Psychosomatic Medicine; Relaxation Techniques; Somatic Disorders; Surgery and Stress.

Further Reading

- Berggren, U., Hakeberg, M. and Carlson, G. (2000). Relaxation vs. cognitively oriented therapies for dental fear. *Journal of Dental Research* 79, 1645–1651.
- Burton, R. C. (1969). The problem of facial pain. *Journal of the American Dental Association* 79, 93–101.
- Clark, M. S. and Brunick, A. L. (1999). *Handbook of nitrous oxide and oxygen sedation*. St. Louis, MO: Mosby.
- Fábián, T. K. and Fábián, G. (1998). Stress of life, stress of death: anxiety in dentistry from the viewpoint of hypnotherapy. *Annals of the New York Academy of Sciences* 851, 495–500.
- Fábián, T. K., Tóth, Z., Fejérdy, L., et al. (2004). Photoacoustic stimulation increases the amount of 70 kDa heat shock protein (Hsp70) in human whole saliva: a pilot study. *International Journal of Psychophysiology* 52, 211–216.
- Green, C. S., Olson, R. E. and Laskin, D. M. (1982). Psychological factors in the etiology, progression, and treatment of MPD syndrome. *Journal of the American Dental Association* 105, 443–448.
- Kent, G. G. and Blinkhorn, A. S. (1991). *The psychology of dental care* (2nd edn.). Oxford: Wright-Butterworth-Heinemann.
- Marxkors, R. and Müller-Fahlbusch, H. (1981). Zur Diagnose psychosomatischer Störung in der zahnärztlich-prothetischen Praxis. *Deutsche Zahnärztliche Zeitschrift* 36, 787–790.
- Meechan, J. G., Robb, N. D. and Seymour, R. A. (1998). *Pain and the anxiety control for the conscious dental patient*. Oxford: Oxford University Press.
- Mehrstedt, M. and Wikström, P.-O. (eds.) (1988). *Hypnosis in dentistry*. Munich: Milton Erickson Society.
- Milgrom, P., Fiset, L., Melnick, S., et al. (1988). The prevalence and practice management consequences of dental fear in a major US city. *Journal of the American Dental Association* 116, 641–647.

- Pilling, L. F. (1983). Psychiatric aspects of diagnosis and treatment. In: Laney, W. R. & Gibilisco, J. A. (eds.) *Diagnosis and treatment in prosthodontics*, pp. 129–141. Philadelphia, PA: Lea & Febiger.
- Schmierer, A. (1997). *Einführung in der zahnärztliche Hypnose* (2nd edn.). Berlin: Quintessenz Verlags.

- Sergl, H. G. (ed.) (1996). *Psychologie Psychosomatik in der Zahnheilkunde*. München: Urban & Schwarzenberg.
- Staats, J. and Krause, W.-R (1995). *Hypnotherapie in der zahnärztlichen Praxis*. Heidelberg: Hüthig Ltd.

Depersonalization: Systematic Assessment

M Steinberg

Northampton, MA, USA

© 2007 Published by Elsevier Inc.

Definition and Characteristics

Etiology

Assessment with the SCID-D-R

Case Study

Conclusions

Glossary

<i>Amnesia</i>	A specific and significant block of past time that cannot be accounted for by memory.
<i>Depersonalization</i>	Detachment from one's self, e.g., a sense of looking at one's self as if one were an outsider.
<i>Derealization</i>	A feeling that one's surroundings are strange or unreal. Often involves previously familiar people.
<i>Dissociation</i>	Disruption in the usually integrated functions of conscious memory, identity, or perception of the environment. The disturbance may be sudden or gradual, transient or chronic.
<i>Identity alteration</i>	Objective behavior indicating the assumption of different identities, much more distinct than different roles.
<i>Identity confusion</i>	Subjective feelings of uncertainty, puzzlement, or conflict about one's identity.

Depersonalization is characterized by a sense of detachment from the self. The symptom itself may manifest in a variety of axis I or axis II psychiatric disorders. The *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (DSM-IV) prefers to define depersonalization not in functional or nosologic terms, but phenomenologically, as in an alteration

in the perception or experience of the self. The sense of detachment itself may be experienced in various ways. Commonly it appears as out-of-body experiences giving a sense of division into a participating and an observing self, resulting in the sense of going through life as though one were a machine or robot. In some cases, there exists a feeling that one's limbs are changing in size or are separated from the body.

It is important to distinguish between recurrent severe depersonalization that is characteristic of the dissociative disorders, including depersonalization disorder, and mild or moderate episodic depersonalization sometimes observed in patients with other nondissociative axis I or II disorders, and the isolated episode experienced by persons in the healthy population (normal controls).

Definition and Characteristics

Although depersonalization was first described in 1872, it was not named until 1898, when Dugas contrasted the feeling of loss of the ego with a real loss. In 1954, Ackner remedied the lack of clear-defined boundaries of the symptom by describing the four salient features: (1) feeling of unreality or strangeness regarding the self, (2) retention of insight and lack of delusional elaboration, (3) affective disturbance resulting in loss of all affective response except discomfort over the depersonalization, and (4) an unpleasant quality that varies in intensity inversely with the patient's familiarity with the symptom. Steinberg defined depersonalization as one of five core symptoms of dissociation (see above), the other four consisting of amnesia, derealization, identity confusion, and identity alteration. Each of the five dissociative disorders has characteristic symptom profiles of these core dissociative symptoms. For this reason, it is essential that the symptom of depersonalization is evaluated within the context of the other dissociative symptoms and not as an isolated symptom.

Episodes of depersonalization accompany or even may precipitate panic attacks and/or agoraphobia; they may also be associated with dysphoria. Chronic depersonalization frequently results in the patient's acceptance of the symptoms, in a manner of resignation. Patients experience difficulty putting their experience into words, but often compare their feelings to such states as being high on drugs, seeing themselves from the outside, or floating in space and watching themselves. Other descriptions of depersonalization include feelings of being unreal, or in severe cases, include the feeling of being numb or dead, or the lack of all feeling, which may be attributed to and/or misdiagnosed as depression.

Depersonalization has been reported to be the third most common complaint among psychiatric patients, after depression and anxiety. Incidence of actual depersonalization has been difficult to determine because of (1) the relative strangeness of the symptoms, (2) the difficulty patients experience in communicating them, and (3) the lack, until recently, of diagnostic tools for the systematic assessment of depersonalization. Detection is further complicated by the fact that depersonalization is not accompanied by altered external or social behavior, but by an altered state of perception on the part of the patient.

Etiology

Various biological and psychodynamic theories have been advanced for the etiology of depersonalization: (1) physiological or anatomical disturbance, with feelings of depersonalization produced by temporal

lobe function and various metabolic and toxic states, (2) the result of a preformed functional response of the brain to overwhelming traumata, (3) a defense against painful and conflictual affects such as guilt, phobic anxiety, anger, rage, paranoia, primitive fusion fantasies, and exhibitionism, (4) a split between the observing and the participating self, allowing the patient to become a detached observer of the self, and (5) the result of a child's being raised in an environment that systematically fails to know some part of the child, who then experiences that part as tentative and as a result is unable to accurately assess the self.

Depersonalization has been reported to be a normal reaction to life-threatening events, such as accidents, serious illnesses, near-death experiences, anaphylactic reactions, and complications of surgery. Depersonalization is frequent among victims of sexual abuse, political imprisonment, torture, and cult indoctrination. Symptoms of depersonalization are often associated with hypnosis, hypnogogic and hypnopompic states, sleep deprivation, sensory deprivation, hyperventilation, and drug or alcohol abuse.

The development of specialized tests for the assessment of dissociation has led to an increase in investigations furthering our understanding of depersonalization. Depersonalization, as a brief, isolated symptom, is nonspecific and not necessarily pathognomonic of any clinical disorder. Research indicates that it is the persistence and nature of depersonalization that differentiates depersonalization in normal subjects versus persons with dissociative and nondissociative disorders. [Table 1](#) is useful for distinguishing between normal and pathological

Table 1 Distinguishing between normal and pathological depersonalization

<i>Common mild depersonalization</i>	<i>Transient depersonalization</i>	<i>Pathological depersonalization</i>
Context		
Occurs as an isolated symptom	Occurs as an isolated symptom	Occurs within a constellation of other dissociative or nondissociative symptoms or with ongoing interactive dialogue
Frequency		
One or few episodes	One or few episodes that are transient	Persistent or recurrent depersonalization
Duration		
Depersonalization episode is brief; lasts seconds to minutes	Depersonalization of limited duration (minutes to weeks)	Chronic and habitual depersonalization lasting up to months or years
Precipitating factors		
• Extreme fatigue • Sensory deprivation • Hypnogogic and hypnopompic states • Drug or alcohol intoxication • Sleep deprivation • Medical illness / toxic states • Severe psychosocial stress	• Life-threatening danger. This is a syndrome noted to occur in 33% of individuals immediately following exposure to life-threatening danger, such as near-death experiences and auto accidents (Noyes et al., 1977) • Single, severe psychological trauma	• Not associated with precipitating factors in column 1. • May be precipitated by a traumatic memory. • May be precipitated by a stressful event, but occurs even when there is no identifiable stress. • Occurs in the absence of a single immediate severe psychosocial trauma.

Table 2 The spectrum of depersonalization on the SCID-D-R

<i>DID and DDNOS</i>	<i>Non-dissociative and personality disorders</i>	<i>No psychiatric disorder</i>
Depersonalization questions elicit descriptions of identity confusion and alteration	No spontaneous elaboration	No spontaneous elaboration
Includes interactive dialogues between individual and depersonalized self	No interactive dialogues	No interactive dialogues
Recurrent–persistent	None–few episodes	None–few episodes

Note: DID = Dissociative Identity Disorder, DDNOS = Dissociative Disorder, Not Otherwise Specified.

Reprinted with permission from: Steinberg M: Interviewer's Guide to The Structured Clinical Interview for DSM-IV Dissociative Disorders – Revised. Washington, D.C. American Psychiatric Press, Second Edition, 1994.

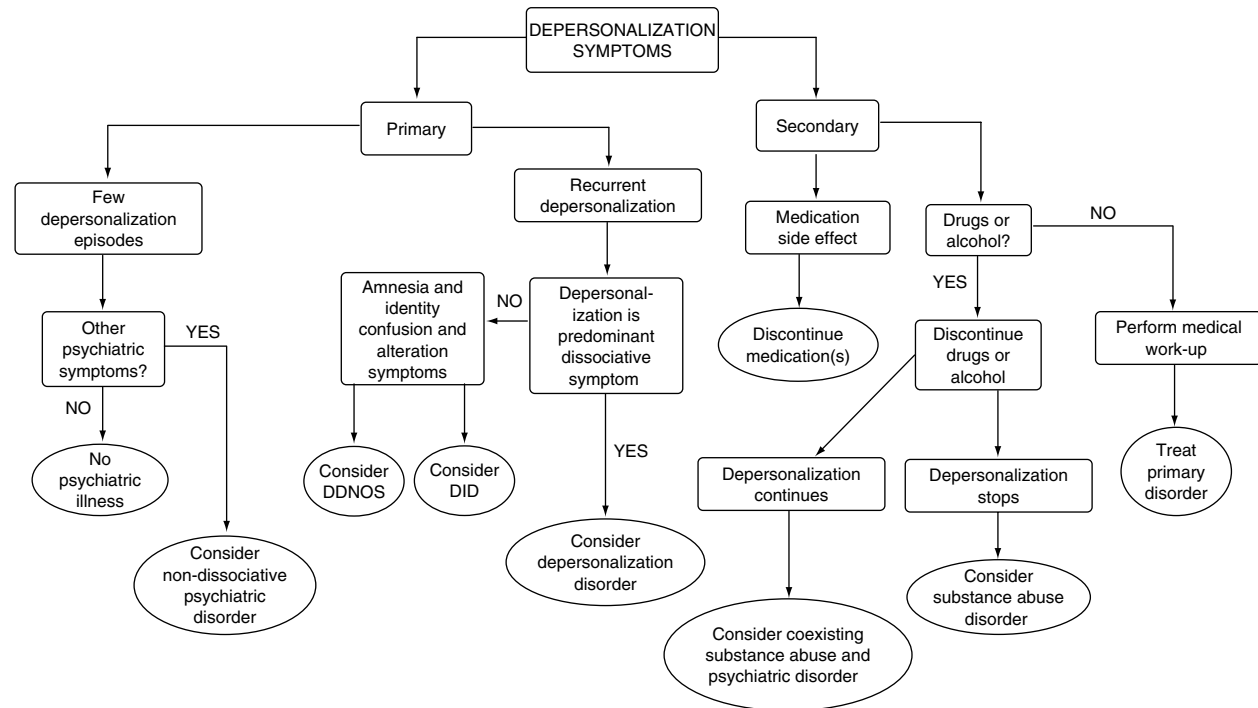


Figure 1 Differential diagnosis decision tree of depersonalization. Reprinted with permission from Steinberg, M (1994). *Interviewer's Guide to the Structured Clinical Interview for DSM-IV Dissociative Disorders – Revised*. Washington, DC: American Psychiatric Press.

depersonalization. [Table 2](#) summarizes the spectrum of depersonalization.

The differential diagnosis tree of depersonalization ([Figure 1](#)) illustrates procedures for distinguishing between depersonalization disorder and other disorders that may resemble it. The differential diagnosis of patients experiencing recurrent or persistent depersonalization should include the dissociative disorders, various other psychiatric disorders, and possible medical disorders/organic etiology, most commonly acute head trauma, seizure disorders, and migraines.

Assessment with the SCID-D-R

The Structured Clinical Interview for DSM-IV Dissociative Disorders – Revised (SCID-D-R) is a diagnostic

tool for the comprehensive assessment of dissociative symptoms and disorders, including the systematic identification of depersonalization. Developed in 1985 and extensively field tested, it is the only diagnostic instrument enabling a clinician to detect and assess the presence and severity of five core dissociative symptoms and the dissociative disorders (dissociative amnesia, dissociative fugue, depersonalization disorder, dissociative identity disorder, and dissociative disorder not otherwise identified) as defined by DSM-IV criteria. The SCID-D-R is a semistructured diagnostic interview with good-to-excellent interrater and test-retest reliability and discriminant validity. Guidelines for the administration, scoring, and interpretation of the SCID-D-R are reviewed in the *Interviewer's Guide to the SCID-D-R*. Severity rating definitions were developed to allow clinicians

to rate the severity of symptoms in a systematic manner and are included in the guide.

The SCID-D-R can be used for symptom documentation and for psychological and forensic reports. Early detection of dissociative disorders, including depersonalization disorder, can be realized from the use of this specialized instrument, the format of which includes open-ended questions designed to elicit spontaneous descriptions of endorsed dissociative symptoms. The SCID-D-R has been demonstrated to be a valuable tool in differential diagnosis with patients of different ages (it can be used in adolescents as well as adults), backgrounds, previous psychiatric histories, and presenting complaints. It also plays a useful role in treatment planning, patient follow-up, and symptom monitoring.

Correct diagnosis is vital to proper treatment of the disorder. If the depersonalization is secondary to an underlying primary disorder, the symptom may be alleviated by treatment of the underlying illness. In instances in which patients experience only occasional episodes of depersonalization in the context of other nondissociative symptoms, the clinician may consider a diagnosis of nondissociative psychiatric disorder. The presence of depersonalization disorder itself will be characterized by recurrent depersonalization.

Case Study

The process of differential diagnosis of depersonalization may best be illustrated by presenting a case study. The study demonstrates the utility of the SCID-D-R in diagnostic assessment, patient education, and treatment planning. For space reasons, conventional formatting and content have been abbreviated.

Sample SCID-D-R Psychological Evaluation

Demographic information and chief complaint: Susan Walker is a 31-year-old administrative assistant at a community college who presented with the complaint of feeling detached from herself since adolescence.

Past psychiatric history: Although the patient had no history of hospitalization for psychiatric disturbance, she began treatment for an episode of depression that interfered with her employment and social relationships. Although admitting to past casual use of marijuana, she had never been in treatment for substance abuse disorder.

Family history: Susan had a younger sibling; both children grew up in an intact but emotionally unsupportive family. Patient reported that both parents suffered mood swings and unpredictable temper outbursts.

Mental status exam: Susan answered questions with relevant replies; although she seemed slightly depressed, her affect appeared full range. She denied hallucinations, both auditory and visual, and evidenced no psychotic thinking. She denied acute suicidal or homicidal ideas.

SCID-D-R evaluation: The SCID-D-R was administered to systematically evaluate the patient's dissociative symptoms and was scored according to prescribed guidelines. Significant findings from the SCID-D-R interview follow. Susan denied experiencing severe episodes of amnesia, but endorsed a persistent sense of depersonalization, resulting in distress and interference with occupational and personal functioning. This feeling of depersonalization had been chronic and occurred all the time rather than episodically. Although the feeling varied in intensity with her overall stress level, the experience of depersonalization was always characterized by a general sense of detachment from life, rather than by disturbances in body image or a split between participating and observing parts of the self. Only a single isolated out-of-body experience had occurred. Susan experienced feelings of derealization that varied in intensity with the depersonalization, but she reported the depersonalization as the most distressing symptom. She described recurrent anxiety and panic episodes triggered by the depersonalization; it was the combination of depersonalization and panic attacks that led to the depression that brought her into therapy. Susan reported that the depersonalization has eroded her sense of control over her occupational functioning and other significant areas of her life, but she did not attribute feelings of loss of control to identity confusion or alteration. She denied having internal dialogues, feelings of possession, or acquiring unexplained possessions or skills. Her descriptions of internal struggle were focused on her feelings of unreality, not on conflicts between different aspects of her personality or different personalities within herself.

Assessment: Susan's symptoms are consistent with a primary diagnosis of a dissociative disorder based on DSM-IV criteria and ICD-10 criteria. Specifically, in the absence of substance abuse disorder or other organic etiology, her severe chronic feelings of unreality toward herself and the accompanying dysfunction (in the absence of other dissociative symptoms such as identity confusion and alteration) are consistent with a diagnosis of depersonalization disorder.

Recommendations: Although detailed discussion of treatment for depersonalization disorder is beyond the scope of this article, it would be standard practice to conduct a follow-up interview to review the findings of the SCID-D-R evaluation, to educate

the patient regarding her symptoms, and to begin the process of individual psychotherapy.

Conclusions

Recent advances in the development of reliable diagnostic tools allow for early detection and accurate differential diagnosis of depersonalization. Research based on the SCID-D-R indicates that depersonalization occurs in individuals without psychiatric illness who experience none to few brief episodes, as well as in individuals with dissociative disorders who experience recurrent to ongoing episodes. In addition to the frequency of the depersonalization, the nature, severity, and context also distinguish cases of dissociative disorder from other nondissociative disorders. Further research is necessary in the form of controlled double-blind studies comparing the efficacy of pharmacotherapeutic agents as well as psychotherapeutic techniques. As the SCID-D-R allows for the assessment of the severity of depersonalization based on operationalized criteria, pharmacotherapy trials can now be systematically performed and can evaluate baseline and postmedication severity levels of depersonalization. Given the frequency of misdiagnosis in patients suffering from depersonalization and other dissociative symptoms, earlier detection of dissociative disorders using the SCID-D-R can allow for rapid implementation of effective treatment.

See Also the Following Articles

Amnesia; Dissociation.

Further Reading

Ackner, B. (1954). Depersonalization I.: aetiology and phenomenology. *Journal of Mental Science* **100**, 838–853.
 American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th edn.). (DSM-IV). Washington, D.C.: American Psychiatric Association.

Baker, D., Hunter, E., Lawrence, E., et al. (2003). Depersonalization disorder: clinical features of 204 cases. *British Journal of Psychiatry* **182**, 428–433.
 Bowman, E. S. and Markand, O. (1996). Psychodynamics and psychiatric diagnoses of pseudoseizure subjects. *American Journal of Psychiatry* **153**(1), 57–63.
 Dugas, L. (1898). Un cas de depersonalization (A case of depersonalization). *Revue Philosophique* **45**, 500–507.
 Lambert, M., Sierra, M. and Phillips, M. D. (2002). The spectrum of organic depersonalization: a review plus four new cases. *Journal of Neuropsychiatry and Clinical Neuroscience* **14**(2), 141–154.
 Noyes, R., Jr. and Kletti, R. (1977). Depersonalization in response to life-threatening danger. *Comprehensive Psychiatry* **18**, 375–384.
 Simeon, D., Gross, S., Guralnik, O., Stein, D. J., Schmeidler, J. and Hollander, E. (1997). Feeling unreal: 30 cases of DSM-III-R depersonalization disorder. *American Journal of Psychiatry* **154**, 1107–1113.
 Steinberg, M. (1994). *The structured clinical interview for DSM-IV dissociative disorders – revised (SCID-D)* (2nd edn.). Washington, D.C.: American Psychiatric Press.
 Steinberg, M. (1994). *The interviewers' guide to the structured clinical interview for DSM-IV dissociative disorders – revised* (2nd edn.). Washington, D.C.: American Psychiatric Press.
 Steinberg, M. (1995). *Handbook for the assessment of dissociation: a clinical guide*. Washington, D.C.: American Psychiatric Press.
 Steinberg, M. (1995). Advances in the clinical assessment of dissociation: the SCID-D-R. *The Bulletin of the Menninger Clinic* **59**, 221–231.
 Steinberg, M. and Hall, P. (1997). The SCID-D diagnostic interview and treatment planning in dissociative disorders. *Bulletin of the Menninger Clinic* **61**(1), 108–120.
 Steinberg, M. and Schnall, M. (2001). *The stranger in the mirror: dissociation—the hidden epidemic*. HarperCollins: New York.
 Steinberg, M., Rounsaville, B., Buchanan, J., Raakfeldt, J. and Cicchetti, D. (1994). Distinguishing between multiple personality and schizophrenia using the Structured Clinical Interview for DSM-IV Dissociative Disorders. *Journal of Nervous and Mental Disorders* **Sept. 1994**, 495–502.

Depression and Coronary Heart Disease

F Lespérance

Centre Hospitalier de l'Université de Montréal Research Center, University of Montreal, and Montreal Heart Institute Research Center, Montreal, Canada

N Frasure-Smith

Centre Hospitalier de l'Université de Montréal Research Center, University of Montreal, McGill University, and Montreal Heart Institute Research Center, Montreal, Canada

© 2007 Elsevier Inc. All rights reserved.

Introduction

Definitions of Depression

The Course and Consequences of Depression in Coronary Heart Disease Patients

Mechanisms

Depression Treatment Trials in Coronary Heart Disease Patients

Conclusion

Glossary

<i>Autonomic nervous system</i>	The system that controls internal organs without a person being consciously aware. There are two parts to this system: the sympathetic nervous system and the parasympathetic (or vagal) nervous system.
<i>Cognitive-behavior therapy (CBT)</i>	A structured, usually time-limited, form of psychotherapy with active therapist involvement involving exercises and homework to alter dysfunctional thought patterns that are believed to underlie and perpetuate depression.
<i>Endothelial dysfunction</i> <i>Endothelium</i>	One of the earliest signs of the development of coronary heart disease. The lining of the arteries; responsible for arterial dilation and constriction in response to systemic demands and chemical stimuli.
<i>Heart-rate variability</i>	The degree to which the length of time between heart beats differs from beat to beat. Low heart-rate variability is a sign of inadequate vagal protection in the autonomic nervous system and predicts cardiac death in individuals with coronary heart disease.
<i>Parasympathetic nervous system</i>	Part of the autonomic nervous system that controls relaxation and promotes digestion; also called the vagal nervous system.
<i>Sympathetic nervous system</i>	Part of the autonomic nervous system that directs the fight-or-flight response to stress.

Sympathetic-vagal balance

The balance between the sympathetic nervous system and the parasympathetic (or vagal) nervous system; this balance is altered in depression.

Ventricular function

A measure of the heart's ability to pump blood; it is often reduced following myocardial infarction because of the damage that occurs to the heart muscle.

Introduction

Depression and coronary heart disease (CHD) are among the most prevalent disabling medical conditions worldwide. The Global Burden of Disease Study predicted that, by 2020, depression and CHD will be the two leading causes of early death and disability in the world. The term depression most frequently refers to a mental state with sustained sadness or loss of interest lasting at least several weeks, coupled with other symptoms such as fatigue, reduced sleep, and appetite and leading to significant occupational or interpersonal impairment. CHD includes all clinical manifestations associated with damage to the myocardium due to acute or chronic ischemia, which itself is most commonly caused by atherosclerotic narrowing of the coronary arteries. Interestingly, psychological stress, such as interpersonal conflicts, perceived or real losses, and work stress, are thought to contribute to the acute exacerbation of depression as well as of CHD.

Definitions of Depression

The *Diagnostic and Statistical Manual of Mental Disorders* (4th edn.; DSM-IV) of the American Psychiatric Association defines major depression as at least 2 weeks of daily (lasting most of the day) sadness or loss of interest plus at least five of the following symptoms: changes in appetite, changes in sleep patterns, psychomotor agitation or retardation, loss of energy or fatigue, feelings of worthlessness, inability to concentrate or make decisions, and thoughts of suicide or death.

Studies conducted in hospitalized cardiac patients have found that about one in six patients meet the criteria for a major depressive episode. This high peak prevalence during the hospitalization subsides, with approximately 10% of CHD patients depressed at 3 months after discharge. However, many patients do not meet the full criteria for major depression, but nevertheless present elevated levels of depressive

symptoms. These patients may have a variety of clinical diagnoses. Some may suffer from a major depression in partial remission or dysthymia (chronic low-level depression); others may suffer from an adjustment disorder. There are many different self-report scales available to measure the overall intensity of depressive symptoms. The Beck Depression Inventory (BDI) is the most commonly used in research with CHD patients.

Even if major depression is not present, however, having high levels of depressive symptoms increases a patient's risk of cardiovascular events. Although major depression and elevated depressive symptoms are approximately three times as common in CHD patients as in the general community, the prevalence of depression in CHD patients is similar to that in patients with other chronic medical conditions.

The Course and Consequences of Depression in Coronary Heart Disease Patients

As for depressed patients without medical illnesses, the course of depression in CHD patients is frequently chronic. More than half of those who experience major depression during hospitalization after a myocardial infarction (MI, or heart attack) remain depressed or have elevated symptoms of depression 1 year later.

However, as important as its chronicity is the fact that depression is associated with a worse cardiac prognosis. Several reviews and meta-analyses have concluded that both major depression and high levels of depressive symptoms predict cardiac mortality in patients with established CHD, as well as the incidence of CHD among previously healthy subjects. The impact of depression is independent of other prognostic variables; that is, it is not explained by other possible confounding factors. This means, for example, that depression predicts cardiac mortality even after statistical control for other predictors of mortality such as age, ventricular function, hypertension, and diabetes. In addition to its independent impact on prognosis, the size of the impact of depression on CHD events is as large as that of these other important prognostic factors. Depression at least doubles the risk of cardiac mortality over time, and the risk increases with the severity of depressive symptoms.

INTERHEART, a very large case-control study conducted in 52 countries across all continents compared levels of modifiable risk factors in more than 11,000 survivors of a first MI and more than 13,000 control subjects. The study found that the population attributable risk (PAR) for depression is 9%, a risk as large as that of diabetes. The PAR takes into account the prevalence and degree of risk associated with a

particular risk factor and is interpreted to mean that, if depression could be eliminated at the population level, we would observe a worldwide reduction of 9% in the incidence of first MIs.

The INTERHEART study also evaluated the contribution of different dimensions of stress to the incidence of MI. It found that all four dimensions of stress measured in the study (stress at work, stress at home, financial stress, and general stress) predicted the incidence of MI, with the PARs for stress ranging between 8 and 12%. In addition, the study found that a measure of locus of control, evaluating how much a person feels that he or she can control life circumstances, also predicted the incidence of MI, with a PAR of 16%. Finally, the occurrence of stressful life events in the last year, such as the death of a spouse or other family member, marital separation or divorce, or major familial conflict, was also more frequent among first MI patients than controls. In summary, this landmark study found that, on a worldwide basis, the contribution of psychosocial stress and depression to the incidence of MI was just behind that of smoking and was greater than that of abdominal obesity. INTERHEART also confirmed that the impact of depression and other psychosocial stresses is independent of age, smoking status, and other cardiovascular risk factors. It underscores the importance of the emotional and behavioral factors in physical health, and it should stimulate research to reduce depression, stress, and their consequences.

Mechanisms

Many biological pathways may be involved in the link between depression and CHD. Depression is similar to a state of chronic, severe psychological stress, and many of the biological systems involved in the stress response are dysregulated in depressed patients regardless of whether or not they have CHD. The sympathetic-vagal balance is tipped toward more noradrenergic activity, heart-rate variability is reduced, and there is an increased risk of ventricular arrhythmias. There is also an augmentation of the innate inflammatory response with more pro-inflammatory cytokines. Platelet activation is increased, facilitating blood-clot formation. There is hyperactivity of the hypothalamic-pituitary-adrenal (HPA) axis, which influences lipid and carbohydrate metabolism. In addition, it has been recently shown that depressed young adult patients without CHD have endothelial dysfunction, one of the early markers of atherosclerotic burden.

Because of the obvious role of the selective serotonin reuptake inhibitor (SSRI) antidepressants in mood regulation, there has been a great deal of interest in their potential cardiovascular benefits. Platelets contain serotonin granules that are secreted during the

platelet activation phase, a process that facilitates blood-clot formation through the activation of the platelet's own serotonin receptors. In addition, platelets have a specific serotonin transporter, which acts to move (reuptake) the secreted serotonin back inside the platelet. SSRIs block this transporter, therefore gradually depleting serotonin in the platelets. Consequently, after several weeks of treatment, SSRIs seem to have blood-thinning effects. However, it is unknown whether this effect is of enough clinical relevance to prevent thrombotic cardiovascular events.

It is also possible that some environmental or genetic factors contribute to the physiopathology of CHD and depression at the same time. For example, a lower dietary intake of long-chain omega-3 fatty acids has been shown to predict the incidence of CHD and a wide variety of mood disorders. There is hope to identify common genetic predispositions for mood disorders and CHD. Specifically, the genes involved in the regulation of the biological systems mentioned thus far (i.e., the sympathetic-parasympathetic system, the inflammatory response, the HPA axis, platelet activation, and thrombus formation) can be profiled and used to evaluate whether specific genetic variations predict the development of CHD, depression, or both. For example, the genes regulating serotonin synthesis, metabolism, and reuptake, such as the serotonin transporter-linked polymorphic region (5-HTTLPR), may be associated with an increased risk of depression because of its role in the brain and with an increased risk of cardiovascular events because of its role in platelet aggregation.

Depression Treatment Trials in Coronary Heart Disease Patients

With the current deluge of mega-trials in cardiology testing new drugs and devices to prevent cardiovascular events among CHD patients, it is very disappointing that only one trial was specifically designed to evaluate whether an intervention for depression has an impact on CHD events. This is, at least partially, because of lack of funding resources; a depression treatment-CHD trial would necessarily depend on peer-reviewed public funding, whereas most of the large phase III drug trials in cardiology are currently supported by the pharmaceutical industry.

ENRICH (Enhancing Recovery in Coronary Heart Disease) is the only trial carried out to date to evaluate the impact of depression treatment on cardiac events. ENRICH was a National Heart, Lung, and Blood Institute (NHLBI)-sponsored, multicenter trial that assessed the potential benefit of 6 months of CBT in some 2400 depressed post-MI patients. Although patients receiving CBT experienced a significantly greater improvement in depression than

those in the usual-care group, this study was unable to demonstrate that CBT prevented the recurrence of cardiac events. The physicians of patients in the usual-care group were informed that their patients were depressed, and there was a high rate of prescription of antidepressant medications in the usual-care group, reducing the contrast between the groups. Other earlier controlled clinical trials designed to treat psychological distress, rather than depression *per se*, were also not successful in demonstrating that treating distress reduces cardiac events.

Conclusion

Despite the extensive epidemiological data showing that an increased risk of cardiac events is associated with depression and considerable animal and human research suggesting that various physiopathological changes, similar to those observed in chronic severe stress, accompany depression, the data from controlled clinical trials do not support the use of any specific interventions for depression or psychological distress to prevent cardiac events in CHD patients.

See Also the Following Articles

Health Behavior and Stress; Heart Disease/Attack.

Further Reading

- Carney, R. M., Blumenthal, J. A., Freedland, K. E., et al. (2005). Low heart rate variability and the effect of depression on post-myocardial infarction mortality. *Archives of Internal Medicine* 165, 1486–1491.
- ENRICH Investigators. (2003). Effects of treating depression and low perceived social support on clinical events after myocardial infarction. *Journal of the American Medical Association* 289, 3106–3116.
- Fraser-Smith, N. and Lesperance, F. (2005). Reflections on depression as a cardiac risk factor. *Psychosomatic Medicine* 67(supplement 1), S19–S25.
- Fraser-Smith, N., Lesperance, F. and Julien, P. (2004). Major depression is associated with lower omega-3 fatty acid levels in patients with recent acute coronary syndromes. *Biological Psychiatry* 55, 891–896.
- Hibbeln, J. R. and Makino, K. K. (2002). Omega-3 fats in depressive disorders and violence: the context of evolution and cardiovascular health. In: Skinner, E. R. (ed.) *Brain lipids and disorders in biological psychiatry*, pp. 67–111. Amsterdam: Elsevier Science.
- Joynt, K. E., Whellan, D. J. and O'Connor, C. M. (2003). Depression and cardiovascular disease: mechanisms of interaction. *Biological Psychiatry* 54, 248–261.
- Kuper, H., Marmot, M. and Hemingway, H. (2002). Systematic review of prospective cohort studies of psychosocial factors in the etiology and prognosis of coronary heart disease. *Seminars in Vascular Medicine* 2, 267–314.

- Lespérance, F., Frasure-Smith, N., Theroux, P., et al. (2004). The association between major depression and levels of soluble intercellular adhesion molecule 1, interleukin-6, and C-reactive protein in patients with recent acute coronary syndromes. *American Journal of Psychiatry* 161, 271–277.
- Murray, C. J. L. and Lopez, A. D. (1997). Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. *Lancet* 49, 1436–1442.
- Musselman, D. L., Evans, D. L. and Nemeroff, C. B. (1998). The relationship of depression to cardiovascular disease: epidemiology, biology and treatment. *Archives of General Psychiatry* 55, 580–592.
- Rosengren, A., Hawken, S., Ounpuu, S., et al. (2004). Association of psychosocial risk factors with risk of acute myocardial infarction in 11,119 cases and 13,648 controls from 52 countries (the INTERHEART study): case-control study. *Lancet* 364, 953–962.
- Schiepers, O. J., Wichers, M. C. and Maes, M. (2005). Cytokines and major depression. *Progress in Neuropsychopharmacology and Biological Psychiatry* 29, 201–217.
- Serebruany, V. L. (2006). Selective serotonin reuptake inhibitors and increased bleeding risk: are we missing something? *American Journal of Medicine* 119, 113–116.
- Suls, J. and Bunde, J. (2005). Anger, anxiety, and depression as risk factors for cardiovascular disease: the problems and implications of overlapping affective dispositions. *Psychological Bulletin* 131, 260–300.
- Vale, S. (2005). Psychosocial stress and cardiovascular diseases. *Postgraduate Medical Journal* 81, 429–435.

Depression and Manic-Depressive Illness

R T Rubin and B J Carroll

VA Greater Los Angeles Healthcare System,
Los Angeles, CA, and Pacific Behavioral Research
Foundation, Carmel, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition
article by R T Rubin, volume 1, pp 666–674,
© 2000, Elsevier Inc.

Introduction

Clinical Characteristics of Major Depression and Manic-Depression

Biological Characteristics of Depression and Manic-Depression

Neuroimaging Studies in Depression and Manic-Depression

Brain Circuit Theories of Depression and Manic-Depression

Psychological Testing in Depression and Manic-Depression

Treatment of Major Depression and Manic-Depression

Conclusion

Glossary

<i>Affect</i>	The objective (observable) component of emotions (appearing depressed, angry, elated, etc.), in contrast to mood.
<i>Anti-depressants</i>	Drugs that are used in the treatment of depression.
<i>Bipolar illness (disorder)</i>	Manic-depressive illness (disorder).

Cerebral cortex The outer layer of the brain, composed mainly of nerve cells (neurons).

Cerebral ventricles Cerebrospinal fluid-filled spaces within the brain.

Electroconvulsive therapy (ECT) A treatment for severe depression or bipolar disorder; a brief electric current is passed through the brain to induce a grand mal convulsive seizure.

Mood The prevailing subjective (experienced) emotional tone (sadness, happiness, elation, anger, etc.), in contrast to affect. (Mood is to affect as climate is to weather.)

Neurotransmitters Chemicals that carry messages from one nerve cell to another.

Introduction

The term depression covers a spectrum of disorders. These range from temporarily feeling down about something that has gone wrong in one's life to long and severe depressions that have a genetic and biological basis and that often are incapacitating to the afflicted person in his or her occupational and social life. The term manic-depression likewise covers a spectrum of disorders from temperamental instability of mood (i.e., cyclothymia) to a severe psychotic disorder with recurrent episodes of mania and depression. These are inappropriate swings of mood and behavior in opposite directions – excessive elation as well as significant depression.

Depression and manic-depression are psychiatric disorders that currently are considered functional, in that there is no clearly understood central nervous

system (CNS) pathophysiology underlying them. This is a nomenclature in transition, however – as knowledge about the pathology of the brain in these and other psychiatric illnesses increases, the adjective “functional” conveys less and less meaning. Furthermore, when these mood disorders are so severe that they impair the daily life of the individual, somatic (physical) treatments are almost always required, including antidepressants, mood-stabilizing drugs such as lithium, and, when drug therapy fails, electroconvulsive therapy (ECT).

This article focuses on the more severe disorders of mood, in particular major depression (also called unipolar depression) and bipolar disorder. The CNS substrates of these illnesses are being examined in several ways, and their genetic underpinnings are slowly being defined. Structural and functional imaging studies of the brain are providing some anatomical localization, limited by the resolution of the imaging techniques and the particular imaging method used.

Several classes of antidepressant drugs are useful for treating major depression, but the interactions among chemical systems (neurotransmitters) in the brain make it difficult to predict which antidepressant will work best for a particular patient. Lithium is the most broadly effective treatment for bipolar illness, but a number of anticonvulsant and antipsychotic drugs, as well as antidepressants, also may be helpful in various clinical circumstances. Stress can play a major role in the severity and duration of both major depression and bipolar disorder. Moreover, the illnesses themselves, given the incapacitating nature of their symptoms, can be very stressful to afflicted people. In both disorders, one's psychological pain and despair can be so severe that suicide may be contemplated and even attempted.

Clinical Characteristics of Major Depression and Manic-Depression

Major depression and manic-depression are classified among the mood disorders. Major depression (unipolar depression) consists of one or more depressive episodes, defined as 2 weeks or more of depressed mood or loss of interest or pleasure in nearly all activities, with at least four additional symptoms from the following: increased or decreased appetite and/or weight; increased or decreased sleep; change in psychomotor activity (retardation or agitation); decreased energy; feelings of worthlessness or guilt; difficulty thinking, concentrating, or making decisions; and repeated thoughts of death, including suicidal thoughts, plans, or attempts. The symptoms must be severe enough to cause significant distress or to interfere with the person's occupational or social

functioning. Finally, the major depressive episode must not be a direct result of a drug or an underlying medical condition.

Based on these criteria for the diagnosis of major depression, we can see that it is a very broad category of mental illness. The diagnostic criteria are disjunctive in that there are few mandatory symptoms, and individual patients may have few symptoms in common. To address this concern about heterogeneity, subtypes of major depressive disorder have been defined that require a more restricted list of symptoms. In particular, melancholic major depression focuses on physiological symptoms. Its diagnosis requires having the criteria for major depression plus either loss of pleasure in all (or almost all) activities or lack or reactivity to usually pleasurable stimuli, along with at least three additional symptoms from the following: a distinct quality to the depressed mood (different from a feeling of grief), the depressive symptoms being regularly worse in the morning, early morning awakening (at least 2 hours earlier than usual), marked psychomotor retardation or agitation, significant loss of appetite and/or weight, and excessive or inappropriate guilt. These also are disjunctive criteria, associated with significant heterogeneity. Melancholic major depression usually requires antidepressant medication, and sometimes ECT, for effective treatment.

Another feature of major depression may be the presence of psychotic features (delusions that are usually mood-congruent, or hallucinations). This form of depression may require treatment with antipsychotic as well as antidepressant drugs. The timing of major depressive episodes also may be linked to the season of the year, particularly to the light-dark cycle – seasonal affective disorder, as this condition is known, usually occurs in the winter when the days are short. Exposure to very bright light for several hours each morning may be useful as an adjunct to antidepressant drugs in the treatment of seasonal affective disorder.

A manic episode consists of a period of abnormally and persistently elevated, expansive, or irritable mood lasting for at least 1 week (less if the person is hospitalized), with at least three of the following additional symptoms: inflated self-esteem or grandiosity, decreased need for sleep, pressure of speech, flight of ideas or subjective feeling that his or her thoughts are racing, being easily distracted, increased activities (work, social, sexual) or psychomotor agitation, and excessive involvement in pleasurable activities that are potentially harmful (e.g., excessive alcohol intake, buying sprees, sexual indiscretions, or foolish business decisions).

Although these diagnostic criteria also are disjunctive, the heterogeneity of manic patients is generally much less than that of depressed patients. As with

major depression, the symptoms must be severe enough to cause significant distress or to interfere with the person's occupational or social functioning, and the manic episode must not be a direct result of a drug or an underlying medical condition. And, as with major depression, psychotic features also may be part of the clinical picture. Finally, although mania and depression are conceptualized as polar opposite forms of mood disorder (hence the term bipolar disorder), coexisting manic and depressive symptoms are very common. When depressive and/or anxiety symptoms are prominent in a manic episode, the term mixed bipolar disorder is applied. Analyses of symptoms in mania have consistently revealed dimensions of depression and anxiety, elation-grandiosity, pressured behavior, psychosis, and irritable hostility. These separate components may display differential responsiveness to treatments; for example, elation-grandiosity may respond to lithium, irritable hostility to anticonvulsant mood stabilizers, and psychosis to antipsychotic drugs.

Bipolar I disorder is defined as a clinical course of one or more manic or mixed episodes, the latter being defined as criteria for both manic episode and major depression when they have been met nearly every day for at least 1 week. The most common form of bipolar illness is manic episodes alternating with major depressive episodes, which may or may not be separated by intervals of normal mood. Mixed episodes are less common, and the rarest pattern is to have manic episodes only, with no depressive features. Occasionally, several major depressive episodes may occur before the first manic episode becomes evident. Bipolar II disorder is defined as recurrent major depressive episodes with hypomanic episodes. Hypomanic episodes must meet the criteria for manic episodes, except they can be of shorter duration and lesser severity.

Epidemiologic studies indicate that the lifetime prevalence of major depression is about 10–25% for women and about half that for men; thus, it is one of the commonest psychiatric disorders. Fewer than half these survey-defined cases come to clinical attention, however. About twice as many adolescent and adult females are affected as males. In children, the prevalence of major depression is about equal between boys and girls. Thus, the hormonal changes of puberty may play a role in its greater incidence in adolescent and adult females. In contrast to widespread clinical belief a few decades ago, there is no increase in major depression in women at the time of menopause.

Approximately one-third of clinical major depressions improve with treatment and do not recur; approximately one-third show remission with treatment but do recur a number of times throughout the life

of the individual; and approximately one-third may improve with treatment, but there remains a chronic, underlying depression of lesser severity. This chronic depression may meet the criteria for dysthymic disorder, which requires, among other criteria, depressed mood for more days than not, for at least 2 years. Major depressive episodes occurring against a background of dysthymia has been called double depression. The term refractory depression is applied to patients who fail to respond to at least two adequate trials of treatment; these make up approximately 15% of clinical major depressive episodes.

Major depression can co-occur with many other psychiatric illnesses such as anxiety disorders, eating disorders, drug abuse and alcoholism, dementias, and psychotic disorders, as well as in the context of many medical illnesses. In these instances, it is referred to as secondary major depression, in contrast to primary major depression, in which the depressive illness itself predominates. Primary major depression also may have some of these illnesses as secondary diagnoses; the determining factor is which came first. For example, anxiety symptoms sufficient to meet a diagnosis of generalized anxiety disorder occur with some frequency in major depression and are considered secondary because the depressive illness preceded them. Regardless of whether the depression is primary or secondary, these comorbid conditions have significant adverse effects on the clinical course of patients with major depression.

Although the classical descriptions of major depression and mania do not focus on cognitive function, it is now clear that pervasive, subtle impairments of executive function and declarative memory do occur during episodes of mood disturbance and that these features sometimes persist between episodes, especially in bipolar subjects. In elderly patients, more striking cognitive impairment may occur as part of a depressive episode. In particular, depressive pseudodementia (false dementia) should be considered in elderly depressed patients who have significant impairment of cognitive function and who might otherwise be diagnosed as having a dementia such as Alzheimer's disease. The distinction between depressive pseudo-dementia and true dementia is extremely important, because pseudo-dementia resolves with successful treatment of the underlying depression, whereas true dementia generally does not respond to currently available treatments.

Not only is major depression an important illness today, but within 15 years it is predicted to become the second most incapacitating illness throughout the world, just a small percentage behind cardiovascular disease. Thus, major depression is fast becoming a significant public health problem. As global

populations age, a newly recognized form of major depression in the elderly, termed vascular depression, is increasing in prevalence. This condition results from disease of the small arteries supplying the brain. It has an unusually late age of onset for unipolar depression (after age 50), and it is recognized by the magnetic resonance imaging (MRI) features of deep white-matter hyperintensities; subcortical gray matter hyperintensities, especially in the caudate nuclei; and micro-infarcts (tissue death from blocked blood supply). All these features occur as well in vascular dementia, which is a significant long-term outcome of vascular depression. Risk factors for vascular depression include hypertension, diabetes mellitus, and abnormal plasma lipid profiles.

The lifetime prevalence for bipolar I illness is about 1% and for bipolar II illness about 0.5%. These illnesses usually begin in adolescence or young adulthood and are equally prevalent in males and females. As mentioned, recurrent alternating manic and depressive episodes represent the most common form; if four or more manic or depressive episodes occur within a year, the illness is considered rapidly cycling and may require frequent hospitalizations and certain medication combinations for its control.

As with the other severe psychiatric disorders, both major depression and manic-depression can be made worse by stressful life situations. Major depression not uncommonly is preceded, if not precipitated, by a significant loss in a person's life, be it death of a significant other, job loss, or financial reversal. As mentioned, the psychological pain and turmoil of a severe depression, coupled with additional physical factors such as insomnia, can in turn be extremely stressful to the sufferer, and, on occasion, suicide may be considered the only way out. Fortunately, almost every major depression is treatable, and suicidal thoughts that are present in the depths of the depressive episode almost always disappear with treatment.

Bipolar disorder is more internally driven (less affected by external stressors), but the illness itself can have extremely stressful consequences for the individual. It is a challenging condition for family members as well, as evidenced by the fact that the rate of marital failure in bipolar disorder exceeds that in any other psychiatric condition. During a manic episode the person may run afoul of the law and be arrested; may spend his or her life savings foolishly and be reduced to poverty; or may behave recklessly and suffer physical injury, for example by reckless or drunken driving with involvement in an auto accident or sexual promiscuity with the contraction of a disease such as AIDS. Bipolar disorder, being a recurrent disorder, can result in many psychiatric hospitalizations over an individual's lifetime.

Biological Characteristics of Depression and Manic-Depression

There appears to be a genetic contribution to both major depression and bipolar disorder. Both illnesses more frequently coexist in monozygotic (single-egg) than in dizygotic twins, but the difference is much more pronounced for bipolar disorder. Similarly, there is a higher incidence of these disorders in the first-degree relatives of patients with both major depression and bipolar disorder; again, the relationship is stronger for bipolar disorder. A consistent observation is that the most common psychiatric disorder in the relatives of subjects with bipolar disorder is unipolar depression. This has led to suggestions of a two-gene vulnerability in bipolar cases. For both of these illnesses, however, as for all the major psychiatric disorders, the modes of inheritance have not been established. Except for a few bipolar disorder families, the modes of inheritance do not appear to fit either an autosomal or an X-linked pattern but most likely are multifactorial, involving several genes of varying influence. Nongenetic contributing factors, as well as different genetic loadings in different families, may interact to produce the same phenotypic outcome.

A particularly intriguing area of research is the interaction of gene polymorphisms and life stresses in the onset of major depressive episodes. A particular polymorphism in the serotonin transporter gene, the presence of two short alleles, appears to confer a greater vulnerability to depressive episodes following mild stressful life events than the presence of either one or two long alleles. There is only a weak overall relationship between the serotonin transporter polymorphism and occurrence of major depression, so life stresses apparently are the precipitating factors. The genetic polymorphism confers vulnerability to developing depression following less severe stresses.

There are several neurochemical theories of the etiology of depression and bipolar disorder, all of which involve abnormal function of CNS neurotransmitters and/or the chemical changes they activate in nerve cells (neurons). Neurotransmitters (for example serotonin, norepinephrine, and dopamine) are chemicals in the brain that are released by neurons and carry a chemical message across a short space (synapse) to neighboring neurons that have receptors for that neurotransmitter. Deficiencies of norepinephrine and serotonin neurotransmission and excessive acetylcholine neurotransmission are among the oldest theories of depression and best supported by research studies.

In both unipolar major depression and the depressive phase of bipolar disorder, reductions in the metabolic products of norepinephrine and serotonin have

been found in blood, urine, and cerebrospinal fluid, suggesting that both these neurotransmitter systems may be underactive, but the reductions in metabolites have been relatively small. In manic patients, the administration of physostigmine, a drug that inhibits the breakdown of acetylcholine, immediately stops the mania and produces a depressionlike state. Physostigmine also produces a depressionlike state in normal subjects. These findings suggest an overactivity of cholinergic neurotransmission in depression, and other studies have shown depressed patients to be supersensitive to cholinergic drugs in their sleep patterns, constriction of the pupils of their eyes, and so forth.

The most compelling evidence for the neurotransmitter theories of depression comes from the effects of drugs used to treat major depression and bipolar disorder. There are several classes of antidepressant drugs, and many of them block the cell membrane transporters for norepinephrine and/or serotonin, which recycle the neurotransmitters from the synapse back into the nerve cells that released them. By blocking transporter uptake, antidepressants increase synaptic neurotransmitter concentrations. Studies have shown that, in the intact human being, even very specific neurotransmitter uptake inhibitors eventually influence multiple neurotransmitter systems because of the physiological interactions among them. If one neurotransmitter system is perturbed by a drug acting on it specifically, the other systems tend to change to come back into balance with it.

More recent theories of antidepressant drug action emphasize follow-on actions of the drugs beyond inhibiting neurotransmitter reuptake by monoaminergic neurons. For example, an intracellular cascade occurs involving the activation of intracellular messengers (adenylyl cyclase and tyrosine protein kinases, the transcription factor cAMP response element binding protein (CREB), and increased mRNA for neurotrophins such as brain-derived neurotrophic factor (BDNF)).

The time course for full clinical response to all antidepressant drugs is 3–6 weeks, even though their specific pharmacological effects occur within 12–24 h. What is common to almost all antidepressants is that downregulation of postsynaptic noradrenergic receptors in the CNS occurs over the same number of weeks as required for the clinical response. Noradrenergic receptor downregulation may be an ultimate result of the increased availability of norepinephrine as a synaptic neurotransmitter, signaling that noradrenergic neurotransmission has increased to such an extent that the number and sensitivity of postsynaptic receptors need to be reduced in compensation. This finding suggests that a common neurochemical

pathology in depressed patients may be a relative underactivity of noradrenergic neurotransmission, which is rectified by drug treatment. The restoration of noradrenergic transmission by antidepressant treatments, including ECT, may be linked to the trophic action of BDNF in restoring functional synapses between noradrenergic neurons and their target neurons in prefrontal cortex and limbic system sites.

There are disturbances in several physiological systems that occur as part of depression and the depressive phase of bipolar disorder. Sleep disturbance has been mentioned as part of the diagnostic criteria. Electrophysiological studies of sleep (polysomnography) have revealed typical changes, including less total sleep time, fragmentation of the orderly cycling of sleep stages with more frequent awakenings throughout the night, and early occurrence of rapid eye movement sleep (REM or dreaming sleep). In normal people, the first REM episode occurs about 60–90 min after sleep is established (REM latency). This interval may be severely shortened in major depression, so much so that the person may go directly into REM sleep when sleep is first established (sleep-onset REM). The occurrence of REM sleep is influenced by CNS cholinergic neurotransmission; as mentioned, the administration of cholinergic neurotransmission-stimulating drugs to depressed people and control subjects while asleep produces a significantly shorter REM latency in the depressives than in the controls. This is one of the findings supporting the cholinergic overdrive hypothesis of major depression.

Another important physiological disturbance in major depression and bipolar disorder is increased activity of the hypothalamic-pituitary-adrenal cortical (HPA) axis. This endocrine axis consists of cells in the hypothalamus of the brain that secrete corticotropin-releasing hormone (CRH) and vasopressin (AVP), which are carried to the anterior pituitary gland and stimulate the secretion of adrenocorticotrophic hormone (ACTH) from the anterior pituitary into the blood stream. In turn, ACTH stimulates the adrenal cortex to secrete glucocorticoids, mineralocorticoids, and adrenal androgens into the bloodstream. Glucocorticoid hormones increase glucose production in the liver and promote lipid breakdown in fat tissue, thereby increasing circulating glucose available for muscle and other tissue use. The principal glucocorticoid in humans is cortisol (hydrocortisone). Mineralocorticoid hormones reduce the excretion of sodium and enhance the excretion of potassium and hydrogen ions by the kidney; the principal mineralocorticoid in humans is aldosterone. Adrenal androgenic hormones have weak male sex hormonelike effects.

The HPA axis is very responsive to both physical and psychological stressors. The secretion of CRH is

regulated by several of the CNS neurotransmitters considered important in depressive illness, including norepinephrine, serotonin, and acetylcholine. The increased activity of this endocrine axis that occurs in 30–50% of patients with major depression and in the depressive phase of bipolar disorder is likely caused by an underlying neurotransmitter dysfunction in the brain. Increased HPA axis activity can be documented by the finding of increased ACTH and cortisol in blood, urine, and cerebrospinal fluid and by ACTH and cortisol resistance to suppression with the potent, synthetic adrenal glucocorticoid, dexamethasone. Dexamethasone acts primarily at the pituitary gland to suppress the secretion of ACTH, which, in turn, suppresses the secretion of hormones from the adrenal cortex. Resistance to dexamethasone suppression in depressed patients originally was interpreted as evidence of overdrive from the brain on the pituitary. However, it is now known that dexamethasone may be cleared from the body more rapidly in the nonsuppressing subjects, with resulting lower plasma dexamethasone concentrations. This factor has to be considered when the test is being used (see **Dexamethasone Suppression Test (DST)**).

When stimulated for a period of time, endocrine glands often hypertrophy (grow in size), and both increased pituitary size and increased adrenal gland size have been reported in major depression. Adrenal size reverted to control values following successful antidepressant treatment. These adrenal gland changes are similar to those occurring during and after chronic stress in animals. The long-term effects of chronic exposure to high cortisol levels in depressed patients include loss of bone density, impairment of immune function, increased incidence of adult-onset diabetes mellitus and cardiovascular disease, and premature mortality. Collectively, these are termed indicators of high allostatic load in depression.

Neuroimaging Studies in Depression and Manic-Depression

With the advent of neuroimaging techniques, it now is possible to probe the structure and function of the living human brain. Brain structure can be viewed with X-ray computed tomography (CT) and MRI. Magnetic resonance spectroscopy (MRS) can determine the concentrations of certain chemical substances in the CNS. Nuclear medicine techniques such as single-photon computed tomography (SPECT) and positron-emission computed tomography (PET) allow the visualization and quantitation of regional cerebral blood flow, glucose metabolic rate, and neurotransmitter receptor occupancy, which are indirect probes of regional neuronal function. SPECT and PET use

radiolabeled compounds as tracers; as these compounds decay, high-energy photons are emitted, which are counted by external detectors. The distribution of the tracer molecules is computed, and cross-sectional images of the brain are created in which image brightness is proportional to the underlying physiological process being measured. The newest technique, functional MRI (fMRI), also indicates blood flow in different areas of the brain.

Studies of brain structure with CT and MRI in patients with major depression and bipolar disorder have demonstrated a number of abnormalities, but it is not clear how specific these changes are. Enlarged cerebral ventricles in relation to overall brain size have been reported in a number of studies, but similar findings have been reported in patients with other psychiatric illnesses, such as schizophrenia. Different groups of patients (unipolar and bipolar), different ways of measuring and reporting ventricular volume, and different control groups have been used, but meta-analysis of many studies has pointed toward a significant enlargement of the ventricles in unipolar and bipolar patients compared to controls. The difference is more marked in elderly subjects. There appears to be no consistent relationship between structural changes in the brain and severity of the illness. Reduced volume of the subgenual anterior cingulate gyrus has been found in familial, pure, primary, unipolar depression, and this feature has been correlated with glial cell loss in the same region.

Reduced volume of the hippocampus also has been reported inconsistently in major depression. Initial suggestions that this feature is related to elevated cortisol exposure have not been confirmed. Moreover, hippocampal volume reduction also has been reported in nondepressed patients with cardiovascular disease. The finding of reduced hippocampal volume in depression, therefore, may simply reflect the high load of cardiovascular disease in these patients.

Elderly patients with vascular depression may show subcortical hyperintensities. These are areas within the white matter (neuron fiber tracts) and the subcortical grey-matter nuclei (neuron cell bodies) that are bright on T2-weighted MRI scans. They are associated with risk factors for cerebrovascular disease and may be areas of demyelination, local brain swelling, and small, old strokes. Elderly patients with major depression and bipolar patients of all ages have more extensive (in total area) subcortical white-matter hyperintensities than do age-matched healthy controls. The presence of subcortical hyperintensities in elderly depressed patients correlates with a greater sensitivity to the side effects of antidepressant medications, including delirium, a greater risk of developing tardive dyskinesia (a movement disorder)

from neuroleptic (antipsychotic) medication, and a longer recovery time from ECT.

MRS has been used to study several compounds in the brain in major depression and bipolar disorder. Phosphomonoesters, a component of cell membranes, have been found to be both elevated and decreased. Phosphocreatine, a high-energy phosphate, has been more consistently decreased. MRS of fluorine-containing drugs, such as some antidepressants, and of lithium provides brain concentrations of these compounds in patients under treatment and may eventually be helpful in guiding dosage adjustments to achieve maximum therapeutic response. SPECT and PET functional neuroimaging studies of major depression and bipolar disorder have shown decreased prefrontal cortical and anterior cingulate gyrus blood flow and glucose metabolism, although these findings have not been consistent. In addition, there have been inconsistent reversals of these abnormalities in treated patients; however, it does appear that successful antidepressant pharmacotherapy is associated with increased blood flow and/or metabolism in the prefrontal cortex, cingulate cortex, and/or the basal ganglia (caudate nucleus, globus pallidus, and putamen).

Brain Circuit Theories of Depression and Manic-Depression

Brain circuit theories of major depression and bipolar disorder all represent variations on the postulated involvement of certain areas of the frontal parts of the brain (frontal lobes) and their functional interconnections with deeper parts of the brain – the cingulate gyrus, basal ganglia, thalamus, and other structures, including hippocampus and amygdala. Several of these structures make up the limbic system, a phylogenetically older part of the brain that influences emotional coloration and behavioral reactivity. In a general sense, the frontal lobes of the brain mediate executive functions; that is, an individual's decision-making and behavioral actions based on a detailed evaluation of environmental demands, along with an appreciation of their historical context, and a coordinated affective/emotional component. Particular areas of the frontal cortex mediate executive behavior, social behavior, and motivation.

The underlying theme of the brain circuit theories of major depression and the depressive phase of bipolar disorder is that the cognitive dorsal prefrontal cortex cannot effectively modulate the ventral orbitofrontal and medial prefrontal cortex and the limbic system components of emotions, so that what in a normal individual would be an expression of sadness becomes magnified and prolonged into a major

depression when the interactive circuitry is not functioning properly. Ascending monoamine projections from the brainstem involving dopamine, norepinephrine, and serotonin set the tone in the emotional circuits comprising the prefrontal cortex, ventral striatum, mediodorsal thalamus, and projections back to the prefrontal cortex. When these delicate, unmyelinated, slow-conducting monoamine projections are damaged or rendered dysfunctional, monoamine control of the emotional circuits becomes abnormal.

Several distinct pathologic processes may cause such disconnection of the monoaminergic projections from their target neurons in prefrontal cortex and subcortical sites. These include loss-of-function polymorphisms of the BDNF gene, with resultant vulnerability to stress or glucocorticoids through reduced terminal arborization of monoaminergic axons; vascular lesions of the monoaminergic fibers in their passage through the midbrain and internal capsule; degenerative changes in brainstem monoaminergic cell nuclei as happens in Parkinson disease, Huntington disease, and Alzheimer disease; and functional impairment of the monoaminergic neurons by drugs such as reserpine, which prevent vesicular storage of the monoamine neurotransmitters, with resultant depletion of tissue monoamine concentrations. No matter which of these processes is operative, the final effect is loss of monoamine control over activity in the prefrontal-subcortical emotion circuits and the appearance of a primary or secondary depressive syndrome. In these syndromes, the expressed symptoms are very similar, regardless of the underlying pathology. The aforementioned reductions in resting prefrontal cortical blood flow and glucose metabolism, as determined by functional neuroimaging with SPECT, PET, and fMRI support this brain circuit theory. Likewise, increased activity in these same regions is seen when negative emotional processing tasks are performed by depressed patients.

Psychological Testing in Depression and Manic-Depression

Many general psychological tests contain questions pertaining to mood, and these tests usually give a score that indicates the degree of depression. Because depression and manic-depression are often recognizable clinically, however, psychological testing is usually not done unless there are ancillary questions, such as the possibility of associated dementia or psychosis. Psychological testing also may be helpful in differentiating depressive pseudo-dementia from primary dementia. In a patient presenting with dementia, an indication of depressive features by psychological testing, which may not have been apparent clinically,

usually signals the need for antidepressant treatment. In such cases, following successful treatment of the depression, the dementia also should have improved, which can be determined by repeat psychological testing.

There also are some rating scales designed specifically to quantitate depression. The Hamilton Rating Scale for Depression (HAM-D), which is scored by the person examining the patient, is the gold standard for clinical trials of antidepressant efficacy. Self-rated scales such as the Beck Depression Inventory (BDI) and the Carroll Depression Scale (CDS) also are used. The HAM-D tends to emphasize the physiological aspects of depression, such as sleep and appetite disturbance and physical symptoms of anxiety. The BDI emphasizes the cognitive component of depression; that is, how the person thinks about him- or herself when depressed. Short, self-rated depression screening scales also are used in epidemiologic studies and in primary medical care settings.

Treatment of Major Depression and Manic-Depression

The first issue in the treatment of patients with major depressive and bipolar disorder is to assess their immediate personal safety in light of the severity of their illness. For depressed patients, this always includes an assessment of their suicide potential. For manic patients, it includes not only an assessment of their suicide potential, especially in mixed bipolar disorder, but also whether their behavior is harmful (e.g., excessive alcohol intake, buying sprees, sexual indiscretions, or foolish business decisions, as mentioned earlier). The most certain way to interrupt behavioral indiscretions and protect the individual from suicide attempts is hospitalization in a specialized psychiatric unit whose staff is trained in the care of such patients. If the patient resists hospitalization, an involuntary hold may be medically and legally justified. A locked psychiatric unit, which prevents the patient from leaving, may be necessary until the illness is sufficiently under control that the patient's better judgment returns.

The treatment of major depression and bipolar disorder almost always requires drug therapy with antidepressants and/or mood-stabilizing drugs. There are several chemical classes of antidepressant drugs, and they have specific pharmacological activities in the CNS. As previously mentioned, many of them block the transporters for norepinephrine and/or serotonin, which recycle the neurotransmitters from the synapse back into the nerve cells that released them. By blocking transporter uptake, antidepressants increase synaptic neurotransmitter concentrations.

The time course for a full clinical response to antidepressant drugs is 3–6 weeks, even though their pharmacological effects occur within 12–24 h. Often, improved sleep may be an early sign of response to medication, especially with antidepressants that have sedative side effects. Objective signs of improvement usually precede the patient's feeling better; for example, the person may be sleeping and eating better, may have more energy and a higher level of activities, and may be speaking more cheerfully, but he or she still may be complaining about feeling as depressed as before. The subjective depressed mood is often the last aspect of the illness to improve. Because the subjective experience of depression is so psychologically painful and stressful, it is very important that a patient with suicide potential not be released from the hospital until he or she is in sufficient behavioral control to no longer be a suicide risk after discharge.

As mentioned, some antidepressants are specific uptake inhibitors of norepinephrine and others of serotonin, but it is not possible to predict which depressed patient will respond to which antidepressant. This is most likely on the basis of the physiological interactions of neurotransmitter systems and, as previously indicated, the fact that almost all antidepressants result in downregulation of postsynaptic noradrenergic receptors and induction of BDNF over the same time course as clinical improvement occurs. Antidepressants therefore are usually chosen on the basis of their side effects and cost, those under patent being more expensive than those available in generic form. Many of the older antidepressants have prominent side effects, such as causing dry mouth, blurred vision, and especially changes in the electrical conduction system of the heart. This last side effect can be particularly dangerous in accidental or deliberate drug overdose.

The class of antidepressant drugs currently in greatest use is the serotonin uptake inhibitors (SUIs). The first SUI accepted for clinical use in the United States was fluoxetine (Prozac). The SUIs may not be quite as effective as the original tricyclic and monoamine oxidase inhibitor antidepressant drugs, especially in severe depression, but they have fewer side effects, especially cardiac, which makes them generally safer drugs to use. They do have other side effects that must be considered, including weight gain, decreased sexual drive, akathisia (inner restlessness), and some increased risk of suicidal thinking or gestures. Although it has not been established that antidepressant drugs provoke completed suicides, regulatory warnings emphasize that patients must be followed closely for this risk during the early weeks of treatment.

Recent studies have shown that the efficacy of the SUIs in the broad, heterogeneous group of patients diagnosed with major depression is modest at best. The Number Needed to Treat (NNT) is a standard therapeutics measure that denotes the number of patients who must receive a drug for one drug-attributable therapeutic outcome to be achieved, that is; over and above the placebo response rate. For the SUIs, the NNT ranges from five to twelve, whereas the NNT for the early tricyclic antidepressant drugs in severe, hospitalized depressed patients was three. Recent studies also confirm that there is no significant difference in response or remission rates between SUIs and some other newer antidepressants vs. placebo in mild depression. The National Institute for Clinical Excellence (NICE) in Britain therefore has recommended that nondrug treatments be used first in mild depression.

A new class of antidepressant drugs that block synaptic reuptake of both norepinephrine and serotonin (e.g., venlafaxine and duloxetine) has recently appeared. The dual action of these drugs recapitulates the pharmacodynamic profile of original antidepressant agents such as imipramine, amitriptyline, and phenelzine, but with a greatly reduced side-effect profile. These dual-action drugs appear to be more effective than the SUIs in treating major depression. For example, in direct comparisons, the NNT for venlafaxine to produce remission is five, compared with ten for the SUIs.

The other major class of drugs used in mood disorders, especially in bipolar disorder, is the mood stabilizers. For manic patients, lithium is the most effective. Lithium is a metal ion, in the same class in the periodic table of elements as sodium and potassium. Lithium has a number of effects on neurotransmission in the CNS. It has a relatively narrow therapeutic index; that is, the blood concentrations at which lithium exerts toxicity are not very far above the concentrations required for its therapeutic effect. Therefore, patients taking lithium require frequent measurement of their circulating lithium concentrations, especially at the outset of treatment, to determine the daily dose necessary to achieve a therapeutic blood level. Lithium takes several weeks to achieve its full antimanic effect. It is a mood stabilizer rather than a pure antimanic compound, so bipolar patients who switch from mania into depression are usually continued on their lithium if an antidepressant is added.

For major depressive and bipolar patients, additional dimensions of their illnesses may suggest the need for other medications in addition to antidepressants and lithium. For example, prominent psychotic

features in either illness may call for an antipsychotic medication to be used concomitantly. For the treatment of depression, hormone supplements such as estrogens in women and thyroid hormone may be helpful. For the treatment of bipolar disorder, a number of anticonvulsant medications have been shown to be effective, often as augmentation of lithium treatment. Compounds such as carbamazepine, valproate, and lamotrigine are currently in use, and several newer anticonvulsants are being tested for their mood-stabilizing properties.

After the successful drug treatment of a first lifetime episode of major depression, continuation treatment at full dosage is advised for 9 to 12 months to prevent relapse. The slow reduction of the medication then is attempted. Patients who experience recurrent depression, particularly those who have had three or more lifetime episodes, require long-term antidepressant maintenance treatment at full dosage to prevent recurrences. Controlled trials have established that, even after 3 years of successful preventive drug treatment, 50% of such patients will have a recurrence within 6 months of stopping their medication.

Bipolar patients need to have medication adjustments made according to the frequency of their manic and depressive mood swings; it may take years for the timing of these cycles to be clearly understood. Both disorders should be viewed as chronic illnesses, often relapsing over a person's lifetime. If strenuous attempts at drug treatment of either disorder fail, ECT often will provide definitive relief of symptoms. Usually, 10 treatments are given, three per week. The patient may suffer some memory loss during and following the treatments, but this is short-lived, whereas the therapeutic effect can be remarkable, especially in patients resistant to drug therapy. Unilateral, brief pulse application of electrical current to the nondominant hemisphere can reduce these memory changes significantly. Maintenance ECT, often one treatment every month or so, can be useful to keep the person in remission from his or her illness.

In addition to pharmacotherapy and ECT, psychotherapies of different types, such as cognitive-behavioral therapy, often are useful to help the person change his or her lifestyle and manner of thinking about adversity. An improvement in self-esteem often results, which may protect against future episodes of depression or at least may reduce their severity.

Conclusion

Major depression and manic-depression (bipolar disorder) represent serious psychiatric illnesses. In

addition to interfering with occupational and social functioning, these illnesses put the sufferer at personal risk – a potential for suicide attempts in the case of depression and impulsive behaviors such as excessive alcohol intake, buying sprees, sexual indiscretions, and foolish business decisions in the case of a manic episode. Severe life stressors often precede a depressive episode, and both illnesses create their own stresses in the sufferer, including the psychological pain of depression and the untoward consequences of unrestrained manic behavior. Fortunately, there are effective pharmacological treatments that can produce remission in the majority of patients, and for those who do not respond to drug treatment, ECT is an excellent therapeutic modality. The neurological circuitry involved in mood disorders is being revealed, with the result that new treatment approaches are emerging.

Acknowledgments

This research is supported by NIH grant MH28380. Dr. Rubin has received no commercial company support. Dr. Carroll has had past consulting, speaking, or research support from the following companies: Abbott, Astra Zeneca, Becton-Dickinson, Cyberonics, Glaxo Smith Kline, Janssen, Johnson and Johnson, Lilly, Novartis, Pfizer, Roche, Warner Lambert-Parke Davis, and Wyeth.

Further Reading

- Akiskal, H. S. (2000). Mood disorders: introduction and overview. In: Sadock, B. J. & Sadock, V. A. (eds.) *Kaplan and Sadock's comprehensive textbook of psychiatry* (7th edn., pp. 1284–1298). Philadelphia: Lippincott Williams & Wilkins.
- Akiskal, H. S. (2000). Mood disorders: clinical features. In: Sadock, B. J. & Sadock, V. A. (eds.) *Kaplan and Sadock's comprehensive textbook of psychiatry* (7th edn., pp. 1338–1377). Philadelphia: Lippincott Williams & Wilkins.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th edn.). Washington, DC: American Psychiatric Association.
- Baldessarini, R. J. (2001). Drugs and the treatment of psychiatric disorders: depression and anxiety disorders. In: Hardman, J. G. & Limbird, L. E. (eds.) *Goodman & Gilman's the pharmacological basis of therapeutics* (10th edn., pp. 447–483). New York: McGraw-Hill.
- Baldessarini, R. J. and Tarazi, F. I. (2001). Drugs and the treatment of psychiatric disorders: psychosis and mania. In: Hardman, J. G. & Limbird, L. E. (eds.) *Goodman & Gilman's the pharmacological basis of therapeutics* (10th edn., pp. 485–520). New York: McGraw-Hill.
- Blazer, D. (2000). Mood disorders: epidemiology. In: Sadock, B. J. & Sadock, V. A. (eds.) *Kaplan and Sadock's comprehensive textbook of psychiatry* (7th edn., pp. 1298–1308). Philadelphia: Lippincott Williams & Wilkins.
- Bloom, F. E. (2001). Neurotransmission and the central nervous system. In: Hardman, J. G. & Limbird, L. E. (eds.) *Goodman & Gilman's the pharmacological basis of therapeutics* (10th edn., pp. 293–320). New York: McGraw-Hill.
- Davis, K. L., Charney, D., Coyle, J. T. and Nemeroff, C. (eds.) (2002). *Psychopharmacology: the fifth generation of progress*. Philadelphia: Lippincott Williams & Wilkins.
- Cummings, J. L. (1995). Anatomic and behavioral aspects of frontal-subcortical circuits. *Annals of the New York Academy of Sciences* 769, 1–13.
- Gabbard, G. O. (2000). Mood disorders: psychodynamic aspects. In: Sadock, B. J. & Sadock, V. A. (eds.) *Kaplan and Sadock's comprehensive textbook of psychiatry* (7th edn., pp. 1328–1338). Philadelphia: Lippincott Williams & Wilkins.
- Hirschfeld, R. M. A. and Shea, M. T. (2000). Mood disorders: psychotherapy. In: Sadock, B. J. & Sadock, V. A. (eds.) *Kaplan and Sadock's comprehensive textbook of psychiatry* (7th edn., pp. 1431–1440). Philadelphia: Lippincott Williams & Wilkins.
- Kelsoe, J. R. (2000). Mood disorders: genetics. In: Sadock, B. J. & Sadock, V. A. (eds.) *Kaplan and Sadock's comprehensive textbook of psychiatry* (7th edn., pp. 1308–1318). Philadelphia: Lippincott Williams & Wilkins.
- Kitayama, I., Yaga, T., Kayahara, T., et al. (1997). Long-term stress degenerates but imipramine regenerates noradrenergic axons in the rat cerebral cortex. *Biological Psychiatry* 42, 687–696.
- Koschack, J. and Irle, E. (2005). Small hippocampal size in cognitively normal subjects with coronary artery disease. *Neurobiology of Aging* 26, 865–871.
- Krishnan, K. R. R. and Doraiswamy, P. M. (eds.) (1997). *Brain imaging in clinical psychiatry*. New York: Marcel Dekker.
- Kupfer, D. J., Frank, E., Perel, J. M., et al. (1992). Five-year outcome for maintenance therapies in recurrent depression. *Archives of General Psychiatry* 49, 769–773.
- Nibuya, M., Morinobu, S. and Duman, R. S. (1995). Regulation of BDNF and trkB mRNA in rat brain by chronic electroconvulsive seizure and antidepressant drug treatments. *Journal of Neuroscience* 15, 7539–7547.
- Post, R. M. (2000). Mood disorders: treatment of bipolar disorders. In: Sadock, B. J. & Sadock, V. A. (eds.) *Kaplan and Sadock's comprehensive textbook of psychiatry* (7th edn., pp. 1385–1430). Philadelphia: Lippincott Williams & Wilkins.
- Rush, A. J. (2000). Mood disorders: treatment of depression. In: Sadock, B. J. & Sadock, V. A. (eds.) *Kaplan and Sadock's comprehensive textbook of psychiatry* (7th edn., pp. 1377–1385). Philadelphia: Lippincott Williams & Wilkins.
- Singh, V., Muzina, D. J. and Calabrese, J. R. (2005). Anticonvulsants in bipolar disorder. *Psychiatric Clinics of North America* 28, 301–323.

Thase, M. E. (2000). Mood disorders: neurobiology. In: Sadock, B. J. & Sadock, V. A. (eds.) *Kaplan and Sadock's comprehensive textbook of psychiatry* (7th edn., pp. 1318–1328). Philadelphia: Lippincott Williams & Wilkins.

Thase, M. E., Entsuah, A. R. and Rudolph, R. L. (2001). Remission rates during treatment with venlafaxine or selective serotonin reuptake inhibitors. *British Journal of Psychiatry* 178, 234–241.

Depression and Stress, Role of n-3 and n-6 Fatty Acids

C Song

University of Prince Edward Island, Charlottetown, Canada

B E Leonard

University of Maastricht, Maastricht, Netherlands, and National University of Ireland, Galway, Ireland

© 2007 Elsevier Inc. All rights reserved.

Introduction

Stress and Fatty Acids in Human Studies

Effects of the n-3 and n-6 Fatty Acids on Stress and Anxiety-like Behavior, Corticosterone, and Neurotransmission in Animal Experiments

Clinical Investigation of Fatty Acid Composition in Major Depression

The n-3 Fatty Acids in the Treatment of Depression

Glossary

<i>Cytokines</i>	A group of pleiotropic proteins produced from leukocytes in the immune system and from microglia and astrocytes in the brain in response to a wide range of stimuli such as antigens, microbial soluble factors, endotoxins, stress, neurotoxicity, and injury.
<i>Elevated plus-maze</i>	A maze that comprises two open and two closed arms, elevated from the floor. Rodents with anxiety will increase entries into and time spent in closed arms.
<i>Limbic system</i>	The brain area involved with involuntary functions, emotion, and behavior; includes the hypothalamus, hippocampus, amygdale, fornix, and stria terminalis.
<i>Mitogens</i>	Foreign substances to the body; induce macrophage and lymphocyte proliferation and cytokine releases.
<i>Monoamine</i>	General name for catecholamine and indoleamine neurotransmitters.
<i>n-3 and n-6 fatty acids</i>	Types of fatty acids, named for the position of the first double bond in the

carbon chain, starting at the methyl end of the molecule.

Neurotransmitters

Open field test

T helper 1 (Th1) cells and pro-inflammatory cytokines

T helper 2 (Th2) cells and anti-inflammatory cytokines

Chemical messengers released from neurons to excite or inhibit adjacent neurons. A test that provides a novel and stressful environment for animal behavior.

A subtype of T lymphocytes; produce cytokines that can trigger inflammatory response.

A subtype of T lymphocytes; produce cytokines, some of which can suppress Th1 functions and suppress inflammation.

Introduction

Cytokines, Stress, and Depression

It is well known that stress is a trigger for depression and that increases in stress- and anxiety-type symptoms are common in depressed patients. In the past 15 years, many studies have demonstrated that an increase in inflammatory response is associated with stress exposure and depressive illness. In human subjects, psychological stress significantly increases pro-inflammatory (but inhibit anti-inflammatory) cytokine production in patients responding to stress and anxiety. In depressed patients, increases in macrophage activity and the production of pro-inflammatory cytokines, complement, and some acute-phase proteins have been consistently reported. Conversely, more than 70% of volunteers or nondepressed patients showed severe depressive symptoms after receiving tumor necrosis factor (TNF)- α or interferon (IFN)- γ treatment. Furthermore, animal experiments have demonstrated that pro-inflammatory cytokines, such as interleukin (IL)-1 β , IL-6, and TNF- α can stimulate the hypothalamus to release corticotropin releasing hormone (CRH), which, via

adrenocorticotrophic hormone (ACTH), induces glucocorticoid (GC) secretion. Excessive secretion of GC may downregulate GC receptors in the hippocampus, which impairs the GC feedback system. Similar neuroendocrine changes also occur in depressed patients. From the neurotransmitter perspective, pro-inflammatory cytokines have been found to reduce both serotonin and norepinephrine availability to the brain, which is similar to levels observed in the depression. Thus, inflammation and cytokines may play an important role in stress and depression

The n-3 and n-6 Fatty Acids, the Immune System, and the Central Nervous System

Long-chain polyunsaturated fatty acids (PUFA) are synthesized from dietary precursors such as α -linolenic (n-3) and linoleic (n-6) fatty acids. The n-3 fatty acids include docosahexaenoic acid (DHA) and eicosapentaenoic acid (EPA); the n-6 fatty acids include arachidonic acid (AA) and dihomo- γ -linolenic acid (DGLA). These fatty acids are important components of membrane phospholipids in neurons, glial, and immune cells and are involved in many functions of the immune and central nervous system for the following reasons. First, changes in membrane fatty acid components may change receptor, enzyme, and peptide functions in the central nervous system (CNS) and immune system. Because a high proportion of proteins in the cell is actually embedded in the membrane, the quaternary structures of proteins and the final modeling and folding often depend on the precise nature of the lipid environment of that protein. Second, fatty acids can influence signal transduction. Neurotransmitters, hormones and cytokines hit a target and induce functional changes through activating phospholipases that then generate a wide range of cell signaling or signal transduction. Third, fatty acids and other lipids can switch on and off many different genes. In particular, by binding to peroxisome proliferator-activated receptors, fatty acids can switch on and off whole genetic programs involving the interaction of many different genes. Lipids or carbohydrates also modulate heat shock proteins, which aid the expression of mRNA and the synthesis of proteins. Fourth, fatty acids influence ongoing metabolic regulation. Several studies have revealed that phospholipids are undergoing constant remodeling, with key fatty acid components having half-lives of a few minutes.

In the immune system, the onset of autoimmune and inflammatory diseases has been related to an imbalance in the intake of n-3 and n-6 fatty acids. Inflammation is an important component of the early immunological response, whereas inappropriate

or dysfunctional immune responses underlie chronic inflammatory and autoimmune diseases. The n-6 PUFA AA is the precursor of eicosanoids that produce prostaglandins (PG), leukotrienes, and related compounds to activate macrophages, produce pro-inflammatory cytokines, and shift the response of T-helper type (Th)1 and Th2. Th1 cells trigger pro-inflammatory response, whereas Th2 cells suppress Th1 response. By contrast, high intakes of long-chain n-3 fatty acids (in fish oils) inhibit certain immune functions (e.g., antigen presentation, adhesion molecule expression, Th1 responses, and production of eicosanoids and pro-inflammatory cytokines). Moreover, a diet enriched with fish oil has been shown to ameliorate the symptoms of autoimmune disease in animal models and to protect against the inflammatory effects of endotoxins. Clinical studies have reported that oral fish oil supplementation has beneficial effects in rheumatoid arthritis and some asthmatics, supporting the idea that the n-3 fatty acids in fish oil are anti-inflammatory.

In the brain, 60% of weight is made up of lipids, of which the essential fatty acids (n-3 and n-6) are important membrane components. Free fatty acids that are released into the blood and then pass through the blood-brain barrier can act at specific binding sites in the brain. Changes in the phospholipids content of the neuronal membranes directly affect membrane viscosity and fluidity, which may cause abnormalities in neurotransmitter binding and reuptake, receptor density and affinity, and hormone secretion. In the CNS, n-3 and n-6 fatty acids have been shown to fulfill different roles and to act as antagonists and agonists of one another. The n-6 fatty acid AA enhances the release of glutamate neurotransmitter, inhibits neurotransmitter uptake, stimulates stress hormone secretion, and enhances synaptic transmission. In the immune system, AA can trigger an inflammatory response and increase oxidants in the brain. Therefore, AA may contribute to inflammatory and oxidative toxicity in neurodegenerative diseases. By contrast, n-3 fatty acids have been found to compete with n-6 fatty acids. EPA can protect neurons from inflammation and oxidants. However, n-3 and n-6 fatty acids also interact with and synergize one another.

Abnormal concentrations of or metabolism of n-3 and n-6 fatty acids has been found in depression and follow stress exposure. The most recent findings have shown that n-3 fatty acids can modulate many inflammation-induced changes in the CNS. This article reviews the changes in fatty acids in stress and depression, clinical treatment with n-3 fatty acids in depressed patients, and the mechanism by which n-3 fatty acids modulate the behavior,

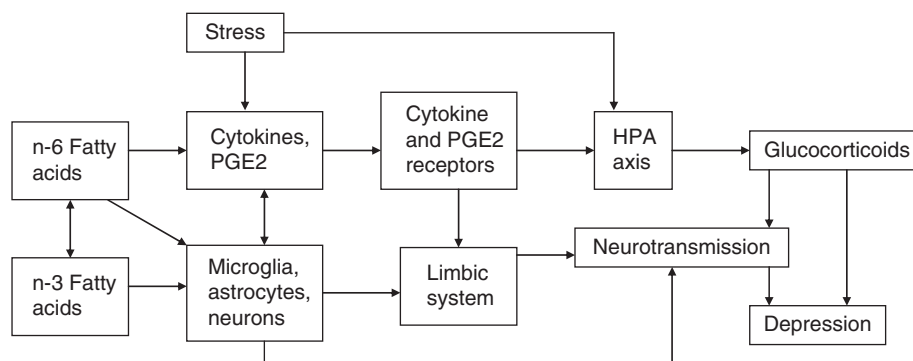


Figure 1 The relationship between stress and neuroendocrine system. The n-6 fatty acids are precursors of prostaglandin E2 (PGE2) and pro-inflammatory cytokines that can stimulate the HPA axis to release stress hormone glucocorticoids. Many studies have demonstrated that this hormone can change neurotransmission and cause stress/depressive symptoms. In the brain, n-6 fatty acids may activate microglia/astrocytes and increase inflammatory response, which may also change neurotransmission. The n-3 fatty acids can compete with the effect of n-6 fatty acids. Several studies have shown that n-3 fatty acid intake can reduce n-6 fatty acid contents in the brain. Furthermore, n-3 fatty acids have been reported to have anti-inflammatory effects, protect neurons, and modulate neurotransmission, such as serotonergic and noradrenergic systems. EPA has been also found to suppress glucocorticoid release, which may result from EPA inhibition of PGE2 and pro-inflammatory cytokine productions.

hypothalamic-pituitary-adrenal axis, immune response, and neurotransmitter systems in experimental studies. The relationship among fatty acids, stress, and depression is shown in [Figure 1](#).

Stress and Fatty Acids in Human Studies

Several studies in university students have shown that psychological stress can induce the release of pro-inflammatory cytokines, such as IFN- γ , TNF- α , and IL-6 from mitogen-stimulated blood. In addition, increased pro-inflammatory cytokine production is positively correlated with stress and anxiety scores, whereas the synthesis of the anti-inflammatory cytokines IL-10, IL-4, and IL-5 are negatively correlated with these scores. As already mentioned, the n-3 and n-6 fatty acids differentially modulate an inflammatory response. For example, an imbalance between the n-6 to n-3 fatty acid ratio may predict a greater production of pro-inflammatory cytokines in response to psychological stress. In support of this, a study measured cytokine release from mitogen-stimulated whole blood culture and the concentration of n-3 and n-6 fatty acids in the serum of university students a few weeks before, a few weeks after, and 1 day before a difficult oral examination. Academic examination stress markedly increased stimulated the production of IFN- γ , TNF- α , and IL-10 and increased the ratio of IFN/IL-5 (a marker of Th1 or Th2 cell activity) in subjects with a lower serum concentration of n-3 fatty acids on the day before the examination compared to pre- and post-examination conditions. The increases in these pro-inflammatory cytokines were negatively correlated with the ratio of n-6 to n-3 fatty acids, and

decreases in pro-inflammatory cytokine production occur in the subjects with a low n-6 to n-3 ratio. These results demonstrate that psychological stress induces pro-inflammatory and Th1-like response in subjects with higher stress/anxiety scores. An imbalance in the levels of n-3 and n-6 fatty acids appears to predispose humans toward an exaggerated pro-inflammatory and Th1-like response.

Antistress effects have been also found in DHA-treated human subjects. For example, 41 students took either DHA-rich oil capsules containing 1.5–1.8 g DHA day⁻¹ or control oil capsules containing 97% soybean oil plus 3% fish oil for 3 months in a double-blind fashion. They took a psychological test (P-F Study), Stroop test, and dementia-detecting test at the start and end of the study. The study started at the end of summer vacation and ended in the middle of a period of mental stress, such as the final examination. In the control group, aggression against others in P-F Study was significantly increased at the end of the study as compared with that measured at the start, whereas it was not significantly changed in the DHA group. DHA supplementation did not affect the Stroop and dementia-detecting tests. Thus, DHA intake prevented aggressive behavior from increasing at times of mental stress. This finding might help to understand why fish oils prevent coronary heart disease.

Sympathoadrenal activation is postulated to be involved in the pathogenesis of heart diseases and may be inhibited by n-3 fatty acids. The effects of a diet supplemented with n-3 fatty acids on the stimulation of the sympathetic nervous system and of stress hormones elicited by a mental stress were evaluated.

Seven human volunteers were studied on two occasions, before and after 3 weeks of supplementation with 7.2 g day⁻¹ fish oil. On each occasion, the concentrations of plasma cortisol, catecholamines, energy expenditure (indirect calorimetry), and adipose tissue lipolysis (plasma nonesterified fatty acid concentrations) were monitored in basal conditions followed by a 30 min mental stress period (mental arithmetics and Stroop test) and a 30-min recovery period. In control conditions, mental stress significantly increased heart rate, mean blood pressure, and energy expenditure. Stress also markedly increased plasma adrenaline, cortisol and nonesterified fatty acids. After 3 weeks of a diet supplemented with n-3 fatty acids, these changes induced by mental stress were all significantly attenuated. This result suggests that n-3 fatty acids inhibit the adrenal activation elicited by stress, presumably through effects exerted at the level of the CNS.

Effects of the n-3 and n-6 Fatty Acids on Stress and Anxiety-like Behavior, Corticosterone, and Neurotransmission in Animal Experiments

The effect of fatty acids on stress in rodents has been investigated in the following experiments. In the n-3 fatty acid-deficient mice, the role of n-6 fatty acids in stress response was compared to mice fed a normal diet. Mice fed 1.2% linoleic acid (a major dietary n-6 fatty acid) for two generations showed a markedly decrease in exploration in open field test and a decrease in time spent in open arms of the elevated plus-maze. A significant reduction in the n-3 fatty acid concentration was also found in the brain of these mice. After feeding with 0.5% α -linolenic acid (a source of n-3 fatty acids) for 8 weeks, the mouse behaviors were partially reversed.

The n-3 fatty acid DHA has been also found to attenuate the effect of stress in an animal model. Rats subjected to a psychosocial stress, such as isolation or an intermittent-feeding schedule, show a significant increase in heart rate, cardiac contractility, and cardiac norepinephrine concentration, which are blocked by DHA-enriched diet. The fatty acid composition of adrenals was significantly related to the fatty acid intake, particularly the neutral lipid fraction, which incorporated a large amount of DHA. Another experiment determined whether DHA affects stress responses in rats on several behavioral tests. Female rats were fed a diet deficient in n-3 fatty acid from mating through pregnancy and lactation. Male pups fed the same diet as their dams were then studied. The effects of dietary (n-3) fatty acid

deficiency and supplementation with DHA on psychological stress (as assessed by behavior in the elevated plus-maze) and conditioned-fear stress were investigated. The n-3 fatty acid-deficient rats spent significantly less time in the open arms of the plus-maze, suggesting increased anxiety; after 1 week of supplementation with DHA, they showed a significant improvement. The paired effects of DHA and CRH on stress were then evaluated. An intracerebro-ventricular (ICV) infusion of CRH under resting conditions was shown to have stress-inducing effects on behavior, such as decreases in rearing, smelling, and feeding and increases grooming in a novel environment open field. The DHA diet significantly improved these stress-induced behaviors. Finally, conditioned fear was induced by 40-min forced exposure to a cage in which the rat had experienced foot shocks 1 day before. Freezing behavior was dramatically suppressed by the supplementation of DHA, even 48 h after the conditioning treatment. Furthermore, the effect of DHA on the conditioned fear stress response was maintained over a long-term period. The ICV pretreatment of rats with bicuculline, a GABA_A receptor antagonist, enhanced the conditioned-fear-induced freezing time in a dose-dependent fashion in the n-3 fatty acid-deficient animals. Significantly, the DHA-supplemented group was not affected by the pretreatment with bicuculline. From these findings, it is concluded that the involvement of DHA in stress responses may act via a GABA_A receptor-mediated mechanism.

As mentioned previously, pro-inflammatory cytokines can stimulate the HPA axis to produce glucocorticoids and also induce stress and anxiety-like behavior. In the novel and stressful environment of the open field apparatus, the ICV administration of IL-1 β in rats reduced locomotor activity, exploratory behavior, and central zone entry but increased grooming behavior. In elevated plus-maze, IL-1 β decreased number of entries into and time spent in the open arms, which indicate an increase in anxiety-like behavior. These changes were correlated with the increase in the blood concentration of corticosterone. Feeding rats for 7 weeks with a diet containing rat basal food, supplemented with 0.2 or 1% n-3 fatty acid EPA led to significant increases in locomotor activity and exploration following the ICV administration of IL-1 β ; the grooming score was also decreased in the treatment with IL-1 followed by EPA. IL-1-induced anxiety-like behavior in the elevated plus-maze was also reversed by the diets enriched with 0.5 or 1% EPA. The elevated corticosterone level induced by IL-1 β was attenuated by 0.5 and 1% but not by 0.2% EPA.

With regard to changes in neurotransmission, several investigations have reported that following acute central administration of IL-1 β , changes in monoamine neurotransmitters and their metabolites were similar to changes observed after stress exposure. These changes include increases in the metabolism of serotonin, dopamine, and norepinephrine. If stress consistently increases the metabolism of these monoamines, a depletion of these neurotransmitters and reduction of these neurotransmitter functions may occur. Indeed, subchronic and central administration of IL-1 β has been found to reduce norepinephrine and serotonin concentrations in the limbic system, a similarity that has been reported to occur in depression. In addition, lower cerebrospinal fluid (CSF) concentrations of serotonin metabolite 5-hydroxyindoleacetic acid (5-HIAA) and the dopamine metabolite homovanillic acid (HVA) are correlated with depressive mood and suicidal ideation in depressed patients. However, what causes the neurotransmitter deficiency is remain elusive. Recent research has shown that the catabolism of tryptophan (a precursor of serotonin) is stimulated under the influence of stress and inflammation by the induction of the enzymes tryptophan pyrrolase and indoleamine 2,3-dioxygenase (IDO). Thus, the reduction in blood concentration of tryptophan under these circumstances leads a reduction of formation of brain serotonin. Furthermore, low levels of n-3 fatty acids and a high n-6 to n-3 fatty acid ratio were correlated with low CSF concentrations of 5-HIAA and HVA. In the clinical study, EPA has been found to significantly improve depressive symptoms, which may be related to the modulation of n-3 fatty acids on these neurotransmitters. Indeed, animal experiments have shown that EPA can increase serotonin and dopamine synthesis, and attenuate the IL-1-induced changes in the noradrenergic and serotonergic systems. These results provide neurochemical mechanism by which EPA may effectively treat depressive illness.

As previously mentioned, the n-3 and n-6 fatty acids are agonist and antagonize one another to maintain the homeostatic balance. Different combinations and ratios of n-3 and n-6 fatty acids have been studied in inflammation or stress-induced rat models. For example, the effect of 0.5% EPA, n-6 fatty acid γ -linolenic acid (GLA), AA, the combination of EPA and AA, and EPA and GLA on changes in IL-1-treated rats has been studied. The GLA had little effect on the IL-1-induced changes in behavior and the rise in the corticosterone concentration. Similarly, the n-6 fatty acid AA had no effect on these changes induced by IL-1. However, AA alone enhanced the basal inflammatory response, raised the serum corticosterone

concentrations, and induced anxiety-like behavior in the elevated plus-maze. Even though the prostaglandin E2 (PGE2), derived from AA, was suppressed significantly, the combination of EPA and GLA did not attenuate the behavioral and corticosterone changes following the administration of IL-1. Neither EPA and AA combination attenuated both the inflammatory and behavioral changes induced by IL-1. Following the ICV administration of IL-1, the decreased norepinephrine (NE) and serotonin, and increased dopamine (DA) and serotonin turnovers in several brain regions were attenuated by EPA, whereas the EPA and GLA combination reversed some, but not all, of these changes. These results also showed that only the combination of EPA and GLA in a one-to-one ratio has stronger anti-inflammatory effects.

In another experimental study, the effect of a diet of the n-3 and n-6 fatty acids was reported in two stress models, namely cortisol injection and immersion in a 10°C saline bath (cold stress). Compared to control rats (saline injection or immersion in 20°C saline bath), the cortisol injection or cold stress largely increased the blood concentration of cortisol and cholesterol. The mixed diet containing α -linolenic (n-3 fatty acid; 0.92 g ml⁻¹) and linoleic (n-6 fatty acid; 0.90 g ml⁻¹) for 3 weeks reversed these increases induced by cortisol or cold stress. It has been also reported that the most effective ratio of n-6 to n-3 fatty acids is four to one. The interaction between n-3 and n-6 fatty acids, and the mechanism whereby some combinations have different effects, needs to be further studied.

Clinical Investigation of Fatty Acid Composition in Major Depression

Epidemiological studies in various countries suggest that decreased n-3 fatty acid consumption correlates with the increasing rates of depression. However, so far there has been no direct measurement of PUFAs in the postmortem brains of depressed patients. Thus, current information depends on the data from measurements in plasma and erythrocyte membranes. In a study, containing 36 patients with major depression, 14 patients with minor depression and matched healthy controls, it was shown that patients with major depression had significantly lower total n-3 fatty acids in cholesteryl esters and lower EPA in both cholesteryl esters and phospholipids. There was also a higher ratio of AA to EPA in patients with major depression. In addition, the level of EPA in cholesteryl esters and plasma was negatively correlated with the score on the Hamilton Depression Rating scale (HAM-D).

In another clinical investigation, the erythrocyte membrane phospholipid level and dietary intake of polyunsaturated fatty acids were studied in depressed patients and their matched controls. A significant depletion of erythrocyte membrane n-3 fatty acids was found. Both erythrocyte membrane levels and diet intake of n-3 fatty acids were negatively correlated with the severity of depression. However, antidepressant treatments had no significant effect on the fatty acid levels of the cholesteryl esters. There is also evidence that pregnancy leads to a depletion of maternal serum n-3 fatty acid DHA and that, after delivery, the maternal serum DHA steadily declines further. One study investigated whether the postpartum fatty acid profile of maternal serum and cholesteryl esters differs in women who develop postpartum depression compared to controls. The fatty acid composition was compared shortly after delivery in 10 women who developed postpartum depression and 38 women who did not. After delivery, DHA and the sum of the n-3 fatty acids in phospholipids and cholesteryl esters were significantly lower in the group of mothers who developed a postpartum depression. The ratio of n-6 to n-3 fatty acids in phospholipids was significantly higher in the postpartum depressed group compared to the controls. Thus, the abnormalities in fatty acid status previously observed in major depression are now also confirmed in postpartum depression. These results indicate that pregnant women who are at risk of developing postpartum depression may benefit from a prophylactic treatment with n-3 PUFAs, such as a combination of EPA and DHA.

The n-3 Fatty Acids in the Treatment of Depression

Clinical investigations have demonstrated that EPA can effectively treat depression. In a double-blind placebo-controlled trial, EPA was added to ongoing antidepressant therapy. Patients with major depression were allocated at random to take either 2 g daily ethyl-EPA or placebo for 4 weeks. The treatment effects were assessed using HAM-D. There was a strong and highly significant treatment effect favoring EPA over the placebo. More recently, a dose-ranging study of ethyl-EPA was conducted in primary-care patients who suffered from treatment-unresponsive depression. The patients were randomly allocated to take ethyl-EPA (1 g, 2 g, or 4 g daily, or a matching placebo) for 12 weeks. Mood was assessed using the HAM-D, Montgomery-Asberg Depression Rating Scale (MADRS), and Beck Depression Inventory (BDI). The greatest improvement was shown in the

EPA 1 g day⁻¹ group compared to the 2 or 4 g groups. All three factors of the HAM-D, 9 of 10 items on the MADRS, and 20 items of BDI showed significant differences between groups taking ethyl-EPA and the placebo, with a particularly large treatment effect on BDI symptoms of sadness, pessimism, inability to work, sleep disturbance, and loss of libido. Adverse effects were equally distributed between the placebo and 1 g EPA. However, in two clinical studies, DHA at a dose of 1.15 or 2 g was found to be no better than a placebo. Nevertheless, these results strongly suggest that EPA may be an effective antidepressant that lacks many adverse effects that are associated with conventional antidepressants.

In conclusion, the n-3 and n-6 fatty acids play different roles in inflammation, stress, and depression; the n-3 fatty acids decreased them, whereas the n-6 fatty acids increased them. The understanding of this new research area is still limited, and most mechanisms are unknown. Future studies should consider (1) the interaction between lipids and neurotransmitter receptor functions, (2) the interaction between membrane n-3 and n-6 fatty acids, (3) the ideal balanced diet for the n-3 and n-6 fatty acids in normal people and in patients with various diseases.

See Also the Following Articles

Cytokines; Cytokines, Stress, and Depression; Genetic Factors and Stress; Immune Cell Distribution, Effects of Stress on; Immune Function, Stress-Induced Enhancement; Depression, Immunological Aspects.

Further Reading

- De Vriese, S. R., Christophe, A. B. and Maes, M. (2003). Lowered serum n-3 polyunsaturated fatty acid (PUFA) levels predict the occurrence of postpartum depression: further evidence that lowered n-PUFAs are related to major depression. *Life Science* 73, 3181–3187.
- Peet, M. and Stokes, C. (2005). Omega-3 fatty acids in the treatment of psychiatric disorders. *Drugs* 65, 1051–1059.
- Peet, M., Glen, I. and Horrobin, D. F. (2003). *Phospholipid spectrum disorder in psychiatry and neurology* (2nd edn.). Lancashire, UK: Marius Press.
- Song, C. and Leonard, B. E. (2000). *Fundamentals of psychoneuroimmunology*. Chichester, UK: John Wiley & Sons.
- Song, C., Li, X. W., Leonard, B. E., et al. (2003). Effects of dietary n-3 or n-6 fatty acids on interleukin-1 β -induced anxiety, stress and inflammatory responses in rats. *Journal of Lipid Research* 44, 1984–1991.
- Song, C., Phillips, A. G., Leonard, B. E., et al. (2004). Ethyl-eicosapentaenoic acid ingestion prevents corticosterone-mediated memory impairment induced by central

- administration of interleukin-1 β in rats. *Molecular Psychiatry* 9, 630–638.
- Takeuchi, T., Iwanaga, M. and Harada, E. (2003). Possible regulatory mechanism of DHA-induced anti-stress reaction in rats. *Brain Research* 964, 136–143.

- Yehuda, S., Rabinovitz, S., Carasso, R. L., et al. (2000). Fatty acid mixture counters stress changes in cortisol, cholesterol, and impair learning. *International Journal of Neuroscience* 101, 73–87.

Depression Models

K Matthews and C Stewart

University of Dundee, Dundee, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by K Matthews, volume 1, pp 675–681,

© 2000, Elsevier Inc.

Validity and Utility of Depression Models

Stress and Depression

Neurotoxic Effects of Stress-Related Hormones

Depression-Related Brain Pathology

Behavioral Correlates of Depression

Anatomy of Reward

Stress Models of Depression

Conclusion

Glossary

<i>Anhedonia</i>	The loss of the capacity to experience pleasure, a core symptom of depression.
<i>Intracranial self-stimulation</i>	An experimental technique whereby a small electrical charge can be delivered directly into the substance of the brain of an awake, freely moving subject via an implanted electrode.
<i>Limbic system</i>	An arrangement of anatomically related brain structures that are hypothesized to play a critical role in the production and regulation of emotion.
<i>Mesolimbic dopamine projection</i>	A grouping of dopamine-secreting neurons that arise in the midbrain to innervate a range of higher brain structures that includes the nucleus accumbens in the ventral striatum.
<i>Validity</i>	The authenticity of a presumed relationship between two separate events or observations.

Validity and Utility of Depression Models

Depression models serve many different purposes and the inherent value of each is determined by the

specific aims and utility of the individual model. It is therefore critical to establish the explicit purpose intended for that preparation in order to determine its validity and utility. There are several scholarly and comprehensive reviews of procedural details and the relative merits of different animal models of depression that achieve this purpose (see Further Reading). Although depression models can be categorized and evaluated according to almost any aspect of their construction, there is general agreement on the utility of addressing different aspects of validity separately. The validity spectrum includes construct validity and the related concepts of discriminant and convergent validity, predictive validity, face validity, and etiological validity.

The construct validity of a model refers to the accuracy with which the model replicates the key abnormalities or phenomena under study within the clinical condition. For example, if a model somehow encapsulated a blunting of a capacity to experience pleasure, it could be considered to demonstrate substantial construct validity with respect to anhedonia, one of the defining features of clinical depression. Of course, to comment on construct validity there is an implicit requirement that the target phenomenon be well defined, either behaviorally or neurobiologically, within the clinical state. For many aspects of depressive disorders, this remains problematic. Nevertheless, judgments can be made as to whether observations in a model appear to bear a meaningful relationship to those of the clinical disorder. Models with strong construct validity represent attempts to replicate core clinical phenomena in a laboratory context and are driven by specific hypotheses concerning the etiology, clinical features, and pathophysiology of depression. Such models are frequently labeled theoretical models.

A closely related concept to construct validity is etiological validity, whereby the circumstances or events that lead to development of the target phenomenon in the model are substantially similar to those generating the same phenomenon in the clinical disorder. Etiological validity can usefully be considered a

refinement of the more general concept of construct validity. For many depression models, the experimenter-initiated intervention is considered to mimic an environmental event of causal significance. However, such assumed relationships are problematic because the causal antecedents of depressive disorders, including those that are best studied, genetic vulnerability and stress, remain poorly defined.

The predictive validity of a model refers to its ability to correctly discriminate between interventions that are known to have effective antidepressant properties. Typically, these interventions are drug treatments. Depression models with strong predictive validity are widely used within the pharmaceutical industry to screen new molecules for potential antidepressant activity, hence the term assay model. The most potent models within this category are capable of discriminating pharmacologically diverse chemical antidepressant treatments with few false positive or negative results. For a strong assay model, there is no requirement that aspects of the preparation bear meaningful relationship with features of the clinical disorder.

The face validity of a model is whether specific features of the model bear resemblance to discrete aspects of the clinical disorder, that is, whether a phenomenological similarity exists, irrespective of whether there is a plausible or established theoretical relationship between the two observations. In this respect, face validity can be differentiated from construct validity. When two or more depression models are being compared with one another, regardless of whether they are theoretical or assay, it is possible to comment on the degree to which observations from each appear to reflect aspects of the same construct. Hence, two or more different models may demonstrate convergent or discriminant validity. Convergent validity is the degree to which the same aspect of a single phenomenon or construct can be accessed and quantified by different models, whereas discriminant validity is the capacity of a model to measure different aspects of a single target phenomenon from those measured by related models. All models can be judged on the degree to which the target phenomena can be reproduced within and between studies – the reliability of the model. In the absence of reliability, no model can be considered to have utility. Hence, both validity and reliability are essential for any animal model of depression.

Stress and Depression

An enduring clinical literature suggests that individual vulnerability to stress and subsequent predisposition to develop certain disease states, notably

depression, are related, at least in part, to a history of early environmental adversity. Exposure to early trauma, for example sexual and physical abuse or other types of early disadvantage, can increase several-fold the risk of being diagnosed with a depressive illness in adulthood. Similarly, the onset and recurrence of adult depression can reliably be predicted by the presence of environmental stressors, often labeled life events. Some individuals may have a genetic propensity to select themselves into high-risk environments, but epidemiological studies using identical and nonidentical twins have shown that there is still a substantial causal relationship between stressful life events and depression. However, this relationship is greatest for the first episode and seems less important in recurrent episodes.

Neurotoxic Effects of Stress-Related Hormones

In humans and other animals, hypothalamic-pituitary-adrenal (HPA) axis activation leads to the release of glucocorticoid hormone (cortisol in humans and corticosterone in other species) from the cortex of the adrenal glands. Glucocorticoid hormones are released in response to a wide variety of psychological and physical stressors and represent a core component of adaptive responsiveness to environmental change and threat. However, high levels of circulating glucocorticoids may have a detrimental effect on brain circuitry and function particularly at certain stages of development. Clinical studies have consistently implicated abnormalities in the regulation of key neuroendocrine responses to stress in a proportion of patients with depression, with a hyperactivity of the HPA axis that is probably driven by hypersecretion of the hypothalamic peptide corticotropin-releasing hormone (CRH). Certain areas of the brain, including parts of the hippocampal formation, are more sensitive to damage from high levels of glucocorticoids.

Depression-Related Brain Pathology

Depression has been associated with abnormal function and/or pathological changes in several cortical, limbic, and striatal brain structures. There is an extensive amount of evidence suggesting alterations in the anterior cingulate cortex in mood disorders and also the orbital and dorsolateral regions of the prefrontal cortex. Two other structures within the temporal lobe, the hippocampal formation and the amygdala, show consistent reductions in size in major depression. There are close anatomical associations among all these structures, which form part of

Maclean's concept of the limbic system, which acts to integrate emotional expression with visceral and endocrine responses. Although support for the notion of a specific neural basis for emotion and motivation has accrued steadily from subsequent experimental, accidental, and therapeutic lesion studies in animals and humans, it is less clear whether the changes seen in the brains of patients with a diagnosis of depression can be specifically related to different symptoms that they show. The reduction in hippocampal gray matter may correlate with the severity and duration of illness and may also be responsible for some of the cognitive impairments seen in severe depression. However, some studies suggest that a decrease in hippocampal volume may be a specific marker of depression diagnosed in patients with a history of early trauma. Abnormalities in areas of the prefrontal cortex have been linked with the failure to anticipate positive outcomes or increased sensitivity to negative outcomes.

Behavioral Correlates of Depression

Depression is a very heterogeneous disorder, and those affected may display very different symptom clusters. Some patients may have pervasive symptoms of negative affect and anxiety, whereas others have a more prominent lack of positive affect. This becomes important when we consider which animal behaviors to assess when trying to model the disorder. Much emphasis has been placed on the capacity for stress to reduce or impair the performance of behaviors that serve to bring animals into contact with key stimuli or appetitive reinforcers within their environment. Engagement with such stimuli is inferred to convey pleasant or pleasurable hedonic consequences. Reduced engagement is generally interpreted as evidence that the stimuli no longer confer such consequences, at least to the same degree, and this has been deemed comparable to anhedonia, a core feature incorporated in current psychiatric diagnostic and classification systems of depression.

Anatomy of Reward

Research during the 1970s and 1980s identified the mesolimbic dopamine projection as a key component of neural circuitry mediating the psychological construct of reward. This pathway arises from cells bodies located in the ventral tegmentum in the midbrain, projecting upward and forward to the amygdala, lateral septum, bed nucleus of the stria terminalis, hippocampus, and nucleus accumbens (a component of the ventral striatum). On its journey forward, the

mesolimbic dopamine system passes through the lateral hypothalamus via a thick cable of fibers known as the medial forebrain bundle. Appetitive reinforcers (a term used to describe pleasurable hedonic consequences), including biologically meaningful stimuli such as food, sex, and drugs of abuse (e.g., cocaine and heroin), are dependent on the functional integrity of the mesolimbic projection system for their behavioral effects. Such appetitive reinforcers elicit approach behavior and contact. Thus, the nucleus accumbens, a structurally defined subregion of the ventral striatum, where ascending dopaminergic neurons synapse with intrinsic striatal neurons and descending glutamatergic projects from limbic cortical structures (such as the hippocampus, amygdala, and prefrontal cortex), represents a critical locus for the processing and execution of goal-directed behavior. Interference with dopaminergic transmission, either by lesioning the cell bodies in the midbrain or by the local administration of dopamine receptor antagonist drugs, leads to a profound disruption of a range of rewarded behaviors, including intracranial self-stimulation (ICSS), the self-administration of psychostimulant drugs of abuse (e.g., cocaine and amphetamine), and responding for primary and secondary reinforcers. In addition, the local infusion of dopamine receptor agonist drugs into the nucleus accumbens and the mesolimbic dopamine system are widely believed to represent the critical site through which drugs of abuse exert their reinforcing and addictive effects.

Stress Models of Depression

There are many experimental demonstrations of the capacity for stress to disrupt, or at least modify, behavior directed toward positively reinforcing stimuli. In this article, we focus on those that have held the greatest promise as valid animal models of depressive disorder.

Learned Helplessness

This is perhaps the best-known stress model of depression but it is covered in detail elsewhere (see *Learned Helplessness*).

Inescapable Foot Shock and Intracranial Self-Stimulation

Early research supporting the existence of specific reward or pleasure circuits in the brain came from studies showing that electrical stimulation of areas of the hypothalamus led to a powerful augmentation of approach behavior and vigorous self-administration of electrical current. Subsequent studies identified the

mesolimbic dopamine system, which passes through the lateral hypothalamus as the medial forebrain bundle, as the key component of this system. A specific hypothesis that linked human depressive behaviors with dysfunction of brain reward pathways was formulated. This hypothesis has been extensively tested in a series of experiments evaluating the interactions among responding for rewarding electrical brain stimulation, stress, and antidepressant drug administration. Rats demonstrate regional specificity within the brain for stress-induced decrements in responding to ICSS, most notably within the terminal fields of the mesolimbic and mesocortical dopamine projections. Exposure to unpredictable electric foot shock induced a profound and enduring reduction in responding when ICSS electrodes were sited in the ventral tegmentum, nucleus accumbens, and medial prefrontal cortex, but not in the adjacent dorsal striatum. This regional specificity exhibited a close anatomical concordance with structures known to respond to stress with alterations in dopamine function. Furthermore, chronic administration of the prototypical chemical antidepressant drug desmethylimipramine (DMI) prior to stress exposure significantly minimized the subsequent ICSS performance deficits. These data have been interpreted as demonstrating beneficial, or protective, effects of antidepressant drugs on the neural responses to stress, specifically in areas that are critically involved in the integration of behavioral responses to reward. However, additional studies employing different drug treatments, different electrode placements, and different strains of rat and mice have generated somewhat variable results.

Behavioral Despair

Also known as the Porsolt test, the forced-swim test (FST) involves forcing rats (or mice) to swim in a tall water-filled cylinder from which there is no escape. The water is sufficiently deep to preclude the animals from resting its limbs or tail on the bottom. When placed in this apparatus for the first time, the animal swims vigorously to maintain its head above water. With appreciation of the inescapable nature of its predicament, attempts to escape from the cylinder rapidly cease. Thereafter, the animal adopts an immobile posture that has been interpreted as a theoretical analog of hopelessness with respect to the possibility for escape. With repeated testing, the time to the onset of immobility decreases in control animals, and this has been interpreted as a theoretical analogue of learned helplessness. Exposing animals to the FST has also been reported to cause some neurochemical changes that are consistent with a depressogenic phenotype. For example, it can transiently reduce

serotonin and norepinephrin levels in both cortical and limbic structures and alter the sensitivity to CRH. Although there are obvious problems with the construct validity of this model, behavioral despair enjoys one of the most impressive records of all depression models with respect to pharmacological sensitivity and specificity. Hence, this model enjoys impressive predictive validity. The question of what exactly this model measures remains unanswered. Whereas the other stress-based models described here demonstrate altered responding toward rewarding stimuli, this test reflects something quite different. Although an operational definition of the phenomenon is simple, a description of the relationship to human affective disorder is not.

Chronic Unpredictable Mild Stress

In an attempt to design models with greater face and etiological validity, several different procedures have been devised that employ the sequential presentation of relatively mild stressors to rats or mice. In these models, the effects of stress are usually quantified by the degree to which they alter the intake of drinking solutions that are preferred over standard drinking water under normal circumstances. Katz and colleagues first explored the effect of unpredictable exposure to different types of stressors, including electric shock, cold swim, water deprivation, and shaker stress, and observed that rats failed to increase fluid intake when saccharin was added to their drinking water. This work was extended by Willner and colleagues, who put the emphasis on the sequential exposure to what they considered mild stressors such as periods of strobe lighting, food and water deprivation, cage tilting, and enforced group housing. These paradigms were reported to produce a loss of preference for or reductions in the absolute intake of sucrose, saccharin, and saline solutions. The behavioral deficits seen with chronic mild stress are reputed to respond to the chronic administration of chemical antidepressant drugs. However, like some other animals models of depression, there have been significant problems establishing these procedures reliably in different laboratories. The behavioral effects are at best modest, and large numbers of subjects are required. Many studies have also failed to include appropriate control measures.

Social Stress Models

This group of models relies on a different approach, disturbing the social environment either during early or later life. These manipulations are potent but naturalistic stressors. They would seem to have considerable etiological validity, although there is still some

debate over the degree to which stress has a causal role in depressive illness.

The disruption of mother–infant bonding or of peer bonding in young social animals generates a constellation of acute and chronic behavioral changes that resemble human depression and vulnerability to depression, respectively. Across a broad range of species, responses to separation follow a predictable time course with definable phases. Initially, the infant displays so-called protest behaviors that include increased locomotor activity and vocalization. This subsides in time and is generally replaced by a second phase that is characterized by locomotor inactivity and an apparent disinterest in motivationally salient external environmental stimuli, so-called despair. Depending on the species studied, there may also follow a third phase of detachment, during which the infant displays indifference to being reunited with its mother and/or peers. Social behaviors in adulthood are often profoundly abnormal in such animals. There is considerable cross-species generality of this bi- or triphasic response to social separation, suggesting that it may be hardwired in the brains of many social animals.

Nonhuman primates In the light of their phylogenetic proximity to humans, their advanced cognitive capacity, and their complex social structures, studies in monkeys have provided the bulk of the behavioral and neurobiological data in support of social separation models of depression. After numerous studies demonstrated the pathological effects of prolonged separation (weeks or months) of infant monkeys from their mothers, Hinde and colleagues reported enduring effects of brief separation (6 days) during infancy on subsequent behavioral development when tested at 12 and 24 months. Alterations were observed with respect to the separated monkeys' responses to novelty and their independent exploration of the immediate environment. Thus, even brief periods of maternal separation appeared capable of inducing enduring alterations in behavior. However, there is no convincing evidence that either the acute or the chronic effects of early separation in monkeys bear meaningful relationships to human depression.

In naturally occurring adverse social circumstances, such as falling in rank, monkeys can exhibit a prolonged withdrawal from social interaction and a suppression of aggression. Thus, some nonhuman primate analogs of the common precipitants of human depressive episodes can induce despairlike reactions and behavioral changes reminiscent of human depression, thus conferring a degree of face validity. In turn, these responses are strongly influenced by

genes (species differences), early social experience, social support networks, and other physical aspects of environment. Also, these despair responses are associated with physiological changes such as alterations in sleep architecture, phenomena that frequently accompany depressive episodes in humans.

Mindful of the value of predictive validity in an animal model of depression, the effects of psychotropic drugs have also been tested in separated monkeys. Several groups have reported an attenuation of behavioral disruption following maternal and peer separation with chronic chemical antidepressant treatment. However, the pharmacological specificity of these effects is questionable because chlorpromazine (an antipsychotic drug with weak, if any, clinical antidepressant activity), alcohol, and diazepam have each been reported to have similar effects, depending on the dose administered and actual measures recorded. However, repeated electroconvulsive stimulation, an analog of the highly effective antidepressant treatment electroconvulsive therapy (ECT), has also been shown to have an antidespair profile of effects.

Nonprimate models Recent work suggests that manipulation of the early social environment of the rat and other rodents may represent useful depression models. Repeated neonatal maternal separation in the rat leads to a robust impairment of the control of behavior by primary and conditioned reward in adulthood, with reduced responses to environmental novelty, psychostimulant drugs, and changes in reward magnitude. Some of these early experience-induced changes may model the hedonic changes found in human depression. To date, the only studies of the effects of antidepressant drugs on long-term consequences of maternal separation have examined their effects on measures of HPA function and ethanol consumption.

Psychosocial stress paradigms have been used successfully in naturally territorial nonprimate species such as the tree shrew. If adults males are kept in visual and olfactory contact in the laboratory after a period of conflict, a stable dominant–subordinate relationship develops and the subordinate exhibits sleep disturbance, reduced grooming behavior, neuroendocrine changes reminiscent of those seen in some depressed patients, and both structural and functional changes to the hippocampal formation. Some of these effects can be ameliorated by treatment with a tricyclic antidepressant or with the atypical agent tianeptine. Social defeat procedures have also been used in rodents, but it is less clear whether they consistently produce any changes in reward-related behavior.

Conclusion

Depression models fulfill many different purposes and can be evaluated according to different criteria. They have contributed to our understanding of the neural basis of reward and how stress affects brain function. Hypotheses generated in the clinical have stimulated model design and information from models has informed clinical practice. The inherent difficulties in attempting to construct valid animal models of human depression are largely offset by the potential utility of the approach. Depression models offer the opportunity to control and manipulate variables that epidemiological and clinical research suggests may be influential, for example, genotype; early social experience; and magnitude, type, and duration of environmental stressors. Cautious interpretation of model data can advance our understanding of a complex and devastating mental disorder.

See Also the Following Articles

Affective Disorders; Antidepressant Actions on Glucocorticoid Receptors; Depression and Manic-Depressive Illness; Dopamine, Central; Glucocorticoids – Adverse Effects on the Nervous System; Hippocampus, Corticosteroid Effects on; Learned Helplessness; Maternal Deprivation.

Further Reading

- Akiskal, H. S. and McKinney, W. T., Jr. (1973). Depressive disorders: toward a unified hypothesis. *Science* **182**, 20–29.
- Chapman, D. P., Whitfield, C. L., Felitti, V. J., et al. (2004). Adverse childhood experiences and the risk of depressive disorders in adulthood. *Journal of Affective Disorders* **82**, 217–225.
- Davidson, R. J., Pizzagalli, D., Nitschke, J. B., et al. (2002). Depression: perspectives from affective neuroscience. *Annual Review of Psychology* **53**, 545–574.
- Everitt, B. J. and Keverne, E. B. (1979). Models of depression based on behavioural observation in experimental animals. In: Paykel, E. S. & Coppen, A. (eds.) *Psychopharmacology of affective disorders*, pp. 41–59. Oxford: Oxford University Press.
- Fibiger, H. C. and Phillips, A. G. (1988). Mesocorticolimbic dopamine systems and reward. *Annals of the New York Academy of Sciences* **537**, 206–215.
- Fuchs, E. and Flugge, G. (2003). Chronic social stress: effects on limbic brain structures. *Physiology and Behavior* **79**, 417–427.
- Heim, C., Owens, M. J., Plotzky, P. M., et al. (1997). Persistent changes in corticotropin-releasing factor systems due to early life stress: relationship to the pathophysiology of major depression and post-traumatic stress disorder. *Psychopharmacological Bulletin* **33**, 185–192.
- Hinde, R. A., Spencer-Booth, Y. and Bruce, M. (1966). Effects of 6-day maternal deprivation on rhesus monkey infants. *Nature* **210**, 1021–1033.
- Holmes, P. V. (2003). Rodent models of depression: re-examining validity without anthropomorphic inference. *Critical Reviews in Neurobiology* **15**(2), 143–174.
- Katz, R. J. (1981). Animal models and human depressive disorders. *Neuroscience and Biobehavioral Reviews* **5**, 213–246.
- Katz, R. J. (1982). Animal model of depression: pharmacological sensitivity of a hedonic deficit. *Pharmacology, Biochemistry and Behavior* **16**, 965–968.
- Kendler, K. S., Karkowski, L. M. and Prescott, C. A. (1999). Causal relationship between stressful life events and the onset of major depression. *American Journal of Psychiatry* **156**, 837–841.
- Koob, G. F., Sanna, P. P. and Bloom, F. E. (1998). Neuroscience of addiction. *Neuron* **21**, 467–476.
- Kraemer, G. W. (1992). A psychobiological theory of attachment. *Behavioral and Brain Sciences* **15**(3), 493–511.
- Kramer, M., Heimke, C. and Fuchs, E. (1999). Chronic psychosocial stress and antidepressant treatment in tree shrews: time-dependent behavioural and endocrine effects. *Neuroscience and Biobehavioral Reviews* **23**, 937–947.
- Maclean, P. D. (1949). Psychosomatic disease and the visceral brain; recent developments bearing on the Papez theory of emotion. *Psychosomatic Medicine* **11**, 338–353.
- Matthews, K., Wilkinson, L. S. and Robbins, T. W. (1996). Repeated maternal separation of preweanling rats attenuates behavioral responses to primary and conditioned incentives in adulthood. *Physiology and Behavior* **59**, 99–107.
- McKinney, W. T. and Bunney, W. E. (1969). Animal model of depression: review of evidence and implications for research. *Archives of General Psychiatry* **21**, 240–248.
- McKinney, W. T., Suomi, S. J. and Harlow, H. F. (1971). Depression in primates. *American Journal of Psychiatry* **127**, 49–56.
- Olds, J. and Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative Physiology and Psychology* **47**, 419–427.
- Paykel, E. S. and Hollyman, J. A. (1984). Life events and depression – a psychiatric view. *Trends in Neurosciences* **7**(12), 478–481.
- Porsolt, R. D., LePichon, M. and Jalfre, M. (1977). Depression: a new animal model sensitive to antidepressant treatments. *Nature* **266**, 730–732.
- Robbins, T. W. and Everitt, B. J. (1992). Functions of dopamine in the dorsal and ventral striatum. *Seminars in the Neurosciences* **4**, 119–127.
- Sadowski, H., Ugarte, B., Kolvin, I., et al. (1999). Early life family disadvantages and major depression in adulthood. *British Journal of Psychiatry* **174**, 112–120.
- Suomi, S. J., Seaman, S. F., Lewis, J. K., et al. (1978). Effects of imipramine treatment of separation-induced social disorders in rhesus monkeys. *Archives of General Psychiatry* **35**, 321–325.

- Willner, P. (1984). The validity of animal models of depression. *Psychopharmacology* 83, 1–16.
- Willner, P., Muscat, R. and Papp, M. (1992). Chronic mild stress-induced anhedonia: a realistic animal model of depression. *Neuroscience and Biobehavioral Reviews* 16(4), 525–534.

- Zacharko, R. M. and Anisman, H. (1991). Stressor-induced anhedonia in the mesocorticolimbic system. *Neuroscience and Biobehavioral Reviews* 15, 391–405.

Depression, Immunological Aspects

M R Irwin

UCLA Semel Institute for Neuroscience, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction

Immune System

Biological Connections between the CNS and Immune System

Depression Influences on Immunity

Clinical Moderating Variables

Behavioral Mechanisms: Role of Insomnia

Depression Treatment: Effects of Antidepressant Medications

Cytokines Influences on the CNS and Behavior: Implications for Depression

Clinical Implications of Psychoneuroimmunology

Glossary

<i>Cytokine</i>	A soluble protein that is produced by lymphocytes, monocytes, and macrophages, as well as by tissue cells and cells of the central nervous system; cytokines are released from a cell to influence the activity of other cells by binding a specific receptor and activating cell function.
<i>Immune system</i>	A variety of interactive cells and soluble molecules that provide for the body's defense against invading external pathogens; the immune system is composed of two components: humoral or antibody responses and cell-mediated responses.
<i>Inflammation</i>	A immune response that occurs to rid the body of an infectious agent or to remove and repair damaged tissue.
<i>Major depressive disorder</i>	A disorder characterized by a period of at least 2 weeks of depressed mood or loss of interest in nearly all activities with the presence of at least four other symptoms drawn from a list that includes

changes in appetite or weight, sleep, or activity; decreased energy; feelings of worthlessness or guilt; difficulty thinking or concentrating; and recurrent thoughts of death or suicidal ideation.

A research field that investigates the physiological systems that integrate behavioral and immunological responses and examines the interactions between the nervous system and the immune system in the relationship between behavior and health.

Psychoneuroimmunology

Introduction

Many immunological changes reliably occur in patients with major depressive disorder, as recently described in comprehensive meta-analyses of over 180 studies with more than 40 immune measures. This article provides a brief overview of the immune system and the pathways that mediate connections between the central nervous system (CNS) and the immune system. A detailed review of the various immune findings that occur in major depression is presented with consideration of the behavioral correlates and biological mechanisms that might contribute to immune changes in major depression. Finally, the clinical implications of immune alterations in depression for infectious disease risk and inflammatory disorders are addressed.

Immune System

The immune system is the body's defense against invading external pathogens such as viruses and bacteria and from abnormal internal cells such as tumors. Innate immunity refers to the body's resistance to pathogens that operates in a nonspecific way without recognition of the different nature of various pathogens, whereas specific immunity is acquired in response to the identification of non-self molecules called antigens. Macrophages and granulocytes are

examples of nonspecific immune cells that react to tissue damage by consuming debris and invading organisms. Natural killer (NK) cells are another example of nonspecific immunity that acts to kill virally infected cells in a nonspecific way without need for prior exposure or recognition. In contrast, each T cell or B cell is genetically programmed to attack a specific target by secreting antibodies (B cell) or by killing cells of the body that harbor a virus (T cell). Both innate and specific immunity are orchestrated by the release of interleukins or cytokines from immune cells; cytokines are protein messengers that regulate the immune cells. This cytokine network aids in the differentiation of the immune response and in the coordination of its magnitude and duration. For example, there are two main classes of cytokines secreted by the T cells. One class of cytokines, T helper type 1 (Th1) cytokines, supports T cell responses (e.g., the ability of T cells to kill virally infected cells), whereas another class of cytokines, T helper type 2 (Th2), supports an antibody-mediated humoral immune response. However, these immunoregulatory processes cannot be fully understood without taking into account the organism and the internal and external milieu in which innate and specific immune responses occur.

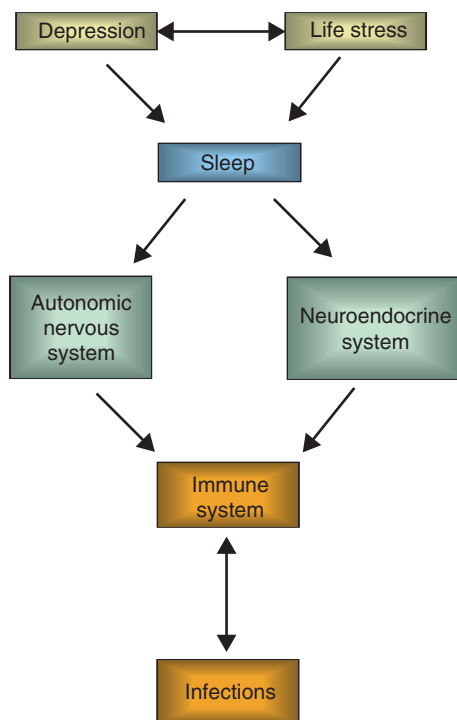


Figure 1 Hypothesized mediators of the effects of major depression on immunity and infectious disease risk in humans.

Biological Connections between the CNS and Immune System

Autonomic Nervous System

The CNS and the immune system are linked by two major physiological pathways: the hypothalamic-pituitary-adrenal (HPA) axis and the autonomic nervous system, composed of sympathetic and parasympathetic branches. The sympathetic nervous system (SNS) is a network of nerve cells running from the brain stem down the spinal cord and out into the body to contact a wide variety of organs, including the eyes, heart, lungs, stomach and intestines, joints, and skin. In organs where the immune system cells develop and respond to pathogens (e.g., bone marrow, thymus, spleen, and lymph nodes), sympathetic nerve terminals make contact with immune cells. Thus, sympathetic release of norepinephrine and neuropeptide Y, together with receptor binding of these neurotransmitters by immune cells, serve as the signal in this hard-wire connection between the brain and the immune system. In addition, sympathetic nerves penetrate into the adrenal gland and cause the release of epinephrine into the bloodstream, which circulates to immune cells as another sympathetic regulatory signal.

Many immune system cells change their behavior in the presence of neurotransmitters. Under both laboratory and naturalistic conditions, sympathetic activation has been shown to suppress the activity of diverse populations of immune cells including NK cells and T lymphocytes. In contrast, other aspects of the immune response can be enhanced. For example, catecholamines can increase the production of antibodies by B cells and the ability of macrophages to release cytokines and thereby signal the presence of a pathogen. Additional studies indicate that sympathetic activation can also shunt some immune system cells out of circulating blood and into the lymphoid organs (e.g., spleen, lymph nodes, thymus) while recruiting other types of immune cell into circulation (e.g., NK cells). In general, SNS activation can reduce the immune system's ability to destroy pathogens that live inside cells (e.g., viruses) via decreases of the cellular immune response, while sparing or enhancing the humoral immune response to pathogens that live outside cells (e.g., bacteria). Together, these observations are a cornerstone for understanding fundamental, neuroanatomic signaling between the autonomic nervous and immune systems.

Neuroendocrine Axis

The other way in which the brain can communicate with the immune system is the HPA system. This

process begins in the hypothalamus, an area of the brain that governs basic bodily processes such as temperature, thirst, and hunger. Following the release of neuroendocrine factors from the brain, the endocrine glands secrete hormones into the circulation, which reach various organs and bind to hormone receptors on the organs. Under conditions of psychological or physical stress, for example, the hypothalamus increases its release of corticotropin-releasing hormone (CRH) into a small network of blood vessels that descends into the pituitary gland. In response to CRH, the pituitary gland synthesizes adrenocorticotrophic hormone (ACTH), which travels through the bloodstream down to the adrenal glands and triggers the release of a steroid hormone called cortisol from the outer portion of the adrenal glands. Cortisol exerts diverse effects on a wide variety of physiological systems and also coordinates the actions of various cells involved in an immune response by altering the production of cytokines or immune messengers. Similar to sympathetic catecholamines and neuropeptide Y, cortisol can suppress the cellular immune response critical to defending the body against viral infections. Indeed, a synthetic analog of cortisol is often used to suppress excessive immune system responses (e.g., in autoimmune diseases such as arthritis or in allergic reactions such as the rash produced by poison oak). Cortisol can also prompt some immune cells to move out from circulating blood into lymphoid organs or peripheral tissues such as the skin. What is even more remarkable about the interactions between the neuroendocrine and immune system is that immune cells can also produce neuroendocrine peptides (e.g., endorphin, ACTH), which suggests that the brain, neuroendocrine axis, and immune system use the same molecular signals to communicate with each other.

Central Modulation of Immunity

Together, this converging evidence of brain-immune system interactions legitimizes the possibility that the brain has a physiological role in the regulation of immunity. Indeed, one key peptide involved in integrating neural and neuroendocrine control of visceral processes is CRH. Release of this peptide in the brain alters a variety of immune processes, including aspects of innate immunity, cellular immunity, and *in vivo* measures of antibody production. Relevant to immune alterations in depression, CRH is elevated in the cerebrospinal fluid of depressed patients. Hence, the brain through the endogenous release of CRH controls immune cells in lymphoid tissue in the same manner it controls other visceral organs, namely, by coordinating autonomic and neuroendocrine pathways; when these

pathways are blocked by specific factors that bind to sympathetic or hormone receptors, the effects of CRH or brain stimulation on immune function are also blocked.

Depression Influences on Immunity

Enumerative Measures

Evidence for increases in the total number of white blood cells and in the numbers and percentages of neutrophils and lymphocytes was among the first immunological changes identified in depressed persons. Further evaluation of lymphocyte numbers in depression has used phenotype-specific cell surface markers to enumerate lymphocyte subsets and found that depression is negatively related to the number and percentage of lymphocytes (B cells, T cells, T helper cells, and T suppressor/cytotoxic cells) as well as the NK cell phenotype.

Functional Measures

For the evaluation of the function of the immune system in depressed patients, a majority of studies have relied on results from assays of nonspecific mitogen-induced lymphocyte proliferation, mitogen-stimulated cytokine production, and NK cytotoxicity. More than a dozen studies have been conducted on lymphocyte proliferation in depression, and there is a reliable association between depression and lower proliferative responses to the three nonspecific mitogens, including phytohemagglutinin (PHA), concanavalin-A (Con A), and pokeweed (PWM). In addition, a number of independent laboratories have confirmed the finding of reduced NK activity in major depression.

Stimulated Cytokine Production

Studies of stimulated cytokine production have not yielded consistent findings. For example, increased lipopolysaccharide stimulated production of interleukin-1 β (IL-1 β) and IL-6 in depressed patients, but resulted in no change in the expression of another pro-inflammatory cytokine, tumor necrosis factor α (TNF- α). There are also reports of a shift in the relative balance of Th1 vs. Th2 cytokine production with increases in the capacity of lymphocytes to produce interferon in depression. However, no difference in the stimulated production of IL-2 has been found. These negative findings cannot be ascribed to differences in depressed samples, as depressed patients who show declines in NK activity show no difference in IL-2 production.

Inflammation and Circulating Levels of Inflammatory Markers

The presence of immune activation in major depression has also been evaluated by examining circulating levels of inflammatory markers, as well as Th1 cytokines; one study reported increases of plasma levels of IL-12 in a large cohort of depressed patients. Meta-analyses indicate that depression is associated with an increase in circulating levels of the pro-inflammatory cytokine IL-6. It is hypothesized that increases in circulating levels of pro-inflammatory cytokines are due to activation of monocyte populations, and increases in circulating levels of other pro-inflammatory cytokines such as TNF- α and IL-1 β have also been reported in depressed patients. Additional studies have extended these observations and assayed markers of systemic inflammation such as acute phase proteins and/or levels of soluble interleukin-2 receptor (sIL-2R), although findings are mixed. Nevertheless, increases in C-reactive protein have been found in association with depression with elevated levels in healthy depressed adults, as well as in those depressed patients with acute coronary syndrome. In turn, systemic immune activation is thought to lead to endothelial activation in depression with increases in the expression of soluble intercellular adhesion molecule.

Dissociation between Declines of Innate Immunity and Inflammatory Markers

Little attention has been given to the potential relationship between measures of innate immunity, such as NK activity, and levels of inflammatory markers in the context of major depression. In a recent study, levels of NK activity, circulating levels of IL-6, sIL-2R, and acute phase proteins were measured in patients with current major depressive disorder. Whereas patients with major depressive disorder showed lower NK activity and higher circulating levels of IL-6, levels of NK activity were not correlated with IL-6 or with other markers of immune activation including acute phase proteins or sIL-2R. Such findings have implications for understanding individual differences in the adverse health effects of major depressive disorder. Some depressed persons show reductions of cellular and innate immune responses that are associated with infectious disease susceptibility, whereas other studies report that depression is linked to immune activation that is associated with risk of inflammatory disorders such as rheumatoid arthritis and cardiovascular disease.

Viral-Specific Immune Measures

Extension of these nonspecific measures of immunity to viral-specific immune response has suggested a

functional decline in memory T cells that respond to at least one virus, namely, the varicella zoster virus, which is thought to be a surrogate marker for herpes zoster risk. Psychological stress is also associated with decline in specific immune responses to immunization against viral infections, although extension of this work to major depression has not yet been conducted.

Assays of *in Vivo* Responses

Basic observations in animals have raised the possibility that depression can alter *in vivo* immune responses, as administration of chronic stress suppresses the delayed type hypersensitivity (DTH) response. Translation of these findings suggests that suppression of the DTH response to a panel of antigenic challenges occurs in depression.

Clinical Moderating Variables

Heterogeneity in the effects of depression on immunity can be accounted for by a number of factors such as age, gender, ethnicity, adiposity, and health behaviors (e.g., smoking, alcohol consumption). Older adults show declines in cellular immunity, and the presence of comorbid depression appears to magnify further age-related immune alterations. Gender of the subject exerts differential effects on pituitary-adrenal and immune systems by modulating the sensitivity of target tissues, and women undergoing laboratory stress show exaggerated expression of cytokines that lead to inflammation as compared to men. Such inflammatory responses to stress may place depressed women at increased risk for autoimmune disorders. In contrast, declines of T cell and NK cell response appear to be more prominent in depressed men than in depressed women. Regarding ethnicity, African American ethnicity interacts with a history of alcohol consumption to exacerbate immune abnormalities, but the effects of ethnicity on depression-related immune alterations are not known. Increases in body mass index and the presence of obesity are associated with increases in markers of inflammation such as circulating levels of C-reactive protein and IL-6. It has been suggested that adiposity and greater body mass partially mediate the increase of inflammatory markers in depression, although other data indicate that elevated levels of inflammatory markers may occur only above a threshold of adiposity or in those with obesity. Finally, depressed patients who are comorbid for alcohol abuse or tobacco smoking show exaggerated declines of NK activity.

Behavioral Mechanisms: Role of Insomnia

Insomnia is one of the most common complaints of depressed subjects, with potential mediating effects on immune alterations in depression. Disordered sleep and loss of sleep are thought to adversely affect resistance to infectious disease, increase cancer risk, and alter inflammatory disease progression. Recent epidemiological data show that self-reported difficulty initiating sleep is a predictor of cardiovascular disease mortality, and objective measures of difficulty initiating sleep (i.e., prolonged sleep latency) yield a twofold elevated risk of death in a healthy older adult population. Animal studies further show that sleep deprivation impairs influenza viral clearance and increases rates of bacteremia, with translational data showing that acute sleep loss reduces immune response to immunization against influenza and hepatitis B.

In humans, normal sleep is associated with a redistribution of circulating lymphocyte subsets, increases of NK activity, increases of certain cytokines (e.g., IL-2, IL-6), and a relative shift toward Th1 cytokine expression that is independent of circadian processes. Conversely, sleep deprivation suppresses NK activity and IL-2 production, although prolonged sleep loss has been found to enhance measures of innate immunity and pro-inflammatory cytokine expression.

In depressed patients, subjective insomnia correlates with NK activity in depression, but not with other depressive symptoms. Recent studies in bereaved persons show that insomnia mediates the relationship between severe life stress and a decline of NK responses. Furthermore, in patients with primary insomnia, prolonged sleep latency and fragmentation of sleep are associated with nocturnal elevations of sympathetic catecholamines and declines in daytime levels of NK cell responses, similar to the abnormalities found in depression. Finally, EEG sleep measures in depressed patients showed that prolonged sleep latency and increases of REM density correlated with elevated levels of IL-6 and sICAM and fully accounted for the association between depression and IL-6.

Depression Treatment: Effects of Antidepressant Medications

Only a limited number of studies have investigated the clinical course of depression and changes of immunity in relation to antidepressant medication treatment and symptom resolution. In a longitudinal case-control study, depressed patients showed increases in NK activity during a 6-month course of

tricyclic antidepressant medication treatment, although improvements in NK activity correlated with declines in symptom severity and not medication treatment status at the follow-up. However, both *in vivo* and *in vitro* treatment with fluoxetine, a selective serotonin reuptake inhibitor, results in enhanced NK activity along with changes in depressive symptoms, consistent with the effects of a number of other selective serotonin reuptake inhibitors on NK responses. Other studies have focused on pro-inflammatory cytokine expression and inflammation and concluded that antidepressants decrease pro-inflammatory cytokine (e.g., IL-1 β , TNF- α , and IL-6) and induce a shift toward Th2 cytokine expression. However, such alterations in the production of cytokines appear to be confined to medication responders, suggesting that symptom resolution is a relatively more important predictor of the cytokine changes than antidepressant medication status.

Cytokines Influences on the CNS and Behavior: Implications for Depression

Not only does the brain participate in the regulation of immune responses, but also the CNS receives information from the periphery that an immune response is occurring with consequent changes in both electrical and neurochemical activity of the brain. During immunization to a novel protein antigen, the firing rate of neurons within the brain (e.g., ventromedial hypothalamus) increases at the time of peak production of antibody; this part of the brain controls autonomic activity. Cytokines released by immune cells are increasingly implicated as messengers in this bidirectional interaction, and the release of IL-1 following activation of macrophages with virus or other stimuli induces alterations of brain activity and changes in the metabolism of central brain chemicals and neurotransmitters such as norepinephrine, serotonin, and dopamine in discrete brain areas. Much recent data have focused on how these cytokines signal the brain given their large molecular size and inability to cross readily the blood-brain barrier. It is now known that IL-1 and possibly other inflammatory cytokines communicate with the brain by stimulating peripheral nerves such as the vagus that provide information to the brain. In sum, the immune system acts in many ways like a sensory organ, conveying information to the brain, which ultimately regulates neuroendocrine and autonomic outflow and the course of the immune response.

Immune activation also leads to changes of peripheral physiology and behaviors that are similar to a stress response. With peripheral immune activation, pro-inflammatory cytokines are expressed in the

CNS, CRH is released by the hypothalamus, and there is an induction of a pituitary adrenal response and autonomic activity. Coincident with these physiological changes, animals show reductions in activity, exploration of novel objects, social interactions, food and water intake, and a willingness to engage in sexual behaviors. Taken together, this pattern of behavioral changes (i.e., sickness behaviors) is similar to that found in animals exposed to fear or anxiety-arousing stimuli and can be reproduced by the central or peripheral administration of IL-1. In contrast, central administration of factors that block IL-1 antagonizes these effects. These cytokine-brain processes are also implicated in increased sensitivity to pain stimuli that is found following nerve or tissue injury.

Translation of these data to clinical samples suggest that physiological activation of the immune system by bacterial products with the release of pro-inflammatory cytokines leads to increases of depressed mood and anxiety, and decreases of verbal and non-verbal memory function. Moreover, large doses of cytokines, given as immunotherapy for cancer or hepatitis C commonly induce depression-like symptoms such as depressed mood, inability to experience pleasure, fatigue, poor concentration, and disordered sleep, which can be effectively treated by giving antidepressant medications. Cytokines can also alter sleep in humans. Expression of the Th2 or anti-inflammatory cytokine IL-10 prior to sleep predicts amounts of delta sleep during the nocturnal period, whereas levels of pro-inflammatory cytokine IL-6 are associated with declines of delta sleep and increases of REM sleep. Likewise, pro-inflammatory cytokine activation is implicated in daytime fatigue, with links identified in cancer survivors as well as healthy older adults.

Clinical Implications of Psychoneuroimmunology

The factors that account for individual differences in the rate and severity of disease progression are not fully understood, although increasing evidence suggests that behavioral and multisystem physiological changes that occur during depression or stress come together to exacerbate the course of many chronic diseases. In the following sections, several pertinent diseases examples are discussed.

Cardiovascular Disease

Atherosclerosis is now thought to be an inflammatory process that involves a series of steps, each of which appears to be impacted by depression. Activated macrophages within the vascular space secrete

pro-inflammatory cytokines, which in turn leads to expression of adhesion molecules. With recruitment of immune cells to the vascular cell wall or endothelium and the release of inflammatory cytokines, the vascular endothelium expresses adhesion molecules that facilitate further binding of immune cells. Importantly, psychological and physical stressors increase both release of pro-inflammatory cytokines and expression of adhesion molecules that tether (slow down) and bind immune cells to the vascular endothelium. Moreover, it appears that depression is associated with activation of the endothelium. Acute coronary patients who are depressed show an increased expression of an adhesion molecule that is released following activation of the vascular endothelium (i.e., soluble intracellular adhesion molecule). Importantly, this molecular marker of endothelial activation, as well as IL-6, predicts risk of future myocardial infarction, independent of cholesterol levels, smoking status, and obesity. In prospective studies, both depressed mood and inflammatory markers contribute independently to the risk for coronary heart disease, particularly in men.

HIV

HIV infection shows a highly variable course, and depression, bereavement, and maladaptive coping responses to stress (including the stress of HIV infection itself) have all been shown to predict the rate of immune system decay in HIV patients. For example, major depression in women with HIV infection was associated with lower NK activity, as well as increases in the numbers of activated CD8 lymphocytes and viral load, which suggested that declines of killer lymphocytes in association with depression may increase risk of HIV disease progression in women. Immune system decline and HIV replication are particularly rapid in patients living under chronic stress (e.g., gay men who conceal their homosexuality by living in the closet) and in patients with high levels of SNS activity (e.g., socially inhibited introverts). Tissue culture studies have shown that SNS neurotransmitters and glucocorticoids can accelerate HIV replication by rendering T lymphocytes more vulnerable to infection and by suppressing production of the antiviral cytokines that help cells limit viral replication.

Depression and Rheumatoid Arthritis: Neuroimmune Mechanisms

In a negative feedback loop, pro-inflammatory cytokines stimulate the HPA axis that results in the secretion of glucocorticoids, which in turn suppresses the immune response. However, in autoimmune

disorders such as rheumatoid arthritis, it is thought that the counterregulatory glucocorticoid response is not fully achieved. In animals that are susceptible to arthritis, there is a central hypothalamic defect in the biosynthesis of CRH, blunted induction of ACTH and adrenal steroids, and decreased adrenal steroid receptor activation in immune target tissues that together contribute to weak HPA response, one that is not sufficient to suppress the progression of an autoimmune response. Rheumatoid arthritis patients also show a relative hypofunctioning of the HPA axis despite the degree of inflammation. Stress and depression can lead to HPA axis activation and to increases of pro-inflammatory cytokines, and recent data suggest that stressful events, particularly those of an interpersonal nature, provoke symptoms of disease such as greater pain and functional limitations. Moreover, the presence of depression in rheumatoid arthritis patients undergoing stress is associated with exaggerated increases of IL-6, a biomarker predictive of disease progression. Conversely, administration of a psychological intervention that decreases emotional distress produced improvements in clinician-rated disease activity in rheumatoid arthritis patients, although immunological mediators were not measured. Likewise, in the case of another autoimmune disorder, psoriasis, a stress reduction intervention, mindfulness meditation, was found to induce a more rapid clearing of the psoriatic lesions.

Cancer

Experimental studies conducted in animal models have shown that exposure to acute stress leads to decreases in NK cell function and facilitates the metastatic spread of NK-sensitive tumors. However, translation of these data has been challenging; clinical findings supporting a link between depression and cancer have been mixed, with limited evidence to support the contribution of immune mechanisms. Immunogenic tumors have rarely been investigated in the context of depression and psychoneuroimmunology. In addition, other physiological systems, such as the endocrine system, may also play a role; for example, dysregulated cortisol rhythm is associated with both reduced NK activity and increased mortality in metastatic breast cancer patients. Behavioral interventions in the setting of cancer recovery appear to impact disease outcomes such as recurrence and survival. In metastatic breast cancer patients, group psychotherapy led to improvements in mood and increased survival time, controlling for initial staging and medical care during the follow-up period. Among patients with malignant melanoma, group psychotherapy was

associated with decreases in distress, increases in active coping, and increases in NK cytotoxicity, as well as a higher rate of survival. Both baseline NK cytotoxicity and improvements in coping behavior were associated with disease outcomes in this study.

Acknowledgments

This work was supported in part by grants A13239, DA16541, MH55253, AG18367, T32-MH19925, M01-RR00865, M01 RR00827, General Clinical Research Centers Program, and the Cousins Center for Psychoneuroimmunology.

See Also the Following Articles

Depression and Manic-Depressive Illness; Depression Models; Immunity; Major Depressive Disorder.

Further Reading

- Ader, R., Felten, D. and Cohen, N. (eds.) (2001). *Psychoneuroimmunology*. San Diego, CA: Academic Press.
- Capuron, L. and Dantzer, R. (2003). Cytokines and depression: the need for a new paradigm. *Brain, Behavior, and Immunity* 17, S119–S124.
- Empana, J. P., Sykes, D. H., Luc, G., et al. (2005). Contributions of depressive mood and circulating inflammatory markers to coronary heart disease in healthy European men: the Prospective Epidemiological Study of Myocardial Infarction (PRIME). *Circulation* 111, 2299–2305.
- Evans, D. L., Ten Have, T. R., Douglas, S. D., et al. (2002). Association of depression with viral load, CD8 T lymphocytes, and natural killer cells in women with HIV infection. *American Journal of Psychiatry* 159, 1752–1759.
- Friedman, E. M. and Irwin, M. (1997). Modulation of immune cell function by the autonomic nervous system. *Pharmacology Therapy* 74, 27–38.
- Irwin, M. (2002). Psychoneuroimmunology of depression: clinical implications (Presidential Address). *Brain, Behavior, and Immunity* 16, 1–16.
- Irwin, M., Pike, J. and Oxman, M. (2004). Shingles immunity and health functioning in the elderly. *Evidence Based Complementary and Alternative Medicine* 1, 223–232.
- Jung, W. and Irwin, M. (1999). Reduction of natural killer cytotoxic activity in major depression: Interaction between depression and cigarette smoking. *Psychosomatic Medicine* 61, 263–270.
- Kronfol, Z. and Remick, D. G. (2000). Cytokines and the brain: implications for clinical psychiatry. *American Journal of Psychiatry* 157, 683–694.
- Lesperance, F., Frasere-Smith, N., Theroux, P. and Irwin, M. (2004). The association between major depression and levels of soluble intercellular adhesion molecule 1,

- interleukin-6, and C-reactive protein in patients with recent acute coronary syndromes. *American Journal of Psychiatry* 161, 271–277.
- Miller, G. E., Cohen, S. and Herbert, T. B. (1999). Pathways linking major depression and immunity in ambulatory female patients. *Psychosomatic Medicine* 61, 850–860.
- Miller, G. E., Freedland, K. E., Duntley, S. and Carney, R. M. (2005). Relation of depressive symptoms to C-reactive protein and pathogen burden (cytomegalovirus, herpes simplex virus, Epstein-Barr virus) in patients with earlier acute coronary syndromes. *American Journal of Cardiology* 95, 317–321.
- Motivala, S. J., Sarfatti, A., Olmos, L. and Irwin, M. R. (2005). Inflammatory markers and sleep disturbance in major depression. *Psychosomatic Medicine* 67, 187–194.
- Raison, C. and Miller, A. H. (2003). When not enough is too much: the role of insufficient glucocorticoid signaling in the pathophysiology of stress-related disorders. *American Journal of Psychiatry* 160, 1554–1565.
- Schleifer, S. J., Keller, S. E. and Bartlett, J. A. (1999). Depression and immunity: clinical factors and therapeutic course. *Psychiatry Research* 85, 63–69.
- Segerstrom, S. C. and Miller, G. E. (2004). Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry. *Psychological Bulletin* 130, 601–630.
- Vedhara, K. and Irwin, M. R. (2005). *Human psychoneuroimmunology*. Oxford, UK: Oxford University Press.
- Zautra, A. J., Yocum, D. C., Villanueva, I., et al. (2004). Immune activation and depression in women with rheumatoid arthritis. *Journal of Rheumatology* 31, 457–463.
- Zorrilla, E. P., Luborsky, L. and McKay, J. R., et al. (2001). The relationship of depression and stressors to immunological assays: a meta-analytic review. *Brain, Behavior, and Immunity* 15, 199–226.

Depth, Effects of See: Pressures, Effects of Extreme High and Low.

Dermatological Conditions

M A Gupta

University of Western Ontario and Mediprobe Research Inc., London, Canada

© 2007 Elsevier Inc. All rights reserved.

Overview of Stress and Dermatology

Role of Stress over the Life Cycle

The Role of Stress in a Wide Range of Dermatological Symptoms

Stress and Some Specific Dermatological Disorders

Glossary

Alopecia areata An immunologically mediated disorder that results in hair loss; it may present in three degrees of severity: (1) single or multiple patches of well-demarcated hair loss, often on the scalp; (2) total or near total loss of scalp hair (termed alopecia totalis); and (3) generalized loss of body hair (termed alopecia universalis). Alopecia areata may occur at any age; the

Atopic dermatitis

peak incidence is during the third to fifth decades.

A chronically relapsing disorder, marked by periods of exacerbations and remission, characterized by a rash manifesting from patches of red irritated skin and raised papules to areas where the skin is thickened, with accentuation of the skin markings, usually due to repeated scratching. In any one individual, the skin lesions can vary; itching is a central feature of the disorder and some patients may present with dry scaly skin. It appears in early infancy, childhood, and adolescence, frequently associated with elevated serum immunoglobulin E levels and a family history of atopic dermatitis, allergic rhinitis, and asthma; also called atopic eczema.

Hyperhidrosis

An increased, above normal sweat production.

Pruritus

Itching; this symptom is the predominant feature of inflammatory skin diseases; however, it may be a symptom of a wide range of disorders, from dry skin

	to an occult cancer. Repeated scratching and rubbing can lead to thickening of the skin with accentuated skin markings and a clinical entity known as lichen simplex chronicus.
<i>Psoriasis</i>	A chronic relapsing disorder characterized by skin lesions of variable appearance. The skin lesions often present as raised circular plaques with redness and scales on the elbows, knees, and the scalp behind the ears; however, the lesions may be generalized and affect the entire body or may be localized to just the palms and soles. Pitting of the nails is often associated with psoriasis. Itching and the shedding of the skin scales are some of the most bothersome symptoms.
<i>Rosacea</i>	A chronic disorder of the face, often the nose, characterized by redness and superficial blood vessels, with or without papules. It is a disorder of adults, usually between ages 30 and 50 years. The onset of rosacea is insidious.
<i>Urticaria</i>	Hives that present as circumscribed, typically itchy, red evanescent areas of swelling. The individual lesions arrive suddenly and rarely persist for longer than 48 h; however, they may continue to recur for indefinite periods. Immunological factors such as allergies play a central role in acute urticaria; however, in the majority of cases of chronic urticaria (i.e., when the lesions are present beyond 6–8 weeks) the underlying cause is not determined.

Overview of Stress and Dermatology

Psychosocial stress has long been recognized as being associated with the onset and/or exacerbation of a wide range of dermatological disorders. Most of the literature on stress and skin disease refers to atopic dermatitis or atopic eczema and psoriasis. The term stress has been used to address three major areas in dermatology: (1) major stressful life events such as the death of a loved one, divorce, or loss of livelihood; (2) stress resulting from the social stigma associated with a cosmetically disfiguring skin disorder such as psoriasis or interpersonal stress resulting from having to provide care to someone with a chronic and recurring disorder such as childhood eczema, which in turn may have an adverse impact on the course of a stress-reactive skin disorder; and (3) traumatic life events that represent situations that overwhelm the coping capacity of the individual – unlike commonly experienced situations (such as simple bereavement),

traumatic events such as maternal neglect, severe emotional abuse, and sexual abuse generally are events that are outside the range of normal human experience and are significantly distressing to almost everyone.

Stress may influence dermatological symptoms in three major situations: (1) in primary stress-reactive skin disorders, such as psoriasis, atopic dermatitis or eczema, urticaria, acne, and alopecia areata, in which neuroendocrine and psychoneuroimmunological factors are believed to play an important role; (2) in disorders that represent an accentuated physiological response such as hyperhidrosis (or excessive perspiration) and blushing; and (3) in skin disorders, such as dermatitis artefacta, neurotic excoriations, or trichotillomania, that are secondary to an underlying psychiatric condition such as obsessive-compulsive disorder, posttraumatic stress disorder, or major depressive disorder. Nutritional deficiencies encountered in eating disorders such as anorexia nervosa and sometimes bulimia nervosa can result in diffuse hair loss or alopecia. Stress exacerbates the dermatological symptoms by aggravating the underlying psychiatric disorder.

Various neurobiological factors have been implicated in stress-mediated skin disorders. Stress-induced analgesia is observed in animals faced with inescapable stressors. Fear activates the secretion of endogenous opioid peptides, and this effect is blocked by naloxone in animal models. This may be a factor in itching, scratching and self-injury that occur during extremely stressful situations.

Psychosocial stress has been associated with the activation of the hypothalamic-pituitary-adrenocortical (HPA) axis and the sympathetic and adrenomedullary (SAM) system. Stressful emotional experiences disturb the regulation of the HPA and the SAM systems; that is, in the face of stress, physiological systems may operate at higher levels, resulting in higher glucocorticoid levels, or lower levels than normal. Recent studies have shown that psychological stress, such as examination-related stress, is associated with derangements in the epidermal (most superficial layer of the skin) permeability barrier function and that the alterations were proportional to the severity of the stressor. It has been proposed that this is mediated by increased endogenous glucocorticoids. Both psoriasis and atopic dermatitis demonstrate increased transepidermal water loss and deterioration of barrier function. Higher anxiety levels in acne patients have been associated with high blood catecholamine levels, which decreased with the treatment of the acne, suggesting that the anxiety in acne is associated with significant physiological stress for the patient and activation of the SAM system.

It has further been observed that there is increased responsiveness of the HPA axis in response to a heel prick stressor in newborns with a family history of atopy or elevated levels of cord blood immunoglobulin E (IgE). Patients with atopic dermatitis also show elevated eosinophil counts and elevated IgE expression in response to stress. Atopic patients with serum IgE levels greater than 100 IU ml^{-1} demonstrated significantly higher levels of excitability and less adequate coping with stress than did patients with lower IgE levels. IgE is considered to be important in atopic dermatitis because it mediates hypersensitivity reactions by stimulating mast cells and basophils.

Chronic stress may induce a state of hyporesponsiveness of the HPA axis whereby glucocorticoid secretion is decreased, leading to an increased secretion of mediators of inflammation such as the cytokines, which are normally regulated by cortisol. Animal data show that a blunted responsiveness of the HPA axis to stress is closely linked to an increased susceptibility to inflammatory disease. Stress-induced hyporesponsiveness of the HPA system has also been proposed to be a factor in some chronic inflammatory dermatoses such as atopic dermatitis.

Role of Stress over the Life Cycle

The interface between psychosocial stress and dermatological disorders begins early in the life cycle and persists throughout adult life because of the role of the skin as a vital organ of communication. The role of stress in dermatological disorders therefore has to be assessed from a developmental perspective. The skin plays an integral role as an organ of communication right from birth because the skin is the primary organ of attachment. The newborn infant's initial physical experience is largely tactile, and the child requires secure holding and hugging to develop physically, neurologically and psychosocially. Psychosocial stressors, such as maternal deprivation, neglect, or abuse during early life and the lack of adequate tactile nurturance in institutions such as orphanages, can impact a child's entire body, and the sequelae of such types of stressors are often manifested in a wide range of dermatological symptoms. Spitz observed that the lack of a nurturing mother-child relationship in an institutional setting was associated with the development of certain syndromes such as rocking, fecal play, and infantile eczema. In this institutionalized population, 15% of the infants had infantile eczema compared with a 2–3% prevalence in the general population. A child's need to be held and physically nurtured may be neglected in cases in which he or she has a severe skin disease; alternatively a child's chronic recurring skin disorder

may place inordinate stress on the caregiver and affect his or her capacity to nurture the child. In some instances, this interpersonal stress may in turn have an adverse impact on stress-reactive skin conditions such as eczema in a vicious cycle.

The onset of a cosmetically disfiguring skin disorder during adolescence often renders the patient very vulnerable to stress because he or she is also dealing with emerging hormonal and physical changes of puberty, body-image issues, and dealing with other developmental tasks of adolescence such as peer pressure, dating, and career choice. Adolescence is also associated with a high incidence of mood disorders such as major depressive disorder. The peak incidence of acne is during adolescence. In addition to the rise in the level of androgens that is the primary trigger for acne, the psychosocial stresses of this life stage can also contribute to flare-ups of the acne. Therefore, psychosocial stress may contribute to flare-up of acne and the acne in turn may result in significant stress for the patient. It is important to recognize that the adolescent patient may be especially vulnerable and may react to even relatively minor acne with a serious psychiatric disorder such as major depression and suicidal ideation. Alternatively, there have been incidences of violent acting out by adolescents who were having difficulty coping with the psychosocial, particularly peer-related, stresses associated with disfiguring acne. In young adulthood and later life, the emergence of stress-reactive disorders such as psoriasis may have a significant impact on the patient's social and occupational functioning, and the stress resulting from the stigmatization may in turn cause flare-ups of the psoriasis.

The Role of Stress in a Wide Range of Dermatological Symptoms

Some general stress-related factors may play a role in a wide range of skin disorders.

Pruritus or Itching

Pruritus is a feature of a wide range of skin disorders to varying degrees, including the common disorders that are exacerbated by psychological stress such as psoriasis, atopic dermatitis, and urticaria. Pruritus has been reported to be exacerbated by stress in up to 86% of cases in large surveys. Stress may also increase scratching behavior, which in turn can trigger the itch-scratch cycle and, in some instances, cause further flare-ups of the underlying skin condition, such as atopic dermatitis. Pruritus may further be a source of significant distress for the patient; it is often rated as the most bothersome feature of the skin disorder and has been associated with suicide.

Stress-induced scratching and itching may lead to the development of conditions such as lichen simplex chronicus, in which the superficial layer of the skin thickens as a result of repetitive scratching.

Disease-Related Stress

The impact of a skin disorder on quality of life, especially due to the cosmetic disfigurement and social stigma that are typically associated with a wide range of skin disorders, can be a source of significant stress for the patient. In some chronic stress-reactive dermatoses such as psoriasis and atopic dermatitis, this disease-related stress may in turn cause the underlying skin condition to flare up and may become an important confounding factor.

Stress-Related Self-Injury/Manipulation of the Skin

When faced with stressful situations, patients may pick at minor irregularities or existing lesions in their skin, scratch themselves excessively, or pluck their hair. Injury to the skin may exacerbate a primary skin condition such as psoriasis secondary to the Koebner phenomenon or may cause flare-ups of acne as a result of the underlying inflammatory process.

The skin is frequently the focus of tension-reducing behaviors, both because of its easy access and because of the primary role of the skin in early attachment. Attachment-related trauma occurs when there is abuse or neglect in early life, and this is often associated with self-injury to the body, especially the skin. Posttraumatic stress disorder (PTSD) associated with a history of sexual, physical, and/or emotional abuse results in a dysregulation of internal emotional states. The PTSD patients may excessively manipulate their skin or hair in an attempt to regulate their affect. The tension-reducing behavior and self-injury may manifest as the self-induced dermatological conditions such as neurotic excoriations, dermatitis artefacta, neurodermatitis, and trichotillomania. In cases in which high levels of stress and dissociation are present, patients may not even recall self-inducing their lesions as a result of dissociative amnesia and may be mistaken as malingers.

Stress and Some Specific Dermatological Disorders

Acne

Emotionally stressful events such as examinations are known to exacerbate acne; furthermore, increasing stress has been shown to be correlated with an increase in acne severity. The stress resulting from the impact of acne on the quality of life of the patient

is comparable to chronic disorders such as diabetes and asthma, and it often does not correlate with the clinical severity of acne – even mild acne may be associated with significant stress.

Acne excoriee is a condition that results from repetitive picking and excoriation of the acne lesions. A stress-related syndrome that is often underrecognized in acne excoriee is PTSD, usually secondary to abuse and/or neglect during early life; patients with PTSD often injure themselves, and this can manifest as recurrent acne excoriee. In some of these patients, the problems caused by the recurrent self-excoriation are more problematic than the primary acne lesions that were excoriated. Some PTSD patients dissociate when excoriating themselves and may not have full recollection of the times when they excoriated their acne lesions.

Alopecia Areata

Psychosocial stress has been reported to play a role in the onset and exacerbation of alopecia areata; however, this association is less robust. Various studies indicate a lack of a direct relationship between stress and the severity of alopecia areata, suggesting stress may be a factor that just triggers the onset of the disease process. One study reported mental stress within 3 months before the onset of alopecia areata by 23% of 114 patients surveyed, and an additional 22% reported suffering from stress that was not related to the appearance of the alopecia. In one study of 52 patients, all patients reported “a very unhappy or stressful life,” whereas in another study only 6.7% of 178 patients reported a severely disturbing event 6 months before the onset of symptoms. One study reported that patients whose alopecia areata is stress-reactive may suffer from depressive illness.

Atopic Dermatitis or Atopic Eczema

Stressful life events may precede the onset of atopic dermatitis or eczema in up to 70% of cases. Itching is a central feature of atopic dermatitis. Disease-related stress, interpersonal stress, and stress resulting from the family environment are important predictors of symptom severity in atopic dermatitis. Atopic patients have been shown to scratch more readily in response to an itch stimulus, suggesting that they develop a conditioned scratch response more readily than controls. Perceived itch may be enhanced in atopic patients in response to mental stress. The scratching can trigger the itch-scratch cycle, which can become a problem in cases of nocturnal itching and scratching. Some earlier studies reported high psychophysiological reactivity in atopic patients; however, this has not been supported by more recent studies that

carried out psychophysiological measures such as blood pressure, heart rate, and skin conductance.

Hyperhidrosis or Excessive Perspiration

Hyperhidrosis has been associated with stress in up to 100% of cases. Over 25% of patients with social phobia or social anxiety disorder, in which the patients experience heightened stress and physiological arousal in social situations, report hyperhidrosis to be a significantly distressing symptom.

Psoriasis

Various studies have reported an association between stress and psoriasis. A study that retrospectively examined the role of stress in the onset of a wide range of skin conditions reported that patients with psoriasis were more likely to report that a stressful experience predated the onset and exacerbation of their condition than patients with other skin diseases. In a study of 132 psoriasis patients, 39% recalled a significant stressful life event (e.g., interpersonal stress within the family, death or hospitalization of close relatives, accidents, sexual assault, and examinations) within 1 month before the first episode of psoriasis, in contrast to 10% of patients with conditions that are generally not associated with stress. A stressful life event was reported 1 month before the onset of psoriasis in 72% of 179 patients with psoriasis. In a survey of 245 children with psoriasis, stress was observed to be a provocative factor among 90% of the children. The severity of the stressor does not typically correlate with the time to onset or the exacerbation of psoriasis, possibly because of considerable individual variation in coping skills and the importance of the emotional meaning rather than the absolute intensity of the life event. PTSD has been associated with a long-term impact on physical health and a wide range of disorders affecting the immune system, including psoriasis. A recent study of 2490 Vietnam war veterans reported an adjusted odds ratio of 4.7 (95%, CI 1.9–11.7) for psoriasis among veterans with comorbid PTSD. The autonomic dysregulation associated with PTSD may be an underlying perpetuating factor in patients with recurring or treatment-resistant psoriasis.

Rosacea

Factors such as overwhelming life stress, an anxious and immature personality associated with excessive feelings of guilt and shame, and social anxiety secondary to easy blushing have all been implicated in rosacea. Psychological stress has been

implicated as a factor among 20–94% of patients with rosacea.

Urticaria

Increased mental tension and fatigue have been spontaneously reported as the main exacerbating or precipitating factors by 77% of 43 patients. Out of 100 patients with chronic urticaria and/or angioedema, 51% reported that stressful life situations were associated with the onset of symptoms, compared with 8% of surgical controls. Earthquakes and other catastrophic life events have been associated with the onset of urticaria. There have been case reports of patients with PTSD secondary to severe physical abuse, who developed linear urticarial lesions in the same region of the body where they had experienced the beatings in their early life.

See Also the Following Article

Psychosomatic Medicine.

Further Reading

- Boscarino, J. A. (2004). Posttraumatic stress disorder and physical illness. Results from clinical and epidemiologic studies. *Annals of the New York Academy of Sciences* 1032, 141–153.
- Buske-Kirschbaum, A. and Hellhammer, D. H. (2003). Endocrine and immune responses to stress in chronic inflammatory skin disorders. *Annals of the New York Academy of Sciences* 992, 231–240.
- Garg, A., Chren, M., Sands, L. P., et al. (2001). Psychological stress perturbs epidermal permeability barrier homeostasis: implications for the pathogenesis of stress-associated skin disorders. *Archives of Dermatology* 137, 53–59.
- Gavenda, A. (ed.) (2005). Psychocutaneous disease. *Dermatologic Clinics* 23, 4.
- Griesemer, R. D. (1978). Emotionally triggered disease in a dermatological practice. *Psychiatric Annals* 8, 49–56.
- Gupta, M. A. and Gupta, A. K. (1996). Psychodermatology: an update. *Journal of the American Academy of Dermatology* 34, 1030–1046.
- Gupta, M. A. and Gupta, A. K. (2003). Psychiatric and psychological co-morbidity in patients with dermatologic disorders. *American Journal of Clinical Dermatology* 4, 833–842.
- Koblenzer, C. S. (1987). *Psychocutaneous disease*. Orlando, FL: Grune & Stratton.
- Koo, J. Y. M. and Lee, C. S. (eds.) (2003). *Psychocutaneous medicine*. New York: Marcel Dekker Inc.
- Panconesi, E. (ed.) (1984). Stress and skin diseases: psychosomatic dermatology. *Clinics in Dermatology* 2(4).
- Spitz, R. A. (1965). *The first year of life: a psychoanalytic study of normal and deviant development of object relations*. New York: International Universities Press.

Desensitization

F A Antoni

Centre of Integrative Physiology, University of Edinburgh, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by F A Antoni, volume 1, pp 682–683, © 2000, Elsevier Inc.

Desensitization of the Stress Response
Cellular Mechanisms of Desensitization

Glossary

<i>Heterotrimeric G proteins</i>	A group of proteins that consist of three subunits (α , β , γ) and couple cell surface receptors to their effector enzymes. On activation of cell surface receptors by their ligands, heterotrimeric G proteins dissociate into the α subunit and the β - γ complex, both of which regulate the activity of effector enzymes in the cell membrane.
<i>Neuroendocrine motoneurons</i>	Hypothalamic nerve cells that produce and secrete neurohormones into the hypophyseal portal circulation and thus regulate the synthesis and release of anterior pituitary hormones.
<i>Protein kinases</i>	Enzymes that tag proteins with phosphoryl residues derived from ATP. The activity of protein kinases may be regulated by a variety of intracellular messengers, protein-protein interactions, or phosphorylation by other protein kinases.

Desensitization is a widespread phenomenon in biological systems whereby the response to persistent or invariant stimuli is reduced.

Desensitization of the Stress Response

As if to react selectively to new or stronger stimuli, biological systems diminish their responses to persistent or stable stimuli in a process termed desensitization, adaptation, or habituation. The stress response is no exception to this rule. A reduction in the activation of the hypothalamic-pituitary-adrenocortical (HPA) axis and behavioral arousal elicited by the same and repeatedly applied stressor occurs in most cases. This is called homotypic desensitization. The degree and speed of desensitization may depend on the frequency and the length of exposure to the

stressor as well as on the genetic make-up of the experimental subjects. Heterotypic desensitization, a blunting of the response to several types of stress induced by repeated exposure to a single stressor, is less common but has been observed. More frequently, the application of a heterotypic stressor in subjects that have undergone homotypic desensitization elicits an exaggerated HPA response to the heterotypic stimulus when compared with nondesensitized controls. This indicates that homotypic desensitization affects the stressor-specific mechanisms of the HPA response as well as the final common pathways of neuroendocrine regulation. Indeed, chronic stressors are known to increase the expression of vasopressin dramatically in the neuroendocrine motoneurons of the hypothalamic paraventricular nucleus, whereas the expression of corticotropin releasing hormone (CRH) remains essentially unchanged. This is likely to augment the pituitary response to stimuli that impinge on the neuroendocrine motoneurons and may well explain the enhanced pituitary adrenocorticotropin response to heterotypic stressors in subjects with homotypic desensitization.

A general attenuation of the HPA response to several stressors occurs in lactating animals. This has been attributed to the suppression of the HPA axis at the level of the paraventricular nucleus by oxytocin. By contrast, stressor-specific habituation does not take place at the level of the paraventricular neuroendocrine motoneurons.

Cellular Mechanisms of Desensitization

At the cellular level, desensitization manifests itself in biological processes as diverse as bacterial chemotaxis and mammalian neurotransmission. Desensitization affects signal-transduction pathways by which cells detect and decode changes in the extracellular milieu. The molecular and cellular details of desensitization are particularly well explored with respect to seven-transmembrane domain receptors (7-TMRs) coupled to heterotrimeric G-proteins. Significantly, plasma membrane receptors for CRH, vasopressin, and catecholamines are in this category.

On binding their respective agonist ligands, receptors activate heterotrimeric G-proteins, which leads to the stimulation of various intracellular effector enzymes that generate intracellular second-messenger molecules. Within a few seconds, cells can diminish or virtually eliminate their agonist-evoked responses through a process that involves the phosphorylation

of the 7-TMRs on one or more intracellular domains. The cellular response to a given agonist may be desensitized by cellular exposure to that agonist itself in a process called homologous desensitization. Alternatively, the desensitization of the response to a given agonist may be caused by agonists for distinct receptor signaling systems in a process termed heterologous desensitization. Potentially affecting multiple receptor systems, heterologous desensitization involves the phosphorylation of 7-TMRs by second-messenger-dependent kinases, such as cAMP-dependent protein kinase and protein kinase C (PKC). In most cells, a further elaborate system exists to trap activated 7-TMRs – the G-protein-coupled receptor kinases and arrestins. These are distinct and independent of second-messenger-stimulated kinases; further, they are specific for activated 7-TMRs and are involved in homologous desensitization.

The common feature of all of these mechanisms is the phosphorylation of 7-TMRs and their consequent uncoupling from G-proteins. Interestingly, uncoupling from G-proteins in many cases leads to the activation of alternative signaling pathways for these receptors involving arrestins as the coupling proteins. Thus, although one aspect of receptor action (e.g., the generation of cAMP) may be desensitized, signaling through another pathway (e.g., the stimulation of microtubule-activated protein kinase) may be turned on. The receptors tagged by phosphorylation may be endocytosed and recycled on dephosphorylation or eventually degraded, depending on the requirements of the cell.

Even longer-term control of signal transduction pathways may involve the alteration of the transcription of receptor genes or their effector systems,

potentially entirely reprofiling the cellular response as a result of prolonged exposure to receptor ligands.

See Also the Following Articles

Infection; Leukocyte Trafficking and Stress; Natural Killer (NK) Cells.

Further Reading

- Aguilera, G. (1994). Regulation of pituitary ACTH-secretion during chronic stress. *Frontiers in Neuroendocrinology* **15**, 321–350.
- Antoni, F. A. (1993). Vasopressinergic control of anterior pituitary adrenocorticotropin secretion comes of age. *Frontiers in Neuroendocrinology* **14**, 76–122.
- Dhabhar, F. S., McEwen, B. S. and Spencer, R. L. (1997). Adaptation to prolonged or repeated stress – comparison between rat strains showing intrinsic differences in reactivity to acute stress. *Neuroendocrinology* **65**, 360–368.
- Flugge, G. (1995). Dynamics of central nervous 5-HT_{1A}-receptors under psychosocial stress. *Journal of Neuroscience* **15**, 7132–7140.
- Hall, R. A., Premont, R. T. and Lefkowitz, R. J. (1999). Heptahelical receptor signaling: beyond the G protein paradigm. *Journal of Cell Biology* **145**, 927–932.
- Ma, X. M. and Lightman, S. L. (1998). The arginine vasopressin and corticotropin-releasing hormone gene transcription responses to varied frequencies of repeated stress in rats. *Journal of Physiology* (London) **510**, 605–614.
- Pitcher, J. A., Freedman, N. J. and Lefkowitz, R. J. (1998). G protein-coupled receptor kinases. *Annual Review of Biochemistry* **67**, 653–692.
- Wang, Q., Zhao, J., Brady, A. E., et al. (2004). Spinophilin blocks arrestin actions *in vitro* and *in vivo* at G protein-coupled receptors. *Science* **304**, 940–1944.

Dexamethasone Suppression Test (DST)

R T Rubin and B J Carroll

VA Greater Los Angeles Healthcare System,
Los Angeles, CA, and Pacific Behavioral Research
Foundation, Carmel, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition
article by R T Rubin, volume 1, pp 684–687,
© 2000, Elsevier Inc.

Introduction

Increased Central Nervous System Driving of the
Hypothalamic-Pituitary-Adrenal Axis
Cushing's Syndrome

Glossary

<i>Adrenal cortex</i>	The outer layer of the adrenal glands, which secretes several steroid hormones, including glucocorticoids, mineralocorticoids, and adrenal androgens. The adrenal glands lie just above the kidneys.
<i>Adrenal androgens</i>	Hormones produced by the adrenal cortex that have weak male sex hormone-like effects.
<i>Aldosterone</i>	The principal mineralocorticoid produced by the adrenal cortex in humans.
<i>Amygdala</i>	A group of nerve cells in the temporal lobes of the brain that stimulates the secretion of corticotropin-releasing hormone and, in turn, the rest of the hypothalamic-pituitary-adrenal cortical axis.
<i>Corticotropin (adrenocorticotrophic hormone, ACTH)</i>	A hormone produced by cells in the anterior pituitary gland that is carried by the bloodstream to the adrenal cortex, where it stimulates the secretion of glucocorticoids, mineralocorticoids, and adrenal androgens.
<i>Corticotropin-releasing hormone (CRH)</i>	A hormone produced by neuroendocrine cells of the hypothalamus that is transported down the pituitary stalk to the anterior pituitary gland, where it stimulates the secretion of ACTH.
<i>Cortisol (hydrocortisone)</i>	The principal glucocorticoid produced by the adrenal cortex in humans. Under normal circumstances, cortisol feeds back to the pituitary gland, hypothalamus, and other brain areas to reduce the secretion of CRH and ACTH, thereby reducing the secretion of cortisol and other adrenal cortical hormones.
<i>Dexamethasone</i>	A synthetic glucocorticoid having 25 times the potency of cortisol.
<i>Glucocorticoids</i>	Hormones produced by the adrenal cortex that increase glucose production in the liver, increase glucose utilization

by tissues, inhibit glycogen formation by body tissues, and promote lipid breakdown in fat tissue. The principal glucocorticoid in humans is cortisol. When administered in high therapeutic doses or when present in excess due to overproduction by the adrenal glands, glucocorticoids suppress immunological function, reduce inflammation, and decrease connective tissue and new bone formation.

A group of nerve cells in the temporal lobes of the brain that inhibits the secretion of CRH and, in turn, the rest of the hypothalamic-pituitary-adrenal cortical axis.

A hormone axis consisting of cells in the hypothalamus of the brain that secrete CRH and vasopressin (AVP), which stimulate the secretion of ACTH from the anterior pituitary gland into the bloodstream. In turn, ACTH stimulates the adrenal cortex to secrete glucocorticoids, mineralocorticoids, and adrenal androgens into the bloodstream.

The area at the base of the brain that controls vital body processes including the production of hormones that stimulate and inhibit the secretion of anterior pituitary hormones.

Hormones produced by the adrenal cortex that reduce the excretion of sodium and enhance the excretion of potassium and hydrogen ions by the kidney. The principal mineralocorticoid in humans is aldosterone.

A gland, connected to the base of the brain by the pituitary stalk, that secretes several hormones into the bloodstream that stimulate the adrenal cortex, the thyroid gland, the gonads (testes and ovaries), and other tissues of the body.

A hormone produced by cells in the hypothalamus that is transported down the pituitary stalk to (1) the anterior pituitary gland, where, along with CRH, it stimulates the secretion of ACTH, and (2) the posterior pituitary gland, from which it is carried by the bloodstream to the kidneys, where it reduces the excretion of water.

Hippocampus

Hypothalamic-pituitary-adrenal cortical (HPA) axis

Hypothalamus

Mineralocorticoids

Pituitary

Vasopressin (AVP; antidiuretic hormone, ADH)

Introduction

The HPA axis is regulated by two processes. One is the closed-loop negative feedback of cortisol to hormone receptors in the hippocampus, hypothalamus, and pituitary gland. This suppresses the secretion of CRH, ACTH, and cortisol itself. This process is

analogous to control of heat by a thermostat; at a certain temperature, the heat signals the thermostat to shut down the heater until the temperature drops, then the heater goes on again, and the cycle repeats itself. The second process is the open-loop driving of the HPA axis by the central nervous system (CNS). Areas of the brain, including the amygdala, hippocampus, and hypothalamus, stimulate and inhibit the HPA axis to different degrees depending on the time of day, season of the year, and physical and environmental stressors.

The HPA axis has a circadian (24-h) rhythm. The secretion of CRH, ACTH, and adrenal hormones is greatest at 7–8 a.m., an hour or so after awakening, and at its lowest about 2–3 a.m. The difference in blood concentrations of ACTH and adrenal glucocorticoids such as cortisol is approximately fourfold between the nadir (low point) and the peak several hours later. If a person shifts his or her sleep–wake cycle, it takes approximately 2 weeks for the HPA axis to resynchronize its circadian rhythm.

The HPA axis is very stress-responsive. Both physical and psychological stressors can cause increased activity of this hormone axis. In particular, novel stressors (stressors that are new experiences) cause HPA axis activation; with a person's repeated encountering of the stressor, there is less and less HPA axis response. This is particularly true of demanding tasks, which initially provoke an HPA axis response but do not after the person achieves mastery of the task through training and experience. This change is called neuroendocrine adaptation.

Pathological (abnormal) function of the HPA axis can occur from several causes. These include repeated uncontrollable environmental stressors; functional psychiatric illnesses such as major depression and schizophrenia; and disorders of the HPA axis itself, such as hormone-secreting tumors of the pituitary gland and adrenal cortex. The dexamethasone suppression test (DST) is one of several endocrine tests that can help diagnose the particular pathological process resulting in abnormal HPA axis activity.

Increased Central Nervous System Driving of the Hypothalamic-Pituitary-Adrenal Axis

As mentioned, repeated, uncontrollable environmental stressors, functional psychiatric illnesses such as major depression and schizophrenia, and other conditions can increase CNS stimulation of the HPA axis to such a degree that it is relatively insensitive to dexamethasone suppression. Major depression is the best-studied psychiatric illness in this regard; 30–50% of major depressives have mildly to moderately increased HPA axis activity, consisting of

increased blood concentrations of ACTH and cortisol at all times of the day and night, with preservation of the circadian rhythm of these hormones. Correlated with this increase is nonsuppression of ACTH and cortisol on the low-dose DST. In the low-dose DST, 1–2 mg of dexamethasone is given orally at 11–12 p.m., and blood samples are taken at 8, 16, and 24 h thereafter for measurement of circulating cortisol concentrations. Normally, cortisol is suppressed to low levels (less than 50 ng ml⁻¹) for a full 24 h. Depressed patients with increased HPA axis activity often show early cortisol (and ACTH) escape from dexamethasone suppression. That is, there may be early suppression of these hormones the morning (8 h) following dexamethasone administration, but hormone concentrations are prematurely increased in the 16- and/or 24-h blood samples. The rapidity and degree of cortisol escape in the DST is an indication of the strength of CNS driving of the HPA axis. It also may be a good indication of recovery from major depression.

Depressed patients who have an initially abnormal DST, who are treated with antidepressant medication, and who have both a good clinical response and a reversion of their DST to normal are more likely to remain in remission when their medication is discontinued. On the other hand, patients who have an apparently good clinical response but who persist with an abnormal DST are more likely to experience relapse, even when continuing antidepressant treatment. In some patients with recurrent depressive episodes, the DST may become abnormal again before clinical symptoms emerge, signaling that a relapse may be imminent and allowing the physician to reinstitute treatment at the earliest possible time. This finding of a neuroendocrine change preceding any clinical change has been well documented in patients with rapidly cycling bipolar disorder. They generally have normal DST suppression of cortisol when manic, but dexamethasone resistance appears up to a week before they switch to the next depressive phase. In patients with mixed bipolar episodes, when both manic and depressive features coexist, the DST is usually abnormal. Because of its change with clinical status, the DST is considered a state marker rather than a trait marker of major depression.

Major depression is a heterogeneous disorder that includes patients with a wide range of clinical profiles (see **Depression and Manic-Depressive Illness**). Abnormal DST results are seen most commonly in those with melancholic depression, psychotic depression, and mixed bipolar episodes. Advancing age and global severity, especially incapacitation, are other important correlates of abnormal DST results in depression. Abnormal DSTs have been reported in depressed prepubertal children and in depressed and bipolar adolescent inpatients. Abnormalities are seen least often in young patients with milder depressions,

especially when these are of short duration and are precipitated by adverse life events. In uncomplicated bereavement and adjustment disorder with depressed mood, the DST is normal.

When used in patients with a well-established diagnosis of major depression, an abnormal DST signifies that the patient is very likely to benefit from somatic treatment such as antidepressant medication and electroconvulsive therapy. In controlled clinical trials of antidepressant drugs in heterogeneous depressed populations, the rates of response to placebo and active drug when the DST is abnormal are approximately 10 and 70%, respectively. In patients with normal DST results, the corresponding response rates are 55 and 65%, respectively. Thus, the drug-attributable, specific response rates are 60% in patients with an abnormal DST and 10% in patients with a normal DST. When cognitive psychotherapy has been used to treat depressed inpatients who have abnormal DSTs and/or elevated urinary free cortisol excretion, without concomitant use of antidepressant drugs, only 44% responded, compared with 92% of patients whose HPA axis biomarkers were normal. Other, long-term follow-up studies have found that depressed patients with abnormal DST results, in comparison to those with normal test results, have an eightfold risk of switch from unipolar major depression to bipolar disorder, an eightfold risk of completed suicide, and an eightfold risk of rehospitalization. Severely depressed patients who require ECT have a very high rate of positive DST results.

When used as a diagnostic test, the DST has several problems. Even for melancholic depression, test sensitivity or true positive rate is only approximately 50%. This means that a negative test result does not rule out the diagnosis. The specificity of the test, or true negative rate, varies with the populations studied. When careful exclusions were applied for medical and drug confounds, specificity of 90% or greater was found in early reports. With wider use, however, lower specificity has been observed. The most common reasons in practice for false-positive DST results are certain drugs, especially anticonvulsant agents that accelerate the metabolism of dexamethasone; conditions marked by notable weight loss, such as anorexia nervosa or cachexia associated with cancer; Alzheimer disease, possibly due to loss of hippocampal inhibitory control on the HPA axis; and the marked physiological stress associated with acute, uncontrolled psychosis. Thus, in accordance with Bayesian principles, the predictive value of a positive DST as a confirmatory test is high when there is strong suspicion of major depressive illness in a given patient, as determined by clinical examination and other tests, but the predictive value is considerably reduced when the DST is incorrectly used as a screening test for major depression (that is, when

there is a low base rate of the illness in a given population).

A final problem, which is largely responsible for the present clinical disuse of the DST, is that it is now agreed that measurement of plasma dexamethasone concentrations is required for a valid interpretation of DST results. When low doses of dexamethasone (1–2 mg) are used, with sampling times up to 24 h later, the plasma dexamethasone concentrations achieved in normal subjects are just above the threshold needed to suppress the HPA axis. Any factor that accelerates the rate of clearance of dexamethasone from plasma therefore will be associated with false-positive results. Low plasma dexamethasone concentrations may be caused by genetic variability of drug clearance or by the factors previously listed that are associated with reduced test specificity. This becomes of particular concern when very low doses of dexamethasone (0.25–0.5 mg) have been used to attempt to enhance the sensitivity of the test.

It also should be mentioned that underactivity of the HPA axis, along with enhanced suppression of ACTH and cortisol by dexamethasone, has been noted in patients with posttraumatic stress disorder (PTSD), such as in holocaust survivors, combat veterans, and resulting from childhood sexual abuse. The reason for reduced, rather than increased, HPA axis activity in the face of repeated psychological reexperiencing of traumatic incident(s), as occurs in PTSD, has been hypothesized to be related to enhanced CNS glucocorticoid receptor feedback sensitivity. For example, patients with PTSD have been described as being supersensitive to very low dexamethasone doses.

Cushing's Syndrome

Cushing's syndrome is the group of clinical signs and symptoms resulting from increased circulating glucocorticoids of long duration. There are many clinical changes in Cushing's patients, including obesity of the face and trunk, weakness and atrophy of limb muscles, increased blood pressure, imbalance of glucose metabolism, and psychological changes. There are two main types of Cushing's syndrome, ACTH-dependent and ACTH-independent. ACTH-dependent Cushing's syndrome results from increased pituitary secretion of ACTH, usually from a pituitary tumor (Cushing's disease); inappropriate ACTH secretion by nonpituitary tumors, often in the lungs; and inappropriate CRH secretion by nonhypothalamic tumors, in turn stimulating excessive pituitary ACTH secretion. These conditions, all involving excess ACTH production, cause enlargement of the adrenal glands and excessive cortisol secretion. ACTH-independent Cushing's syndrome is caused by primary tumors or abnormalities of the adrenal cortex itself, resulting in excessive cortisol secretion and suppression of ACTH production by the

pituitary. Prolonged administration of glucocorticoids for the treatment of certain illnesses also may cause ACTH-independent Cushing's syndrome.

The DST is one of several endocrine tests used to diagnose the different causes of Cushing's syndrome. Indeed, the DST was originally developed to aid endocrinologists in the differential diagnosis of hypercortisolemic and Cushingoid states. In the differential diagnosis of Cushing's syndrome, dexamethasone is administered in different dosage strengths and for different periods of time in a series of low- and high-dose tests. The route of administration may be oral or intravenous. The relative resistance of ACTH and cortisol to dexamethasone suppression is noted, along with baseline circulating ACTH and cortisol concentrations and their response to stimulation by administered CRH. The excessive pituitary production of ACTH and adrenal production of cortisol in Cushing's disease are only partially suppressible by low-dose dexamethasone but are more suppressible by higher doses. In contrast, patients with nonpituitary sources of ACTH production will rarely show suppression of ACTH and cortisol by dexamethasone. Similarly, patients with ACTH-independent Cushing's syndrome will not show suppression of cortisol, even with high doses of dexamethasone, because their ACTH already is suppressed by the independently high circulating cortisol concentrations and dexamethasone does not act to suppress the adrenal cortex directly.

It should be emphasized that, in diagnosing the causes of Cushing's syndrome, the DST must be interpreted in the light of other endocrine findings. For example, one important distinction between the increased ACTH and cortisol production in some patients with major depression versus patients with Cushing's disease is that the former have preservation of their circadian rhythms of ACTH and cortisol, whereas the latter often have high circulating hormone concentrations throughout the entire 24 h. Major depressives also do not have the clinical changes seen in Cushing's syndrome, because their body tissues are not exposed to equivalent, continuously high circulating cortisol concentrations. The mean 24-h plasma cortisol concentration in Cushing's disease exceeds 250 ng ml⁻¹, whereas in psychotic depression it is around 100–120 ng ml⁻¹ and in normal subjects around 60–80 ng ml⁻¹. The differences in tissue exposure to free cortisol are even greater, because the cortisol-binding transport protein in plasma, transcortin, is reduced in Cushing disease but normal in depression. The DST therefore must be considered as just one in a series of clinical examinations and laboratory studies used to determine the causes of increased HPA axis activity.

Acknowledgment

This work is supported by NIH grant MH28380.

Further Reading

- Carroll, B. J. (1989). Diagnostic validity and laboratory studies: rules of the game. In: Robins, L. N. & Barrett, J. E. (eds.) *The validity of psychiatric diagnosis*, pp. 229–245. New York: Raven Press.
- Carroll, B. J., Feinberg, M., Greden, J. F., et al. (1981). A specific laboratory test for the diagnosis of melancholia: standardization, validation, and clinical utility. *Archives of General Psychiatry* 38, 15–22.
- De Kloet, E. R. (1997). Why dexamethasone poorly penetrates in brain. *Stress* 2, 13–20.
- Griffin, M. G., Resick, P. A. and Yehuda, R. (2005). Enhanced cortisol suppression following dexamethasone administration in domestic violence survivors. *American Journal of Psychiatry* 162, 1192–1199.
- Orth, D. N., Kovacs, W. J. and DeBold, C. R. (1998). The adrenal cortex. In: Wilson, J. D., Foster, D. W., Kronenberg, H. M. & Larsen, P. R. (eds.) *Williams textbook of endocrinology* (9th edn., pp. 517–664). Philadelphia: WB Saunders.
- O'Sullivan, B. T., Cutler, D. J., Hunt, G. E., et al. (1997). Pharmacokinetics of dexamethasone and its relationship to dexamethasone suppression test outcome in depressed patients and healthy control subjects. *Biological Psychiatry* 41, 574–584.
- Ritchie, J. C., Belkin, B. M., Krishnan, K. R. R., et al. (1990). Plasma dexamethasone concentrations and the dexamethasone suppression test. *Biological Psychiatry* 27, 159–173.
- Rubin, R. T. (1994). Neuroendocrine aspects of stress in major depression. In: Liberman, R. P. & Yager, J. (eds.) *Stress in psychiatric disorders*, pp. 37–52. New York: Springer.
- Rubin, R. T. and Poland, R. E. (1984). The dexamethasone suppression test in depression: advantages and limitations. In: Burrows, G. D., Norman, T. R. & Maguire, K. P. (eds.) *Biological psychiatry: recent studies*, pp. 76–83. London: John Libbey.
- Schimmer, B. P. and Parker, K. L. (2001). Adrenocorticotrophic hormone; adrenocortical steroids and their synthetic analogs; inhibitors of the synthesis and actions of steroid hormones. In: Hardman, J. G. & Limbird, L. E. (eds.) *Goodman & Gilman's the pharmacological basis of therapeutics* (10th edn., pp. 1649–1677). New York: McGraw-Hill.
- Rush, A. J., Giles, D. E., Schlessler, M. A., et al. (1996). The dexamethasone suppression test in patients with mood disorders. *Journal of Clinical Psychiatry* 57, 470–484.
- Stein, M. B., Yehuda, R., Koverola, C., et al. (1997). Enhanced dexamethasone suppression of plasma cortisol in adult women traumatized by childhood sexual abuse. *Biological Psychiatry* 42, 680–686.
- Thase, M. E., Dube, S., Bowler, K., et al. (1996). Hypothalamic-pituitary-adrenocortical activity and response to cognitive behavior therapy in unmedicated, hospitalized depressed patients. *American Journal of Psychiatry* 153, 886–891.
- Tran, H. A. and Petrovsky, N. (2005). Dexamethasone infusion testing in the diagnosis of Cushing's syndrome. *Endocrine Journal* 52, 103–109.

DEX-CRH Test*

N C Schommer and I Heuser

Charité – Campus Benjamin Franklin,
Berlin, Germany

© 2007 Elsevier Inc. All rights reserved.

Introduction

Procedures

Impact Factors

DEX/CRH Test in Depressive Disorders and Other
Psychiatric Diseases

Conclusion

Glossary

<i>Adrenal cortex</i>	Part of the adrenal gland secreting different steroid hormones such as glucocorticoids, mineralocorticoids, and androgens after stimulation with releasing hormones from the pituitary gland.
<i>Adrenocorticotropic hormone (ACTH)</i>	A hormone secreted from the anterior pituitary gland that induces glucocorticoid secretion from the adrenal cortex.
<i>Corticotropin-releasing hormone (CRH)</i>	A neuropeptide hormone mainly secreted by the paraventricular nucleus of the hypothalamus. One main effect is the induction of ACTH secretion at the level of the anterior pituitary gland. Additionally, CRH is well known to be involved in anxious and depressive mood and behavior.
<i>Dexamethasone (DEX)</i>	A synthetic glucocorticoid with a 25 times higher binding affinity to glucocorticoid receptors of the pituitary than cortisol itself. Since DEX is not able to cross the blood–brain barrier, the negative-feedback sensitivity of the HPA system can be assessed at the level of the pituitary gland.
<i>Glucocorticoid receptor (GR)</i>	This receptor type is widely distributed in the central nervous system, such as the hypothalamus, the limbic system, including the hippocampus, and the cortex. In concert with the hippocampal MR, it is mainly involved in the negative

Glucocorticoids

Hypothalamic-pituitary-adrenal (HPA) axis

Mineralocorticoid receptor (MR)

Vasopressin

feedback regulation in case of stimulated HPA-activity. During unstimulated conditions only 5–10 percent of central GRs are occupied by the ligand.

Steroid hormones, secreted by the adrenal cortex, with cortisol as the principle human glucocorticoid. Their main effects are induction of liver glucose production, inhibition of glucose metabolism in body tissues and lipid decomposition in fat tissue. In higher concentrations, glucocorticoids results in suppression of immunological function with anti-inflammatory effects. Cortisol, which readily crosses the blood–brain barrier serves as a negative-feedback signal on hypothalamic CRH and vasopressin as well as pituitary ACTH secretion.

A hormone system encompassing the hypothalamus that secretes the neuropeptides CRH and vasopressin and which in turn stimulates the anterior pituitary to release ACTH. This pro-opiomelanocortin-derived hormone then induces glucocorticoid, e.g., cortisol, release at the level of the adrenal cortex. The HPA system constitutes a long-loop, negative-feedback system, with cortisol and ACTH inhibiting further ACTH and CRH release via hippocampal glucocorticoid and mineralocorticoid receptor occupancy. Activation of the HPA system occurs mainly via afferents from the brain stem and cortical areas.

Centrally, this receptor type is only found in the CA-3 layer of the hippocampus. There it is mainly co-localized with glucocorticoid receptors and regulates HPA-system activity. Hippocampal MRs bind GRs with a 6–10-fold higher affinity than GRs and are almost completely occupied by the ligand under resting conditions, whereas the GR is only occupied at times of stress activation and/or the circadian peaks of the HPA system.

A neuropeptide hormone that regulates in concert with CRH HPA system drive. Vasopressin- and CRH-expressing neurons are to a large part co-localized in the paraventricular nucleus.

* This DEX/CRH test should be distinguished from the endocrine DEX/CRH test employed in the differential diagnosis of patients with Cushing syndrome from those with Pseudo-Cushing syndrome, i.e., obese patients with hypercortisolism due to psychological factors or chronic active alcoholism that cannot be differentiated otherwise. This test entails administration of a higher dose of dexamethasone (0.5 mg 4 times a day for 2 days prior to the administration of CRH). Presence of an ACTH and cortisol response to CRH denotes Cushing syndrome.

Introduction

To study the functioning of the hypothalamic-pituitary-adrenal (HPA) system in humans there are

four possible options: (1) assessing basal, circadian secretory profiles; (2) using psychological stressors, such as free speech in front of an audience; (3) using physical stressors, e.g.; treadmill exercise; or (4) applying pharmacological challenge tests. Psychological stressors activate the HPA system by stimulating the limbic and cortical areas (prefrontal cortex (PFC), hippocampus, amygdala), and the paraventricular nucleus (PVN) of the hypothalamus. A more direct pathway for the PVN to secrete CRH is reached by both physical stressors and pharmacological challenges. The dexamethasone (DEX)-suppression test, the CRH test, and the combined DEX/CRH test act at the level of the pituitary and indirectly, at the adrenal level and allow the assessment of feedback sensitivity of the HPA system.

A low dose of dexamethasone (1.5 mg), a synthetic glucocorticoid with a high binding affinity to glucocorticoid receptors of the pituitary, given in the late evening results in a suppression of ACTH- and subsequently cortisol release the following day in healthy, unstressed subjects. In depressed patients or stressed individuals, however, low-dose dexamethasone is insufficient to suppress ACTH and cortisol adequately, a phenomenon known as escape from DEX-suppression or DST-nonsuppression. By administering a physiological dose of CRH to healthy, unstressed humans, the pituitary responds with markedly increased ACTH secretion, whereas in depressed patients and in stressed individuals who are hypercortisolemic, the same test will result in a blunting of ACTH release and no measurable changes in prechallenge cortisol concentrations. The reason for this outcome is due to the fact that the higher basal cortisol levels in these subjects feedback negatively to the pituitary, putting a brake on ACTH release after CRH. When exposed to the combined DST plus CRH test (DEX/CRH test) depressed patients frequently show excessive, and not blunted, hormonal (ACTH and cortisol) responses, indicating that both CRH and vasopressin drive HPA-system activity. This conclusion is indirectly derived from the following observations: first, in depression and in experimental animals under chronic stress conditions a shift to a gradually intensifying vasopressinergic regulation of the pituitary-adrenocortical system occurs. Second, the low dose of DEX used in the DEX/CRH test acts primarily on the pituitary and thereby suppresses ACTH and cortisol. In the brain, however, the depletion of endogenous cortisol is not compensated by low-dose DEX since it is unable to cross the blood-brain barrier. Third, as a result, in a depressed DEX-pretreated patient, secretion of ACTH in response to exogenous CRH is greater than in healthy control subjects because of the synergistic actions of the

administered CRH bolus with a large amount of vasopressin present at the corticotrophs. This notion is supported by findings in experimental animals where excessive hormonal responses in the DEX/CRH test could be normalized by pretreatment with a vasopressin antagonist. Also, increased numbers of CRH- and vasopressin-expressing neurons in the PVN have been reported in postmortem studies of depressed patients.

Procedures

CRH test

In endocrine challenge procedures timing is of utmost importance since most hormonal systems have a circadian rhythm. In order to achieve measurable hormone responses and to avoid ceiling effects it is preferable to perform challenge tests at the lowest circadian hormone activity. Also, possible confounding factors need to be controlled.

The CRH stimulation test consists of an intravenous (IV) bolus injection of 100 µg human CRH (hCRH) or 1 µg kg⁻¹ ovine CRH (oCRH), reconstituted in 1 ml 0.02% hydrochloric acid in 0.9% saline, in the late afternoon or early evening when the HPA system is relatively quiescent. Subjects should rest in a supine, comfortable position in a single room after having had a light lunch at least 2 h prior to the procedure. An IV forearm catheter should be put in place 30 min before blood sampling starts to allow for adaptation. In most laboratories the catheter is then connected to a long tubing system which is passed through a sound-proof lock (through-the-wall-technique) into the adjacent laboratory where the blood sampling is performed. Between blood samplings the IV line is kept patent by saline infusion at a rate of 50 ml h⁻¹. Thirty minutes and immediately before bolus injection of CRH, blood samples for baseline levels of ACTH and cortisol are drawn. During the first 2 h after CRH administration blood should be sampled every 15 min, thereafter every 30 min for a further 1 h. In healthy controls, ACTH peak responses occur after 15–30 min while cortisol peaks about 30–60 min after CRH administration. CRH stimulation results in at least a twofold ACTH response with a fairly high interindividual variance and the cortisol response is usually greater than 270 nmol l⁻¹. However, in subjects with hypothalamic dysfunctions (e.g., depression) hormonal responses to CRH may be augmented with delayed peaks. In cases where higher doses of CRH are used, a biphasic ACTH response can be observed caused by cortisol feedback effects. Also, due to a longer half-life, oCRH has a prolonged effect on the pituitary.

Side effects of CRH are normally very mild and transient and may include heat flashes in the area of thorax, neck, or face, thoracic constriction, shortness of breath, and/or quickening of breathing and heart rate.

DEX/CRH Test

Subjects receive 1.5 mg dexamethasone orally at 23.00 p.m. The following day, after a light lunch around 12.00 p.m. they are fitted with an IV forearm catheter at 14.30 p.m. (same procedure as described above for the CRH test). The first blood sample is drawn at 15.00 p.m. (postDEX, but preCRH – referred to as basal). At 15.05 p.m. CRH is injected, and thereafter blood is sampled every 15 minutes until 16.15 p.m.

Results in the DEX/CRH tests are usually reported as hormone concentrations at the predetermined time points. In addition, two derivative parameters, the DELTA-value (peak concentration after 15.05 p.m. minus basal concentration at 15.00 p.m.) and the AUC value (area under the concentration time curve minus linear background, determined by the trapezoidal rule, arbitrary units).

Impact Factors

Factors that impact on the outcome of both the CRH test and the combined DEX/CRH test are manifold, but the most important in the context of stress and depression research are time of day, age, gender, phase of menstrual cycle, and medication. Obviously, any medical, especially endocrine disorder, can have major effects on the test results.

Age

Most DEX/CRH test studies found that elderly patients have higher DELTA and AUC of ACTH and cortisol values in comparison to those in young subjects. Together with the observation that baseline, resting cortisol levels – as assessed by repeated, circadian blood or saliva sampling – are elevated in older age, these findings suggest that the activity of the HPA system increases and feedback inhibition might decrease during aging. To date, it is still unclear and a matter of great debate, whether this increasing HPA-system activity in aging constitutes a physiological or adaptive response to the accumulation of daily hassles of life or whether this is an indication of gradually failing compensatory endocrine mechanisms.

Gender

Women, especially elderly women, release more ACTH and cortisol after DEX/CRH challenge than elderly men. Also, DEX-pretreated, basal hormone concentrations tend to be higher in women compared with men. However, this gender difference is not pronounced in younger (aged < 30 years) individuals. This age-gender interaction in DEX/CRH test outcome may be due in part to the larger reductions in gonadal steroids (estradiol) in older, postmenopausal women.

Menstrual-Cycle Phase

Menstrual-cycle phase as well as estrogen administration (e.g., oral contraceptive use or hormonal replacement therapy (HRT)) are consistently found to have a major impact on HPA reactivity. Women show a reduced HPA reactivity during the follicular phase of the menstrual cycle as opposed to the luteal phase. In elderly women who received HRT, ACTH and cortisol responses to DEX/CRH stimulation were dampened in comparison to placebo-treated elderly females. To the best of our knowledge, studies in men with regard to testosterone and challenge test outcomes are not yet available.

Medication

A variety of different drugs has been shown to alter HPA-system activity and thus, endocrine challenge-test results. Antidepressants but also benzodiazepines and antipsychotic drugs, for example, are known to alter HPA-system activity. Thus it is advisable to perform challenge tests in drug-free individuals, unless, the effect of certain medications are of interest. Note, that alcohol, caffeine, and nicotine increase HPA-system activity, and, unfortunately, withdrawal from these substances has the same effect!

DEX/CRH Test in Depressive Disorders and Other Psychiatric Diseases

The DEX/CRH test is one of the most sensitive neuroendocrine function tests to assess HPA dysregulation in psychiatric disorders, including dementia, exceeding greatly the sensitivity of the DST (at least in the case of major depression). However, it should be emphasized that DEX/CRH test results are non-specific, that is, they are not indicative for any mental syndrome.

Several studies used repeated DEX/CRH testing to monitor HPA system functioning during the course of antidepressant treatment. It was shown that initially abnormal HPA responses gradually disappeared in those patients responding favorably to treatment. Noteworthy is the fact that normalization of HPA-system functioning preceded the resolution of psychopathological symptoms. In contrast, patients whose HPA system imbalance did not sufficiently normalize during the course of treatment, had a greater liability for an early relapse.

With respect to bipolar disorder enhanced HPA reactivity after DEX/CRH challenge is also well known and is more pronounced compared to patients with unipolar depression. Furthermore, patients with remitted bipolar disorder have higher responses compared to those with unipolar remitted disorder and healthy controls.

To tackle the issue whether HPA dysregulation is a disease-predisposing, pre-existing vulnerability factor or whether it is a consequence of depression, high-risk, but healthy probands, with at least one depressed, first-degree relative underwent DEX/CRH testing. The high-risk probands, who themselves had never suffered from depression, showed higher cortisol responses compared with controls. However, HPA responses to DEX/CRH challenge in high-risk probands who developed an affective disorder during the 4-years follow-up period did not differ in terms of DEX/CRH results from those high-risk subjects who had remained healthy during this period. These findings suggest that subtle changes of HPA-system function are a genetic trait that increases the risk of developing depression and possibly other stress-related disorders.

An enhanced release of ACTH and cortisol following DEX/CRH compared with healthy controls is also described in other psychiatric disorders such as schizophrenia, panic disorder, and borderline-personality disorder, especially in case of early sexual or physical abuse.

Conclusion

In conclusion, HPA-system dysbalance is a frequent neuroendocrine sign in patients with mental disorders or in individuals subjected to stressful conditions. To assess activity of the HPA-system basal measurements of cortisol concentrations in saliva, blood, or 24-h urine may be used, each mode of collection having its specific advantages and disadvantages. To study HPA-system regulation, the DST has been the most widely used procedure since it is simple to apply,

relatively inexpensive, and benign for the individual. During the last decade, the combined DEX/CRH-stimulation test has gained increasing attention, mostly in research settings, due to its high sensitivity for stress-related conditions and mental disorders. With the emergence of the DEX/CRH test as a laboratory test in psychological and psychiatric research, the usage of the simple CRH test in this context has markedly declined. However, no single test is superior to any other, their applicability, reliability and validity depending solely on the research questions asked.

Further Reading

- Baghai, T. C., Schüle, C., Zwanzger, P., et al. (2002). Evaluation of salivary based combined dexamethasone/CRH test in patients with major depression. *Psychoneuroendocrinology* 27, 385–399.
- Deuschle, M., Gotthardt, U., Schweiger, U., et al. (1997). With aging in humans the activity of the hypothalamus-pituitary-adrenal system increases and its circadian rhythm flattens. *Life Sciences* 61, 2239–2246.
- Deuschle, M., Schweiger, U., Gotthardt, U., et al. (1998). The combined dexamethasone/CRH-stimulation test is more closely associated with features of diurnal activity of the hypothalamus-pituitary-adrenal system than the dexamethasone suppression test. *Biological Psychiatry* 43, 762–766.
- Hatzinger, M., Hemminger, U. M., Baumann, K., et al. (2002). The combined DEX-CRH test in treatment course and long-term outcome of major depression. *Journal of Psychiatric Research* 36, 287–297.
- Heuser, I., Gotthardt, U., Schweiger, U., et al. (1994). Age-associated changes of pituitary-adrenocortical hormone regulation in humans: importance of gender. *Neurobiology and Aging* 15, 227–231.
- Heuser, I., Yassourides, A. and Holsboer, F. (1994). The combined dexamethasone/CRH test: a refined laboratory test for psychiatric disorders. *Journal of Psychiatric Research* 4, 341–356.
- Ising, M., Lauer, C. J., Holsboer, F., et al. (2005). The Munich vulnerability study on affective disorders: pre-morbid neuroendocrine profile of affected high-risk probands. *Journal of Psychiatric Research* 39, 21–28.
- de Kloet, E. R., Joels, M. and Holsboer, F. (2005). Stress and the brain: from adaptation to disease. *Nature Reviews – Neuroscience* 6, 463–475.
- Kudielka, B. M., Schmidt-Reinwald, A. K., Hellhammer, D. H., et al. (1999). Psychological and endocrine responses to psychosocial stress and DEX-CRH in healthy postmenopausal women and young controls: the impact of age and a two-week estradiol treatment. *Neuroendocrinology* 70, 422–430.
- Lammers, C.-H., Garcia-Borreguero, D., Schmider, J., et al. (1995). Combined dexamethasone/corticotropin-

- releasing hormone test in patients with schizophrenia and in normal controls: II. *Biological Psychiatry* 38, 803–807.
- Otte, C., Hart, S., Neylan, T. C., et al. (2005). A meta-analysis of cortisol response to challenge in human aging: importance of gender. *Psychoneuroendocrinology* 30, 80–91.
- Rinne, T., DeKloet, E. R., Wouters, L., et al. (2002). Hyper-responsiveness of hypothalamic-pituitary-adrenal axis to combined dexamethasone/corticotropin releasing hormone challenge in female borderline personality disorder subjects with a history of sustained childhood abuse. *Biological Psychiatry* 52, 1102–1112.
- Schmider, J., Lammers, C.-H., Gotthardt, U., et al. (1995). Combined dexamethasone/corticotropin-releasing hormone test in acute and remitted manic patients, in acute depression, and in normal controls: I. *Biological Psychiatry* 38, 797–802.
- Watson, S., Gallagher, P., Ritchie, J. C., et al. (2004). Hypothalamic-pituitary-adrenal axis function in patients with bipolar disorder. *British Journal of Psychiatry* 184, 496–550.
- Yanovski, J. A., Cutler, G. B., Chrousos, G. P., et al. (1993). Corticotropin-releasing hormone stimulation following low-dose dexamethasone administration: a new test to distinguish Cushing's syndrome from pseudo-Cushing's states. *JAMA* 269, 2232–2238.
- Zobel, A. W., Nickel, T., Sonntag, A., et al. (2001). Cortisol response in the combined dexamethasone/CRH test as predictor of relapse in patients with remitted depression. A prospective study. *Journal of Psychiatric Research* 35, 83–94.

DHEA

J Herbert

Cambridge University, Cambridge, UK

© 2007 Elsevier Inc. All rights reserved.

Glossary

<i>γ-Aminobutyric acid (GABA)</i>	A major transmitter in the brain; known to be implicated in anxiety.
<i>Addison's disease</i>	A condition resulting from the failure of the adrenal glands; treated by replacing the missing hormones.
<i>Adrenal cortex</i>	The outermost part of the adrenal gland, which secretes a variety of steroid hormones including cortisol and DHEA.
<i>Adrenarche</i>	A stage that occurs at around 8 years in humans; characterized by a rapid increase in DHEA levels in the blood.
<i>Adrenocorticotrophic hormone (ACTH)</i>	A hormone secreted by the pituitary gland that controls the activity of part of the adrenal cortex (e.g., cortisol secretion).
<i>Allosteric interaction</i>	The interaction of two molecules because of their shape.
<i>Androgen</i>	A set of steroids having masculine-like actions on the body; testosterone is the major one in humans.
<i>Anoxia</i>	The lack of oxygen.
<i>Blood-carrier (binding) protein</i>	A protein to which substances that are carried in the blood are bound; it affects the way they are transported into the brain and other tissues.

Cerebrospinal fluid (CSF)

The fluid surrounding the brain and also present inside it in cavities called ventricles; compounds entering or leaving the brain may travel through the CSF.

Downstream gene expression
Estrogens

The alteration of the way in which some genes are activated caused by steroids binding to DNA.

The major steroid hormones secreted by the ovaries and, during pregnancy, by the placenta.

Femoral neck

The upper part of the thigh bone that joins it to the pelvis; a common site of fracture in the elderly.

Glucocorticoids

One class of steroids secreted by the adrenals, including cortisol; so-named because, among other actions, they increase blood levels of glucose.

Hippocampus

An area of the brain associated with the formation of certain types of memory.

Immune function

The ability to resist infection; also associated with resistance to some forms of cancer.

Neurogenesis

The process of forming new nerve cells in the brain; occurs mostly during development but is now known to continue in some areas throughout adult life.

Neurosteroid

A steroid formed locally in the brain, as opposed to one synthesized and secreted by a peripheral endocrine gland. Some authors use the term to mean any steroid that has an action on the brain.

<i>Polycystic ovaries</i>	A largely inherited condition in which the ovaries contain large abnormal egg follicles; also associated with insulin resistance and obesity.
<i>Posttraumatic stress disorder (PTSD)</i>	A persistent disorder that sometimes follows the experiencing of life-threatening events or disasters; characterized by nightmares and persistent intrusive memories of the event, set off by a sight or sound associated with that event.
<i>Pregnenolone</i>	The first of a series of steroids made by the adrenal glands (and other organs) from cholesterol.
<i>Pyramidal neurons</i>	Large nerve cells found in several parts of the brain, including the hippocampus.
<i>Rodents</i>	A class of mammals that includes rats and mice.
<i>Serotonin (5-HT)</i>	A chemical used as a transmitter throughout the brain; implicated in depression and the target of a major class of antidepressants (the selective serotonin reuptake inhibitors, SSRIs).
<i>Toxins</i>	Substances that poison or damage cells.
<i>Ungulates</i>	A class of mammals that includes sheep and cows.

Dehydroepiandrosterone (DHEA) is an adrenal-derived steroid present in very high concentrations in the blood of primates. It is secreted mainly from the innermost zone of the adrenal cortex (the zona reticularis) which, in the human fetus, is hypertrophied and called the X zone. This reflects the fact that DHEA in the fetus is a major source of other steroids (e.g., estrogens). Because the human fetus itself lacks the enzymes to process DHEA, the formation of daughter steroids (e.g., estrogens and androgens) occurs in the placenta. At birth the X zone atrophies and blood DHEA falls to very low levels. Levels remain low until around age 8, when they begin to increase progressively, a process termed adrenarche. The stimulus that underlies adrenarche remains speculative (no pituitary controlling factor has so far been conclusively identified), as does its physiological significance, although DHEA can promote the growth of pubic hair. Levels continue to increase throughout puberty, reaching apogee in early adult life (around 20 years). In adulthood, DHEA is present in the blood largely (around 90%) as its sulfated derivative (DHEAS). There is about 10 times more DHEA(S) in the blood as cortisol, the next most abundant steroid. From its peak value, levels then decline progressively but variably with age, so that by around age 60 they are only approximately one-third or less of those during the twenties. As for

adrenarche, ignorance of the factors controlling the adrenal production of DHEA(S) prevents further understanding of the cause of this decline. Infusions of adrenocorticotrophic hormone (ACTH), the major controlling signal from the pituitary in the regulation of cortisol, also increase DHEA(S), but the very different life-span trajectory of DHEA compared to cortisol argues strongly against this being the only, or even the major, control system. For example, ACTH (and cortisol) do not increase during adrenarche, and levels of cortisol increase, if anything, with age. The functions of DHEA(S) in humans remain puzzling. Conversion to other steroids with known functions (e.g., estrogens and androgens) is quite small, and, in any case, seems unnecessary. Old World monkeys have a similar pattern of DHEA(S), but rats and ungulates have only very low levels (around 5% of those in humans) throughout the life span.

The rodent brain can make steroids locally from the precursor molecule cholesterol, including pregnenolone, the immediate precursor of DHEA. Such steroids are, consequently, called neurosteroids. However, it is uncertain whether the adult human brain (in contrast to the fetal one) has the enzymes required to convert pregnenolone into DHEA. In any case, the high amounts of DHEA(S) in the blood in humans seem to indicate that most DHEA(S) in the brain comes from the blood. Nevertheless, there are those who consider that local production of neurosteroids may play a significant role in regulating brain function, including mood and emotions. The direct measurement of DHEA and DHEAS in the cerebrospinal fluid (CSF), which gives a good indicator of the concentrations to which neurons are exposed, shows that DHEA levels are approximately 5% of those in the blood in humans and that DHEAS (a less lipid-soluble steroid) levels are approximately 0.2%. Levels in the saliva, a convenient way of measuring DHEA(S) status, are similar. In both cases, it is not clear what prevents more of the blood-borne steroid from entering the brain. No blood-binding proteins, which could limit the access of lipid-soluble steroids to the brain, have been identified for DHEA, yet the relative amounts in the blood and CSF correspond to other steroids (e.g., cortisol) that do have a blood-carrier (binding) protein. DHEA(S) levels in the blood are abnormally high in women with polycystic ovaries.

Stressful events, such as intercurrent illness, tend to lower DHEA(S) levels in humans (in contrast to cortisol, which is elevated). It is not clear how this relates to the age-dependent decline, that is, whether this occurs more or less in aged people or whether current (baseline) levels influence the response to

stress. People with lower baseline levels may be at increased risk for stress-related disorders such as coronary heart disease or major depression, although the literature is not unanimous on this. Interest in the DHEA response to stress lies in the possibility that it represent a resilience factor against some of the adverse results of chronic or intermittent stress. DHEA supplements are being taken by large numbers of older people as an anti-aging remedy – but mostly under unregulated circumstances, which prevents the accurate assessment of possible beneficial or adverse effects. Those with Addison's disease (adrenal failure) have virtually no DHEA in their blood, and restoring DHEA to near-normal values has beneficial effects on fatigue and, in some cases, mood. Supplementary DHEA in adrenal-intact older people has similar effects, but so far there have been no convincing reports of beneficial actions on cognitive functions such as memory. There is some evidence that age-related bone loss (e.g., in the femoral neck) may benefit from supplementary DHEA. There are no reports of people taking DHEA to counteract stressful events (such as a loss or excessive demand) or to prevent the adverse effects of life-threatening events (e.g., posttraumatic stress disorder, PTSD).

Experimental studies on DHEA have been mainly conducted on rodents, and it is important to remember that in these species DHEA administration is a pharmacological rather than physiological maneuver. Nevertheless, supplementary DHEA has been found to improve immune function (particularly in older animals). DHEA(S) may moderate the actions of glucocorticoids (cortisol in humans, corticosterone in rats), which may indicate a significant role for this steroid in stress-related responses in humans, in which cortisol levels can reach very high levels. Because excess cortisol has a variety of damaging actions on the body (including the immune system, metabolism, and brain function), individual or age-related differences in DHEA may influence the likelihood of adverse reactions to chronic or uncontrollable stress. DHEA counteracts the degeneration of the thymus that otherwise occurs after excess glucocorticoids (e.g., poststress) and has been found to improve survival rates after experimental infections in older mice. One of the measurable actions of glucocorticoids on the brain occurs in the hippocampus, an area that is curiously sensitive to the damaging actions of toxins and anoxia and that shows some of the earliest age-related changes in the brain, including those typical of incipient Alzheimer's disease. Corticosterone damages the pyramidal neurons in the CA3 region in rats and also reduces the proliferation

rates of progenitor (stem) cells in the dentate gyrus (part of the hippocampus), which continues to produce new neurons throughout adult life in rats and in humans. Neurogenesis declines with age; however, the functional significance of neurogenesis in the dentate gyrus (e.g., for certain forms of memory) is still uncertain. DHEA has some moderating actions on these glucocorticoid-dependent actions in the brain. It may also improve the survival of embryonic neurons and thus play a hitherto unsuspected role in neural development.

Depression is a well-known sequela to an acute stress (a severe adverse life event), and higher cortisol levels are a risk factor for subsequent depression. Cortisol is elevated during depression in some cases, and the cortisol/DHEA ratio has been shown to predict the persistence of or recovery from depression in adolescents (in whom DHEA is relatively high and rising). Recently, it has been proposed that altered neurogenesis in the hippocampus may underlie either the onset of depression or the ability of serotonin (5-HT)-acting drugs such as selective serotonin reuptake inhibitors (SSRIs) to alleviate this condition. If this is so, it points to another possible association between DHEA, which also has been shown to be an effective antidepressant in middle-age people, and the response to stress.

Understanding the way that DHEA has its effects is hampered by a lack of information about its actions at a cellular level. Unlike other steroids, no receptor in the interior of the cell has been identified for DHEA. This makes the way that DHEA alters cell function a mystery because these receptors are responsible for many of the effects of other steroids on cell function or downstream gene expression. Nor is the binding of corticosterone to its receptor altered by DHEA in rats, so there is no immediate explanation of the way that DHEA moderates glucocorticoids. In the brain, DHEA may interact allosterically with membrane-bound receptors for γ -aminobutyric acid (GABA_A) and this may go some way to explaining its actions on the central nervous system (CNS). What is clear is that defining DHEA(S) in humans as a weak androgen is not only inaccurate but has discouraged the investigation of the role of this prominent steroid in humans. Its unique age-related pattern of secretion and its possible role in intercurrent stressful events such as illness make such studies overdue.

See Also the Following Article

Androgen Action.

Further Reading

- Auchus, R. J. and Rainey, W. E. (2004). Adrenarche – physiology, biochemistry and human disease. *Clinical Endocrinology* (Oxford) **60**, 288–296.
- Bauer, M. E. (2005). Stress, glucocorticoids and ageing of the immune system. *Stress* **8**, 69–83.
- Blouin, K., Despres, J. P., Couillard, C., et al. (2005). Contribution of age and declining androgen levels to features of the metabolic syndrome in men. *Metabolism* **54**, 1034–1040.
- Butterfield, M. I., Stechuchak, K. M., Connor, K. M., et al. (2005). Neuroactive steroids and suicidality in posttraumatic stress disorder. *American Journal of Psychiatry* **162**(2), 380–382.
- Cleare, A. J., O’Keane, V. and Miell, J. P. (2004). Levels of DHEA and DHEAS and responses to CRH stimulation and hydrocortisone treatment in chronic fatigue syndrome. *Psychoneuroendocrinology* **29**, 724–732.
- Dillon, J. S. (2005). Dehydroepiandrosterone, dehydroepiandrosterone sulfate and related steroids: their role in inflammatory, allergic and immunological disorders. *Current Drug Targets – Inflammation & Allergy* **4**, 377–385.
- Fonda, S. J., Bertrand, R., O’Donnell, A., et al. (2005). Age, hormones, and cognitive functioning among middle-aged and elderly men: cross-sectional evidence from the Massachusetts Male Aging Study. *Journal of Gerontology A: Biological Sciences and Medical Sciences* **60**, 385–390.
- Genazzani, A. R., Inglese, S., Lombardi, I., et al. (2004). Long-term low-dose dehydroepiandrosterone replacement therapy in aging males with partial androgen deficiency. *Aging Male* **7**, 133–143.
- Hajszan, T., MacLusky, N. J. and Leranth, C. (2004). Dehydroepiandrosterone increases hippocampal spine synapse density in ovariectomized female rats. *Endocrinology* **145**, 1042–1045.
- Herbert, J., Goodyer, I. M., Grossman, A. B., et al. (2006). Do corticosteroids damage the brain? *Journal of Neuroendocrinology* **18**, 393–411.
- Hunt, P. J., Gurnell, E. M., Huppert, F. A., et al. (2000). Improvement in mood and fatigue after dehydroepiandrosterone replacement in Addison’s disease in a randomized, double blind trial. *Journal of Clinical Endocrinology and Metabolism* **85**, 4650–4656.
- Huppert, F. A. and VanNiekerk, J. K. (2001). Dehydroepiandrosterone (DHEA) supplementation for cognitive function. *Cochrane Database of Systematic Reviews* **2**, CD000304.
- Labrie, F. (2004). Adrenal androgens and intracrinology. *Seminars in Reproductive Medicine* **22**, 299–309.
- Libe, R., Barbetta, L., Dall’Asta, C., et al. (2004). Effects of dehydroepiandrosterone (DHEA) supplementation on hormonal, metabolic and behavioral status in patients with hypoadrenalism. *Journal of Endocrinological Investigation* **27**, 736–741.
- McEwen, B. S. (2003). Interacting mediators of allostasis and allostatic load: towards an understanding of resilience in aging. *Metabolism* **52**(supplement 2), 10–16.
- Perrini, S., Laviola, L., Natalicchio, A., et al. (2005). Associated hormonal declines in aging: DHEAS. *Journal of Endocrinological Investigation* **28**, 85–93.
- Pico-Alfonso, M. A., Garcia-Linares, M. I., Celda-Navarro, N., et al. (2004). Changes in cortisol and dehydroepiandrosterone in women victims of physical and psychological intimate partner violence. *Biological Psychiatry* **56**(4), 233–240.
- Rasmusson, A. M., Vythilingam, M. and Morgan, C. A., III (2003). The neuroendocrinology of posttraumatic stress disorder: new directions. *CNS Spectrum* **8**, 651–656.
- Santoro, N., Torrens, J., Crawford, S., et al. (2005). Correlates of circulating androgens in mid-life women: the study of women’s health across the nation. *Journal of Clinical Endocrinology and Metabolism* **90**, 4836–4845.
- Schmidt, P. J., Daly, R. C., Bloch, M., et al. (2005). Dehydroepiandrosterone monotherapy in midlife-onset major and minor depression. *Archives of General Psychiatry* **62**, 154–162.
- Tannenbaum, C., Barrett-Connor, E., Laughlin, G. A., et al. (2004). A longitudinal study of dehydroepiandrosterone sulphate (DHEAS) change in older men and women: the Rancho Bernardo Study. *European Journal of Endocrinology* **151**, 717–725.
- Widstrom, R. L. and Dillon, J. S. (2004). Is there a receptor for dehydroepiandrosterone or dehydroepiandrosterone sulfate? *Seminars in Reproductive Medicine* **22**, 289–298.

Diabetes, Type 1

A Riazi and C Bradley

Psychology Department, Royal Holloway,
University of London, Surrey, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by A Riazi and C Bradley, volume 1, pp 688–693,
© 2000, Elsevier Inc.

Stress and Type 1 Diabetes Onset

Stress and Type 1 Diabetes Control

Stress Management Training and Type 1 Diabetes Control

Glossary

<i>B or β cells</i>	The cells in the islets of Langerhans in the pancreas that produce insulin.
<i>Diabetes mellitus</i>	A heterogeneous group of disorders, characterized by hyperglycemia, and disturbances of carbohydrate, fat, and protein metabolism that are associated with absolute or relative deficiencies of insulin secretion and/or insulin action.
<i>Glycosuria</i>	Sugar in the urine.
<i>Hyperglycemia</i>	Excessive levels of glucose in the blood; a feature of untreated or undertreated diabetes mellitus.
<i>Hypoglycemia</i>	Abnormally low levels of glucose in the blood. Symptoms are idiosyncratic but may include trembling, faintness, sweating, palpitations, mental confusion, slurred speech, headache, loss of memory, and double vision. Severe untreated hypoglycemia may lead to fits or coma and, on rare occasions, death. It can be caused by an overdose of insulin.
<i>Islet cell antibodies (ICAs)</i>	Antibodies that are present in the blood when the body develops an autoimmune reaction to the pancreatic islet cells. The presence of ICAs indicates that the process of destruction has started that will eventually lead to type 1 diabetes. This process appears to take several years before diabetes is clinically manifested.
<i>Insulin</i>	A peptide hormone produced in the β cells of the islets of Langerhans in the pancreas. Insulin facilitates and accelerates the movement of glucose and amino acids across cell membranes. It also controls the activity of certain enzymes within the cells concerned with carbohydrate, fat, and protein metabolism.

Type 1 diabetes mellitus, previously known as insulin-dependent diabetes mellitus (IDDM), juvenile-onset diabetes, or ketosis-prone diabetes, is the most common form of diabetes occurring in children and young adults of European origin. The age at clinical onset of the condition is usually under the age of 40 years and often under 30 years. People who have type 1 diabetes lose the ability to produce insulin. Exogenous insulin, usually delivered by injection, is needed continuously throughout the lifetime of a person with type 1 diabetes, with the possible exception of a brief honeymoon period that may occur within a year of clinical onset when endogenous insulin production is temporarily restarted. The onset of type 1 diabetes is abrupt, with severe thirst, excessive urination, and dramatic weight loss. Individuals usually present to the doctor with one or more of these symptoms and an elevated blood glucose level.

Once it has developed, type 1 diabetes can be managed by balancing a combination of insulin injections, intake of carbohydrates in the diet, and energy expenditure. The goal of treatment is to maintain blood glucose levels as close to the normal range as possible in order to reduce the risk of chronic complications while also avoiding the dangers of blood glucose falling to hypoglycemic levels.

Stress and Type 1 Diabetes Onset

Although it has not been possible to determine the exact pathogenesis involved in the expression of type 1 diabetes, it is clear that genetics play an important role. The genetic component of type 1 diabetes is indicated by its increased prevalence of 5% in first-degree relatives, compared with less than 1% in the general population. However, genetic susceptibility is not sufficient to cause diabetes because the majority of people with disease-associated alleles do not develop type 1 diabetes. Environmental factors play an important role in the overt expression of type 1 diabetes, although the mechanisms are still unclear. Psychological stress may have a role in increasing vulnerability to viral infection or impairing defense mechanisms against infection, thereby facilitating the progression of the hidden pathological process. For example, stress-related changes in immune function may increase the likelihood of viral or bacterial disease, which may provide the initial insult to the B (or β) cells. A second mechanism whereby stress may be implicated in diabetes onset may occur around the time when the diabetes becomes symptomatic.

Stress-related counterregulatory hormone activity may aggravate the metabolic disturbance that has already developed. In fact, if the already-elevated blood glucose level increases to beyond the renal threshold, glycosuria causes dehydration, which may then produce the first symptoms of overt diabetes. Many anecdotal accounts and descriptive reports of life stresses occurring synonymously with symptomatic type 1 diabetes onset may reflect this second mechanism in which stress triggers overt manifestation of symptoms.

Animal Studies

It has been shown that stress may be associated with the onset of type 1 diabetes in the genetic model for this type of diabetes. In the diabetes-prone BB Wistar rat, 30 to 70% of these animals have impaired glucose tolerance with hypoinsulinemia and hyperglycemia by 150 days of age. It has been shown that these rats developed diabetes earlier, when exposed to restraint and crowding. It has also been found in the BB rat that chronic stress significantly increases the incidence of the phenotypic expression of the gene for type 1 diabetes. It was found by Lehman et al. in 1991 that 80% of the stressed males and 70% of the stressed female animals developed diabetes, compared with 50% in both control groups. However, because of the range of additional immune and endocrinological abnormalities evident in the BB rats, the generalizability of these findings to humans is limited.

Human Studies

Research has shown that people who go on to develop type 1 diabetes are more likely to suffer a major family loss or an increase in other stressful life events before diagnosis. Some of these studies have been poorly controlled and were also prone to recall bias, in which the pattern of life events reported may reflect an attempt to find an explanation for diabetes onset rather than a difference in the events actually encountered. However, a more carefully controlled study of life events that avoided problems of recall bias by Robinson and Fuller in 1985 has also shown that people with diabetes had significantly more severe life events within the 3 years before diagnosis than either nondiabetic siblings or matched controls. It is possible that life events experienced over longer time periods may play a role in the etiology of type 1 diabetes. It is clear from prospective studies of islet cell antibodies (ICAs) that many years may elapse between the actions of the possible stress-related causal agents that initiate cell damage and the appearance of symptomatic diabetes.

Clear evidence of the effect of stress on diabetes onset is difficult to establish. Some difficulties include methodological problems and interpretation of the findings that involve expensive, large, prospective studies or a reliance on retrospective methods to identify the relationship between stressful experiences and the onset of the disease. However, even in a seemingly large trial, it can be difficult to delineate the effects of stress if a broad age range of people is recruited into the trial. As discussed by Lloyd et al. in 2005 it is often the case that differences in the type and intensity of life changes are associated with different ages. The relationship between stress and diabetes control is more amenable to study.

Stress and Type 1 Diabetes Control

It is widely recognized by people with diabetes and their clinicians that psychological stress can impair glycogenic control. Theoretically, stress-related hyperglycemia should be greater in type 1 patients (compared to type 2 patients) because type 1 patients have little or no endogenous insulin to offset increased blood glucose levels. Stress may affect diabetes control in at least two ways:

1. A direct psychophysiological effect via sympathetic and pituitary activity resulting in the elevation of catabolic hormone levels and the suppression of anabolic hormones. In people with diabetes, this may result in increased blood glucose levels, although, for a small minority, less readily understood decreases in blood glucose levels result.
2. A behavioral mechanism whereby stress leads to behavioral changes capable of disrupting self-care. For example, the occurrence of unexpected, frustrating events may disrupt the diabetes self-care routines.

Acute Stress and Blood Glucose Control

Pioneering work examining the relationship between stress and diabetes was most often conducted in the laboratory, and the stresses induced were normally acute. In the early 1950s, Hinkle and others induced stress in some of their patients by a psychiatric interview and found changes in the patients' blood ketone levels that remitted when the stress was removed. These researchers carried out a series of studies that showed that individuals without diabetes and individuals with diabetes (both type 1 and type 2 patients were studied) have a metabolic response to psychological stress that included changes in urine glucose, blood glucose, and blood ketones, but the response of those with diabetes was greater. Research in the 1960s included investigations of stress under hypnosis and

examination stress, which resulted in a decrease in blood glucose levels. However, the early studies have been criticized on methodological and conceptual grounds. There was concern that the stressors used were not sufficiently potent or reproducible. Furthermore, the grouping together of heterogeneous patients with type 1 and 2 diabetes and patients with different degrees of blood glucose control in studies of very small groups were also a focus of criticism.

Several seemingly well-controlled acute stress studies conducted in the 1980s reported no significant changes in blood glucose control in response to potential stressors such as mental arithmetic and public speaking. However, in an attempt to meet the criticisms leveled at the earlier studies, these later studies overlooked the possibility that individual differences in response to stress might be real and interesting and not a reflection of methodological inadequacies. More recent studies of experimental stress have examined the physiological mechanisms hypothesized to mediate the relationship between stress and blood glucose control. There is evidence that reduced blood flow to the insulin injection site and insulin resistance over several hours may cause increased blood glucose levels in individuals with type 1 diabetes in response to acute laboratory stressors. There is also evidence that the level of blood glucose rises in some patients and falls in others in response to a laboratory stressor (Stroop test) and that these changes can be largely explicable in terms of changes in injection-site blood flow. Vasodilation at the subcutaneous insulin injection site may in some cases lead to a paradoxical hypoglycemic effect during acute stress via an increased rate of absorption of insulin, an effect that is then counterbalanced to a greater or lesser degree by increases in counterregulatory hormones. More commonly, the absorption of insulin may fall during stress as a result of vasoconstriction and reduced skin blood flow and contribute to hyperglycemia.

Idiosyncratic blood glucose responses that were reliable across a 12-week period in individuals with type 1 diabetes have been reported in response to caffeine and to the competitive playing of a video game. Researchers have therefore begun to consider individual differences in response to stress to be real and interesting and not a reflection of methodological inadequacies. The inconsistent findings of previous research may have arisen at least in part because of differences in the stress responsiveness of the individual patients recruited.

Life Stress and Diabetes Control

Major life events Studies of major life events and diabetes control have produced fairly consistent results

that have suggested that increased life events (over past several months to a year) are associated with raised blood glucose levels, usually measured by glycosylated hemoglobin (GHb) or the related measures of hemoglobin A1 (HbA1) and hemoglobin A1c (HbA1c). These measures reflect average blood glucose levels over the previous 6–8 weeks. In some studies, more life events were reported by groups showing a stronger association between life events and HbA1c levels. This could be due to a perceptual bias that may have led people who experienced a greater disturbance in association with life events to be more likely to recall the life events. It could be that the subgroups did not actually differ in the number of life events reported, but only in their recollection of the events. However, a two-way causal link in which life events cause disruptions in diabetes control, which in turn cause increased life events has been suggested (see [Figure 1](#)). There is evidence that glycemic fluctuations themselves can contribute to behavioral changes. It is now well-established that both hypo- and hyperglycemia can impair cognitive functioning. Cognitive impairment may in turn cause inadvertent behavioral changes (e.g., in poor self-care responses to feedback from blood glucose monitoring), which can affect blood glucose control. Poorly controlled diabetes may also result in mood changes, causing interpersonal conflict and thereby increasing stress levels and associated physiological reactivity. Physical symptoms are also caused by extreme blood glucose levels (e.g., fatigue is often associated with high blood glucose levels).

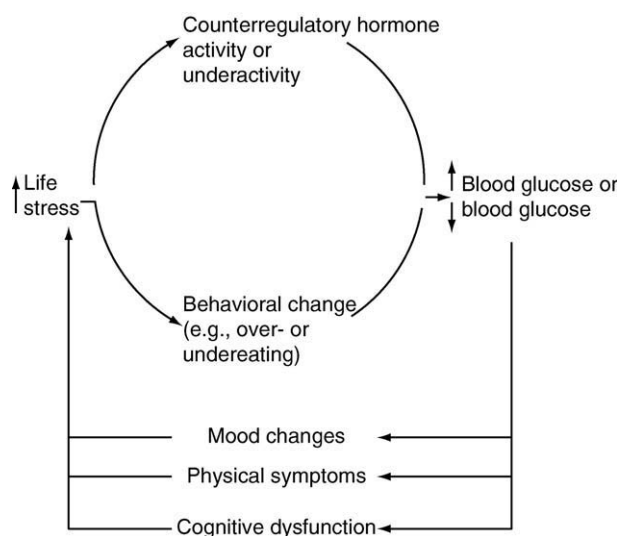


Figure 1 An elaborated two-way causal model of relationships between life stress and diabetes control. From Bradley, C. (1988). Stress and diabetes. In: Fisher, S. & Reason, J. (eds.) *Handbook of life stress, cognition and health*, pp. 383–431. Chichester, UK: John Wiley.

Hypoglycemic comas may have many consequences, including accidents, job loss, and relationship problems. Such associations are described in the stress–blood glucose model shown in [Figure 1](#).

The practice of aggregating life stresses and blood glucose levels across time, incorporated by the majority of the life event studies, assumes minimal within-individual variation in stress and in blood glucose response. Such assumptions are not supported by common experience or by the literature, which suggests that the stress experienced varies in individuals from day to day. Information about the variability of blood glucose levels is lost when measures of GHb are used that give only an average of blood glucose control. Hypoglycemia is not reflected in GHb measures because of its transient nature. Multiple observations of stress are required, together with serial blood glucose measurement. Recent research has focused on the relationship between minor daily events and diabetes control (using serial blood glucose measurements) in order to overcome these methodological limitations.

Minor daily events Investigations of minor events have provided more interpretable results with studies reflecting individual differences in response to stress that were also seen in some of the laboratory-based acute stress studies. Although the overall picture suggests that increased daily stress correlates with increased blood glucose levels in people with type 1 diabetes, an examination of individuals' blood glucose responses to stress indicates that some people display blood glucose reactivity to stress and some people do not. In a 1990 study by Halford et al., approximately one-half of the 15 participants with type 1 diabetes had significant associations between stress and blood glucose levels independent of the effects of diet and exercise self-management. In a small 2000 study by Kramer et al., two of the nine participants had significant correlations between stress and HbA1c, whereas six participants exhibited significant associations between stress and daily level or variability of glucose readings. In a larger 2004 study by Riaz et al. with 54 participants, in approximately one-third of the sample stress was significantly associated with either same- or next-day blood glucose levels. In those people who display stress reactivity, most show an increase in blood glucose but there is also evidence to suggest that a smaller proportion of people show a decrease in blood glucose. Therefore, in order for people with type 1 diabetes to take appropriate action to prevent or correct stress-related disruptions in diabetes control, they need to discover empirically how their own blood glucose levels respond to different kinds of stress.

Investigation of Correlates of Stress Reactivity

There is some evidence to suggest that people with poor control of their diabetes may be more stress reactive (stress being either laboratory-induced or in the form of life events). This supports the model in [Figure 1](#). There is also evidence from Stabler et al. in 1986 that children with type 1 diabetes who are classified as having type A personality characteristics show elevated blood glucose levels after playing a stressful video game, whereas type B children with type 1 diabetes showed a decrease in blood glucose levels. However, there is also evidence that internality and self-esteem did not relate to stress reactivity in the group of women with type 1 diabetes studied by Aikens and colleagues in 1994. It is likely that individual differences in stress response are mediated by both physiological and psychological processes. For example, Riaz et al. in 2004 found that stress-reactive individuals, compared to non-stress-reactive individuals, had both high HbA1c values and significant relationships between emotion-focused coping and blood glucose levels. However, further research is necessary to clarify the associations between these variables and stress reactivity.

Stress Management Training and Type 1 Diabetes Control

The overly simple model of life stress causing raised blood glucose levels has inspired the use of stress management training as an aid to diabetes control. There is now self-help guidance available in a book by Surwit and Bauman published in 1994 for people with diabetes who are interested in achieving better diabetes control through stress management training, in particular, relaxation training. Relaxation is thought to decrease the levels of both cortisol and catecholamines and, hence, is expected to prevent stress-induced increases in blood glucose levels.

The research to date supports the view that stress management techniques may be valuable for aiding diabetes management in some people but not in others. It has been suggested that relaxation techniques are unlikely to do harm except when blood glucose is already tightly controlled and the insulin dosage is not appropriately reduced to balance the effects of relaxation on insulin requirements (which may well be reduced) and/or when used at a time when blood glucose is already low ($<4 \text{ mmol l}^{-1}$) and there is a risk of hypoglycemia. It is important to take the precaution of measuring blood glucose immediately before all relaxation training or practice sessions. Twenty minutes of relaxation can lead to a drop in blood glucose by as much as 3 mmol l^{-1} .

Such a substantial fall is likely only if blood glucose is well above the normal range of 4–6 mmol l⁻¹ to start with and probably results from a suppression of catecholamine secretion. It is recommended that relaxation training should not be conducted at blood glucose levels below 4 mmol l⁻¹ because of the risk of hypoglycemia. Relaxation training might be continued if the anticipated reduction in blood glucose were counteracted with the pretraining intake of slow-release carbohydrate (e.g., an apple). The further precaution of measuring the blood glucose after a relaxation session, before leaving the therapist's office, is also recommended to avoid the risk of hypoglycemia while traveling home.

Several studies have reported some success with improvements in diabetes control following relaxation training (sometimes involving electromyograph, EMG, biofeedback) in people with type 1 diabetes. There is evidence to suggest that relaxation training is least useful for those people whose glycogenic control is good to start with and is most useful when used by people who not only have poor control of their diabetes but who feel that stress disrupts their diabetes control and who are currently experiencing stressful events. Although some studies have found no significant differences between relaxation and control groups in HbA1c levels or insulin requirements after relaxation training, information about the variability of the individual patients was rarely reported. In addition, the lack of selection of patients likely to benefit from relaxation training may be responsible for the nonsignificant findings of some studies.

What is emerging from the studies of relaxation training in people with type 1 diabetes is that it is most demonstrably useful for those who are showing stress-related disturbance of blood glucose control at the time. An assessment of patients' stress reactivity before embarking on relaxation training in future studies will help in the identification of patients most likely to show benefit from relaxation training. To date, most studies have incorporated biofeedback in the relaxation training. Relaxation instructions alone would be more practical and less costly. The evaluation of low-technology forms of relaxation training in patients shown to be stress reactive would be a useful contribution to the state of present research. Such an evaluation needs to take account of the fact that the benefits of relaxation training for blood glucose control will only be apparent during periods of stress when stress-reactive relaxation-trained individuals are expected to have more stable blood glucose control than stress-reactive individuals who have not been trained. Improvements in blood glucose control will be more apparent if recruitment

to such a study is restricted to individuals who show increased blood glucose under stress.

Further Reading

- Aikens, J. E., Wallander, J. L., Bell, D. S. H., et al. (1994). A nomethic-idiographic study of daily psychological stress and blood glucose in women with Type 1 diabetes mellitus. *Journal of Behavioural Medicine* 17, 535–548.
- Bradley, C. (1988). Stress and diabetes. In: Fisher, S. & Reason, J. (eds.) *Handbook of life stress, cognition and health*, pp. 383–431. Chichester, UK: John Wiley.
- Bradley, C. (1994). *Handbook of psychology and diabetes: a guide to psychological measurement in diabetes research and practice*. Chur, Switzerland: Harwood Academic.
- Bradley, C. and Gamsu, D. S. (1994). Guidelines for encouraging psychological well-being: report of a working group of the World Health Organization Regional Office for Europe and International Diabetes Federation European Region St. Vincent Declaration Action Programme for Diabetes. *Diabetic Medicine* 11, 510–516.
- Bradley, C., Pierce, M., Hendrieckx, C., et al. (1998). Diabetes mellitus. In: Bellack, A. S. & Hersen, M. (eds.) *Comprehensive clinical psychology* (vol. 8), pp. 277–304. Oxford: Pergamon.
- Feinglos, M. N., Hastedt, P. and Surwit, R. S. (1987). Effects of relaxation therapy on patients with type 1 diabetes mellitus. *Diabetes Care* 10(1), 72–75.
- Halford, W. K., Cuddihy, S. and Mortimer, R. M. (1990). Psychological stress and blood glucose regulation in type 1 diabetic patients. *Health Psychology* 9, 516–528.
- Kramer, J. R., Ledolter, J., Manos, G. N., et al. (2000). Stress and metabolic control in diabetes mellitus: methodological issues and an illustrative analysis. *Annals of Behavioral Medicine* 22, 17–28.
- Lehman, C., Rodin, J., McEwen, B. S., et al. (1991). Impact of environmental stress on the expression of insulin-dependent diabetes mellitus. *Behavioral Neuroscience* 105, 241–245.
- Lloyd, C., Smith, J. and Weinger, K. (2005). Stress and diabetes: a review of the links. *Diabetes Spectrum* 18, 121–127.
- Riazi, A., Pickup, J. and Bradley, C. (2004). Daily stress and glycaemic control in type 1 diabetes: individual differences in magnitude, direction, and timing of stress-reactivity. *Diabetes Research and Clinical Practice* 66, 237–244.
- Robinson, N. and Fuller, J. H. (1985). Role of life events and difficulties in the onset of diabetes mellitus. *Journal of Psychosomatic Research* 29, 583–591.
- Shillitoe, R. (1994). *Counselling people with diabetes*. Leicester, UK: British Psychological Society.
- Stabler, B., Morris, M. A., Litton, J., et al. (1986). Differential glycaemic response to stress in type A and type B individuals with IDDM. *Diabetes Care* 9, 550–551.
- Surwit, R. S. and Bauman, A. (2004). *The mind body diabetes revolution: a proven new program for better blood sugar control*. New York: Free Press.

Diet and Stress, Non-Psychiatric

J Wardle

University College London, London, UK

E L Gibson

Roehampton University, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J Wardle, volume 1, pp 694–698, © 2000, Elsevier Inc.

Introduction

Animal Research into Stress and Eating Behavior

Human Research into Stress and Food Intake

Mechanisms Relating Stress to Eating

Summary

Glossary

<i>Dieting</i>	Deliberate modification of food intake in order to lose weight.
<i>Disinhibition</i>	Acute loss of cognitive control over food intake.
<i>Emotional eating</i>	Eating in response to emotional cues.
<i>Hyperphagia</i>	Eating more than usual.
<i>Hypophagia</i>	Eating less than usual.
<i>Restrained eating</i>	The chronic tendency to try to exert control over food intake. It was first measured with the Restraint Scale, which in addition to items reflecting attempts to diet, included loss of control over eating and weight fluctuation. More recent measures of restraint have placed greater emphasis on the dieting component, conceptualizing the episodes of loss of control as forms of disinhibition of restraint, rather than part of the core concept.

Introduction

There is a widespread belief that stress influences eating behavior, but considerable uncertainty about the direction of the effect, i.e., whether stress increases or decreases food intake. An analysis of the physiology of stress would lead us to expect decreased eating, because stress slows gastric emptying and increases energy substrate levels, which should reduce appetite. Stress also promotes behaviors designed to cope with or escape from the source of stress, and under these circumstances, eating might be accorded a lower priority. In contrast, the literature on human reactions to stress suggests that it can be associated with increased food intake or a shift toward a higher fat diet.

Our knowledge of stress and eating comes from diverse areas of research, including clinical studies on the etiology of obesity and eating disorders, laboratory studies on dieting and the regulation of food intake, surveys of stress and health-related behaviors, and animal work on stress responses. The diversity of this literature means that there is great variation in the sources and intensity of the stressors that have been examined, the circumstances in which food is consumed, the quality of dietary information that can be gathered, and the populations that have been studied. On this basis, it may not be surprising to find that the links between stress and eating appear to be complex.

Animal Research into Stress and Eating Behavior

Animal research appears to offer the ideal opportunity to examine the general, physiologically mediated effect of stress on food intake. Animals can be exposed to stress or a control condition; stressors can vary in quality, intensity, or duration; the animal's deprivation state can be manipulated; and the type of food supplied can be varied. However, to date, this research has produced an extraordinarily inconsistent set of results. Much of the early work used rats as the experimental subjects and tail pinch as the stressor and typically found that food intake was increased by this procedure. However, tail pinch is now thought to represent an atypical stressor, if it is even stressful at all, and increased eating may be part of a general increase in oral behaviors, rather than a specific behavioral response. Other stressors have included electric shock, noise, immobilization, isolation, housing changes, cold water swims, and defeat in fights, with varied effects on food intake from study to study, ranging from substantial increases in food intake in some studies to substantial decreases in other studies. Attempts to relate the direction of the effect on eating to the characteristics of the stressor have so far been unsuccessful, partly because many of the stressors have been used in only a handful of studies, whereas electric shock, which has been used often, shows inexplicable variation. To date, the only possible consistencies are that chronic stressors appear to be more likely to have a hyperphagic effect and social stressors may be more hyperphagic than physical stressors.

Recent animal research has provided some more fruitful results by considering individual differences in susceptibility to stress. For example, rats that are resistant to dietary-induced obesity seem to be

hyperresponsive to stress, at both the neurohormonal and behavioral level, compared to those rats that become obese when allowed to overeat. Intriguingly, it has recently been shown that it is the ability to choose from separate sources of energy-dense palatable foods that confers stress resilience, rather than an increase in energy-dense food intake per se.

Human Research into Stress and Food Intake

Human research on stress and food intake has come from several research traditions. Naturalistic studies usually investigate community samples and attempt to gather information on food intake at high- and low-stress periods of life. The focus in these studies, as in most of the animal research, has been on the general effect of stress on eating over the short or medium term. Laboratory studies usually administer a stressful procedure in parallel with an eating task, then covertly assess food intake. The emphasis therefore is on the acute effects of stress. Most of the laboratory studies have taken an individual differences perspective, examining stress-related eating in relation either to weight or dietary restraint.

Naturalistic Studies

There is a variety of naturally occurring, but predictable, stressful circumstances that can provide an experimental context in which to study stress-related variations in diet. School or university examinations have been used as the stressor in several studies, and periods of high work stress in others. Food intake is recorded either in a diary record kept by the study participants or with a 24-h recall procedure. The food records are then analyzed to obtain estimates of intake of energy and nutrients. Biological measures such as weight or blood lipids have been included in some studies to provide another indicator of dietary change.

Among the published studies there is one element of consistency, namely, none of them has found lower food intake, on average, during the higher stress time. Likewise, there is little evidence for lower weight or lower serum cholesterol under stress. Otherwise, the results have been divided. In some of the studies, both overall energy intake and intake of fat are higher at the high-stress time. In other studies, there is enormous individual variability, but no average difference in energy intake between low- and high-stress periods. In a recent study, mood and eating were sampled at random intervals in students awaiting an exam in a few days. They reported more negative moods, together with an increase in reported eating for

distraction, relative to an exam-free control group. However, no differences were found for food intake or choice. In another study, which unlike many naturalistic studies was not confounded by the passage of time, high workload in department store workers was associated with higher fat, sugar, and total energy intake, but only in people who habitually restrained their food intake.

In a recent survey of a large number of U.S. teenagers, depressive symptoms were associated with perceived barriers to healthy eating, meal skipping, and more disordered eating, although the only significant change in diet appeared to be increased consumption of soft drinks in more depressed school children. In a Finnish population-based study of adults, stress-driven eaters ate more energy-dense high-fat foods and had higher body mass indices. Other studies have suggested a link between stressful life events and higher body weight. However, this does not necessarily imply increased eating, since reduced physical activity may also be an important contributor. For example, parents of a child recently diagnosed with cancer gained more weight over 3 months compared to parents with a healthy child. However, since they also reported both less physical activity and lower energy intake, it is likely that most of the weight gain was attributable to a stress-induced reduction in activity.

The balance of evidence so far has to be that stress in everyday life rarely appears to have a hypophagic effect, and it may, in some people and some circumstances, have a hyperphagic effect.

One explanation for the variability in results is the poor validity and reliability of measures of food intake. Keeping a food diary is an onerous task, so compliance is poor, and even among those who return a diary there are likely to be errors: people forget foods, are not able to assess quantities accurately, and often complete the diary well after the eating event. Diary keeping is also reactive, i.e., it is likely to influence the behavior being monitored, because of either the increased attention or the wish to present a good record to the experimenter. Unscheduled 24-h recalls avoid the problem of reactivity, but are still susceptible to forgetting and misreporting, as well as providing a much more restricted snapshot of food intake. Methodological problems also arise from translating typical information on food intake (e.g., "I had a medium-sized cheese sandwich, but I am not sure if the bread was spread with butter or margarine, or what type of cheese it was") into any meaningful value for energy and nutrient intake. Even expert dietitians are reduced to estimating and guessing a good deal of the time. The poor data quality means that any results must be evaluated cautiously, and

results from small samples could easily give a misleading impression.

A second explanation for variability in the results is the different kinds of stressor. The animal literature suggests that physical stressors might be more likely to elicit hypophagia, and social stressors, hyperphagia. In human experiences, the stress of surgery for a hernia, which was used in one study, is qualitatively and quantitatively different from the stress of an examination or a period of long working hours used in others, and this difference could well have an effect on food intake. Some stressors could occur together with changes in circumstances that enforce a change in eating behavior even if there is no direct effect of the stress on appetite. However, at present there are too few human studies to make any systematic analyses of the differential effects of different types of stressor.

The third explanation is that individuals might vary in their response to stress, and the admixture of response types could differ from study to study. Few naturalistic studies have taken an individual difference perspective, but those that do indicate that there are important individual differences that need to be considered in examining the effects of stress. In one study, middle-aged men and women kept daily stress diaries over several weeks and at the same time recorded whether, on that day, they had eaten more, the same, or less than usual. The pattern of results showed that most individuals were consistent in their reactions to stress, with some consistently eating more on higher stress days, and others consistently eating less. Over the sample as a whole, the hypophagic response predominated. Similarly, in surveys that ask respondents about how stress affects their eating, the majority report some effect of stress, with approximately equal proportions saying that they eat more and eat less while under stress.

Laboratory Studies

Most laboratory studies have been based on the idea that the biologically natural response to stress is hypophagia, but that individuals who are either overweight or highly restrained eaters are unresponsive to their internal signals. Participants are therefore characterized according to these features and hypothesized to respond differently to stress.

Laboratory studies have used a range of stressors, including unpleasant films, false heart rate feedback, and threat of public speaking. In the typical design, participants are exposed either to the stressor or to a control procedure, and food intake is assessed covertly, often disguised as a taste test in which participants are asked to taste and rate some flavors of a palatable

food such as ice cream. The amount that is eaten is recorded accurately by preweighing then reweighing the food containers.

Obesity and stress-induced eating The clinical interest in stress-related eating stemmed from the psychosomatic theory of obesity, which suggested that for the obese, eating met emotional rather than nutritional needs, and that the tendency toward emotional eating explained why they had become obese in the first place. Eating was hypothesized to provide reassurance, and hence stress was predicted to trigger higher-than-usual food intake, so-called emotional eating. There is no doubt that many obese people report that they eat more under stress, but these clinical reports need to be examined in controlled studies, both to establish their validity and to see whether stress-related eating is specific to obesity.

In practice, laboratory studies on stress-related eating in the obese have produced mixed results, with some finding higher intake under stress in the obese and others not. There has also been variability in naturalistic studies, but on balance, there is probably enough evidence to conclude that obesity is an indicator of a higher risk of stress-induced hyperphagia, or at least a lower likelihood of stress hypophagia.

The other side of the psychosomatic theory was that eating would have an anxiolytic effect, and this has received less support. Clinically, many obese people admit that any solace they derive from eating is transient and rapidly followed by shame and regret at not having showed more self-control. In laboratory studies there has been no evidence that eating successfully reduces emotional arousal. These observations have largely discredited the basic idea of the psychosomatic theory, the idea that obesity represents a disorder of emotional reactions. However, there is still interest in the idea that the obese have a reduced sensitivity to internal satiety cues, and the fact that they do not show stress hypophagia may be a consequence of this lack of internal responsiveness.

Stress, restraint, and emotional eating In 1972, a radical alternative to the psychosomatic theory of obesity was proposed, namely, that any abnormalities of eating observed in the obese were not pre-existing tendencies, but a consequence of the steps that they were taking to reduce their weight. At first the emphasis was on the effects of maintaining a body size below the hypothesized set-point. This was superseded by Restraint Theory, which proposed that one of the important determinants of food intake regulation was the tendency toward restrained eating – the combination of concern about eating and weight fluctuation. Most obese people and a significant proportion of normal-weight adults, particularly women,

are constantly trying to restrict their food intake in order to reduce their weight. The habit of trying to restrict food intake could have the effect of changing people's relationship with food, such that the usual cues to hunger and fullness become less effective in regulating their eating behavior. Restrained eaters were found to limit their food intake at times when external or emotional pressures were low, but at other times, they would abandon restraint and eat to capacity, so-called disinhibition.

One of the early observations on restrained eating was that, among individuals who experienced anxiety in response to a stressor, food intake was increased among the restrained eaters and decreased among unrestrained eaters. This general pattern, involving an interaction between stress and restraint in predicting food intake, has proved extremely robust. Across many different studies, with a wide range of stressors, restrained eaters almost always eat more in the stressed than in the unstressed condition. In contrast, unrestrained eaters show more variation, sometimes eating the same amount and sometimes eating less while under stress. The results of these laboratory studies have been strikingly consistent in an area in which most of the work is notable for its inconsistency.

There is now a growing interest in examining the role of restraint in predicting individual differences in responses to stress in real-life studies, and so far the results are consistent with the laboratory studies in showing that restrained eaters are more likely to report stress-induced hyperphagia, whereas unrestrained eaters are more likely to report hypophagia. These findings are supported by results from quantitative studies using psychometric measures of restraint and emotional eating, which found that higher restraint was associated with higher levels of emotional eating.

Recently there has been a resurgence of interest in the role of negative affect and emotional eating as predictors of problematic eating and poor control of weight. One reason for this has been the realization that earlier psychometric measures to some extent conflated restrained and emotional eating. Dietary restraint was measured in most studies by the Restraint Scale, which includes the tendency for eating to be disinhibited by emotional states. In one laboratory study, when separate scales were used to measure restraint and emotional eating (Dutch Eating Behaviour Questionnaire), the latter was the better predictor of stress-induced eating. That study tested the effects of public-speaking stress on intake of a variety of foods from sweet, salty, and bland taste categories, and in addition, high- and low-fat examples within those sensory groups. Analyzing effects by taste category may be important because stress and mood have been shown to affect taste perception. The food was

presented as a buffet-style meal during preparation of the speech task. Stress did not alter overall intake; however, stressed emotional eaters ate more sweet, high-fat foods (chocolate and cake), and a more energy-dense meal, than either unstressed emotional eaters or nonemotional eaters in either condition. This supports the survey findings that sweet, fatty foods such as chocolate may be preferentially sought during stress or negative affect, at least in a subgroup of susceptible individuals (see next section).

Emotional eating may underlie previous reports that dietary restraint or female gender predicts stress-induced eating. Moreover, emotional eaters may be more susceptible to effects of stress: women who ate more from a selection of snack foods after a stressful task also showed the greatest release of the stress-sensitive hormone, cortisol, and more stress-induced negative affect. These high reactors also showed a preference for sweet foods. So it seems that emotional eaters may be more likely to experience mood disturbance when challenged.

Emotional and disinhibited eating and dietary restraint could interact: the latter seems to contribute to stress-induced eating of sweet, fatty foods by women classified as disinhibited eaters. Recently, it was shown that highly restrained/low emotional eaters ate more chocolate after both ego-threatening and cognitively demanding tasks, whereas low restraint/high emotional eaters ate more chocolate only after the ego threat.

Emotional responses to food may also be sensitive to expectations. In a laboratory study, women were asked to rate various emotions immediately after eating small amounts (5 g) of nine different foods, three being low in energy, three medium, and three high in energy (in counterbalanced order). Intensity of negative moods (sad, ashamed, anxious, sleepy) increased with increasing energy density of the foods, and more so for overweight than for normal-weight women. Moreover, medium- and high-energy foods were rated less healthy and more dangerous than low-energy foods. These effects were independent of rated pleasantness of the foods. It is probable that these effects were psychological rather than physiological in nature, given the small amounts of food eaten and the immediacy of the ratings. The negative effects of the high-energy foods presumably reflect concerns about their impact on health and weight gain. Interestingly, though, stronger increases in negative mood were seen for women reporting greater tendencies to eat in response to emotional state.

These results are similar to the finding that self-identified chocolate addicts felt guiltier after eating chocolate than a control group. The chocolate addicts

also reported worse mood before eating the chocolate. By contrast, in another study in healthy men, experimental induction of sadness decreased appetite, whereas when the men were cheerful, the chocolate tasted more pleasant and stimulating and more of it was eaten. This gender difference is likely to be confounded by dispositional and attitudinal differences.

Stress and Food Choice

Human studies on stress and meal patterns/food choice Studies of stress and eating usually focus on the amount of food consumed, but stress might also affect food choices or meal patterns. In the naturalistic studies that showed increased energy intake, there was also an increase in the proportion of energy from fat. This could reflect an increased preference for fatty foods, or alternatively a different meal pattern. In most Western countries the proportion of fat in meal-type foods tends to be lower than in snack-type foods. If stress induced a shift from meals to snacks, this would be reflected in a higher fat intake and might also increase the total energy intake, since higher fat foods have a higher energy density.

In most naturalistic studies, it is not possible to distinguish meal pattern changes from food choice changes, because the data are presented in terms of daily intake of nutrients, without any information either on foods eaten or timing of consumption. Most laboratory studies have also failed to address the food choice issue, because they present only a single type of food. However, there is some evidence that sweet, high-fat foods are the ones most likely to show an increase (see previous section), whereas fruit and vegetables appear to be avoided during stress. In a study of health behaviors in 11- to 13-year-old schoolchildren in London, greater perceived stress was associated with more fatty food intake, less fruit and vegetable intake, more snacking, and a reduced likelihood of daily breakfast consumption.

A survey of possible changes in food choice during stress among 212 students revealed an interesting pattern of effects of stress, which was partly independent of whether participants were grouped as reporting eating more, the same, or less overall when stressed. That is, sweets and chocolate were reported to be eaten more under stress by all groups, even those eating less overall; conversely, intakes of fruit and vegetables, and meat and fish, were reported as less or unchanged under stress in all groups. The changes for the staple food, bread, matched the overall group self-perceptions of changes in eating due to stress. These data imply that mechanisms governing effects of stress on food choice may be somewhat separate from those influencing overall appetite under stress

and that foods such as sweets and chocolate may be particularly selected in stressful circumstances.

Chocolate: A unique mood-enhancing food? Chocolate does seem to have a special place in the relationship between food choice and mood: it is a widely liked food, and people often report choosing to eat chocolate during stress. Indeed, experimenters frequently choose chocolate or chocolate-containing foods as the test foods in studies looking at stress-induced eating, reflecting a cultural awareness of its value as a mood enhancer. There are popular notions, perpetuated both in the lay and scientific media, that chocolate contains psychoactive chemicals, such as phenylethylamine and cannabinoids, that are responsible for its effect on mood. However, for reasons of quantity and availability to the brain, these compounds are not likely to alter mood when obtained from chocolate. Nevertheless, it has recently been shown that the methylxanthines in chocolate, caffeine, and theobromine can indeed improve mood (hedonic and energetic dimensions) as well as mental function. Perhaps the effects of these compounds underlie a curious recent report showing that mothers who said they ate chocolate every day during pregnancy subsequently rated their babies' temperaments more positively when aged 4–9 months. These mothers did not appear to have distinctive personality traits that might otherwise explain the rating difference.

Chocolate has other characteristics that could be mood enhancing. It is high in fat and sugar and so would be likely to activate opioidergic hedonic and dopaminergic reward pathways (see next section). It is also low in protein (3–6% of energy), so if eaten in sufficient amounts on an empty stomach might conceivably enhance mood via increased synthesis of the brain neurotransmitter serotonin, as outlined in a later section. Finally, chocolate could also acquire secondary reinforcing or mood-enhancing properties by association with its use as a treat or gift and convivial social interaction. However, it should be remembered that in those who may consider chocolate to be a dangerous temptation that threatens their weight control, eating it can have negative consequences for mood.

Mechanisms Relating Stress to Eating

Neurohormonal Pathways

Eating food activates brain pathways involved in reward, and so might be expected to have a positive effect on mood. Furthermore, some of the ingested nutrients could alter the function of neurohormonal systems involved in coping with stress. We now

consider the main pathways that appear to underlie such effects (see **Food Intake and Stress, Human; Neuroendocrine Systems**).

Opioids Endogenous opioid neuropeptides are released during stress and are known to be important for adaptive effects such as tolerance of pain. They are also involved in motivational and reward processes in eating behavior, such as stimulation of appetite by palatable foods. This would tend to suggest that there would be a link between opioid action, stress, and food choice. Perhaps the best evidence for opioid involvement in an interaction between stress and eating is the finding that, in animals and human infants, ingestion of sweet and fatty foods, including milk and even a nonnutritive sweetener, alleviates crying and other behavioral signs of distress. This stress-reducing effect of sweet taste has been shown to depend on opioid transmitter systems.

In human neonates, sucking on a pacifier (dummy) or brief exposure to breast milk is also very effective in reducing signs of pain or spontaneous crying. Interestingly, as babies grow older, sweet taste becomes less effective at calming than pacifier sucking, which might reflect a maturational separation of taste and emotion or a difference in opportunities to learn the instrumental emotional value of the two experiences.

The extent to which adults (whether rat or human) retain an analgesic effect of sweet taste remains controversial. Using the cold pressor test (holding the hand in very cold water), pain threshold latency was extended by concurrent sweet taste in 8- to 11-year-old children. In adults, pain tolerance, but not threshold, was increased by prior tasting of a sucrose solution. However, perceived palatability (liking) may be important, since in one study it was found that only highly liked foods (chocolate chip cookies), and not bland or disliked foods, were able to increase pain tolerance in female students when eaten beforehand. It is tempting but speculative to conclude that adults may select sweet, fatty, palatable foods for opioid-mediated relief of stress.

Dopamine The neurotransmitter dopamine is thought to be the major chemical messenger for reward in the brain. The availability of dopamine receptors (D_2) in part of the brain reward pathway, the striatum, has been shown to be inversely correlated to body mass index. This finding has led to the suggestion of a neurochemical trait that elicits overeating of palatable foods so as to enhance dopamine release – a neural predisposition for comfort eating. In support of this, energy-dense snack foods appear to reinforce greater effort to obtain them in obese than in non-obese women. Young children of obese parents

also show greater enjoyment of food as well as higher preference for high-fat energy-dense foods than offspring of non-obese parents. It may also be relevant that binge eating and other eating disorders are associated with greater risk of substance abuse, as well as susceptibility to negative affect, possibly linked by the fact that dopamine also mediates stress sensitivity, depression, and reinforcement of drug-taking habits.

Serotonin The neurotransmitter serotonin has been implicated in mood disorders, anxiety, and stress coping, as well as in eating behavior. Synthesis of serotonin (or 5-hydroxytryptamine [5-HT]) depends on dietary availability of the precursor essential amino acid, tryptophan (TRP), due to a lack of saturation of the rate-limiting enzyme, tryptophan hydroxylase, which converts TRP to the intermediate compound 5-hydroxytryptophan.

An important complication is that TRP competes with several other amino acids, the large, neutral, primarily branched chain amino acids (LNAA), for the same transport system from blood to brain. If the protein content of a meal is sufficiently low, such as 5% or less total energy as protein, then relatively few amino acids will be absorbed from the food in the gut. At the same time, insulin will stimulate tissue uptake of competing amino acids from the circulation, and the plasma ratio of TRP to those amino acids (TRP/LNAA) will rise, favoring more TRP entry to the brain. Conversely, a high-protein meal, which would be less insulinogenic, results in absorption of large amounts of competing amino acids into the blood, especially the branched-chain amino acids, leucine, isoleucine, and valine. On the other hand, TRP is scarce in most protein sources and is readily metabolized on passage through the liver; thus, the plasma ratio of TRP to competing amino acids falls after a protein-rich meal.

The possibility that a carbohydrate-rich, low-protein meal could raise 5-HT function gave rise to the proposal that some depressed people may self-medicate by eating high proportions of carbohydrate. It was suggested that this would lead to increased 5-HT release in a manner reminiscent of antidepressant drugs, which enhance aspects of 5-HT function by inhibiting removal of 5-HT from the synaptic cleft between nerve cells. For the most part, early behavioral and pharmacological evidence for such a phenomenon was not very convincing.

Nevertheless, recent research provides some further support for beneficial effects of carbohydrate-rich, protein-poor meals on mood and emotion in some people. When participants were divided into high or low stress-prone groups, as defined by a questionnaire measure of neuroticism, carbohydrate-rich,

protein-poor meals (which raised plasma TRP/LNAA ratios) prior to a stressful task were found to block task-induced depressive feelings and release of the glucocorticoid stress hormone cortisol, but only in the high stress-prone group. This finding was replicated using high- vs. low-TRP-containing proteins (α -lactalbumin and casein, respectively). It was argued that because stress increases 5-HT activity, the poor stress coping of this sensitive group might indicate a deficit in 5-HT synthesis that is improved by this dietary intervention.

Neuroendocrine systems A critical neurohormonal system mediating the effects of stress on energy flux and appetite is the hypothalamic-pituitary-adrenocortical (HPA) axis. Neural stimulation of selected nuclei in the hypothalamus of the brain leads to corticotropin-releasing hormone (CRH) acting on the pituitary gland to release adrenocorticotrophic hormone (ACTH), which in turn stimulates release of the steroid glucocorticoid hormone cortisol from the adrenal glands. In animals, CRH is known to mediate the suppression of food intake by severe stressors. On the other hand, exogenous glucocorticoids stimulate appetite by acting on the hypothalamus to inhibit CRH release.

The involvement of the HPA axis in the influence of stress on eating behavior seems likely for several reasons, including (1) glucocorticoid hormones substantially affect energy substrate mobilization and metabolism, increasing lipolysis, proteolysis, and gluconeogenesis while protecting hepatic glycogen stores; (2) the circadian rhythm of HPA axis responsiveness is dependent on patterns of food ingestion; (3) normal meals containing at least 10% protein as energy activate the HPA axis and release cortisol in humans; and (4) acute food deprivation suppresses the HPA axis response to stress.

Furthermore, most animal models of obesity, genetic or otherwise, depend on an intact HPA axis, and disruption of its function may underlie at least some human obesity, in particular increased visceral adiposity, especially among obese patients suffering from anxiety and depression. The eating disorders bulimia and anorexia nervosa are also associated with hypercortisolemia and disrupted HPA axis activity even in the absence of undernutrition. An inability to cope with stress, coupled with chronic activation of appetite-suppressing CRH, is one psychobiological model proposed for anorexia nervosa.

Another hormone that may be involved in integrating HPA axis activity and food intake is the adipocyte hormone leptin. Leptin is synthesized in adipose tissue, and plasma levels correlate strongly with fat mass in humans, with consistently higher levels in

women than in men independent of body composition. In genetic rodent models, obesity has been shown to be associated with the absence of either leptin (ob/ob mice) or its receptor (db/db mice; fa/fa rats). Glucocorticoids administered to humans cause gradual and sustained increases in plasma leptin, and leptin in turn seems to inhibit the activity of the HPA axis. Indeed, normal release of leptin and activity of the HPA axis are inversely related. Leptin may contribute to the regulation of energy balance in part by inhibiting activity of neuropeptide Y, which is known to stimulate feeding via an action in the hypothalamus. However, acute physiological changes in neither glucocorticoids nor short-term food intake seem directly to alter circulating leptin levels. Instead, leptin may be important for longer term regulation of food intake. Moreover, chronic stress may disrupt the balance between leptin and the HPA axis, leading to abnormal eating. The precise nature of the change in eating under chronic stress – under- or overeating – may depend on individual predispositions (see next section).

In stressed people, raised levels of cortisol contribute to abdominal obesity, which in turn promotes insulin resistance. However, insulin resistance may increase the likelihood that high-carbohydrate, low-protein foods would raise brain TRP and 5-HT levels because of increased levels of plasma free fatty acids, which, by competing for binding to albumin, could result in more unbound TRP in plasma. This might underlie recent findings that insulin-resistant people are less prone to suicide and depression, both of which are believed to be increased by low 5-HT function. However, animal studies suggest that this could also depend on normalization of HPA axis function and release of cortisol by feedback from increased abdominal adiposity and intake of high-carbohydrate, energy-dense foods, with a possible resultant improvement in mood. Similarly, in humans, patients with seasonal affective disorder (a form of depression) show increased insulin resistance in the winter, together with a greater predilection for sugar-rich foods, both of which might help alleviate their mood symptoms. Unfortunately, despite this possible protective effect, insulin resistance is a substantial risk to health by promotion of cardiovascular disease.

Biobehavioral Pathways

Theoretical considerations and empirical results indicate that the mechanisms linking stress to changes in food intake are far from straightforward. **Figure 1** illustrates the three principal pathways that have been implicated in the work in this area. In the center is the simple biological pathway whereby

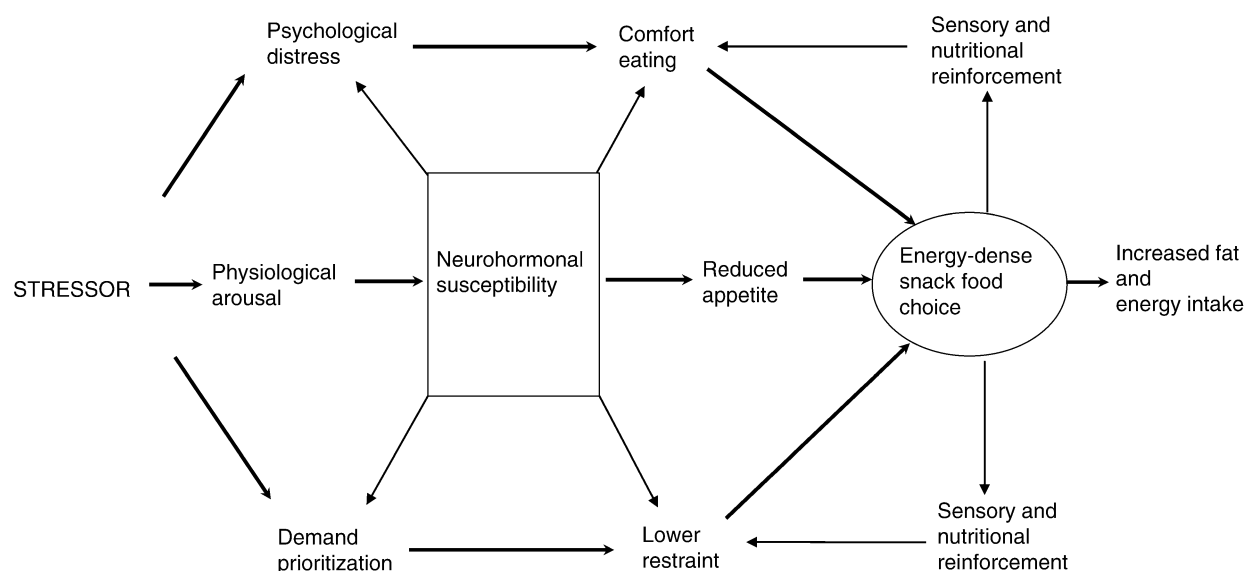


Figure 1 Biobehavioral pathways linking stress and diet.

the physiological effects of stress inhibit the physiological features of hunger (or perhaps mimic the effects of satiety) and hence modify appetite and food intake. Animal research has generally been used to test this pathway since it is assumed that animals' food intake is more directly controlled by the basic drives of hunger and satiety. From present evidence it would seem that some stressors probably induce hypophagia and others hyperphagia, implicating a more complex pathway from stress to food intake even in animals. Humans have the advantage as research subjects that they can tell us about their subjective experience, yet few studies have incorporated measures of hunger or satiety to address this issue. At present, therefore, there is little direct evidence for the hypothesized physiological depression of hunger. The best evidence available from human studies comes from the studies of food intake among normal-weight, nonrestrained eaters. In the laboratory, unrestrained eaters either are unaffected by stress or eat slightly less, which is somewhat supportive of a reduction in appetite, at least in some situations, and for some foods.

A second pathway (shown at the bottom of [Figure 1](#)) results from consideration of the broader effects that stressors have across a range of aspects of life. In real life, stress motivates people to deploy resources toward dealing with the source of stress (so-called problem-focused coping). This may mean that longer term goals like dieting or healthy eating are temporarily set aside, with consequent effects on food choices. Food intake should increase most in those who were normally restrictive, so women, healthy eaters, and dieters should

show the strongest hyperphagic response to stress. This idea is given some support from the higher food intake in stressed restrained eaters, the higher fat and energy intake under stress among women in some of the naturalistic studies, and the observation that energy intakes increase most among those who normally have a low energy intake.

The third pathway, shown at the top of [Figure 1](#), relates to the fact that stress also comes with an emotional coloring, so coping efforts are deployed to deal not only with the source of stress, but also with the emotional state (emotion-focused coping). For some people, especially those who are usually self-denying with food, food may have a significant reward value and may be used to provide emotional comfort. Women in particular often describe food in terms of treats and rewards, and the potential comfort to be derived from sweet foods is deeply enshrined in contemporary culture. Clearly this idea has echoes of the discredited psychosomatic theory of obesity, but it should be remembered that it was rejected not because of the absence of emotional eating, but because eating did not appear to be a successful anxiolytic. Other emotional coping strategies are equally unsuccessful in the longer term (e.g., alcohol, smoking), but that does not prevent the smokers, drinkers, and stress eaters from feeling that their habits serve a comforting role. In the field of eating disorders, excessive eating has been hypothesized to provide escape from self-awareness, which is especially valued in response to ego-threatening stressors, so the idea that eating has a role, functional or dysfunctional, as an emotional coping strategy may have

been rejected prematurely. If emotional eating is part of a stress-coping repertoire, it should be more common in restrained eaters, for whom food has a higher emotional value, and that of course is observed in all of the laboratory studies.

Figure 1 also depicts a likely final outcome of stress on diet, as discussed in the previous sections. That is, for multiple reasons, including sensory pleasure, modulation of neurotransmitter control of mood, and stress coping, susceptible stressed people may end up choosing energy-dense, sweet, fatty foods that will lead to a less healthy diet and possibly weight gain.

Finally, in real life, stresses are not usually single, transient, and novel, as they are in the laboratory. They are prolonged or repeated, they wax and wane, and they can become a familiar part of the landscape. This means that stress reactions are likely to be modified over time. Adaptational processes should be triggered to protect body fat stores; if stress induces hypophagia in the short term, then in the longer term, the loss of body fat should upregulate appetite and restore food intake to normal levels. Few human studies have looked at the effect of repeated stress exposures on eating, but animal studies show that stress-induced hypophagia diminishes with repeated stress exposures. The inclusion of this dynamic perspective adds yet another dimension to an already complex process.

Summary

In summary, the relationship between stress and eating is a complex one, moderated in humans not only by the type of person, the type of stress, and the types of foods that are available, but also by the dynamic relationship between the internal milieu and behavior over time. There is increasing evidence to support the view that some people are particularly susceptible to eating more sweet, fatty, energy-dense foods during stress. Possible explanations can be found at both the psychological and physiological level, but in either case it is probable that such a change in eating behavior helps the individual to cope with stress. Unfortunately, this change in diet may

also result in weight gain, particularly abdominal obesity, and increased risk of cardiovascular disease.

This fascinating area is likely to challenge clinicians and researchers for some time to come, but it also provides one of the richest areas for demonstrating the importance of integrating biology and psychology to understand human behavior.

See Also the Following Articles

Diet and Stress, Psychiatric; Eating Disorders and Stress; Food Intake and Stress, Human; Food Intake and Stress, Non-Human; Negative Affect; Neuroendocrine Systems; Obesity, Stress and.

Further Reading

- Cartwright, M., Wardle, J., Steggle, N., et al. (2003). Stress and dietary practices in adolescents. *Health Psychology* 22, 362–369.
- Dallman, M. F., Pecoraro, N., Akana, S. F., et al. (2003). Chronic stress and obesity: a new view of “comfort food”. *Proceedings of the National Academy of Sciences USA* 100, 11696–11701.
- Greeno, C. G. and Wing, R. R. (1994). Stress-induced eating. *Psychological Bulletin* 115, 444–464.
- Heatherton, T. F. and Baumeister, R. F. (1991). Binge-eating as escape from self-awareness. *Psychological Bulletin* 110, 86–108.
- Herman, C. P. and Polivy, J. (1984). A boundary model for the regulation of eating. In: Stunkard, A. J. & Stellar, E. (eds.) *Eating and its disorders*, pp. 141–156. New York: Raven Press.
- Markus, R., Panhuysen, G., Tuiten, A. and Koppeschaar, H. (2000). Effects of food on cortisol and mood in vulnerable subjects under controllable and uncontrollable stress. *Physiology and Behavior* 70, 333–342.
- Robbins, T. W. and Fray, P. J. (1980). Stress-induced eating: fact, fiction or misunderstanding. *Appetite* 1, 103–133.
- Rogers, P. J. (1995). Food, mood and appetite. *Nutrition Research Reviews* 8, 243–269.
- Slochow, J. A. (1983). *Excessive eating: The role of emotions and environment*. New York: Human Sciences Press.
- Stone, A. R. and Brownell, K. D. (1994). The stress-eating paradox: multiple daily measurements in adult males and females. *Psychology and Health* 9, 425–436.

Diet and Stress, Psychiatric

V March and M H Fernstrom

University of Pittsburgh Medical Center, Pittsburgh,
PA, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction

Biological Basis of Stress Eating

Practical Treatment of Stress Eating

Conclusion

Glossary

<i>Behavioral modification</i>	A specific set of counseling techniques using specific tools to change an individual's behavior by altering habitual thoughts and activities in specific situations.
<i>Body mass index (BMI)</i>	The formula of weight (in kilograms)/height (in meters) squared that is used to quantitate medical overweight and obesity. Healthy BMI = 18.5–24.9; Overweight = 25–29.9; Class I Obesity = 30–34.9; Class II Obesity = 35–39.9; Class III Obesity = 40 and above.
<i>Central adiposity</i>	Body fat concentrated in the abdomen; even in the absence of overweight or obesity, this may alone increase cardiovascular risk.
<i>Lifestyle changes</i>	The stepwise substitutions of unhealthy behaviors with more healthful ones, with the goal of improving health, including weight management.
<i>Readiness</i>	An index of individual willingness and motivation to make lifestyle changes leading to better health.
<i>Stress</i>	A biological condition altering brain and metabolic functions as the body responds to a noxious or life-threatening stimulus.
<i>Stress eating</i>	Overeating with a perceived lack of control in response to behavioral stressors; also called emotional eating.

Introduction

Over the past 3 decades, the surging numbers of overweight and obese individuals worldwide have prompted the use of the term obesity epidemic to describe this phenomenon. Not only adult but also young children and adolescents are affected, putting these populations at risk for obesity-related comorbidities, including sleep apnea, hypertension, and type II diabetes. Many hypotheses have been proposed to explain the increasing rate of obesity, including both biological

and environmental factors. Although larger portions, readily available food, and lack of activity all contribute to weight gain, the perception of a stressful environment is often raised by patients, both adults and children, as a major reason for overeating and weight gain. The term stress eating, also called emotional eating, has been used to describe eating behavior in which an individual responds to a mentally stressful situation with uncontrolled eating of high-calorie foods, even in the absence of hunger.

Stress can be loosely defined as a physical or emotional response by which the organism attempts to protect itself. A stressor, whether environmental (exogenous) or biological (endogenous), can cause certain physiological changes to occur. These changes, both biochemical and hormonal, permit an organism to adapt in survival mode. The adrenal hormones epinephrine, norepinephrine, and cortisol modulate the metabolic stress response in animals. During this fight-or-flight reaction, these hormones modify blood flow by increasing circulation to the heart, lungs, and skeletal muscle to facilitate escape while decreasing circulation to nonessential areas, including the splanchnic vasculature supplying the gastrointestinal tract. This classic stress response is not compatible with eating.

However, excessive cortisol secretion alone seems to stimulate appetite in certain individuals and has been shown to be a factor in visceral and abdominal fat distribution in humans. This is illustrated in the extreme in (1) Cushing's syndrome and Cushing's disease, disorders that both involve an overproduction of cortisol by the adrenal glands, and (2) people receiving high-dose corticosteroid therapy. Because of this pattern of appetite increase and fat deposition, some investigators have postulated that stress in animals (including humans) causes high cortisol levels, which, in turn, lead to stress-eating responses. Data supporting this hypothesis have been inconsistent, which is not surprising given the complexity of hypothalamic and gut signaling of hunger and fullness control, as well as the newly discovered role of adipocytes in appetite regulation.

In this article, we summarize the most recent research on stress eating in animal models and in humans. We also discuss environmental and psychosocial factors contributing to overeating and weight gain. Our clinical experience serves as a resource to identify practical strategies for the clinical management of stress eating in overweight and obese patients.

Biological Basis of Stress Eating

Current medical literature is replete with articles on both the psychological and physiological responses to stress. The animal literature typically exposes animals to noxious stimuli (including isolation, tail pinch, electric shock, cold water immersion) and examines food consumption. The results of these studies have been diverse. Following the stressors, both eating response and the weight of subjects have been variable, either increasing or decreasing compared to controls. A recent series of studies in rats using restraint as a stressor measured caloric intake, caloric efficiency, body weight, intra-abdominal fat, and hormonal levels including adrenocorticotrophic hormone (ACTH), insulin, and leptin. The results indicated that, although weight tended to vary after stress, the deposit of abdominal fat and the consumption of comfort foods (lard and sucrose-sweetened beverages) in proportion to chow increased. The authors suggested that the ingestion of comfort foods appeared to decrease stress hormone responses. They further speculated that the same type of response might occur in humans.

Human studies have typically addressed the interaction of stress and overeating by examining the hypothalamic-pituitary-adrenal axis. In controlled laboratory studies, stress is induced with either physiological or environmental manipulations, and hormonal measurements and/or eating behavior are measured. A blunting of the usual adrenal cortisol response to the pituitary-stimulating hormone ACTH is often reported. Other studies have indicated that laboratory stressors produced increases in serum cortisol response that could be positively correlated with the increased consumption of sweet and high-fat foods.

The discordant conclusions of both the animal and human studies have prompted more questions than answers. Nonetheless, there is general acknowledgment that stress-eating behavior exists and that it may pose a significant barrier to the treatment of obesity. If we are to treat obesity effectively, practical recommendations for treatment of stress eating are needed even before a complete understanding of the neurological and metabolic basis for this phenomenon is reached.

Practical Treatment of Stress Eating

A key element of stress eating is the consumption of food in the absence of physical hunger, often to the point of being uncomfortably full. Frequently, stress eating results in a sensation of uncomfortable fullness followed by feelings of guilt, self-hatred, and self-disgust.

The pattern of stress-related overeating is heterogeneous both among and within individuals and may

include constant grazing, binge-eating, and nighttime eating. During an episode of stress eating, significant quantities of comfort foods are consumed, uniformly described as palatable and most often sugar-fat or salty-fat combinations. Left untreated, regular stress eating may contribute to the obesity.

Following an incident of stress eating, an individual may attempt to compensate by consciously curtailing food intake. This restrained eating, in turn, may trigger more stress eating. Stress eating may be continuous over variable stretches of time or brief, depending on the intensity and duration of the stressor. Many environmental factors that increase anxiety may contribute to stress eating in both children and adults. These include frightening world events, financial worries, and occupational or domestic demands. Habitual exercise, which could help offset stress-eating through both regular calorie expenditure and stress reduction, is often avoided because of time constraints, concerns for personal safety, reductions in gym classes for youths, and increasing dependence for entertainment on more passive activities, such as television, computer and video games.

Despite the increased incidence of obesity, there remains the societal pressure to be thin. The quest for thinness at any cost abounds, but the ideal body type promoted in the media is not realistic for most people. A vicious cycle may occur when the failure of attempts at quick weight loss lowers an individual's self-esteem and increases frustration, inciting further episodes of stress eating. In order to break this cycle, active intervention is required. What is effective intervention? The health-care community must help refocus our society's expectations by helping the individual set realistic goals for achieving a healthy weight. People tend to do well when very small yet incremental behavioral changes are adopted initially. The goal of better health, not a specific weight, should be emphasized; as little as a 5% reduction in weight can improve health significantly. Motivational interviewing, a form of nonthreatening history-taking, enables the practitioner to identify unhealthful eating behaviors in a patient while simultaneously permitting a collaborative interaction in which the patient determines the course of action and the clinician serves as a facilitator. To best assist a patient, the health-care provider should become familiar with some basic behavioral modification techniques.

Overview of Lifestyle Change

Stress eating may be considered one of several types of nonhunger eating, that is, eating that is unrelated to physiological hunger. Rarely do stress-eaters complain of hunger symptoms (e.g., sweating, headache, poor

concentration, irritability, lightheadedness, racing pulse, and hypotension) prior to food consumption. In fact, most of these individuals acknowledge no feelings of hunger when feeling the urge to stress eat. Although there is no consensus about treatment, the National Weight Control Registry has been tracking a cohort of 3000 individuals who have lost at least 30 pounds and kept it off for at least 1 year. Started in 1994, the Registry has identified five behaviors contributing to long-term weight loss: (1) maintaining high levels of physical activity; (2) following a low-fat, low-calorie high-carbohydrate diet; (3) monitoring food intake and weight; (4) possessing coping skills enabling the continuation of weight-loss strategies despite stressful situations or cravings; and (5) eating breakfast. Although it is not known whether these recommendations will help for stress eating by itself, it is reasonable to use these tools to manage weight over the long term.

Overcoming Therapeutic Barriers

Many health-care professionals, even when provided with materials, do not discuss weight control with their patients. The reasons for this are unclear, but there are probably at least two: (1) The discussion of lifestyle change is time-consuming and cannot easily be accommodated by the typical 10- to 20-min patient visit and (2) clinicians may hesitate to broach the subject of weight because many of them are overweight or obese themselves, they are afraid they might insult the patient, and they may feel incompetent to discuss weight management because the medical education curricula today generally do not incorporate courses related to weight control. Nonetheless, it is quite possible to include brief patient education sessions about weight management into a primary-care office visit, especially in the context of a longitudinal clinician-patient relationship. For a clinician who cannot or will not spend time on this topic during the office visit, important weight management concepts can be assimilated into patient handouts that can be distributed without discussion or left in the examination and waiting rooms for patients to peruse at their leisure. In addition, the use of a physician-extender can be considered to help guide the weight-loss process.

Use of the Prescription Pad

The clinician's role as medical expert and prescriber of treatment can be a powerful catalyst of change. Patients are likely to take a written prescription more seriously than other modes of medical advice, including verbal instruction and guidance. The use of the prescription pad for dietary and exercise recommendations can be a very useful tool because it lends

further credibility to lifestyle change as an actual treatment recommendation. (When necessary, more traditional prescriptions for laboratory testing, weight-loss medications, and/or a surgical consultation can be added.) Moreover, the patient is more likely to remember what has been written down, particularly if there are multiple instructions.

Basic Recommendations for the Primary-Care Clinician

Recommendations, such "eat less; exercise more," although logical, are unhelpful because they are both condescending and too general. The majority of obese individuals are well aware that in order to lose weight, they must eat less and exercise more; however, they do not know how to accomplish this. Therefore, the clinician's suggestions should be specific and individualized.

A clinician can first teach a patient to become more cognizant of the physical signs of hunger and satiety by making the following suggestions.

1. Use a three-point hunger scale to rate feelings of hunger (1 = hungry; 2 = content; 3 = stuffed).
2. Determine when the desire to eat coincides with higher stress levels, with signs of physical hunger, or both.
3. Attempt to delay eating during high levels of stress, particularly if no physical hunger is sensed.
4. Eat slowly – it takes 20 min for the stomach to signal fullness to the brain.
5. After a 100- to 200-Cal snack, find a distraction to help postpone further eating for at least 1 h (encourage noneating, stress-reducing activities).
6. Repeat the phrase, "I can always have more later"; this can reassure an individual who becomes anxious if food is unavailable.
7. Avoid liquid calories; this is a crucial piece of advice because patients often ignore beverages in their calculations of daily calories consumed.

A sense of deprivation often occurs when people eliminate favorite foods from their daily diet. Historically, such foods have been judgmentally referred to as illegal, off-limits, not allowed, and even bad. The avoidance of these foods, in and of itself, can cause mental stress while the forbidden nature of these foods may increase their desirability and thus provide a powerful trigger for stress eating. The uncontrolled eating of allowed foods is a common way to overcompensate for not being permitted to eat the desired food, but this most often results in an overconsumption of total daily calories because the mental hunger has not been fulfilled. Instead, stress-eaters should be given permission to eat their favorite foods in the context of a healthful diet. If no foods

are excluded from stress-eaters' diets, compensatory overeating is more likely to be better managed. A caloric allowance for favorite foods should be considered in a day's food intake equation.

Nearly all patients are well aware that insufficient physical activity presents a significant barrier to weight control. Regular physical activity, because it can relieve stress, helps to reduce the need to stress eat. Activity also modestly increases the expenditure of calories, making long-term weight loss success more likely. After medical clearance, the exercise prescription for all sedentary individuals should emphasize a gradual increase in duration and intensity of activity. For sedentary individuals reluctant or unable to exercise, increasing the activities of daily living can provide an alternate route for augmenting calorie expenditure. Activity goal-setting can be accomplished with the use of a pedometer, a small digital instrument which, when clipped to an individual's waistband, records the number of steps taken per day. By demonstrating that simple changes can make a significant difference in overall activity level, a pedometer can also be a motivational tool. The daily step number may be increased with such modifications as parking farther away from work or the store, choosing a more distant bus stop, taking the stairs instead of the elevator, using the manual channel-changer instead of the remote, pacing while talking on the phone, and walking for 5 min during a lunch break.

Other Behavioral Techniques

Many weight management programs, whether commercial or medical center-based, apply the principles of cognitive-behavioral therapy directed at lifestyle change. In the case of lifestyle change related to weight control, these principles can help people strike a more healthful balance in their behaviors relating to both eating and activity. These include self-monitoring, bartering, cognitive restructuring, preplanning, positive self-talking, imaging, and asserting.

- Self-monitoring, one of the most important tools for the stress-eater, has been proven to lead to sustained success. The technique involves recording periodic (i.e., daily or weekly) weights, daily activity, and food intake. Despite the proven efficacy and relative ease of this method, many individuals seem to have difficulty sustaining the effort involved in logging.

- Bartering in behavioral modification terms means making choices. Substitutions, trade-offs, and banking calories are different types of bartering.

- Cognitive restructuring is a technique that allows replacing eating and rumination about food with

another behavior. As a cognitive-restructuring exercise, the patient may generate a list of stress-reducing activities. Such activities might include taking a bubble bath, talking to a friend, watching a movie, reading, getting a massage, participating in community or other volunteer work, becoming occupied with a hobby, taking a class, cleaning out drawers, shopping, or rekindling intimacy with a partner.

- Preplanning is a method of preparing for events in which stress eating is likely to occur. Through this process, many patients have been able to retrieve a sense of control over their eating.

- Positive self-talk is another effective way to interrupt the stress-eating pattern. Negative thoughts seem to exacerbate stress eating and should be discouraged; the switch to a more affirmative cognitive process often takes practice. One exercise incorporates patients' transcribing negative thoughts in one column and positive replacement statements in an adjacent column. A variation on this method is for patients to speak aloud or write down all their accomplishments, recent or past, or to congratulate themselves on accomplishments related to weight loss, for example, to say "I am a size smaller than I was 2 months ago" or "I rode the stationary bike for half an hour three times this week." Repeating these sorts of sentences may not only change the patients' thought processes but also help relieve stress, both of which will subsequently promote behavior change.

- Imaging is mental rehearsal. By envisioning an upcoming stressful scenario, individuals may also imagine an appropriate response to the stressor that can later be put into action. Another strategy is to recommend imagining a relaxing situation (e.g., lying on a beach or getting a massage) during a stressful event, a process that, through mitigating the feelings that induce stress eating, may minimize the odds of an actual stress-eating response.

- Asserting is an invaluable tool for the stress-eater. Because individuals stress eat in order to avoid confrontation, teaching strategies to gradually increase self-assertiveness are important. The phrase "I need" is central to learning self-assertiveness. For example, people can be given the homework of becoming assertive in a relatively nonthreatening situation, such as informing a demanding young child, "I need to go take a shower now, and I need privacy while I am doing this."

Understanding and Addressing Failure in a Stress-Eater

The breaking of any unhealthful habit (or addiction) carries with it the risks of lapse, relapse, and collapse.

In fact, when the habit is ingrained, lapses are almost inevitable and should be considered a normal part of the habit-changing process.

A lapse is an isolated episode of recurrence of the old habit or behavior. Lapses often occur when people have not planned for an unusual occasion, but they can happen even with the most detailed calculation. A relapse is several episodes in succession of the old pattern of behavior. Vacation commonly provokes a relapse, which, for some, is difficult to reverse. Collapse, a complete abandonment of healthful behaviors, is less likely if the stress-eater, after a single lapse, is able to get right back on the track of a palatable and manageable weight-loss plan. In the event of a lapse or relapse, the stress-eater should be encouraged to bring all behavioral methods to bear, return for weight management visits in order to get the necessary reinforcement, and garner support from friends and family.

Other Therapies: Pharmacotherapy and Surgery

Weight-loss medication should be considered a tool to make the lifestyle effort easier. It is not a replacement for lifestyle, nor should it be viewed as a cure for obesity. Some prescription medications can be used in stress-eaters who exhibit great difficulty with behavior change related to eating and activity. Medications such as sibutramine and bupropion (in depressed obese patients) exert more subtle effects on hunger and satiety and can therefore be employed as teaching devices. However, even with the medications, the signals can be ignored and an individual can eat through the effects of the medications. Therefore, the importance of a complete explanation by the practitioner about the use of these or other weight-loss medications cannot be overemphasized. Sibutramine is approved for long-term use in weight loss, and although active weight loss seems to last for approximately 6–9 months for responders, the weight maintenance effects of the drug may continue for at least 2 years. Bupropion, which is FDA-approved as an antidepressant and as a smoking-cessation adjunct – and not FDA-approved for weight loss – has sometimes been used for weight management in depressed patients. Sibutramine, used either intermittently or continuously, can help promote both weight loss and maintenance. Another medication, orlistat, a pancreatic lipase inhibitor, also has behavioral effects, but not by acting directly on the brain; this medication works by inhibiting the absorption of dietary fat from the intestine. When orlistat is taken as directed and a high-fat meal is consumed, bloating, diarrhea, flatulence, and leakage of foul-smelling stools will occur as a consequence of the decreased dietary fat absorption. The medication allows weight loss in part due

to a reduction in the absorption of calories but primarily via its role as a negative reinforcer because the consumption of high-fat comfort foods results in unpleasant side effects. Because this medication has no effect on carbohydrate or protein absorption, total calorie intake may not be reduced.

Bariatric surgery, including both the roux-en-y gastric bypass and laparoscopic gastric banding, is initially considered by many patients to be the cure for obesity. When a severely obese individual (BMI 40 or greater) has been unable to maintain weight loss even with sustained effort, weight loss surgery should be considered. For these patients, either restriction (gastric banding) or a combination of restriction and malabsorption (gastric bypass) can be the missing tool in the weight management toolbox. Obesity surgery allows the motivated patient to sustain a lifestyle effort that results in long-term weight-loss success. The lifestyle after surgery is quite demanding, and patients need to be taught that even after surgery it is possible to regain weight with poor eating habits – including grazing throughout the day and consuming high-calorie liquids. The stress-eater must be taught new behavioral modification techniques to manage stress without food, with these concepts being reinforced after surgery for at least 2–3 years.

Conclusion

Stress can produce both biochemical and behavioral responses in certain individuals, often resulting in stress eating and subsequent weight gain. Stress eating may be an important contributor to the current obesity epidemic. Although the understanding of stress eating is rudimentary at this time, certain behavioral techniques seem to be effective treatments, at least for some patients; many of these therapies are outlined in this article. These strategies can be easily initiated by a wide range of health-care providers during routine outpatient office visits and need not be time-consuming. Although all patients will not respond, many patients will benefit from a brief review of treatment options.

Changing eating behavior is difficult, but it is possible when a stress-eater is motivated to change and is willing to put in the effort. Assistance from health-care providers may facilitate lifestyle changes sufficient for many patients to achieve long-term weight-loss success. A basic understanding of simple behavioral modification will allow a more effective treatment of obesity in the primary-care setting. In certain individuals who are refractory to such methods alone, referral to a comprehensive weight management program may be necessary, particularly for the coordination of the more complex modalities

of cognitive-behavioral methods, diet and exercise education, and the judicious use of medications and surgery.

Further Reading

- Dallman, M. F., Pecoraro, N., Akana, S. F., et al. (2003). Chronic stress and obesity: a new view of "comfort food." *Proceedings of the National Academy of Sciences USA* 100, 11696–11701.
- Epel, E., Lapidus, R., McEwen, B., et al. (2001). Stress may add bite to appetite in women: a laboratory study of stress-induced cortisol and eating behavior. *Psychoneuroendocrinology* 26, 37–49.
- Habib, K. E., Gold, P. W. and Chrousos, G. P. (2001). Neuroendocrinology of stress. *Endocrinology & Metabolism Clinics of North America* 30, vii–viii, 695–728.
- Jessop, D. S., Dallman, M. F., Fleming, D., et al. (2001). Resistance to glucocorticoid feedback in obesity. *Journal of Clinical Endocrinology and Metabolism* 86, 4109–4114.
- Klem, M. L., Wing, R. R., McGuire, M. T., et al. (1997). A descriptive study of individuals successful at long-term maintenance of substantial weight loss. *American Journal of Clinical Nutrition* 66, 239–246.
- Kouvonen, A., Kivimäki, M., Cox, S. J., et al. (2005). Relationship between work stress and body mass index among 45,810 female and male employees. *Psychosomatic Medicine* 67, 577–583.
- Lichtman, S. W., Pisarska, K., Berman, E. R., et al. (1992). Discrepancy between self-reported and actual caloric intake and exercise in obese subjects. *New England Journal of Medicine* 327, 1893–1898.
- Lowe, M. R. and Levine, A. S. (2005). Eating motives and the controversy over dieting: eating less than needed versus less than wanted. *Obesity Research* 13, 797–806.
- Pecoraro, N., Reyes, F., Gomez, F., et al. (2004). Chronic stress promotes palatable feeding, which reduces signs of stress: feedforward and feedback effects of chronic stress. *Endocrinology* 145, 3754–3762.
- Stafford, R. S., Farhat, J. H., Misra, B., et al. (2000). National patterns of physician activities related to obesity management. *Archives of Family Medicine* 9, 631–638.
- Yin, Z., Davis, C. L., Moore, J. B., et al. (2005). Physical activity buffers the effects of chronic stress on adiposity in youth. *Annals of Behavioral Medicine* 29, 29–36.

Disaster Syndrome

P Valent

Melbourne, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Valent, volume 1, pp 706–708, © 2000, Elsevier Inc.

Features of Disaster Syndrome

The Place of Disaster Syndrome in Disasters and among Similar Syndromes

Sense and Purpose of Disaster and Associated Syndromes

Unfolding of Disaster Phases and of Disaster Syndromes

Treatment

Glossary

- Disaster* Natural calamities such as floods, earthquakes, and fires; also human-made calamities and other traumatic situations.
- Disaster phases* Time periods around disasters. These are (1) preimpact, the period before the disaster; (2) impact, when the disaster occurs; (3) recoil, the period immediately after impact; (4) postimpact, the

period days to weeks after impact; and (5) recovery and reconstruction, the period months or years after impact.

Biopsychosocial templates that evolved to enhance survival in specific circumstances. Examples are fight, flight, adaptation, and attachment.

An experience in which a person's life is grossly threatened and from which a variety biological, psychological, and social wounds and scars result.

External situations in which trauma occurs.

Survival strategies

Trauma

Traumatic situations

Features of Disaster Syndrome

Wallace described the responses of Petun warriors who in 1649 returned to their village to find their kin had all been slain or captured. They sat down in the snow, mute, and no one stirred or spoke for half a day. Wallace noted that such responses are ubiquitously found in disasters and are variably described as shock, stupor, and being dazed, stunned, and numbed. He added that, as well as sitting, victims may stand motionless or wander aimlessly. Emotion

such as pain, grief, fear, and anger are missing, and people are docile. The state may last up to hours or even days, injured people remaining dazed longer. The features of disaster syndrome are opposite to the anger and fear (fight and flight) sympathetic nervous system-associated arousal in posttraumatic stress disorder (PTSD) and some other anxiety disorders. This suggests that there may be a variety of emotional and biochemical survival responses in trauma that can develop into a variety of later symptoms.

The Place of Disaster Syndrome in Disasters and among Similar Syndromes

The behavioral features of disaster syndrome described by Wallace are still valid in descriptions today and are called by the original name. However, various psychological, behavioral, and physiological lines of research emphasize different aspects of what is behaviorally disaster syndrome and call them by different names. The psychological equivalent has been called psychic shock and the physiological equivalent general adaptation syndrome; conservation-withdrawal syndrome spans the all biopsychosocial aspects.

Psychic Shock

Psychic shock is the subjective view of being stunned. It has been most studied in bereavement and dying and is seen there as the first stage of the grieving process.

Subjective feelings of psychic shock as a response to being told the news of someone's death use physical metaphors such as an assault, a blow on the face, head or the guts, being knocked out, or winded. Psychological terms include being overwhelmed, not being able to take it, the world shattering around one, and sinking into a black hole.

Defenses against such sensations include denial voiced as disbelief and dissociation, described variably as feeling numb or as experiencing a sense of unreality of the world or of oneself. A common description is looking at oneself and the events as if they were a film.

General Adaptation Syndrome

Selye described general adaptation syndrome as a ubiquitous triad of physiological stress responses that were associated with psychic shock. The stress responses are the enlargement of the adrenal cortex and increased levels of cortisol; shrinking of the thymus gland, spleen, and lymphatic structures (along with compromising immunocompetence); and deep

ulcers in the stomach. Selye noted that these responses help organisms to adapt to overwhelming situations.

Conservation-Withdrawal Syndrome

Conservation-withdrawal syndrome, described by Engel and Schmale, enlarges on the features of disaster syndrome by emphasizing immobility, quiescence, and unresponsiveness, accompanied by sagging muscles, feelings of extreme fatigue, constricted attention, and detachment.

Adding to general adaptation syndrome, conservation-withdrawal is associated with parasympathetic, trophotropic activity, which may manifest in diminished heart rate to the point of arrhythmia and possibly asystole and even death. Gastric secretions have been noted to diminish, possibly because food intake does not occur in this state. The syndrome has been observed at all ages, ranging from neonates to adults, and in various situations in which no definite course of action was possible.

Sense and Purpose of Disaster and Associated Syndromes

Initially it may be difficult to see these responses of being overwhelmed as adaptive. However, just as physical shock helps survival, so may psychosocial shock. For instance, Darwin noted that immobility made animals less likely to be seen and be attacked and that the deathlike state could stimulate predators to release their grip. Immobility could also help others to scoop the unresisting person to safety and help members of the group to find its previously abandoned members. Further, the associated psychosocial limpness or docility may facilitate cooperation with helpers and their directions. Last, a type of psychological encystment and conservation of energy could provide a buffer space for the replenishment of reserves.

Selye and Engel and Schmale also saw their respective syndromes as primary regulatory organismic templates that enhanced the survival of the species. They were at the opposite end of the spectrum to fight and flight, and they occurred in situations in which fight and resistance could be fatal but in which rolling with the punches and surrendering old goals and finding new ones was advisable.

Valent integrated the biopsychosocial features and purposes of these syndromes into a survival strategy that he called adaptation. The strategy adapts the organism to situations requiring initial surrender, but through processes such as grieving it facilitates the organism's turning to new hopeful situations and bonds. If this does not occur, physiological symptoms, depression, and other illnesses can occur.

Unfolding of Disaster Phases and of Disaster Syndromes

The features of later disaster phases and the unfolding of the disaster syndromes (or the adaptation survival strategy) follow.

Ensuing Disaster Phases

In the postimpact phase (Wallace's stages 2 and 3) of disasters, survivors emerge from their cocoons. They are grateful to be alive, help others, and reconnect with family and friends. Survivors drop their usual reserves to become a cohesive altruistic community. The accompanying optimism has been called the post-disaster euphoria. At the same time, anger is often directed at outsiders whose help is seen as unempathic. In addition to Wallace's original stages are the recovery and reconstruction phases, hard prolonged times of rebuilding the physical environment and internal lives. The names, timing, and contents of the phases are flexible, indicating fluctuating progression.

Similar phases occur in other traumatic situations, although different contents are emphasized in the literature. For instance, in bereavement psychological features are emphasized. After shock come the phases of searching, then anger, guilt, and working through, followed by acceptance and turning to new bonds.

Unfolding of Disaster Syndrome

Physiological stress response derivatives of disaster syndrome (or adaptation survival strategy) may manifest clinically as hypotension, dizziness, tiredness, fatigue, and the sense of being ill. Prolonged compromised immunocompetence may help to account for the increased rates of infections, autoimmune diseases, and cancers following disasters, bereavements, and other traumas. Thus, this specific response may be

significantly involved in the frequent stress-induced maladies. Maladaptive psychological unfolding includes hopelessness, despair, giving in, unresolved and chronic grief, and clinical depression.

Treatment

The treatment of psychic shock, like physical shock, includes physical and psychological warmth, comfort, and support. Physical contact with another human, a reassuring voice, explanations of what is happening, hope, and connection with loved ones are helpful. Later treatment depends on the nature of the evolving symptoms and illnesses. The best treatment is prevention or mitigation of circumstances in which people are overwhelmed.

Further Reading

- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th edn.). Washington DC: American Psychiatric Association.
- Engel, G. L. and Schmale, A. H. (1972). Conservation-withdrawal: a primary regulatory process for organismic homeostasis. In: *Ciba Foundation symposium: physiology of emotion and psychosomatic illness*, pp. 57–85. New York: Elsevier.
- Raphael, B. (1984). *The anatomy of bereavement*. London: Hutchinson.
- Raphael, B. (1986). *When disaster strikes*. London: Hutchinson.
- Selye, H. (1946). The general adaptation syndrome and the diseases of adaptation. *Journal of Clinical Endocrinology* 6, 117–196.
- Valent, P. (1998). *From survival to fulfillment: a framework for the life-trauma dialectic*. Philadelphia: Taylor and Francis.
- Wallace, A. F. C. (1957). Mazeway disintegration: the individual's perception of socio-cultural disorganization. *Human Organization* 16, 23–27.

Disasters and Mass Violence, Public, Effects of

G Stevens, B Raphael and M Dobson

Centre for Mental Health, New South Wales, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
G Stevens, B Raphael and M Dobson, volume 1, pp 699–705,
© 2000, Elsevier Inc.

Disaster Stressors

Stressor Elements and Psychosocial Effects

Vulnerabilities and Special Populations

Determining Who Is Affected

Recovery Environment

Glossary

<i>Disasters</i>	A class of events that can overwhelm individuals and their communities. Disasters elicit biological, psychological, and social stress reactions in individuals and affect the functioning of a community's social systems.
<i>Disaster stressors</i>	A range of events experienced by individuals in the course of their exposure to a disastrous situation. Disaster stressors may occur prior to, during, and in the aftermath of the disaster. The impact of these stressors is mediated by factors such as individuals' perceptions of their role in the disaster and the community's awareness of its scope.
<i>Mass violence</i>	A range of actions intended to injure, kill, or terrorize groups or communities. These may range from individual impulsive acts to group or community-based actions calculated to promote fear, civil disruption, or cultural-political change.

Disasters, by definition, are overwhelming events both for the communities they affect and for the individuals involved. The stress associated with disasters affects the functioning of social systems as well as creating biological, psychological, and social reactions in the people affected. As an event, a disaster may be brief or long lasting, anticipated or unexpected, humanmade or the result of natural forces, or any combination of these. Such elements, as well as background strengths and vulnerabilities, may contribute to how the events are resolved and to the outcomes: positive or adverse.

A range of models have been considered in describing disaster response. These include those based on

linear, systemwide reactive processes. A phase model, for example, may include warning, threat, impact, inventory, rescue, and recovery. This is one relatively simple model for a single-incident disaster. Both individual and social system reactions may be relevant in each phase. There is also a range of person–environment interactive models, which conceptualize the way that individuals and groups process the features of the event and their own capacities in their response to disaster events.

There is recognition that disaster scenarios are increasingly complicated by political, cultural, and military factors. A number of authors have outlined ways of conceptualizing and planning responses to these complex emergencies. Modern disaster planning, it is argued, must encompass the multiple and compounding disasters of famine, conflict, social disintegration, and the degradation of the environment. These scenarios more often reflect experiences within developing countries and cannot be easily conceptualized within the usual frameworks.

Disaster Stressors

The stresses experienced by individuals depend not only on the objective nature of the disaster event itself but also on the previous experience, perceived coping, available resources, and response readiness of those affected. For instance, a hurricane or flood may have been anticipated through recent warnings. Similar previous events may have prompted the creation of effective early warning and response systems. Alternatively, these natural occurrences may be totally unexpected, affecting particularly vulnerable human habitats with few counter-disaster systems in place. This latter scenario was seen in several regions, with devastating consequences, following the Indian Ocean earthquake and tsunami in December 2004. This interplay of factors means that those involved may experience effects ranging from the destruction of their homes, families, or communities to involvement in an exciting and unique challenge that has been successfully overcome.

The type of disaster, its intensity, and the availability of response resources may all be significant mediators of the psychosocial response. It has been generally observed that, as the level of exposure to disaster-related stressors increases, the mental health symptoms of those exposed similarly increase. This increase in impact associated with proximity, duration, or frequency of exposure is commonly referred to as the dose–response effect and has been observed across

different types of disasters and other traumatizing events.

Adverse impacts on mental health are possibly the most prevalent consequence of disasters. In a broad review of 121 studies examining disaster mental health, Norris found that adverse disaster effects included, in order of frequency, specific psychological problems (notably posttraumatic stress disorder, PTSD; major depression; and generalized anxiety disorder), nonspecific distress, physical health problems, chronic problems of living, resource loss, and problems specific to youth. The most common specific disaster stressors include bereavement, life threat and injury to self or family members, horror, property damage, and financial loss.

Disasters involving mass violence are found to consistently produce higher levels of distress and psychological impairment. Terrorism with conventional weapons typically produces immediate health effects in the form of injuries and deaths. For the perpetrators, these effects may be incidental to the wider aim of creating cultural or political change through civil disruption. In this sense, the intended effects of terrorism, as the term implies, are primarily psychological – to foster pervasive insecurity through specific acts of violence, create behavior changes disruptive to public infrastructure, and engender a loss of confidence in public institutions. Terrorism involving biological, chemical, or radiological (BCR) agents represents a particularly insidious, ongoing form of assault. Biological agents, for example, are typically invisible and odorless, produce delayed illness (and uncertainty), and may be associated with grotesque forms of death.

In one of the broadest reviews of the disaster literature to date (225 studies), Norris observed that when other variables are controlled, those exposed to mass violence and terrorism experience severe impairment at almost twice the rate of those subjected to natural or technological disasters. Acts defined as mass violence range from incidents such as the Columbine High School shootings to the September 11 attacks and the London underground bombings of 2005. Technological disasters are associated with human failure or negligence and may include pilot error or the industrial accident and chemical release at the Union Carbide factory in Bhopal, India. In general, the review findings indicated that those who experience technological disasters do not experience greater distress or morbidity than those exposed to natural disasters.

Disasters that are the direct result of human malevolence may be particularly difficult for survivors to comprehend and assimilate and may increase the likelihood of intrusion symptoms. In contrast, the

perceived absence of malevolent intent associated with natural disasters and technological accidents may allow the resolution of these events to occur more readily. An exception to this may be those humanmade disasters associated with chronic and invisible threats, such as nuclear radiation and toxic waste. Distress may be further heightened by supposed mismanagement or neglect, which allowed the exposure to occur. Cycles of blame may ensue. The actual and perceived risks may include damage to health, children, genetic/reproductive health, and otherwise unknown effects. These events often generate a different pattern of response characterized by chronic stress and associated physical health problems rather than discrete mental health conditions.

Location and the availability of economic resources may also be significant predictors of stress and other mental health effects. Natural disasters in developing countries have more severe consequences than either natural or technological disasters in developed countries. This is largely due to a combination of more severe natural events and a limited resource base, including poor housing and warning/recovery systems. It has been observed that following a disaster, individuals from developing countries experience severe mental health outcomes at more than twice the rate of those from developed countries.

Although multiple classifications of disaster have been attempted, they do not provide an easy guide to their consequences or to the appropriate actions to deal with the stress or mental health effects they may generate.

Stressor Elements and Psychosocial Effects

Life Threat and Traumatic Injury

Methodological inconsistencies make it difficult to determine, in comparative terms, which elements of disaster stressors have the greatest psychosocial impacts. Debate has ensued in the literature about which disaster exposures are more pathogenic (i.e., produce greater psychiatric morbidity). Despite these inconsistencies, the more comprehensive studies have generally found that threats to life and injury during disaster have stronger or longer-lasting effects on mental health.

Individuals who experience threat to their lives, especially when this is associated with shock and helplessness, are at increased risk of a range of reactive processes. These include heightened arousal, numbing of emotions, avoidance behavior, and the reexperiencing of traumatic events. Intense arousal and distress may trigger specific psychological and

physiological reactions, including dissociation or heightened arousal with prolonged elevation in heart rate. It has been shown that these reactions are more likely to develop into a pattern of psychiatric morbidity such as acute stress disorder (ASD), PTSD, or other specific conditions such as depression and generalized anxiety disorder. Studies of disasters such as the Mount St. Helen volcanic eruption have shown the specificity of perceived life threat as a stressor associated with the development of PTSD in adults. Pynoos et al. observed a similar relationship with children affected by a sniper shooting at a school.

Trauma events occasioning physical injury are also associated with greater mental health impacts. Ten years after the 1988 Piper Alpha oil platform disaster (in Scotland), 21% of workers directly exposed still met stringent criteria for PTSD, with features such as physical injury and survivor guilt associated with significantly higher morbidity. Consistent with other dose-response effects, the extent of the mental health impact also appears to increase with the severity of the injury. This phenomenon has also been observed in adolescents injured during a massive industrial accident in Toulouse, Italy. Injury effects also appeared to be cumulative, with PTSD rates being even higher in those with injury to family members and personal injury. However, the broad findings in relation to injury have been questioned. O'Donnell et al. critically reviewed the traumatic injury literature and suggested that methodological discrepancies and confounds in this field must be resolved if mental health associations are to be more accurately determined.

When deaths are massive and gruesome, the levels of PTSD-related morbidity may be considerably higher, particularly for those who may have experienced dissociation in relation to this experience. Lifton described the phenomena of death immersion (when the levels of death are massive and horrific), as occurred in Hiroshima. Unfortunately, genocides in modern conflict, war, and mass killings can also provoke such an effect. Massive deaths, or the deaths of key figures, may threaten the functioning of the community and its survival. This was exemplified in the 1966 disaster where a slagheap destroyed the school at Aberfan (in Wales). Although some of the effects of this tragedy have been long-lasting, the community did indeed survive and renewed itself. Death counts may provide an indicator of the degree to which a community may be affected by a disaster, but they are not the only stressor effect.

It must be remembered that, for the majority of individuals, reactive phenomena of the type noted normally diminish over time. Residents of lower Manhattan, for example, reported significant acute

mental health effects immediately after the September 11 attacks (measured as rates of PTSD and depression), followed by a pattern of strong recovery within 4 months. At a population level, there is a consistent trend for most survivors to return to their predisaster mental health status within several months. This occurs across disaster types, but it is rarely a linear decline.

Losses

Personal losses Personal losses may range from the intense and personal loss of a primary attachment figure (a parent, child or partner) to the loss of other family members, friends, and neighbors. Losses may extend to the deaths of community members and leaders, affecting the sense of community structure and leadership, key resources on which both an ongoing sense of security and practical recovery management may be based.

When personal losses relate to the deaths of others, bereavement represents the normal reactive psychological process for those dealing with these disaster outcomes. Grief and mourning occur for those to whom the individuals were attached or for whom they cared or valued. There is often an initial period of disbelief, which may be more pronounced when the death is unexpected and untimely. This is especially the case in the midst of the chaos and threat frequently observed in a major disaster. This shock, denial, and sense that things are unreal may affect whole communities as well as individuals. Although bereavement represents a normal reaction to traumatic loss, some individuals may be at increased risk of complex grief and other morbidity in the longer term. Risk factors for this may include a belief in survivor responsibility, the nature of the loss (e.g., sudden losses or the death of a child), and anxious personal coping styles.

The nature of disasters and some terrorist attacks is such that there may be prolonged periods of uncertainty about the fate of loved ones. The remains of some of the deceased may never be located. Others may be grossly disfigured and identifiable only through forensic examination. In these instances formal disaster victim identification (DVI) procedures are used, such as DNA and dental record analysis. Evidence gathering for criminal investigation may also be required. Although providing certainty and supporting a process of justice, the delay and apparent complexity of these procedures is often intensely stressful for those who are bereaved. Raphael, Dunsmore and Wooding described the complex DVI and support processes following the terrorist bombing on the Indonesian Island of Bali in 2002. They

described key support processes during this period, including viewings of the deceased, which may prevent or reduce negative mental health effects. They argued that it is not always apparent who the affected groups may be and that further research is needed to determine optimal methods of outreach and clinical support at such times.

Grief processes may be complicated by co-occurring experiences of personal life threat and traumatic stress reactions that may occur during disasters. Grief and trauma are distinct phenomena, the former characterized by distress, yearning, and an orientation toward the absent person and the latter characterized by heightened fear and arousal, vigilance, and a reactive orientation away from the feared event or its reminders. Galea et al. noted the specificity of responses to the September 11 attacks, stating that “experiences involving exposure to the attacks [and perceived life threat] predicted current PTSD and that loss as a result of the attack predicted current depression” (2002: 346). Similar specificity of response has been observed with natural disaster exposures.

Whereas grief, depression, and PTSD do represent distinct processes, losses during disasters or single-incident events (e.g., a motor vehicle accident) may result in co-occurring grief and trauma. These events cut across the ambivalence of everyday life and may particularly impact those in very dependent relationships, leading to patterns of complicated grief. The traumatic circumstances of disaster-related deaths often produces a complex mix of traumatic stress and bereavement phenomenology, which is more properly called a traumatic bereavement or traumatic grief. It remains unclear whether deaths resulting from particularly horrific exposures (e.g., the September 11 attacks) produce unique complications beyond traumatic bereavement or PTSD as they are currently conceptualized. However, the wider literature on complicated grief (traumatic and nontraumatic loss) indicates that certain groups may experience greater vulnerability. For example, the loss of a close family member (e.g., a parent, child, or spouse) is the best predictor of complicated grief. Preloss dependent or ambivalent relationship styles are also associated with enhanced risk. Early life experiences such as abuse or parental loss may similarly increase risk, possibly by promoting insecure attachment styles. Subsequent disaster losses may then exacerbate abandonment fears and affect mood regulation, increasing the risk of complicated grief.

The death of a loved one through acts of human malevolence, such as terrorism, has been shown to produce higher levels psychological morbidity than other types disaster. The arbitrary and calculated nature of these losses may be difficult for survivors

to comprehend and are associated with higher levels of intrusion and avoidance symptoms. There is anecdotal evidence that such events may also act as a significant and complicating stressor for those bereaved. Preoccupation with the reasons for the death (i.e., its meaning), the perpetrators’ motivation, and thoughts of revenge and justice may all act as a defense against the trauma. These preoccupations may ultimately inhibit acceptance of the loss and impede its resolution.

Clinicians working with populations so affected increasingly recognize the need to deal separately and specifically with the effects of both sets of stressor reactions to facilitate resolution. Traumatic stress effects (i.e., those related to life threat) must often be dealt with first, before it is possible to grieve. Regehr and Sussman suggested that an integrated evidence-based approach is likely to entail (1) cognitive restructuring and management of the arousal, avoidance, and reexperiencing of symptoms associated with trauma stressors and (2) phased (or parallel) relational therapies that resolve relationship issues and enable a new sense of the relationship and the self.

Community and resource losses In addition to personal losses of life, homes, and treasured possessions (e.g., family photographs), disasters frequently entail the loss of workplaces, neighborhoods, and community networks. Sociologists such as Erickson have studied these community impacts. Erickson described a massive flood affecting a mining valley and showed that there were not only the types of stressful experiences already but also the disruption of neighborhoods and social networks when people were displaced to new locations of safety and shelter. Although community losses may produce modest effects in themselves, their additive effects may be significant. Phifer and Norris observed that flood survivors from the United States who fared most poorly were those who experienced both high personal loss and high community destruction.

Community losses may not only produce grief but also lead to ongoing deficits and disruptions. These may include the loss of family and social structure and also the loss of opportunities for income and meaningful work. The Conservation of Resources stress theory predicts that psychological stress occurs when there is a threat of resource loss, an actual resource loss, or the absence of a resource gain following an investment of resources. Resource types identified within this theory include objects (e.g., homes and furniture), energy (e.g., money and insurance coverage), conditions (e.g., marriages and work roles), and personal characteristics (e.g., self-esteem).

Several studies have indicated that resource loss is an important predictor of psychological distress following natural disasters. Freedy et. al. found that 4–7 months after the 1991 Sierra Madre Earthquake (in Los Angeles County) resource loss was a more significant predictor of psychological distress than perceived life threat, prior exposure to traumatic events, and other stressful events. In a cross-national study 1 month after Hurricane Georges (Caribbean and U.S. gulf coast, 1998), Sattler et al. found that psychological distress (as measured by ASD symptoms) was most strongly predicted by location, resource loss, and the availability of social support. Approximately 25% of respondents in Puerto Rico and the Dominican Republic reported distress levels consistent with ASD, whereas only 10% of those in the United States (Alabama) and the U.S. Virgin Islands were affected in this way. Rates of depression were also higher in the first two countries. Although countries differed in reported vulnerabilities, the personal characteristics component of resource loss accounted for the greatest symptom variance overall. Basic object loss (food, water, and money) was greatest in Puerto Rico. Condition loss (family and work relationships) and object loss (sentimental items, furniture, etc.) accounted for the greatest resource distress in the Dominican Republic and the U.S. regions, respectively.

The limited resource base, greater exposure, and poor perceived support of developing countries produce significant mental health burdens. However, these resource spirals (continuing and compounding loss of resources) may be arrested through practical actions of the disaster response. For example, recent United Nations programs involving the priority reopening of schools, local markets, and factories have actively countered the prolonged sense of despair that may affect some postdisaster communities.

Separation and Dislocation

As noted, there are many aspects of disaster that separate family members and dislocate people from their homes and communities. In the aftermath of disasters, intense distress and searching behaviors are common as affected individuals look for their loved ones, even placing their own lives at further risk to do so. Information as to their whereabouts and opportunities for reunion may be the only influences that will lessen these behaviors. It may be, of course, that loved ones have died and grief will supervene; being able to see the body of the loved one may be a very important factor in resolution of the loss.

Mutual support of those so affected may be important for resolution, as may be the support of the

community in recognition and commemoration, perhaps of the site or in ceremony or memorial. As well as helping to reestablish a supportive community network, these rituals may support the resolution of these experiences by helping to construct a sense of meaning about what has occurred. This psychological process, sometimes referred to as an individual or community narrative, has been shown to be associated with better adjustment to traumatic events.

Dislocation is a process that relates more specifically to the loss of homes and the move to shelter. It may involve repeated moves to new places of safety or places to live until a new home is found. These may be key stressors for communities affected by natural disasters, but may profoundly impact societies subjected to bombings and warfare. This is also one of the powerful stressors affecting refugees.

Unlike losses through death or injury, dislocation stressors appear to affect mental health and well-being at a later stage and often in more chronic ways. This was first observed in studies of the effects of a cyclone that devastated the northern Australian city of Darwin. Parker found that an initial rise of distress measured in a population of evacuees reflected a mortality or life-threat stress. A subsequent rise several months later reflected dislocation/relocation stressor effects. Other studies of those affected by this disaster have shown similar longer-term stress effects related to dislocation.

Chronic Stressors

Disaster responses in communities often follow predictable cycles. There is typically an initial honeymoon phase, when there is an enormous positive and often altruistic response by surrounding communities, nationally and even internationally. This is driven by a wish to undo, or to make right, what has happened. Usual social boundaries diminish and there are powerful affiliative tendencies at a local level. However, this lasts only for the first few days or weeks. Then the reality of what has happened, what it will take to recover, who will pay, and how long it will take creates a very different environment. Anger and scapegoating, chronic distress, and frustration may occur. Confrontation with bureaucracies that have returned to their predisaster levels of functioning is common. All these factors may mean that there is an environment of disillusionment. This may contribute further to the distress of individuals or may adversely influence the functioning of the community, leading to cynicism and disappointment.

Fortunately most communities, like individuals, build on their resources and resilience and recover. The experience of the catastrophe becomes part of a

history of mastery and moving on. This is not always the case, however, and tragically, for some communities, the stigma of a particular catastrophe may remain. The community becomes identified with the disaster and the disaster with the community, often in the most negative ways—for instance the Chernobyl nuclear incident or the Port Arthur massacre.

Some disaster stresses, by their nature, are chronic and ongoing. Slow disasters include drought and deforestation and may provide the preconditions for starvation and the dislocation of communities. As noted, these environmental effects may coincide with chronic civil or ethnic conflict or the dispossession associated with years spent in refugee camps. The degradation of the environment through industrial accidents or toxic waste disposal can produce similar chronic effects.

Chronic stress may affect physical health via the immune system, neuroendocrine response, altered health behaviors (e.g. alcohol abuse), or changed mood states. Four months after Hurricane Andrew (in the United States) Ironson et al. found evidence that survivors were experiencing altered physiological functioning, suggestive of stress-related immune system suppression. Severe disaster stressors have even been known to lead to sudden death, as has been demonstrated in a number of different studies of earthquakes.

Posttraumatic Growth

For some individuals, the process of dealing effectively with loss or psychological trauma may ultimately provide a growth experience and enhance their personal maturity and relationships. This is well exemplified in the contributions of Victor Frankel and other Holocaust survivors. More recently, there has been increasing interest and study of the phenomena of posttraumatic growth. Tedeschi and Calhoun reviewed models of these phenomena and suggest that the provision of sensitive support and the opportunity for disclosure following trauma or disaster may allow significant cognitive restructuring in relation to the event. The experience of surviving and overcoming trauma may promote the development of a coherent and reempowered self-identity. In effect, the disaster experience and response may become a vehicle through which powerful changes to personal and group narratives may occur.

Vulnerabilities and Special Populations

Previous Exposure

Previous exposure to specific disaster stressors is an important factor affecting individual and community

response. In some people, this has built their strength, enhancing their capacity to overcome the new threat while also inoculating them against its stress effects. Conversely, exposure may have the effect of sensitizing some individuals, leading to poorer coping with subsequent events. For some, vulnerabilities associated with unresolved trauma or grief may surface again, triggered by the current crisis. Although there may be opportunities to rework those earlier experiences, they are often not recognized. Experiences such as childhood sexual abuse or wartime suffering may not clearly link to a current disaster context but are known to represent significant risk factors for further trauma.

Communities that have been exposed in the past may develop a disaster subculture. This may influence how they respond to warnings and to the disaster itself. This subculture reflects a belief system about what will happen, what can be done, and what the outcomes may be. This may influence responses in either positive or negative ways, depending on the adaptability of the beliefs and their fit with the particular incident.

Displaced Persons

Other vulnerable populations may be those who are already dislocated from familiar settings, for instance, refugees, new immigrants, or minority groups. They may lose their papers, documents of identification. Their links to their new community may still be fragile. They may not have the language, knowledge, or understanding needed to respond to the incident. Alternatively, they may have integrated these events in terms of earlier experiences of threat. They may experience intense fears for their safety because the chaos of the disaster is similar to the chaos they have known. Those who have suffered oppression in their home country may be fearful of government or law enforcement officials and emergency service workers. Displaced persons have been identified as vulnerable groups in many disaster responses and may need special attention to facilitate their recovery.

People with a history of mental illness are also a group that is vulnerable to the effects of disasters. The stressors associated with disasters are commonly recognized as risk factors that may exacerbate a range of preexisting conditions. Following traumatic events, individuals with prior or ongoing psychiatric conditions are more likely to experience a relapse or an exacerbation of these conditions, as well as the development of other forms of psychopathology, including PTSD. In addition to the stress effects associated with common exposures, specific elements of the event and its aftermath may trigger key

vulnerabilities. For example, bereavement or the loss of a support network or job may result in a relapse of major depression. The intense activity and officialdom of the recovery may heighten paranoid ideation in those experiencing psychotic conditions.

Indigenous peoples may face particular difficulties. Their backgrounds may have already involved life threat through premature mortality and racist violence. Other losses may involve family disruptions, deaths, and loss of land, culture, identity, and meaning. Such losses may entail both literal and symbolic dislocation from place, tribe, or community. Socioeconomic, educational, and other disadvantages may all contribute further to this background of traumatization. Disaster stressors superimposed on this context further increase the likelihood of adverse outcomes for these individuals and communities.

Children

Children and young people may be vulnerable through separation or by seeing effects on parents. The stresses of life threat, loss, separation, and dislocation can impact not only their present well-being but also their psychological and social development. This has been well demonstrated by Terr's 1991 in-depth studies showing that the effects may appear in repetitive traumatic play and behavioral disruptions that are both aggressive and withdrawing. Pynoos et al. studied the effects of a school-based sniper attack and were able to show that the mental health effects reflected the nature of the stressor experiences each child faced. For instance, those who experienced life threat showed PTSD phenomena, whereas those who experienced bereavement showed phenomena of grief and later depression. Those who experienced separation from their siblings showed separation distress phenomena.

The effects may be delayed. McFarlane reported that disaster-affected children's pattern of morbidity at first appeared less than that of the broader community and was particularly good. This had the apparent effect of forming a buffer between the children and the effects of what had happened. Their distress appeared later in the first year after a forest fire disaster and was also linked to ongoing symptomatology in their parents. Other studies of children support these findings and emphasize the need to recognize both behavioral responses and developmental impact. The centrality of the family for children, and their affection and security needs, must also be recognized if these stressors are to be dealt with.

The school is key environment of care and development for children following disaster. Teachers who have themselves been stressed may not identify the

effects on children. Some behavioral reactions may appear only in the school setting. For example, trauma-related cognitive and behavioral effects may become manifest as disruptive behavior or poor learning. Some school-based interventions have been shown to lessen the risk of adverse outcomes. In a study by McDermott, children used a workbook to facilitate the processing of traumatic experiences related to a forest fire.

The Elderly

Older people have been frequently described as vulnerable to disaster stresses. As with other studies in the mental health field, no specific increase in psychosocial morbidity has been found among the elderly. There may even be a decrease in later years. There is no good evidence that the elderly run a greater risk of mental health morbidity, although the patterns may be different. It has been observed that older people experience higher arousal but lower intrusive (re-experiencing) symptoms than young adults similarly exposed and that maturity and general life experience may buffer the stress of disasters. However, the elderly are also more likely to die during disasters. This may be due to several factors including recognition of warnings, the physical attributes necessary for escape, and a greater acceptance of death. As such, there is a need to ensure that individuals who may be vulnerable in this way are encompassed in disaster-response planning.

Determining Who Is Affected

Determining those who are affected by a disaster may be difficult. The registration of victims provides a formal mechanism, and this is often the role of organizations such as the Red Cross or Red Crescent or of formal institutions such as police services. But this really describes only the direct victims. A number of researchers have developed classifications of the different levels of victimhood in disasters. In broad terms, these include those directly experiencing the incident (who are often described as primary victims), those who were not there but are profoundly affected (such as bereaved family members), and those formally involved in rescue and helping efforts. The last group also experiences exposure effects and may be seen as secondary or indirectly affected victims. The distinction between victim and helper and the roles inherent in these terms may present a dichotomy that is a problematic. Victim implies weak and helpless, whereas helper implies strong and powerful. However, there are numerous examples of disasters in which survivors (the authors' preferred term) were both

victims and helpers in the early stages. As such, this distinction does not reflect the practical fluidity of response or the shifting of roles and exposures that often occurs.

It is of importance scientifically that a dose-response effect, as previously noted, may be demonstrated for the levels of stressor experience. In a number of studies, this relationship has been observed with both the frequency of exposure and its intensity or closeness to the epicenter of the disaster incident. The excellent 1989 study by Weisaeth of a workforce affected by a paint factory explosion and fire clearly shows such a relationship. The highest levels of morbidity were found in those closest to the explosion, with lesser levels in those further away. Significantly, however, a lesser but significant effect was also observed in workers who were not on shift but who would have been directly impacted had they been. This vicarious traumatization is a compelling indicator that the adverse effects of disaster and trauma need not be limited to direct physical exposure to the event. They may be powerfully linked to its indirect exposures and losses, including appraisals associated with personal risk.

Affected individuals and communities may experience any combination of the stressors identified. These may occur against a backdrop of ongoing life difficulties or specific events, but may be attributed to the distress of the disaster experience. They may be affected, or even secondarily injured, by the failure of others to acknowledge their experience or their suffering. The perception that social support is available and offered may be a central factor in reducing both disaster-related stress and longer-term mental health effects. Studies of Hurricanes Hugo and Andrew found that the deterioration in perceived social support (deterioration path) was associated with adverse mental health effects. Conversely, high levels of received support (actual postdisaster help) in the aftermath (mobilization path) were able to reduce these negative perceptions and offset some of these adverse outcomes. Despite findings that those with the greatest direct losses are often most strongly affected, these findings indicate that tangible support may bolster resilience; both may act as significant mediators of recovery.

Recovery Environment

The recovery environment consists of those natural and introduced forces that may influence the adjustment of individuals and communities following their experience of disaster. These include the natural processes of resilience that are vital for survival and the various adaptations that may support these, such

as specific coping strategies, personal resources, and social interaction. The more positive recovery environments are those in which the experience is acknowledged and help is offered but in which the strengths of individuals and communities to overcome disaster effects are recognized and supported. This approach, which focuses recovery activities on community resources and participation, is now formally advocated by the World Health Organization (Sphere Project).

The response to disaster-affected people and communities is often intense, and there may be a convergence of those wishing to help, including those offering practical and medical aid, debriefing, and subsequently welfare assistance. The altruism driving this wish to help represents the best of human values. Unfortunately, in this convergence there may be an excess of offers by the unskilled or those who may in turn become victims themselves. Debriefing is frequently part of such a response; however there is no evidence that it is beneficial to either emergency workers or the wider population. Some controlled studies have even shown that debriefing (i.e., critical incident stress debriefing) led to slightly but significantly worse outcomes. It was on the basis of findings such as these that the recent U.S. Consensus Conference on Mass Violence recommended against its use with the general population.

Cognitive-behavioral interventions have been shown to be effective in the treatment of specific trauma and are likely to benefit those with established PTSD conditions following disasters. Psychological first aid is an important helping response that focuses on the provision of immediate safety, reunion with loved ones, shelter, and support and the triaging of mental health and other needs. This approach promotes the recognition of and response to a postevent hierarchy of need and is now internationally recognized. Focused counseling for traumatization or grief may also be of value, but this needs to be based on evidence of effective interventions and delivered by skilled and knowledgeable personnel. The aim of any intervention may be "first to do no harm."

Unfortunately there remains a shortage of systematic investigations of the effectiveness of interventions postdisaster. This is largely due to the chaos of these events and their aftermath, as well as the methodological and ethical constraints on research at such times. As with the direct response, a primary aim of disaster researchers should also be "to do no harm" to affected populations. Some beneficial effects have been found for systematic intervention in environments such as schools or in postdisaster outreach and support programs as part of a disaster mental health response. There is a pressing need for

integrated research to ensure that human resources and mobilized personnel contribute to common goals, are effective in enhancing social and health outcomes, and are powerful and positive forces of normal recovery.

See Also the Following Article

Hiroshima Bombing, Stress Effects of.

Further Reading

- Alexander, D. (2005). Early mental health intervention after disasters. *Advances in Psychiatric Treatment* 11, 12–18.
- Bisson, I. J., Jenkins, P. L., Alexander, J., et al. (1997). A randomised controlled trial of psychological debriefing for victims of acute harm. *British Journal of Psychiatry* 171, 78–81.
- Brewin, C. R., Andrews, B. and Valentine, J. D. (2000). Meta-analysis of risk factors for posttraumatic stress disorder in trauma-exposed adults. *Journal of Consulting and Clinical Psychology* 68(5), 748–766.
- Briere, J. and Elliott, D. M. (2000). Prevalence, characteristics, and long-term sequelae of natural disaster exposure in the general population. *Journal of Traumatic Stress* 13, 661–679.
- Birmes, P., Brunet, A., Carreras, D., et al. (2003). The predictive power of peritraumatic dissociation and acute stress symptoms for posttraumatic stress symptoms: a three-month prospective study. *American Journal of Psychiatry* 160(7), 1337–1339.
- Chung, M., Dennis, I., Easthope, Y., et al. (2005). A multiple-indicator multiple-cause model for posttraumatic stress reactions: personality, coping, and maladjustment. *Psychosomatic Medicine* 67, 251–259.
- Dew, M. and Bromet, E. (1993). Predictors of temporal patterns of psychiatric distress during 10 years following the nuclear accident at Three Mile Island. *Social Psychiatry and Psychiatric Epidemiology* 28, 49–55.
- Erikson, K. (1994). *A new species of trouble: explorations in disaster, trauma and community*. New York: Norton.
- Foss, O. T. (1994). Mental first aid. *Social Science & Medicine* 38, 479–482.
- Frankel, V. (1962). *Man's search for meaning*. New York: Simon and Schuster.
- Freedy, J. R., Saladin, M. E., Kilpatrick, D. G., et al. (1994). Understanding acute psychological distress following natural disaster. *Journal of Traumatic Stress* 7, 257–273.
- Galea, S., Ahern, J., Resnick, H., et al. (2002). Psychological sequelae of the September 11 terrorist attacks in New York City. *New England Journal of Medicine* 356, 982–987.
- Galea, S., Vlahof, D., Resnick, H., et al. (2003). Trends of probable post-traumatic stress disorder in New York City after the September 11 terrorist attacks. *American Journal of Epidemiology* 158, 514–524.
- Godeau, E., Vignes, C., Navarro, F., et al. (2005). Effects of a large-scale industrial disaster on rates of symptoms consistent with posttraumatic stress disorders among schoolchildren in Toulouse. *Archives of Pediatric & Adolescent Medicine* 159, 579–584.
- Gray, M., Prigerson, H. and Litz, B. (2004). Conceptual and definitional issues in complicated grief. In: Litz, B. (ed.) *Early intervention for trauma and traumatic loss*, pp. 65–84. New York: Guilford Press.
- Hobfoll, S. E. (1989). Conservation of resources: a new attempt at conceptualizing stress. *American Psychologist* 44, 513–524.
- Hobfoll, S. E. (1998). *Stress, culture, and community: the psychology and philosophy of stress*. New York: Plenum Press.
- Hull, A. M., Alexander, D. A. and Klein, S. (2002). Survivors of the Piper Alpha oil platform disaster: long term follow-up study. *British Journal of Psychiatry* 181, 433–438.
- Ironson, G., Wynings, C., Schneiderman, N., et al. (1997). Posttraumatic stress symptoms, intrusive thoughts, loss, and immune function after Hurricane Andrew. *Psychosomatic Medicine* 59(2), 128–124.
- Leor, J., Poole, W. and Kloner, R. (1996). Sudden cardiac death triggered by an earthquake. *New England Journal of Medicine* 334(7), 413–419.
- Lifton, R. (1979). *The broken connection: on death and the continuity of life*. New York: Simon and Schuster.
- Litz, B. and Gray, M. (2004). Early intervention for trauma in adults: a framework for first aid and secondary prevention. In: Litz, B. (ed.) *Early intervention for trauma and traumatic loss*, pp. 87–111. New York: Guilford Press.
- Litz, B., Gray, M., Bryant, R., et al. (2002). Early interventions for trauma: current status and future directions. *Clinical Psychology: Science and Practice* 9, 112–134.
- Lundin, T. (1984). Morbidity following sudden and unexpected bereavement. *British Journal of Psychiatry* 144, 84–88.
- Maes, M., Mylle, J., Delmire, L., et al. (2000). Pediatric morbidity and comorbidity following accidental man-made traumatic events: incidence and risk factors. *European Archives of Psychiatry and Clinical Neuroscience* 250(3), 156–162.
- McCarroll, J., Fullerton, C., Ursano, R., et al. (1996). Posttraumatic stress symptoms following forensic dental examination: Mt Carmel, Waco, Texas. *American Journal of Psychiatry* 153, 778–782.
- McDermott, B. M., Lee, E., Judd, M., et al. (2005). Post traumatic stress disorder and general psychopathology in children and adolescents following a wildfire disaster. *Canadian Journal of Psychiatry* 50, 137–143.
- McFarlane, A. C. (1987). Post traumatic phenomena in a longitudinal study of children following a natural disaster. *Journal of the American Academy of Child and Adolescent Psychiatry* 26, 764–769.
- Morgan, L., Scourfield, J., Williams, D., et al. (2003). The Aberfan disaster: 33-year follow-up of survivors. *British Journal of Psychiatry* 182, 532–536.
- National Institute of Mental Health (2002). *Mental health and mass violence: evidence-based early psychological*

- interventions for victims/survivors of mass violence: a workshop to reach consensus on best practices.* NIH Publication No. 02-5138. Washington DC: US Government Printing Office.
- Norris, F., Friedman, M., Watson, P., et al. (2002). 60,000 disaster victims speak. Part 1: An empirical review of the empirical literature, 1981-2001. *Psychiatry* 65, 207-239.
- Norris, F. and Kaniasty, K. (1996). Received and perceived social support in times of stress: a test of the social support deterioration deterrence model. *Journal of Personality and Social Psychology* 71, 498-511.
- North, C., Nixon, S., Hariat, S., et al. (1999). Psychiatric disorders among survivors of the Oklahoma City bombing. *Journal of the American Medical Association* 282, 755-762.
- O'Donnell, M. L., Creamer, M., Bryant, R. A., et al. (2003). Posttraumatic disorders following injury: an empirical and methodological review. *Clinical Psychology Review* 23(4), 587-603.
- O'Toole, B. I., Marshall, R. P., Schureck, R. J., et al. (1998). Posttraumatic stress disorder and comorbidity in Australian Vietnam veterans: risk factors, chronicity and combat. *Australian & New Zealand Journal of Psychiatry* 32, 132-142.
- Parker, G. (1975). Psychological disturbance in Darwin evacuees following cyclone Tracy. *Medical Journal of Australia* 21, 650-652.
- Pfefferbaum, B., North, C. S., Flynn, B. W., et al. (2001). The emotional impact of injury following an international terrorist incident. *Public Health Reviews* 29(2-4), 271-280.
- Phifer, J. and Norris, F. (1989). Psychological symptoms in older adults following disaster: nature, timing, duration and course. *Journal of Gerontology* 44, 201-217.
- Prigerson, H., Ahmed, I., Silverman, G., et al. (2002). Rates and risks of complicated grief as a risk factor among psychiatric clinic patients in Karachi, Pakistan. *Death Studies* 26, 781-792.
- Prigerson, H., Shear, M., Jacobs, S., et al. (1999). Consensus criteria for traumatic grief: a preliminary empirical test. *British Journal of Psychiatry* 174, 67-73.
- Pynoos, R., Frederick, C., Nader, K., et al. (1987). Life threat and post-traumatic stress in school age children. *Archives of General Psychiatry* 44, 1057-1063.
- Pynoos, R. and Nader, K. (1988). Psychological first aid and treatment approach to children exposed to community violence: research implications. *Journal of Traumatic Stress* 1, 445-473.
- Raphael, B. (1977). The Granville train disaster: psychological needs and their management. *Medical Journal Australia* 1(9), 303-305.
- Raphael, B. (1986). *When disaster strikes: how individuals and communities cope with catastrophe*. New York: Basic Books.
- Raphael, B., Dunsmore, J. and Wooding, S. (2004). Terror and trauma in Bali: Australia's mental health response. *Journal of Aggression, Maltreatment and Trauma* 9(2), 245-256.
- Raphael, B., Martinek, N. and Wooding, S. (2004). Assessing loss, psychological trauma and traumatic bereavement. In: Wilson, J. (ed.) *Assessing psychological trauma and PTSD* (2nd edn.) pp. 492-510. New York: Guilford Press.
- Raphael, B., Meldrum, L. and McFarlane, A. (1995). Does debriefing after psychological trauma work? *British Medical Journal* 310, 1479-1480.
- Raphael, B. and Wilson, J. P. (1993). Theoretical and intervention considerations in working with victims of disaster. In: Wilson, J. P. & Raphael, B. (eds.) *International handbook of traumatic stress syndromes*, pp. 58-70. New York: Plenum Press.
- Raphael, B. and Wooding, S. (2004). Early mental health interventions for traumatic loss in adults. In: Litz, B. (ed.) *Early intervention for trauma and traumatic loss*, pp. 147-178. New York: Guilford Press.
- Regehr, C. and Sussman, T. (2004). Intersections between grief and trauma: toward an empirically based model for treating traumatic grief. *Brief Treatment and Crisis Intervention* 4, 289-309.
- Rose, S., Bisson, J. and Wessely, S. (2001). Psychological debriefing for preventing post traumatic stress disorder (PTSD). *The Cochrane Library* 4 [online].
- Sattler, D. N., Preston, A. J., Kaiser, C. F., et al. (2002). Hurricane Georges: a cross-national study examining preparedness, resource loss, and psychological distress in the U.S. Virgin Islands, Puerto Rico, Dominican Republic, and the United States. *Journal of Traumatic Stress* 15, 339-350.
- Salama, P., Spiegel, P., Talley, L., et al. (2004). Lessons learned from complex emergencies over past decade. *Lancet* 364(9447), 1801-1813.
- Shalev, A. Y., Sahar, T., Freedman, S., et al. (1998). A prospective study of heart rate responses following trauma and the subsequent development of posttraumatic stress disorder. *Archives of General Psychiatry* 55, 553-559.
- Shore, J. and Tatum, L. (1986). Psychiatric reactions to disaster: the Mount St. Helens experience. *American Journal of Psychiatry* 143, 590-595.
- Silverman, G. K., Johnson, J. G. and Prigerson, H. G. (2001). Preliminary explorations of the effects of prior trauma and loss on risk for psychiatric disorders in recently widowed people. *Israeli Journal of Psychiatry and Related Sciences* 38, 202-215.
- Tedeschi, R. and Calhoun, L. (2004). Post-traumatic growth: conceptual foundations and empirical evidence. *Psychological Inquiry* 15, 1-18.
- Terr, L. (1991). Childhood traumas: an outline and overview. *American Journal of Psychiatry* 148, 10-20.
- Ursano, R. J., Fullerton, C. S. and McCaughey, B. G. (1994). Trauma and disaster. In: Ursano, R. J., McCaughey, B. & Fullerton, C. S. (eds.) *Individual and community responses to trauma and disaster: the structure of human chaos*, pp. 3-27. Cambridge, UK: Cambridge University Press.
- Ursano, R. J., Fullerton, C. S. and Norwood, A. E. (eds.) (2004). *Bioterrorism: psychological and public health interventions*. London: Cambridge University Press.

- Ursano, R. J., Fullerton, C. S., Tzu-Cheng, K., et al. (1995). Longitudinal assessment of posttraumatic stress disorder and depression after exposure to traumatic death. *Journal of Nervous and Mental Disease* 183(1), 36–42.
- van Doorn, C., Kasl, S. V., Beery, L. C., et al. (1998). The influence of marital quality and attachment styles on traumatic grief and depressive symptoms. *Journal of Nervous and Mental Disease* 186, 566–573.
- Weisaeth, L. (1989). The stressors and the post-traumatic stress syndrome after an industrial disaster. *Acta Psychiatrica Scandinavica* 80(supplement), 25–37.

Relevant Websites

- Norris, F. (2005). Range, magnitude and duration of the effects of disaster on mental health: review update 2005, Dartmouth College research education review article. www.redmh.org.
- Sphere Project (2004). Humanitarian charter and minimum standards in disaster response, Geneva, <http://www.sphereproject.org>.

Disease, Stress Induced

H S Willenberg, S R Bornstein and G P Chrousos
University Hospital, Duesseldorf, Duesseldorf, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by H S Willenberg, S R Bornstein and G P Chrousos, volume 1, pp 709–713, © 2000, Elsevier Inc.

Sympathetic-adrenomedullary system (SA system)
Stress
Stressor

Part of the autonomic neural system.

State of threatened homeostasis.

Forces that threaten to destroy the state of homeostasis.

Anatomy and Physiology of the Stress System
Stress and Disease
Disease and Stress

Glossary

<i>Adrenocorticotrophic hormone (ACTH)</i>	Is known under its scientific term “corticotropin”. Regulates adrenal steroidogenesis, specifically glucocorticoid synthesis and cortisol release.
<i>Adaptive response</i>	Complex set of an organism’s reactions to reestablish homeostasis.
<i>Apoptosis</i>	Cell death initiated by several mechanism in a regulated process. It is also termed “programmed cell death.” Cells with this fate undergo a sequence of changes typical for this process.
<i>Arginine-vasopressin (AVP)</i>	Peptide hormone, which stimulates pituitary ACTH release
<i>Corticotropin-releasing hormone (CRH)</i>	Also termed “corticotropin-releasing factor” (CRF) or “corticoliberin.” Main regulator of ACTH.
<i>Homeostasis</i>	State of dynamic equilibrium.
<i>LC/NE system</i>	Locus ceruleus / norepinephrine system.
<i>Proopiomelanocortin (POMC)</i>	Polypeptide secreted by pituitary cells which is the basis for several hormones of the pituitary that are yielded by alternative processing, including corticotropin, melanocyte-stimulating hormones (MSH), and endorphins.

The continuation of life depends on the ability of an organism to maintain a state of dynamic equilibrium or homeostasis. Homeostasis is constantly disturbed by entropic forces, the stressors; the state of threatened homeostasis is called stress. A complex set of behavioral and physical reactions, the adaptive response, is employed by the organism to reestablish balanced physiological conditions. At the beginning of the twentieth century, Walter Cannon and Hans Selye developed these modern concepts of stress and discussed the relation between stress and disease. In antiquity, Hippocrates regarded health as the harmonious balance of the four key elements of life – water, fire, air, and earth – and defined states of imbalance among these elements as disease. He also described the adaptive response as the “healing power of nature.” (Chrousos et al.)

The actual term stress was originally adopted from physics by Hans Selye. In physics, it describes the resistance of a body to applied pressure, following Hook’s law of elasticity. An acting force may distort a body elastically in a linear fashion or, if applied excessively in quality, quantity, or time, may produce nonlinear deformations. Similarly, stress may be a transient and time-limited state, well balanced by the adaptive responses of the organism, whereas excessive and prolonged stress may alter the physiological and behavioral defense mechanisms, rendering them inefficient or deleterious.

Anatomy and Physiology of the Stress System

Structures that include components of a network of neural and endocrine responses that are activated adaptively during stress have been collectively called the stress system. On the cellular level, molecules such as the highly conserved heat shock proteins (hsp) are primitive key elements of the cellular stress response. They comprise several families of proteins that are found across species and are classified depending on their molecular size (kilodaltons) into hsp16–30, hsp60, hsp70, and hsp90. They have been shown to regulate cell and tissue homeostasis and to be protective when this is threatened by stressors.

The central components of the stress system are located in the hypothalamus and the brain stem, the phylogenetically oldest parts of the brain. These centers receive information inputs of internal and external origins and compute the appropriate responses, which are effected by two main pathways: the endocrine hypothalamic-pituitary-adrenal (HPA) axis and the neural systemic sympathetic-adrenomedullary (SA) and parasympathetic systems. Their central regulators and peripheral end products, corticotropin releasing hormone (CRH), glucocorticoids, and catecholamines, are key hormones in the reestablishment and maintenance of cardiovascular, metabolic, immune, and behavioral homeostasis.

The Hypothalamic-Pituitary-Adrenal Axis and Stress

CRH, the principal central regulator of the HPA axis, is mainly produced by parvocellular neurons of the hypothalamus in the paraventricular and parabrachial nuclei of the medulla and in the central nucleus of the amygdala. Its action on the pituitary gland is supported by arginine vasopressin (AVP), also secreted by parvocellular neurons of the paraventricular hypothalamic nuclei (PVN). CRH induces the production and secretion of adrenocorticotrophic hormone (ACTH) from the pituitary, a hormone whose main target is the cortex of the adrenal gland. The intracerebroventricular administration of CRH causes the activation of the stress system and behavioral patterns similar to those observed during stress.

The adrenal gland, as the end organ of the HPA axis, is subject to functional changes of the stress system. It responds quite rapidly to sustained stimulation with hypertrophy and hyperplasia, which can be monitored on the macroscopic, microscopic, ultrastructural, and molecular levels. Whereas in acute stress high blood concentrations of glucocorticoids and ACTH can be found, in chronic stress high levels of glucocorticoids and relatively low levels

of ACTH are frequently observed. This dissociation between the central activation of the HPA axis and adrenal cortex function represents an adaptation of the stress system whereby medullary input and hypertrophy-hyperplasia of the zona fasciculata of the adrenal cortex sustain elevated glucocorticoid secretion in the presence of low normal ACTH concentrations. In addition, although ACTH is a potent secretagogue for adrenal steroids, this peptide does not seem to mediate the hyperplasia of the adrenal cortex during stress. Other pituitary factors are under discussion, especially N-terminal fragments of proopiomelanocortin (POMC).

Glucocorticoids exert their pleiotropic effects through ubiquitously distributed intracellular receptors in almost all tissues of the body. In addition, they participate in their own regulation of secretion through a classic negative feedback mechanism by binding to high-affinity mineralocorticoid and low-affinity glucocorticoid receptors in the frontal cortex, amygdala, hippocampus, hypothalamus, and pituitary. Exposure to pathologically elevated levels of glucocorticoids causes neurons to slow down their metabolism, although these neurons may undergo death by apoptosis. Therefore, this loop of HPA axis regulation serves to terminate or suppress an excessive response of the stress system.

The Systemic Sympathetic-Adrenomedullary and Parasympathetic Systems

This axis originates from mostly norepinephrine cell groups of the locus ceruleus, medulla, and pons (LC/NE). Efferent projections of these neurons terminate throughout the brain and spinal cord as well as in the peripheral ganglia of the autonomic nervous system. Through a complex neural network, a wide spectrum of physiological functions can be controlled. These include cardiovascular tone, respiration, skeletal muscle and adipose tissue metabolic activity, and gastrointestinal function. The adrenal medulla is an effector organ of this system, supplying the systemic circulation with epinephrine and norepinephrine. These catecholamines also potentiate ACTH-stimulated cortisol production by the adrenocortical cells.

Coordination and Communication of Hypothalamic-Pituitary-Adrenal Axis and LC/NE System

The organization and implementation of a qualitatively and quantitatively appropriate stress response depends on the fine-tuning of both the regulatory centers of the stress system and the response of target organs. Indeed, a complex network of neurons, neuronal and hormonal factors, and other cell mediators is at interplay at each level of interactions. The secretion of CRH and central catecholamines is controlled

by the neurons of the frontal cortex, mesocortical/mesolimbic system, and the amygdala–hippocampus complex. Neuronal pathways that produce γ -aminobutyric acid (GABA), substance P, and β -endorphin inhibit CRH-producing neurons. CRH is not only responsible for the activation of the HPA axis but also for the stimulation of LC/NE neurons in the brain stem.

The adrenal gland unites the HPA and SA axes under a common capsule. A great degree of intermingling of cortical and medullary cells increases the area of contact between these systems. In addition to the stimulation through ACTH, ACTH-independent mechanisms seem to be involved in the adaptation to chronic stress. The stimulation of chromaffin cells leads to the potentiation of glucocorticoid secretion, whereas glucocorticoids enhance catecholamine production by medullary cells.

Interplay of the Stress System with Major Endocrine, Immune, and Other Systems

The activation of the stress system helps the organism to overcome the influence of stressors and, therefore, postpones all functions that may interfere with the chance of the individual to survive. The stress system stimulates the mesocorticolimbic system, influencing motivation and reward phenomena. The stress system also stimulates the amygdala and hippocampus. The former generates fear and anxiety, whereas the latter participates in the negative feedback control of the HPA axis; both recall implicit emotion-laden memories. Neurons of the arcuate nucleus of the hypothalamus produce POMC-derived peptides. Two of these, α -melanocyte-stimulating hormone (-MSH) and β -endorphin inhibit the activity of the stress system. Projections to the brain stem and spinal cord elevate the gain threshold, producing so-called stress-induced analgesia.

The activation of the stress system leads to the inhibition of a number of endocrine axes. These include the gonadal, growth, and thyroid axes. The direct inhibition of central neural circuits is supported by the peripheral suppressive actions of glucocorticoids.

Other profound consequences of stress system activation are observed in the immune system. However, the adrenal cells express toll-like receptors, making the adrenal cortex sensitive to registering the activations of the innate immune system. Further important participants in the immunological homeostasis are cytokine signals from cells of the immune tissue. All three major inflammatory cytokines – tumor necrosis factor α , interleukin 1, and interleukin 6 – produced at sites of inflammation in a cascade-like fashion, exert stimulatory effects on the HPA axis. Their actions are directed

to all three levels of the HPA axis. The most potent mediator among them is interleukin 6; it causes elevation of CRH, AVP, ACTH, and the glucocorticoids. Indirect actions of the inflammatory cytokines on the HPA axis are also attained through the stimulation of the noradrenergic centers in the brain stem. In chronic hypoactivation of the HPA axis, increased serum levels of interleukin 6 can be observed.

The stress system uses additional mechanisms to affect immune homeostasis. In addition to the suppressive effect of glucocorticoids on the production of tumor necrosis factor α , interleukin 1, and interleukin 6, the glucocorticoids also inhibit the target tissue responses to cytokines through the suppression of nuclear transcription factor- κ B and activator protein 1 (AP-1). Glucocorticoids also cause a switch from type 1 T-helper cytokine profile to type 2 T-helper cytokine profile by inhibiting interleukin 12 and by stimulating interleukin 10. In this way, a shift from a cellular to a humoral immune response is promoted. In addition, the glucocorticoids initiate apoptosis in the eosinophils and in some lymphocytes.

Stress and Disease

Stress plays an important role in individual and species survival, but it also participates in the development and aging of every individual. The quality, intensity, and duration of stress are important in determining the positive or negative effects on the organism.

There are two main conditions of the stress system that lead to pathology: an excessive and prolonged or a defective adaptive response to a stressor. In both cases, the dose–response relation between the potency of a stressor and the adaptive response of the organism is shifted to the left or right, respectively. Excessive or defective reactions to stressors that are associated with pathological or physiological states are summarized in [Table 1](#). More than half of the variance determining the stress response is due to inherited factors, whereas the remaining half is due to early-life constitutional factors and concurrent environmental input.

In the syndrome of major melancholic depression, a disproportionate activation of the generalized stress response is chronically manifested. Central activation with consequent elevation of CRH is followed by a blunted ACTH response of the pituitary to exogenous CRH, probably due to the downregulation of CRH receptors and/or increased glucocorticoid feedback. At autopsy, these patients have an increased number of CRH and AVP neurons in the paraventricular nucleus of the hypothalamus and hippocampal

Table 1 Pathological conditions associated with altered activity of the HPA axis

<i>Increased HPA activity</i>	<i>Decreased HPA activity</i>	<i>Disrupted HPA activity</i>
Severe chronic disease	Atypical depression	Cushing's syndrome
Melancholic depression	Seasonal depression	Glucocorticoid deficiency
Anorexia nervosa	Chronic fatigue syndrome	Glucocorticoid resistance
Obsessive-compulsive disorder	Hypothyroidism	
Panic disorder	Adrenal suppression	
Chronic excessive exercise	Obesity (hypoenergetic forms)	
Malnutrition	Nicotine withdrawal	
Diabetes mellitus	Vulnerability to inflammatory disease (Lewis rat)	
Hyperthyroidism	Rheumatoid arthritis	
Central obesity	Premenstrual tension syndrome	
Childhood sexual abuse	Postpartum mood and inflammatory disorders	
Pregnancy	Vulnerability to alcoholism	

atrophy. In addition, imaging and morphometric studies have shown that the pituitary and adrenals are enlarged. The sequelae of hypercortisolism include pseudo-Cushing's syndrome manifestation, osteoporosis, and opportunistic infections. Similar observations were made in patients with chronic active alcoholism, with anorexia nervosa, who were sexually abused in their youth, and with obsessive-compulsive disorders.

Another example of hyperactivation of the stress system is critical illness. When entering a chronic state, interleukin 6 and possibly other cytokines synergize with ACTH to maintain increased stimulation of the adrenal cortex. In this condition, a dissociation between ACTH and cortisol levels occurs. As in melancholic depression, these patients escape full HPA axis suppression following dexamethasone.

In the autoimmune-inflammatory syndrome of rheumatoid arthritis, a hypoactive HPA axis has been reported. It appears that a defective glucocorticoid response to inflammation is an important factor in the pathogenesis of this disease.

Disease and Stress

There are pathological conditions of the stress system leading to disease and, *vice versa*, diseases that alter the activity of the stress system in a pathological fashion.

Stress System Disorders

The adaptive response to stress has three components. First, stress has to be sensed. Sensory neural and hormonal pathways have developed to connect the outside world and the rest of the body to the central nervous system. Hypersensitivity or hyposensitivity to stressors may cause disproportionate responses, resulting in pathology. Therefore, disorders characterized by altered perception of signals to

the stress system cause inefficient responses to natural challenges, with the consequent development of secondary diseases.

The second component covers the calculation and integration of the information input, and the third component covers the response. Disorders of the central nervous system that influence these responses interfere with the stress system and develop pathology related to this interference.

The activation of the stress system is necessary to overcome homeostasis-threatening events, but excessive and prolonged activation may be self-destructive. Transient hypercortisolemia is necessary during surgery, sepsis, and other severe conditions, but hypercortisolemia also leads to hyperglycemia, hyperlipidemia-increased catabolism, and immune suppression – all states that, if chronically present, have major adverse effects on the body.

Diseases That Directly Influence the Stress System

Diseases generally disturb homeostasis and should be regarded as stressors. Conversely, every stressful influence that alters homeostasis of the organism in a long-term fashion could be regarded as a process that may cause disease.

Cushing's syndrome or Addison's disease leads to the disintegration of the stress system, which then fails to mount an adaptive response. In addition, severe sepsis leads directly to the activation of the stress system. Patients with these conditions develop multisystem disorders that result from excessive or defective cortisol effects on systems, whose proper function depends on optimal glucocorticoid activity.

Further Reading

Bornstein, S. R. and Chrousos, G. P. (1999). Adrenocorticotropin (ACTH)- and non-ACTH-mediated regulation of the adrenal cortex – neural and immune inputs.

Journal of Clinical Endocrinology and Metabolism 84 (5), 1729–1736.

Bornstein, S. R. and Ehrhart-Bornstein, M. (1999). Interactions of the stress system in the adrenal: basic and clinical aspects. In: Bolis, C. L. & Licinio, J. (eds.) *Stress and adaptation: from Selye's concept to application of modern formulations*, pp. 89–108. Geneva: World Health Organization.

Chrousos, G. P. (1998). Stressors, stress, and neuroendocrine integration of the adaptive response: the 1997 Hans Selye memorial lecture. *Annals of the New York Academy of Sciences* 851, 311–335.

Chrousos, G. P., Loriaux, D. L. and Gold, P. W. (1988). The concept of stress and its historical development. *Mechanisms of physical and emotional stress. Advances in Experimental Medical Biology* 245, 3–7.

Dissociation

J R Maldonado

Stanford University, Stanford, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J R Maldonado, volume 1, pp 714–722, © 2000, Elsevier Inc.

Introduction

Models of Dissociation

The Dissociative Disorders

Treatment

Hypnosis as a Treatment Tool in the Psychotherapy of
Dissociative Disorders

Legal Ramifications of Memory Work

Glossary

<i>Absorption</i>	The process of absorbing or being absorbed; great interest; entire occupation of the mind.
<i>Dissociation</i>	The act of dissociating or state of being dissociated; separation. In psychology, term used to describe the separation of whole segments of the personality (as in multiple personality) or of discrete mental processes (as in the schizophrenias) from the mainstream of consciousness or of behavior.
<i>Hypnosis</i>	The state of being hypnotized; abnormal sleep. There is no formal definition in the dictionary for hypnotizability or hypnotic capacity. This could properly be defined as a subject's ability to use a natural state of mind known as hypnosis.
<i>Suggestibility</i>	The ability to influence by suggestion.
<i>Trauma</i>	Bodily injury caused by violence; emotional shock (psychic trauma) with a lasting effect; a disordered psychic or behavioral state resulting from mental or emotional stress or physical injury.

Introduction

The word dissociation is defined as the act of dissociating or the state of being dissociated. It can also stand for the splitting off of a group of mental processes from the main body of consciousness. In psychiatry, the term is used to describe the reversible separation of elements of “identity, memory or consciousness” (APA, 2000) from the mainstream of consciousness or of behavior, as in dissociative amnesia, fugue, depersonalization, or dissociative identity disorder. The word comes from the Latin verb *dissociate*, meaning to separate, to disunite; *dis*, meaning asunder; and *sociare*, meaning to unite.

Thus, dissociation can be understood as the mechanisms by which we are able to separate mental processes and contents. Ordinarily, these are processed together. Normally, the (nonpathological) mechanism of dissociation allows us to carry on more than one complex task or action simultaneously by keeping out of consciousness routine experiences or tasks. For example, it allows a person to watch a television program while knitting or cooking. Similarly, it allows individuals to successfully drive their car while listening and learning from a book on tape.

Unconsciously, the mechanism of dissociation can also be elicited as an attempt to preserve some sense of control, safety, and identity in the face of overwhelming stress. Dissociation is usually elicited as a response to trauma that threatens personal physical integrity or the safety of a loved one, or as a response to extreme isolation and abandonment. Unfortunately, dissociative defenses give victims a false sense of control while immediately provides temporary relief from the full impact of the traumatic experience. Because the use of dissociation allows people to separate themselves from the awareness of danger, some trauma victims act as if the event is not happening. Yet others act as if it had never happened.

Pathological dissociation is usually the result of trauma. The form of trauma may vary widely, ranging from abandonment to overt physical trauma. Trauma may involve natural disasters (e.g., earthquakes, tornadoes, floods), man-made accidental disasters (e.g., war, nuclear plant explosions), large-scale man-made intentional disasters (e.g., the World Trade Center attack of 9/11 in New York City, the Oklahoma City bombing, school shootings, terrorism), or more personal traumas, such as rape, car accidents, or even a traumatic medical procedure. Probably among the most publicized of all forms of trauma are the victims of childhood physical and sexual abuse.

The traumatic experience invariably objectifies its victim. The sense of helplessness engendered by such situations creates sudden challenges to normal ways of processing perception, cognition, affect, and relationships. Trauma in the form of both natural disaster and human assault causes disruption to normal cognitive and affective processes and forces its victims to reorganize mental, psychological, and physiological processes in an attempt to prevent the immediate impact of the trauma. This reorganization may take the form of various forms of dissociative processes, fostering separation from painful surroundings and realities (derealization), or from the victim's own body (depersonalization). Even though such defenses may initially be adaptive, directed at maintaining control at times of overwhelming stress, some trauma victims develop persistent dissociative, amnesic, and anxiety-like symptoms. The ultimate sequelae of trauma for a substantial minority of its victims may be the development of traumatic stress disorders or dissociative disorders.

Dissociative disorders have been described in many cultures and settings, with women making up the majority of cases. Nearly 90% or more of cases reported in the literature are women. The most common dissociative disorder diagnosis is the not otherwise specified category, both in the United States and in non-Western countries, where dissociative trance and possession trance are the most common dissociative disorder diagnoses.

Dissociation as a pathological (i.e., psychiatric) phenomena was initially described by the French philosopher and psychologist Pierre Janet. Janet used the French term *désagrégation*, which carries with it a different and more accurate nuance than its English translation, dissociation. *Désagrégation* implies a separation of mental contents despite their general tendency to aggregate or be processed together. According to Janet, the problem experienced by patients suffering from dissociative phenomena is a difficulty in

integration of psychic processes, rather than a proliferation of components of consciousness, memory, identity, or perception. Janet (1989) viewed *désagrégation* as a purely pathological process, "a malady of the personal synthesis." Janet was also the first to study psychological trauma as a principal cause of dissociative phenomena. Unfortunately, during the twentieth century Janet's work and theories of dissociation as a pathological process were eclipsed by the emerging psychoanalytic theories proposed by Sigmund Freud. Freud's teachings and methods emphasized the mechanism of repression, rather than dissociation, as the driving force of his patients' problems. Nevertheless, early writings by Freud and his colleague Josef Breuer explored unconscious phenomena through an examination of seemingly similar dissociative phenomena. Infact, expanding on the theories on hysteria of their mentor, the French neurologist Jean Marie Charcot, Breuer and Freud suggested that the dissociative symptoms exhibited by their patients could be attributed to the patients' capacity to enter spontaneous and uncontrolled hypnoid states.

Based on observations of combat soldiers and former concentration camp inmates, Hilgard developed a neodissociation theory that revived interest in Janetian psychology and psychopathology. Hilgard's neodissociation model conceived of a mental structure with divisions that were horizontal rather than vertical, different from Freud's topographic model, which was composed of deeper and deeper layers of unconscious mental processing. Unlike Freud's system, Hilgard's model allowed for immediate access to consciousness of any of a variety of warded-off memories. In his model, amnesia was the crucial mediating mechanism that provided the barriers that divided one set of mental contents from another. Thus, he proposed the flexible and reversible use of amnesia as a key defensive tool.

Carl Jung had a unique insight in recognizing that the dissociability of the psyche is a fundamental process that extends along the continuum from normal mental functioning to abnormal states. Thus, Jung viewed the process of dissociation as a universal and necessary psychic activity for the development of personality through the differentiation of functions. On the other hand, he postulated that when the cohesion of consciousness was shattered by extreme traumatic experiences (e.g., childhood physical or sexual abuse), the natural process of differentiation of function was intensified, creating extreme dissociative splits between autonomous forces in the psyche, as observed in the cases of several psychiatric disorders, e.g., borderline personality disorder (BPD) or dissociative identity disorder (DID).

Models of Dissociation

Several models have been proposed to explain the process of dissociation. One such model, the neurological model, suggests that some underlying neurological process, such as hemispheric disconnection or epilepsy, causes the manifested clinical symptoms of dissociation. Several forms of epilepsy, particularly temporal lobe seizures, can manifest with rather complex behaviors (e.g., automatism), of which the subjects have little or no recollection. In other forms of epilepsy (e.g., petit mal seizures) subjects may experience depersonalization-like phenomena, reminiscence of dissociative phenomena.

The second proposed model is the role enactment model, also known as the social role demand theory. This model suggests that the symptoms associated with dissociative disorders represent an artificial social construct rather than a true psychiatric disorder. The proponents of this model suggest that dissociation is a clinical artifact created out of conscious or unconscious needs on the part of patients to capture the attention of their doctors. Some have even proposed a direct iatrogenic etiology in which therapists suggest the existence of symptoms (e.g., personality states). Finally, the autohypnotic model recognizes the connection between trauma, dissociation, and hypnotizability.

The third proposed model is the autohypnotic model. This theory recognizes and reconciles the connection between traumatic events, dissociative experiences, and hypnotizability. Thus, the autohypnotic model suggests that trauma victims dissociate or forget traumatic memories, using self-hypnotic mechanisms in order to keep them out of consciousness because reality is too painful to be faced. Memories may be stored at conscious and/or unconscious levels. The degree of amnesia varies, with some patients who have always remembered. Some remember only parts of the trauma. Yet in certain cases, some memories are transformed and others are interspersed with fantasy (e.g., cases of ritual abuse or alien abduction). Later on, some memories slowly leak into the conscious mind. Because full memories are not available it is difficult for patients to make sense of them. Some victims experience leaked memories in the form of flashbacks or dreams. Other trauma victims isolate themselves due to the shame they feel in relation to the suspected trauma.

Even though dissociated memories may be temporarily unavailable to consciousness, they may continue to influence conscious (or unconscious) experiences and behavior. Clinical experience suggests that dissociated information continues to affect patients' moods, behavior, and cognitive processes. It is common for

patients suffering from dissociative disorders to feel that something is wrong but to be unable to identify what. Many patients will experience a sense of discomfort, or sometimes even panic, when in specific situations. In fact, many of the symptoms experienced by trauma victims (e.g., flashbacks, psychosomatic symptoms) can be explained by the influence exerted by dissociated (forgotten or inaccessible) memories.

Nevertheless, it is important to remember that not everyone exposed to trauma will develop dissociative symptoms. In fact, studies conducted with war veterans reported that less than 25% of soldiers exposed to trauma during combat go on to develop severe dissociation or posttraumatic syndrome. This suggests that other factors may be involved in the production of dissociative and anxiety symptoms. A leading factor may involve the developmental stage during which the trauma occurs. Available data suggest that, as a rule, the earlier in life the trauma occurs, the greater the psychological sequelae caused. This is probably due, in part, to the defense mechanisms and coping styles that the victim has had an opportunity to develop prior to the occurrence of the trauma, which in turn will modulate the way in which the victim perceives the experience and adapts to its impact and consequences.

Putnam proposed the discrete behavioral states model of dissociative identity disorder. Forrest described a possible biological mechanism to support Putnam's theory by proposing the involvement of the orbital frontal cortex in the development of DID and suggesting a potential neurodevelopmental mechanism responsible for the development of multiple representations of the self. The proposed orbital frontal model integrates and elaborates on theory and research from four domains: the neurobiology of the orbital frontal cortex and its protective inhibitory role in the temporal organization of behavior; the development of emotion regulation, the development of the self, and experience-dependent reorganizing of neocortical processes. The hypothesis proposed by Forrest establishes that the experience-dependent maturation of the orbital frontal cortex in early abusive environments, characterized by discontinuity in dyadic socio-affective interactions between the infant and the caregiver, may be responsible for a pattern of lateral inhibition (i.e., dissociation or inhibition of conflicting subsets of self-representations that are normally integrated into a unified self). His basic idea is that discontinuity in the early caretaking environment is manifested in the discontinuity in the organization of the developing child's self.

Brenner has suggested that dissociation may be seen as a complex defense. Furthermore, he believes

that DID may be thought of as a lower level dissociative character, and that there is a unique psychic structure, which he calls the dissociative self, whose function is to create alter personalities out of disowned affects, memories, fantasies, and drives. As many other experts on dissociation have suggested, Brenner believes that this dissociative self must be dissolved during the therapeutic process in order for integration of alter personalities to occur.

In fact, recent scientific literature has focused on the characterological features of DID, thus extending the state versus trait debate to the realm of the dissociative disorders. This has led to the development of theories suggesting that DID may be a variant of, on a continuum with, or comorbid with the diagnoses of narcissistic and BPDs. Brenner suggested that DID would be best considered as a distinct characterological entity. He presented two theories, which described DID as a personality disorder whose predominant defense is dissociation. The more developed model, which possibly has more explanatory value, is the dissociative character. In this schema, DID would be considered a lower-level dissociative character, utilizing primitive forms of dissociation in which splitting is enhanced by an autohypnotic defensive altered state of consciousness. These patients experience altered states, which originate in response to the overstimulation stemming from external trauma, but which are reactivated in response to here-and-now intrapsychic conflicts.

Universality of Dissociation as a Syndrome

Dissociative symptoms have been reported in virtually every major psychiatric disorder and, in less severe forms, even in nonpatient (normal) populations. In the United States general (nonclinical) population, 6.3% of adults have reported three to four dissociative symptoms. Studies have suggested that dissociative disorders might comprise 5 to 10% of the psychiatric populations. Similarly, dissociative disorders have been described in many cultures and settings. Studies demonstrating prevalence rates have been conducted in Turkey, the Netherlands, Switzerland, Australia, Ethiopia, and Germany.

The Dissociative Disorders

The *Diagnostic and Statistical Manual*, 4th edition (DSM-IV) recognizes five different dissociative disorders. These include depersonalization disorder, in which patients feel detached from their own mental processes or body and which patients describe as a robot-like or out-of-body experience. Dissociative amnesia is a disorder in which patients exhibit an inability to recall important personal information.

The memory loss (amnesia) may apply to a part of the trauma (localized) or all (selective) of the traumatic experience. In extreme cases the amnesia may include personal information regarding the patient's entire life (generalized amnesia). Patients suffering from dissociative fugue suddenly go away, usually traveling far from home and developing a new identity while remaining amnesic to their past personal history. Patients suffering from DID exhibit two or more identities or personality states accompanied by frequent memory gaps (i.e., amnesia). Finally, there is a category called dissociative disorder not otherwise specified (NOS), by which patients exhibiting the core of dissociative symptoms but not fulfilling the diagnostic criteria of the previous four syndromes are classified. An example that may fit in this category is patients suffering from trance states.

Dissociative Amnesia (Psychogenic Amnesia)

The hallmark of this disorder is the inability to recall important personal information, usually of a traumatic or stressful nature, that is too extensive to be explained by ordinary forgetfulness. This is probably the most common of all dissociative disorders. In fact, amnesia is not only a psychiatric disorder by itself, but also a symptom, commonly found in a number of other dissociative and anxiety disorders. The highest incidence of dissociative amnesia has been described in the context of war and other natural and man-made disasters. Data suggest a direct relationship between the severity of the exposure to trauma and the incidence of amnesia.

Cases of dissociative amnesia usually share the following characteristics. First, the memory loss is episodic. That is, patients usually suffer the first-person recollection of certain personal events rather than knowledge of procedures (i.e., a person may not remember that he or she is a mechanic by trade or training, but is still able to intuitively and competently handle the tools). The memory loss involves a discrete period of time ranging from minutes to years. There seems to be a clear, dense amnesia, not a vagueness or inefficient retrieval of memories. There is usually no difficulty in learning new episodic information. The memory loss generally encompasses events of a traumatic or stressful nature.

In cases of dissociative amnesia patients are usually aware of their memory loss. Cases of dissociative amnesia can also be distinguished from amnesia of neurological origin by patients' intact cognition and capacity to learn new information. The primary problem in amnesia patients is a difficulty in retrieval rather than encoding or storage; thus, memory deficits are usually reversible. In fact, once the amnesia has cleared, normal memory function is resumed.

The epidemiology of this disorder is unknown. Nevertheless, dissociative amnesia is considered to be the most common of all dissociative disorders. The etiology of amnesic disorders is presumed to be posttraumatic, since they generally occur within the context of severe psychosocial stress. The duration of the disorder varies from a few days to a few years.

Even though it is possible to experience a single episode of amnesia, many patients have experienced several episodes during their lifetime. This is more common when the stresses that caused the initial episode are not resolved. The spontaneous resolution of the symptoms is rather common. In fact, many patients experience spontaneous recovery without more treatment than a protective environment. Sometimes more systematic treatment may be necessary.

The differential diagnosis of dissociative amnesia includes epilepsy, brain malignancy, head trauma, medication side effect (e.g., benzodiazepines use), chronic drug abuse and acute intoxication, cardiovascular and metabolic abnormalities, other dissociative disorders, an organic brain syndrome, a factitious disorder, and malingering.

Dissociative Fugue (Psychogenic Fugue)

Dissociative fugue is characterized by the sudden, unexpected travel away from home or one's customary place of daily activities, with inability to recall some or all of one's past. As in the previous disorder, amnesia is present, causing a sense of confusion about personal identity. On occasion, patients assume a new identity.

In contrast to dissociative amnesia cases, patients suffering from fugue states appear normal to the lay observer. Patients usually exhibit no signs of psychopathology or cognitive deficit. Fugue patients differ from those with dissociative amnesia in that the former are usually unaware of their amnesia. Only upon resumption of their former identities do they recall past memories, at which time they usually become amnesic for experiences during the fugue episode. Often, patients suffering from fugue states take on an entirely new (and often unrelated) identity and occupation. In contrast to patients suffering from DID, in fugue states the old and new identities do not alternate.

Not much is known regarding the etiology of this disorder. Nevertheless, the underlying motivating factor appears to be a desire to withdraw from emotionally painful experiences. Clinical data suggest that predisposing factors include extreme psychosocial stress such as war or natural and man-made disasters, personal and/or financial pressures or losses, heavy alcohol use, and intense and overwhelming stress

such as assault or rape. The onset of some fugue episodes may occur during sleep or be associated with sleep deprivation.

As in cases of acute dissociative amnesia, the onset of the disorder is usually associated with a traumatic or overwhelming event accompanied by strong emotions such as depression, grief, suicidal or aggressive impulses, or shame. Dissociative fugue is the least understood dissociative disorder. This may be due to the fact that most of these patients do not present for treatment. Usually they do not come to the attention of medical personnel until they have recovered their identity and memory and return home. Typically, patients seek psychiatric attention once the fugue is over and they are seeking to recover their original identity or retrieve their memory for events that occurred during the fugue.

Dissociative Identity Disorder (Multiple Personality Disorder)

DID is defined by the presence of two or more distinct identities or personality states that recurrently take control of behavior. This disorder represents the failure to integrate various aspects of identity, memory, and consciousness. Characteristics of this disorder are memory disturbances and amnesia. In contrast to other dissociative disorders, the degree of amnesia experienced in DID is usually asymmetrical. That is, it selectively involves different areas of autobiographical information, i.e., alters (personality states or identities) differ in the degree of amnesia for the experiences of other alters and the access to autobiographical information.

Usually there is a primary or host personality that carries the patient's given name. Often the host is not completely aware of the presence of alters. Because of the presence of amnesic barriers, different personalities may have varying levels of awareness of the existence of other personalities. On average there are 2 to 4 personalities present at the time of diagnosis, and usually up to 13 to 15 personalities are discovered during the course of treatment.

The symptoms that usually prompt patients or their families to seek treatment include memory deficits, moodiness, erratic and unpredictable behavior, depression, self-mutilation, suicidal ideation or attempts, and the overt manifestation of an alternate personality. Transition from one personality to another is usually sudden and is commonly triggered by environmental/interpersonal factors.

Alter identities may have different names, sexes, ages, and personal characteristics and often reflect various attempts to cope with difficult issues and problems. Alters can have a name and well-formed

personalities, e.g., Rose, an 8-year-old girl, or can be named after their function or description, e.g., the Angry One.

The factors that can lead to the development of DID are quite varied, but most authors seem to agree that physical and sexual abuse during childhood is the most commonly found etiological factor in these patients. In fact, a history of sexual and/or physical abuse has been reported in 70–97% of patients suffering from DID, with incest being the most common form of sexual trauma (68%). Other forms of childhood trauma that are associated with later development of DID include physical abuse other than sexual abuse (75%), neglect, confinement, severe intimidation with physical harm, witnessing physical or sexual abuse of a sibling, witnessing the violent death of a relative or close friend, traumatic physical illness on self, and near-death experiences.

The actual incidence and prevalence of this disorder are unclear. The estimated prevalence of DID in the general population has been reported to range from 0.01 to 1%. The average time from the appearance of symptoms to an accurate diagnosis is 6 years. The average age at diagnosis is 29 to 35 years. It has been described to be more common in women than in men by a ratio of 3–9:1. Female patients are also reported to present more personalities (average of 15) than men (average of 8).

There is a high incidence of comorbid psychiatric and medical syndromes. Of the psychiatric disorders, depression is the most common (85–88%), followed by posttraumatic stress disorder, BPDs, and substance abuse. There are a number of other psychiatric symptoms common to patients with DID, including insomnia, suicide attempts or gestures, self-destructive behaviors, phobias, anxiety, panic attacks, auditory and visual hallucinations, somatization, conversion reactions, and psychotic-like behavior.

As in cases of dissociative amnesia and fugue, the differential diagnosis of dissociative disorders includes an organic condition (e.g., temporal lobe epilepsy, brain malignancy, head trauma, medication side effect, drug abuse, and intoxication), other dissociative disorders, psychotic disorders (e.g., schizophrenia), factitious disorder, and malingering.

Depersonalization Disorder

Depersonalization is characterized by persistent or recurrent episodes of feelings of detachment or estrangement from one's self. Depersonalization disorder is primarily a disturbance in the integration of perceptual experience. Individuals commonly report feeling like a robot or as if they are living a dream or a movie. Patients suffering this condition describe

their experience as if they were an outside observer of their own mental processes and actions. In contrast to delusional disorders and other psychotic processes, reality testing is intact. The phenomena associated with depersonalization are not uncommon. In fact, depersonalization is seen in a number of psychiatric and neurological disorders, including agoraphobia and panic disorder, acute and posttraumatic stress disorder, schizophrenia, other dissociative disorders, personality disorders, acute drug intoxication or withdrawal, psychotic mood disorders, epilepsy, Ménière's disease, sensory and sleep deprivation, hyperventilation, and migraine headaches. In fact, people with no psychiatric condition can transiently experience symptoms of depersonalization. Because of this it is important that the diagnosis is applied only when the presence of symptoms causes severe impairment in functioning or marked distress.

The incidence and prevalence of this condition are unknown. The symptom of depersonalization has been described as being the third most common psychiatric symptom, after depression and anxiety. It is believed that under severe stress, up to 50% of all adults have experienced at least one brief episode of depersonalization. On the other hand, up to 12–46% of normal college students, nearly 30% of individuals exposed to life-threatening danger, up to 40% of hospitalized psychiatric patients, and about 69% of patients with panic disorder have experienced transient episodes of depersonalization. Episodes of depersonalization may also occur as a symptom of alcohol and drug abuse, as a side effect of prescription medication, and during stress and sensory deprivation. The sex distribution is unknown.

Theories for the etiology of this disorder range from the completely physiological, such as anatomical defects similar to epilepsy, to the purely psychological, such as a defense against painful and conflictual affects or the split between observing and participating ego/self, to combinations of all of the above as the result of a preformed functional response of the brain as an adaptation to overwhelming trauma. In any event, exposure to traumatic experiences seems to be the common etiological factor in this disorder. The course of the illness is usually chronic, with episodes of exacerbation usually following exposure to real or perceived stress.

Treatment

To date there are no controlled studies addressing the treatment of any of the dissociative disorders. All the information available reflects the experience and case reports of clinicians and specialized treatment centers. No single treatment modality has been

systematically studied in this patient population. Similarly, there are no established pharmacological treatments except for the use of benzodiazepines or barbiturates for drug-assisted interviews.

Dissociative Amnesia (Psychogenic Amnesia)

Before treatment is started it is necessary to establish that the amnesia or fugue state is of dissociative (psychogenic) origin, ruling out the possibility of a neurological disorder such as temporal lobe epilepsy. Usually the initial step in the treatment is to provide a safe environment. Simply removing the person from the threatening situation and providing security and protection has allowed for the spontaneous recovery of acute episodes. At times additional help may be needed to obtain the necessary biographical information or to facilitate the patient's recall. Among the adjuvants commonly used, hypnosis and barbiturate- or benzodiazepine-facilitated recall are the most popular and better described.

No studies have addressed the efficacy of hypnosis in the treatment of dissociative amnesia. Nevertheless, most researchers in this area agree that hypnosis is a very useful tool for the recovery of repressed and dissociated memories. Once the amnesia has been reversed, treatment should be directed at restructuring the events and defining the factors that led to the development of the amnesia. This is followed by the establishment of appropriate defenses and mechanisms to prevent further need for dissociation. This is best done within the context of more extensive therapeutic work.

Dissociative Fugue (Psychogenic Fugue)

The treatment should involve provision of rest and assurances of safety, development of a trusting therapeutic relationship, recovery of personal identity, review of triggers or factors associated to the onset of the fugue, reprocessing of traumatic material, reintegration of traumatic memories into personal history, and returning the patient to his or her previous life.

Hypnosis and drug-facilitated interviews have commonly been used during the stages of recovery of personal identity and memories associated with the onset of the fugue. Except for the facilitation of memory retrieval (narcosynthesis) or diminution of anxiety related to the therapeutic process, no pharmacotherapeutic agents have been systematically studied in the treatment of this condition. The treatments for acute dissociative amnesia and fugue states have been generally similar. Traditionally, hypnosis and amytal narcosynthesis were the tools of choice

for the recovery of dissociated material in subjects suffering from both amnesia and fugue states. In fugue cases, treatment should be undertaken as quickly as possible, while the repressed material is more readily accessible and before the memories have an opportunity to consolidate into a nucleus, as this may increase the possibility of future flight episodes. Patients may also experience spontaneous memory recovery upon removal from the stressful situation, when exposed to cues from their past, or when they feel psychologically safe. Psychodynamic psychotherapy may help to address the conflicts that precipitated the amnesia or fugue, thereby reducing subsequent dissociation under stress.

Dissociative Identity Disorder (Multiple Personality Disorder)

To date, there are no studies demonstrating the efficacy of pharmacological agents in the treatment of DID. As in the other dissociative disorders, pharmacological agents do not seem to address the disorder, but may serve as adjunct treatment of accessory symptoms (e.g., comorbid insomnia, anxiety and depressive symptoms).

Almost every psychotherapeutic and psychopharmacological treatment modality has been applied to patients suffering from DID, from group psychotherapy to family therapy, from antidepressants to antipsychotics to anticonvulsants to β -blockers, from hypnosis to eye movement desensitization and reprocessing (EMDR) to electroconvulsive therapy (ECT). All of them have been shown to have various degree of success.

In general terms, the treatment of DID involves (1) development of a therapeutic relationship based on safety and trust, (2) negotiation with the patient about cooperation with treatment, (3) development of a contract against harm to self or others, (4) history taking and understanding personality structure, (5) abreaction and working through of traumatic experiences and frequently repressed or dissociated material, (6) negotiating and modulating conflicts among aspects of identity and personality states, (7) development of mature and more appropriate, nondissociative defenses, and (8) working toward integration of alters.

Techniques such as hypnosis can facilitate control over dissociative episodes and integration of traumatic memories (see next section for details). Efforts at development of a social network and support system are helpful. Once integration has been achieved, further work is needed to deal with residual or renewed dissociative responses to external stress or internal conflicts and to further integration with society.

The treatment of DID is usually arduous, painful, and prolonged. Even though the ultimate goal is the achievement of integration, in some cases a reasonable degree of conflict-free collaboration among the personalities is all that can be achieved. Among clinicians, the most common treatment modality used is individual psychotherapy facilitated by the use of hypnosis. The average DID patient is seen twice a week for a period of about 4 years before integration is achieved. Clinical experience seems to suggest that patients suffering from DID do not experience spontaneous remission of their illness if left untreated.

Depersonalization Disorder

Episodes of depersonalization are usually transient and thus remit without formal treatment. Recurrent or persistent depersonalization may require intervention. Treatment modalities include paradoxical intention, record keeping and positive reward, flooding, psychodynamic psychotherapy, psychoeducation, psychostimulants, antidepressants, antipsychotics, anticonvulsants, benzodiazepines, ECT, and hypnosis. In considering somatic treatments for depersonalization, it is important to determine whether the complaint represents the primary condition or whether it is secondary to another disorder. When an additional psychiatric diagnosis (e.g., anxiety or depression) is present and successfully treated, the symptoms of depersonalization usually improved or resolve. Thus, the most important aspect of the treatment of depersonalization disorder is careful assessment of possible psychiatric comorbidity and treatment of those conditions.

As a symptom, depersonalization responds to self-hypnosis training. Often, hypnotic induction will induce transient depersonalization symptoms in patients. This is a useful exercise because having a structure for inducing the symptoms provides patients with a context for understanding and controlling them. Individuals for whom this approach is effective can be taught to induce a pleasant sense of floating lightness or heaviness in place of the anxiety-related somatic detachment. Often, the use of an imaginary screen to picture problems in a way that detaches them from the typical somatic response is also helpful.

Hypnosis as a Treatment Tool in the Psychotherapy of Dissociative Disorders

Traumatic memories may be elicited or spontaneously emerge during the course of psychotherapy without the utilization of any technique for memory enhancement. Nevertheless, there may be instances when the judicious use of hypnosis as adjuvant to

psychotherapy is recommended. For instance, hypnosis may facilitate access to repressed memories that have not emerged using other methods. Clinical experience suggests that many trauma victims respond to the traumatic event by using dissociative-like defenses during or after the trauma. In cases of repeated trauma, some victims learn how to trigger these dissociative responses (self-hypnosis-like defenses) in order to avoid further suffering. The structured use of a therapeutic hypnotic induction may elicit dissociative phenomena, as seen during the administration of the hypnotic induction profile. If needed, hypnosis can be helpful in facilitating controlled access to dissociated personalities and can be used to simply call up different identities or personality states as needed during the therapeutic process. More importantly, teaching patients to use self-hypnotic techniques to access alter states helps patients understand that they can have control over traumatic memories and experiences, leading to a way to control the episodes of spontaneous dissociation (e.g., flashbacks, personality states).

Studies show that patients suffering from dissociative disorders are highly hypnotizable. Thus, if hypnotic-like states are elicited during traumatic experiences and some patients unknowingly used them in order to dissociate from their surroundings, it makes sense that the very entry into this same state could lead to the retrieval of memories and effects associated with the original trauma, as would be predicted by the theory of state-dependent memory. Hypnosis then can be useful both as a diagnostic tool and as a powerful therapeutic technique. Properly done, hypnotic techniques can effectively facilitate symbolic restructuring of the traumatic experience.

The condensed hypnotic approach using hypnosis as a facilitator has two major treatment goals: to make conscious previously repressed traumatic memories and to develop a sense of congruence between memories associated with the traumatic experience and patients' current realities and self-images. These goals can be achieved by the use of six consecutive and interdependent steps or stages. Each of them is designed to help patients work through previously repressed memories, while enhancing control over their dissociative mental processes. The six stages are confrontation, condensation, confession, consolation, concentration, and control.

During the first stage, confrontation of the trauma, patients come to terms with the factors and events associated to the trauma. The therapist's role at this time is that of a supportive, nonjudgmental listener, who avoids suggesting or implanting facts.

During the second stage, hypnosis is used to facilitate a condensation of the traumatic memories. There is little need to force patients to recount all the details

of every traumatic episode. Thus, therapists can use hypnosis to define segments or episodes that summarize the traumatic experience without allowing patients to compulsively relive the entire trauma.

This is followed by the stage of confession of feelings and experiences that patients are profoundly ashamed of and may have never told anyone before. During this stage the therapist should convey a sense of being present for the patient, while remaining neutral.

The recovery and confession of traumatic memories is often accompanied by an immense sense of shame and sorrow. Thus, during the next stage, consolation, the therapist is more active and emotionally available to patients. It is appropriate to make empathic comments about the impact the experience must have had on them. Because of the frequent incidence of boundary violations during the episodes of abuse, therapists are warned to be extra careful regarding the ways in which care and empathy are expressed. Therapists are reminded of the possibility of the development of traumatic transference relationship at this point. That is, patients may interpret the therapist's interest in their past as a pretext for making them suffer again. Because of the mobilization of painful emotions, therapists are experiencing as hurting rather than helping them.

The intense concentration characteristic of hypnosis facilitates therapeutic work on selective traumatic memories. The structure of the hypnotic trance allows patients to turn on memories in the secure environment of the therapy session, while allowing them to shut them off once the intended work has been completed at the end of the session. Hypnosis also provides the flexibility to work on one aspect of the memory without requiring patients to recall the entire trauma. Therefore, the hypnotic process promotes concentration on the desired goal while helping patients remain in control. Using the structured experience of the therapeutic trance, patients learn how to think about the trauma in a constructive fashion.

Finally, patients should be helped to achieve an enhanced level of control, while restoring a sense of mastery and order in their lives. In order to achieve that goal, therapists guide the therapeutic interaction in such a way that patients' sense of control over their memories is enhanced. Patients are allowed to remember as much as can safely be remembered now, rather than pushed to remember the entire event at the therapist's will or in a compulsive fashion. The therapeutic use of self-hypnosis teaches patients that they are in charge of the experience, thus they can regain control of their memories. It allows patients to learn to trust their own feelings, perceptions, and

judgments. An important aspect in enhancing patients' sense of mastery is for them to learn not only how to control memories and symptoms, but also, equally important, when to ask for help.

Legal Ramifications of Memory Work

Every treatment involving potential repressed memories or childhood abuse must be preceded by a thorough explanation of the therapeutic and legal ramifications and a complete informed consent. Because of these potential problems some courts have restricted the admissibility of the testimony of hypnotized witnesses. Patients need to be warned about the fallibility surrounding hypnotically (or indeed any, even nonhypnotically) recovered memories, including the possibilities of confabulation, concreting, and the difficulty of differentiating between fantasized or remembered memories. It is difficult to avoid potential contamination using any memory enhancement (e.g., imagery work, hypnosis, EMDR, and conventional psychotherapy); thus, the therapist must use a neutral tone throughout the entire treatment. If hypnosis is to be used, therapists may guide patients through the experience, but at all times must avoid using leading or suggestive questions. The goal is to avoid contamination by introducing or suggesting information during the course of treatment. Indeed, hypnosis and any other nonhypnotic method of memory enhancement, including police interrogation, can distort memory by any of three ways: confabulation (the creation of pseudomemories, which are then reported as real); concreting (an unwarranted increased sense of confidence with which hypnotized individuals report their memories, whether these are true or false), or as the result of an additional recall trial.

More recently, there has been concern about the creation of false memories. Nevertheless, we must remember that the mind is not an accurate tape recorder that registers every experience we have. Because of the way people register, encode, and recover filed memories, false memories are by no means rare. In fact, most people probably are in doubt about certain aspects of their past. At times it may be difficult to differentiate things that we have seen, said, or done from those we have dreamed or just imagined. It is very likely that the most frequent source of false memory is the accounts we give to others of our experiences. Usually, such accounts are almost always made both more simple and more interesting than the truth. In fact, memory can be understood as a capacity for the organization and reconstruction of past experiences and impressions in the service of present needs, fears, and interests. Thus, therapists must be aware that not every memory recovered

with the use of hypnosis (or any other method of memory enhancement) is necessarily true. Hypnosis can facilitate improved recall of both true as well as confabulated material.

If memories of abuse are recovered, we do not encourage our patients to take legal action. There is no scientific evidence indicating that confrontation with alleged perpetrators of childhood abuse provides any therapeutic benefit to patients. The same is true for the pursuit of legal action or retribution toward perpetrators. It is impossible for therapists or the courts to be certain of the veracity of the memories recovered by either conventional therapy or hypnosis. Without objective external confirmation, therapists cannot distinguish among memories that are real, those that are the result of confabulation, and those that result from a combination of both. Furthermore, clinical experience in abuse cases suggests that not only would it be difficult to substantiate many of the allegations, but also little can be done in order to protect patients from the embarrassment, humiliation, and further trauma that would be imposed on them by virtue of the legal proceedings.

Several professional organizations have established clear guidelines regarding the use of hypnosis in instances where hypnosis is used as a method of memory enhancement. The guidelines suggest that when hypnosis or any other memory enhancement method is being used for forensic purposes or in the context of working out traumatic memories, especially those related to childhood physical and/or sexual abuse, multiple steps should be applied to minimize the possibility of memory contamination and maximizing the integrity of original memory.

See Also the Following Articles

Hypnosis; Memory Impairment; Posttraumatic Stress Disorder – Clinical; Posttraumatic Stress Disorder, Delayed; Posttraumatic Stress Disorder, Neurobiology of; Posttraumatic Therapy; Trauma and Memory.

Further Reading

Anderson, G., Yassenik, L. and Ross, C. A. (1993). Dissociative experiences and disorders among women who identify themselves as sexual abuse survivors. *Child Abuse and Neglect* 17, 677–686.

- Brenner, I. (1999). Deconstructing DID. *American Journal of Psychotherapy* 53, 344–360.
- Brown, G. R. and Anderson, B. (1991). Psychiatric morbidity in adult inpatients with childhood histories of sexual and physical abuse. *American Journal of Psychiatry* 148, 55–61.
- Cattell, J. P. and Cattell, J. S. (1974). Depersonalization: psychological and social perspectives. In: Arieti, S. (ed.) *American Handbook of Psychiatry*, pp. 767–799. New York: Basic Books.
- Chu, J. A., Frey, L. M., Ganzel, B. L., et al. (1999). Memories of childhood abuse: dissociation, amnesia, and corroboration. *American Journal of Psychiatry* 156, 749–755.
- Foote, B., Smolin, Y., Kaplan, M., Legatt, M. E. and Lipschitz, D. (2006). Prevalence of dissociative disorders in psychiatric outpatients. *American Journal of Psychiatry* 163, 623–629.
- Hammond, D. C., Garver, R. B., Mutter, C. B., et al. (1995). *Clinical hypnosis and memory: Guidelines for clinicians and for forensic hypnosis*. Bloomington, IL: American Society of Clinical Hypnosis Press.
- Horevitz, R. P. and Braun, B. G. (1984). Are multiple personalities borderline? An analysis of 33 cases. *Psychiatr Clin North Am* 7, 69–87.
- Kluft, R. P. (1991). Multiple personality disorder. In: Tasman, A. & Goldfinger, S. M. (eds.) *American Psychiatric Press Review of Psychiatry* (vol. 10), pp. 161–188. Washington, D.C.: American Psychiatric Press.
- Li, D. and Spiegel, D. (1992). A neural network model of dissociative disorders. *Psychiatr Ann* 22, 144–147.
- Maldonado, J. R. and Spiegel, D. (1998). Trauma, dissociation and hypnotizability. In: Marmar, C. R. & Bremner, J. D. (eds.) *Trauma, memory and dissociation*, pp. 57–106. Washington, D.C.: American Psychiatric Press.
- Maldonado, J. R., Butler, L. D. and Spiegel, D. (2000). Treatment of dissociative disorders. In: Nathan, P. & Gorman, J. M. (eds.) *Treatments that work*, pp. 463–493. New York: Oxford University Press.
- Maldonado, J. R. and Spiegel, D. (2007). Dissociative disorders. In: Talbot, J. & Yudofsky, S. (eds.) *Textbook of Psychiatry* (Fifth Edition). Washington, DC: American Psychiatric Press (in press; last edition, 4th info: pp. 709–742, published 2002).
- Orne, M. T., Axelrad, A. D., Diamond, B. L., et al. (1985). Scientific status of refreshing recollection by the use of hypnosis. *Journal of the American Medical Association* 253, 1918–1923.
- Spiegel, D. (1990). Hypnosis, dissociation and trauma: hidden and overt observers. In: Singer, J. L. (ed.) *Repression and dissociation*, pp. 121–142. Chicago, IL: University of Chicago.

Distress

G Matthews

University of Cincinnati, Cincinnati, OH, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G Matthews, volume 1, pp 723–729, © 2000, Elsevier Inc.

Introduction

Assessment of Distress

Influences on Distress

Personality and Distress

Psychological Concomitants of Distress

Glossary

<i>Affect</i>	Emotional or feeling states, including specific emotions such as anxiety and depression.
<i>Neuroticism</i>	A personality trait associated with a predisposition to experience negative affect.
<i>Self-regulation</i>	The motivational and cognitive processes controlling goal-directed personal adaptation to the external environment.
<i>Stress</i>	The processes and responses associated with adaptation to demanding or challenging environments.
<i>Trait</i>	A personal disposition or characteristic showing long-term stability and influencing behavior in a range of situations

Introduction

Distress is an imprecise term that typically refers to unpleasant subjective stress responses such as anxiety and depression. It is also sometimes used to describe behaviors and medical symptoms (e.g., somatic distress). The concept of distress derives from Selye's general adaptation syndrome (GAS), the generalized temporal sequence of physiological and psychological stress responses that may be elicited by noxious or threatening life events. Often, stress responses are characterized by difficulties in adapting to the external stressor (distress), although stress may sometimes have a stimulating, energizing effect (eustress). Distress may thus be conceptualized as the internal strain provoked by an external stressor. However, this simple metaphor has been largely superseded by the view of Lazarus that stress responses reflect dynamic person–environment relationships. From this perspective, distress signals

that adaptation to environmental demands is taxing or unsuccessful.

This article reviews the assessment of distress, the roles of situational factors and personality in generating distress, and its psychological significance, including clinical implications. Contemporary research on distress is dominated by two themes. The first is the extent to which a global distress concept is to be preferred to more specific negative affective responses such as anxiety, depression, and anger. The second theme is that distress must be understood within a wider framework of the person's active attempts to adapt to the physical and social environment.

Assessment of Distress

Distress is a broad label for a variety of stress responses, and so there is no single, generally accepted means of assessment. Typically, the measurement of distress is questionnaire-based, although, especially in animal and child research, distress may be operationalized as behaviors such as vocalizations. Approaches to assessment vary in two respects: (1) the time scale over which distress is measured, and (2) the breadth of distress symptoms encompassed. Distress is typically assessed over the following time scales:

1. A transient state lasting for a few minutes, that is, the person's immediate state of mind.
2. An episodic condition lasting for weeks or months, such as the distress provoked by a life event or a short-lived clinical disorder.
3. A personality trait that may show stability over decades, such as trait anxiety.

The narrowest operationalizations of distress focus on negative affects such as anxiety and depression. Spielberger pioneered questionnaires that assess these constructs at the state and trait levels. Depending on the context, other indices of distress may be used in tandem with scales for negative affect. Distress may also describe cognitive responses, such as worry, lowered self-esteem, and lack of perceived control. In life event research, indices of distress may include difficulties in social functioning and health problems. Specialized measures are used in specific contexts, such as scales for marital distress, trauma symptoms, and health concerns. Somatic distress measures have been developed on the basis that self-reported medical symptoms tend to intercorrelate. Hence, different measures of distress probably relate to different underlying constructs.

Distress may be defined more precisely through multivariate psychometric research. Studies of mood have identified orthogonal dimensions of negative affect and positive affect, such that negative affect is a superordinate category relating to tension, unhappiness, irritability, and other negative affects, whereas positive affect covers states such as happiness and energy. The tripartite model of Mineka, Clark, and Watson proposes that a general factor of distress is common to mood and anxiety disorders. Specific disorders are differentiated by two further factors of anhedonia (characteristic of depression) and hyperarousal (elevated in panic disorders). The model has generally been well supported in studies of patient groups and by confirmatory factor analyses, although some studies have failed to confirm its predictions regarding the characteristics of specific disorders.

Other research has explored the expression of distress as a transient state. Most simply, distress may be identified with negative affect; many studies of mood discriminate negative and positive affect as fundamental dimensions. However, the state model of Matthews defines distress more precisely as an affective-cognitive factor, defined by tension, unpleasant mood, and cognitions of lack of control and low confidence. Distress is distinct from the two further broad state factors of task engagement (e.g., energy and task motivation) and worry (e.g., self-focused attention and intrusive thoughts).

Influences on Distress

Distress reflects both situational influences, such as life events, and intrapersonal influences, such as personality traits. This section reviews situational influences, investigated either through controlled experiments or correlational studies of real-life stressors. Events that threaten or damage the person's well-being often provoke distress. Environmental factors promoting distress include (1) traumatic events, (2) physical factors such as loud noise, (3) social factors such as criticism by others, and (4) ill health. The adverse effects of agents of these kinds are well documented, although there is considerable variation in their effects across individuals and across different occasions. However, there are competing physiological, cognitive, and social accounts of the mechanisms that link these external stressors to psychological stress responses.

Physiological Influences

Historically, the distress concept has been important to physiological research, for example, as a concomitant of Selye's GAS. Physiological studies of distress focus primarily on the brain mechanisms that may

output and regulate negative affect. Evidence for a biological basis for distress is provided most directly by studies of brain damage in humans and animals. For example, damage to the amygdala may provoke extreme emotional responsivity. Learning studies link the central nucleus of this structure to fear conditioning, such that it may function as the brain's emotional computer. In humans, damage to the frontal lobes seems to provoke disinhibition syndromes in which the disruption of emotional response is accompanied by the loss of behavioral control. Pharmacological studies allow negative affects to be linked to the neurotransmitters sensitive to the drug concerned. Different negative emotions may be controlled by different brain systems: anxiolytic drugs operate through benzodiazepine receptors, and anti-depressants such as fluoxetine (Prozac) operate through serotonergic pathways. Indeed, brain-imaging studies show that distress correlates with the activation of multiple cortical and subcortical areas. The key issue is the functional organization of these different sites, but it remains controversial whether distress can be linked to a single brain system associated with response to punishment. Some animal models of basic emotions replace the concept of general distress with an account of multiple brain systems for specific emotions including anger, fear/anxiety, and disgust.

Davidson links general distress to a withdrawal system supported by multiple brain structures including the right prefrontal cortex and amygdala. However, anxiety and depression are differentiated by other frontal systems, so that anxiety relates to increased right anterior activity and depression relates to increased right frontal activity. These hypotheses have been tested, with mixed success, in studies using the electroencephalogram (EEG). Another line of evidence is provided by psychophysiological studies that link anxiety states to the fight-or-flight reaction and to arousal of the sympathetic branch of the autonomic nervous system.

Contemporary studies aim just to establish correspondences between subjective affect and brain processes, but they also aim to establish that affective responses are embedded within a functional system that controls the organism's adaptation to a particular type of stimulus or event. Gray and his colleagues developed a detailed animal model that distinguishes three such systems. In the most recent version of this reinforcement sensitivity theory (RST), the systems are as follows. A fight-flight-freeze system (FFFS) is activated by signals of punishment and nonreward and may elicit various avoidance behaviors. A behavioral activation system (BAS) is activated by reward and nonpunishment, producing approach behaviors.

A behavioral inhibition system (BIS) is activated by conflict between FFFS and BAS, generating anxiety and heightened attention to threat. Hence, subjective distress is a concomitant of the adaptive processes linked to both the FFFS and BIS, which tend to operate jointly in managing response to punishment. There is extensive evidence on the effects of anxiolytic drugs in animals that supports the model, but its application to human emotional distress and anxiety remains contentious.

Cognitive Influences

Cognitive models of stress are supported by experimental studies demonstrating that both the psychological and physiological impact of stressors are moderated by the person's beliefs and expectancies. For example, negative moods may be induced by suggestive techniques such as reflecting on negative events or making statements that one is unhappy. Distress in the workplace reflects not just work demand but also perceived control and decision latitude. Appraisal theories of emotion seek to link affect to the person's evaluation of external stimuli and their personal significance. Anxiety may then relate to appraisals of personal threat and uncertainty, and depression may relate to uncontrollable personal harm. Appraisal is not necessarily accessible to consciousness, and there may be several distinct appraisal processes operating in tandem. Appraisal theories tend to focus on the differing cognitive antecedents of specific emotions rather than on undifferentiated distress. Studies of transient distress measured as a unitary state show that it relates to appraisals of threat, loss, and uncontrollability of events and to use of emotion-focused coping strategies such as self-criticism.

The transactional stress theory of Lazarus embeds appraisal within a wider matrix of person-environment interaction within which negative emotions are tied to core themes describing the person-environment relationship. Core relational themes also relate to action tendencies. Transactional theory emphasizes the person's active attempts to cope with threatening or damaging events. In general, distress is likely to develop when the people appraise themselves as failing to cope adequately and as lacking in personal control over significant events. Contemporary theory also emphasizes the differing relational themes associated with different negative emotions. The dynamic nature of the stress process is supported by evidence from longitudinal studies for reciprocal relationships between distress and life events and between distress and health perceptions. In cases of chronic distress, negative cognitions drive distress responses that in turn feedback into further

negative cognitions. This harmful feedback loop may be accentuated by metacognitions that maintain the focus of attention onto stress symptoms and reinforce negative self-referent beliefs.

Social Influences

The disruption of social relationships associated with bereavement, marital discord, and unemployment are among the most potent factors eliciting distress. Conversely, the availability of social support often functions to alleviate stress responses. Social psychologists characterize as inadequate attempts to locate distress solely within the individual. Instead, distress may reflect people's discourses about their problems and their interpersonal interactions within a social and cultural context. For example, Oatley has attributed depression to loss of socially defined roles. Social psychologists also emphasize the importance of the expression and display of emotion that reflects social motivations and norms.

Personality and Distress

Neuroticism and Negative Affect

Some individuals appear to be generally more distress-prone than others. Psychometric and experimental studies identify stable personality traits that relate to the person's predisposition to experience negative emotion. Eysenck identified a dimension of neuroticism associated with a low threshold for emotional arousal. A similar dimension corresponds to one of the factors described by the five factor model of personality. Studies of temperament in infants and children, based on observations of behavior, also identify negative affectivity or distress-proneness as a fundamental dimension.

Neuroticism predicts negative moods such as depression and anxiety, although the strength of the relationship varies with situational stressors. It also predicts vulnerability to worry and disturbances of cognition. The association between neuroticism and distress is substantiated by both controlled experimental studies and field studies of everyday mood using diary or experience-sampling methods. Neuroticism is also elevated in patients suffering from emotional disorders. Longitudinal studies suggest that neuroticism is a factor predisposing clinical disorder, although there is some reciprocity between the personality trait and disorder. Elevated neuroticism may also be an outcome or a scar of mental illness. In addition, both the Eysenck and five factor models include an extraversion-introversion factor. Extraversion relates primarily to happiness and positive affect but also correlates negatively with distress

measures to some extent. One influential view is that neuroticism is essentially a dimension of negative affectivity, whereas extraversion may be identified with positive affect, although this perspective may be oversimplified.

Neuroticism also relates to self-reported somatic distress and medical symptoms, including various psychosomatic conditions. It remains controversial whether more neurotic individuals are genuinely more prone to illness or whether they are just prone to complain about physical symptoms. However, a causal role for neuroticism is suggested by growing evidence that various negative affects, including depression, anger, and anxiety, operate as risk factors for illnesses, including coronary heart disease, although the effect sizes for the association between distress and illness are often small.

Current personality models generally adopt an interactionist or diathesis–stressor perspective, such that the distress response depends on the interaction of personal and situational factors. Hence, although more neurotic individuals are more vulnerable to distress, the extent to which they experience greater distress than less neurotic people on a given occasion varies with situational factors. A more subtle view of neuroticism is also suggested by life event research. Consistent with the negative affectivity hypothesis, more neurotic individuals seem to experience life events as more distressing. However, neurotic people also seem to experience a higher frequency of life events, which may reflect behavioral difficulties in adapting to life circumstances. For example, neuroticism relates to a higher frequency of interpersonal conflicts. Thus, there are at least three causal paths that may link neuroticism to distress: (1) high neurotics may be generally distress-prone, independent of external circumstances; (2) high neurotics may react to external events with elevated distress; and (3) high neurotics may behave in ways that elicit a higher frequency of distressing events. The dynamic nature of distress response is demonstrated also in studies of temperament in children. Overly distress-prone infants may be perceived as whiny or clingy, eliciting parental annoyance or neglect, which provokes further dysregulative responses from the child.

Other Traits

Various more narrowly defined traits are associated with vulnerability to distress. Spielberger's anxiety and depression measures are among the best known. According to Spielberger, trait anxiety represents a predisposition to experience transient state anxiety in threatening environments, but, consistent with interactionism, the trait-anxious individual does not experience state anxiety on all occasions.

Experimental studies suggest that trait anxiety correlates with state anxiety under conditions of ego threat, such as personal criticism. Negative affect is central to the definition of the Spielberger traits. Other traits such as optimism/pessimism, self-focus of attention, self-efficacy, and attributional style are cognitively defined, but correlate nevertheless with distress indices.

Distress-related traits tend to intercorrelate with one another and also with broader measures of neuroticism and negative affectivity. One view is that the various traits predict distress because they overlap with neuroticism and that the narrower trait constructs may add little to the broad trait. It is plausible that neglect of neuroticism as a confound of other traits has added unduly to the complexity of the field. Nevertheless, there is evidence for discriminative validity for traits such as optimism/pessimism and self-focus of attention; distress is best predicted from multiple traits.

Contextualized Traits

Neuroticism and other traits appear to act as generalized factors predisposing distress across a variety of contexts. However, in many cases people possess traits linked to specific contexts that predict distress over and above more generalized traits. A paradigmatic case is test anxiety. Sarason has developed scales that ask respondents specifically about their thoughts and feelings when required to take tests and examinations. People show stable individual differences in vulnerability to test anxiety that predict how they feel and perform in examination settings. Researchers have focused on evaluative anxiety in various other settings, using measures of social anxiety, math anxiety, computer anxiety, and so forth. Such measures typically correlate with neuroticism, but have better predictive validity in the appropriate context. The contextualized trait may reflect both neuroticism and the person's more specific stable cognitions of the context concerned. Much of this work has focused on anxiety as a distress symptom, but it is likely that contextualized traits predict other distress symptoms in a similar way. For example, individuals vulnerable to driver stress are prone to experience both anxiety and unhappiness during vehicle operation.

Psychological Explanations for Distress Vulnerability

There are competing explanations for individual differences in vulnerability to distress. Biological and cognitive models have been most fully articulated. Social psychological explanations are also relevant, although progress has been retarded by the reluctance

of social psychologists to accept the validity of traits. Biological explanations are bolstered by demonstrations from kinship and adoption studies that distress-related traits are partially inherited. The tendency for distress vulnerability to run in families relates more to genetic predisposition than to the influence of the common family environment. There is also an extensive literature on psychophysiological and biochemical correlates of neuroticism and other traits, although results are often weak or inconsistent. Hence, the pathways connecting genes for distress-prone personality to specific brain mechanisms are poorly understood, although various theories have been advanced. Eysenck claimed that neuroticism related to the sensitivity of the limbic system to emotional arousal; Gray's RST links trait anxiety more specifically to the BIS and (possibly) FFFS systems previously described.

Cognitive explanations are supported by evidence that traits relate to the constructs of the transactional model of stress, especially appraisal and coping. Many studies have shown that neuroticism, trait anxiety, and trait depression relate to a distinctive style of cognition of environmental demands. Individuals with these characteristics are prone to a variety of negative self-appraisals; they appraise the world as demanding and threatening and appraise themselves as ineffective in dealing with external demands. More neurotic individuals are prone to coping with potentially stressful events through emotion-focused strategies such as ruminating on their problems or criticizing themselves. Such strategies are frequently maladaptive and may sometimes even exacerbate the problem and contribute to clinical disorder. To a lesser degree, neuroticism also relates to avoidance of the problem and to reduced task-focused coping.

Psychological Concomitants of Distress

Distress has various behavioral correlates including the impairment of objective performance, bias in selective attention, and overt clinical symptoms. These findings are explained by theories of self-regulation pioneered by Carver and Scheier and elaborated as a general theory of distress by Wells and Matthews. In these models, self-regulation operates homeostatically to reduce the discrepancies between actual and preferred status. Distress signals self-discrepancy and the need for coping efforts to reduce self-discrepancy and maintain adaptation to the external environment. The regulatory system may have physiological as well as psychological expressions. It also operates within a social context in that interactions with others provide strong cues toward actual status and in that societal and cultural norms are reflected in part by

ideal status. Experimental studies of attention and performance identify biases in the cognitions supporting self-regulation, which may be the source of clinical disorders associated with distress.

Performance Impairment

Trait and state distress measures are associated with slower and/or less accurate performance on a variety of tasks. Both anxiety and depression tend to be associated with performance impairment. Contextualized distress measures are also typically associated with performance impairments within the context concerned; for example, test anxiety predicts poorer examination performance. Anxiety studies show that it is the cognitive rather than the emotional aspects of the anxiety state that interfere with performance. Self-referent worries appear to be especially damaging because attention or working memory may be diverted from the task at hand to personal preoccupations. The detrimental effects of depression have been similarly ascribed to worry and rumination. Loss of performance may reflect loss of attentional resources, controlled processing, or working memory.

The impact of distress on performance depends on motivational factors. Performance may be maintained or even enhanced when good performance is appraised as instrumental in relieving distress through compensatory effort. Conversely, decrements are more likely when motivation and effort are low. The maintenance of performance effectiveness through effort is more likely in anxiety than in depressed states. It appears that impairment is not a fixed cost of distress but a consequence of a style of self-regulation that promotes worry, self-focus of attention, and emotion-focused coping at the expense of performance goals.

Cognitive Bias

Distress is often associated with enhanced processing of negative stimuli. Such effects have been demonstrated in studies of experimentally induced moods, in studies comparing the performance of distressed patients with controls, and in studies comparing non-clinical groups selected for high and low distress. Some of the most reliable effects are found in studies of anxiety and selective attention. Anxious individuals are slow to name the color in which threat-related words are printed (the emotional Stroop effect), and they also appear to attend preferentially to threat stimuli in other paradigms. Other distressed groups are similarly biased toward negative stimuli congruent with their source of distress. These effects may reflect either an automatic bias toward threat stimuli in vulnerable individuals, deliberately chosen strategies, or both voluntary and involuntary processes. Biases in

memory are most reliable in studies of depression and negative mood, although comparable anxiety effects are also documented. Depressed individuals tend to show enhanced recall for negative events or stimuli at the expense of positive events. These effects seem most reliable when the test of memory is explicit rather than implicit and when some active processing of the memorized material is required. Various other tasks, including judgment and decision making, also show cognitive bias. From the self-regulative perspective, distress may relate to the extent to which the person balances the goal of maintaining awareness of threat against a focus on the immediate task at hand.

Clinical Disorder

Distress is a central symptom of both mood and anxiety disorders. A simple view of clinical distress is that it represents the upper extreme of a continuum of vulnerability to negative emotion. However, clinical disorder involves qualitative abnormalities of behavior in addition to distress, such as avoidance of challenging situations, self-harm, and difficulties in social interaction. Distress in clinical depression and anxiety is bound up with problems in behavioral adaptation to everyday life. Self-regulative approaches to understanding distress contribute to understanding the maladaptive nature of these disorders. Central to both anxiety and depression are stable but false negative beliefs about the self, which may be represented as schemas. The anxious person believes him- or herself to be especially vulnerable to threats of certain kinds, whereas the depressed patient experiences cognitions of lack of self-worth and hopelessness. Cognitions and behaviors differ across clinical conditions, but, in general, unrealistic negative self-beliefs feed into the faulty appraisal of the person's place in the external world, choice of ineffective coping strategies such as worry, and consequent distress.

The view of distress as a sign of maladaptive self-regulation has implications for therapy. In mild sub-clinical distress conditions, it may be sufficient for individuals to learn specific coping skills that permit them to manage the specific situations that provoke distress. For example, individuals with mild social anxiety may be helped by explicit instruction in social skills. Distress is alleviated both through individuals' increased confidence and positive self-appraisal and through the likelihood that the individuals' greater skill will lead to more positive outcomes in the problematic situation. In more severe cases, stress management techniques such as skills training fail to address the underlying cognitive distortions that drive both distress and behavioral problems. Cognitive-behavior therapy seeks to uncover and modify faulty cognitions through a variety of techniques.

Empirical studies show that clinical improvement in behavior and affect is accompanied by a decline in cognitive bias.

See Also the Following Articles

Alarm Phase and General Adaptation Syndrome; Anxiety; Depression and Manic-Depressive Illness.

Further Reading

- Carver, C. S. and Scheier, M. F. (1981). *On the self-regulation of behavior*. New York: Cambridge University Press.
- Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Springer.
- Matthews, G., Campbell, S. E., Falconer, S., et al. (2002). Fundamental dimensions of subjective state in performance settings: task engagement, distress and worry. *Emotion* 2, 315–340.
- Matthews, G., Deary, I. J. and Whiteman, M. C. (2003). *Personality traits* (2nd edn.). Cambridge, UK: Cambridge University Press.
- McNaughton, N. and Corr, P. J. (2004). A two-dimensional neuropsychology of defense: fear/anxiety and defensive distance. *Neuroscience and Biobehavioral Reviews* 28, 285–305.
- Oatley, K. (1992). *Best-laid schemes: the psychology of emotions*. Cambridge, UK: Cambridge University Press.
- Rothbart, M. K. (2004). Temperament and the pursuit of an integrated developmental psychology. *Merrill-Palmer Quarterly* 50, 492–505.
- Selye, H. (1976). *The stress of life*. New York: McGraw-Hill.
- Shankman, S. and Klein, D. N. (2003). The relation between depression and anxiety: an evaluation of the tripartite, approach-withdrawal and valence-arousal models. *Clinical Psychology Review* 23, 605–637.
- Spielberger, C. D., Reheiser, E. C., Owen, A. E. and Sydeman, S. J. (2004). Measuring the psychological vital signs of anxiety, anger, depression, and curiosity in treatment planning and outcomes assessment. In: Maruish, M. E. (ed.) *The use of psychological testing for treatment planning and outcomes assessment. Vol. 3: Instruments for adults* (3rd edn., pp. 421–447). Mahwah, NJ: Lawrence Erlbaum.
- Suls, J. and Bunde, J. (2005). Anger, anxiety, and depression as risk factors for cardiovascular disease: the problems and implications of overlapping affective dispositions. *Psychological Bulletin* 131, 260–300.
- Thayer, R. E. (1996). *The origin of everyday moods*. Oxford: Oxford University Press.
- Watson, D., Gamez, W. and Simms, L. J. (2005). Basic dimensions of temperament and their relation to anxiety and depression: a symptom-based perspective. *Journal of Research in Personality* 39, 46–66.
- Wells, A. and Matthews, G. (1994). *Attention and emotion: a clinical perspective*. Hove, UK: Erlbaum.
- Zeidner, M. (1998). *Test anxiety: the state of the art*. New York: Plenum Press.

Diurnal-Nocturnal Stress in the Workplace *See: Sleep Loss, Jet Lag, and Shift Work.*

Diving, Stress Effects of *See: Pressures, Effects of Extreme High and Low.*

Divorce, Children of

K N Hipke

Wisconsin Psychiatric Institute and Clinics, Madison, WI, USA

S A Wolchik and I N Sandler

Arizona State University, Tempe, AZ, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by K N Hipke, I N Sandler and S A Wolchik, volume 1, pp 730–733, © 2000, Elsevier Inc.

Divorce-Related Stressors

Mental Health Consequences

Risk and Resilience

Implications for Prevention

Glossary

<i>Effect size</i>	A statistical indicator, expressed in standard deviation units, of the magnitude of the relation between two variables.
<i>Externalizing</i>	A constellation of psychological symptoms indicative of acting-out behavior, such as aggression and conduct problems.
<i>Internalizing</i>	A constellation of psychological symptoms experienced internally, such as depression and anxiety.
<i>Meta-analysis</i>	A quantitative method of literature review in which measures of central tendency (e.g., the mean) and variation are obtained for outcomes examined across multiple studies.
<i>Moderator</i>	A variable that affects the direction and/or strength of the relation between an independent or predictor variable (e.g., divorce) and a dependent or criterion variable (children's mental health problems).

Divorce-Related Stressors

Parental divorce has become a common experience for children, particularly in the United States, where divorce is more common than in any other developed country. The rate of divorce in the United States has increased over the course of the twentieth century, with a particularly rapid rise beginning in the decades of the 1960s and 1970s followed by a modest decline at the close of the century. Currently, close to half of all marriages end in divorce, affecting approximately 1.5 million children a year. Although divorce rates vary among families of different demographic profiles, between 30 and 40% of all children are expected to live in a divorced, single-parent home by the age of 16.

Parental divorce is not a singular event. Rather, it marks a multitude of potential stressful changes and disruptions in children's social and physical environments that occur before, during, and after parental separation. For example, interparental conflict often begins many months and/or years prior to separation, can escalate around separation and/or divorce, and can continue well after separation as parents negotiate new parenting-related issues such as child custody, visitation arrangements, and child support. As in the case of interparental conflict, many of the stressful changes and experiences faced by children from divorced families occur in the family domain, involve interpersonal loss and/or conflict, and are not easily controllable by children.

Several of these changes occur in children's relationships with parents. Despite an increasing emphasis on joint custody in many states, over 80% of children from divorced families still reside primarily with their mothers. Consequently, parental separation often means the loss of daily contact with fathers. Moreover, for most children contact with non-custodial fathers becomes increasingly infrequent

over time. Children with continued father involvement may face the challenge of residing in two households, with different rules and discipline practices. The quality of relationship between fathers and children may also change following divorce, as fathers try to adapt to a parenting role when they do not reside with their children. Considerable research indicates that a positive relationship, including warm and supportive interactions as well as appropriate discipline, is an important factor affecting children's well-being following divorce.

Children's relationships with their mothers also change. Custodial mothers report high levels of stress as they negotiate the dual task of adjusting to the dissolution of marriage as well as to a new lifestyle with increased responsibilities. Juggling single parenthood, work-related demands, legal matters, and the re-negotiation of social relationships places custodial mothers at increased risk for psychological distress and physical illness. Maternal stress and distress, in turn, have been linked to disruptions in parenting behavior. Following divorce, many custodial mothers are less supportive, affectionate, and consistent in their discipline practices and engage in more conflict with children.

Many newly divorced families also face financial stress. Both mothers and fathers may experience a decline in standard of living following divorce. Income loss can result in potentially stressful changes for children, including moving to a different home, changing neighborhoods, and changing schools. While these transitions involve the loss of material possessions, children also face the loss of friends, teachers, and other important members of their social support networks.

When asked about the various changes that arise following divorce, children, parents, and clinicians who work with children from divorced families agreed that the most stressful experiences children face occur in the nuclear family domain. These include being blamed for the divorce, verbal and/or physical interparental conflict, and derogation of one parent by another. Other experiences perceived as highly stressful by children include derogation of parents by relatives or neighbors, having to give up possessions, and maternal psychological distress.

For many children, parental divorce is the first of several family transitions, each of which brings a new set of stressful changes. Forty percent of divorced couples remarry before their youngest child reaches age 18. Given that remarriages are also more likely than first marriages to end in divorce, a sizeable number of children face multiple family transitions over the course of their childhood. Although

remarriage can improve the financial and emotional support provided to parents, it requires complex family reconfiguration. For children, this includes the addition of a stepparent, as well as other stepkin relations such as stepsiblings, grandparents, and other extended family members. When additional children are born into the family following remarriage, half-sibling relationships are created. The incorporation of new family members alters family dynamics, including relationships with biological parents. In addition, children may face physical transitions such as an additional residential move or sharing of a room with other children.

Mental Health Consequences

The findings from five decades of empirical research have consistently demonstrated that children from divorced families are at increased risk for adjustment or mental health problems relative to children from continuously married families. When this body of literature is summarized using meta-analysis, the effect sizes are small to moderate, with the largest group difference occurring in the domain of externalizing problems or conduct problems. Children from divorced homes, on average, engage in more misbehavior, are more aggressive, and/or have higher rates of delinquency. Children from divorced homes also show greater difficulties in the areas of psychological adjustment (i.e., depression, anxiety, happiness), social adjustment, self-concept, and school achievement.

The small to moderate effect sizes have caused some researchers to question the significance of parental divorce as a risk factor for adjustment problems. However, there are several reasons why divorce should be considered an important risk factor with significant population-level implications. First, given the high prevalence of divorce in society, small to moderate effects on child maladjustment have a substantial impact on population rates of problem outcomes for children from divorced families. Second, trends in studies from the 1990s show that effect sizes have increased from the 1980s. Third, a sizeable subset of children from divorced families have significant mental health problems. Approximately 20–35% of children from divorced families experience clinically significant mental health problems, a rate at least twice that of children from two-parent families. Fourth, adolescents from divorced homes engage in more high-risk behavior with potentially long-ranging consequences. Adolescents from divorced homes are more likely than those from two-parent homes, for example, to drop out of school, become

sexually active or pregnant at an earlier age, and use drugs or alcohol.

Finally, the effects of divorce for many children and adolescents are not transitory. Longitudinal studies indicate that children from divorced families are at elevated risk for difficulties several years after parental separation, with problems continuing throughout adulthood. Across studies, adults who experienced divorce in childhood or adolescence have been found to function more poorly on a range of indicators including depression, life satisfaction, marital quality, educational attainment, income, occupational prestige, and health outcomes such as illness and mortality. Adults who experienced divorce as children or adolescents are also more likely to divorce themselves. Researchers have speculated that these long-term consequences are the result of disruptions in normal development during childhood and adolescence.

Risk and Resilience

Although children from divorced families are at elevated risk for adjustment problems, there is tremendous variability in response. The majority of children in divorced families show no adverse effects relative to children raised by continuously married families. Thus, research on children from divorced families has shifted over time from examining the mental health and adjustment consequences of divorce toward understanding why some children are more adversely affected by divorce than others. Researchers have identified several risk and protective factors that act as moderators between divorce and children's mental health, including stress, parent-child relationships, and child characteristics.

Exposure to Stressors

Greater exposure to stressful divorce-related stressors is associated with increased mental health problems among children. The frequency of divorce-related changes faced by children and families, as assessed by divorce events checklists, relates positively to difficulties in the domains of internalizing and externalizing symptoms, self-concept, and social and academic competence. Children exposed to more divorce-related stressors also show maladaptive cognitions, including a reduced sense of their ability to affect their environments.

Studies have also examined the relations between specific types of divorce stressors and children's mental health. More limited post-divorce economic resources have been consistently related to increased maladjustment. Exposure to interparental conflict

is one of the most potent determinants of children's post-divorce adjustment problems. A positive association between intense marital conflict and child mental health problems has been well documented. Researchers increasingly believe that children in families marked by intense, chronic marital conflict and/or violence may be better off in the long run if their parents divorce and the degree of conflict is reduced (although a decrease in conflict does not always occur). Interestingly, in divorcing families in which overt conflict was very low or absent prior to separation, children are also at risk for poor adjustment, perhaps due to the unexpected nature of the dissolution. In either scenario, post-divorce conflict that involves putting a child in the middle by, for example, one parent denigrating the other parent or asking the child to carry hostile messages, is associated with higher levels of child depression and anxiety.

Parent-Child Relationships

The degree to which the quality of children's relationships with their parents is disrupted following divorce is predictive of adjustment problems. Children whose custodial mothers provide high levels of warmth, affection, and effective, consistent discipline following divorce exhibit fewer behavior problems, higher self-esteem, and better academic performance than children with low-quality mother-child relationships. Experimental research has shown that intervention-induced increases in maternal warmth and effective discipline mediate program effects to reduce children's mental health problems. Moreover, high-quality mother-child relationships serve as a buffer, mitigating the negative effects of divorce-related stressors on children's adjustment problems.

Because fathers are more likely to be noncustodial parents, research initially focused upon whether the amount of father-child contact affected children's adjustment after divorce. Findings from this line of research were inconclusive. It is now understood that it is the quality and not quantity of father-child contact that is important to divorce adaptation. High-quality father-child relationships are related to better child post-divorce adjustment than low-quality relationships.

Child Characteristics

Research examining the association between child demographics, such as age and gender, and post-divorce adjustment has revealed few consistent findings. Although some studies, particularly older studies or those examining conduct problems, indicate less advantageous outcomes for boys relative to girls,

meta-analyses show that divorce is associated with a range of poor outcomes for both boys and girls. Similarly, while meta-analysis has also shown that children in middle childhood and adolescence appear to be more negatively impacted by divorce than pre-schoolers or college students, far fewer studies have examined divorce adjustment in these latter two age groups, making it difficult to draw firm conclusions regarding the effect of age on children's adaptation to divorce.

Recently, a body of literature has begun to emerge that shows that the manner in which children interpret and cope with divorce-related stressors has implications for their adjustment. Children who lack an internal sense of control over their environment or who have a tendency to interpret divorce-related stressors in a characteristically negative style or as threatening to their well-being are more likely to experience post-divorce adjustment problems. Also, children who use positive cognitive or behavioral coping strategies, such as active problem solving or positive cognitive restructuring, experience fewer psychological adjustment problems than children who employ avoidant coping strategies. Theoretically, researchers agree that the adaptive value of specific coping strategies depends on context. Children may benefit most, for example, by cognitively reframing a situation that is largely outside of their control, while problem solving may be an adaptive strategy for situations that are more controllable. There is also growing evidence that the effects of using appropriate coping strategies may be mediated by increasing children's sense of efficacy to deal with the problems they encounter in their lives.

Children's temperament, which theoretically modulates the expression of activity, emotionality, reactivity, and behavior, has also been implicated as important to children's post-divorce adjustment, although this area of research is currently small. Findings indicate, for example, that children rated as more emotionally reactive are at increased risk for maladjustment compared to children who are less reactive and that this effect may be mediated by leading to an increase in negative appraisals of the threat involved in stressful situations.

There is also evidence that the effects of parent and environmental factors are not independent of child factors, but rather are mediated through their effects on children's beliefs about their relations to the world and their ability to cope. For example, the effects of poor parenting quality on children's mental health appears to be partially mediated by increasing children's fears that they will not be cared for by anyone. Similarly, the effects of interparental conflict (a stressor over which children have little control) on

children's mental health is in part mediated by decreasing children's coping efficacy beliefs.

Implications for Prevention

Efforts to facilitate adaptation to divorce and prevent maladjustment have focused on altering modifiable contextual, family, and/or child variables demonstrated to affect children's divorce adjustment. The most widely available preventive interventions are child-focused programs that are often provided in school settings. Such interventions generally seek to increase children's accurate understanding of divorce and teach new skills for coping with divorce-related stressors. Preventive programs that target custodial parents focus on teaching effective parenting skills to enhance the quality of parent-child relationships and help parents shield children from stressful experiences such as interparental conflict. Among programs that have been rigorously evaluated using experimental and/or quasi-experimental trials, results have been positive. Child-focused programs have been successful in improving child mental health outcomes, including internalizing and externalizing symptoms, with program effects still present at follow-up studies 1 to 2 years later. Similarly, prevention programs for custodial mothers have been successful in improving the quality of the mother-child relationship, effective discipline, and child mental health outcomes. Recent evidence shows that program effects last well into adolescence, leading to reductions in mental health problems, drug and alcohol use, school failure, and risky sexual behavior and improvements in academic performance and competence.

See Also the Following Articles

Adolescence; Childhood Stress; Marital Conflict; Marital Status and Health Problems; Posttraumatic Stress Disorder in Children.

Further Reading

- Amato, P. R. (2000). The consequences of divorce for adults and children. *Journal of Marriage and the Family* 62, 1269-1287.
- Amato, P. R. (2001). Children of divorce in the 1990's: an update of the Amato and Keith (1991) meta-analysis. *Journal of Family Psychology* 15, 355-370.
- Amato, P. R. and Keith, B. (1991). Parental divorce and adult well-being: a meta-analysis. *Journal of Marriage and the Family* 53, 43-58.
- Grych, J. H. and Fincham, F. D. (1997). Children's adaptation to divorce: from description to explanation. In: Wolchik, S. A. & Sandler, I. N. (eds.) *Handbook of children's coping: linking theory and intervention*. New York: Plenum Press.

- Haine, R. A., Sandler, I. N., Wolchik, S. A., Tein, J.-Y. and Dawson-McClure, S. R. (2003). Changing the legacy of divorce: evidence from prevention programs and future directions. *Family Relations* 52, 397–405.
- Hetherington, E. M., Bridges, M. and Isabella, G. M. (1998). What matters? What does not? Five perspectives on the association between marital transitions and children's adjustment. *American Psychologist* 53, 167–184.
- Kelly, J. B. and Emery, R. E. (2003). Children's adjustment following divorce: risk and resilience perspectives. *Family Relations* 52, 352–362.
- Pedro-Carroll, J. (2005). Fostering resilience in the aftermath of divorce: the role of evidence-based programs for children. *Family Court Review* 43, 52–63.
- Sandler, I. N., Wolchik, S. A., Davis, C., Haine, R. and Ayers, T. (2003). Correlational and experimental study of resilience in children of divorce and parentally bereaved children. In: Luthar, S. L. (ed.) *Resilience and vulnerability. Adaptation in the context of childhood adversities*. Cambridge, UK: Cambridge University Press.
- Wolchik, S. A., Sandler, I. N., Millsap, R. E., et al. (2002). Six-year follow-up of a randomized, controlled trial of preventive interventions for children of divorce. *Journal of the American Medical Association* 288, 1874–1881.
- Wolchik, S. A., Sandler, I. N., Winslow, E. and Smith-Daniels, V. (2005). Programs for promoting parenting residential parents: moving from efficacy to effectiveness. *Family Court Review* 43, 65–80.

Domestic Violence

B Donohue, H Hill and T Maier-Paarlberg

University of Nevada at Las Vegas, Las Vegas, NV, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by T Maier-Paarlberg and B Donohue, volume 1, pp 734–738, © 2000, Elsevier Inc.

Description

Consequences

Remediation Strategies

Historical Perspectives

Glossary

<i>Domestic violence</i>	Any act of maltreatment occurring in the home environment of the victim perpetrated by a person who assumes the role of a caregiver or who is in an established relationship with the victim.
<i>Neglect</i>	The omission of appropriate caretaking functions that leads to a lack of self-preservation and growth in the dependent individual.
<i>Physical abuse</i>	A harmful act carried out on an individual with the intention of causing physical pain or injury.
<i>Sexual abuse</i>	Illegal, nonconsensual, or socially inappropriate contact or exposure of genitalia.
<i>Verbal/emotional/mental abuse</i>	The use of derogatory statements that cause the victim psychological pain or feelings of devaluation.

Description

Domestic violence includes all aspects of abuse and neglect that are perpetrated within the victim's home environment by a person who assumes the role of a caregiver or who is in an established relationship with the victim. Abuse and neglect have been around since the beginning of civilization; however, various forms of abuse have only recently been defined and studied extensively. Although definitions of abuse vary considerably, neglect and three types of abuse (physical abuse, sexual abuse, and psychological/mental/emotional/verbal abuse) have consistently been recognized in the domestic violence literature. This article provides an overview of domestic violence, including its risk factors, its prevalence, and a delineation of neglect and abuse.

Neglect

In general, neglect refers to the omission of appropriate caretaking functions that results in harm to the dependent individual. Victims of neglect are therefore dependent on a guardian or caregiver for self-preservation and growth. Most often, victims of neglect include children, adults who evidence severe mental or physical disabilities (e.g., mental retardation and Parkinson's disease), and elderly people with medical complications or other problems associated with old age. Prevalence studies consistently indicate that neglect is the most frequently occurring form of maltreatment. For instance, the National Clearinghouse on Child Abuse

and Neglect reported that an estimated 906,000 children were determined to be victims of child abuse in 2003 and that, of those claims, more than 60% experienced neglect. Although risk factors are varied, they most often include substance abuse, physical disabilities, social isolation, psychiatric illness, social isolation, stress, poor parenting skills, and unrealistic expectations of the child's development. People of lower economic status appear to be overrepresented in abuse samples, perhaps because abuse is more common for these individuals but also because they may be reported and charged with abuse more often due to reasons associated with poverty (e.g., relatively inadequate representation by legal counsel). Incidents of neglect include interfering with (or delaying) medical protocol, leaving small children unattended for extended periods of time in potentially dangerous situations (e.g., being left in an automobile while the caregiver drinks alcohol inside a bar), restricting or failing to provide sufficient nutrients or shelter, failing to repair home hazards (e.g., exposed electrical wires and broken steps), exposing victim to toxins or bacteria (e.g., failing to clean dog feces from the kitchen floor where an infant crawls), failing to provide adequate affection, and restricting children from attending school. Indicators of neglect include dirty and unkempt bodies, rotten teeth, clothes that are too small or are in need of repair, and untidy and disheveled home environments.

Psychological Abuse

Psychological abuse, also known as verbal, mental, or emotional abuse, typically refers to derogatory statements that cause harm to or interfere with the psychological adjustment of the victim. These statements are most often attacks on the victim's competence ("You can't do anything right; you're just a big dummy") or character (e.g., "You're a lazy, unpopular, good for nothing"). Psychological abuse is usually associated with obscenities, negative voice tones, exploitation, encouraging corruption and delinquent behavior, excessive teasing, harmful threats, ridicule, or derogatory statements about the victim or people whom the victim likes. There do not appear to be significant gender differences regarding the prevalence of psychological abuse because studies have consistently indicated that the vast majority of the population has experienced some level of psychological abuse from an adult family member in their lifetime. Interestingly, people often underestimate the severity of psychological abuse relative to other forms of maltreatment because its consequences are subtle and there are few, if any, legal sanctions that

mitigate against its use. Nevertheless, severe consequences of verbal abuse have been found, including depression, antisocial behaviors, low self-esteem, intellectual deficits, academic difficulties, health problems, shyness, problem-solving deficits, anxiety, and difficulties with conflict resolution and other social skills. Moreover, psychological abuse is often a precursor of physical abuse by the perpetrator and of retaliatory aggression by the victim.

Physical Abuse

Physical abuse is usually defined as any harmful act carried out on an individual with the intention of causing pain or injury (e.g., kicking, spanking severe enough to cause bruises, whipping with an extension cord, punching, biting, or threatening with a knife or a gun). Victims are usually people who are emotionally or financially dependent on the perpetrator (e.g., spouses, lovers, children, or the elderly). Although neglect is estimated to be the most frequently occurring form of maltreatment, physical abuse has probably received greater attention in the scientific literature. According to a 1998 Commonwealth Fund Survey, approximately 3 million women are physically abused each year by their intimate companions, sometimes resulting in fatalities. In 2003, the Bureau of Justice Statistics Crime data indicated that, on average, more than three women are murdered by their husbands or boyfriends each day in the United States. Approximately one-half of the adult victims of physical abuse retaliate with physical violence of their own due to self-defense or escalation of violence, lending support to the notion that physical abuse in adults is often reciprocal. Indeed, males in dependent relationships also suffer from physical abuse; however, it is relatively less often reported to authorities and the maltreatment is usually not as severe. Interestingly, the prevalence rates of physical abuse among homosexual adult relationships approximate those of heterosexual adult relationships.

Children are prime targets for physical abuse. Indeed, estimates of physical-abuse victimization in children are usually around 30%. The perpetrators of child physical abuse are often of lower economic status, substance abusers, single parents, socially isolated, and depressed; evince inadequate coping skills; and report a history of being maltreated as a child. Other risk factors include low birth weight, physical and mental disabilities, immaturity, child noncompliance and misconduct, impulsiveness, deficits in family cohesion, family conflict, history of abuse, limited positive interactions, unrealistic expectations and negative perceptions of children, inconsistent

parenting styles, lack of familiarity with the developmental norms of children, stress, lack of empathy, and antisocial personality. The number of studies conducted with elderly victims of physical abuse and other forms of maltreatment has recently increased, more so than in other age groups. Although this increased attention is due to multiple factors, several are fairly obvious. The human life span has increased, resulting in more elderly individuals who are living with age-associated medical problems and physical limitations. Changes in health care, among other factors, have forced older adults to reside with family who are often ill-equipped to tolerate the added stress associated with their care, which often results in frustration for both the victim and perpetrator. A recent study of nursing home residents revealed that aggression from residents toward the staff was the strongest predictor of elder abuse. Investigations of the risk factors for elderly abuse are consonant with those of other age groups.

Sexual Abuse

Sexual abuse is generally defined as any illegal, non-consensual, or inappropriate contact or exposure of genitalia, including sexual humiliation. However, legal definitions of sexual abuse vary considerably. For instance, some agencies (i.e., Federal Bureau of Investigation) have defined rape as attempted or completed vaginal intercourse with a female by force and against her will, whereas states have adopted unique qualifiers for this term, such as gender neutrality; anal or oral penetration; insertion of objects; and being unable to give sexual consent because of mental illness, mental retardation, or intoxication. Other common patterns of sexual abuse include, but are not limited to, being forced to observe or perform pornography and fondling or kissing the genitalia of a child or nonconsenting adult. Approximately 10% of marriages have involved rape without physical violence, and estimates from the National Clearinghouse on Child Abuse and Neglect have indicated that as much as 10% of reported abuse cases in the United States involved inappropriate sexual contact or exposure. Sexual abuse victimization may occur as early as infancy, and it is characteristically progressive; approximately half of sexually abused children will continue to be abused until they leave the abusive home. The perpetrators of domestic sexual abuse are very often in-laws, foster parents, nonrelatives, older siblings, or cousins, but they may also include other relatives living in the home of the victim. The majority of perpetrators of sexual abuse are male. However, female perpetrators may be more common than studies suggest, perhaps because their abuse is

not reported more often than that of males. A large percentage of rapes in college populations are committed by someone the victim knows. Factors that have been found to influence sexual abuse include neurological disorders, particularly frontal lobe abnormalities; sexual abuse of or repeated witnessing of pornography by the perpetrator during childhood; mental illness; substance abuse; and dysfunctional family systems.

Consequences

The consequences of abuse and neglect are, as might be inferred, extremely debilitating for the victim, and these negative consequences inevitably affect close relatives and friends of the victim. Consequences that have been identified include low self-worth, anxiety and mood disorders, problems maintaining pleasant functional relationships, problems in toileting, antisocial conduct and maltreatment of others, suicide, poor social and coping skills, problems in school or employment, substance abuse, and personality disorders. Malnutrition is a hallmark of neglect. Sexual abuse victims may demonstrate inappropriate or promiscuous sexual activity, including prostitution and other illegal sexual undertakings. A particular concern is the recapitulation of abuse because, more often than not, the victim becomes a perpetrator with the passage of time.

Remediation Strategies

Professionals (i.e., teachers, police officers, ministers, and health-care workers) are mandated in most states to report suspected incidents of domestic violence. Familiarity with the risk factors and consequences of domestic violence are of great assistance in the identification of abuse. However, professionals are often not trained, or able, to recognize abuse, particularly because victims are motivated to deny or minimize assaultive behaviors because their perpetrators are usually people whom they often love, fear, or respect. Moreover, once identified, domestic violence is difficult to remediate. Initial strategies focus on assessing the victim's environment to determine if it is (1) safe for the victim to remain in the home with the perpetrator, (2) necessary to remove the victim from the home (to live with other relative, in a foster-care home, or in a shelter), or (3) necessary to mandate that the perpetrator leave the home. The separation of the victim from the perpetrator is certainly the safest strategy for terminating the abuse. However, this approach is often not possible because relatives are frequently unavailable and shelters and

foster-care homes usually operate with inadequate funding, causing them to sometimes be full, dangerous, or overcrowded. Moreover, familial separation is marked with changes that bring about stress (e.g., new school placements, feelings of abandonment and rejection, and financial losses). Therefore, separation is typically sought only when cases of maltreatment are severe (e.g., sexual abuse or repeated instances of physical abuse). Adult victims and perpetrators of domestic violence are usually forced to fund their own therapy. However, state child protective service agencies will provide psychologically based intervention for child victims and relevant family members at no cost, when indicated. Interventions may be focused on reunifying and improving the function of the existing family system or facilitating a comfortable and healthy transition for the victim without the perpetrator. Therapy is ordinarily mandated by the legal system for child victims and their perpetrators when the goal is family reunification.

Of the interventions that have been evaluated, the cognitive-behavioral and stress-reduction strategies appear to be most promising. Cognitive-behavioral interventions most often include the recognition of the early signs of abuse, learning strategies to improve safety and prevent the escalation of abuse, restructuring faulty cognitions and beliefs about the abuse (e.g., victims may blame themselves for abuse; perpetrators might think the role of a woman is to serve her spouse), learning to improve appropriate social/assertion skills, anger-management skills training, contingency contracting (obtaining rewards that are contingent on behavioral improvement, e.g., good grades in school, compliance to the directives of parents, and making statements of affection), learning to eliminate home hazards and improve home beautification, behavioral substance-abuse counseling (e.g., controlling urges to use substances, restructuring the environment so that drug use is less likely, and contingency contracting), vocational skills training, learning positive parenting techniques (e.g., reinforcing the child for desired behaviors, positive practice, and learning about child development and nutrition), marital counseling, multisystemic family therapy, and communication skills training. Stress-management strategies include bringing children to day care or babysitters to allow the caregiver a relaxing break at home, learning problem-solving and coping skills, attending Parents Anonymous groups (an organization established to support the victims and perpetrators of abuse), planning pleasant activities with the family, and encouraging/facilitating participation in community groups (e.g., church and boys' and girls' clubs).

Historical Perspectives

Perceptions of domestic violence have changed markedly throughout history and across cultures. Indeed, it was not until the past century that laws were initiated to protect children and wives from being beaten by their parents and husbands; for instance, as recently as 1968 many states did not mandate professionals to report their knowledge of abusive treatment to abuse registries. Nevertheless, there have been radical improvements in public sentiment and the legal system regarding the eradication of domestic violence. In ancient times, individuals with mental and physical disabilities were scorned or killed to rid the state and parents of the economic burden of caring for them. Although the treatment of children and women improved somewhat in most societies during the Middle Ages, neglect, ridicule, slavery, and harsh physical punishment continued. It is also interesting to note that children in nineteenth-century industrialized nations were forced to work up to 14 h a day in hazardous conditions.

The improvements in the rights of children and women were partially a result of humanitarians who wrote books about the injustice of domestic violence. For instance, during the Renaissance Thomas Phaire wrote a book addressing the problems of children called *The Book of Chyldren*, and in the late 1600s and early 1700s Locke and Rousseau wrote books that mentioned issues pertaining to the development of children. Erin Pizzey's 1974 book *Scream Quietly or the Neighbors Will Hear* was one of the first manuscripts in the literature that specifically addressed remediation procedures for spousal abuse.

National organizations, governmental meetings, and activists also influenced the passing of laws protecting against domestic violence. The Society for the Prevention of Cruelty to Children in the late 1800s was formed to protect rights and welfare of children. In 1909, President Theodore Roosevelt organized the first meeting in the White House on children's welfare, which supported the Children's Bureau in 1912 to protect the welfare and rights of children. Pizzey was responsible for developing the first refuge for female victims of spousal abuse in England in 1971; the first refuge for women and their children in the United States was initiated by the Women's Advocates in 1972.

Although domestic violence continues today at relatively high rates, research and public opinion and protection laws against domestic violence are at their highest level of sophistication. Indeed, all 50 states currently require professionals to report suspected instances of child maltreatment, and most states

have explicitly allocated significant funding to programs for the prevention and treatment of domestic violence. Indeed, law enforcement personnel now routinely receive specialized training for managing situations involving domestic violence, including victim-sensitivity training. Violence treatment programs have also become increasingly more evidence-based while addressing contemporary issues such as violence among same-sex couples, date rape, violence prevention within school settings, and comorbid conditions (e.g., substance abuse).

See Also the Following Articles

Aggression; Caregivers, Stress and; Child Abuse; Male Partner Violence; Marital Conflict; Sexual Assault.

Further Reading

- Bergen, R. (ed.) (1998). *Issues in intimate violence*. Thousand Oaks, CA: Sage.
- Donohue, B., Ammerman, R. T. and Zelis, K. (1998). Child physical abuse and neglect. In: Watson, T. S. & Gresham, F. M. (eds.) *Handbook of child behavioral therapy*, pp. 183–202. New York: Plenum.
- Feindler, E. L., Rathus, J. H. and Silver, L. B. (2003). *Assessment of family violence: a handbook for researchers and practitioners*. Washington, DC: American Psychological Association.
- Lutzker, J. R. (2006). *Preventing violence: research and evidence-based intervention strategies*. Washington, DC: American Psychological Association.

Dopamine, Central

G D Stanwood

Vanderbilt University, Nashville, TN, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G D Stanwood and M J Zigmond, volume 1, pp 739–745, © 2000, Elsevier Inc.

Introduction to Nerve Cells and Synaptic Transmission
Overview of Dopaminergic Pathways in the Central Nervous System
Responses of Dopaminergic Systems to Acute Stressful Stimuli
Responses of Dopaminergic Systems to Chronic Stress
Interactions between Dopamine and other Neurochemical Systems Altered by Stress
Developmental Modulation

Glossary

<i>Agonist</i>	A drug that mimics the actions of a neurotransmitter at its receptor.
<i>Antagonist</i>	A drug that blocks the actions of a neurotransmitter at its receptor.
<i>Catecholamine</i>	A chemical containing a catechol nucleus (a benzene ring with two adjacent hydroxyl groups) and an amine group.
<i>Central nervous system (CNS)</i>	A general term referring to the entire brain and spinal cord.
<i>Dopamine (DA)</i>	A catecholamine used as a neurotransmitter in the central nervous system; also known as dihydroxyphenylethylamine.

Hypothalamic-pituitary-adrenocortical (HPA) axis
Mesencephalon

A principle component of the stress response that regulates the secretion of glucocorticoid hormones from the adrenal gland.

The midbrain; many DA cells have their soma, or cell bodies, located in two regions of the mesencephalon, the substantia nigra (SN) and the ventral tegmental area (VTA).

Neuron/nerve

The fundamental cellular unit of the nervous system.

Neurotransmitter

A chemical released by a neuron at a synapse to provide a signal to another neuron.

Nucleus accumbens (NAC)

A subcortical brain region thought to be an important interface for motivation and motor outputs; it consists of two subregions, the core and the shell.

Prefrontal cortex

A cortical region located towards the front of the brain involved in planning and other cognitive functions; it receives a dopaminergic input from the ventral tegmental area and projects to the striatum and nucleus accumbens.

Striatum (or neostriatum)

A region of the basal ganglia that is important for movement; the striatum receives a dopaminergic projection from the substantia nigra.

Synapse

A specialized contact zone that is the site of chemical communication between neurons.

Tyrosine hydroxylase (TH)

The rate-limiting enzyme in the synthesis of dopamine; it converts the amino acid tyrosine to dihydroxyphenylalanine (DOPA).

Dopamine is a chemical transmitter in the central nervous system, where it acts to modulate the actions of other neural signaling molecules. Physical or psychological stress results in an activation of nearly all neurons that use dopamine as their transmitter, causing an increase in dopamine release. Those dopamine neurons projecting from the mesencephalon to the prefrontal cortex are particularly sensitive to such stress. The mechanisms by which dopamine release is increased in response to stress appear to include (1) an increase in the firing rate of the dopamine neurons, (2) changes in the inputs to the dopamine terminals from adjacent cells, and (3) an increase in the activity of tyrosine hydroxylase, the rate-limiting enzyme in dopamine synthesis. The precise roles that these changes in dopaminergic neurotransmission play in the generation of stress and/or coping responses are not yet clear. However, the study of these mechanisms in animal models may lead to a more complete understanding of the biological basis of a number of psychiatric disorders.

Introduction to Nerve Cells and Synaptic Transmission

Nerve cells, or neurons, are the elementary signaling units within the nervous system. The highly complex behaviors of humans and other animals result from

the specific interconnections of a vast number of these neurons. Neurons have four main regions: (1) the dendrites, which receive most of the information coming into the neuron, (2) the cell body, or soma, which processes that information and also contains the mechanisms of protein synthesis, (3) the axon, which carries an electrical signal called an action potential from the cell body to the nerve terminal, and (4) the terminal, which can respond to signals received from the axon and from adjacent neurons by releasing a chemical transmitter. Neurons communicate with one another through minute gaps between them known as synapses. The chemical transmitters released by the presynaptic terminal of a neuron diffuse across this synaptic space to act on specialized receptors on the dendrites and/or cell body of a neighboring postsynaptic neuron (Figure 1).

Overview of Dopaminergic Pathways in the Central Nervous System

The nervous system uses many different transmitters, each with different functional characteristics. However, they can be broadly divided into two groups. The first contains the fast-acting transmitters that provide primary information. Many of the neurons that use acetylcholine or the amino acids glutamate and γ -aminobutyric acid (GABA) fall into this

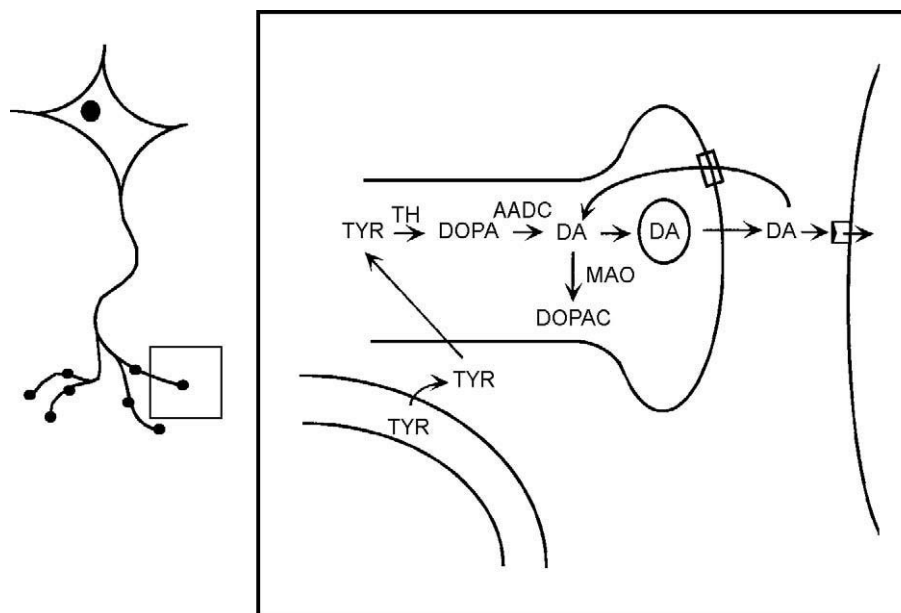


Figure 1 Schematic representation of the life cycle of dopamine. Upon accumulation of tyrosine (Tyr) by the neuron, tyrosine is sequentially metabolized by tyrosine hydroxylase (TH) and aromatic acid decarboxylase (AADC) to dopamine (DA). DA is then accumulated into vesicles by the vesicular transporter, where it is poised for release and protected from degradation. Once DA is released, it can interact with postsynaptic dopaminergic receptors or different types of presynaptic autoreceptors that regulate transmitter release, synthesis, or firing rate. The accumulation of DA by the high-affinity membrane DA transporter (DAT) terminates the extracellular actions of DA. Once accumulated by the neuron, DA can be metabolized to inactive species by key degradative enzymes such as monoamine oxidase (MAO) or taken back up by the vesicular transporter.

category. The second group is composed of the more slowly acting transmitters that have a modulatory influence on the actions of the first group. Dopamine (DA) is an important member of this second, modulatory group of transmitters. DA has been implicated in a variety of functions including motor control, cognition, endocrine regulation, emotion, and cardiovascular regulation. Abnormalities in DA systems can lead to major neurological and psychiatric dysfunction.

DA is a catecholamine, a term that refers to its biochemical structure. Other catecholamines include norepinephrine (NE) and epinephrine, also known as adrenaline. DA is synthesized from tyrosine through the actions of two enzymes. First, tyrosine hydroxylase (TH), the rate-limiting enzyme in the process, converts L-tyrosine to dihydroxy-L-phenylalanine (L-DOPA). Then, aromatic acid decarboxylase (also known as DOPA decarboxylase) immediately converts L-DOPA to DA (Figure 1). In some cells, DA is further converted to the neurotransmitter NE or even further to epinephrine, but in many neurons DA itself serves as an active neurotransmitter.

The rate at which DA is synthesized can be controlled in many ways. DA can inhibit tyrosine hydroxylase (TH), a process termed end-product inhibition. Additional influences can be exerted by changes in the number or structure of the TH molecules or by changes in the availability of necessary cofactors for tyrosine hydroxylation. Because of these regulatory processes, neurons are usually able to match the rate of DA synthesis to the rate of DA utilization, thereby avoiding either the build up or the depletion of the transmitter.

DA release occurs in response to an influx of calcium into the nerve terminal, which is triggered by the arrival of an action potential. Like DA synthesis, DA release is regulated by several processes. For example, DA can act back on the terminal from which it was released to inhibit subsequent release. Such influences represent negative feedback loops and act to keep the rate of DA utilization within relatively narrow limits.

DA induces a wide range of cellular and biochemical effects in neurons by way of its actions on its specific postsynaptic receptor proteins. These effects include relatively rapid (seconds) modulation of biochemical events in the target cell resulting in changes in the responsiveness of the cell to other neuronal inputs, as well as more gradual (minutes to hours) alterations in gene expression. The sensitivity of a target cell to DA is regulated: a period during which DA release is reduced causes an increase in the number of postsynaptic receptors and thus a supersensitivity of the neuron to DA, whereas an excess of DA leads to a decrease in receptors and in the sensitivity

of the target. Intracellular components of signal transduction responses are also actively regulated in the brain, providing many potential mechanisms by which signaling can be altered.

The neurotransmitter actions of released DA are terminated by diffusion from the receptor site. It can then be transported into presynaptic terminals by high-affinity uptake sites. Once taken up, DA can be metabolized in the nerve terminal by the monoamine oxidase (MAO) to dihydroxyphenylacetic acid (DOPAC) or further sequestered into storage vesicles for later reuse. Released DA can also be metabolized extraneuronally by catechol-O-methyltransferase (COMT) to 3-methoxytyramine (3-MT). The product of MAO plus COMT is homovanillic acid (HVA).

There are several major dopaminergic pathways (Figure 2). The two main mesencephalic regions containing DA cell bodies are the substantia nigra (SN) and ventral tegmental area (VTA). Axons of the DA cells in the SN form the nigrostriatal projection and provide the dopaminergic innervation of the caudate and putamen, or striatum. This region is part of the basal ganglia, a group of nuclei involved in motor and cognitive functions. Degeneration of the nigrostriatal DA neurons is the principal cause of Parkinson's disease.

The mesolimbic and mesocortical DA systems arise from the VTA, which is located just medially to the SN. The mesolimbic DA system innervates subcortical regions such as the nucleus accumbens (NAC), olfactory tubercle, and amygdala. The mesocortical system provides dopaminergic afferents to prefrontal, anterior cingulate, and entorhinal cortex. Disruption of dopaminergic neurotransmission in these areas has been associated with neuropsychiatric illnesses such as schizophrenia. The tuberoinfundibular DA

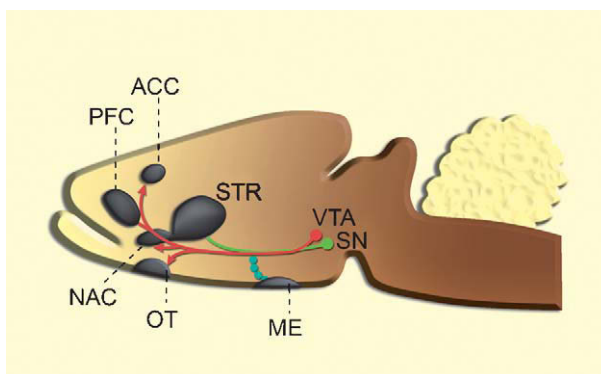


Figure 2 Schematic illustrating some of the major dopaminergic neurons and their projection areas in the rodent. The mesocortical limbic pathway is illustrated in red, nigrostriatal tract in green, and tuberoinfundibular in blue. Abbreviations: ACC, anterior cingulate cortex; ME, median eminence; NAC, nucleus accumbens; OT, olfactory tubercle; PFC, prefrontal cortex; SN, substantia nigra; STR, striatum; VTA, ventral tegmental area.

system is another important dopaminergic pathway; it connects the hypothalamus and pituitary gland. These DA neurons are important in the regulation of the hormone prolactin from the anterior pituitary.

Responses of Dopaminergic Systems to Acute Stressful Stimuli

Most antidepressant drugs have direct pharmacological effects on catecholamine neurotransmission. Tricyclic antidepressants, for example, block the uptake of DA and other catecholamines, thus increasing the biochemical effects of these neurotransmitters. Based on the rationale that people suffering from anxiety and depressive disorders may have dysfunctions of brain catecholamine systems, changes in dopaminergic activity in response to stressful stimuli have been studied in a number of animal models. Some interesting effects have been observed, as is described below. However, it must also be emphasized that nearly all antidepressant drugs have very prominent effects on NE and serotonin (5-hydroxytryptamine, 5-HT) systems, and it is clear that multiple neurotransmitter systems are involved in the pathophysiology of stress disorders and depression.

Rodent models of environmental stress include tail pinch and/or shock, immobilization, forced swim, exposure to cold, foot shock, and social defeat. These stress-inducing procedures produce heterogeneous effects on the various dopaminergic projection systems. For example, the activity of tuberoinfundibular DA neurons is decreased by stressful stimuli, which results in an increase in plasma prolactin levels. This increase in prolactin likely plays a role in stress-induced activation of the immune system.

In contrast to the tuberoinfundibular DA neurons, the dopaminergic neurons that ascend from the midbrain are activated by stress. Mesocortical DA neurons projecting from the VTA to the prefrontal cortex appear to be particularly sensitive to acute stress so that relatively mild stressors lead to pronounced activation of DA neurotransmission in the prefrontal cortex. Stress-induced activation of DA neurotransmission also occurs in the mesolimbic and nigrostriatal pathways, but to lesser extents than in the prefrontal cortex. Further heterogeneity is observed within the mesolimbic projections to the NAC. Stressors appear to activate dopaminergic activity to a greater extent in the NAC_{shell} (the limbic NAC) subregion than in the NAC_{core} (the motor NAC) and in the striatum.

Consistent with these differences in activation of DA terminal regions, it has been observed that VTA neurons increase their rate of firing during stress, but SN neurons respond only briefly if at all.

Moreover, acute exposure to stress has been shown to increase TH activity in DA terminal regions, and the rank order of this effect is identical to that of the neurochemical measurements (prefrontal cortex > nucleus accumbens > striatum). Behavioral stress also induces neuroadaptations in dopaminergic neurons that are very similar to those induced by drugs of abuse.

Collectively, these data demonstrate that increases in DA neurotransmission occur during exposure to acute stress. But what functional significance do these changes in DA activity represent? There are several lines of evidence suggesting that DA systems play critical roles in stress responses. First, anti-anxiety drugs such as benzodiazepines can reduce stress-induced increases in DA release, suggesting that the therapeutic actions of these compounds may be due at least in part to reversing increased dopaminergic activation. Second, lesions of midbrain DA neurons lead to a partial suppression of endocrine stress responses (see “Dopamine and the HPA Axis” for a more complete discussion). Third, enhancement of DA signaling in the amygdala attenuates stress-induced ulcer formation, and stress-induced increases in dopaminergic activity in the prefrontal cortex have been shown to be associated with behavioral coping responses. These and other data have been interpreted as indicating that stress-induced increases in dopaminergic activity are important components of coping responses to environmental stressors rather than being a reflection of anxiety.

Yet, in trying to understand the functional significance of the activation of brain DA systems, it must be kept in mind that DA neurotransmission can be increased by a variety of environmental and behavioral challenges that are nonaversive. Such challenges include feeding, sexual activity, and exposure to drugs of abuse. Indeed, the dopaminergic pathway from the VTA to the NAC is critical for the reinforcing effects of many addictive drugs, including cocaine, morphine, and nicotine, as well as direct brain stimulation. These two seemingly contradictory observations – DA release is increased by stress and by positively reinforcing stimuli – can be reconciled if one assumes that any stimulus that causes behavioral arousal leads to an increase in DA neurotransmission. DA activation is probably essential for attention and survival. However, at some point the amount of activation becomes detrimental. For example, a rat trained to lever press in order to receive injections of cocaine will respond at a very steady rate. Such an animal will exhibit elevated rates of DA release in the NAC and other brain regions. If the dose of cocaine per injection is suddenly increased, however, the animal will reduce its rate of lever pressing such

that only enough cocaine is present to maintain the previous level of DA release. In other words, if given the choice, the animal will not allow the extent of DA activation to exceed some optimal level. It is likely that this optimum varies both from individual to individual and from physiological state to state. A similar inverted-U relationship between stress-induced DA activation and working memory has been well documented within the frontal cortex (Figure 3).

Responses of Dopaminergic Systems to Chronic Stress

Animals are often confronted with repeated or prolonged exposure to stressors, and this can lead to

additional neurobiological and behavioral changes. Under certain conditions, habituation can result – the animals cease responding to the stimulus. But under other conditions sensitization can occur. In this case, the animal's biological and/or behavioral responses increase to the next presentation of the same stimulus or to a different stimulus. Whether or not a given paradigm leads to habituation or sensitization is likely determined by the exact nature and duration of the stressors used, the species, and/or the brain region examined. These processes likely contribute to stress-related disorders such as anxiety, depression, and fibromyalgia.

Chronic stress can also lead to cross-sensitization to drugs of abuse such as cocaine, amphetamine, and opiates at both the behavioral and cellular level.

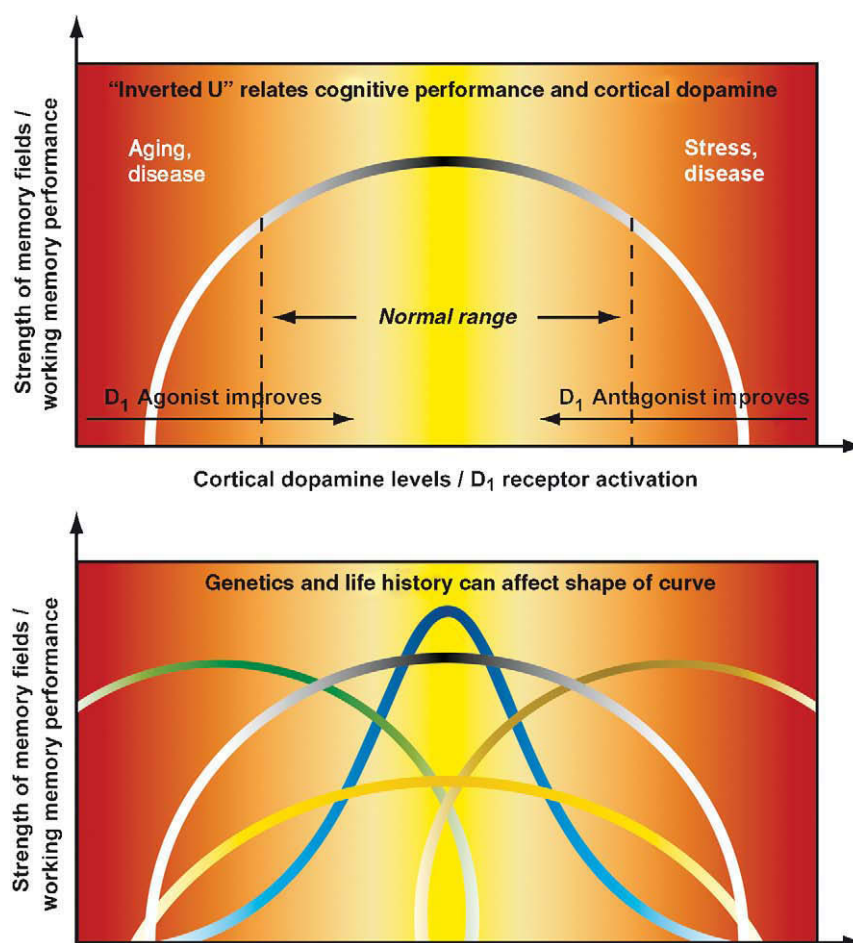


Figure 3 The inverted-U model for dopaminergic regulation of prefrontal function, which was originally proposed by Patricia Goldman-Rakic and colleagues. Top: A generalized model for DA function is depicted that assumes that, under normal conditions, there is a range of DA concentration over which signal processing in prefrontal neurons operates efficiently and working memory performance is sufficient to the task. At the center of this range is an optimum for cortical DA level, and D1 receptor stimulation in particular, at which performance is maximal. The exact shape and slope of this curve is likely task specific, such that complex tasks may be more sensitive to a narrower range of DA concentrations. Bottom: The curve may be altered in shape or slope (blue and yellow gradients) or be shifted to the left or right (green and brown gradients, respectively) depending on genotype, life history, and/or conditions. Individuals who suffer from deficiency of prefrontal DA transmission may benefit from DA receptor stimulation by agonists, whereas those who suffer from excessive endogenous transmission may benefit from treatment with DA receptor antagonists.

It is possible that sensitization plays a role in the development of posttraumatic stress disorder, panic attacks, and psychosis. Furthermore, because DA can oxidize to form potentially toxic metabolites, it has been hypothesized that the prolonged elevations in DA observed with repeated stress may lead to impairments of cellular functioning, although this possibility has not yet been rigorously examined.

Many cellular and molecular changes have been described in the development of neurochemical and behavioral sensitization following chronic stress. For example, chronic intermittent or sustained stress has been shown to induce changes in the number of DA receptors in several brain regions, including the prefrontal cortex and NAC. Chronic stress also leads to changes in TH expression in the VTA and in the activity of several components of DA receptor signaling in the NAC, dorsal striatum, and amygdala.

Interactions between Dopamine and other Neurochemical Systems Altered by Stress

Dopamine and the HPA Axis

Many neurochemical and hormonal systems are activated by stress. One of the best-characterized physiological responses to stress is the activation of the hypothalamic-pituitary-adrenocortical (HPA) axis. The HPA system controls secretion of glucocorticoids by the adrenal gland. Interestingly, several dopaminergic regions in the brain express glucocorticoid receptors, and the density of the receptors in these regions parallels the ability of stress to increase the activity of DA neurons: receptor density is higher in VTA than SN. There is reason to believe that the correlation between glucocorticoid receptors and stress-induced changes in DA release has functional significance. Administration of glucocorticoids can increase extracellular DA. Moreover, the suppression of stress-induced corticosterone secretion causes a reduction in the dopaminergic response to stress.

Glucocorticoid regulation of midbrain DA neurons also appears to be a biological substrate of the effects of stress on the propensity to develop drug abuse. Conversely, midbrain DA neurons also seem capable of regulating the HPA axis since loss of DA decreases the basal and stress-induced secretion of corticosterone.

Dopamine and Excitatory Amino Acids

Stress has been shown to increase the release of glutamate and other excitatory amino acids in many brain regions, including the prefrontal cortex, striatum, NAC, VTA, and hippocampus. The presence of glucocorticoids appears to be necessary for this response,

as adrenalectomy abolishes the stress-induced increase of extracellular glutamate, and corticosterone replacement restores this response. Glutamatergic neurons in the cortex project to the striatum and NAC as well as SN and VTA. Furthermore, glutamate agonists appear to increase the synthesis and release of DA in each of these regions, and it has been hypothesized that stress-induced activation of glutamate may contribute to stress-induced increases in DA. The local administration of glutamate antagonists into the region of the midbrain containing DA cell bodies can block stress-induced increases in extracellular DA in the striatum. Additionally, stress-induced increases in DA synthesis can be blocked by local administration of a glutamate antagonist directly into the striatum itself. Glutamate antagonists can also block stress-induced increases in extracellular DA in the prefrontal cortex. Intracellularly, glutamatergic and dopaminergic signals can converge on a single protein, the dopamine- and cAMP-regulated phosphoprotein (DARPP-32), a protein whose activation is a crucial component of neural signaling responses.

Dopamine and Norepinephrine

NE is formed from DA and shares many of its basic characteristics. Acute exposure to stress causes an activation of NE neurons in the locus coeruleus and an increase in both the synthesis and release of NE. The exact relationship between these events and stress-induced changes in DA is unclear. NE terminals in the prefrontal cortex can modulate the dopaminergic responses to stress. The NAC and striatum, on the other hand, exhibit increases in dopaminergic activity with stress but are NE-poor regions; presumably NE-DA interactions have little significance in these regions. However, novel roles for the NE-rich bed nucleus of the stria terminalis in mediating polysynaptic responses to stress and drugs of abuse have recently been implicated and are a current area of intense scrutiny.

Dopamine and Serotonin

The involvement of 5-HT in animal models of stress and depression has received great attention due to the antidepressant effects of selective 5-HT reuptake inhibitors (SSRIs) such as fluoxetine (Prozac). Acute exposure to stress leads to dramatic increases in 5-HT release in the striatum and other brain regions. Local administration of 5-HT agonists into the striatum can increase DA release, suggesting that it is possible that stress-induced increases in DA release are at least partially mediated by increases in 5-HT. Anatomical and electrophysiological evidence also suggests prominent interactions between 5-HT and DA in both cell body

and terminal regions. These effects appear to be mediated via multiple families of 5-HT receptors, including the 5-HT_{1A}, 5-HT_{2C}, and 5-HT₄ subtypes.

Dopamine and Dopamine

In developing an understanding of the role of DA neurotransmission in stress, one must take into consideration the functional differences among distinct brain DA systems. In particular, it must be remembered that mesocortical DA neurons inhibit glutamatergic cortical pyramidal neurons, which in turn can stimulate dopaminergic activity in nigrostriatal and mesoaccumbens neurons. Therefore, the less pronounced effects of acute stress on subcortical DA systems may be due in part to the suppression of subcortical DA systems by activated mesocortical DA neurons. Consistent with this hypothesis, depletion of DA in the prefrontal cortex enhances the stress-induced increase in dopaminergic transmission in the NAC_{shell}. Furthermore, injection of a DA receptor antagonist into the prefrontal cortex has been shown to block the DA stress response in the NAC. Last, under certain conditions, chronic stress produces neurochemical tolerance in the prefrontal cortex but sensitization in subcortical areas.

Interactions between cortical and subcortical DA systems have also been proposed in the pathophysiology of schizophrenia. A dysfunction of dopaminergic signaling in the prefrontal cortex can produce secondary disruptions in subcortical DA systems and vice versa. In fact, decreases in the expression of TH and DA receptors in the prefrontal cortex have recently been described. The effects of cortical DA depletion on subcortical DA stress responses may thus be of particular clinical importance given that stress is known to exacerbate psychotic symptoms in schizophrenic patients.

Developmental Modulation

Finally, recent studies have demonstrated major effects of altering neurodevelopmental trajectory on subsequent functioning of brain DA systems. For example, prenatal exposure to stress or drugs of abuse can permanently alter the anatomical organization and cellular responses of both midbrain DA neurons and their forebrain targets. The quality of parental care has also been observed to alter these relationships in both animal models and humans. For example, repeated periods of maternal separation during early life decreases DA transporter expression and increases behavioral and brain DA responses to stress. DA systems are also uniquely sensitive to environmental, genetic, and/or pharmacological modifications

during adolescence. Responses of brain DA systems to stress, therefore, are critically impacted by genetic and environmental risk factors. Elucidation of these risk factors, their specific effects on brain maturation, and the development of therapeutic methods to normalize developmental trajectory in their presence are all important areas of active research.

Acknowledgments

G. D. S. receives support from DA17957, DA11165, and NICHD core grant P30HD15052.

See Also the Following Articles

Antidepressant Actions on Glucocorticoid Receptors; Catecholamines; Serotonin in Stress; Renal and Adrenocortical Actions of Dopamine; Hypothalamic-Pituitary-Adrenal; Chronic Social Stress: GR Sensitivity in Leukocytes; Excitatory Amino Acids; Serotonin.

Further Reading

- Abercrombie, E. D., Keefe, K. A., DiFrischia, D. and Zigmond, M. (1989). Differential effect of stress on *in vivo* dopamine release in striatum, nucleus accumbens and medial frontal cortex. *Journal of Neurochemistry* **52**, 1655–1658.
- Deutch, A. Y. and Roth, R. H. (1999). Neurotransmitters. In: Zigmond, M. J., Bloom, F. E., Landis, S. C., Roberts, J. L. & Squire, L. R. (eds.) *Fundamental neuroscience*, pp. 193–234. San Diego, CA: Academic Press.
- Finlay, J. M. and Zigmond, M. J. (1997). The effects of stress on central dopaminergic neurons: possible clinical implications. *Neurochemical Research* **22**, 1387–1394.
- Goldman-Rakic, P. S., Muly, E. C. 3rd and Williams, G. V. (2000). D(1) receptors in prefrontal cells and circuits. *Brain Research Reviews* **31**, 295–301.
- Horger, B. A. and Roth, R. H. (1996). The role of mesoprefrontal dopamine neurons in stress. *Critical Reviews in Neurobiology* **10**, 395–418.
- Lucas, L. R., Celen, Z., Tamashiro, K. L., et al. (2004). Repeated exposure to social stress has long-term effects on indirect markers of dopaminergic activity in brain regions associated with motivated behavior. *Neuroscience* **124**, 449–457.
- Moghaddam, B. and Jackson, M. (2004). Effect of stress on prefrontal cortex function. *Neurotoxicological Research* **6**, 73–78.
- Pezze, M. A. and Feldon, J. (2004). Mesolimbic dopaminergic pathways in fear conditioning. *Progress in Neurobiology* **74**, 301–320.
- Piazza, P. V. and LeMoal, M. (1998). The role of stress in drug self-administration. *Trends in Pharmacological Sciences* **19**, 67–74.
- Pruessner, J. C., Champagne, F., Meaney, M. J. and Dagher, A. (2004). Dopamine release in response to a

psychological stress in humans and its relationship to early life maternal care: a positron emission tomography study using [^{11}C] raclopride. *Journal of Neuroscience* **24**, 2825–2831.

Stanwood, G. D. and Levitt, P. (2004). Drug exposure early in life: functional repercussions of changing neuropharmacology during sensitive periods of brain development. *Current Opinion in Pharmacology* **4**, 65–71.

Svenningsson, P., Nishi, A., Fisone, G., Girault, J. A., Nairn, A. C. and Greengard, P. (2004). DARPP-32: an integrator

of neurotransmission. *Annual Review of Pharmacology and Toxicology* **44**, 269–296.

Teicher, M. H., Andersen, S. L., Polcari, A., Anderson, C. M., Navalta, C. P. and Kim, D. M. (2003). The neurobiological consequences of early stress and childhood maltreatment. *Neuroscience and Biobehavioral Reviews* **27**, 33–44.

Wood, P. B. (2004). Stress and dopamine: implications for the pathophysiology of chronic widespread pain. *Medical Hypotheses* **62**, 420–424.

Drosophila Genes and Anoxia

G G Haddad

Montefiore Medical Center, Bronx, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition article by G G Haddad, volume 1, pp 746–752, © 2000, Elsevier Inc.

Introduction

Genetic Approach: Mutagenesis Screen

Reverse Genetic Approach: Differential Display

Known Genes: Genetic Networks

Conclusion

Glossary

<i>Anoxia-inducible genes</i>	Genes that are activated by low oxygen.
<i>Anoxia tolerant</i>	Resistant to injury from anoxia.
<i>Differential display</i>	A technique used to screen for differentially expressed mRNA.
<i>Hypoxia-inducible factor (HIF1)</i>	A factor found to enhance the activation of transcription of the Epo gene as well as other genes by binding to DNA.
<i>Mutagenesis screen</i>	A screen for a phenotype after mutagenesis of genome.

This article addresses the environmental stresses of oxygen and glucose deprivation and their consequences, especially in the central nervous system and in sensitive organs in early life. Such stresses can be rather mild or moderate and can activate mechanisms that allow the organism or a specific tissue to adapt and survive the stress. More severe stresses may be injurious to the organism and may lead to cell death, tissue death, and ultimately death of the whole organism.

Introduction

It is clear from a large body of literature that organisms vary greatly in their response to lack of oxygen. Some, such as mammals, are generally susceptible to this stress, although mammalian newborns are more resistant to such a stress than adults, which have more differentiated organs and tissues. In addition, tissues within an organism, cells within an organ, and even cell types within a region of an organ can be differentially sensitive to the lack of oxygen and substrates. Consider and compare, for example, the tolerance to a lack of oxygen in muscle versus the brain, the hippocampal neurons versus the glial cells, and the CA1 layer versus the dentate gyrus.

Because the turtle is an organism that is extremely tolerant to oxygen deprivation, it has been used extensively and is an excellent animal model for the study of tolerance to oxygen deprivation. However, several problems inherent to this model have prevented major progress in understanding the basis of the turtle tolerance; for example, a genetic approach is precluded. Since the late 1980s, we have searched for a model in which multiple approaches can be used, including molecular and genetic techniques. In this search, we have found that *Drosophila melanogaster* is extremely tolerant to oxygen deprivation. It can resist, without apparent injury, several hours of complete oxygen deprivation or anoxia. Clearly the advantages of using *Drosophila* are multiple: (1) genetic approaches can be easily adapted and used, (2) there are hundreds of markers and mutations on the *Drosophila* chromosomes that facilitate mapping of mutations, (3) there are only three chromosomal pairs (the fourth is extremely small) in the *Drosophila* genome, (4) physiological, molecular biological, and genetic approaches are feasible in the fruit fly, and

(5) whole-fly studies including transgenic fly work is rather easily performed. We have therefore developed this model further and performed a considerable number of investigations using the fruit fly. We have used three different approaches to address the genetic basis of the response of the fruit fly to low or absent environmental oxygen. In this article, we summarize our findings to date using these different approaches.

Genetic Approach: Mutagenesis Screen

In order to perform a mutagenesis screen, we characterized the wild-type response to hypoxia using a number of behavioral and physiological assays. We first subjected flies to extremely low oxygen concentrations (0.01%) and studied their physiological and behavioral responses during anoxia and after the start of reoxygenation. Flies were first exposed to anoxia (complete lack of oxygen) for periods of 5–240 min. After 1–2 min in anoxia, *Drosophila* lost coordination, fell down, and became motionless. However, they tolerated a complete nitrogen atmosphere for up to 4 h, following which they recovered. In addition, a nonlinear relation existed between the time spent in anoxia and the time to recovery. Extracellular recordings from flight muscles in response to giant fiber stimulation revealed complete recovery of muscle-evoked response, a response that was totally absent during anoxia (Figure 1). Mean oxygen consumption per gram of tissue was substantially reduced in low oxygen concentrations (20% of control). We concluded from these studies that (1) *D. melanogaster* is very resistant to anoxia and can be useful in the study of mechanisms of anoxia tolerance and (2) the profound decline in metabolic rate during periods of low environmental oxygen levels contributes to the survival of *Drosophila*.

Genetic Screen for Anoxia-Sensitive Mutants

We focused our screen for mutations on the X chromosome because the mutant phenotype can be observed in the immediate next generation without the need for single-pair matings. Mutagenized (x ray, 4000 rad) C-S males, which were crossed to attached-X females [*c(1)yw*], transmitted their X chromosome to the male offspring. More than 20,000 flies, carrying mutagenized chromosomes, were screened. A threshold, which is close to the ninety-sixth percentile of the wild-type distribution, was used to identify and isolate mutants. Eight mutations were identified and found to alter the distribution of recovery times profoundly (Figure 2). The marked delay in recovery after anoxia displayed by these mutant flies suggested that they were much more sensitive to a lack of oxygen. We have therefore termed these mutants *hypnos* to describe (1) their sensitivity to oxygen deprivation (*hypoxia*, *anoxia*, sensitive) and (2) the phenotype of delayed recovery and sleepiness.

Genetic Mapping

The mapping of the induced mutations was performed with X-chromosomal markers and complementation tests. Several markers were used, including *y*, *cv*, *v*, *f*, *car*, and *su(f)*. Complementation testing was done on several X-linked recessive mutations obtained. A number of these mutations were mapped, and they seem to be spread in a number of locations on the X chromosome. For example, one of them is on the right side of *f* (*forked*), another is between *y* and *cv*, and yet another is between *cv* and *f*. One of the mutations has been refined in terms of the mapping, and we have localized it to a rather narrow region between *y* and *cv* using deficiencies.

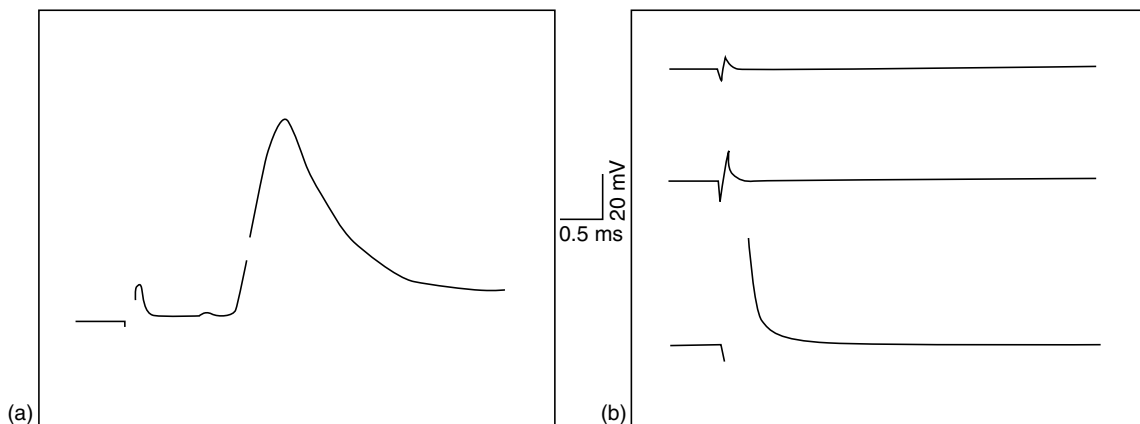


Figure 1 Effect of anoxia on evoked dorsal longitudinal muscle (DLM) response. a, DLM response; b, lack of DLM response during anoxia, even after a 10-fold increase in duration and intensity of stimulus.

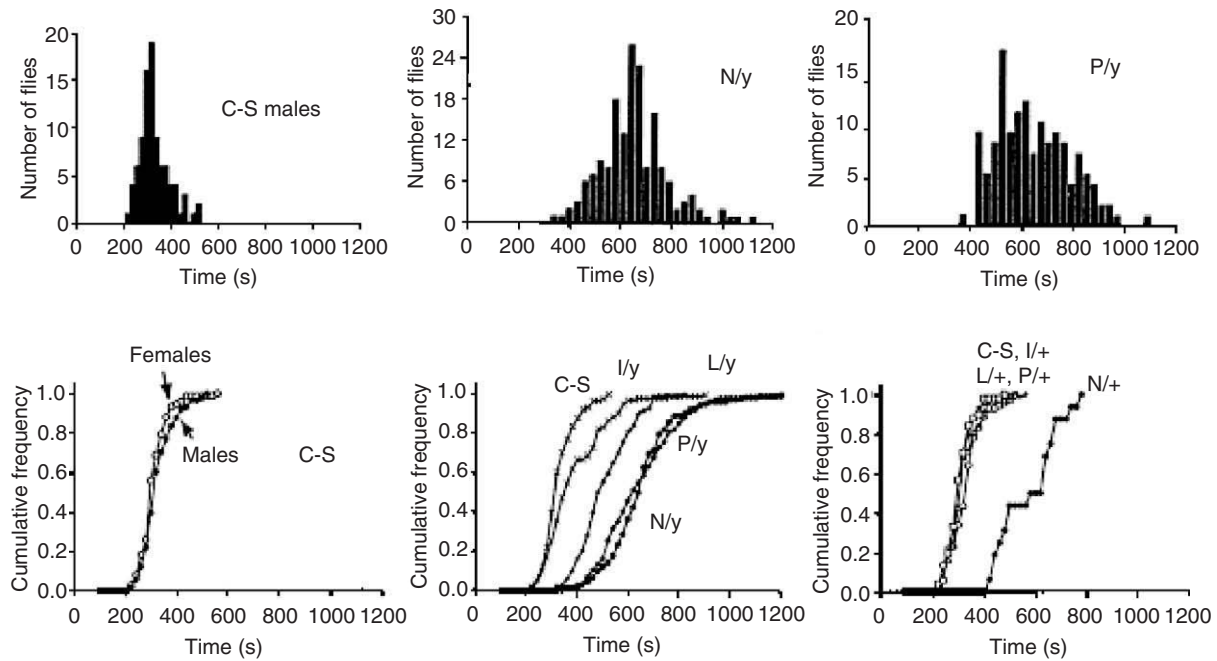


Figure 2 Distribution of cumulative frequency of recovery time in wild-type (C-S) and mutant flies.

Hypnos Mutations Affect Central Nervous System Recovery from Anoxia

The behavioral testing, which showed delayed recovery from anoxia, led us to believe that the *hypnos* mutations affected the central nervous system. To further our understanding of the mutations we obtained, we directly examined the effect of these mutations on central nervous system function. Identified neurons that can be studied electrophysiologically in *Drosophila* are those of the Giant Fiber system. The stimulation of eye neurons, using low voltages, evoked spikes in the dorsal longitudinal muscle (DLM) with a long latency (4 ms). Experiments were carried out in three of the eight mutations (*hypnos*^{2^{PL}}, *Hypnos*^{1^N}) and on wild-type flies. In one mutant (*hypnos*^{2^P}), which had a severe phenotype, long-latency-evoked potentials could not be obtained, irrespective of the voltage and duration used. Flies having an allele of this locus (*hypnos*^{2^L}) or with the other severe mutation (*Hypnos*^{1^N}) had a baseline long-latency recording that was similar to that of C-S flies. However, during recovery, whereas the DLM of wild-type flies started to respond after 2 min into recovery, flies with these two mutations had a much longer time to the first evoked potential, with some mutant flies requiring up to 25–30 min for the first evoked response to occur.

The current genetic screen is not saturated and is restricted to the X chromosome. However, we had to screen more than 20,000 mutagenized flies to obtain

hits in the same locus (two loci hit twice). This suggests that a limited number of genes can be mutated in *Drosophila* to produce similar phenotypes. Because these mutations profoundly disrupt the recovery from anoxia, we believe that this approach can be used effectively to dissect the genetic basis of responses to low oxygen.

Reverse Genetic Approach: Differential Display

The idea behind the reverse genetic approach is to identify those genes that are differentially expressed during a condition or a stress. In our case, we used very low levels of oxygen and assessed whether there were up- or downregulated mRNA during the period of hypoxia compared to naive or nonexposed flies. RNA differential display was performed as described by Liang and Pardee, with minor modifications. The DNase-treated total RNA was used for reverse transcription. The cDNAs generated from polymerase chain reaction (PCR) were blotted; the bands of interest were located, cut out, and reamplified and cloned into the PCRII vector. To determine the relative size of selected candidate transcripts, northern analysis was performed. To obtain full-length cDNA, we screened the fly head λ gt11 library using the cDNA fragments that were isolated from the differential display. In addition, 5' and 3' rapid amplification of cDNA ends (RACE) PCR were also performed.

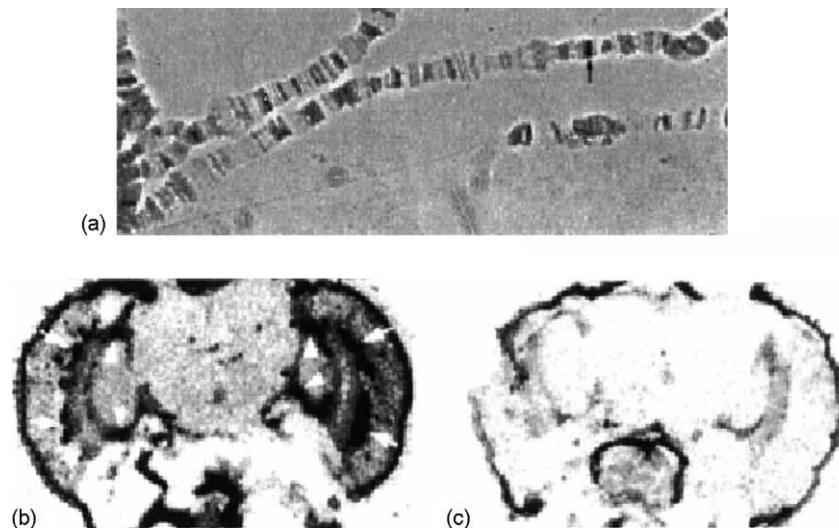


Figure 3 *In situ* hybridization of gene *fau*. a, Polytene chromosomes; b, In the central nervous system of *Drosophila* (antisense); c, In the central nervous system of *Drosophila* (sense).

In addition to the cloning of this particular gene, we have localized it on the X chromosome using *in situ* hybridization. This is usually helpful for developing mutants with P elements and therefore examines the role of that particular gene in the context of the whole organism or cell (Figure 3a). Furthermore, we have also localized the mRNA of *fau* (*fly*, *anoxia*, *upregulated*) in the central nervous system of the fly as seen in Figure 3b–c. This figure shows that the *fau* mRNA is localized in certain areas such as the cortex.

Gene Expression Following Anoxia

The data from our differential display and northern analysis clearly show that certain genes were up-regulated while others were downregulated during anoxia. From the PCR reactions, we found that the expression level of several transcripts was visibly affected by anoxia. We have selected one transcript that was markedly upregulated to focus on and study. We termed this transcript *fau*. The *fau* cDNA and its deduced protein sequence have several interesting and potentially important characteristics. Seven ATTT motifs, for example, were found in its 3' UTR. The AUUUA motifs reportedly play a role in the stability and translocation of transient mRNAs. Therefore, the ATTT motifs found in *fau* cDNA may play a similar role in hypoxia-induced mRNAs and could regulate their expression. In addition, there are two (TA)_{9–10} stretches in the 3' UTR of this cDNA. Although their function is not clear, taken together, these unique sequences in the 3' UTR with high GC content in the *fau* cDNA open reading frame may define a functional anatomy of transient or stress-induced mRNAs. The deduced protein sequence of *fau* cDNA also has a high

number of phosphorylation sites, which makes it an appropriate substrate for phosphorylation. It is possible then that the *fau* protein can participate in metabolic pathways. Because our computerized search did not reveal a significant homology with published sequences, the deduced *fau* protein is most likely a novel phosphorylated one.

Role of *fau* Protein

In spite of the fact that we have no direct evidence that this gene is involved in anoxic survival in *D. melanogaster*, we believe that *fau* plays an important role during anoxic stress based on the following. (1) The expression of this gene was upregulated when the overall reduction of protein synthesis in general occurs during the lack of oxygen. (2) The overexpression of *fau* in transgenic flies prolonged their recovery from anoxia. (3) The putative protein encoded by this gene is probably highly regulated by phosphorylation, suggesting that it is active during O₂ deprivation. (4) The richness of the AUUU motif in 3' UTR of this mRNA could make it a transient one, with a relevant function during O₂ deprivation.

There is a debate in the literature as to whether cells, particularly nerve cells, die from anoxia via mechanisms that induce active processes or via others that shut off pathways and protein synthesis. It is clear from a number of studies that survival and adaptation during anoxia or cell death resulting from anoxia involves the synthesis of proteins, although the overall net balance is a marked reduction in protein synthesis. Numerous studies, for example, have shown that there are genes that are selectively turned on or upregulated by oxygen deprivation. Our previous data have demonstrated that

anoxia remarkably upregulated the expression of a heat shock protein (HSP70) gene, whereas it suppressed the expression of ubiquitin 4 in the fruit fly. A hypoxia-inducible factor 1 (HIF-1) was found to enhance the activation of transcription of the Epo gene as well as other genes by binding to DNA. Hence, we believe that during anoxic stress-tolerant cells such as those in *Drosophila* depress their oxygen consumption and requirements by lowering the metabolic rate and reducing protein synthesis. However, a protein hierarchy within these cells exists for survival; that is, there are certain proteins that are either newly induced or protected from proteolysis, whereas the majority of proteins have a decreased or negligible synthetic rate and are downregulated. It is possible then that proteins such as HIF are overly expressed in order to render cells better adapted to lack of oxygen. However, why would an increase Fau protein occur during hypoxia if an increase in this protein increases by itself the sensitivity to anoxia? At present we cannot be certain because we do not have a comprehensive picture of the various genes and gene products that are important during such stress. For example, we do not know the relation between Fau and other proteins.

Known Genes: Genetic Networks

This approach relies on genes known to be important in the mammal's response to oxygen deprivation. This approach then aims at cloning and studying fruit flies in order to better delineate the genetic pathways and molecular mechanisms underlying the resistance of fruit flies to lack of oxygen. It is assumed that this challenge is better met in organisms such as *Drosophila* than in mammals because of the flexibility and feasibility of molecular approaches in this organism.

Expression of Known Genes

We analyzed mRNA expression of heat shock proteins (HSP26 and HSP70), ubiquitins (UB, UB3, and UB4), cytochrome oxidase I (COX1), and superoxide dismutase (SOD) using slot-blot analysis. The expression of HSP genes, especially HSP70, was remarkably upregulated (up to 1000-fold), whereas those of UB4 and COX1 were downregulated (10–60%) in response to the anoxic stress. The expression of UB3 gene was upregulated by 1.5 times, and the expression of the SOD gene was not significantly affected. In response to heat shock stress, the expression of HSP genes increased by up to several thousandfold, the expression of UB genes increased modestly (23–91%), and the expression of SOD and COX1 genes was reduced by approximately 25%. Furthermore, the expression patterns of HSP genes under anoxia and heat shock

were clearly different. The expression of HSP genes peaked by 15 min into anoxia and then declined but stayed above the baseline. In contrast, their expression increased as a function of time during heat exposure.

Known Pathways

One of the transcription factors that has been shown to be crucial for sensing the lack of oxygen is HIF1. This transcription factor has two subunits and, when they dimerize, they induce the expression of a number of genes, including *Nitric oxide synthase* and *Vascular endothelial growth factor* (VEGF). We and other investigators have cloned the HIF1 β -subunit from *Drosophila*. The α -subunit has been cloned and both the α - and the β -subunits have relatively high homology to those of mammals. At present, we are examining the role of these genes in *Drosophila*. In addition, the approach here aims at examining the pathways that these genes are involved in.

Conclusion

Stress as defined in this article involves the lack of oxygen. Generally, this type of stress has been equated with a pathophysiological signal, that is, with a condition that can lead to cell injury or cell death. However, it is apparent that, although this can certainly be true, especially in moderate or severe stress, it may not be so in milder stresses. In addition, the timing of the stress is exceedingly important. Even though it can be severe, hypoxic conditions in development or during organ formation can be tantamount to signals that are important in the induction of certain genes such as VEGF, which is crucial for organ development and embryogenesis. A failure of response to hypoxia in knockouts such as the HIF knockout can lead to the failure of embryogenesis and demise of the embryo. Indeed, fetal demise is observed in knockouts of both subunits of HIF and in that of VEGF.

See Also the Following Article

Drosophila Studies.

Further Reading

- Akashi, M., Shaw, G., Hachiya, M., et al. (1994). Number and location of AUUUA motifs: role in regulating transiently expressed RNAs. *Blood* 83, 3182–3187.
- Clegg, J. S., Jackson, S. A., Liand, P., et al. (1995). Nuclear-cytoplasmic translocations of protein p26 during aerobic-anoxic transition in embryos of *Artemia franciscana*. *Experimental Cell Research* 219, 1–7.

- Clegg, J. S., Jackson, S. A. and Warner, A. H. (1994). Extensive intracellular translocations of a major protein accompany anoxia in embryos of *Artemia franciscana*. *Experimental Cell Research* **212**, 77–83.
- Haddad, G. G. and Jiang, C. (1993). O₂ deprivation in the central nervous system: on mechanisms of neuronal response, differential sensitivity and injury. *Progress in Neurobiology* **40**, 277–318.
- Haddad, G. G., Sun, Y. A., Wyman, R. J., et al. (1997). Genetic basis of tolerance to O₂ deprivation in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences USA* **94**, 10809–10812.
- Krishnan, S. N., Sun, Y. A., Mohsenin, A., et al. (1997). Behavioral and electrophysiologic responses of *Drosophila melanogaster* to prolonged periods of anoxia. *Journal of Insect Physiology* **43**(3), 203–210.
- Liang, P. and Pardee, A. B. (1992). Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257**, 967–971.
- Ma, E. and Haddad, G. G. (1997). Anoxia regulates gene expression in the central nervous system of *Drosophila melanogaster*. *Molecular Brain Research* **46**, 325–328.
- Ma, E., Xu, T. and Haddad, G. G. (1999). Gene regulation by O₂ deprivation: an anoxia-regulated novel gene in *Drosophila melanogaster*. *Molecular Brain Research* **63**, 217–224.
- Wang, G. L. and Semenza, G. L. (1993). Characterization of hypoxia-inducible factor 1 and regeneration of DNA binding activity by hypoxia. *Journal of Biological Chemistry* **268**, 21513–21518.
- Wang, G. L. and Semenza, G. L. (1995). Purification and characterization of hypoxia-inducible factor 1. *Journal of Biological Chemistry* **270**, 1230–1237.

Drosophila Studies

M Allikian and J Tower

University of Southern California, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition article by M Allikian and J Tower, volume 1, pp 753–754, © 2000, Elsevier Inc.

Stress Response Genes Including the hsp's Are Induced in Response to Heat and Other Stresses and Can Confer Stress Tolerance

Different Environmental Stresses or Laboratory Selection for Altered Stress Resistance Can Affect Gene Allele Frequencies Including Stress Response Genes

Drosophila Mounts an Immune Response to Biotic Stress
Life Span and Aging Phenotypes Correlate with Stress Resistance and a Characteristic Pattern of hsp Expression

Glossary

<i>Chromosome puff</i>	An expansion or decomposition of a polytene chromosome band indicative of altered chromatin structure and high-level transcriptional activity.
<i>P element</i>	The class II DNA transposable element which was engineered to carry modified or foreign genes into the <i>Drosophila</i> genome and allow germ-line transformation.
<i>Polytene chromosome</i>	Lengthwise alignment and coherence of the multiple copies of the chromosome in specialized polyploid cells which allows for light microscope visualization of banding patterns.

Powerful genetic and molecular tools combined with ease of culture have made *Drosophila* a leading model organism for studies of stress. For example, stress response gene induction was first discovered in the early 1960s in *Drosophila* as a characteristic pattern of polytene chromosome puffs in cells subjected to heat or oxidative stress. In the late 1970s *Drosophila* hsp genes were among the first eukaryotic genes cloned. The development of P element-mediated germ-line transformation in 1982 made *Drosophila* one of the first transgenic multicellular organisms. These and other discoveries led to pioneering and ongoing functional studies of stress responses and stress tolerance. A subset of the vast literature on *Drosophila* stress studies is discussed.

Stress Response Genes Including the hsp's Are Induced in Response to Heat and Other Stresses and Can Confer Stress Tolerance

In *Drosophila* six major hsp's are induced in response to heat stress, hsp83, –70, –27, –26, –23, and –22, and several additional proteins exhibit smaller inductions. Hsp83 is constitutively expressed and upregulated severalfold during heat stress. Hsp83 is a member of the *hsp90* gene family which is highly conserved through evolution. Hsp83 functions include regulating the activity of proteins in signal transduction pathways whose functions in turn involve conformational changes. Hsp70 is the most highly conserved through

evolution, and in *Drosophila* is induced over 1000-fold during heat stress. Based on studies of hsp70 family members in *Drosophila* and other organisms, hsp70 is thought to decrease protein denaturation and aggregation, facilitate refolding of partly denatured proteins, facilitate entry of damaged proteins into proteolytic pathways, and protect the organism from additional stress: experimental induction of hsp70 transgenes can confer thermotolerance to cultured *Drosophila* cells, and *Drosophila* transgenic for extra copies of hsp70 exhibit increased thermotolerance at certain stages of development. The *Drosophila* small hsps (–22, –23, –26, and –27) are related to each other and to small hsps from other organisms by a conserved α -crystallin protein domain and also appear to be molecular chaperones involved in stress resistance. The *hsr-omega* gene is also induced by heat stress, but produces no detectable protein product and may function in stress tolerance by an unusual mechanism.

The mechanism of *Drosophila* heat shock gene transcriptional induction by heat stress has been studied in detail, particularly for the *hsp70* gene. Binding sites for the GAGA and TBP (TATA-binding protein) transcription factors and other promoter sequences generate a primed promoter chromatin structure in unstressed cells. This structure includes a transcriptionally engaged RNA polymerase paused ~25 nucleotides downstream of the start site for transcription. This arrangement is thought to facilitate rapid induction upon stress. Heat stress causes trimerization and activation of the constitutively expressed heat-shock transcription factor (HSF), which is required for heat-shock gene induction. Heat-shock transcription factor binds to the heat shock response elements (HSEs) which are evolutionarily conserved promoter elements essential for transcriptional induction during heat stress. Heat-shock transcription factor binding to the HSEs results in release of paused polymerase and high-level transcription of the heat-shock gene.

Induction of the hsps also involves posttranscriptional regulation, as heat-shock gene RNAs are preferentially translated in heat-stressed cells and are more stable during heat stress.

Different Environmental Stresses or Laboratory Selection for Altered Stress Resistance Can Affect Gene Allele Frequencies Including Stress Response Genes

Drosophila isolated from different natural stress environments often exhibit correlated stress response phenotypes. For example, *Drosophila* isolated from hotter environments are generally more resistant to thermal stress. Similarly, laboratory populations of

Drosophila selected over many generations for survival of a particular stress can exhibit increased resistance to that stress and exhibit correlated changes in allele frequencies in specific stress-response genes. Studies of this type provide a model for the relationship between environmental stress and evolution.

Drosophila Mounts an Immune Response to Biotic Stress

The response to septic injury includes rapid induction of antibacterial and antifungal peptides. Induction is mediated in part by the evolutionarily conserved Toll/Dorsal signal transduction pathway, which is homologous to the interleukin-1 receptor NF- κ B immune response pathway in mammals.

Life Span and Aging Phenotypes Correlate with Stress Resistance and a Characteristic Pattern of hsp Expression

Drosophila populations can be genetically selected for increased life span, and these populations exhibit increased resistance to various stresses, including desiccation, starvation, and oxidative damage. Conversely, selection for increased resistance to specific stresses can result in increased life span. A subset of hsps are induced during aging, primarily hsp22 and hsp70. Mild heat stress which induces hsps can cause small increases in the life span of *Drosophila*. Slightly larger heat-induced increases in life span were achieved in *Drosophila* transgenic for extra copies of the *hsp70* gene. Finally, oxidative stress resistance and aging phenotypes have been altered in *Drosophila* transgenic for additional (or modified) oxygen radical defense genes, such as *catalase* and *superoxide dismutase*.

See Also the Following Articles

Drosophila Genes and Anoxia; Heat Shock Genes, Human.

Further Reading

- Belvin, M. P. and Anderson, K. V. (1996). A conserved signalling pathway: The *Drosophila* toll-dorsal pathway. *Annual Review of Cell and Developmental Biology* 12, 393–416.
- Bijlsma, R. and Loeschcke, V. (1997). *Environmental Stress, Adaptation and Evolution*. Basel: Birkhauser-Verlag.
- Hoffmann, A. A. and Parsons, P. A. (1991). *Evolutionary Genetics and Environmental Stress*. Oxford: Oxford University Press.
- Meister, M., Lemaitre, B. and Hoffman, J. A. (1997). Antimicrobial peptide defense in *Drosophila*. *Bioessays* 19, 1019–1026.

Morimoto, R. I., Tissieres, A. and Georgopoulos, C. (1990). *Stress Proteins in Biology and Medicine*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Morimoto, R. I., Tissieres, A. and Georgopoulos, C. (1994). *The Biology of Heat Shock Proteins and Molecular Chaperones*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Parsell, D. A. and Lindquist, S. (1993). The function of heat-shock proteins in stress tolerance: degradation and reactivation of damaged proteins. *Annual Review of Genetics* 27, 437–456.

Tatar, M. (1999). Transgenes in the analysis of lifespan and fitness. *American Naturalist* 154, S67–S81.

Tower, J. (1996). Aging mechanisms in fruit flies. *Bioessays* 18, 799–807.

Drug Use and Abuse

J R Mantsch

Marquette University, Milwaukee, WI, USA

© 2007 Elsevier Inc. All rights reserved.

Stress and Drug Addiction

Clinical Evidence

Animal Studies

Effects of Drug Abuse on Stress Responses

Neurobiological Mechanisms

Treatment Implications

Glossary

<i>Addiction</i>	A chronically relapsing neurobiological disease characterized by an inability to control drug use.
<i>Craving</i>	An intense, consuming, and often uncontrollable desire to use a drug.
<i>Neuroplasticity</i>	The ability of the brain to physically change in response to stimuli such as drugs.
<i>Relapse</i>	The recurrence of drug use after a period of drug abstinence.
<i>Withdrawal</i>	A broad range of adverse symptoms that emerge on the termination or reduction of drug use.

Stress and Drug Addiction

Drug addiction is a chronically relapsing condition that is characterized by an inability to self-regulate drug consumption. In an addicted individual, drug use is compulsive and occurs in spite of negative social, legal, financial, and medical consequences. Despite its prevalence and tremendous cost to society, little progress has been made in the development of effective addiction treatment strategies. One of the issues that makes addiction so difficult to treat is that not all addicts use drugs for exactly the same

reasons. Thus, the factors that contribute to drug use by one individual may be quite different from those that underlie drug self-administration by another. In many individuals, drug use is a stress-driven behavior. This is especially problematic because the life of a drug addict is a particularly stressful one; even though many individuals initially turn to drugs in order to help them cope with stress, drug use by itself can generate and/or aggravate stress responses. The result of this circular relationship between stress and drug use appears to be a vicious cycle within which the onset of stress promotes drug use, which in turn generates more stress and therefore leads to further drug-seeking behavior. Breaking this self-perpetuating cycle by minimizing the influence of stress will probably provide a key to the successful treatment of many drug addicts.

People use drugs because of the effects that they have on the brain. Acutely, drugs produce their euphoric effects through the activation of the neural circuits that underlie natural reward. Such effects tend to offset the negative feelings typically associated with stress. With repeated drug exposure, changes in this neurocircuitry begin to emerge such that the brain of a drug addict is functionally different from that of a non-drug user or someone who only occasionally uses drugs. Significantly, when a drug is used repeatedly, the incentives and neurobiological events that drive drug use change such that the drug user enjoys the drug less but needs the drug more. These changes in brain function that emerge with repeated drug use are collectively referred to as neuroplasticity and appear to be long-lasting, if not permanent. The consequence of such long-lasting neuroplasticity is a diseaselike condition that is associated with a high risk for relapse in many individuals. For this reason, understanding how neuroplasticity that is pathogenic for addiction develops and is expressed is critical for the prevention and treatment of the condition. Similar to its role in the addiction cycle, the

involvement of stress in addiction-related neuroplasticity is bidirectional. Stress promotes the onset of drug-induced neuroplasticity, the expression of which includes aggravated behavioral and hormonal stress responses.

Clinical Evidence

Clinical evidence for a role for stress in drug addiction is largely anecdotal and correlative. In general, stress is cited as a causative factor for drug use, and the inability to effectively cope with stress is predictive of drug relapse. In addition, there is a high incidence of stress-related comorbidity, including posttraumatic stress disorder (PTSD), panic attacks, and depression in drug-dependent populations. For example, it has been reported that up to 43% of cocaine-dependent individuals meet the *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn. rev.; DSM III-R) criteria for lifetime PTSD. In one study, 95.5% of subjects with concurrent PTSD and cocaine dependence reported a functional relationship between cocaine use and their PTSD symptoms, with 86.4% indicating that their PTSD symptoms worsened with drug use.

More direct evidence that stress can lead to drug use is provided by clinical laboratory studies demonstrating that the personalized stress imagery scripts can increase subjective measures of craving for cocaine and alcohol in recovering addicts. In fact, these studies have shown that scripts describing stressful events can be just as effective at eliciting drug craving as scripts describing actual drug use. The ability to directly assess stress-induced craving in human addicts within a laboratory setting should provide a useful tool for testing the effectiveness of novel treatment approaches aimed at preventing stressor-induced drug use.

Animal Studies

The study of drug abuse has relied heavily on preclinical animal research. In particular, drug self-administration procedures in which rats are surgically implanted with intravenous catheters and required to perform a behavioral task, such as pressing a response lever, in order to receive drug infusions have been valuable for the preclinical study of drug abuse and addiction. The validity of such procedures arises from the observation that, with few exceptions, drugs that are abused by humans are also self-administered by rats. Variations of the self-administration procedure have been used to define a role for stress in different aspects of the addiction process. In particular, research has focused on the effects of stressors on the acquisition, escalation, and reinstatement of

drug-seeking behavior. A brief summary of this research is provided next.

Acquisition

Not everyone who tries a drug for the first time will like it and will start using the drug on a regular basis. Individual differences in initial responsiveness to drugs exist and can be influenced by a number of variables. Some individuals are genetically predisposed toward drug abuse, whereas others are influenced by environmental factors such as stress. It is likely that for many, individual susceptibility to addiction arises from a complex interplay among genes, drug effects, and environment such that the genetic potential for addiction is realized only if the drug is used in the wrong environmental context.

Preclinically, factors that increase the likelihood that drug use will be initiated can be studied using acquisition self-administration models in which predisposition to engage in drug use is defined according to how readily drug-naïve subjects acquire the behavior of drug self-administration. Using such models, it has been reported that a variety of stressful stimuli ranging from very basic stressors (such as repeated tail pinch) to more complex stressors (such as social defeat) facilitate the acquisition of drug self-administration by rats. These effects of stress do not appear to be short-lived. In fact, it has been reported that when rats are exposed to a stressor during the first weeks of life (e.g., daily isolation) more rapid acquisition of drug self-administration can be observed months later during adulthood, suggesting that the onset of stress, especially at critical phases of life, may permanently increase individual susceptibility to drug addiction. Interestingly, it is not just stress but the ability to control stress that is important for its effects on acquisition. Rats exposed to electric foot-shock stress acquire cocaine self-administration at lower dose than nonshocked controls when they are unable to control shock delivery but not when the same amount of shock is delivered under response-contingent conditions. These findings parallel clinical reports that it is not the onset of stress *per se* that predicts drug use but rather the coping strategy that is used when an individual is confronted with a stressor. Understanding how stress serves as a determinant of the initiation of drug use will permit the identification of populations that are at risk for addiction and should aid efforts to prevent drug abuse.

Escalation

Just because someone uses a drug regularly does not mean that that person is addicted. It is the loss of control over drug use and its negative impact on

other facets of a drug user's life that define addiction. This aspect of drug addiction has been studied using rodent drug self-administration models in which progressively escalating patterns of drug intake emerge when daily access to a drug is prolonged. The onset of escalating use patterns is thought to reflect the transition from controlled drug use to out-of-control drug addiction. For this reason, further characterization of the variables that promote escalating self-administration patterns in rodents should provide insight into factors that accelerate the onset of drug addiction. Recently, our laboratory has begun to examine the impact of chronic stress delivered across an extended period of ongoing cocaine self-administration. To date, our studies have shown that exposure to stress at the time of daily cocaine self-administration progressively escalates drug intake over a 14-day period, suggesting that, like prior exposure to cocaine, repeated stress can accelerate the onset of addiction. Significantly, these findings imply that in addition to acutely generating drug use, stress can aggravate addiction by promoting drug-induced neuroplasticity.

Reinstatement

In drug addicts, craving, or the intense desire to use a drug, typically precedes drug use. The onset of craving is often unpredictable and can suddenly occur even after months or years of drug abstinence, making it a formidable obstacle to the effective long-term management of drug addiction. Drug craving can be studied preclinically using reinstatement self-administration procedures in which the abilities of various stimuli to reinstate extinguished drug-seeking behavior are examined. With these procedures, drug self-administration is extinguished by replacing the drug with an inert substance. As a result, responding previously maintained by presentation of the drug declines. The restoration or reinstatement of extinguished behavior under various experimental stimuli is thought to reflect the onset of craving for that drug and therefore can be used to identify the risk factors for drug relapse. Using such procedures, it has been demonstrated that acute exposure to stressors can reinstate responding previously reinforced by a variety of abused drugs including cocaine, ethanol, heroin, and nicotine without increasing responding not previously associated with drug delivery. These findings parallel those from clinical studies showing that subjective stress imagery can increase measures of drug craving in recovering addicts and suggest that the inclusion of pharmacological and behavioral strategies aimed at reducing stress during drug abstinence is probably critical to the success of any addiction treatment program.

Effects of Drug Abuse on Stress Responses

In addition to acting as a causative factor for drug-seeking behavior, stress also emerges as a consequence of drug use. It is well known that for most drugs of abuse, intense stress and anxiety are prominent features of acute drug withdrawal. In animal studies, these withdrawal symptoms can be observed as exaggerated anxiety-like behaviors measured using a variety of behavioral tests following withdrawal from repeated drug exposure. Significantly, altered anxiety responses appear to persist long after other acute withdrawal symptoms dissipate. These augmented anxiety-like responses during late withdrawal probably translate into increased susceptibility to drug relapse during periods of stress. One way to examine the effects of prior drug use on relapse is to compare stressor-induced reinstatement during withdrawal in rats with different histories of drug exposure. When rats have a history of self-administering greater amounts of drug, they also display higher levels of reinstatement in response to stress, suggesting that the amount of prior drug use can alter responsiveness to stress in a way that contributes to further drug-seeking behavior. These findings are paralleled by the results of clinical studies showing that stress-induced cravings, anxiety, and physiological responses in abstinent cocaine users are augmented in individuals with a history of high-frequency cocaine use compared to individuals with a history of using the drug less frequently. Our own data indicate that the behavioral responses to stressors are not simply exaggerated as a consequence of drug self-administration but, rather, are qualitatively different such that rats with a history of cocaine self-administration appear to be more assertive when confronted with a stressful situation, possibly reflecting a maladaptive coping strategy during stress that may result from drug use.

Neurobiological Mechanisms

Like all factors that influence drug-seeking behavior, stressors exert their effects on drug use by altering the activity of the neurobiological systems within the brain that underlie motivated behavior. In particular, a critical role for a pathway in the brain consisting of nerve projections from the medial prefrontal cortex that release the amino acid neurotransmitter glutamate into the nucleus accumbens has been identified for drug-seeking behavior. This neural subcircuit is closely modulated by the neurotransmitter dopamine, released from nerves that originate in the ventral tegmental area (VTA) of the midbrain. Like other stimuli that acutely generate drug use, stress provokes

drug-seeking behavior through the activation of this pathway. The mechanism through which this activation occurs during stress is not entirely clear, but it appears to require the stimulation of dopaminergic neurons in the VTA that project to the medial prefrontal cortex and almost certainly involves the neuropeptide corticotropin releasing hormone (CRH). Significantly, stress can intensify the responsiveness of this motivational pathway to subsequent stimulation by inducing a form of cellular neuroplasticity in dopaminergic cells in the VTA called long-term potentiation (LTP), thereby increasing vulnerability to the effects of abused drugs.

Corticotropin Releasing Hormone

In addition to its role as an initiator of the hormonal stress response through its actions in the pituitary gland, CRH exerts its effects within the brain, where it functions as a neuropeptide mediator of stress-related behavioral responses and anxiety. CRH administered directly into a number of brain regions, including the VTA, reinstates extinguished drug-seeking behavior, whereas CRH antagonists prevent stress-induced reinstatement. The exact mechanism through which CRH stimulates dopamine cells in the VTA is unclear, but it appears to require the association of CRH with its binding protein and the activation of a specific subtype of the CRH receptor, the CRH₂ receptor, which in turn facilitates cellular activation by the neurotransmitter glutamate. The likely role for CRH as a mediator of stress-induced drug-seeking behavior makes the peptide an obvious target for the development of pharmacotherapy aimed at promoting drug abstinence by preventing relapse during periods of stress. Although promising, clinical evaluation of the usefulness of drugs that interfere with the actions of CRH awaits the availability of selective nonpeptide CRH-receptor-blocking drugs that can readily penetrate the blood-brain barrier to reach sites of CRH action in the brain.

Glucocorticoids

Although the ability of acute stress to precipitate drug craving and use is problematic for addicts, it is chronic stress that can have an especially deleterious impact on the addiction process by producing or facilitating neuroplasticity that is pathogenic for addiction. Whereas the acute effects of stress on drug-seeking behavior appear to be mediated primarily through the actions of CRH in the brain, many of the effects of chronic stress on drug use appear to involve the release of glucocorticoid hormones from the adrenal gland. Glucocorticoids (such as cortisol in humans and corticosterone in rats) are critical mediators of

the physiological and behavioral responses that enable an organism to adapt to and cope with stress. Such responses are beneficial for short-term adaptation to stressful stimuli, but, with constant activation, they can be detrimental to an organism, a consequence of chronic stress referred to as allostatic load. These allostatic effects appear to include adaptations in the motivational neurocircuitry of the brain that promote illicit drug use. Although glucocorticoids alone are capable of producing addiction-related neuroplasticity, it appears that in many cases these hormones work in concert with other stress-related molecules such as CRH to influence the addiction process. It is likely that many such interactions occur in the VTA, where stressor-induced neuroplasticity involving dopamine cells appears to be mediated by CRH and may require glucocorticoid receptor activation. Further research is needed to determine whether glucocorticoids and their receptors are viable targets for the pharmacotherapeutic management of drug addiction. A summary of one putative neurobiological pathway through which stress influences drug-seeking behavior is provided in [Figure 1](#).

Treatment Implications

The recognition that stress is a key contributor to drug abuse and addiction should highlight the need for the development and implementation of novel therapeutic strategies aimed at minimizing the contribution of stress to the addiction process, especially in subpopulations of addicts whose drug use is stress-driven. Identifying which addicts will benefit most from such approaches poses a challenge to drug-dependence treatment providers and will probably require the establishment of new assessment tools that permit the determination of the role of stress in drug use on an individual basis. Although available, cognitive and behavioral therapies aimed at helping recovering addicts manage their stress are underused and understudied. Further, many drugs that are traditionally used for the treatment of stress-related psychiatric conditions (e.g., antidepressants and benzodiazepines) are not commonly prescribed to treat addiction and, in many cases, are avoided even though they may be beneficial if used properly in the right individuals. The eventual development and approval of new drugs that block the actions of CRH in the brain will, it is hoped, provide important tools for the management of drug addiction. At the same time, basic preclinical research must continue to examine the neurobiological mechanisms through which stress influences drug addiction so that new targets (e.g., the orexins/hypocretins) can be discovered and more effective drugs can be developed.

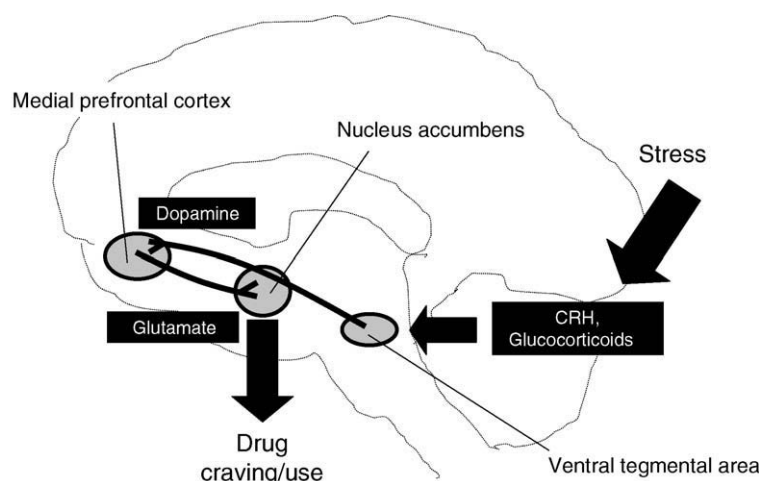


Figure 1 One putative neurobiological pathway through which stressful stimuli may influence drug-seeking behavior and addiction. The results of recent animal studies indicate that stressful stimuli increase brain corticotropin releasing hormone (CRH) levels, which, in concert with elevated glucocorticoids, exert effects on cells in the ventral tegmental area of the brain to increase dopaminergic neurotransmission in the medial prefrontal cortex and nucleus accumbens and facilitate the activity of neurocircuitry that underlies motivation and natural reward.

See Also the Following Articles

Alcohol, Alcoholism and Stress: A Psychobiological Perspective; Alcohol and Stress: Social and Psychological Aspects; Allostasis and Allostatic Load; Corticotropin Releasing Factor (CRF); Glucocorticoids, Role in Stress; Interactions between Stress and Drugs of Abuse; Opioids.

Further Reading

- Back, S. E., Brady, K. T., Jaanimagi, U., et al. (2006). Cocaine dependence and PTSD: a pilot study of symptom interplay and treatment preferences. *Addictive Behaviors* 31, 351–354.
- Breese, G. R., Chu, K., Dayas, C. V., et al. (2005). Stress enhancement of craving during sobriety: a risk for relapse. *Alcoholism: Clinical and Experimental Research* 29, 185–195.
- Fox, H. C., Talih, M., Malison, R., et al. (2005). Frequency of recent cocaine and alcohol use affects drug craving and associated responses to stress and drug-related cues. *Psychoneuroendocrinology* 30, 880–891.
- Goeders, N. E. (2002). Stress and cocaine addiction. *Journal of Pharmacology and Experimental Therapeutics* 301, 785–789.
- Kalivas, P. W. and McFarland, K. (2003). Brain circuitry and the reinstatement of cocaine-seeking behavior. *Psychopharmacology* 168, 44–56.
- Koob, G. F., Ahmed, S. H., Boutrel, B. P., et al. (2004). Neurobiological mechanisms in the transition from drug use to drug dependence. *Neuroscience and Biobehavioral Reviews* 27, 739–749.
- Kreek, M. J., Nielsen, D. A., Butelman, E. R., et al. (2005). Genetic influences on impulsivity, risk taking, stress responsivity and vulnerability to drug abuse and addiction. *Nature Neuroscience* 8, 1450–1457.
- Marinelli, M. and Piazza, P. V. (2002). Interaction between glucocorticoid hormones, stress and psychostimulant drugs. *European Journal of Neuroscience* 16, 387–394.
- Saal, D., Dong, Y., Bonci, A., et al. (2003). Drugs of abuse and stress trigger a common synaptic adaptation in dopamine neurons. *Neuron* 37, 577–582.
- Sarnyai, Z., Shaham, Y. and Heinrichs, S. C. (2001). The role of corticotropin-releasing factor in drug addiction. *Pharmacological Reviews* 53, 209–243.
- Shalev, U., Grimm, J. W. and Shaham, Y. (2002). Neurobiology of relapse to heroin and cocaine seeking: a review. *Pharmacological Reviews* 54, 1–42.
- Sinha, R. (2001). How does stress increase risk of drug abuse and relapse? *Psychopharmacology* 158, 343–359.
- Vocci, F. J., Acri, J. and Elkashef, A. (2005). Medication development for addictive disorders: the state of the science. *American Journal of Psychiatry* 162, 1432–1440.
- Volkow, N. D. and Li, T. K. (2004). Drug addiction: the neurobiology of behaviour gone awry. *Nature Reviews Neuroscience* 5, 963–970.
- Wang, B., Shaham, Y., Zitzman, D., et al. (2005). Cocaine experience establishes control of midbrain glutamate and dopamine by corticotropin-releasing factor: a role in stress-induced relapse to drug seeking. *Journal of Neuroscience* 25, 5389–5396.

E

Early Environment and Adult Stress See: Stress Hyporesponsive Period; Eclampsia and preeclampsia; Fetal Stress; Pregnancy, Maternal and Perinatal Stress, Effects of; 11 β -Hydroxysteroid Dehydrogenases.

Earthquakes, Stress Effects of

M Livanou

King's College London, London, UK, and Hellenic
Institute of Psychotraumatology, Athens, Greece

M Başoğlu

King's College London, London, UK and Istanbul Center
for Behavior Research and Therapy, Istanbul, Turkey

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by V J Carr, volume 2, pp 1–3, © 2000, Elsevier Inc.

Psychological Consequences

Factors Associated with Earthquake-Related Traumatic
Stress

Treatment of Earthquake Survivors

Prevention

*Prospective
cohort study
Randomized
controlled
clinical trial*

*Selective sero-
tonin reuptake
inhibitors
(SSRIs)*

avoidance of stimuli associated with the traumatic stressor, numbing of emotional responsiveness, and increased arousal. A study in which the subjects are identified and then followed forward in time. A clinical trial in which investigators randomly assign eligible subjects to groups to receive or not receive one or more interventions that are being tested or compared. It is a design widely considered reliable because it eliminates a variety of biases that compromise the validity of medical research.

A class of antidepressants that increase serotonin levels by inhibiting serotonin reuptake.

Glossary

*Cognitive-
behavioral
treatment*

A psychological treatment involving the modification of maladaptive thinking patterns and systematic exposure to anxiogenic stimuli until the associated anxiety subsides.

*Earthquake
epicenter
Earthquakes*

The point on Earth's surface directly above the focus of an earthquake.

Natural phenomena that involve abrupt shaking and vibration of the Earth's crust caused by the release of stress accumulated along geologic faults or by volcanic activity.

*Longitudinal
study*

A study that involves observations of the same items over a period of time.

*Posttraumatic
stress disorder
(PTSD)*

An anxiety disorder that develops after exposure to a traumatic stressor and involves symptoms of persistent reexperiencing of the traumatic stressor,

Earthquakes are common natural disasters that can cause widespread devastation and casualties and expose large numbers of people to bereavement, injury, loss of property, homelessness, and displacement. Considering only the earthquakes that killed more than 1000 people, approximately 1.8 million people died in 108 earthquakes in the twentieth century. Earthquakes often cause greater devastation and casualties in developing countries because of the generally low quality of buildings, lack of disaster preparedness, and inadequate rescue and relief efforts. Indeed, 91 of the 108 major earthquakes (with a death toll over 1000) in the twentieth century occurred in developing countries, accounting for 83% of 1.8 million deaths worldwide.

Exposure to earthquakes is associated with psychological distress, PTSD, and depression. Although the extent of devastation and casualties caused by a disaster is thought to relate to the severity of the

subsequent psychological distress, recent studies suggest that even earthquakes of relatively low magnitude or those that cause limited devastation can lead to extensive posttraumatic stress reactions. Consequently, earthquakes pose a mental health hazard not only in developing countries but also in industrialized countries, even when people suffer relatively less damage and fewer casualties.

Psychological Consequences

The Earthquake Experience

The violent tremors of a major earthquake can cause serious damage to solid structures. People may be killed, injured, trapped under rubble or may witness the injury or death of others and the loss of their property. The impact of the earthquake experience is associated with many factors, including earthquake characteristics (magnitude, duration, etc.), proximity to the epicenter of the earthquake, strength of the buildings, and the nature of the ground that supports the buildings.

Fear and helplessness are among the most common emotional reactions that people experience during an earthquake. Intense fear during earthquakes has also been linked to deaths from cardiac arrests and extreme behaviors that reflect panic (e.g., jumping from the window of a tall building).

Acute Phase

In the first few days and weeks after a major earthquake, the affected communities focus on rescuing those who have been trapped under rubble, caring for the injured, assessing the extent of the damage, and planning the management of the problems caused by the earthquake. Bereaved survivors face the task of burying their loved ones, and homeless survivors seek temporary accommodation. Ongoing aftershocks may cause further damage and casualties.

In the acute phase, the regular activities of the affected communities may be disrupted. Some survivors avoid entering their relatively undamaged or mildly damaged homes either because they are uncertain about their safety or because of fear of further earthquakes. Governments and relief organizations often set up tent camps and prefabricated housing sites to accommodate homeless survivors. In addition, the fearful anticipation of further earthquakes is prevalent.

Relatively little is known about the psychological status of survivors in the acute-phase aftermath of a major earthquake. The few existing studies that involved assessments 1–4 weeks postearthquake were based on convenience or clinical samples and reported

increased rates of symptoms of acute stress disorder (dissociative symptoms, reexperiencing, avoidance of trauma reminders, and anxiety or increased arousal) and depression. In cases in which the earthquake caused casualties, the acute phase may also involve bereavement reactions.

Medium-Term Effects

In the year that follows a major disaster, many survivors face the difficult task of coming to terms with bereavement, injury, material loss, homelessness, disruption of activities, relocation, and other earthquake-related stressors. Those living in shelters (tents, makeshift barracks, and prefabricated houses) may be deprived of basic needs and face disruption in their daily routine. Ongoing aftershocks may cause additional traumatic stress.

Studies examining the psychological status of adult and child survivors 1–18 months postearthquake show considerable variation in methodology (e.g., sampling, severity of earthquake exposure and related devastation, assessment instruments, and time between the earthquake and assessment) and, consequently, in the reported rates of psychiatric problems. Most studies (Table 1) reported high rates of PTSD and depression among the survivors. In addition, although little is known about the incidence of other psychiatric disorders in earthquake survivors, there is evidence that earthquakes may lead to increased rates of generalized anxiety disorder and alcohol abuse/dependence. The psychological effects of earthquakes

Table 1 Range of estimated rates of PTSD and major depression^a

Type of sample ^b	PTSD ^c (%)	Major depression ^c (%)
Probability or consecutive	2–43	7–52
Convenience	9–95	13–42
Clinical	63–74	22–42

^aPTSD, posttraumatic stress disorder.

^bIn probability sampling there is random selection of study subjects, thus ensuring that each member of the population has equal probabilities of being chosen. In consecutive sampling, each household is targeted for assessment within a particular temporary housing site (tent-camp or prefabricated housing site) in the disaster region. In convenience sampling, the selection of subjects is based on easy availability and/or accessibility. In clinical sampling, the participants are drawn from among patients or treatment-seekers.

^cThe wide range of estimated reported rates mainly reflects differences in methodology (sampling, assessment instruments, time lapse between earthquake and assessment, etc.) and in the severity of the earthquake that the subjects of each study experienced.

may be less severe in professional rescuers, possibly because they tend to be more psychologically prepared for disasters and their consequences.

Long-Term Effects

Relatively few studies have examined the psychological status of survivors beyond 18 months post-earthquake. In a prospective cohort study based on survivors with high exposure to the 1989 earthquake in Armenia, the rates of PTSD 18 and 54 months after the disaster were 87% and 73%, respectively. Although symptoms of depression tended to subside over time, the severity of PTSD in survivors with high earthquake exposure did not show a significant reduction. In another study following the 1989 Newcastle earthquake, 48% of the survivors who had PTSD at 6 months postearthquake still had PTSD 2 years later. Earthquake-related morbidity showed a decline in the months after the earthquake, but stabilized at approximately 18 months postdisaster. Finally, in a longitudinal study based on survivors of the 1998 earthquake in northern China, the rates of lifetime PTSD at 3 and 9 months postearthquake were 19% and 24%, respectively.

The long-term consequences of earthquakes were also reflected in a cross-sectional survey of survivors of the 1999 earthquake in Turkey. This study showed that 39% of the survivors living in shelters had PTSD and 19% had depression 20 months after the disaster. Similarly, other cross-sectional studies that involved assessments more than 3 years after an earthquake reported increased incidence of psychological distress and traumatic stress reactions in the survivors.

Factors Associated with Earthquake-Related Traumatic Stress

The lowest rates of psychiatric morbidity tend to be reported by survivors of earthquakes that caused relatively less devastation and fewer casualties, whereas survivors of more devastating earthquakes often show substantially higher rates of psychopathology. In general, existing studies suggest a strong positive association (dose-response relationship) between the degree of earthquake exposure and subsequent traumatic stress responses. Therefore, high exposure to earthquakes tends to relate to greater postearthquake psychological distress. High exposure to earthquakes involves exposure not only to more intense tremors but also to various other related traumatic stressors, such as damage to or collapse of the house, being trapped under rubble, physical injury, participation

in rescue work, witnessing grotesque scenes, loss of close ones, loss of property, and being relocated. On the other hand, low earthquake exposure involves experiencing relatively low-intensity tremors, with less exposure to additional stressors.

Recent studies suggested that the intensity of fear during an earthquake is a much stronger predictor of earthquake-related PTSD than other traumatic events, such as the extent of damage to house, being trapped under rubble, participation in rescue efforts, and being inside a building at the time of the earthquake. In addition, in most of these studies PTSD related to fear during the earthquake and other traumatic events, whereas depression related to loss (bereavement and material loss) and personal characteristics (e.g., history of psychiatric illness, marital status, and age). Other predictors of PTSD included female gender, personal or family history of psychiatric problems, and lower educational level.

Treatment of Earthquake Survivors

The interventions that have been used in the treatment of earthquake survivors include play or art therapy for children, family therapy, a didactic experiential group instruction program, debriefing, and brief crisis-oriented psychotherapy. Because most of these treatments were not tested by randomized controlled studies, their effectiveness has not yet been established.

In a randomized controlled study involving child survivors, 10 sessions of group play therapy reduced anxiety, but had no significant effect on depression and overall life adjustment. In another randomized controlled trial, six sessions of a combination of various interventions (trauma reconstruction and reprocessing, problem solving, relaxation techniques, and grief counseling) led to only a partial improvement in PTSD and had no effect on depression.

The usefulness of pharmacological agents in earthquake-related PTSD has not yet been adequately examined. In general, the PTSD literature contains relatively few randomized placebo-controlled trials of psychotropic drugs. The treatment effect sizes obtained in these studies were fairly small. Although there are some studies suggesting that SSRIs may be useful in reducing PTSD, most of them have not included a treatment-free follow-up period. Thus, we do not know whether the treatment gains were maintained in the long term in these studies. Furthermore, considering that relapse on discontinuation of medication is fairly common in anxiety disorders, the use of psychotropic medications in earthquake-related PTSD appears to be of dubious value.

Brief Modified Behavioral Interventions

The effectiveness of cognitive-behavioral treatment for PTSD is well established, and there is consensus among experts that it is the psychological treatment of choice for PTSD. Although cognitive-behavior therapy is considered a brief treatment (usually delivered in about 10 sessions), it is still too long in post-disaster circumstances in which thousands of people may require help. Furthermore, other factors, such as postdisaster hardships and demographic mobility, often interfere with the survivors' ability to attend treatment sessions regularly. Such circumstances require much briefer interventions, preferably involving no more than one or two sessions.

A series of studies in Turkey examined whether cognitive-behavioral treatment could be shortened without compromising its effectiveness. In an open clinical trial, 231 survivors with chronic PTSD symptoms received a modified version of behavioral treatment and were assessed weekly to determine the minimum number of sessions required for significant clinical improvement. The treatment involved no systematic cognitive restructuring and focused only on reducing behavioral avoidance by giving instructions for systematic self-exposure to feared and avoided situations. The rationale of the treatment was not based on habituation that emphasizes reduction in distress during exposure (e.g., "Stay in the situation until your anxiety subsides") but on enhancement of sense of control over fear and distressing trauma reminders (e.g., "Stay in the situation until you gain control over your anxiety/distress").

The survivors were given a mean of 4.3 sessions and a survival analysis showed that 76% of them improved after one session and 88% after two sessions. The mean number of sessions required for significant improvement was 1.7. The results revealed a patholytic treatment effect, with significant improvement in PTSD, depression, and overall adjustment. Improvement was maintained at the 3- and 9-month follow-ups; of the 75 survivors who had a 9-month follow-up assessment, only one showed relapse.

This study suggested that modified behavioral treatment may be effectively delivered in one session. To examine this hypothesis, a randomized waitlist controlled trial was conducted, involving 59 earthquake survivors who were given only one session of modified behavioral treatment. Blind assessment at week 6 showed significant treatment effects on all study measures, including PTSD, depression, fear of earthquakes, and overall adjustment. The overall improvement rate was 49% at week 6, 80% at week 12, 85% at week 24, and 83% at 1–2 years posttreatment.

In another clinical study, 10 survivors with chronic PTSD symptoms received a single session of exposure treatment in an earthquake simulator (a small furnished house resting on a shake table that simulates tremors on nine intensity levels). A control switch inside the simulator allowed survivors to stop and start the tremors and increase their intensity whenever they felt ready for it. The treatment was aimed at enhancing the survivors' sense of control over simulated tremors and anxiety-evoking memories of trauma. Assessments at pre- and postsession and at 2, 4, 8, and 12 weeks posttreatment revealed significant improvement in all measures of psychopathology. Eight patients were markedly and two slightly improved at follow-up. The effectiveness of this treatment was further ascertained in a randomized controlled treatment that involved 30 subjects.

These studies also provided some evidence that fear reduction through modified behavioral treatment may have protective effects against the traumatic effects of future earthquakes. Many of the survivors were treated during the first year after the disaster when the aftershocks were continuing. Despite the additional traumatic effects of these aftershocks, relapse after recovery was rare, suggesting that the treatment had a protective effect. The survivors often reported that they no longer panicked during real earthquakes as they used to in similar occasions before treatment. Behavioral treatment involving exposure to simulated tremors in an earthquake simulator was particularly effective in reducing fear and enhancing a sense of control over future earthquakes, suggesting that this intervention might have a potential use in increasing psychological preparedness for earthquakes.

Dissemination of treatment to large numbers of survivors after natural disasters is a critical issue, particularly in developing countries with limited mental health-care resources. A pilot study in Turkey examined the usefulness of a highly structured self-help manual in delivering modified behavioral treatment to survivors. When the manual was delivered to 89 survivors with PTSD in the community, one in four survivors read the manual, complied with the treatment instructions, and reported significant improvement in their problems.

These studies show that brief and modified versions of behavioral treatment are both practical and effective in postearthquake settings. These interventions lead to significant and sustainable improvement while reducing therapist time to a minimum. In addition, the dissemination of the treatment through self-help manuals to large numbers of survivors may be a cost-effective way of dealing with the widespread posttraumatic stress responses in the community.

Further research is needed to examine whether the development of PTSD can be prevented by brief modified behavioral treatment delivered in the early aftermath of an earthquake. In addition, randomized controlled studies are needed to confirm the effectiveness of various methods of treatment dissemination (e.g., through manuals, computers, videocassettes, and television).

Prevention

Given the inevitability of earthquakes, preparedness for future earthquakes is of utmost importance. Undoubtedly, the first preventive measure against the mental health hazard posed by earthquakes involves measures designed to reduce the extent of devastation and casualties caused by earthquakes, such as stricter building regulations, effective emergency response policies, and training of the public in self-protection. Accordingly, much of the effort has concentrated on this aspect of earthquake preparedness. What is less clear, however, is how people can be prepared psychologically against earthquakes. Recent work in Turkey provided some insight into effective methods of enhancing psychological preparedness in people at risk of earthquakes. Considering that fear plays an important role in the development of PTSD in earthquake survivors, interventions designed to enhance a sense of control over fear might be expected to increase psychological resilience against the traumatic effects of earthquakes. Evidence suggests that this could be effectively achieved by behavioral interventions, particularly by those involving the use of an earthquake simulator. Further research in this area appears to be worthwhile.

Further Reading

- Armenian, H. K., Morikawa, M., Melkonian, A. K., et al. (2000). Loss as a determinant of PTSD in a cohort of adult survivors of the 1988 earthquake in Armenia: implications for policy. *Acta Psychiatrica Scandinavica* 102, 58–64.
- Başoğlu, M., Kiliç, C., Şalcioğlu, E., et al. (2004). Prevalence of posttraumatic stress disorder and comorbid depression in earthquake survivors in Turkey: an epidemiological study. *Journal of Traumatic Stress* 17, 133–141.
- Başoğlu, M., Livanou, M., Şalcioğlu, E., et al. (2003). A brief behavioral treatment of chronic PTSD in earthquake survivors: results from an open clinical trial. *Psychological Medicine* 33, 647–654.
- Başoğlu, M., Livanou, M. and Şalcioğlu, E. (2003). A single-session with an earthquake simulator for traumatic stress in earthquake survivors. *American Journal of Psychiatry* 160, 788–790.
- Başoğlu, M., Şalcioğlu, E. and Livanou, M. (2002). Traumatic stress responses in earthquake survivors in Turkey. *Journal of Traumatic Stress* 15, 269–276.
- Başoğlu, M., Şalcioğlu, E., Livanou, M., et al. (2005). Single-session behavioral treatment of earthquake related posttraumatic stress disorder: a randomized waitlist controlled trial. *Journal of Traumatic Stress* 18, 1–11.
- Carr, V. J., Lewin, T. J., Webster, R. A., et al. (1995). Psychosocial sequelae of the 1989 Newcastle earthquake. I: Community disaster experiences and psychological morbidity 6 months post-disaster. *Psychological Medicine* 25, 539–555.
- Carr, V. J., Lewin, T. J., Webster, R. A., et al. (1997). Psychosocial sequelae of the 1989 Newcastle earthquake. II: Exposure and morbidity profiles during the first 2 years post-disaster. *Psychological Medicine* 27, 167–178.
- Goenjian, A. K., Najarian, L. M., Pynoos, R. S., et al. (1994). Posttraumatic stress disorder in elderly and younger adults after the 1988 earthquake in Armenia. *American Journal of Psychiatry* 151, 895–901.
- Goenjian, A. K., Steinberg, A. M., Najarian, L. M., et al. (2000). Prospective study of posttraumatic stress, anxiety, and depressive reactions after earthquake and political violence. *American Journal of Psychiatry* 157, 911–916.
- Livanou, M., Başoğlu, M., Şalcioğlu, E., et al. (2002). Traumatic stress responses in treatment-seeking earthquake survivors in Turkey. *Journal of Nervous and Mental Disease* 190, 816–823.
- Livanou, M., Kasvikis, Y., Başoğlu, M., et al. (2005). Earthquake-related psychological distress and associated factors four years after the Parnitha earthquake in Greece. *European Psychiatry* 20, 137–144.
- McMillen, J. C., North, C. S. and Smith, E. M. (2000). What parts of PTSD are normal: intrusion, avoidance, or arousal? data from the Northridge, California, earthquake. *Journal of Traumatic Stress* 13, 57–75.
- Noji, E. K. (1997). Earthquakes. In: Noji, E. K. (ed.) *The public health consequences of disasters*, pp. 135–178. New York: Oxford University Press.

Relevant Websites

National Earthquake Information Center (2000). Earthquakes with 1,000 or more deaths from 1900. www.neic.cr.usgs.gov.

Eating Disorders and Stress

D C Jimerson

Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 4–8, © 2000, Elsevier Inc.

Clinical Overview of Eating Disorders

Anorexia Nervosa

Bulimia Nervosa

Etiological Influences in Eating Disorders

Glossary

<i>Amenorrhea</i>	A symptom of anorexia nervosa in adolescent and adult women, characterized by the failure of menstrual cycles to occur during three or more consecutive expected cycle intervals.
<i>Binge-eating episode</i>	consumption of an abnormally large amount of food in a discrete episode, during which eating feels out of control.
<i>Hypothalamus</i>	A region in the lower part of the brain involved in the regulation of eating behavior, as well as the modulation of other metabolic and hormonal systems.
<i>Neuro-transmitter</i>	A naturally occurring neurochemical which participates in the selective transmission of signals in neuronal pathways modulating behavior and physiological functions.
<i>Purging behaviors</i>	In patients with eating disorders, symptoms intended to control body weight by eliminating ingested foods and liquids; examples include self-induced vomiting and misuse of laxatives.
<i>Restricting behavior</i>	In patients with eating disorders, a pattern of strict limitation of food intake in order to control body weight.
<i>Serotonin</i>	A naturally occurring neurochemical which acts as a neurotransmitter in the central nervous system, including pathways modulating meal size and mood.

Clinical Overview of Eating Disorders

In recent years there has been increasing awareness of the prevalence of anorexia nervosa and bulimia nervosa, and the personal and social costs associated with these disorders. This awareness has been reflected in attention to these problems in the lay press and in professional publications. Clinical investigations have

yielded new insights into the prevalence and symptom patterns of anorexia nervosa and bulimia nervosa. Although the etiology of these disorders is unknown, potential psychosocial and biological risk factors have been identified, and new treatment approaches have been developed. This article provides an overview of clinical characteristics of these disorders, and highlights several directions of recent research advances. Consultation with a medical and/or mental health professional is important for an individual seeking specific guidance regarding clinical evaluation and treatment.

Anorexia Nervosa

Clinical Characteristics

Anorexia nervosa is widely recognized as a syndrome of weight loss which is associated with fears of unwanted weight gain, and is not a consequence of other medical illness. Symptoms of anorexia nervosa were first described in the medical literature more than 100 years ago. Specific diagnostic features of the syndrome are outlined in the *Diagnostic and Statistical Manual of Mental Disorders, 4th edn. (DSM-IV)*. The most prominent features of anorexia nervosa include weight loss below 85% of an individual's expected body weight, associated with an intense fear of excessive weight gain and abnormal perception of body-weight status. An additional characteristic of the disorder in adolescent and adult females is a period of amenorrhea, usually thought to be a consequence of low body weight. Approximately half of individuals with anorexia nervosa exhibit binge-eating or purging behaviors associated with efforts to maintain a low weight, while others engage predominantly in food restriction. Based on DSM-IV diagnostic categories, these individuals are grouped into binge eating-purging type and restricting type, respectively.

Anorexia nervosa is associated with a range of medical problems which are thought to be primarily related to malnutrition and weight loss. Among the more common problems are anemia, abnormalities in blood pressure and heart rhythm, and alterations in blood electrolyte levels. Activity in the hypothalamic-pituitary-adrenal axis is often increased, with malnutrition-related elevations in blood levels of the stress-related hormone, cortisol. Consistent with physiological adaptations to conserve energy, functional activity in the hypothalamic-pituitary-thyroid axis and the sympathetic nervous system are frequently decreased. Reduction in bone mineral density with increased risk of bone fractures is often a

long-term consequence of severe malnutrition in anorexia nervosa.

Naturalistic follow-up studies indicate that at 5–10 years after initial assessment, about 40–60% of patients achieve stable weight recovery, while 10–20% of patients continue to have symptoms of the disorder. In general, it is difficult to predict an individual's long-term outcome based on initial symptomatology. The estimated mortality associated with anorexia nervosa is approximately 6% during the first 10 years after initial assessment, with variability in findings most likely reflecting the severity of initial symptomatology.

Epidemiology and Psychiatric Comorbidity

The average age of onset of anorexia nervosa is approximately 18 years, although a significant number of patients experience an early age of onset (i.e., at 8–14 years). Anorexia nervosa is approximately 10 times more prevalent in females than in males. Among young women, the group at highest risk for the disorder, the prevalence of anorexia nervosa has been estimated as approximately 0.5%. Although epidemiological data are limited, it is known that anorexia nervosa can occur in individuals of various ethnic backgrounds and socioeconomic classes.

Individuals with anorexia nervosa often have a history of additional psychiatric symptoms, particularly depression and anxiety disorders, including obsessive-compulsive disorder. It is of note that marked weight loss in healthy volunteers has been shown to result in depressive symptoms and preoccupation with food. Thus, for some anorexic patients, depressive symptoms may improve with nutritional stabilization and weight restoration.

Overview of Treatment Approaches

For patients with acute weight loss with life-threatening medical and psychiatric symptoms, hospitalization may be necessary. Following medical stabilization, inpatient treatment often includes a behaviorally oriented weight-restoration program with a gradual increase in caloric intake. Individual, group, and family psychotherapy (especially for younger patients) are commonly employed. For less severely ill patients, day hospital treatment or outpatient therapy is usually recommended. Available medication treatments appear to be of only limited benefit in achieving weight restoration, although they may be particularly helpful for symptoms of depression and anxiety. Following weight restoration, a significant number of individuals experience a relapse to low-weight episodes. Again, available medication

treatments appear to have only limited benefit in preventing relapse.

Bulimia Nervosa

Clinical Characteristics

The hallmark symptom of bulimia nervosa is the binge-eating episode. In the *DSM-IV*, a binge-eating episode is characterized as consumption of an abnormally large amount of food in a discrete episode, during which eating feels out of control. Although binge eating had previously been recognized in patients with anorexia nervosa, during the 1970s clinicians noted a syndrome of binge eating that occurred in individuals in a normal weight range. Current diagnostic criteria for bulimia nervosa specify that binge-eating episodes, along with compensatory behaviors designed to avoid weight gain, occur on average at least twice a week over a period of at least 3 months. Additionally, body shape and weight play an excessively large role in influencing the individual's sense of self-worth. For some patients, compensatory behaviors include purging activities such as self-induced vomiting or laxative misuse, while for others, fasting or excessive exercise are the main weight control strategies.

Although nutritional abnormalities may result in medical symptoms in patients with bulimia nervosa, these are usually less marked than in anorexia nervosa. It has been recognized, however, that purging behaviors can contribute to significant alterations in blood levels of potassium and other electrolytes, potentially leading to serious heart arrhythmias. A range of other serious medical complications have also been reported. Abnormalities in neuroendocrine hormone levels are not infrequent in bulimia nervosa, being associated with such symptoms as abnormal menstrual cycle patterns in up to 50% of patients.

Follow-up studies ranging up to 10 years following initial assessment indicate that approximately 50–60% of patients with bulimia nervosa will have recovered from the disorder. Approximately one-third of individuals will have significant persisting symptoms, and 10–15% will still meet criteria for the disorder. It is not uncommon for individuals to have one or more relapses, particularly during the first 6 months following their initial recovery from symptoms.

Recently there have been a number of clinical studies of a syndrome provisionally identified as binge-eating disorder. Binge-eating disorder is similar to bulimia nervosa in that affected individuals have recurrent episodes of binge eating. In contrast to bulimia nervosa, individuals with binge-eating disorder do not

engage in recurrent inappropriate compensatory behaviors designed to counteract the caloric intake associated with binge eating. Thus, binge-eating disorder is commonly associated with obesity. Further research is needed to clarify the role of psychosocial environment, stressful life situations, and biological characteristics as risk factors for binge-eating disorder.

Epidemiology and Psychiatric Comorbidity

Similar to anorexia nervosa, the average age of onset of bulimia nervosa is approximately 18 years, with a female to male ratio of approximately 10:1. Based on current diagnostic criteria, bulimia nervosa is estimated to occur in 1–3% of young women.

Approximately 50% of patients with bulimia nervosa have an episode of major depression, either concurrent with, or at a time separate from, the eating disorder. Other relatively common comorbid psychiatric symptoms in patients with bulimia nervosa include substance-use disorders and anxiety disorders.

Overview of Treatment Approaches

Patients with bulimia nervosa are most often treated in an outpatient setting, although life-threatening psychiatric symptoms (such as active suicidal ideation) or medical complications may lead to hospitalization. Clinical psychotherapy trials have focused primarily on structured short-term therapies (e.g., cognitive behavioral treatment or interpersonal treatment) lasting 4–5 months. These treatments generally achieve greater than a 50% decrease in binge frequency, with a substantial number of patients achieving abstinence from binge episodes. Additionally, placebo-controlled trials have shown that a range of antidepressant medications can significantly decrease frequency of bulimic episodes for many individuals. Current treatment research includes comparative trials of stepped-care approaches such as an initial trial of self-help guided by a health-care provider, and studies of sequential treatments for individuals who do not show evidence of improvement as psychotherapy proceeds.

Etiological Influences in Eating Disorders

Psychosocial and Biological Risk Factors

Although clinical investigations have identified important leads, specific etiological factors responsible for anorexia nervosa and bulimia nervosa remain unknown. Cultural factors leading to increased emphasis on slender appearance are generally thought to

contribute to an increased prevalence of eating disorders. Stresses associated with school transitions and occupational demands have been suggested as possible contributors to the increased risk for the eating disorders among adolescents and young adults.

Clinical investigations have demonstrated increased ratings of perfectionistic character traits and depressive symptoms in anorexia nervosa. Research on risk factors suggests that perfectionism and low self-evaluation may be associated with the later development of anorexia nervosa.

The onset of bulimia nervosa is often preceded by extended periods of recurrent dieting occurring in the context of other psychosocial stressors. Other behavioral characteristics that have been identified in patients with bulimia nervosa include impulsivity and mood lability, and it is possible that these traits may contribute to the onset or perpetuation of symptoms in this disorder. Several psychological models of binge-eating behavior have been proposed. In one model, for example, an individual attempting to follow a reduced calorie diet may experience an abstinence violation effect following ingestion of modest amounts of snack foods, leading to a transient inclination to abandon dietary restraint altogether. Factors that may lead to dieting, such as parental or childhood obesity, have been identified as potential risk factors for the development of this disorder.

Psychobiological Models

Family studies have shown that there is an increased rate of eating disorders in first-degree relatives of individuals with anorexia nervosa and bulimia nervosa. Similarly, twin studies have shown a higher concordance for the eating disorders in monozygotic twins in comparison to dizygotic twins. These studies suggest that heritable biological characteristics contribute to the onset of the eating disorders, although the potential role of familial environmental factors must also be considered.

Biological factors are likely to contribute to one or more dimensions of symptomatology in eating disorders. Neurotransmitters, neuropeptides, and related neuromodulators have been shown to act in the central nervous system, particularly in the hypothalamus, to modulate hunger and satiety. Recent research has shown that peptides released from the gastrointestinal tract help to regulate eating behavior. For example, preclinical studies have shown that naturally occurring neurochemicals such as ghrelin can substantially increase food intake and body weight, while other neuropeptides such as cholecystokinin can decrease food intake. Thus, binge-eating episodes

characteristic of bulimia nervosa may reflect the interaction of stressful psychosocial events with abnormalities in the release of meal-regulating peptides such as ghrelin and cholecystokinin. There has also been recent interest in the metabolic signaling protein, leptin, which is produced in adipose tissue and released into the circulation. Dieting and weight loss is associated with a decrease in leptin levels. It has been hypothesized that a marked reduction in blood leptin levels may play a role in menstrual cycle abnormalities commonly associated with anorexia nervosa.

The neurotransmitter serotonin has been the focus of considerable research in patients with anorexia nervosa and bulimia nervosa. Laboratory studies have shown that patients with eating disorders often experience abnormal patterns of hunger and satiety over the course of a meal. Serotonin plays an important role in postingestive satiety, and appears to be important in regulation of mood and anxiety-related symptoms. Preliminary findings suggest that impaired function in central nervous system serotonergic pathways may contribute to binge eating and mood instability in bulimia nervosa. Dieting behaviors may tax the adaptive capacities of serotonergic pathways. Therapeutic effects of antidepressant medications in bulimia nervosa are thought to be related to their capacity to restore more normal signaling patterns in serotonergic pathways.

Recent studies have also explored whether abnormalities in metabolic signals related to energy metabolism contribute to symptoms in the eating disorders. Several studies have suggested that patients with bulimia nervosa may have a lower rate of energy utilization (measured as resting metabolic rate) than healthy individuals. Thus, a biological predisposition toward greater than average weight gain could lead to preoccupation with body weight and food intake in bulimia nervosa.

See Also the Following Articles

Adolescence; Diet and Stress, Non-Psychiatric; Familial Patterns of Stress; Nutrition; Obesity, Stress and; Waist-Hip Ratio.

Further Reading

- Agras, W. S., Walsh, T., Fairburn, C. G., et al. (2000). A multicenter comparison of cognitive-behavioral therapy and interpersonal psychotherapy for bulimia nervosa. *Archives of General Psychiatry* 57, 459–466.
- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders, 4th edn., text revision*. Washington, DC: American Psychiatric Association.
- American Psychiatric Association Workgroup on Eating Disorders (2006). Practice guideline for the treatment of patients with eating disorders. 3rd edn. *American Journal of Psychiatry* 163, 1–54.
- Eckert, E. D., Halmi, K. A., Marchi, P., et al. (1995). Ten-year follow-up of anorexia nervosa: clinical course and outcome. *Psychological Medicine* 25, 143–156.
- Fairburn, C. G., Cooper, Z., Doll, H. A., et al. (1999). Risk factors for anorexia nervosa: three integrated case-control comparisons. *Archives of General Psychiatry* 56, 468–476.
- Fairburn, C. G., Welch, S. L., Doll, H. A., et al. (1997). Risk factors for bulimia nervosa. A community-based case-control study. *Archives of General Psychiatry* 54, 509–517.
- Jimerson, D. C. and Wolfe, B. E. (2004). Neuropeptides in eating disorders. *CNS Spectrums* 9, 516–522.
- Jimerson, D. C., Wolfe, B. E. and Naab, S. (2006). Eating disorders. In: Coffey, C. E., Brumback, R. A., Rosenberg, D. R., et al. (eds.) *Pediatric neuropsychiatry*, pp. 307–320. Philadelphia: Lippincott, Williams & Wilkins.
- Keel, P. K., Mitchell, J. E., Miller, K. B., et al. (1999). Long-term outcome of bulimia nervosa. *Archives of General Psychiatry* 56, 63–69.
- Kendler, K. S., MacLean, C., Neale, M., et al. (1991). The genetic epidemiology of bulimia nervosa. *American Journal of Psychiatry* 148, 1627–1637.
- Kishi, T. and Elmquist, J. K. (2005). Body weight is regulated by the brain: a link between feeding and emotion. *Molecular Psychiatry* 10, 132–146.
- Stoving, R. K., Hangaard, J., Hansen-Nord, M., et al. (1999). A review of endocrine changes in anorexia nervosa. *Journal of Psychiatric Research* 33, 139–152.
- Sunday, S. R. and Halmi, K. A. (1996). Micro- and macroanalyses of patterns within a meal in anorexia and bulimia nervosa. *Appetite* 26, 21–36.
- Zhu, A. J. and Walsh, B. T. (2002). Pharmacologic treatment of eating disorders. *Canadian Journal of Psychiatry* 47, 227–234.

EBV (Epstein-Barr Virus) See: Herpesviruses; Immune System, Aging; Cytotoxic Lymphocytes; Antibody Response.

Eclampsia and Pre-Eclampsia

A Makrigiannakis and G Petsas

University of Crete, Heraklion, Greece

G P Chrousos

Athens University Medical School, Athens, Greece

© 2007 Elsevier Inc. All rights reserved.

Maternal and Perinatal Outcome

History

Epidemiology

Pathophysiology

Prediction of Pre-Eclampsia

Prevention of Pre-Eclampsia

Management of Pre-Eclampsia

Glossary

<i>Eclampsia</i>	Occurrence of seizures in a woman with pre-eclampsia that cannot be attributed to other causes.
<i>Invasive trophoblast</i>	The inner trophoblast lineage which differentiates into tumorlike cells that invade the lining of the pregnant uterus and the maternal uterine spiral arteries.
<i>Oxidative stress</i>	A disequilibrium between antioxidant defenses and production of reactive oxygen species in favor of the latter.
<i>Pre-eclampsia</i>	A pregnancy-specific syndrome that occurs after mid-gestation, defined by the <i>de novo</i> appearance of hypertension (systolic blood pressure of ≥ 140 mmHg or diastolic blood pressure of ≥ 90 mmHg), accompanied by new-onset proteinuria (defined as ≥ 300 mg per 24 h).
<i>Pregnancy-induced hypertension</i>	<i>De novo</i> hypertension arising after mid-gestation distinguished from pre-eclampsia by the absence of proteinuria.

Pre-eclampsia is a multisystem disorder of unknown etiology, which is unique to pregnancy. Women with pre-eclampsia usually develop hypertension and proteinuria, but the condition is also associated with abnormalities of the blood coagulation system, disturbed liver function, renal failure, and cerebral ischemia. The incidence of pre-eclampsia is 2–10%, depending on the population studied and definitions of pre-eclampsia. Eclampsia, the occurrence of one or more convulsions superimposed on the syndrome of pre-eclampsia, occurs less frequently. In the Western world, the reported incidence of eclampsia ranges from one pregnancy in 2000 to one in 3448.

Maternal and Perinatal Outcome

Pre-eclampsia is a major obstetric problem leading to substantial maternal and perinatal morbidity and mortality worldwide, especially in developing countries. In general, maternal and perinatal outcomes are usually favorable in women with mild pre-eclampsia developing beyond 36 weeks of gestation. By contrast, maternal and perinatal morbidities and mortalities are increased in women who develop the disorder before 33 weeks of gestation, in those with pre-existing medical disorders, and in those from developing countries. Eclampsia compared to pre-eclampsia carries a much higher risk of death and serious morbidity for the woman and her baby. In the UK, for example, one in 50 of the women who have eclampsia die.

History

Eclampsia was defined by Celsus in 100 AD as seizures during pregnancy that abated with delivery. For the ensuing 1700 years, eclampsia was considered a pregnancy-specific seizure disorder. It was not until the mid-1800s that the similarity of the edematous eclamptic women and the dropsic patients with Bright's disease (acute glomerulonephritis) stimulated clinicians to determine whether women with eclampsia, like individuals with Bright's disease, had protein in their urine. Protein was indeed present and 50 years later an association of increased blood pressure and eclampsia was recognized. It was soon evident that hypertension and proteinuria during pregnancy, even without seizures, identified a woman with the potential for a rapidly progressive life-threatening disorder and a fetus at increased risk of stillbirth.

Epidemiology

Several risk factors have been identified with increased risk of pre-eclampsia. These factors include a previous history of pre-eclampsia, nulliparity, maternal age over 40 years, family history of pre-eclampsia, limited sperm exposure, multiple pregnancies, obesity, and chronic medical conditions such as chronic hypertension, renal disease, insulin-dependent diabetes (IDDM), autoimmune disease, and antiphospholipid syndrome. Paradoxically, smoking during pregnancy has been associated with a reduced risk of pre-eclampsia.

Studies conducted to determine the relation between job strain and hypertensive complications during pregnancy showed that work-related psychosocial stress increased the risk of pre-eclampsia. Indeed stressful job characteristics did show associations with pregnancy-induced hypertension. In particular, gestational hypertension was associated with low decision latitude and low job complexity among women in lower-status jobs and with job pressures/low control among women in higher-status jobs.

Pathophysiology

Reduced Placental Perfusion, Oxidative Stress, and Pre-Eclampsia

Although the causes of pre-eclampsia remain still unknown, it is generally agreed that pre-eclampsia arises from the placenta and/or the maternal response to placentation. The clinical syndrome is characterized by secondary systemic circulatory disturbances that can be ascribed to generalized maternal endothelial dysfunction. Pre-eclampsia is appropriately divided into two stages: alterations in placental perfusion and the maternal syndrome. In placental pre-eclampsia, the problem arises from a placenta that is under hypoxic conditions and oxidative stress. Once defective placentation is established, reduced perfusion appears to interact with maternal factors to result in the maternal syndrome. These factors are posited to be genetic, behavioral, or environmental.

The placenta seems to be the key component of pregnancy that leads to pre-eclampsia. During the first half of normal human pregnancy, uteroplacental arteries undergo a series of pregnancy-specific changes that include the replacement of endothelial and media smooth muscle cells by invasive trophoblast cells, loss of elasticity, dilatation to incontractile tubes, and loss of vasomotor control, which allows a substantial increase in blood supply to the growing fetus. In pre-eclampsia, spiral arteries undergo superficial or no remodeling and this results in reduced placental perfusion.

A principal question is how reduced perfusion of the placenta can result in the maternal syndrome. Oxidative stress has been proposed as the link between the two phases of pre-eclampsia. The oxidative stress hypothesis proposes that hypoxia at the fetal-maternal interface results in the generation of free radicals that may lead to oxidative stress by the following potential mechanism: hypoxia stimulates xanthine oxidase, an important source of superoxide generation. As a result, the placenta releases what can be described as trophoblastic debris into the maternal circulation. This comprises syncytiotrophoblast

membrane microparticles, cytokeratin fragments, soluble ribonucleic acid (RNA) and deoxyribonucleic acid (DNA) of fetal origin and even cytotrophoblast cells. Syncytiotrophoblast microvesicles, normally present in the circulation in pregnancy, are increased in pre-eclampsia and have been directly linked to activation of maternal neutrophils, which in turn may contribute to endothelial cell activation.

In pre-eclampsia, the hypoxia/reperfusion injury leads also to increased expression of nicotinamide adenine dinucleotide phosphate (NAD(P)H) oxidase and hence increased superoxide generation in placental tissue. Potential additional stimuli to activation of NAD(P)H oxidase in pre-eclampsia include raised feto-placental vascular shear stress, elevation of maternal plasma cytokine concentrations, and enhanced angiotensin II sensitivity.

Abundant evidence of oxidative stress in placenta, tissues, and blood of women with pre-eclampsia support the oxidative stress hypothesis. These oxidative damage biomarkers include higher placental levels of markers for lipid peroxidation such as the F_2 -isoprostanes and malondialdehyde and for non-lipid markers, including increased nitrotyrosine residues in fetal blood vessels. Supportive of a role for oxidative stress in the maternal circulation is the maternal plasma elevations of lipid peroxidation products, including malondialdehyde (MDA), conjugated dienes, F_2 -isoprostanes (usually 8-epi-prostaglandin $F_{2\alpha}$), and antibodies against oxidatively modified low-density lipoprotein (LDL). Besides the elevation of lipid peroxidation markers, another indication of the presence of oxidative stress is the decreased antioxidant capacity in the maternal circulation. Indeed, women with pre-eclampsia have lower plasma concentrations of glutathione, a major intracellular water-soluble antioxidant.

Stress and Pre-Eclampsia

The hypothalamic-pituitary-adrenal (HPA) axis, along with the arousal and sympathetic nervous systems (SNS), constitute the stress system. Activation of the stress system leads to behavioral and peripheral changes that improve the ability of the organism to adjust homeostasis and increase its chances for survival. The principal regulators of the HPA axis are corticotropin-releasing hormone (CRH) and arginine-vasopressin (AVP), both produced by parvocellular neurons of the paraventricular nucleus of the hypothalamus and secreted into the hypophyseal portal system. CRH and AVP synergistically stimulate pituitary adrenocorticotrophic hormone (ACTH) secretion and subsequently cortisol secretion by the adrenal cortex.

CRH and its receptors have been identified in several female reproductive organs, including the ovaries, endometrial glands, decidualized endometrial stroma, and placenta. Placental CRH is produced by the syncytiotrophoblast, placental decidua, and fetal membranes. In pregnant women, the placenta is the main source of circulating CRH, as supported by undetectable levels of circulating CRH levels in non-pregnant woman. The biologic activity of placental CRH is attenuated by the presence of a circulating CRH-binding protein (CRH-BP), a 37-kD protein of 322 amino acids that binds circulating CRH and modulates CRH actions on pregnant target tissues during pregnancy. To optimize placental perfusion during pregnancy, the umbilical-placental vessels are normally maintained in a state of physiological dilatation by the action of locally produced vasodilators. CRH is a potent dilator of placental resistance vessels, acting via production of nitric oxide (NO). Maternal plasma CRH concentrations are significantly elevated in complicated pregnancies.

The mean umbilical cord plasma CRH in pre-eclamptic pregnancies is significantly higher than that from normotensive pregnancies. This is also reflected by maternal plasma CRH levels that are significantly higher in hypertensive women than those with uncomplicated pregnancies. Moreover, in women with pregnancy-induced hypertension, CRH-BP levels are significantly lower in patients who progressed to pre-eclampsia. CRH-BP may be involved in the paracrine regulation of vasodilator CRH effects in human placenta. In this regard, the reduced secretion of CRH-BP may enhance and amplify the effects of CRH by increasing its biologically active free fraction. Taken together these findings suggest that in pre-eclampsia or eclampsia the placenta takes part in a stress syndrome by releasing CRH, which may help improve uterine perfusion thereby protecting the fetus from anoxia and starvation. Indicative of a stress-responsive compensatory mechanism in the human placenta is the fact that the concentration of CRH in the fetal circulation is significantly increased in pregnancies complicated by abnormal umbilical artery flow-velocity waveforms.

In pregnancies complicated by pre-eclampsia, the plasma levels of CRH are elevated, together with a concomitant reduction in CRH receptor type 1 α (CRHR1 α) and type 2 (CRHR2) expression. The dampening of CRH-induced vasodilatation in pre-eclamptic placentas could be attributable to loss of CRHRs. Although in normal placenta there is a balance between the vasodilatory action of CRH and the opposing actions of vasopressor agents, such as angiotensin II, in pre-eclampsia upregulation of

angiotensin II receptors (AT₁R) and formation of heterodimers with bradykinin B₂ receptors (B₂R) alter this balance and further potentiate the vasoconstriction caused by angiotensin II. Enhanced angiotensin II activity induces secretion of placental CRH, which, in turn, can downregulate its own receptors leading to abnormal vascular resistance and the clinical sequelae of pre-eclampsia.

One of the signaling pathways mediating CRH actions during pregnancy appears to be the nitric oxide (NO)/cyclic guanosine monophosphate (cGMP) pathway. This signaling pathway appears to play an important role in vascular adaptation and placental physiology by mediating the vasodilatory effects of agonists in resistance vessels, which helps in the maintenance of low vascular resistance in the fetoplacental circulation. In pre-eclampsia the placental CRH-R down-regulation and the associated inability of CRH and CRH-related peptides to stimulate the NO/cGMP pathway might result in the disturbance of the balance controlling vascular tone towards vasoconstriction. In addition to CRH, other hormones with vasodilatory actions involved in the stress response, such as ACTH and cortisol, are increased in the fetuses of pre-eclamptic pregnancies, underscoring the role of the placental stress syndrome in the disease.

The role of CRH in the regulation of pituitary adrenocorticotropin secretion and the stress response is well established. In recent years, association of abnormally high placental expression of CRH with pregnancy complications, such as pre-eclampsia and intrauterine growth restriction (IUGR), suggest that CRH output from the placenta is specifically increased in response to fetal stress. Placental CRH activates a number of potentially useful adaptive responses for the stressed fetus. These include maximization of placental blood flow, acceleration of fetal organ development, and the early initiation of labor when fetal survival is threatened.

Besides the activation of the HPA axis, stress is associated with the activation of the SNS and hence with elevated serum levels of catecholamines. Enhanced release of catecholamines is postulated to aggravate the already existing vasoconstriction in pre-eclampsia, whereas, in normal pregnancy, loss of vasomotor control keeps vessels relatively insensitive to vasoactive substances.

Prediction of Pre-Eclampsia

A variety of biochemical markers, based primarily on the above rationale implicated in the pathophysiology of pre-eclampsia have been proposed to predict the development of pre-eclampsia later in pregnancy.

Only markers related to placental stress syndrome are discussed in this section.

Although there is an inverse relation between reduced plasma CRH-BP levels and increased CRH levels in the maternal circulation of patients with pregnancy-induced hypertension, the measurement of these polypeptides in maternal plasma does not predict the development of pre-eclampsia because these hormonal changes do not occur before the onset of disease. Measurements, however, of CRH and CRH-BP in pregnant women who are at-risk for pre-eclampsia may add significant prognostic information for predicting pre-eclampsia, as the probability of the disease was shown to be remarkably higher among women with positive than in women with negative double-hormone tests.

In women with pre-eclampsia, the combination of elevated lipid peroxidation markers and decreased antioxidant capacity has raised the possibility that markers of oxidative stress might prove useful in the prediction of the disease. However, data for the reliability of these markers in suggesting pre-eclampsia have been inconsistent because of the difficulties and variability in the methods of measurement. As pre-eclampsia is of multifactorial origin, the determination of a single biochemical marker is unlikely to provide adequate predictive power.

At present, the best predictive test involves assessment of the velocity of uterine artery blood flow in the second trimester, however this test is still not recommended for routine screening of pregnant women for pre-eclampsia.

Prevention of Pre-Eclampsia

During the past decade, several trials reported the use of various methods to reduce the rate or severity of pre-eclampsia. Although prevention trials have been disappointing to date, the evidence for oxidative stress in pre-eclamptic women indicates that there may be a role for antioxidant vitamins C and E for prophylaxis. Large multicenter trials are now underway and will determine whether antioxidant prophylaxis may be used routinely in the prevention of pre-eclampsia.

Management of Pre-Eclampsia

The most important factor in the management of pre-eclampsia is adequate and proper prenatal care. After diagnosis, subsequent treatment will depend on the results of initial maternal and fetal assessment. Mother always takes the first priority in the management of pre-eclampsia and thus delivery remains the

ultimate treatment. When possible, vaginal delivery is preferable to avoid the added physiologic stressors of a cesarean section. Since delivery is not always the best choice for a very premature fetus, the decision between delivery and expectant management depends on the severity of maternal condition, the fetal status, and fetal gestational age at the time of assessment.

Further Reading

- Cunningham, F. G., Gant, N. F., Leveno, K. J., et al. (2003). Hypertensive disorders in pregnancy. In: Seils, A., Noujaim, S. R. & Davis, K. (eds.) *Williams obstetrics*, (21st edn, pp. 567–618). New York: McGraw-Hill Companies.
- Duckitt, K. and Harrington, D. (2005). Risk factors for pre-eclampsia at antenatal booking: systematic review of controlled studies. *British Medical Journal* 330, 565–571.
- Duley, L. (2003). Pre-eclampsia and the hypertensive disorders of pregnancy. *British Medical Bulletin* 67, 161–176.
- Florio, P., Imperatore, A., Sanseverino, F., et al. (2004). The measurement of maternal plasma corticotropin releasing factor (CRF) and CRF-binding protein improves the early prediction of preeclampsia. *Journal of Clinical Endocrinology and Metabolism* 89, 4673–4677.
- Kalantaridou, S. N., Makriganakis, A., Zoumakis, E., et al. (2004). Stress and the female reproductive system. *Journal of Reproductive Immunology* 62, 61–68.
- Karteris, E., Goumenou, A., Koumantakis, E., et al. (2003). Reduced expression of corticotropin-releasing hormone receptor type-1 α in human preeclamptic and growth-restricted placentas. *Journal of Clinical Endocrinology and Metabolism* 88, 363–370.
- Karteris, E., Vatish, M., Hillhouse, E. W., et al. (2005). Preeclampsia is associated with impaired regulation of the placental nitric-oxide-cyclic guanosine monophosphate pathway by corticotropin-releasing hormone (CRH) and CRH-related peptides. *Journal of Clinical Endocrinology and Metabolism* 90, 3680–3687.
- Landsbergis, P. A. and Hatch, M. C. (1996). Psychosocial work stress and pregnancy-induced hypertension. *Epidemiology* 7, 346–51.
- Paarlberg, K. M., Vingerhoets, A. J., Passchier, J., et al. (1995). Psychosocial factors and pregnancy outcome: a review with emphasis on methodological issues. *Journal of Psychosomatic Research* 39, 563–595.
- Petraglia, F., Florio, P., Benedetto, C., et al. (1996). High levels of corticotropin-releasing factor (CRF) are inversely correlated with low levels of maternal CRF-binding protein in pregnant women with pregnancy-induced hypertension. *Journal of Clinical Endocrinology and Metabolism* 81, 852–856.
- Raijmakers, M. T. M., Dechend, R. and Poston, L. (2004). Oxidative stress and preeclampsia: rationale for antioxidant clinical trials. *Hypertension* 44, 374–380.

- Redman, C. W. and Sargent, I. L. (2005). Latest advances in understanding preeclampsia. *Science* 308, 1592–1594.
- Roberts, J. M., Pearson, G., Cutler, J., et al. (2003). Summary of the NHLBI Working Group on research on hypertension during pregnancy. *Hypertension* 41, 437–445.

- Sibai, B. M. (2005). Diagnosis, prevention, and management of eclampsia. *Obstetrics and Gynecology* 105, 402–410.
- Sibai, B., Dekker, G. and Kupferminc, M. (2005). Pre-eclampsia. *Lancet* 365, 785–799.

Economic Factors and Stress

R A Catalano

University of California, Berkeley, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R A Catalano, volume 2, pp 9–14, © 2000, Elsevier Inc.

Introduction: The Economy as a Population Stressor

The Economy's Effect on Hazard Avoidance

The Economy's Effect on Coping Assets

The Economy's Effect on Tolerance for Coping

Socioeconomic Status and Stress

Policy Implications and Conclusions

Glossary

<i>Discount rate of pain</i>	A measurement of the difference between fear of pain experienced now and of that experienced in the future.
<i>Expected value of pain</i>	The product of multiplying the likelihood of experiencing a painful event by the amount pain it would inflict.
<i>Reservation wage</i>	The wage at which a worker with chronic illness will apply for disability benefits rather than compete for a job.
<i>Social support</i>	Tangible and intangible help from family, friends, and social networks to cope with stressors.

Introduction: The Economy as a Population Stressor

We intuitively separate the stressors associated with physiological and behavioral disorders into several groups. These groups include undesirable as well as intuitively desirable job and financial events. As might be expected, the incidence of undesirable job and financial stressors increases during periods of economic contraction. The experience of undesirable job and financial events by a principal wage earner also increases the risk of disorder for his or her spouse and family. Undesirable job and financial experiences,

moreover, increase the likelihood of subsequent stressful nonjob, nonfinancial experiences for the wage earner and his or her family. The contraction of a regional economy can, in other words, increase the experience of undesirable job and financial events that, in turn, increase the risk of experiencing yet other undesirable experiences not intuitively related to the economy. These undesirable events raise the risk of disorder not only for those who experience them, but also for spouses and other members of the family.

Researchers have focused on job loss more than any other economic stressor. While disagreement remains over the virulence of job loss as a pathogen, there is agreement that forced job loss increases the risk of depressed mood, alcohol abuse, and antisocial behavior. More controversy arises from claims of a connection with somatic illness, but job loss reportedly increases the risk of stress-related illnesses, including those associated with compromised endocrine and immune responses.

Persons exposed to undesirable job and financial events early in life may exhibit elevated risk of succumbing to future stressors, whether those are economic in nature or not. Research on those who experienced the Great Depression of the 1930s, for example, suggests that economic stressors can affect behavior well into the future.

Classic and recent theoretical literature concerned with the economy as a source of stressors posits that economies expanding more quickly than the rate to which a population has habituated should be pathogenic. This work argues that movements away from the expected value of macroeconomic performance, regardless of direction, should increase the incidence of stress-related illness. Individual-level research has not found desirable job and financial events to be as virulent or as contagious as undesirable events, but reports of ecological associations between rapid economic growth and the incidence of trauma and alcohol-related pathology appear in the literature. Fear of job loss, moreover, appears to reduce risk-taking behavior such as alcohol use and antisocial

behavior. The pure income effects of changing economies could also affect the incidence of pathology in that more income makes risk taking more likely, while less income should suppress risk taking that requires outlay.

The known and suspected health effects of economic expansion raise an important issue for those who would estimate the health and behavioral cost of economic perturbations. Such estimates will be of the net effect of change and therefore require more empirical research than we now have into pathology induced by economic growth.

The Economy's Effect on Hazard Avoidance

The classic and contemporary stress literature connects environmental stressors to illness through the presumed effects of stressors on the endocrine and immune systems. Stressors, however, may affect health through other avenues, such as the effects of stressors on hazard avoidance.

We have a fixed capacity to assess, manage, or reduce hazards in our physical and social environments. We budget this capacity in that not all hazards receive equal attention. Candidates for less attention include hazards with the least potential for pain. Potential for pain is the product of the expected value of pain and our discount rate. The expected value, in turn, is the product of the likelihood of succumbing to the hazard and the pain that doing so would inflict. The discount rate gauges the difference in fear of pain experienced now compared to that experienced in the future. We most likely neglect hazards that promise relatively little pain, have a low probability of occurring, and occur farthest in the future.

If a contracting economy confronts us with new and immediate hazards, those already low on our list for attention may fall off the list all together. Screening for the early signs of treatable disease or compliance with treatment regimens, for example, may decline among persons attending to hazards posed by lost income.

Attending to new hazards at the expense of old might not change the incidence of illness in a population if hazards came into and left our lives for reasons peculiar to each of us. Some of us would be experiencing new hazards while others were losing old ones, leaving the net effect in the population at or near zero. Incidence would, however, vary over time if the environments we share became more or less hazardous. We all share the constantly shifting economic environment.

The early detection of breast cancer provides an example of how the economic environment may affect the incidence of serious illness by causing shifts in attention across hazards. Ninety-three percent of

women discovered with local breast cancer survive 5 or more years, but only 18% of those discovered with remote tumors do so. It appears that living in a contracting economy reduces the likelihood that women will discover tumors in the local stage. Women apparently seek less screening when coping with the sequelae of economic contraction. They do not, in other words, have sufficient capacity to attend to both the new hazards inflicted upon them by economic contraction and the old hazards farther down the list of priorities.

As noted earlier, changing economies have income effects such that expansion enables individuals to consume more while contraction tends to reduce consumption. Some individuals may use added income to pursue activities that expose them to more hazards than would otherwise be the case. Indeed, ecological associations between rapid economic growth and the incidence of trauma may reflect the risk taking enabled by economic expansion.

The Economy's Effect on Coping Assets

The stress literature posits, without great controversy, that coping assets mediate an individual's response to stressors. These assets can be genetic or acquired through interactions with the environment. The latter can be further separated into those that are unintentionally acquired (e.g., immunities induced by naturally occurring exposures to infectious pathogens) and those sought out (e.g., vaccines). Intentionally acquired coping assets include those purchased with money and those acquired through social arrangements analogous to mutual aid societies or insurance pools. While some have speculated regarding the effect of economic forces on unintentionally acquired, and even genetically endowed, coping assets, the empirical work focuses on those purchased with money or obtained through social arrangements.

Coping assets purchased with money include goods and services explicitly marketed as means to bolster one's capacity to deal with new or chronic stressors. These assets include professional medical care, leisure activities, organized exercise activity, and dietary supplements. Other goods and services can also help persons cope with stressors. Advertisers may not often cast nutritious food, decent housing, and entertainment, for example, as stress buffers, but much of the value of these products in the market may well arise from this function. The performance of the economy obviously affects our ability to acquire these resources because it affects how much money we have to spend. Money, in effect, is a generalized coping asset.

The performance of an economy also affects the availability of coping assets acquired outside the market. The literature pays much attention to social

support and social capital as mediators of the stress response. These terms refer to all sorts of tangible and intangible resources obtained by stressed persons from their families, social networks, and public and private institutions.

The insurance pool metaphor conveys the effect of the economy on social support and capital. Social networks can be thought of as informal insurance pools to which participants not coping with stressors contribute surplus coping assets, and from which participants coping with stressors draw assets. As with all insurance pools, the arrangement works only if demands do not exceed the pooled resources. Participants in such pools often underestimate how many resources a pool needs because we intuitively make the actuarial assumption that the incidence of stressors remains relatively constant although those who suffer them may vary. If, however, an ambient shock stresses many persons in the pool, the demand for resources can be unexpectedly high and deplete the pool. A contracting economy acts as an ambient stressor that causes an unexpectedly large number of individuals and their families to suffer stressful losses and to resort to social networks for coping resources. Other members of the network who fear such a loss withdraw contributions to the pool, believing they themselves will need them. The incidence of stress-related illness therefore rises because social support cannot be gotten from an actuarially insufficient pool. Persons dealing with chronic stressors or stressful events unrelated to the economy also exhibit elevated incidence of illness because they cannot retrieve their usual support from a depleted pool. The effect of a contracting economy on the incidence of illness in populations embedded in contracting economies thereby grows beyond that expected from studies of individuals who, for example, are forced out of work.

Recent years have seen an increasing interest in the role of economic inequality on the coping process. Social reformers have traditionally drawn an analogy between income disparity and ambient pathogens such as air pollution. This disparity presumably emits a morally numbing pathogen that makes us not only less helpful to the needy, but also less willing to contribute to social support pools that include socioeconomic peers.

Contemporary reformers claim a causal chain in which widening income disparity erodes social cohesion. The loss of social cohesion makes it more difficult for everyone, not just the poor, to cope with biological and behavioral hazards. Failure to cope then manifests in illness and death itself.

The reformers' argument, although intended primarily to influence the policy debate over the

distribution of wealth, has implications for the epidemiology of stress-related illness. If income disparity weakens social cohesion, it will likely reduce the effectiveness of social support networks. Less effective social support increases the likelihood that economic and other stressors will yield illness.

The Economy's Effect on Tolerance for Coping

The economy apparently affects the diagnosis of illness by affecting community tolerance for coping. Society applies the label illness to physical and behavioral characteristics that reduce a person's ability to perform day-to-day functions. The tolerance of a community for changes in performance, therefore, becomes an important determinant of whom it judges ill.

Sociologists and psychologists report that the economy affects tolerance for deviance from performance standards. Ecological psychologists have, for example, noted that understaffed organizations (i.e., those with a high ratio of roles, or functions, to participants) tolerate poor performance by members more than overstaffed organizations. Maintaining an understaffed organization supposedly requires members to be tolerant of the shortcomings, real or imagined, of the relatively few persons available to perform needed functions. Persons in understaffed organizations whose coping strategies include physiological and behavioral adaptations that could be labeled as illness (e.g., alcohol abuse) will go undiagnosed because the label may, by rule, disqualify them from positions for which less deviant candidates cannot be found.

Overstaffed organizations, on the other hand, can choose from among many candidates for positions and can set relatively high standards for acceptable functioning. Persons adapting to stressful events may not function as well as other candidates. This performance deficit makes them the least likely to find positions. Accepting the label ill gives them a socially acceptable explanation for their lack of a position.

Economists have developed the theory of reservation wage to describe the individual choice to seek work or income transfers. The theory implies a continuum of physical and behavioral fitness among those who compete for jobs. Society decides where on this continuum it will make income transfers. If too high on the continuum, otherwise productive workers will seek transfers rather than compete for work. If too low, persons with illness will be forced to compete for jobs and become even sicker.

Wherever on the continuum political institutions draw the income transfer line, workers near it have the option of competing for jobs or seeking benefits. Economists refer to the wage at which someone no

longer competes and applies for income transfers as his or her reservation wage.

Sociologists have applied the staffing and reservation wage concepts to the community at large. A contracting economy implies relatively many overstaffed organizations, while an expanding economy implies the opposite. The fraction of the labor continuum that can exceed its reservation wage by holding a job goes up with understaffing and down with overstaffing. The number of persons seeking the label ill thereby changes with the ability of the economy to provide jobs with relatively high wages. The incidence of diagnosed illness moves inversely over time with the economy, not only because the labor market affects true incidence but also because it affects community tolerance for performance deficits.

Socioeconomic Status and Stress

The most frequently replicated findings in epidemiology include the inverse association between measures of socioeconomic status and health. Considerable controversy remains over the cause of this association. As might be expected, some argue that the relationship results from the fact that sick persons cannot compete well in the labor market with healthy persons. The argument that illness induces relative poverty justifies, in part, the extensive array of disability compensation programs in the industrial world. Ill persons unable to compete for jobs presumably need income transfers to avoid sinking into poverty.

The rival argument assumes that market economies create a distribution of income in which those with relatively little inherited wealth remain poor because gaining highly compensated skills and making investments requires wealth. Being relatively poor, moreover, means that a person can purchase relatively few of the goods and services that shelter us against stressors (e.g., decent housing in safe neighborhoods) or bolster our coping capacity (e.g., leisure, wholesome food). Being poor, therefore, increases the likelihood of becoming ill.

The relative contribution of each mechanism to the statistical association between socioeconomic status and health remains controversial. Researchers who study this issue also contribute to the economic stress literature because they document that persons most likely to avoid economic stressors also have coping assets and enjoy high community tolerance for their coping. We know from empirical study that persons with wealth suffer fewer stressful events during economic contraction than those in the middle and lower classes. It does not require empirical study to conclude that persons with more money can purchase more coping resources than can those with less

money. It also appears that physical and behavioral deviance among the poor more likely leads to being labeled ill than similar deviance among the rich. Being relatively poor, therefore, makes a person particularly vulnerable to the stressors of economic contraction.

Policy Implications and Conclusions

Any discussion of economic stressors and their effects must at least allude to economic policy because such stressors result mostly from public and private decisions regulated by political institutions. Regulatory institutions intervene in these decisions only when costs appear to exceed benefits. While researchers may study the health effects of economic perturbations to satisfy their curiosity, they typically claim that the work has applied value in that it could help regulators better account the social costs of public and private decisions. The claim implies that public policy could, indeed, should reduce the frequency or virulence of economic stressors such as job loss.

Much controversy has arisen in Europe and North America over whether regulators have gone too far or not far enough in their attempts to reduce the stress of economic change. Policies now in place appear to assume that we previously went too far and discouraged private investment by depressing the return to capital. Defenders of the current policy acknowledge that increasing the return to private investment induces economic restructuring and its attendant pain in our personal and communal experience. They predict, however, that lowering the return to investors would drive capital to less regulated and less taxed economies. This would supposedly displace more labor in the developed world than restructuring, and inflict much unregulated stress and untreated illness in developing countries.

The assumptions underlying current policy remain controversial. The epidemiology of economic stress, therefore, will likely remain an important issue in public health as well as in the debate over economic policy.

In summary, the performance of an economy can affect the health of the population it supports. Economic contraction increases the number of persons coping with undesirable job and financial events. These events increase the risk of experiencing other stressors not intuitively connected to the economy. The adverse effects of these stressors spread to family and friends.

Rapid economic growth also induces adaptations that should, according to classic theory, increase the incidence of stress-related illness. Research on work-related trauma and alcohol consumption, for example, supports this connection.

A contracting economy affects our capacity to cope with stressors. Persons who lose income cannot purchase as many coping resources as they once did. This reduces their ability to deal not only with new economic stressors but also with chronic stressors previously buffered with purchased coping resources. The tangible and intangible coping resources gotten from social networks are also more difficult to obtain when the economy contracts. This is true because there are fewer surplus resources to contribute to the common pool at the same time that there are more demands upon it.

Economic stressors can also increase the incidence of illnesses not typically thought to be stress related, because coping with such stressors can leave us with fewer resources to avoid risk factors for, or detect early signs of, illnesses unrelated to stress. The economy can, moreover, affect the tolerance of society for persons coping with stressors. We know that overstaffed communities, as indicated by high unemployment rates, reduce competition for scarce jobs by increasing the diagnosis of disability.

Economic contraction inevitably increases the number of persons who are poor. Being poor is a risk factor for stressors of all sorts and, by definition, means that access to coping resources is relatively constrained.

See Also the Following Articles

Coping Skills; Health and Socioeconomic Status; Industrialized Societies; Job Insecurity: The Health Effects of a Psychosocial Work Stressor; Social Status and Stress.

Further Reading

- Brenner, M. (1973). *Mental illness and the economy*. Cambridge, MA: Harvard University Press.
- Catalano, R. (1979). *Health, behavior and the community: an ecological perspective*. New York: Pergamon Press.
- Catalano, R. (1991). The health effects of economic insecurity. *American Journal of Public Health* 81, 1148–1152.
- Catalano, R., Novaco, R. and McConnell, W. (1997). A model of the net effect of job loss on violence. *Journal of Personality and Social Psychology* 72, 1440–1447.
- Catalano, R., Satariano, W. and Ciemins, E. (2003). Unemployment and the detection of early stage breast tumors among African Americans and non-Hispanic whites. *Annals of Epidemiology* 13, 8–16.
- Catalano, R., Bruckner, T., Anderson, B. and Gould, J. (2005). Fetal death sex ratios: a test of the economic stress hypothesis. *International Journal of Epidemiology* 34, 944–948.
- Fryer, D. (1998). Special issue on mental health consequences of economic insecurity, relative poverty, and social exclusion: community psychological perspectives on recession. *Journal of Community and Applied Social Psychology* 8.
- Keiselbach, T. (1997). Special issue on job loss, unemployment, and social injustices. *Social Justice Research* 10.
- Neumayer, E. (2004). Recessions lower (some) mortality rates: evidence from Germany. *Social Science and Medicine* 58(6), 1037–1047.
- Ruhm, C. J. (2004). Healthy living in hard times. *Journal of Health Economics* 24(2), 341–363.
- Warr, P. (1987). *Work, unemployment, and mental health*. Fair Lawn, NY: Oxford University Press.
- Wilkinson, R. (1996). *Unhealthy societies: the afflictions of inequality*. London: Routledge.

Education Levels and Stress

J Mirowsky and C E Ross

University of Texas at Austin, Austin, TX, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A V Ranchor and R Sanderman, volume 2, pp 15–18, © 2000, Elsevier Inc.

What Education Levels Indicate

Lower Exposure to Stressful Situations

Better Response in Stressful Situations

Higher Socioeconomic Status

Learned Effectiveness and the Sense of Control

Life-Course Compounding
Signs of Lower Stress

Glossary

<i>Acute stressor</i>	An undesirable and uncontrollable event or transition requiring personal adaptation.
<i>Chronic stressor</i>	A situation characterized by prolonged discrepancy between goals and means.
<i>Human capital</i>	Productive capacity developed, embodied, and stored in humans themselves.
<i>Learned effectiveness</i>	The ability to produce desired outcomes gained through study, practice, and experience.

<i>Resource substitution</i>	Using one thing in place of another, and finding ways to achieve ends with whatever materials, relationships, and circumstances present themselves.
<i>Sense of control</i>	A learned and generalized sense of directing one's own life, ranging by degrees from fatalism and a deep sense of helplessness to instrumentalism and a firm sense of mastery.

What Education Levels Indicate

Formal education serves three functions in modern societies: (1) developing abilities through progressive instruction and training, (2) grading individual levels of development and gating advancement, and (3) regulating access to occupations and jobs. Developed nations all measure attainment by years and degrees, in a manner increasingly comparable across nations. The years measure progressive levels, each attainable by most students within one school year of effort. The degrees certify the completion of specific multi-year programs. Education level refers both to years and degrees; it indicates the level of ability developed, the exposure to progression and selection, and the opportunities regulated as a consequence. Together these influence an individual's exposure and response to stressors throughout adulthood.

Lower Exposure to Stressful Situations

The abilities and opportunities provided by higher levels of education reduce exposure to stressful situations. Researchers generally categorize stressors as events (acute stressors) or conditions (chronic stressors.) Life-change events are transitions that require personal adaptation. They can be desirable (e.g., getting married) or undesirable (e.g., becoming widowed). They also can be controllable (e.g., quitting a job) or uncontrollable (e.g., getting fired). The more undesirable and uncontrollable the events, the more distress they produce. Higher levels of education shift the balance of events. Controllable and desirable ones become more likely; uncontrollable and undesirable ones become less likely. An event such as a job promotion produces some stress as the individual adapts to the new demands, but the positive implications moderate the stress and make it stimulating rather than disturbing. In addition, it marks a transition to new circumstances with a more favorable balance of events and greater social and economic resources. An event such as getting fired or laid off demands as much or more adaptation, but also has the demoralizing implications of inadequacy and impending hard times.

In addition, it often marks a transition into prolonged unemployment, economic hardship, family strife, lower social standing, and diminished opportunity. Those situations are stressful, and they shift the likely balance of subsequent events unfavorably.

Higher levels of education help individuals avoid chronic stressors, characterized by a prolonged discrepancy between goals and means. Chronic stressors include prolonged unemployment; economic hardship and poverty; work that is repetitious, closely supervised, or boring; neighborhood disorder and decay; unsupportive or conflict-ridden relationships; parenting with no partner or an unhelpful one; solitary caring for an aging parent or other impaired dependent; and conflicting demands of work and family.

Chronic stressors are especially harmful to emotional and physical health. Their persistence erodes emotional, physical, and social resources and discourages corrective action. For example, when people are unemployed, they need or want paid work but are not able to find an acceptable position. Being unemployed for months can use up personal and household savings. The shortage of money can lead to economic hardship, which is not being able to buy things the household needs, such as food, clothes, transportation, medicine, and medical care, and not being able to pay the rent or other bills. Repeated failures to find a job become increasingly demoralizing and discouraging. Each failure makes people's emotional need for success greater while making success on the next try seem less likely. The need for comfort and relief can encourage people to overeat sugary and fatty foods, drink heavily, abuse drugs, engage in sexual indiscretions, or escape obsessively into television, movies, novels, games, and the like. Meanwhile, the economic shortfalls get larger with time.

Supportive relationships can ease the individuals' emotional and economic strains, but they do it by spreading the burden to others. The more intense and persistent the hardship, the greater the stain put on relationships. Tensions mount, expressed in irritation, blaming, anger, and strife that may include violence. The others become unsupportive and perhaps even hostile. Relationships break up.

The higher the level of education, the less likely individuals will get into situations of chronic stress, the better they cope if in them, and the quicker they get out of or correct the situation.

Better Response in Stressful Situations

Individuals with higher levels of education respond better in stressful situations. They solve problems

better, control their neuroendocrine stress responses better, and use their mobilized energy more constructively. People with higher levels of education take a more pragmatic approach to problems, looking for solutions rather than just getting anxious, angry, or depressed. They also have greater cognitive flexibility, which is the ability to imagine and consider a variety of possible solutions or viewpoints. They are better at communicating and negotiating, have better sources of information, and have more skill at judging and applying information. These abilities make individuals more effective and also more self-assured, which reduces neuroendocrine reactivity.

Individuals with higher levels of education also make better use of the energy mobilized in the stress response. The mental and physical activity directed at problem solving helps use the energy constructively. In addition, adults with higher levels of education are more likely to have habits designed to channel the excess energy mobilized by stress constructively, such as jogging, bicycling, walking, gardening, weight lifting, swimming, playing active games such as tennis, or taking fitness classes such as aerobics or yoga.

These better events, situations, and responses associated with higher levels of education result from its two primary consequences: socioeconomic status and learned effectiveness.

Higher Socioeconomic Status

Social scientists view education level as a major element of achieved social status, along with occupational status, management level, earnings, household income, and wealth. The level of education achieved acts as a structural element of a person's life, like a structural beam in a building. Many other aspects of life depend on a person's level of education and take shape with respect to it, including the other achieved statuses. In the advanced industrial nations, education level is increasingly important to status attainment (how high a person rises socially and economically) and status transmission (how many advantages in one generation are passed to the next). Higher status gives an individual more money, authority, influence, and freedom. It shifts life events toward more controllable and desirable ones, and it reduces the risk of a prolonged or severe discrepancy between goals and means.

The higher statuses achieved through education generally reduce stress. The chief exception is managerial responsibility. Management requires a responsibility to others for the performance of others; a manager is responsible to those above for the accomplishments of those below and is responsible to those below for the rewards and resources allocated by

those above. But no one can completely control the behavior of others. This creates stress. Work organizations compensate for the stress of greater responsibility by reducing other strains on managers. Higher salaries reduce managers' household economic strains and allow the purchase of goods and services that relieve other demands on time and effort. More important, however, managers generally are given more freedom to decide what they do and how and when they do it, which reduces the stress associated with job demands. They also generally are given more resources and opportunities to do things they find challenging and interesting, which helps make job demands stimulating rather than distressing.

Learned Effectiveness and the Sense of Control

During the twentieth century, formal education came to be viewed as a system for developing human capital. Economists had long viewed capital as material wealth in the form of money or property that is or can be used to produce more material wealth. However, the growth of wealth in the United States and other nations exceeded what could be attributed solely to accumulating monetary and physical capital. Economists revived Adam Smith's concept of human capital as productive capacity developed, embodied, and stored in humans themselves. Levels of formal education are the most important measure of human capital, along with work experience. Formal education develops the skills and abilities of general productive value (such as reading ability or punctuality), as distinct from ones of value in a particular job (such as knowing which forms to fill out or how to operate a particular machine).

Education reduces stress partly by enhancing material productivity but mostly because it develops general problem-solving abilities. The learned effectiveness makes people better at avoiding or solving problems and more confident, reducing stress throughout life. Formal education teaches people to learn. It develops the skills and habits of communication: reading, writing, inquiring, discussing, looking things up, and figuring things out. It develops analytic skills of broad use such as mathematics, logic, and, on a more basic level, observing, summarizing, synthesizing, interpreting, and experimenting. The more years of schooling, the greater the cognitive development, characterized by flexible, rational, complex strategies of thinking. Education teaches people to think logically and rationally, see many sides of an issue, analyze problems, and design strategies of personal action.

Education also develops broadly effective habits and attitudes such as dependability, good judgment,

motivation, effort, trust, and confidence. It instills the habit of meeting problems with attention, thought, action, and persistence. In school, people encounter and solve problems that are progressively more difficult, complex, and subtle. This process develops persistence and self-assurance as well as skill.

When individuals control their own lives, this means they exercise authority and influence over it by directing and regulating it themselves. People vary in the control felt over their own lives. Some feel they can do just about anything they set their minds to. They see themselves as responsible for their own successes and failures and view misfortunes as the results of personal mistakes. Others feel that any good things that happen are mostly luck – fortunate outcomes they desire but do not design. They feel personal problems mostly result from bad breaks or the callous selfishness of others and feel little ability to regulate or avoid the bad things that happen. The sense of control varies by degree, ranging from fatalism and a deep sense of helplessness to instrumentalism and a firm sense of mastery.

The sense of personal control is a learned and generalized expectation. As such it acts as a cognitive-behavioral accumulator, integrating past experience and bringing it into the present. In many ways, a sense of control is to successful action as wealth is to profitable investment – an accumulated product returning subsequent advantage. Perceptions of control grow out of the interaction between intention and outcome. The occurrence of something desired, planned, or attempted reinforces a sense that individuals' choices and actions have consequences. This, in turn, encourages attention to setting goals, directing actions toward the goals, evaluating apparent consequences, and revising efforts.

The sense of control is more than seeing things occurring as individuals want. It is individuals seeing themselves as the authors and editors of the choices and actions that link their preference to occurrence. Within contexts that support its success, critical self-direction sharpens their ability and encourages effort. The resulting effectiveness and resilience strengthen the sense of control in a beneficial developmental spiral. The system of formal education serves as the chief social institution for developing this sense of control.

A number of social and behavioral sciences recognize the importance of a sense of personal control. The concept appears in a number of related forms with various names, including internal locus of control, mastery, instrumentalism, self-efficacy, and personal autonomy; at the other end of the continuum (lack of control), this is fatalism, perceived helplessness, and perceived powerlessness. A low or negative

sense of control corresponds to learned helplessness, a behavioral state of suppressed attention and action that induces biological stress in mammals.

Life-Course Compounding

Many things in life produce effects that fade over time as individuals adjust to their circumstances, tastes and times change, enthusiasms wane, and new experiences intervene. Education's effects work the opposite way, growing with time. Education transforms people, putting individuals' lives on a different track. Because education develops resources that inhere in people, its consequences are present in all aspects of life throughout their entire lifetime. Education affects virtually all aspects of life, including habits, interpersonal relationships, family responsibilities, occupational exposures and opportunities, economic sufficiency and security, neighborhood qualities, autonomous and creative activities, and people's sense of controlling their own lives.

Education tends to speed or advance beneficial accumulations and to slow or delay detrimental ones. The consequences accumulate on many levels. They include job security, pay level, occupational status, and wealth on the socioeconomic level; habits such as exercise or smoking or relationships such as marriage on the behavioral level; body mass, atherosclerosis, or hippocampal mass on the anatomic level; aerobic capacity and blood pressure on the physiological level; and glucose tolerance or mitochondrial damage on the intracellular level. Undesirable accumulations typically can be reversed, even after a crisis, but generally only over a period of time as a result of concerted and multifaceted effort. Education helps individuals to avoid undesirable accumulations that need correction, to correct undesirable accumulations before they precipitate a crisis, and, failing that, to heed the implications of the crisis and take the difficult but necessary corrective action.

Many of education's consequences influence one another or regulate one another's effects. The feedback among accumulators amplifies the long-term effects of unchanging attributes such as sex, race, and year of birth and of persistent ones such as occupation and wealth. It also amplifies the effects of short-term random shocks to each accumulator.

Education makes individuals more adept at resource substitution (using one thing in place of another) and finding ways to achieve ends with whatever materials, relationships, and circumstances present themselves. Higher education makes individuals better at acquiring whatever they need and better at using whatever they find available. As a result, more education tends to increase an individual's store of

the society's standard resources while improving the individual's ability to improvise resources. A greater capacity for resource substitution makes the absence of any one standard resource less harmful for the better educated. Conversely, less education leaves individuals less adept at acquiring and inventing resources, increasing the individuals' dependence on any standard resource they have.

Stressors often occur in cascading sequences. The relative lack of resources and resourcefulness among the poorly educated exacerbates the outcomes at each step. Ineffective individuals often move through a cascading sequence of corrosive situations made worse at each step by the predisposing traits and conditions that led to those situations. The capacity for resource substitution helps the better educated to avert problems or ameliorate outcomes at each step.

Signs of Lower Stress

A variety of observations indicate that people with higher levels of education generally have lower levels of stress. They encounter fewer undesirable events and situations, as summarized earlier. They also report fewer of the symptoms and health consequences of stress. They report fewer signs of autonomic arousal such as sweaty palms, queasy stomach, dry mouth, faintness, or short of breath and a racing heart when not exercising or working hard. They report fewer days of feeling worried, anxious, and tense. They also report less hostility and anger toward others. (However, when worried, anxious, or tense, they are more likely to report anger along with it.) The better educated have fewer symptoms of depression such as feeling sad, lonely, or worthless and being sleepless, listless, or unable to face the day. They have fewer stress-related health conditions such as high-blood pressure, poor blood sugar control, or frequent colds and influenzas; fewer of the stress-related diseases such as coronary artery disease or type 2 diabetes; and fewer of the stress-related medical crises such as heart attack or stroke.

The better educated have lower scores on bio-marker indexes of allostatic load, which is the cumulative biological degradation from stress. There are two kinds of markers. One set measures blood, urine, or saliva levels of the hormones involved in the stress response, such as high overnight urinary excretion of cortisol, epinephrine, and norepinephrine and low serum dihydroepiandrosterone sulfate (DHEAS). The other set measures the cumulative effect of exposure to those hormones, such as high plasma levels of glycosylated hemoglobin (indicating chronic poor blood sugar control), high systolic and diastolic

blood pressure, high serum levels of total and of high-density cholesterol, high waist-to-hip ratio (measuring central adiposity), low serum albumin and high fibrinogen and C-reactive protein (indicating chronic inflammation), and indicators of impaired organ function such as low creatinine clearance in the kidneys and low peak air flow in the lungs. Differences in allostatic load scores predict mortality risk in older Americans and statistically account for approximately one-third of the differences in mortality risk across degree levels of education.

See Also the Following Articles

Aerobic Exercise and Stress Reduction; Allostasis and Allostatic Load; Anxiety; Autonomic Nervous System; Control and Stress; Coping and Stress: A Lens and Filter Model; Distress; Economic Factors and Stress; Health and Socioeconomic Status; Health Behavior and Stress; Learned Helplessness; Social Status and Stress; Social Support; Workplace Stress.

Further Reading

- Kristenson, M., Eriksen, H. R., Sluiter, J. K., et al. (2004). Psychobiological mechanisms of socioeconomic differences in health. *Social Science and Medicine* 58, 1511–1522.
- Mirowsky, J. and Ross, C. E. (1998). Education, personal control, lifestyle and health: a human capital hypothesis. In: O'Rand, A. (ed.) *Special issue on education over the life course, Research on Aging*, 415–449.
- Mirowsky, J. and Ross, C. E. (2003). *Education, social status and health*. Somerset, NJ: Aldine Transaction.
- Mirowsky, J. and Ross, C. E. (2003). *Social causes of psychological distress* (2nd edn.). Somerset, NJ: Aldine Transaction.
- Mirowsky, J. and Ross, C. E. (2005). Education, cumulative advantage and health. *Aging International* 30, 27–62.
- Ross, C. E. and Mirowsky, J. (1999). Refining the association between education and health: effects of quantity, credential, and selectivity. *Demography* 36, 445–460.
- Ross, C. E. and Mirowsky, J. (2001). Neighborhood disadvantage, disorder, and health. *Journal of Health and Social Behavior* 42, 258–276.
- Ross, C. E. and Van Willigen, M. (1997). Education and the subjective quality of life. *Journal of Health and Social Behavior* 38, 275–297.
- Ross, C. E. and Wu, C. (1995). The links between education and health. *American Sociological Review* 60, 719–745.
- Seeman, T. A., Crimmins, E., Huang, M.-H., et al. (2004). Cumulative biological risk and socio-economic differences in mortality: MacArthur Studies of Successful Aging. *Social Science & Medicine* 58, 1985–1998.

- Taylor, S. E. and Repetti, R. (1997). What is an unhealthy environment and how does it get under the skin? *Annual Review of Psychology* **48**, 411–447.
- Turner, R. J. and Avison, W. R. (2003). Status variations in stress exposure: implications for the interpretation of research on race, socioeconomic status, and gender. *Journal of Health and Social Behavior* **44**, 488–505.
- Turner, R. J., Wheaton, B. and Lloyd, D. A. (1995). The epidemiology of social stress. *American Sociological Review* **60**, 104–125.
- Wheaton, B. (1985). Models for the stress-buffering functions of coping resources. *Journal of Health and Social Behavior* **24**, 208–229.

Effort–Reward Imbalance Model

J Siegrist

University of Duesseldorf, Duesseldorf, Germany

© 2007 Elsevier Inc. All rights reserved.

Theory
Research Evidence
Implications for Intervention

Glossary

<i>Contract</i>	An agreement that defines a norm of equivalence by specifying obligations, benefits, rights, and duties in interpersonal exchange (e.g., work contract).
<i>Overcommitment</i>	A motivational pattern of work-related striving toward high achievement that is associated with increased health risks in the long run.
<i>Prediction error theory</i>	A theoretical model developed in neuroscience to predict neural activation and deactivation in relevant areas of the brain according to the principles of economy and learning.
<i>Social reciprocity</i>	A fundamental principle of social exchange that guarantees equivalence of give and take between two individuals or parties.
<i>Social role</i>	A set of expectations or norms (duties and options) directed toward people who hold important social positions (e.g., work role, family role, and volunteer role).

Theory

One of the important tasks of medical sociology consists in explaining how the social environment affects human health. Theoretical models are instrumental in identifying those aspects within the complex social reality that accounts for increased or reduced health risks in populations. These models are then translated into measures with the help of social science research

methods that meet the criteria of adequate reliability and validity and are tested in the framework of epidemiological or experimental study designs.

One such model, labeled effort–reward imbalance, is concerned with adverse effects on health produced by the violation of one of the fundamental, evolutionarily stable principles of human societies – social reciprocity. According to this principle, any action or service provided by person A to person B that has some utility to B is expected to be returned by person B to A. Exchange expectancy does not implicate a full identity of the service in return, but it is essential that this activity meet some agreed-on standard of equivalence. To secure equivalence of return in crucial transactions, social contracts have been established as a universal societal institution. A contract defines a norm of equivalence by specifying obligations, benefits, rights, and duties in interpersonal exchange. Trade, work and employment, marriage, and intergenerational transfer are examples of contractual exchange. These contracts may vary considerably according to the specificity of their regulations, the sanctions expected in case of deviance, or the time frame of exchange. Yet, in all instances, contracts are instrumental in providing members of a society with a sense of security by creating trust. Trust is a mental state motivating people to engage in social exchange even if the trade-off is highly uncertain. Expectancy of reciprocity is the driving force of trust.

The principle of reciprocity not only is rooted in human evolution, but plays a significant role in ontogenesis as well. Research on attachment formation in infancy has demonstrated the importance of reciprocal exchange between infant and caregiver in early postnatal life as one of the preconditions of normal human development. For these reasons, breaking the norm of reciprocity in contractual exchange is expected to have adverse consequences for the health and well-being of disadvantaged people.

The model of effort–reward imbalance states that recurrent or long-lasting nonreciprocity in contractual

exchange increases the risk of stress-related disorders in exposed people, due to the powerful role of this evolutionarily old grammar of interpersonal cooperation. This model has been developed and tested with regard to work and employment, and more recently has been expanded to less formalized types of contractual exchange, in particular marital or partnership relationship, exchange between parents and children, and voluntary work.

The principle of social reciprocity lies at the core of the employment contract, which defines distinct obligations or tasks to be performed in exchange for equitable rewards. Yet, according to the effort–reward imbalance model, nonsymmetric contractual exchange is expected to occur frequently under specific conditions (see [Figure 1](#)). In this case, high efforts spent at work are not reciprocated by equitable rewards in terms of money, esteem, and career opportunities, including job security. The model of effort–reward imbalance claims that lack of reciprocity between the costs and gains (i.e., high-cost, low-gain conditions) elicits strong negative emotions with a special propensity to sustained autonomic and neuroendocrine activation and their adverse long-term consequences for health.

There are three important conditions that increase the probability of recurrent effort–reward imbalance at work: dependency, strategic choice, and overcommitment. Dependency reflects the structural constraints observed in certain types of employment contracts (e.g., unskilled or semi-skilled workers and elderly employees) when no alternative choice in the labor market is available. In modern economies characterized by a globalized labor market and organizational downsizing, the lack of an alternative choice of workplace is relatively frequent.

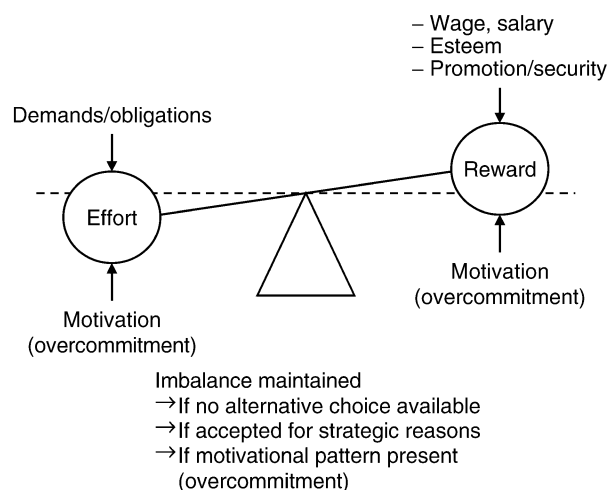


Figure 1 The model of effort–reward imbalance at work. Modified from Siegrist (1996).

In strategic choice, people accept high-cost, low-gain conditions of their employment for a certain time, often without being forced to do so, because they tend to improve their chances of career promotion and related rewards at a later stage. This pattern is frequently observed in early stages of professional careers and in jobs that are characterized by heavy competition.

The third condition, overcommitment, reflects the psychological reasons for a recurrent mismatch between efforts and rewards at work. People characterized by the motivational pattern of excessive work-related overcommitment may strive toward continuously high achievement because of their underlying need for approval and esteem at work. Although these excessive efforts often are not met by adequate rewards, overcommitted people tend to maintain their level of involvement. Work-related overcommitment is elicited and reinforced by a variety of job environments and is often experienced as self-rewarding over a period of years in occupational trajectories. However, in the long run, overcommitted people are susceptible to exhaustion and adaptive breakdown.

In summary, the model of effort–reward imbalance at work maintains that nonreciprocal contractual exchange is frequent under these structural and personal conditions, and that people experiencing dependency, strategic choice, or overcommitment, either separately or in combination, are at elevated risk of suffering from stress-related disorders.

Research Evidence

Several sources of information on the associations between nonreciprocal contractual exchange and health are available, such as data from cross-sectional and case-controlled studies, from prospective epidemiological observational investigations, from studies using ambulatory monitoring techniques or experimental designs, and from intervention trials. The prospective epidemiological observational study is considered a gold standard approach in this field because of its temporal sequence, its sample size, and the quantification of the subsequent disease risk following exposure.

Up to now, the model has been tested in approximately a dozen prospective epidemiological reports and, in addition, in many cross-sectional and case-controlled studies. Associations of either the full model or its main components with increased health risks were found in a majority of these studies. The relatively strongest and consistent effects concern cardiovascular diseases – effort–reward imbalance at work is associated with a doubling of risk over a mean

8-year observation period. In addition, nonreciprocity at work increases the probability of suffering from depression, psychosomatic symptoms, sickness absence, and poor self-rated health. With respect to health-adverse behaviors, and especially the intensity of daily cigarette smoking, alcohol dependence, and weight gain, associations with effort–reward imbalance at work were observed more consistently in men than in women. In addition to health-adverse effects, this type of chronic stress at work increases the likelihood of changing or giving up one's job, of retiring prematurely, and, according to first results, of deviating from regulations at work.

In summary, there is solid evidence indicating that failed reciprocity in a core social role, the work role, represents an independent risk factor for a variety of highly prevalent diseases. Moreover, it undermines work-related motivations and behaviors. Results are derived from a wide range of different occupations and professions, and they concern both genders, although, in general, the effects are stronger in men. In several studies, this imbalance was found to be more frequent among lower socioeconomic groups, and its effects on health are generally stronger in these groups.

So far, the bulk of evidence comes from investigations conducted in Europe, the United States, and Canada. Given the fact that the norm of contractual reciprocity is rooted in human evolution, we would expect recurrent violations of this norm to have similar effects on health across societies and cultures. According to recent studies, effort–reward imbalance at work is associated with several indicators of reduced health in Asian countries such as China, Taiwan, and Japan.

An additional line of research is devoted to the study of nonreciprocal exchange in marital or partnership relationships, in relationships between parents and children, and in other types of cooperative activities, such as voluntary work. The recurrent experience of effort–reward imbalance in these types of social exchange was found to reduce mental health and well-being in exposed people.

What are the neural and physiological correlates of the recurrent experience of effort–reward imbalance? Several studies investigated heart rate, heart rate variability, and blood pressure during work and non-work days, using ambulatory monitoring techniques. In general, cardiovascular activity differed between employees scoring high on measures of effort–reward imbalance and those with low scores on these measures. One study, in addition, assessed salivary cortisol excretion and found a higher mean level in employees with high scores on the scale measuring work-related overcommitment.

The neural correlates of reward frustration are still poorly understood. Yet, in an investigation using functional magnetic resonance imaging with a monetary reward paradigm, neural activation in reward-sensitive prefrontal mesolimbic dopamine projection sites was not reduced when expected rewards were frustrated in otherwise healthy subjects with an extensive history of work-related effort–reward imbalance. This was different in subjects without chronic stress experience, in whom activation in these areas was reduced, as hypothesized by the prediction error theory.

Implications for Intervention

Measures to improve the balance between effort and reward and, hence, to improve reciprocity and contractual fairness at work can be implemented at three levels. The first level relates to the individual worker. Increasing awareness of failed reciprocity at work among employees, informing them about possible health effects, and providing cognitive-behavioral interventions in high-risk groups to reduce the intensity of stressful experiences (relaxation response, stress inoculation, self-instruction, and reducing high levels of overcommitment) are examples of this approach.

A related second level of health promotion at work, the interpersonal level, concerns training in leadership skills, the improvement in the handling of conflicts, or the improvement of communication and cooperation in everyday work settings. Providing appropriate esteem and recognition was shown to be an important component of the balance between effort and reward at work, and this target can be met by focused leadership training.

However, to produce a lasting impact, these human-relations measures need to be supplemented by evidence-based organizational and structural changes in the work environment. Such changes concern the division of work; its quantity and quality; work schedules and their flexibility; monetary incentives; tailored promotion opportunities, including investment in training and requalification on the job; and, most important, enhanced job security. The probability that these and related measures are realized will increase with available evidence of the economic benefits for companies and organizations that invest their increased efforts into health-promoting quality of work. Such measures are not limited to paid work, but may be extended to other types of contractual social exchange, in particular voluntary work. In conclusion, it remains to be seen how this new scientific evidence can be transferred into health-promoting policy measures that diminish the burden of societal stress.

See Also the Following Articles

Economic Factors and Stress; Health Behavior and Stress; Psychosocial Factors and Stress; Social Status and Stress; Workplace Stress.

Further Reading

- Gouldner, A. W. (1960). The norm of reciprocity. *American Sociological Review* 25, 161–178.
- Marmot, M., Siegrist, J. and Theorell, T. (2006). Health and the psychosocial environment at work. In: Marmot, M. & Wilkinson, R. G. (eds.) *Social determinants of health* (2nd edn., pp. 97–130). Oxford: Oxford University Press.

- Siegrist, J. (1996). Adverse health effects of high effort – low reward conditions at work. *Journal of Occupational Health Psychology* 1, 27–43.
- Siegrist, J. (2005). Social reciprocity and health: new scientific evidence and policy implications. *Psychoneuroendocrinology* 30, 1033–1038.
- Siegrist, J., Starke, D., Chandola, T., et al. (2004). The measurement of effort-reward imbalance at work: European comparisons. *Social Science & Medicine* 58, 1483–1499.
- Tsutsumi, A. and Kawakami, N. (2004). A review of empirical studies on the model of effort – reward imbalance at work: reducing occupational stress by implementing a new theory. *Social Science & Medicine* 59, 2335–2359.

Elder Abuse

C P Holstege

University of Virginia, Charlottesville, VA, USA

H Holstege

Calvin College, Grand Rapids, MI, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by C P Holstege and H Holstege, volume 2, pp 19–22,

© 2000, Elsevier Inc.

Demographics of Elder Abuse

Characteristics of the Abused

Characteristics of the Abuser

Theories about the Cause of Elder Abuse

Institutional Abuse

Signs and Symptoms of Elder Abuse

Glossary

<i>Domestic elder abuse</i>	The abuse of an elderly person by someone closely associated within that elder's home or by somebody in the home of the person caring for that elder.
<i>Elder abandonment</i>	The desertion or willful forsaking of an elder by any person having the care or custody of that elder under circumstances in which a reasonable person would continue to provide care.
<i>Elder neglect</i>	The failure of a caregiver to fulfill his or her obligation or duty to care for an elder.
<i>Elder self-abuse</i>	A neglectful or abusive behavior by an elder that is dangerous to his or her own safety or health.

Financial elder abuse

Any theft or misuse of an elder's financial funds, property, or other resources by a person in a position of trust with an elder.

Institutional elder abuse

The abuse that occurs by paid staff in establishments designated to care for the elderly, including foster homes, group homes, and nursing homes.

Physical elder abuse

The willful infliction of physical force against an elder that results in tangible injury, pain, or impairment by a person who cares for or has custody of or who stands in a position of trust with that elder.

Psychological elder abuse

The deliberate infliction of mental or emotional suffering upon an elder by use of insults, humiliation, intimidation, threats, or other verbal or nonverbal abusive conduct.

Sexual elder abuse

The nonconsensual sexual contact of an elderly person.

Demographics of Elder Abuse

In the United States, the National Center on Elder Abuse reported that nearly 400 000 adult/elder abuse reports were investigated by Adult Protective Services in 2000. Of these, 48.5% were substantiated by Adult Protective Services. The majority of reports involved the domestic setting (60.7%), with only 8.3% occurring in the institutional setting. Perpetrators were most commonly family members (61.7%), especially spouses or intimate partners (30.2%) and adult children (17.6%). Evidence points toward the validity of the iceberg theory of elder abuse, which states that

only the most visible types of elder abuse and neglect are reported to official sources (e.g., Adult Protective Services) and that a large number of other incidents are unidentified and unreported.

Characteristics of the Abused

Women are disproportionately victimized, representing two-thirds of the victims of physical abuse, three-fourths of the incidents of psychological abuse, and greater than 90% of the cases of financial abuse. Elderly living alone have lower rates of abuse than elderly living with other persons. Increasing age is associated with a higher incidence of abuse and neglect. Elderly in poor health are nearly four times more likely to be abused than comparably aged elderly in good health. The elderly who are widowed, divorced, or never married are less likely to be abused.

Characteristics of the Abuser

Men are slightly more likely than women to abuse elders. Approximately 90% of alleged elder abusers are related to their victims. The majority of perpetrators of elder abuse are either the elder's spouse or adult children. Since families are frequently the primary caregivers for elderly relatives in domestic settings, the fact that family members are the primary perpetrators of elder abuse is not surprising. Most elder abusers are younger than their victims. This fact is most striking in the area of financial abuse, where nearly half of the abusers are 40 years old or younger.

Theories about the Cause of Elder Abuse

Numerous studies have examined the possible etiologies of elder abuse. Although the causes of most cases of elder abuse are multifactorial, six widely accepted theories have been reported in the literature.

Psychopathology in the Abuser

The psychopathology of the abuser theory focuses on the abuser's mental derangement as the primary cause of elder abuse. Specific conditions, such as psychiatric illness, alcohol and drug addiction, dementia, and mental retardation, lead to inadequacy in the abuser. As a result, the abuser is unable to positively relate to the elder. This subsequently leads to the abuse. Approximately 30% of abusers have documented psychiatric illness, and nearly 40% have a history of substance abuse.

Transgenerational Violence

According to the transgenerational violence theory, the abuser learns abusive behavior from other family

members, and this dysfunctional behavior is passed from one generation to the next. A cycle of violence then ensues. For these families, violence is perceived as a normal behavior pattern. For example, a person who was abused as child becomes the abuser when his or her parents become elderly and more dependent. Child, spouse, and elder abuse may be seen in these extended families.

Caregiver Stress

The caregiver stress theory emphasizes the stress of the abuser as the predominant factor that leads to elder mistreatment. The obligations associated with providing care for the elderly may place overwhelming demands on providers. Frequent falls, wandering, incontinence, disrobing, and verbal abuse by elders are examples of occurrences that place undue stress upon the caregiver. External stresses upon the caregiver (e.g., unemployment, personal illness, and financial hardship) may lead to his or her lashing out at the elder in frustration. A good relationship between the caregiver and the care recipient prior to illness and disability has been shown to minimize stress, even in the face of heavy caregiver demands.

The Web of Dependency

The web of dependency theory suggests that the frailty of an elder in and of itself is a cause for abuse and neglect. Increasing physical and mental impairment in the elder leads to an increasing dependency on a caregiver for the activities of daily living. As the caregiver's obligations increase, the burden and stress on him or her is also increased. As a result of this increased dependency, abuse occurs. Greater impairment and subsequent dependency of the elder may also diminish his or her ability to defend him- or herself from abuse or to escape the situation.

Caregiving Context

The caregiving context theory places emphasis on the situation of the caregiving. Elderly who are socially isolated are more prone to be abused. Families in which abuse occurs remain secluded to avoid detection. Elderly who share living space with their caregiver are also more prone to be abused. Tensions and daily conflicts are more difficult to avoid when the parties live together.

The Sociocultural Climate

Cultural and ethnic diversity may influence the prevalence of abuse. In cultures in which elders are highly esteemed or in which a strong value is placed on

Table 1 Possible signs of elder abuse and neglect

<i>Physical abuse</i>	<i>Physical neglect</i>	<i>Sexual abuse</i>	<i>Emotional abuse</i>
Bruises Unexplained In various stages of healing In regular patterns In shape of the article used Bite marks In an unusual location	Consistent hunger Poor hygiene Inappropriate dress Soiled clothing Weight loss Dehydration Urine burns Pressure sores	Difficulty walking Difficulty sitting Genitalia itching/pain Bruised/bloody genitalia Bruises around breasts Vaginal bleeding/tears Anal tears/bruising Stained/bloody underwear Sexually transmitted diseases	Habit disorder Sucking Biting Rocking Conduct disorder Antisocial Destructive Neurotic traits Sleep disorders Speech disorders Inhibition of play Psychoneurotic reaction Hysteria Obsession Compulsion Phobias Hypochondria
Burns In the shape of cigarette/cigar Immersion burns Patterned burns (e.g., from an iron) Rope burns Caustic burns	Over- or undermedication Hypothermia Lice Lack of functional aids, e.g., glasses, dentures, hearing aids, walking aids		
Fractures Unexplained In various stages of healing Multiple Spiral			
Lacerations Unexplained In an unusual location			
Internal injuries			

Table 2 Signs of financial or material exploitation

Abrupt changes in a will or other financial documents
Inclusion of additional names on an elder's bank signature card
Unauthorized withdrawal of funds from an elder's financial account
Unexplained disappearance of money or other valuables
Forging of an elder's signature for financial transactions
Sudden appearance of previously uninvolved relatives claiming rights to an elder's possessions
Sudden transfer of the elder's assets to another individual
Unnecessary services being rendered by an elder (such as unnecessary home improvements)
Unpaid bills or inability to purchase basic provisions despite availability of adequate financial resources
Sudden changes in a financial account or the banking practice of an elder
Elder's report of financial exploitation

family responsibility, caring for an elder member of the family is expected. Failure to fulfill this caregiver role may cause shame and frustration. Cultural sanctions against revealing family problems to outsiders may subsequently prevent families from seeking help. If a family has relocated to a foreign country, adaptation to that culture may create additional stress and conflict as to how to care for their elders.

Institutional Abuse

Abuse occurs not only in domestic settings, but also in institutions such as nursing homes, mental institutions, and adult foster homes. At least half of all

nursing home patients suffer from some form of dementia or mental illness. This group of patients is more prone to be abused and is less likely to report the abuse. In institutional settings, physical elder abuse has been reported to occur in one-third of patients, with psychological abuse occurring in up to 80% of elders. The causes for this abuse in institutions are multifactorial. The lack of understanding of the cause of the elder's behavior, dissatisfaction among the institution's staff, patient-staff conflict, and burnout among staff have been cited as causes for institutional elder abuse. Education of the staff is the most important factor to decrease elder abuse in institutions. In addition, more psychiatric services need to be provided for elders residing within institutions.

Signs and Symptoms of Elder Abuse

The signs and symptoms of elder abuse can range from subtle to grossly apparent. Professionals who routinely work with the elderly are typically not well trained in evaluating signs and symptoms of elder abuse. As a result of this lack of training, an underreporting of abuse has been clearly documented. Possible signs and symptoms of elder abuse are demonstrated in [Table 1](#). Elder abandonment also occurs and can manifest as the desertion of an elder at a hospital, nursing facility, shopping center, or other public location. Financial or material exploitation can manifest in a number of ways, with examples given in [Table 2](#).

Self-neglect is difficult to determine, and, depending on where the elder lives, specific laws exist pertaining to when the elder's self rights become superseded by the government. Signs of self-neglect include unsanitary living conditions (e.g., animal or insect infestation, no functioning toilet, fecal or urine smell), hazardous living quarters (e.g., improper wiring, no heat or running water), inappropriate or inadequate clothing, lack of necessary medical aids (e.g., eyeglasses, hearing aids, dentures), grossly inadequate housing or homelessness, poor personal hygiene, and improperly attended medical conditions. Social workers, the medical profession, law enforcement, and the general public must be educated to increase awareness and thereby decrease the occurrence of elder mistreatment.

See Also the Following Articles

Aging and Psychological Stress; Aging and Stress, Biology of; Alzheimer's Disease; Caregivers, Stress and; Child Abuse.

Further Reading

- Jones, J. S., Holstege, C. P. and Holstege, H. (1997). Elder abuse and neglect: understanding the causes and potential risk factors. *American Journal of Emergency Medicine* 15, 579–583.
- Kosberg, J. I. and Jordan, L. G. (1995). *Elder abuse: international and cross-cultural perspectives*. Binghamton, NY: Hawthorne Press.
- Lachs, M. S. and Pillemer, K. (2004). Elder abuse. *Lancet* 364(9441), 1263–1272.
- Riekse, R. J. and Holstege, H. (1996). *Growing older in America*. New York: McGraw-Hill.

Relevant Website

<http://www.elderabusecenter.org> – National Center on Elder Abuse (NCEA), U.S. Administration on Aging 2005.

Electrodermal Activity

G Turpin and T Grandfield

University of Sheffield, Sheffield, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G Turpin and L Harrison, volume 2, pp 23–27,

© 2000, Elsevier Inc.

Introduction

Methodology

Biological Mechanisms

Psychological Applications

Glossary

Electrodermal lability

An individual difference dimension of electrodermal activity. High labile individuals tend to demonstrate greater rates of nonspecific responses and may also have higher skin conductance levels and slower habituation rates.

Habituation

The process whereby the repetition of discrete stimuli leads to the decline in response amplitude or occurrence of electrodermal responses to successive stimulus presentations.

Nonspecific responses

Frequency counts of spontaneous or nonspecific responses (e.g., nonspecific skin conductance responses), which occur in the absence of external stimulation. Rates of nonspecific responding are said to reflect individual differences in electrodermal arousal levels and may also be related to task demands for energy mobilization, together with responses to internally mediated thoughts or emotions.

Phasic responses and tonic levels

Phasic responses consist of short-lived changes in electrodermal activity (EDA) elicited by discrete external stimuli and are characterized as either skin conductance responses (SCRs) or skin resistance responses (SRRs). Tonic levels are longer-term absolute measures of either skin conductance level (SCL) or skin resistance level (SRL). Typical ranges are 0.05–0.5 μS for SCR and 1–30 $\mu\text{S}/\text{cm}^2$ for SCL.

Skin conductance (SC) and skin resistance (SR)
Sweat glands

Units of electrodermal activity, which are expressed in either conductance (microsiemens or micromhos) or resistance (micro-ohms).

Two types of sweat gland are located under the skin surface: the apocrine

glands, which are usually associated with the hair follicles and distributed around the axilla and pubic regions, and the eccrine glands, which are widely distributed over the body surface and serve a thermoregulatory function. However, the latter are also more closely distributed (approximately 1000 glands/cm²) on the palms of the hands and the soles of the feet and are sympathetically innervated.

Introduction

At the turn of the nineteenth century, several researchers (e.g., Féré and Tarchanoff) demonstrated changes in the electrical resistance of the skin, which were elicited by either sensory or emotional stimuli. With the advent of electronic amplification and recording equipment in the mid-twentieth century, electrical recording from the skin as a correlate of psychological activity was becoming a common measure within psychological research. Early psychological researchers such as Darrow investigated the mechanisms underlying electrodermal responding, together with its psychological correlates. The use of the galvanic skin response as a measure of classical conditioning or psychological stress and arousal received widespread acceptance within the 1960s. There followed a period of consolidation whereby the appropriate methodologies were standardized and models of electrodermal responding were elucidated. Alongside other noninvasive measures of autonomic activity such as heart rate and blood pressure, electrodermal measures were commonplace within psychophysiological publications.

Electrodermal responses are frequently interpreted as reflecting changes in attention since they are considered major components of the orienting response. More generally, measures of electrodermal activity (EDA) are considered as reflecting changes in autonomic reactivity elicited by psychosocial events associated with emotion, task demands, motor responding, workload, stress, etc. Although contemporary psychophysiological research tends to rely more on centrally derived measures such as evoked potentials or brain imaging, peripherally recorded autonomic measures such as EDA still have their place as sensitive and noninvasive physiological indices of psychological processes and states. Finally, it should be acknowledged that there are a range of nonelectrical methods for quantifying sweat gland activity. One such method is the Palmar Sweat Index, which consists of taking an imprint of the

skin in order that the frequency of active sweat glands can be identified; it has frequently been used by stress researchers.

Methodology

Recording Techniques

There are a number of electronically derived measures of EDA, which can be obtained from the skin surface depending on the design and characteristics of the recording amplifiers. First, endogenous skin potentials (SPs) can be recorded directly using an appropriate amplifier and recording electrodes. This technique is referred to as endosomatic EDA. In contrast, exosomatic EDA consists of changes in an externally applied voltage or current, which is impressed through the skin via a pair of bipolar recording electrodes. The most common techniques rely on direct current (DC) measurement and consist of either measuring changes in conductance by keeping the voltage constant or changes in resistance by keeping the current constant. Alternatively, skin impedance may be measured using alternating current (AC). The most widely accepted technique is the measurement of skin conductance using a constant voltage applied between the electrodes of around 0.5 V. The techniques may also be applied outside of laboratory situations using ambulatory methods. More recently, circuits have been designed to measure EDA concurrently with brain imaging equipment and techniques.

The preparation of the skin surface and placement of electrodes are important methodological considerations. Since only some skin areas contain eccrine sweat glands that are psychologically reactive, EDA tends to be restricted to the palmar surfaces of the hands, and in particular, to the first and second fingers of the nondominant hand. In order to record EDA, specifically prepared electrodes (usually silver/silver chloride) need to be fixed in pairs to the skin and used in conjunction with specially made electrolytes. Researchers wishing to use EDA need to pay particular attention to published accounts of standardized procedures.

The recorded signal, whether it is conductance or resistance, is either displayed using a chart recorder or digitally sampled using an analog-to-digital converter and computer. Care should be taken to ensure that sufficient amplification and digital sampling rates are employed, so that the overall conductance level can be assessed, together with sufficient resolution so that phasic fluctuations can also be discriminated. In the case of traditional amplifiers and chart recorders, this will necessitate the use of a device that backs off a

calibrated proportion of the skin conductance level prior to high gain amplification.

It is essential that any research intended for publication and using electrodermal measures strictly conforms to published methodological guidelines.

Quantification

Electrodermal activity is usually expressed in conductance units (microsiemens or micromhos), even if a measure of skin resistance was originally obtained (i.e., conductance = $1/\text{resistance}$). Researchers usually discriminate between tonic and phasic electrodermal measures. Tonic measures refer to long-term fluctuations in EDA and are best characterized by changes in skin conductance level (SCL) and the frequency of nonspecific or spontaneous fluctuations (see Figure 1). They are brief changes in EDA that are not specifically elicited by external stimuli but that arise due to internal thoughts, situational demands,

emotional states, and individual differences in electrodermal lability. In contrast, phasic measures are specific skin conductance responses (SCRs), which have been elicited by an identified and usually externally presented stimulus. Phasic SCRs are usually quantified as an amplitude change from a baseline level value (prior to presentation of a stimulus) to the peak of the response (see Figure 1). In addition to amplitude measures, temporal indices such as response latency, rise time, and half-recovery time may also be assessed. Skin conductance responses are frequently obtained within what is known as an habituation paradigm, whereby simple sensory stimuli are presented to the participant and the autonomic responses measured to subsequent repetitions. When habituation paradigms are used, further measures of change in response amplitude across successive trials are also calculated using regression techniques.

Prior to analysis, electrodermal data are frequently transformed in order to achieve more normally distributed samples for parametric statistical analysis. There also exists a statistical dependency between phasic response amplitude and individual differences in tonic levels. Researchers will frequently use a variety of statistical transformations (e.g., log. and ratio measures) and range-correction methods to adjust for this response level dependency, which is sometimes termed the Law of Initial Values.

Biological Mechanisms

Fluctuations in either conductance or resistance have been attributed to variations in sweat gland activity. Essentially, alterations in conductance of the sweat gland are said to arise from a variety of different mechanisms involving changes in both sweat conductivity and physical level of sweat within the duct. The former is said to be associated with an active ionic reabsorption membrane located within the duct wall, and the latter with the mechanical extrusion of sweat from the gland. From a psychological perspective, the most important influence on sweat gland activity is through the autonomic cholinergic innervation to the glands. A variety of brain regions are implicated in the central control of EDA, including the lateral frontal and dorsolateral cortex, together with various cortico-limbic connections including the anterior hypothalamus and anterior limbic and infralimbic cortices. More recently, several brain-imaging studies have been conducted that looked at the functional location of EDA. These studies suggested separate mechanisms for the maintenance of electrodermal level and the activation of responses.

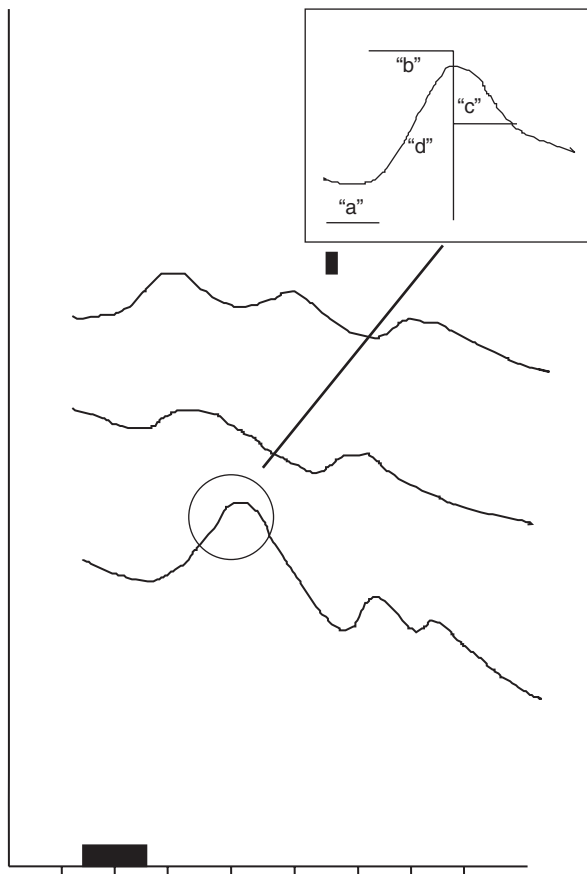


Figure 1 Schematic representation of skin conductance activity. Three individual skin conductance curves are represented with differing skin conductance levels and numbers of nonspecific responses (ns). Inset displays enlarged skin conductance response to specific stimulus (shaded block) and indicates (a) response latency, (b) response rise time, (c) response half-recovery time, and (d) response amplitude.

Psychological Applications

Attention and Information Processing

Phasic measures of EDA are frequently adopted as indices of attention and information processing. Skin conductance responses obtained within habituation and classical conditioning paradigms are interpreted in terms of the orienting response. Variations in SCR amplitudes are associated with changes in stimulus intensity, novelty, and significance. According to Ohman, the presence of an SCR indicates that a novel or significant stimulus has been detected by the information processing system and that a call has been made for further controlled processing of the incoming stimulus. The sizes of SCR amplitudes, therefore, are frequently interpreted as measures of attention.

Emotion

Attempts have been made to interpret SCRs with respect to either attentional or emotional responding. Lang and colleagues have argued, however, that whereas other peripheral psychophysiological measures such as heart rate, potentiated startle, and electromyographic activity might distinguish between the positive and negative aspects of emotional experience, SCRs appear sensitive only to the intensity or arousal dimension and not to the direction or valence of the emotion involved.

Stress

Both phasic and tonic measures have been argued to reflect peripheral arousal responses to stressful stimuli. In particular, EDA is said to be sensitive to psychosocial aspects of social interactions, workload, and occupational strain and has also been argued to reflect changes in anxiety. As well as the amplitude of SCRs, increases in the absolute SCL and the frequency of nonspecific response are often interpreted as reflecting increasing response to stress. Indeed, it is argued that a reliable individual difference dimension known as electrodermal lability might moderate the relationship between autonomic arousal and psychosocial influences. Similarly, EDA is frequently assessed in studies of psychopathology and has been widely investigated in clinical populations, including depression, anxiety, schizophrenia, hypertension, and diabetes.

See Also the Following Articles

Psychological Stressors, Overview; Regional Blood Flow, Stress Effects; Stress Effects, Overview.

Further Reading

- Boucsein, W. (1992). *Electrodermal activity*. New York: Plenum Press.
- Critchley, H. (2002). Electrodermal responses: what happens in the brain. *The Neuroscientist* 8, 132–142.
- Dawson, M., Schell, A. M. and Filion, D. L. (2000). The electrodermal system. In: Cacioppo, J. T. & Tassinary, L. G. (eds.) *Principles of psychophysiology: physical, social, and inferential elements* (2nd edn., pp. 200–223). New York: Cambridge University Press.
- Fowles, D. C., Christie, M. J., Edelberg, R., et al. (1981). Publication guidelines for electrodermal measurement. *Psychophysiology* 18, 232–239.
- Hugdahl, K. (1995). *Psychophysiology: the mind-body perspective*. Cambridge, MA: Harvard University Press.
- Jacobs, S. C., Friedman, R. and Parker, J. D. et al. (1994). Use of skin conductance changes during mental stress testing as an index of autonomic arousal in cardiovascular research. *American Heart Journal* 128, 1170–1177.
- Lang, P. J., Greenwald, M. K., Bradley, M. M. and Cuthbert, B. N. Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology* 30, 261–273.
- Nagai, Y., Critchley, H. D., Featherstone, E., Trimble, M. R. and Dolan, R. J. (2004). Activity in ventromedial prefrontal cortex covaries with sympathetic skin conductance level (SCL): a physiological account of a “default mode” of brain function. *NeuroImage* 22, 243–251.
- Ravaja, N. (2004). Contributions of psychophysiology to media research: review and recommendations. *Media Psychology* 6, 193–235.
- Schell, A. M., Dawson, M. E., Nuechterlein, K. H., Subotnik, K. L. and Ventura, J. (2002). The temporal stability of electrodermal variables over a one-year period in patients with recent-onset schizophrenia and in normal subjects. *Psychophysiology* 39, 124–132.
- Shastri, A., Lomarev, M., Nelson, S., George, M. S., Hozwarth, M. R. and Bohning, D. E. (2001). A low-cost system for monitoring skin conductance during functional MRI. *Journal of Magnetic Resonance Imaging* 14, 187–193.
- Turpin, G. (1990). Ambulatory clinical psychophysiology: an introduction to techniques and methodological issues. *Journal of Psychophysiology* 4, 299–304.
- Turpin, G. and Clements, K. (1993). Electrodermal activity and psychopathology: the development of the Palmar Sweat Index as an applied measure for clinical settings. In: Roy, J. C., Boucsein, W., Fowles, D. & Gruzelier, J. (eds.) *Advances in electrodermal research*, pp. 49–60. New York: Plenum Press.
- Venables, P. H. and Christie, M. J. (1980). Electrodermal activity. In: Martin, I. & Venables, P. H. (eds.) *Techniques in psychophysiology*, pp. 3–67. Chichester: Wiley.
- Vetrugno, R., Liguori, R., Cortelli, P. and Montagna, P. (2003). Sympathetic skin response: basic mechanisms and clinical applications. *Clinical Autonomic Research* 13, 256–270.

Emergency Personnel, Stress in

D S Weiss

University of California, San Francisco, San Francisco, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D S Weiss, volume 2, pp 28–31, © 2000, Elsevier Inc.

Recognition of the Problem

Predictors

Biological Factors

Interventions

Glossary

<i>Acute stress disorder</i>	A mental disorder involving anxiety, dissociation, and reexperiencing that leads to clearly impaired functioning; arises after exposure to a traumatic stressor in a 1-month time frame.
<i>Compassion fatigue</i>	The reduction of empathy and presence of burnout in helpers who have been overexposed to the suffering from traumatic stressors of others.
<i>Debriefing</i>	An intervention conducted soon after exposure to a traumatic stressor, in a group or individually, designed to lessen the psychological impact of the event on future functioning.
<i>Hippocampal volume</i>	An index of the size of a brain structure involved in both stress response and memory processes as measured by magnetic resonance imaging.
<i>Hypothalamic-pituitary-adrenal (HPA) axis</i>	The hormone system in the body, comprising the hypothalamus and the pituitary and adrenal glands, that regulates many basic functions, including the fight-or-flight response.
<i>Peritraumatic dissociation</i>	The experience of altered time sense, altered sense of self, or altered sense of place during and immediately following exposure to a traumatic stressor.
<i>Posttraumatic stress disorder (PTSD)</i>	A mental disorder involving the reexperiencing and avoidance of aspects of a traumatic stressor, accompanied by symptoms of increased arousal occurring at least 1 month after traumatic exposure.
<i>Routine occupational stressors</i>	The nontraumatic events that derive from organizational forces or other aspects of emergency work, such as contact with the public or lack of necessary equipment.

Traumatic stressor

The occurrence of an event that involves the actual or threatened death or serious injury or the threat to the physical integrity of oneself or a family member as either a victim or as a witness.

Exposure to traumatic stress in emergency personnel follows from the nature of the work, being primarily involved with witnessing or being the potential victim of serious injury, death, or threat to physical integrity and frequently involving danger, sometimes potentially lethal, to oneself. The stress reaction is similar to that of other groups exposed in this way, such as soldiers in combat, and can include the reexperiencing of the traumatic event, avoidance and numbing, and manifestations of hyperarousal such as sleep disturbance, impulsive anger and irritability, trouble concentrating, or hypervigilance. Although the response may be moderated or may habituate after simulations in training or repeated exposures over the course of years on the job, it does not seem to disappear. First-responders typically include police, firefighters, ambulance personnel, search-and-rescue teams, and the like. In addition to the effects of exposure to traumatic stress on emergency services personnel, most groups experience and react to routine stressors such as paperwork, media attention, insufficient equipment, and departmental politics. It is not yet definitively known what the relative impact of these two types of stressors may be on an individual's functioning on the job, but accumulating evidence suggests that both types of stressors are implicated in deleterious effects on psychological adjustment and biological processes.

Recognition of the Problem

The introduction in 1980 of posttraumatic stress disorder (PTSD) into the psychiatric nomenclature via the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn.) codified scattered knowledge about the effects of exposure to traumatic stressors from studies of combat veterans from World War I, II, Korea, and Vietnam; women who had been sexually assaulted; victims of the Holocaust; and children who had been sexually molested or physically abused. This codification gave structure to the investigation of the effects of traumatic stress on those who are exposed to such events, either accidentally or by

design. Traumatic stress can be defined as an event or series of events that threaten life, injury, safety, or bodily integrity. The event typically results in the experience of fear, horror, or a similar emotional response in those initially exposed. Obviously, one important difference between emergency service personnel and ordinary citizens is that, over the course of their careers, emergency services personnel are exposed repeatedly to such events. Consequently, these groups have become of considerable interest in the attempt to understand the effects of traumatic stressors on functioning.

The sequelae of exposure to a traumatic stressor varies enormously, all the way from no discernible effect to severe and disabling psychiatric symptoms. The formal psychiatric diagnoses that can result are acute stress disorder and PTSD. Acute stress disorder is diagnosed after exposure to a traumatic stressor when the individual has, within a month of exposure, a sufficient number of four types of symptoms that cause significant disruption in ongoing daily functioning: (1) dissociative symptoms such as depersonalization or derealization; (2) intrusive or reexperiencing symptoms such as nightmares or intrusive mental images of the event; (3) avoidance symptoms such as shunning talking about the event or any reminder of it; and (4) anxiety or arousal symptoms such as poor sleep, irritability, or exaggerated startle response. Acute stress disorder appears to be quite rare in emergency services workers because screening selects out those with obvious vulnerability and training familiarizes and acclimates workers and by so doing is thought to reduce the shocking nature of the events.

When there is the presence of a formal psychiatric disorder, it is more commonly PTSD. PTSD is diagnosed only after at least 1 month has passed since the traumatic event and is characterized by symptoms of three types: (1) intrusion or reexperiencing, (2) avoidance and numbing, and (3) increased arousal. The types of symptoms involved are very similar to those of acute stress disorder, with the exception of dissociative symptoms; these do not need to be present to diagnose PTSD.

It was only after PTSD appeared in the nomenclature that researchers began to systematically study the effects of their occupation on emergency services personnel from the perspective of traumatic stress. There is now a relative consensus that some percentage of workers in these occupations develop psychological symptoms, sometimes to a disabling degree, as a result of their exposure to some of the horrific and repeated incidents with which they are faced. Disasters and catastrophes since the year 2000 have also helped bring these issues to greater

public awareness. Concomitantly, there is a greater recognition that the normal course of reaction to traumatic stressors involves many of the same psychological and biological systems that are implicated in PTSD symptoms.

There have now been studies of police and firefighters in the United States, the United Kingdom, the Netherlands, Australia, and New Zealand, to name a few. The results of these studies are in agreement that PTSD symptoms are present in a subset of employees. They are in less agreement about whether the severity of symptoms warrants a diagnosis of full PTSD or whether partial or subclinical PTSD is more common. As well, there is a wide range of results for the percentage of workers who do meet criteria for a diagnosis. These range from rates of approximately 30% to less than 5%. The high end of these rates are higher by tenfold than the findings of the National Comorbidity Study Replication, a national prevalence study in the United States of ordinary citizens. The 12-month prevalence of PTSD was 3.5%. The severity of symptoms was roughly one-third each for serious, moderate, and mild.

Part of the heterogeneity of the results may be due to the issue of how much time has elapsed since the traumatic event that serves as the trigger for the diagnosis to be made because the natural course of reactions is for a considerable diminution of symptoms over a period of 6 months to 1 year. Another factor may be that some studies have assessed symptoms in relation to exposure to critical incidents on the job without requiring the specification of a distinct event, especially when an event is prolonged in time (e.g., rescue workers at the September 11 site).

PTSD symptoms, whether at the level of formal diagnosis or only of partial or subclinical proportions, are problematic because they interfere with functioning. Moreover, some attempts to deal with the symptoms can be harmful (e.g., the excessive use of alcohol) and exacerbate the symptoms or cause additional difficulties. In the last 5–10 years, awareness and education have helped reduce the stigma associated with sustaining psychological injuries from exposure to traumatic stress, but the attitude that the symptoms are a sign of weakness continues to be relatively prevalent. Those who are affected are frequently not met with compassion or understanding for incurring these difficulties. Consequently, the minimization or concealment of problems from others, and perhaps even from oneself, can lead to not getting help. This tendency may also partially contribute to the wide range of estimates of the degree of symptoms.

Another problem for emergency services personnel is the impact on family members when the worker

does not want to burden his or her loved ones with the stories of difficult or horrific outcomes. Although understandable, not doing so can eliminate one of the most powerful antidotes that helps overcome traumatic stress – social support, a result now documented in two recent meta-analytic studies.

Better documented has been the effect of nontraumatic stressors on first-responders. Stressors such as relations with the media, bureaucratic or political issues within organizations, interactions with supervisors or superiors, and inadequate equipment or backup do not customarily produce intrusion, avoidance, and hyperarousal but, instead, have been shown to be related to more general psychiatric symptoms such as anxiety, depression, and hostility. As well, there are studies that have documented a relationship between hostility and cardiovascular disease, a relationship that has repeatedly been found in ordinary citizens. Unknown at this time is whether there is a synergistic or interactive effect of exposure to both kinds of stressors that is greater than would be expected from the effects of each alone. This area remains ripe for investigation.

Predictors

There has been a growing body of recent work on the predictors of whether a person will have difficulty after exposure to traumatic stress. Because many workers are exposed and only a minority ever develop significant problems, research has focused on identifying factors that might explain why some develop problems whereas others do not. The bulk of the findings come from retrospective studies, with their inherent weaknesses, and have involved mostly men. Gender has been found in some studies to moderate relationships, but the main effect is the well-known result that women report more symptoms. Longitudinal prospective studies of police and firefighters are currently underway, some large scale, but the results of such work has yet to be reported.

From retrospective work, one clear factor that appears to be related to an individual's developing problems is the level of exposure to a critical incident, although the relationship is not as strong as might be assumed. For example, all other things being equal, a police officer who is shot at several times is more at risk than one who is assaulted by a suspect without a weapon. Similarly a firefighter who narrowly escapes a flash-over is at greater risk than a colleague who is just missed by a falling beam. The level of exposure as a predictor has been found in other groups as well (e.g., combat veterans and rape victims). Across all groups, the two meta-analyses found that life threat produced an average weighted effect of $r = 0.26$ and

$r = 0.23$. Some studies, however, have shown that the individual's appraisal of how serious an event was, as opposed to the broad consensus of different levels of threat, is also a factor in determining who develops PTSD symptoms.

A more psychological factor that has been identified is peritraumatic dissociation. Peritraumatic dissociation is the experience of altered time sense, altered sense of self, altered sense of place, disorientation, and panic during and/or immediately following exposure to a traumatic stressor. In a number of retrospective studies of emergency services personnel involved in a large natural disaster, those with the highest levels of symptoms were those who also described having dissociative experiences during or immediately after the trauma. In a group of civilians, a prospective study confirmed the importance of peritraumatic dissociation. The meta-analysis that examined it found that peritraumatic dissociation was the strongest identified predictor of PTSD, with an average weighted effect size of $r = 0.35$.

Preexisting psychiatric difficulties or a family history of psychiatric disorder shows some ability to predict symptomatic level, but the degree is less than more proximal factors. The meta-analyses showed an average weighted effect size of $r = 0.17$ and $r = 0.12$ for prior trauma, $r = 0.17$ and $r = 0.11$ for prior adjustment, and $r = 0.17$ and $r = 0.13$ for family history of psychopathology.

A recent 2-year prospective study of firefighters identified high hostility and low self-efficacy as predictive of PTSD and general psychiatric symptom development. Other studies have shown that cumulative exposure to critical incidents over a career is not as salient a factor as might have been expected.

Biological Factors

In the last 5 years, there has been a marked increase in attention to biological factors in stress response, with a special growth of work in the consequences of traumatic stress on the brain and neuroendocrine systems. Earlier work had established that sleep was disturbed in those exposed to traumatic stressors, but recent findings have differentiated these observations. At least one large-scale study has suggested that in police officers lifetime cumulative exposure to critical incidents was clearly related to the frequency of nightmares but less strongly related to sleep quality. The main factor affecting sleep quality (other than shift-work issues) was the more common stress from the work-environment problems described previously. Studies in the sleep laboratory have tended to confirm that those with PTSD do have diminished sleep quality, but it remains an open question as to whether

general stressors or traumatic stressors exert a larger effect on sleep quality in first-responder groups.

The first biological area examined in the study of the effects of exposure to traumatic stress in emergency services personnel, as well as in research in PTSD more broadly, was the response of the hypothalamic-pituitary-adrenal (HPA) axis to stressors. The role of the hormone cortisol has long been known to be involved in the stress response. Advances in the ease of measurement of cortisol and accumulating findings from many groups exposed to traumatic stress, including emergency services personnel, have shown abnormalities. The most well-established finding is a lowered baseline cortisol on awakening and higher PTSD symptoms. Another fairly well-established finding is relatively more suppression of cortisol following the administration of the synthetic hormone dexamethasone, a finding indicating increased feedback sensitivity. At this point, findings involving the HPA axis are rapidly accumulating, and more careful studies are showing more differentiated results.

A second area of work has been in the area of brain imaging: magnetic resonance imaging (MRI), functional MRI, and positron emission computed tomography (PET). An initial set of findings in combat veterans and victims of sexual assault began to appear in the last decade and tended to show that the volume of the hippocampus was reduced in those who had developed PTSD compared to those various controls. Because the hippocampus is part of the limbic system and is implicated in the formation, storage, and processing of memory, these initial findings were theoretically interesting (recall that intrusive memories of the traumatic event are a unique symptomatic feature). Subsequent studies tended to find reduced volume, sometimes on both sides and sometimes on only the left or right side of the brain, but several found no volume reduction. These ensuing studies were helpful in illuminating the complexity of these brain changes. Currently, it appears that reduction is more common than not and that this is true in police samples as well, but whether smaller volume is a result of traumatic exposure or a risk factor for it is unclear. Newer studies are examining the functioning of the amygdala, which is known to be activated by fear and stress, and the prefrontal cortex, a brain area that plays a role in inhibiting the activity of the amygdala and is involved with complex emotional and cognitive functioning.

Interventions

Given the increased recognition of the psychological consequences of exposure to traumatic stress, many

agencies have instituted programs for debriefing their personnel soon after traumatic exposure. Debriefing as a term is now used in a variety of ways, although its origin is military and was directed at clarifying actual events and improving operations rather than assisting with psychological reactions. Debriefing is now used to refer to some kind of structured intervention designed to assist in the psychological processing of a traumatic event and/or to serve as a preventive intervention. Consumers of debriefings have been extended from those who respond to an event to individuals who can be described as victims or witnesses of the event.

The advent of an evidence base to justify the provision of interventions, most commonly referred to as evidence-based medicine, has had an impact on interventions after traumatic stressors. The practice of evidence-based interventions is predicated on combining clinical judgment with knowledge of external evidence from rigorous systematic research. Several recent summaries of the effects of debriefing for victims have shown equivocal results, with one or two studies concluding that there may even be deleterious effects. These summaries have had a profound impact on the conclusions of the scientific community about the use of debriefing, even though the main summary was based on only nine studies. One fairly recent study of interventions in the workplace after September 11 showed positive effects up to 2 years after the collapse of the World Trade Center towers.

Whether intentionally or due to lack of knowledge, the literature as of this writing has taken the cautions against the efficacy of debriefing based on studies of single-session individual interventions for victims as applying to single-session group interventions for first-responders. This has led to some confusion and a debate about how to help emergency personnel deal with exposure to traumatic stressors.

The most widely used approach in emergency services personnel is critical incident stress debriefing (CISD). This intervention, developed by Mitchell, was developed specifically for emergency services personnel. It, or variations of it made by individual organizations, is still currently widely used and is typically evaluated as helpful and useful, but studies of its actual efficacy have been sorely lacking. Only a single controlled randomized study could be located, albeit with a positive result. It typically comprises seven stages: introduction, facts, thoughts and impressions, emotional reactions, normalization, planning for the future, and disengagement. Widely overlooked in the debate, however, is that CISD was designed as one element of a larger critical incident stress management (CISM) system that includes

precrisis preparation, demobilization and briefings, defusing, CISM, individual crisis intervention, family crisis intervention and organizational consultation, follow-up and referral for assessment, and, if needed, treatment.

Evaluating CISM from an evidence-based perspective is daunting, and it is unclear how many first-responder organizations use the whole CISM approach as opposed to the CISM component only. Given the lack of systematic evidence of efficacy, the persistence of CISM or modifications of it need to be explained. One possibility is that debriefing serves a needed function even if eventually it is shown not to reduce psychological symptoms more than not being debriefed.

Debriefing acknowledges in an organizational framework that reactions to traumatic stress are part of the hazards of the occupation, in the same way that dosimeters symbolize the hazards of exposure to radiation for those who work in such an environment. Unlike radiation, however, whose effects are understood to be out of the control of the individuals exposed, the effects of traumatic stress are often viewed as being under the control of those exposed. Debriefing explicitly acknowledges that this is only partly so.

The goals of psychological debriefing include a decrease of overwhelming emotions and cognitive disorganization and an increase in self-efficacy through the process of giving meaning to the traumatic event and legitimizing emotional reactions. An important component of debriefing is emotional disclosure, something that involves risk for those who see themselves or who think they are seen as people who do not react emotionally. Notwithstanding this view, which is consistent with either avoidance and denial or appropriate positive coping, the very essence of social interaction revolves around telling about what has happened, telling one's story. A debriefing is a formal opportunity to tell one's story and to tell it in a way that brings out aspects that might otherwise be overlooked or ignored, such as having been afraid, ashamed, fearful, confused, overwhelmed, distraught, anxious, revolted, enraged, or any one of many other emotions that arise during emergencies. Even for those who do not volunteer much personal information, information about the natural process of responding to traumatic stress and adaptive coping strategies is often presented.

For the small percentage of first-responders who do develop PTSD and could benefit from professional intervention, the evidence base is more substantial. Antidepressant medications from the broad class of selective serotonin reuptake inhibitors (SSRIs) have

established efficacy and in the United States sertraline and paroxetine have been approved by the Food and Drug Administration (FDA). There is less evidence for the efficacy of other agents (e.g., anticonvulsants or tricyclic antidepressants), and the literature suggests their use is more limited. Both U.S. and UK treatment guidelines endorse an SSRI as the medication of choice.

There is good evidence that trauma-focused psychotherapeutic approaches are also effective. The largest body of evidence supports cognitive behavioral psychotherapeutic approaches, with or without some explicit prolonged exposure. In the largest study of group therapy for veterans with chronic PTSD, exposure therapy was not found to be distinctly superior to a present-centered group that did not focus on the trauma.

There is considerably less work on treatment explicitly for first-responders. Part of the explanation for this is that many individuals seek treatment privately, outside their organization or agency, to avoid the potential stigma or impact on career advancement. Concerted efforts to reconceptualize symptoms that result from exposure to traumatic stressors as psychological injuries, akin to symptoms of physical injury, may go a long way toward mitigating the stigma associated with such responses.

Further Reading

- Brewin, C. R., Andrews, B. and Valentine, J. D. (2000). Meta-analysis of risk factors for posttraumatic stress disorder in trauma-exposed adults. *Journal of Consulting and Clinical Psychology* 68, 748–766.
- Keane, T. M., Marshall, A. D. and Taft, C. T. (2006). Posttraumatic stress disorder: etiology, epidemiology, and treatment outcome. *Annual Review of Clinical Psychology* 2, 161–197.
- Mitchell, J. D. and Everly, G. S. (2001). *Critical incident stress debriefing: an operations manual for CISM, defusing and other group crisis intervention services*. Ellicott City, MD: Chevron Publishing.
- National Institute for Clinical Excellence. (2005). *Clinical guideline 26 post-traumatic stress disorder (PTSD): the management of PTSD in adults and children in primary and secondary care*. London: NICE.
- Ozer, E. J., Best, S. R., Lipsey, T. L., et al. (2003). Predictors of posttraumatic stress disorder diagnosis and symptoms in adults: a meta-analysis. *Psychological Bulletin* 129, 52–73.
- Rose, S., Bisson, J., Churchill, R., et al. (2002). Psychological debriefing for preventing post traumatic stress disorder (PTSD). *Cochrane Database of Systematic Reviews* 2, no. CD000560.
- Smith, A. and Roberts, K. (2006). Interventions for post-traumatic stress disorder and psychological distress

in emergency ambulance personnel: a review of the literature. *Emergency Medicine Journal* 20, 75–78.

Ursano, R. J., Bell, C., Eth, S., et al. (2004). Practice guideline for the treatment of patients with acute stress disorder and posttraumatic stress disorder. *American Journal of Psychiatry* 161(supplement), 3–31.

Wagner, S. L. (2005). Emergency response service personnel and the critical incident stress debriefing debate. *International Journal of Emergency Mental Health* 7, 33–41.

Emotional Inhibition

H C Traue

The University of Ulm, Ulm, Germany

R M Deighton

The Cairnmillar Institute, Melbourne, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by H C Traue and R M Deighton, Volume 2, pp 32–38, © 2000, Elsevier Inc.

Emotion and Inhibition

Domains of Emotional Inhibition

Pathways from Emotional Inhibition to Health Disorders and Illness Behaviors

Rituals and Therapeutic Interventions

Glossary

Alexithymia

A constellation of cognitive and affective characteristics including difficulty identifying and communicating subjective feelings, a restricted imaginative life, and a concrete and reality-oriented style of thinking.

Behavioral inhibition

A concept encompassing several behaviors in about 15% of otherwise healthy children in response to unfamiliar social events, including reduced spontaneity, subdued emotional expressiveness, shyness, social avoidance and several peripheral physiological hyperactivities. It is thought to be related to limbic-hypothalamic arousal, in response to socially stressful events.

Emotional intelligence

The ability of an individual to adaptively and effectively regulate his or her emotional behavior in a social context. This encompasses the ability to recognize subjective feelings, to manage emotions, to transform emotions into expressiveness and action, to react empathetically, and to shape relationships.

Myogenic pain

Pain stemming from dysfunctional muscular activity as part of motor behavior in relation to stress, posture, movement, and emotion (e.g., low back pain, tension type headache, repetitive strain injury, and myofascial pain disorder).

Socialization

The process by which an individual gradually becomes integrated into the norm and value system of a social group or a society. It is based on the assumption of interaction between the biological organism and the social environment during psychological development.

Torture

One of the most severe stressors motivated mainly by political ideologies and perpetrated in about 60 states worldwide. The adverse effects of torture stem from the man made nature of these stressors and induce specific and nonspecific physical and mental disorders of long duration because they shatter victims' basic assumptions about human benevolence.

Emotion and Inhibition

Emotions are essentially transactions between individuals and their social environment. They give personal meaning to external and internal stimuli and communicate meaning from the individual to others. Emotions are composed of interpretations of intero- and exteroceptive stimuli, intentions, physiological patterns of arousal, and motor behavior including overt emotional expressiveness. The interaction of these different components in the individual and the social and physical environment are mediated by the central nervous system. From a system regulation point of view, emotional expressiveness has two important functions: first, it serves a communicative function in that it facilitates the regulation of person–environment transactions and, second, the feedback function of behavioral expressions controls

the intraindividual regulation of emotion. This means that active responding toward the environmental trigger may influence an experience indirectly through the attenuation of a negative emotional stimulus or directly through self-regulation. Thus, expressive behavior can serve simultaneously as a component of emotional processes and as a coping response. Three prominent scientists of the turn of the century, all of whom were active at a time of major discoveries in neurophysiology, contributed to important developments in the concept of inhibition. The neurophysiology of C. S. Sherrington (born 1857), the theory of the higher nervous system of I. P. Pavlov (born 1849), and the psychoanalysis of S. Freud (born 1856) transformed the principle of inhibition into a key concept in neurophysiology (in the case of Sherrington) and higher mental functioning (in the cases of Pavlov and Freud).

For many years, an inverse relationship between expressive behavior and autonomic responsivity has been documented, such that the inhibition of overt emotional expressiveness can lead to an autonomic overreaction. This has been considered to be a significant factor in the etiology and maintenance of psychosomatic disorders. A number of early researchers in the first two decades of the century reported measurements of high physiological activity in subjects suppressing emotional expression. These studies led toward the concept of internalization and externalization, wherein two behavioral coping styles for dealing with psychological tension were discerned: behaviorally, outwardly directed, or physiologically, within the individual. Following this concept, the term *internalizer* has been used to describe a person exhibiting a low level of overt expressiveness under stress yet a high level of physiological excitation, whereas an *externalizer* is characterized by high expressiveness and a low level of physiological expressiveness in social situations. Temoshok proposed a model of internalizing and externalizing coping styles integrating the severity of stressors which intended to predict the occurrence of mental disorders (dependent on degree of externalizing coping) and somatic disorders (dependent on degree of internalizing coping). Another inhibition theory, put forward by Pennebaker, summarizes the process by which failure to confront traumatic events results in poorer health. The principal assumption of this theory is that inhibiting ongoing behavior, thoughts, and feelings requires physiological work. It has been suggested that the increased autonomic responses of internalizers may reflect the work of behavioral inhibition. Over time, the work of inhibition acts as a low-level cumulative stressor. As with all cumulative stressors, sustained inhibition is linked to increases in stress-related diseases and various other disorders

such as cardiovascular and skin disorders, asthma, cancer, and also pain.

Domains of Emotional Inhibition

A model (see [Figure 1](#)) of how emotional stress, under a given social situation, can trigger or modulate health disorders is described later. Health disorders and illness behavior are considered as different, but related, processes, and distinct mechanisms may contribute dependently and independently to different aspects of a given disorder and its behavioral consequences.

On a phenomenological level, emotional stress can occur on a severity dimension ranging from daily stressors, through more traumatic life events, to more chronic or severe psychotrauma. A common underlying factor among such situations is that each needs to be coped with by the individual. Emotional stress can be seen as being processed by way of an inhibition-implosion dimension (to implode means to collapse or cause to collapse inward in a violent manner as a result of external pressure), modulated by dispositional factors (innate, personality, and socialization).

Innate and socialization factors are of particular importance. Possible individual differences in limbic mechanisms for opioidergic pathways corresponding with increased vulnerability to stress induction have been suggested. Emotional processing relating to the inhibition-implosion idea has been discussed in relation to several topics like control, suppression, type C personality, repression, alexithymia, or ambivalence. Each of these concepts covers different aspects of overt emotional expressiveness from a personality or coping perspective.

Inhibition is the most general term for the incomplete processing of emotional stress when these stressors induce bodily changes (physiological, endocrinological, or immunological) and the cognitive, emotional, and behavioral processing are dysfunctional such that subjective experience and spontaneous expression of emotions and action tendencies are separately or simultaneously attenuated and intra- and interpersonal regulation is disturbed. This process may be a product of innate and/or acquired behavior.

A classification of involved mechanisms differentiates among the following: genetic inhibition, repressive inhibition, suppressive inhibition, and deceptive inhibition. All four classes of inhibition occur in every state of psychophysiological arousal produced by emotional stimulation.

Genetic inhibition reflects the genetically determined basis of behavioral inhibition. Studies working with young children have classified those children

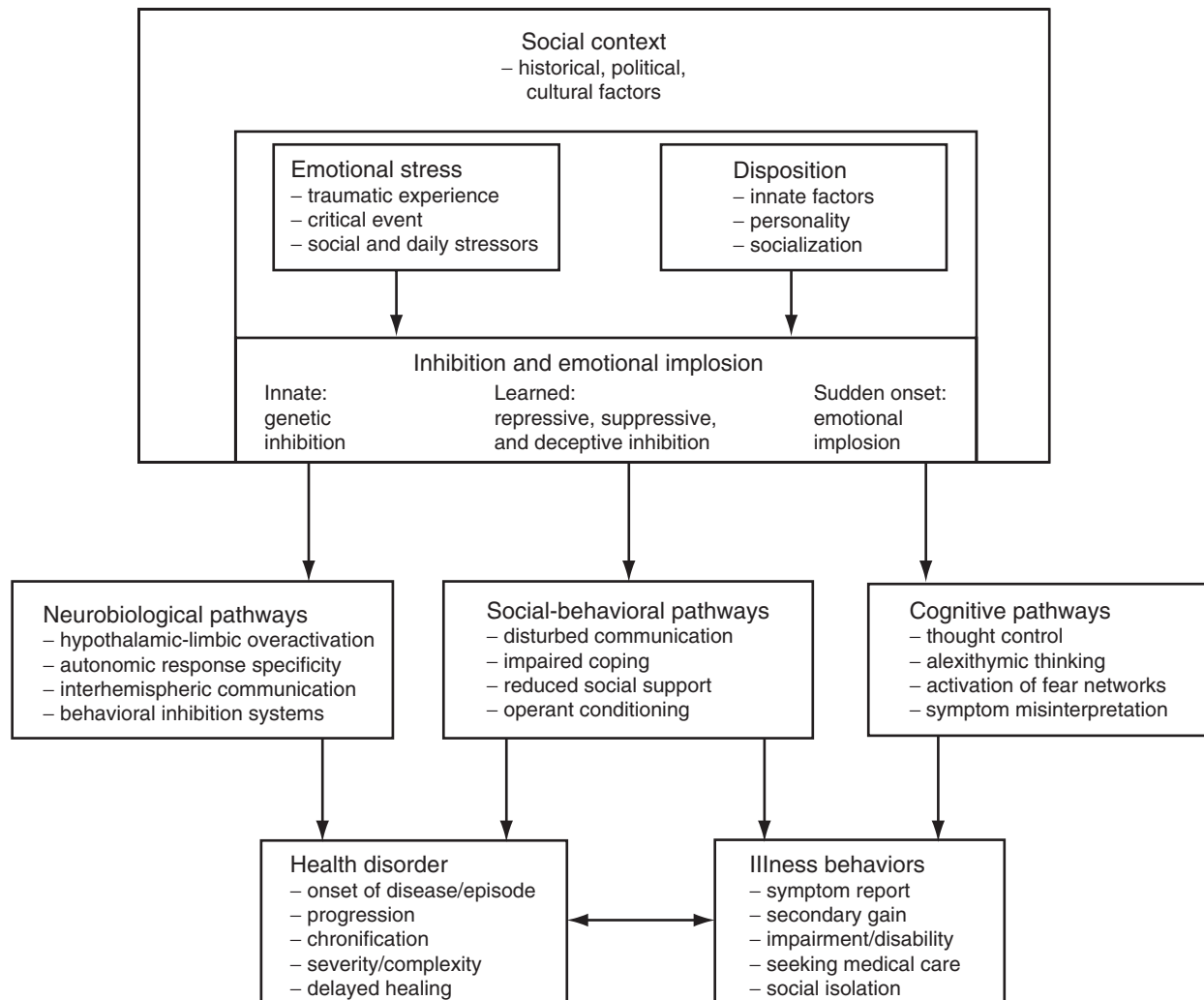


Figure 1 Psychological pathway model of emotional inhibition with neurobiological, social-behavioral, and cognitive pathways between stress/disposition, emotional processing, and subsequent health disorders or illness behaviors (adapted from Traue, 1998).

who were least able to initiate interaction in a social situation with other children and adults as behaviorally inhibited. Most children's degree of behavioral inhibition has been shown to be stable over a period of 5 years. In inhibited children, increased levels of arousal, norepinephrine and salivary cortisol have been found.

Repressive inhibition is defined as emotional processing with attenuated subjective experience of emotional arousal. Emotional expressive responses in repressive inhibition can be based solely on cognitive interpretation of the situation and are nonspontaneously organized. Because the individual is unable to feel his or her own arousal, insufficient response information is experienced, which in turn decreases the need to express emotions or cope with an emotional stressor. In addition, the cognitive interpretation of the situation without the emotional component may

be wrong or misleading. Prolonged bodily arousal and impaired coping could result. Repressive coping style is the best known model for repressive inhibition.

Suppressive inhibition is best circumscribed by suppression of emotional arousal. The emotional arousal is recognized by the individual but spontaneous expressive and cognitive behaviors are involuntarily suppressed. Suppressive inhibition of emotions could result from interactions between innate factors and socialization. For example if individuals show increased responses under stressful encounters, they are prone to socialization conditions of punishment and negative reinforcement, initiating a learning history with decreases in spontaneous expressiveness and increases in bodily reactivity.

Finally, individuals under emotional stress, aware of their bodily reactions and their urge for expressiveness, can voluntarily suppress this need or try to

present a false response to a receiver, called deceptive inhibition. Whether an individual is poker faced or displays a false emotional response, such inhibition consumes cognitive capacity, reducing the individual's coping capacity and providing additional stress.

In reviewing the psychophysiological and psychosomatic data, one is led to the conclusion that emotional inhibition is potentially harmful. However, it should be noted that although the correlations between bodily processes and the above four forms of inhibition support such a notion, in certain circumstances inhibition can be beneficial for the individual and his/her relationship with the social environment. Inhibition becomes toxic when it is related first to physiological, endocrinological hyperarousal or immunological dysfunction, second to longstanding dysregulation of emotions within the individual on a cognitive and behavioral basis, and third if inhibition disturbs the individual's social relations.

Inhibition constitutes a risk factor for health under normal stressors. If the severity of stressors is dramatically high, the mental and physical health consequences are inevitable. In traumatic stress situations like rape, criminal bodily attacks, or torture, the individual may well lose control over strong emotional responses. The emotional responses of horror, panic, and loss of control, could literally cause a violent breakdown in the mental and bodily systems. Such an emotional implosion is visible in the symptom pattern of posttraumatic stress disorder (PTSD): cognitive and behavioral avoidance of trauma stimuli, numbing of general and emotional responsiveness (emotional anesthesia), detachment from other people, and persistent symptoms of arousal such as disturbed sleep, exaggerated startle response, and somatic complaints. In addition, persons with PTSD suffer an increased risk of social phobia and major depressive and somatizing disorders.

The psychological, physical, and social symptoms in PTSD are a form of emotional processing that describes an extreme form of inhibition. While inhibition generally develops over a long time span through interaction between innate and socialization factors, implosion can occur in a very short time as a result of a single event.

Pathways from Emotional Inhibition to Health Disorders and Illness Behaviors

Emotional stress modulated through innate, personality, and socialization factors can trigger, maintain, or worsen health disorders and related illness behavior through neurobiological, social-behavioral, and cognitive pathways. With respect to illness behaviors,

the pathways include biases in symptom reporting, secondary gain by presented symptoms, subjective feeling of being impaired, pressure to seek medical help, and social isolation.

Neurobiological Pathways

It can be assumed that emotional inhibition is strongly neurobiologically based. The behavioral inhibition system and the behavioral activation system have been discussed as possible neurobiological structures. Empirical evidence from between-subject studies shows that inhibited, repressed, or suppressed emotional expressiveness is linked to greater autonomic arousal, both under conditions of emotion induction and voluntary deception. There is rich empirical evidence for neurobiological correlates of inhibition in respiratory, cardiovascular, muscular, digestive, endocrine, and immune functions. Immune functioning is of particular interest because it is the immune system that may be relevant in all sorts of infectious, allergic, and neoplastic illness processes. Inhibited style of processing upsetting events can compromise immune functions, resulting in higher serum antibody titers, decreased monocyte counts, and poorer natural killer cell activity. Other areas of research relevant to the neurobiological pathways of inhibition include hypothalamic-limbic overactivation, prolonged activation of physiological response specificity, hemispheric brain lateralization of emotion processing and faulty interhemispheric communication, the neuroregulation of action, and the behavioral activation and behavioral inhibition systems.

Social-Behavioral Pathways

A variety of social-behavioral pathways connect inhibition-implosion to health disorders and illness behaviors. First, neurobiologically innate factors (shyness, behavioral inhibition, hypersensitivity, introversion) are superimposed by classical and operant conditioning in the socialization of an individual. Since individuals with these characteristics in early childhood are more easily conditioned, the process of socialization involves greater vulnerability to them than it does for others. Under critical developmental conditions, the gap between emotional expressiveness and physiological hyperactivity may increase. A lack or deficit in emotional expressiveness will hinder interpersonal communication. It is implied that deficits in interpersonal communication disturb the development of emotional competence which is important for sharing experiences, maintaining psychological and physical contact, and adapting to the social environment. These are the deficiencies in healthy coping competencies that Salovey and

Mayer termed emotional intelligence. Other consequences of inhibited emotional expressiveness are disturbed social relations resulting in social isolation and a disrupted social support network.

Normally, persons respond to emotion-evoking stimuli with emotional expression, and such reactive expression is realized through facial muscle activity and movements with reafferent neuronal signals in the central nervous system, which contribute to the individual's emotional experience. However, subjective emotional experience does not depend mainly on this nervous input as argued in the facial feedback hypotheses, but feedback does contribute positively to sensitivity toward the physiological aspects of emotion. If this sensitivity is disrupted, an individual will not perceive adequately increased muscle tension or other autonomic nervous system reactions caused by stress and consequently will not initiate healthy relaxing behavior. Bischoff *et al.* demonstrated that the hypothesis of deficient perception of muscle tension holds for myogenic pain. Patients with this kind of pain were reliably less able to judge the extent of their muscle tension than were controls.

It is conceivable that, under unfavorable circumstances, expressive behavior of mainly negative emotions like anger and aggressiveness is punished socially and thus justifiably avoided. The suppression of expressive behavior can be realized by an additional increase in muscle activity. Such avoidance behavior or inhibition is very adaptable in the short term, and it helps to modify a socially stressful situation. This reduction of emotional expressiveness is conditioned by the learning mechanism of negative reinforcement (the avoidance of punishment).

Cognitive Pathways

Memories and thoughts of emotional stressors are generally unpleasant. Increasing severity of the stressful encounter makes imagining the event painful or even unbearable in the case of traumatic experience. Most individuals attempt to suppress or inhibit the thoughts surrounding the events. As soon as the inhibition work begins, the urge to distract oneself and the mental energy put into this process fuels the images. Consequently, triggered intrusions and unwanted thoughts make life more stressful than before. In addition, thought control interferes with natural ways of coping (e.g., sharing the experience with important others and thinking through the event). Therefore inhibition may be dangerous because it hampers the individual who has suffered a critical life event from resolving the stressful experience cognitively and behaviorally.

Other facets of problematic cognitive processing include an alexithymic or low-level thinking style.

As part of the inhibition process, individuals may tend to exclude the emotional content of the stressful encounter from their language representation of the event. Although this may help to avoid negative emotionality in the short term, it impairs one's own complete processing and integration of the stressful experience. The lack of integration into the self-concept makes an individual prone to activation of fear networks. Finally, impaired or unfinished cognitive processing makes an individual prone to a misinterpretation of bodily symptoms. Instead of understanding bodily reactions as part of emotional responses, the individual conceptualizes the bodily reactions as symptom patterns and seeks medical help for illnesses. Badly advised medical treatment procedures result in iatrogenic diseases, trapping the individual in a vicious cycle. Cognitive appraisal of a situation depends partly on facial feedback as a source of emotional information. When the expressive components of emotional reactions are systematically repressed by inhibition, the individual unlearns accurate assessment of stressful circumstances. This learning mechanism occurs since the estimated load of a stress situation is dependent not only on external features of the situation, but also on the subjective experience of stress-conditioned reactions. When, however, an inhibited person takes bodily reactions into account in the evaluation of a situation, his or her judgment will be impaired when the original physiological components of mainly negative emotions are interpreted as symptoms.

In clinical studies, psychosomatic patients (e.g., suffering headache) have been found to report significantly lower stress levels than control groups, but showed nearly twice as much neck muscle tension as the controls. Although the arousal and muscle tension data indicated higher levels of stress, patients were unable or not willing to report those stressors. It appears that patients tend to interpret their stressors in terms of bodily symptoms rather than as underlying levels of stress.

Rituals and Therapeutic Interventions

There is at least some implicit knowledge in most societies that emotional inhibition has negative health implications. The conflicts resulting from the need for emotional regulation on the one hand and the need for disclosure, sharing, and catharsis on the other lead to a variety of cultural phenomena to overcome these adverse consequences. These include older universal cultural rituals (such as rituals of grief or lament) or religious acts such as confessions. The Western (Wailing) Wall in Jerusalem, where Jews have been going for centuries to deliver a written

prayer, is possibly an example of an ancient disclosure phenomenon. Today people can deliver their prayer to the Western Wall via the Internet, and similar services are offered in connection to Christian confession. Contemporary western societies have also introduced psychotherapy for enhancing emotional expressiveness. Here, talking or writing about emotions is encouraged as well as acting out emotions, in role plays. Assertiveness training aims at effective expression of emotion, and catharsis-based techniques like confrontation are modern remedies for anxiety, posttraumatic stress disorders, and the like. All of these techniques seem to have in common that they are directed at the construction of meaning from emotional experience. Different cultures tend to construe emotional experience (including stress) in different ways, such that many non-Western cultures tend to emphasize the somatic components of emotional suffering, whereas Western cultures focus on the psychological components. Such cultural differences could influence the pathways between emotional inhibition and illness. Hence, clinicians should take into account the relevant cultural conceptions of emotional experience and behavior a client has been exposed to when choosing interventions for emotional inhibition. If a client emphasizes the somatic components of his or her reaction to a very stressful event, it may not be (dysfunctional) emotional inhibition, but rather a cultural construal of emotional suffering. Only when culture-sensitive exploration reveals lack of insight into relevant psychosocial factors should the possibility of dysfunctional emotion processing be considered.

See Also the Following Articles

Anger; Emotions: Structure and Adaptive Functions; Grieving; Social Support; Cardiovascular Disease, Stress and; Torture.

Further Reading

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th edn.). Washington, DC: American Psychiatric Association.

Bischoff, C., Traue, H. C. and Zenz, H. (eds.) (1989). *Clinical perspectives on headache and low back pain*. Hogrefe & Huber. Toronto/Lewiston/Gottingen/Bern.

Deighton, R. M. (2003). *Culture, emotional inhibition & somatization*. Ann Harbour, MI: Proquest Digital Dissertations (www.umi.com).

Deighton, R. and Traue, H. C. (2005) Emotional inhibition and somatization across cultures, *International Review of Social Psychology* 18(1/2): 109–140.

Kagan, J., Reznick, J. S. and Snidman, N. (1988). Biological bases of childhood shyness. *Science* 240(4849), 167–171.

Mauss, I. B. and Gross, J. (2004). Emotional suppression and cardiovascular disease: is hiding your feelings bad for your heart? In: Nykliček, I., Temoshok, L. & Vingerhoets, A. (eds.) *Emotional Expression and Health. Advances in theory, assessment, and clinical applications*. Hove/New York: Brunner-Routledge.

Nykliček, I., Temoshok, L. and Vingerhoets, A. (2004). *Emotional Expression and Health. Advances in theory, assessment, and clinical applications*. Hove/New York: Brunner-Routledge.

Pennebaker, J. W. (1995). Emotion, disclosure, and health: an overview. In: Pennebaker, J. W. (ed.) *Emotion, disclosure, and health*, pp. 3–10. Washington, DC: American Psychological Association.

Rimé, B., Herbette, G. and Corsini, S. (2004). The social sharing of emotion: Illusory and real benefits of talking about emotional experiences. In: Nykliček, I., Temoshok, L. & Vingerhoets, A. (eds.) *Emotional Expression and Health. Advances in theory, assessment, and clinical applications*. Hove/New York: Brunner-Routledge.

Salovey, P. and Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition, Personality* 9, 185–211.

Schwartz, G. E. and Kline, J. P. (1995). Repression, emotional disclosure and health. In: Pennebaker, J. W. (ed.) *Emotion, Disclosure, and Health*. Washington, DC: American Psychological Association.

Smith, C. E., Fernengel, K., Holcroft, C., Gerald, K. and Marien, L. (1994). Meta-analysis of the association between social support and health outcomes. *Behavior and Medicine* 16, 352–362.

Temoshok, L. (1983). Emotion, adaption, and disease. In: van Dyke, C. & Zegans, L. S. (eds.) *Emotions in Health and Illness*. New York: Grune and Stratton.

Traue, H. C. (1998). *Emotion und Gesundheit. Die psychobiologische Regulation durch Hemmungen*. Heidelberg: Spektrum.

Traue, H. C. (2001). Emotional Inhibition and Health. In: Smelser, N. J. & Baltes, P. B. (eds.) *The international encyclopaedia of the social and behavioural sciences*, pp. 4449–4454.

Traue, H. C. and Pennebaker, J. W. (eds.) (1993). *Emotion, Inhibition and Health*. Toronto/Lewiston/Gottingen/Bern: Hogrefe & Huber.

Emotions: Structure and Adaptive Functions

R J Contrada and H Leventhal
Rutgers, New Brunswick, NJ, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition
article by H Leventhal, volume 2, pp 39–46,
© 2000, Elsevier Inc.

The Relationship between Psychological and Physiological
Models of Emotion
Two Questions about the Relationship of Emotion and
Cognition
Primary Emotions Theory
The Function and Structure of Emotion
Biology, Emotion, Stress, and Health
Concluding Comment

Glossary

Autonomic nervous system The nervous system, consisting of three branches, two of which, the sympathetic (energy mobilizing) and parasympathetic (energy restorative), play a critical role in the stress response. Sympathetic effects are widespread; parasympathetic effects are localized to particular organs. The third branch, the enteric system is localized in the gastrointestinal tract and affects the gut during stress.

Conspecifics The members of one's own species. Much emotional behavior has evolved in interaction with members of one's own species and is designed, therefore, for regulating behavior within the social group.

Cortisol A hormone produced by the outer layer of cells (cortex) of the adrenal glands. Its multiple functions include the conversion of fats and proteins to sugars usable for energy expenditure, and moderating the immune response. Secreted as part of the hypothalamic-pituitary-adrenocortical stress response.

Cytokines Protein molecules formed by immune cells that communicate with and activate other subsets of immune cells and have direct effects on the central nervous system, producing sickness behaviors, including fatigue, malaise, and anhedonia.

Dopamine A neurotransmitter important for addictions and the activation of physical movement. Severe deficits lead to physical freezing (Parkinson's disease) and are

Epinephrine

Emotion circumplex

Hierarchical mechanisms

Insulin

Moods

Primary emotions

Schemata

treated with dopamine replacement therapy (L-Dopa).

A hormone that has widespread effects throughout the body. It is produced by the inner part of the adrenal gland (medulla) and plays a major role in the stress response, increasing heart rate and blood pressure, activating motor activity, releasing sugars for work, and stimulating immune responses. It is the end product in the sympathetic-adrenal-medullary stress response; also known as adrenaline. A circular representation of affective space defined by two dimensions: pleasant (love; mirth; happiness) to unpleasant (anger; determination) emotions and activated (surprise; suffering; fear) to relaxed (contempt; disgust) emotions. The names of the dimensions and specific emotion terms vary slightly from study to study.

Psychological and physiological processes described in terms of a hierarchy of behavioral control mechanisms. The output of each level produces reactions that affect activity at both lower (automatic action) and higher (conscious thought) levels.

A hormone found in various forms in plants, bacteria, and invertebrate and vertebrate animals; in humans it is formed in the pancreas. It is critical for the conversion (and hence the storage) of sugars as fats. Insulin deficits lead to excess blood sugar, called type 1 diabetes in children (caused by destruction of insulin-producing cells) and type 2 in adults (typically caused by excess weight), which can result in destruction of small blood vessels and lead to blindness, loss of limbs, kidney damage, and heart disease.

Affective states that vary in quality, as do emotions, but describe milder fluctuations in feeling. Moods are typically longer lasting than emotions and are presumed to involve the same underlying psychological and physiological systems. A finite set of six emotional reactions (happiness, surprise, fear, suffering, anger, and disgust) that have been identified in repeated studies and are presumed to be innate and potentially identifiable in people in all cultures at all times.

Organized neural structures formed by associative learning (repeated pairing of stimulus and response) that integrate

stimulus inputs with interpretations (meaning) and motor reactions. Schemata operate automatically, without conscious thought or voluntary effort, creating conscious emotional experience (fear, anger, and joy) and overt responses (flight, fight, and approach).

Self-conscious emotions

Emotions that can appear only after the infant has formed a central neural representation of him- or herself as distinct from other humans.

Subcortical areas

Collections of neural cells or nuclei that are important for emotional behaviors such as fear and positive affect. These nuclei lie deep within the brain and are less accessible and harder to study than the cortex. Ongoing research continually expands our understanding of their function (e.g., the amygdala is involved in learned fear behaviors and fear memory).

Emotional experience and emotional disorders have occupied the minds of philosophers, physicians, and laymen through much of recorded history. Beginning with Plato in the fourth century BC, emotion has been set apart from rational thought and typically viewed as a potential source of danger and dysfunction. In the second century AD, the Greek physician Galen identified four emotional temperaments – sanguine, choleric, melancholic, and phlegmatic – which he linked to four bodily humors or fluids. Historians of the Middle Ages described how moments of rage led kings to slay close friends and grieve them the following day. And since the late nineteenth century, physicians have hypothesized that anger and hostility generated by the stress of competitive, rapidly paced modern life explain the ascendance of cardiovascular disease as a major source of death in industrialized nations. Today, most Westerners believe that emotional distress is a primary cause of diseases ranging from life-threatening cancers and cardiovascular disorders to the annoying common cold.

The prominence of emotion in historical and medical thought and everyday life contrasts with the scanty attention given emotion in psychological research during the first half of the twentieth century. Only in the past 3 decades has emotion emerged as a focus of research in psychology and neuroscience, although the quantity of research on emotion has been dwarfed by the vastly greater number of studies on cognitive processes. It has proved far easier for computers to simulate perception, thought, decision processes, and learning than to describe and make sense of emotions. Disagreements in defining emotion, the subjectivity of human emotion, and unwarranted confusion about the relationship between

psychological and physiological analyses have been responsible for this state of affairs.

The Relationship between Psychological and Physiological Models of Emotion

William James's elegant prose stimulated a massive search for patterns of peripheral autonomic nervous system activity (e.g., heart rate and sweat gland activity) when he argued that the commonsense view that "... we meet a bear, are frightened and run ..." is a sequence that "... is incorrect, that the one mental state (fright) is not immediately induced by the other (seeing the bear), that the bodily manifestations must first be interposed between, and that we feel ... afraid because we tremble ..." (1890/1950: 450, 451). Although the strong form of James's somatopsychic model of affect has not been supported, there is growing knowledge regarding visceral influences whose effects on higher brain processes may be relevant to understanding emotion. Nonetheless, decades of study have failed to generate compelling evidence of clear patterns of autonomic activity associated with specific feelings, due perhaps to a failure to attend to the more complex features of James's writing. One lesson to be taken from this search is that data on physiological activity cannot, by themselves, illuminate our understanding of emotion unless these data are related either to behavioral manifestations of emotion such as overt action or to self-reports of private experience. This is just as true of real-time images of brain activity now available to affective neuroscientists as it was of signs such as perspiration and an irregular heart beat on which the early physicians first based the link between emotion and bodily processes. Psychological analyses of subjective feelings, cognition, and action and physiological analyses of central neural activity and peripheral physiology are parallel descriptions of unfolding emotional events. One description does not substitute or explain the other. A full understanding of ongoing emotional processes requires elaboration of both psychological and physiological models and the analysis of their linkages. Each type of analysis informs, defines problems, and sets limits or constraints for the other.

Two Questions about the Relationship of Emotion and Cognition

Psychologists and neurobiologists, not always in opposition to one another, have taken contrasting positions on the often heatedly debated question: Are emotions independent of cognition? Those arguing no assert that emotions always follow cognitive

appraisals; that is, the bear must be understood as endangering one's life before its presence elicits fear. Those arguing for independence assert that emotions can be provoked without prior thought or reasoning and support their position by pointing out that stimuli processed by subcortical areas of the brain can evoke emotion before these stimuli are enriched by cerebral cognition. The heat of debate is often unaccompanied by light because the advocates of both sides are seldom clear as to the meaning of cognition and the meaning of emotion. If we recognize that we are describing a common process from two separate vantage points, we can decompose the argument into separate questions.

Does Cognition (Elicit) Precede Emotion?

An answer to the timing question depends on the meanings of cognition and emotion. If cognition is defined as complex logical or deliberate reasoning, the answer is: cognition seldom precedes emotion. If cognition includes the automatic, sometimes preconscious representations of a stimulus, the likely answer is: very often. An explosive sound to one's immediate rear or the unexpected sight of a scorpion resting on the bathroom ceiling of one's vacation suite is likely to evoke startle and an abrupt shudder of fear. Being cut off by an aggressive driver can provoke anger, and the sight of one's grandchild emerging from an airliner readily elicits joy. Emotions are often provoked by recognition processes that occur far more swiftly and with less involvement of working memory than that which is needed for abstract logical thought. Different levels of cognition, from minimum perception through recognition to more complex cognition, can be antecedents of emotion, with the quality of the emotion elicited depending on features of the information (surprise, threat, etc.).

If cognition is broadly defined, it is likely that most emotional experiences are preceded and elicited by cognitive activity, a point with which James would agree. But what is usually the case need not always be so. Emotional reactions may also be the direct product of neurochemical changes in the brain and the body. Cytokines produced by the immune system's response to infection act on the nervous system and appear to produce depressed mood, whereas exercise elevates levels of cerebral peptides to create positive mood. Although it is likely that emotional episodes are more frequently initiated either by external events or by images and thoughts about such events rather than by detached neurochemical changes, we ought not regard the sequence as rigid because cognitive and emotional processes interact with one another during ongoing emotional episodes.

Can There Be Emotion without Cognition?

The second emotion-cognition question concerns whether one can have an emotion without cognition. At a time when the literature was nearly entirely devoid of evidence linking specific patterns of peripheral autonomic response to specific subjective emotions, Schachter and Singer hypothesized that emotional behavior and experience require the combination of unpatterned peripheral physiological arousal with the perception of a meaningful contextual-environmental cognition. The perception of unpatterned autonomic activity is necessary for an emotional episode, but the contextual or environmental context defines the meaning of the episode and gives it its particular quality (i.e., fear, anger, or joy). Thus, the scorpion on the bathroom ceiling identifies the perceived physiological activity as fear, the aggressive driver defines it as anger, and the grandchild's smile causes it to be experienced as joy. Advocates of this proposition were not mixing psychological and physiological explanations because their hypothesis proposes that it is the perception of physiological activity that is necessary, but not sufficient, for emotional experience and action.

Schachter and Singer's classic 1962 experiment illustrates the nature of this combinatorial hypothesis. Subjects were recruited for a study of the effects of a pharmacological agent on vision and were told they would have to wait 20 min after an injection of the agent before the visual tests could begin. For some subjects, the injected agent was epinephrine, which activates the autonomic system and provides the somatic arousal hypothesized to be necessary for emotion; other participants were inoculated with an inert placebo. Approximately one-third of the participants given epinephrine were informed about its actual side effects, that they might notice shakiness in their hands, warmth in the face, and heart palpitations. These informed participants had a nonemotional cognition to explain their somatic activation. The other epinephrine-injected participants were either uninformed or misinformed about the somatic effects of the injection. The misinformed group was given to expect somatic cues irrelevant to epinephrine as a control for focusing attention on the body.

The experimenters then added the second cognitive component presumed to be necessary for the experience of emotion; this component was not expected to elicit emotion in the accurately informed or the placebo subjects. The cognition was supplied by a second participant, actually a collaborator of the experimenters, who enacted one of two routines during the 20-min wait before the visual tests were to begin. For some subjects, the collaborator enacted

fun-generating activities such as making and flying paper airplanes; for others, the collaborator enacted anger-generating activities such as expressing irritation with the questionnaires and finally refusing to participate and marching angrily from the room. As expected, the epinephrine uninformed and misinformed subjects emulated the happy or angry actions that supplied a meaning for their somatic arousal. Less emotional behavior was apparent from the participants given a placebo because they lacked the necessary bodily arousal and from the informed epinephrine-injected participants because they were accurately informed and had an explanation for their somatic arousal.

The combination hypothesis has not fared well over time. It faltered for several reasons. First, it could not account for the presence of fear-motivated avoidance seen in many phobic subjects, particularly males, in the absence of strong signs of physiological arousal. Second, later experiments failed to replicate the findings, particularly with regard to anxious emotion as distinct from fear or anger; epinephrine might elicit anxious upset without the need for a cognitive cue. Most important, however, the hypothesis that meaningful cognition was necessary to provide the quality of emotional behavior identified the hypothesis with social constructionist theories, all of which ran afoul of accumulating evidence that a set of discrete or primary emotions are part of our innate constitutional makeup. The Schachter and Singer study does, however, make an important, often ignored, point – under ambiguous conditions, environmental cues play an important role in shaping emotional reactions.

Primary Emotions Theory

Seemingly in sharp opposition to social constructionist positions are theoretical models postulating a set of primary emotions that are universal and inborn. Specific problems with this proposal need to be kept in mind as we review the evidence in its favor: There is disagreement as to which emotions should be included in the set of primaries, and there are various shades of meaning as to what is meant by universal and inborn.

The Primary Emotions

Expressive behavior and primary emotions The use of facial, vocal, and postural expressions as the pathway for the identification and study of emotion was stimulated by Charles Darwin's treatise on *The expression of the emotions in man and animals* (1872). Patterned facial expressions, body postures, and vocal tones were seen as indicators of emotional

states and as ways of communicating emotional states to others and, more important perhaps, as the source of emotional experience in the actor. Thus, in contrast to William James, who emphasized the role of visceral somatic activity as the source of emotional experience, theorists of primary emotions believe that the expressive face and voice tell both others and ourselves what we are feeling.

Early studies of judgments of facial expression conducted through the first half of the twentieth century generated a picture of emotional life with specific emotions such as joy, contentment, anger, fear, and sadness arranged about the circumference of a circular plane defined by two axes: pleasant–unpleasant and active–passive. The circumplex emerged from statistical summaries of the judgments made of photographs of actors posing specific emotions. These enacted expressions (to look happy, contented, angry, fearful, disgusted, sad, etc.) were appropriately classified by judges ranging from newspaper readers who mailed in their ratings of the pictures printed in local newspapers to the more usual undergraduate psychology students. The judgments mirrored what the actors had been instructed to pose and the circular model provided an economical description of the data.

Critics of the primary emotion hypothesis pointed to the lack of agreement among the lists of affects proposed by different theorists. This criticism has been countered, effectively in our view, by identifying primary emotions about which all agree, that is, joy; contentment; sadness or depression, disgust, fear, and anger. Disagreements about these primary emotions have been more apparent than real, reflecting differences in labels used to refer to the same emotion. Disagreement about expressions of surprise and interest are less easily reconciled because they are labeled as emotions by some theorists and not by others.

A second issue regarding the concept of primary emotions concerns the role of facial expressions in subjective emotional experience. Some theorists argue that the subjective experience of the primary emotions is produced by sensory feedback from facial muscles, although there is disagreement as to whether it matters if the expression is spontaneous and genuine as opposed to voluntary and faked. Although it seems reasonable to assume that behavioral activity, including that from the face, can intensify emotional experience, there is reason to doubt that peripheral motor feedback is necessary for subjective emotion.

The primary emotions are universal Although Ekman and Friesen were not the first to examine emotional behaviors cross-culturally, the methods they used in their excellent cross-cultural work on

the judgments of emotional expression provide strong evidence in support of primary emotions. This work was important because it undermined the social constructionist position that cultural factors (i.e., learned behavioral rules and cognition) governed the emergence of all emotional expression and subjective feelings. Rather than asking their participants, New Guinea natives who had no contact with Western culture and Californians who had no contact with New Guinea culture, to provide emotional labels for photographs of facial expressions, participants were asked to judge which of several photographed expressions fit a specific social situation. The New Guinea natives picked the smiling face of a Californian as the face that would resemble that of a valued relative being greeted at the door of their hut, and Californians made similar selections from the photos of New Guinea natives. Expressions were matched to the appropriate situation on both sides of the Pacific. Primary emotions seem to be universal.

Are primary emotions innate? Studies of the emergence of emotion in infants are a second source of evidence in support of the contribution of innate primary emotions. Words must be used with care in discussing infant emotion because we cannot question the 2-h-, 2-week-, or 2-month-old infant about his or her emotional state; indeed, the questions we can ask of the 2-year-old are also limited. We can, however, observe infant expressive behavior and infer emotional experience and emotional states, although with due caution. Studies conducted during the 1930s suggested that infants could display little more than contented smiles and undifferentiated states of distress. The failure to find differentiation appears to have been due to the coarseness of the stimuli used to elicit emotion (e.g., dropping and catching the infant after a 1- to 2-foot free fall). Studies conducted since then made clear that, although expressive behavior in infancy is less differentiated than that in adults, it is not simply a two-dimensional affair, that is, one dimension of sleep-active and another of smiles-distress. Indeed, infants less than 1 week old mimic the facial expressions of adults, such as smiles or looks of surprise or sadness, and respond with angry distress when the delivery of a sweet-tasting liquid is discontinued.

Although the data suggest the presence of innate patterns of facial expression for primary emotions, we can infer, but not prove, that these expressive patterns are associated with subjective feelings. This does not mean, however, that these states and/or their overt motor expressions are free of cognition. Mimicry of facial expressions is dependent on perceptual processes. The infant must perceive critical features of

the face if he or she is to mimic the adult's expression. The angry response to withdrawal of reward is contingent on the infant's acquiring the expectation that his or her sucking response will deliver a sweetened substance; learning is required for there to be a disconfirmed expectation. It is clear that the emotional expressions of infants are embedded in perceptual and cognitive processes, although the cognitive processes surrounding these behaviors are relatively primitive in the earliest days of life. Affects linked to such cognition (e.g., joy, sadness, disgust, anger, and surprise) emerge during the first 6 months of life. The cognitive context for emotional elicitation becomes more complex during the second half of the first year. Among the most interesting of these changes is the emergence of self-conscious emotions following the growth of awareness of an objective sense of self. Lewis and his colleagues have shown that affects such as empathy and envy appear only after the infant clearly recognizes him- or herself as a physical entity. The evidence for self-recognition is the infant's reaching toward a mark drawn on his or her forehead while looking in a mirror. An infant who cannot identify his or her mirror image cannot experience self-conscious emotions; emotional reactions are indeed embedded in cognition.

The Function and Structure of Emotion

Functions of Emotion

Findings showing that primary emotional motor patterns are embedded in perceptual and cognitive processes have important implications for the functions and structure of the psychological system producing emotional behavior. The communicative and motivational properties of emotion are two that stand out.

Emotional communication and social regulation

Darwin suggested that the function of most expressive behavior is to communicate action-readiness to conspecifics. Expressions communicate readiness to fight (anger), to flee (fear), to avoid a harmful substance (disgust), to withdraw (depression), to share pleasure (joy and love), and to take in, imbibe, and/or attend to an event (interest). Emotional communication is the glue that bonds infant to parent, and unchecked infant expressions of distress can ignite abuse. Despite the common view that emotions represent sources of irrationality and dysfunction, other theorists since Darwin also have emphasized the adaptive value and organizing function of emotion.

Ethologists observing animal behavior in the wild and anthropologists and sociologists observing

behavior in humans have identified the many ways in which emotional communication creates in-groups and out-groups, provides internal structure in organizations and face to face groups, and moderates the flow of interpersonal behavior in dyadic interaction. Fear vocalizations in chimpanzees communicate the presence of enemies to other members of the local clan, allowing all to escape predation. Sadness, with its depressive postures and self-abnegating cognition, and fear, with its associated flight, defend against attack from higher-status members of the tribe; aggressive anger establishes who is lord. The importance of emotion in regulating social behaviors is consistent both with the appearance of emotional expression early in infancy and with the intimate connections between expressive responses and cognitive processes. Most of the perceptual events that activate emotional reactions early in life are interpersonal, for example, infant and parental smiles, infant distress signaling parental assistance, and self-awareness as the basis for self-referential emotion. Affectively based commitments to institutions and lifetime partners in later life involve yet more elaborate integrations of multiple primary emotions with complex cognition. Emotional communication therefore requires the integration of affective and cognitive processes.

Emotional motivation Emotions have an energizing or activating role in behavior, and they seem to instigate movement toward or away from environmental stimuli. This function calls on systemwide resources for the support of overt action. Thus, neurotransmitters of the central nervous system that are linked to specific emotions are involved in complex chains of neurochemical activity for marshaling the sugars needed for energy expenditure. Emotions contribute more to motivation, however, than the mere energizing of behavior: They direct attention to specific cues, define particular goals for action (person to attack vs. person with whom to share resources), and shape coping procedures (i.e., thoughts and actions to meet these goals). This emphasizes once again the importance of the network of perceptual, cognitive, and motor processes associated with emotional behaviors.

The Structure of Emotion

A variety of frameworks have been suggested to integrate the array of issues we have covered. Those that seem best able to manage this task postulate a hierarchical mechanism underlying the creation of emotional activity. The primitive, innate primary expressions of emotion that are elicited by specific environmental stimuli form the base of the hierarchy.

The repeated elicitation of these primary emotions connects them with a wider range of eliciting cues (smiles of a particular parent) and with varied responses on the output side (reaching to be held). This history creates schematic structures that enrich the primary affects, connecting them to an increasingly wide range of stimuli and behavioral responses. Because schemata develop prior to the availability of a rich verbal system, much schema development involves nonverbal motor learning and this system remains automatic and subject to limited volitional control. A consequence is that control over emotional behavior is a central problem for managing life's adversities and temptations and is central to contemporary developments in psychotherapy.

Emotional life also becomes increasingly complex and mature through the integration of multiple primary affects in schematic structures. Reflective melancholy, pleasure linked to sadness about a valued but lost past, emerges from combinations of previously separate primary affects and their specific schematic structures. And just as cognitive growth and the awareness of self allow for the onset of emotions of shame and guilt in infancy, affective changes in later life and in response to chronic illness involve major cognitive shifts in the self system and alterations in the biological substrate underlying the primary affects at the base of the emotion hierarchy.

Biology, Emotion, Stress, and Health

Two final points merit brief comment. The first focuses on the nature of the biological systems involved in emotional behavior and the second on connections between emotion and health.

Biological Basis of Emotion

Panksepp marshaled an impressive array of neuroscience data in support of the position that specific neurobiological systems underlie the primary affects; the base for the primary emotions lies deep in the upper portion of the brain, not in the peripheral autonomic system. The chemistries of the body (e.g., epinephrine, dopamine, a vast array of peptides, and notably insulin, cortisol, and other agents involved in the storage and availability of sugars for energy expenditure) affect this central system because the brain and body are in bidirectional communication with one another. It is interesting to note that these chemistries regulate communication and social regulation even in single-cell bacteria. Insulin equivalents that transport sugars into the cell communicate the presence of food to conspecifics and regulate the behavior of the colony. This may be seen as a parallel to the action of the mammalian brain, which facilitates

intraspecies emotional communication by connecting emotional reactions to a broad cognitive base. A bacterium will not release a sufficient amount of this insulin-like factor to capture a nutrient unless other members of its local community also detect the nutrient. Communicating with one another precludes the premature use of this resource in the presence of a limited quantity of nutrient. Although these functions do not depend on the presence of a central nervous system, as in multicellular organisms, the action of such chemical systems may provide insight into the origins and function of complex emotion systems.

The neurobiological substrates for affective processes are currently the subject of intense investigation. Accumulating evidence has informed theory regarding the role of systems involving the prefrontal cortex, hippocampus, and amygdala in generating affective experience, emotional communication, and other behavioral expressions of affect and motivation. This work has also stimulated the rethinking of the distinction between cognition and emotion given that common as well as separate neurobiological structures have been implicated in both functions. In sum, the biological description of the emotion system is being linked to functional analyses in showing that emotions are critical sources of information for the regulation of resources, interpersonal behavior, and survival.

Emotion, Health, and Illness

The virtual elimination of life-threatening respiratory and gastrointestinal infections, the extension of life into the seventh and eighth decades, and the emphasis on autonomy and self-control in industrialized Western nations confront us with the chronic diseases of old age. One result has been the need for effective forms of self-management for prevention and control of chronic conditions and the revival of, it is hoped, life-improving folk medicine. Billions of dollars are spent annually on natural herbal products, mega-vitamins, special diets, stress-control programs, biofeedback, and so forth for the prevention and treatment of disease. The reduction of stress and avoidance of negative affects are culturally accepted, alternative routes for prevention and treatment. With the exceptions perhaps of cardiovascular disease and the common cold, however, there is scant evidence for such a direct, causal role of stress or negative emotion on disease, although the evidence for an indirect role is abundant (e.g., through the motivation of risk behaviors such as cigarette smoking and excessive alcohol consumption). Life-style behaviors related to emotional and motivational processes are critical not only for the prevention of chronic illness but also for their control. For example,

exercise, dietary control, and weight loss are nearly twice as effective as medication for preventing the onset of diabetes among individuals at very high risk. Despite the evidence, many people seem more concerned about reducing stress than quitting smoking, exercising, and reducing saturated fats in the diet, which may be far more likely to reduce the possibility of serious disease.

What lends credence to the commonsense belief that stress causes disease? Clearly stress makes us feel bad and stress reduction makes us feel better. Because physical illness and emotional states are bidirectionally related, fatigued and depressed mood often both precede and follow the onset of clinical illnesses such as the common cold. The immunological responses that defend us against disease have direct effects on the central nervous system; they cause fever and shivering, inhibit motor behavior and eating, and create fatigue and its associated feeling state of depressed mood. Thus, rather than seeing emotional states as directly causing disease, it might be closer to the truth to see emotional states as indicators of active disease processes and of reduced resistance to disease. Just as our thermometers tell us the temperature without causing the weather, so too might emotions indicate that something is amiss without directly causing illness.

Most important, emotions are responsive to a wide range of events. From a psychometric perspective, emotions are poor indicators of disease states; they may be sensitive indicators of illness, but they are nonspecific. Thus, if one is actively connected to one's social world, one will feel emotions of joy and distress, and distress may indicate the presence of chronic, life-threatening illness in elderly spouses, family members, and friends, and not just oneself. Because being connected enhances survival, distress caused by illness in a spouse or family member may be positively associated with function and good health. Because they are nonspecific, measures of emotional traits and states are poor predictors of health-relevant behaviors; for example, generalized anxiety is a poor predictor of management of diabetes complications. Prediction of self-management behaviors improves when affect is assessed in association with specific perceptions; for example, a fear of hypoglycemia predicts poor adherence to the use of insulin and elevated blood glucose, and fear of foot amputation predicts efforts at effective self-management. In short, an examination of the implications of emotion for physical health agrees with functional and neurobiological analyses in pointing to the interaction and integration of emotional process with perception, cognition, and contextualized action.

Concluding Comment

Both psychological and physiological accounts of emotional processes are advancing rapidly, the two types of description complementing one another and deepening our understanding of emotion. Neither alone is sufficient – physiological analysis does not inform us about emotion unless it can be related to a theoretically and empirically clear picture of the behavioral psychology of emotion. As both sides of the emotional equation are extended and elaborated, we can expect to deepen our understanding as to how these processes connect to our ecology, both social and physical, and to our genetic makeup. We hope that such knowledge will make us less susceptible to false nostrums and more accepting of life styles that maximize well-being and that help us to cope with and make use of the angers, fears, and depressions that are intrinsic to life.

Further Reading

- Berntson, G. G., Sarter, M. and Cacioppo, J. T. (2003). Ascending visceral regulation cortical affective information processing. *European Journal of Neuroscience* 18, 2103–2109.
- Cosmides, L. and Tooby, J. (2000). Evolutionary psychology and the emotions. In: Lewis, M. & Haviland, J. M. (eds.) *Handbook of emotions* (2nd edn.). New York: Guilford.
- Darwin, C. (1872). *The expression of the emotions in man and animals*. London: Murray.
- Ekman, P. (1982). *Emotion in the human face* (2nd edn.). New York: Cambridge University Press.
- Gray, J. A. (1990). Brain systems that mediate both emotion and cognition. *Cognition and Emotion* 4, 269–288.
- Izard, C. E. (1991). *The psychology of emotions*. New York: Plenum.
- James, W. (1950/1890). *The principles of psychology*. New York: Dover/Holt.
- Knowler, W. C., Barrett-Connor, E., Fowler, S. E., et al. (2002). Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *New England Journal of Medicine* 346, 393–403.
- Leventhal, H. (1984). A perceptual-motor theory of emotion. *Advances in experimental social psychology* 17, 117–182.
- Leventhal, H., Patrick-Miller, L., Leventhal, E. A., et al. (1997). Does stress-emotion cause illness in elderly people? In: Schaie, K. W. & Lawton, M. P. (eds.) *Annual review of gerontology and geriatrics. Vol. 17: Focus on emotion and adult development*, pp. 138–184. New York: Springer.
- Leventhal, H. and Patrick-Miller, L. (1993). Emotion and illness: the mind is in the body. In: Lewis, M. & Haviland, J. (eds.) *Handbook of emotion research*, pp. 365–379. New York: Guilford.
- Leventhal, H. and Scherer, K. R. (1987). The relationship of emotion to cognition: a functional approach to a semantic controversy. *Cognition & Emotion* 1, 3–28.
- Lewis, M. (1993). The emergence of human emotions. In: Lewis, M. & Haviland, J. M. (eds.) *Handbook of emotions*, pp. 223–235. New York: Guilford.
- Lewis, M. (1993). Self-conscious emotions: embarrassment, pride, shame and guilt. In: Lewis, M. & Haviland, J. M. (eds.) *Handbook of emotions*, pp. 563–573. New York: Guilford.
- Panksepp, J. (1998). *Affective neuroscience: the foundations of human and animal emotions*. New York: Oxford University Press.
- Plutchik, R. (1984). Emotions: a general psychoevolutionary theory. In: Scherer, K. & Ekman, P. (eds.) *Approaches to emotion*, pp. 197–219. Hillsdale, NJ: Lawrence Erlbaum.
- Scherer, K. R. (1988). *Criteria for emotion-antecedent appraisal*. Dordrecht: Kluwer Academic.
- Schachter, S. and Singer, J. E. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review* 69, 379–399.
- Teasdale, J. D. and Barnard, P. J. (1993). *Affect, cognition, and change: remodelling depressive thought*. Hove, UK: Lawrence Erlbaum.
- Tomkins, S. S. (1962). *Affect, imagery, consciousness. Vol. 1: The positive emotions*. New York: Springer-Verlag.
- Tomkins, S. S. (1963). *Affect, imagery, consciousness. Vol. 2: The negative emotions*. New York: Springer-Verlag.
- Woodworth, R. S. and Shlosberg, H. (1955). *Experimental psychology* (rev. edn.). New York: Henry Holt & Co.

Employee Assistance and Counseling

M E Mor Barak

University of Southern California School of Social Work
and Marshall School of Business, Los Angeles, CA, USA

D J Travis

University of Southern California School of Social Work,
Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

History

Rationale

EAPs and Stress

Program Characteristics

Challenges and Future Directions

Glossary

Counseling services

The provision of professional advice and guidance utilizing psychological methods or various techniques of the personal interview to guide an employee in a constructive direction that will lead to the development and realization of the individual's potential.

Employee assistance program (EAP)

Employer-sponsored counseling, educational, and mental health services that are available to all employees on a confidential and self- or supervisory referral basis.

In-house vs. external employee assistance program (EAP)

Services provided by counselors who are employees of the company vs. those provided by an external firm.

Work-family conflict

Incompatibility in some aspects of the simultaneous pressures from the work and family domains that makes it difficult for employees to meet the demands of one or both domains.

Employee assistance and counseling are resources utilized by work organizations to manage employee stress and enhance workplace effectiveness through preventing, identifying, and resolving personal and work-related problems. Employee assistance programs (EAPs), specifically, are employer-sponsored counseling, educational, and mental health services that are available to all employees on a confidential and self- or supervisory referral basis. These programs can offer assessments, referrals, and clinical, preventive, and educational services to employees and their families to aid in managing work and life challenges.

History

EAPs grew out of the 1940s occupational alcoholism prevention programs (OAPs) that focused on alcohol abuse in the workplace. In the 1970s, the scope of services was broadened to include mental health and drug abuse problems (often referred to as broad brush services), hence the development of EAPs. During that time, employers began adopting measures to assess the costs at the individual and organizational level related to alcohol and drug abuse as well as occupational stress and individual and family problems.

In the late 1980s, work organizations also began to consider the overall mental, physical, and emotional well-being of their employees by incorporating preventive mechanisms into the structure of traditional EAPs as well as by developing employee enhancement programs (EEPs). Preventive models focused on stress management and other addictions such as overeating and smoking.

EAPs now tackle employees' work and life challenges and overall well-being involving workplace stress and conflict, relationships and family problems, financial and legal problems, childcare, eldercare, and career counseling. A recent estimate of EAP utilization indicates that more than 80 million individuals within the United States have access to an EAP, an approximate 250% increase from only a decade ago.

Rationale

Recent research reveals that workers' job performance and workplace productivity can be affected by work-related stress, family conflicts, and personal challenges. As an illustration, depression costs \$44 billion annually in lost productivity, and substance abuse and mental illness cost \$100 billion annually, or \$3000 per employee. Hence, employee assistance and counseling provide supports to the employee and the employer, resulting in shared benefits including enhanced employee mental, emotional, and physical well-being, improved job performance, and improved organizational functioning (see [Figure 1](#)). Employees utilize EAPs and equivalent counseling programs as resources for managing work-related or family/personal stress. For the employer, EAPs help (1) contain health-care costs, (2) provide resources that supplement current organizational support services, (3) demonstrate care for their employees and enhance the work organization's image, (4) prevent or contain potential legal problems, and (5) address labor/organization relations.

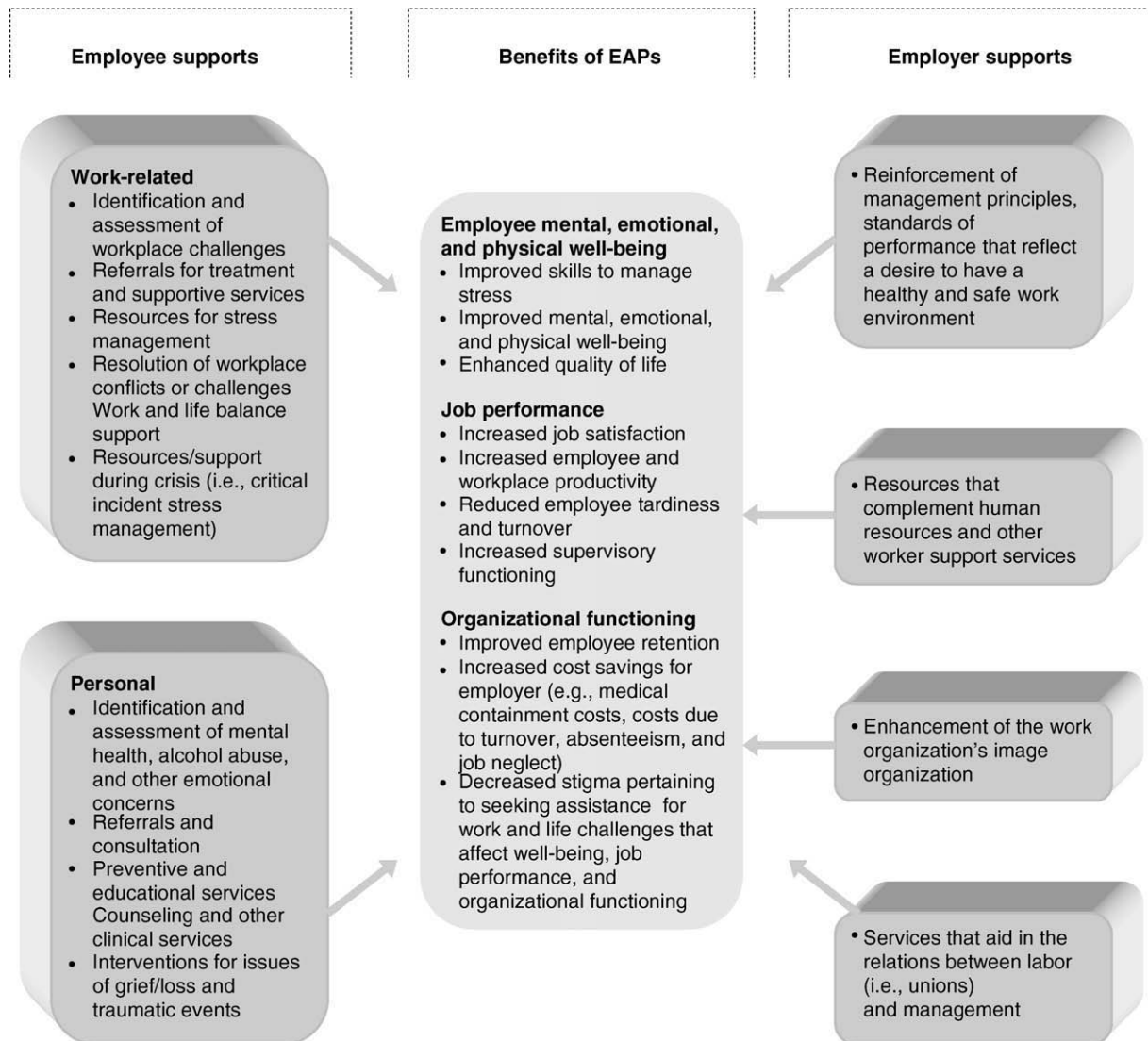


Figure 1 Value of employee assistance and counseling.

EAPs and Stress

EAPs and Job Stress

Employee assistance and counseling services are available to aid in managing stress that is associated with employees' work and life experiences. Counselors and therapists who provide those services (typically social workers, psychologists, and professional counselors) generally operate under the assumption that stress is not necessarily bad and that some stress is a normal part of work and life, but that excessive stress could be detrimental to both emotional and physical health. As such, employees can experience stress in the workplace in several areas, most notably in areas related to their job role and perception of social exclusion. These forms of work-related stress

can adversely affect worker outcomes, such as lost job productivity, counterproductive work behaviors, and preventable turnover.

EAPs often address stress on the job by offering individual counseling, stress reduction seminars, or counseling management on how to restructure the work environment or work processes to reduce stress. The experience of job stress is the psychological state that represents an imbalance between peoples' perceptions of the demands placed on them and their ability to cope with those demands. Typically, employees experience too much work (role overload), a conflict between their expectations of their jobs and the actual reality of those jobs (role conflict), and lack of clarity regarding their supervisors' or management's expectations from them (role ambiguity).

However, it is not only excessive demands that are the source of job stress but also understimulation, represented by work tasks that are boring, too simple, or not challenging.

Employee assistance services can also offer supportive resources to help employees negotiate perceived stress-related feelings of exclusion within the organization. The inclusion–exclusion continuum reflects the extent to which individuals feel part of important organizational processes that affect their jobs, have access to the organizational decision-making process, and have access to its information networks. Research over the past two decades indicates that exclusion from organizational information networks and from important decision-making processes is a significant problem facing today's diverse workforce. Perceptions of exclusion may play a critical role in explaining the connection between those employees who experience themselves as different from the corporate mainstream and their job stress, well-being, and other worker outcomes (e.g., job satisfaction, productivity, and turnover).

EAPs and Work–Family Life Stress

In the past two decades a new field of workplace service delivery, as well as a new field of scholarly inquiry, have emerged that focus on the stress associated with the need to balance work and family and, more generally, work and personal life. Many workers struggle to negotiate the demands of work with the expectations from their families and life outside of work, which include familial responsibilities such as child care and elder care, activities related to individuals' personal fulfillment, and activities related to community contributions and civic duty. Balancing or integrating the demands of work, family, and life can be associated with specific stress-related outcomes for workers, based on the presence and absence of conflict between their roles. The research on work–family and work–life balance has focused predominantly on the United States and Europe, but in recent years there has been more research on other countries and even some comparative studies that examine the similarity and differences in challenges faced by workers worldwide.

More and more work organizations are introducing work–family or work–life programs as part of, or in addition to, their EAP services. These services focus on the conflicts that employees experience in reconciling their work and family roles. They include individual and family counseling to help employees cope with the stresses associated with, for example, having to care for a sick parent or taking care of a new baby. In addition, work–family or work–life programs include new policies that allow employees to

take periods of paid leave for family or personal reasons. More recently, workplace flexibility options have emerged as a way to address conflicts between work and life outside of work, including flexibility in scheduling full-time hours (e.g., compressed work weeks), flexibility in the number of hours worked (e.g., reduced work hours per week or working part of the year), career flexibility with multiple points of entry, exit, and re-entry into the workforce, and the ability to address unexpected and ongoing personal and family needs.

Program Characteristics

Service Delivery Models

There are several different EAP service delivery models, described in the following sections.

Alcohol and drug abuse prevention and treatment vs. broad brush Alcohol and drug abuse prevention programs focus on providing referrals and counseling for employees who are addressing substance abuse problems that can impact the workplace. Broad brush programs essentially focus on an expansive realm of services that aid employees in managing their marriage and family, emotional, financial, and legal problems.

In-house EAP vs. contracted EAP services Some companies choose to provide EAP services internally, by counselors who are themselves employees of the organization, while other companies choose to contract the services from external EAP providers. The latter are firms that provide the entire EAP staff. The advantages of the in-house service delivery model include the counselors having a greater familiarity with the organization's structure and the ability to provide organizational-level interventions when needed (e.g., when the cause of a problem is at the work unit or department level). The advantages of the external EAPs is that they provide a perceived (though not necessarily real) greater level of confidentiality, and sometimes they are less expensive than the in-house services.

Consortium and affiliate models When organizations lack the number of employees (typically fewer than 2000) to financially justify having their own EAP, a consortium model offers an alternative. An EAP consortium is a mutual agreement among organizations that collaborate to maximize resources as opposed to contracting individually. Affiliate models occur when a company has several offices and contracts services; the contractor may use affiliates and subcontractors. The affiliates, in this model, are those

EAP providers based on a fee-for-service or case-by-case basis. The goal with the affiliate model is to ensure that all employees are covered as deemed necessary.

Types of Interventions

EAPs provide a range of services from traditional alcohol and drug abuse treatment to holistic interventions geared toward stress management, other addictions (e.g., eating disorders, smoking cessation), work-life balance, and general well-being. These services can include assessment and referral models (one or two meetings), short-term intervention models (in which the number of meetings may be limited to 6–12 meetings per year), and long-term models (in which the number of meetings is not limited and depends on the needs of the client). Many of the services can be conducted in conjunction with other services to effectively address the employee's problems. [Table 1](#) provides examples of such interventions and services.

Challenges and Future Directions

The development, management, and delivery of employee assistance and counseling services continue to evolve and reflect the challenges of the contemporary workplace. Cultural and economic shifts, which influence the types and prevalence of employee problems, will affect the future configuration and operation of EAPs. Ultimately, these types of contemporary issues may affect worker productively and performance.

Employers are faced with new challenges related to the increased diversity in the workforce (e.g., gender, age, race, ethnicity, religion, sexual orientation,

disability). This diverse workforce has the potential to be more creative and productive than a homogeneous one, but employers have to overcome potential problems related to distrust and miscommunication that are often associated with diversity. EAPs can help by offering organizational diversity assessment and training and team dispute resolution.

The exponential growth in global work arrangements such as international mergers, subsidiaries, call centers, and subcontracting in the past few decades has introduced yet another set of challenges: how do families and communities cope with these new stressful conditions? How do they adjust to the pressures presented by the 24/7 economy? How do individuals cope with the blurring boundaries between work and private time? EAPs will be challenged with expanding their services to adjust to the demands of rapidly changing family and work structures.

Finally, the future of EAPs will also be affected by two conflicting trends in health and mental health care. One trend is motivated by the urgent need to contain the ever-rising health-care costs through contracting out services and using managed care systems that work diligently to contain costs by limiting types of services and supervising the service provision process. The other trend advocates a holistic health-care approach that focuses on treating the whole person by attending to the individual's mental, emotional, and physical well-being. The latter may require more in-house services and, in the short term, may be more costly, though in the long term its preventive aspects may contribute to cost containment. The resolution of these two trends may affect the scope and methods of service delivery that EAPs will provide in the future.

Table 1 EAP interventions and services

<i>Scope</i>	<i>Examples of EAP interventions and services</i>
Assessment	Diagnosis of presented problems and treatment plan
Informational	Referrals to appropriate services outside of the EAP
	Consultation and suggestions of solutions
Therapeutic	Counseling
	Therapy (short-term or long-term and family, couple, or individual)
	Clinical services
Preventive	Stress management
	Health and fitness programs
Educational	Training
	Personal development
	Lunchtime seminars
	Newsletters, pamphlets, brochures
Crisis intervention	Critical incident stress debriefings
	Grief/loss/traumatic events
Organizational interventions	Trainings for work groups and leaders/supervisors
	Team building

See Also the Following Articles

Job Insecurity: The Health Effects of a Psychosocial Work Stressor; Work–Family Balance; Workplace Stress.

Further Reading

- Akabas, S. H. and Kurzman, P. A. (2005). *Work and the workplace*. New York: Columbia Press.
- Allen, T. D., Herst, D. E. L., Bruck, C. S. and Sutton, M. (2000). Consequences associated with work-to-family conflict: a review and agenda for future research. *Journal of Occupational Health Psychology* 5, 278–308.
- Arthur, A. R. (2000). Employee assistance programmes: the emperor's new clothes of stress management? *British Journal of Guidance and Counseling* 28(4), 549–559.
- Berridge, J., Cooper, C. I. and Highley-Marchington, C. (1997). *Employee assistance programmes and workplace counseling*. Chichester, UK: Wiley.
- de Croon, E. M., Sluiter, J. K., Blonk, R. W., Broersen, J. P. and Frings-Dresen, M. H. (2004). Stressful work, psychological job strain, and turnover: a 2-year prospective cohort study of truck drivers. *Journal of Applied Psychology* 98(3), 442–454.
- Emmer, W. G., Hutchison, J. and Richards, M. A. (eds.) (2003). *Employee assistance programs: wellness enhancement programming* (3rd edn.) Springfield, IL: Charles C. Thomas Publisher, LTD.
- Kossek, E. E. and Lambert, S. J. (eds.) (2005). *Work and life integration: organizational, cultural, and individual perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Mor Barak, M. E. (2005). *Managing diversity: toward a globally inclusive workplace*. Thousand Oaks, CA: Sage.
- Mor Barak, M. E. and Bargal, D. (eds.) (2000). *Social services in the workplace*. New York: Haworth.
- Mor Barak, M. E. and Cherin, D. A. (1998). A tool to expand organizational understanding of workforce diversity. *Administration in Social Work* 22(1), 47–65.
- Mor Barak, M. E. and Levin, A. (2002). Outside of the corporate mainstream and excluded from the work community: a study of diversity, job satisfaction and well-being. *Community, Work & Family* 5(2), 133–157.
- Mor Barak, M. E., Nissly, J. A. and Levin, A. (2001). Antecedents to retention and turnover among child welfare, social work, and other human service employees: what can we learn from past research? A review and meta-analysis. *Social Service Review* 625–661.
- Spector, P. E., Cooper, C. L., Poelmans, S., Allen, T. D., O'Driscoll, M., Sanchez, J. I., et al. (2004). A cross-national comparative study of work-family stressors, working hours, and well-being: China and Latin America versus the Anglo world. *Personnel Psychology* 57, 119–142.
- Stewart, W. F., Ricci, J. A., Chee, E., et al. (2003). Cost of lost productive work time among US workers with depression. *Journal of the American Medical Association* 289(23), 3135–3144.
- Williams, E. S., Konrad, T. R., Scheckler, W. E., Pathman, D. R., et al. (2001). Understanding physicians' intentions to withdraw from practice: the role of job satisfaction, job stress, mental and physical health. *Health Care Management Review* 26(1), 7.

Endocrine Systems

G Fink

Mental Health Research Institute, Melbourne, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G Fink, volume 2, pp 47–48, © 2000, Elsevier Inc.

The endocrine system is one of three control systems of the body, along with the nervous and immune/inflammatory systems. The classical definition of endocrinology was the discipline that dealt with hormones and the glands and target tissues involved in hormone secretion and action. Hormones were traditionally defined as chemical compounds synthesized and secreted by glands and transported by the bloodstream to their target organs. In this respect, a

hormone (from the Greek, meaning to excite) differs from a neurotransmitter, which is a chemical substance released from nerve terminals at synapses or neuromuscular junctions and which exerts its effects at postsynaptic or postjunctional receptors without entering the bloodstream. However, as illustrated extensively by the hypothalamic-pituitary-adrenal axis, some nerve cells also liberate their neurotransmitters into the bloodstream. Neurotransmitters released from nerve cells at neurohemal junctions are termed neurohormones (see **Neuroendocrine Systems**).

The endocrine system is under the control of the nervous system by way of hypothalamic control of the pituitary gland. In turn, hormones exert potent effects on brain differentiation and function. The immune and inflammatory system, too, is under endocrine as well as nervous control. In turn, leukocytes

and macrophages secrete factors termed cytokines (which include the interleukins, interferons, and tumor necrosis factor) that exert effects on the brain and the endocrine system.

The same chemical compound can have several different functions, serving as a neurotransmitter in the nervous system as well as a hormone through its transport by the circulation. It has long been accepted that the evolution of hormones is predominantly about their functions, which depend upon tissue-selective expression of hormone-specific receptors. Many hormones, neurohormones, and neurotransmitters exert local or paracrine functions – that is, they reach their target cells by regional diffusion to nearby cells. The recognition of the different mechanisms by which a molecule might act in the body (i.e., hormone, neurohormone, neurotransmitter, paracrine transmitter) and

the complex interaction between the three control systems mediated by chemical signaling have led to an explosion of knowledge in the past 100 years. This biomedical renaissance has been made even more exciting by the precision added through molecular genetic analysis of endocrine and other signaling mechanisms.

See Also the Following Article

Neuroendocrine Systems.

Further Reading

Reed Larsen, P., Kronenberg, H. M., Melmed, S. and Polonsky, K. S. (eds.) (2003). *Williams textbook of endocrinology* (10th edn.). Philadelphia, PA: WB Saunders.

Endometriosis

S N Kalantaridou and T Vrekoussis

University of Ioannina Medical School, Ioannina, Greece

A Makrigiannakis

University of Crete Medical School, Heraklion, Greece

G P Chrousos

University of Athens Medical School, Athens, Greece

© 2007 Elsevier Inc. All rights reserved.

Etiology

Endometriosis and Pelvic Pain

Endometriosis and Infertility

Endometriosis, Immune Function and Stress Response

Peripheral Corticotropin Releasing Hormone and Endometriosis

Diagnosis and Treatment of Endometriosis

Conclusion

Glossary

<i>Dyspareunia</i>	Pelvic pain during intercourse.
<i>Dysmenorrhea</i>	Excessive menstrual pain.
<i>Endometriosis</i>	An estrogen-dependent disease characterized by the presence of endometrial tissue in ectopic sites outside the uterus.
<i>Fibromyalgia</i>	Chronic widespread pain involving all four quadrants of the body and the presence of 11 of 18 tender points on examination.
<i>Interleukins (ILs)</i>	A class of peptides with immune actions. IL-15 and IL-18 activate macrophages

Natural killer (NK) cells

T helper (Th) cells

and are responsible for cell-mediated immunity and phagocyte-dependent protective responses. IL-4, IL-5, IL-6, IL-10, and IL-13 (together with granulocyte-macrophage colony stimulating factor), produced by T helper 2 cells, are responsible for antibody production, eosinophil activation, and inhibition of several macrophage functions.

Large granular lymphocytes with cytolytic and cytokine/chemokine production activities. NK cells are involved in the control of microbial infections and tumor development. The role of NK cells in autoimmunity or in inflammatory conditions is unclear.

CD4+ T lymphocytes that, by producing cytokines, play a key role in regulating immune response. Naïve Th cells can be differentiated into two specialized subsets: Th1 and Th2. Th1 cells produce cytokines, such as interferon (IFN)- γ , tumor necrosis factor (TNF)- α , and IL-1, IL-2, and IL-12. Th2 cells produce IL-4, IL-5, IL-6, IL-10, IL-13, and granulocyte-macrophage colony stimulating factor.

Endometriosis, the presence of a functional endometrium outside the uterine cavity, is a common chronic gynecological disorder occurring in 4–10% of women of reproductive age. Endometriosis is a recurrent

disease that may significantly impair the patient's quality of life. It is characterized by the presence of uterine endometrial tissue on the pelvic peritoneum, the ovaries, and the rectovaginal septum and more rarely in the pericardium, pleura, and even the brain. In the United States, endometriosis is the third most common gynecological disease that requires hospitalization and a leading cause of hysterectomy.

Although endometriosis was described over 300 years ago, its pathogenesis remains an enigma. At the time of clinical presentation, most women already have established disease and, therefore, it is very difficult to investigate the early developmental stages of the disorder. Ectopic endometrial tissue growth and inflammation are two pathophysiological features responsible for endometriosis-associated pelvic pain and infertility. Ectopic endometrial tissue responds to cyclic changes in gonadal hormones by proliferation and differentiation and by the local production of autocrine and paracrine factors. Therefore, endometriosis is an estrogen-dependent condition, with lesions undergoing regression during states of cessation of ovarian function, such as hypogonadal amenorrhea and menopause. Ectopic endometrium, however, appears to behave differently from its eutopic counterpart. Thus, endometriotic lesions are characterized by high local aromatization to estradiol biosynthesis and low estradiol inactivation compared with eutopic endometrium. Using laser-capture microdissection and a cDNA microarray with 9600 genes/expressed sequence tags (ESTs), a recent study showed that 904 genes/ESTs are differentially expressed in ectopic and eutopic endometrial tissue.

Pelvic endometriosis can be subdivided into three distinct entities: superficial peritoneal (and ovarian) endometriosis, cystic ovarian endometriosis, and deeply infiltrating endometriosis. Women with endometriosis and pelvic pain may also have chronic generalized pain and fatigue (i.e., fibromyalgia and/or chronic fatigue syndrome), as well as atopic and autoimmune disorders. It has been suggested that endometriosis may be associated with increased risk of ovarian cancer, non-Hodgkin lymphoma, and breast cancer. However, further epidemiological studies are necessary for the investigation of the possible association between endometriosis and cancer.

Etiology

The etiology of endometriosis is unclear; the most widely accepted hypothesis for its development is retrograde menstruation into the peritoneal cavity via the fallopian tubes and the subsequent attachment of endometrial cells to the peritoneal/ovarian surface. An alternative explanation is that the endometriosis

implants result from *in situ* metaplasia. Whichever mechanism is responsible, there are likely to be additional factors that determine whether a woman develops endometriosis. Indeed, susceptibility to endometriosis is thought to depend on the interaction of genetic, immunological, hormonal, and environmental factors. Therefore, it appears that endometriosis is a multifactorial genetic disorder, inherited as a complex genetic trait in which multiple genes conferring disease predisposition (including cancer-susceptibility genes and genes coding for cytochrome 450 enzymes, nuclear receptors, and immunological mediators) interact with one another and the environment to facilitate development of the disease.

Endometriosis and Pelvic Pain

Pelvic pain associated with endometriosis usually is cyclical; however, the pain may become continuous as the disease progresses. Indeed, it is common for any type of regional pain syndrome to eventually spread to become more systemic and involve the entire body (i.e., transform into fibromyalgia).

Little is known about the association between the endometriotic lesions and pain. Medical therapies aimed at the downregulation of ovarian function result in the elimination or reduction in size of endometriotic implants and the reduction of endometriosis-related pain in women. The surgical removal of the ectopic implants may alleviate pain symptoms in some women; however, the surgery may also fail to alleviate the pain.

It has been reported that pain is greater in women with deeply infiltrating endometriotic lesions in highly innervated areas than in women with more superficial lesions. Also, it has been suggested that the nerve fibers are closer to the endometriotic implants in women with pelvic pain than in women without pain, implicating the nervous system in endometriosis-related pain.

Endometriosis and Infertility

The mechanisms of endometriosis-associated infertility are not well understood. In advanced-stage endometriosis, pelvic adhesions distort normal pelvic anatomy and impair tubo-ovarian function. However, it seems that even mild disease may negatively affect oocyte quality and implantation. A meta-analysis of assisted reproduction studies showed that the pregnancy rate in women with endometriosis is approximately one-half that in women with tubal factor infertility. In women with endometriosis undergoing *in vitro* fertilization, the presence of auto-antibodies is associated with significantly lower pregnancy rates.

Endometriosis, Immune Function and Stress Response

There is increasing evidence that immunological factors play an important role in the pathogenesis of endometriosis. Women with endometriosis have an altered function of the peritoneal macrophages, natural killer cells, and lymphocytes, as well as changes in growth factors and inflammatory mediators in the peritoneal fluid.

Endometriosis is associated with the suppression of cell-mediated immunity and the stimulation of humoral immunity. Impaired natural killer cell activity resulting in the inadequate removal of refluxed menstrual debris may have a role in the development and maintenance of endometriotic implants. In endometriosis, macrophages are found in increased numbers in the peritoneal fluid, and although they would be expected to facilitate ectopic endometrial cell clearance, they appear in fact to support endometriotic lesions by producing growth factors and pro-inflammatory cytokines. These cytokines – interleukin (IL)-1, IL-8, tumor necrosis factor (TNF)- α , and interferon (INF)- γ – may in turn activate chemotactic factors, inducing infiltration of the peritoneum by T lymphocytes and macrophages. The immune mediators produced by these cells may promote the implantation and growth of ectopic endometrium by inducing its proliferation and angiogenesis. It is thus possible that in endometriosis a vicious circle is activated, with ectopic endometrial lesions causing inflammation and inflammatory mediators stimulating endometrial lesion growth.

Interestingly, decreased cellular immunity along with increased humoral immunity is also seen in patients with chronic fatigue syndrome and fibromyalgia, in people in other chronic stressful conditions, such as students taking examinations, prisoners from concentration camps, and spouses of Alzheimer's patients, and in animal studies of chronic stress (Table 1). Interestingly, a study of a symptomatic woman with endometriosis and her family, using

electroencephalography, digital skin temperature, electrodermal response, and electromyography, showed that endometriosis symptoms were associated with prolonged stress reactions. The study revealed that facts of family history and relationships for three generations set the stage for such stress reactions. The relation between endometriosis and stress is intriguing and has been poorly investigated. It is tempting to speculate that endometriosis may be associated with dysregulation of the stress system.

That endometriosis may have an autoimmune basis, associated with stress response dysregulation, is further supported by the finding that atopic diseases, such as allergies, asthma, and eczema are more frequent in women suffering from endometriosis. In addition, a higher incidence of autoimmune inflammatory disorders (rheumatoid arthritis, systemic lupus erythematosus, Sjögren's syndrome, multiple sclerosis, and autoimmune thyroid disease) is found in women with endometriosis. Of note, women with endometriosis are more likely to have allergies, asthma, and eczema if they also have fibromyalgia and/or chronic fatigue syndrome. In addition, family members of women with endometriosis have autoimmune inflammatory diseases, endocrine diseases, and chronic fatigue states, indicating that this spectrum of disorders may share a common genetic predisposition and underlying pathogenic mechanisms.

Peripheral Corticotropin Releasing Hormone and Endometriosis

In endometriosis, the implantation of ectopic endometrial tissue is associated with a local inflammatory reaction, including macrophage activation and the elevation of cytokines and growth factors. Corticotropin releasing hormone (CRH), the principal regulator of the hypothalamic-pituitary-adrenal axis, as well as its receptors, have been identified in various systems, such as the immune and female and male reproductive systems. Indeed, CRH-like immunoreactivity has been found in peripheral inflammatory sites and in a number of reproductive organs, including the ovaries and the endometrial glands. Therefore, immune and reproductive CRHs are forms of CRH found in peripheral tissues. Immune CRH plays a direct immunomodulatory role as an autocrine/paracrine mediator of inflammation. One of the early effects of immune CRH is the degranulation of mast cells and the release of histamine and several inflammatory cytokines. High numbers of activated mast cells are present in endometriosis sites, which are also strongly positive for CRH. Further studies are needed to clarify the possible role of peripheral CRH in the pathogenesis of endometriosis.

Table 1 Common immune changes in endometriosis, fibromyalgia, chronic fatigue syndrome, and other chronic stressful conditions^a

Diminished cellular immune responses	Decreased NK cell activity Decreased response to mitogenic activity Decreased Th1 activity
Increased humoral immune responses	Increased immunoglobulin production Increased Th2 activity

^aNK, natural killer; Th, T helper.

Diagnosis and Treatment of Endometriosis

The symptoms of endometriosis include dysmenorrhea, dyspareunia, pelvic pain, and infertility. Adhesions, which are often associated with endometriosis, can play a role in painful symptoms. On the other hand, endometriosis is detected in 2–50% of cases during laparoscopy in women with no symptoms.

The diagnosis of endometriosis is usually made by direct visualization of the pelvis during diagnostic laparoscopy because noninvasive diagnostic imaging techniques have a relatively low sensitivity. For this reason, the precise frequency of endometriosis in the general population is unknown, given that an unknown percentage of affected women are asymptomatic. The reported surgically diagnosed annual incidence is 1.3–1.6 per 1000 women ages 15–49.

The therapeutic strategy for endometriosis must be determined for each patient individually. The principal goal in treating endometriosis is the relief of symptoms, the amelioration of infertility, and the prevention or delay of disease progression. For women with pain, surgery usually provides temporary relief; symptoms recur in up to 75% of women within 2 years. Therefore, many women require repeated medical and surgical therapy to control the symptoms.

Pharmacological treatment options include non-steroidal anti-inflammatory agents and agents that suppress ovarian function and limit the growth and activity of endometriosis, such as gonadotropin-releasing (GnRH) analogs, androgenic/progestagen agents (danazol), combined oral contraceptives, and progestagens. However, currently available medical therapies are unsatisfactory, and they cannot be used for prolonged periods of time because of severe secondary side effects.

There is a need to develop new drugs to provide more efficient therapeutic alternatives that eliminate endometriotic lesions and prevent recurrences. Under investigation are new agents, such as aromatase inhibitors, selective estrogen receptor modulators, and progesterone receptor modulators.

The medical treatment for pain is not useful for infertility management. Infertility management includes controlled ovarian hyperstimulation and intra-uterine insemination or *in vitro* fertilization and embryo transfer.

Conclusion

Endometriosis is a common, estrogen-dependent, chronic gynecological disorder. It is characterized by the presence of uterine endometrial tissue outside

the uterine cavity. Endometriosis may present as superficial and/or deep pelvic peritoneal implants, adhesions, and ovarian cysts (endometriomas). The symptoms of endometriosis include pelvic pain and infertility. Affected women are at higher risk than the general population for developing fibromyalgia, chronic fatigue syndrome, autoimmune inflammatory diseases, and atopic diseases. The association between endometriosis and stress is an intriguing but poorly investigated issue. However, it is possible that endometriosis may be associated with the dysregulation of the stress system. Endometriosis requires multidisciplinary care and long-term follow-ups for the surveillance of associated disorders that may develop in susceptible women.

See Also the Following Articles

Pain; Reproduction, Effects of Social Stress on; Reproductive Dysfunction in Primates, Behaviorally Induced; Fibromyalgia; Stress Effect of Assisted Reproduction.

Further Reading

- Barnhart, K. T., Dunsmoor-Su, R. and Coutifaris, C. (2002). Effect of endometriosis on in-vitro fertilization. *Fertility and Sterility* 77, 1148–1155.
- Berkley, K. J., Rapkin, A. J. and Papka, R. E. (2005). The pains of endometriosis. *Science* 308, 1587–1589.
- Blumenthal, R. D., Samoszuk, M., Taylor, A. P., et al. (2000). Degranulating eosinophils in human endometriosis. *American Journal of Pathology* 156, 1581–1588.
- Brinton, L. A., Gridley, G., Pesson, I., et al. (1997). A cancer risk after a hospital discharge of endometriosis. *American Journal of Obstetrics and Gynecology* 176, 572–579.
- Clauw, D. J. and Chrousos, G. P. (1998). Chronic pain and fatigue syndromes: overlapping clinical and neuroendocrine features and potential pathogenic mechanisms. *Neuroimmunomodulation* 4, 134–153.
- Chrousos, G. P. and Gold, P. W. (1992). The concepts of stress and stress system disorders. *Journal of the American Medical Association* 267, 1244–1252.
- Giudice, L. C. and Kao, L. C. (2004). Endometriosis. *Lancet* 364, 1789–1799.
- Harrison, V., Rowan, K. and Mathias, J. (2005). Stress reactivity and family relationships in the development and treatment of endometriosis. *Fertility and Sterility* 83, 857–864.
- Kalantaridou, S. N., Makrigiannakis, A., Zoumakis, E., et al. (2004). Stress and the female reproductive system. *Journal of Reproductive Immunology* 62, 61–68.
- Kempuraj, D., Papadopoulou, N., Stanford, E. J., et al. (2004). Increased numbers of activated mast cells in endometriosis lesions positive for corticotropin-releasing hormone and urocortin. *American Journal of Reproductive Immunology* 52, 267–275.

Missmer, S. A. and Cramer, D. W. (2003). The epidemiology of endometriosis. *Obstetrics and Gynecology Clinics of North America* 30, 1–19.

Seli, E. and Arici, A. (2003). Endometriosis: interaction of immune and endocrine systems. *Seminars in Reproductive Medicine* 21, 135–144.

Sinaii, N., Cleary, S. D., Ballweg, M. L., et al. (2002). High rates of autoimmune and endocrine disorders,

fibromyalgia, chronic fatigue syndrome and atopic diseases among women with endometriosis: a survey analysis. *Human Reproduction* 17, 2715–2724.

Wu, Y., Kajdacsy-Balla, A., Strawn, E., et al. (2006). Transcriptional characterizations of differences between eutopic and ectopic endometrium. *Endocrinology* 147, 232–246.

Endorphin See: Beta-Endorphin.

Enkephalins See: Opioids.

Enuresis

S K Anand and C D Berkowitz

Harbor-UCLA Medical Center, Torrance, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S K Anand and C D Berkowitz, volume 2, pp 49–52, © 2000, Elsevier Inc.

during daytime while awake (10%), or both. The focus of this article is on nocturnal enuresis; diurnal enuresis is only minimally included. Although enuresis is a benign condition, many school-age children who experience bed wetting find it socially embarrassing and stressful. In some children, it may also lead to negative self-esteem.

Prevalence and Epidemiology

Etiology

Evaluation

Management

Prognosis

Glossary

Enuresis

Involuntary passage of urine in bed or clothes in children whose age and development would suggest achievement of bladder control.

Enuresis refers to involuntary passage of urine in bed or clothes in children whose age and development would suggest achievement of bladder control. It may be nocturnal, only during sleep (80%), diurnal,

Prevalence and Epidemiology

At age 6 years, 10–15% children wet the bed at least once monthly, with some wetting nightly. Beyond 6 years there is a spontaneous resolution rate of approximately 15% per year. In adolescents older than 15 years and in adults, the prevalence of enuresis is reported to be between 1 and 2%. The disorder is more common in boys than girls. Enuresis is prevalent in all areas of the world with similar frequency. The disorder is termed primary enuresis in 80% of enuretic children if they have never achieved sustained dryness, and secondary enuresis in the remainder when wetting recurs after 6 months of dryness. There is a strong genetic predisposition for enuresis. The probability of enuresis is nearly 75% if both parents are enuretic, and 45% if one parent is affected. The concordance for identical twins is 68% and for fraternal

twins, 36%. Enuresis is more common in children with low IQ scores, attention deficit disorder, fecal incontinence, low socioeconomic status, large family size, or single-parent families, and in institutionalized children.

Etiology

The etiology of nocturnal enuresis is unknown except for the genetic predisposition described previously. Most children with primary enuresis are biologically and psychologically normal except that during sleep they are unable to recognize the sensation of a full bladder and fail to awaken to urinate in the toilet. Several factors have been reported to contribute to enuresis; these include delayed maturation of physiological control, urodynamic changes, sleep disturbance, alterations in nocturnal vasopressin secretion, psychological factors, and organic medical or urological disorders.

Delayed Maturation

Until the age of 5 years, variability or delay in maturation of urinary control is often considered the basis of enuresis. Many experts believe that beyond age 5 years, however, this is an unsatisfactory explanation, especially if a child can achieve daytime control. Nevertheless, others believe that a disorder that affects 10–15% of otherwise normal children and that has no known long-term medical consequences except social embarrassment must be considered in the realm of variability or delay in micturition maturation.

Urodynamic Changes

In the majority of enuretic children, the actual bladder capacity is normal; however, the functional bladder capacity (volume at which bladder empties itself during the night) is reduced in some. Many children with diurnal enuresis have unstable bladder (due to involuntary bladder contractions), but bladder instability is not an important factor in nocturnal enuresis.

Sleep Disturbance

The fact that enuresis occurs during sleep suggests that a disturbance in normal sleep may cause enuresis. Also, parents of enuretic children often report that the child sleeps too deeply and is difficult to awaken during the night. Some studies have reported that enuretic children have had impaired arousal to loud sounds during sleep. However, objective findings in most studies demonstrate normal sleep pattern and do not document that either the children sleep too deeply or are more difficult to awaken.

Nocturnal Vasopressin Deficiency/Polyuria

Normal children have an increase in vasopressin (AVP) secretion during the night, which results in decreased urine volume. In some children with enuresis, AVP secretion during the night does not increase, thus leading to larger volume of urine and enuresis. Desmopressin (DDAVP) may be especially useful in treating such children.

Psychological Factors

During acute stressful states such as the loss of a parent or addition of a sibling, transient recurrence of enuresis is often observed. The vast majority of children with nocturnal enuresis, however, as stated earlier, are psychologically normal. The occasional emotional problems noted in such children are generally due to the consequences of enuresis (e.g., social embarrassment or punitive measures to control enuresis) rather than the cause of enuresis. Once enuresis is cured, these emotional problems and the child's self-esteem generally improve and do not deteriorate or result in substitute symptoms as might be expected if enuresis were the primary symptom of an emotional disorder.

Organic Medical or Urological Disorders

Organic medical or urological disorders are the cause of enuresis in less than 5% of affected children and include urinary tract infection, diabetes mellitus, diabetes insipidus, sickle cell anemia, cystic kidney disease, and other urological disorders. Even enuretic patients with these disorders eventually gain nocturnal bladder control by adolescence.

Miscellaneous Causes

Enuresis in children and adults may accompany obstructive sleep apnea; resolution usually follows relief of the obstruction. Caffeine intake is sometimes associated with enuresis in adults and should be reduced or stopped before initiating other treatments.

Evaluation

A thorough history should be obtained, especially for genetic predisposition and medical or urological disorders (including dysfunctional voiding) that may result in enuresis. History should be sought for previous efforts to stop enuresis and reasons for consultation at this point in time (e.g., ruling out serious medical problems, child would not sleep overnight at a friend's house). Finally, the motivation of the child and family to implement various measures to stop enuresis should be assessed.

A complete physical examination should be performed, including growth parameters and blood pressure. Abdomen should be examined for organomegaly, bladder distention, and fecal impaction. Genitalia should be examined for any genitourinary abnormalities such as labial adhesion. Rectal examination may be needed to evaluate rectal tone. The lower back should be examined for dimples, sinuses, a tuft of hair, or any spinal abnormalities. If necessary, the child's neurodevelopment should be assessed.

Urinalysis of a freshly voided, first morning specimen is the only laboratory test usually necessary; the specimen should be assessed for specific gravity, protein, blood, and abnormal cellular elements. Blood tests (e.g., for renal function, electrolytes) are usually unnecessary. Similarly most children do not require any radiological studies. In children with diurnal enuresis, ultrasound of the kidneys and bladder as well as cystometry may be sometimes indicated. Children with urinary tract infection as well as other neurological or medical disorders may require special studies.

Management

Treatment begins with the reassurance and education that enuresis is not a serious renal or urological problem, that it is not the child's fault or due to laziness, and that eventually, even untreated, the disorder will resolve in the vast majority of patients. The rationale for treatment is that cure of enuresis is often accompanied by improvement in child's self-esteem. The two main methods of management are based on behavior modification (i.e., conditioning therapy) or drug treatment, and these modalities may be used in conjunction with each other. Encouraging the child to take responsibility for solving the problem is essential for success. This implies the need for motivation on the part of the child to learn to get out of bed and go to the toilet in the middle of the night, as well as a commitment to help with laundry and changing sheets. A reward system, like using a star chart, may be useful but must be used in addition to other interventions. Fluid restriction alone is ineffective; nevertheless, limiting fluid intake for 2 h before bedtime, urinating before going to bed, and improving access to a toilet are sensible, useful measures. Although various management plans discussed previously and in further detail in the following sections have been shown to be effective in a reproducible fashion, evaluation of various measures as a recommendation for any individual child has been difficult because there is a spontaneous cure rate of 15% every year. In addition, in many children, there

is cessation of enuresis after the first physician visit and reassurance.

Behavior Modification, i.e., Conditioning Therapy

The principle behind behavior therapy is that the child must get up during the night and use the toilet, before or as soon as urination starts. For behavior therapy to succeed, the child and parents must be motivated to implement the techniques discussed here; otherwise, the therapy is doomed to failure.

Self-awakening or parent-awakening program The technique of self-awakening consists of rehearsing a sequence of steps: (1) the child is asleep and it is the middle of night; (2) the child perceives that the bladder is full and that he or she needs to go to bathroom; (3) the child goes to the bathroom; (4) the child sits on the toilet and urinates; (5) the child goes back to bed and falls asleep. The steps can be rehearsed by getting in the bed sometime before regular sleep hours or even during the day. The parent-awakening program is indicated if self-awakening does not succeed; this technique still requires the child's participation. The parent's role is in awakening the child at a set hour. The child still must complete steps 2 to 5. This technique is most useful in the adolescent and older elementary school children.

Enuresis alarms Enuresis alarms are used as an adjunct to self-awakening. They are triggered by wetness, when only a few drops of urine have been voided. The child must inhibit further urination, get up, and complete steps 3 to 5 described in the preceding section. With practice, bladder distention does not lead to an immediate urge to urinate. This gives the child time to complete the steps. Newer versions of alarms that are transistorized and/or vibrate instead of giving a sound signal are available at relatively modest prices.

Conditioning therapy including self-awakening and alarm systems has the highest likelihood of long-term success in treating enuresis, with rates reported to be as high as 50–80%. Acupuncture, hypnosis, and psychotherapy have been reported to be of benefit, but there are no good trials demonstrating their effectiveness.

Drug Therapy

Pharmacological therapy for nocturnal enuresis includes tricyclic antidepressants (TCA) and DDAVP. Anticholinergics such as oxybutynin are more effective for diurnal enuresis than for nocturnal enuresis. Drug therapy is generally safe and well tolerated, but

like all medications has potential side effects and complications.

TCA Imipramine has been the principal TCA used for enuresis, but other TCAs are also effective. The mechanism of action is unknown. The starting dose of imipramine is 25 mg (which may be increased to a maximum of 50 mg for 8- to 12-year-olds and 75 mg for >12 years old) 1 h before bed time. A trial period of 2 to 4 weeks is adequate to assess effectiveness. The initial success rate is about 40–60% with a relapse rate of 50–70% upon discontinuation.

DDAVP DDAVP, a synthetic vasopressin analog (1-(3-mercaptopropionic acid)-8-D-arginine vasopressin monoacetate (salt) trihydrate), works by reducing the amount of urinary output during the night. The overall effectiveness is about 25–33% (range of 12–86%) for children becoming totally dry for 14 or more consecutive nights. DDAVP more often reduces urine volume and the number of wet nights per week, but does not make children totally dry. Most children (up to 90%) relapse after discontinuation of DDAVP. It may be administered intranasally or as a tablet; both appear equally effective.

Prognosis

The prognosis for children with enuresis is generally good. Reassurance to the family that enuresis is not due to a serious renal, medical, or emotional problem and that it has a spontaneous cure rate of 15% per year is very helpful. Medical management improves symptoms in more than 70% of children.

See Also the Following Articles

Childhood Stress; Sleep, Sleep Disorders, and Stress.

Further Reading

- Berkowitz, C. D. (2000). Enuresis. In: Berkowitz, C. D. (ed.) *Primary care pediatrics* (2nd edn., chap. 33, pp. 131–134). Philadelphia, PA: W. B. Saunders.
- Butler, R. J. (2002). Impact of nocturnal enuresis on children and young people. *Scandinavian Journal of Urology and Nephrology* 35, 169–176.
- Glazener, C. M. and Evans, J. H. C. (2005). Simple behavioral and physical interventions for nocturnal enuresis in children. *Cochrane Database of Systematic Reviews* CD 003637, 1–31.
- Glazener, C. M., Evans, J. H. C. and Cheuk, D. K. L. (2005). Complementary and miscellaneous interventions for nocturnal enuresis in children. *Cochrane Database of Systematic Reviews* CD 005230, 1–38.
- Glazener, C. M., Evans, J. H. C. and Peto, R. E. (2005). Complete behavioral and educational interventions for nocturnal enuresis in children. *Cochrane Database of Systematic Reviews* CD 004668, 1–46.
- Koff, S. A. and Jayanthi, V. R. (2002). Non-neurogenic lower urinary tract dysfunction. In: Walsh, P. C. (ed.) *Campbell's Urology* (8th edn., Ch. 64, pp. 2261–2283). Philadelphia, PA: W.B. Saunders.
- Moffatt, M. E. K. (1989). Nocturnal enuresis: psychologic implications of treatment and non-treatment. *Journal of Pediatrics* 114, 697–704.
- Moffatt, M. E. K. (1994). Nocturnal enuresis: is there a rationale for treatment? *Scandinavian Journal of Urology and Nephrology* 163(supplement), 55–67.
- Nield, L. S. and Kamat, D. (2004). Enuresis: how to evaluate and treat. *Clinical Pediatrics* 43, 409–415.
- Schmidt, B. D. (1997). Nocturnal enuresis. *Pediatrics in Review* 18, 183–190.

Environmental Factors

W R Avison

University of Western Ontario, London, ON, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by W R Avison, volume 2, pp 53–59, © 2000, Elsevier Inc.

Historical Background

The Impact of Environmental Factors on Stressors

Differential Exposure and Vulnerability to Environmental Stressors

Glossary

Differential exposure hypothesis

The contention that specific statuses or roles are associated with greater or lesser levels of stressors that arise out of the conditions of life and which, in turn, affect levels of psychological distress.

<i>Differential vulnerability hypothesis</i>	Argument that elevated levels of distress among individuals in disadvantaged statuses reflect their greater responsiveness to stressors; that is, in two social groups exposed to the same levels of stressors, distress will be higher among those people in the disadvantaged statuses.
<i>Environmental factors</i>	Social or economic circumstances that affect exposure to stressors.
<i>Social causation</i>	A model that suggests that people in socially disadvantaged statuses will be exposed to more stressors that, in turn, result in more distress or disorder.
<i>Social selection</i>	A model that suggests that people with high levels of stress in their lives are more likely to experience mental health problems and consequently are more likely to occupy disadvantaged statuses.
<i>Stress process paradigm</i>	A model that describes how stressors manifest themselves in distress or disorder. It postulates the existence of three factors that either mediate or moderate the effects of stressors on distress or disorder: psychosocial resources, coping resources and strategies, and social supports. The model also specifies how environmental factors influence all other variables in the paradigm.

Historical Background

There is a vast body of research literature on the impact of social and economic factors on physical and mental health problems. In general, these studies have demonstrated that individuals' social locations in society, in terms of their social statuses and social roles, have important implications for both the kinds and levels of stressors to which they are exposed. Indeed, one of the most important contributions of the sociology of mental health has been to demonstrate conclusively that stressors are not experienced randomly by individuals but, rather, that there is a social distribution of stressors. This perspective has emerged out of a number of theoretical and empirical developments.

Much of the work on the social distribution of stressors has its intellectual origins in the work of Bruce and Barbara Dohrenwend in the 1970s. They observed that exposure to stressful life events was correlated positively with symptoms of mental illness. They argued that the higher rates of psychological impairment found among members of socially disadvantaged groups might be explained by their greater exposure to life events. The underlying assumption of this explanatory model was that social causation processes were operating: differences in social statuses

and social roles produce variations in the experience of life events that result in different rates of mental health problems among these social groups.

For the next decade, social scientists exhaustively studied the relationships among environmental factors, stressful life events, and psychological distress and disorder. At least three major themes emerged from this work. First, a conceptual paradigm emerged for examining the interplay among environmental factors, stressors, mediators, and mental health outcomes. Second, it became clear that stressful life events constituted only one kind of stressor and that research needed to identify and study other types. Third, there was general agreement with the notion that the distribution of stressors is socially patterned.

The Stress Process

Several researchers working independently in the 1980s developed conceptual models that enable us to better understand how socially induced stressors manifest themselves in psychological distress, symptoms of psychiatric disorder, or in other dysfunctions or health problem. All of these models focus on the association between stressful life events and their distressful manifestations (psychological distress, psychiatric disorder, etc.). These models also postulate the existence of three important factors that either mediate or moderate the effects of stressors on distress or disorder. These include psychosocial resources, coping resources and strategies, and social supports. In recent years, this model has been elaborated by specifying the kinds of environmental factors – primarily social and economic variables associated with social statuses and roles – that influence all other variables in the paradigm. This paradigm has provided social scientists and epidemiologists with a conceptual framework for studying the effects of environmental factors on stressors as well as on all other components of the model.

The Stress Universe

Throughout the 1970s and early 1980s, most of the research on environmental sources of stressors focused on life-change events. The most widely used measures of life stress were events checklists that usually contained from 30 to 100 events that represented potential changes in people's lives. These included events such as the death of a loved one, marital separation or divorce, job loss, and geographical moves. Over time, however, two developments substantially altered the study of environmental stressors. First, it became clear to researchers who were studying the stress process that the association between eventful stressors and distress was modest at best. For many

researchers, this suggested that stressors were not being measured comprehensively. Accordingly, they expanded life events inventories to include events that were not life changes, but which included other difficulties such as family conflicts, work conflicts, and financial difficulties. In part, the inclusion of these events reflected the realization among many researchers that stress often arises in the context of individuals' work and family roles. Second, investigations of intra-event variability in stressful life events revealed that life events inventories included ongoing or chronic stressors (such as family conflict, marital difficulties, and financial problems) as well as eventful stressors.

These considerations led researchers to develop more extensive strategies to measure a broader array of life stressors than had previously been the case. Perhaps the most comprehensive consideration of the many dimensions of life stress was presented by Wheaton. He distinguished among various concepts of stressors, including chronic stress, daily hassles, macro events, and traumatic events. In doing so, he presented a scheme for arraying these different stressors on a continuum ranging from sudden traumatic experiences and life change events at the discrete end to chronic stressors that are more continuous in nature. Wheaton demonstrated two important properties of these various dimensions of the stress universe: (1) they are relatively independent of one another and therefore are unlikely to be empirically confounded with one another and (2) they each have significant independent effects on various measures of physical and mental health outcomes. These findings suggest that any comprehensive attempt to estimate the impact of life stressors on health outcomes requires that we measure an array of stressors that includes both discrete life experiences and more chronic or ongoing stressors.

The Social Distribution of Stressors

Early research by Bruce and Barbara Dohrenwend and by George Brown and his colleagues on life events clearly established that socially induced stressors exerted important influences on various measures of mental health outcomes. What seemed particularly interesting was that the socially disadvantaged groups with the highest levels of distress and disorder also appeared to be exposed to the most stressful life events. This led to the hypothesis that social differences in distress and disorder might be accounted for by variations in either exposure or vulnerability to social stressors.

Subsequently, in a seminal article on the sociological study of stress, Pearlin argued for the need to systematically investigate the ways in which social

structure affects individuals' exposure to stressors. Essentially, Pearlin asserted that the roles and statuses that people occupy in their everyday lives have important consequences for the kinds of stressors to which they are exposed and the frequency with which this occurs. In short, he argued that accounting for the impact of environmental factors is crucial for understanding the ways in which individuals experience social stressors.

This point has been made even more salient by Turner, Wheaton, and Lloyd in their paper on the epidemiology of social stress. Their findings from a large community study reveal that younger people are more exposed to a variety of stressors than are older respondents. They also find this pattern among women compared to men, unmarried people compared to married people, and individuals in lower socioeconomic status (SES) positions compared to those in higher SES positions. In their view, this clearly indicates that stressors are not experienced randomly in the population. Quite the contrary, they asserted that their results reveal that there is a social distribution of stressors that is characterized by greater exposure among members of disadvantaged social groups.

The Impact of Environmental Factors on Stressors

In thinking about the social distribution of stressors, social scientists have focused primarily on three major environmental determinants: social statuses, social roles, and the ambient environment. Most formulations of the stress process model take the view that individuals' locations in the structure of society place them at greater or lesser risk of encountering stressors. These locations are defined by the various statuses that individuals hold and by the various social roles they occupy. In addition, there are more ambient characteristics of the social environment that are not specific to statuses or roles but that may generate stressful experiences for individuals.

Social Statuses

In the literature, individuals' positions in the structure of society are often defined by six major status characteristics: age, gender, marital status, race/ethnicity, employment status, and SES. Each of these environmental factors is associated with differential exposure to stressful experiences.

Age There appears to be substantial agreement among researchers that exposure to stressful experiences declines with age. Whether the stressors in ques-

tion are life events or chronic role strains, younger people report significantly more stress than do the elderly. Moreover, economic hardship tends to decline with age as does marital conflict in marriages that stay intact. Despite these findings, it is not yet clear whether the impact of age on exposure to stressors is a function of maturation processes, birth cohort or generational effects, or life-cycle processes.

Gender Studies of gender differences in exposure to stressors have generated inconsistent findings. Some research concludes that women experience more stressful life events than men, whereas other research finds no differences. Turner and Avison conducted one of the more comprehensive studies of gender exposure to various dimensions of stress. Women report more stressful life events than men, but they experience fewer lifetime traumatic events and less discrimination stress. Women and men do not differ in their exposure to chronic stressors. When these various dimensions of stress were considered cumulatively, they found that men experience greater exposure compared to women but that this difference is relatively small. Turner and Avison concluded that these small gender variations in exposure to stressors are unlikely to account for gender differences in psychological distress.

Although other researchers reported that women are more exposed than men to chronic strains due to financial hardship, workplace difficulties, and role overload associated with parenting and work, these differences are difficult to separate from the gendered roles that women and men play in the workplace and in the household division of labor. Indeed, one of the lessons of stress research in the sociology of mental health is that gender conditions the effects of a variety of roles and statuses on the exposure to stress.

Marital status One of the more robust findings in the study of social stress is the observation that married individuals experience considerably fewer stressors than either never married or formerly married individuals. This pattern can be observed for chronic strains as well as for stressful life events. Unlike age and gender, however, the causal direction of the relationship between marital status and stressful experience is open to competing interpretations. A social selection interpretation suggests that people with high levels of stress in their lives are less likely to marry or, if they do so, they are more likely to separate and divorce. A social causation interpretation suggests that people who have never married and those who have separated or divorced are more likely to experience an array of life events and chronic strains than are the married.

Some longitudinal research indicates that divorce leads to elevated levels of depression and that this change is accounted for by a decline in living standard, economic difficulties, and a reduction in social support. Even if the end of a marriage provides some escape from a stressful situation, divorce is accompanied by life stress that has depressive consequences. Other studies also report significant increases in psychological distress among the marital-ly disrupted over and above their predivorce levels, but find virtually no evidence that changes in financial stressors, changes in role demands, or changes in geographic location mediate the relationship between marital disruption and distress. These are reflective of longitudinal studies on marital disruption and remarriage insofar as they generate few consistent findings other than the observations that marital disruption is associated with elevated psychological distress and that remarriage results in only a partial reduction in this elevation.

Race and ethnicity Studies of racial and ethnic variations exposure to stressors are relatively recent in the literature. There is general agreement that experiences of discrimination constitute an important set of stressors for African Americans, Hispanic Americans, and American Indians. In addition, stressors arising out of the social disadvantages of these groups have also been observed. Among Asian Americans, the experience of discrimination and the stress of migration have been noted as important threats to their mental health.

Turner and Avison systematically examined differences between African Americans and Whites in the United States in exposure to stressors. Across five different dimensions (recent life events, chronic stressors, total lifetime major events, lifetime major discrimination, and daily discrimination), African Americans experience significantly more stress than do Whites. Indeed, the cumulative difference between African Americans and Whites in exposure to stress is more than 0.5 standard deviation.

It is interesting to note, however, that these elevated levels of stressors among racial and ethnic minority members do not necessarily translate into higher rates of distress or disorder for all groups. In their review of this issue, Williams and Harris-Reid concluded that African Americans have lower prevalence rates of psychiatric disorders than do Whites. The data are inconsistent when comparing Whites with Hispanic Americans or Asian Americans. What little data exist about the epidemiology of mental health among American Indians suggests that their rates of disorder, especially depression and alcohol abuse, are elevated.

It is difficult to interpret the meaning of these findings. Lower rates of disorder in the face of elevated levels of distress suggests that some racial and ethnic group may be less vulnerable than others to these stressors or that countervailing effects of psychosocial resources may reduce the impact of stressors on their mental health. Turner and Avison suggested that African Americans may exhibit a response tendency in which they underreport infrequent or mild experiences of distress, thus leading to an underestimation of their levels of distress. They demonstrate that once this tendency is taken into account, there is clear evidence that African Americans' elevated exposure to stress manifests itself in significantly higher distress.

Employment status Research leaves little doubt that the unemployed experience more negative mental health outcomes than do the employed. The evidence of this correlation is most clear for outcomes such as symptoms of depression and anxiety and measures of psychological distress. Longitudinal studies support the conclusion that job loss results in higher levels of mental health symptoms.

Studies of the factors that intervene between individuals' job losses and their health problems have identified at least two major sources of environmental stressors that mediate this relationship. Some researchers have shown that job loss and unemployment create financial strains that lead to mental health problems. Others have examined the mediating role of marital and family conflict. These studies report that unemployment leads to increasing conflicts between the unemployed worker and other family members. Some researchers have suggested that the elevated levels of distress observed among women whose husbands are experiencing job-related stress may be consistent with the costs of caring hypothesis described earlier.

Socioeconomic status Although there is general agreement that SES and mental illness are inversely correlated and that exposure to stressors are a major determinant of mental health problems, there has been surprisingly little consensus among researchers about the SES-stress relationship. Whether SES is measured by some combination of education and income or by occupational prestige among those with jobs, contradictory results emerge. These inconsistent findings have led some researchers to argue that it is not stress exposure that produces higher rates of mental illness among individuals with lower SES. Rather, they suggest that the impact of stressors on mental illness is more substantial among low-SES than high-SES individuals. In other words, they argue that lower-class individuals are differentially vulnerable to stressors. It seems clear that this debate with

respect to the influence of SES has not been resolved. Nevertheless, more recent studies suggest that when stressors are comprehensively measured, there is a significant social-class gradient in exposure to stress.

Some of the most compelling evidence of the effects of economic hardship on marital relationships has been presented in studies of families during the Great Depression of the 1930s. This work clearly indicates that economic difficulties increase marital tensions in most families and especially in those that were most vulnerable prior to the economic hardship. Similar findings have been reported in studies of families whose lives were affected by the farm crises of the 1980s.

Social Roles

A central focus of much research on the stress process has been on the ways in which the social roles that individuals occupy expose them to stressors. Pearlin described these role strains in rich detail. For Pearlin, several types of stress may arise from role occupancy: excessive demands of certain roles, inequities in rewards, the failure of reciprocity in roles, role conflict, role captivity, and role restructuring. These various types of stress are important sources of stress that may manifest themselves in symptoms of distress or disorder. These kinds of environmental stressors have been studied most intensively in investigations of family roles and work roles.

Family roles Recent research on the family and mental health has focused on family structure in terms of the intersection of marital status and parenthood. In this context, there has been intense interest in the impact of single parenthood on symptoms of distress. Results consistently show higher levels of distress among single mothers than among married mothers. These studies of single parenthood clearly reveal that family structure and the roles embedded in that structure are important determinants of women's mental health.

Research indicates that separation and divorce trigger chronic stressors such as income reduction and housing relocation. In addition, the divorced experienced more life events than the married, particularly negative events involving loss. When children are involved, there may be additional strains associated with separation or divorce. The custodial parent, usually the mother, assumes many household, financial, and emotional responsibilities previously shared by two parents.

The work role In addition to examining how differences in employment or work status influence exposure to stressors and subsequent mental health

outcomes, social scientists have become aware of the importance of understanding how experiences in the work role are related to stress and health. Despite the observation that being employed generally has positive psychosocial consequences for individuals, not all employment circumstances are the same. Indeed, there are important variations in the stressors associated with any particular work situation. Work exposes individuals to various kinds of stressful experiences and provides different kinds of rewards for different people – financial rewards, self-esteem, a sense of control over one's life, and so on. It seems, then, that the net effect of paid employment for any individual will depend on the balance of these costs and benefits.

A number of recent contributions to the study of work and mental health have provided some useful models for this kind of research. In her comprehensive review of the literature on the interplay between work and family, Menaghan demonstrated clearly that work-related stressors are significantly associated with the mental health of family members. Other researchers have convincingly documented how role overload and the sense of personal power have important implications for the effects of women's employment on their mental health. Some investigations have shown how job characteristics such as full-time versus part-time work and substantive complexity have significant effects on mental health.

The intersection of work and family roles Research on work and family stress among women suggests a number of ways in which work and family roles interact in their effects on psychological distress and depression. These studies highlight the importance of considering both family stressors to which women are exposed and work-related stress.

Relatively few studies have examined how single parenthood and paid employment interact in their impact on mental health problems. The studies that have investigated this issue conclude that differences between single-parent families and two-parent families in role obligations and opportunities may have significant effects on the balance between work and family responsibilities. For married mothers, paid employment may be more easily integrated into daily family life. The presence of a spouse provides the opportunity to share some of the child-care responsibilities. Alternatively, dual-income families have more funds for outside child care or paid assistance in the home.

Employment for single-parent mothers may represent more of a pressing responsibility than an opportunity for achievement and development. Most single-parent families live in poverty or near poverty.

In such circumstances, although paid employment may alleviate some of the most pressing financial strains, other strains persist. Moreover, when single mothers obtain employment, it is common for them to also bear the continued sole responsibility for the care and nurturance of their children. Such dual demands may generate a cost to employment – role overload.

Under these circumstances, it seems probable that fewer psychosocial rewards associated with employment (greater self-esteem, self-efficacy, or social support) will accrue to single mothers. This is all the more likely to be the case because single parents may be constrained to select jobs that are not their first choice but that, instead, are near their homes or have hours of work that fit with their children's schedules or with the availability of child care.

Ambient Strains

Not all stressors are associated with statuses and roles. Ambient strains that are not attributable to a specific role but, rather, are diffuse in nature and have a variety of sources. These include experiences such as living through an economic recession, living in unsafe housing, or living in a dangerous neighborhood. Some studies have demonstrated that ambient strains associated with the neighborhoods in which adolescents live magnify the impact of other stressors on their mental health. Other studies make the same point with respect to the ambient effect of economic environments.

Differential Exposure and Vulnerability to Environmental Stressors

An important issue has been to test the relative importance of differential exposure and differential vulnerability to stressors as explanations of the differences in psychological distress by social status or social role. The differential exposure hypothesis contends that specific statuses or roles are associated with greater or lesser levels of stressors that arise out of the conditions of life and which, in turn, affect levels of psychological distress. For example, applying this argument to marital status, the hypothesis is that the transition from marriage to separation or divorce brings with it significant increases in exposure to financial strains, role overload, and other types of stressors. This increase in the burden of stress experienced by the separated or divorced people translates into elevated levels of mental health problems. Conversely, individuals who remain married experience far fewer stressful circumstances and, accordingly, have lower levels of psychological distress.

The competing explanation, the differential vulnerability hypothesis, argues that elevated levels of distress among individuals in certain statuses reflect their greater responsiveness to stressors. Such increased responsiveness or vulnerability to stressors has been attributed to a number of different sources. These include dimensions of personal and social competence, such as self-efficacy, that contribute to individuals' greater or lesser resilience in the face of stress and access to coping resources that moderate stressful experiences. The most recent work on this debate has been conducted by Turner, Wheaton, and Lloyd. Their investigation of age, gender, marital status, and SES reveals little evidence that any of the differences associated with these social statuses in distress or disorder can be attributed to differential vulnerability. Instead, they found that observed variations in mental health outcomes are due largely to different levels of exposure to stressors. They and others argue that the use of a more comprehensive measure of stressors that includes life events and chronic strains is likely to better estimate actual exposure and to avoid the error of attributing unmeasured differential exposure to differential vulnerability.

See Also the Following Articles

Familial Patterns of Stress; Gender and Stress; Life Events Scale; Marital Status and Health Problems; Social Status and Stress; Workplace Stress; Ethnicity, Mental Health; Work–Family Balance.

Further Reading

Aneshensel, C. S. and Phelan, J. C. (eds.) (1999). *Handbook of the sociology of mental health*. New York: Kluwer Academic/Plenum Publishers.

- Brown, G. and Harris, T. (1989). *Life events and illness*. New York: Guilford Press.
- Dohrenwend, B. S. and Dohrenwend, B. P. (eds.) (1974). *Stressful life events: Their nature and effects*. New York: John Wiley.
- Menaghan, E. G. (1994). The daily grind: work stressors, family patterns, and intergenerational outcomes. In: Avison, W. R. & Gotlib, I. H. (eds.) *Stress and mental health: contemporary issues and prospects for the future*, pp. 115–147. New York: Plenum Press.
- Mirowsky, J. and Ross, C. E. (2003). *Social causes of psychological distress*, 2nd ed. Hawthorne, NY: Aldine de Gruyter.
- Pearlin, L. I. (1989). The sociological study of stress. *Journal of Health and Social Behavior* 30, 241–256.
- Pearlin, L. I., Lieberman, M. A., Menaghan, E. G. and Mullan, J. T. (1981). The stress process. *Journal of Health and Social Behavior* 22, 337–356.
- Turner, R. J. and Avison, W. R. (2003). Status variations in stress exposure: implications for the interpretation of research on race, socioeconomic status, and gender. *Journal of Health and Social Behavior* 44, 488–505.
- Turner, R. J., Wheaton, B. and Lloyd, D. A. (1995). The epidemiology of social stress. *American Sociological Review* 60, 104–125.
- Wheaton, B. (1994). Sampling the stress universe. In: Avison, W. R. & Gotlib, I. H. (eds.) *Stress and mental health: contemporary issues and prospects for the future*, pp. 77–114. New York: Plenum Press.
- Williams, D. R. and Harris-Reid, M. (1999). Race and mental health: emerging patterns and promising approaches. In: Horwitz, A. V. & Scheid, T. L. (eds.) *A handbook for the study of mental health: social contexts, theories, and systems*, pp. 295–314. New York: Cambridge University Press.

Environmental Stress, Effects on Human Performance

G R J Hockey

University of Sheffield, Sheffield, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G R J Hockey, volume 2, pp 60–65, © 2000, Elsevier Inc.

Theories of Effects of Stressors on Performance
Patterns of Stressor Impairment
Compensatory Control under Stress
Latent Degradation under Stress

Glossary

<i>Compensatory control</i>	A general adaptive response to stress serving to protect high-priority goals from disruption.
<i>Costs</i>	Effort or other resources used to achieve task goals.
<i>Environmental stressors</i>	Identifiable external conditions that pose a threat to the performance of subjective state, such as noise, heat, and time pressure.
<i>Human performance</i>	The effectiveness of behavior in relation to specific task demands and goals, assumed to reflect the operation of underlying mental processes.
<i>Latent degradation</i>	The effects of performance impairment not necessarily manifested in disruption of primary task activities.
<i>Short-term memory (STM)</i>	General label used to describe memory for very recent events, referring to the storage of information over brief periods (typically 10–20 s).
<i>Working memory (WM)</i>	Cognitive system acting as a “mental workspace” for manipulating information necessary for carrying out current processing and problem solving tasks, comprising both storage and attention/control components.

Human performance refers to the effectiveness of task activity in relation to goals. This assumes a focus on the primary mental processes that underlie controlled human action, particularly in tasks requiring the use of memory, attention, decision making, and perceptual-motor skills. A concern with the effects of stress on human performance has a long history in psychology, driven by both practical and theoretical issues. It has been known since the 1970s that the nature of the task is a strong determinant of how performance will be affected; it has also recently become clear that there is a general problem that affects the performance of all tasks. The key to understanding this appears to be to consider the goals that underlie behavior. It is not surprising that goals change under extreme stress away from the current task and toward a concern with bodily states or emergency reactions. In more everyday contexts, this implies a need to examine the strategies used to manage goal-relevant information and the factors that determine changes in goals. This has not always been done in research in this area. In what follows, I discuss the theoretical background to research on stress and performance, review the evidence for different patterns of disruption across various task and performance criteria, and show how an adaptive regulatory perspective enhances our understanding of the nature of the problem.

Theories of Effects of Stressors on Performance

Stressors have traditionally been considered as having one of two kinds of effect on performance: distraction or arousal. Early research on the effects of what we now consider stressors (e.g., loud noises) assumed a broadly distraction kind of theory. In general, apart from transient effects, such experiments were spectacularly unsuccessful in demonstrating any effects on complex performance – including tests of intelligence, memory, and skill. Not until the mid-1950s, with the application of information processing concepts to psychology, did theory have much of a part to play in this kind of research. More recent treatments (triggered by the seminal work of Donald Broadbent and his colleagues) has recognized that the effects of stressors may be masked, or compensated for, by the built-in redundancy and strategy options available to the cognitive system. Broadbent argued that it should be possible to demonstrate the effects of stressors by designing tasks that effectively stretched this regulatory process, for example, by presenting information rapidly and without breaks or by making critical events rare and unpredictable in time and place (e.g., serial reaction and vigilance tasks).

In the distraction theory, environmental events were thought to compete for attention through either their strong stimulus qualities (noise or heat) or their impact on bodily states (causing anxiety and other emotional reactions). In its modern form, the distraction view was developed most fully in Broadbent's influential filter theory. For example, noise was thought to impair task-relevant operations by capturing selective attention. This involuntary attention to the noise source interrupted the effective intake of relevant task information and gave rise to processing errors. The distraction theory is best supported by studies of intermittent noise, which impairs performance in the few seconds following each burst. In general, however, the distraction view has been thought to offer a less convincing explanation of the effects of both continuous noise and of other stressors, such as sleep deprivation, not so readily identified with external events.

As an alternative to distraction theory, the arousal theory of stress and performance was developed in the 1960s following the generally accepted view of the day that arousal was a unitary, nonspecific brain process. The application of arousal theory to stress effects was based mainly on the systematic studies of combined effects of stressors carried out in Cambridge, UK. This made extensive use of the five-choice serial reaction task, in which subjects were required to respond to a continuous stream of lights coming on by tapping one of the five metal plates

corresponding to the lights. This task showed impairment under noise and sleep deprivation and improvement when subjects were given incentives to do well. However, performance under noise improved somewhat when subjects were also sleep deprived, whereas incentives made their effects worse. These interactions between stressors were regarded as supporting the arousal theory of stressors because they appeared to show that impairment could be caused by both understimulation (sleep deprivation) and overstimulation (noise). However, the assumptions concerning the nature of arousal could not be convincingly demonstrated independently of the task itself. For example, there is no direct evidence that noise is arousing in a physiological sense, except in terms of the transient startle pattern at the onset. Increased arousal under noise is typically found only when people are carrying out cognitive tasks, and then only when they are concerned to prevent performance decrements occurring. In addition, it is clear from developments in physiological theory that arousal is more complex than previously assumed – more specific and linked to task processes. In sum, it has proved to be inadequate as a general theory of effects of stress.

Patterns of Stressor Impairment

The central approach to the understanding of stress effects presented in this article focuses on the problem of managing the general threat to task goals through adaptive changes in behavior. However, as I mentioned earlier, different stressors also appear to pose specific threats for tasks, depending on their information processing requirements. An analysis of stress effects on performance carried out in the 1980s found specific patterns of decrement across different indicators of performance and strategy. Some stressors were more likely to impair performance on one kind of task and others on a different kind. For example, loud noise typically impairs performance on tasks that require accuracy, short-term memory (STM), or problem solving. Sleep deprivation causes impairment on tasks that require accuracy, speed, and a high level of selective attention, as well as having more general effects on memory. Both noise and sleep deprivation have effects that are more pronounced under fatigue conditions (when tasks involve long periods of work without breaks). Working in hot conditions has widespread effects on most aspects of performance, especially tasks involving more complex decision making. The effects are related to the exposure time and effective temperature, but, unlike noise, do not appear to increase with time at work.

The set of indicators used in this analysis included general alertness, selectivity of attention, speed versus

accuracy, and STM. With changes in cognitive theory over the intervening period and the experience of the effects in real work tasks, some changes to this list are needed. General alertness no longer appears to have strong diagnostic value because it is likely to be involved in all active regulatory behavior. The mechanisms underlying STM have undergone considerable evolution with the development of working memory (WM) theory by Baddeley and colleagues. In addition, there is a need for a new analysis, which takes into account differential effects of regulatory activity. However, even with these caveats, changes in these performance indicators may be seen as a profile of the sorts of information-processing problems that different stress conditions may give rise to. The most general pattern of decrement is associated with environmental factors, such as noise, danger, or social evaluation, that give rise to subjective states of threat or anxiety. This may be regarded as the modal stress pattern involving a subjective state of high activation, high selectivity of attention, a preference for speed over accuracy, and reduced WM function. Decrements are more common on tasks of long duration, especially where the continued use of WM is central to maintaining the flow of the work. Selective attention is normally very effective, unless response is required to a number of different events or subtasks, in which case only the most important may be maintained. A familiar effect of such stressors is narrowed attention, in which the high-priority features of tasks are maintained and secondary aspects are neglected. Such an effect has been observed for a wide range of stress states, including noise, high workload, threat of shock, danger, and most forms of induced anxiety. Other stressors are associated with different kinds of changes. For example, WM appears relatively stable under hot working conditions or with extended work periods. In all cases, however, it has become clear that we cannot separate the underlying effects on cognitive processes from those relating to changes in performance goals or strategies. An increase in reliance on one kind of process may be the result of a strategic reduction in the use of another. Because of this, patterns of stressor effects cannot be discussed without reference to an understanding of what the performer is trying to do when carrying out a task and of what conflicts exist between different goals.

Compensatory Control under Stress

Modern treatments of psychological stress emphasize the cognitive transactions that mediate between stressful events and the adaptive response to them. This appraisal process evaluates the implications of the stressor for both current activities and personal

well-being. In terms of performance tasks, this may mean focusing information processing resources more strongly on the task (performance protection) or withdrawing resources in order to combat the stressor itself. This latter reaction is likely to be more effective in reducing the effects on bodily or emotional states, but it inevitably leads to a loss of performance goals. Performance protection is the usual response in everyday situations in which the individual is highly skilled, the task sufficiently important, and the stressor familiar and manageable. Serious disruption is rare for high-priority activities and is usually associated with traumatic events. This is because a compensatory process operates to maintain the primary task goals under the increased threat of disruption, resulting in a reduced response to the control of the emotional state and other competing goals. The increased effort underlying compensatory control is considered to reflect the involvement of the central executive functions responsible for the maintenance of high-level cognitive behavior, as observed in problem solving, reasoning, and all goal maintenance activity. As such, it is a limited resource that inevitably attracts costs when it is overemployed.

On the other hand, we know that decrements are relatively common, especially in laboratory studies or where skill and motivation are low. Within this framework, the specific patterns of decrement outlined earlier may be considered a baseline or default pattern of decrement under different stressors – how performance might be expected to suffer in the absence of compensatory control activity. As an example, consider a pair of studies carried out in Stockholm in the 1980s. They showed that noise impaired performance on an arithmetic task on one occasion but not on another. How can this be understood? The answer is related to motivational factors such as compensatory effort. The investigators also measured the physiological and subjective costs associated with having to work on the task under noise. In the study in which performance was unimpaired, they observed a marked increase in adrenaline and ratings of subjective effort. However, in the case in which performance was disrupted by noise, no such changes were observed. The most satisfactory explanation of this (and other similar) findings is that noise imposes an additional load on our capacity to maintain adequate orientation toward the task. If we can make an additional effort under such circumstances, performance may be protected against disruption, although only at the cost of increased strain in other areas. Alternatively, we may be unwilling (or unable) to make such an effort, in which case we will experience less strain but inevitably suffer a decrement in task performance. Such trade-offs are the routine consequences of having to manage stress

and other environmental demands while still carrying out our commitments to external task goals.

Latent Degradation under Stress

Although stress does not always result in any obvious reduction in performance, this should not be taken to mean that there is no threat to task goals. There is now considerable evidence of knock-on effects of performance protection to secondary aspects of behavior, related to both performance and costs; I have referred to these as latent degradation. By reducing the safe working margins of the adaptive control process, these may threaten the integrity of performance – for example, strategies that work only if there are no new problems to deal with. Four kinds of latent degradation may be identified: two performance indices (secondary task decrements and strategy changes) and two cost indices (psychophysiological activation and fatigue aftereffects).

Secondary Task Decrements

Decrements in the secondary aspects of performance are commonly observed in studies of the effects of high workload, providing an indirect measure of increasing load on primary tasks. Such effects have been studied less systematically in assessing threats from environmental stressors, although they are, in fact, also common. One of the best-documented forms of secondary task decrement under stress is the narrowing of attention found in spatially complex tasks. For example, although a central tracking task may be carried out effectively under noise, the detection of signals in the visual peripheral may be impaired. Similar attentional narrowing has been found under both laboratory and field conditions and for a wide range of environmental conditions (noise, heat, anxiety associated with deep sea diving, and threat of shock). This type of decrement may be related to strategy changes because they depend on a shift of priority between task elements.

Strategy Changes

Strategy changes are (usually adaptive) changes in the way that tasks are carried out under stress. One obvious way is to minimize the disruption to primary activities by reducing the time spent on secondary tasks. However, there may also be more subtle changes, involving a shift to less resource-intensive modes of task control, reducing dependency on effort-demanding processes such as WM, which is known to be impaired under stress conditions. Despite their obvious diagnostic value, such effects have not been well studied, partly because of the complex task environments

necessary to analyze strategy changes. It has been known for some time that, under periods of difficulty or stress, industrial process operators may shift from an attention-demanding open-loop control (in which whole sequences of actions are guided mainly by the operator's internal mental model) to a simpler closed-loop strategy (in which actions are carried out one at a time, paying more attention to feedback). This may slow the process or fail to make optimal use of available options, but it minimizes the likelihood of serious errors.

A good example of this is a well-known study of French air-traffic controllers, carried out in the 1970s. The controllers adopted a simplified method of dealing with aircraft contacts when they exceeded a comfortable number, but under very high workload they switched from individual plane-by-plane routing instructions to a fixed procedure for all contacts. By minimizing the demands for planning and aircraft management, they reduced the need to involve the vulnerable WM system. The strategy change is adaptive in that secondary goals such as airport schedules and passenger comfort are compromised in the service of the primary goal of safety. A second example is the work of Schönplug's group in Berlin, which used a simulated office environment to examine decision making in stock control. Under normal conditions, participants typically held background information (on prices, stock levels, etc.) in WM while making a sequence of decisions. However, under time pressure or loud noise, they tended to check lists containing such information before making each decision. Reducing the load on WM helped people to keep decision errors to a minimum, although at the expense of increased time costs. Again, the change is adaptive because accuracy matters more than speed in such work. However, in situations in which speed is also important, the hidden loss of efficiency represents a genuine stress-induced impairment.

Psychophysiological Activation

One of the most reliable costs of the use of increased effort to protect performance is the observation of increased levels of activation. This is particularly true of the physiological systems involved in emergency reactions (e.g., sympathetic and musculoskeletal responses, and responses of the neuroendocrine stress systems). These effects are typically accompanied by changes in subjective reports of emotional and mood states reflecting the affective response to emergency and sustained coping effort. These may be thought of as the unwanted side-effects of the compensatory behavior that helps to maintain

primary performance under threat from environmental conditions.

The effect is illustrated in an early study of sleep deprivation, in which decrements in arithmetic computation following a night without sleep were smaller for participants found to have increased muscle tension (interpreted as evidence of greater effort to combat sleepiness and maintain orientation toward the task). This performance-cost trade-off is seen more clearly in several studies of noise effects using more meaningful psychophysiological measures. Noise has been shown to increase heart rate, blood pressure, adrenaline, and subjective effort in tasks in which performance decrement is forestalled. In the Swedish study referred to earlier, two different patterns of arithmetic performance and costs were observed in different experiments. In one, performance was impaired by noise, with no change in levels of adrenaline or effort. In the other, performance was maintained, but adrenaline and effort levels were both greater. Unfortunately, because of the difficulty of obtaining psychophysiological measures under such circumstances, there are few studies within real work contexts, although another Swedish study found that the absence of decrement in work output during an intense period of organizational change was, again, accompanied by a compensatory increase in adrenaline and cognitive effort.

Such effects illustrate the role of compensatory regulation in the protection of performance and may be seen as a trade-off between the protection of the primary performance goal and the level of mental effort that has to be invested in the task. They also indicate that the regulation of effort is at least partially under the control of the individual rather than being an automatic feature of task or environmental conditions.

Fatigue Aftereffects of Stress

A final form of latent degradation is one that appears only after set tasks have been completed, in terms of decrements on new (and less critical) tasks. Such after-effects have also been studied very little, and then normally within a workload-fatigue paradigm. However, they are equally appropriate as a response to the sustained effort required to maintain effective levels of work under stressful environmental conditions.

Given its long-recognized importance, work fatigue has been studied extensively since the early days of psychology, although it has proved surprisingly difficult to demonstrate carry-over effects of this kind. Even intensive research programs carried out by the U.S. army failed to find any marked fatigue effects of periods of up to 60 h of continuous work. Holding

and others have showed that there are methodological difficulties in the analysis of this apparently straightforward problem. As with the compensatory response to stressors, participants in such experiments appear able to work harder (make more effort) for brief periods to respond to the challenge of any new test, effectively compensating for any reduction in capacity. However, when tired people are provided with alternative ways of carrying out the postwork test they are more likely to choose one requiring low effort, even though it entails more risk of error. Similar results of high workload and stressful jobs have been found for driving examiners, bus drivers, and junior doctors. This approach to fatigue reveals it to be a state in which there is a shift toward preferring activities requiring less effort or use of WM. So far, there has been little direct research on this form of decrement with laboratory stressors, although similar effects have been found for noise and high workload.

The link between stress and fatigue is a very strong one. It is likely that actively managing stress in order to protect performance leads directly to fatigue, so that recovery is necessary before we can function even when the stressor is no longer present. Recent work carried out in the Netherlands suggests that fatigue impairs the effectiveness of the executive control system that maintains the activation of tasks in working memory, triggering both the withdrawal of effort and compensatory changes in information processing strategy. At present, we have no direct evidence of the

brain processes involved in this stressor → control/effort → fatigue chain, but it is currently being addressed in a number of laboratories. Clearly, a better understanding of the physiological basis of the control of stress during task performance will help us manage work and other tasks more effectively, as well as informing our approach to the design and management of the working environment.

Further Reading

- Frankenhaeuser, M. (1986). A psychobiological framework for research on human stress and coping. In: Appley, M. H. & Trumbell, R. (eds.) *Dynamics of stress: physiological, psychological and social perspectives*, pp. 101–116. New York: Plenum.
- Hancock, P. A. and Desmond, P. A. (eds.) (2001). *Stress, workload, and fatigue*. Mahwah, NJ: Lawrence Erlbaum.
- Hockey, G. R. J. (1986). Changes in operator efficiency as a function of environmental stress, fatigue and circadian rhythms. In: Boff, K., Kaufman, L. & Thomas, J. P. (eds.) *Handbook of perception and performance* (vol. 2), pp. 1–44. New York: John Wiley.
- Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: a cognitive energetical framework. *Biological Psychology* 45, 73–93.
- Hockey, G. R. J., Gaillard, A. W. K. and Burov, O. (eds.) (2003). *Operator functional state: the assessment and prediction of human performance degradation in complex tasks*. Amsterdam: IOS Press.

Epilepsy

C J Schramke and K M Kelly

Allegheny General Hospital and Drexel University
School of Medicine, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by K M Kelly and C J Schramke, volume 2, pp 66–70,
© 2000, Elsevier Inc.

Epilepsy and Seizures
Stress and Seizures
Epilepsy as a Stressor
Interventions

Glossary

Electroencephalogram (EEG)
Epilepsy

A study of the electrical activity of the brain.

A condition characterized by recurrent (two or more) unprovoked seizures.

Epileptic seizure

An abnormal electrical discharge of brain neurons that results in a transient alteration in a person's normal level of consciousness and/or motor, sensory, autonomic, or psychic functioning.

Non-epileptic seizure-like event (NESLE)

An episode that behaviorally resembles an epileptic seizure but that is not associated with epileptic activity on EEG monitoring.

Provoked seizure

A seizure caused by and occurring in close temporal association with nonepileptic conditions such as metabolic abnormalities, infections, intoxications (e.g., from alcohol and illicit drugs), brain injury,

sleep deprivation, and various emotional states and experienced typically as an isolated event.

Epilepsy and Seizures

Patients with epilepsy frequently report that stress and stressful life events influence the frequency and severity of their seizures and that epilepsy itself can cause significant stress. A survey of more than 1500 patients with epilepsy that was completed in 2005 indicated that emotional stress, sleep deprivation, and tiredness were the most frequently reported precipitants of seizures. Another survey, completed in 2003, found that 64% of patients believe stress increased seizure frequency and that 32% had specifically tried stress-reducing techniques to help with their epilepsy. Despite the apparent relationship between stress and epilepsy, there is a relatively poor understanding of the specific means by which stress affects the occurrence of seizures. Clinical studies have provided strong anecdotal information and evidence of stress-related seizure activity, but these studies are relatively few and inconclusive. Basic science research has delineated many of the relevant anatomical pathways of the brain and the potential physiological mechanisms that could be affected by stress, including how changing concentrations of various hormones and neurotransmitters can affect seizure threshold. However, the information from these studies remains incomplete and further research is necessary. Despite the lack of specific knowledge about the complex interplay between stressful stimuli and the seizures that apparently occur in response to them, some basic understanding of these relationships does exist.

To facilitate an understanding of how stress can influence the occurrence of seizures, a description of the various types of epilepsy and the seizures that characterize them is important. Epilepsy is classified as of unknown cause (idiopathic, primary), of known cause (symptomatic, secondary), or of unknown cause but presumed to be symptomatic (cryptogenic). Idiopathic epilepsy is considered to have a hereditary cause (e.g., childhood and juvenile absence epilepsy and juvenile myoclonic epilepsy), whereas symptomatic epilepsy is considered to be a consequence of a known or suspected brain injury (e.g., perinatal hypoxia, meningitis, and trauma). The different types of epilepsy are characterized by a variety of seizure types that are classified as either partial (clinical or EEG evidence that indicates a focal onset in one cerebral hemisphere) or generalized (no evidence of focal onset, thus appearing to begin simultaneously in both cerebral hemispheres).

Partial seizures are described as simple (without impairment of consciousness) or complex (with impairment of consciousness). Simple partial seizures can include the following types of symptoms: motor (jerking of an arm), sensory (numbness, tingling, autonomic (piloerection, pupillary dilatation), and psychic (deja vu, fear, structured hallucinations). Simple partial seizures are typically associated with an abnormal electrical discharge in a restricted area of the contralateral cerebral cortex corresponding to the body area involved. Aura has been used traditionally to describe the sensory, autonomic, or psychic symptoms perceived by the patient before the onset of impaired consciousness and/or a motor seizure. Actually, an aura is a simple partial seizure that may or may not progress to another seizure type. Prodromal symptoms of prolonged mood changes, uneasiness, or premonitions are usually not auras. A simple partial seizure may progress to a complex partial seizure, which also can occur spontaneously. A complex partial seizure includes impairment of consciousness, which refers to the patient's abnormal awareness of and responsiveness to environmental stimuli. The initial features of the seizure include an arrest reaction or motionless stare usually followed by automatisms, which are relatively coordinated motor activities occurring during the period of impaired consciousness (lip smacking, fumbling with clothing). The seizure is usually brief, lasting from seconds to minutes, and there is a period of postictal (after the seizure) confusion. These seizures are often referred to as temporal lobe seizures or psychomotor seizures. A partial seizure, simple or complex, can evolve into a generalized tonic-clonic seizure. The EEG shows focal discharges in one cerebral hemisphere evolving into generalized, bilaterally synchronous discharges.

Generalized seizures are divided into convulsive (major motor) or nonconvulsive (brief loss of consciousness or minor motor) types. Generalized convulsive seizures include tonic, clonic, and tonic-clonic types. Generalized tonic seizures usually occur in childhood and typically include impaired consciousness, muscle contraction of the face and trunk, flexion of the upper extremities, flexion or extension of the lower extremities, and postictal confusion. Generalized clonic seizures usually begin in childhood and include impaired consciousness, bilateral limb jerking, and postictal confusion. Generalized tonic-clonic seizures, commonly referred to as grand mal seizures, are characterized by a sudden loss of consciousness and tonic and clonic phases. Following this, there is muscular flaccidity and there may be incontinence. Consciousness returns gradually and the patient awakens in a confused state. Fatigue and headache are common. The EEG shows abnormal synchronized discharges

from both cerebral hemispheres in the tonic and clonic phases and diffuse slowing in the postictal period. Generalized nonconvulsive seizures include absence, myoclonic, and atonic types. Generalized absence seizures, commonly referred to as petit mal seizures, typically are 5- to 10-s episodes of impaired consciousness characterized by staring and unresponsiveness. Clonic movements, changes in postural tone, automatisms, and autonomic phenomena frequently accompany absence seizures. The patient quickly resumes normal consciousness, has no postictal confusion, and is generally unaware of the episode. Generalized myoclonic seizures are bilaterally synchronous jerks that can be single or repeated in trains. The muscles involved may be few and restricted to a body part (e.g., face) or extensive, involving all limbs. Most myoclonic seizures occur with no impairment of consciousness. Generalized atonic seizures, commonly referred to as drop attacks, consist of a sudden loss of tone in postural muscles often resulting in a fall. There is brief, mild impairment of consciousness and little postictal confusion.

When partial or generalized seizures represent a chronic condition and have no identifiable cause or a cause that cannot be cured by specific treatment, the seizure is caused by epilepsy. This distinguishes partial or generalized provoked seizures, which are not caused by epilepsy but by transient reversible causes (e.g., low blood sugar) or by conditions that can be cured by definitive therapy of an underlying disease state (e.g., removal of a brain tumor). Provoked seizures are usually isolated events but can be recurrent. There are many diverse causes of provoked seizures, including metabolic disorders (uremia, liver failure, and hypoglycemia), electrolyte disorders (hyponatremia and hypocalcemia), acid-base disorders, connective tissue and inflammatory disorders (systemic lupus erythematosus, rheumatic fever, and vasculitis), endocrine disorders, infections (meningitis and encephalitis), vascular disorders (transient ischemic attack and stroke), toxic disorders (medications, substances of abuse, and environmental toxins), head trauma, and pregnancy (eclampsia). It is important to distinguish provoked seizures from epileptic seizures because treatment of the underlying cause of provoked seizures will usually prevent further seizure occurrence, thereby obviating the need to treat the seizures with medication.

Stress and Seizures

Stress can be thought of as a diverse set of stimuli or conditions, intrinsic or extrinsic, that can perturb the physiological state of an organism. Categories of stressors include physiological (sleep deprivation,

hyperventilation, overexertion, dehydration, fever, illness, and menstrual cycle), pharmacological (medications and substances of abuse), environmental (extremes of temperature, noise, lighting, and smells), and psychological (situations resulting in fear, sadness, and anger). These stressors, alone or in combination, can alter the homeostatic mechanisms maintaining a normal physiological state, lower the seizure threshold, and result in the occurrence of a seizure. The specific means by which stress precipitates a seizure is not known. The subjective intensity of a stressor frequently does not necessarily correlate with the likelihood of seizure occurrence. Similarly, the temporal onset of a stressor(s) is often difficult to determine, and the latency-to-seizure occurrence is frequently variable. However, in many circumstances, a specific stressor can be identified and its intensity and temporal relationship to the seizure strongly suggest a direct relationship. A recent study in Croatia found that children with epilepsy in a war-affected area had more seizures than children in a non-war-affected area.

Some patients with well-controlled epileptic seizures experience an unexpected seizure after an unusual event has altered their typical routines and activities. Sleep deprivation is frequently the only identified change in a patient's routine that is associated with a breakthrough seizure; patients report having had little or no sleep the night before a seizure has occurred. This may have been due to simply going to bed later than usual, studying late for an exam, or awakening earlier than usual in the morning for an appointment. Often these circumstances can include other potential stressors such as caffeine, alcohol, and anxiety, which compound the clinical identification of only one stressor associated with the seizure occurrence. Hyperventilation, the increased frequency or depth of respirations, can be a reaction to stress and is used as a standard procedure in most routine EEGs to determine whether some forms of seizure activity can be induced. For example, generalized absence seizures and, less commonly, complex partial seizures at times can be easily precipitated with hyperventilation. Vigorous physical activity is usually associated with hyperventilation and can be the setting in which an unexpected seizure occurs. Seizures often occur during illness when fever and dehydration may be present. Correspondingly, overexertion with dehydration and the elevation of core body temperature can also be associated with seizure occurrence.

In addition to these circumstances that are unusual for the patient's day-to-day experience, certain physiological changes occur cyclically and may be related to seizure occurrence. The hormonal fluctuations of the menstrual cycle can contribute to seizure break-

through in susceptible women at the time of ovulation, perimenstrually, and throughout the luteal phase of the menstrual cycle (the time between ovulation and menstruation). Animal studies that assess gender-based differences in the association of stress and seizure risk (e.g., those of Chadda and Devaud) support the concept that the relationship between stress and seizure risk may be different in men and women. Pharmacological stressors associated with seizures in patients are often prescribed or over-the-counter medications. The seizures may be idiosyncratic or a result of an interaction with the metabolism of antiepileptic or other drugs that the patient takes chronically. As mentioned earlier, common substances such as caffeine and ethanol can be associated with seizure occurrence, as can other substances such as cocaine, amphetamines, and withdrawal from barbiturates or benzodiazepines.

Environmental stressors are at times difficult to identify, but an increased level of subjective stress can be associated with extremes of temperature, persistent and offending noises, lighting that can cause glare and subsequent headache, and particularly offensive smells. These situations are distinguished from reflex epilepsy, which is characterized by seizures evoked by specific stimuli and usually time-locked with the stimulus, such as in stroboscopic illumination. Psychological stressors, such as sudden anger, marked fear, and unexpected mourning, can result acutely in the occurrence of seizures. More chronic psychological factors, such as depression, anxiety, or hypomania, may fluctuate in intensity and predispose an individual to seizure occurrence.

The nature of stress-related seizures is such that attempts to study the relationship of a stressor and potential changes in the electrical brain activity of patients are very difficult. Although standard activating procedures such as photic stimulation and hyperventilation are used in routine EEG studies to precipitate ictal discharges and clinical events, the relationship of other stressors to seizure occurrence cannot be studied in this fashion. Patients undergoing evaluation for epilepsy surgery can be studied with a variety of electrode types, sometimes indwelling in the brain, and electrical and clinical seizures are monitored in a more comprehensive way. However, even in these very structured settings, the common stressors that have been discussed are not easily evaluated. The difficulty of quantifying the relationships between stress severity and latency to seizure onset is no less in these settings than it is in a typical clinical situation.

Basic science research into the functional anatomy of the brain involved in stress and seizure activity includes the amygdala, hippocampus, frontal cortex,

and dorsal raphe nucleus. These same brain structures are known to be involved in the propagation of electrical seizure activity in both animal models and in humans. Many studies have focused on the effects of hormones and neurotransmitters, including corticotropin releasing hormone (CRH), catecholamines, and serotonin, many of which can have effects on seizure threshold. Although a growing body of evidence implicates these neuromodulators or neurotransmitters as being involved in the expression of stress-related seizures, their precise mechanisms of action are not fully elaborated and clear clinical confirmation of their effects are not easily obtained. For example, in Klein and Sahoo's pilot study of patients with partial seizures and subjectively perceived stress-related seizures, the injection of adrenocorticotrophic hormone (ACTH) induced hypercortisolemia in all patients (cortisol is known to promote seizure activity in animal models), but was not associated with increases in interictal spike discharges on EEG or seizures in any patient. Despite the experimental difficulties inherent in this field of study, numerous studies are ongoing that will undoubtedly increase the knowledge base in both basic science and clinical areas.

Epilepsy as a Stressor

People with epilepsy frequently are faced with limitations in educational and employment opportunities, they may be restricted or prevented from driving, and they may regularly confront societal prejudices against them. Even well-meaning family members, friends, and coworkers may alter how they behave toward the person with epilepsy after witnessing a seizure or learning of this diagnosis. Feelings of loss of control are common, particularly for the patient with seizures that are incompletely controlled by medications. The high cost of medical care combined with limited employment opportunities increases the risk of financial hardship, which can increase life stress. In addition, epilepsy may be accompanied by cognitive deficits, decreased interpersonal skills, and psychiatric problems, which can cause increased stress in an individual's life and result in decreased resources for coping with life stress. All this probably contributes to the finding in a 2003 study that people with epilepsy have an average rate of suicide as high as 12%, compared to less than 1.5% in the general population.

NESLEs, or pseudoseizures, are generally understood to be psychological or psychiatric in origin rather than physiological. Suggested causes include affective and anxiety disturbances, panic disorder, posttraumatic stress disorder, conversion disorder, dissociative disorder, factitious disorder, psychosis, malingering,

hypochondriasis, and somatoform disorder. NESLEs are not uncommon, with as many as 400 000 individuals in the United States at risk for developing them in their lifetime. Many patients with epilepsy will also experience NESLEs, which can make the diagnosis even more challenging. Differentiating NESLEs from epileptic seizures is frequently performed by a multidisciplinary team in epileptology with expertise in neurology, neuropsychology, psychiatry, and nursing. Video-EEG monitoring can demonstrate that some behavioral events do not have an associated ictal discharge on the EEG and are often consistent with the diagnosis of NESLEs.

Interventions

Epileptic seizures are usually treated with antiepileptic drugs (AEDs). An AED is chosen based on the patient's type of epilepsy and/or seizure(s). Several of the available AEDs are effective in treating both partial and generalized seizures; however, most AEDs have specific indications for use and may be ineffective in the treatment of particular seizure types. An AED may be prescribed as monotherapy when it is effective in controlling seizures. When one AED is inadequate in controlling seizures, combinations of AEDs (polytherapy) are used. AEDs used commonly in the treatment of partial seizures include phenytoin, carbamazepine, lamotrigine, topiramate, oxcarbazepine, levetiracetam, and zonisamide; phenobarbital and valproic acid are also used for the treatment of partial seizures. Convulsive generalized seizures are treated with phenytoin, phenobarbital, valproic acid, and benzodiazepines. Nonconvulsive generalized seizures are treated with valproic acid and ethosuximide. The treatment of generalized seizure disorders will often include many of the AEDs used for treatment of partial seizures.

When seizure disorders are inadequately controlled with the use of AEDs, epilepsy surgery and the vagus nerve stimulator (VNS) may be used. Epilepsy surgery is performed for some seizure disorders, typically when the seizures have a focal onset in an area of the brain determined to be safe for surgical intervention. The VNS is a device implanted in the chest and connected to the vagus nerve by a wire; it is designed to deliver current pulses at specific intervals to decrease the likelihood of seizure onset. Patients can also activate the device to deliver a current pulse when they recognize the onset of a seizure in an attempt to abort or lessen the severity of the seizure. The VNS is programmed by the patient's epileptologist and adjusted as required by the frequency and severity of the seizures.

Helping patients cope with life stress and the stress that results from having epilepsy may decrease the likelihood of seizures and enhance quality of life. Multiple treatment modalities can be effective for managing stress, including progressive relaxation, biofeedback, and meditation. However, this relationship may be more complicated than merely working to reduce stress and more research on which treatments are most helpful clearly needs to be done. For example, Jaseja suggested that meditation may result in brain changes that can increase the risk of developing epilepsy and the severity and frequency of attacks. This is in conflict with the results of other studies with various behavioral therapies that suggest that these therapies can reduce seizure frequency and/or increase satisfaction with other aspects of life. In addition, one study demonstrated that patients, particularly those who were described as highly reactive to stress, were able to reduce seizure frequency using a behavioral self-management program that taught self-regulation of slow cortical potentials. Patients may also benefit from cognitive behavioral therapy, which encourages them to alter their thinking and behaviors that may be contributing to anxiety, depression, and feeling overwhelmed. Interpersonal psychotherapy and assertiveness training can assist patients in learning how to interact in healthier ways with family members, coworkers, and peers and may positively impact levels of perceived stress. Medication for psychiatric problems can result in better control of anxiety and of affective and thought disorder symptoms and thus reduce life stress. A recent study found that the antidepressant fluoxetine reduced seizures in mice that were and were not exposed to stressful conditions. For patients with specific problems related to hyperventilation, relaxation training may be helpful. Patients with sleep disturbances may benefit from learning basic sleep hygiene, increasing activity levels, and treating anxiety and affective disorders that may be disrupting sleep. Gupta's study suggested that training that specifically addresses cognitive deficits in patients with epilepsy can be helpful in improving cognitive performance, which may in turn reduce life stress. Support groups, family therapy, psychoeducation, and supportive psychotherapy are beneficial for patients coping with epilepsy and the daily challenges and difficulties that attend it. Encouraging patients to make use of special programs and organizations specifically developed to assist people with epilepsy (e.g., the Epilepsy Foundation) can be helpful.

The treatment of patients with NESLE is determined by the suspected cause. Patients frequently are found to have problems with depression, anxiety,

or thought disorders, and treatment with antidepressants, anxiolytics, or antipsychotic medication, as well as psychotherapeutic and behavioral intervention, may be useful. Psychological and personality testing, combined with an evaluation of environmental factors that may be contributing to this disorder, contribute in formulating appropriate intervention strategies. For patients with both epileptic seizures and NESLEs, different strategies may be suggested for treating the seizures and the NESLEs. Identifying stressors that may be precipitating or maintaining the NESLEs may be particularly effective in reducing both the frequency and the impact of these events on the patients' lives.

See Also the Following Articles

Psychotherapy; Stress Management and Cardiovascular Disease; Stress Management, CAM Approach; Quality of Life.

Further Reading

- Betts, T. (1992). Epilepsy and stress. *British Medical Journal* 305, 378–379.
- Bosnjak, J., Vukovic-Bobic, M. and Mejaski-Bosnjak, V. (2002). Effect of war on the occurrence of epileptic seizures in children. *Epilepsy and Behavior* 3, 502–509.
- Chadda, R. and Devaud, L. L. (2004). Sex differences in effects of mild chronic stress on seizure risk and GABA_A receptors in rats. *Pharmacology, Biochemistry, and Behavior* 78, 495–504.
- Gates, J. R. (1998). Diagnosis and treatment of nonepileptic seizures. In: McConnell, H. W. & Snyder, P. J. (eds.) *Psychiatric comorbidity in epilepsy: basic mechanisms, diagnosis, and treatment*, pp. 187–204. Washington D.C.: American Psychiatric Press.
- Goldstein, L. H. (1997). Effectiveness of psychological interventions for people with poorly controlled epilepsy. *Journal of Neurology, Neurosurgery & Psychiatry* 63, 137–142.
- Gupta, A. (2003). Cognitive retraining in epilepsy. *Brain Injury* 17, 161–174.
- Haut, S. R., Vouyiouklis, M. and Shinnar, S. (2003). Stress and epilepsy: a patient perception survey. *Epilepsy and Behavior* 4, 511–514.
- Homayoun, H. and Dehpour, A. R. (2004). Differential contribution of cholecystokinin receptors to stress-induced modulation of seizure and nonception thresholds in mice. *Pharmacology, Biochemistry, and Behavior* 78, 209–215.
- Jaseja, H. (2005). Meditation may predispose to epilepsy: an insight into the alteration in brain environment induced by meditation. *Medical Hypotheses* 64, 464–467.
- Jones, J. E., Hermann, B. P., Barry, J. J., et al. (2003). Rates and risk factors for suicide, suicidal ideation, and suicide attempts in chronic epilepsy. *Epilepsy and Behavior* 3(supplement), S31–S38.
- Klein, P. and Sahoo, S. (2005). Effect of ACTH-induced hypercortisolemia on the EEG in patients with stress-related epilepsy. *Epilepsy and Behavior* 6, 187–190.
- Lai, C. W. and Trimble, M. R. (1997). Stress and epilepsy. *Journal of Epilepsy* 10, 177–186.
- Nakken, K. O., Solaas, M. H., Kjeldsen, M. J., et al. (2005). Which seizure-precipitating factors do patients with epilepsy most frequently report? *Epilepsy and Behavior* 6, 85–89.
- Pericic, D. and Svob, S. D. (2005). Anticonvulsant effects of acute and repeated fluoxetine treatment in unstressed and stressed mice. *Brain Research* 1033, 90–95.
- Strehl, U., Kotchoubey, B., Trevorrow, T., et al. (2005). Predictors of seizure reduction after self-regulation of slow cortical potentials as a treatment of drug-resistant epilepsy. *Epilepsy and Behavior* 6, 156–166.

Epinephrine See: Adrenaline.

Erectile Dysfunction *See: Impotence, Stress and.*

Estrogen

S B Miller, E Neumark and A Sita
Concordia University, Montreal, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by S B Miller, E Neumark and A Sita, volume 2, pp 71–73,
© 2000, Elsevier Inc.

The Stress Response: Sympathetic Nervous System and the
Hypothalamic-Pituitary-Adrenocortical Axis

Estrogen and the Sympathetic Nervous System

Estrogen and the Hypothalamic-Pituitary-Adrenocortical
Axis

Sex Differences in the Stress Response: The Effects of
Estrogen on Cognitive Function

Implications for Health

Glossary

<i>Amenorrhea</i>	The absence of menses.
<i>Endothelium</i>	The innermost layer of blood vessels; made up of squamous epithelial cells.
<i>Genomic mechanisms</i>	Mechanisms involving the regulation of gene expression.
<i>Gonadotropins</i>	The pituitary hormones that stimulate the ovaries and testes (i.e., luteinizing hormone and follicle stimulating hormone).
<i>Gonadotropin releasing hormone</i>	A hormone secreted by the hypothalamus that acts on the pituitary to control the secretion of the gonadotropins.
<i>High-density lipoprotein (HDL) cholesterol</i>	A lipoprotein cholesterol complex that transports cholesterol from cells in peripheral tissues to the liver for excretion from the body.
<i>Low-density lipoprotein (LDL) cholesterol</i>	A lipoprotein cholesterol complex that delivers cholesterol to cells in peripheral tissues in order to support metabolic and structural needs.

Estrogen is a class of steroid hormones produced mainly by the ovaries. It is often referred to as the

female sex hormone; however, this is a misnomer because estrogen is also found in males, albeit in much smaller amounts than in females. In premenopausal women, the most common and potent form of estrogen is estradiol, followed by estrone and estriol. Estradiol secretion rates vary according to the phases of the menstrual cycle; it is, 0.07 mg day⁻¹ in the early follicular phase, 0.6 mg day⁻¹ preceding ovulation, and 0.25 mg day⁻¹ during the middle of the luteal phase. In men, the estradiol production rate is approximately 0.05 mg day⁻¹. Estrogens have numerous well-documented effects on the female reproductive system. They facilitate the growth of the ovarian follicles, increase the motility of the uterine tubes, maintain the cyclic changes of the uterus, stimulate ductal growth in the breasts, inhibit the secretion of follicle stimulating hormone, and play a major role in the development of female secondary sex characteristics. Estrogens exert the majority of their effects via intracellular receptors that bind to DNA; however, recently it has been shown that estrogens may also act through nongenomic mechanisms (e.g., alterations in cell membrane permeability).

The Stress Response: Sympathetic Nervous System and the Hypothalamic-Pituitary-Adrenocortical Axis

In addition to their well-known effects on the reproductive system, estrogens interact with the two major systems involved in the stress response – the sympathetic nervous system and the hypothalamic-pituitary-adrenocortical (HPA) axis. The activation of the sympathetic nervous system results in the release of noradrenaline from the various organ systems it innervates, with the exception of the adrenal medulla, which releases adrenaline. The release of noradrenaline and adrenaline causes a number of bodily changes including an increase in heart rate and blood pressure and a rise in serum cholesterol and free-fatty acid levels. With regards to the HPA axis, when a stressor is encountered, corticotropin releasing hormone

(CRH) is secreted from the hypothalamus. CRH stimulates the pituitary to release adrenocorticotrophic hormone (ACTH), and, in turn, ACTH stimulates the adrenal gland to release glucocorticoids. Cortisol is the main glucocorticoid in humans, whereas corticosterone is the main glucocorticoid in rodents. Approximately 95% of circulating glucocorticoids are bound to corticosteroid-binding globulin (CBG), a protein synthesized by the liver. The unbound 5% constitutes the biologically active portion of glucocorticoids.

Estrogen and the Sympathetic Nervous System

There is a significant body of animal studies and, to a lesser extent, human studies demonstrating that estrogens affect sympathetic activity. Animal studies suggest that estrogens decrease β receptor sensitivity to catecholamines and may diminish the secretion of epinephrine from the adrenal medulla. Estrogens modify α receptor sensitivity and noradrenergic synthesizing and degrading enzymes. In both animal and human studies, estrogens have been found to affect cardiovascular activity. For example, estrogens increase blood flow by increasing vasodilation. How this is achieved is at present unclear; however, a number of possible mechanisms exist. Estradiol may enhance endothelium-dependent and -independent relaxation, and it may have a direct effect on the vascular smooth muscle. Estrogens may also affect cardiac function. Estradiol receptors have been found in the rat heart and estradiol has been found to affect both heart rate and contractility. Estrogens also affect cholesterol levels. Following menopause, women's estradiol levels are low and a decrease in HDL cholesterol and an increase in LDL cholesterol levels are observed. Moreover, estrogen replacement in postmenopausal women has been found to partially reverse these changes in cholesterol levels.

The effects of estrogens on cardiovascular, noradrenergic, and lipid responses to stress are at present unclear. To date, only a few studies have examined this relation. On the one hand, endogenous estrogen levels have been found to be negatively related to heart rate, cardiac output, and systolic blood pressure responses to stress in healthy premenopausal women. On the other hand, exogenous estrogens have been found to either decrease or have no effect on postmenopausal women's cardiovascular responses to stress and to increase premenopausal women's cardiovascular responses to stress. Similarly, estrogens have been found to increase, decrease, or have no effect on adrenaline responses to stress in humans. The equivocal results may be due to the type of estrogen used and

the method of administration. For example, some studies have used only estradiol, whereas other studies have used preparations containing estradiol and other estrogens. There is evidence that estradiol may have different physiological effects when compared to other types of estrogens. Also, transdermally applied estradiol may be more likely to cause a decrease in blood pressure than orally administered estrogens, although recent research has found decreases following oral administration. Increased plasma levels of an oxytocin intermediate peptide (oxytocin-glycine), resulting from oral estrogen replacement in postmenopausal women, has been linked to reduced blood pressure and vascular resistance during a speech stressor task. Acute transdermal estrogen administration to healthy postmenopausal women, however, had no statistically significant differential effects on heart rate, blood pressure, or noradrenaline spillover, during a mental stressor (Stroop test) or physical stressor (cold pressor) task.

The effect of estrogens on lipid responses to stress has been directly and indirectly examined in only a few studies. For example, there is evidence that estrogen administration attenuates free fatty acid and glycerol responses to stress. Also, women who have undergone a surgical menopause have been found to exhibit greater LDL cholesterol responses to stress than women with intact ovaries.

Estrogen and the Hypothalamic-Pituitary-Adrenocortical Axis

There is a bidirectional relation between the HPA axis and the reproductive system. The activation of the HPA axis inhibits the functioning of the hypothalamic-pituitary-gonadal axis in mammals, and there is some evidence for this effect in humans as well. Glucocorticoids have been found to inhibit the secretion of gonadotropin releasing hormone and the gonadotropins, which leads to a decrease in estrogen synthesis. Glucocorticoids also inhibit estradiol-stimulated uterine growth. In turn, there is evidence from animal studies that estrogens affect HPA axis function.

In rats, estrogen administration increases baseline corticosterone and ACTH secretions and corticosterone and ACTH responses to stress. Furthermore, corticosterone and ACTH levels are at their peak during proestrus, when circulating levels of estrogen are high. Immobilization stress during a 2-h period has been shown to result in an elevation in plasma ACTH in non-estradiol-treated rats and a lesser elevation in estradiol-treated rats. In this study, blood pressure remained elevated during the 2-h period in the control group; however, the elevation was reduced after 70 min in the treated rats. Other stress

responses were either unaffected, for example, an equivalent rise in tyrosine hydroxylase mRNA in both the adrenal medulla and the nucleus of the solitary tract of the control and treated rats, or reversed, for example, a decrease in the dopamine β -hydroxylase mRNA levels in the locus coeruleus. The administration of diethylstilbestrol to ovariectomized rats suppressed c-Fos accumulation in prolactin-releasing peptide, producing A2 noradrenergic neurons, while the rats were in a restraint stress condition. These A2 neurons located in the medulla oblongata are known to mediate the stress response.

In humans, the effects of estrogens on the HPA axis response to stress are less clear. Only a few studies have been conducted. Although estrogen administration has been found to increase cortisol levels in most studies, this has not been a uniform finding. Estrogens affect the HPA axis at multiple levels. Estrogens increase HPA axis activity by modulating corticosterone negative feedback in the hippocampus, binding to CRH cells in the hypothalamus, increasing the secretion of corticosterone by the adrenals, and increasing anterior pituitary sensitivity to CRH. Estrogens also stimulate the production of CBG by the liver. With higher levels of CBG, more glucocorticoids must be secreted in order to maintain the same amount of glucocorticoids in the unbound biologically active state. Thus, this may account for the higher levels of circulating glucocorticoids in females. Sex differences in HPA axis function may be the result of both central differences and adrenal sensitivity to ACTH. Kajantie and Phillips have proposed a framework to explain the discrepancies found with regard to estrogen and its role in the psychosocial stress response. Attenuation may be the result of a decrease in arginine vasopressin (AVP) mediated by estrogen β receptors and an increase in CRH mediated by estrogen α receptors.

Sex Differences in the Stress Response: The Effects of Estrogen on Cognitive Function

Stress has been shown to have effects on both animal and human cognitive functioning, with sex differences being reported. Restraint-stressed female rats showed improved spatial memory; however, male rats showed small impairments. Estrogen may not only moderate the HPA axis and alter corticosterone release, it may counteract the effects of corticosterone. Given the lower level of estrogen in males, they may be more negatively impacted by high levels of stress-induced corticosterone.

Findings related to spatial learning and memory in nonstressed rats have been controversial, and recently researchers reported that physiological levels of estra-

diol did not appear to have any effect and that, at higher levels, performance impairments have been observed. Similar impairment has been found with tasks that depend on reference memory. Hippocampal neurons that are sensitive to estrogen(s) may play a role in influencing estrogen-level-mediated task-solving cognitive strategies in female rats. There is limited evidence that the circulating level of estradiol is an important component of sex differences in both the stress response and its impact on cognitive function. Stress-induced anxiety in ovariectomized rats has recently been shown to be attenuated following 17 β -estradiol replacement. Using a benzodiazepine inverse agonist (FG7142), which activates stress systems and typical stress responses including elevated corticosterone release, catecholamine turnover, heart rate, and blood pressure, researchers found that female rats were more sensitive to pharmacologically induced stress than male rats, but only under conditions of high estrogen, showing prefrontal cortex (PFC) impairment. This sex difference was also observed during mild levels of natural stress, with high-estrogen-level females showing PFC impairment. Behaviors consistent with reduced stress levels, such as social investigation by female rats, were shown to increase following long-term estrogen replacement, along with improved social recognition and reduced corticosterone stress response.

Implications for Health

The interaction of estrogens with the systems involved in the stress response has a number of implications for health. For example, it is well known that severe chronic stress interferes with the reproductive system and may cause amenorrhea. Estrogen is also known to protect women from cardiovascular disease, as evidenced by the increase in cardiovascular disease postmenopausally. Although it is known, for example, that the favorable effect of estrogens on resting cholesterol levels is one mechanism by which premenopausal women are protected, estrogen may also lower risk by decreasing the cardiovascular, noradrenergic, and lipid responses to stress. Estradiol's anti-inflammatory effects may act as a neuronal protector by moderating inflammatory factors and reducing the ability of inflammatory cells to enter the central nervous system (CNS). In addition, estrogens mitigate the effects of neuronal oxidative stress. Collectively, these CNS estrogenic effects are thought to play a protective role in several neurodegenerative pathologies, including Alzheimer's disease and Parkinson's disease.

Conversely, stress-induced lower levels of estrogen have recently been hypothesized to have a positive health effect. Higher levels of self-reported daily

stress were associated with a reduced risk of primary breast cancer, especially among women receiving hormone therapy. Overall, however, the chronic dysregulation of the stress system may have greater negative than positive health impact.

Recently, Kajantie and Phillips suggested that evolutionary forces led to a stress-response attenuation system that functions during pregnancy. Prenatal stress has been shown to have negative effects on the fetus, effects that have continued developmental sequelae. Estrogen, which rises during pregnancy to levels that are on the order of 100 times higher than peak nonpregnancy levels, clearly acts protectively through its effects on the autonomic nervous system. What is less clear is its effects on the HPA axis; however, during pregnancy AVP levels are low, which may attenuate the HPA axis response. During pregnancy AVP appears to have a greater role in regulating cortisol secretion than does CRH, which normally regulates cortisol. Further research in this area may help elucidate the recent findings related to the seemingly trimester-dependent effects of maternal prenatal stress and developmental sequelae.

See Also the Following Articles

Amenorrhea; Androgen Action; Menopause and Stress; Steroid Hormone Receptors; Sympathetic Nervous System.

Further Reading

- Adachi, S., Mochiduki, A., Nemoto, H., et al. (2005). Estrogen suppresses the stress response of prolactin-releasing peptide-producing cells. *Neuroscience Letters* 380(3), 311–315.
- Amantea, D., Russo, R., Bagetta, G., et al. (2005). From clinical evidence to molecular mechanisms underlying neuroprotection afforded by estrogens. *Pharmacological Research* 52(2), 119–132.
- Carey, M. P., Deterd, C. H., de Koning, J., et al. (1995). The influence of ovarian steroids on hypothalamic-pituitary-adrenal regulation in the female rat. *Journal of Endocrinology* 144, 311–321.
- Farhat, M. Y., Lavigne, M. and Ramwell, P. W. (1996). The vascular protective effects of estrogen. *FASEB Journal* 10, 615–624.
- Handa, R. J., Burgess, L. H., Kerr, J. E., et al. (1994). Gonadal steroid hormone receptors and sex differences in the hypothalamic-pituitary-adrenal axis. *Hormones and Behavior* 28(4), 464–476.
- Kajantie, E. and Phillips, D. I. W. (2006). The effects of sex and hormonal status on the physiological response to acute psychosocial stress. *Psychoneuroendocrinology* 31, 151–178.
- Korol, D. L. (2004). Role of estrogen in balancing contributions from multiple memory systems. *Neurobiology of Learning and Memory* 82, 309–323.
- Light, K. C., Grewen, K. M., Amico, J. A., et al. (2005). Oxytocinergic activity is linked to lower blood pressure and vascular resistance during rest and stress in postmenopausal women on estrogen replacement. *Hormones and Behavior* 47, 540–548.
- Lunga, P. and Herbert, J. (2004). 17-Oestradiol modulates glucocorticoid, neural and behavioural adaptations to repeated restraint stress in female rats. *Neuroendocrinology* 16(9), 776–785.
- Magiakou, M. A., Masmnkos, G., Webster, E., et al. (1997). The hypothalamic-pituitary-adrenal axis and the female reproductive system. *Annals of the New York Academy of Sciences* 816, 42–56.
- McEwen, B. S. (1991). Non-genomic and genomic effects of steroids on neural activity. *Trends in Pharmacological Sciences* 12, 141–147.
- Nielsen, N. R., Zhang, Z. F., Kristensen, T. S., et al. (2005). Self reported stress and risk of breast cancer: prospective cohort study. *British Medical Journal* 331, 548–550.
- Sand, P. M. (1990). Ovarian hormones and the circulation. *Maturitas* 590, 287–290.
- Serova, L. I., Maharjan, S. and Sabban, E. L. (2005). Estrogen modifies stress response of catecholamine biosynthetic enzyme genes and cardiovascular system in ovariectomized female rats. *Neuroscience* 132, 249–259.
- Shansky, R. M., Glavis-Bloom, C., Lerman, D., et al. (2004). Estrogen mediates sex differences in stress-induced prefrontal cortex dysfunction. *Molecular Psychiatry* 9, 531–538.

- Skafar, D. F., Xu, R., Morales, J., et al. (1997). Female sex hormones and cardiovascular disease in women. *Journal of Clinical Endocrinology and Metabolism* 82(13), 3913–3918.
- Sofowora, G. G., Singh, I., He, H. B., et al. (2005). Effect of acute transdermal estrogen administration on basal, mental stress and cold pressor-induced sympathetic responses in postmenopausal women. *Clinical Autonomic Research* 15, 193–199.

- Tang, A. C., Nakazawa, M., Romeo, R. D., et al. (2005). Effects of long-term estrogen replacement on social investigation and social memory in ovariectomized C57BL/6 mice. *Hormones and Behavior* 47(3), 350–357.
- Varga, H., Németh, H., Tóth, T., et al. (2002). Weak if any effect of estrogen on spatial memory in rats. *Acta Biologica Szegediensis* 46, 13–16.
- Wren, B. G. (1992). The effect of oestrogen on the female cardiovascular system. *Medical Journal of Australia* 157, 204–208.

Ethanol and Endogenous Opioids

D K Sarkar

Rutgers University, New Brunswick, NJ, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D K Sarkar and N Boyadjieva, volume 2, pp 74–78, © 2000, Elsevier Inc.

whereas others endure extremely severe reactions that can be life-threatening.

Pharmacokinetics of Ethanol

Alcohol Reinforcement

Alcoholism

Alcohol Effects on the Developing Brain

Glossary

<i>Alcohol withdrawal syndrome</i>	Symptoms experienced by alcoholics when they stop drinking, including profound anxiety, body shakes, intense hyperactivity, sleep disturbances, hallucinations, and seizures.
<i>Alcoholism</i>	A disease that develops when a person loses control over his or her drinking of alcohol and develops tolerance for and dependence on it; also called alcohol dependence.
<i>Craving</i>	The hunger for alcohol before drinking begins.
<i>Impaired control over drinking</i>	The difficulty that an alcoholic experiences in stopping once his or her drinking has started.
<i>Physical dependence</i>	An adaptive state manifested by intense physical disturbances that occur when drinking is discontinued. Dependence necessitates continued drinking to relieve the discomfort of alcohol withdrawal syndrome. During withdrawal, some alcoholics experience only mild symptoms,

Ethanol is an organic chemical that belongs to the chemical class of alcohols; it is also called ethyl alcohol, grain alcohol, methyl carbinol, and ethyl hydrate. The molecular formula of ethanol is C_2H_6O . Ethanol is used as a food, a drug, and in the manufacture of industrial and consumer products. The use of alcohol as a beverage at social events and festivals dates back to the ancient Egyptian, Mesopotamian, Greek, and Roman civilizations. Moderate alcohol use, up to two drinks per day for men and one drink per day for non-pregnant women and older people, is not harmful for most adults.

Pharmacokinetics of Ethanol

Ethanol is rapidly absorbed on ingestion; approximately 80–90% of ethanol is absorbed within 30–60 min, although food may delay the absorption for a longer period of time (4–6 h). Ethanol is both water- and lipid-soluble, easily penetrates the blood-brain barrier, and distributes into the brain and total body fluid. Approximately 90–98% of any consumed ethanol is oxidized by the liver, and 5–10% is excreted by the kidneys and lungs.

Ethanol can be a central nervous system (CNS) depressant. Mild ethanol intoxication (blood alcohol concentration, BAC, 0.05–0.15%) leads to the impairment of visual acuity, changes in mood and personality, decreased reaction time, and muscular incoordination. Severe ethanol intoxication (BAC 0.3–0.5%) exacerbates these problems and in addition brings about hypothermia, vomiting, nausea, hypoglycemia and convulsions, depressed reflexes, and death from

respiratory or circulatory failure. Alcohol consumption during pregnancy can lead to a congenital malformation of the offspring, which is known as fetal alcohol syndrome. Affected infants often show mental deficiency, microcephaly, and irritability.

Alcohol Reinforcement

Alcohol's basic action on the CNS causes pleasant subjective effects or a high feeling. This mild euphoria produced by alcohol often reinforces alcohol drinking. Alcohol also relieves anxiety from stressful life situations by reducing the aversive physiological stimuli associated with anxiety and stress. The reinforcing effects of alcohol involve chemical communication between various neuronal systems, including opiodergic neurons in the CNS.

Alcoholism

Alcoholism, or alcohol dependence, normally refers to when a person loses control over his or her alcohol drinking and develops a tolerance to and dependence on alcohol. Whereas moderate alcohol consumption may relieve stress and improve psychological well-being, alcohol dependence and depression are a prevalent combination of psychiatric disorders among patients seeking treatment. The highest psychiatric comorbidity for people with alcohol dependence appears to be affective anxiety and antisocial personality disorders.

Behavioral Aspects

The stress-relieving action of alcohol has been well documented in human subjects. This anxiolytic action of alcohol led to the belief that the motivation for abusing this substance may be prompted by individuals' needs to cope with (relieve) stress. Stress is often defined as a process involving the perception of, interpretation of, and response to harmful, threatening, and challenging events. Psychosocially, it is defined as tensions or challenges that an organism faces; physiologically, it is the biological responses upregulated by exposure to stressors. People often use drugs to enhance mood and alleviate emotional distress, and the motivation to enhance mood is great in acute and chronic stress states. There is a link between the maladaptive stress response and drug addiction in vulnerable individuals who are exposed to stress. The maladaptive stress response includes high or low reactivity and sensitivity to stress stimuli, a slow biological recovery after the initiation of stress, and poor cognitive and behavioral coping. Individuals at high risk for alcoholism show a heightened physiologi-

cal and subjective sensitivity to the stress-reducing effects of ethanol. In addition, genetic and individual vulnerability factors may influence the maladaptive stress response to increase the use of abusive substances. Clinical studies have suggested a link between genetically related risks for alcoholism and an alteration in the hypothalamic-pituitary-adrenal (HPA) axis. It has been shown that the capacity of the HPA axis to respond to stress is more labile in family-history-positive individuals, and this may promote their use of ethanol to relieve stress.

Biochemical Aspects

Endogenous opioid peptide mediation of ethanol action Opioid deficiency may be partly responsible for the maladaptive stress response that may promote alcohol abuse in vulnerable humans following stress. A slow biological recovery of the HPA axis after the initiation of stress may be explained by the opioid deficiency hypothesis. After the perception of stress, the neuronal input converges on the corticotropin releasing hormone (CRH) neurons. These neurons secrete CRH into hypophyseal portal circulation. CRH then enters the pituitary and acts primarily on corticotrophs (and melanotrophs) to stimulate the release of adrenocorticotrophic hormone (ACTH) and β -endorphin from the pituitary gland. ACTH acts on the adrenal gland to release glucocorticoids. When the glucocorticoids reach threshold levels, it prevents CRH secretion and brings about homeostasis. The CRH-producing neurons receive signals through three major neurotransmitter systems: stimulatory input from serotonergic-producing neurons and inhibitory inputs from gamma-aminobutyric acid (GABA) neurons and β -endorphin neurons. Because of the powerful inhibitory input of β -endorphin neurons on CRH release, an acquired or inborn abnormality in β -endorphin activity can be an important determinant of the magnitude of the stress response. More β -endorphin activity allows the constrained stress response, whereas less β -endorphin activity allows for a more labile stress response. If individuals at high risk for alcoholism have reduced functional β -endorphin neurons, these opioid differences may explain why these individuals show slow biological recovery after the initiation of stress.

Three major groups of endogenous opioid peptides have been isolated and characterized: the endorphins from the β -endorphin/ACTH precursor known as proopioidmelanocortin, the enkephalins from the pro-enkephalin precursor, and the dynorphins and neo-endorphins from the prodynorphin precursor. The 31-amino-acid peptide β -endorphin, with a reasonably high affinity for the mu and delta forms of the

opiate receptors, is present in high concentrations in the hypothalamus. The β -endorphin-producing perikarya are located mainly in the ventromedial arcuate nucleus region, which projects to widespread brain structures, including many areas of the hypothalamus and the limbic system, where this opioid peptide has been proposed to function as a neurotransmitter or neuromodulator regulating a variety of brain functions. These brain functions include psychomotor stimulation; positive reinforcement; adaptive processes; drinking, eating, and sexual behaviors; pituitary function; thermoregulation; nociception; and mood. It is generally believed that the endogenous opioid system mediates some of the reinforcing properties of ethanol. Neurobiological studies indicate that alcohol alters opioid peptide systems. Acute alcohol administration increases endorphin and enkephalin gene expression in discrete brain regions and increases the release of these peptides from the brain and pituitary in rodents. Chronic or binge alcohol administration decreases proopiomelanocortin gene expression and alters the diurnal rhythm of proopiomelanocortin gene expression, β -endorphin release, hypothalamic levels of β -endorphin, and opioid receptor affinity and binding. Opioid-receptor antagonists have been shown to decrease ethanol consumption. The extended amygdala is considered to be a site of the opioid action because a high proportion of cells in the medial nucleus of the amygdala and the bed nucleus of the stria terminalis express μ - and δ -opioid receptors. The extended amygdala is also a major brain area involved in excessive ethanol consumption.

The abnormal expression of opioid receptors in the brain is connected with the volitional ethanol consumption. μ -Opioid receptor densities have been shown to be higher in the extended amygdala in alcohol-preferring rat and mouse strains. Selective μ -opioid receptor antagonists reduced ethanol drinking in selectively bred alcohol-preferring rats and normal rats. Targeted gene mutation (knockout) strategies have produced mice that lack either μ - or δ -opioid receptors. μ -Opioid receptor knockout mice avoid alcohol, but δ -opioid receptor knockout mice self-administer more alcohol. The excess alcohol administration in δ -opioid receptor knockout mice is considered to be due to increased μ -opioid receptor activity, because μ - and δ -opioid receptor systems may oppose one another.

Cellular action of ethanol on hypothalamic opioid peptides Using primary cultures of hypothalamic neurons, it has been shown that low concentrations of ethanol acutely stimulate β -endorphin release from cultured hypothalamic neurons. Similarly, a single administration of ethanol acutely stimulated the plasma levels of β -endorphin in male rats. However, an inhibitory effect of chronic ethanol on β -endorphin secretion has been demonstrated in both *in vivo* and *in vitro* studies. Hypothalamic proopiomelanocortin gene expression also increases following acute ethanol, but it decreases following chronic ethanol in both *in vivo* and *in vitro* systems. These functional changes in β -endorphin neurons following ethanol treatments parallel many behavioral changes observed following ethanol use in humans. It is proposed that by enhancing opioid activity, ethanol could compensate for

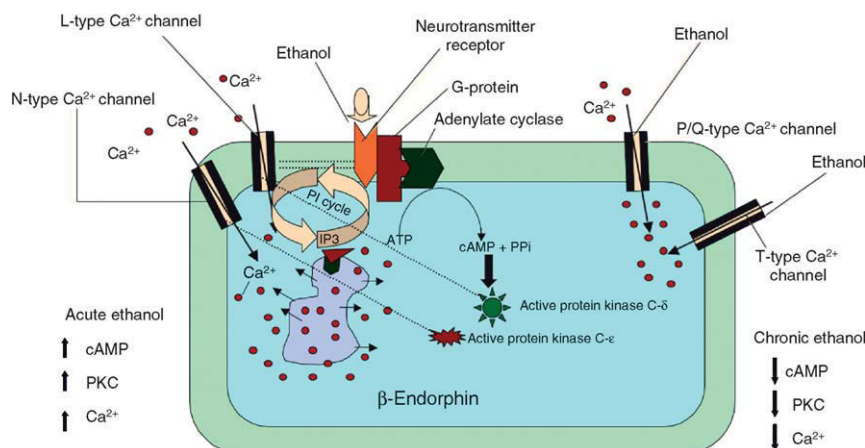


Figure 1 The effects of single or moderate ethanol administration (acute ethanol) and binge or long-term ethanol administration (chronic ethanol) on the β -endorphin system. Acute ethanol enhances and chronic ethanol suppresses β -endorphin synthesis or release by altering the activity of key intracellular transducers (cAMP, Ca^{2+} , and PKC). Up arrows symbolize upregulation; down arrows symbolize downregulation; IP₃, inositol triphosphate; PI cycle, phosphatidylinositol cycle; PPI, inorganic pyrophosphate.

constitutive deficiencies of endogenous opioids that may contribute to ethanol self-administration. In increasing the sensitivity of the opioid peptide as well as other neurotransmissions, ethanol could increase the rate of ethanol self-administration. Ethanol may produce tolerance and physical dependence, in part, by altering opioid neurotransmissions.

Studies using hypothalamic cells in primary cultures have elucidated primary transduction pathways for the ethanol action on β -endorphin release. These studies suggest that acute ethanol stimulates β -endorphin secretion, possibly by activating the cAMP cascade, protein kinase C (PKC), and Ca^{2+} -dependent mechanisms (Figure 1). The ethanol action on the cAMP system appears to involve the inhibition of adenosine uptake, leading to increasing extracellular levels of adenosine and the activation of membrane adenosine receptors and, consequently, to an increase in cAMP production and β -endorphin secretion. Acute ethanol action on β -endorphin release also involves calcium entry via the activation of voltage-dependent calcium channels. Furthermore, acute ethanol stimulates the expression and translocation of δ - and ϵ -PKC isoforms to activate β -endorphin release. In contrast to this, chronic ethanol causes the development of tolerance and desensitization of β -endorphin neurons due to reduction in the production and translocation of δ - and ϵ -PKC isoforms and the production of intracellular levels of cAMP. The inhibitory effect of chronic ethanol on cAMP is a result of the development of heterologous desensitization of adenosine, prostaglandins, and adrenergic receptors on these neurons. Hence, the changes in signal transduction could be critical for the adaptation of β -endorphin neurons to chronic ethanol exposure (Figure 1).

Genetic Aspects

Findings from twin, adoption, and cross-fostering studies and pedigree analyses have indicated that genetic and environmental factors play an important role in determining an individual's vulnerability to alcoholism. Although a multitude of biological differences have been found as a function of a family history of alcoholism, a number of studies suggest that diminished endogenous hypothalamic opioid activity and opioid receptor polymorphisms may increase an individual's vulnerability to becoming an alcoholic.

Genetic studies have shown that a genetic predisposition toward alcohol drinking is accompanied by an altered plasma β -endorphin response to alcohol. These studies have shown an altered opioid regulation of the HPA axis in family-history-alcoholism-positive individuals. Polymorphisms in opioid receptor genes and other opioid-regulated neurotransmitters has shown to be partly involved in the altered opioid regulation

of the HPA axis in family-history-positive individuals. A common A118G nucleotide exchange in exon 1 of the μ -opioid receptor has been shown to cause an Asn40Asp substitution polymorphism in the extracellular N-terminal domain of the μ -opioid receptor. It is believed that opioid receptor polymorphisms decrease dopamine secretion, which contributes to the altered capacity of the HPA axis to respond to stress, and that this may promote the use of ethanol to relieve stress.

Alcohol Effects on the Developing Brain

Alcohol abuse during pregnancy has been shown to be deleterious to the normal structural development of the fetal brain. Animal studies have also shown that alcohol alters the number and the shape of neurons in the developing brain. The alcohol effects on the developing brain are believed to cause the behavioral abnormalities seen in alcohol-exposed offspring. Animal models have proven useful in determining how alcohol exposure during development affects behavior. Using rat as an animal model, research has shown that the neurotransmitter system that regulates neuroendocrine and autonomic responses to stress is especially vulnerable to ethanol during the developmental period. Behavioral and neurochemical studies indicate that defects in the ability of these rats to respond appropriately to stress appears to be due to alterations in the function of hypothalamic neurons, including β -endorphin neurons and CRH neurons. The biological basis for the abnormalities in β -endorphin neuronal functions in animals exposed to ethanol during the fetal period is unknown, but it presumably involves an alteration in β -endorphin neuron growth and differentiation.

How ethanol affects β -endorphin neuron growth, differentiation, and neuronal connection in the CNS is not well understood. Some recent studies suggested the possibility that ethanol may reduce the production of a neurotrophic factor necessary for β -endorphin neuron growth and differentiation. One of the neurotrophic agents that affects β -endorphin neuronal growth in cultures is cAMP. Using cells from the rat fetal hypothalamic tissues, it has been shown that cAMP increases β -endorphin cell numbers and neurite growth by preventing programmed cell death. Ethanol reduces the production of cellular levels of cAMP and increases the secretion of transforming growth factor β 1 and consequently upregulates the expression of proapoptotic genes and induces programmed cell death of these neurons. The loss of β -endorphin neurons during the developmental period can lead to permanent opioid deficiency in the hypothalamus and abnormality in stress axis function.

G-proteins are membrane-bound proteins that regulate cellular adenylyl cyclase activity and maintain the production of intracellular cAMP. Recently, abnormalities in G-protein expression and adenylyl cyclase activity have been observed in lymphocytes and erythrocytes derived from alcohol-dependent individuals. Similar abnormalities in the cAMP-signaling system have been documented in high-drinking lines of rats. In nonalcoholic children of alcoholics, an enhanced expression of the stimulatory $G_s\alpha$ protein was observed in erythrocyte and lymphocyte membranes. It has been postulated that adenylyl cyclase and G-protein expression may be markers for vulnerability to alcoholism.

Further Reading

- Ahmed, F. E. (1995). Toxicological effects of ethanol on human health. *Critical Review of Toxicology* 25, 347–367.
- Anton, R. F. (1996). Neurobehavioural basis for the pharmacotherapy of alcoholism: current and future directions. *Alcohol & Alcoholism* 31(supplement 1), 43–53.
- Boyadjieva, N. I., Chaturvedi, K., Poplawski, M. M., et al. (2004). Opioid antagonist naltrexone disrupts feedback interaction between μ - and δ -opioid receptors in splenocytes to prevent alcohol inhibition of NK cell function. *Journal of Immunology* 173, 42–49.
- Charness, M. E. (1992). Molecular mechanisms of ethanol intoxication, tolerance, and physical dependence. In: Mendelson, J. H. (ed.) *Diagnosis and treatment of alcoholism*, pp. 155–199. New York: McGraw-Hill.
- Chen, C. P., Kuhn, P., Chaturvedi, K., et al. (2006). Ethanol induces apoptotic death of developing beta-endorphin neurons via suppression of cyclic adenosine monophosphate production and activation of transforming growth factor-beta1-linked apoptotic signaling. *Molecular Pharmacology* 69, 706–717.
- De, A., Boyadjieva, N., Pastoric, M., et al. (1994). Cyclic AMP and ethanol interact to control apoptosis and differentiation in hypothalamic β -endorphin neurons. *Journal of Biological Chemistry* 269, 26697–26705.
- Froehlich, J. C. (1995). Genetic factor in alcohol self-administration. *Journal of Clinical Psychiatry* 56, 15–23.
- Gianoulakis, C. A., DeWale, J. and Thavundayil, J. (1996). Implication of the endogenous opioid system in excessive ethanol consumption. *Alcohol* 13, 19–23.
- Gordon, A. S., Collier, K. and Diamond, I. (1986). Ethanol regulation of adenosine receptor-stimulated cAMP levels in a clonal neuronal cell line: an *in vitro* model of cellular tolerance to ethanol. *Proceedings of the National Academy of Sciences USA* 83, 2105–2108.
- Koob, G. F. and Roberts, A. J. (1999). Brain reward circuits in alcoholism. *CNS Spectrums* 4, 23–37.
- O'Brein, C. P., Volpicelli, L. A. and Volpicelli, J. R. (1996). Naltrexone in the treatment of alcoholism: a clinical review. *Alcohol* 13, 35–39.
- Sayette, M. A. (1999). Does drinking reduce stress? *Alcohol Research & Health* 23, 250–255.
- Sinha, R. (2000). How does stress increase risk of drug abuse and relapse? *Psychopharmacology* 158, 343–359.
- Wand, G. S., McCaul, M., Yang, X., et al. (2002). The mu-opioid receptor gene polymorphism (A118g): alters HPA axis activation induced by opioid receptor blockade. *Neuropsychopharmacology* 26, 106–114.
- Wand, G. S., Waltman, C., McCaul, M. E., et al. Differential expression of GTP-binding proteins in individuals with high and low risks for the future development of alcoholism. *Journal of Clinical Investigations* 94, 1004–1011.
- West, J. R., Chen, W. J. and Pantazis, N. J. (1994). Fetal alcohol syndrome: the vulnerability of the developing brain and possible mechanisms of damage. *Metabolic Brain Disease* 9, 291–322.

Ethnicity, Mental Health

K Iley

Canterbury Christ Church University, Canterbury, UK

J Y Nazroo

University of Manchester, Manchester, UK

© 2007 Elsevier Inc. All rights reserved.

Introduction

Ethnicity

Ethnicity and Mental Health in the United Kingdom

Problems with Existing Sources of Data

Explaining Key Findings and Contradictions

Conclusions

Glossary

Census

A census is a count or survey of all households and people within a particular location. In the United Kingdom a census has been carried out every 10 years since 1841. The 1991 Census was the first one to ask individuals to define their ethnicity.

<i>Ethnicity</i>	The identification, by self or others, with a social group and or social collectivity on the basis of shared values, beliefs, customs, traditions, language, or lifestyle.
<i>Patriarchy</i>	Traditionally means rule of the father and used to describe the dominance of men over women.
<i>Race</i>	A contentious concept within British sociology, which is normally associated with a group connected by a common biological/genetic origin, typically associated with skin color.

Introduction

In this article we consider the key findings that have emerged on ethnic differences in mental health, with particular reference to the United Kingdom. This will be followed by a brief discussion of ethnic minority people's experiences of mental health services and the implications this has for healthcare provision and practice. We will conclude with a discussion of the conclusions we can draw on ethnic inequalities in mental health and the implications this has for our wider understanding of the experiences of ethnic minority people in United Kingdom and how these relate to stress and mental health. But we begin by introducing the concept of ethnicity.

Ethnicity

Approaches to ethnicity have been concerned with understanding how ethnicity relates both to social structures and how it relates to social relationships and identities, allowing us to provide a sensitive and contextual understanding of ethnicity, rather than resort to explanations based on stereotypes. So, research in the United Kingdom has demonstrated the social and economic inequalities faced by ethnic minority people and how economic inequalities and racism relate to ethnic inequalities in physical and mental health. In addition, importance is placed on the notion of ethnicity as an identity that reflects self-identification with cultural traditions, and which provide both meaning and boundaries between groups. So, although there is strong evidence to show that socioeconomic disadvantage contributes to ethnic inequalities in health, it is suggested that there remains a cultural component to ethnicity that could play a defining role.

When considering the relationship between culture, ethnicity, and mental health, however, it is vital to avoid reducing ethnic differences in mental health to stereotyped notions of fixed cultural or biological difference. Ethnicity is considered to reflect identification

with sets of shared values, beliefs, customs, and life-styles and has to be understood dynamically, as an active social process. In particular, the influence of an ethnic identity on individuals and their health depends on the wider context in which that identity is lived.

Ethnicity and Mental Health in the United Kingdom

The difference in rates of mental illness among different ethnic groups in the United Kingdom is probably one of the most controversial issues in the health inequalities field. Given the topic, which potentially allows the alignment of mental disorder with ethnic minority status, the controversial nature of the field is not surprising. And this controversy is aggravated by the complexity of conducting research on ethnic differences in mental illness and the consequent disputed nature of research findings. Much of the controversy has focused on the high rates of hospital treatment for schizophrenia and other forms of psychosis among the African-Caribbean population, where the alignment of mental disorder with ethnicity is most apparent. Interestingly, these findings are reflected in research in the United States and the Netherlands, where black populations are also shown to have high rates of psychotic illness. In the United Kingdom, evidence suggesting low rates of mental illness among the South Asian population, but high rates of suicide and attempted suicide among young South Asian women has also caused controversy. The following provides a summary of key findings from the United Kingdom.

Psychotic Illness

Psychotic illnesses, which include schizophrenia, are relatively infrequent; they are thought to affect around one person in 250 in the United Kingdom, but often result in severe disability. Typically they involve a fundamental disruption of thought processes, where the individual suffers from a combination of distressing delusions and hallucinations.

Most research on ethnic differences in psychotic illnesses has been based on treatment rates, and over the past three decades such studies in the United Kingdom have consistently shown elevated rates of schizophrenia among African-Caribbean people compared with the white population. African-Caribbean people are typically reported to be three to five times more likely than whites to be admitted to hospital with a first diagnosis of schizophrenia. These findings have been repeated in studies that have looked at first contact with all forms of treatment, rather than just hospital services, although the rates in one such study were only twice those of the white population. Some of the more recent of these studies have also looked at

those of African ethnicity and have reported similarly raised rates of psychotic illness in this group. Explorations of the demographic characteristics of black people admitted to hospital with a psychotic illness suggest that these illnesses are particularly common among young men, and some studies have suggested that the rates are very high among young African-Caribbean people who were born in the United Kingdom, reporting that rates of first contact with psychiatric services for psychotic illness among this group are 18 times the general population rate.

Given the consistency of the evidence based on treatment statistics, it is somewhat surprising that the only national community based studies of mental illness among ethnic minority groups, the Fourth National Survey of Ethnic Minorities (FNS) and the Ethnic Minority Psychiatric Illness Rates in the Community (EMPIRIC) study, produced rather different findings. Overall, they found that Caribbean people had rates of psychotic illness that were at most twice as high as those in the general population. And when differences were considered across gender, age, and migrant/nonmigrant groups it was found that the prevalence of psychotic illness among: men; young men; and nonmigrant men; was no greater than that for equivalent white people. For example, in the FNS the annual prevalence of psychotic disorder among Caribbean men was estimated as 10 per 1000, while among white British men it was estimated as eight per 1000.

Findings on rates of psychotic illness among South Asian people are even more mixed. Studies of hospital-based treatment suggest that rates of admission for psychotic illness among South Asian people are similar to those among white people. This has been confirmed by a more comprehensive prospective study of first contact for schizophrenia with *all* treatment services in one area of London, but an earlier study using the same methods in another London district suggested that rates of psychotic illness among South Asian people were raised to similar levels to those found among black Caribbean people. Indeed, there is evidence suggesting that rates of psychotic illness among each of the ethnic minority groups defined by the 1991 Census categories are similarly raised in comparison with a white group. This, of course, suggests that it is misleading to maintain an exclusive focus on those of African-Caribbean origin when examining ethnic differences in psychotic illness in the United Kingdom.

In contrast to the findings for contact with treatment services, the community based FNS and EMPIRIC prevalence studies suggested that rates of psychotic illness might not be raised for South Asian people, and may be lower for Bangladeshi people,

than those for white British people. In support of the conclusions drawn by some researchers, these community surveys also suggested a high rate of psychosis among white people who were not of British origin.

Depression

Neurotic disorders, which include depression, are much more common than psychotic disorders. A national survey in the United Kingdom suggested that in the week before interview about one person in 16 was affected by such a disorder. They are usefully separated into two categories, anxiety and depression that, although common, do involve considerably more than a sense of anxiety or sadness. Here we will focus on depression, because of the lack of data on ethnic differences in anxiety.

According to treatment statistics, rates of depression among African-Caribbean people appear to be markedly lower than those for white people, and rates of depression among South Asian people appear to be a little lower than those for white people. For South Asian people, these findings have been confirmed in community studies, including the FNS. However, the more recently conducted, and more thorough, EMPIRIC study suggested few differences between South Asian groups and white people, with higher rates of depression among the Pakistani population.

In comparison with the high rates of treatment for psychotic illness among African-Caribbean people, the low rates of treatment for depression are a puzzle. Most factors that might be implicated in the higher rates of psychotic disorders should also lead to a higher rate for other mental illnesses. In addition, in contrast to the low treatment rates, evidence from the FNS suggests that the prevalence of depression among Caribbean people in the community is, in fact, more than 50% higher than that among whites. Moreover, that study also suggested that despite this higher prevalence, rates of treatment for depression among Caribbean people were very low.

In contrast to the low rates of hospital admission for depression among those born in the Caribbean and South Asia, it has been reported that rates of admission for depression among those living in England but born in Northern Ireland and Ireland were markedly higher than for those born and living in England. The high rate of depression among white people who were not of British origin was confirmed in the community based FNS, which reported that this group had rates that were two-thirds higher than the white British group, and was confirmed for Irish men in EMPIRIC.

Suicide

In contradiction to the apparent lower overall relative rates of mental illness among South Asian people, analyses of immigrant mortality statistics show that mortality from suicide are higher for young women born in South Asia, and this is particularly the case for very young women (aged 15–24 years), where the rate is two to three times the national average. In contrast to the findings for young South Asian women, these studies also showed that men and older women (aged 35 years or more) born in South Asia had lower rates of suicide. Analysis of the most recent data on immigrant mortality has been more detailed, because it could be coupled with the 1991 Census, which included a question on ethnicity as well as country of birth. This confirmed the overall pattern just described, but showed that the high rates appear to be restricted to those born in India and East Africa, rather than Pakistan or Bangladesh.

In terms of morbidity, rather than mortality, again the picture is mixed with studies of treatment suggesting very high rates of suicidal ideation, attempted suicide, and suicide among both male and female South Asians. In contrast, the only studies to look at suicidal thought among national community populations (the FNS and EMPIRIC) found that the rates of suicidal ideation were not raised among South Asian groups and, if anything, lower among South Asian women than white women regardless of age.

Problems with Existing Sources of Data

Although mental illness is a relatively common condition, it is difficult to measure and the defining characteristics are contested. Both the definition and the measurement of mental illness depend on the presence of clusters of psychological symptoms that indicate a degree of personal distress, or that lead to behaviors that cause such distress to others. The clusters of symptoms associated with particular forms of mental illness are clearly defined by psychiatrists, although the elicitation of these symptoms for diagnostic or research purposes can be difficult, particularly in cross-cultural studies.

One of the central problems arises from the reliance of most work on data based on contact with treatment services. Contact with treatment services, even when access is universal as in the United Kingdom's National Health Service, reflects illness behavior (i.e., the way that symptoms are perceived, evaluated, and acted upon), rather than illness *per se*. This raises a number of linked problems, particularly as illness behavior is likely to be affected by a number of fac-

tors that vary by ethnicity, such as socioeconomic position, health beliefs, expectations of the sick role, and lay referral systems. And these problems become particularly important for work on rates of psychosis, where contact with services might be against the patient's wishes. So, despite the consistency of research findings showing that African-Caribbean people have higher rates of treatment for psychosis, some commentators have not accepted the validity of the interpretation of these data and continue to suggest that a higher incidence (rather than a higher treatment rate) remains unproven, because of the serious methodological flaws with the research that has been carried out.

Briefly, the kind of problems that are focused on include: underestimation of census-based denominators for African-Caribbean people, leading to overestimation of admission rates (although not to the extent that could explain the many times higher rates of admission); overestimation of first onsets of psychosis, because of under-identification of previous episodes for African-Caribbean people (because of high geographical mobility leading to records of previous admissions being missed and a reluctance to disclose previous diagnoses because of a concern about the impact of these on how they might be managed); variations in the routes of admission and treatment by ethnicity, with African-Caribbean people over-represented among patients compulsorily detained, more likely to have been in contact with the police or forensic services prior to admission (despite them being both less likely than whites to display evidence of self-harm and no more likely to be aggressive to others prior to admission), and more likely to have been referred to these services by a stranger rather than by a relative or neighbor. Taken together, these comments suggested that there are a variety of potential problems with existing work and, consequently, that there must remain some doubt about the higher rates of psychosis reported among the African-Caribbean group.

It is equally possible that the reported differences between white and South Asian groups and the inconsistencies found in these could, as for the African-Caribbean group, be a result of the methodological limitations of studies in this area. In addition to the difficulties of relying on treatment data, outlined above, the lower rates of mental illness among South Asian people could reflect language and communication difficulties, or a general reluctance among South Asian people to consult with doctors over mental health problems. More fundamentally, they may reflect a difference in the symptomatic experience of South Asian people with a mental illness compared

with white people. In particular, it has been suggested that South Asian people may experience particular culture-bound syndromes, that is a cluster of symptoms which is restricted to a particular culture, such as sinking heart, and consequently not be identified as mentally ill.

Explaining Key Findings and Contradictions

There a number of explanations considered for ethnic differences in mental health and these are similar to those considered for other ethnic inequalities in health in the epidemiological literature. These include migration, genetic differences, culture, racism, and socioeconomic position.

Migration

First, different rates of mental health across migrant and nonmigrant groups could be a consequence of factors related to the actual process of migration. Social selection into a migrant group could have favored those with a higher or lower risk of developing illness, or the stresses associated with migration might have increased risks. There is evidence to both support and counter these suggestions. In the case of the apparently higher rates of schizophrenia among African-Caribbean people in the United Kingdom, investigations of the rates of schizophrenia in Jamaica and Trinidad suggest that they are much lower than those for African-Caribbean people in the United Kingdom and, in fact, similar to those of the white population of the United Kingdom. This would suggest that the higher rates are either a consequence of factors related to the migration process, or of the greater stresses surrounding the lives of ethnic minority people in the United Kingdom.

However, there is evidence that shows that rates of schizophrenia for African-Caribbean people born in the United Kingdom are even higher than for those who migrated, suggesting that factors relating directly to the process of migration may not be involved, although these data (like most work in this area) are dependent on a very small number of identified cases.

Genetic Differences

As with all work on ethnicity and health, there has been discussion of the possibility that differences may be a consequence of a genetic factor that correlates with ethnic background, but little supporting evidence has been marshaled. Evidence suggests large differences in risk within ethnic categories, implying that any genetic basis for mental illness does not

correlate closely with ethnic background. For example, the evidence on schizophrenia cited in the previous section, which shows that there are important differences between African-Caribbean people who stayed in Jamaica or Trinidad (who do not have raised rates), those who migrated to the United Kingdom (who appear to have raised rates), and those who were born in the United Kingdom (who appear to have markedly raised rates), suggests that the higher rates cannot be straightforward consequence of ethnic differences in genetic risk.

Culture

Not surprisingly, most research that has focused on cultural explanations for ethnic differences in mental health has based the cultural argument on speculative and stereotyped characterizations of the cultures of ethnic minority groups, which do not acknowledge the dynamic nature of culture.

It has been suggested that overall lower rates of mental illness among South Asian people could be a consequence of a strong and protective Asian culture, which provides extended and strong communities with protective social support networks. In contrast, in the attempt to explain the high mortality rates of suicide among young women born in South Asia and living in the United Kingdom, South Asian communities are portrayed as constraining, demanding, and conflictual, rather than supportive and cohesive, and so contributing to the high suicide rates.

Of course, such stereotypes will not hold when closely examined. For example, despite the focus on patriarchal South Asian families, there are, in fact, great similarities between the motives of white and South Asian patients for their suicidal actions. So, although it is worthy of considering how culture informs our understanding of ethnic minority people and their experiences of illness, it is necessary to question the stereotypes we use and to understand ethnic identities as dynamic and dependent on context.

Racism and Socioeconomic Position

Different rates in mental health across different ethnic groups might be a consequence of stress resulting from the different forms of discrimination and racism that ethnic minority people face in the United Kingdom. This could be a direct result of the experience of discrimination and harassment, or a result of the social disadvantages that racism leads to. It would not be surprising if the poor run-down, inner city environments and poor housing that many ethnic minority people live in, and their poorer employment prospects and standards of living, led to greater mental distress. As elsewhere in the ethnicity and health

field, there has been considerable criticism of the failure to take into account explanatory variables related to social disadvantage in research that links ethnicity to poor mental health, as there is a strong possibility that these underlie the relationship. And there is a growing body of evidence linking racism and discrimination directly, and indirectly, to adverse mental health outcomes.

Conclusions

This article suggests that many basic questions concerning the relationship between ethnicity and mental health remain controversial. There remains a question of whether the use of western psychiatric instruments for the cross-cultural measurement of psychiatric disorder is valid and produces a genuine reflection of the differences between different ethnic groups. This has been raised particularly in relation to the low detection and treatment rates for depressive disorders among South Asians, but may apply to other disorders and other ethnic minority groups. It is also likely that treatment-based statistics do not accurately reflect the experiences of the populations from which those in treatment are drawn.

Perhaps, as other writers have pointed out, the most important conclusion to draw is that it is vital to avoid reducing ethnic differences in mental health to stereotyped notions of fixed cultural or biological difference, there is a need to explore the factors associated with ethnicity that may explain any relationship between ethnicity and mental health, such as the various forms of social disadvantage and consequent stressors that ethnic minority people face. And it remains important to explore how racism and the social disadvantages that this leads to structure the experiences of ethnic minority people when they come into contact with mental health services. Despite this, the focus on biological rather than social explanations for ethnic differences in mental illness continues in both the United States and the United Kingdom.

Further Reading

Adebimpe, V. R. (1995). Race, racism, and epidemiological surveys. *Hospital and Community Psychiatry* 45, 27–31.

American Psychiatric Association (1995). *Diagnostic and statistical manual IV*. Washington DC: American Psychiatric Association.

Crawford, M. J., Nur, N., McKenzie, K., et al. (2005). Suicidal ideation and suicide attempts among ethnic minority groups in England: results of a national household survey. *Psychological Medicine* 9, 1369–1377.

Harrison, G., Owens, D., Holton, A., et al. (1988). A prospective study of severe mental disorder in Afro-Caribbean patients. *Psychological Medicine* 18, 643–657.

Iley, K. and Nazroo, J. Y. (2001). Ethnic inequalities in mental health: a critical examination of the evidence. In: Culley, L. & Dyson, S. (eds.) *Sociology, ethnicity and nursing practice*, pp. 67–89. Basingstoke: Palgrave.

Iley, K. and Nazroo, J. (In press). Sociology of mental health and illness. In: Bhui, K. & Bhugra, D. (eds.) *Textbook of culture and mental health disorder*. London: Hodder Arnold.

Karlsen, S., Nazroo, J. Y., McKenzie, K., et al. (2005). Racism, psychosis and common mental disorder among ethnic minority groups in England. *Psychological Medicine* 35, 1795–1803.

King, M., Coker, E., Leavey, G., et al. (1994). Incidence of psychotic illness in London: comparison of ethnic groups. *British Medical Journal* 309, 1115–1119.

King, M., Nazroo, J., Weich, S., et al. (2005). Psychotic symptoms in the general population of England: a comparison of ethnic groups (The EMPIRIC study). *Social Psychiatry and Psychiatric Epidemiology* 40, 375–381.

Kleinman, A. (1987). Anthropology and psychiatry: the role of culture in cross-cultural research on illness. *British Journal of Psychiatry* 151, 447–454.

McGovern, D. and Cope, R. (1987). First psychiatric admission rates of first and second generation Afro-Caribbeans. *Social Psychiatry* 22, 139–149.

Nazroo, J. Y. (2001). *Ethnicity, class and health*. London: Policy Studies Institute.

Sashidharan, S. and Francis, E. (1993). Epidemiology, ethnicity and schizophrenia. In: Ahmad, W. I. U. (ed.) *Race and health in contemporary Britain*. Buckingham, UK: Open University Press.

Selten, J. P., Slaets, J. P. and Kahn, R. S. (1997). Schizophrenia in Surinamese and Dutch Antillean immigrants to the Netherlands: evidence of increased incidence. *Psychological Medicine* 27, 807–811.

Weich, S., Nazroo, J., Sproston, K., et al. (2004). Common mental disorders and ethnicity in England: the EMPIRIC study. *Psychological Medicine* 34, 1543–1551.

Evolutionary Origins and Functions of the Stress Response

R M Nesse, S Bhatnagar and E A Young

University of Michigan, Ann Arbor, MI, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R M Nesse, S Bhatnagar and E A Young, volume 2, pp 79–83, © 2000, Elsevier Inc.

Utility of the Stress Response

Phylogeny of the Stress Response

How Does the Stress Response Help?

Adaptive Regulation of Stress Responsiveness

Mismatch between Ancestral and
Modern Environments

Glossary

<i>Defense</i>	A trait that is latent until aroused by threatening situations in which it is useful.
<i>Natural selection</i>	The process by which genes that provide a fitness advantage become more common from generation to generation and those that decrease fitness become less common, thus shaping adaptive traits, including defenses.
<i>Phylogeny</i>	The evolutionary history of a trait or a species.
<i>Trade-offs</i>	The fitness costs and benefits of a trait whose net effects yield a selective advantage.

Evolution is the process in which traits such as the capacity for the stress response are shaped by natural selection. Understanding the evolutionary history of a trait, how it gives a selective advantage, and the costs it imposes can help to illuminate its design and regulation and can guide research into its mechanisms and control. The stress response has been shaped by natural selection to increase the ability of organisms to cope with situations that require action or defense. Stress-related mechanisms emerged early in the history of life. Like all traits, they have costs as well as benefits. Because the stress response is so often associated with negative events, its utility has often been neglected. In particular, the release of glucocorticoids, which is often thought to be the hallmark of the stress

response, may in fact exist, in part, to protect against other aspects of defensive systems.

Utility of the Stress Response

The vast bulk of research on stress has investigated its causes, mechanisms, and effects. An evolutionary approach instead addresses two very different and relatively neglected questions: (1) How does the stress system give a selective advantage and (2) what is the evolutionary history of the stress system? The answers to these questions provide a foundation in Darwinian medicine for understanding why the stress response is the way it is and why it causes so much suffering and disease. The first and most important contribution of an evolutionary perspective on stress is a clear focus on its utility. The stress system is a complex, sophisticated, and carefully regulated adaptation that has been shaped by natural selection because it gives a selective advantage. That advantage must be substantial in order to outweigh its huge costs. The idea that stress is useful is by no means new. In fact, the very phrase Hans Selye chose to describe it, the general adaptation syndrome, emphasizes its utility. Despite this early emphasis on its benefits, as the idea of stress entered the popular imagination there was a tendency to emphasize its dangers so that the fundamental fact of the utility of the stress response was often forgotten.

Stress and Other Defenses

Other defenses are also often confused with the problems they protect against. The capacities for pain, fever, vomiting, cough, and inflammation are often thought of as medical problems, although a moment's thought reveals that they are useful protective reactions. The ubiquity of the illusion that defenses are abnormalities arises from several sources. First, defenses are often associated with some kind of suffering and therefore seem maladaptive. Unfortunately, however, discomfort is itself probably one aspect of a mechanism that makes it useful. Second, they are reliably associated with disadvantageous situations, so the association bias makes it seem as if they are the problem. Finally, it is possible to use drugs to block the expression of many defenses with very little harm, completing the illusion that defenses are useless. In fact, blocking

a defense can be harmful. For instance, suppressing cough in a patient with pneumonia makes it harder to clear the infection and may lead to death, and stopping the diarrhea of a person with a serious intestinal infection may lead to complications. Blocking fever, however, usually has little effect on the speed of recovery from a cold. When blocking a defense is not dangerous, this is because the body has backup protective mechanisms and because the regulation mechanism seems to be set to a hair trigger that expresses the defense at the slightest hint of danger.

Situations in Which Stress Is Useful

Stress, like fever and pain, is useful only in certain situations. Such traits lie latent until aroused by the particular circumstances in which they are useful. This means that the evolutionary explanation for such traits cannot be summarized in a single function. Instead, the inducible defenses give an advantage by changing multiple aspects of the body that increase its ability to cope effectively with the adaptive challenges that arise in a particular situation. One defense may have many aspects that serve many functions. So, the first step in understanding the adaptive value of stress is not to try to specify its function but to understand the exact situations in which the stress response is useful. To do that, we need to go back to the very origins of complex life forms 600 million years ago. If a very primitive organism had only two states, what are they? The answer is quite straightforward: activity and rest. This is a fundamental divide, one that is maintained even in our biochemical and nervous systems. Biochemical pathways are divided into the catabolic, in which energy is used, and the anabolic, in which energy is stored and tissues are repaired. Parallel to this division are the two arms of the autonomic nervous system. The sympathetic system, which is activated as part of the stress response, increases arousal, blood pressure, heart rate, respiratory rate, and physical activity and institutes other endocrine and physiological changes necessary for action. The other half of the autonomic nervous system, the parasympathetic, inhibits muscular activity, stores energy, and shunts blood to digestion and bodily repair. Is stress, then, the same as arousal for action? Not exactly. As soon as a generic state of arousal was well established, natural selection probably began to differentiate it into subtypes to better meet the demands of different kinds of challenges. Here again, the main bifurcation is quite clear. Arousal is useful in two different situations: threats and opportunities. This division is also represented in our nervous systems. As Gray and others have pointed out, the brain seems to have moderately distinct systems for behavioral inhibition and for reward

seeking. The corresponding behaviors are said to be defensive or appetitive and are associated with feelings of fear/pain or pleasure. In psychology, the same division is recognized in the distinct cognitive states promotion versus prevention.

Phylogeny of the Stress Response

Cross-Species Comparisons

Comparisons among different species can help to reconstruct the phylogeny of the stress response. All vertebrates have the proopiomelanocortin (POMC) molecule that gives rise not only to adrenocorticotrophic hormone (ACTH) but also to opiate-like peptides. It is intriguing to note that these molecules, with their related functions, are derived from the same parent molecule. All vertebrates also make corticosteroids. Peptide sequences very similar to those of human ACTH are found not only in mammals but also in amphibians and reptiles and even in insects, mollusks, and marine worms. Interestingly they are usually associated with immune cells, equivalent to macrophages, where they set defensive processes in motion. ACTH has long been closely associated with other signaling molecules such as corticotropin releasing hormone (CRH), biogenic amines such as epinephrine and norepinephrine, steroids such as cortisol, cytokines such as interleukin-1, and nitric oxide. All these substances are crucial to defensive systems. The remarkable thing is that genetic sequences for these molecules have not only been conserved over hundreds of millions of years but they continue to serve closely related defensive functions. Why have they changed so little? If a single molecule has several essential functions, this creates a strong selective force against mutations that change the sequence. By contrast, mutations that result in the differentiation of different classes of receptors in target tissues can slowly specialize the responses of that tissue to the signal molecule. And they have, judging from the proliferating classes and subclasses of receptors that are now being discovered.

Cost-Benefit Trade-Offs

Why is the stress system not better? It could provide more effective protection against danger – but only at a still greater cost. *Drosophila* have been bred to resist the stress of food shortage. After 60 generations of such selection, the new strain is 80% more resistant to starvation. However, the larvae are more likely to die, and development is slowed. Similar results have been found for selection of resistance to other stressors and in other species. Like everything else in the body, stress responses are shaped by trade-offs, some-

times with benefits and costs occurring in different parts of the life cycle.

The mechanisms that regulate the responsiveness of the stress system are shaped by the trade-off between the long-term costs and the immediate benefits of a relatively quick or intense or prolonged stress response. It has been hypothesized that the individuals who are most resilient or resistant to the effects of stress on physiology or behavior are the ones least vulnerable to stress-related diseases. However, are these individuals resilient to all stress-related disorders, or are there situations in which resilience to some disorders means vulnerability to others? The answer is not known, but given the large number of physiological systems affected by stress and/or glucocorticoids, it is unlikely that resilience to one stress-related disorder necessarily protects against all.

Another trade-off reflects the benefits and costs of habituation. For instance, rats exposed to some repeated stressors, particularly those that are mild and cognitive in nature, habituate to that stressor. The hypothalamic-pituitary-adrenal (HPA) response is lower to the n th exposure than to the first exposure. Such habituation seems adaptive for most situations; if the stressor is known and can be easily coped with, then the HPA response should be moderated. Such habituation would, at the very least, conserve resources. However, it would be maladaptive to habituate to stressors that do present some danger. These trade-offs have shaped the brain mechanisms that regulate habituation.

There are interesting sex differences in the habituation of the stress response, with some evidence indicating that habituation occurs in male but not in female rats. Such findings suggest that the selection forces acting on male and female rats may have differed enough to shape distinctly different patterns of habituation. Research is now addressing whether habituation occurs with some stressors in males but with others in females.

Stress responses in adult animals are profoundly affected by early environmental events such as prenatal stress and variations in maternal care. The effects of variations in maternal care are transmitted across generations, with offspring that experience high maternal care exhibiting lower stress responses and providing high maternal care themselves. Such effects seem adaptive in that offspring are likely to experience an environment similar to that of their parents. So, for example, mothers providing low maternal care have high stress responses and so do their offspring when they become adults. However, when cross-fostered to other mothers, the offspring show patterns of stress responsivity similar to that of their foster mother and not their biological mother. Such

results suggest that stress responsivity and maternal care are not simply genetically transmitted but are also regulated by early experiences. Such regulation is seen in other mammals and even plants. Some of such transmission across generations may arise from facultative mechanisms that evolved to adjust the system based on early life experiences, and some may arise from more general learning mechanisms.

Difficulties in Defining Stress

The human mind seems wired to try to make neat categories with sharp boundaries, perhaps because we communicate with words and this requires dividing the world up into categories even when that is unnatural. This leads to a tendency to try to make sharp distinctions between different states that may, in fact, overlap considerably. States of defensive arousal, for instance, are different from states of arousal for seeking food, but there is no reason to expect that the differentiation is complete. For instance, cortisol secretion is aroused by opportunities as well as threats. In fact, cortisol is even involved in reward mechanisms. Thus, any attempt to define the stress response in terms of cortisol arousal is doomed. For that matter, any attempt to define stress or the stress response is liable to be an exercise in frustration for the evolutionary reason that the system does not have sharp boundaries or a single function. The closest we can come to a defining characteristic is the kinds of situations in which stress has given a selective advantage, and those situations are not sharply defined. The stress system was, after all, not designed by an engineer but shaped by a process of tiny tinkering changes. The long unsatisfying history of attempts to define stress and the wish, expressed by many researchers, that the term would go away, arise from this difficulty. Even after defensive arousal was differentiated considerably from appetitive arousal, there were undoubtedly advantages to further differentiating subtypes of stress responses to match specific challenges. Thus, different situations – a predator, a high place, injury, infection, starvation, loss of a status battle, and speaking in public – all seem to have shaped somewhat different defensive responses. These responses are only partially differentiated from a more generic response, so they have overlapping characteristics with functions in common. Attempts to sharply distinguish different kinds of anxiety disorders are as frustrating as attempts to define stress itself and for the same reasons. New attempts to study anxiety disorders in the context of normal anxiety will be helpful.

How Does the Stress Response Help?

Immediate Response

These difficulties notwithstanding, a stress response is a coordinated pattern of changes that is useful in situations in which the organism is faced with possible damage or a loss of resources. The next question is: How is it useful? Even before Selye, Walter Cannon provided some answers. In situations that might require fight or flight, he observed the utility of increased heart rate and contractility to speed circulation, increased rate and depth of breathing to speed gas exchange, sweating to cool the body and make it slippery, increased glucose synthesis to provide energy, shunting of blood from gut and skin to muscles, increased muscle tension to increase strength and endurance, and increased blood clotting in preparation for possible tissue damage. More recently, others have demonstrated faster reaction times and cognitive benefits as a result of sympathetic arousal. These immediate responses are mostly mediated by the sympathetic nervous system and the associated release of epinephrine from the adrenal medulla.

Adrenal Cortical Response

However, the stress system also includes a more delayed-response release of cortisol from the adrenal cortex, although this system does have more rapid effects (such as fast negative feedback) that are probably mediated by putative membrane steroid receptors. This is initiated by neural signals to the hypothalamus, which releases CRH, which in turn results in the secretion of ACTH from the anterior pituitary gland on the bottom of the brain. The ACTH induces cortisol synthesis and release from the adrenal gland. The whole system is called the HPA system because the signal acts via the hypothalamus, the pituitary, and the adrenal glands. Many actions of the HPA system seem, like those of the sympathetic system, well designed for acute action. It changes physiology so the liver breaks down glycogen into glucose, and it alters cells so glucose can get in more readily. Through different populations, CRH not only releases ACTH, but it also directly increases anxiety and arousal and activates cells in the locus coeruleus, the brain center where the cell bodies for most noradrenergic neurons are located. All in all, the system seems admirably designed to get the organism ready for action. Indeed, both branches of the system are readily aroused by exercise, and trained athletes, far from having low levels of cortisol, have chronic high levels – just the thing for a person who frequently exerts him- or herself.

Association with Negative Events

So, why should the stress system as a whole be associated so closely with bad events instead of positive ones? To answer that question, we need to understand why the components of the stress system are carefully packaged. If the stress is so useful, why is stress not expressed all the time? There are at least three good reasons why it is not. First, it is expensive in terms of calories. No organism can afford to waste effort. Second, it interferes with other adaptive behaviors. A vigilant organism has less time for finding food and eating, to say nothing of mating. Finally, and most important, some changes that give an advantage in the face of threats may also cause tissue damage. For this reason, they need to be carefully sequestered, except in those few circumstances in which the costs are outweighed by the benefits. This helps to explain why some aspects of the stress response are associated more with negative than positive arousal. The benefits of a stress response that increases the likelihood of catching prey may sometimes be worth it, but if a stress response prevents being caught as prey, this is always worth it, even if substantial damage results from the stress response itself. This helps to explain why the stress response, despite its costs, is so ubiquitous. An optimal regulatory mechanism will express a stress response whenever, on average, it is worth it. Given the uncertainties of environmental cues and the potential life-saving effects of the stress response, there will be many instances where the expression of stress is worthwhile, even though there is only a small chance that danger is actually present. The global conclusion is that the damage caused by stress responses is not necessarily from abnormal stress. Some components of the stress response may be a part of the response specifically because they are too damaging to be expressed except when they protect against great danger. There is every reason to think that normal stress, like every other bodily trait, has costs as well as benefits. This idea is expressed in the concept of allostasis, as proposed by McEwen and colleagues, which emphasizes the short-term benefits and the long-term costs.

That a normal stress response might be crucial for optimal physiological functioning has implications for recent notions about pharmacological therapies for reducing perceived stress. For example, if a drug such as a CRH inhibitor blocks all stress response at the brain level, then how does the body react to what may be real requirements for increases in energy use and how do systems that are generally opposed in functions to those of glucocorticoids, counterregulatory systems such as the regulation of insulin release, stay in check? Furthermore, there are many indica-

tions in both human and animal studies of mismatches between perceived stress or behavioral indices of stress and HPA activation. The extreme of an absent stress response is, of course, Addison's disease. Thus, although the cognitive nature of many current human stressors results in costs disproportionate to actual threats, from an evolutionary point of view the general inhibition of stress responses is by no means optimal. It would be ironic as well as tragic if the history of the excessive use of cortisone were to repeat itself with a new generation of drugs that block the stress response.

Cortisol as Protection against Other Aspects of Stress

If some aspects of the stress response cause harm, has selection shaped systems to protect against this damage? In 1984, Munck and colleagues reviewed the actions of cortisol and said, "We propose that stress-induced increases in glucocorticoids levels protect, not against the source of stress itself, but rather against the body's normal reactions to stress, preventing those reactions from overshooting and themselves threatening homeostasis." They noted that many inflammatory diseases had been attributed to overproduction of cortisol until 1941, when adrenal steroids were shown to decrease inflammation. Subsequent demonstrations showed that steroids inhibit the production of cytokines, prostaglandins, and other mediators of the immune response, thus decreasing immune function. This is just the opposite of what would make sense as protection from danger, but it is entirely consistent with a role in protecting against damage from immune system activation induced by other changes. It is now clear that the effects of glucocorticoids on immune function are much more complicated than originally thought. Added to this are recent novel findings regarding the effects of glucocorticoids on the brain versus the body on the regulation of energy balance and fat deposition that suggest that our assumptions about the physiological and neural functions of glucocorticoids will continue to be significantly challenged.

Adaptive Regulation of Stress Responsiveness

The recognition that a moderately responsive stress system is optimal is quickly leading to a consideration of the adaptive significance of variations in stress responsiveness. Several lines of thinking addresses the adaptive significance of innate individual differences in stress responsiveness. The first attempts to explain the growing evidence for a bimodal pattern

of stress response in some organisms, with some individuals responding quickly and strongly and others showing a much more restrained response. It has been suggested that these distinct patterns reflect distinct patterns of adaptation; a hawk pattern that is optimal in crowded situations and a dove pattern that is optimal when population density is low. The different strategies are hypothesized to be the products of frequency-dependent selection. There is also increasing evidence that maternal effects on an offspring's stress responsiveness may be adaptive. Mothers exposed to stressful environments give birth to offspring with especially responsive stress systems that may give them an advantage in harsh environments. The same evolutionary forces may have shaped mechanisms that explain the connection between early abuse or neglect and increased stress vulnerability.

Mismatch between Ancestral and Modern Environments

Much has been made of the differences between our environment and that of our ancestors. In the case of stress, this argument comes in several flavors. Some suggest that life is more stressful now than it was for our predecessors. Special aspects of our environment do cause new kinds of stress. Working in a bureaucracy is tedious and political at best. Driving to work, living in a ghetto, running a corporation, working in a factory – these all arouse the stress system. Despite the amount of stress we experience, however, our ancestors almost certainly experienced more. With no police, no food reserves, no medicine, no laws, rampant infections, and prevalent predators, danger could come at any time. True, social groups were closer, kin networks were stronger, and people spent all their time with one another and none alone reading books. Still, life was hard. Perhaps in that environment, where stressors were more often physical, the stress response was more useful than it is now. Today, we mainly face social and mental threats, so the actions of the HPA system may yield net costs. This is plausible and supports the many efforts to reduce stress and to find drugs that block the stress response. This brings us back to the very concept of stress as a mismatch between the demands made on an individual and that individual's ability to meet those demands. We think of these demands as coming from the outside, and sometimes they do, as when we are attacked by the proverbial tiger. But most stresses in modern life arise not from physical dangers or deficiencies but from our tendency to commit ourselves to personal goals that are too many and too high and to ruminate about them. When our efforts

to accomplish these goals are thwarted or when we cannot pursue all the goals at once and must give something up, the stress reaction is expressed. In short, much stress arises, ultimately, not from a mismatch between our abilities and the environment's demands, but from a mismatch between what we desire and what we can have.

Further Reading

- Bhatnagar, S., Lee, T. and Vining, C. (2005). Prenatal stress differentially affects habituation to repeated restraint in adult male and female rats. *Hormones and Behavior* 47(4), 430–438.
- Cannon, W. B. (1929). *Bodily changes in pain, hunger, fear, and rage: researches into the function of emotional excitement*. New York: Harper and Row.
- Charney, D. S. (2004). Psychobiological mechanisms of resilience and vulnerability: implications for successful adaptation to extreme stress. *American Journal of Psychiatry* 161(2), 195–216.
- Dallman, M. F. (2003). Fast glucocorticoid feedback favors “the munchies.” *Trends in Endocrinology and Metabolism* 14(9), 394–396.
- Frankenhauser, M. (1976). The role of peripheral catecholamines in adaptation to understimulation and overstimulation. In: Serban, G. (ed.) *Psychopathology and human adaptation*, pp. 173–191. New York: Plenum.
- Gray, J. A. (1987). *Fear and stress* (2nd edn.). Cambridge, UK: Cambridge University Press.
- Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist* 52(12), 1280–1300.
- Marks, I. M. and Nesse, R. M. (1994). Fear and fitness: an evolutionary analysis of anxiety disorders. *Ethology and Sociobiology* 15(5–6), 247–261.
- McEwen, B. S. (2003). Interacting mediators of allostasis and allostatic load: towards an understanding of resilience in aging. *Metabolism* 10(supplement 2), 10–16.
- Munck, A., Guyre, P. M. and Holbrook, N. J. (1984). Physiological functions of glucocorticoids in stress and their relation to pharmacological actions. *Endocrine Review* 5(1), 25–44.
- Korte, S. M., Koolhaas, J. M., Wingfield, J. C., et al. (2005). The Darwinian concept of stress: benefits of allostasis and costs of allostatic load and the trade-offs in health and disease. *Neuroscience and Biobehavioral Review* 29, 3–38.
- Nesse, R. M. (2005). Natural selection and the regulation of defenses: a signal detection analysis of the smoke detector principle. *Evolution and Human Behavior* 26, 88–105.
- Nesse, R. M. and Williams, G. C. (1994). *Why we get sick: the new science of Darwinian medicine*. New York: Times Books.
- Rosen, J. B. and Schulkin, J. (1998). From normal fear to pathological anxiety. *Psychological Review* 105(2), 325–350.
- Young, E. and Vasquez, D. (1996). Hypercortisolemia, hippocampal glucocorticoid receptors, and fast feedback. *Molecular Psychiatry* 1(2), 149–159.
- Zhang, T. Y., Parent, C., Weaver, I., et al. (2004). Maternal programming of individual differences in defensive responses in the rat. *Annals of the New York Academy of Sciences* 1032, 85–103.

Excitatory Amino Acids

J V Nadler

Duke University Medical Center,
Durham, NC 27710

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
J V Nadler, volume 2, pp 84–89, © 2000, Elsevier Inc.

Synthesis and Release
Glutamate Receptors
Termination of Action
Excitotoxicity

Glossary

Excitatory postsynaptic potential (EPSP)
Excitotoxicity

The electrophysiological response evoked by glutamate released from mammalian nerve terminals.

Ionotropic receptor

Cellular injury caused by compounds, such as glutamate, that are both neuroexcitatory and neurotoxic.

Long-term depression
Long-term potentiation
Metabotropic receptor

An ion channel with binding affinity for a particular ligand, in which ligand binding promotes channel opening.

A long-lasting, stable reduction in the size of the postsynaptic response.

A long-lasting, stable increase in the size of the postsynaptic response.

A macromolecule with binding affinity for a particular ligand, in which ligand binding produces a physiological effect

- Munck, A., Guyre, P. M. and Holbrook, N. J. (1984). Physiological functions of glucocorticoids in stress and their relation to pharmacological actions. *Endocrine Review* 5(1), 25–44.
- Korte, S. M., Koolhaas, J. M., Wingfield, J. C., et al. (2005). The Darwinian concept of stress: benefits of allostasis and costs of allostatic load and the trade-offs in health and disease. *Neuroscience and Biobehavioral Review* 29, 3–38.
- Nesse, R. M. (2005). Natural selection and the regulation of defenses: a signal detection analysis of the smoke detector principle. *Evolution and Human Behavior* 26, 88–105.
- Nesse, R. M. and Williams, G. C. (1994). *Why we get sick: the new science of Darwinian medicine*. New York: Times Books.
- Rosen, J. B. and Schulkin, J. (1998). From normal fear to pathological anxiety. *Psychological Review* 105(2), 325–350.
- Young, E. and Vasquez, D. (1996). Hypercortisolemia, hippocampal glucocorticoid receptors, and fast feedback. *Molecular Psychiatry* 1(2), 149–159.
- Zhang, T. Y., Parent, C., Weaver, I., et al. (2004). Maternal programming of individual differences in defensive responses in the rat. *Annals of the New York Academy of Sciences* 1032, 85–103.

Excitatory Amino Acids

J V Nadler

Duke University Medical Center,
Durham, NC 27710

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
J V Nadler, volume 2, pp 84–89, © 2000, Elsevier Inc.

Synthesis and Release
Glutamate Receptors
Termination of Action
Excitotoxicity

Glossary

<i>Excitatory postsynaptic potential (EPSP)</i>	The electrophysiological response evoked by glutamate released from mammalian nerve terminals.
<i>Excitotoxicity</i>	Cellular injury caused by compounds, such as glutamate, that are both neuroexcitatory and neurotoxic.
<i>Ionotropic receptor</i>	An ion channel with binding affinity for a particular ligand, in which ligand binding promotes channel opening.
<i>Long-term depression</i>	A long-lasting, stable reduction in the size of the postsynaptic response.
<i>Long-term potentiation</i>	A long-lasting, stable increase in the size of the postsynaptic response.
<i>Metabotropic receptor</i>	A macromolecule with binding affinity for a particular ligand, in which ligand binding produces a physiological effect through coupling to a GTP-binding (G) protein.
<i>Receptor desensitization</i>	An agonist-induced reduction of the response transduced by a receptor.

Transporter

A membrane-bound macromolecule that selectively binds a particular ligand, transfers it across the cell membrane (both with and against its concentration gradient), and releases it on the other side.

Several naturally occurring amino acids possess neuroexcitatory properties and may be transmitters. A large amount of aspartate ($\sim 3 \mu\text{mol g}^{-1}$ wet weight) is present in the brain, as are small amounts of other excitants, including cysteine sulfinic acid and quino-
linic acid. However, current evidence favors glutamate as the transmitter used by nearly all excitatory pathways in the mammalian central nervous system. These pathways include many that are known to be involved in responses to stress, such as the outflow from the hippocampus to the hypothalamus. Therefore, only glutamate mechanisms are covered here.

Synthesis and Release

The cerebral cortex and other higher brain centers contain more glutamate than any other part of the body, approximately $12 \mu\text{mol g}^{-1}$ wet weight. This is more than twice the glutamate content of the spinal cord and much higher than the glutamate content of the peripheral organs. Immunocytochemical labeling methods show an excess of glutamate in excitatory neurons and their axon terminals. The terminals contain small, clear, round synaptic vesicles. The postsynaptic element is usually a dendritic spine with a broad postsynaptic density. The postsynaptic density contains glutamate receptors and their associated signaling molecules. Approximately 60–70% of all synapses in the brain appear to be glutamate synapses (Figure 1).

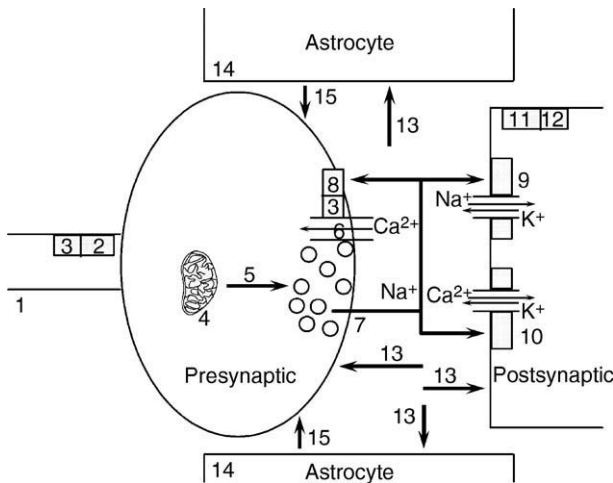


Figure 1 Diagram of a glutamate synapse. (1) Presynaptic action potential. (2) Type II metabotropic receptor. (3) G_o, G_i. (4) Glutaminase. (5) Vesicular transport. (6) Voltage-dependent Ca²⁺ channel. (7) Exocytotic release. (8) Type III metabotropic receptor. (9) AMPA receptor. (10) NMDA receptor. (11) Type I metabotropic receptor. (12) G_q. (13) Plasma membrane transport. (14) Glutamine synthetase. (15) Diffusion and terminal uptake of glutamate.

Glutamate can be synthesized by at least six different metabolic routes. However, most of the glutamate that is used as a transmitter is synthesized from glutamine. This biosynthetic process involves cooperation between nerve terminals and the adjacent astrocytes. Glutamate is taken up by astrocytes and converted to glutamine by glutamine synthetase, an enzyme expressed by glia but not by neurons. The glutamine synthetase reaction requires ATP. Astrocytic end feet, enriched in glucose transporters, cover virtually all capillary walls in the brain. Glutamate transport into the astrocytes stimulates glucose uptake by these cells. Glutamine produced by the astrocytes diffuses into the extracellular space and is taken up by the nerve terminals. Glutamate nerve terminals are enriched in phosphate-stimulated glutaminase, a mitochondrial enzyme that hydrolyzes glutamine to glutamate. This synthetic pathway is referred to as the glutamate–glutamine cycle. The glutamate–glutamine cycle provides a mechanism for coupling glutamate transmission with glucose use. The energy demands of this cycle account for approximately 85% of total brain glucose use!

The expression of vesicular glutamate transport determines that a particular synaptic terminal will release glutamate. The vesicular transporters are very specific for glutamate and exclude even closely related amino acids, such as aspartate. Transport is driven by an inside-positive vesicular membrane potential and a pH gradient, which are both created by vacuolar H⁺-ATPase activity. Three cDNAs that encode distinct vesicular glutamate transporters have been cloned. The transporters are expressed by different neuronal

populations. Two of the vesicular transporters were originally identified as inorganic phosphate transporters located on the plasma membrane of the nerve terminal. It is now known that the transporter is incorporated into the vesicular membrane in an orientation that allows for the inward transport of glutamate. After exocytosis, the transporter is transiently incorporated into the plasma membrane in the opposite conformation (inside-out relative to the synaptic vesicle). In that conformation, it can transport inorganic phosphate into the terminal.

Glutamate is released on nerve terminal depolarization by Ca²⁺-dependent exocytosis. In the mature brain, the opening of P/Q-type Ca²⁺ channels in the nerve terminal membrane provides the bulk of the Ca²⁺ required to drive this process. Glutamate can regulate its further release by feeding back onto terminal autoreceptors. Glutamate release is also regulated by the previous activity of the synaptic terminal, by other transmitters that may be present locally in the extracellular fluid, and by products of cell metabolism, such as adenosine and arachidonic acid.

Glutamate Receptors

Glutamate acts on several structurally and pharmacologically different types of membrane receptors, both ionotropic and metabotropic. The ionotropic glutamate receptors, for historical reasons, are named for structural analogs of glutamate that selectively activate that particular receptor: (RS)- α -amino-3-hydroxy-5-methyl-4-isoxazolepropionate (AMPA) and *N*-methyl-D-aspartate (NMDA). The expression pattern, density, subunit composition, and posttranslational modification of glutamate receptors are modulated by ontogenic development and by changes in physiological and pathological state.

AMPA Receptor

Nearly all glutamate-activated fast EPSPs depend on the activation of AMPA receptors. AMPA receptors have the multisubunit structure characteristic of all ligand-gated ion channels. For AMPA receptors in particular and ionotropic glutamate receptors in general, the receptor is composed of (probably) four glycopolypeptide chains arranged like the staves of a barrel with a channel in the middle through which ions can flow. The binding of glutamate to the receptor initiates a conformational change that opens the channel and allows ion flow. AMPA receptors typically bind glutamate with low affinity ($K_D \sim 100 \mu\text{M}$) and desensitize rapidly. That is, the channel current declines to a lower steady-state level while glutamate is still bound.

The AMPA receptor forms what is known as a small cation channel. The channel is permeable to both Na^+ and K^+ . When glutamate opens the channel, Na^+ flows down its electrochemical gradient into the postsynaptic cell and K^+ flows down its electrochemical gradient out of the cell. Because the Na^+ flux exceeds the K^+ flux ($E_{\text{Na,K}} \sim 0 \text{ mV}$), the net effect is a depolarization – the fast EPSP.

Four cDNAs that encode AMPA receptor subunits, designated GluR1–4 (or GluRA–D), have been cloned. AMPA receptor subunits can form a channel either by themselves (homomeric) or with other subtypes (heteromeric). Each combination of subunits forms a channel with slightly different properties. Moreover, each of the AMPA receptor subunits can exist in flip and flop versions, depending on alternative splicing. The involvement of flip versus flop versions of the subunit in channel formation affects the extent of receptor desensitization. Molecular biological studies of the ionotropic glutamate receptors led to two surprising and unique findings. First, the ionic selectivity of the channel is determined by a single amino acid in the putative second transmembrane (actually an intramembrane) domain. This amino acid can be glutamine, asparagine, or arginine. If a glutamine is present (as in GluR1, -3, or -4), a homomeric channel has a fairly high permeability to both Ca^{2+} and Mg^{2+} . If an arginine is present (as in GluR2), any channel that includes the subunit exhibits low permeability to Ca^{2+} and Mg^{2+} . (If an asparagine is present, as in NMDA receptor subunits, the channel exhibits high Ca^{2+} permeability and low Mg^{2+} permeability). Because the great majority of AMPA receptors expressed by excitatory neurons have low Ca^{2+} permeability, they must include at least one GluR2 subunit. In contrast, many AMPA receptors expressed by inhibitory neurons lack a GluR2 subunit and thus exhibit high Ca^{2+} permeability. Second, the presence of arginine at the critical site depends on RNA editing. All AMPA receptor genes code for a glutamine residue there. Selected mRNAs are altered after transcription to code for arginine (an adenine residue is converted to inosine).

NMDA Receptor

The NMDA receptor usually coexists with the AMPA receptor in the postsynaptic membrane. However, some synapses have only AMPA receptors. Immature (or silent) synapses have only NMDA receptors. (Silent synapses show no postsynaptic response at typical resting membrane potentials, -60 to -80 mV). Like the AMPA receptor, the NMDA receptor probably consists of four glycopolypeptide subunits. cDNAs for five subunits, designated NMDAR-1 and NMDAR-2A–D, have been cloned. Native NMDA receptors appear

to consist of the NMDAR-1 subunit and at least one of the NMDAR-2 subunits. Each combination of subunits forms a channel with slightly different properties. The development of the brain involves a change in the subunit composition of NMDA receptors. During the early development of the hippocampus, for example, NMDA receptors consist predominantly of the NR1–NR2B combination. Maturation involves replacement of NR2B subunits with NR2A. The physiological effect of this substitution is a reduction in the amplitude and duration of NMDA receptor-mediated postsynaptic currents.

The extent to which NMDA receptor activation contributes to the EPSP depends on the resting membrane potential of the postsynaptic cell; its contribution increases as the membrane potential becomes less negative. At most synapses, the NMDA receptor plays an insignificant role in the EPSP evoked by a single impulse. In general, the NMDA receptor serves as a mechanism by which experience alters the synaptic transmission and properties of the postsynaptic cell for a period of hours to years. Such diverse examples of neuronal plasticity as certain forms of memory and learning, opiate tolerance, kindling-induced epileptogenesis, and differentiation of the visual system have in common the requirement of NMDA receptor activation. The NMDA receptor has several unique properties that allow it to serve this function.

First, at normal membrane resting potential, the channel is blocked by Mg^{2+} ions present in the extracellular fluid of brain. The binding of Mg^{2+} is voltage-dependent; it binds much less strongly when the membrane is depolarized. Thus, NMDA receptors are activated only when glutamate release is paired with postsynaptic depolarization. In response to a single impulse, glutamate is released from the synaptic terminal and produces a depolarization by activating AMPA receptors. However, this EPSP is too short-lived to relieve much of the Mg^{2+} block. (The EPSP is short-lived because the AMPA receptor desensitizes rapidly when activated by glutamate; the AMPA receptor has a low affinity for glutamate, causing bound glutamate to dissociate from the receptor rapidly; and the following inhibitory postsynaptic potential, IPSP, returns the membrane potential toward its resting value.) Thus, the NMDA receptor at most synapses contributes little to the EPSP. During repetitive activity ($5 \text{ impulses s}^{-1}$ or more), the postsynaptic depolarization is more prolonged and Mg^{2+} block is effectively relieved. In addition, repetitive activity reduces γ -aminobutyric acid (GABA) inhibition. Under these conditions, the NMDA receptor contributes substantially to the EPSP. Compared to the AMPA receptor-mediated component, the NMDA component is of longer duration and exhibits slow

kinetics. High-frequency activity is associated with experiences that result in memory and learning; high-frequency activity can also be pathological, as in seizures.

Second, the NMDA receptor forms a small cation channel that is permeable not only to Na^+ and K^+ but also to Ca^{2+} . Like the AMPA receptor, it is the large flux of Na^+ and K^+ through the open channel that accounts for the NMDA receptor-mediated component of the EPSP. In contrast, the smaller Ca^{2+} influx is responsible for altering the properties of the synapse and the postsynaptic cell. Because the intracellular Ca^{2+} concentration at rest is very low, even a small influx of Ca^{2+} can increase the local intracellular concentration substantially. At this higher concentration, Ca^{2+} transiently activates a number of key enzymes, including protein kinases and enzymes that synthesize the diffusible modulators, nitric oxide and arachidonic acid. The concerted activation of these enzymes leads, by a mechanism that is only partially understood, to a long-lasting (days to weeks) enhancement of the EPSP. This process is known as long-term potentiation (LTP). There is growing evidence that the initial phase of NMDA receptor-dependent LTP involves the phosphorylation of AMPA receptors by Ca^{2+} /calmodulin-dependent protein (CaM) kinase II followed by the insertion of these receptors into the postsynaptic membrane. The subunit composition of postsynaptic AMPA receptors changes. In the basal state, these receptors are composed mainly of GluR2 and GluR3 subunits. The activation of CaM kinase II results in the phosphorylation of cytoplasmic GluR1 subunits, causing GluR1–GluR2 receptors to be inserted into the postsynaptic membrane. The Ca^{2+} that enters through the NMDA channel also activates transcription factors that alter gene expression. Proteins synthesized as a result of this Ca^{2+} influx stabilize the potentiated state of the synapse. A larger or longer-lasting influx of Ca^{2+} can also enhance NMDA receptor function and initiate changes in gene expression that are potentially damaging (e.g., that lead to epileptogenesis or apoptosis). Conversely, a smaller Ca^{2+} flux through the NMDA channel (such as from stimulation at 1 Hz for 10 min in immature rats) triggers a reverse process known as long-term depression (LTD). The mechanism of LTD involves the Ca^{2+} -dependent activation of phosphatases. Both LTP and LTD are thought to be important for information storage in the brain. During the development of the nervous system, LTP appears to be part of the mechanism for strengthening synaptic connections, whereas LTD leads to disconnection of the synapse. It should be noted that the strength of glutamate transmission can also be modulated at some sites by

processes that do not include the activation of NMDA receptors.

A third unique characteristic of the NMDA receptor is that glutamate and glycine (or the closely related amino acid D-serine) serve as co-agonists. That is, both amino acids must bind before the channel can open. Glycine is present in the extracellular fluid of brain at a concentration of 1–3 μM , enough to saturate its binding site on the NMDA receptor, and there is no evidence for a specific glycine-release process at glutamate synapses. The glycine binding site on the NMDA receptor is unrelated to the glycine receptor that mediates transmission at some inhibitory synapses. The NMDA receptor also has modulatory sites for Zn^{2+} (which inhibits noncompetitively), polyamines (which act as open channel blockers at normal resting potential and as potentiators at depolarized membrane potentials), oxidation-reduction (oxidation reduces and reduction enhances channel activity), H^+ (mild acidification reduces channel activation), and dissociative anesthetics/psychotomimetic drugs such as ketamine and phencyclidine (which act as open channel blockers).

Kainate Receptor

The kainate receptor resembles the AMPA receptor with respect to structure, regulation, and ionic selectivity. Kainate activates both AMPA and kainate receptors; there is considerable overlap in agonist action on the two receptors. In general, kainate receptors are most abundant in pathways in which NMDA receptors are least abundant. Postsynaptic kainate receptors appear to play their greatest role in synaptic transmission onto glutamate neurons mainly during repetitive activity. Kainate receptors are also present on GABA neurons, where their activation can both stimulate and depress inhibitory transmission. The mechanism by which these contrasting effects occur is not entirely clear. Finally, kainate receptors are present on some glutamate preterminal axons and on portions of the synaptic terminal outside the synapse. Both low- and high-affinity receptors may be present. The activation of low-affinity kainate receptors depolarizes the terminal, reducing subsequent glutamate release by limiting the Ca^{2+} influx triggered by action potentials. Where high-affinity kainate receptors are present, their activation enhances subsequent glutamate release. These receptors also work through presynaptic depolarization. The lesser depolarization produced by the activation of high-affinity kainate receptors prolongs the action potential-induced depolarization and enhances Ca^{2+} entry into the terminal. Whether Ca^{2+} entry is enhanced or reduced by kainate receptor activation thus depends on the magnitude and duration of presynaptic

depolarization. Through the regulation of Ca^{2+} influx, kainate receptors play a crucial role in presynaptic forms of synaptic facilitation at some synapses.

cDNAs that encode three low-affinity kainate receptor subunits (GluR5–7) and two high-affinity subunits (KA1–2) have been cloned. The native receptors are composed of either a combination of low- and high-affinity subunits or exclusively of low-affinity subunits. The GluR5 and GluR6 subunits are usually subjected to RNA editing, like GluR2 AMPA receptor subunits, which has the effect of reducing Ca^{2+} permeation through any channel in which they are included.

Metabotropic Glutamate Receptors

The metabotropic glutamate receptors have been subdivided into three groups, based on their agonist affinities and their preferred signal transduction pathways. Type I metabotropic glutamate receptors (mGluR1 and mGluR5) signal mainly through the activation of phospholipase C. Type I receptors are located postsynaptically but outside the synaptic cleft. They are activated mainly during high-frequency activity when glutamate overflows the synapse. The activation of these receptors typically produces a slow depolarization due to K^+ -channel inactivation. This action may convert single action-potential firing into multiple firing. The strong co-activation of AMPA and type I metabotropic receptors (but not of either one alone) can raise intracellular Ca^{2+} enough to activate phospholipase A_2 (leading to arachidonic acid production) and, in some cases, to provoke LTP. A selective agonist for these receptors is 3,5-DHPG.

The activation of type II metabotropic glutamate receptors (mGluR2 and mGluR3) can reduce adenylate cyclase activity, open a K^+ channel and/or inhibit the opening of voltage-dependent Ca^{2+} channels. Type II receptors are typically located on the preterminal portion of the axon and on the synaptic terminal membrane outside the synapse. They are activated when glutamate overflows the synapse. The activation of these receptors reduces subsequent glutamate release. DCG-IV is a selective agonist for these receptors.

Type III metabotropic glutamate receptors are particularly sensitive to the glutamate analog L-AP4. A type III receptor serves as the postsynaptic receptor for glutamate at the synapse in the retina between the photoreceptor cell and the bipolar cell (mGluR6). At this site, the receptor is coupled to the activation of cyclic nucleotide phosphodiesterase and reduced opening of cGMP-dependent cation channels. Light reduces the spontaneous release of glutamate from the photoreceptor, leading to the enhanced opening of the cGMP-dependent cation channels with consequent

bipolar cell depolarization. The signal transduction mechanisms of the other type III metabotropic glutamate receptors (mGluR4, mGluR7, and mGluR8) are the same as those of type II receptors. Type III metabotropic receptors are present on some glutamate terminals, where they are located near the active zones. Receptor activation reduces subsequent glutamate release, probably by inhibiting the opening of voltage-dependent Ca^{2+} channels.

Termination of Action

High-affinity ($K_T = 1\text{--}10\ \mu\text{M}$) transport mechanisms remove glutamate from the extracellular space. They are dependent on Na^+ , K^+ , and energy. The transporters are glycopolypeptides expressed on the plasma membrane of glia, neuronal cell bodies and dendrites, glutamate nerve terminals, and even some cells outside the central nervous system. Plasma membrane glutamate transporters serve several functions. First, they recover transmitter for reuse (nerve terminals) or as a precursor for glutamine (astrocytes). Second, they bind glutamate after its synaptic release, thus clearing the synaptic cleft in preparation for the next impulse and limiting the diffusion of glutamate away from the cleft. The latter effect inhibits glutamate from acting on the extrasynaptic receptors or receptors located at nearby synapses. Because NMDA receptors have a high affinity for glutamate, they would be inappropriately activated if much glutamate were to escape the synaptic cleft routinely. Third, they maintain the extracellular glutamate concentration below the level that will chronically activate the postsynaptic receptors. The maintenance of a low extracellular glutamate concentration is necessary to avoid excitotoxicity, and this is largely the function of glial transporters. Indeed, glutamate transporter 1 (GLT-1; excitatory amino acid transporter 2, EAAT2), the more heavily expressed glial transporter, accounts for 1% of total cell protein! Cloning studies have isolated cDNAs for six transporters, two glial and four neuronal. GLT-1, although expressed predominantly by astrocytes, is also expressed by most glutamate neurons and has been localized to the plasma membrane of the synaptic terminal. Synaptic terminal expression of GLT-1 presumably accounts for the robust glutamate transport observed in synaptosome preparations.

Excitotoxicity

Glutamate and other excitatory amino acids can kill neurons when they are present in high concentration for a prolonged period. This occurs mainly through

the excessive influx of Ca^{2+} . When present in the cytoplasm for an extended period of time, high Ca^{2+} concentrations can activate hydrolytic enzymes that result in cellular autolysis. More important, Ca^{2+} activates phospholipase A_2 , producing arachidonic acid; converts xanthine dehydrogenase to xanthine oxidase; and activates nitric oxide synthase, producing NO. The end result of these enzymatic cascades is the production of cytotoxic oxygen and hydroxyl free radicals. Ca^{2+} can also induce programmed cell death by activating transcription factors. Ca^{2+} enters the cell through the NMDA channel, through Ca^{2+} -permeable AMPA channels, through voltage-dependent Ca^{2+} channels, and through exchange mechanisms. Excitotoxicity is probably responsible for most of the neuronal degeneration associated with stroke, epilepsy, and head trauma. Glutamate has been implicated in the pathology of amyotrophic lateral sclerosis (through the loss of plasma membrane transporter function). The pathology of Alzheimer's disease, Huntington's disease, and other neuropathological states probably also involves excitotoxicity.

Further Reading

- Collingridge, G. L., Isaac, J. T. and Wang, Y. T. (2004). Receptor trafficking and synaptic plasticity. *Nature Reviews Neuroscience* 5, 952–962.
- Freneau, R. T., Voglmaier, S., Seal, R. P., et al. (2004). VGLUTs define subsets of excitatory neurons and suggest novel roles for glutamate. *Trends in Neuroscience* 27, 98–103.
- Gillessen, T., Budd, S. L. and Lipton, S. A. (2002). Excitatory amino acid neurotoxicity. *Advances in Experimental Medicine and Biology* 513, 3–40.
- Huang, Y. H. and Bergles, D. E. (2004). Glutamate transporters bring competition to the synapse. *Current Opinion in Neurobiology* 14, 346–352.
- Huettnner, J. E. (2003). Kainate receptors and synaptic transmission. *Progress in Neurobiology* 70, 387–407.
- Javitt, D. C. (2004). Glutamate as a therapeutic target in psychiatric disorders. *Molecular Psychiatry* 9, 984–997.
- Kelly, A. and Stanley, C. A. (2001). Disorders of glutamate metabolism. *Mental Retardation in Developmental Disabilities Research Reviews* 7, 287–295.
- Maragakis, N. J. and Rothman, J. D. (2004). Glutamate transporters: animal models to neurologic disease. *Neurobiology of Disease* 15, 461–473.
- Mattson, M. P. (2003). Excitotoxic and excitoprotective mechanisms: abundant targets for the prevention and treatment of neurodegenerative disorders. *NeuroMolecular Medicine* 3, 65–94.
- Mayer, M. L. and Armstrong, N. (2004). Structure and function of glutamate receptor ion channels. *Annual Review of Physiology* 66, 161–181.
- Moghaddam, B. and Wolf, M. E. (eds.) (2003). *Special issue on glutamate and disorders of cognition and motivation. Annals of the New York Academy of Sciences* 1003.
- Nicoll, R. A. (2003). Expression mechanisms underlying long-term potentiation: a postsynaptic view. *Philosophical Transactions of the Royal Society of London: Biological Sciences* 358, 721–726.
- Pin, J. P. and Acher, F. (2002). The metabotropic glutamate receptors: structure, activation mechanism and pharmacology. *Current Drug Targets – CNS Neurological Disorders* 1, 297–317.
- Schousboe, A. (2003). Role of astrocytes in the maintenance and modulation of glutamatergic and GABAergic neurotransmission. *Neurochemical Research* 28, 347–352.
- Shulman, R. G., Rothman, D. L., Behar, K. L., et al. (2004). Energetic basis of brain activity: implications for neuroimaging. *Trends in Neuroscience* 27, 489–495.

Excitotoxins

M P Mattson

National Institute on Aging Intramural Research Program, Baltimore, MD, USA

Published by Elsevier Inc.

Introduction

Excitotoxic Mechanisms

Evidence That Physiological and Psychological Stress Can Endanger Neurons

Stress Hormones and Excitotoxicity

Anti-excitotoxic Effects of Mild Neuronal Stress
Environmental and Genetic Risk Factors for Stress-Mediated Excitotoxic Neuronal Degeneration

Glossary

Excitotoxicity Cell injury or death caused by the excessive activation of excitatory transmitter receptors, particularly glutamate receptors. Neurons become vulnerable to excitotoxicity under conditions of reduced energy (glucose) availability and increased oxidative stress.

Shulman, R. G., Rothman, D. L., Behar, K. L., et al. (2004). Energetic basis of brain activity: implications for neuroimaging. *Trends in Neuroscience* 27, 489–495.

Excitotoxins

M P Mattson

National Institute on Aging Intramural Research Program, Baltimore, MD, USA

Published by Elsevier Inc.

Reactive oxygen species

One of several oxygen-derived molecules that have the potential to damage cellular proteins, nuclei acids, and lipids; examples include superoxide anion radical, hydroxyl radical, peroxynitrite, and hydrogen peroxide.

Introduction

Excitotoxic Mechanisms

Evidence That Physiological and Psychological Stress Can Endanger Neurons

Stress Hormones and Excitotoxicity

Anti-excitotoxic Effects of Mild Neuronal Stress

Environmental and Genetic Risk Factors for Stress-Mediated Excitotoxic Neuronal Degeneration

Glossary

Excitotoxicity Cell injury or death caused by the excessive activation of excitatory transmitter receptors, particularly glutamate receptors. Neurons become vulnerable to excitotoxicity under conditions of reduced energy (glucose) availability and increased oxidative stress.

Hippocampus A structure in the temporal lobe of the brain that plays a major role in learning and memory processes and that is a focus of nerve cell degeneration in several prominent disorders including stroke, epilepsy, and Alzheimer's disease.

Neurodegenerative disorders Disorders characterized by progressive nerve cell degeneration in the brain or spinal cord; examples include Alzheimer's disease, Parkinson's disease, and amyotrophic lateral sclerosis.

Neuronal calcium homeostasis The ability of the nerve cell to regulate intracellular free calcium levels during physiological conditions and in response to stress. Membrane-associated ion channels, calcium transporters, and cytosolic calcium-binding proteins participate in this process.

Introduction

The major excitatory neurotransmitter in the nervous system is glutamate. Most, if not all, neurons in the brain and spinal cord express one or more types of glutamate receptors. Glutamate receptors are classified as ionotropic (ion channels) and metabotropic (linked to GTP-binding protein signaling pathways). Ionotropic glutamate receptors include N-methyl-D-aspartate (NMDA) receptors, which flux high levels of Ca^{2+} , and kainate and α -amino-3-hydroxy-5-methylisoxazole-4-propionate (AMPA) receptors, which flux mainly Na^+ . Glutamatergic synapses play major roles in most functions of the nervous system, including learning and memory and motor and sensory processing. However, the overactivation of glutamate receptors, particularly under conditions of reduced energy availability and increased oxidative stress, can damage and kill neurons. The excitotoxic mechanism, which is reviewed here, involves the massive influx of Ca^{2+} through NMDA receptors and voltage-dependent Ca^{2+} channels and the subsequent generation of reactive oxygen species (ROS) and mitochondrial dysfunction. Classic examples of excitotoxic neuronal degeneration in humans include severe epileptic seizures and stroke. There is also considerable evidence that excitotoxicity contributes to the neurodegenerative process in chronic neurodegenerative conditions, including Alzheimer's disease, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis (ALS). This article considers the emerging evidence that the neuroendocrine stress response, and cellular stress responses in the nervous system, might modify excitotoxic neurodegenerative processes in neurodegenerative disorders.

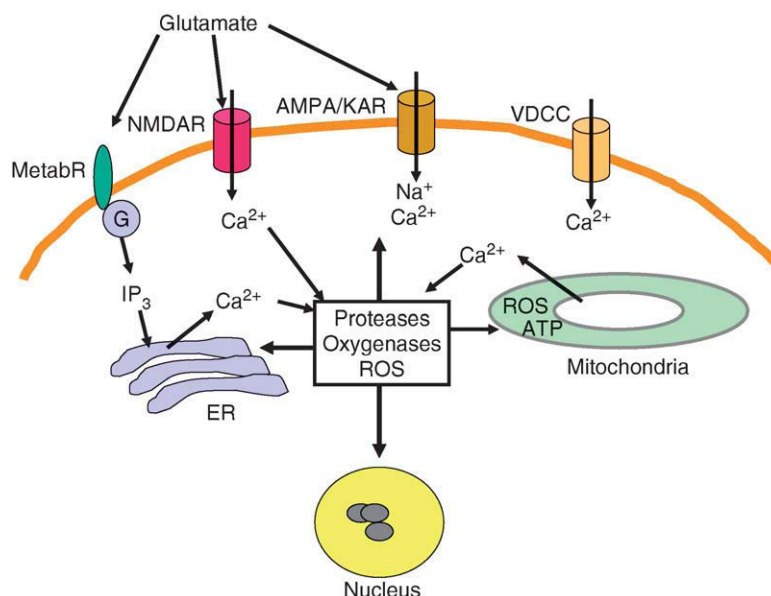


Figure 1 Excitotoxic mechanisms. The binding of glutamate to AMPA receptors and kainate receptors (KAR) opens the receptor channels, resulting in Na⁺ influx and consequent membrane depolarization and opening of voltage-dependent Ca²⁺ channels (VDCC). Some forms of AMPA receptor are also permeable to Ca²⁺. The binding of glutamate to NMDA receptors (NMDAR) under depolarizing conditions opens the NMDA receptor channel, resulting in large amounts of Ca²⁺ influx. The activation of metabotropic glutamate receptors (MetR) induces IP₃ production and the activation of IP₃ receptor and ryanodine receptor channels in the endoplasmic reticulum (ER) membrane, resulting in the release of Ca²⁺ from the ER into the cytoplasm. The increases in cytoplasmic Ca²⁺ levels in response to glutamate receptor activation can induce Ca²⁺ uptake into the mitochondria which, if excessive, can induce the production of reactive oxygen species (ROS) and inhibit ATP production. By activating proteases and inducing oxidative stress, Ca²⁺ is a key mediator of excitotoxic cell death. Modified from Mattson, M. P. (2003), Excitotoxic and excitoprotective mechanisms: abundant targets for the prevention and treatment of neurodegenerative disorders, *Neuromolecular Medicine* 3, 65–94.

Excitotoxic Mechanisms

Excitotoxicity is a complex process triggered by glutamate receptor activation that results in the degeneration of dendrites and in cell death. All subcellular compartments are affected by the excitotoxic process, with changes in the cytosol, mitochondria, endoplasmic reticulum (ER), and nucleus being pivotal. Normal amounts of glutamate receptor activation can damage neurons under conditions of metabolic and oxidative stress, which occur after a stroke or traumatic brain injury or in age-related neurodegenerative disorders.

Excitotoxicity is triggered by the overactivation of glutamate receptors, resulting in Na⁺ and Ca²⁺ influx across through the plasma membrane as the result of opening of glutamate receptor (AMPA–kainate and NMDA) channels and voltage-dependent Ca²⁺ channels (Figure 1). In addition, the activation of GTP-binding-protein-coupled metabotropic glutamate receptors stimulates inositol trisphosphate (IP₃) production and the release of Ca²⁺ from the ER. Depending on the particular type of neuron, its developmental stage, and various environmental factors, the relative contributions of AMPA–kainate, NMDA,

and metabotropic receptors and voltage-dependent Ca²⁺ channels to excitotoxicity may differ. An antagonist that selectively blocks one of the different glutamate receptors or Ca²⁺ channels may therefore exhibit differential effectiveness in protecting different populations of neurons against excitotoxicity. A neuron's abilities to remove and buffer Ca²⁺ are also important determinants of its susceptibility to excitotoxicity. Examples of such excitoprotective mechanisms are Ca²⁺-ATPases in the plasma membrane and ER, Na⁺/Ca²⁺ exchangers, the mitochondrial Ca²⁺ uniporter, and various Ca²⁺-binding proteins.

Both the magnitude and the duration of the increase of the intracellular Ca²⁺ concentration after glutamate receptor activation are important determinants of whether the neuron degenerates. Under resting conditions, the cytoplasmic Ca²⁺ concentration in neurons is typically approximately 100 nM. Large elevations of intracellular Ca²⁺ (low micromolar concentrations) can be tolerated as long as they are transient (recovering within seconds to minutes). On the other hand, even a lesser increase of the cytoplasmic Ca²⁺ concentration (to 500 nM, for example) that is sustained for more than 20–30 min can kill the neuron. Ca²⁺ damages dendrites and kills neurons,

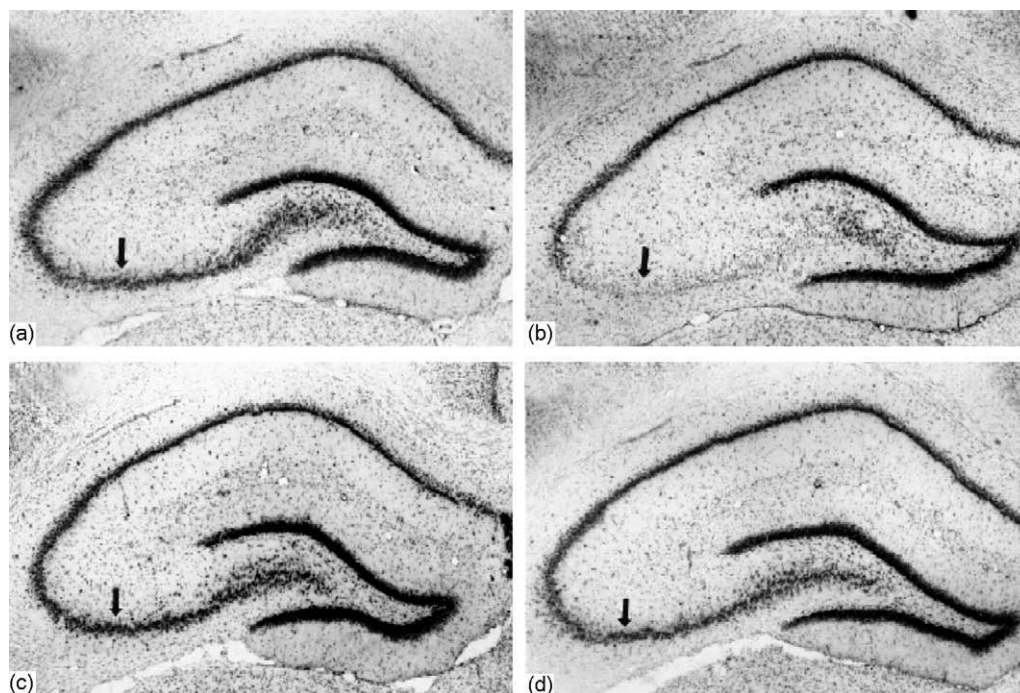


Figure 2 Cresyl violet-stained coronal sections of hippocampus showing that bacterial alkaloids that activate neurotrophic factor signaling pathways protect hippocampal neurons against excitotoxic death. a, Section from a sham-operated rat not receiving the excitotoxin kainic acid; b, section from a rat administered saline for 7 days and then injected with kainic acid; c, section from a rat administered K252a for 7 days and then injected with kainic acid; d, section from a rat administered K252b for 7 days and then injected with kainic acid. Note the extensive loss of neurons in the saline-treated mouse compared to the mice treated with K252a and K252b. Modified from Smith-Swintosky, V. L., Kraemer, P. J., Bruce, A. J., et al. (1996), Bacterial alkaloids mitigate seizure-induced hippocampal damage and spatial memory deficits. *Experimental Neurology* **141**, 287–296.

in part, by activating cysteine proteases called calpains that degrade a variety of substrates, including cytoskeletal proteins, membrane receptors, and metabolic enzymes.

Oxidative stress involves the production of ROS such as superoxide anion radical, hydrogen peroxide, and hydroxyl radical. ROS-mediated damage to proteins, membranes, and DNA plays a key role in excitotoxic damage to neurons. Glutamate-induced Ca^{2+} influx can cause ROS production by activating cyclooxygenases and lipoxygenases, perturbing mitochondrial metabolism and inducing membrane lipid peroxidation. Membrane lipid peroxidation may render neurons susceptible to excitotoxicity by impairing the function of membrane ion-motive ATPases and of glucose and glutamate transporters. The importance of oxidative stress in excitotoxicity has been demonstrated in cell culture and animal models in which antioxidants such as vitamin E, lipoic acid, and glutathione protect neurons.

Because of the high amounts of energy (ATP) required to maintain and restore ion gradients, neurons are particularly vulnerable to excitotoxicity when under conditions of reduced energy availability such as hypoglycemia and ischemia. Glutamate receptor antagonists can prevent the death of neurons under

such conditions, as can agents such as creatine that promote the maintenance of ATP levels. Damage to the DNA of neurons, such as occurs under conditions of oxidative stress or exposure to certain toxins, can render neurons vulnerable to excitotoxicity. DNA damage activates the enzyme poly-(ADP-ribose) polymerase (PARP), a nicotinic adenine dinucleotide-dependent enzyme that can consume large amounts of ATP. DNA damage can also induce the production and activation of a protein called p53, which can trigger apoptosis.

Excitotoxic neuronal death can occur rapidly as the result of massive influx of Na^+ through AMPA-kainate receptors and voltage-dependent Na^+ channels, resulting in cell swelling and lysis, a process called necrosis. Another, subtler form of neuronal death called apoptosis is characterized by cell body shrinkage, the formation of cell surface membrane blebs, and nuclear chromatin condensation and fragmentation. Excitotoxic apoptosis involves pivotal mitochondrial changes including increased membrane permeability and the release of cytochrome c. Cytochrome c then binds to a protein called Apaf-1, which mediates the activation of caspases 9 and 3. Drugs that stabilize mitochondrial membrane, as well as caspase

inhibitors, can protect neurons against excitotoxic death.

Many different signaling mechanisms can protect neurons against excitotoxicity. These include those activated by neurotrophic factors such as brain-derived neurotrophic factor (BDNF), basic fibroblast growth factor, and insulin-like growth factors. Two transcription factors that are known to induce the expression of excitoprotective genes are nuclear factor (NF)- κ B and cAMP response element binding protein (CREB). Examples of such protective genes are those encoding the antioxidant enzyme manganese superoxide dismutase and the anti-apoptotic protein Bcl-2. These endogenous protective pathways can be activated by certain drugs. For example, bacterial alkaloids that activate neurotrophic factor signaling pathways can protect hippocampal neurons against seizure-induced excitotoxic damage (Figure 2).

Evidence That Physiological and Psychological Stress Can Endanger Neurons

Early studies of the impact of stress on the brain focused on the action of sustained elevations of glucocorticoids. It was shown that damage to hippocampal neurons induced by the excitotoxin kainic acid was reduced in adrenalectomized rats and exacerbated in rats administered corticosterone. The suppression of endogenous glucocorticoid production, effected by administration of the 11- β -hydroxylase inhibitor metyrapone, significantly reduced kainic acid-induced damage to hippocampal neurons. Metyrapone administration also significantly reduced brain damage in two different rat models of stroke, a middle cerebral artery occlusion model of focal cerebral ischemia and a transient global forebrain ischemia paradigm that results in the selective loss of CA1 hippocampal neurons. The latter studies showed that both seizures and ischemia resulted in a massive stress response (i.e., large increases in plasma corticosterone levels) that was effectively suppressed by metyrapone. Consistent with a role for the elevation of glucocorticoids in the endangerment of neurons in human neurodegenerative conditions are data showing that plasma glucocorticoid levels are increased in patients with temporal lobe epilepsy; stroke; and Alzheimer's, Parkinson's, and Huntington's diseases. Not only were levels of circulating glucocorticoids shown to be elevated in stroke patients and in patients experiencing severe epileptic seizures, but increased levels and duration of elevation were correlated with worse outcome. Despite these and other compelling data, clinical trials of agents that

suppress glucocorticoid production in these disorders have not yet been pursued.

Recent animal studies have shown that physiological stress (physical stressors such as cold temperature or pain and psychological stressors such as crowding or exposure to dominant males) can endanger neurons in the brain. The subjection of adult rats to chronic stress resulted in the atrophy of apical dendrites of hippocampal CA3 pyramidal neurons. The subjection of adult rats to an intermittent stress regimen resulted in an increase in vulnerability of their hippocampal neurons to seizure-induced injury.

Stress Hormones and Excitotoxicity

Studies of primary hippocampal and cortical neurons in cell culture have provided direct evidence that glucocorticoids can increase neuronal vulnerability to excitotoxicity. Based on analyses of the effects of glucocorticoids on glucose uptake and metabolic parameters (e.g., ATP levels) in cultured neurons, Sapolsky and coworkers proposed that a primary mechanism underlying the endangering effects of glucocorticoids involves impaired energy availability. This hypothesis is consistent with many studies showing that neuronal vulnerability to excitotoxicity is increased under conditions of reduced energy availability. Further supporting the metabolic compromise hypothesis are data showing reduced energy availability to neurons in Alzheimer's disease. The exposure of cultured hippocampal neurons to corticosterone resulted in a reduction in the level of glucose transport. It has not been established whether the adverse effect of glucocorticoids on glucose transport is a transcription-dependent process mediated by classic steroid receptors. Indeed, an alternative mechanism of action is suggested by data showing that the inhibition of glucose transport occurs rapidly and requires higher concentrations of corticosterone than would be expected for a receptor-mediated process.

Measurements of intracellular calcium levels in hippocampal neurons exposed to glutamate and other excitatory amino acids have shown that glucocorticoids disrupt cellular calcium homeostasis and promote calcium overload. It is not known whether this calcium-destabilizing action of glucocorticoids can be explained solely on the basis of impaired glucose uptake and ATP depletion or whether it involves one or more of the recently described actions of glucocorticoids on ion channel function. Interestingly, estrogen protects cultured hippocampal neurons against excitotoxic, metabolic, and oxidative insults by a mechanism involving the suppression of oxidative stress. In particular, estrogen suppresses mem-

brane lipid peroxidation, apparently via the direct inherent antioxidant activity of this steroid. Indeed, estrogen can prevent the impairment of ion-motive ATPases and glucose and glutamate transporters in cortical synaptosomes, a preparation lacking nuclei.

Other possible endangering mechanisms of glucocorticoids include the suppression of the production of neurotrophic factors and the suppression of antioxidant defense mechanisms. Indeed, glucocorticoids increase the vulnerability of cultured hippocampal neurons to oxidative insults, whereas neurotrophic factors protect neurons against such insults. Amyloid β -peptide is a 40- to 42-amino-acid peptide that forms insoluble aggregates (plaques) in the brain in Alzheimer's disease; considerable evidence indicates that amyloid β -peptide may directly damage neurons and/or increase their vulnerability to excitotoxicity. Glucocorticoids significantly increase the vulnerability of hippocampal neurons to cell death induced by amyloid β -peptide. Measurements of levels of ROS in hippocampal neurons have shown that amyloid β -peptide increases peroxide accumulation and membrane lipid peroxidation and that corticosterone exacerbates such oxidative stress. Membrane lipid peroxidation is believed to play a major role in the neurodegenerative process in Alzheimer's disease by impairing the function of membrane ion-motive ATPases (Na^+/K^+ -ATPase and Ca^{2+} -ATPase) and glucose transporters.

Collectively, the available data suggest that excessive physiological and psychological stress can increase the vulnerability of at least some populations of neurons in the brain to excitotoxicity. The mechanism may involve metabolic compromise secondary to the suppression of glucose transport and/or alterations in cellular calcium homeostasis. Excitotoxic neurodegeneration involves oxidative stress and mitochondrial dysfunction, and glucocorticoids can exacerbate such oxidative stress either by directly affecting antioxidant pathways or by indirectly enhancing oxyradical production by disturbing calcium and energy homeostasis.

Although much attention has been given to the roles of glucocorticoids in exacerbating excitotoxicity, there are many other stress-induced hormones that may also influence a neuron's vulnerability to excitotoxicity. For example, corticotropin releasing hormone (CRH) and a related neuropeptide called urocortin have been shown to affect the vulnerability of neurons to oxidative, metabolic, and excitotoxic injury. CRH and urocortin act on receptors coupled to cAMP production. The activation of these receptors can protect cultured neurons against oxidative stress and excitotoxicity. Opioid peptides are another

class of stress-related hormones that can affect the excitotoxic process. For example, dynorphin and nociceptin can exacerbate excitotoxicity, whereas σ receptor ligands can have excitoprotective actions.

Anti-excitotoxic Effects of Mild Neuronal Stress

The kinds of data described so far provide strong evidence that excessive activation of the neuroendocrine stress axis is detrimental to neurons in certain regions of the brain. However, emerging findings suggest that moderate levels of stress may enhance neuronal plasticity and increase the resistance of neurons to various insults, including excitotoxic conditions. It has been repeatedly demonstrated that exposure of neurons, *in vivo* and in cell culture, to a moderate (subtoxic) level of stress can protect those neurons against a subsequent intense insult that would otherwise be neurotoxic. For example, the exposure of cultured hippocampal or cortical neurons to subtoxic levels of heat shock or excitotoxins renders them resistant to excitotoxicity. Moreover, the extent of the brain damage caused by cerebral ischemia in adult rodents is greatly reduced when the animals are subjected to a mild alchemic episode prior to exposure to severe ischemia. Such preconditioning hormesis may be mediated by the upregulation of heat shock proteins and antiapoptotic genes. The increased expression of heat shock proteins occurs in both acute neurodegenerative conditions, such as stroke and severe epileptic seizures, and in chronic neurodegenerative disorders, such as Alzheimer's disease. The available data suggest that these stress-responsive gene products serve neuroprotective functions.

In acute neurodegenerative conditions such as stroke, epileptic seizures, and traumatic brain injury, there is a robust and quite rapid increase in the expression of several different cytokines and neurotrophic factors. Considerable data suggest that this particular aspect of the stress response represents an attempt of brain cells to prevent neuronal death and promote recovery. Indeed, cell culture and *in vivo* studies have shown that the administration of some neurotrophic factors (e.g., basic fibroblast growth factor, nerve growth factor, BDNF, and insulin-like growth factors) and cytokines (e.g., tumor necrosis factor, TNF) can protect neurons against excitotoxic, metabolic, and oxidative insults. An example of a stress-responsive signaling pathway that increases neuronal resistance to excitotoxicity involves the activation of the transcription factor NF- κ B by intercellular signals such as TNF and secreted β -amyloid precursor protein (APP), as well as by increased levels

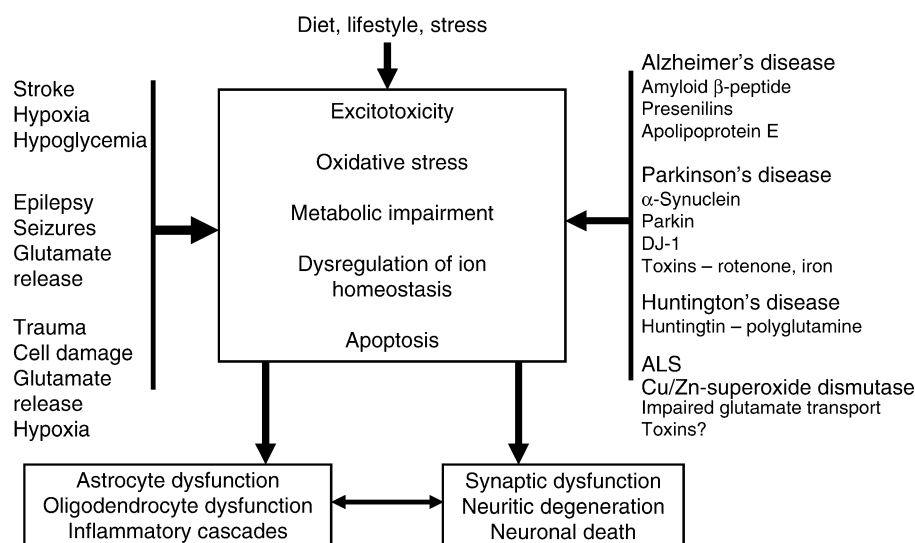


Figure 3 Excitotoxicity is a convergence point in the neurodegenerative cascades of each of the major acute and chronic neurodegenerative disorders. The genetic and environmental factors that initiate the neurodegenerative process may differ among disorders. For example, stroke is caused by the atherosclerotic occlusion of a cerebral blood vessel; Alzheimer's disease can result from mutations in the β -amyloid precursor protein or presenilins or by age-related increases in oxidative and metabolic stress, resulting in increased production of neurotoxic forms of amyloid β -peptide; and Huntington's disease is caused by polyglutamine expansions in the huntingtin gene. Despite such differences in initiating factors, each disorder results in similar neurodegenerative cascades that involve increased oxidative stress, metabolic impairment, and overactivation of glutamate receptors, resulting in excessive Ca^{2+} influx and excitotoxic cell death. Stress, dietary factors, and certain aspects of lifestyle can influence the risk of both the acute neurodegenerative conditions listed on the left and the chronic neurodegenerative disorders listed on the right. Modified from Mattson, M. P. (2003), Excitotoxic and excitoprotective mechanisms: abundant targets for the prevention and treatment of neurodegenerative disorders, *Neuromolecular Medicine* 3, 65–94.

of oxidative stress. The NF- κ B induces the expression of several cytoprotective gene products including Mn-superoxide dismutase.

A final example of a scenario in which moderate stress may be beneficial for the brain comes from recent studies of the impact of food restriction on aging of the brain and on the brain's response to metabolic and excitotoxic insults. A well-established method for increasing the life span of laboratory rodents and reducing the incidence of age-related cancers and immune alterations is to reduce their caloric intake. A decrease in levels of cellular oxidative stress appears to play an important role in the beneficial effects of food restriction. Recent studies have provided evidence that food restriction retards age-related alterations in the brain such as increases in glial reactivity and impaired performance on learning and memory tasks. With respect to interactions between stress and excitotoxicity, it was recently reported that food restriction in adult rat (alternate day feeding regimen for 2–4 months) results in resistance of hippocampal neurons to excitotoxin-induced degeneration and of striatal neurons to degeneration induced by the mitochondrial toxins 3-nitropropionic acid and malonate. Food restriction dramatically

reduced excitotoxin-induced deficits in learning and memory and also prevented the mitochondrial toxin-induced impairment of motor function. Although the mechanism underlying the beneficial effect of chronic food restriction in these rodent models of excitotoxic brain injury has not been established, it is possible that the food-restriction regimen subjects the body and brain to a moderate level of stress resulting in the upregulation of cellular defense mechanisms. In support of this scenario, data show that levels of activation of the hypothalamic-pituitary-adrenal axis and levels of heat shock proteins in some tissues are increased in food-restricted rodents.

Environmental and Genetic Risk Factors for Stress-Mediated Excitotoxic Neuronal Degeneration

An increasing number of environmental factors (e.g., diet, exercise, and exposure to toxins) and genetic factors are being identified that can affect the vulnerability of neurons to excitotoxicity (Figure 3). Several different environmental toxins have been identified that can induce nervous system damage and behavioral deficits that are remarkably similar to those seen

in human neurodegenerative conditions. As already described, kainic acid (produced by seaweed) can induce seizures and hippocampal damage reminiscent of human temporal lobe epilepsy. Similarly, domoic acid (produced by algae and concentrated in shellfish) can cause hippocampal neuron degeneration and severe memory loss. I-Methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) damages dopaminergic neurons by a mechanism involving, in part, the excessive activation of glutamate receptors under conditions of increased oxidative stress. The ingestion of a mitochondrial toxin (3-nitropropionic acid) produced by a mold results in selective damage to striatal neurons and motor deficits similar to those seen in Huntington's disease. An excitotoxin in the cycad seed may play a role in the widespread neurodegeneration in brains and spinal cords of individuals suffering from ALS–Parkinsonism–dementia complex of Guam. There is evidence that both the neuroendocrine stress response and neuronal stress responses can modulate the adverse effects of such excitotoxins.

Genetic defects have been identified that cause several major neurodegenerative disorders in which excitotoxicity is believed to play a role. For example, mutations in APP and the presenilins (PS-1 and PS-2) are responsible for some cases of early-onset autosomal-dominant familial Alzheimer's disease. Mutations in APP may render neurons vulnerable to age-related metabolic and oxidative stress by altering the proteolytic processing of APP in a manner that increases the production of neurotoxic amyloid β -peptide and decreases the levels of a neuroprotective secreted form of APP. Presenilin mutations appear to place neurons at risk by perturbing calcium regulation in the ER, which, in turn, leads to enhanced oxidative stress and mitochondrial dysfunction when neurons are subjected to metabolic and excitotoxic insults. Glucocorticoids have been shown to exacerbate the neurodegenerative process initiated by amyloid β -peptide and may also interact with presenilin mutations to disturb neuronal calcium homeostasis. Conversely, estrogen protects neurons against the death-enhancing effects of PS-1 mutations.

ALS, a disease in which the lower motor neurons degenerate, is believed to involve increased oxidative stress and excitotoxicity. Mutations in the antioxidant enzyme Cu/Zn-superoxide dismutase are responsible for some cases of inherited autosomal-dominant ALS. Huntington's disease is a genetic disorder in which the aberrant gene (called huntingtin) exhibits expansions of a trinucleotide repeat encoding the amino acid glutamine. Although it has not been established how the trinucleotide repeats promote the degeneration of

striatal neurons, available data suggest an alteration that increases neuronal vulnerability to excitotoxicity.

In each of these genetic neurodegenerative disorders, the defective gene is widely expressed in various cell types throughout the body and nervous system. The available data suggest that a major reason neurons are selectively vulnerable in these disorders is because of their high metabolic demands and unique vulnerability to excitotoxicity. For the same reason, various extrinsic and endogenous stressors (including the aging process itself) can have profound effects in promoting excitotoxic neuronal degeneration.

See Also the Following Articles

Alzheimer's Disease; Brain Trauma; Epilepsy; Excitatory Amino Acids.

Further Reading

- Beal, M. F. (1992). Does impairment of energy metabolism result in excitotoxic neuronal death in neurodegenerative illnesses? *Annals of Neurology* 31, 119–130.
- Bruce-Keller, A. J., Umberger, G., McFall, R., et al. (1998). Food restriction reduces brain damage and improves behavioral outcome following excitotoxic and metabolic insults. *Annals of Neurology* 45, 8–15.
- Fassbender, K., Schmidt, R., Mossner, R., et al. (1994). Pattern of activation of the hypothalamic-pituitary-adrenal axis in acute stroke: relation to acute confusional state, extent of brain damage, and clinical outcome. *Stroke* 25, 1105–1108.
- Mattson, M. P. (2003). Excitotoxic and excitoprotective mechanisms: abundant targets for the prevention and treatment of neurodegenerative disorders. *Neuromolecular Medicine* 3, 65–94.
- McEwen, B. S. (2004). Protection and damage from acute and chronic stress: allostasis and allostatic overload and relevance to the pathophysiology of psychiatric disorders. *Annals of the New York Academy of Sciences* 1032, 1–7.
- Pedersen, W. A., Wan, R., Zhang, P., et al. (2002). Urocortin, but not urocortin II, protects cultured hippocampal neurons from oxidative and excitotoxic cell death via corticotropin-releasing hormone receptor type I. *Journal of Neuroscience* 22, 404–412.
- Pomara, N., Greenberg, W. M., Branford, M. D., et al. (2003). Therapeutic implications of HPA axis abnormalities in Alzheimer's disease: review and update. *Psychopharmacology Bulletin* 37, 120–134.
- Sapolsky, R. M. (1996). Stress, glucocorticoids, and damage to the nervous system: the current state of confusion. *Stress* 1, 1–19.
- Smith-Swintosky, V. L., Kraemer, P. J., Bruce, A. J., et al. (1996). Bacterial alkaloids mitigate seizure-induced hip-

pocampal damage and spatial memory deficits. *Experimental Neurology* 141, 287–296.

Stein-Behrens, B., Mattson, M. P., Chang, I., et al. (1994). Stress exacerbates neuron loss and cytoskeletal pathol-

ogy in the hippocampus. *Journal of Neuroscience* 14, 5373–5380.

Exercise

P Khatri and J A Blumenthal

Duke University Medical Center, Durham, NC USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Khatri and J A Blumenthal, volume 2, pp 98–102, © 2000, Elsevier Inc.

General Methodological Issues
 Cardiorespiratory Responses to Exercise
 Psychological Responses to Exercise
 Health Risks of Exercise
 Mechanisms
 Summary and Future Directions

Glossary

<i>Acute exercise</i>	A single session or bout of exercise.
<i>Aerobic activities</i>	Activities using the large muscle groups at mild to moderate intensities that permit the body to use oxygen to supply energy and to maintain a steady state for more than a few minutes. Examples include walking, jogging, swimming, and cycling.
<i>Anaerobic activities</i>	Activities using the muscle groups at high intensities that exceed the body's capacity to use oxygen to supply energy and that create an oxygen debt by using energy produced without oxygen. Examples include weight lifting and sprinting.
<i>Cardiorespiratory endurance</i>	The ability to continue or persist in strenuous tasks involving the large muscle groups for extended periods of time. The ability of the circulatory and respiratory systems to adjust to and recover from the effects of whole-body exercise or work. Also called aerobic endurance.
<i>Chronic exercise</i>	Exercise usually performed for 6–8 weeks or longer.
<i>Physical fitness</i>	A dynamic state of energy and vitality that enables one to carry out daily tasks, to engage in active leisure-time

pursuits, and to meet unforeseen emergencies without undue fatigue.

The impact of exercise training and improved physical fitness on stress including psychological, social, and physical functioning has received considerable attention, particularly in light of the increasing popularity of exercise as a leisure-time activity. Exercise may vary by mode (i.e., kind of exercise), duration, intensity, and frequency. There are two primary kinds of exercise that are classified on the basis of how energy to perform work is derived: aerobic exercise and anaerobic exercise. Most of the research on stress and exercise has focused on the psychological benefits of chronic aerobic exercise (i.e., exercise performed several times per week over a period of weeks or months).

General Methodological Issues

The empirical evidence for the relation between exercise and stress is derived from multiple disciplines (psychology, exercise physiology, psychophysiology, physical therapy, etc.) that may approach the topic from different perspectives and conceptual models. The type of experimental design (e.g., cross-sectional, longitudinal, or interventional) is an important consideration. Cross-sectional studies typically compare stress responses of fit versus unfit groups, athletes versus nonathletes, or people who engage in one form of exercise with individuals who engage in different exercise (e.g., joggers vs. weight lifters). Because of the correlational nature of cross-sectional studies, however, causal attributions are not possible. Longitudinal designs also are correlational and typically follow subjects who engage in different kinds of exercise over time. Interventional studies typically involve the assignment of individuals to either exercise or nonexercise conditions and compare the groups over time. Random group assignment is necessary to

attribute differences in outcomes to the intervention. Control groups, another important design consideration, should be matched as closely as possible to the experimental group on nonspecific variables (e.g., demand characteristics, experimental contact, and social support). Other important issues include methods for assessment of aerobic fitness operationally and procedures for measuring stress. Although there are a number of ways to measure aerobic fitness (e.g., questionnaires, 10-min walk/run, step test, and treadmill test), the most precise assessment of aerobic fitness involves a multistage maximal exercise test on a treadmill or cycle ergometer with direct measurements of oxygen consumption. Stress responses can be measured by self-report questionnaires or may be measured physiologically. For example, heart rate and blood pressure are commonly used measures of stress response, most likely due to their ease of measurement and sensitivity to differences in fitness levels. The selection of laboratory stressors is important because the stressor must elicit a strong psychophysiological or subjective response that is measurable and potentially modifiable. The most commonly used stressors include cognitive tasks such as mental arithmetic, the Stroop Test, or public speaking; performance tasks such as the Mirror Trace Task or video games; or physical tasks such as the hand dynamometer test or cold-pressor test. Ideally, a number of stressors, varying in intensity and patterns of physiological response, may be used most effectively in identifying changes in stress reactivity. Other physiological indices of stress, and particularly of sympathetic nervous system activation, such as secretion of urinary catecholamines, also may be used to provide an overall physiological index of stress. Adherence is another important issue in exercise research, particularly because of the high rates of nonadherence. It has been estimated that more than 50% of people who initiate an exercise program will terminate the program within the first several months. This issue is especially relevant in light of the fact that exercise must be maintained in order to achieve its physical and mental health benefits. Such strategies for improving compliance have been the focus of much research.

Cardiorespiratory Responses to Exercise

The ability to engage in physical exercise depends on the ability of the circulatory system to deliver oxygen to exercising muscles. When exercise is initiated, the heart rate is increased, first by withdrawal of the parasympathetic nervous system and then by the activation of the sympathetic nervous system. Stroke volume (i.e., the amount of blood pumped from the heart with each contraction)

increases, along with cardiac output. The heart rate may increase up to threefold the resting values, and the stroke volume may increase by 15%. Systolic blood pressure typically increases proportional to exercise intensity, whereas diastolic blood pressure normally remains unchanged or may increase slightly.

During exercise, there is a gradual increase in oxygen demand by the exercising muscles. As the oxygen delivered in the arterial blood is used by the muscle, the oxygen level in the returning venous blood is decreased, resulting in a widening of the arterial-venous oxygen difference. These changes in oxygen demand and use are reflected by increased oxygen consumption (VO_2). Repeated bouts of aerobic exercise produce training effects, increased maximal VO_2 and lowered heart rate at rest and during submaximal exercise workloads.

Psychological Responses to Exercise

The psychological effects of chronic exercise on personality functioning and mood have been widely studied. Although there do not appear to be any specific personality types associated with specific kinds of exercise (e.g., joggers vs. weight lifters), active individuals tend to have higher self-esteem and greater self-confidence than sedentary individuals. Athletes (with the exception of marathon runners) are more likely to be extroverted and less neurotic than nonathletes. Athletes also tend to have less anxiety, depression, anger, and fatigue and more vigor than their nonathletic counterparts. Interventional studies have shown exercise to improve self-concept, self-esteem, and self-efficacy.

The type A personality is a behavioral complex that has been shown to be associated with increased risk for premature cardiovascular disease. Type A individuals are more competitive, hard driving, impatient, and aggressive than their more easy-going type B counterparts. Although type As are not more likely to be physically active than type Bs, type A behavior has been shown to be reduced by exercise training, including both jogging and weight lifting.

There is increasing evidence that chronic exercise results in improved mood and psychological well-being. Cross-sectional studies have demonstrated that active individuals report less anxiety and depression and better emotional well-being relative to their sedentary counterparts. Longitudinal studies also have reported that negative mood states can be reduced with regular exercise. A number of interventional studies have demonstrated that increasing physical activity results in reduced levels of stress, anxiety, and depression. It has been shown that sedentary individuals are more likely to be depressed

than active people and that individuals who increase their activity levels have no greater risk for depression than people who maintain higher activity levels. Moreover, active people who become sedentary may increase in their risk for depression by 50%.

Studies also have investigated the relation between exercise and physiological stress responses. A number of cross-sectional studies have demonstrated that physically fit or physically active individuals exhibit lower cardiovascular responses to physical and mental stressors compared to unfit or sedentary individuals. Because regular exercise results in physiological adaptations that may affect responses to psychological challenges, interventional studies also have examined the effects of exercise training on cardiovascular and neuroendocrine responses to laboratory stressors such as mental arithmetic and public speaking. The results of these studies have shown that heart rate and blood pressure responses to these mental stressors are attenuated after 12–16 weeks of exercise and that aerobic exercises (e.g., jogging or biking) are generally more effective than anaerobic exercises (e.g., strength training) in reducing psychophysiological stress responses. The extent to which these altered physiological responses to laboratory tasks generalize to stressors during daily life, however, has not been examined.

The acute effects of exercise also have received attention. It has been hypothesized that the chronic adaptations associated with exercise result from the accumulated effects of individual bouts of exercise over time. Consequently, investigators also have examined the acute effects of exercise on various measures of stress. Studies of the short-term effects of exercise provide evidence for the mood-enhancing benefits of acute aerobic exercise. Correlational and experimental studies have demonstrated reduction in depression, anxiety, tension, and anger following a single bout of exercise. However, the intensity level at which exercise is beneficial is unclear. Some studies have found benefits from low- to moderate-intensity exercise and negative consequences of high-intensity exercise (e.g., greater increases in tension and fatigue); however, other studies have demonstrated reductions in psychological tension following moderate to intense exercise but not mild exercise. Acute aerobic exercising also has been shown to attenuate cardiovascular responses to laboratory stressors in several studies, although the data are too limited to draw definite conclusions.

Health Risks of Exercise

The potential adverse effects of exercise are important considerations in evaluating exercise as a stress-management technique. It is important to note that

Table 1 Stress-related changes with chronic aerobic exercise

	<i>Change</i>
<i>Physiological</i>	
Resting heart rate	↓
Resting blood pressure	↓
Catecholamines	↓
Vagal tone	↑
<i>Psychological</i>	
Anxiety	↓
Depression	↓
Self-esteem	↑
<i>Psychophysiological</i>	
Blood pressure	↓
Heart rate	↓
Catecholamines	↓

although exercise is widely regarded as a beneficial activity, exercise may be associated with untoward events in susceptible individuals. However, even in cardiac patients, the potential adverse effects of exercise such as arrhythmias and myocardial infarction are relatively rare and occur in less than 1 of 10 000 exercise tests and at a rate of less than 1 per 100 000 h of exercise training. Minor muscle aches and joint pains are common – either as a result of overuse or acute trauma – but are usually self-limited. Major orthopedic problems are very uncommon in supervised exercise settings, but may be more common in unsupervised programs or in individuals who fail to comply with exercise prescriptions. The annual rate of exercise-related injury in adults is roughly 10%, with only 5% requiring medical care. These injuries often result from a sudden unexpected trauma, such as twisting an ankle. Proper warm-up and stretching are useful preventative measures. Overuse injuries, on the other hand, are a result of chronic, accumulated stress on the musculoskeletal system. This type of injury is more common in high-impact, weight-bearing activities such as jogging and includes such problems as tendonitis, stress fractures, and ruptured disks in the lower back. Gradual progression of exercise, attention to environmental factors (e.g., heat and humidity, altitude), and the use of proper equipment all serve to reduce the risk of injury.

Exercise addiction is a term related to overuse and is generally described as an addiction of psychological and/or physiological nature to regular physical activity, characterized by withdrawal symptoms after 2 to 3 days without exercise. Although exercise as a compulsive behavior has been characterized as a positive addiction that promotes well-being, the potential for abuse has been recently recognized. Individuals may experience depression, anxiety, and irritability if they are unable to exercise. There is some evidence that exercise addiction can have an adverse impact on

interpersonal and occupational functioning, as well as on self-esteem. There is also some suggestion of overlap between excessive exercise and eating disorders. Obligatory running in men, for example, has been likened to anorexia nervosa in women. However, there is little empirical support for the notion that either the etiology or level of psychopathology is comparable between anorexics and compulsive exercisers. Eating disorders also may be more prevalent among women, particularly among participants in certain forms of exercise (e.g., ballet, gymnastics, and figure skating).

Mechanisms

The mechanisms by which exercise reduces stress are not known. A number of psychological mechanisms have been proposed, including increased self-efficacy, a sense of mastery, positive thoughts, distraction from negative thoughts, and enhanced self-concept, which in turn can lead to improved mood. Exercise may also be a form of meditation that triggers a more relaxed state or a form of biofeedback that teaches people to regulate their autonomic arousal. Exercise may also be associated with social reinforcement. There is not, however, definitive evidence to support any one explanation.

A number of biological mechanisms also have been suggested, although there is no consensus as to the precise mechanism whereby stress is reduced. There are data to suggest that the antidepressant effect of chronic exercise may be mediated by increased central norepinephrine neurotransmission. Changes in the hypothalamic-pituitary-adrenocortical axis may also be related to improved mood due to the lessened cortisol response to submaximal exercise following exercise training. Increased levels of β -endorphins may also contribute to the antidepressant effects of exercise. However, most studies have failed to employ adequate

methodologies to accurately measure circulating endorphin levels in the brain.

Summary and Future Directions

Both acute exercise and chronic exercise have been shown to reduce stress responses measured behaviorally, psychologically, and physiologically (see [Table 1](#)). Research efforts are currently directed at identifying those individuals who are most likely to benefit from exercise, at examining the mechanisms by which exercise reduces stress, and at developing strategies for improving long-term adherence to exercise and for motivating people to initiate exercise behavior.

See Also the Following Articles

Aerobic Exercise and Stress Reduction; Cardiovascular System and Stress; Relaxation Techniques; Type A Personality, Type B Personality.

Further Reading

- Brosse, A. L., Sheets, E. S., Lett, H. S. and Blumenthal, J. A. (2002). Exercise and the treatment of clinical depression in adults: recent findings and future directions. *Sports Medicine* 32, 741–760.
- Fillingim, R. G. and Blumenthal, J. A. (1993). The use of aerobic exercise as a method of stress management. In: Lehrer, P. M. & Woolfolk, R. L. (eds.) *Principles and Practice of Stress Management*, (2nd edn.). New York: Guilford, 443–462.
- Lawlor, D. A. and Hopker, S. W. (2001). The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomised controlled trials. *British Medical Journal* 322, 1–8.
- Leon, A. S. (ed.) (1997). *Physical activity and cardiovascular health*. Champaign, IL: Human Kinetics.

Expression Profiling of Stress Responsive Gene Patterns

N A Datson and M C Morsink

Leiden University, Leiden, The Netherlands

© 2007 Elsevier Inc. All rights reserved.

Introduction

Corticosteroids Modify Gene Expression
Effects of Stress on the Hippocampus
Large-Scale Gene Expression Profiling
Model Systems of Stress
Stress-Responsive Gene Patterns in Brain
Future Perspectives

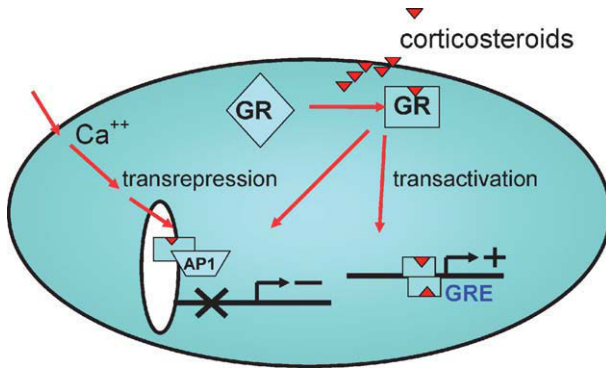


Figure 1 Molecular mechanisms of GR action on gene expression. Transactivation: after binding to corticosteroids, GRs translocate from the cytoplasm to the cell nucleus and bind to glucocorticoid-responsive elements (GREs) on the DNA, thereby influencing the transcription rates of target genes. Transrepression: corticosteroid-activated GRs bind to other transcription factors (such as AP-1), inhibiting their transcriptional actions.

Glossary

<i>Corticosteroids</i>	Steroid hormones secreted at high levels during stress that regulate carbohydrate metabolism and influence the inflammatory response and cognition, mood, and behavior. Also known as glucocorticoids.
<i>Glucocorticoid receptor (GR)</i>	An intracellular receptor that binds corticosteroids and then translocates to the nucleus, where it influences gene expression.
<i>Hippocampus</i>	A complex neural structure (shaped like a sea horse) consisting of gray matter and located on the floor of each lateral ventricle; intimately involved in motivation and emotion as part of the limbic system; has a central role in the formation of memories.
<i>Large-scale gene expression profiling</i>	A general name for molecular biological methods aimed at measuring expression levels of large numbers of genes (>100–30,000) simultaneously. Also known as transcriptomics or (pharmaco)genomics.
<i>Limbic system</i>	A group of interconnected deep brain structures, common to all mammals, and involved in olfaction, emotion, motivation, behavior, and various autonomic functions.
<i>Mineralocorticoid receptor (MR)</i>	An intracellular receptor that binds aldosterone and then translocates to the nucleus, where it influences gene expression. This receptor, which is colocalized with glucocorticoid receptor in several limbic brain structures, also has a high affinity for corticosteroids.

Introduction

The hypothalamic-pituitary-adrenal (HPA) axis is involved in regulation of stress responses in the organism. Under normal conditions, the adrenals secrete basal concentrations of corticosteroids into the bloodstream in a circadian manner. When the organism experiences a stressor, the activity of the HPA axis increases, which results in increased concentrations of corticosteroids. These corticosteroids (CORT) target many organs, exerting effects on blood glucose levels and immune function. Furthermore, corticosteroids are able to pass the blood–brain barrier, affecting brain functions such as cognition, behavior, and mood.

This article focuses on the effects of corticosteroids on gene expression in brain. These stress hormones have a profound effect on brain function. Moreover, exposure to chronic stress can precipitate the onset of central nervous system disorders such as depression.

Corticosteroids Modify Gene Expression

There are two endogenous receptors for CORT that are colocalized in several limbic brain structures such as the hippocampus and amygdala. The mineralocorticoid receptor (MR), which is expressed at high levels in the hippocampus, has a high affinity for its ligand and is already occupied under basal concentrations. The glucocorticoid receptor (GR) is ubiquitously expressed, has a tenfold lower affinity, and becomes occupied during rising concentrations of corticosteroids, for instance, during stress.

Both MR and GR belong to the superfamily of ligand-regulated nuclear receptors and are able to modify gene expression via two different mechanisms: (1) transactivation, in which the receptors bind to glucocorticoid-responsive elements (GREs) on the DNA and influence the transcription rate of the target genes, and (2) transrepression, in which the receptors interact with other transcription factors (such as activator protein-1 [AP-1] and nuclear factor- κ B [NF- κ B]), thereby inhibiting their transcriptional actions (Figure 1).

Effects of Stress on the Hippocampus

One of the major targets in the brain is the hippocampus, a structure that plays a crucial role in learning, memory, emotion, and regulation of the stress system. Several aspects of hippocampal cell function, such as neurotransmission and energy metabolism, are under the tight control of corticosteroids.

MR and GR, which are colocalized in hippocampal neurons, exert differential effects on hippocampal cell function. For instance, under low, basal concentra-

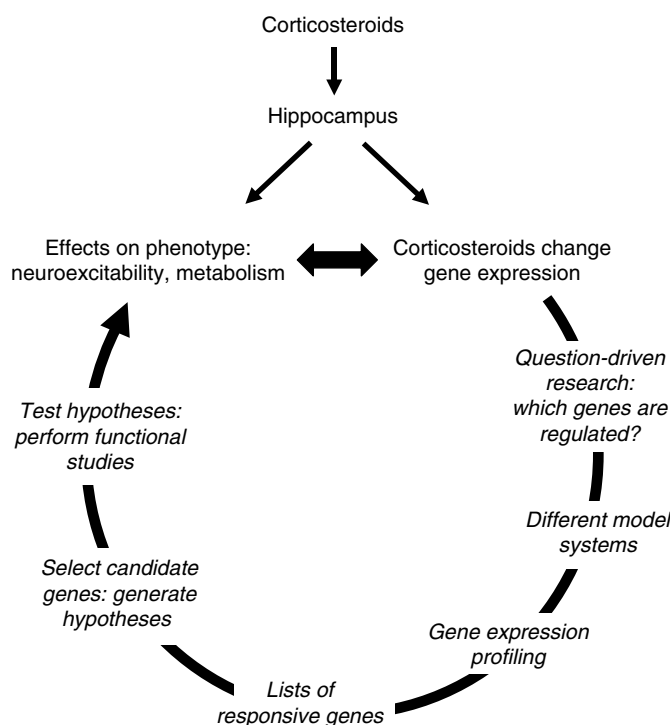


Figure 2 Outline for a large-scale gene expression profiling strategy. In order to elucidate the molecular mechanisms underlying the corticosteroid-mediated effects on the hippocampal phenotype, large-scale gene expression profiling techniques are used to assess corticosteroid-responsive genes. Candidate genes are selected and hypotheses are generated as to how these genes may play a role in the observed effects of corticosteroid on the hippocampus. These hypotheses are then tested in functional studies.

tions in which only MR is occupied, neuroexcitability in hippocampal neurons is high, whereas under stress conditions, in which both MR and GR are occupied, neuroexcitability is low. Thus, MR is involved in maintenance of neuronal excitability and basal activity of the stress system, while GR activation results in reduced hippocampal output and is involved in negative feedback to restore homeostasis. Inadequate corticosteroid input as well as chronic occupation of GR results in degeneration of hippocampal neuronal circuits, accompanied by deficits in cognition and maladaptation to stress. A balanced activation of MR/GR is therefore an important determinant of neuronal excitability, stress responsiveness, and neuronal health.

Since (1) MR and GR are able to modify gene expression and (2) the majority of the effects of corticosteroids on hippocampal cell function develop in a delayed manner, it can be hypothesized that changes in gene expression underlie these CORT-induced effects. Therefore, by determining the CORT-induced changes in gene expression, the molecular mechanisms responsible for the actions of corticosteroids can be investigated. This entails measuring the expression levels of thousands of genes simultaneously, known as large-scale expression profiling, transcriptome analysis, or pharmacogenomics.

Large-Scale Gene Expression Profiling

In [Figure 2](#) the general outline for a large-scale gene expression profiling strategy is depicted. Instead of starting with a hypothesis, this strategy raises the question which genes are regulated by the hormone. Therefore, this type of research can be described as question-driven as opposed to hypothesis-driven research, and its primary goal is to generate gene expression profiles. When these profiles are established, candidate genes can be selected and hypotheses are generated as to how the selected candidate genes may play a role in the observed effects of corticosteroids on the hippocampal phenotype. Finally, these hypotheses are tested in functional studies using *in vitro* systems such as cell lines or *in vivo* systems such as transgenic animals.

Large-scale gene expression profiling can be performed in a number of ways. Two of the more commonly used techniques are serial analysis of gene expression (SAGE) and DNA microarrays.

In SAGE, gene expression profiles are established by sequencing and counting 10-base-pair-long SAGE tags that are derived from a defined position within the 3'-untranslated region of each transcript. By sequencing and counting sufficient transcript-specific tags, a representative expression profile is obtained of

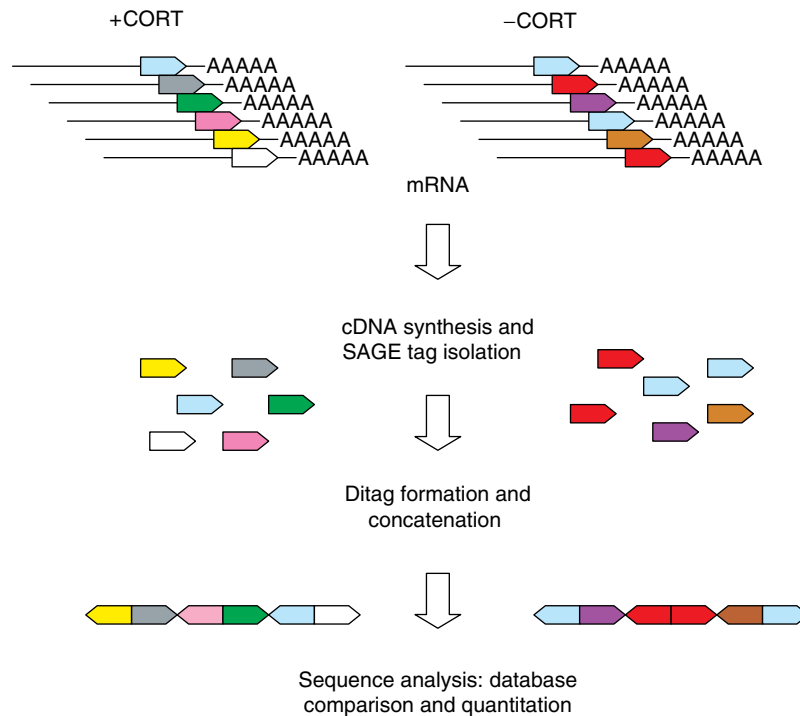


Figure 3 Serial analysis of gene expression (SAGE). In SAGE, 10-bp-long tags are generated from mRNA pools of different experimental treatment groups (for instance, corticosteroid-treated [+CORT] and control [-CORT]) groups). These mRNA-specific tags are ligated to produce concatemers, which are cloned and sequenced. By counting the tag abundancies, mRNA expression levels can be estimated and compared between the different experimental groups.

a tissue under conditions of interest, for example, after hormone treatment (Figure 3).

DNA microarrays are glass microscope slides or silica chips on which a large number of DNA sequences representing a part of each transcript are deposited at a very high density. Depending on the type of microarray, the number of represented transcripts can range from 1000 to more than 20 000. Expression levels of each transcript are measured by hybridizing labeled RNA from a tissue of interest to the microarray, thus generating a hybridization signal (mostly fluorescent) that is a measure for the expression level of the corresponding transcript.

The use of DNA microarrays has been described as following a closed expression profiling strategy, since only the expression of a predefined set of genes (i.e., the sequences that are represented on the microarray) is profiled. In contrast, by using SAGE, no selection of transcripts is made beforehand; thus, this method can be considered an open expression profiling strategy.

Model Systems of Stress

The first step of an expression profiling strategy is the choice of the model system. Several model systems in which corticosteroid levels are pharmacologically manipulated can be used, although one must be aware

that this only represents the hormonal part of the stress response and is not the same as stress. For instance, surgically removing the adrenals from rats results in complete absence of endogenous corticosteroids and lack of hippocampal MR and GR occupancy. Adrenalectomized rats can be further manipulated by (1) replacing them with low concentration corticosteroid-secreting pellets, resulting in occupancy of hippocampal MRs, and (2) administering high-concentration corticosteroid injections in order to additionally occupy hippocampal GRs. After isolating the hippocampi from these animals, gene expression profiles can be generated. Additionally, *ex vivo* hippocampal slices, which are derived from isolated hippocampi and are kept alive in carbogenated buffer, have also been used as a model to study the effects of corticosteroids on gene expression. In these slices, corticosteroid receptors can be directly activated by adding the hormone to the solution, thereby circumventing peripheral effects of injections into animals and removing the influence of projections from other brain areas to the hippocampus.

Another way of looking at the effects of stress on gene expression is to subject intact animals to a stressor, causing physiological activation of their HPA axis and a rise in plasma CORT levels. Several stress paradigms have been described in the literature, both for

acute stress and for chronic stress. For example, placing a rat or mouse in a novel environment is sufficient to elicit a stress response, or placing a rat on top of the cage of a mouse will give rise to a stress response in the mouse. Purely physical stressors, such as cold stress or shaking stress, can be distinguished from psychological stressors such as psychosocial defeat and chronic unpredictable stress, in which experimental animals are subjected to a variety of different stressors in an unpredictable manner over several weeks.

A disadvantage of the physiological approach is that it can be difficult to control for interindividual differences in HPA response. Pharmacological manipulation of CORT levels in combination with a relevant psychological stressor may therefore be a good choice, depending on the focus of the study.

Stress-Responsive Gene Patterns in Brain

Different combinations of model systems and gene expression profiling techniques have been applied to analyze stress or corticosteroid-dependent changes in gene expression in brain. The vast majority of these studies focus on the hippocampus, because the activity of hippocampal neurons is modulated by elevated corticosteroid levels and it is the region with the highest expression of CORT receptors in the brain. Most of the studies have focused on GR-dependent gene expression. So far, not much is known about the genes regulated by MR as well as its molecular mode of action. It has been shown, using SAGE to analyze CORT-dependent gene expression in hippocampus of rats in which either MR or both MR and GR were occupied by pharmacological manipulation of CORT levels, that MR and GR activate or repress distinct yet partially overlapping sets of genes. This suggests that largely different signaling pathways are involved in maintenance of the basal activity of the stress system (MR dependent) and in the stress response (GR dependent).

In a study in tree shrews subjected to chronic psychosocial stress, reduced expression of four genes involved in cell differentiation in response to stress was found. This finding is in agreement with previous findings that this chronic psychosocial stress paradigm in tree shrews leads to structural alterations in hippocampal neurons such as dendritic retraction and impaired neurogenesis.

In a chronic psychosocial stress study performed in mice selected for attack latency (long attack latency [LAL]), hippocampal expression profiles were analyzed using DNA microarrays after 25 days of continuous sensory contact with an aggressive short attack latency (SAL) male. LAL mice but not SAL

mice have been shown to display symptoms of chronic stress, such as increased CORT levels, decreased body weight, and lower hippocampal MR expression, using this paradigm. Interestingly, only downregulated genes were observed in response to this chronic psychosocial stress paradigm. The identified genes encoded a heterogeneous group of different proteins, the majority showing a subtle (less than twofold) downregulation. Several transcripts were identified to be involved in regulation of NF- κ B signaling, suggesting that NF- κ B may be involved in stress-induced functional changes in hippocampal neurons, thereby increasing their vulnerability.

Another study using *ex vivo* hippocampal slices investigated the gene expression response in time following GR activation. This revealed that corticosteroid effects are subtle and transient, with several subsequent waves of gene expression occurring after a single corticosteroid-pulse. Furthermore, within 5 h after activation of GR the gene expression response was comparable to the control group (non-activated GR).

Taking all these studies together, a general conclusion can be drawn that CORT-dependent gene expression encompasses a large variety of functional gene classes involved in energy metabolism, cell adhesion, synaptic transmission, oxidative stress, metabolism, neurogenesis, the cell cycle, and structural plasticity. In addition, the observed responses in hippocampus are subtle in magnitude. Therefore, it is very difficult to conclude which signaling pathways in the brain are most important and most profoundly affected by stress. It appears that corticosteroid stress hormones confer a pleiotropic signal to the neuron, affecting its excitability, connectivity to other neurons, metabolic rate, strength of its synaptic connections, differentiation from newborn neuron into mature neuron, and its vulnerability to metabolic damage. What exactly this means for neuronal function and vulnerability of the individual for stress-related brain disorders remains to be clarified.

The current expression profiling methodology has several limitations. It is still not possible today to perform genome-wide screening for genes regulated by stress due to the fact that annotation of the rat and mouse genome is not fully completed. Furthermore, so far low abundant genes such as neurotransmitter receptors and neurotrophic factors have been underrepresented due to sensitivity issues. Both these factors have also contributed to the currently incomplete picture of the effects of stress and corticosteroid stress hormones on hippocampal gene expression.

Future Perspectives

Since the hippocampus consists of many different cell types, it can be expected that by profiling whole hippocampus, CORT-dependent changes in expression that occur in a specific subpopulation of cells will be diluted to undetectable levels. Therefore, the feasibility of reducing the complexity of the tissue by using laser-microdissected hippocampal subregions has recently been evaluated. A high degree of differential expression between two different hippocampal subregions was identified, indicating that by isolating specific subregions of the hippocampus, detection efficiencies of transcripts with a subregion-specific expression or regulation may be improved. This type of approach combined with use of more genome-wide DNA microarrays and better bioinformatics tools to analyze this type of complex expression data should extend our knowledge on stress-dependent gene patterns in brain.

In conclusion, the application of large-scale expression profiling strategies has revealed a number of interesting candidate genes that may underlie some of the effects of corticosteroids on hippocampal cell function. Additionally, the use of laser-microdissected hippocampal subregions for expression profiling may increase the detection of low abundant genes, resulting in a more refined view on corticosteroid-responsive genes.

See Also the Following Articles

Corticosteroid Receptors; Corticosteroids and Stress; Hippocampus, Corticosteroid Effects on.

Further Reading

Alfonso, J., Pollevick, G. D., van der Hart, M., Flügge, G., Fuchs, E. and Frasch, A. C. C. (2004). Identification of

genes regulated by chronic psychosocial stress and antidepressant treatment in the hippocampus. *European Journal of Neuroscience* 19, 659–666.

Datson, N. A., van der Perk, J., de Kloet, E. R. and Vreugdenhil, E. (2001). Identification of corticosteroid-responsive genes in rat hippocampus using serial analysis of gene expression. *European Journal of Neuroscience* 14, 675–689.

Datson, N. A., Meijer, L., Steenbergen, P. J., et al. (2004). Expression profiling in laser microdissected hippocampal subregions in rat brain reveals large subregion-specific differences in expression. *European Journal of Neuroscience* 20, 2541–2554.

Feldker, D. M., Morsink, M. C., Veenema, A. H., et al. (2006). The effect of chronic exposure to highly aggressive mice on hippocampal gene expression of non-aggressive subordinates. *Brain Research* 1089, 10–20.

Joëls, M. (2001). Corticosteroid actions in the hippocampus. *Journal of Neuroendocrinology* 13, 657–669.

Karst, H., Karten, Y. J., Reichardt, H. M., de Kloet, E. R., Schutz, G. and Joëls, M. (2000). Corticosteroid actions in hippocampus require DNA binding of glucocorticoid receptor homodimers. *Nature Neuroscience* 3, 977–978.

de Kloet, E. R., Joels, M. and Holsboer, F. (2005). Stress and the brain: from adaptation to disease. *Nature Reviews Neuroscience* 6, 463–475.

McEwen, B. S. (1999). Stress and hippocampal plasticity. *Annual Reviews of Neuroscience* 22, 105–122.

Morsink, M. C., Steenbergen, P. J., Vos, J. B., et al. (2006). Acute activation of hippocampal glucocorticoid receptors results in different waves of gene expression throughout time. *Journal of Neuroendocrinology* 18, 239–252.

Oitzl, M. S. and de Kloet, E. R. (1992). Selective corticosteroid antagonists modulate specific aspects of spatial orientation learning. *Behavioral Neuroscience* 106, 62–71.

Vreugdenhil, E., de Kloet, E. R., Schaaf, M. and Datson, N. A. (2001). Genetic dissection of corticosterone receptor function in the rat hippocampus. *European Neuropsychopharmacology* 11, 423–430.

VOLUME ONE

ENCYCLOPEDIA OF STRESS

SECOND EDITION

EDITOR-IN-CHIEF
GEORGE FINK



DEDICATION

For Ann Elizabeth

EDITOR-IN-CHIEF

George Fink

Professorial Research Fellow (formerly Director)
Mental Health Research Institute of Victoria
Parkville, Melbourne, Victoria
Australia
Formerly Director, MRC Brain Metabolism Unit
Edinburgh, Scotland, UK.

ASSOCIATE EDITORS

Bruce McEwen

The Rockefeller University
Laboratory of Neuroendocrinology
1230 York Avenue, Box 165
New York NY 10021

E. Ronald de Kloet

Universiteit Leiden
Natuurwetenschappen Wiskunde en
LACDR Medical Pharmacology
Einsteinweg 55, kamer HB 902
Leiden 2333 CC
The Netherlands

Robert Rubin

Department of Psychiatry (116A)
VA Greater LA Healthcare System
11301 Wilshire Blvd.
Los Angeles, CA 90073

George Chrousos

First Department of Pediatrics,
Athens University Medical School,
Aghia Sophia Children's Hospital,
115 27 Athens, Greece

Andrew Steptoe

University College London
Department of Epidemiology and Public Health
Psychology
1-19 Torrington Place London WC1E 6BT
UK

Noel Rose

Johns Hopkins University
Center for Autoimmune Disease Research
Molecular Microbiology and Immunology
615 North Wolfe Street, Rm E5014
Baltimore MD 21205

Ian Craig

SGDP Research Centre
Institute of Psychiatry, Molecular Genetics Group
De Crespigny Park, Box P082
Denmark Hill
London SE5 8AF
UK

Giora Feuerstein

Merck Research Laboratories 42-209
Department of Cardiovascular Diseases
770 Sumneytown Pike
West Point PA 19486

CONTENTS

<i>Contents by Subject area</i>	<i>xxiii</i>
<i>Preface to First Edition</i>	<i>xxxi</i>
<i>Preface to the Second Edition</i>	<i>xxxiii</i>
<i>Guide to Encyclopedia</i>	<i>xxxv</i>
<i>Foreword</i>	<i>xxxvii</i>

VOLUME 1

A

Acute Stress Disorder and Posttraumatic Stress Disorder	1
R. Yehuda and C. M. Wong	
Acute Stress Response: Experimental	7
K. Pacak and R. McCarty	
Acute Trauma Response	15
W. C. Chiu, D. E. Carlson and M. P. Lilly	
Adenylyl Cyclases and Stress Response	21
F. A. Antoni	
Adjustment Disorders	24
M. Dascalu and D. Svrakic	
Adolescence	28
G. N. Swanson	
Adolescent Suicide	36
M. Berk, R. Suddath and M. Devich-Navarro	
Adrenal Cortex	38
G. P. Vinson, B. J. Whitehouse and J. P. Hinson	
Adrenal Insufficiency	47
H. S. Willenberg, S. R. Bornstein and G. P. Chrousos	

Adrenal Medulla	52
R. Kvetnansky and R. McCarty	
Adrenaline	60
T. M. Pollard	
Adrenocortical Function, Factors Controlling Development Thereof	64
T. Else and G. D. Hammer	
Adrenocorticotrophic Hormone (ACTH)	69
M. E. Rhodes	
Aerobic Exercise and Stress Reduction	73
E. J. C. de Geus and J. H. Stubbe	
Affective Disorders	78
D. F. MacKinnon	
Aggression	84
E. F. Coccaro and E. C. Manning	
Aggressive Behavior	89
J. M. Koolhaas	
Aging and Adrenocortical Factors	92
C. Lord and J. C. Pruessner	
Aging and Psychological Stress	96
B. W. J. Penninx	
Aging and Stress, Biology of	102
M. A. Horan, R. N. Barton and G. J. Lithgow	
AIDS	108
M. H. Antoni and D. G. Cruess	
Airline Accidents	114
G. Li	
Alarm Phase and General Adaptation Syndrome	119
R. McCarty and K. Pacak	

Bereavement	317	Caregivers, Stress and	416
P. J. Clayton		S. H. Zarit, K. Bottigi and J. E. Gaugler	
Beta-Adrenergic Blockers	323	Catecholamines	419
M. B. Hamner and G. W. Arana		U. Lundberg	
Beta-Endorphin	332	Central Stress Neurocircuits	424
M. Lee and S. L. Wardlaw		S. Kollack-Walker, H. E. W. Day and H. Akil	
Blood Pressure	335	Cerebral Metabolism, Brain Imaging	432
A. Sherwood and R. A. Carels		K. P. Ebmeier, C. L. Donaghey and N. J. Dougall	
Blood-Brain Barrier, Stress and	342	Chaperone Proteins and	
S. N. Malaeb and B. S. Stonestreet		Chaperonopathies	438
Borderline Personality Disorder	348	A. J. L. Macario and E. Conway de Macario	
H. W. Koenigsberg and L. J. Siever		Chaperonopathies	444
Brain and Brain Regions	351	A. J. L. Macario and E. Conway de Macario	
A. G. Watts		Chemical Warfare	449
Brain Natriuretic Peptide (BNP)	357	J. Berberich	
P. Pervanidou and G. P. Chrousos		Chernobyl, Stress Effects of	452
Brain Trauma	360	A. Tønnessen and L. Weisæth	
B. Pentland		Child Abuse	457
Breast Cancer	364	C. C. Swenson and L. Saldana	
A. Moyer		Child Physical Abuse	460
Burnout	368	C. C. Swenson	
C. Maslach and M. P. Leiter		Child Sexual Abuse	463
		J. A. Cohen	
		Childbirth and Stress	467
		S. Ayers and E. Ford	
		Childhood Stress	472
		S. Sandberg	
		Cholesterol and Lipoproteins	478
		C. M. Stoney	
		Chronic Fatigue Syndrome	484
		A. J. Cleare and S. Wessely	
		Chronic Social Stress: GR Sensitivity in	
		Leukocytes	493
		A. Weizman and B. Rotberg	
		Circadian Clock Genes as Modulators of	
		Sensitivity to Genotoxic Stress	496
		M. P. Antoch and R. V. Kondratov	
		Circadian Rhythm Effects on	
		Cardiovascular and Other Stress-Related	
		Events	500
		R. Manfredini, B. Boari, R. Salmi, A. M. Malagoni and F. Manfredini	
Cardiovascular Disease, Stress and	410		
G. P. Chrousos and G. Kaltsas			

C

Calbindin	373
F. A. Antoni	
Calcium, Role of	374
F. A. Antoni	
Calcium-Dependent Neurotoxicity	375
J. S. Kelly	
Cancer	378
D. Spiegel	
Cancer Treatment	384
F. I. Fawzy, A. L. Canada and N. W. Fawzy	
Captivity, Adaptation to	388
R. H. Rahe	
Captivity, Recovery from	392
R. H. Rahe	
Cardiovascular System and Stress	396
P. Hjelm Dahl	
Cardiovascular Disease, Stress and	410
G. P. Chrousos and G. Kaltsas	

Circadian Rhythms, Effects of Prenatal Stress in Rodents	505	Corticosteroids and Stress	613
S. Maccari and O. Van Reeth		A. Munck	
Circadian Rhythms, Genetics of	508	Corticotropin Releasing Hormone (CRH)	620
F. W. Turek and M. H. Vitaterna		A. T. Lim	
Cognition and Stress	513	Corticotropin Releasing Factor Receptor Deficiency in Mice	623
M. W. Eysenck		N. J. Justice and K.-F. Lee	
Cognitive-Behavioral Therapy	515	Corticotropin Releasing Factor-Binding Protein	627
G. A. Fava		P. J. Lowry, C. F. Kemp and R. J. Woods	
Combat Reaction, Chronic	518	Corticotropin-Releasing Factor Circuitry in the Brain – Relevance for Affective Disorders and Anxiety	630
R. H. Rahe		D. A. Gutman and C. B. Nemeroff	
Combat Stress Reaction	524	Corticotropin-Releasing Factor (CRF) Family of Neuropeptides – Role in Inflammation	635
M. Dobson		A. Gravanis and A. N. Margioris	
Combat, Acute Reactions to	529	Corticotropin-Releasing Factor Receptors	641
R. H. Rahe		D. E. Grigoriadis	
Common Cold and Stress	533	Cortisol Awakening Response	649
A. Smith		A. Steptoe	
Community Studies	536	C-Reactive Protein	653
C. J. Holahan, R. H. Moos and L. M. Groesz		W. J. Kop and A. A. Weinstein	
Comorbid Disorders and Stress	542	Crime Victims	659
S. Sundram and Avril Pereira		I. Robbins	
Comparative Anatomy and Physiology	549	Crisis Intervention	662
A. G. Watts		D. Hamaoka, D. Benedek, T. Grieger and R. J. Ursano	
Concentration Camp Survivors	553	Critical Thermal Limits	667
J. D. Kinzie		J. Roth	
Congenital Adrenal Hyperplasia (CAH)	556	Crowding Stress	669
A. Solomon and P.-M. G. Bouloux		L. Kovács and P. Csermely	
Conservation of Resources Theory	562	Cultural Factors in Stress	672
S. E. Hobfoll and J. S. Ford		J. W. Berry and B. Ataca	
Control and Stress	568	Cultural Transition	678
A. Steptoe		M. S. Kopp	
Coping and Stress: A Lens and Filter Model	574	Cushing's Syndrome, Medical Aspects	682
R. H. Rahe		S. R. Bornstein, M. Gruber, H. S. Willenberg, C. A. Stratakis and G. P. Chrousos	
Coping Skills	578	Cushing's Syndrome, Neuropsychiatric Aspects	688
A. DeLongis and E. Puterman		M. N. Starkman	
Corticosteroid Receptor Genes: Functional Dissection in Mice	584	Cytokines	692
F. Tronche		G. D. Marshall	
Corticosteroid Receptors	594		
O. C. Meijer, E. R. de Kloet and B. S. McEwen			
Corticosteroid-Binding Globulin (Transcortin)	605		
B. E. P. Murphy			

Cytokines, Chronic Stress, and Fatigue S. Jain and P. J. Mills	698	Diet and Stress, Non-Psychiatric J. Wardle and E. L. Gibson	797
Cytokines, Stress, and Depression B. E. Leonard and C. Song	705	Diet and Stress, Psychiatric V. March and M. H. Fernstrom	806
Cytotoxic Lymphocytes M. A. Fletcher and N. G. Klimas	711	Disaster Syndrome P. Valent	811
D		Disasters and Mass Violence, Public, Effects of G. Stevens, B. Raphael and M. Dobson	814
Death Anxiety R. Kastenbaum	717	Disease, Stress Induced H. S. Willenberg, S. R. Bornstein and G. P. Chrousos	824
Defensive Behaviors D. C. Blanchard, M. Yang, M. Hebert and R. J. Blanchard	722	Dissociation J. R. Maldonado	828
Demand–Control Model T. Theorell	727	Distress G. Matthews	838
Dental Stress T. K. Fábán, G. Fábán and P. Fejérdy	733	Divorce, Children of K. N. Hipke, S. A. Wolchik and I. N. Sandler	844
Depersonalization: Systematic Assessment M. Steinberg	736	Domestic Violence B. Donohue, H. Hill and T. Maier-Paarlberg	848
Depression and Coronary Heart Disease F. Lespérance and N. Frasure-Smith	741	Dopamine, Central G. D. Stanwood	852
Depression and Manic–Depressive Illness R. T. Rubin and B. J. Carroll	744	<i>Drosophila</i> Genes and Anoxia G. G. Haddad	859
Depression and Stress, Role of n-3 and n-6 Fatty Acids C. Song and B. E. Leonard	754	<i>Drosophila</i> Studies M. Allikian and J. Tower	864
Depression Models K. Matthews and C. Stewart	760	Drug Use and Abuse J. R. Mantsch	866
Depression, Immunological Aspects M. R. Irwin	766	E	
Dermatological Conditions M. A. Gupta	773	Earthquakes, Stress Effects of M. Livanou and M. Başoğlu	871
Desensitization F. A. Antoni	778	Eating Disorders and Stress D. C. Jimerson	876
Dexamethasone Suppression Test (DST) R. T. Rubin and B. J. Carroll	780	Eclampsia and Pre-Eclampsia A. Makrigiannakis, G. Petsas and G. P. Chrousos	880
DEX-CRH Test N. C. Schommer and I. Heuser	784	Economic Factors and Stress R. A. Catalano	884
DHEA J. Herbert	788	Education Levels and Stress J. Mirowsky and C. E. Ross	888
Diabetes, Type 1 A. Riazi and C. Bradley	792		

Food Shift Effect F. K. Stephan	88	Glucocorticoid Receptor Mutations and Polymorphisms J. W. Koper	183
Freud, Sigmund D. J. Lynn	90	Glucocorticoids – Adverse Effects on the Nervous System R. M. Sapolsky	185
G		Glucocorticoids, Effects of Stress on J. C. Buckingham	190
GABA (Gamma Aminobutyric Acid) J. D. C. Lambert	97	Glucocorticoids, Overview B. E. Pearson Murphy	198
Gastrointestinal Effects R. Murison and A. M. Milde	109	Glucocorticoids, Role in Stress J. C. Buckingham	210
Gender and Stress R. J. Handa and W. C. J. Chung	115	Glucose Transport A. L. McCall	217
Gene–Environment Interactions in Early Development I. S. P. Davis and R. Plomin	122	Glycobiology of Stress G. Lauc and M. Flögel	222
Genetic Factors and Stress C. A. Koch and C. A. Stratakis	128	Gonadotropin Secretion, Effects of Stress on M. Ferin	228
Genetic Polymorphisms in Stress Response I. W. Craig	135	Graves' Disease (Thyrotoxicosis) W. M. Wiersinga	234
Genetic Predispositions to Stressful Conditions D. Blackwood and H. Knight	141	Grieving A. Ray and H. Prigerson	238
Genetic Testing and Stress V. Senior and M. Cropley	146	Group Therapy N. Wong	242
Genetic Variation of HPA Axis Activity and Function in Farm Animals P. Mormède	150	Gulf War Syndrome, Psychological and Chemical Stressors H. Soreq	248
H			
Ghrelin and Stress Protection T. Brzozowski, M. Pawlik, D. Drozdowicz, Z. Sliwowski, S. J. Konturek, W. W. Pawlik and P. C. Konturek	153	Health and Socioeconomic Status T. Chandola and M. Marmot	255
Glia or Neuroglia G. W. Bennett and D. E. Ray	161	Health Behavior and Stress A. Steptoe	262
Glucocorticoid Effects on Memory: the Positive and the Negative O. T. Wolf	166	Heart Disease/Attack G. J. Baker, S. Suchday and D. S. Krantz	266
Glucocorticoid Negative Feedback M. F. Dallman	172	Heart Failure, Stress Effects M. Alevizaki and G. P. Chrousos	272
Glucocorticoid Receptor Mutant Mice as Models for Stress-Induced Affective Disorders P. Gass	176	Heart Rate J. R. Jennings	274
		Heat Resistance A. J. L. Macario and E. Conway de Macario	278

Heat Shock Genes, Human	284	Hyperthermia	381
A. J. L. Macario and E. Conway de Macario		J. Roth	
Heat Shock Proteins: HSP60 Family Genes	288	Hyperthyroidism	388
H. Kubota		I. M. Lesser and D. L. Flores	
Heat Shock Response, Overview	292	Hyperventilation	390
A. J. L. Macario and E. Conway de Macario		C. Bass	
Hemostasis and Stress	300	Hypnosis	394
Roland von Känel		W. G. Whitehouse, E. C. Orne and M. T. Orne	
Herpesviruses	305	Hypocortisolism and Stress	400
D. A. Padgett, M. T. Bailey and J. F. Sheridan		C. M. Heim and U. M. Nater	
Hippocampal Neurons	311	Hypoglycemia	408
R. L. Spencer and S. T. Bland		B. M. Frier	
Hippocampus, Corticosteroid Effects on	321	Hypotension, Hypovolemia, and Septic Shock	413
M. Joëls and H. Karst		A. Beishuizen, A. B. Johan Groeneveld and I. Vermes	
Hippocampus, Overview	327	Hypothalamic-Pituitary-Adrenal Axis	421
M. J. Meaney and S. J. Lupien		M. F. Dallman, S. Bhatnagar and V. Viau	
Hiroshima Bombing, Stress Effects of	332	Hypothermia	428
R. J. Lifton		M. J. Taylor	
HIV Infection/AIDS	336	Hypothyroidism	439
B. W. Dixon		R. T. Joffe	
Holocaust Survivors, Experiences of	339	Hysteria	442
P. Valent		K. Pajer	
Holocaust, Stress Effects of	342		
P. Valent			
Homeostasis	347		
B. S. McEwen			
Homosexuality, Stress and	348	Immobilization Stress	445
J. Drescher		R. Kvetnansky and R. McCarty	
Hostility	354	Immune Cell Distribution, Effects of Stress on	449
L. H. Powell and K. Williams		F. S. Dhabhar	
HPA Alterations in PTSD	359	Immune Function, Stress-Induced Enhancement	455
R. Yehuda		F. S. Dhabhar	
Hurricane Katrina Disaster, Stress Effects of	364	Immune Response	462
C. Piotrowski		P. J. Delves	
11β-Hydroxysteroid Dehydrogenases	368	Immune Suppression	470
J. R. Seckl		P. Prolo and F. Chiappelli	
Hyperreactivity (Cardiovascular)	372	Immune Surveillance – Cancer, Effects of Stress on	477
A. Georgiades		D. Spiegel and F. S. Dhabhar	
Hypertension	376	Immune System, Aging	481
A. Steptoe		M. A. Horan	

Immunity	485	L	
F. Chiappelli			
Impact of Terrorism on the Development of Mental Health Symptoms	493	Learned Helplessness	567
R. Yehuda		D. M. Isaacowitz and M. E. P. Seligman	
Impotence, Stress and	497	Learning and Memory, Effects of Stress on	571
M. R. Gignac, G. M. Rooker and J. K. Cohen		M. Lindau, O. Almkvist and A. H. Mohammed	
Impulse Control	500	Left Ventricular Mass	577
I.-M. Blackburn		T. G. Pickering	
Incest	502	<i>Leishmania</i>, Stress Response in	579
A. P. Mannarino		M. Shapira	
Income Levels and Stress	506	Leptin, Adiponectin, Resistin, Ghrelin	584
S. V. Subramanian and I. Kawachi		A. Gavrilu, D. Barb and C. S. Mantzoros	
Indigenous Societies	511	Leukocyte Trafficking and Stress	592
W. W. Dressler		S. Hong, M. U. Goebel and P. J. Mills	
Industrialized Societies	517	Life Events and Health	599
J. Siegrist		P. Surtees and N. Wainwright	
Infection	521	Life Events Scale	603
H. Friedman and S. H. Pross		E. Wethington	
Inflammation	530	Lockerbie Air Crash, Stress Effects of	608
G. Z. Feuerstein, R. R. Ruffolo, C. Coughlin, J. Wang and D. Miller		H. Livingston and M. Livingston	
Instinct Theory	535	Loss Trauma	612
R. Gardner Jr.		A. Bifulco	
Integrative Medicine (Complementary and Alternative Medicine)	537	Lymph Nodes	616
D. Spiegel		T. L. Whiteside	
Interactions Between Stress and Drugs of Abuse	540	Lymphocytes	622
P. V. Piazza and M. Le Moal		N. R. Rose	
J		M	
Job Insecurity: The Health Effects of a Psychosocial Work Stressor	549	Macrophage Antimycobacterial Activity, Effects of Stress on	629
J. E. Ferrie and P. Martikainen		C. S. Boomershine and B. S. Zwilling	
K		Macrophages	634
Kidney Function	557	W. P. Fehder, F. Tuluc, W.-Z. Ho and S. D. Douglas	
W. J. Welch		Major Depressive Disorder	640
Korean Conflict, Stress Effects of	563	A. B. Negrão and P. W. Gold	
C. A. Goguen and M. J. Friedman		Male Partner Violence	645
		M. Ingram, N. P. Yuan and M. P. Koss	
		Marital Conflict	651
		P. T. McFarland and A. Christensen	

Marital Status and Health Problems	653	Mitochondria	754
I. M. A. Joung		I. Manoli, S. Alesci and G. P. Chrousos	
Marriage	660	Monoamine Oxidase	761
A. C. Yoneda and J. Davila		P. Huezo-Diaz and I. W. Craig	
Maternal Deprivation	667	Motor Vehicle Accidents, Stress Effects of	764
R. L. Huot, C. O. Ladd and P. M. Plotsky		T. C. Buckley and E. B. Blanchard	
Medical Profession and Stress	674	Mucosal Secretory Immunity, Stress and	768
K. G. Power and V. Swanson		J. A. Bosch and D. Carroll	
Meditation and Stress	678	Multi Drug Resistance P Glycoprotein and other Transporters	774
J. L. Kristeller		E. C. M. de Lange	
Membrane Glucocorticoid Receptors	686	Multiple Personality Disorder	783
P. J. Gasser and M. Orchinik		R. P. Kluft	
Memory and Stress	693	Multiple Sclerosis	790
S. J. Lupien and F. S. Maheu		A. T. Reder	
Memory Impairment	699	Multiple Trauma	795
A. J. Parkin [†]		P. N. Soucacos and E. O. Johnson	
Menopause and Stress	703	Musculoskeletal Problems and Stress	800
N. E. Avis		S. Svebak	
Menstrual Cycles and Stress	706	Myopathy	807
R. Suri and L. Altshuler		G. A. Small	
Mental Stress Testing	712		
P. G. Saab, K. A. Kline and J. R. McCalla			
Metabolic Syndrome	717	Natural Killer (NK) Cells	815
L. Keltikangas-Järvinen		T. L. Whiteside, M. Boyiadzis and R. B. Herberman	
Metabolic Syndrome and Stress	721	Negative Affect	822
R. Rosmond		A. A. Stone and A. A. Gorin	
Metals, Oxidative Stress, and Brain Biology	724	Neighborhood Stress and Health	825
D. I. Finkelstein, T. Lynch, S. Wilkins, R. A. Cherny and A. I. Bush		A. V. Diez Roux	
Metastasis	726	Nelson's Syndrome	828
M. K. Demetrikopoulos		A. Stathopoulou, K. Dimitriou and G. Kaltsas	
Metyrapone: Basic and Clinical Studies	730	Neural Stem Cells	832
E. A. Young		U. S. Sohur, J. G. Emsley, B. D. Mitchell and J. D. Macklis	
Migraine	733	Neurodegenerative Disorders	840
N. M. Ramadan		M. F. Mendez and A. M. McMurtry	
Mineralocorticoid Receptor Polymorphisms	744	Neurodevelopmental Disorders in Children	844
R. H. DeRijk and E. R. de Kloet		N. J. Rinehart and B. J. Tonge	
Minorities and Stress	748	Neuroendocrine Systems	851
I. Mino, W. E. Profit and C. M. Pierce		G. Fink	
		Neurogenesis	865
		P. Tanapat and E. Gould	

[†]Deceased.

Pheromones	119	Prison	217
A. Kumar, C. A. Dudley, S. Chakravarty and R. L. Moss		D. L. Whitehead and A. Steptoe	
Pituitary Regulation, Role of	127	Prisoners of War	223
F. A. Antoni		C. Tennant	
Police, Stress in	131	Problem-Solving Skills Training	227
R. J. Burke		A. M. Nezu, C. M. Nezu and T. J. D'Zurilla	
Posttraumatic Stress Disorder – Clinical	135	Prolactin and Stress	231
N. C. Feeny, L. R. Stines and E. B. Foa		G. Tolis, G. Rombopoulos, D. Kaltsas, E. Katounda, V. Kaltzidou and N. Angelopoulos	
Posttraumatic Stress Disorder – Neurobiological basis for	140	Pro-opiomelanocortin (POMC)	233
M. Barad		A. B. Bicknell	
Posttraumatic Stress Disorder in Children	145	Prostaglandins	240
A. M. La Greca		S. Moshonov, U. Zor and Z. Naor	
Posttraumatic Stress Disorder, Delayed	150	Proteases in Prokaryotes and Eukaryotic Cell Organelles	247
A. Holen		E. Conway de Macario and A. J. L. Macario	
Posttraumatic Stress Disorder, Neurobiology of	152	Proteases in the Eukaryotic Cell Cytosol	252
J. D. Bremner		E. Conway de Macario and A. J. L. Macario	
Posttraumatic Therapy	157	Protein Synthesis	258
A. M. Rasmusson, C. M. Monson and P. A. Resick		M. A. Brostrom and C. O. Brostrom	
Pregnancy – Maternal and Perinatal Stress – Effects of	165	Proteosome	266
P. Smirnaki and M.-A. Magiakou		M. Maldonado and J. Wang	
Premenstrual Dysphoric Disorder	173	Psoriasis	271
S. Nowakowski, P. Haynes and B. L. Parry		A. B-. Kirschbaum	
Pre-pulse Inhibition	180	Psychoanalysis	274
M. van den Buuse		P. Roazen	
Pressure, Effects of Extreme High and Low	184	Psychological Stressors, Overview	278
R. J. Værnes		S. M. Monroe and G. M. Slavich	
Primate Hierarchies and Personality	194	Psychoneuroimmunology	284
R. M. Sapolsky		R. Dantzer	
Primate Models, Behavioral– Immunological Interactions	199	Psychosocial Factors and Stress	288
M. L. Laudenslager and S. Tiefenbacher		J. Siegrist	
Primate Models, Cardiovascular Disease	204	Psychosomatic Heart Disease: Role of Sympathetic and Sympathoadrenal Processes	292
J. R. Kaplan		G. W. Lambert and T. Dawood	
Primate Models, Overview	211	Psychosomatic Medicine	296
D. M. Lyons		T. Theorell	
Primates: Rearing and Effects of Stress on Primate CNS Function	214	Psychotherapy	302
T. K. Newman and C. S. Barr		G. C. Smith	

Psychotic Disorders J. Ventura	308	Revenge Fantasies M. Horowitz and S. Meffert	391
Q		Rheumatic Disorders A. T. Masi and J. C. Aldag	395
Quality of Life S. M. Skevington	317	S	
R		Salivary Cortisol C. Kirschbaum and D. H. Hellhammer	405
Racial Harassment/Discrimination D. R. Williams and S. A. Mohammed	321	Salt Appetite R. S. Weisinger and N. Chen	409
Recovery from Stress S. Sonnentag and C. Fritz	327	Schizophrenia M. van den Buuse and D. Copolov	418
Reductive Stress J. P. Kehrer	331	School Stress and School Refusal Behavior C. A. Kearney, L. C. Cook and G. Chapman	422
Reenactment Techniques N. Wong	336	School Violence and Bullying J. Juvonen	425
Refugees, Stress in J. D. Kinzie	338	Seasonal Changes in Stress Responses R. J. Nelson and L. B. Martin II	427
Regional Blood Flow, Stress Effects W. W. Blessing	341	Seasonal Rhythms L. M. Romero	432
Relaxation Techniques W. G. Whitehouse, E. C. Orne and M. T. Orne	345	Secretagogue G. Fink	435
Religion and Stress S. Packer	351	Selective Serotonin Reuptake Inhibitors (SSRIs) M. Gitlin	440
9/11, Religion and Stress S. Packer	357	Self-Esteem, Stress, and Emotion D. Roger	443
Remodelling of Neuronal Networks by Stress E. Fuchs and G. Flügge	364	Selye, Hans B. Tuchweber and P. Bois	448
Renal and Adrenocortical Actions of Dopamine B. C. Williams, Y.-C. Lo and S. W. Walker	371	Sepsis, Acute Respiratory Distress Syndrome, and Glucocorticoid Resistance G. U. Meduri	450
Reproduction, Effects of Social Stress On C. A. Shively	374	Serotonin B. C. Williams and Y.-C. Lo	457
Reproductive Dysfunction in Primates, Behaviorally Induced J. L. Cameron	380	Serotonin in Stress S. Kusljic and M. van den Buuse	461
Resistance L. M. Zabarenko	386	Serotonin Transporter Genetic Modifications K. Sugden	465
Restraint Stress R. J. Servatius, G. Salameh, K. M. Coyle and W. P. Paré	389	Sex Differences in Human Stress Response B. M. Kudielka, D. H. Hellhammer and C. Kirschbaum	469

Sex Steroids, Response to Stress and Susceptibility to Depression	474	Startle Response	561
M. V. Seeman and L. E. Ross		C. O. Ladd, P. M. Plotsky and M. Davis	
Sex-Specific Effects of Early Social Stress in Mammals: A Study in Guinea Pigs	479	Steroid Hormone Receptors	568
S. Kaiser and N. Sachser		M. Beato and J. Klug	
Sexual Assault	484	Steroid Hydroxylases	581
N. C. Feeny, T. J. Linares and E. B. Foa		F. H. de Jong	
Sexual Dysfunction	490	Strain Differences in Stress Response in Rodents	585
W. T. O'Donohue and L. Woodward Tolle		P. Mormède	
Sexual Offenders	494	Stress and Anxiety: Treatment with 2nd Generation Antipsychotic Drugs	587
F. M. Saleh and H. M. Malin		J. S. Ballon, J. A. Boyd and D. A. Wirshing	
Sickle Cell Disease and Stress	502	Stress and CNS Arousal: Genomic Contributions	591
K. Midence		A. C. Ribeiro and D. W. Pfaff	
Sleep Loss, Jet Lag, and Shift Work	504	Stress Effect of Assisted Reproduction	596
E. Van Cauter		F. M. Helmerhorst, R. Sibug and E. R. de Kloet	
Sleep, Sleep Disorders, and Stress	506	Stress Effects, Overview	599
A. N. Vgontzas, S. Pejovic and M. Karataraki		A. Steptoe	
Smoking and Stress	515	Stress Generation	601
F. J. McClernon and D. G. Gilbert		J. E. Roberts and J. A. Ciesla	
Social Capital	520	Stress Hyporesponsive Period	606
I. Kawachi		S. Levine, E. R. de Kloet, G. Dent and M. S. Schmidt	
Social Networks and Social Isolation	523	Stress in University Students	612
L. F. Berkman		B. Andrews and J. Hejdenberg	
Social Status and Stress	528	Stress Induced Anovulation	615
D. de Ridder		S. L. Berga and T. L. Loucks	
Social Stress, Animal Models of	533	Stress Management and Cardiovascular Disease	631
J. Bugajski, A. Gądek-Michalska and A. J. Bugajski		D. Lane and D. Carroll	
Social Support	539	Stress Management, CAM Approach	636
T. C. Antonucci, J. E. Lansford and K. J. Ajrouch		K. Krebs	
Social Support in Trauma	542	Stress of Self Esteem	640
K. O'Donnell and A. Steptoe		J. C. Pruessner, S. Wuethrich and M. W. Baldwin	
Somatic Disorders	545	Stress System Balance Hypothesis	646
F. Creed		E. R. de Kloet	
Space, Health Risks of	548	Stress, Beneficial Effects of	650
V. A. Convertino		S. Joseph and P. A. Linley	
Spinal Cord Injury, Physical Stress of	554	Stress, Definitions and Concepts of	653
S. Bhatia and M. Quigley		B. S. McEwen	
Spinal Cord Injury, Psychological Stress of	557		
S. D. Schnakenberg-Ott and M. R. Lovell			

Stress, Insulin Resistance, and Type II Diabetes	654	Torture	749
C. Tsigos and I. Kyrou		I. Genefke, H. Marcussen and O. V. Rasmussen	
Stress, NPY, and Cardiovascular Diseases	660	Transport-Related Stress	756
Z. Zukowska		R. G. Smart	
Suicide Terrorism, Genesis of	667	Trans-sexualism	760
A. Speckhard		R. A. Allison	
Suicide, Biology of	677	Trauma and Memory	765
M. A. Oquendo, L. Giner and J. J. Mann		B. A. van der Kolk	
Suicide, Psychology of	684	Trauma Group Therapy	767
D. Lester and R. L. Walker		J. Dwyer and L. G. Martin	
Suicide, Sociology of	689	Traumatic Stress and Posttraumatic Stress Disorder, the Israeli Experience	771
D. Lester and R. M. Fernquist		E. Klein and J. Zohar	
Surgery and Stress	693	Trier Social Stress Test	776
C. Vögele		B. M. Kudielka, S. Wüst, C. Kirschbaum and D. H. Hellhammer	
Survivor Guilt	695	Type A Personality, Type B Personality	782
P. Valent		W. S. Shaw and J. E. Dimsdale	
Sympathetic Nervous System	697		
D. S. Goldstein			
Synthetic Glucocorticoids	704		
A. M. Karssen and E. R. de Kloet			
Systemic Lupus Erythematosus	708		
D. J. Wallace			
		U	
		Ulceration, Gastric	787
		R. Murison and A. M. Milde	
		Ultradian Rhythms	791
		S. L. Lightman	
		Understimulation/Boredom	794
		V. J. Sutherland	
		Unemployment, Stress, and Health	797
		M. Bartley	
		Urocortins	804
		É. M. Fekete and E. P. Zorrilla	
		V	
Teaching and Stress	713	Vaccination	813
E. R. Greenglass		V. E. Burns, A. C. Phillips and K. M. Edwards	
Temperature Effects	717	Vasoactive Peptides	817
J. Roth		W. K. Samson and M. M. White	
Terrorism	719	Vasopressin	824
P. Bell		L. P. Renaud	
Thermal Stress	723	Vietnam Veterans, Postwar Experiences and Health Outcomes	830
C. M. Blatteis		J. A. Boscarino	
Thermotolerance, Thermoresistance, and Thermosensitivity	726		
S. Gatti McArthur, D. Alberati and T. Bartfai			
Three Mile Island, Stress Effects of	735		
A. L. Dougall and A. Baum			
Thymus	738		
M. S. Vacchio			
Thyroid Hormones	743		
T. J. Visser and E. Fliers			

Violence	838	War, Suicide and Sacrifice	860
E. K. Englander		V. Hazboun	
Viral Virulence and Stress	842	War-Related Posttraumatic Stress Disorder, Treatment of	865
A. J. L. Macario and E. Conway de Macario		L. B. Slone and M. J. Friedman	
Viruses and Stress	850	Work-Family Balance	868
G. Z. Feuerstein, R. R. Ruffolo and B.-N. David		J. G. Grzywacz and A. B. Butler	
		Workplace Stress	871
		U. Lundberg	

W

Waist-Hip Ratio	853
R. Rosmond	
War Stress in the Former Yugoslavia	855
M. Flögel, S. Šupraha Goreta and G. Lauc	

VOLUME 4

<i>Contributors</i>	1
<i>Subject Index</i>	21

CONTENTS BY SUBJECT AREA

ANIMAL STUDIES AND MODELS

- Animal Models (Nonprimate) for Human Stress
- Circadian Rhythms, Effects of Prenatal Stress in Rodents
- Fish, Stress in
- Immobilization Stress
- Pre-pulse inhibition
- Primate Hierarchies and Personality
- Primate Models, Behavioral-Immunological Interactions
- Primate Models, Cardiovascular Disease
- Primate Models, Overview
- Primates: Rearing and Effects of Stress on Primate CNS Function
- Sex-Specific Effects of Early Social Stress in Mammals: a Study in Guinea Pigs

CONFLICT, WAR, TERRORISM

- Chemical Warfare
- Gulf War Syndrome, Psychological and Chemical Stressors
- Hiroshima Bombing, Stress Effects of
- Holocaust, Stress Effects of
- Impact of Terrorism on the Development of Mental Health Symptoms
- Korean Conflict, Stress Effects of
- Lockerbie Air Crash, Stress Effects of
- Northern Ireland, Post Traumatic Stress Disorder in
- Nuclear Warfare, Threat of
- Oklahoma City Bombing, Stress Effects of
- Persian Gulf War, Stress Effects of
- Prisoners of War
- 9/11, Religion and Stress
- Suicide Terrorism, Genesis of
- Terrorism
- Torture

- Traumatic Stress and Posttraumatic Stress Disorder; the Israeli Experience
- Vietnam Veterans, Postwar Experiences and Health Outcomes
- War Stress in the Former Yugoslavia
- War, Suicide and Sacrifice

DISASTERS

- Airline Accidents
- Chernobyl, Stress Effects of
- Disasters and Mass Violence, Public, Effects of
- Disaster Syndrome
- Earthquakes, Stress Effects of
- Floods, Stress Effects of
- Hurricane Katrina Disaster, Stress Effects of
- Motor Vehicle Accidents, Stress Effects of
- Three Mile Island, Stress Effects of

DIURNAL, SEASONAL AND ULTRADIAN RHYTHMS

- Circadian rhythm effects on cardiovascular and other stress-related events
- Circadian Rhythms, Genetics of
- Night Shiftwork
- Seasonal Changes in Stress Responses
- Seasonal Rhythms
- Sleep Loss, Jet Lag, and Shift Work
- Sleep, Sleep Disorders, and Stress
- Ultradian Rhythms

DRUGS (EFFECTS)

- Alcohol, Alcoholism and Stress: A Psychobiological Perspective
- Drug Use and Abuse
- Interactions between Stress and Drugs of Abuse
- Opioids

DRUGS (TREATMENT)

- Antipsychotic Drugs and Stress
- Anxiolytics
- Benzodiazepines
- Beta-Adrenergic Blockers
- Pharmacological Treatments of Stress
- Selective Serotonin Reuptake Inhibitors (SSRIs)
- Stress and Anxiety: treatment with 2nd generation antipsychotic drugs

GENERAL CONCEPTS AND MODELS

- Alarm Phase and General Adaptation Syndrome
- Allostasis and Allostatic Load
- Behavior, Overview
- Conservation of Resources Theory
- Control and Stress
- Demand–Control Model
- Effort–Reward Imbalance Model
- Environmental Factors
- Evolutionary Origins and Functions of the Stress Response
- Fight-or-Flight Response
- Instinct Theory
- Life Events Scale
- Psychological Stressors, Overview
- Remodelling of neuronal networks by stress
- Selye, Hans
- Stress, Definitions and Concepts of
- Stress Effects, Overview
- Stress System Balance Hypothesis
- Stress, Beneficial Effects of

GENETICS AND GENOMICS

- Circadian clock genes as modulators of sensitivity to genotoxic stress
- Corticosteroid Receptor Genes: Functional Dissection in Mice
- Corticotropin Releasing Factor Receptor Deficiency in Mice
- Drosophila Genes and Anoxia
- Expression Profiling of Stress Responsive Gene Patterns
- Gene Environment Interactions in Early Development
- Genetic Factors and Stress
- Genetic Polymorphisms in Stress Response
- Genetic Predispositions to Stressful Conditions
- Genetic Testing and Stress
- Genetic variation of HPA axis activity and function in farm animals
- Glucocorticoid Receptor Mutant Mice as Models for Stress-Induced Affective Disorders

- Glucocorticoid Receptor Mutations and Polymorphisms
- Mineralocorticoid Receptor Polymorphisms
- Neuroticism Genetic Mapping of
- Neuroticism Response to Stress, Genetic Mapping of Mice
- Serotonin Transporter Genetic Modifications
- Strain Differences in Stress Response in Rodents
- Stress and CNS Arousal; Genomic Contributions

HUMAN COGNITION, EMOTION AND BEHAVIOR

- Adolescence
- Aggression
- Aggressive Behavior
- Aging and Psychological Stress
- Anger
- Anxiety
- Bereavement
- Burnout
- Captivity, Adaptation to
- Captivity, Recovery from
- Caregivers, Stress and
- Childbirth and Stress
- Cognition and Stress
- Combat, Acute Reactions to
- Combat Reaction, Chronic
- Combat Stress Reaction
- Concentration Camp Survivors
- Coping Skills
- Coping and Stress: a Lens and Filter Model
- Cortisol Awakening Response
- Death Anxiety
- Dental Stress
- Depersonalization: Systematic Assessment
- Diet and Stress, Non-Psychiatric
- Dissociation
- Distress
- Emergency Personnel, Stress in
- Emotional Inhibition
- Emotions: Structure and Adaptive Functions
- Fatigue and Stress
- Fear
- Fear and the Amygdala
- Firefighters, Stress in
- Grieving
- Holocaust Survivors, Experiences of
- Homosexuality, Stress and
- Hostility
- Impotence, Stress and
- Impulse Control
- Industrialized Societies
- Learning and Memory, Effects of Stress on
- Marriage

Medical Profession and Stress
 Memory Impairment
 Menopause and Stress
 Neuroimaging and Emotion
 Nightmares
 Optimism, Pessimism and Stress
 Pain
 Parenting, Stress of
 Peacekeeping
 Police, Stress in
 Prison
 Psychosomatic Medicine
 Recovery from Stress
 Refugees, Stress in
 Religion and Stress
 Self-Esteem, Stress, and Emotion
 Somatic Disorders
 Stress Generation
 Stress in University Students
 Suicide, Psychology of
 Suicide, Sociology of
 Surgery and Stress
 Survivor Guilt
 Teaching and Stress
 Trauma and Memory
 Type A Personality, Type B Personality
 Understimulation/Boredom

HUMAN HEALTH AND PHYSICAL ILLNESS

Amnesia
 Adrenal Insufficiency
 AIDS
 Alzheimer's Disease
 Arthritis
 Asthma
 Atherosclerosis
 Attention-Deficit/Hyperactivity Disorder, Stress and
 Breast Cancer
 Cancer
 Cardiovascular Disease, stress and
 Chaperonopathies
 Chronic Fatigue Syndrome
 Common Cold and Stress
 C-Reactive Protein
 Cytokines, Chronic Stress, and Fatigue
 Cushing's Syndrome, Medical Aspects
 Cushing's Syndrome, Neuropsychiatric Aspects
 Dermatological Conditions
 DHEA
 Diabetes, Type I
 Disease, Stress Induced
 Eclampsia and preeclampsia
 Endometriosis
 Enuresis

Epilepsy
 Fibromyalgia
 Graves' Disease (Thyrotoxicosis)
 Health Behavior and Stress
 Heart Disease/Attack
 Heart Failure, Stress Effects
 Hemostasis and Stress
 Hyperreactivity (Cardiovascular)
 Hypertension
 Hyperthermia
 Hyperthyroidism
 Hyperventilation
 Hypoglycemia
 Hypotension, Hypovolemia, and Septic Shock
 Hypothermia
 Hypothyroidism
 Hysteria
 Metastasis
 Metastasis
 Migraine
 Multiple Trauma
 Nelson's Syndrome
 Neurodegenerative Disorders
 Organ Transplantation, Stress of
 Parkinson's Disease
 Perinatal Dexamethasone
 Pregnancy, Maternal and Perinatal Stress, Effects of
 Premenstrual Dysphoric Syndrome
 Prolactin and stress
 Psychosomatic heart disease: role of sympathetic
 and sympathoadrenal processes
 Rheumatic Disorders
 Sickle Cell Disease and Stress
 Space, Health Risks of
 Spinal Cord Injury, Physical Stress of
 Spinal Cord Injury, Psychological Stress of
 Stress Effect of Assisted Reproduction
 Stress Induced Anovulation
 Stress, Insulin Resistance and Type II Diabetes
 Stress, NPY and Cardiovascular Diseases
 Ulceration, Gastric

HUMAN MENTAL HEALTH AND PSYCHOPATHOLOGY

Acute Stress Disorder and Posttraumatic Stress
 Disorder
 Adjustment Disorders
 Adolescent suicide
 Affective Disorders
 Antisocial Disorders
 Anxiety
 Arthritis – Psychological
 Borderline Personality Disorder
 Child Abuse
 Child Sexual Abuse

Comorbid Disorders and Stress
Corticotropin-Releasing Factor Circuitry in the
Brain – Relevance for Affective Disorders
and Anxiety
Defensive Behaviors
Depression and Coronary Heart Disease
Depression and Manic-Depressive illness
Depression and Stress, Role of n-3 and n-6
Fatty Acids
Depression, Immunological Aspects
Depression Models
Diet and Stress, Psychiatric
Disaster Syndrome
Dissociation
Eating Disorders and Stress
Elder Abuse
Ethnicity, Mental Health
HPA alterations in PTSD
Hypocortisolism and Stress
Incest
Learned Helplessness
Life Events and Health
Loss Trauma
Male Partner Violence
Major Depressive Disorder
Multiple Personality Disorder
Negative Affect
Neurodevelopmental Disorders in Children
Neurosis
Obsessive–Compulsive Disorder
Panic Disorder and Agoraphobia
Paranoia
Posttraumatic Stress Disorder – Clinical
Posttraumatic Stress Disorder, Delayed
Posttraumatic Stress Disorder in Children
Posttraumatic Stress Disorder – neurobiological
basis for
Posttraumatic Stress Disorder,
Neurobiology of
Psychotic Disorders
Revenge Fantasies
Schizophrenia
Sexual Assault
Sexual Offenders
Social support in trauma
Stress of Self Esteem
Suicide, Biology of
Trans-sexualism
Violence

IMMUNOLOGY, INFECTION AND INFLAMMATION

Antibody Response
Autoimmunity

Corticotropin-releasing factor (CRF) family of
neuropeptides – role in inflammation
Cytokines
Cytokines, Stress, and Depression
Cytotoxic Lymphocytes
Herpesviruses
HIV Infection/AIDS
Immune Cell Distribution, Effects of Stress on
Immune Function, Stress-Induced Enhancement of
Immune Response
Immune Suppression
Immune Surveillance – Cancer, Effects of Stress on
Immune System, Aging
Immunity
Infection
Inflammation
Leishmania, Stress Response in
Leukocyte Trafficking and Stress
Lymphocytes
Mucosal Secretory Immunity, Stress and
Multiple Sclerosis
Natural Killer (NK) Cells
Neuroimmunomodulation
Neuroinflammation
Psoriasis
Psychoneuroimmunology
Sepsis, Acute Respiratory Distress Syndrome and
Glucocorticoid Resistance
Systemic Lupus Erythematosus
Vaccination
Viral Virulence and Stress
Viruses and Stress

LABORATORY STUDIES AND TESTS

Antimineralocorticoid Challenge
Dex-CRH Test
Dexamethasone Suppression Test (DST)
Mental Stress Testing
Metyrapone: Basic and Clinical Studies
Trier Social Stress Test
Waist–Hip Ratio

THERAPIES

Aerobics, Exercise and Stress Reduction
Behavior Therapy
Cancer Treatment
Desensitization
Exercise
Family Therapy
Group Therapy
Hypnosis
Integrative Medicine (Complementary and
Alternative Medicine)

Meditation and Stress
 Problem-Solving Skills Training
 Psychosomatic Medicine
 Reenactment Techniques
 Relaxation Techniques
 Stress Management and Cardiovascular Disease
 Stress Management, CAM Approach
 Trauma Group Therapy

PHYSIOLOGICAL, PHARMACOLOGICAL AND BIOCHEMICAL ASPECTS

Acute Stress Response: Experimental
 Acute Trauma Response
 Adenylyl Cyclases and Stress Response
 Adrenal Cortex
 Adrenaline
 Adrenal Medulla
 Adrenocorticotrophic Hormone (ACTH)
 Adrenocortical Function, factors controlling development thereof
 Aging and Stress, Biology of
 Aging and Adrenocortical Factors
 Alcohol and Stress: Social and Psychological Aspects
 Aldosterone and Mineralocorticoid Receptors
 Ambulatory Blood Pressure Monitoring
 Amenorrhea
 Amygdala
 Anatomy of the HPA Axis
 Androgen Action
 Angiotensin
 Angiotensin receptors
 Annexin
 Anti-CRF
 Antidepressant Actions on Glucocorticoid Receptors
 Apoptosis
 Arterial Baroreflex
 Autonomic Nervous System
 Autotolerance
 Avoidance
 Beta-Endorphin
 Blood–Brain Barrier, Stress and
 Blood Pressure
 Brain and Brain Regions
 Brain Natriuretic Peptide (BNP)
 Brain Trauma
 Calbindin
 Calcium-Dependent Neurotoxicity
 Calcium, Role of
 Cardiovascular System and Stress
 Catecholamines
 Central Stress Neurocircuits
 Cerebral Metabolism, Brain Imaging
 Chaperone Proteins and Chaperonopathies
 Cholesterol and Lipoproteins
 Comparative Anatomy and Physiology
 Congenital Adrenal Hyperplasia (CAH)
 Corticosteroid-Binding Globulin (Transcortin)
 Corticosteroid Receptors
 Corticosteroids and Stress
 Corticotropin Releasing Factor (CRF)
 Corticotropin Releasing Factor-Binding Protein
 Corticotropin Releasing Factor Receptors
 Critical Thermal Limits
 Dopamine, Central
 Drosophila Studies
 Electrodermal Activity
 Endocrine Systems
 Estrogen
 Ethanol and Endogenous Opioids
 Excitatory Amino Acids
 Excitotoxins
 Familial Patterns of Stress
 Febrile Response
 Feedback Systems
 Fibrinogen and Clotting Factors
 Feeding Circuitry (and Neurochemistry)
 Fetal Stress
 Food Intake and Stress, Human
 Food Intake and Stress, Non-Human
 Food Shift Effect
 GABA (Gamma Aminobutyric Acid)
 Gastrointestinal Effects
 Gender and Stress
 Ghrelin and Stress Protection
 Glia or Neuroglia
 Glucocorticoid Negative Feedback
 Glucocorticoid Effects on Memory: the Positive and Negative
 Glucocorticoids – Adverse Effects on the Nervous System
 Glucocorticoids, Effects of Stress on
 Glucocorticoids, Overview
 Glucocorticoids, Role in Stress
 Glucose Transport
 Glycobiology of Stress
 Gonadotropin Secretion, Effects of Stress on
 Heart Rate
 Heat Resistance
 Heat Shock Genes, Human
 Heat Shock Proteins: HSP60 Family Genes
 Heat Shock Response, Overview
 Hippocampal Neurons
 Hippocampus, Corticosteroid Effects on
 Hippocampus, Overview
 Homeostasis
 11 β -Hydroxysteroid Dehydrogenases
 Hypothalamic-Pituitary-Adrenal Axis

Instinct Theory
Kidney Function
Left Ventricular Mass
Leptin, Adiponectin, Resistin, Ghrelin
Lymph Nodes
Macrophage Antimycobacterial Activity, Effects of
Stress on
Macrophages
Maternal Deprivation
Membrane Glucocorticoid Receptors
Memory and Stress
Menstrual Cycles and Stress
Metabolic Syndrome
Metals, Oxidative Stress and Brain Biology
Mitochondria
Monoamine Oxidase
Multi Drug Resistance P Glycoprotein and other
Transporters
Musculoskeletal Problems and Stress
Myopathy
Neural Stem Cells
Neuroendocrine Systems
Neurogenesis
Neuropeptide Y
Neuropeptides, Stress-Related
Nitric Oxide
Nutrition
Obesity, Stress and
Orexin
Oxidative Stress
Oxidative Stress and Acidosis, Molecular
Responses to
Oxidative Stress and Aging
Oxytocin
Paraventricular Nucleus
Peptides
Personality Processes
Pheromones
Pituitary Regulation, Role of
Pressure, Effects of Extreme High and Low
Pro-opiomelanocortin (POMC)
Prostaglandins
Proteases in Prokaryotes and Eukaryotic Cell
Organelles
Proteases in the Eukaryotic Cell Cytosol
Protein Synthesis
Proteosome
Reductive Stress
Regional Blood Flow, Stress Effects
Renal and Adrenocortical Effects of
Dopamine
Reproduction, Effects of Social Stress on
Reproductive Dysfunction in Primates,
Behaviorally Induced

Resistance
Restraint Stress
Salivary Cortisol
Salt Appetite
Secretagogue
Serotonin
Serotonin in Stress
Sex Differences in Human Stress Response
Sex Steroids, Response to Stress and Susceptibility
to Depression
Sexual Dysfunction
Smoking and Stress
Startle Response
Steroid Hormone Receptors
Steroid Hydroxylases
Stress Hyporesponsive Period
Sympathetic Nervous System
Synthetic Glucocorticoids
Temperature Effects
Thermal Stress
Thermotolerance, Thermoresistance, and
Thermosensitivity
Thymus
Thyroid Hormones
Urocortin
Vasoactive Peptides
Vasopressin

PSYCHOLOGICAL THERAPY

Cognitive Behavioral Therapy
Freud, Sigmund
Posttraumatic Therapy
Psychoanalysis
Psychotherapy
War-Related Posttraumatic Stress Disorder,
Treatment of

PSYCHOSOCIAL AND SOCIOECONOMIC ASPECTS

Childhood Stress
Child Physical Abuse
Chronic Social Stress: GR Sensitivity
in Leukocytes
Community Studies
Crowding Stress
Crime Victims
Crisis Intervention
Cultural Factors in Stress
Cultural Transition
Divorce, Children of
Domestic Violence
Economic Factors and Stress
Education Levels and Stress

Employee Assistance and Counseling
 Environmental Stress, Effects on Human
 Performance
 Health and Socioeconomic Status
 Income Levels and Stress
 Indigenous Societies
 Job Insecurity; the health effects of a psychosocial
 work stressor
 Marital Conflict
 Marital Status and Health Problems
 Minorities and Stress
 Neighborhood Stress and Health
 Psychosocial Factors and Stress

Quality of Life
 Racial Harassment/Discrimination
 School Violence and Bullying
 School Stress and School Refusal Behavior
 Social Capital
 Social Networks and Social Isolation
 Social Status and Stress
 Social Stress, Animal Models of
 Social Support
 Transport-Related Stress
 Unemployment, Stress and Health
 Work–Family Balance
 Workplace Stress

PREFACE TO FIRST EDITION

“Stress” remains one of the most frequently used but ill-defined words in the English language. Stress is a phenomenon that has quite different meanings for the politician, social scientist, physician, nurse, psychotherapist, physiologist, molecular biologist, and perhaps you and me. This diversity of meanings was one impetus for creating the *Encyclopedia of Stress*, the aim being to derive a definition of stress from a variety of expert descriptions. The second impetus was the obvious need for an up-to-date compendium on one of the most important social, medical, and psychological phenomena of our age. We were fortunate in attracting stars for our Editorial Board and a set of most distinguished contributors for the 400 or so entries—indeed, the list of contributors is a Who’s Who in stress research.

We anticipate that the diversity of our readers will equal the diversity of the topics covered. They will find that the coverage of the Encyclopedia extends well beyond the general adaptation theory of Hans (Janos) Selye and the fight-or-flight response of Walter Cannon. Nevertheless, the general principles annunciated by these two great pioneers in the field still underpin our understanding of the biology of the stress phenomenon. That is, stress is a real or perceived challenge, either endogenous or exogenous, that perturbs body equilibrium or “homeostasis.” The stressor may range from overcrowding, traffic congestion, violence, bereavement, redundancy, or unemployment to physical, chemical, biological, or psychological insults. Whether the person can adapt to or cope with the stress will depend on the nature and severity of the stressor and the person’s physical and mental state, which in turn depends on genetic, experiential, social, and environmental factors. These issues are discussed in depth in the Encyclopedia, as are the mechanisms of coping and the impact of stress on health and predisposition to diseases such as

cancer, infection, rheumatoid arthritis, heart disease, high blood pressure, and mental disorder.

Aggression remains a hallmark of human behavior, even as we move into the third millennium, and so the Encyclopedia covers several topical areas that have only recently been analyzed systematically. These include war and specific wars, posttraumatic stress disorder (PTSD; formerly thought of vaguely as “shell shock”), rape, torture, marital discord and spousal abuse, and the Holocaust. In tackling these topics we accept that our entries may not include all the nuances that are necessary for a full understanding of what these phenomena are all about and described so graphically and sensitively in Tolstoy’s *War and Peace* or Pat Barker’s monumental *Regeneration* trilogy on the horrific psychological traumas of the First World War. Nevertheless, an important start has been made in that we now accept that PTSD is not just a lack of “bottle” (courage or “guts”), but rather a syndrome that needs to be and can be understood within the framework of medicine and psychology.

Biologically, the stress response reflects a set of integrated cascades in the nervous, endocrine, and immune defense systems. As in most areas of biology, molecular genetics has made a significant difference in the precision with which we now understand the physiopathological processes of the stress response. And so the adage formerly applied to diabetes mellitus may now apply equally to stress: “Understand stress and you will understand medicine.”

In summary, we hope that this first *Encyclopedia of Stress* will indeed define the phenomenon and at the same time provide a valuable source of information on a phenomenon that affects us all. In setting out on this adventure we were aware that there is nothing new under the sun and that stress has been around since the first biological particles, bacteria, or even viruses competed for the same mechanisms of replication.

There is a tendency for each generation to imagine that stress and its untoward effects are uniquely harsh for them, but it is not this misconception that underlies this work. Rather, the stress of “stress” itself—the massive accumulation of knowledge—made it seem propitious to bring the information together in a systematic manner that allows ready access to all who need or wish to understand the phenomenon.

The idea of producing this Encyclopedia was conceived at an Academic Press reception in San Diego held in conjunction with the annual meeting of the Society for Neuroscience in 1996. I am deeply indebted to Erika Conner for enabling conversion of the idea to a concept and then a project and to

Jennifer Wrenn, Christopher Morris, and Carolan Gladden, all of Academic Press, for their enthusiasm, encouragement, and herculean efforts that converted the concept into a reality. For giving generously of their intellect, expertise, sound advice, and unstinting work, I am greatly indebted to my friends and colleagues on the Editorial Board, who made the project such a satisfying experience. To all our contributors go our profound thanks for taking time out from wall-to-wall schedules to produce their entries, which together have made this Encyclopedia.

George Fink
August 1999

PREFACE TO THE SECOND EDITION

The popularity of the first edition of the Encyclopedia of Stress, the several major stressful conflicts that have occurred since 1999, and the significant advances in our knowledge of stress prompted the production of this second edition. In addition to more than 140 new articles, most of the original 400 first edition articles have been updated for the second edition to reflect, for example, advances in our understanding of corticotropin releasing factor (CRF) and urocortin receptors, ideas about the role of central CRF in the overall control of the psychological/mental as well as neuroendocrine response to stress, and the novel discoveries of the complex neuropeptide regulation of hunger and satiety. The effects of maternal and perinatal stress on subsequent body development and function in the adult, the adverse effects of stress on brain (and particularly hippocampal memory function), the role of the amygdala in fear and aggression, and the role of stress in the etiology of obesity and the metabolic syndrome all feature prominently in this new edition. 9/11 and the many new conflicts since 1999 have engendered new articles on terrorism, suicide bombers and their effects, as well as a retrospective on previous human apocalypses such as Hiroshima. Coverage of the whole field of post traumatic stress disorder (PTSD) has been expanded, reflecting the greater recognition and understanding of PTSD, and possibly related disorders such as *combat fatigue* and *burnout*. Stress and depression and anxiety, already touched on in the first edition, now feature prominently. On the conceptual side, *allostasis* (modified from Hans Selye's *heterostasis*), or stability through centrally regulated change in physiological set points, receives greater emphasis than it did in the first edition.

Overall, our understanding of stress mechanisms in the human has benefited from two major quantum leaps in technology. First, the advances in molecular genetics and genomics and sequencing of the human

genome (published after the first edition had appeared) have increased the rigor and precision of our understanding of the molecular neurobiology of stress. Thus, new to the second edition are a series of articles on the genetic factors that play a role in susceptibility to stress and in various components of the stress response. Secondly, functional brain imaging has enhanced our understanding of the neurobiology of stress in the human. Many of the articles in the second edition reflect the positive impact of these two advances.

Accessibility to the second edition of the Encyclopedia of Stress will be facilitated by virtue of the fact that it is published online as well as in hard copy. My profound thanks for their selfless work go to our team of distinguished authors and to my friends and colleagues on the Editorial Board. Production of the Encyclopedia would not have been possible without our colleagues at Elsevier, and in particular Bob Donaldson and Hilary Rowe whose dedication, commitment and excellent work I am pleased to acknowledge.

George Fink
November 2006

Note on Terminology of Corticotropin Releasing Factor/Hormone and the Catecholamines

The central nervous regulation of the anterior pituitary gland is mediated by substances, mainly peptides, which are synthesized in the hypothalamus and transported to the gland by the hypophysial portal vessels. Because these compounds are transported by the blood, the term "hormone" gained acceptance in the neuroendocrine literature. The major hypothalamic peptide involved in the stress response is the

41-amino acid corticotropin releasing factor (CRF). The Endocrine Society (USA), following convention, adopted the term corticotropin releasing hormone (CRH). However, this nomenclature has been challenged. Hauger et al. (*Pharmacol Rev* 55: 21–26, 2003), in liaison with the International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification, argued that the CRF's function extends well beyond the biology of a hormone, and that therefore it should be termed corticotropin releasing factor (CRF) rather than hormone. Since the terminology of CRF versus CRH has yet to be resolved, the two terms and abbreviations are here used interchangeably, depending on author preference.

Adrenaline and *noradrenaline* are catecholamines that play a pivotal role in the stress response. These

terms are synonymous with *epinephrine* and *norepinephrine*, respectively. Both sets of terms are used interchangeably in the endocrine, neuroendocrine and stress literature, and this principle has been adopted for the Encyclopedia. Style has depended on author preference, but wherever possible within-article consistency has been ensured.

George Fink
November 2006

Hauger, R. L., Grigoriadis, D. E., Dallman, M. F., Plotsky, P. M., Vale, W. W., and Dautzenberg, F. M. (2003). International Union of Pharmacology. XXXVI. Current status of the nomenclature for receptors for corticotropin-releasing factor and their ligands. *Pharmacological Reviews* 55(1), 21–26.

FOREWORD

This is the second edition of the *Encyclopedia of Stress*, the three heavy volumes that first appeared in 2000 under the overall editorial guidance of George Fink. From its original coining in reference to rather unusual and unpleasant situations, the word *stress* is now quasi synonymous with *life*, i.e. modern life! This second edition of the *Encyclopedia of Stress* presents several hundreds of articles written by experts in that particular field, especially for this series, discussing, clarifying, explaining to the average-reader as well as to anyone more educated in one field or another, the extraordinary involvement of stress and our normal or abnormal reactions to it, both in our physical body and its accompanying mind. The table of contents is truly amazing, from the stress of being born to the stress of dying and so many circumstances in between, with clear presentations of the pertinent, current knowledge about them. This is truly an encyclopedia where, surely, you will find some explanation and answer to your own stresses (and those of others).



Roger Guillemin, MD, PhD
Nobel laureate in Physiology & Medicine

F

Familial Patterns of Stress

A Bifulco

University of London, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A Bifulco, volume 2, pp 103–107, © 2000, Elsevier Inc.

Measurement of Stressors

Stress and Disorder

Shared Stressors and Disorders in Families

Transmission of Stress Intergenerationally

Glossary

Childhood stressors

Adverse events or difficulties impacting on an individual in childhood; they typically involve specific experiences impinging on the child (e.g., parental loss; neglect, and physical, sexual, or psychological abuse), adverse family circumstances for the child (e.g., poverty, parental conflict, or parental psychiatric or criminal disorder), and child life events and difficulties. Childhood stressors can have an immediate impact or a longer-term impact in adult life.

Family environment

Experiences potentially common to the whole family unit. Thus, shared environment refers to experiences held in common, particularly within the household; nonshared environment refers to experiences that are different for various family members, specifically siblings, including the school and social arena.

Family studies

Studies that assess family members to look at the concordance of both experience and psychological disorders, as well as the intergenerational transmission of risk; they include parent to child comparisons, intrasibling and twin comparisons, and nontraditional family arrangements

Intergenerational transmission

Psychological disorder

Sibling relationships

Stress agency

such as step-parent households, lone parents, and adoptive parents and siblings.

The passing of the risk of disorder and adversity from parent to child and within families across generations; several mechanisms are likely to be involved, including social (perpetuation of high rates of adversity/stressors), psychological (poor parenting and negative interpersonal style), and biological (genetic or constitutional) factors.

A disorder that results from the negative impact of stressors on psychological health (i.e., being stressed), such as major depression and anxiety disorders; these disorders have been studied in relation to familial stress in both adults and young people.

Relationships between monozygotic (identical) twins, dizygotic (fraternal) twins, full siblings, half-siblings, and adoptive siblings. Studies of these relationships are central to behavioral genetic studies of families and include designs to show cascade effects when comparing similarities in characteristics and behavior in siblings and studies examining similarities and differences in siblings brought up together and apart. Likely genetic contributions can be inferred from the known genetic similarity of different types of sibling pairs and the similarity of their measured characteristics.

The quality of bringing about stressful situations through poor coping with adversity and mismanagement of interpersonal interactions. People with high agency in producing stress are more likely to have externalizing disorders, such as delinquency, antisocial personality disorder, substance abuse, and criminal behavior, which impact negatively on close others and family members. Assessing agency or independence of life events

	is key to establishing causal direction of the impact of stressors.
<i>Stress carriers</i>	Individuals who are both the recipients and bearers of high levels of stressors, which, although having no direct agency in their generation, can nevertheless transmit the stress to other family members. Mechanisms may be through passive and helpless coping styles, which perpetuate or increase difficulties. This is more common among those with the internalizing disorders such as depression and anxiety. For example, the choice of a partner with antisocial behavior may lead to a higher experience of stressful circumstances for both parents and children, even though specific difficulties are only indirectly brought about by the individual concerned.
<i>Stressors</i>	Adverse events or difficulties impacting on an individual.
<i>Vulnerability factors</i>	Characteristics that increase an individual's susceptibility to the impact of stressors; they interact with stressors to increase the risk of the onset of psychiatric disorder for particular individuals. Such factors can be psychological (e.g., low self-esteem and insecure attachment style), social (e.g., interpersonal conflict and lack of close confidants), or biological (e.g., irregular cortisol patterns). Vulnerability factors can be recent and ongoing or stem from early life experiences such as neglect or abuse in childhood.

Measurement of Stressors

The psychiatrist Adolf Meyer was the first modern practitioner in the 1950s to suggest that life events relate to psychopathology, developing a life chart system for recording the temporal relationships of life experiences and disorder. This method was developed further in the 1960s by Holmes and Rahe, with a questionnaire based on a summation of the quantity of environmental change in a given period in terms of a list of items representing both normative and non-normative changes such as moving one's residence or bereavement. The approach was then revolutionized in the 1970s both in psychiatry (by Paykel) and in social science (by Brown) by measuring the contextual threat or unpleasantness of events in more detail and by using more sophisticated assessment techniques to incorporate meaning into experience. This involved attending to the characteristics of events (such as whether they were undesirable or involved losses, entrances, or exits) by taking the context of the event into account when assessing its likely

unpleasantness, by examining the duration of the stressor (i.e., both chronic and acute), and finally by having the researcher assess the severity of the event according to benchmarked examples to circumvent the biases raised by self-report. Relevant context is crucial; for example, in the case of a pregnancy it is necessary to collect factual details of the experience including whether it was planned, the state of the marriage/partnership, the health of the individuals, and their financial and housing situation in order to assess its negative and potentially depressogenic characteristics. Although more attention was paid to life events because of their clear dating in relation to disorder, a role for chronic stressors or difficulties was also found. These involved problematic situations that continued over time and that frequently generated new events, for example, a financial difficulty involving debts that fluctuates in severity as new financial events occur. Severe events, matched to areas of ongoing marked difficulties, were found to be particularly pathogenic.

Studies in inner-London in the 1970s showed that severe life events (those that are highly threatening or unpleasant) were relatively common in the community with approximately one in three individuals experiencing at least one in a year. They were also found to be more common in working-class groups, in inner-city locations, and among single-parent families. Among events judged to be severe, three-quarters involved close relationships and these more often led to depression outcomes. The extension of this methodology to children was undertaken in the 1980s simultaneously by Goodyer and by Sandberg, following the Paykel and Brown traditions, respectively. A similar role for stressor and disorder was found in children but with chronic difficulties given a more prominent role. Parallel work looked at stressors simultaneously in parents and children and showed a large degree of familial patterning.

Stress and Disorder

Prior to the 1970s a relationship between the occurrence of life events and relapse of schizophrenia had been confirmed, and in the 1970s and 1980s the main development was the exploration in adults of the relationship of life events and the onset of affective disorder, particularly major depression. The relationship between severe life events and the onset of major depression was confirmed largely in samples of women and among mothers, in whom higher rates of depression are typically found.

A model was investigated prospectively, looking at the vulnerability characteristics of individuals who later succumbed to affective disorder when

encountering stressors. In prospective community studies, it was possible to identify interpersonal and cognitive factors that placed an individual at higher risk of affective disorder when later experiencing a severe life event. For adults, vulnerability factors encompassed family conflict (negative interaction with a partner or child), lack of a close confidant outside the home, and low self-esteem. When negative family interactions, low self-esteem, and severe life events all occurred, the risk for the onset of major depression was raised more than fourfold (47% versus 10% in the community at large). When lifetime experience was examined, the presence of neglect or abuse in childhood was highly related both to disorder and adult vulnerability, with one-third of women with adverse childhoods becoming depressed in a 12-month study period. A similar model is emerging for children, whereby family conflict and low self-esteem increase a child's susceptibility to disorder on encountering severe stressors. Stress can also have negative biological impacts, for example on stress hormones. Thus, abnormal cortisol patterns are associated with stressors and with a range of psychological and physical illnesses. Children experiencing chronic neglect and abuse have been shown to have abnormal diurnal cortisol patterns associated with behavior problems and poor functioning.

Lifespan models of disorder have examined linkages in early life adversity with later life impacts on disorder. In particular, neglect or abuse in childhood has been linked with adverse social trajectories, including unplanned home-leaving, teenage pregnancy, lone parenthood, high rates of adult life events and difficulties, and lower social class. The same early life experiences have been linked with negative psychological impacts at different life stages, including low self-esteem, helplessness, poor coping, and insecure attachment styles. These provide causal links between early-life family adversity and later-life stressors, which often impact subsequent family formation.

Shared Stressors and Disorders in Families

Studies have consistently shown that both stressors and disorder congregate within families and that these provide simultaneous, as well as later, life risks for both parents and children. Not only do family conflicts and severe events increase the risk of disorder in both, but there is also evidence that disorder in parents provides additional risk for children, and vice versa. Although most studies have shown links between mothers' and children's disorder, associations have also been shown with the fathers' psychological disorder. Disorder in both parents increases the risk for

the children substantially. Contributory risk factors from mothers include a history of affective disorder, current depressed symptoms, vulnerability characteristics, and ongoing chronic stressors, all of which play a part in the children's disorder. Although chronic stressors predict children's disorder independently of mental illness in the parents, the accompanying maternal impairment and poor parenting are associated with higher negative outcomes. This is likely because common stressors (e.g., marital conflict) affect the mothers and children and, in addition, mothers who are depressed are emotionally unavailable as a protective influence to help the children to cope. Thus, without the buffering effect of maternal support, the children are additionally susceptible to stress and can become symptomatic. The role of types of family stressors, family structure, sibling stressors, parenting, and parental disorder in familial patterning of stressors is now examined.

Types of Family Stressors

The types of family stressors that have proved to be important in predicting disorder in children involve family conflict and parental loss and bereavement. Children in families with parental affective disorder experience more acute stressful events, particularly those of an interpersonal nature, than children with parents with other problems, such as medical illnesses. However, chronic stressors are equally high in both. The mothers' disorder itself also acts as stressor for the children, thereby increasing risk. Children's responses to stressors is another contributor to risk, with negative self-concept being not only more common in children with higher numbers of stressors but also increasing the risk of disorder, particularly depression. These negative cognitions are associated with criticism and noninvolvement by the mothers. Although little work has been done on the agency of the stressors, aspects of both stress agency and stress carrying are likely to be present in the parents and children, consistent with the disorders and associated vulnerability.

Family Arrangements

Nontraditional family structures following the loss of a biological parent have been examined in relation both to stress and parenting and to its differential effects on siblings raised together and raised apart explored. Siblings brought up apart in different family arrangements show the greatest difference in adjustment, although even siblings raised together show different responses to family conflict and stressors. Studies agree that the loss of parents affects children differentially, even when objective aspects of the loss

are assessed. Poor parenting in terms of neglect and physical abuse are more common in certain family structures, such as when a biological and step-parent are present.

Sibling Stressors

Siblings' experiences have mainly been studied in the context of genetic studies to identify the sources of shared and nonshared environments to help explain the differences and similarities between siblings. These have used a variety of research strategies involving family studies, adoption studies, and twin studies to look at variations between siblings in personality, disorder, and experience of parenting. Interestingly, siblings have been shown to be largely different from one another in most assessments, and thus a large focus of such studies has been on assessing the role of the nonshared environment in explaining experience, personality, and psychopathology. The reasons for differences in siblings' environments include their relative ages, position in the family, gender, extra-familial experience such as school and social experience, and more random accidental or idiosyncratic experience. Although less work has been conducted on the degree to which stressors are shared by siblings, studies that do exist show a high degree of sharing of life events but less sharing of the negative impact of these events. Thus only approximately one-third of events held in common have a similar degree of negative impact for different family members. Twin studies have shown there to be a high degree of concordance for severe life events and a greater impact of such events in monozygotic than dizygotic twins.

In the late 1990s, an approach opposing the importance of family stressors emphasized the importance of children's experiences outside the family. Nonshared environmental influences in terms of school and social life were postulated as the critical environmental influences on child development and disorder, with familial influences being accounted for by genetic inheritance. Thus, sibling comparisons examining nonshared and usually nonfamily-based stressors, as well as positive experiences, have also been investigated.

Parenting

The quality of parenting has mainly been assessed around issues of the care and control of children. Poor parenting involves low levels of care and high control; at one extreme, this can involve the neglect and abuse of children. Parenting styles have also been categorized as authoritative, authoritarian, permissive, and disengaged, with the first of these being the most adaptive and related to positive child development. At the more extreme levels of poor parenting,

involving neglect and physical abuse, there is a high degree of shared experience among same-sex siblings. But where cold or critical parenting is involved, differential treatment by parents is much more common, with both favoritism and scapegoating of individual offspring occurring quite frequently. A process model of parenting shows the effects of the parents' own early development and personality, adult work history, marital relationship, and social network on parenting. This, in turn, impacts on child development and characteristics.

Parent Disorder

Psychiatric disorder in the parents can act as a magnet for familial stress, both in provoking the disorder and in perpetuating it. There is a strong relationship between concurrent disorder in parents (particularly mothers) and children. However, the chronicity and severity of the impairment by the disorder are more highly related to child's adjustment than to the actual parental diagnosis. Although the fathers' diagnoses exert weaker effects than mothers', the presence of disorder in both parents substantially increases the likelihood of disorder in the offspring. Because adults with depression have a higher likelihood of being partnered by people who have a psychiatric disorder, the likelihood of two parents being affected increases. Thus, 25% of husbands of women with depression have been shown to have disorders themselves, and as many as 41% of wives of men with a psychiatric diagnosis were similarly affected. This is argued to occur not only through assortative pairing, whereby individuals with a disorder appear to select partners with a similar impairment but also through the subsequent development of a disorder in the partner of an affected individual. This may be due to existing spousal difficulties or to the psychiatric disorder in both being caused by prior conditions in the family, such as poverty or conflict.

Transmission of Stress Intergenerationally

Because disorders and interpersonal difficulties run in families, three mechanisms of intergenerational transmission from parent to child have been investigated, with most studies suggesting that all three are likely to play a role.

Genetic Transmission

The genetic investigation of family influences on psychiatric disorder is engaged in the disentangling of genetic and environmental influences. A case has consistently been made for gene-environment correlations,

with children appearing to inherit some biological qualities that later result in depression and some environments that shape and pull their experiences in ways that lead to depression and other disorders. Three types of gene–environment correlations have been identified, and the current challenge in the field lies in assessing the role of each of these in the intergenerational transmission of both disorders and susceptibility to a disorder.

1. Passive genotype–environment correlations occur when children passively inherit an environment from parents that is correlated with their genetic predisposition.
2. Reactive genotype–environment correlations occur when children actively evoke environments associated with their genetic endowment.
3. Active genotype–environment correlations occur when the children actively seek out environments correlated with their genetic endowment.

Some progress is being made in seeking molecular genetic modes of transmission for adversity and disorder. For example, genetic deficiencies in monoamine oxidase A (MAOA) activity, responsible for metabolizing neurotransmitters such as norepinephrine (NE), serotonin (5-HT), and dopamine (DA), have been linked with aggression. The examination of violent offending using the Dunedin longitudinal study of children followed from birth showed a functional polymorphism in the gene encoding MAOA, which moderated the effects of childhood maltreatment on subsequent violent behavior in boys. An interaction between the level of maltreatment in childhood and low MAOA activity predicted a threefold higher conduct disorder rate and a tenfold higher violent convictions rate. This is an example of genetic factors playing a moderating or protective role. Thus, the low activity of MAOA and maltreatment accounted for only 12% of one male cohort studied but for 44% of the cohort's violent convictions. Eighty-five percent of cohort males who had the low-activity MAOA genotype and who were also maltreated developed some form of antisocial behavior.

Direct Impact of Parental Disorder

Another likely mechanism of transmission of stressor and susceptibility to disorder from parents to children is through the direct effect of the parents' disorder. In addition to providing additional stress for the children, the parents' disorder also presents the children with a poor role model for coping and reduces the possibility of any buffering effect from parental support. Three generations of influence have been identified, with women raised in dysfunctional families

likely to have impaired parenting skills from their parents (i.e., the grandparent generation), in turn increasing the risk of disorder in the children (i.e., the grandchild generation). However, evidence that individuals with disorders are more likely to be partnered with others with disorders means that this does not operate exclusively through the maternal line; disorder in both parents has been shown to provide the greatest risk for children.

The mother–child interaction is, however, particularly critical in intergenerational transmission, and this can trigger episodes either in the mothers or children, with evidence of influence in both directions. Depressed mothers have been shown to relate differentially to their children, with more negativity to dysfunctional children than to a normally functioning ones. In contrast, nondepressed mothers are less likely to acknowledge dysfunctionality in their children and therefore treat them more similarly. Thus, characteristics of both the parents and the children play a part.

Correlates of Parental Disorder

For children, as for their parents, psychiatric disorder is a consequence of adversity, but it also leads to the perpetuation of adversity. The most common family adversity associated with disorder is interpersonal, involving both marital conflict and conflict with children, although poverty and social isolation are also implicated. Susceptible individuals are thus capable of either being the agents of stress for close others or being the recipients and carriers of such stress and passing it on to others. Where both parents in a family have disorder, the accumulation of stressors for all family members is likely to be great. Thus negative family interaction can be a function both of parents' vulnerability characteristics (e.g., poor interpersonal relating, poor support, and low self-esteem) and of symptoms of psychological disorder (e.g., depression, anxiety, and antisocial behavior). Thus, intergenerational studies with mothers and young adult offspring interviewed independently show an important role for neglect/abuse in the offspring in disorder but with separate causal linkages from maternal depression and maternal vulnerability. It is of note that the perpetrators of physical abuse in children are rather more likely to be the fathers or surrogate fathers, despite the intergenerational impact observed for mothers. Thus, the mothers' role can be both a carrier and agent in stress, with both being conveyed to the offspring.

The reasons for familial stress patterning are therefore varied. The congregation of stressors within family groupings is due to both common and individual

influences. Negative interpersonal functioning, poor parenting, poor coping, and family disadvantage are key factors. Biological factors involving the genetic transmission of risk and the physical impacts of stress are increasingly being discovered as contributing to final explanatory models.

Further Reading

- Belsky, J. and Vondra, J. (1989). The determinants of parenting: a process model. In: Cicchetti, D. & Carlson, V. (eds.) *Child maltreatment – theory and research on the causes and consequences of child abuse and neglect*, pp. 153–203. Cambridge, UK: Cambridge University Press.
- Bifulco, A. and Moran, P. M. (1998). *Wednesday's child: research into women's experience of neglect and abuse in childhood and adult depression*. London: Routledge.
- Bifulco, A., Moran, P. M., Ball, C., et al. (2002). Childhood adversity, parental vulnerability and disorder: examining inter-generational transmission of risk. *Journal of Child Psychology and Psychiatry* 43, 1075–1086.
- Brown, G. W. and Harris, T. O. (1978). *Social origins of depression*. London: Tavistock.
- Caspi, A. (2002). Role of genotype in the cycle of violence in maltreated children. *Science* 297, 851–854.
- Eley, T. C. and Stevenson, J. (2000). Specific life events and chronic experiences differentially associated with depression and anxiety in young twins. *Journal of Abnormal Child Psychology* 28, 383–394.
- Goodyer, I. (1990). *Life experiences, development and childhood psychopathology*. Chichester, UK: John Wiley.
- Hammen, C. (1991). *Depression runs in families: the social context of risk and resilience in children of depressed mothers*. New York: Spring Verlag.
- Harris, J. R. (1998). *The nurture assumption: Why children turn out the way they do*. New York: The Free Press.
- Kendler, K. S., Kessler, R. C., Walters, E. E., et al. (1995). Stressful life events, genetic liability, and onset of an episode of major depression in women. *American Journal of Psychiatry* 152, 833–842.
- Plomin, R. (1994). *Genetics and experience: the interplay between nature and nurture*. Thousand Oaks, CA: Sage Publications.
- Reiss, D., Plomin, R., Hetherington, E. M., et al. (1994). The separate worlds of teenage siblings: an introduction to the study of the nonshared environment and adolescent development. In: Hetherington, E. M. & Reiss, D. (eds.) *Separate social worlds of siblings: the impact of nonshared environment on development*, pp. 63–109. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rutter, M. (1989). Intergenerational continuities and discontinuities in serious parenting difficulties. In: Cicchetti, D. & Carlson, V. (eds.) *Research on the consequences of child maltreatment*, pp. 317–348. New York: Cambridge University Press.

Family Therapy

B Jalali

West Los Angeles Veterans Administration Medical Center and UCLA School of Medicine, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Stress and the Family

Theoretical Models of Family Therapy

What Family Therapy Can Achieve

Indications for Family Therapy

Glossary

Differentiation of self The degree of emotional separation between a person and his or her family of origin. More differentiated individuals have a better awareness of the influence of their emotions, behavior, and actions and are less dependent on approval of others for their decisions.

Double bind

Communication characteristics of a system that produces conflicting definitions of incongruous messages between members of a family, leading to stress and anxiety.

Enmeshment/disengagement

Enmeshment is a type of family relationship in which the family members are fused and are unable to define their roles and boundaries. Subsystems are not well defined, and the hierarchy of executive functions is unclear. Disengagement refers to a lack of relatedness and binding between family members, who appear uninvolved and distant.

Homeostasis

Steady state of any living system that is regulated and maintained as it faces the external environment. If the system faces stress and breaks down, it attempts to repair itself so that a state of constancy and balance is maintained.

Identified patient

The person who bears the symptoms and is the reason for the family to seek treatment. However, from the family system

	point of view, individual psychopathology is maintained and or amplified by the family, if not a consequence of family pathology.
<i>Pseudo-mutuality</i>	A facade of harmony and pretending to share same behavior, feelings, and attitude in a family at the expense of lack of individuation in its members. Anything that threatens to deviate from the sameness of the structure and roles provokes a great deal of anxiety.
<i>Rubber fence</i>	A specific state of relationship between the family system, its members, and the social environment. It protects the family from any intrusion or departure from its norm and incorporates items that fit and rejects those that appear noncomplimentary.
<i>Scapegoat</i>	A person who is recruited by the family system to become the problem, instead of the family system attempting to resolve its conflicts.
<i>Schism and skew</i>	Schism is a pattern of marital relationship in which the members of a couple fail to reach an agreement on role reciprocity and therefore undercut or coerce each other. Skew in a marital relationship means that one member of the couple has significant pathology, but the other goes along with it, and all conflicts, while present, are hidden.
<i>Triangle</i>	When conflict or anxiety rises in a dyadic or a two-person relationship and a third person is included to diffuse the conflict. The pathological form of triangle is a rigid triangle wherein conflicts are never resolved in the dyad, and the third person is put in a position of either conflict of loyalty or alliance with one member against the other.

Stress and the Family

In a living system, nothing is static. Different psychological, biological, and social forces are continuously impinging on the system to synchronize and change. Stress on the family system disturbs its homeostatic equilibrium. In response, the system must mobilize its internal and external resources, become flexible enough to produce alternative modes of adaptation, be able to exchange roles and functions across its subsystems, reorganize to meet the demands of the stress, and redefine itself after adjustment to the stress.

The family system must be assessed as a whole, with a set of functions, relationships, and roles, which are continuously changing as the social structure

surrounding the family system is changing. Therefore, in evaluating the effects of stress on the family system, one must be aware of the cultural, social, and historical makeup of the family, as it influences the experience of, meaning of, and response to stress.

The types of stress on the family are mainly the following:

1. The stress of transitions in the family's developmental stages (becoming a couple, birth of first child, children transitioning into adolescence, first child leaving home, last child leaving home, retirement and empty nest, sudden illness and death of a family member, chronic illness in a family member, threat of dissolution of the family such as separation and divorce, remarriage and blended family, illness or death of grandparents, etc.)
2. Outer stressors, such as loss of job, a move to a new location, changing schools, economic loss, catastrophic events such as flood and earthquake, and crime.

Regardless of the type of stress, the family system must first recognize the stress, then attempt to adjust and reorganize in order to cope with it. Adjustment and reorganization include changes in the structure, function, and roles of the system, which will be new and unfamiliar to the members and may bring anxiety and uncertainty with them.

If the family characteristics and functioning were problematic and conflictual to begin with, including poor communication skills, unclear boundary definition, poor parental skills, and marital conflicts, then the family structure may skew further in the face of stress and will have difficulty in reorganizing and readjusting. This may result in either poor adaptation or development of symptoms in vulnerable members (biologically or constitutionally), such as substance abuse, depression, suicide attempt, school or behavioral problems, eating disorder, infidelity, and threats of separation or divorce.

These families may mobilize pathological coping mechanisms, such as denial, blame, scapegoating, detouring of conflicts, and developing rigid triangulations, which may result in an increase of perceived stress in its members. Although denial can serve as a protective defense mechanism at first, if it persists it will impede the family's coping, because the family will resist change and continue to limp along as if no modifications were required. However, their established patterns will not be sufficient, and the stress may plunge the family into a state of crisis. The functioning of the family system will halt, and the system will be incapacitated, in a state of disequilibrium, and unable to function.

Lack of flexibility in gender roles and division of labor can cause rigidity and restrictions in the system, which will reduce resources necessary for coping. Families that are isolated from their social network and community also have restricted coping resources. The family system can recover from the state of crisis with the help of family therapy.

Family members can grow from the experience of stress and learn how to manage future stresses. Family therapy can help them to identify the stress, become more cohesive, communicate better among themselves and with the outer environment, recruit social and community resources, respect interpersonal boundaries, and avoid blame.

Theoretical Models of Family Therapy

Overview

Family therapy is a type of psychotherapy in which the unit of treatment is the family (whether nuclear or extended), and not the individual. It assumes that problems and symptoms emanate from unhealthy interactions in the family system and not solely from the psychological or subconscious problems of the individual. The theory uses circular rather than linear thinking; each individual is part of the larger group, with continuous interaction and feedback processes between the two.

In general, family therapy is based on family system theory. It examines family interactions, interpersonal relationships, and communications and aims to change unhealthy patterns. This theoretical model was introduced in late 1940s and early 1950s, emanating from the study of schizophrenic patients and the link between exaggerations of their symptoms and their family interactions. Family system theory assumes that the family system is in a constant state of change and balance, and it attempts to achieve homeostasis and adaptation to change by mobilizing alternative interactions. When the system can not adapt successfully due to stress, rigidity of patterns, and/or psychological features, problems and symptoms ensue in different members. Consequently, adaptation is inappropriate, and a state of dysfunctional homeostasis settles in.

The early family therapists hypothesized that whenever the family system's homeostasis is threatened, symptomatic behavior in one member may become a way to preserve the family's equilibrium. Therefore, the concepts of scapegoating and the identified patient were introduced.

Early family therapists also hypothesized that many of the preceding phenomena occur through a feedback process and nonverbal communications, through

sequence of interactions. At first they attempted to examine the communication of schizophrenic patients and their families, and concluded that there are series of contradictory messages transmitted to the patient. The concept of double bind was introduced, which was considered the essential ingredient in the child-parent interaction, specifically the mother (the concept of the schizophrenogenic mother), in which a family member receives two contradictory messages that he or she can not obey or disobey. Later the concept of double bind was extended from a two-person system to a three-person system, as attention was drawn to fathers as well as mothers. Interactions in the marital relationship were found to be dysfunctional and pathological, and the concepts of schism and skew were introduced. They highlighted patterns of hostility, undercutting, domination, passivity, and distortion.

Other family therapists studied communication patterns and family roles in schizophrenia and identified pseudomutuality as a feature of families with a schizophrenic member. These families appear cohesive on the surface but are in deep conflict underneath. They may also erect a boundary around the family that will extrude any input that will not fit their norm (a rubber fence) and will only allow input that will fit their norm.

These theories were later applied to other family systems with a member identified as carrying a psychiatric diagnosis, such as an eating disorder, depression, or substance abuse. Although all family therapists agree on the basic principle of how a dysfunctional family can lead to stressing a vulnerable member(s) so that symptoms develop and are maintained and amplified, the techniques that have been developed to treat families are quite different from each other. However, each strives to establish a less stressful state in the family.

The length of a course of family therapy is usually shorter than individual therapy, ranging from several sessions to several months. On the whole, the therapist attempts to change the dysfunctional structure and functioning of the system, rather than address changes in the individual member. When therapists attend to symptoms, they are treated as part of a complex system of interaction and feedback processes in which symptoms have become faulty resolutions for other family problems. The therapist strives for a change in rigid behavioral patterns, increased communication, healthier interactions, and improved problem solving among all members.

Examples of Theoretical Models

Major theoretical orientations of family therapy include psychoanalytic (Ackerman), systems (Bowen), structural (Minuchin), strategic (Haley, Madanes),

experiential (Whitaker), communicational (Satir, Bateson, Watzlawick), object relations (Framo, Nagy), and a few others, such as constructivist and Milan's system of counter paradoxical family therapy.

Psychoanalytic family therapy Ackerman theorized that a healthy family is in a state of constant homeostatic balance, which is dynamic and changing as the family grows in its developmental stages and as new stresses occur to which the family needs to adapt. The dysfunctional family cannot adapt to the required changes, and therefore the various family functions become faulty and lead to problems and symptoms in its members and possibly to disintegration of the family. Therefore, change can occur through implementing improved problem solving, learning healthier communication, and diminishing the scapegoating of its members.

Family Systems Therapy

Bowen developed family systems therapy, theorizing that the basis of human relationships is not a two-person system but a three-person one. He stated that when conflicts and anxiety develop in a relationship, invariably a third person becomes involved who will side with one of the other two. In this threesome, two individuals are close, and the third one is on the outside. This is the basis for the concept of triangles in the family. The triangle may appear to diffuse the anxiety, but conflicts remain unresolved.

Triangles are constantly shifting and unstable and have a tendency to interlock with other triangular relationships. Individuals learn from their family of origin how to operate in triangles. When the adaptive level of functioning in a family is pathological, the individuals in that family are more likely to resolve anxiety with triangulation and are less differentiated. The concept of triangles and triangulating was later applied to larger systems such as organizations and small groups.

Bowen stated that each family is part of an undifferentiated ego mass, and each member strives to differentiate from it, to develop one's own sense of self, and to detriangulate oneself. He also stated that patterns of relationships and handling stress are a part of a process that is transmitted across generations. When getting caught in triangulated relationships with their family becomes too intense and stressful, members may try to distance themselves emotionally or physically, or they may try to cut themselves off from the family. However, similar patterns will be recreated and repeated in other relationships that the individual develops.

This model also introduced the concept of family projection process, which means that the

emotional processes are transmitted, for example, to a child, who as a result may develop difficulties in differentiation of self.

The family system theory of Bowen states that historical patterns are factors in symptom production as well as change; therefore, the main tool of treatment is using a genogram or diagram of the family's three generations in detail, demonstrating sets of relationships, conflict areas, and triangular relationships, and then exploring methods of unlocking and changing them. Bowen encouraged families to resolve issues with their families of origin and attempt to achieve separation and differentiation from them.

The goal of treatment is understanding the three-generation system of involvement in the symptoms. The main therapeutic technique is helping and teaching the family to differentiate and detriangulate themselves from their family of origin ego mass, which will lead to healthier relationships on the whole.

Structural family therapy Structural family therapy (Minuchin) presents a clear model of what is considered a normal family system. This is a system with a clear hierarchy of power, clear boundaries between members and the family subsystems (subsystem of parents and children), no cross-generational coalition, and clear cross-generational boundaries.

Minuchin presented a model of pathological family enmeshment in which boundaries are blurred, the hierarchy of power is unclear, cross-generational alliances and coalitions are present, and executive functions of the family are hampered. He stated that a dysfunctional family cannot adapt to developmental and extrafamilial stresses and it therefore recruits a vulnerable member who becomes symptomatic.

Minuchin applied his methods to severe problems of childhood and later to eating disorders. Structural family therapy's major contribution to the field was in the treatment of psychosomatic illnesses such as asthma and illnesses in which an emotional component was detrimental to control and compliance, such as diabetes. Research with diabetic and asthmatic patients done by structural family therapists clearly showed that family stress levels correlated with induction of an asthmatic attack or a rise in blood glucose level of the patient, and when the stress subsided, the asthma subsided and the blood glucose level returned to normal. Conflicts in the family correlated with exacerbation of medical symptoms, which was supported by laboratory data. Minuchin demonstrated that family therapy and reducing the family's conflicts and stress can indeed reduce the number of emergency visits for both the asthmatic and the diabetic family member.

Structural family therapy was also applied to patients with eating disorders, ranging from bulimia to anorexia, and to patterns of lack of conflict resolution and rigidity that were dysfunctional and therefore maintained and amplified the preceding symptoms. Once these patterns were addressed and changed, the eating disorder improved. Techniques used in this theoretical framework include joining, establishing clear subsystems, expanding the alternative ways of problem solving, strengthening the boundaries between generations, relabeling symptoms, creating intensity leading to change, increasing flexibility, modifying dysfunctional interactions, clarifying the hierarchy of power, and restructuring.

Strategic family therapy Strategic family therapy and the problem-oriented approach (Haley) also adhere to a systemic model of family interaction, stating that symptoms maintain homeostasis. Strategic family therapy studies the sequence of interactions among family members that leads to symptoms and then attempts to change this sequence to a better-functioning one and a healthier hierarchy. The symptom is also a solution, although an unsatisfactory one. It is both supported and opposed by other members' behavior. Therefore, the therapist has to identify the sequence of interactions that both reinforces and exacerbates it.

In this model, a historical perspective of the family is not studied, nor is the larger context or where the meaning of the symptom is embedded. The focus is on the presenting problem or symptom and not on the family. However, in order for the problem to improve, the organizational sequence that maintains the problem must change. Change is accomplished by prescribing strategies, positive reframing, changing dyadic relationships that are involved with the symptom, reversal of roles, and use of paradoxical assignments and interventions. Symptoms are changed through changes in hierarchy, coalitions, and communication. This model clearly adheres to a briefer period of therapy.

Experiential family therapy The experiential model (Whitaker) focuses on process and growth in therapy. The family is challenged by being offered a variety of alternative ways of interacting, including absurd ones, as a method to unstick the family. Change is induced by creating new experiences, personal growth, and marking boundaries between generations. Individual and family creativity and growth are promoted. The problems or symptoms are viewed as systems problems. In this model, an attempt is made to join the system in the beginning, but later in treatment the therapist distances and becomes an outside

consultant to the system, encouraging the family to take the initiative in making changes.

Counter paradoxical family therapy The counter paradox model of the Milan group suggests that in a dysfunctional family, the members' response to each other is often paradoxical and illogical. This model attempts to counter the paradox by the use of a team of therapists, circular questioning (to gather more information), positive connotations, remaining neutral, increasing the time interval between the sessions, prescribing family rituals, and systemic prescriptions.

What Family Therapy Can Achieve

Family therapy can achieve the following:

- Help the family to understand the patterns of their interactions and relationships as a whole, instead of focusing on one member as the scapegoat.
- Identify conflicts and defects within the family structure and help members to develop new methods of resolving problems.
- Identify internal and external resources and strengths.
- Provide support, education, and guidance in handling inner and outer stresses.

The main rule of thumb in conducting family therapy is that all members who are identified as key members must participate; otherwise, it will not be effective.

Indications for Family Therapy

Indications that family therapy is needed include the following:

- Substance abuse in one or more family member
- Chronic psychosomatic disorder or other chronic medical illness in one member
- Physical or emotional abuse in the family
- Marital conflict
- Poor parental skills
- Eating disorder in one member
- Diagnosis of major psychiatric disorder in one member such as schizophrenia or bipolar disorder
- Depression in the family
- Any behavioral problems or psychological symptoms in a child or adolescent
- A dysfunctional adult offspring in the family
- A traumatic event experienced by the family

See Also the Following Articles

Depression and Manic-Depressive Illness; Eating Disorders and Stress; Familial Patterns of Stress; Schizophrenia.

Further Reading

- Bateson, G. (1977). *The steps to an ecology of mind*. New York: Ballantine Books.
- Boscolo, L., Cecchin, G., Hoffman, L. and Penn, P. (eds.) (1987). *Milan systemic family therapy: conversations in theory and practice*. New York: Basic Books.
- Broderick, C. B. (1993). *Understanding family process: basics of family system theory*. California: Sage Publishers.
- Bumberry, W. (1998). *Reshaping family relationships: the symbolic therapy of Carl Whitaker*. New York: Brunner Mazel.
- Framo, J. L. (1992). *Family-of-origin therapy. An inter-generational approach*. New York: Brunner Mazel.
- Gurman, A. S. (1981). *Handbook of family therapy*. New York: Brunner Mazel.
- Haley, J. and Richeport-Haley, M. (2003). *The art of strategic therapy*. New York: Brunner-Routledge.
- Hoffman, L. (1981). *Foundations of family therapy*. New York: Basic Books Publishers.
- Lidz, T., Fleck, S. and Cornelison, A. R. (1965). *Schizophrenia and the family*. New York: International Press.
- Kerr, M. I. and Bowen, M. (1988). *Family evaluation*. New York: W.W. Norton Press.
- McGoldrick, M., Gerson, R. and Shellenberger, S. (1999). *Genograms: assessment and interventions*. New York: W.W. Norton.
- McGoldrick, M., Giordano, J. and Garcia-Preto, N. (eds.) (2006). *Ethnicity and family therapy*. New York: Guilford Press.
- Minuchin, S., Rosman, B. L. and Baker, L. (1978). *Psychosomatic families*. Cambridge, MA: Harvard University Press.
- Minuchin, S., Wai-Yung and Simaon, G. M. (1996). *Mastering family therapy: journey of growth and transformation*. New York: John Wiley and Sons.
- Minuchin, S., et al. (2006). *Assessing families and couples: from symptoms to system*. New York: Allyn and Bacon.
- Nichols, M. P. (1984). *Family therapy: concepts and methods*. New York: Gardner Press.
- Wynn, E., Ryckoff, J. M., Day, J. and Hirsch, S. I. (1958). Pseudomutuality in the family treatment of schizophrenia. *Psychiatry*, **21**: 205–220.
- Zeig, J. L. and Haley, J. (2001). *Changing directives: the strategic therapy of Jay Haley*. Phoenix, AZ: Milton Erickson Foundations Press.

Fatigue and Stress

A Appels

University of Maastricht, Maastricht, Netherlands

W J Kop

Uniformed Services University of the Health Sciences, Bethesda, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A Appels, volume 2, pp 108–110,

© 2000, Elsevier Inc.

Psychiatric Approach

Organizational and Occupational Health Approach

Psychophysiological Approach

Psychosomatic Approach

Conclusion

Glossary

Burnout

A state characterized by emotional exhaustion that may occur after exposure to prolonged stress.

Hypothalamic-pituitary-adrenal (HPA) axis

Part of the neurohormonal stress system. Both increased and decreased activity of the HPA axis can be observed, depending on the nature of the disorder.

Neurasthenia

A state characterized by fatigue, impaired cognitive functioning, and increased emotional sensitivity.

Vital exhaustion

A state characterized by unusual fatigue and loss of energy, increased irritability, and feelings of demoralization; often precedes cardiac events. In the original conceptualization of the exhaustion construct, the term vital was included to reflect the far-reaching consequences of this condition on daily life function (similar to vital depression).

Fatigue is a common symptom that can substantially interfere with daily functioning and quality of life. A large literature indicates that episodes of extreme fatigue are precipitated and maintained by emotional distress. A substantial number of individuals (5–20% of the general population) suffer from persistent fatigue that interferes with routine daily life activities, and fatigue is one of the most common presenting complaints in primary care. Prevalence estimates of fatigue depend on the measurement instruments, cut-off points for presence versus absence, and the persistence of the complaint. Patients generally regard

fatigue as important because it is disabling, whereas physicians often do not place a strong emphasis on fatigue because it is diagnostically nonspecific for most if not all diseases.

Several theories have been developed to conceptualize and investigate fatigue. These theories describe fatigue as a pollution of the *milieu interieur*, as an energy problem, as a breakdown in nervous system functioning, or as an evaluative motivational assessment of an individual's psychophysiological condition in relation to exogenous demands, resulting in a decision to either continue or discontinue an activity. These theories have strengths and limitations, and reflect the way in which scientists organize and operationalize their approaches to investigating a coherent set of symptoms. There are overlaps and distinctions between the approaches to the origins and treatment of fatigue, depending on the investigator's fields of interest and scientific background. In the following sections we provide a selective review of these approaches.

Psychiatric Approach

Since the late 1800s, much debate has arisen addressing the question of whether fatigue states should be distinguished from depressive mood disorders. There is little doubt that depressed individuals often feel tired and that fatigue is one of the diagnostic symptoms for depression. However, not all individuals who feel depleted of energy suffer from mood disturbances, lack of self-esteem, or guilt feelings. Therefore, psychiatric constructs have been developed to describe a condition that is mainly characterized by profound fatigability. In 1869, George M. Beard introduced the concept neurasthenia to describe an organically based disorder characterized by fatigability of the body and mind, largely resulting from environmental factors. Neurasthenia was primarily treated by rest and reportedly common among well-educated and professional individuals. Other concepts that have been developed to describe fatigue states include effort syndrome, neurocirculatory asthenia, or postinfectious fatigue syndrome.

More recently, the chronic fatigue syndrome (CFS), myalgic encephalomyelitis (ME), and also posttraumatic stress disorder (PTSD) have been introduced as conditions in which fatigue and distress play a primary role. CFS is an operationally defined syndrome characterized by a minimum of 6 months of severe physical and mental fatigue and fatigability made worse by minor exertion. Because of the similarities among the prior neurasthenia construct, CFS, and ME, it has been questioned whether these more recent

concepts are mainly "old wine in new bottles." PTSD describes the sequels of exposure to severe psychological distress, such as occur during major disasters, war, and other life-threatening events. Increased fatigue and fatigability belong to the major characteristics of PTSD. Detailed research has revealed that some symptoms are relatively specific for PTSD and CFS, justifying their heuristic value. There is, nonetheless, substantial overlap among the symptoms that constitute these postulated syndromes, which may suggest that stress is a common factor.

Psychiatric classifications and diagnoses were originally based on the presumed pathophysiological origins of a complaint. In contrast, contemporary diagnostic classifications, such as the *Diagnostic and Statistical Manual of Mental Disorders* (4th edn.; DSM-IV), are predominantly based on the description and classification of symptomatology. DSM-IV includes PTSD, but does not include any of the other fatigue-related constructs or syndromes. DSM-IV lists decreased energy, tiredness, and fatigue among the symptoms of depressive mood disorders or dysthymia. Thus, the most commonly used classification system for mental disorders does not make a distinction between depression and other fatigue states.

Organizational and Occupational Health Approach

In occupational medicine and occupational psychology, fatigue is mainly approached as the result of an imbalance between demand and supply or as a breakdown in adaptation to stress. During the last decades, much attention has been given to burnout as the end point of long-lasting work-related stress. Burnout is a negative state of physical, emotional, and mental exhaustion that is the result of a gradual process of disillusionment. The concept of burnout was first used by Freudenberger in the mid-1970s to describe a common psychological characteristic observed in health-care professionals. Freudenberger observed that many of the volunteers with whom he was working experienced a gradual emotional depletion and a loss of motivation and commitment. Burnout has three main components: emotional exhaustion, depersonalization, and reduced personal accomplishment.

There has been substantial debate about the optimal definition of the burnout construct. However, this debate has not prevented fruitful empirical research, including Maslach's successful development of a reliable scale to measure burnout. This scale assesses the three dimensions of burnout (emotional exhaustion, depersonalization, and reduced personal accomplishment) and has become the gold standard

for the assessment of burnout. Empirical research has shown that the emotional exhaustion component is related to depression, whereas the relationships between depersonalization and personal accomplishment with depression are less strong.

The burnout literature is more concerned with situational antecedents than with individual differences as the determinants of burnout, probably because of the occupational health origins of this construct. The extension of the burnout concept to nonoccupational domains (e.g., feelings of emotional exhaustion caused by marital conflicts) has been controversial, especially because the concepts of depersonalization and reduced personal accomplishment are less applicable outside the realm of occupational settings. Therefore, the contributions of research on burnout to fatigue lie primarily in the identification of social and occupational determinants.

Psychophysiological Approach

Psychophysiological approaches address behaviors, emotions, and cognitions that correspond with physiological reactions to normal and abnormal, internal or external stimuli. For example, psychophysiologicals and other behavioral scientists investigate the influence of fatigue on inflammation, as well as how the mental state is influenced by inflammation. The stress system receives much attention in these types of psychophysiological research designs.

The autonomic nervous system and neurohormonal processes play a crucial role in the human stress response. The autonomic nervous system involves two branches: the sympathetic nervous system (involved in the acute stress response) and the parasympathetic nervous system (which is commonly more active in periods of rest and, in general, counterbalances the sympathetic nervous system). The neurohormonal stress response is complex and involves the sympathetic-adrenomedullary (SAM) system and the HPA system. The stimulation of the SAM system is characterized by increased secretion of the catecholamines epinephrine and norepinephrine, as well as increased heart rate and blood pressure, sweating, palpitations and other vegetative symptoms (i.e., related to autonomic nervous system activation). Most of the autonomic nervous system and neurohormonal responses in response to stress exposure are normal, enabling the organism to adapt to environmental changes.

The HPA axis response displays important divergent concomitants of increased versus decreased activity. Increased HPA axis activity is observed in response to acute stress; increased HPA axis activity characterized by increased secretion of cortisol is also well documented among individuals with

melancholic depression, panic disorder, central obesity, and Cushing's syndrome. These elevated cortisol levels may coincide with the suppression of inflammation. In contrast, decreased HPA axis activity has been observed in chronic stress, persistent fatigue, atypical depression (i.e., hyperphagia and hypersomnia), and nicotine withdrawal. These conditions are associated with a decreased production of cortisol and purportedly with the activation of immune-mediated inflammation.

During inflammation, the body produces cytokines, which play a pivotal signaling role in immune system activation. Some of these cytokines reach the brain or elicit the release of cytokines in the brain, resulting in what has been labeled sickness behavior, characterized by locomotor retardation, general malaise, anorexia, and inhibition of sexual behavior. Sickness behavior is part of the defensive response to infection or inflammation. Evidence consistently demonstrates that fatigue and inflammation are interrelated. Thus, there is probably a bidirectional or circular association between inflammation and fatigue, which is mediated in part by the central and autonomic nervous systems. The psychophysiological approach makes a unique contribution to the study of fatigue by investigating which symptomatic and physiological reactions belong to normal health-protecting behaviors and which reactions are markers of a breakdown in adaptation to environmental challenges.

Psychosomatic Approach

In psychosomatic medicine, fatigue is approached in several ways, depending on the nature of the coinciding medical disorder. For example, among patients with cancer, fatigue is considered as a (still poorly understood) consequence of the disease and its treatment. Programs are mainly directed at coping with the fatigue. In patients with multiple sclerosis, fatigue is particularly disabling symptom, distinct from depressed mood or physical weakness.

In coronary artery disease, fatigue and exhaustion are of particular interest because it is well established that exhaustion (i.e., feelings of unusual fatigue, loss of energy, and increased irritability and demoralization) is an important precursor of myocardial infarction and sudden cardiac death. Exhaustion and depression also predict recurrent cardiac events among high-risk populations, and more research is needed to further document the divergent validity of these two partially overlapping constructs.

Based on the predictive value of exhaustion and depression for future adverse cardiovascular outcomes, two different intervention approaches have

been investigated recently. Many scientists interpret the premonitory symptoms of extreme fatigue as manifestations of a clinical or subsyndromal depression; consequently, Beck's cognitive therapy has been used to treat coronary disease patients with depressive symptomatology. Other investigators have approached the same symptoms as manifestations of a sustained state of exhaustion resulting from long-term exposure to uncontrollable emotional distress. This perspective emphasizes the similarity between exhaustion and the psychological consequences of decreased HPA activity following chronic stress. Interventions targeting exhaustion use group therapy, including relaxation and reducing stressors that can lead to exhaustion. The effects of both types of interventions on the risk of new coronary events have been tested in well-designed, randomized, controlled intervention trials. Neither of the two strategies succeeded in demonstrating that a reduction of depression or exhaustion reduced the risk of a new coronary event. However, the interventions were successful in subsamples (determined *a posteriori*). It is reasonable to assume that future trials will profit from the lessons learned in these pivotal studies. Such trials may demonstrate that a behavioral treatment of fatigue and depressive symptomatology will not only improve quality of life but also reduce the risk of recurrent cardiac events.

Conclusion

Fatigue is a ubiquitous phenomenon. Less than 10% of patients presenting with fatigue in primary care present with a disease that plays a direct causal role in this symptom. The law of parsimony has accompanied fatigue research and will accompany future work. It has been debated whether concepts such as neurasthenia, burnout, and vital exhaustion are needed and whether the procrustean bed of DSM-IV does more harm than good in stress research. Each of the aforementioned approaches has contributed to our obtaining more insight into the origins of fatigue, its biological and physiological correlates, and the possibilities and limitations of helping individuals who suffer from unusual fatigue. The domain of fatigue research has unique epistemological problems, which require further scientific attention. More research is needed on the distinction between fatigue as a health-protecting factor and fatigue as a marker of overtaxing of the body. The precipitating and maintaining factors of chronic fatigue conditions are complicated because some of the original etiological factors may have become undetectable (e.g., the long-term consequences of viral infections). It will be difficult to answer the question of whether chronic fatigue can be

best approached as a form of psychological adaptation to prolonged and uncontrollable environmental challenges or whether fatigue is a manifestation of depression and/or a dysfunctional neurohormonal system. Future multidisciplinary research will continue to add to our understanding of the role of stress and other psychological factors in the origins, mental and somatic consequences, and treatment of fatigue, particularly in conditions in which fatigue is a known predictor of adverse medical prognosis.

See Also the Following Articles

Burnout; Chronic Fatigue Syndrome; Depression Models; Sleep Loss, Jet Lag, and Shift Work; Sleep, Sleep Disorders, and Stress; Workplace Stress; Night Shift-work; Posttraumatic Stress Disorder – Clinical; Posttraumatic Stress Disorder – Neurobiological basis for.

Further Reading

- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th edn.). Washington DC: American Psychiatric Association.
- Appels, A. (1990). Mental precursors of myocardial infarction. *British Journal of Psychiatry* 156, 465–471.
- Appels, A. (1997). Depression and coronary heart disease: observations and questions. *Journal of Psychosomatic Research* 43(5), 443–452.
- Appels, A., Bar, F., van der Pol, G., et al. (2005). Effects of treating exhaustion in angioplasty patients on new coronary events: results of the Randomized Exhaustion Intervention Trial (EXIT). *Psychosomatic Medicine* 67(2), 217–223.
- Berkman, L. F., Blumenthal, J., Burg, M., et al. (2003). Effects of treating depression and low perceived social support on clinical events after myocardial infarction: the Enhancing Recovery in Coronary Heart Disease Patients (ENRICH) Randomized Trial. *Journal of the American Medical Association* 289(23), 3106–3116.
- Chrousos, G. P. (1995). The hypothalamic-pituitary-adrenal axis and immune-mediated inflammation. *New England Journal of Medicine* 332(20), 1351–1362.
- Kop, W. J. (1999). Chronic and acute psychological risk factors for clinical manifestations of coronary artery disease. *Psychosomatic Medicine* 61(4), 476–487.
- Kop, W. J. (2003). The integration of cardiovascular behavioral medicine and psychoneuroimmunology: new developments based on converging research fields. *Brain, Behavior & Immunity* 17(4), 233–237.
- Krupp, L. B. (2003). Fatigue in multiple sclerosis: definition, pathophysiology and treatment. *CNS Drugs* 17(4), 225–234.
- Maslach, C. and Jackson, S. E. (1986). *Maslach burn-out inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Sharpe, M. and Wilks, D. (2002). Fatigue. *British Medical Journal* 325(7362), 480–483.

Siegrist, J. (1991). Contributions of sociology to the prediction of heart disease and their implications for public health. *European Journal of Public Health* 1(1), 10–29.

Skapinakis, P., Lewis, G. and Mavreas, V. (2003). Unexplained fatigue syndromes in a multinational primary care sample: specificity of definition and prevalence and

distinctiveness from depression and generalized anxiety. *American Journal of Psychiatry* 160(4), 785–787.

Wagner, L. I. and Cella, D. (2004). Fatigue and cancer: causes, prevalence and treatment approaches. *British Journal of Cancer* 91(5), 822–828.

Wessely, S. (1990). Old wine in new bottles: neurasthenia and “ME.” *Psychology and Medicine* 20(1), 35–53.

Fear

A Öhman

Karolinska Institutet, Stockholm, Sweden

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A Öhman, volume 2, pp 111–115, © 2000, Elsevier Inc.

Components of Fear

Measures of Fear

Fear Stimuli

Fear Learning

Pathological Fear: Phobias

The Neurophysiology of Fear

Glossary

<i>Agoraphobia</i>	An intense fear and avoidance of situations in which one would be helplessly exposed in case of a panic attack.
<i>Amygdala</i>	A collection of interconnected nuclei in the anterior medial temporal lobe, which is the hub of the brain network controlling the activation of fear responses.
<i>Autonomic nervous system</i>	A part of the peripheral nervous system concerned with the metabolic housekeeping of organisms. In general, the activation of its sympathetic branch results in the expenditure of energy and the activation of the parasympathetic branch results in the restoration of energy.
<i>Fear</i>	The emotional state associated with attempts to cope with threatening events.
<i>Social phobia</i>	An intense fear and avoidance of being socially evaluated or scrutinized.
<i>Specific phobia</i>	An intense fear and avoidance of specific objects or situations.

Fear is an activated, aversive emotional state that serves to motivate attempts to cope with events that provide threats to the survival or well-being of

organisms. The coping attempts are typically centered on defensive behaviors such as immobility (freezing), escape, and attack.

Components of Fear

A fear response comprises several partially independent components, such as subjective feelings (accessible through verbal reports), peripheral physiological responses, and overt behavior. In humans, the phenomenological quality of fear is best described as an aversive urge to get out of the situation. This is a familiar feeling for every human being and a frequent target for artistic representation.

Fear is closely associated with the activation of the autonomic nervous system. Depending on situational constraints, the direction of this activation may differ. If the threat is not imminent and appears stationary, the typical response is one of freezing or immobility. This is associated with enhanced attentiveness toward the environment and the potential threat stimulus and with a vagally mediated deceleration of the heart. If the threat is imminent or approaching, there is a pervasive mobilization of the sympathetic branch of the autonomic nervous system, including heart rate acceleration and increases in blood pressure and circulating catecholamines (primarily epinephrine) from the adrenal medulla. These responses lay a metabolic foundation for taxing overt reactions of flight or fight.

In terms of behavior, there is an important distinction between expressive and instrumental overt acts. Expressive behavior includes automatic tendencies to withdraw from the threat and a typical facial expression of fear. The latter is composed of elevated eyebrows, wide open eyes, and a mouth that is either slightly opened or shut with depressed mouth corners. Instrumental behavior primarily concerns escape from and avoidance of the fear stimulus.

Measures of Fear

Measures of fear can be readily derived from the various components of fear. The intensity of the subjective component of fear can be directly assessed through ratings that may be anchored, for example, at a zero level of no fear at all, with a maximal level of fear corresponding to the most intense fear ever experienced by the subject.

Many fear indices have been derived from effectors innervated by the autonomic nervous system. Some of these measures, such as heart rate deceleration and skin conductance responses, are primarily related to the increased attentiveness associated with the initial stages of fear. Other measures, such as increases in heart rate or blood pressure, reflect the sympathetic mobilization that supports active coping attempts. Endocrine indices are available from the adrenal medulla and the adrenal cortical hormones. Measures among the former, such as circulating epinephrine, are often assumed to index successful coping, whereas measures from the latter, such as cortisol, are regarded as related to failing coping attempts.

Many behavioral measures have been used to assess fear. Typically they focus on avoidance behavior. For example, in a standard behavioral avoidance test to assess specific human fears, subjects are encouraged to approach their feared object as close as they dare, and the minimal achieved distance (touching included) is taken as inversely related to avoidance. In animals, the strength of escape behavior or the duration of freezing responses can be measured, depending on the experimental situation. In rodents, who typically freeze when fearful, the effect of a fear stimulus can be assessed by its interfering effect on a regular background behavior, such as operant lever pressing for food rewards on a variable interval schedule.

On the premise that defensive reflexes are primed by an induced fear state, the modulation of such reflexes by fear stimuli can provide accurate information about fear both in animals and humans. Typically, the startle reflex is studied using whole-body startle in rodents and the eyeblink component of startle in humans. The typical result is an enhancement of the startle reflex to a standard startle probe stimulus (e.g., a white noise with abrupt onset) when it is presented against a background of a fear stimulus compared to when the probe is presented alone.

It is important to realize that different measures of fear are not necessarily highly intercorrelated, even when assessing the effects of a common fear stimulus. This is because the fear response is better conceptualized as a loosely coupled ensemble of partially independent response components, sensitive to various

modulating parameters, than as a unitary internal state mechanically elicited by the appropriate fear stimulus.

Fear Stimuli

There are innumerable events and situations that are feared by humans. Loosely speaking, they have in common that they provide a threat to the integrity of the individual, either in a physical or in a psychological sense. Many of the stimuli that are feared by humans are feared by other mammals as well. They include, for instance, loud noises, predators, and dominating conspecifics.

Classification of Fear Stimuli

In general, behaviors can be classified as communicative or noncommunicative depending on whether they elicit an active response from the environment. Communicative behavior is directed toward other living creatures, whereas noncommunicative behavior is directed toward the physical environment. Communicative behavior can be further subdivided into behavior directed toward members of another species, such as in predator-prey relationships, and members of the own species in what is commonly regarded as social behavior. Applied to fear, this classification system distinguishes among fear of physical stimuli, fear of animals, and social fears.

Physical stimuli Humans and other animals fear many types of physical stimuli, particularly those that may inflict tissue damage, thus inducing pain. A primary stimulus dimension is that of intensity. Highly intense stimuli of any modality may induce pain, and certainly they elicit fear and associated attempts to escape in most animals. Complex events that incorporate high-intensity stimulation, such as lightning and thunder, often evoke intense fright. But there are other complex events or situations whose fear-eliciting power is less directly dependent on simple stimulus dimensions. The effectiveness of such situations can be understood only in relation to the recurrent threat they have provided throughout evolution. Examples that come to mind include heights, small enclosures, wide-open spaces, and darkness or light (for species primarily active in daylight or at night, respectively).

Animal stimuli Species typically share ecological niches, and thus the presence of other animals has been a shaping force in evolution. Species compete for similar food supplies, and in predation, one species

provides the food supply for another. In this latter case, avoiding capture as prey is a prerequisite for reproduction; therefore, potential prey species fear their predators. There is also a widespread fear of potentially poisonous animals, such as snakes, spiders, and insects, and these fears may be better represented as fear of (and disgust for) contamination rather than as fear of the animal itself.

Social stimuli No animal is more dangerous to humans than other humans; the most dangerous predators are humans who are ready to use violence to exploit the resources of other humans. Conflicts with fellow humans, however, are not restricted to fights about tangible resources. Typically they deal with something more abstract, but also more pervasive – power. Like other primates, human groups are structured in terms of dominance, that is, some members dominate others. Fear is part of the submissiveness shown by the dominated group members when confronting a dominant conspecific. This fear is not automatically connected to escape or (least of all) to attack, but it is shown in a readiness to emit signals of submissiveness and in refraining from competition. In humans (and other primates), it denotes a fear of being negatively evaluated, of losing face in front of the group, rather than of physical harm.

Moderating Factors

Several general dimensions modify the fear elicited by these types of fear stimuli. One of them is closeness. In general, the closer the fear stimulus, the stronger the fear response. This may, in fact, provide an explanation for the effect of stimulus intensity on fear – an intense stimulus is likely to be very close. Prey animals (e.g., gazelles) show a minimal reactive distance before they overtly take notice of a predator (e.g., a lion). If the predator is far away or appears to be resting, it is monitored only by increased attention. If it gets closer or appears to be hunting, the attention enhancement is accompanied by defenses such as immobility. When the predator gets dangerously close, finally, there is active defense such as flight.

A second moderating factor is movement of the stimulus and the direction of the movement. Approaching objects, in general, elicit more fear than stationary objects or objects moving away. For example, more fear is generated by a fearsome authority figure who is heading in our direction than by one who is heading in another direction.

A third class of moderating factors involves predictability and/or controllability of the fear stimulus. An abruptly occurring stimulus is, by definition, not predicted and elicits immediate fear. Fear stimuli that

are predictable may also be behaviorally controlled, for example, through avoidance. Less predictable or controllable stimuli elicit more fear. However, it is only reasonable to talk about increasing fear as long as the uncontrollability does not completely undermine the coping attempts. When the situation is too uncontrollable to support active attempts to cope, fear is replaced by anxiety, and when the organism eventually gives up and becomes helpless, anxiety is replaced by depression (see **Anxiety**).

Fear Learning

Events such as obstructed breathing, physical constraints, and rapidly approaching large objects can be regarded as innate fear stimuli and may be called natural fear triggers. However, even though evolution has equipped us with defenses for a number of events that have threatened our survival in the long past of our species, modern humans, for good reason, fear many stimuli that simply were not around during our evolution. Thus, these stimuli must have acquired their fear-eliciting power through learning. They may be called learned fear triggers.

Pavlovian conditioning is the central mechanism for associative fear learning. Through Pavlovian conditioning, a natural fear trigger may transfer its potential to a new, previously neutral stimulus, thus turning it into a learned trigger. The procedure for achieving this is simply to present the two stimuli together, so that the to-be-learned trigger serves as a signal for the natural trigger. This is direct Pavlovian conditioning, but the procedure also works in a social arrangement in which one individual sees another individual express intense fear of the to-be-learned trigger. In this way, fear may transfer to new stimuli and circumstances, and, particularly for a species that has access to language, with its associative structure, this means that fear may come to be elicited from large classes of new stimuli only remotely associated with the original natural trigger.

An interesting possibility is that fear is more easily transferred to some stimuli that by themselves do not elicit fear, even though throughout evolution they have occurred in threat-related contexts. However, because of this long historic association with threats to survival, they may be evolutionarily prepared to enter easily into association with fear after only minimal aversive experience. For example, even though rhesus monkeys do not show an innate fear of snakes, they easily learn such fear after seeing conspecifics in fearful interactions with snake-related stimuli. Similar easy fear learning to neutral stimuli such as flowers is not obvious.

Pathological Fear: Phobias

Fear may sometimes be excessive to the extent that it interferes with normal adaptive functioning. Intense, involuntary, and rationally unfounded fear of a specific object or situation that provokes maladaptive avoidance is called a phobia. Phobias are classified into three categories depending on the feared object or situation.

Specific Phobias

Specific phobias concern circumscribed objects or situations, such as knives, other sharp objects, and dental treatments. One subgroup of specific phobias involves nature fears, such as fears of water or thunderstorms. Another important category centers on fear of animals; snakes, spiders, dogs, cats, and birds are typical examples. A third category involves fear of blood, taking injections, and mutilated bodies. Contrary to other anxiety disorders, which often incorporate a fear of fainting, this type of phobia is the only one actually associated with fainting as a result of a pronounced vasovagal response to exposure to the phobic situation.

Social phobias Some people tremble at the mere thought of meeting new people or having to formally address a group. In general, they fear social situations that involve being scrutinized or evaluated by others and the associated risk of being socially humiliated. Social phobia is more debilitating than specific phobia because social situations are central to human adjustment. For example, consistently avoiding such situations or enduring them only with intense dread may be detrimental both to academic and vocational careers.

Agoraphobia Agoraphobia (from the Greek *agora*, which means marketplace) denotes an intense fear of being out among people in crowded places such as in supermarkets or on public transportation vehicles. However, the fear is less concerned with this particular situation than with an intense fright of being overwhelmed by fear or anxiety in a situation in which no help or assistance is available; the common denominator is that the person would be left helpless without any escape route back to safety in the event of a sudden fear attack. Safety typically is defined as being at home, and thus many agoraphobics become captives in their homes, only able to leave if accompanied by a trusted companion.

Preparedness and Phobias

It is reasonable to assume that phobias derive primarily from learned fear triggers. Even though phobias are fairly common in the population, the

overwhelming majority does not show a phobia-level fear of, say, snakes, knives, blood, underground trains, or airplanes. Thus, the assumption that phobics somehow have associated fear to the phobic situation is readily invoked. If this were the case, we would expect a correlation between common ecological traumas and phobias. However, there is an obvious discrepancy between the distribution of phobic situations and the distribution of traumas in our environment. For example, many people have aversive experiences with broken electrical equipment, but there are few bread-toaster phobics needing clinical assistance. On the other hand, clinicians encounter many spider and bird phobics, even in environments completely lacking poisonous spiders or threatening birds. This discrepancy between traumas and phobias can be accounted for by the preparedness hypothesis because phobic objects and situations appear more obviously related to threats in an evolutionary than in a contemporary perspective. In fact, most phobic situations (animals, heights, enclosures, dominant conspecifics, lack of escape routes, etc.) have provided recurrent threats to humans and our predecessors throughout evolution. As a result, humans may have become biologically predisposed easily to associate such situations with fear even after only minimal trauma. Nonprepared situations, such as cars, on the other hand, may not come to elicit fear even after having been paired with excessive traumas.

The Neurophysiology of Fear

Research during the last 2 decades has delineated a neural network in the brain that controls fear responses and fear learning. This network is centered on the amygdala, a small set of nuclei in the anterior temporal lobe that has long been identified as a limbic structure associated with emotion.

A primary role for the amygdala appears to be the evaluation of input to the brain in terms of its potential threat. This purpose is achieved by the lateral nucleus, which receives fully processed sensory information from the cortex as well as only preliminarily processed information from subcortical structures such as the thalamus. These two routes to the amygdala have been described as the high and the low routes, respectively. Because the low route depends on monosynaptic linkage between the thalamus and the amygdala, the incompletely processed information conveyed by this route reaches the amygdala faster than the polysynaptically wired information conveyed by the high route. As a consequence, the amygdala can start recruiting defense responses even before the veridicality of the threat stimulus is confirmed by the full cortical analysis.

The high route, therefore, is not necessary for eliciting fear, but the low route may be necessary both for fear elicitation and fear learning. Fear may be elicited and learned in animals even after the ablation of the relevant sensory cortices, and in humans, fear responses can be elicited and learned via stimuli that are prevented from conscious recognition through backward masking. Thus, fear responses are recruited after a quick and nonconscious analysis of the stimulus, which explains the automatic, nonvoluntary character of intense fear, such as in phobias.

After threat evaluation in the lateral and basolateral nuclei of the amygdala, the information is conveyed to the central nucleus, which controls various efferent aspects of the fear response. Neural pathways to the lateral hypothalamus activate sympathetically controlled responses, such as heart-rate acceleration and skin conductance responses, and parasympathetically dominated responses, such as heart-rate decelerations, are influenced through the vagal motor nucleus and the nucleus ambiguus of the brain stem, which also can be controlled from the amygdala. Through connections between the central nucleus and the paraventricular nucleus in the hypothalamus, corticotropin releasing hormone can be released, which, via adrenocorticotrophic hormone from the anterior pituitary, activates corticosteroid stress hormones from the adrenal cortex. Paths to the tegmentum, including the locus coeruleus, activate the dopaminergic, noradrenergic, and cholinergic arousal systems of the forebrain, resulting in electroencephalogram (EEG) activation and increased vigilance. Overt motor responses associated with fear such as freezing, fight, and flight are activated by connections between the central nucleus of the amygdala and the dorsal

and ventral periaqueductal gray of the brain stem. The facial expressions of fear are activated through the facial motor nucleus and the seventh cranial nerve, the nervus facialis. The modulation of the startle reflex, finally, is accomplished through neurons connecting the central nucleus of the amygdala with the central nucleus for the startle reflex, the nucleus reticularis pontis caudalis in the brain stem. This circuit accounts both for peculiarities of fear activation such as its independence of conscious recognition of the fear stimulus and the complex efferent organization of the fear response. Furthermore, its sensitization through various procedures may contribute to the understanding of pathological fear and anxiety.

See Also the Following Articles

Anxiety; Fear and the Amygdala.

Further Reading

- Davis, M. and Whalen, P. J. (2001). The amygdala: vigilance and emotion. *Molecular Psychiatry* 6, 13–34.
- Lang, P. J., Bradley, M. M. and Cuthbert, B. N. (1997). Motivated attention: affect, activation, and action. In: Lang, P. J., Simons, R. F. & Balaban, M. T. (eds.) *Attention and orienting: sensory and motivational processes*, pp. 97–135. Mahwah, NJ: Lawrence Erlbaum.
- LeDoux, J. E. (1996). *The emotional brain*. New York: Simon & Schuster.
- Öhman, A. and Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological Review* 108, 483–522.
- Rosen, J. B. and Schulkin, J. (1998). From normal fear to pathological anxiety. *Psychological Review* 105, 325–350.

Fear and the Amygdala

R Norbury

Warneford Hospital and University of Oxford, Oxford, UK

G M Goodwin

University of Oxford, Oxford, UK

© 2007 Elsevier Inc. All rights reserved.

Amygdala Anatomy

Afferent and Efferent Connections

The Amygdala and Fear Conditioning in Animals

The Amygdala and Fear in Humans

Glossary

Amygdala

A complex brain structure with numerous subnuclei located deep within the temporal lobe of the brain.

Backward masking paradigm

A paradigm in which an initial emotional face (e.g., fear) is presented for a very short duration (17–33 ms) and then immediately replaced by a neutral face presented for a longer duration (~200 ms). In this situation, the initial image is subliminal; subjects report that they have seen only the neutral expression.

<i>Excitatory postsynaptic potential (EPSP)</i>	A depolarization of the postsynaptic membrane.
<i>Fear</i>	An unpleasant, powerful emotion caused by anticipation or awareness of potential or actual danger.
<i>Fear conditioning</i>	A form of emotional learning in which an emotionally neutral or conditioned stimulus acquires the ability to elicit behavioral and physiological responses associated with fear after being associated with an aversive unconditioned stimulus.
<i>Long-term potentiation (LTP)</i>	The long-lasting strengthening of the connection between two neurons.
<i>Neuroimaging</i>	A number of powerful noninvasive techniques, including positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), that allow the examination of regional patterns of brain activation in the living human brain with considerable spatial resolution, but limited temporal resolution, in the order of seconds as determined by vascular responses associated with neural activity.
<i>Startle response</i>	An innate reflex observed in nearly all animals, including humans, that manifests as an involuntary motor response to any unexpected noise, touch, or sight. In studies of emotion, it has been used to measure the aversiveness of emotive stimuli.

Amygdala Anatomy

The amygdala is an almond-shaped structure located deep within the temporal lobe of higher animals. It was first identified by Burdach in the early nineteenth century, who described a group of cells, or nuclei, now referred to as the basolateral complex. Subsequently, however, the amygdala was shown to be both more complex and more extended, comprising more than twelve subnuclei. In rats, current nomenclature divides the amygdala nuclei into three main groups: (1) the basolateral complex, which includes the lateral nucleus, the basal nucleus, and accessory basal nucleus; (2) the cortical nucleus, which includes the cortical nuclei and the lateral olfactory tract; and (3) the centromedial nucleus, comprising the medial and central nuclei.

Afferent and Efferent Connections

The amygdala receives input from all sensory systems: olfactory, somatosensory, gustatory, visceral,

auditory, and visual. Olfactory inputs arise at the olfactory bulb and project to the lateral olfactory tract. Somatosensory inputs pass via the parietal insular cortex in the parietal lobe and via thalamic nuclei to the lateral, basal, and central nuclei. Primary gustatory and visceral sensory areas project to the basal nucleus and central nucleus. In contrast, auditory and visual information, thought to be important in fear conditioning, arise from association areas rather than from the primary sensory cortex.

The amygdala has widespread efferent connections to cortical, hypothalamic, and brain-stem regions. The basolateral complex projects to the medial temporal lobe memory system (e.g., the hippocampus and perirhinal cortex), and the basal nucleus has a major projection to prefrontal cortex, nucleus accumbens, and the thalamus. Thus, the anatomy of the amygdala is consistent with its role in fear processing. The amygdaloid complex receives inputs from all sensory modalities and activates brain regions important to measurable neurobehavioral correlates of fear. How it works offers an insight into the nature of emotion in humans and animals.

The Amygdala and Fear Conditioning in Animals

Much of the scientific interest in the amygdala stems from its established role in fear conditioning, research that has been carried out mostly in rats. Classical Pavlovian fear conditioning is a type of emotional learning in which an emotionally neutral conditioned stimulus (CS), often a tone, is presented in conjunction with an aversive unconditioned stimulus (US), typically a small electric shock to the foot of the animal. After one or more pairings, the emotionally neutral stimulus (CS) is able to elicit a constellation of species-specific conditioned responses (CRs) that are characteristic of fear, such as freezing or escape behavior, autonomic responses (elevated heart rate and blood pressure), potentiated acoustic startle to aversive acoustic stimuli, and increased neuroendocrine responses (release of stress hormones). Fear conditioning therefore allows new or learned threats to activate ways of responding to threat that have been long established in evolution.

Numerous studies have demonstrated that lesions to the amygdala impair the acquisition and expression of conditioned fear in rats. The basolateral complex of the amygdala is a substrate for sensory convergence from both the cortical and subcortical areas, and it is considered a putative locus for CS-US association during fear conditioning. Thus, its cells encode this emotional learning. By contrast, the central nucleus of the amygdala projects to brain

regions implicated in the generation of fear responses, such as the hypothalamus; it may therefore act as a common output pathway for the generation of fear-conditioned responses. Consistent with this hypothesis, lesions to either the basolateral or central nucleus of the amygdala impair the both the acquisition and expression of conditioned fear.

The amygdala pathways involved in fear conditioning have also been studied using electrophysiological studies (see [Figure 1](#)). During fear conditioning, the convergence of CS and US inputs to the basolateral complex results in sustained enhancement, or long-term potentiation, of EPSPs evoked by the CS ([Figure 1](#)). Thus, the amygdala is able to both integrate and associate sensory information and influence the motor and physiological responses associated with fear conditioning.

The Amygdala and Fear in Humans

Evidence from Lesion Studies

Damage to the amygdala, or areas of the temporal lobe that include the amygdala, also produces deficits in fear processing in humans. In a recent study, a patient with Urbach–Wiethe disease (S.M.), a rare congenital lipid storage disease that results in the bilateral degeneration of the amygdala, underwent fear conditioning with either visual or auditory CSs and a loud noise as the US. Compared to normal control subjects, S.M. showed no evidence of fear conditioning (as measured by galvanic skin response). Notably, S.M.'s recall of events associated with fear conditioning was intact. These data support the hypothesis that the amygdaloid complex plays a key role in the acquisition of Pavlovian fear conditioning, whereas the hippocampus is important in acquiring declarative knowledge of the conditioning

contingencies. In a further study with the same patient, it was demonstrated that bilateral amygdala damage impaired the recognition of fearful facial expressions. Other patients with Urbach–Wiethe disease have also been shown to be impaired in recognizing fearful facial expressions. S.M., however, had no difficulty in recognizing people by their faces or in learning the identity of new faces. These results suggest that the human amygdala is directly involved in the processing of emotion but not in recognizing aspects of facial appearance. There is also evidence to suggest that the impairment of the perception of expressive signs of fear in patients with amygdala damage extends beyond facial expressions; bilateral amygdala damage has also been associated with impairment in the recognition of vocalic expressions of threat.

The effect of amygdala damage on the startle response has also been studied in humans. A recent study compared the startle response in a 32-year-old male with a localized lesion in the right amygdala with that of eight age-matched controls. The control subjects displayed the well-documented effect of aversive stimuli potentiating startle magnitude. In the patient with the right-amygdala lesion, no startle potentiation was observed in response to aversive versus neutral stimuli. Together, studies of patients with amygdala damage suggest that this structure plays a key role in the perception and production of negative emotions, particularly fear.

Evidence from Neuroimaging Studies

During the early 1990s, researchers began to explore the role of the amygdala in fear processing using neuroimaging (e.g., positron emission tomography, PET; and functional magnetic resonance imaging, fMRI) and continue to do so today. So far, amygdala activity has been assessed predominantly using two

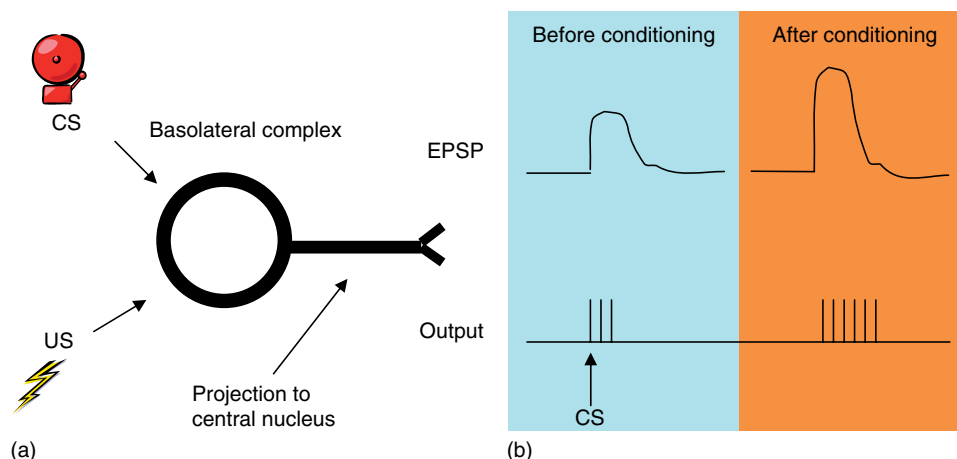


Figure 1 During fear conditioning convergence of inputs (CS and US) induce long-term potentiation of EPSPs evoked by the CS. a, Schematic; b, graphs. CS, conditioned stimulus; EPSPs, excitatory postsynaptic potentials; US, unconditioned stimulus.

basic paradigms: fear conditioning and presentation of emotional facial expressions.

Functional MRI experiments have demonstrated increased amygdala activity during both the early acquisition and early extinction of fear conditioning to a visual stimulus. Indeed, it has been suggested that the amygdala's role is limited to early conditioning or early extinction, when response contingencies change. That is, the amygdala is particularly important for forming new associations as relationships – emotional learning and unlearning.

The amygdala may also be responsible for generating coordinated reflexive behavioral responses to highly aversive stimuli. The animal literature shows that the amygdala, specifically the central nucleus, generates conditioned autonomic, behavioral, and endocrine responses in acute stress paradigms. CSs, which may be interpreted as changing the animal's emotional set, increase the response in startle paradigms, for example. Clearly, this kind of response may be quite primitive, but lesioning experiments demonstrated definite mediation by the amygdala. It is usually assumed the amygdala may be involved in the selection of more purposive motor-behavioral responses in response to more subtle aversive conditions.

The second major neuroimaging protocol for assaying amygdala activity is the presentation of faces expressing different emotions. It is difficult to over-emphasize the importance of facial expressions in social communication. Facial expressions act at a number of levels to signal important information; expressions of disgust enable the avoidance of the ingestion of harmful substances, and fearful facial expressions rapidly communicate the presence of imminent threat. More subtly, cues from the faces

of others are continuously informing us of our social impact and acceptability, and we, in turn, communicate our feelings and intentions through our own facial gestures and expressions.

It is no surprise, therefore, that facial expressions have provided a useful experimental tool to measure fear-related amygdala activity. Many studies have demonstrated that the amygdala is preferentially activated during the presentation of fearful facial expressions. Moreover, even when subjects are unaware of seeing a fearful face, by the use of a masking protocol or when attention is directed away from the fearful face, the amygdala is still activated. Such findings suggest that the extraction of potentially threatening information within the amygdala may not be sufficient for conscious face perception but it may well be necessary. The amygdala may act to direct attention toward emotionally salient events that are ambiguous or require further processing. Indeed, anatomical studies in rats suggest that sensory information travels to the amygdala via two distinct pathways: a short, rapid thalamic route, and a longer cortical pathway. It is proposed that the short thalamic pathway rapidly prepares the animal for a potentially aversive encounter independent of conscious processing, which occurs later via the slower cortical route. A similar two-way route to the amygdala has yet to be anatomically defined in humans; however, evidence from a number of well-designed fMRI experiments point to the existence of these two pathways ([Figure 2](#)). First, as already described, the amygdala responds to fearful faces even when presented outside conscious awareness. Second, in an extension of the masking paradigm, Whalen and colleagues presented degraded versions of fearful and happy expressions

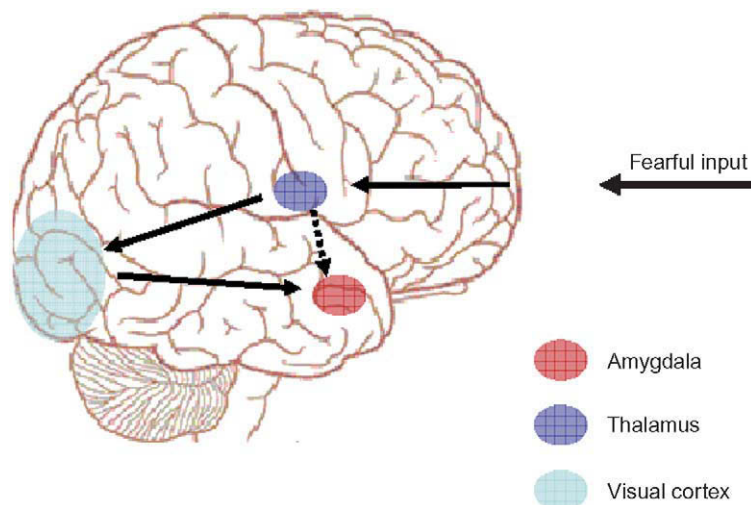


Figure 2 Two-pathway hypothesis for amygdala activation to fear. The short, rapid thalamic route is shown by the dashed arrow; the longer cortical pathway is shown by bold arrows. (For clarity, subcortical structures are shown overlaid on brain surface.)

by displaying only the eye whites immediately masked by a complete image of a neutral expression. Fearful faces are typically associated with the enlarged exposure of the sclera, and it was observed that the crude information provided by the wide eyes of fearful expressions was sufficient to stimulate the amygdala. Analogously, Vuilleumier and coworkers used fMRI to compare amygdala and cortical responses to rapidly presented (200 ms) fearful and neutral faces that had been filtered to extract either low-spatial-frequency information (giving a blurry image) or high-spatial-frequency information (giving a finely detailed, sharp image) (Figure 3). As detected by fMRI, the finely detailed images elicited a greater response in a region of the cortex called the fusiform face area, a brain region widely implicated in the conscious recognition of faces and facial expressions. The amygdala was relatively unresponsive to these images even if they were fearful in nature. Low-spatial-frequency fearful faces, however, produced a robust response in the amygdala. Although these findings could imply the existence of a two-pathway route to the amygdala in humans, with dissociation between neural responses in amygdala and fusiform gyrus across different spatial frequency ranges for face stimuli, they do not prove it. However, they suggest that the amygdala does not require very complex configurational stimuli, such as presented by the complete image of a face, in order to respond to emotion. It is obviously possible that crude fragmentary information related to emotionally salient signals could be rapidly communicated via a subcortical route. Extreme facial expressions of emotion may be well adapted to attract immediate attention via this mechanism.

Given the involvement of the amygdala in fear processing, there is increased interest in the role of this structure in anxiety disorders, such as post-traumatic stress disorder (PTSD), and obsessive-compulsive disorder (OCD), and in depression. For example, patients with PTSD show increased blood flow in the right amygdala (as measured by PET) in response to the presentation of images designed to activate traumatic memories, compared to neutral images. Using fMRI, it has also been demonstrated that individuals with OCD have significantly increased activity bilaterally in the amygdala in response to ideographically tailored stimuli designed to provoke their symptoms.

Functional MRI can also be used to investigate how pharmacological treatment for anxiety and depression modulates amygdala activity. A number of studies have demonstrated that the increased amygdala response to negative facial expressions seen in depression is reduced following effective treatment with serotonergic antidepressants. Although these results suggest an important role for the amygdala in the recovery from depression, it remains unclear if this normalization of amygdala response to fearful faces with time is a direct action of the drug or, rather, a reflection of the current symptom state (the scans of patients when depressed and when they had recovered were compared). Work from our laboratory, in healthy volunteers, demonstrated that the short-term administration (7 days) of either serotonergic (e.g., citalopram) or norepinephrine (e.g., reboxetine) antidepressants reduces the amygdala response to the masked expressions of fear compared to placebo controls. Significantly, these effects on the amygdala were observed in the absence of changes in mood. These

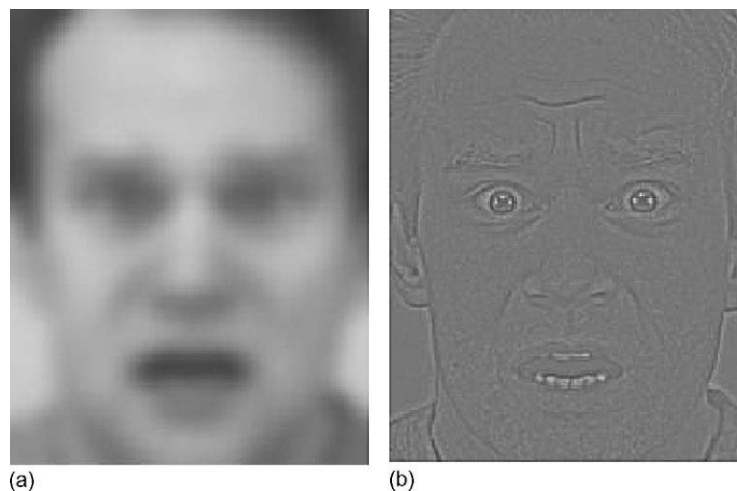


Figure 3 Processing of images by the amygdala and cortical regions. a, Low-frequency image; b, finely detailed image. The rapid subcortical pathway may allow low-frequency inputs to reach the amygdala rapidly, whereas finely detailed images are further processed in cortical regions (e.g., fusiform gyrus).

results suggest that antidepressants have rapid, direct effects on the amygdala. Moreover, if similar effects are seen in clinical populations, drug modulation of amygdala activity in response to negative stimuli may be a key component in patients' recovery from depression and anxiety.

See Also the Following Articles

Anxiety; Fear; Posttraumatic Stress Disorder, Delayed; Posttraumatic Stress Disorder – Clinical; Posttraumatic Stress Disorder – Neurobiological basis for.

Further Reading

Harmer, C. J., Mackay, C. E., Reid, C. B., et al. (2006). Antidepressant drug treatment modifies the neural

processing of nonconscious threat cues. *Biological Psychiatry* 59, 816–820.

Ledoux, J. (2003). The emotional brain, fear, and the amygdala. *Cellular and Molecular Neurobiology* 23, 727–738.

Phelps, E. A. and LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behaviour. *Neuron* 48, 175–187.

Sah, P., Faber, S. L., de Armentia, L., et al. (2003). The amygdaloid complex: anatomy and physiology. *Physiological Reviews* 83, 803–834.

Vuilleumier, P., Armony, J. L., Driver, J., et al. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience* 6, 624–631.

Whalen, P. J., Kagen, J., Cook, R. G., et al. (2004). Human amygdala responsivity to masked fearful eye whites. *Science* 306(5704), 2061.

Febrile Response

S Gatti McArthur

F. Hoffmann-LaRoche, Basel, Switzerland

T Bartfai

The Scripps Research Institute, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S Gatti McArthur and T Bartfai, volume 2, pp 116–123, © 2000, Elsevier Inc.

Introduction

Thermal Control in Endotherms

Endogenous Pyrogens and Fever Induction

Stress and Fever Induction

Glossary

<i>Anterior hypothalamus/preoptic area</i>	Hypothalamic area representing the major center of control of the temperature set point of the body.
<i>Basal metabolic rate</i>	The minimal (or steady state) energy use, keeping the body temperature constant under conditions of no net work carried out.
<i>Cytokines and interleukins</i>	Small proteins with regulatory functions produced and secreted mostly by immunocompetent cells.
<i>Endogenous pyrogens</i>	Proinflammatory cytokines or endogenous compounds that function as humoral mediators of fever induction.

Hypothalamic-pituitary-adrenocortical axis

The central and peripheral system sustaining the physiological response to stressors.

Intraperitoneal and intracerebroventricular injections

Usual ways to administer exogenous pyrogens, i.e., lipopolysaccharide or endogenous pyrogens to study their peripheral and central proinflammatory effects.

Lipopolysaccharide

A fraction of the *Escherichia coli* bacterial wall used in different experimental models of gram-negative bacterial infection.

Thermoneutral zone

A range of ambient temperatures compatible with a control of the core body temperature achievable solely by cardiovascular changes in the peripheral compartment of the body.

Introduction

Fever is a transient change of the hypothalamic temperature (HT) set point causing the rise of the core body temperature (BT) of about 2–4°C in humans. This physiological response to inflammatory stimuli is always associated with peripheral and central events: massive hepatic production of acute-phase proteins, sleep induction, and anorexia, defined globally as the acute-phase response. Fever is triggered by the peripheral and central production of interleukin-1β (IL-1β),

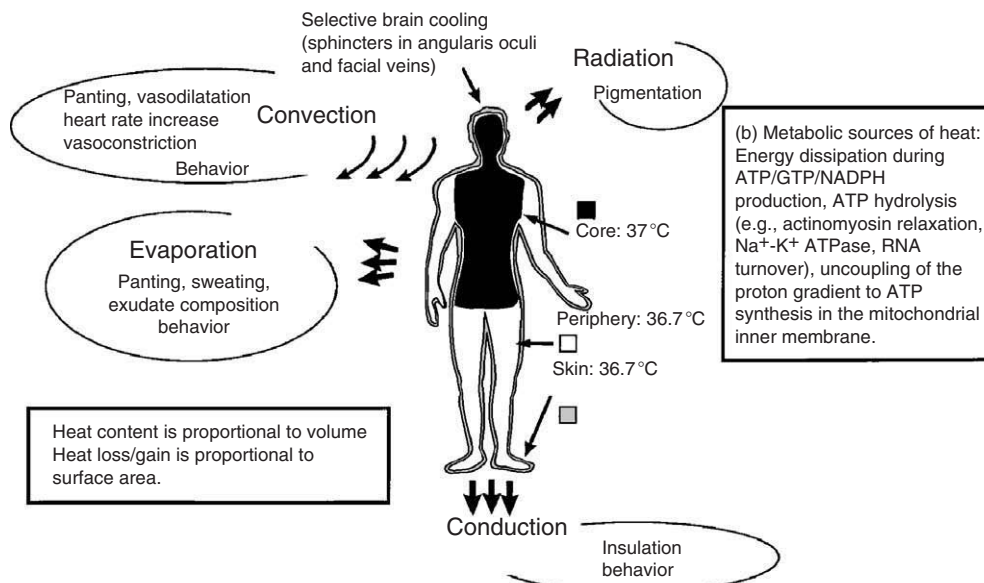
tumor necrosis factor α (TNF- α), and interleukin-6 (IL-6) and is caused by the local production of prostaglandins (PGE₂) in the anterior hypothalamus/preoptic area (AH/POA). The rise of the BT during febrile response is achieved by reducing heat loss (vasoconstriction) and by increasing simultaneously the heat production (facultative thermogenesis) via the activation of the sympathetic system. The sudden heat production associated with acute stress, usually described as hyperthermia, is also triggered by the activation of the sympathetic system. It is a common observation that subchronic, chronic, or exhaustive stresses are associated either with fever-like episodes or with an impairment of the whole inflammatory response. Experimental data demonstrate the endogenous production of IL-1 β in rat hypothalamus during unescapable stress, pointing to the tight molecular control exerted by glucocorticoids (GCs) and corticotropin-releasing factor (CRF) on the production and effects of pyrogenic cytokines.

Thermal Control in Endotherms

Heat production is a sign of metabolic activity. Heat storage/heat exchange mechanisms are strikingly different in living organisms, resulting from the evolution of specialized systems (e.g., enzymes, organs, and cell types) required to survive and acclimate in

different environmental conditions. **Figure 1** shows a summary of the mechanisms controlling thermal balance in endotherms. The core BT is the result of both heat production and heat exchange with the environment under the control of different pools of thermal sensors present in skin, viscera, main cardiovascular districts, spinal cord, and brain. The whole body can be divided in three main compartments (core, periphery, and skin), which regulate the exchange of heat mainly by modifying blood flow in the different districts. Cardiovascular, respiratory, and perspirative effects are always accompanied by behavioral changes in order to obtain an effective increase or drop of the BT. The main metabolic exothermic reactions contributing to heat production are listed in **Figure 1b**. The brown fat tissue (BAT), which is very rich in mitochondria, can produce heat by uncoupling the electron transport/proton gradient in the inner mitochondrial membrane thanks to the presence of uncoupling protein-1 (UCP-1) (nonshivering thermogenesis). Other gene products of the same family, UCP-2 and UCP-3, are expressed in white fat tissue/liver and in the muscular tissue, respectively. The contribution of UCP-2 to thermogenesis during fever is under debate.

Transient (or prolonged) temperature changes, observed during fever, have a deep metabolic impact, being associated with kinetic changes in all enzymatic reactions, in membrane conductance, and change



ATP, adenosine triphosphate; GTP, guanosine triphosphate; NADPH, nicotinamide adenine dinucleotide phosphate; RNA, ribonucleic acid.

Figure 1 Schematic view of the mechanisms of heat production and heat exchange in endotherms. (a) Conduction, convection, radiation, evaporation, and metabolic heat production contribute in an additive manner to the definition of the temperature in the different districts of the body: skin, periphery, and core. The mechanism of selective brain cooling controls the facial venous flow in order to locally control the brain temperature by circulatory means. (b) The main exothermic metabolic reactions.

of permeability of ionic channels. Furthermore, to actively increase the BT of a human by 1°C within 1 h requires the net heat storage of about 3.48 J kg^{-1} with an energy cost for a human of 70 kg (in thermoneutral conditions) of about 243.6 J. GCs do support the metabolic heat requirements of the febrile response by enhancing several thermogenic reactions. Endogenous pyrogens (EPs) are also responsible for the metabolic changes observed during the acute-phase response. In particular leptin, a protein controlling white fat tissue metabolism and energy expenditure, exerts a necessary role in sustaining facultative thermogenesis during fever. Considering the energy required to mount a febrile response, it becomes understandable why starvation, cachexia, extreme ambient temperature, or other causes of metabolic distress would reduce the magnitude of the febrile response if not abolish it completely.

Temperature Set Point

Figure 2 shows effector systems activated in endotherms to keep the core BT constant while the ambient temperature changes progressively. During cold exposure, shivering is controlled by the activation of the motor system, whereas changes in superficial microcirculation and nonshivering thermogenesis are signs of the activation of the sympathetic system. The energetic requirement of the cold-induced thermal

response is reduced progressively as the ambient temperature reaches the thermoneutral zone. Within this narrow range of temperatures, energy requirements to keep the BT constant (in absence of stress and motor activity) are met by the basal metabolic rate (BMR). The response to a warm environment is conversely characterized by vasodilatation, panting, and sweating. Shivering frequency and BMR are inversely correlated on a logarithmic scale with the body mass. Furthermore, the BMR is constantly adapted in different individuals in response to climatic conditions, diet, and motor activity under the control of the thyroid hormones. The BT in the different compartments and the range of temperatures triggering the effector systems in order to increase alternatively heat storage or heat loss are defined by the integration of thermoafferent signals and their comparison with a temperature set point defined in the AH/POA. Among the signals integrated in AH/POA, it is important to point at the local HT, reported by several studies as the stimulus able to overcome peripheral inputs. The local HT cannot fluctuate freely more than $0.1\text{--}0.2^{\circ}\text{C}$ in a narrow range of temperatures that is defining the set point. Whatever change of the local temperature outside the limits of this range would trigger the effector response to a warm or cold environment (Figure 2). During fever the hypothalamic set point is shifted upward with a

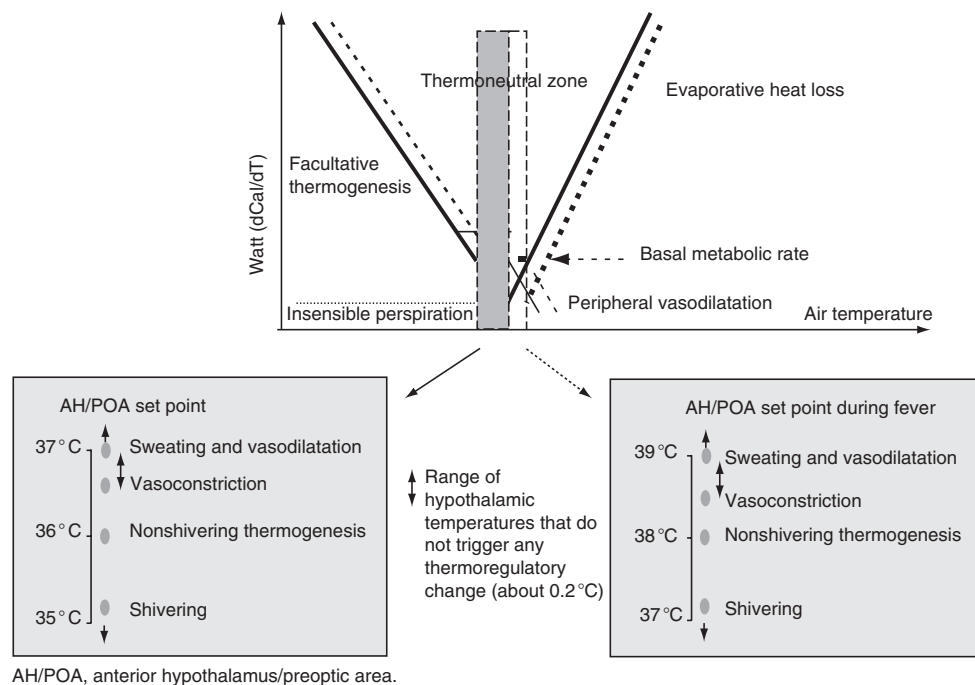


Figure 2 Thermal control in endotherms: the sympathetic and parasympathetic systems are alternatively activated by changes of the ambient temperature exceeding the thermoneutral zone in order to keep the core BT constant. The thermoneutral range is defined by the HT set point. The whole system is transiently shifted upward during fever.

reset of the range of HTs triggering the different effector responses. Because the body would need increased heat storage and heat production to reach the new temperature in the different districts, fever, in its initial phase, would cause a thermal response mimicking the normal reaction to a cold environment.

Thermosensitive Neurons

There are no known molecular markers for the neurons involved in the central control of the temperature set point. The only information available about the neuronal populations involved is that they share the common feature of being thermosensitive. Neuronal thermosensitivity (TS) can be characterized by measuring the neuronal firing rate in a parapsycho-physiologic range of temperatures and calculating the thermal coefficient for each cell. The range of thermal sensitivity of AH/POA neurons may vary quite extensively and it is generally linear within a narrow range of temperatures. Two distinct populations of TS neurons have been characterized in the hypothalamus: cold-sensitive and warm-sensitive neurons. Cold-sensitive neurons are less studied, mostly because they represent less than 10% of the TS neurons in AH/POA. Warm sensitivity has been studied quite extensively using hypothalamic slices. These studies demonstrate that TS is an intrinsic cellular property. Warm-sensitive neurons are present in several brain regions: in AH/POA, in lower centers in the brain stem, and in the spinal cord. The presence of TS neurons in different brain regions may suggest temperature differences in the cerebral versus peripheral compartment. However, it has been demonstrated that thermal curves are similar both in time course and in absolute values in the cerebral cortex and the peritoneum under basal conditions, during fever, and pharmacological hypothermia.

Fever and Thermosensitivity

Fever depresses the TS of neurons whose firing rates had been increased previously by local AH/POA warming. Furthermore, local injections in the AH/POA of EPs or prostaglandin E₂ (PGE₂) cause a reduction of the tonic firing of warm-sensitive neurons similar to the one observed when lowering locally the HT. The plasticity of warm-sensitive neurons and their responsivity to bath-applied IL-1 or IL-6 has yet to be characterized properly, and moreover, the molecular components of neuronal activity changes in AH/POA during fever have yet to be determined. Few studies reporting intracellular recordings from warm-sensitive neurons are opening up a key path of research in this respect.

Endogenous Pyrogens and Fever Induction

IL-1 β was among the first cytokines described, and its biological properties as EP, the leukocytic pyrogen, were described long before the purification of the protein. In general, the concept of a humoral mediator of a fever response is older than the discovery of pro-inflammatory cytokines and was initially related to the observation that bacterial debris lipopolysaccharide (LPS) could cause fever with almost a 2-h delay from the time of the injection. The LPS effect was also characterized by priming and tolerance, further suggesting the presence of endogenous mediators. The progressive discovery of several pyrogenic cytokines present in the plasma only during an inflammatory reaction, and the study of their synergism and production in hypothalamus during a febrile response, further clarified the definition of EP. All EPs cause fever when administered by the intracerebroventricular (ICV) route and in the periphery and they are released into the circulatory system during the fever response to exogenous pyrogens. The magnitude and time course of the fever response could be related to the plasma and hypothalamic levels of the EP only in the case of IL-6. IL-1 β or TNF- α seems to trigger the febrile response more than sustaining it because several endogenous antagonistic systems (such as IL-1 receptor antagonist (IL-1ra), in the case of IL-1 α and IL-1 β) are buffering or suppressing *in vivo* the initial EPs effects both peripherally and centrally (Figure 3a). The list of known EPs would then contain interferons (IFN- α , - β , and - γ), IL-1 α / β , TNF- α / β , and IL-6. IL-1 β is the most potent: a 30-ng kg⁻¹ intravenous injection is sufficient to cause fever in humans. IL-6 and chemokines (IL-8, macrophage inflammatory protein-1) would trigger the fever response only when injected centrally into the AH/POA or ICV. The generation of mice carrying null mutations for the different EPs or their receptors has been extremely important in demonstrating the presence of a hierarchy of EP in the generation of the fever response, with IL-6 acting probably downstream of IL-1 β and TNF α .

The most relevant results of the studies can be summarized as follows:

- Studies on IL-1 type 1 receptor (IL-1RI) knockout (KO) mice and IL-1R accessory protein (IL-1Rap) KO mice have demonstrated that IL-1RI mediates the fever response induced by an intraperitoneal (IP) injection of IL-1 α and IL-1 β . The fever response to LPS is rather normal in these animals, suggesting that TNF and other EPs can bypass the IL-1 system in the generation of the febrile response.

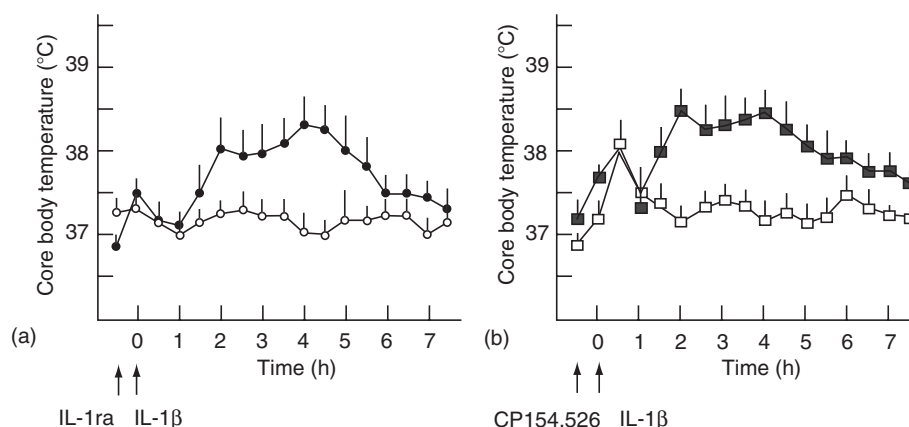


Figure 3 (a) Fever response in rats injected IP with $5 \mu\text{g kg}^{-1}$ of recombinant human IL-1 β . The febrile response is blocked by the injection of recombinant human IL-1ra (2.5 mg kg^{-1}). (b) Fever response in rats injected IP with human recombinant IL-1 β ($5 \mu\text{g kg}^{-1}$). The effect is blocked by the previous injection of the nonpeptidic CRF receptor antagonist CP154,526 (0.16 mg kg^{-1} , IP). Data are means $\pm$ standard error means. Data were recorded using a telemetry system and averaged every 30 min. Open circles, IL-1ra plus IL-1 β ($n = 5$), closed circles, IL-1 β ($n = 3$), open squares, CP154 526 plus IL-1 β ($n = 4$), and closed squares, IL-1 β ($n = 8$).

- IL-1 β KO mice show hyperresponsive fever mostly in terms of magnitude of the effects in response to IL-1 α and IL-1 β IP injection.
- IL-6 KO mice are unable to mount a febrile response on injection of IL-1 β or LPS in the periphery. Only the parallel co-injection of IL-6 IP seems to restore a sort of fever response to LPS IP injection. Very interesting genetic studies on different autoimmune familiar disorders like Mediterranean fever are pointing future research towards proteins such as pyrin and cryopyrins which are possibly upstream EP production.

Exogenous Pyrogens

The most used experimental model of fever response in the presence of an endogenous stimulus is represented by LPS injection in the periphery. This model mimics a response to gram-negative bacterial infections. Fever is also present during gram-positive bacterial infections because macrophages can produce somnogenic and pyrogenic muramyl peptides during the digestion of staphylococci fractions. Furthermore, agents such as turpentine are broadly used as an experimental model of aseptic inflammation, which is almost uniquely mediated by IL-1 β . However, the broadest spectrum of exogenous pyrogens is offered by viral proteins 2 or by the etiological agents of parasitic infections such as schistosomiasis or malaria. In case of viral infections, interferons (IFNs) are the principal responsible of the observed febrile response with IL-1 β contributing as EP. EP and Toll-like proteins responsive to exogenous pyrogens are probably present in all vertebrates. Equally ancient from the

phylogenetic point of view is the complement cascade of proteins which interact with EPs and EP production mainly via anaphylatoxins.

Expression of Endogenous Pyrogens and their Receptors in the Hypothalamus

IL-1 α/β , TNF- α , and IL-6 gene expression can be induced in AH/POA on peripheral immune challenge (with LPS or IL-1 β for instance). The peripheral administration of LPS induces IL-1 β production in meninges, choroid plexus macrophages, some perivascular cells, and microglial cells shortly after IP injection (1–8 h). Constitutive expression of IL-1 β and IL-6/IL-6R in neuronal cell types has been reported by some authors, but the matter is still controversial. Receptors for the different EPs, particularly IL-1RI, are expressed constitutively in AH/POA. Their expression is modulated by GCs and EP, particularly IL-1 β . At the same time, pharmacological doses of corticosterone block the LPS-induced fever response and IL-1 α/β and TNF- α mRNAs expression in the hypothalamus, suggesting a causal role for the hypothalamic production of cytokines in fever induction. The control by GC and EPs on the expression of EPs/EP receptors is mostly transcriptional. The promoter regions of EP genes contain an unusually high number of GC receptor response elements. In the case of IL-1 β and TNF- α , posttranscriptional control of GC at the mRNA and proteic level has been demonstrated to be extremely important in controlling the cellular production of these EPs in immunocompetent cells and the activation/secretion of the mature form of these cytokines.

Humoral Versus Neuronal Mediators of Fever Induction

How does the brain sense EPs? This key question has several answers. In fact, circulating pro-inflammatory cytokines can reach the brain via the interaction with saturable carriers in the plexus chorioideus. However, this mechanism would account for the entrance of a limited amount of protein, eventually able to trigger local cytokine production AH/POA. Several pathologic conditions cause damage to the blood-brain barrier. In this case the amount of circulating EPs reaching the hypothalamus could be higher. There is also a specialized circumventricular organ, the organum vasculosum lamina terminalis (OVLT), located along the margin of the ventricular system that has fenestrated capillaries and, therefore, direct contact with the circulatory system. The role of the OVLT has been demonstrated in the case of peptides such as angiotensin II and cholecystokinin and is probably also essential for the transduction of the plasmatic levels of EPs into central effects. A third mechanism of immune signaling, essential in the case of peritoneal immune challenge, is the peripheral stimulation of sensory nerves, particularly the afferent part of the vagus nerve. IL-1Rs are present in vagal paraganglia, chemosensory bodies, all along the course of the vagal nerve in the cervical region, and in thoracic and abdominal cavities. Subdiaphragmatic vagotomy is able to block the central production of IL-1 β in response to the IP injection of IL-1 β or LPS. The immediate c-fos mRNA induction has been used to trace the neuronal pathway activated in response to the peripheral loading of LPS, and several studies would point to the presence of a selective activation of neurons in the hypothalamus (paraventricular nucleus, ventromedial POA), brain stem (A1 and A2 region), nucleus of the solitary tract, and ventrolateral medulla. In some cases the activated pathways represent antipyretic responses that are initiated during fever, such as the activation of neurons that contain α -melanocyte-stimulating hormone (α -MSH) and arginine-vasopressin.

Role of Prostaglandins in Fever

The role of prostaglandins (PG) of the E series in fever induction was first described by Milton and colleagues in 1970. PGs of the F class can also cause fever. Their effect is, however, mediated by the release of CRF. Both PG series are thus EPs. The fever response triggered by pro-inflammatory cytokines is always mediated by the local production in AH/POA of PGs. In fact, the pyrogenic effect of peripheral LPS and EPs can be blocked by cyclo-oxygenase inhibitors injected locally into AH/POA. A dense distribution

of PGE₂-binding sites is present in the anterior wall of the third ventricle surrounding the OVLT in the rat brain. Moreover, a study on mice carrying a null mutation for the different PG receptors demonstrated that receptor EP3 selectively mediates the PGE₂-induced fever response in the presence of IL-1 β or LPS. It is important to note that PG synthesis inhibitors are excellent antipyretic agents but have few or no effects on stress. So the picture is not complete. A missing link for fast hypothalamic events triggered by EPs could be the C2-ceramide pathway.

Stress and Fever Induction

Stress-induced fever can be measured using several experimental protocols of stress in rodents, e.g., unescapable stress or restriction. In some cases, it should be noted, however, that the stress response would modify the circadian rhythms of BT and that the results obtained should be normalized considering possible shifts of the circadian rhythm of the BT of the animals. The development of a fever response to certain kinds of stress is also confirmed by the increased expression in the hypothalamus and hippocampus of EPs; particularly IL-1 β mRNA and protein. The effects of IL-1 β on all the levels of the hypothalamus-pituitary-adrenocortical (HPA) axis are significant and very well characterized by several studies. IL-1 β can activate CRF synthesis and release and adrenocorticotrophic hormone (ACTH) secretion acting locally at the hypothalamic and pituitary level, respectively. Hence, IL-1 β is one of the few known agents able to induce CRF release, although the direct effect of IL-1 β on the production/release of GC is still controversial. In contrast, the GC effect on cytokine production is dual and depends on the circulating levels of GC. In fact, physiological levels of GC would support the metabolic requirement of the acute-phase reaction, leading also to increased GC production as the HPA is activated. As a result, the increased GC release during inflammation buffers the pro-inflammatory cytokine effects by restricting their expression. A special note should, however, be taken while discussing the effect of GC on IL-6. In rodents, IL-6 mRNA expression is higher in the hypothalamus after chronic treatment with high doses of corticosterone, and mice overexpressing IL-6 in astrocytes exhibit higher plasmatic levels of corticosterone, reduced ACTH response, and adrenal hyperplasia. This may suggest a direct effect of IL-6 at the level of the adrenal glands. Furthermore, both IL-6 and GC are responsible for the production of different sets of hepatic proteins carrying GC-responsive elements and nuclear factor-IL-6 response elements in the promoter region of their genes, and

they are both probably involved in controlling the duration and the magnitude of the acute-phase response. Increased plasmatic cytokine levels during stress have been explained considering the possible bacterial translocation from the gastrointestinal tract to the mesenteric lymph nodes.

The presence of this phenomenon is, however, seldom considered when planning an experimental protocol, and the impact of bacterial translocation on central and peripheral IL-1 production has not yet been characterized clearly. When examining the activation of the HPA axis during fever response, it is important to keep in mind that a central role is played by CRF and vasopressin. Central injection of CRF causes fever in rats and it also activates thermogenesis in the BAT. Furthermore, neutralizing antibodies to CRF or synthetic CRF antagonists block the pyrogenic response associated with IL-1 β or TNF injection ([Figure 3b](#)). CRF plays a necessary but indirect role in the development of the fever response. Vasopressin is an endogenous antipyretic hormone. New synthetic selective antagonists will possibly allow in the future a better understanding of which receptor is mediating fever and HPA-axis effects.

Finally, the role of the sympathetic system in the development of stress or the cytokine-induced thermal response has to be redefined on the basis of the demonstrated presence of the constitutive production of IL-1 β and IL-6 in the sympathetic ganglia and in adrenal medulla. *In vivo* during an inflammatory reaction and fever induction, the presence of EP and EP receptors at different levels can ensure a tight control on the activation of the sympathetic system in terms of magnitude and duration of the effect based on the presence of a local or systemic inflammatory reaction.

See Also the Following Articles

Homeostasis; Hyperthermia; Hypothermia; Prostaglandins.

Further Reading

Chai, Z., Alheim, K., Lundkvist, J., et al. (1996). Subchronic glucocorticoid pretreatment reversibly attenuates IL-1 beta induced fever in rats; IL-6 mRNA is elevated

while IL-1 alpha and IL-1 beta mRNAs are suppressed, in the CNS. *Cytokine* 8, 227–237.

Dinareello, C. (1998). Interleukin-1, interleukin-1 receptors and interleukin-1 receptor antagonist. *International Review of Immunology* 16, 457–499.

Elmqvist, J., Scammell, T. and Saper, C. (1997). Mechanisms of CNS response to systemic immune challenge: The febrile response. *Trends in Neurosciences* 20, 565–570.

Kluger, M. (1991). Fever: role of pyrogens and cryogens. *Physiological Reviews* 71, 93–127.

Lundkvist, J., Chai, Z., Teheranian, R., et al. (1996). A non peptidic corticotropin releasing factor receptor antagonist attenuates fever and exhibits anxiolytic-like activity. *European Journal of Pharmacology* 308, 195–200.

Nguyen, K. T., Deak, T., Owens, S. M., et al. (1998). Exposure to acute stress induces brain interleukin-1beta protein in the rat. *Journal of Neuroscience* 18, 2239–2246.

Rolfe, D. and Brown, G. (1997). Cellular energy utilisation and molecular origin of standard metabolic rate in mammals. *Physiological Reviews* 77, 731–758.

Sanchez-Alavez, M., Tabarean, I. V., Behens, M. M., et al. (2006). Ceramide mediates the rapid phase of febrile response to IL-1beta. *Proceedings of the National Academy of Sciences USA* 103, 2904–2908.

Schobitz, B., Reul, J. and Holsboer, F. (1994). The role of the hypothalamic-pituitary-adrenocortical system during inflammatory conditions. *Critical Reviews in Neurobiology* 8, 263–291.

Schonbaum, E. and Lomax, P. (eds.) (1990). *International encyclopedia of pharmacology and therapeutics, sect. 131, Thermoregulation: physiology and biochemistry*. New York: Pergamon Press.

Kluger, M. J., Kozak, W., Leon, L. R., et al. (1998). Fever and antipyresis. In: Sharma, H. S. & Westman, J. (eds.) *Brain function in hot environment. Progress in brain research* (vol. 115), pp. 465–475. Amsterdam: Elsevier Science.

Sundgren-Andersson, A., Gatti, S. and Bartfai, T. (1998). Neurobiological mechanisms of fever. *Neuroscientist* 4, 113–121.

Sundgren-Andersson, A. K., Ostlund, P. and Bartfai, T. (1998). Simultaneous measurement of brain and core temperature in the rat during fever, hyperthermia, hypothermia and sleep. *Neuroimmunomodulation* 5, 241–247.

Ting, J. P., Kastner, D. L. and Hoffman, H. M. (2006). CATERPILLERS, pyrin and hereditary immunological disorders. *Nature Reviews Immunology* 6, 183–195.

Ushikubi, F., Segi, E., Sugimoto, Y., et al. (1998). Impaired febrile response in mice lacking the prostaglandin E receptor subtype EP3. *Nature* 392, 281–284.

Feedback Systems

G Fink

Mental Health Research Institute, Parkville, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G Fink, volume 2, pp 124–137, © 2000, Elsevier Inc.

Introduction

Overview

Interaction between Negative Feedback and Circadian Rhythm in the HPA System

Corticosteroid Feedback on the Hypothalamus and Pituitary Gland: Phase Differences

Glucocorticoid Feedback Effects on Stress Neurohormone Biosynthesis

Role of Atrial Natriuretic Peptide in Glucocorticoid Inhibition

Role of Hippocampus and Amygdala in Glucocorticoid Negative Feedback

Glucocorticoid Negative Feedback at the Pituitary Level

Possible Role of 11β -Hydroxysteroid Dehydrogenase

Functional Importance of Glucocorticoid Negative Feedback

Clinical Manifestations of Disordered Glucocorticoid Feedback Regulation of the HPA System

Allostasis and Allostatic Load

Conclusion

Glossary

<i>Allostasis</i>	Maintaining stability (or homeostasis) through change; term introduced by Sterling and Eyer in 1988 to describe cardiovascular adjustments to resting and active states.
<i>Homeostasis</i>	The maintenance of equilibrium, or constant conditions, in a biological system by means of automatic mechanisms (generally feedback systems) that counteract influences (stressors) that tend toward disequilibrium (term introduced by Walter Cannon in 1932).
<i>Hypercortisolemia</i>	Excessively high concentrations of adrenal corticosteroids in plasma.
<i>Hypophyseal portal vessels</i>	A group of small venules that connect a primary plexus of capillaries in the median eminence of the hypothalamus (base of the brain) with a secondary plexus of sinusoids in the pituitary gland. Blood flowing in these vessels,

which run down the pituitary stalk, transports hypothalamic-pituitary regulatory neurohormones, released at nerve terminals in the median eminence of the hypothalamus, to the anterior pituitary gland, where they stimulate or inhibit the release of anterior pituitary hormones.

Limbic system

An extensive brain region that includes the cingulate and parahippocampal cortex, hippocampus, amygdala, hypothalamus, septal nuclei, and other structures. The precise interaction of the several components of the system is poorly understood, but since the proposal by Papez, the limbic system has been implicated in emotion. The connections between the limbic system and the neocortex, hypothalamus, and the olfactory bulb via the olfactory tract suggest that it serves as an important analyzer-integrator of signals, which it conveys from the neocortex to the hypothalamus. The input of olfactory information to the limbic system has led to its alternative name, the rhinencephalon (olfactory brain), and possibly reflects the strong emotive and signaling effects of smell in many animals. The structure and relative size of the limbic system have remained remarkably constant through evolution; in humans it is overshadowed by the massive development of the neocortex.

Neurohormones

Chemical neurotransmitters released from nerve terminals into the hypophyseal portal vessels or the systemic circulation at neurohemal junctions and conveyed to their target cells by the blood stream. This contrasts with classical neurotransmitters that reach receptors on target cells by crossing synaptic clefts or neuromuscular junctions.

Suprachiasmatic nuclei

Two small hypothalamic nuclei located immediately above (dorsal to) the optic chiasm that are responsible for regulating most circadian rhythms of the body.

Introduction

Feedback control systems are fundamental for the normal physiological functioning of the body. There are two types of feedback, negative and positive, of which the former is the more common. Removal of the main pituitary target glands, the adrenals, gonads, and thyroid, the most reproducible and reproduced

experiment in classical endocrinology, demonstrates that the secretion of pituitary adrenocorticotropin (ACTH), gonadotropins, and thyrotropin is controlled by negative feedback exerted by the adrenal corticosteroids, gonadal steroids, and thyroid hormones, respectively. Interruption of the negative feedback loop caused, for example, by surgical or pharmacological adrenalectomy, gonadectomy, or thyroidectomy, or an enzymatic defect in steroid or thyroid hormone biosynthesis, results in massive hypersecretion of pituitary ACTH, gonadotropin, or thyrotropin.

Crucial for homeostasis, negative feedback control mechanisms are composed of a system in which the output moderates the strength of the controller to a predetermined set point level (Figures 1 and 2). The set point–stimulator complex contains a comparator (error detector), which compares the strength of the feedback signal with a preset level. An increase in the strength of the feedback signal above the preset level reduces the output of the stimulator, whereas a decrease in the strength of the feedback signal below the preset level results in an increase in output of the feedback signal and corrects the error. Negative feedback control operates widely throughout the body at the molecular (e.g., end product inhibition of enzyme activity), cellular, and whole system/body levels. The mechanisms by which the set point is determined vary between systems: in the case of the hypothalamic-pituitary-adrenocortical (HPA) feedback system, the set point is regulated by the central nervous system. GABAergic and glutamatergic neural projections from the limbic system and other brain

regions to the paraventricular nuclei (PVN) of the hypothalamus may play a key role in HPA set point regulation.

Positive feedback, whereby the output of a system increases the output of the stimulator (gain in the system) (Figure 2b), is far less common than negative feedback, possibly because taken to its logical conclusion, a positive feedback system, uncontrolled, will eventually self-destruct. Some systems loosely termed positive feedback are, in fact, servomechanisms. A servomechanism is a closed loop control system, which increases significantly the power of a small signal. Positive feedback is exemplified by (1) estrogen stimulation of gonadotropin secretion, which results in ovulation, and (2) the release of oxytocin induced during parturition by pressure of the fetal head on the uterine cervix. The latter triggers volleys of impulses that, by way of a multisynaptic pathway to the hypothalamus, stimulate oxytocin release. Oxytocin stimulates further contraction of the uterus, which results in a further increase in pressure on the uterine cervix. This vicious cycle is broken only when the fetus is expelled. Servomechanisms are illustrated by the way that, just before ovulation in humans as well as in other spontaneously ovulating mammals, elevated plasma estrogen concentrations increase the responsiveness of the anterior pituitary gland to gonadotropin-releasing hormone (GnRH) by 20- to 50-fold. This estrogen effect, reinforced by the priming effect of GnRH, whereby a small pulse of GnRH further potentiates pituitary responsiveness to itself, ensures the occurrence of a massive ovulatory gonadotropin

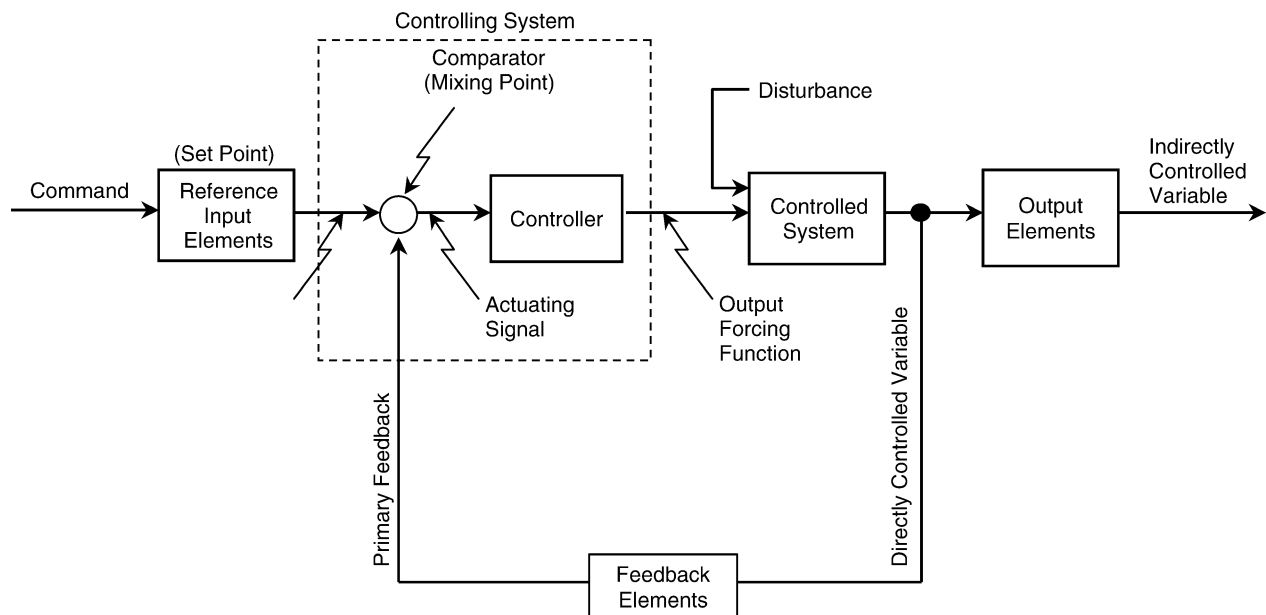


Figure 1 A generalized feedback control system. Modified from Milhorn, H. T. J. (1966). *The application of control theory to physiological systems*, with permission from Saunders, Philadelphia, PA.

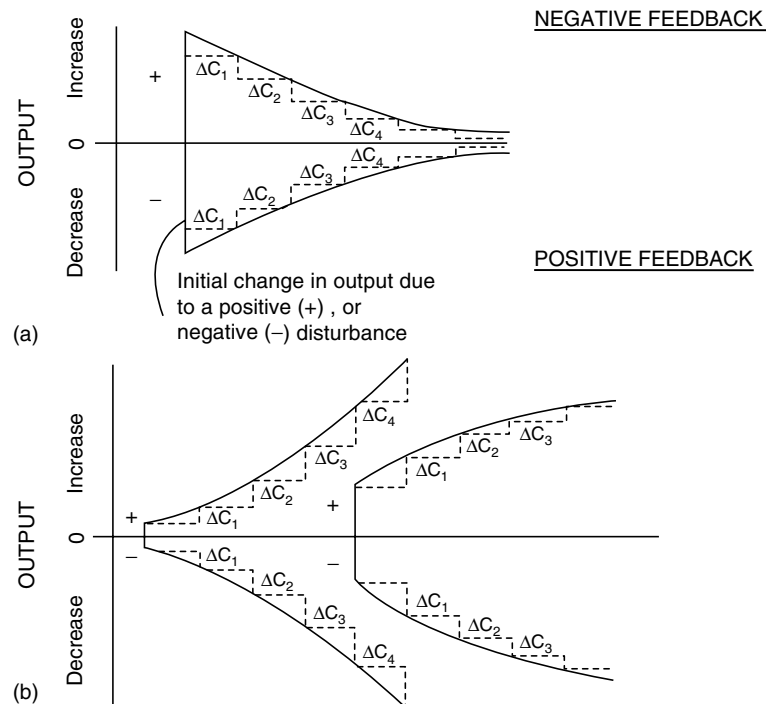


Figure 2 (a) Negative feedback minimizes the disturbance to a regulator, resulting in a system in which the output tends to remain constant. In this case, the ratios of the decrements of the controlled variable C (C_2/C_1 , C_3/C_2) are less than unity. (b) In positive feedback (left-hand curves) an initial disturbance results in a continuous increase in output (vicious cycle). The increments of the controlled variable (C_2/C_1 , C_3/C_2) are greater than unity. When the response does not result in a vicious cycle, the ratios of C_2/C_1 , C_3/C_2 , and so on are less than unity (see right-hand curves in B). From Milhorn, H. T. J. (1966). *The application of control theory to physiological systems*, with permission from Saunders, Philadelphia, PA.

surge in response to a small surge or increased pulse frequency of GnRH.

This article focuses on the principles of negative feedback control in the HPA system with emphasis on neurohormone and hormone release. While the discussion concentrates on negative feedback maintenance of homeostasis, the same mechanisms are involved in allostasis that is brought about by a change in set point so as to enable the organism to deal with the physiological challenge or stress.

Overview

In the normal state, with the negative feedback loop intact, the set point in the brain-pituitary module maintains the secretion of ACTH within a relatively narrow bandwidth. Within this bandwidth, the basal secretion of ACTH is pulsatile. Plasma ACTH concentrations also show a circadian rhythm, with a peak in the morning and a trough approaching a nadir around midnight. In nocturnal animals, such as rodents, the phase of this rhythm is reversed so that plasma ACTH concentrations reach a peak just before the onset of darkness. The circadian rhythm of ACTH results in a circadian rhythm in the plasma concentrations of adrenal corticosteroids: cortisol in

humans and corticosterone in rodents. A key feature in both diurnal and nocturnal animals is that corticosteroid plasma concentrations reach a peak just before the animal is due to wake from sleep; the levels are lowest just before or during sleep.

Glucocorticoids (cortisol in the human; corticosterone in rodents) secreted by the adrenal cortex exert their inhibitory action both on the brain and on the pituitary gland (Figures 3 and 4). The PVN of the hypothalamus contain the final common pathway neurons that mediate the neural control of pituitary ACTH synthesis and release (see **Secretagogue**). The PVN receive stimulatory and inhibitory neural inputs and an important projection from the suprachiasmatic nuclei (SCN), the key circadian oscillator in brain. PVN drive is mediated by stress neurohormones, corticotropin releasing factor-41 (CRF-41) and arginine vasopressin (AVP), which are released into hypophyseal portal vessel blood. ACTH is released into the systemic circulation and stimulates adrenal corticosteroid synthesis and release. In keeping with revised standard nomenclature, CRF-41 will be abbreviated to CRH (corticotropin-releasing hormone).

The release of ACTH is pulsatile in humans and is cleared rapidly from the blood by both metabolic degradation and distribution into several body

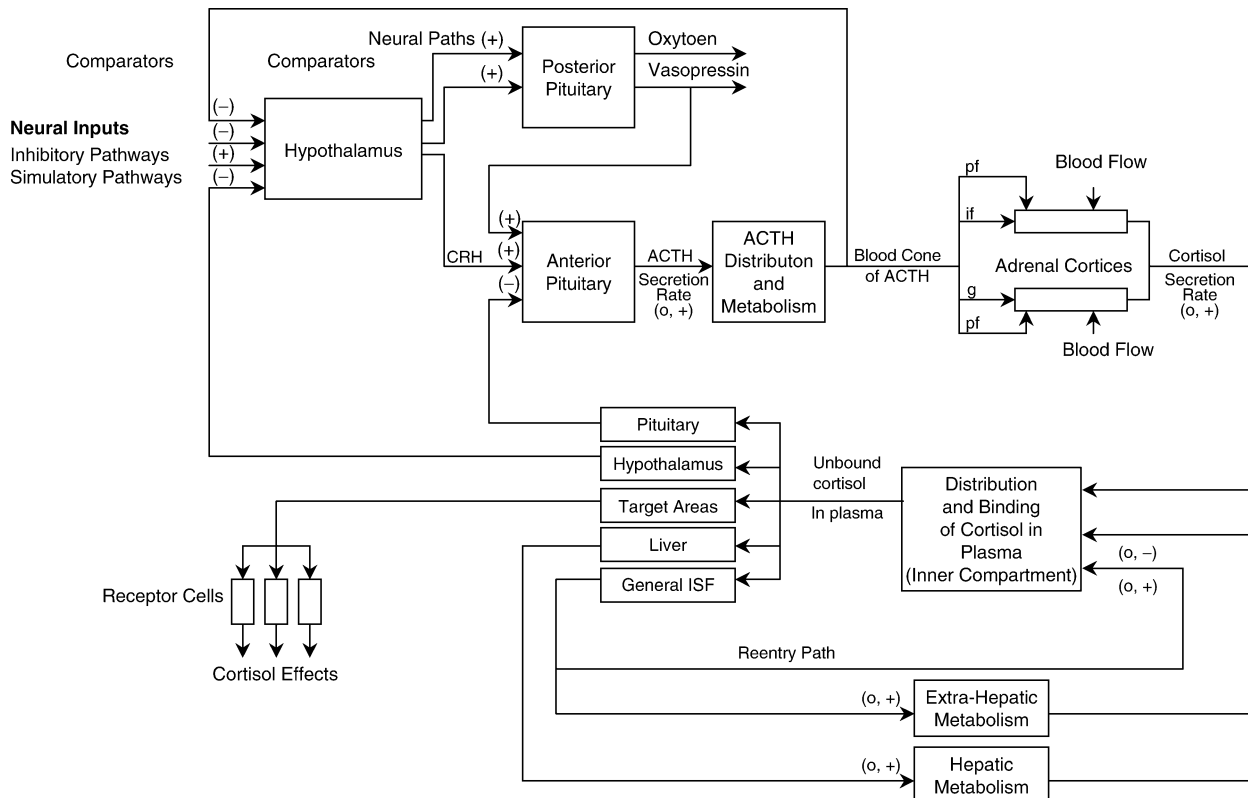


Figure 3 Block diagram of the hypothalamic-pituitary-adrenal glucocorticoid control system. Abbreviations: if, input forcing of adrenal by ACTH; pf, parametric forcing of adrenal (hypertrophic effect) caused by ACTH over a longer time period. The parametric effect of changes in adrenal blood flow is also indicated. The designators 0, + and 0, - indicate that signals in pathways are restricted in values (e.g., there are no negative masses or frequencies, and removal processes or inhibitors are negative in effects). From Yates, F. E. and Maran, J. W. (1975). In: Knobil, E. & Sawyer, W. H. (eds.) *Handbook of physiology* (pp. 367–404). Washington, D.C.: American Physiological Society, with permission.

compartments. The glucocorticoids are bound rapidly by albumin and a specific corticosteroid-binding protein, transcortin, and are metabolized by several organs, especially the liver and kidney (Figure 3). Only unbound (free) glucocorticoids inhibit ACTH release, and therefore the degree of glucocorticoid binding and metabolism, as well as the magnitude of adrenal secretion (Figure 3), determines the strength of the negative feedback signal. The HPA feedback system might also be influenced by a CRH-binding protein in plasma (first discovered by Orth and Mount in 1987). The physiological role of the CRH-binding (glyco) protein, which is synthesized in the liver, awaits to be determined – in humans, CRH-binding protein concentrations in plasma increase during the third trimester of pregnancy (see **Corticotropin Releasing Factor-Binding Protein**).

Interaction between Negative Feedback and Circadian Rhythm in the HPA System

The circadian rhythm of ACTH and (the consequent) glucocorticoid secretion, which has a peak at the end

and a nadir at the beginning of the sleep phase, is driven by a neural mechanism mediated mainly by the stress neurohormones. The twofold increase in the ACTH signal between the nadir and the peak of the circadian rhythm in the rat results in a ninefold increase in corticosterone due to an increase in the responsiveness of the adrenal cortex.

In the unstressed state, the HPA system operates in an approximately linear domain with all the loop variables (Figure 3) showing circadian periodicity. Ultradian rhythms in plasma cortisol concentration, independent of ACTH concentration, occur in humans and rhesus monkeys, in which they are highly synchronized between animals with a predominant periodicity of 85–90 min.

The circadian rhythm of ACTH secretion is driven by the SCN. In addition to the diurnal rhythm of corticosterone, the SCN are involved in other key rhythms in the body, such as the sleep–wake cycle, motor activity, thermoregulation, pineal *N*-acetyl transferase activity, and the regular occurrence of ovulation. If any one of these functions is abnormal, then there is a high probability that the others will

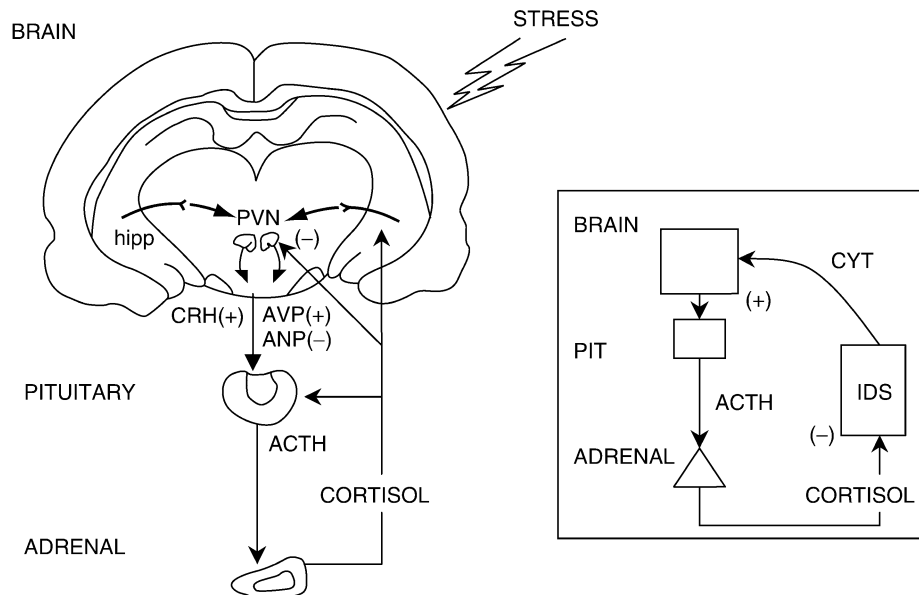


Figure 4 Schematic diagram of the elements involved in glucocorticoid negative feedback. The paraventricular nuclei (PVN) contain the main stress (final common pathway neurons) secreting both CRH and AVP. The negative feedback action of cortisol (corticosterone in rodents) is exerted mainly on the PVN and the pituitary corticotropes. However, long-term effects mediated through the hippocampus (Hipp) and amygdala (not shown) cannot be excluded. There are several possible indirect connections between the hippocampus and PVN. Neural control of ACTH secretion may also be mediated by a corticotropin inhibitory peptide, atrial natriuretic peptide (ANP). The inset shows the important inhibitory action of cortisol on the immune defense system (IDS). The IDS produces cytokines (CYT), which act on the brain to stimulate ACTH secretion. Cytokine secretion is inhibited by the negative feedback effect of glucocorticoids.

also be abnormal. The corticosterone rhythm is especially sensitive in that even partial lesions of the SCN, which have no effect on any of the other circadian functions, do disrupt the adrenal rhythm.

The SCN receive afferent projections from (1) the retina (direct as well as indirect after a relay in the ventral lateral geniculate nucleus), (2) the 5-hydroxytryptamine (5-HT) raphe neurons, and (3) the hippocampus by way of the medial corticohypothalamic tract. Each of these inputs to the SCN is likely to affect the periodicity of the pacemaker: the retinal input affects the time of the light–dark cycle; the hippocampal input may be related to the sleep–wake cycle; the raphe input may be related to paradoxical sleep. The raphe input also plays a major role in determining the amplitude of diurnal ACTH oscillations.

Although the central action of glucocorticoids is mainly on the PVN, the 5-HT neurons of the dorsal raphe nuclei, which project to the SCN, have high concentrations of glucocorticoid receptors. The SCN have direct projections to the PVN and may also affect PVN function by way of their projections to the subparaventricular zone. While circadian rhythmicity of the HPA is modulated by glucocorticoids and other factors, the intrinsic rhythmicity of the SCN circadian pacemaker is primarily under genetic control (see **Circadian Rhythms, Genetics** of).

Corticosteroid Feedback on the Hypothalamus and Pituitary Gland: Phase Differences

The seminal modeling studies of Eugene Yates and associates suggested that corticosteroid negative feedback on ACTH release occurs in three phases: a fast and immediate rate-sensitive phase of about 30 min, followed by a level-sensitive phase that occurs 2–3 h after the start of corticosteroid administration, followed by a long-term chronic phase. Fink and associates investigated the effect of corticosteroids on the release of stress neurohormones into hypophyseal portal blood (see **Neuroendocrine Systems**) in the intermediate and chronic phase. They found that adrenalectomy induced a three- to fourfold increase in the release of both CRH and AVP into hypophyseal portal blood. Administration of the synthetic glucocorticoid dexamethasone 2.5 h before portal blood collection significantly reduced the release of AVP, but not CRH (**Figure 5**), and also blocked the ACTH response to CRH (**Figure 6**). These findings suggest that intermediate-delayed (2–3 h) glucocorticoid feedback is mediated by the blockade of pituitary responsiveness to CRH and a reduction in AVP output into hypophyseal portal blood. That is, in this situation AVP is the regulatory or signaling neurohormone,

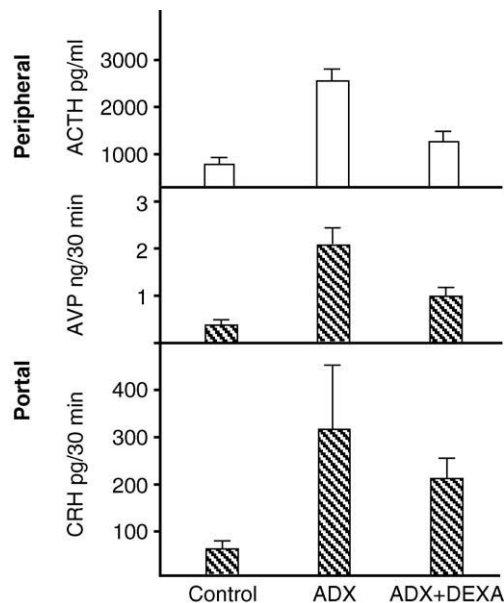


Figure 5 The delayed negative feedback action of glucocorticoid in adult female Wistar rats anesthetized with sodium pentobarbitone. ACTH was measured in plasma taken immediately before sectioning of the pituitary stalk for the collection of hypophyseal portal vessel blood. Adrenalectomy (ADX) resulted in a fourfold increase in ACTH concentration and a similar increase in the output of AVP and CRH release relative to that in intact, untreated animals (control). The administration of dexamethasone (DEXA) significantly reduced the levels of ACTH and AVP, but not CRH. Modified from Fink, G., Robinson, I. C. A. F. and Tannahill, L. A. (1988). Effects of adrenalectomy and glucocorticoids on the peptides, CRF-41, AVP and oxytocin in rat hypophyseal portal blood. *Journal of Physiology* **401**, 329–345, with permission of the authors and Cambridge University Press.

whereas CRH is the permissive neurohormone. In contrast, studies on the long-term effects of corticosterone administered by a subcutaneous pellet suggested that long-term negative glucocorticoid feedback is due to a decreased release of CRH, as well as AVP, into portal blood. That is, both stress neurohormones are sensitive to the long-term effects of corticosteroids.

These findings agree broadly with those of Plotsky and associates, which showed that pharmacological adrenalectomy with metyrapone and aminoglutethimide (which block glucocorticoid biosynthesis) resulted in a significant increase in the release of both CRH and AVP into hypophyseal portal blood. The intravenous infusion of corticosterone inhibited nitroprusside-evoked CRH release into portal blood, but had no effect on AVP release.

Data in the rat are complemented by those obtained in the sheep. Thus, the concentrations of CRH and AVP in hypophyseal portal blood collected from conscious sheep were (1) similar to those in the rat, (2) increased by volume depletion, fear-associated audiovisual stimuli, and insulin-induced hypoglycemia, and (3) inhibited by dexamethasone.

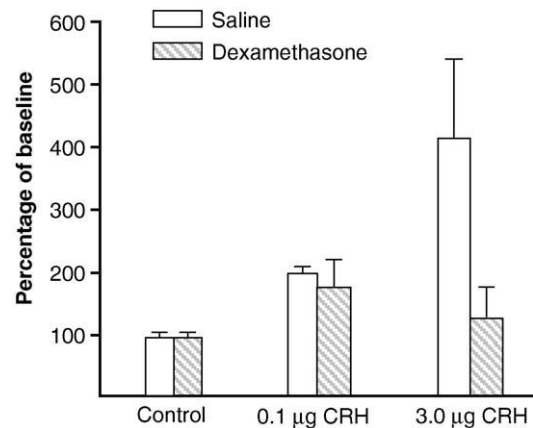


Figure 6 Mean ($\pm$ SEM, $n = 5$) percentage increase over the basal concentration of ACTH after the injection of saline, 0.1 µg CRH, or 3 µg CRH. Female Wistar rats that had been adrenalectomized 3 weeks earlier were treated with either saline or dexamethasone 3 h before injection of CRH. Note that dexamethasone blocked the ACTH response to 3.0 µg of CRH. From Fink, G., Robinson, I. C. A. F. and Tannahill, L. A. (1988). Effects of adrenalectomy and glucocorticoids on the peptides, CRF-41, AVP and oxytocin in rat hypophyseal portal blood. *Journal of Physiology* **401**, 329–345, with permission of the authors and Cambridge University Press.

Glucocorticoid Feedback Effects on Stress Neurohormone Biosynthesis

Glucocorticoids have potent inhibitory effects on CRH biosynthesis as well as release. Thus, adrenalectomy is followed by a significant increase in CRH mRNA levels in the parvocellular PVN, and this increase can be reduced by either corticosterone or dexamethasone. As assessed by CRH-intron *in situ* hybridization, the stimulation of CRH gene transcription can be detected as early as 15 to 30 min after the injection of the glucocorticoid synthesis inhibitor metyrapone. This increase in CRH gene transcription in the PVN was associated with a coincident increase in *c-fos* mRNA in the PVN. An increase in the levels of CRH mRNA in the PVN after metyrapone injection was delayed by about 60 min, possibly a function of the high resting levels of CRH mRNA and the time taken to assemble mRNA from the CRH primary transcript.

The effects of corticosterone and dexamethasone are dose dependent and can be demonstrated by systemic administration of the steroid, as well as by implantation of steroid pellets into the brain. This effect of glucocorticoids is cell specific in that the glucocorticoid-induced decrease in CRH mRNA was localized to the dorsomedial parvocellular neurons of the PVN, the major source of CRH fiber projections to the median eminence. In contrast, glucocorticoid increased the levels of CRH mRNA in parvocellular

neurons that project to the brain stem and the spinal cord. The implantation of dexamethasone micropellets in cerebral cortex, dorsal hippocampus, ventral subiculum, lateral septum, or amygdala had no effect on CRH mRNA levels in the PVN.

The molecular mechanism by which glucocorticoids regulate CRH gene expression remains unclear. Recent data show that the repressor isoform of cAMP response element modulator (CREM) is involved in CRH gene regulation, and that, *in vivo*, stress-induced glucocorticoids do not limit CRH gene transcription, in spite of the fact that the CRH promoter has a glucocorticoid receptor response element. The inhibitory effect of glucocorticoids on CRH gene transcription may be mediated, at least in part, by GABAergic, glutamatergic, and monoaminergic projections to the PVN from the forebrain and hindbrain limbic system.

The concentration of AVP mRNA in the dorso-medial parvocellular neurons of the PVN paralleled those of CRH mRNA, that is, levels increased after adrenalectomy and decreased after treatment with glucocorticoids. Corticosterone treatment also produced a modest reduction in the levels of enkephalin mRNA in all four regions of the PVN, but glucocorticoid manipulation had no significant effect on the PVN concentrations of the mRNAs for angiotensin, cholecystokinin, preprotachykinin, and tyrosine hydroxylase. There is, therefore, a close correlation between the synthesis and the release of CRH and AVP. Whether this means that synthesis and release are mechanistically coupled is not established.

Data on the release and synthesis of stress neurohormones suggest that even though corticosteroid receptors are present in high concentration in parts of the brain remote from the hypothalamus, including the hippocampus, amygdala, and the monoaminergic nuclei of the hindbrain, central inhibition of CRH and AVP synthesis and release may be due in large part to corticosteroid action mainly at the level of the PVN. Although a role for mineralocorticoid receptors cannot be excluded, glucocorticoid receptors predominate.

Role of Atrial Natriuretic Peptide in Glucocorticoid Inhibition

Fink and associates showed that atrial natriuretic peptide (ANP) is released into hypophyseal portal blood and, by immunoneutralization, that ANP is possibly an inhibitor of ACTH release. This suggests that the glucocorticoid inhibition of ACTH release may be mediated in part by ANP. A consensus sequence for the glucocorticoid response element is present in the second intron of the ANP gene, and ANP gene expression in and ANP release from cardiac cells are increased significantly by glucocorticoids.

A pathophysiological role for ANP is suggested by multiple trauma and septic shock in humans, where, paradoxically, plasma ACTH concentrations are decreased but plasma cortisol concentrations are increased. These results raise the possibility that under conditions of extreme stress, increased levels of ANP inhibit ACTH release while cortisol secretion is stimulated by a direct action of endothelin-1 on the adrenal gland. ANP inhibition of ACTH release might also explain the dissociation between plasma ACTH and CRH in patients who have undergone cardiac surgery or sustained myocardial infarction. Dissociation between plasma concentrations of ACTH and CRH under extreme, and especially cardiopulmonary, stress is relevant to allostatic override of the normal glucocorticoid negative feedback system.

On the basis of all the present *in vivo* and *in vitro* data, Engler and colleagues conclude that studies with ANP have so far provided the most compelling evidence that a neuropeptide may inhibit ACTH release.

Role of Hippocampus and Amygdala in Glucocorticoid Negative Feedback

It has long been assumed that the hippocampus and the amygdala, major components of the forebrain limbic system, play a key role in the stress response. The forebrain limbic system, together with the thalamus and neocortex and brain stem structures, form an analyzer-integrator system involved in neuroendocrine control. The hypothalamus can be considered an intermediate relay station in the reciprocal circuits between the limbic forebrain and brain stem structures, receiving inputs from both. The forebrain limbic system is thought to be involved in stressors that require analysis by higher brain structures (processive stressors), whereas direct brain stem projections to the PVN subserve stressors that pose an immediate physiological threat (systemic stressors).

Projections to the PVN from the hippocampus are mainly multisynaptic, involving (1) the hippocampal fimbria-fornix system involving the lateral septum, bed nucleus of the stria terminalis (BNST), and anterior hypothalamus, which all project to the parvocellular PVN, and (2) the medial corticohypothalamic tract from the anteroventral subiculum to the ventromedial, arcuate, and suprachiasmatic nuclei of the hypothalamus. The PVN also receive direct projections from the amygdala as well as projections that relay in the BNST.

Many, but not all, studies in which HPA activity was assessed by the assay of corticosterone and, less frequently, ACTH suggest that the hippocampus inhibits HPA activity. Hippocampal inhibition of the

HPA appears to be due mainly to corticosteroid negative feedback inhibition, although there is also evidence that the hippocampus may exert an inhibitory tone on the HPA independently of corticosteroid feedback. However, the hippocampus is neither the only site nor necessarily the major site of corticosteroid negative feedback, since removal of its input to the hypothalamus reduces, but does not abolish, the efficacy of corticosteroid inhibition. Thus, for example, (1) although electrical stimulation of the PVN resulted in a two- to threefold increase in CRH release into hypophyseal portal blood, stimulation of the hippocampus had no effect on the release of CRH, AVP, or oxytocin; (2) transection of the fornix, the major hippocampal-hypothalamic connection, had no significant effect on either basal or stress (nitroprusside-induced hypotension) evoked CRH release into hypophyseal portal blood; and (3) implantation of dexamethasone pellets in hippocampus had no effect on CRH synthesis in the PVN. Nevertheless, fornix section did elevate AVP concentrations in portal blood and blocked corticosterone reduction of elevated CRH, but not AVP, levels in portal blood during hypotensive stress.

Further evidence that the hippocampus normally moderates the synthesis of stress neurohormones is suggested by the finding that lateral fimbria-fornix lesions increased CRH mRNA and AVP mRNA in medial parvocellular PVN, and also plasma ACTH concentrations.

Feldman and Weidenfeld showed that, in freely moving male rats, photic and acoustic stimuli depleted the CRH content of the median eminence with a concomitant increase in plasma ACTH and corticosterone levels. This photic/acoustic activation of the HPA was inhibited by the systemic administration of dexamethasone. HPA inhibition by dexamethasone was in turn inhibited by hippocampal implants of glucocorticoid and, to a lesser extent, mineralocorticoid receptor antagonists. These results suggest that in conscious animals the hippocampus does play a role in corticosteroid feedback inhibition of stress-induced ACTH release.

Herman and associates have underscored the role of the BNST as a nucleus that integrates or processes hippocampal and amygdaloid modulation of the HPA. This hypothesis is based, first, on the anatomical connections of the BNST in that the nucleus receives rich projections from the amygdala and hippocampus and projects to the parvocellular PVN. Second, lesion of the anterior BNST resulted in a 30% decrease in CRH mRNA levels in the PVN, whereas lesion of the posterior BNST resulted in a 13% increase in the level of CRH mRNA in the PVN. On the basis of these data, Herman and colleagues

inferred that the anterior BNST integrates excitatory inputs mainly from the amygdala, whereas the posterior BNST integrates inhibitory inputs mainly from the hippocampus.

In summary, the hippocampus plays an important role in moderating the HPA both by mediating corticosteroid negative feedback and by exerting an endogenous inhibitory tone on the HPA. Corticosteroid negative feedback is also exerted directly on the PVN and the pituitary gland. The hippocampus is a heterogeneous structure; thus, for example, Dunn and Orr found that electrical stimulation of the CA1 region increased corticosterone levels, whereas stimulation of the CA3, dentate, and subiculum decreased plasma corticosterone concentrations. This heterogeneity will need to be considered in the design of further studies on the role of the hippocampus in negative feedback control of the HPA.

Glucocorticoid Negative Feedback at the Pituitary Level

As well as exerting a central effect, corticosteroids inhibit ACTH synthesis and release by an action at the level of the anterior pituitary gland. As in brain, the inhibitory action of corticosteroids on pituitary corticotropes is mediated by glucocorticoid (type II) receptors.

Studies on dispersed pituitary cells with inhibitors of mRNA and protein synthesis have shown that both the rapid and delayed glucocorticoid inhibition of ACTH release depends upon mRNA and protein synthesis. Glucocorticoids also exert potent inhibitory effects on the expression of the ACTH precursor proopiomelanocortin (POMC) in the anterior lobe of the pituitary gland. In male Sprague-Dawley rats, transcription assays showed that dexamethasone inhibited POMC gene transcription by 10-fold within 30 min of a single injection of the glucocorticoid. Inhibition of POMC transcription was paralleled by a dramatic fall in plasma ACTH concentrations. The same study showed that CRH stimulated POMC transcription by nearly twofold within 15 min, which coincided with a massive increase in ACTH release. The action of dexamethasone is cell specific in that the steroid had no effect on the transcription rate of POMC in primary cultures of neurointermediate lobe cells.

Studies in transgenic mice showed that no more than 769 base pairs of the rat POMC promoter are required for cell-specific expression and glucocorticoid inhibition of the POMC gene in the anterior pituitary gland. A good deal is known about the glucocorticoid response elements in the POMC promoter and the transcription factors involved in POMC

expression, but the precise molecular mechanism of glucocorticoid inhibition of POMC transcription remains to be elucidated.

In addition to a direct action on POMC gene expression, which determines the amount of ACTH available for release, glucocorticoids inhibit ACTH release by a mechanism mediated by adenylate cyclase (see Adenylyl Cyclases and Stress Response).

Possible Role of 11 β -Hydroxysteroid Dehydrogenase

Two types of 11 β -hydroxysteroid dehydrogenase play a pivotal role in glucocorticoid synthesis and metabolism. This enzyme is involved in glucocorticoid negative feedback (see 11 β -Hydroxysteroid Dehydrogenases).

Functional Importance of Glucocorticoid Negative Feedback

Glucocorticoid feedback inhibition of ACTH release protects the organism against the deleterious effects of hypercortisolemia (excessive concentrations of cortisol in blood). Whether due to endocrine disorders, such as Cushing's syndrome, or other causes such as trauma or chronic stress, hypercortisolemia is associated with at least three major deleterious effects. First, it suppresses the immune-inflammatory defense system and so incapacitates the animal's ability to respond to infection by pathogenic microorganisms or to chemical or physical insult. Second, persistent hypercortisolemia has major adverse effects on intermediary metabolism, resulting eventually in all the features of Cushing's syndrome, that is, android obesity, diabetes mellitus, hyperlipidemia, hypertension, and osteoporosis. Third, hypercortisolemia and/or stress are associated with reduction in hippocampal volume (atrophy) in several neuropsychiatric disorders, such as depression and posttraumatic stress disorder, as well as in Cushing's syndrome. In all three disorders, the hippocampal atrophy is associated with explicit memory deficits.

Clinical Manifestations of Disordered Glucocorticoid Feedback Regulation of the HPA System

No attempt is made here to give a detailed account of the clinical effects of disruption of feedback control within the HPA system. Rather, this section considers two clinical examples that underscore the principles of negative feedback and illustrate the consequences of disruption of normal HPA feedback control.

The first example is of enzyme defect in the adrenal cortex that results in the absence or deficiency of the afferent glucocorticoid signal of the HPA negative feedback system, whereas the second is probably due to an alteration in the central set point of the negative feedback control system.

Congenital Adrenal Hyperplasia: Failure of Glucocorticoid Negative Feedback

There are several types of inherited enzymatic defects in cortisol synthesis known to result in congenital adrenal hyperplasia (CAH), also known as the adrenogenital syndrome. By far the most common form is due to a deficiency of P450_{c₂₁} (21-hydroxylase; see Figure 7), which leads to a deficiency in cortisol biosynthesis. Excessive androgen secretion results from a failure of glucocorticoid negative feedback and consequent, uncontrolled, high ACTH secretion. Excessive androgen levels may lead to virilization of females *in utero*. About two-thirds of patients also have mineralocorticoid deficiency, resulting in salt wasting. If not obvious during the neonatal period, androgen excess may appear in early infancy, resulting in sexual precocity in boys and clitoral enlargement and pubic hair growth in girls. Excess androgen accelerates linear growth and epiphyseal closure, leading ultimately to diminished adult height. In adult women with untreated CAH, reproductive function is impaired due to (1) the disturbance of normal menstrual cycles as a consequence of the high plasma progesterone and androgen concentrations and (2) labial fusion, which prevents successful coitus. The former can be corrected by glucocorticoid replacement therapy, whereas the latter can be treated surgically.

The P450_{c₂₁} deficiency is transmitted as a single gene autosomal recessive trait linked to the major histocompatibility complex locus on the short arm of chromosome 6. An allelic variable of classical 21-hydroxylase deficiency results in a late-onset type, which frequently presents with clinical features similar to those of polycystic ovarian disease.

Deficiency of P450_{c₁₁} (11 β -hydroxylase) is a much less common cause of CAH. As in the case of P450_{c₂₁} deficiency, it is transmitted as an autosomal recessive disorder but is not linked to the HLA locus. As in the case of P450_{c₂₁} deficiency, a deficiency in P450_{c₁₁} results in impaired glucocorticoid feedback and a consequent hypersecretion of ACTH and adrenal androgens. The condition is treated with glucocorticoid replacement therapy.

Much more rare forms of CAH are produced by deficiencies of 17 α -hydroxylase and 3 β -hydroxysteroid dehydrogenase, which result in defective adrenal androgen, as well as glucocorticoid secretion.

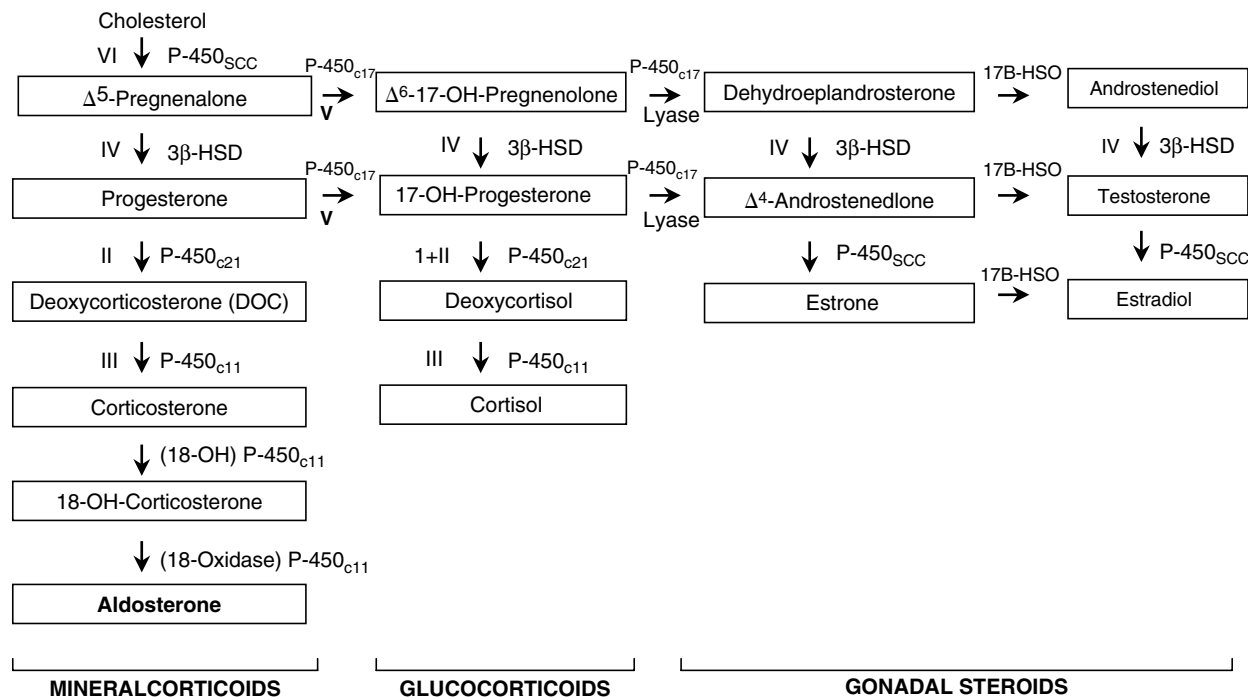


Figure 7 A diagrammatic representation of the steroid biosynthetic pathways in humans. I to VI correspond to numbers for specific biosynthetic defects that result in congenital hyperplasia. P450_{SCC}, cholesterol side chain cleavage; 3 β -HSD, 3 β -hydroxysteroid dehydrogenase/ Δ^4 - Δ^5 -isomerase; P450_{c21}, 21-hydroxylase; 17 β -HSD, 17 α -hydroxysteroid oxidoreductase; P450_{c11} catalyzes 11 β -hydroxylation, as well as 18-hydroxylation and 18-oxidation; P450_{c17} catalyzes 17-hydroxylation and 17,20-lyase activity. Subscripts refer to the steroidogenic enzymatic action of the several cytochrome P450 enzymes. From Grumbach, M. M. and Conte, F. A. (1998). Disorders of sexual differentiation. In: Wilson, J. D., Foster, D. W., Kronenberg, H. M. & Larsen, P. R. (eds.) *Williams textbook of endocrinology* (9th edn., pp.1303–1425). Philadelphia, PA: Saunders, with permission.

Hypercortisolemia in Major Depression: Possibly due to an Altered Set Point in Glucocorticoid Negative Feedback

Major depressive disorder is characterized by a significant increase in plasma cortisol concentrations (hypercortisolemia), which is most prominent at the nadir of the circadian rhythm around midnight. It was first thought that hypercortisolemia and resistance to the suppression of endogenous cortisol secretion by dexamethasone were specific features of major depression, which led to the hope that the dexamethasone suppression test could be used as a specific biological marker of depression. However, extensive studies have shown that hypercortisolemia and resistance to dexamethasone suppression are also associated with other types of psychoses, such as schizoaffective disorder and organic dementia, including Alzheimer's disease.

As shown by Goodwin and associates, hypercortisolemia in major depression is associated with a threefold increase in the mean plasma concentration of β -endorphin. In fact, resistance of β -endorphin to dexamethasone suppression appears to be a more robust marker of major depression than cortisol. Thus, in a study by Young and associates of 73 patients

with major depressive disorder, 39 (53%) showed β -endorphin nonsuppression to dexamethasone, while only 8 patients (11%) showed cortisol non-suppression. These findings suggest that hypercortisolemia in major depression is due to a resistance of the brain-pituitary-ACTH module to glucocorticoid negative feedback, that is, an elevation of the set point for glucocorticoid feedback. The precise mechanism remains to be determined, but decreased responsiveness of the limbic system, PVN, and/or pituitary gland to glucocorticoid negative feedback is a likely explanation. Reduced responsiveness of the PVN could be caused by transynaptic changes triggered by changes in function of the limbic system and frontal cortex. Because the serotonergic raphe neurons determine the amplitude of the circadian excursions of plasma ACTH and corticosterone, it is also conceivable that hypercortisolemia reflects dysregulation of serotonergic function, which seems to occur in major depression.

Allostasis and Allostatic Load

Allostasis and allostatic load are mentioned here to clarify the link between homeostasis and allostasis.

The latter refers to maintenance of stability through change or adaptation and is illustrated by the increase in blood pressure on waking or sex to cope with extra circulatory demands. In normal subjects, the increased blood pressure required to match increase in load does not necessarily reflect all but a transitory override of blood pressure control mechanisms involving baroreceptors (pressure receptors in the carotid sinus) and chemoreceptors (chemical sensors in the carotid body). Rather, there is a shift or adaptation in set point of the blood pressure control system so that it maintains mean arterial blood pressure within narrow limits around the new set point commensurate with the new load. Simply put, this is the same as altering the set point of a thermostat in a room to increase or decrease the mean ambient temperature – the set point has been changed, but the control mechanisms remain the same. The same principle may be applied to the control of plasma cortisol or adrenaline concentrations. The ability to alter the set point of one or more feedback control systems is termed regulation. In the case of the HPA and systemic blood pressure, set point regulation is located within the CNS. The elements that change the controlled variable (e.g., blood pressure, cortisol, adrenaline) toward the new set point are the control mechanisms.

Allostatic load was first introduced and defined by McEwen and Stellar (McEwen 2001) “as the cost of chronic exposure to fluctuating or heightened neural or neuroendocrine response resulting from repeated or chronic environmental challenge that an individual reacts to as being particularly stressful.” That is, allostatic load reflects the hidden cost of chronic stresses to the body that act as a predisposing factor for the adverse effects of acute, stressful life events.

Conclusion

The HPA system, together with the sympathetic-medullary system, plays a pivotal role in the neuroendocrine response to stress. Homeostasis within the HPA is maintained by a precise negative feedback system by which the adrenal glucocorticoids (the afferent inhibitory signal), cortisol in humans or corticosterone in rodents, moderate ACTH synthesis and release (efferent output signal). Allostasis – that is, change in HPA activity to cope with increased stress load – is brought about by change in feedback set point. The major sites of negative feedback are the PVN, where glucocorticoids inhibit CRH and AVP synthesis and release, and the pituitary gland, where they block the ACTH response to CRH and inhibit POMC/ACTH synthesis. The limbic system of the brain, especially the hippocampus and amygdala,

plays an important role in glucocorticoid negative feedback control.

Disruption of the HPA negative feedback system has serious deleterious effects, a point illustrated by the congenital adrenogenital syndrome and hypercortisolemia associated with serious mental illnesses. The adrenogenital syndrome, due to defective or absent cortisol secretion (loss of the afferent feedback signal) consequent on a congenital enzyme defect in the adrenal cortex, results in massive uncontrolled pituitary ACTH release. The latter induces excessive androgen production that in turn causes precocious puberty in males and masculinization of females. Hypercortisolemia, a prominent feature of major depressive disorder and other psychoses and organic dementias, is probably due to elevation of the set point of the HPA negative feedback system. Elucidation of the precise cause of this change in feedback set point may provide insight into the central disorder in depression. Hypercortisolemia may exert adverse effects by (1) inhibiting immune-inflammatory defense mechanisms, (2) disrupting intermediary metabolism, (3) inducing effects akin to Cushing’s syndrome that lead to obesity, diabetes type II, and osteoporosis, and (4) compromising hippocampal structure, neurogenesis, and function that might lead to cognitive impairment.

See Also the Following Articles

11 β -Hydroxysteroid Dehydrogenases; Adenyllyl Cyclases and Stress Response; Allostasis and Allostatic Load; Circadian Rhythms, Genetics of; Corticotropin Releasing Factor-Binding Protein; Food Shift Effect; Glucocorticoid Negative Feedback; Glucocorticoids – Adverse Effects on the Nervous System; Major Depressive Disorder; Neuroendocrine Systems; Pituitary Regulation, Role of.

Further Reading

- Bremner, J. D. and Vermetten, E. (2004). Neuroanatomical changes associated with pharmacotherapy in posttraumatic stress disorder. *Annals of the New York Academy of Science* **1032**, 154–157.
- Christie, J. E., Whalley, L. J., Fink, G., et al. (1986). Raised plasma cortisol concentrations are a feature of drug-free psychotics and not specific for depression. *British Journal of Psychiatry* **148**, 58–65.
- Christie, J. E., Whalley, L. J., Fink, G., et al. (1987). Characteristic plasma hormone changes in Alzheimer’s disease. *British Journal of Psychiatry* **150**, 674–681.
- Coplov, D. L., Rubin, R. T., Stuart, G. W., et al. (1989). Specificity of the salivary cortisol dexamethasone suppression test across psychiatric diagnoses. *Biological Psychiatry* **25**, 879–893.
- Engler, D., Redei, E. and Kola, I. (1999). The corticotropin-release inhibitory factor hypothesis: A review of the

- evidence for the existence of inhibitory as well as stimulatory hypophysiotropic regulation of adrenocorticotropin secretion and biosynthesis. *Endocrine Review* 20, 460–500.
- Feldman, S. and Weidenfeld, J. (1999). Glucocorticoid receptor antagonists in the hippocampus modify the negative feedback following neural stimuli. *Brain Research* 821, 33–37.
- Fink, G. (1979). Feedback actions of target hormones on hypothalamus and pituitary with special reference to gonadal steroids. *Annual Review of Physiology* 41, 571–585.
- Fink, G. (1988). The G. W. Harris Lecture: steroid control of brain and pituitary function. *Quarterly Journal of Experimental Physiology* 73, 257–293.
- Fink, G. (1995). The self-priming effect of LHRH: a unique servomechanism and possible cellular model for memory. *Frontiers in Neuroendocrinology* 16, 183–190.
- Fink, G. (1997). Mechanisms of negative and positive feedback of steroids in the hypothalamic-pituitary system. In: Bittar, E. E. & Bittar, N. (eds.) *Principles of medical biology* (vol. 10A), pp. 29–100. New York: JAI Press.
- Fink, G., Robinson, I. C. A. F. and Tannahill, L. A. (1988). Effects of adrenalectomy and glucocorticoids on the peptides, CRF-41, AVP and oxytocin in rat hypophyseal portal blood. *Journal of Physiology* 401, 329–345.
- Fink, G., Dow, R. C., Casley, D., et al. (1992). Atrial natriuretic peptide is involved in the ACTH response to stress and glucocorticoid negative feedback in the rat. *Journal of Endocrinology* 135, 37–43.
- Grumbach, M. M. and Conte, F. A. (1998). Disorders of sexual differentiation. In: Wilson, J. D., Foster, D. W., Kronenberg, H. M. & Larsen, P. R. (eds.) *Williams textbook of endocrinology* (9th edn., pp. 1303–1425). Philadelphia, PA: Saunders.
- Herman, J. P. and Cullinan, W. E. (1997). Neurocircuitry of stress: central control of the hypothalamo-pituitary-adrenocortical axis. *Trends in Neurosciences* 20, 78–84.
- Malberg, J. E. and Schechter, L. E. (2005). Increasing hippocampal neurogenesis; a novel mechanism for antidepressant drugs. *Current Pharmaceutical Design* 11, 145–155.
- Malkoski, S. P. and Dorin, R. I. (1999). Composite glucocorticoid regulation at a functionally defined negative glucocorticoid response element of the human corticotropin-releasing hormone gene. *Molecular Endocrinology* 13, 1629–1644.
- McEwen, B. S. (2002). Sex, stress and the hippocampus: allostasis, allostatic load and the aging process. *Neurobiology of Aging* 23, 921–939.
- Milhorn, H. T. J. (1966). *The application of control theory to physiological systems*. Philadelphia, PA: Saunders.
- Munck, A., Guyre, P. M. and Holbrook, N. J. (1984). Physiological functions of glucocorticoids in stress and their relation to pharmacological actions. *Endocrine Reviews* 5, 22–44.
- Roth-Isigkeit, A., Dibbelt, L., Eichler, W., et al. (2001). Blood levels of atrial natriuretic peptide, endothelin, cortisol and ACTH in patients undergoing coronary artery bypass grafting surgery with cardiopulmonary bypass. *Journal of Endocrinology Investigation* 24, 777–785.
- Sapolsky, R. M., Armanini, M. P., Sutton, S. W., et al. (1989). Elevation of hypophyseal portal concentrations of adrenocorticotropin secretagogues after fornix transection. *Endocrinology* 125, 2881–2887.
- Sawchenko, P. E. and Swanson, L. W. (1983). The organization of forebrain afferents to the paraventricular and supraoptic nuclei of the rat. *Journal Comparative Neurology* 218, 121–144.
- Schulkin, J. (ed.) (2004). *Allostasis, homeostasis, and the costs of physiological adaptation*, pp. 1–372. Cambridge: Cambridge University Press.
- Shepard, J. D., Liu, Y., Sassone-Corsi, P. and Aguilera, G. (2005). Role of glucocorticoids and cAMP-mediated repression in limiting corticotropin-releasing hormone transcription during stress. *Journal of Neuroscience* 25, 4073–4081.
- Tannahill, L. A., Sheward, W. J., Robinson, I. C. A. F. and Fink, G. (1991). Corticotropin-releasing factor-41, vasopressin and oxytocin release into hypophyseal portal blood in the rat; effects of electrical stimulation of the hypothalamus, amygdala and hippocampus. *Journal of Endocrinology* 129, 99–107.
- Watts, A. G., Tanimura, S. and Sanchez-Watts, G. (2004). Corticotropin-releasing hormone and arginine vasopressin gene transcription in the hypothalamic paraventricular nucleus of unstressed rats: daily rhythms and their interactions with corticosterone. *Endocrinology* 145, 529–540.
- Yates, F. E. and Maran, J. W. (1974). Stimulation and inhibition of adrenocorticotropin release. In: Knobil, E. & Sawyer, W.H. (eds.) *Handbook of physiology*, pp. 367–404. Washington, D.C.: American Physiological Society.
- Young, D. A., Kotun, J., Haskett, R. F., et al. (1993). Dissociation between pituitary and adrenal suppression to dexamethasone in depression. *Archives of General Psychiatry* 50, 395–403.

Feeding Circuitry (and Neurochemistry)

S E La Fleur, J J G Hillebrand and R A H Adan
Rudolf Magnus Institute of Neuroscience, University
Medical Centre Utrecht, Utrecht, The Netherlands

© 2007 Elsevier Inc. All rights reserved.

Neural Circuitry Involved in Food Intake
Interactions with the Stress Circuitry
Effects of Stress on Food Intake
Effects of Feeding (or Fasting) on Stress Responses
Concluding Remarks

Glossary

<i>Anorexigenic</i>	Inhibiting food intake (e.g., POMC (α -MSH), CART, leptin).
<i>Arcuate nucleus</i>	A collection of neurons (nerve cells) in the hypothalamus.
<i>Hypothalamus</i>	Region of the brain linking the nervous system to the endocrine system by synthesizing and secreting neurohormones and neuropeptides. The hypothalamus is also the area of the brain that controls hunger and thirst, body temperature, and circadian cycles.
<i>Orexigenic</i>	Stimulating food intake (e.g., AgRP, NPY, ghrelin).

Neural Circuitry Involved in Food Intake

Food intake involves several aspects of different behaviors, such as hunting for food and decision making. The complexity of feeding behavior is reflected in the number of brain areas involved; several nuclei in the caudal brain stem, hypothalamus, and the corticolimbic area are part of the feeding circuitry. Despite the large number of brain areas involved in feeding behavior, the hypothalamus is regarded as the main center for regulation of homeostatic feeding in the brain. Within the hypothalamus, the arcuate nucleus is central to the regulation of food intake. The arcuate nucleus contains at least two distinct groups of neurons controlling food intake: neurons that contain the orexigenic neuropeptides agouti-related protein (AgRP) and neuropeptide Y (NPY) and neurons that contain the anorexigenic neuropeptides pro-opiomelanocortin (POMC) and cocaine- and amphetamine-regulated transcript (CART). These arcuate neurons contain different receptors of hormones that are secreted in the periphery by adipocytes, the pancreas, and the gut in response to nutrients passing or being

stored. Leptin, for example, is secreted by adipocytes, reflecting the amount of fat stores in the body. When the body becomes fat, leptin will silence orexigenic neurons and trigger anorexigenic neurons in the arcuate nucleus to secrete α -melanocyte-stimulating hormone (α -MSH) (derived from POMC) and CART to decrease food intake. In contrast to leptin, the hormone ghrelin, which is released from the (empty) stomach and which has been associated with the anticipation of meals, activates NPY/AgRP neurons and inhibits POMC/CART neurons in the arcuate nucleus. The neurons of the arcuate nucleus project to several nuclei within the hypothalamus, such as the paraventricular (PVN) and the lateral hypothalamic (LHA) nuclei, which also receive and send information from and to the caudal brain stem and corticolimbic areas. In addition to the anorexigenic and orexigenic neuropeptides in the arcuate nucleus, within the PVN corticotropin-releasing hormone (CRH) and within the LHA melanin-concentrating hormone and orexin(s) are peptides also involved in food intake regulation. Within the brain stem, groups of catecholamine cell groups are localized that project to the hypothalamus and are important for the regulation of feeding behavior, for example, in response to a glucoprivation. In the brain stem, integration of signals from the periphery (such as cholecystokinin [CCK]) takes place, carrying information on gastrointestinal distention and presence of meals in the gastrointestinal tract and signals from the hypothalamus indicating long-term nutritional status. For instance, leptin and α -MSH affect meal size, probably by determining the sensitivity for CCK at the level of the brain stem, which is released when fat-containing foods reach the duodenum. Furthermore, within the corticolimbic area, the amygdala and nucleus accumbens are implicated in the evaluation of taste and rewarding aspects of food. Taken together, this complex neural circuitry holds many brain areas, in which neuropeptides and neurotransmitters convey information important for the regulation of feeding behavior.

The brain, however, does not function on its own. To maintain a stable body weight, food intake needs to be tuned to the needs of the peripheral body. This means that when fat stores are increasing or depleting, satiety or hunger factors (e.g., hormones) are produced by different organs to change feeding behavior. For example, leptin that is secreted from adipose tissue acts on its receptor in the arcuate nucleus to decrease food intake. In addition, glucocorticoids

secreted by adrenals in response to fasting act as factors to adjust feeding behavior.

Interactions with the Stress Circuitry

During stress, several neuropeptides and hormones are affected that are part of the previously described neural circuitry involved in feeding behavior. CRH in the PVN plays an initiating role in hypothalamic-pituitary-adrenocortical (HPA) axis activity and is increased with acute stress and decreased with chronic stress (through feedback effects of released glucocorticoids). In addition, acute stress increases NPY mRNA in the arcuate nucleus. In the amygdala, which is also part of processing taste information, NPY mRNA will decrease whereas CRH mRNA will increase after a stressful event. Thus, during stress, depending on the anatomical site, anorexigenic (CRH) as well as orexigenic (NPY) signaling is increased. This might result in conflicting food intake signals. Interestingly, some eat more while others eat less following stressful experiences. The individual sensitivity and reactivity in these different neural circuits may underlie the response of food intake to a stressful situation. It has also been shown that NPY induces changes in the daily rhythm of circulating corticosterone (which shows a rhythm over the day-night cycle) when rats' food intake is restricted. Interestingly, AgRP, but not NPY, mRNA levels are reduced following a stressful event. Thus, although both orexigenic neuropeptides are colocalized in the arcuate nucleus and their expression appears similarly affected following fasting, their expression is differently regulated following a stressor. Also, α -MSH interacts with the stress circuitry; central administration of α -MSH stimulates grooming and the release of ACTH and corticosterone in rats, and this activation is further increased in stressed rats.

Another overlapping system for food intake and stress responsivity is the catecholaminergic pathway between the brain stem and the hypothalamus, which is involved in feeding behavior as well as in stress responsivity. Physiological evidence that these systems indeed interact is provided by data showing effects of stress on food intake and vice versa.

Effects of Stress on Food Intake

Human studies showed that the psychophysiological response to stress is related to greater food consumption. Stress is the main trigger for obese patients to start binge eating. Moreover, people that release more cortisol due to a stressor are consuming more calories and eat significantly more sweet foods across days. Stress eaters gain more weight and are at greater risk

for cardiovascular disease and type II diabetes. In Cushing's disease as well as in patients treated with corticosteroids for longer periods, a typical visceral form of obesity is induced. Generally, adrenal insufficiency is accompanied by reduced food intake and body weight. Removal of the adrenal glands in rodents results in reduced food intake and body weight, which demonstrates that corticosteroids are essential for maintaining a normal body weight. More severe traumatic environmental stressors (e.g., the loss of a spouse) can have an opposite effect, decreasing food intake and fat stores. Stress is also thought to play a central role in anorexia nervosa. The neurodevelopmental model of anorexia nervosa proposes that chronic stress in predisposed individuals results in (maladaptive) hyperactivity of the HPA axis, resulting in a loss of nutritional homeostasis. Other models, however, state that the principles of anorexia nervosa are starvation and physical activity, and that the consequential activation of the HPA axis is rewarding for the patient.

In animal models, the influence of stress (ranging from physical stress to social stress) on food intake has been studied extensively. The typical response of rodents to a wide range of stressors is to decrease food intake. Only in Syrian hamsters has it been shown that social stress increases food intake and fat stores. Most experiments studying effects of stress on food intake in rats and mice have been done using normal lab chow, which is balanced, healthy, and high in carbohydrates. In daily society, however, humans have choices, and often have easy access to cheap foods with high fat and sugar content. Providing rats with a greater choice of food results in different food intake responses when animals are subjected to a chronic stressor. Rats were given access to normal lab chow, lard, and a bottle of 1 M sugar water and were repeatedly restrained for 5 days, 3 h/day. Rats increased their intake of the more palatable items, lard and sugar, and consumed more of those as compared to the nonstressed animals.

Effects of Feeding (or Fasting) on Stress Responses

Although stress alters food intake regulation, the absence of food (fasting or starvation) is itself a stressful event. For instance, stress hormones and CRH (anorexigenic) are chronically increased in anorexia nervosa patients, which may contribute to the difficulty these patients experience in trying to reverse their illness. Stress from absence of food is, however, different from stressors such as restraint stress. Glucocorticoid receptors (GRs) in the hypothalamus are downregulated following chronic, repeated restraint,

thus decreasing negative feedback on CRH mRNA in the PVN. The decreased inhibition of CRH mRNA results in fewer anorexigenic effects of the PVN CRH during starvation. Furthermore, the ACTH response to stress is altered during fasting.

Not only does the absence of food change stress responsivity, but eating palatable (e.g., sugar- and fat-containing) food also changes the response to stress. Rats that voluntarily choose the percentage of lard in their diets (composing their own high-fat diet) have reduced HPA responses to restraint stress; however, when forced to eat a high-fat diet the response of the HPA axis to restraint is high or similar compared to rats that consume only chow. Some controversy exists regarding the stress responsivity when animals are eating a high-fat diet. Although the choice to eat fat seems important for the response to stress, rats that were force-fed a high-fat diet for a longer time did have decreased responses in activity and body temperature when socially stressed (social defeat).

Concluding Remarks

Stress clearly affects the neural circuitry regulating food intake at several levels. Some people react to stress by overeating, which contributes to the development of obesity, whereas others (e.g., anorexia nervosa patients) reduce their food intake. Differences in these coping strategies may originate in subtle differences between individuals in the neural circuitry underlying food intake and stress. A better insight into how the stress system interacts with the food intake circuitry, as well as into the mechanism underlying individual responsiveness to stress, is necessary to find new strategies to treat eating disorders such as obesity and anorexia nervosa. The type of diet also affects the stress response. Future research is needed to understand the mechanisms underlying how palatable food interacts with the neural circuitry of the stress response.

See Also the Following Articles

Food Intake and Stress, Human; Food Intake and Stress, Non-Human; Food Shift Effect; Oxidative Stress; Leptin, Adiponectin, Resistin, Ghrelin; Orexin.

Further Reading

- Bergh, C. and Sodersten, P. (1996). Anorexia nervosa, self-starvation and the reward of stress. *Nature Medicine* **2**, 21–22.
- Berthoud, H. R. (2004). Mind versus metabolism in the control of food intake and energy balance. *Physiology and Behavior* **81**, 781–793.

- Buwalda, B., Blom, W. A., Koolhaas, J. M., et al. (2001). Behavioral and physiological responses to stress are affected by high-fat feeding in male rats. *Physiology and Behavior* **73**, 371–377.
- Connan, H., Campbell, I. C., Katzman, M., et al. (2003). A neurodevelopmental model for anorexia nervosa. *Physiology and Behavior* **79**, 13–24.
- Conrad, C. D. and McEwen, B. S. (2000). Acute stress increases neuropeptide Y mRNA within the arcuate nucleus and hilus of the dentate gyrus. *Molecular Brain Research* **79**, 102–109.
- Dallman, M. F., Akana, S. F., Bhatnagar, S., et al. (2000). Bottomed out: metabolic significance of the circadian trough in glucocorticoid concentrations. *International Journal of Obesity and Related Metabolic Disorders* **24**(supplement 2), S40–S46.
- Dallman, M. F., Pecoraro, N. C. and la Fleur, S. E. (2004). Chronic stress and comfort foods: self-medication and abdominal obesity. *Brain Behavior and Immunity* **19**, 275–280.
- Donohoe, T. P. (1984). Stress-induced anorexia: implications for anorexia nervosa. *Life Science* **34**, 203–218.
- Epel, E., Lapidus, R., McEwen, B., et al. (2001). Stress may add bite to appetite in women: a laboratory study of stress-induced cortisol and eating behavior. *Psychoneuroendocrinology* **26**, 37–49.
- Foster, M. T., Solomon, M. B., Huhman, K. L. and Bartness, T. J. (2006). Social defeat increases food intake, body mass and adiposity in syrian hamsters. *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology* **290**, R1284–R1293.
- Hillebrand, J. J., De Wied, D. and Adan, R. A. (2002). Neuropeptides, food intake and body weight regulation: a hypothalamic focus. *Peptides* **23**, 2283–2306.
- Inui, A. (2001). Eating behavior in anorexia nervosa – an excess of both orexigenic and anorexigenic signaling? *Molecular Psychiatry* **6**, 620–624.
- Kas, M. J., Bruijnzeel, A. W., Haanstra, J. R., et al. (2005). Differential regulation of agouti-related protein and neuropeptide Y in hypothalamic neurons following a stressful event. *Journal of Molecular Endocrinology* **35**, 159–164.
- la Fleur, S. E., Houshyar, H., Roy, M., et al. (2005). Choice of lard, but not total lard calories, damps adrenocorticotropin responses to restraint. *Endocrinology* **146**, 2193–2199.
- Makino, S., Kaneda, T., Nishiyama, M., et al. (2001). Lack of decrease in hypothalamic and hippocampal glucocorticoid receptor mRNA during starvation. *Neuroendocrinology* **74**, 120–128.
- Solomon, M. R. (2001). Eating as both coping and stressor in overweight control. *Journal of Advanced Nursing* **36**, 563–572.
- Von Frijtag, J. C., Croiset, G., Gispen, W. H., et al. (1998). The role of central melanocortin receptors in the activation of the hypothalamus-pituitary-adrenal-axis and the induction of excessive grooming. *British Journal of Pharmacology* **123**, 1503–1508.

Fetal Stress

M Eleftheriades, P Pervanidou and G P Chrousos
Athens University Medical School, Athens, Greece

© 2007 Elsevier Inc. All rights reserved.

Fetal Programming and Fetal Origins of Adult Disease
Mediators of the Predictive Adaptive Response
Fetal Distress – Clinical Evaluation
Fetal Growth Restriction (FGR)
Preterm Birth

Glossary

Biophysical profile

A noninvasive test that identifies a compromised fetus. It integrates five parameters to determine the biophysical profile score: the nonstress test, ultrasound measurement of the amniotic fluid volume, observation of the presence or absence of fetal breathing movements, gross body movements, and fetal tone.

Cardiotocography (CTG)

A method of monitoring fetal heart rate using either an abdominal transducer or a probe on the fetal scalp; in addition, another transducer placed over the uterine fundus monitors uterine activity (contractions).

Cordocentesis

An invasive procedure that involves the aspiration of fetal blood from the umbilical cord under sonographic guidance for diagnostic (fetal-karyotyping) or therapeutic (blood/platelet-transfusion) purposes.

Doppler ultrasound

A noninvasive technique of blood-flow velocity measurement. The Doppler effect (named after Christian Andreas Doppler) is the apparent change in frequency and wavelength of a wave that is perceived by an observer moving relative to the source of the waves; the total Doppler effect may result from either the motion of the source or motion of the observer. For the velocity measurement of blood, ultrasound is transmitted into a vessel and the sound that is reflected from the blood is detected.

Preeclampsia

A pregnancy-specific multisystem disorder that is characterized by the development of hypertension of 140 mmHg or higher systolic or 90 mmHg or higher diastolic after 20 weeks of gestation in a woman with previously normal blood pressure and proteinuria of 0.3 g or more of protein in a 24-h urine collection (usually corresponds to 1+ or greater on a urine dipstick test).

Programming

The process whereby a stimulus or insult, when applied during a critical or sensitive period of development, results in a long-term or permanent effect on the structure or function of the organism.

Fetal Programming and Fetal Origins of Adult Disease

Experimental and clinical studies have highlighted the importance of the fetal environment in the subsequent growth and development of an individual. Indeed, a poor *in utero* environment elicited by maternal malnutrition or placental insufficiency may increase the susceptibility of the fetus to the later development of metabolic and cardiovascular disease. The fetal origins of adult disease hypothesis originated from the epidemiological research of Barker and colleagues, which demonstrated an association between low birth weight and cardiovascular disease, hypertension, insulin resistance, and dyslipidaemia later in life. In an attempt to explain this association, Hales and Barker proposed the thrifty phenotype hypothesis, according to which the fetus in a poor intrauterine environment maximizes the uptake and conservation of fuel resources by altering its metabolism. This alteration is adaptive in a deprived environment but maladaptive in an environment of plentiful resources.

The fetal programming response to environmental stressors, predictive adaptive response (PAR), however, is much broader than the one suggested by Barker and colleagues, extending well beyond energy use and conservation. PAR uses developmentally plastic processes to set a postnatal physiological and behavioral phenotype that the fetus predicts will offer an optimal chance for survival and reproduction during adulthood. The prediction of an adverse, deprived, risky, and hence stressful environment leads to adaptive changes in body size, body composition, organ size, neurohormonal activity, and many aspects of behavior. These predictive adjustments could be advantageous in a deprived environment but are maladaptive under abundant conditions.

Mediators of the Predictive Adaptive Response

The fetal adaptive response to stressors is mediated largely by the stress system. The hypothalamic-pituitary-adrenal (HPA) axis and the locus ceruleus–norepinephrine–sympathetic nervous system (LC-NE-SNS) and their mediators corticotropin releasing

hormone (CRH), arginine vasopressin (AVP), norepinephrine (NE), epinephrine (E), and cortisol are responsible for both the concurrent adaptive response to the stressor and for causing the long-term organizational changes in the brain and periphery that constitute the PAR. Indeed, CRH and cortisol have been shown to have strong activating and organizational effects.

Fetal Distress – Clinical Evaluation

In perinatology practice, the term fetal distress has been employed as the condition in which the fetal heart rate (FHR) pattern detected by cardiotocography (CTG) is abnormal during antepartum surveillance. The generation of reactive FHR patterns constitutes a complex process involving appropriately for gestational age mature brain-stem centres, normally functioning sympathetic and parasympathetic system, intact electrical pathways to conduct the contractile stimulus, appropriately functioning cardiomyocytes' neurohormone receptors, and normal cardiac partitioning and development.

CTG tracing patterns indicative of fetal hypoxemia and/or acidemia are characterized by fixed FHR baselines, loss of FHR variability, absence of accelerations and are described as 'non-reassuring'. Documentation of spontaneous late decelerations is associated with a significant risk of fetal compromise. CTG patterns can be influenced by constitutional and environmental parameters, such as fetal congenital abnormalities, fetal maturation (depending mainly on gestational age) and movements. Moreover, maternal emotional or physical stress and the use of medication can also generate abnormal FHR patterns. Fetuses exhibit restricted central nervous system maturation and therefore cerebral activity before 28 weeks of gestation. Thus, CTG monitoring before this gestational age usually demonstrates reduced variability and benign spontaneous decelerations. Fetal sleep cycles are frequent and longer monitoring is required to record reactivity.

In addition FHR patterns are correlated to the nature of the hypoxic insult. Acute hypoxemia generates abrupt and profound declines in FHR baseline and variability and fetal activity. Chronic hypoxemia is associated with a more gradual deterioration in these parameters and therefore may remain undiagnosed by CTG monitoring for days to weeks.

Numerous studies have assessed CTG sensitivity, specificity, and predictive values. Most of them have documented relatively high specificity (>90%) but much lower sensitivity (averaging 50%). Positive and negative predictive values are less than 50% and more than 90%, respectively, suggesting that

the non stress test (NTS) is better at excluding than diagnosing fetal compromise.

Despite extensive research, the association between a compromised fetus (invasively traceable by altered biochemical parameters) and a nonreassuring CTG pattern remains an unresolved issue. In addition, the correlation between an acute or chronic intrauterine adversity and the extent of fetal compromise caused by it, is not clearly defined regarding stressor's threshold level, intensity and duration required to impair fetal developmental, physiological and metabolic mechanisms. Furthermore, fetal capability to sustain a stressor is characterized by a wide range of adaptive mechanisms whose efficacy depends on gestational age and other only partially elucidated factors.

Fetal hypoxemia can be categorized, depending on the severity and duration of the *in utero* adversity, as acute or chronic. Acute hypoxemia as a stressor activates the fetal stress system. There is evidence from experimental studies that acute hypoxemia is associated with bradycardia and compensatory tachycardia, elevated blood pressure, reduced fetal breathing and reduced gross body movements, and increased cerebral blood flow, indicating cardiac redistribution. The fetal ovine HPA axis responds to acute hypoxemia with elevated arterial plasma concentrations of ACTH and cortisol. Furthermore, catecholamines and vasopressin concentrations are elevated in the hypoxic ovine fetus. In the sheep fetus during late gestation, episodes of acute stress induce endocrine responses that promote fetal adaptation and survival during the period of adversity.

Chronic hypoxemia is associated with the progressive return of fetal body movements, heart rate, and blood pressure to normal. However, cardiac output redistribution is maintained, as documented by increased cerebral blood flow. In fetal sheep, the partial compression of the umbilical cord for a period of three days causing a 30% reduction in umbilical blood flow can produce reversible mild fetal asphyxia, a transient increase in fetal plasma ACTH concentration, and a progressive and sustained increase in fetal plasma cortisol. Chronic hypoxemia is associated with increased NE concentrations. On the contrary epinephrine concentrations are restored to control levels during the course of hypoxemic period.

Hypoxia can activate a sequence of systemic, cellular, and metabolic responses, allowing tissue adaptation to the adverse effect of the lack of oxygen. Furthermore, hypoxia can induce altered gene expression with long-term detrimental outcome. The underlying molecular mechanisms that determine fetal development in response to hypoxia are not clearly defined, and it is suggested that placental gene upregulation is associated with the inflammatory response

after the hypoxic exposure. Hypoxia may cause placental insufficiency or may induce preeclampsia through an inflammatory response, leading to FGR or even to *in utero* demise. Additionally pro-inflammatory molecules, such as cytokines, are considered responsible for the systemic maternal inflammatory response observed in pregnancies complicated with preeclampsia.

Current research suggests that exposure to inflammatory stressors during the critical developmental windows may induce altered gene expression of the neuroendocrine-immune axis. Evidence of HPA axis and autonomic nervous system (ANS) activation during inflammation supports the hypothesis that inflammatory stress during fetal development induces the programming of both the neuroendocrine and immune systems, influencing vulnerability to disease development later in life. Moreover, both clinical and experimental studies suggest that low-grade systemic inflammation is present in the metabolic syndrome participating in the development of atherosclerosis and cardiovascular disease.

Fetal Growth Restriction (FGR)

Genetically predetermined fetal growth potential influenced by environmental parameters determines fetal growth and birth weight. These parameters can be physiological (fetal and maternal health and placental function), behavioral (maternal emotional stress), social (socioeconomic status, life-style and malnutrition), and ecological (starvation/famine). The interaction between genome and environment in the regulation of fetal growth and development constitutes a complex process. Intrauterine adversities challenge the growing fetus, forcing it to reallocate its energy resources between the adaptation to the challenge and growth in order to ensure survival, reproduction enhancement and genome preservation. Thus, FGR is the evolutionary result of adaptation to a suboptimal intrauterine milieu.

FGR occurs in approximately 5% of all human pregnancies and is associated with a substantially increased risk of perinatal mortality and long-term morbidity. The definition of FGR excludes small for gestational age (SGA) fetuses that are constitutionally small. Based on growth curves, growth-restricted fetuses are considered to be those with estimated weights below the 10th percentile for gestational age, below the 5th percentile for gestational age, or two standard deviations (<3%) below the mean for gestational age. Approximately 80% of fetuses with estimated weights below the 10th percentile for gestational age are constitutionally small (normal small), 15% are starved small due to placental insufficiency,

and 5% are abnormal small. The lower the percentile that defines SGA, the higher the likelihood of FGR.

FGR is further categorized as symmetric or asymmetric. Symmetric FGR is characterized by proportional reduction in growth velocity of both the fetal head circumference (HC) and abdominal circumference (AC), indicating a constitutionally small fetus. Symmetric FGR can also be indicative of *in utero* adversities such as chromosomal abnormalities, genetic syndromes, congenital infections, use of medication or teratogenic agents. Asymmetric FGR is characterized by a disproportional reduction in growth velocity of the fetal HC to AC and is mainly associated with placental insufficiency. However, body proportionality does not always specify the causes of FGR. Snijders et al. performed fetal blood karyotyping on 458 fetuses between 17 and 39 weeks gestation who had been referred for evaluation of intrauterine growth restriction (IUGR) and they demonstrated that the relative shortening of the femur can be found in both chromosomally normal and abnormal fetuses. In the same study, fetuses with triploidy had been found to have severe, early-onset, asymmetric FGR, whereas fetuses with chromosomal abnormalities other than triploidy were symmetrically growth retarded at <30 weeks. Chromosomally abnormal IUGR fetuses diagnosed after 30 weeks are usually asymmetrically growth restricted. The authors suggested that the asymmetry results from superimposed starvation due to placental insufficiency.

Placental insufficiency and fetal abnormalities (chromosomal, structural and congenital infections) are responsible for the majority of FGR cases in singleton pregnancies. According to histopathological studies, placental insufficiency is associated with abnormal trophoblast invasion and villous development that impairs placental vascular function, restricting nutrient and oxygen transfer to the fetus. The severity of placental insufficiency can be manifested in both maternal and fetal vascular compartments, (uterine and umbilical arteries respectively). The discrepancy between fetal growth demands and environmental conditions, mainly defined by placental vascular dysfunction, determines the severity of the compromised growth velocity that may manifest as FGR or even as fetal demise.

Clinical studies have documented the correlation between increased blood flow impedance in the uterine arteries and the subsequent development of preeclampsia, and FGR. Abnormal Doppler waveform in the uterine arteries is associated with a 70% chance of developing preeclampsia and an approximately 30% chance of developing growth restriction. In FGR pregnancies increased Doppler indices in the umbilical arteries are associated with fetal hypoxemia and

acidemia correlated to the severity of the impedance of the umbilical flow. Fetal cardiovascular adaptations to placental insufficiency result in the redistribution of cardiac output. As a consequence, there is an increase in the blood supply to the brain, myocardium, and the adrenal glands at the expense of the kidneys, gastrointestinal tract, and lower extremities. The aforementioned redistribution of cardiac output can be documented by the reduced impedance of the blood flow in the middle cerebral artery. Fetuses exhibiting blood flow redistribution are usually hypoxemic, while acidemia may still not be clinically apparent. Further deterioration of fetal circulatory status is manifested by cardiac decompensation due to the exhaustion of adaptive mechanisms to hypoxia, indicated by right heart failure and abnormal organ autoregulation.

Elegant studies investigating the metabolic status of growth-restricted fetuses by cordocentesis have documented fetal compromise by hypoxemia, hypercapnia, hyperlacticemia, and acidosis. Growth-restricted fetuses are deprived of glucose and amino acids and are hypertriglyceridemic. In addition, some of these fetuses are hypoinsulinemic and their degree of hypoinsulinemia is disproportional to the degree of hypoglycemia, suggesting pancreatic dysfunction. In growth-restricted fetuses, the plasma cortisol concentration is increased and inversely correlated to fetal hypoglycemia. There is evidence of a significant correlation between umbilical cord plasma CRH and both ACTH and cortisol concentrations, as well as a significant negative correlation between CRH and dehydroepiandrosterone sulfate (DHEAS) levels in the growth-restricted fetuses. The umbilical cord plasma CRH level is extremely elevated compared to that of normal fetuses.

Placental CRH synthesis and release, in contrast to hypothalamic CRH, appears to be stimulated by glucocorticoids. In pregnancies complicated by uteroplacental insufficiency, placental CRH production may be enhanced by increased fetal glucocorticoids. In turn, placental CRH may modulate fetal pituitary-adrenal steroidogenesis to favor increased cortisol secretion. Conditions of chronic stress may stimulate hypothalamic as well as placental CRH, modulating fetal pituitary-adrenal function and consequently fetal response to a compromised intrauterine environment.

In FGR fetuses there is delayed maturation of biophysical parameters is apparent in IUGR fetuses. Fetal behavioral responses to placental insufficiency can affect heart-rate control by delaying the physiological decline of the baseline heart rate and maturation of reactivity and by decreasing short- and long-term variability.

Diagnostic Approach to Fetal Growth Restriction

The best method to diagnose the growth-restricted fetus and determine its growth velocity is ultrasound evaluation. The diagnostic approach to the FGR fetus is based on a detailed biometric and anatomic assessment evaluating for growth, chromosomal and structural defects, and congenital infections. Compared to other fetal biometric parameters, the AC is regarded as the most accurate measurement with the highest sensitivity and negative predictive value for identifying FGR. However, assessment of the estimated fetal weight (EFW) requires additional measurement of the HC and femur length.

The assessment of fetal growth velocity by serial ultrasound evaluation of AC and EFW is superior to single estimates of AC and EFW in the prediction of FGR. Moreover, it is important to assess the amniotic fluid volume. Placental insufficiency and fetal hypoxemia cause fetal oliguria and consequently oligohydramnios. Decreased amniotic fluid volume is regarded as an early sign of IUGR and oligohydramnios is defined as the presence of a largest vertical pocket of amniotic fluid less than 2 cm or amniotic fluid index (AFI) less than 5 cm.

The next essential step in the diagnostic process of FGR is based on Doppler blood-flow studies. It is well established that Doppler ultrasound assessment of the uterine and umbilical arteries helps to distinguish growth restricted fetuses due to impaired placentation from constitutionally small fetuses. In FGR fetuses, abnormal blood flow in the umbilical artery usually precedes other cardiovascular adaptations. Abnormal Doppler findings indicative of uteroplacental insufficiency include increased impedance with absent or reversed end-diastolic flow in the umbilical artery.

The further evaluation of the fetal cardiovascular system involves the assessment of the middle cerebral artery and the ductus venosus. When redistribution occurs, ductus venosus blood flow is usually normal and biophysical abnormalities are generally not clinically apparent. Abnormal ductus venosus blood flow indicates a decompensation of the cardiovascular system, which is followed by abnormalities in biophysical parameters. The biophysical profile consists of fetal tone, breathing movements, gross body movements, amniotic fluid volume assessment, and CTG monitoring, and it reflects the fetal behavioral status. The deterioration of the fetal biophysical profile associated with Doppler evidence of right heart failure indicates intrauterine fetal jeopardy. Optimizing the time of delivery, in which iatrogenic intervention should be delayed to avoid the risks of premature birth without increasing the risks of exposure to an

adverse intrauterine environment, remains a challenge for perinatologists.

Perinatal mortality and long-term morbidity, including neurodevelopmental delay, is considerably increased in FGR fetuses with evidence of abnormal ductus venosus blood flow. The delivery decision should not be based merely on the umbilical artery waveform alone but mainly on the combination of abnormal venous Doppler indices and abnormal biophysical profile parameters. Furthermore, if early delivery is required, the gestation- and birth weight-specific charts produced by a study by Draper et al. should be used to determine the likelihood of survival.

Preterm Birth

Stress experienced during pregnancy increases the risk of preterm labor. Preterm birth, i.e. delivery before 37 weeks of gestation, occurs in approximately 10% of all pregnancies constituting the leading cause of neonatal morbidity and mortality. It is noteworthy that the incidence of premature labor remains unaffected by improvements in social health care and socioeconomic level observed in Western countries during the last 50 years. A substantial body of evidence has emphasized the role of stress as a major risk factor for preterm delivery. Furthermore, a correlation has been documented between maternal stress, either physical or emotional and premature delivery and it is proposed that the earlier in pregnancy the adversity is experienced, the earlier parturition occurs.

It has been suggested that premature activation of placental CRH gene expression in response to stressors induces early fetal neuroendocrine maturation and premature labor. Indeed, maternal and placental endocrine dysfunction triggered by a variety of stressors and the consequent occurrence of fetal stress may generate the process leading to preterm delivery. Although the early expulsion of the stressed FGR fetus from an adverse intrauterine environment carries the evolutionary advantage of genome maintenance it has the cost of increased risk of perinatal morbidity and mortality as well as the development of dysmetabolic syndrome in adulthood.

Further Reading

Barker, D. J. (2002). Fetal programming of coronary heart disease. *Trends in Endocrinology and Metabolism* 13(9), 364–368.

Barker, D. J. P., Osmond, C., Forsen, T. J., et al. (2005). Trajectories of growth among children who have coronary events as adults. *New England Journal of Medicine* 353, 1802–1809.

Baschat, A. A. (2004). Pathophysiology of fetal growth restriction: implications for diagnosis and surveillance. *Obstetrical and Gynecological Survey* 59(8), 617–627.

Baschat, A. A., Gembruch, U. and Harman, C. R. (2001). The sequence of changes in Doppler and biophysical parameters as severe fetal growth restriction worsens. *Ultrasound in Obstetrics and Gynecology* 18(6), 571–577.

Challis, J. R. G., Sloboda, D., Matthews, S. G., et al. (2001). The fetal placental hypothalamic-pituitary-adrenal (HPA) axis, parturition and post natal health. *Molecular and Cellular Endocrinology* 185(1–2), 135–144.

Chrousos, G. P. and Gold, P. W. (1992). The concepts and stress and stress systems disorders. *Journal of the American Medical Association* 267, 1244–1252.

Devoe, L. D. and Jones, C. R. (2002). Nonstress test: evidence-based use in high-risk pregnancy. *Clinical Obstetrics and Gynecology* 45(4), 986–992.

Dole, N., Savitz, D. A., Hertz-Picciotto, I., et al. (2003). Maternal stress and preterm birth. *American Journal of Epidemiology* 157, 14–24.

Draper, E. S., Manktelow, B., Field, D. J., et al. (1999). Prediction of survival for preterm births by weight and gestational age: retrospective population based study. *British Medical Journal* 319(7217), 1093–1097.

Gluckman, P. D. and Hanson, M. A. (2006). The consequences of being born small – an adaptive perspective. *Hormone Research* 65(supplement 3), 5–14.

Gluckman, P. D., Hanson, M. A. and Spencer, H. G. (2005). Predictive adaptive responses and human evolution. *Trends in Ecology and Evolution* 20(10), 527–533.

Hales, C. N. and Barker, D. J. P. (2001). The thrifty phenotype hypothesis. *British Medical Bulletin* 60, 5–20.

Huang, S.-T. J., Vo, K. C. T., Lyell, D. J., et al. (2004). Developmental response to hypoxia. *FASEB Journal* 18(12), 1348–1365.

Kofman, O. (2002). The role of prenatal stress in the etiology of developmental behavioural disorders. *Neuroscience and Biobehavioral Reviews* 26(4), 457–470.

Laplanche, D. P., Barr, R. G. and Brunet, A. (2004). Stress during pregnancy affects general intellectual and language functioning in human toddlers. *Pediatric Research* 56(3), 400–410.

Latini, G., De Mitri, B., Del Vecchio, A., et al. (2004). Foetal growth of kidneys, liver and spleen in intrauterine growth restriction: “programming” causing “metabolic syndrome” in adult age. *Acta Paediatrica* 93(12), 1635–1639.

McMillen, I. C. and Robinson, J. S. (2000). Developmental origins of the metabolic syndrome: prediction, plasticity and programming. *Physiological Reviews* 85, 571–633.

Nicolaides, K. H., Bilardo, C. M., Soothill, P. W., et al. (1988). Absence of end diastolic frequencies in umbilical artery: a sign of fetal hypoxia and acidosis. *British Medical Journal* 297(6655), 1026–1027.

Ozanne, S. E. and Hales, C. N. (2002). Early programming of glucose-insulin metabolism. *Trends in Endocrinology and Metabolism* 13(9), 368–373.

- Phillips, D. (2002). Endocrine programming and fetal origins of adult disease. *Trends in Endocrinology and Metabolism* 13(9), 363.
- Seckl, J. R. (2001). Glucocorticoid programming of the fetus; adult phenotypes and molecular mechanisms. *Molecular and Cellular Endocrinology* 185(1–2), 61–71.
- Snijders, R. J. M., Sherrod, C., Gosden, C. M., et al. (1993). Fetal growth retardation: associated malformations and chromosome abnormalities. *American Journal of Obstetrics and Gynecology* 168(2), 547–555.
- Smith, J. F., Jr. and Onstad, J. H. (2005). Assessment of the fetus: intermittent auscultation, electronic fetal heart rate tracing, and fetal pulse oximetry. *Obstetrics and Gynecology Clinics of North America* 32(2), 245–254.
- Soothill, P. W., Nicolaides, K. H. and Campbell, S. (1987). Prenatal asphyxia, hyperlacticaemia, hypoglycaemia, and erythroblastosis in growth retarded fetuses. *British Medical Journal Clinical Research Edition* 294(6579), 1051–1053.
- Ward, A. M., Syddall, H. E., Wood, P. J., et al. (2004). Fetal programming of the hypothalamic-pituitary-adrenal (HPA) axis: low birth weight and central HPA regulation. *Journal of Clinical Endocrinology and Metabolism* 89(3), 1227–1233.

Fever See: Febrile Response.

Fibrinogen and Clotting Factors

E Brunner

University College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by E Brunner, volume 2, pp 138–142, © 2000, Elsevier Inc.

Restriction site polymorphisms

The differences in the DNA base sequence of a gene identified by fragmentation of the gene with a DNA-cutting enzyme.

Thrombus

A blood clot; formed by platelet aggregation and the conversion of soluble fibrinogen to an insoluble mesh of polymerized fibrin.

Molecular Biology
Physiology
Pathophysiology
Epidemiology

Glossary

<i>Acute-phase reactant</i>	A blood constituent (e.g., C-reactive protein, leukocytes, fibrinogen, or albumin) that shows a transient change in concentration in response to infection or inflammation.
<i>Endothelium</i>	The tissue lining the inside surface of blood vessels.
<i>Fibrinolytic activity</i>	A measure of the ability of blood to dissolve a thrombus.
<i>Mendelian randomization</i>	An approach to testing causal associations in epidemiology using the influence of a common genetic polymorphism on the risk factor being studied.
<i>Monocyte/macrophage</i>	A type of phagocytic leukocyte.

Fibrinogen is a large protein synthesized by the liver that makes up some 5% of total plasma protein. Various stimuli, including injury, activate a sequence of clotting factors that cause soluble fibrinogen to be converted enzymically to fibrin, an insoluble polymer matrix that is the structural component of blood clots. Raised fibrinogen levels indicate an increased risk of ischemic heart disease and stroke. It is uncertain whether fibrinogen is a risk factor that causes cardiovascular disease (CVD) or a marker of developing disease. Fibrinogen is an acute-phase reactant, and the circulating level increases in response to infection, chronic inflammation, smoking, and other environmental stressors.

Molecular Biology

Structure

Fibrinogen is a large soluble protein of approximately 340 kDa. It is made up of two tripeptide units, each

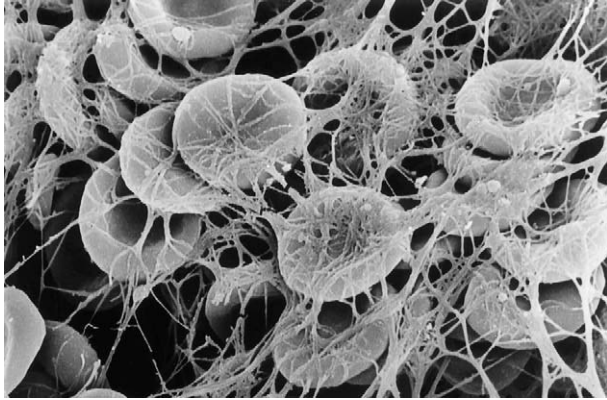


Figure 1 High-resolution electron micrograph of a fibrin polymer matrix. Red blood cells are trapped within the mesh. Courtesy of Dr. D. W. Gregory, University of Aberdeen/Wellcome Trust Medical Photographic Library.

consisting of $A\alpha$, $B\beta$, and γ peptides held together by covalent disulfide bonds. The three-dimensional structure has the form of a rod of dimensions 7×48 nm with a globular domain at each end and a third globular domain at the center of the rod, where the two units are joined together by further disulfide bonds. Free N-terminal residues project radially from the central domain. This negatively charged region prevents the aggregation of fibrinogen molecules by electrostatic repulsion. Fibrin is formed when the four N-terminal peptides are cleaved to release fibrinopeptides A and B. With the loss of the negatively charged peptides, staggered end-to-end polymerization takes place to produce fibrin fiber bundles (Figure 1). The clot is completed by the reinforcement of the fibers with covalent cross-links. More than 20 enzymes and protein cofactors are involved in clot formation and dissolution. Blood platelets interact with and possess receptors for fibrinogen. These nonnucleated bodies originate from bone marrow cells (megakaryocytes) and circulate at a concentration of some $150\text{--}500\,000\text{ mm}^{-3}$.

Biosynthesis and Genetics

Fibrinogen is synthesized primarily by hepatic parenchymal cells and to a lesser extent by megakaryocytes. The fibrinogen peptides are encoded on chromosome 4. A 50-kb span contains the genes for the $A\alpha$, $B\beta$, and γ chains separated by untranslated base sequences. These promoter sequences appear to coordinate the expression of the three genes. After translation and assembly, with the formation of the disulfide bonds, the molecule is glycosylated before release into the circulation from secretory vacuoles.

Serious congenital abnormalities of fibrinogen biosynthesis are uncommon. Absent or very low levels of fibrinogen have been identified, but congenital

hyperfibrinogenemia has not. Major abnormalities are probably due to translation, transcription, or secretion defects rather than gross fibrinogen gene defects. Less dramatic abnormalities are associated with bleeding or thrombosis, or may be asymptomatic. Several restriction site polymorphisms exist on the α and β genes. Common genetic variants have been shown to account for some 15% of within-population differences in plasma fibrinogen levels and may explain susceptibility to CVD in some families. Genetic epidemiology uses this variation in association studies of unrelated individuals to test the causal link with coronary heart disease (CHD), applying the method of Mendelian randomization.

Physiology

The Clotting System

The clotting, or coagulation, system functions to maintain the integrity of blood vessels and is essential to homeostasis. Two mechanisms initiate formation of a thrombus: (1) injury of the vascular lining (endothelium) to expose anionic surfaces causes binding and activation of factor XII, the first step in the intrinsic pathway, and (2) exposure of the transmembrane protein factor III or tissue factor (TF) at the site of injury leads to the initiation of the extrinsic pathway (so called because it relies on a substance not usually present in the circulation).

The main reaction of the extrinsic pathway involves the binding of factor VII and calcium ions (Ca^{2+}) to the exposed tissue factor receptor to form the catalyst TF-VIIa-Ca^{2+} , which converts inactive factor X to active Xa. The cascade of four reactions in the intrinsic pathway also generates an increase in the local concentration of factor Xa. If each step in this pathway activates 100 enzyme molecules, the initial event will produce a signal amplified by 10^6 . The two clotting pathways lead to the formation of the same prothrombinase complex (Va-Xa-Ca^{2+} -phospholipid), which catalyses the hydrolysis of prothrombin to thrombin. Thrombin has three major roles: it splits fibrinopeptides A and B from the $A\alpha$ and $B\beta$ chains of fibrinogen to form fibrin and promotes the aggregation of platelets at the site of injury. The clotting process is self-controlling by means of inhibitory mechanisms in which thrombin and protein C are central.

Blood Platelets

Low platelet counts are associated with capillary bleeding, and high counts are associated with intravascular thrombosis. Platelet aggregation is initiated in two ways. First, thrombin and receptors on damaged endothelial cells cause the exposure of a receptor

for fibrinogen (integrin α_{IIb}/β_3 or glycoprotein GPIIb/IIIa) on the platelet membrane. Fibrinogen binds the activated platelet to other platelets or to collagen. Second, high shear conditions, such as those occurring in narrowed atheromatous vessels, activate a second receptor that binds von Willebrand factor (vWF). This glycoprotein is released by activated platelets to form links between platelets and to anchor the growing thrombus to the vessel wall. Aspirin inhibits the aggregation mechanism involving fibrinogen but not shear-induced aggregation involving vWF.

The Fibrinolytic System

The thrombus is broken down to restore normal blood flow by the proteolytic action of plasmin on the fibrin matrix when, for example, the wound has healed. This enzyme is converted from the inactive form by tissue plasminogen activator (t-PA), which is secreted by endothelial cells of the vessel wall. Fibrin degradation products are detectable in the blood of healthy individuals, indicating that thrombus formation and fibrinolysis proceed continually. Inhibitors of fibrinolysis include lipoprotein (a) and plasminogen activator inhibitor.

Blood Viscosity

The plasma level of fibrinogen is an important influence on the flow properties of blood. Plasma viscosity increases exponentially with fibrinogen concentration. Because blood contains cells, among which the erythrocyte (red blood cell) predominates at some $5 \text{ million mm}^{-3}$, it is a non-Newtonian fluid exhibiting an apparent viscosity depending on local flow conditions. Viscosity is lower when blood flow is turbulent and under high shear conditions; it is higher under the opposite conditions. The blood flow rate is inversely related to viscosity and directly proportional to the fourth power of the vessel diameter. A small change in diameter thus produces a large change in tissue perfusion.

Pathophysiology

Thrombus Formation

Thrombus formation is involved in several important causes of death and disability. Coronary thrombosis may precipitate sudden coronary death or heart attack, whereas cerebrovascular thrombosis leads to stroke. A small proportion of women possess an abnormally sensitive clotting system and an increased tendency to form venous thrombi in response to raised circulating estrogen levels with oral contraceptive use or in pregnancy. Clot-busting or thrombolytic

drugs, such as recombinant t-PA, administered within hours of heart attack are effective in restoring blood supply to the damaged heart muscle and in reducing subsequent mortality.

Atheroma Formation

The clotting system is involved in the slow evolution of atheroma, the plaques that produce arterial narrowing. Fibrinogen and fibrin are found in atherosclerotic lesions. Histological studies show that small thrombi formed at the vessel wall can be incorporated into the evolving plaque. Platelets or platelet fragments are detectable in many plaques. Fibrinogen and its degradation products interact with platelets, endothelial and smooth muscle cells, and monocytes. These processes include the transfer of cholesterol from lipoproteins and platelets to monocytes to form giant foam cells and their incorporation into the growing plaque.

Acute-Phase Reaction

The acute response to infection or inflammation includes an increase in plasma fibrinogen concentration. This is, in part, stimulated by a rise in monocyte secretion of interleukin (IL)-6, one of a group of inflammatory mediators known as the cytokines. An IL-6-responsive element has been identified in the promoter region of the β -fibrinogen gene, and IL-6 thus appears to be a major regulator of fibrinogen synthesis. The size of the inflammatory response appears to be tuned by a vagally mediated anti-inflammatory pathway. Parasympathetic outflow via cholinergic neurons inhibits release of pro-inflammatory cytokines from liver, spleen, and other organs of the reticuloendothelial system. The connection between the nervous and immune systems has been demonstrated in animal models employing direct electrical stimulation of the efferent vagus nerve.

The higher fibrinogen levels seen in smokers may be due to chronic inflammation of the pulmonary epithelium and consequent raised IL-6 production. Long-term low-level inflammatory processes may also be involved in atheroma development as a result of chronic asymptomatic infections, such as gastric *Helicobacter pylori* infection.

Epidemiology

Cardiovascular Epidemiology

It is not clear whether hemostasis is causally involved in early development of coronary atherosclerosis, as opposed to the thrombotic processes that precede an acute CVD event. Internationally, levels of hemostatic

factors reflect underlying rates of CHD. Fibrinogen levels measured around 1980 were high in Czechoslovakia and Finland and low in urban Gambia. Mean fibrinolytic activity was highest in Gambia, intermediate in Scotland, and lowest in Finland. Plasma fibrinogen measured in the United States and Japan in 1987 showed substantially higher levels in White American men than in Japanese men. The difference was evident in nonsmokers; among Japanese nonsmokers mean fibrinogen was 2.2 g l^{-1} , whereas among American nonsmokers it was 2.8 g l^{-1} . These ecological associations, which parallel rates of disease, could be the product of cause or consequence.

At the individual level, a high plasma fibrinogen concentration in adulthood has been shown to predict CVD in a meta-analysis of 31 large cohort studies. Other hemostatic factors, including factor VII, white blood cell count, plasma viscosity, and fibrinolytic activity, predict CHD, although the evidence is less extensive. Fibrinogen level may be a mediator of the relationship between smoking and CVD. Likewise, the alcohol–fibrinogen relationship is consistent with the protective effect of moderate alcohol consumption on heart attack risk, and the fibrinogen level is lower in those taking regular leisure time physical activity. Fibrinogen remains predictive of CVD after taking into account smoking, blood pressure, body mass index, and serum cholesterol level. However, from prospective observational studies this evidence does not constitute proof of causation. The alternative interpretation, that fibrinogen is not a causal factor but a marker of the disease process, remains plausible. One deficiency in the evidence is the lack of fibrinogen-lowering trials that might demonstrate reduction of coronary risk.

Mendelian Randomization

Evidence that fibrinogen, and by implication hemostasis, is not causal for CVD comes from genetic epidemiology. Employing the method of Mendelian randomization, the lifelong influence of a common genetic variant in the promoter region of the β -fibrinogen gene that raises the plasma fibrinogen level was found to be without effect on CHD risk. In the same large case-control study, as in many other observational studies, plasma fibrinogen levels were substantially higher in coronary cases. The conflicting findings for phenotype and genotype suggest that plasma fibrinogen is raised as a result of the presence of developing atherosclerotic plaques. If replicated, the conclusion that raised fibrinogen is a consequence, rather than a cause of disease, becomes more likely.

Infancy and Childhood

Poor intra-uterine and early postnatal development are important elements in the genesis of adult disease. In this context, poor early growth and poor childhood socioeconomic conditions are related to elevated fibrinogen levels in adulthood. In addition, height in adulthood (reflecting early environment and genetic endowment) is inversely related to fibrinogen level and to cardiorespiratory mortality risk. These findings are consistent with the long-term *in utero* and childhood influences on CVD risk. It remains unclear whether the association with hemostatic factors is a direct result of early life conditions or the indirect consequence of accelerated disease processes that raise fibrinogen levels secondary to damage of the arterial wall. If the latter explanation emerges as the correct understanding, the measurement of hemostatic factors will remain important as a marker of subclinical disease in studies of the early origins of CVD.

Social Position and Psychosocial Factors

Plasma fibrinogen level is related to several social and psychosocial factors in adulthood. This is consistent with a contribution of hemostasis to the high rates of coronary disease experienced by people in unfavorable socioeconomic circumstances or, alternatively, with plasma fibrinogen being a marker of socially patterned disease processes. Mean fibrinogen levels are strongly and inversely related to grade of employment in the British civil service (Figure 2) and to social class based on occupation in a stratified sample of the population of England and Wales. Among civil servants, approximately one-half of the inverse association with social position was accounted for by health-related behaviors and degree of obesity.

Low job control, a measure of an adverse psychosocial work environment, has been linked to higher fibrinogen levels in both sexes in several studies. This association has been found when the work environment was assessed either subjectively, using standard self-completion questionnaires, or externally, by personnel managers. Complementing the link with adverse psychological conditions, positive affect is associated with a small plasma fibrinogen response to behavioral stress tasks (color-word and mirror tracing) in the laboratory.

It is plausible that psychosocial factors exert effects on the hemostatic system and thereby influence disease risk. Acute effects include increased platelet stickiness in the fight-or-flight response via increased circulating catecholamines. In a study of astronauts, spaceflight evoked raised urinary IL-6 on the first day

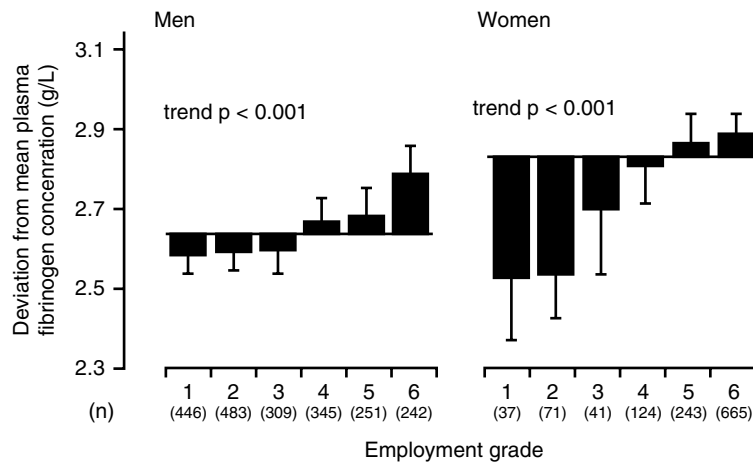


Figure 2 Plasma fibrinogen levels by employment grade (the Whitehall II study). Civil service employment grade classifications: 1, unified grades 1–6; 2, unified grade 7; 3, senior executive officer; 4, higher executive officer; 5, executive officer; 6, clerical/office support. Reproduced with permission from Brunner, E. J., Davey Smith, G., Marmot, M. G., et al. (1996), Childhood social circumstances and psychosocial and behavioural factors as determinants of plasma fibrinogen, *Lancet* **347**, 1008–1013. © The Lancet Ltd.

but not on other days before or after launch, which is consistent with a psychological explanation because infection does not appear to be implicated. Cytokine release in response to psychosocial stimuli, and thereby increased fibrinogen synthesis, is feasible in view of the observation that mRNA for IL-6 is expressed within the brain. Over decades, such small effects may add to the formation of arterial plaques to increase risks of heart disease and stroke.

Further Reading

- Brunner, E. J., Davey Smith, G., Marmot, M. G., et al. (1996). Childhood social circumstances and psychosocial and behavioural factors as determinants of plasma fibrinogen. *Lancet* **347**, 1008–1013.
- Danesh, J., Collins, R., Appleby, P., et al. (1998). Association of fibrinogen, C-reactive protein, albumin, or leukocyte count with coronary heart disease. *Journal of the American Medical Association* **279**, 1477–1481.
- Dang, C. V., Bell, W. R., Baltimore, M. D., et al. (1989). The normal and morbid biology of fibrinogen. *American Journal of Medicine* **87**, 567–576.
- Davey Smith, G., Harbord, R. and Ebrahim, S. (2004). Fibrinogen, C-reactive protein and coronary heart disease: does Mendelian randomization suggest the

associations are non-causal? *Quarterly Journal of Medicine* **97**, 163–166.

Fibrinogen Studies Collaboration (2005). Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant meta-analysis. *Journal of the American Medical Association* **294**, 1799–1809.

Meade, T. W., Ruddock, V., Stirling, Y., et al. (1993). Fibrinolytic activity, clotting factors, and long-term incidence of ischaemic heart disease in the Northwick Park Heart Study. *Lancet* **342**, 1076–1079.

O'Brien, J. R. (1990). Shear-induced platelet aggregation. *Lancet* **335**, 711–713.

Rabbani, L. E. and Loscalzo, J. (1994). Recent observations on the role of hemostatic determinants in the development of the atherothrombotic plaque. *Atherosclerosis* **105**, 1–7.

Stein, T. P. and Schluter, M. D. (1994). Excretion of IL-6 by astronauts during spaceflight. *American Journal of Physiology* **266**, E448–E552.

Steptoe, A., Wardle, J. and Marmot, M. (2005). Positive affect and health-related neuroendocrine, cardiovascular, and inflammatory processes. *Proceedings of the National Academy of Sciences* **102**, 6508–6512.

Vaisanen, S., Rauramaa, R., Penttila, I., et al. (1997). Variation in plasma fibrinogen over one year: relationships with genetic polymorphisms and non-genetic factors. *Thrombosis and Haemostasis* **77**, 884–889.

Fibromyalgia

P B Wood

Louisiana State University Health Sciences Center,
Shreveport, LA, USA and McGill University, Montreal,
Quebec, Canada

© 2007 Elsevier Inc. All rights reserved.

Introduction
Background
Epidemiology
Etiology: The Role of Stress
Experimental Findings
Hypothetical Mechanisms of Abnormal Pain Processing
Future Directions

Glossary

<i>Allodynia</i>	Pain produced by normally nonpainful stimulation such as touch, gentle pressure, cold, or gentle joint movement.
<i>Central sensitization</i>	The development within the central nervous system of an exaggerated, pathological response to a stimulus that was originally innocuous, which depends on plasticity in function of <i>N</i> -methyl-D-aspartate (NMDA) glutamate receptors. Abnormally increased pain sensation.
<i>Hyperalgesia</i>	Class of glutamate receptors that act as ion channels characterized by their affinity for NMDA that are crucial for neural plasticity.
<i>N-methyl-D-aspartate (NMDA)-receptors</i>	The perception of pain.
<i>Nociception</i>	An unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage.
<i>Pain</i>	Discrete areas of hyperalgesia/allodynia where localized pressure will elicit pain.
<i>Tender points</i>	A frequency-dependent increase in the nervous system response to repeat stimulation thought to be dependent on glutamate and NMDA receptors in the dorsal horn of the spinal cord; temporal summation to pain produced by repeat stimulation.
<i>Wind-up</i>	

Introduction

Fibromyalgia is a disorder characterized by chronic widespread pain and tenderness to light palpation, sleep disturbance, and chronic fatigue. A diagnosis of fibromyalgia is made by the presence of diffuse

tenderness to light palpation in nonarticular soft tissues in a characteristic distribution in the absence of specific laboratory or radiological abnormalities, along with a history that typically includes exposure to significant stressors. Fibromyalgia can occur as an isolated condition, sometimes termed primary fibromyalgia, or it may occur in the presence of other painful conditions such as rheumatoid arthritis, a condition known as secondary fibromyalgia. In addition to the core features of chronic widespread pain and fatigue, patients are typically afflicted by a variety of other symptoms involving multiple organ systems, which has led in turn to the development of a more comprehensive construct: the fibromyalgia syndrome.

Background

The current concept of fibromyalgia is primarily derived from the work of Gowers and Stockman, who, in 1904, proposed that chronic widespread pain likely resulted from fibrositis, or diffuse inflammation of fibrous connective tissues. Subsequent analysis of said tissues has failed to demonstrate such changes, and, in fact, no peripheral pathology has been elucidated to date, despite extensive investigations. Given the lack of inflammation, the name was subsequently changed to fibromyalgia, which literally translated means simply pain in the muscles and fibrous tissues. The diagnostic criteria for classification and research purposes were formalized by the American College of Rheumatology, based on findings of the Multicenter Criteria Committee in 1990, which rely on the presence of discrete areas of tenderness to light palpation occurring diffusely in a primarily axial distribution that have been present for a minimum of 3 months (see [Figure 1](#)).

In the original documentation of the disorder, patients with fibromyalgia were noted to experience, in addition to chronic widespread pain, symptoms in multiple systems including disturbed bowel function, genitourinary problems, and allergic phenomena such as sinus drainage and congestion. An increased incidence of affective- and anxiety-related symptoms is also well known. More recently, the Canadian Case Definitions has been published, which places an increased emphasis on a variety of neurological phenomena associated with the disorder, including non-dermatomal paresthesias, headache, temporospatial dysmetria, and restless legs syndrome. In addition, patients frequently complain of cognitive dysfunction (commonly referred to as “fibro fog”) characterized

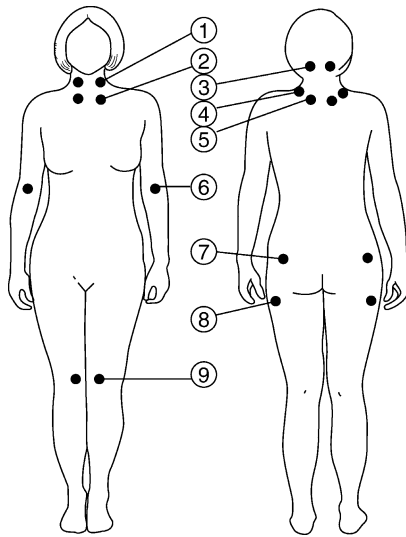


Figure 1 A diagnosis of fibromyalgia is based on (1) the presence of chronic widespread pain of at least 3 months duration in a distribution above and below the waistline, on the left and right sides of the body, and in a primarily axial distribution and (2) pain on palpation with a 4-kg force in 11 of 18 tenderpoints.

by impaired concentration and short-term memory consolidation, impaired speed of performance, inability to multitask, and frequent cognitive overload.

The combination of a lack of readily demonstrable pathology to which symptoms might be ascribed and the dependence on subjective patient report has fueled much of the controversy surrounding fibromyalgia and led to large segments of the medical community dismissing such patients as neurotic complainers aspiring to fashionable diagnoses. However, although the exact pathology underlying the disorder's cardinal feature, namely, chronic widespread pain, remains one of medicine's enduring mysteries, recent advances have shed increasing light on the nature of the disorder. As investigations into potential peripheral pathologies have remained largely fruitless, increasing attention has been focused on mechanisms within the central nervous system. Accordingly, fibromyalgia is increasingly conceptualized as an aberration of information processing in pain-related pathways and thus is considered a form of central sensitization.

Epidemiology

Fibromyalgia is a common condition with an estimated prevalence of 2% of the general population: 3.4% for women and 0.5% for men. Although all ages are affected, the prevalence increases with age, reaching greater than 7% in women aged 60 to 79 years. Fibromyalgia is the second most common diagnosis in rheumatology clinics. Given that the vast majority of fibromyalgia patients are female, many

researchers have suggested that gender-specific factors likely play a role in the development and/or maintenance of symptoms.

Etiology: The Role of Stress

Fibromyalgia has been characterized as a stress-related disorder due to the recognition of its frequent onset after acute or chronic stressors and apparent exacerbation of symptoms during periods of physical or emotional stress. Factors associated with the onset of fibromyalgia include severe infectious illness, physical trauma, and severe emotional distress. In addition, many of the comorbid conditions that affect fibromyalgia patients are likewise considered to be stress-related disorders, including irritable bowel syndrome, chronic fatigue syndrome, interstitial cystitis, and such neuropsychiatric disorders as depression and posttraumatic stress disorder. Patients with fibromyalgia also report a higher incidence of adverse early life events than unaffected individuals. While the experience of adverse early life events has been demonstrated to produce a variety of changes to the neuroendocrine function and to have an adverse impact on the brain limbic system in susceptible individuals, the specific contribution of these events to the etiology of the disorder is not clear, nor is there a strict correlation between the experience of adverse early life events and the development of chronic widespread pain.

While a number of phenomena associated with fibromyalgia have been attributed to the stress of chronic pain, basic science has demonstrated that the experience of chronic stress per se produces harmful sequelae such as disrupted immune and neuroendocrine function, alterations of neurotransmitter release, and damage to the very substance of the brain itself. These observations have in turn informed some of the more recent theoretical considerations concerning the etiopathogenesis of the disorder, which are explored later. In addition, a growing body of evidence supports a component of genetic heritability in the development of the disorder. Taking into account the contribution of both genetic as well as environmental factors, the development of fibromyalgia likely results from a combination of genetic predisposition and subsequent exposure to critical environmental conditions that must reach a critical load in order to initiate the changes that result in the development of chronic widespread pain and tenderness. (One might therefore suggest that what fibromyalgia represents is a rather bizarre form of allostasis.) As such, individuals with high susceptibility would require little or no stress in order for symptoms to develop, while others might endure years of duress prior to symptom onset.

This suggestion raises the possibility of a prodromal period during which subtle changes are occurring within the susceptible nervous system during periods of prolonged stress.

Experimental Findings

An unfortunately intransigent misconception among many clinicians is the notion that the patient with fibromyalgia essentially has nothing wrong with him or her, a position that stems in large part from a frustrating lack of readily demonstrable pathology in the soft tissues or blood. On the contrary, research has demonstrated a variety of abnormalities are in fact associated with the disorder, which are detailed below.

Polymodal Sensitivity

Results from studies examining responses to experimental stimulation have shown that fibromyalgia patients display sensitivity to pressure, heat, cold, and electrical and chemical stimulation. Experiments examining pain regulatory systems have shown that fibromyalgia patients also display a dysregulation of diffuse noxious inhibitory control, an exaggerated wind-up response, and an absence of exercise-induced analgesic response. Together, these results point to dysregulation of the nociceptive system at the central level.

Sleep Disturbances

The first objective findings associated with the disorder were reported in 1975 by Moldofsky and colleagues, who reported the presence of anomalous alpha wave activity (typically associated with arousal states) on sleep electroencephalogram (EEG) in fibromyalgia patients during non-rapid eye movement sleep. In fact, by consistently disrupting stage IV sleep in young, healthy subjects, Moldofsky was able to reproduce a significant increase in muscle tenderness similar to that experienced by fibromyalgia patients but that was resolved when the subjects were able to resume their normal sleep patterns. Since that time a variety of other EEG sleep abnormalities has also been reported in subgroups of fibromyalgia patients. While chronic sleep disruption has been documented to disrupt normal brain function and produce a variety of deleterious neuroendocrine and metabolic consequences, the specific mechanism whereby increased tenderness might result remains to be determined, as does the exact relationship of sleep disruption to the disorder, whether as a contributing factor or rather an additional manifestation of an underlying pathology.

Neuroendocrine Disruption

Patients with fibromyalgia have been demonstrated to have a disruption of normal neuroendocrine function, characterized by mild hypocortisolemia, hyperactivity of pituitary adrenocorticotropin hormone release in response to challenge, and glucocorticoid feedback resistance. A progressive reduction of serum growth hormone levels has also been documented – at baseline in a minority of patients, while most demonstrate reduced secretion in response to exercise or pharmacological challenge. Other abnormalities include reduced melatonin secretion, reduced responsiveness of thyrotropin and thyroid hormones to thyrotropin-releasing hormone, and a mild elevation of prolactin levels with disinhibition of prolactin release in response to challenge and hyposecretion of adrenal androgens. These changes might be attributed to the effects of chronic stress, which, after being perceived and processed by the central nervous system, activates hypothalamic corticotropin-releasing hormone neurons. Thus, the multiple neuroendocrine changes evident in fibromyalgia have been proposed to stem from chronic overactivity of corticotropin-releasing hormone releasing neurons, resulting in a disruption of normal function of the pituitary-adrenal axis and an increased stimulation of hypothalamic somatostatin secretion, which, in turn, inhibits the secretion of a multiplicity of other hormones.

Sympathetic Hyperactivity

Functional analysis of the autonomic system in patients with fibromyalgia has demonstrated disturbed activity characterized by hyperactivity of the sympathetic nervous system at baseline with reduced sympathoadrenal reactivity in response to a variety of stressors including orthostatic challenge and mental stress. Fibromyalgia patients demonstrate lower heart rate variability, an index of sympathetic/parasympathetic balance, indicating sustained sympathetic hyperactivity, especially at night. In addition, plasma levels of neuropeptide Y, which is colocalized with norepinephrine in the sympathetic nervous system, have been reported to be low in patients with fibromyalgia. Incongruent with this finding, there have been inconsistent reports of low, normal, and high circulating levels of epinephrine and norepinephrine. Administration of interleukin-6, a cytokine capable of stimulating the release of hypothalamic corticotropin-releasing hormone, which in turn stimulates the sympathetic nervous system, resulted in a dramatic increase in circulating norepinephrine levels and a significantly greater increase in heart rate over baseline in fibromyalgia patients as compared to healthy controls.

Cerebrospinal Fluid Abnormalities

Perhaps the most reproducible laboratory finding in patients with fibromyalgia is an elevation in cerebrospinal fluid levels of substance P, a putative nociceptive neurotransmitter. While this finding is intriguing, a review of the literature reveals that an elevation in substance P levels is not specific to this disorder, nor has any correlation between these levels and clinical pathology been determined. In addition, metabolites for the monoamine neurotransmitters serotonin, norepinephrine, and dopamine – all of which play a role in natural analgesia – have been shown to be lower, while reports of changes in cerebrospinal fluid concentrations of endogenous opioids (i.e., endorphins and enkephalins) have been contradictory. The mean concentration of nerve growth factor, a substance known to participate in structural and functional plasticity of nociceptive pathways within the dorsal root ganglia and spinal cord, has been reported to be elevated in fibromyalgia patients. In addition, evidence for increased excitatory amino acid release within cerebrospinal fluid has also been reported, with a correlation between levels for metabolites of glutamate and nitric oxide and clinical indices of pain.

Brain Imaging Studies

Evidence of abnormal brain involvement in fibromyalgia has been provided via functional neuroimaging. The first findings reported a decreased blood flow within the thalamus and elements of the basal ganglia and mid-brain (i.e., pontine nucleus) in patients with fibromyalgia. Differential activation in response to painful stimulation has also been demonstrated. Brain centers showing hyperactivation in response to noxious stimulation include such pain-related brain centers as the primary and secondary somatosensory cortex, anterior cingulate cortex, and insular cortex, while relative hypoactivation at subjectively equal pain levels appears to occur within the thalamus and basal ganglia. In addition, patients exhibit neural activation in brain regions associated with pain perception in response to nonpainful stimuli, in such areas as the prefrontal, supplemental motor, insular, and cingulate cortices. Some of the brain abnormalities seem to be amenable to pharmacological intervention, as a few studies have reported changes on neuroimaging in response to effective therapeutic intervention, although replication of such findings is needed. Patients with fibromyalgia also have evidence of hippocampal disruption indicated by reduced brain metabolite ratios and reduced gray matter concentration. Finally, positron emission tomography using radiolabeled DOPA, a precursor to dopamine synthesis,

has demonstrated a diffuse reduction in monoaminergic (i.e., dopamine, norepinephrine, and epinephrine) metabolism specifically in brain areas in which dopamine plays a critical role in natural analgesia, including the anterior cingulate cortex, insula, and thalamus.

In summary, the abnormalities associated with fibromyalgia show evidence of enhanced nociception associated with increased cortical excitability (as evidenced by abnormal arousal patterns on sleep EEG and increased activation in response to stimulation), subtle changes in brain chemistry, and a dysregulation of both arms of the stress response, i.e., the HPA axis and the sympathetic nervous system. As a review of these findings might suggest, one problem facing both patients and clinicians is that the type of abnormalities reviewed herein are generally detected by assays not typically available to primary care or even tertiary referral clinics. Given the lack of more readily demonstrable pathology, the difficulty that the medical culture has had coming to terms with fibromyalgia and related conditions seems to stem from the pervasive philosophical tradition of Cartesian dualism, in which human phenomena are categorized as belonging to one of two domains, i.e., psyche or soma – the mind or the body. Thus, if an individual presents to a clinician with somatic complaints for which no pathological change is evident in the organ from which the symptoms seem to originate, and if stress appears to play a contributory role, then the patient is liable to be dismissed as somatizing or expressing psychic distress by way of physical symptoms. This unfortunate attitude on the part of uninformed clinicians represents what essentially amounts to a medical form of being politically correct in suggesting that a patient's problems are all in his or her head – a phrase that tacitly avoids the pejorative and clinically irrelevant label “crazy”.

Hypothetical Mechanisms of Abnormal Pain Processing

While the preceding findings are intriguing, the fact is that the exact mechanism of enhanced nociception in fibromyalgia remains a matter of debate and continuing inquiry. Accordingly, the following represents a synopsis of some hypothetical considerations regarding the potential pathologies that might contribute to the phenomenon of enhanced nociception.

There exists a school of thought that considers the pain of fibromyalgia to represent a primarily psychic phenomenon. Accordingly, those who ascribe to this perspective consider the enhanced nociception displayed by fibromyalgia patients to represent a form of secondary hyperalgesia, attributed to such

processes as hypervigilance or increased attention to somatosensory stimulation suggesting a perceptual style of amplification. This conclusion is derived in part from the consistent observation of increased activation of the anterior cingulate cortex, which plays a role in both the attentional and affective, or emotional, dimensions of pain perception. Enhanced activation of the anterior cingulate cortex in response to stimulation has also been demonstrated in another stress-related disorder characterized by enhanced somatosensory perception: irritable bowel syndrome.

On the other hand, the hyperalgesia and other forms of polymodal sensitivity that characterize the disorder have been conceptualized as a form of central sensitization in which elements of the central nervous system have become pathologically sensitive to stimulation, thereby making stimuli that were previously innocuous (light touch, for example) excruciating. The phenomenon of central sensitization has been demonstrated to depend on the plasticity in function of *N*-methyl-D-aspartate (NMDA) subtype glutamate receptors, which play a fundamental role in the flow of information within the brain and spinal cord. In support of this notion, it has been demonstrated that a large subgroup of patients respond to very small doses of the anesthetic ketamine, which at sufficient doses antagonizes NMDA receptors, with dramatic reductions in the painful symptoms. It is then intriguing that NMDA receptors play a critical role in pain perception, specifically within the hippocampus and anterior cingulate cortex, wherein a variety of stress-related factors such as corticotropin-releasing hormone, cortisol, norepinephrine, substance P, and histamine contribute to neuronal excitability specifically by enhancing NMDA receptor activity. Thus, the pain of fibromyalgia (and perhaps other related disorders) might stem from stress-related changes within brain centers involved in overlapping mechanisms of nociception (i.e., attentional and affective dimensions), in which stress-related factors reduce the threshold for excitation. As such, these changes provide a credible organic basis for what might otherwise be appreciated as psychic phenomena.

A relative increase in the amount of information flowing out of the body and toward the brain might also result in enhanced pain perception. One naturally expects this in scenarios involving tissue damage, whether related to mechanical causes (as in an injury or a degenerative condition such as arthritis) or a chemical or thermal exposure. Likewise, inflammatory processes result in the release of various chemicals within affected tissues that sensitize peripheral pain receptors (nociceptors), thereby amplifying the flow of information from the area of inflammation.

The adaptive value of such an increased amount of painful information originating from injured tissue is clear: the organism protects the area and allows it to heal. On the other hand, no such peripheral pathology has been described in fibromyalgia; therefore, an area of intense focus regarding the potential source of increased information flowing toward the brain has been potential pathological changes within the dorsal root ganglion and dorsal horn of the spinal column, which collectively represent the interface between the peripheral and central nervous systems. Both of these regions contain high densities of NMDA receptors, which are thought to contribute substantially the phenomenon of wind-up, and both have been demonstrated to undergo pathological transformation under certain conditions, thereby resulting in hyperalgesia and allodynia in the affected area. Intriguingly, recent investigations using rat models have demonstrated that stress exposure results in increased metabolic activity within the dorsal horn. However, the relevance of these findings to fibromyalgia remains unknown.

Somatosensory information processing represents a two-way flow of information within the spinal cord: information concerning the status of the body is transferred from the peripheral nervous system to the central nervous system and ascends the spinal cord toward the brain, while a variety of filtering mechanisms descend from the brain to modulate the amount of information transferred specifically within the dorsal horn. The vast majority of neurons that carry descending inhibition run along tracks within the dorsal spinal column (i.e., the part furthest toward the back). Research has demonstrated that interruption of the dorsal descending systems leads to hyperactivity of pain-related neurons within the spinal cord characterized by an increase in background activity, lowering in stimulation threshold, and increase in response magnitude to noxious stimuli. Collectively, the findings suggest that if such an interruption were to occur in humans, such impairment would likely result in spontaneous deep pain, tenderness to palpation, and hyperalgesia of deep tissues – all of which characterize patients with fibromyalgia. Accordingly, some researchers are focused on determining the potential contribution of restrictive lesions, specifically within the cervical spinal cord, to the experience of chronic widespread pain. Likewise, an increase in descending facilitation involving medullary brain centers might also contribute to the amplification of sensory information.

The apparent dysregulation of the sympathetic nervous system detailed previously has led to the hypothesis that fibromyalgia might represent a sympathetically maintained pain syndrome. In support of this notion,

proponents of the hypothesis observe that, like other so-called sympathetically maintained pain conditions (e.g., complex regional pain syndrome, or reflex sympathetic dystrophy), the pain associated with fibromyalgia may be characterized by its frequent posttraumatic onset and by the presence of chronic stimuli-independent pain accompanied by allodynia, paresthesias, and distal vasomotor changes. Limited reports suggest that fibromyalgia pain may be aggravated by peripheral administration of epinephrine. Historically speaking, it has been generally considered that sympathetically maintained pain is alleviated by maneuvers intended to block sympathetic activity, and there are (again, limited) reports reflecting that fibromyalgia patients respond to these as well. More recently, however, the contribution of sympathetic activity to those conditions traditionally considered to represent sympathetically maintained pain conditions has become a matter of some contention, while the specific mechanism whereby sympathetic activation might contribute to the development of chronic widespread pain remains to be elucidated.

The most recent hypotheses concerning the nature of pain in fibromyalgia have to do with a disruption of natural analgesia within the limbic system involving the neurotransmitter dopamine. Stressful experience has been robustly demonstrated to disrupt the normal function of dopamine neurons, which in turn plays an important role in natural analgesia in such brain centers as the thalamus, basal ganglia, nucleus accumbens, insular cortex, and anterior cingulate cortex. Recent insights regarding the pharmacology of ketamine have demonstrated that the drug activates dopamine D2 receptors, which may then explain its efficacy in relieving symptoms of fibromyalgia when given at very low doses. Intriguingly, many of the brain centers that have been shown to be involved in fibromyalgia using functional neuroimaging are among the same brain centers in which dopamine plays a role in analgesia. Dopamine acts within the thalamus and insular cortex to recruit mechanisms of descending inhibition, while in the anterior cingulate cortex dopamine opposes the activity of NMDA receptors specifically involved in nociception. In addition, the reactivity of dopamine neurons is controlled by excitatory output from the hippocampus, which in turn is exquisitely sensitive to the effects of stress-related factors. A stress-related increase in neuronal excitability within the hippocampus would be predicted to suppress the reactivity of dopamine neurons, thereby producing a relative increase in painful sensations. Other neurotransmitters whose dysfunction has been proposed to contribute to the development of chronic widespread pain include serotonin and norepinephrine. In general,

medications that act chiefly to boost serotonergic activity have met with disappointing results, while those with mixed activity for serotonin and norepinephrine are somewhat more promising.

Future Directions

Among the current concepts driving investigations into the nature of fibromyalgia is the growing awareness that patients with fibromyalgia likely represent a homogenous population in which chronic widespread pain is the common denominator. This idea resonates with the conceptualization of two other categories of central nervous system disorders: epilepsy, which has as its common denominator stereotyped and uncontrollable movement stemming from pathology within a variety of brain centers, and depression, in which patients experience dysphoria stemming from the disturbance of a variety of neurohormonal factors. Thus, efforts to subclassify fibromyalgia patients along both organic and psychological lines are currently under way.

In order to effectively target the symptoms of a disorder for therapy, there must first be an understanding of the underlying pathology. In light of recent advances in our insights into the disorder, we indeed seem to be standing at the threshold of understanding. While the status of fibromyalgia as a bona fide disorder is a matter of continuing debate among the uninformed, the fact that major pharmaceutical companies have begun to develop large-scale trials to investigate the potential efficacy of existing drugs at relieving the various symptoms associated with the disorder is evidence of its increasing acceptance within the medical community as a whole. In general, there seems to be grounds for an increased sense of optimism regarding relief for patients, both from their symptoms and from the stigma previously associated with the diagnosis.

See Also the Following Articles

Chronic Fatigue Syndrome; Dopamine, Central; Hippocampus, Overview; Hypothalamic-Pituitary-Adrenal; Musculoskeletal Problems and Stress; Pain; Posttraumatic Stress Disorder, Neurobiology of.

Further Reading

- Bennett, R. M. (ed.) (2002). *The clinical neurobiology of fibromyalgia and myofascial pain: therapeutic implications*. Binghamton, NY: Haworth Press, Inc.
- Clauw, D. J. (1995). Fibromyalgia: more than just a musculoskeletal disease. *American Family Physician* 52, 843–851, 853–854.
- Clauw, D. J. and Crofford, L. J. (2003). Chronic widespread pain and fibromyalgia: what we know, and what

- we need to know. *Best Practice and Research: Clinical Rheumatology* 17, 685–701.
- Gracely, R. H., Petzke, E., Wolf, J. M., et al. (2002). Functional magnetic resonance imaging evidence of augmented pain processing in fibromyalgia. *Arthritis and Rheumatism* 46, 1333–1343.
- Jain, A. K., Heffez, D. S., Carruthers, B. M., et al. (2003). The fibromyalgia syndrome: a clinical case definition for practitioners. *Journal of Musculoskeletal Pain* 11, 3–108.
- Neeck, G. and Crofford, L. J. (2000). Neuroendocrine perturbations in fibromyalgia and chronic fatigue syndrome. *Rheumatic Diseases Clinics of North America* 26, 989–1002.
- Staud, R. (2002). Evidence of involvement of central neural mechanisms in generating fibromyalgia pain. *Current Rheumatology Reports* 4, 299–305.
- Wolfe, F., Smythe, H. A., Yunus, M. B., et al. (1990). The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia. Report of the Multi-center Criteria Committee. *Arthritis and Rheumatism* 33, 160–172.
- Wood, P. B. (2004). Fibromyalgia: a central role for the hippocampus – a theoretical construct. *Journal of Musculoskeletal Pain* 12, 19–26.
- Wood, P. B. (2004). Stress and dopamine: implications for the pathophysiology of chronic widespread pain. *Medical Hypotheses* 62, 420–424.
- Yunus, M. B. (1992). Towards a model of pathophysiology of fibromyalgia: aberrant central pain mechanisms with peripheral modulation. *Journal of Rheumatology* 19, 846–850.

Fight-or-Flight Response

R McCarty

Vanderbilt University, Nashville, TN, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 143–145, © 2000, Elsevier Inc.

Emotional Stimulation and Epinephrine Secretion
The James–Lange Theory of Emotions
The Cannon–Bard Theory of Emotions
Contemporary Research on the Fight-or-Flight Response

Glossary

Cannon–Bard theory

A theory based on the work of the American physiologist W. B. Cannon and his student P. Bard. Their work indicated that portions of the brain stem and hypothalamus were essential for the expression of emotions and the peripheral physiological manifestations of the emotions.

Fight-or-flight response

A concept developed by W. B. Cannon that linked emotional expression to physiological changes in the periphery. Cannon emphasized the emergency function of the adrenal medulla in such situations. He reasoned that the increased discharge of epinephrine from the adrenal medulla stimulated increases in blood glucose and blood flow to skeletal muscles and a corresponding decrease in blood flow to the viscera.

James–Lange theory

A theory based on the independent work of the American psychologist W. James and the Danish physiologist K. Lange. The theory proposed that emotions resulted from afferent feedback from the periphery to the brain.

Emotional Stimulation and Epinephrine Secretion

Walter Cannon and his students and coworkers conducted an initial series of experiments from 1910 to 1911 that examined the link between emotional experiences and the secretion of epinephrine from the adrenal medulla. Employing an intestinal strip bioassay for the semiquantitative study of epinephrine secretion, he and his students determined that epinephrine secretion increased dramatically when a quiescent animal was exposed to a stressful or threatening stimulus. In addition, Cannon's laboratory used increases in glucose in the urine (glycosuria) as an indirect measure of epinephrine secretion with similar results. The connection between epinephrine secretion and increases in blood glucose led Cannon to write the following entry in his journal on January 20, 1911: "Got idea that adrenals in excitement serve to affect muscular power and mobilize sugar for muscular use – thus in wild state readiness for fight or run!" In time, "fight or run" was transformed into the more familiar fight or flight.

The initial description of the fight-or-flight response encompassed only the peripheral manifestations of the response, including a redirection of blood flow and

Table 1 Sympathetic-Adrenal Medullary Components of Fight-or-Flight Response

	<i>Physiological effects</i>	<i>Physiological consequences</i>
Heart	Increased rate	Increased rate of blood flow
	Dilation of coronary vessels	Increased availability of O ₂ and energy to cardiac muscle cells
Circulation	Dilation of vessels serving skeletal muscle	Increased availability of O ₂ and energy to skeletal muscle cells
	Vasconstriction of vessels serving digestive and related organs	Allows for shunting of blood to skeletal muscle cells
	Contraction of spleen	Increase in delivery of O ₂ to metabolically active tissues
Respiration	Dilation of bronchi	Increased availability of O ₂ in blood
	Increased respiratory rate	Increased availability of O ₂ in blood
Liver	Increased conversion of glycogen to glucose	Increased availability of glucose for skeletal muscle cells and brain tissue

increased delivery of O₂ and glucose to the skeletal muscles as an organism prepared to confront a threatening or stressful situation (Table 1).

It was not until the 1920s that Cannon began a series of studies on the central control of sympathetic-adrenal medullary activity. This work focused initially on studies of sham rage in decorticate cats. The findings pointed clearly to the posterior hypothalamus and basocaudal subthalamus as brain structures important for the elevated adrenal medullary discharge of epinephrine during sham rage. This line of research has continued to the present as investigators have used ever more refined anatomical techniques to define the central neural circuitry underlying emotional expression.

The James–Lange Theory of Emotions

At the turn of the twentieth century, the prevailing physiological theory of emotions was based on the independent efforts of the American psychologist William James and the Danish physiologist K. Lange. The James–Lange formulation was as follows:

1. An emotionally charged event occurs and stimulates sensory receptors.
2. Afferent nerve impulses reach the cerebral cortex and the event is perceived.
3. Efferent nerve impulses are directed to the skeletal muscles and the viscera and alter their levels of activity in preparation for the event.

4. Afferent nerve impulses from the periphery go once again to the cerebral cortex and the emotionally charged aspect of the event is perceived.

James emphasized the importance of visceral afferent signals and emotion, whereas Lange drew attention to the relationship between vascular afferent signals and emotion. This was the prevailing theory when Cannon and coworkers began to examine the neural basis for visceral and behavioral control during stressful experiences.

The Cannon–Bard Theory of Emotions

Cannon recognized that a critical test of the James–Lange theory would involve an examination of emotional expression in a subject with no visceral afferent feedback. In such an experiment, the critical link between visceral changes and the feedback required to stimulate the cerebral manifestations of the emotion would be broken. At least three prevailing lines of evidence pointed to the inaccuracy of the James–Lange theory: Sherrington had reported in 1900 that dogs exhibited intense emotions following complete destruction of the sympathetic and spinal sensory roots; later, Cannon and colleagues noted that cats remained emotionally responsive following surgical destruction of the sympathetic nervous system; and, finally, clinical case studies demonstrated that spinally transected patients retained their emotional responses despite being completely paralyzed below the neck.

Cannon was joined in this research effort by several of his students, most notably Philip Bard. Their efforts, beginning in the 1920s, changed the focus of research on emotions to include the central nervous system as the primary site for receiving sensory information, directing peripheral visceral and vascular changes, and generating the psychological manifestations of the experience.

Contemporary Research on the Fight-or-Flight Response

The study of the neural bases of emotions remains an active and exciting area of research to the present. Researchers continue to examine the neural circuits involved in the classic fight-or-flight responses of laboratory animals. Studies suggest that central command neurons within hypothalamic and brain stem nuclei project via sympathetic outflow systems to both the heart and the adrenal medulla. These collections of central command neurons may direct multiple sympathetically controlled neural and endocrine responses that subserve the fight-or-flight response. The midbrain periaqueductal gray region

has also been implicated in the physiological and behavioral changes attending the fight-or-flight response. In addition, research on the amygdala has revealed the importance of this brain region for memories of emotionally charged events. Finally, with the advent of imaging techniques (e.g., functional magnetic resonance imaging, fMRI) for studying regional brain activation in human subjects, it is now possible to determine which brain areas are activated during emotionally stimulating events in humans.

Cannon introduced the fight-or-flight concept to physiology and conducted some of the seminal studies in this area. It is impressive to note that his original research linking physiology with emotions is still frequently referenced more than 80 years later and that the fight-or-flight response continues to serve as a stimulating topic of inquiry for researchers in a variety of fields, including physiology, neuroanatomy, and psychology.

See Also the Following Articles

Emotions: Structure and Adaptive Functions; Evolutionary Origins and Functions of the Stress Response; Fear.

Further Reading

- Bandler, R. and Shipley, M. T. (1994). Columnar organization in the midbrain periaqueductal gray: modules for emotional expression? *Trends in Neuroscience* 17, 379–389.
- Benison, S., Barger, A. C. and Wolfe, E. L. (1987). *Walter B. Cannon: the life and times of a young scientist*. Cambridge, MA: Belknap Press of Harvard University Press.
- Brooks, C. M., Koizumi, K. and Pinkston, J. O. (1975). *The life and contributions of Walter Bradford Cannon 1871–1945: his influence on the development of physiology in the twentieth century*. Brooklyn, NY: State University of New York Downstate Medical Center.
- Cannon, W. B. (1915). *Bodily changes in pain, hunger, fear and rage: an account of recent researches into the function of emotional excitement*. New York: Appleton.
- Cannon, W. B. (1927). The James-Lange theory of emotions: a critical examination and an alternation. *American Journal of Psychology* 39, 106–124.
- Davis, M. (1992). The role of the amygdala in fear and anxiety. *Annual Review of Neuroscience* 15, 353–375.
- LeDoux, J. E. (1995). Emotion: clues from the brain. *Annual Review of Psychology* 46, 209–235.

Firefighters, Stress in

T L Guidotti

George Washington University, Washington, DC, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by T L Guidotti, volume 2, pp 146–148, © 2000, Elsevier Inc.

Firefighting as an Occupation

Physical Stress

Psychological Stress

Glossary

- 17-Ketosteroids** The metabolites of endogenous steroid hormones produced under sustained stress.
- Aerial ladder** A very high extension ladder on a hook and ladder truck, used for rescue and to reach the fire with water from above.
- Catecholamines** Endogenous hormones produced acutely under stress.

Ischemia

A condition of insufficient blood supply; in the heart, coronary artery ischemia may lead to angina (chest pain), arrhythmia (irregular heartbeat), or a heart attack.

Knockdown measures

Measures to suppress a fire, such as laying on water from a hose carried up a ladder.

Moonlighting

Holding a job or doing work in addition to one's regular employment, usually without the knowledge or approval of the principal employer; firefighters may have several days off between long shifts and often work other jobs to earn extra money.

Myocardial infarction

A heart attack, in which some of the heart muscle dies from lack of oxygen due to ischemia.

Personal protective equipment

The equipment used by firefighters during work to protect the individual workers from a hazard, as distinct from measures used to control the hazard itself; examples include self-contained breathing apparatus, gloves, helmet, and turnout gear (heavy coat and boots).

Firefighting as an Occupation

Firefighting is an intrinsically stressful occupation that attracts an unusual type of person. It is a commonplace saying that firefighters are unique people because they run into burning buildings when other people are fleeing them. The sad human toll of the World Trade Center attack, in which 121 firefighters died, trapped in the North Tower because their radios did not deliver warning messages due to lack of interoperability, demonstrates the risk that firefighters face in their work and their dependence on equipment and organization.

Firefighting, like most occupations involving emergency response, is characterized by prolonged periods of preparation for episodes of intensely stressful activity. Episodically, performance demands require exertion near the limits of human factors and cardiopulmonary conditioning. The risk of injuries is high in firefighting, as indicated by an increased hospitalization rate compared to other occupations for soft tissue injuries, burns, and overexertion. However, much of a firefighter's time is routine, spent in training and maintenance. The psychological stresses of firefighting include long periods of relative inactivity punctuated by highly stressful alarms and extremely stressful situations, such as rescues, as reflected in physiological and biochemical indicators. The literature on the health risks associated with firefighting is large and has been extensively reviewed with particular reference to human factors and ergonomics, chemical hazards, and work-related health outcomes.

The hazards associated with firefighting are changing in nature and magnitude. In the mid-twentieth century, the new composition of interior furnishings introduced new chemical hazards for firefighting, including combustion products from polyvinyl chloride (including phosgene, vinyl chloride) and acrylonitrile (including cyanide). In modern times, commercial and manufacturing facility fires may produce combustion products that are not well characterized and may not even be known. Although firefighting is generally becoming safer due to better and more available equipment and improved procedures, the terrible toll at the World Trade Center illustrates the growing potential for multiple fatalities and serious injuries when things go seriously wrong in ever-larger modern skyscrapers and large facilities.

Firefighters tend to take care of their own but on an informal basis. Firefighting is a close-in activity requiring teamwork under very threatening conditions, and firefighters tend to be very aware of the problems that their fellow firefighters might have in handling difficult situations. Individual firefighters who show early signs of illness are often selectively transferred out of active firefighting positions.

Physical Stress

Energy Requirements

Firefighting is a physically strenuous occupation that is sometimes performed under extreme environmental conditions. There are several components to the physiological demands of firefighting; these include the energy cost of performing firefighting activities, heat stress associated with heat from the fire, encumbrance by personal protection equipment, and psychological stress related to personal safety, protection of property, and protection of lives. The ergonomic demands of firefighting are extreme at peak activity because of high energy costs of activities such as climbing aerial ladders, the positive heat balance from endogenous and absorbed environmental heat, and encumbrance by bulky but necessary protective equipment. The use of personal protection equipment has imposed new physiological demands on firefighters despite the increased level of personal protection.

Physiological Stress

Firefighting is intrinsically stressful. Firefighters increase their urinary excretion of catecholamines (endogenous hormones produced acutely under stress) following workdays compared to days off. Paramedics, for whom job satisfaction was much less, had even higher levels of stress than firefighters. The urinary excretion of 17-ketosteroids (metabolites of endogenous steroid hormones produced under sustained stress) and catecholamines reflects perceived patterns of psychological stress and activity levels at the station.

Cardiovascular Stress

At the sound of an alarm, firefighters experience a degree of immediate anxiety because of the inherent unpredictability of the situation that they are about to encounter. Some investigators believe that the psychological stress experienced at this moment is as great and perhaps greater than any of the stresses that follow during the course of responding to an alarm. Studies of the cardiovascular response to the alarm itself show a marked and highly variable increase in heart rate during the response to a fire alarm, well before physical exertion begins at levels associated with maximal exertion or anxiety. Urinary catecholamine levels are particularly elevated among male firefighters during and after work in the alarm center, an assignment that is highly unpopular. There is also a gender-associated difference in the response, which is less pronounced in women. En route, hazardous traffic maneuvers and high noise levels from the sirens contribute to stress.

Fitness

Numerous studies have evaluated the fitness of firefighters and have shown fairly consistently that most firefighters are as fit or somewhat more fit than the general adult male population. They are not, however, fit to an athletically trained level. Recent studies of job performance requirements suggest that exercise programs for firefighters should be more comprehensive rather than emphasizing cardiovascular fitness and muscle strength, as is now the case.

An increasing number of female applicants are now competing for firefighting jobs. This has forced a re-evaluation of performance tests. As a group, women appeared to demonstrate lower scores on average than men in all performance items, but a subgroup of women performed nearly as well in some tasks. The overall difference in performance was attributed primarily to lower absolute lean body weight, which correlated most strongly and consistently with performance difference.

There is evidence that some firefighters may be stressed to the limit during the exertions of their work. Barnard et al. described an ischemic response in healthy young men, including firefighters, who engaged in sudden vigorous exercise without warm-up. They suggested a transient mismatch in oxygen supply and demand. Although this mechanism is probably not a cause of persistent or cumulative myocardial injury, it may play a role in unusual and emergent situations requiring sudden maximal exertion.

On this basis, it may be concluded that acute myocardial infarction occurring under these circumstances or shortly after maximal exertion in an emergency response may be directly associated with working as a firefighter. It could also be that such an event may happen at work to a firefighter with existing coronary artery disease, in which case it is reasonable to assume an association with firefighting activity. However, it is difficult to attribute myocardial infarction when off-service, hypertension, or chronic manifestations of coronary artery disease to firefighting at this time, given the available evidence.

Because of the high level of physical and psychological stress and the presence of chemical hazards known to injure the cardiovascular system, heart disease is a major occupational health concern in firefighting. Most studies of firefighters show no elevation in mortality from cardiovascular causes or, at most, a small excess risk. The healthy worker effect expected in such a highly selected group may be offset by greater than expected mortality. A healthy worker effect is more readily visible when comparisons are made to similarly selected occupational groups such as police. Although the findings of individual studies may differ, the pattern that emerges

after careful examination is that mortality from cardiovascular causes is not obviously increased in association with firefighting activity (either in terms of years of service or exposure opportunity).

Psychological Stress

There are many extraordinary sources of psychological stress in the life of a firefighter. The expectations for performance placed on firefighters are well beyond that of other civilian occupations, with the possible exception of police. Firefighters accept a level of personal risk beyond the range of normal human experience. Although this risk is controlled to the extent possible with fire equipment and personal protection, the reality of firefighting is that there is much that can go wrong at any fire and the course of a serious fire is often unpredictable. In addition to personal security, the firefighter must be concerned with the safety of others. A failed rescue or the death of a victim is so stressful that some firefighters who have experienced this cannot cope.

Studies of firefighters have shown high levels of job satisfaction and motivation compared to the general population, despite job-related anxiety and adjustment issues related to sleep deprivation, heavy workload or pace of work, constantly being on guard, unsatisfactory public interactions, rigid working conditions, personal risk, inconsistent instructions, and unforeseen situations. On the other hand, firefighters exert a great deal of control over their work in the selection of approaches to knocking down the fire and the use of protective equipment. They are continually faced with new challenges because every alarm presents a different story or situation.

However, firefighters who have experienced particularly threatening situations have demonstrated adverse emotional effects. Chief among these is loss of a life during a rescue attempt or the death or serious injury on the job of a fellow firefighter.

Firefighters have been intensively studied as a model population at risk for posttraumatic stress disorder. Posttraumatic stress disorder can significantly impair in their ability to maintain social and family relationships. Firefighters who do not show the full-blown posttraumatic stress disorder may still be affected by intrusive memories and anxiety. Incident debriefing sessions have been shown to be valuable and effective in assisting firefighters who would otherwise have difficulty coping and in reducing the short-term anxiety of those who would have been able to cope without assistance.

Firefighters are enmeshed in conflicting expectations and numerous technicalities related to the insurance industry, the legal profession, the fire service

administration and politics, labor-management relations, and building codes and inspection. Firefighters may be highly trained in the technicalities of fire suppression, but may become so enmeshed unwillingly in the nontechnical aspects of the occupation that they may experience great frustration. When inevitable mistakes happen or their judgment is questioned, firefighters may feel psychologically threatened because of the pressure to accommodate all these expectations and still manage during an evolving emergency.

On the other hand, firefighters enjoy many positive aspects of their work and social status. Firefighters have a high social standing and are considered to belong to the most highly skilled blue-collar occupation. The work itself is interesting, although highly routine between alarms. There is a high degree of autonomy and a heavy emphasis on teamwork and mutual support. Between shifts many firefighters moonlight (have second jobs). Firefighters enjoy the high esteem of the community, the admiration of other citizens, the aspirations of children as role models, and widespread support as unequivocally positive defenders of the public good.

See Also the Following Articles

Heat Shock Response, Overview; 9/11, Religion and Stress; Temperature Effects; Workplace Stress; Night Shiftwork.

Further Reading

- Barnard, R. J. (1975). Fire fighters – a fit population with ischemic heart disease. *Sports Medicine Bulletin* 10, 7–9.
- Barnard, R. J. and Duncan, H. W. (1975). Heart rate and ECG responses of fire fighters. *Journal of Occupational Medicine* 17, 247–250.
- Barnard, R. J., Gardner, G. W. and Diaco, N. V. (1976). “Ischemic” heart disease in firefighters with normal coronary arteries. *Journal of Occupational Medicine* 18, 818–820.
- Barnard, R. J., Gardner, G. W., Diaco, N. V., et al. (1978). Near-maximal ECG stress testing and coronary artery disease risk factor analysis in Los Angeles City fire fighters. *Journal of Occupational Medicine* 17, 693–695.
- Bos, J., Mol, E., Visser, B., et al. (2004). Risk of health complaints and disabilities among Dutch firefighters. *International Archives of Occupational and Environmental Health* 77, 373–382.
- Byrd, R. and Collins, M. (1980). Physiologic characteristics of fire fighters. *American Corrective Therapy Journal* 34, 106–109.
- Dutton, L. M., Smolensky, M. H., Leach, C. S., et al. (1978). Stress levels of ambulance paramedics and firefighters. *Journal of Occupational Medicine* 20, 111–115.
- Guidotti, T. L. (1992). Human factors in firefighting: ergonomic-, cardiopulmonary- and stress-related issues. *International Archives of Occupational and Environmental Health* 64, 1–12.
- Guidotti, T. L. (1993). Mortality of urban firefighters in Alberta, 1927–1987. *American Journal of Industrial Medicine* 23, 921–940.
- Guidotti, T. L. (1995). Occupational mortality among firefighters: assessing the association. *Journal of Occupational and Environmental Medicine* 37, 1348–1356.
- Guidotti, T. L. (1998). Emergency and security services. In: Stellman, J. M., McCann, M., Warshaw, L., Brabant, C. & Dufresne, C. (eds.) *The ILO encyclopaedia of occupational health* (4th edn. vol. 3), pp. 95.1–95.22. Geneva: International Labor Organization.
- Guidotti, T. L. and Clough, V. (1992). Occupational health concerns of firefighting. *Annual Review of Public Health* 13, 151–171.
- Hytten, K. and Hasle, A. (1989). Fire fighters: a study of stress and coping. *Acta Psychiatrica Scandinavica* 80 (supplement 355), 50–55.
- Kalimo, R., Lehtonen, A., Daleva, M., et al. (1980). Psychological and biochemical strain in firemen’s work. *Scandinavian Journal of Work, Environment & Health* 6, 179–187.
- Kuorinka, I. and Korhonen, O. (1981). Firefighters’ reaction to alarm, an ECG and heart rate study. *Journal of Occupational Medicine* 23, 762–766.
- Lee, D. J., Fleming, L. E., Gomez-Marin, O., et al. (2004). Risk of hospitalization among firefighters: the National Health Interview Survey, 1986–1994. *American Journal of Public Health* 94(11), 1938–1939.
- Markowitz, J. S., Gutterman, A. M., Link, B., et al. (1987). Psychological response of firefighters to a chemical fire. *Journal of Human Stress* 13, 84–93.
- Reischl, U. W. E., Bair, H. S. and Reischl, P. (1979). Fire-fighter noise exposure. *American Industrial Hygiene Association Journal* 40, 482–489.
- Rhea, M. R., Alvar, B. A. and Gray, R. (2004). Physical fitness and job performance of firefighters. *Journal of Strength and Conditioning Research* 18, 348–352.
- Romet, T. T. and Frim, J. (1987). Physiological responses to fire fighting activities. *European Journal of Applied Physiology* 56, 633–638.
- Rosenstock, L., Demers, P., Heyer, N. J., et al. (1990). Respiratory mortality among firefighters. *British Journal of Industrial Medicine* 47, 462–465.
- Selkirk, G. A. and McLellan, T. M. (2004). Physical work limits for Toronto firefighters in warm environments. *Journal of Occupational and Environmental Hygiene* 1, 199–212.
- Smith, D. (1988). *Firefighters: their lives in their own words*. New York: Doubleday.
- Sothmann, M. S., Gebhardt, D. L., Baker, T. A., et al. (2004). Performance requirements of physically strenuous occupations: validating minimum standards for muscular strength and endurance. *Ergonomics* 22, 864–875.
- von Hallmeyer, A., Lingbeil, M., Kohn-Seyer, G., et al. (1981). Physische und psychische Belastungskomponenten in der Tätigkeit des Feuerwehrmannes. *Zeitschrift für die Gesamte Hygiene und Ihre Grenzgebiete* 27, 191–194.

Fish, Stress in

C B Schreck

Oregon State University, Corvallis, OR, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by C B Schreck, volume 2, pp 149–152, © 2000, Elsevier Inc.

Elements of the Stress Response

Factors Affecting the Stress Response

Glossary

<i>Clearance dynamics</i>	The temporal pattern of disappearance of a substance from the circulation, a body part, or the body.
<i>Interrenal</i>	The cells (tissue) located in the fish's head kidney responsible for secretion of corticosteroids; analogous to the adrenal gland of mammals.
<i>Ontogenetic stage</i>	Life history or life cycle stage.
<i>Osmo-regulatory</i>	Pertaining to water balance and related to hydromineral balance, which also includes electrolytes.
<i>Stressor</i>	The factor extrinsic to the fish that causes a stress response.

The physiological responses of fish to stressful situations are very similar to those of the other vertebrates, including humans. Perception of a threat is followed by a neuroendocrine cascade involving catecholamines and the corticosteroid cortisol. Together, these result in secondary stress responses that can be categorized as those concerned with provision of energy, hydromineral balance, the immune system, and pathways that affect behaviors. The psychogenically induced responses tend to be similar, irrespective of the stressful agent; specific stressors can also induce specific responses. While the inter- and intraspecific differences in characteristics of the stress response are ultimately determined by genetics, the animal's history and its extant environment determine the exact nature of the magnitude, duration, and consequences of the response to a particular stressor. Acutely stressful situations result in temporary impairment of necessary functions such as disease resistance, predator avoidance, and learning ability. In addition, prolonged or repeated exposure to stress can have negative effects on development, growth, and reproduction.

Elements of the Stress Response

The general physiological response of a fish to a stressful situation (stressor) is similar to that of the warm-blooded vertebrates. Upon perception of a stressor, the fish's central nervous system initiates a neuroendocrine cascade, resulting in release of catecholamines from chromaffin tissue and secretion of glucocorticosteroids from interrenal tissue (Figure 1). These affect other physiological systems (Figure 2). Of course, the nature of the stress response is quite polymorphic, and specific stressors also have specific consequences.

Stress Response Factors

Fish, including lamprey, produce catecholamines such as adrenaline. These hormones are important for the initial response to the stressor, assisting the fish in terms of energy and stamina. Compared to the steroid hormones, there is relatively little known about the function and dynamics of these peptides. They are secreted almost immediately following the onset of stress and have cardiovascular actions, increasing cardiac output and gill blood flow. Catecholamines have other metabolic actions such as elevating circulating levels of glucose and depleting liver glycogen; their effects can ultimately lead to elevated plasma lactate and altered free fatty acid concentrations. They also affect hydromineral balance, perhaps making the fish more leaky to water because of hypertensive actions, are diuretic, and influence drinking rate.

The jawed fish activate a hypothalamic-pituitary-interrenal axis (HPI) to secrete glucocorticosteroids when stressed. Some literature suggests that these hormones are important in agnathans, but the question is still open. Elasmobranchs produce 1 α -hydroxycorticosterone while the other groups of fish produce cortisol. Cortisol is secreted from interrenal tissue located in the head kidney within minutes after the onset of a stressor upon stimulation by adrenocorticotropin. Upon perception of a stressor, a neural response triggers the synthesis of corticotropin-releasing hormone (CRH) in the hypothalamus of the brain. This peptide is transported neuronally to the pituitary gland to stimulate production and release into the circulation of corticotropin (CTH; ACTH in mammals). There appear to be other brain peptides that may also have a similar action. CTH stimulates the release of cortisol into the blood from the interrenal cells located in the head (anterior) kidney. There is only a small reservoir of stored cortisol, so

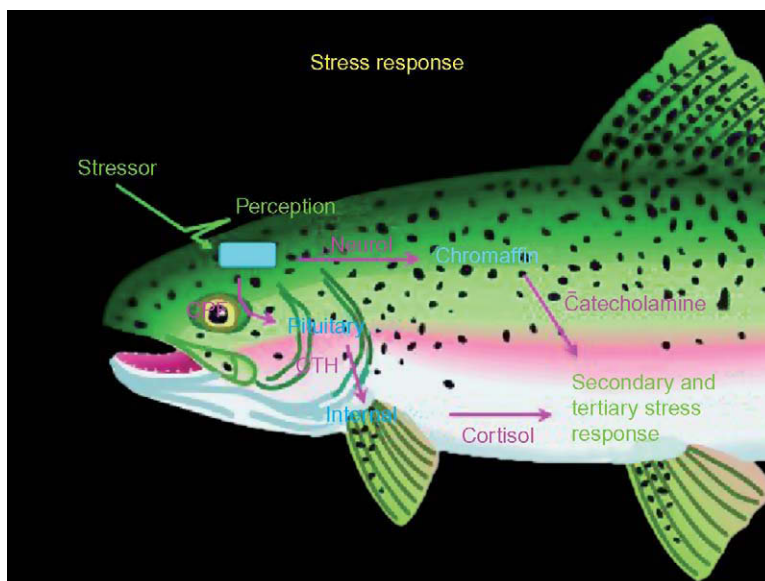


Figure 1 Schematic of the general stress response of teleostean fish leading from perception to the hormones that cause the secondary and tertiary stress responses.

Stress response			
Primary	Cortisol	Catecholamines	
Secondary	<u>Energy</u> Glucose Lactate Cardiovascular FFA, Pr.	<u>Hydromineral</u> Water ↑ ↓ Na ↑ ↓ K Gill vascularization	<u>Immune</u> Redistribution Suppression
Tertiary	<u>Disease Rx</u> AB ↓	<u>Development</u> Retarded	
<u>Behavior</u> Learning ↑ ↓ Predator avoidance ↓ Migration ↓		<u>Growth</u> Hypertrophy ↓ Hyperplasia ↓ Apoptosis ↓	<u>Reproduction</u> Accelerated Inhibited Fecundity ↓

Figure 2 Primary, secondary, and tertiary stress response factors of teleostean fish.

new synthesis is also initiated. The rate of elevation of circulating cortisol varies greatly amongst the species of fish; increase can be seen in less than a minute in some species, while it appears to take hours in others. Cortisol also has a metabolic role during the stress response that serves to provide energy for fight-or-flight functions as well as for osmoregulation. It is essential for the active pumping of electrolytes such as sodium; it could thus help restore the water-ion imbalance consequent to stress by functioning as a mineralocorticoid (fish do not have a mineralocorticoid hormone like mammals). The glucocorticosteroids

might be beneficial to both resistance and recovery from stress. Cortisol is also immunosuppressive. The secondary stress responses are thus regulated via catecholamines and cortisol, acting via specific receptors for these hormones on tissues with target cells. The other elements of the HPI axis mentioned previously can also have direct effects independent of cortisol.

Responding to a stressor is an energetically costly process. That means that there is less energy available for other necessary life functions. The ability of fish to resist pathogens can be affected by stress. It appears that there can be enhancement of certain aspects of

the immune response at some times after exposure to a stressor, but depression of immune system capability tends to be a more general and prolonged response. Cells of the fish's immune system also likely manufacture the stress-related hormones CRH and CTH. Behaviors such as general activity, swimming performance, schooling, thermoregulation, orientation taxis, avoidance, chemoreception, feeding, predation evasion, aggression territoriality, negative conditioning, and learning are all affected by stress, primarily negatively. CRH and several other brain neurotransmitters (e.g., serotonin and dopamine) affect stress-related behavioral responses.

Adult Pacific salmon, *Oncorhynchus* spp., die shortly after spawning. Mature fish of these species have activated HPI axes as encountered during stress and hence elevated circulating levels of cortisol; their ability to clear cortisol from the circulation also appears to be impaired. These fish die from something akin to Cushing's syndrome as described for mammals – i.e., depletion of energy reserves, deterioration of structural tissues, and infection with pathogens resulting from hypercortisolemia.

Nature of the Stress Response

The severity and duration of a stressful encounter influences the magnitude and duration of the physiological stress response. Fish are able to recover from acute, nonlethal stressors. The physiological stress response is initiated within fractions of seconds of when a stressor is first perceived. The physiological events happening soon thereafter (within an hour or so), representing the activation or getting started aspects (alarm phase of stress), are stereotypical (i.e., more predictable) than those of the later stages (resistance and compensation/recovery). Depending on species and environmental conditions, some physiological and behavioral components of the stress response recover quite rapidly following cessation of the stressor, perhaps in minutes to hours. Energy-regarding factors tend to require several hours to return to prestress levels, while those concerned with the immune system, disease resistance, and learning may take days to weeks to recover completely. Magnitudes and durations of response are extremely varied among species as stressor types and durations.

Sequential exposure to acutely stressful events can lead to cumulative physiological responses; it depends on the frequency and severity of the stressor. If the number of repeated exposure to stressors is substantial and frequent enough, fish can compensate and adapt to the experiences. Exposure to less-frequent repeated stressors may not lead to cumulative effects for physiological stress response factors but can have long-term negative effects on a fish's ability to grow,

develop, or reproduce, presumably because there is less energy available for these functions.

Chronic exposure to a stressor can lead to exhaustion (death) or compensation if the stressor is not too severe. Severe but not acutely lethal stressors can result in eventual death via indirect mechanisms such as lowering resistance to pathogens. If fish have adapted to a chronically stressful situation, elements of the physiological stress response will likely not have the identical characteristics (e.g., circulating concentration, clearance dynamics, or rate of response) that they did in a resting state prior to the stressor. Also, the long-term fitness of a fish would be reduced under such circumstances.

When fish are exposed to more than one type of stressor simultaneously (e.g., crowding and temperature elevation), the closer a fish is to its tolerance limits to one stressor, the more severe will be the negative consequences of another concurrent stressor.

In general, the onset of stress in a fish can be determined, but science has not yet been able to establish how to determine when a fish has recovered from exposure to a stressor. In addition, we are not yet very capable of measuring the severity of a stressor. Higher order responses, such as resumption of feeding, appear to be better predictors of the absence of stress than clinical measurements regarding physiology.

Factors Affecting the Stress Response

The response to stressors is ultimately determined by the fish's genetic heritage. However, the stress response phenotype evident during any particular situation is strongly influenced by the fish's prior history, its extant environment, and its developmental stage.

Genetic Effects

There can be major interspecific differences in the nature of the stress response. While the basic elements are generally similar across groups, the magnitude and dynamics of the physiological responses to stressors can be quite different. Some species appear to respond quite slowly, at least in terms of the HPI axis, compared to others; there is a large variation in tolerance of stressors among different groups of fish. There are also genetic differences with reasonably high heritability between individuals within a species in the magnitude of the physiological response to stress. Similarly, there are genetically determined between-individual differences in tolerance of stressors, but it is unknown if these are related to the nature of the fish's physiological stress response.

The physiological stress response factors operate by activating and deactivating specific genes. The

genes affected differ between organs. Functional genomics (which genes respond and what they do) of the fish stress response is a relatively new field of study. The liver, immune system, and brain have received the most attention to date, and are all very strongly affected by the stress response.

Environmental Effects

A fish's prior history and its present environment can affect how it responds to stressors. While the initial stages of the hormonal responses to stress are reasonably invariant in different environments, the persistence of the response and its secondary and tertiary (i.e., whole animal) consequences are very dependent on the nature of the extant environment. Temperature is important in affecting the rates of reactions. Other water quality variables can also be important as controlling or directing factors. Prior experiences with stressors can also harden the fish and hence lessen the apparent severity of subsequent, similar stressors. Nutritional history of the fish can also modify the magnitude and/or dynamics of certain physiological factors of the stress response, particularly with reference to factors concerned with energy.

Importantly, some environmental challenges, albeit lethal, might not necessarily invoke a general HPI stress response. It appears that stressors that are perceived by the central nervous system are those that elicit such a response.

Developmental Effects

Fish at different ontogenetic stages of their life cycle may respond to stressors differently. The basic elements of the physiological stress response, such as the hormones, are present at very early stages of development, and the system appears functional around hatching or soon thereafter. Transitional stages such as the onset of feeding by mouth, metamorphosis, or smoltification and reproduction appear very sensitive to stressors in regard to the increased magnitude of their physiological response and decreased tolerance limits (Figure 3).

Ecological Consequences

Exposure to acute stressors reduces the fish's capacity to resist other immediate threats (Figure 3). Longer-term or repeated exposure to relatively brief stressors depresses the capacity of fish during the time they

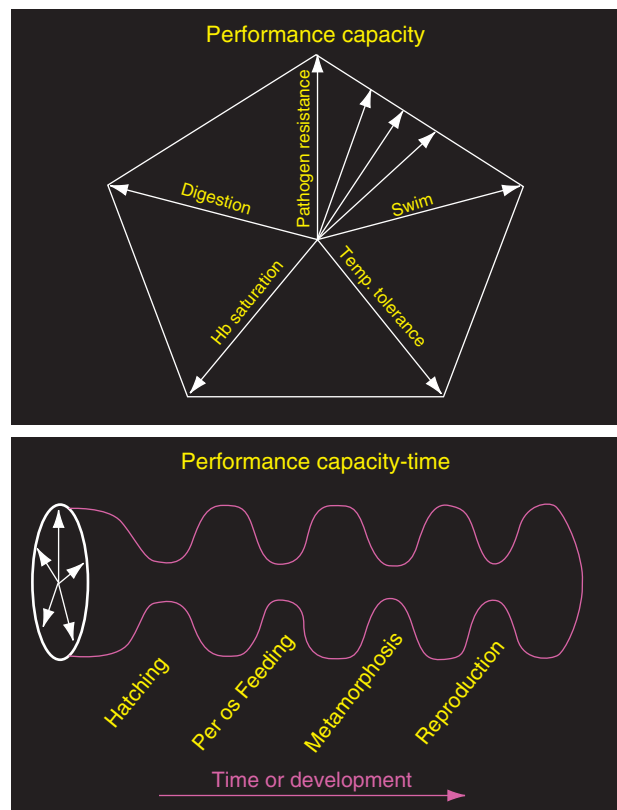


Figure 3 Top panel: Conceptualization of the performance capacity of a fish's necessary life functions that are determined genetically and limited by the environment and negatively affected by stress. The length of each arrow (vector) goes from zero in the center to some level of performance at the point (Hb = hemoglobin). Bottom panel: Cartoon representing the performance capacity of a fish as it changes through time, showing a smaller capacity to resist stress during transitional stages.

are adapting to carry on other necessary life functions. The fish will have less energy to channel into anabolic processes such as growth, development, and reproduction.

It is likely that longer-term stress can change a fish's developmental trajectory and thereby alter its life history strategy. For example, fish that are stressed during certain stages of reproductive maturation may change their spawning strategy. Depending on species of fish and the nature of the stressor, spawning could be accelerated, delayed, or totally inhibited. Stress experienced by females can lead to lower egg quality and subsequent progeny fitness.

While general disturbances and deteriorated water quality can cause a stress response, negative social interactions and low hierarchical status also result in a similar response. Fish of low social rank grow less well and tend to have reduced chances of survival. Submissive fish are stressed by aggressive displays, have activated HPI axes, and suffer its maladaptive consequences.

Temperature is one of the most important environmental controlling factors for fish because the fish are cold-blooded. The rates of biochemical reactions that drive physiological processes are directly dependent on the temperature of the water; in general, for each 10°C increase in temperature, the rates of biological reactions double or triple. Temperature is thus a key variable regarding how fish respond to a stressor, both physiologically and ecologically.

How fish interact with pathogens and parasites is also affected by stress. Stress can increase both the infection rate and pathogenicity of a microorganism.

See Also the Following Article

Animal Models (Nonprimate) for Human Stress.

Further Reading

- Adams, M. S. (ed.) (1990). *Biological indicators of stress in fish*. Bethesda, MD: American Fisheries Symposium 8.
- Ader, R., Felten, D. L. and Cohen, N. (eds.) (2001) *Psychoneuroimmunology* (vols. 1 and 2). San Diego, CA: Academic Press.
- Balm, P. M. (1999). *Stress physiology in animals*. Sheffield, UK: Sheffield Academic Press.
- Barton, B. A. and Iwama, G. K. (1991). Physiological changes in fish from stress in aquaculture with emphasis on the response and effects of corticosteroids. *Annual Review of Fish Diseases* 1, 3–26.
- Fry, F. E. J. (1947). *Effects of the environment on animal activity*. University of Toronto Studies. Biol. Ser. No. 55. Publ. Ontario Fish. Res. Lab. No. 68. Toronto, Canada: Toronto University Press.
- Iwama, G. K., Pickering, A. D., Sumpter, J. P. and Schreck, C. B. (eds.) (1997). *Fish stress and health in aquaculture*. Society for Experimental Biology, Seminar Series 62. Cambridge, UK: Cambridge University Press.
- Pickering, A. D. (ed.) (1981). *Stress and fish*. London: Academic Press.
- Schreck, C. B. (1996). Immunomodulation: endogenous factors. In: Iwama, G. & Nakanishi, T. (eds.) *The Fish Immune System*, pp. 311–337. Fish Physiology Series 15. San Diego, CA: Academic Press.
- Wedemeyer, G. A., Barton, B. A. and McLeay, D. J. (1990). Stress and acclimation. In: Schreck, C. B. & Moyle, P. B. (eds.) *Methods for fish biology*, pp. 451–489. Bethesda, MD: American Fisheries Society.

Flashback See: Posttraumatic Stress Disorder – Clinical.

Floods, Stress Effects of

J O Brende

Mercer University School of Medicine, Macon, GA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J O Brende, volume 2, pp 153–157, © 2000, Elsevier Inc.

Introduction and Background

Postdisaster Coping

Postdisaster Effects on Children

Postdisaster Physical Symptoms

Postdisaster Effects on Emergency and Medical Personnel

Assessing Traumatic Exposure Severity and Symptoms

Introduction and Background

Flooding remains the most common catastrophe worldwide and accounts for an estimated 40% of all natural disasters and 26% of disaster-related deaths. Flash flooding kills more people in the United States than any other natural disaster, accounting for approximately 200 fatalities per year. Flash floods can occur within several seconds to several hours, with little warning and are deadly because they produce rapid rises in water levels and have devastating flow velocities. Several factors can contribute to flash flooding including rainfall intensity, rainfall duration, surface conditions, topography, and slope of the receiving basin. Urban areas are susceptible to flash floods because a high percentage of the surface area is composed of impervious streets, roofs, and parking lots where runoff occurs very rapidly. Mountainous areas also are susceptible to flash floods, as steep topography may funnel runoff into a narrow canyon. Floodwaters accelerated by steep stream slopes can cause the floodwave to move downstream too fast to allow escape, resulting in many deaths. One example of a deadly flash flood occurred after 37.5 cm of rain in 5 h from slow-moving thunderstorms killed 237 people in Rapid City, South Dakota, in 1972.

Floods can occur unexpectedly after excessive rainfall, tropical storms, hurricanes, or failing dams or levees. Deluges of heavy rains can quickly fill narrow ravines, mountain canyons, dry stream beds, or streets in a downriver community. Powerful torrents of water can roll boulders, tear out trees, destroy buildings and bridges, scour out new channels, trigger catastrophic mud slides, sweep away vehicles and houses, drown unsuspecting campers or residents, destroy property, kill or injure friends or family members,

cause homelessness, spread disease, deplete financial resources, and cause emotional and psychological symptoms. Indeed, floods are major disasters and the Federal Emergency Management Agency (FEMA) has warned people to consider these facts:

- Most flood-related deaths are due to flash floods,
- 50% of all flash-flood fatalities are vehicle related. Individuals should not drive through floodwaters,
- 90% of those who die in hurricanes drown. People should heed the warnings and evacuate areas where flooding from storm surges has been predicted,
- Most homeowners' insurance policies do not cover floodwater damage. Flood insurance can be purchased through FEMA's National Flood Insurance Program.

Flooding can also take place from storm surges caused by major hurricanes and are unquestionably the most dangerous parts of hurricanes. Pounding waves create very hazardous flood currents and depths of water that can be greater than 60 m, depending upon the distance of water and the velocity of the wind. Nine out of 10 hurricane fatalities are caused by storm surges. Worst-case scenarios occur when storm surges occur concurrently with high tides. Stream flooding is also much worse inland during storm surges because of backwater effects. In September 1900, the hurricane and storm surge at Galveston, Texas, killed nearly 8000 people, making it the worst natural disaster in the nation's history.

Floods seem to have become more severe and commonplace in recent years. Within the boundaries of the United States, there have been many severe floods. One of the most significant of these was the Great Mississippi Flood of 1927. Rains began in the summer of 1926 and by May of 1927 the Mississippi River had flooded and was 97 km wide below Memphis, Tennessee, and flooded 109 km² in six states: Arkansas, Illinois, Kentucky, Louisiana, Mississippi, and Tennessee. There were 246 people killed in seven states and damages exceeded \$650 million.

The floods of 1993 were also one of the greatest natural disasters ever to hit the United States. Major flooding primarily affected the Mississippi and Missouri rivers and extended into 150 major rivers and tributaries across nine states: North Dakota, South Dakota, Nebraska, Kansas, Minnesota, Iowa, Missouri, Wisconsin, and Illinois. Hundreds of levees failed and 600 river points were above flood stage at the same time. Fifty people were killed and damages approached \$15 billion.

When tropical storm Alberto lingered over the Florida Panhandle on 3 July 1994, the ensuing rains continued until Georgia's flood of the century eventually devastated its central and southwest portions, killed 32 people, displaced 70 000 people, flooded 1214 km² of croplands, caused crop damage of \$100 million, and a total monetary damage exceeding \$500 million.

Two devastating floods inundated much of northern California in January 1995. Three different hurricanes struck the Caribbean and the United States Atlantic Coast in late summer and fall of 1996; Bertha struck in July, Fran in September, and then Hortense washed away southwestern Puerto Rico. From the early months of 1997, unprecedented floods affected eastern, mid-western, and northern United States, particularly devastating communities bordering the Red River which flooded the Minnesota and North Dakota border and continued into Canada in April 1997.

With the onset of El Niño, 1998 became an unusually severe year for flooding. The Pacific Coast was devastated by rain, rising rivers, ocean swells, and killer mud slides in January and February followed by significant flooding in the southeast and Midwest from February until June. Downpours, high ocean swells, and flooding rivers struck the West Coast and the Southeast Coast of the United States with a fury in early February 1998. Southern California was inflicted with the worst wind and rain storms in 5 years. Rain soaked hills and dikes began giving way, causing more flooding and destruction of homes and property.

While blizzards in the Upper Midwest paralyzed highways, roads, and air traffic, rain brought 30–50 cm of rainfall across the south during early March, 1998. The community of Elba, Alabama evacuated 2000 of its 4000 residents when a levee broke and killed two children from torrents of water roaring through the town without warning. Fierce thunderstorms swamped parts of seven states in the northeast. Flooding was so extensive that it required helicopter rescues. Vermont's Mad river, usually docile in summertime, lived up to its name when it jumped the banks to wash out roads and destroy homes, and drove campers in Vermont to higher ground.

On 24 August 1998, flash floods from tropical storm Charlie struck Del Rio, Texas on the Mexico–Texas border. Within the first few days, it was known that 15 people lost their lives although rescuers were continuing to look for missing persons during the following week. South Texas continued to be a target during mid-October as 37–45 cm of rain over a period of several days flooded the Guadalupe and Colorado Rivers, causing the deaths of 24 people.

The Gulf Coast was also stricken by flooding from hurricane Georges which first struck Puerto Rico on

21 September 1998, and devastated the Dominican Republic. It swept across the Caribbean, the eastern coast of Cuba, the Florida Keys, and continued northward through the Gulf of Mexico until its center crossed the coast at Gulfport and Biloxi, Mississippi, 8 days later. Storm surges and flooding caused extensive damage and drove 1.5 million Gulf Coast residents from their homes. In Mobile, Alabama, a 2.7 m storm surge from Mobile Bay sent chest-deep water into many neighborhoods so that more than 300 people had to be evacuated.

In 1998, the worst flooding to hit Central America began on 27 October when hurricane Mitch stalled over the isthmus of Central America. Its 180 mph winds, accompanied by 62 cm of rain over a 6-hour period, first devastated the Island of Guanaja near the coast of Honduras. In 2 days, torrential rain immobilized the Honduran Capital of Tegucigalpa, then moved inland where rainfall estimated to be 63–125 cm caused flooding and mud slides that wiped out complete villages in the mountains of Honduras and Nicaragua. On 30 October 1998, the rim of the Nicaraguan Volcano Casitas burst, releasing a deadly torrent of water and mud which destroyed the villages of Rolando Rodriguez and Posoltega as most of its residents and several thousand others in Nicaragua, were killed. Virtually all of Honduras, from the lowland marshes on the Atlantic coast to the mountains, hills, and plateaus of the interior, along with most of the highways, tens of bridges, and 70% of the national agriculture in Honduras, was destroyed. President, Carlos Flores Facusse, announced: "There are corpses everywhere from victims of landslides or floodwaters. The floods and landslides erased from the map many villages and households as well as whole neighborhoods of cities. We have before us a panorama of death, desolation, and ruin throughout the national territory." The number of dead in Honduras, Nicaragua, El Salvador, Guatemala, Costa Rica, Belize, and Mexico was estimated at 10 000 with 13 225 people missing, four million people homeless, and very little clean water, food, and medicine.

When hurricane Stan struck Central America on 5 October 2005 it brought several days of rain and caused major flooding which impacted 1000 Central American communities which were covered partially or completely by water and mud. Along the Southern Coast and western highlands of Guatemala flooding and landslides caused loss of life, injury, and displaced persons as well as damage to housing and infrastructure in 251 of 331 municipalities in 15 of the country's 22 departments. There were over 900 landslides and considerable damage to a high proportion of roads and bridges. In the highland and south-western departments of Solola and San Marcos,

entire villages were swept away by landslides, with significant loss of life. One village was completely buried under tons of dirt and declared a Mayan mass grave site.

When category 3 hurricane Katrina struck the gulf states of Louisiana and Mississippi on 28 August 2005 there was flooding both from storm surges and heavy rains. The storm surge surpassed 6 m and caused major property destruction. The subsequent flooding of the City of New Orleans was unexpected and particularly devastating. Although hundreds of thousands of residents had left the danger zone prior to 29 August, there were yet thousands more who either refused to evacuate or were unable to leave. As Lake Pontchartrain rose to unprecedented heights, water began pouring into the city through broken levees. Over the next 2 days, the gradual rising water levels drowned the unfortunate who could not escape or forced them to their attics and rooftops.

On 4 September 2005, the Mayor of New Orleans speculated to reporters that the death toll might rise into the thousands after the clean-up was completed. By early November the number of confirmed fatalities in Louisiana reached a toll of 1056, a lower figure than previously estimated. Some fatalities were caused by self-destructive behaviors or violence, including one death at the Superdome evacuation site from a drug overdose and another from suicide. At least seven homicide victims were found in the city according to the medical examiner.

In early October 2005, heavy rain struck the eastern seaboard due to tropical storm Tammy and heavy rains and flooding which impacted five states and created enormous stress, according to the following time-table:

- Sunday, 9 October – Prolonged, heavy rain and flooding from North Carolina to New Hampshire. Hundreds of people evacuated, power lost, dams weakened and roads impassable. At least four people died, including two killed in New Hampshire when a car drove off a washed-out bridge into floodwaters. New Hampshire's Governor John Lynch declared a state of emergency and called in 500 National Guard for assistance.
- Monday, 10 October – At least 10 people dead and about six unaccounted for, including a couple whose house was washed away by a surge of water over Warren Lake dam in Alstead, New Hampshire.
- Tuesday, 11 October – Rescue crews and police dogs searched rivers and woods for people missing in New Hampshire after the weekend flooding. John Lynch said the floods are the worst the state had experienced in a quarter of a century.
- Thursday, 13 October – Steady rain fell around the saturated northeast, swamping roads, stalling airline passengers, and sending streams and rivers over their banks. Flood warnings covered parts of Connecticut, New York, and New Jersey.
- Friday, 14 October – Toilets backed up with sewage, military trucks plowed through headlight-high water to rescue people, and swans glided down the streets as rain fell for an eighth straight day around the waterlogged northeast. Overflowing lakes and streams forced hundreds of people from their homes, tens of thousands of sandbags were handed out in New Hampshire and flood warnings covered parts of New Jersey, New York, and Connecticut.
- Saturday, 15 October – The death toll from a week of driving rain and swelling floods across the Northeast rose to 11 after a 75-year-old Connecticut man was swept away by rushing water at a campground. Flooding also shut down major highways and forced hundreds to evacuate their homes. Massachusetts Governor Mitt Romney declared a state of emergency, following the lead of New Jersey. Flood warnings were in effect for New Jersey, New Hampshire, Connecticut, Massachusetts, Rhode Island, and southern New York.
- Sunday, 16 October – In Connecticut, the body of a woman who fell into the rapids of the Natchaug River in Chaplin was found.
- Tuesday–Saturday, 18–22 October – Taunton, Massachusetts. Population 50 000. Residents prepared for the worst after the 173-year-old, 3.6 m high Whittenton pond dam on the rain-swollen Mill river deteriorated, threatening a wall of water up to 1.8 m deep. Schools and businesses closed and thousands of residents were unable to return home as a weakened timber dam threatened to give way and spill a surge of water into downtown Taunton, Massachusetts. Over the next 4 days there was continuing heavy rain. About 2000 residents were evacuated and workers pumped millions of liters of water from the rain-swollen lake above the dam and begin shoring up the battered structure ahead of the new round of heavy rain that was forecast. Working in the rain, crews finished building a new rock dam and tore down the old wooden one that had buckled after a week of heavy downpours.

Postdisaster Coping

How do survivors cope with impending disasters related to rapidly rising water or flash floods? They have been observed to traverse through five psychological and emotional stress-related phases as follows.

Immediate Coping – Alarm and Outcry

In some cases victims become confused and overwhelmed by fear. However, most survivors respond to the danger with little or no expectation of personal risk, often with heroic efforts, clear thinking, and physical endurance. There are also those who refuse to believe the seriousness of the danger.

Early Adaptation – the Disbelief and Denial Response

Incredulity in the face of an extraordinary event should be perceived as normal, even potentially life-saving. But those who refuse to believe they will ever be overcome by life threatening circumstances can be at risk.

In some cases disbelief or denial can be a fatal response when individuals go too far and will not acknowledge they are in mortal danger. This attitude of omnipotence has been called the delusion of personal invulnerability. There are many examples throughout history. When the dikes broke, and 1835 people were drowned in The Netherlands in the early 1900s, many perished because they did not listen to those who warned of the danger. In somewhat similar circumstances, as the Rio Grande began to flood near Eagle Pass, Texas and Piedras Negras, Mexico in June 1954, warnings were issued for 36 h that a record crest was sweeping down the river. Nevertheless people refused to comprehend the danger and 130 persons were needlessly drowned.

Ideally, individuals will respond during this phase in a dramatic way to save lives. A very recent example occurred in Taunton, Massachusetts on 18 October 2005 when residents feared for the worst after the wooden Whittenton pond dam on the rain-swollen Mill River deteriorated and threatened to collapse (see earlier).

The fact that most survivors in the middle of a disaster are somewhat numb to the reality that they could be killed helps them contend with incredible circumstances as they organize themselves and other survivors in adaptive and innovative ways. Even after the disaster's force diminishes, survivors are likely to maintain an attitude of disbelief for a time until they begin to resume normal activities. In some cases, the demands on rescue workers and other helpers are exceedingly great during this time and the risk of developing significant physical and emotional symptoms is high. Reports from New Orleans' Methodist Hospital after hurricane Katrina had flooded New Orleans indicated that the medical staff worked unendingly in horrendous conditions and some succumbed to dehydration and exhaustion. Patients requiring ventilators were kept alive with hand-powered resuscitation

bags. When the first floor of the hospital flooded the dead were stacked in a second floor operating room. Other hospitals reported even more dire conditions and in some cases the hospital staff were unable to evacuate the severely ill and they did not survive.

Mid-Phase Adaptation – Intrusions

The middle phase is marked by intrusive recollections alternating with periods of denial and numbing as survivors begin to remember and relive their experiences in the form of intrusive thoughts, emotions, and dreams several days or weeks later. Their intrusive symptoms may be intensified when survivors finally realize they may have lost lives, health, or property. This phase may be prolonged or delayed when the work of recovery, reconstruction, or self-preservation becomes all consuming and there may also be delayed resolution by the use of alcohol or drugs.

Survivors often contend with a variety of other stresses at this time including possible injuries, losses, role changes, displacement, cleaning up, repairing property damage, moving into temporary or even new housing, and preparing lengthy health, recovery assistance, and insurance claims. Thus, posttrauma symptoms are related not only to the disaster itself but other related difficulties, which has been called the dripping faucet syndrome. Taken individually, many small stresses seem minor but when taken collectively they can push a person to the edge and cause irritability, withdrawal from relationships, or loss of pleasure in activities.

Disaster workers subjected to prolonged or highly intensive exposure to the disaster tend to remain emotionally numb until the enormity of the disaster finally sinks in. Not until then do they develop sleep disturbances and nightmares but these are frequently resolved within 4–6 weeks.

Late Phase Adaptation – Anger and Physical Symptoms

After a period of time (from 1 to 3 months) survivors are more prone to experience negative and hostile reactions. There is a marked decrease in tolerance, an increase in complaining, a loss of humor, and mistrust of others. This becomes a problem for those trying to help victims of trauma and for the victims themselves, who keep others at a distance. Couples may become more easily aggressive and prone to violent outbursts. During this phase there is also a marked increase in physical complaints including headaches, nausea, diarrhea, vomiting, muscle aches, chest pain, restlessness, tremors, sweating, and general fatigue. These symptoms may also be associated

with secondary victimization befalling those who are exploited by shysters, neglected by thoughtless people, and mistreated by dehumanizing agencies.

Resolution or Symptom Development

Finally, survivors either resolve their posttraumatic symptoms, develop recurring difficulties related to unresolved past stresses, or become victims of prolonged or worsening disorders associated with unresolved guilt, preexisting anxiety or depression, prior traumatization, drug or alcohol addiction, and organic syndromes. New symptoms may develop or prior difficulties may worsen including flashbacks, depression, obsessive-compulsive symptoms, panic attacks, nightmares, sleeplessness, interpersonal alienation, and emotional outbursts alternating with detachment. For some victims, these symptoms may extend indefinitely in the form of a chronic emotional or physical disorder requiring treatment.

Postdisaster Effects on Children

Children are vulnerable to all of the typical responses and symptoms following disasters and need special attention in order to prevent longstanding difficulties. Unfortunately, adults are not always aware of the full impact of traumatic stress on children nor do they always help them resolve their frightening dreams and memories. Most children, especially younger ones, become very fearful of being alone, overdependent, and sometimes clinging. They may also develop strong phobias, nightmares, night terrors, and regression to a previous level of function, including thumb sucking or bed wetting in younger children. Additional symptoms include clinging to parents, reluctance to go to bed, nightmares, fantasies that the disaster never happened, crying and screaming, withdrawal and immobility, refusal to attend school, school problems, inability to concentrate, and being poorly understood by parents and teachers.

Early intervention is particularly important for children. Parents need to recognize that withdrawal, fear, and overdependency will improve with time. Special care should be paid to helping the children resume normal daily routines within a safe and secure environment. Watching television programs about the flood can be helpful if parents are there to talk about what happened and reassure that being afraid is normal. Supervised trauma-related play can help children resolve their fears. They should be reassured with explanations like: "That hurricane is not a monster who took away children's toys and hurt people. The 'eye' will not seek children out or return again."

Postdisaster Physical Symptoms

Disaster survivors not uncommonly have physical complaints which often bring them to their doctor's offices. Following sudden disasters, many of the symptoms are associated with prolonged autonomic nervous system over-reactivity. Palpitations, tightness in the chest, shortness of breath, and feelings of dread may accompany intrusive memories. Headaches and gastrointestinal complaints are not uncommon. While survivors may have physical problems they often do not seek medical help. Researchers have found that only 40% of those reporting physical symptoms after a disaster consulted their doctors while 34% sought help from family members or friends. Most of them, 66%, sought medical help within 4 months while the remaining 33% waited until after 4 months to visit their physicians.

Survivors may also suffer from health problems indirectly linked to the disaster. For example, following the Iowa floods in the early summer of 1993, the national Centers for Disease Control (CDC) conducted a public health investigation. The CDC researchers discovered that survivors in seven Iowan counties had been hospitalized for flood-related illnesses or injuries, including carbon monoxide poisoning, hypothermia, electrocution, wound infections, and exacerbations of chronic illnesses.

Sometimes there are reports of unusual and long-term health problems. Investigators from the New York State Department of Health discovered some unusual health problems following a flood disaster in western New York in 1974. Within a population of 100 000 people exposed to flooding in three river valleys, scientists found a regional rise in spontaneous abortions as well as an increased incidence of leukemias and lymphomas of 4 years' duration, resulting in an excess of 30–35 malignancies. The reasons are not yet entirely clear.

After the flooding of New Orleans began on 29 August 2005, there was concern about a potential outbreak of serious health problems. Flood waters remained at high levels, over some roof tops, for a number of weeks and became toxic and a breeding area for bacteria. In addition to dehydration and food poisoning, New Orleans flood survivors were also at risk to contract hepatitis A, cholera, and typhoid fever, because of contaminated food and drinking water. Survivors also faced longer-term health risks due to prolonged exposure to the petrochemical tainted flood waters and mosquito-borne diseases such as yellow fever, malaria and West Nile Virus. *Escherichia coli* was detected in the water supply within a few days and at least five people died from bacterial infections caused by *Vibrio vulnificus* bacteria in the

toxic waters. Other sources of contaminations included dead bodies lying in city streets and floating in still-flooded sections, causing an advanced state of decomposition of many corpses, some of which were left in the water or sun for days before being collected.

Medical assistance began in earnest within 3 days when an emergency triage center was set up at the regional airport at New Orleans. Helicopters and ambulances brought in the weak, elderly, sick, and injured and makeshift gurneys were used to transport persons from the flight line to a temporary hospital set up in the terminal. The captain in charge described the site as organized chaos but the emergency medical staff assembled from around the country kept pace and were able to handle anything from bruises to critical cases requiring ventilators. They triaged the seriously ill and injured to medical centers in the surrounding states and within 72 h had triaged up to 5000 people.

The potential health problems from flooding affected the residents of the city who had been evacuated. They were unable to return to their homes because of the length of time necessary for the Army Corps of Engineers to drain the flood waters. They managed to accomplish the task within 1 month but parts of the city re-flooded when hurricane Rita struck on 24 September. Many buildings that had withstood the initial storm and the direct force of the flooding were damaged beyond repair by the deleterious effect of long immersion and the rapid propagation of mold.

Hundreds of thousands of New Orleans residents that were displaced by the flooding lost their homes completely and the subsequent emotional effects will not be known for some time.

Postdisaster Effects on Emergency and Medical Personnel

Emergency medical technicians, law enforcement personnel, military medics or corpsmen, fire fighters, emergency room workers, physicians, intensive care nurses, and other crisis workers are exposed routinely to persons who are severely injured, killed, deathly ill, or have expired unexpectedly. These helpers usually adapt to their stress by becoming emotionally numb and detached from their feelings. As a result they can develop a pattern of maintaining emotional distance from people while preferring to throw themselves into their work. This accumulated exposure to traumatic events eventually causes posttraumatic symptoms: emotional detachment, sleep disturbance, irritability, marital problems, alcohol abuse, physiological symptoms, and an addiction to repetitive

stress. Helpers (health professionals, mental health workers, ministers, and even business executives) can feel used up. They eventually find it difficult, if not impossible, to meet the emotional and communication needs of their wives and children. They even ignore their own personal needs for rest and recreation as they often turn to work for gratification rather than their families. Eventually this leads to burnout.

Helpers may find it emotionally stressful and experience intense feelings in response to listening to victims' stories about disasters, assault, sexual abuse, and violent trauma. Those with little experience may even feel confused and even ashamed while empathically taking on the victim's emotional distress. The listener's experience may lead to symptoms of depression, bad dreams, irritability, excessive drinking, and other distressing difficulties caused by a process called vicarious traumatization (VT).

Assessing Traumatic Exposure Severity and Symptoms

The severity of symptoms, should they persist, often are frequently directly correlated with the enormity of the traumatic experience/loss or accumulation of prior unresolved traumatic experiences. Survivors may have experienced a seemingly unending series of stressors. For example, they may have suffered damage or destruction to their homes or personal belongings. They may have witnessed violence, injury, deaths of other victims, and they may have lost loved ones and pets.

When survivors fail to talk openly about the event(s); if they have been separated from family members; if they lack emotional support from friends or family members, the emotional aftermath is more severe. They are at risk to experience many or all of the following symptoms: flashbacks, sleep disorder, anxiety attacks, recurring nightmares or dreams, problems with concentration and memory, intrusive recollections, irritability, emotional outbursts alternating with periods of emotional numbing, misusing drugs or alcohol, and interpersonal conflicts. Clearly, individuals who are experiencing these symptoms should seek the help of an experienced trauma therapist. A successful resolution is possible when survivors find relief from their symptoms and discover new or broadened meaning in their lives.

See Also the Following Articles

Acute Stress Response: Experimental; Anxiety; Adrenaline; Hurricane Katrina Disaster, Stress Effects of; Other Natural Disasters, Illness Epidemics Related to Contamination.

Further Reading

- Brende, J. O. (1998). Coping with floods, assessment, intervention, and recovery processes for survivors and helpers. *Journal of Contemporary Psychotherapy* 28, 107–140.
- Brende, J. O. (1995). *Coping with floods and other disasters: a workbook for physicians, helpers, and survivors*. Columbus, GA: Trauma Recovery Publications.
- Clayer, J. R., Bookless-Pratz, C. and Harris, R. L. (1985). Some health consequences of a natural disaster. *Medical Journal of Australia* 143, 182–184.
- Figley, C. R. (ed.) (1985). *Trauma and its wake, volume 2: traumatic stress: theory, research, and intervention*. New York: Brunner/Mazel.
- Forster, P. (1992). Nature and treatment of acute stress reactions. In: Austin, L. S. (ed.) *Responding to disaster, a guide for mental health professionals*, pp. 2–51. Washington DC: PA Press.
- Kalayjian, A. S. (1995). *Disaster and mass trauma: Global perspectives on post disaster mental health management*. Long Branch, NJ: Vista Publishing.
- McCann, I. L. and Pearlman, L. A. (1990). Vicarious traumatization: a contextual model for understanding the effects of trauma on helpers. *Journal of Traumatic Stress* 3, 131–149.
- Murphy, S. A. (1986). Perceptions of stress, coping, and recovery one and three years after a natural disaster. *Issues in Mental Health Nursing* 8, 63–77.
- Olson, L. (1993). After the flood: the dripping faucet syndrome. *Iowa Medicine* 83, 324–327.
- Pearlman, L. A. and Saakvitne, K. S. (1995). *Trauma and the therapist*. New York: WW Norton.
- Robinson, D. (1959). *The face of disaster*. New York: Doubleday.
- Tierney, K. J. (1986). The social and community contexts of disaster. In: Gist, R. & Lubin, B. (eds.) *Psychosocial aspects of disaster*, pp. 11–39. New York: Wiley.

Food Intake and Stress, Human

K Smith and G M Goodwin

Oxford University, Oxford, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 158–161, © 2000, Elsevier Inc.

Effect of Stress on Food Intake

Effect of Food Composition and Intake on the Ability to Cope with Stress

Glossary

<i>Dietary restraint</i>	The tendency to control food intake consciously in order to prevent weight gain or promote weight loss.
<i>Dieting</i>	A behavior designed to bring about weight loss that demands a relatively severe degree of food restriction.
<i>5-Hydroxy-tryptamine (serotonin)</i>	A neurotransmitter, present in the brain, that is involved in controlling many different behaviors, including food intake, mood, motor output, learning, sleep, circadian pattern, aggression, and sexual behavior.
<i>Macronutrients</i>	Components of food, including carbohydrate, fat, and protein.

Satiety

The feeling of fullness that increases as food is ingested and that increases the drive to stop eating.

Tryptophan

An amino acid that forms one of the essential components of protein.

Effect of Stress on Food Intake

Under a variety of stressful conditions, food intake is altered in both animals and humans. However, different stressors seem to have different effects, with some resulting in an increase and some in a decrease in food intake. In human studies, some subjects appear to be nonresponders to stress, whereas of the responders, some markedly reduce their intake (stress fasters) while others increase their intake (stress eaters). Stress-induced eating is of particular interest because of the role it may play in contributing to obesity. Much over-eating behavior is reported to be stress related, and some obese subjects report a decrease in anxiety after eating.

In everyday life there are a number of specific stressors that have an impact on food intake, particularly in women. Since the 1960s, Western society has placed increasing demands on the population, particularly women, to be slim and to aspire to an ideal of thinness. There has been an associated increase in

body image dissatisfaction, particularly among young women. This has also coincided with a dramatic increase in the prevalence of dieting and of eating disorders such as anorexia and bulimia nervosa. The prevalence of eating disorders seems to be in direct proportion to the prevalence of dieting behavior in a given community, and longitudinal studies of young adults suggest that the risk of developing an eating disorder is significantly higher among dieters than non-dieters. Indeed, most of those suffering from an eating disorder report that their eating problem started as normal dieting that then got out of control.

Low self-esteem is an additional factor associated with dieting, body image dissatisfaction, and alterations in eating pattern, although it is difficult to know whether this is a cause or an effect. Subjectively, low mood and low self-esteem are often reported to be contributing factors to unwelcome weight gain. Experimentally induced stress or dysphoria in volunteers causes an increased food intake in some individuals. This effect occurs particularly in those subjects who are dieting or are restrained eaters, but not in those who are not restricting their food intake. Interestingly, subjects who binge eat often report that this occurs in response to stress or dysphoria and, at least initially, binges seem to relieve anxiety.

Dieting and dietary restraint are intimately bound up with mood and self-esteem. Subjects who restrict their food intake to lose weight record ratings of happiness that are directly related to weight loss or weight gain. In contrast, subjects who are not restricting intake report that weight change has little impact on ratings of happiness. Feedback to subjects about weight change will therefore alter mood, which in turn can affect food intake and weight. In a study where subjects were weighed regularly but the results were fed back to them as consistently heavier than their true weight, restrained eaters reported more mood changes than nonrestrained eaters in response to the feedback, including anxiety, depression, and low self-esteem. Interestingly, these restrained eaters, who experienced lowered mood, then ate more in a subsequent test meal.

Thus, there is growing evidence that stress can alter food intake. However, there is debate about how this effect is mediated. Stressful events can alter a number of different biological variables, which in turn could affect food intake. It has been hypothesized that stress-induced increases in food intake are mediated by endogenous opiates. Stress increases the release of β -endorphin, a naturally occurring opiate that increases food intake, but also reduces the perception of pain and the anxiety associated with stress.

An alternative is that these effects are mediated by the neurotransmitter 5-hydroxytryptamine (5-HT). 5-HT has attracted great interest because its function

seems to be altered in disorders such as depression, which are known to be closely associated with the occurrence of stressful life events, and in eating disorders such as bulimia nervosa. In addition, altering 5-HT function (e.g., by using medication) has a profound effect on eating behavior. In human and animal studies, interventions that increase 5-HT function reduce food intake; conversely, reducing 5-HT function increases food intake.

Could levels of 5-HT be altered by stress? Stress increases the level of a naturally occurring hormone, cortisol. When this occurs, 5-HT function is reduced. Reduced 5-HT function would be expected to lead to increased food intake, and experimental work suggests that this increase would occur particularly in carbohydrate intake, rather than protein. Interestingly, this alteration in the balance of macronutrients could itself affect 5-HT function. This is discussed in more detail in the next section.

Effect of Food Composition and Intake on the Ability to Cope with Stress

Research on the biological and psychological effects of varying food intake has focused on two main areas. First, there has been interest in the behavioral effects of dieting. Second, there has been interest in the effects of dieting on biological parameters, some of which may affect the ability to cope with stress. Research in the latter area has focused on the effects of dieting on the neurotransmitter 5-HT. 5-HT is biologically unusual in that the rate of its synthesis is entirely dependent on the supply of its essential amino acid precursor, tryptophan, in food. Furthermore, the availability of tryptophan to the brain depends on the balance of macronutrients (carbohydrate, protein, and fat).

Dieting is a very common behavior in Western cultures. Approximately 25% of men and 40% of women in the United States report that they are currently trying to lose weight. Dieting seems to be particularly common among adolescent girls, but the rates are higher at all ages in women compared with men.

Behavioral Effects of Dieting

The first direct evidence on the behavioral effects of food restriction came from a study of World War II conscientious objectors who were placed on a 24-week semistarvation diet. The main findings were an increase in lethargy, tiredness, depression, and irritability and a lowering of energy. In contrast, other studies using less extreme dietary restriction have found less dramatic changes in mood.

Several studies have shown that dieters perform less well on cognitive tasks. Dieting impairs sustained attention and short-term recall and results in slower

reaction times. Nondieters with high levels of restraint over eating showed intermediate cognitive performance between dieters and low restraint nondieters. Impairments in cognitive functioning, such as impaired vigilance, have also been noted among subjects with bulimia nervosa and anorexia nervosa, probably due to restricted eating.

Dieting itself affects eating behavior. In a series of well-replicated studies, Polivy and Herman showed that dieters failed to compensate for a modest calorie loading by decreasing their caloric intake at the next meal. They actually ate more after the preload than after no preload; they counterregulated rather than showing the normal compensatory regulation seen in the controls. It has been widely argued that counterregulation is a result of the preload causing disinhibition and abandoning of dietary restraint in the dieter. This cognitive explanation is supported by the finding that dieters also counterregulate when they perceive the preload as high in calories, regardless of the actual calorie content.

Severe dietary restriction in normal controls can result in binge eating. The World War II study of conscientious objectors described earlier also investigated the effects of refeeding. The men returned to their normal weights after the study but exhibited a persistent tendency to binge eat. This type of behavior had not occurred prior to the study. Healthy volunteer studies during short-term calorie-controlled diets show that subjects have increased preoccupation with thoughts of food, urges to eat more frequently, and feelings of being out of control with their eating. Rates of postwar binge eating in World War II combat veterans and prisoners of war show that binge eating is significantly more prevalent in veterans who lost large amounts of weight during their captivity.

Thus, dietary restriction and normal dieting have significant effects on behaviors such as food intake, mood, and cognitive functioning. There has been debate, however, about how these effects are mediated. One possibility is that dieting alters 5-HT function, discussed next.

Dieting and 5-HT

The rate of synthesis of 5-HT in the brain is critically dependent on the amount of its precursor, tryptophan, which is able to enter the brain. Tryptophan is available solely from the diet, and its availability in the brain can be critically affected by alterations in food intake, particularly in the balance of macronutrients. Tryptophan competes for entry to the brain across the blood-brain barrier with other amino acids (components of protein). Any change in food intake that alters this ratio of tryptophan to other amino acids will directly affect the amount of 5-HT synthesized.

Studies of the effects of a standard 3 week, 1000 kcal diet have shown that the diet results in a lowering of tryptophan in the plasma (blood). The key question is whether this affects 5-HT function itself. Studies using a variety of challenge tests for different aspects of 5-HT function show that it does, although the effect is significant only in women. It seems that female dieters can compensate to some extent for the reduction in 5-HT synthesis induced by dieting, but it is likely that dieting results in a reduction in 5-HT function. Because this effect is gender specific, women may be biologically more vulnerable to the effects of dieting on 5-HT function.

What might be the consequences of this lowered 5-HT function after dieting? First, decreased 5-HT function will reduce satiety and increase the drive to eat. This may be one of the reasons why it is often reported that dieting is difficult to sustain. Some subjects who are dieting report craving carbohydrate or even bingeing on high-carbohydrate, low-protein foods. Wurtman and colleagues have proposed that this craving may be explained by the fact that these foods increase 5-HT function indirectly. This occurs because a carbohydrate-rich meal causes secretion of the hormone insulin, increasing the ratio of tryptophan to other amino acids and an increase in the entry of tryptophan to the brain. The end result is an increase in the synthesis of 5-HT, which to some extent may correct the deficit induced by dieting.

It is also possible that the reduction in 5-HT function might affect mood. Experimental work using an artificial mixture of amino acids to temporarily lower 5-HT synthesis has shown that this can induce symptoms of depression, but only in women who have a strong previous history of depression. Animal and human work suggests that a key role of 5-HT is to promote resilience in the face of stressful circumstances. Failure of this system, thus reducing a person's ability to deal effectively with a stressful life event, may be one of the key factors contributing to the onset of depression. If dieting impairs 5-HT function directly, it may contribute to the onset of depression after a stressful life event. However, this may occur only in those people who have a vulnerability to depression because of their heredity or past history. Further research on the effects of dieting, particularly in people at risk of depression or eating problems, will need to be performed before this hypothesis can be confirmed.

See Also the Following Articles

Eating Disorders and Stress; Self-Esteem, Stress and Emotion.

Further Reading

- Cowen, P. J., Clifford, E. M., Williams, C., Walsh, A. E. S. and Fairburn, C. G. (1995). Why is dieting so difficult? *Nature* 376, 557.
- Deakin, J. F. W. and Graeff, F. G. (1991). 5-HT and mechanisms of defence. *Journal of Psychopharmacology* 5, 305–315.
- Hsu, L. (1997). Can dieting cause an eating disorder? *Psychological Medicine* 27, 509–513.

- Polivy, J. and Herman, C. (1993). Etiology of binge eating: psychological mechanisms. In: Fairburn, C. & Wilson, G. (eds.) *Binge eating: nature, assessment and treatment* (pp. 173–205). New York: Guilford Press.
- Smith, K. A., Fairburn, C. G. and Cowen, P. J. (1997). Relapse of depression after rapid depletion of tryptophan. *Lancet* 349, 915–919.
- Wurtman, R. J. and Wurtman, J. J. (1995). Brain serotonin, carbohydrate-craving, obesity and depression. *Obesity Research* 3(Supplement 4), 477s–480s.

Food Intake and Stress, Non-Human

Belinda A Henry and Iain J Clarke
Monash University, Melbourne, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition
article by I J Clarke, volume 2, pp 162–166,

© 2000, Elsevier Inc.

Satiety factor A factor that acts on the appetite-regulating regions of the brain to indicate satiety or to reduce food intake.

Introduction

Nonhormonal Responses to Nutritional Stress

Hormonal Responses to Nutritional Stress

Nutritional Stress and the Hypothalamic-Pituitary-Adrenal Axis

Neuronal Systems Underlying Endocrine Perturbations Induced by Nutritional Stress

Possible Hormonal Signals Involved in Mediating Nutritional Stress

Glossary

<i>Basal metabolic rate</i>	The rate of oxygen consumption, or calculated equivalent heat production, of an awake individual lying at rest who has had no food for at least 12 h.
<i>Gluconeogenesis</i>	The formation of glucose from non-carbohydrate sources such as lactate or fatty acids.
<i>Nutritional stress</i>	A physiological, environmental, or metabolic stress that is imposed on an organism to disrupt metabolic homeostasis, thus causing a neural and/or endocrine response.
<i>Orexigenic factor</i>	A factor that acts to increase appetite and stimulate food intake.

Introduction

Nutritional stress is, in some cases, an environmental stress and, in all cases, a metabolic stress. In particular cases, such as pregnancy and lactation, it could be construed that the nutritional demands of these states impose a physiological stress demanding increased food intake. In nonhuman species, nutritional stress occurs with a reduction in the availability of food or increased metabolic demand, but obesity is also associated with increased levels of stress hormones. In humans, a nutritional stress may be voluntarily imposed or due to psychogenic factors that have poorly defined etiology. The degree of impact varies among species depending on the type of metabolism of the organism (monogastric versus ruminant, for example) and also on the amount of body reserves of the organism.

Nutritional stress causes an alteration in the metabolic status of the animal and leads to altered concentrations of circulating metabolic substrates and metabolic hormones such as insulin and leptin. It is clear that blood-borne substances signal the metabolic status of the animal to the brain and that this leads to the alteration of hormonal output. These are responses of the organism to the stressor, but it is unlikely that a single factor that fulfils this signaling role. The neuroendocrine response to such a stress is probably the result of integrated effects of a composite of blood-borne substances. The gut may also

signal directly to the brain by a neural route and through the altered secretion of ghrelin to regulate brain function and hormonal status.

Nonhormonal Responses to Nutritional Stress

Stress due to undernutrition produces a catabolic state, typified by changes in the levels of various substances in the bloodstream. Changes in metabolic factors, however, can be influenced by other physiological stimuli, so variations in the concentrations of singular components do not often reflect metabolic state. For example, fat is metabolized during a fast to produce nonesterified fatty acids (NEFA) that are used to synthesize carbohydrates via gluconeogenesis. Accordingly, plasma NEFA levels are elevated during a fast. On the other hand, prolonged undernutrition decreases body fat, leading to reduced NEFA levels, and obese subjects have increased NEFA levels, due

to increased body fat. Thus, plasma NEFA concentrations are increased in both fasting and obese states. Other changes indicative of a catabolic state include reduced plasma insulin and leptin levels and increased plasma urea concentrations (due to the breakdown of muscle tissue in an effort to mobilize amino acids for the production of carbohydrates via gluconeogenesis). Chronic undernutrition (and a reduction in gluconeogenic substrates such as NEFA and amino acids) is typically associated with a reduction in plasma glucose concentrations. Plasma levels of these metabolic factors are a reflection of nutritional state (undernutrition versus obesity) and also the extent of catabolic stress (fasting versus prolonged undernutrition).

Hormonal Responses to Nutritional Stress

Endocrine changes occur during food deprivation or food restriction and the effects of fasting are summarized in [Table 1](#).

Table 1 Endocrine changes occur during food deprivation or food restriction^a

<i>Hormone</i>	<i>Change</i>	<i>Comments</i>
Small monogastric animals (e.g., laboratory rat)		
Catecholamines	↑	Reflects the output of adrenalin from the adrenal medulla; catecholamines are lipolytic, mobilizing fat stores in response to reduced energy substrate
Insulin	↓	Insulin facilitates the uptake of sugar from the bloodstream
Growth hormone	↓	Response is opposite to that seen in large animals and also opposite to the response seen in humans
Gonadotropins (LH and FSH)	↓	Fasting for a period of 1–2 days can prevent the preovulatory LH surge in rats
Gonadal steroids (estrogen, testosterone, progesterone)	↓	
Prolactin	↓	With extended period of undernutrition
ACTH	↑	
Cortisol	↑	
TSH	↓	
T ₃	↓	Through reduced conversion of T ₄ to T ₃
Nonhuman primates		
Insulin	↓	
Catecholamines	Not known	
Growth hormone	↑	
Gonadotropins (LH and FSH)	↓	LH is decreased, but there is a slower response in terms of FSH
Prolactin	Not known	
ACTH	Not known	
Cortisol	↑	
TSH	Not changed in the short term	
T ₃	↓	
Ruminants		
Catecholamines	Not known	
Growth hormone	↑	
Gonadotropins (LH and FSH)	↓	LH is decreased, but there is a slower response in terms of FSH
Prolactin	Not changed	
Adrenocorticotropin	Not known	
Cortisol	↑	Morning levels only

^aACTH, adrenocorticotrophic hormone; FSH, follicle stimulating hormone; LH, luteinizing hormone; T₃, triiodothyronine; T₄, thyroxine; TSH, thyroid stimulating hormone.

Nutritional Stress and the Hypothalamic-Pituitary-Adrenal Axis

General Considerations

The hypothalamic-pituitary-adrenal (HPA) axis is commonly activated by nutritional stress, which results in the increased secretion of glucocorticoids from the adrenal cortex. Most information to date has been derived from the laboratory rodent. The degree of activation of the HPA axis is dependent on a variety of factors, including:

- The species studied (monogastric small mammals such as rodents versus larger ruminant mammals such as sheep).
- The degree of intensity of undernutrition.
- The duration of undernutrition.
- The nutritional state of the animal prior to onset of catabolic state.
- The time of day; there is a diurnal pattern of glucocorticoid secretion.

Plasma glucocorticoids are elevated in states of undernutrition and in obese subjects, so both conditions might be regarded as stressors. The degree of change in body condition is an important factor. With undernutrition, the rapidity of change may also be a determinant of the stress response because animals can adapt over time to reduced energy stores. In obesity, the impact on the stress axis (glucocorticoid levels) depends on the obesity phenotype (where fat is distributed in the body).

There are species differences in the stress response to changing nutritional status and effects on plasma glucocorticoid levels are summarized in Table 2. It is important to note that the major glucocorticoid in rodents is corticosterone, whereas larger species including humans produce cortisol. Small monogastric mammals such as rodents exhibit robust activation of the HPA axis in response to fasting, but the same is not true for ruminants, in which digestion of nutrients is slower. Ruminants have a dynamic digestive system, with four stomach compartments (rumen, reticulum, omasum, and abomasum), which enables

them to digest their primary dietary component, cellulose. Ruminant digestion takes a considerable time, so there is a delayed response to fasting. Furthermore, ruminants stringently maintain blood glucose levels by the mobilization of NEFA from fat stores. Small animals have higher basal metabolic rates, which impacts fuel utility. Small animals (rodents) also have very low fat stores (approximately 10%) in comparison to larger species, including both sheep and humans (approximately 20%).

The Effect of Fasting versus Food Restriction on Hypothalamic-Pituitary-Adrenal Axis Activity

The effect of nutritional stress on the HPA axis is dependent on the duration and the intensity of undernutrition. Fasting is often used in the laboratory as a nutritional stressor, and in rodents plasma corticosterone levels are elevated within 24 h. In sheep, the response takes longer and cortisol secretion is increased 48–72 h after the commencement of fasting. In rodents, chronic food restriction (animals fed a 50% maintenance diet) for 4 weeks increases plasma levels of corticosterone, but basal cortisol levels are minimally changed in sheep made lean by long-term food restriction. There is a perturbation of the diurnal pattern of cortisol secretion in fasting sheep, with higher morning levels.

Pregnancy and Lactation

Pregnancy and lactation can be considered physiological conditions that induce nutritional stress by virtue of the high energy demands consigned to the mother. In small mammals, it has been calculated that the energy demands of lactation are four times greater than those of nonpregnant females. The high metabolic demands of lactation are met by profoundly increased food intake, driven by changes in appetite-regulating systems in the brain. Unlike other typical nutritional stressors, however, lactation is associated with hyporesponsivity of the HPA axis; neuroendocrine responses to stress are blunted in lactating animals. adrenocorticotrophic hormone (ACTH) and

Table 2 Species differences in glucocorticoid responses to nutritional stressors

	<i>Laboratory rodent</i>	<i>Sheep</i>	<i>Nonhuman primate</i>	<i>Human</i>
Fasting	↑ Corticosterone within 24–48 h	↑ Cortisol within 48–72 h	↑ Cortisol within 24 h	↑ Cortisol within 24 h
Food restriction	↑ Corticosterone	No change	Unknown	↑ Cortisol
Obesity	↑ Corticosterone in genetic models of obesity such as the fatty Zucker rat and the <i>ob/ob</i> mouse	No change in basal cortisol concentrations ↑ Cortisol response to stress	Unknown	↑ Cortisol (particularly in abdominal obesity)

cortisol levels are maintained during pregnancy, whereas corticotropin releasing hormone (CRH) levels are elevated due to peripheral synthesis and secretion by the placenta. This increase in CRH production is counteracted by increased circulating levels of CRH-binding protein (which sequesters CRH), thus preventing inappropriate secretion of ACTH. This phenomenon, however, occurs only in mammals that produce placental CRH.

Nutritional Stress and Negative Feedback Mechanisms of the Hypothalamic-Pituitary-Adrenal Axis

The HPA axis is tightly regulated by glucocorticoid (and mineralocorticoid) feedback to the hypothalamus and the anterior pituitary gland to regulate the secretion of hypophysiotropic hormones (CRH and arginine vasopressin, AVP) from the hypothalamus and ACTH from the pituitary corticotrophs. It has been hypothesized that abdominal obesity in humans and other species leads to dysfunction of the negative feedback actions of glucocorticoids, causing upregulation of the HPA axis and increased plasma levels of cortisol. In rodents, it has been hypothesized that decreased negative feedback actions within the brain actually precedes the obese state; animals susceptible to diet-induced obesity have decreased levels of glucocorticoid receptor mRNA within the hippocampus compared to diet-resistant animals.

Some work suggests that the sensitivity of the HPA axis to the negative feedback effects of glucocorticoids depends on carbohydrate ingestion. Adrenalectomized rodents suffer various abnormalities within the HPA axis, including increased expression of CRH mRNA in the paraventricular nucleus of the hypothalamus and decreased expression of CRH mRNA in the central nucleus of the amygdala. Interestingly, these perturbations in the adrenalectomized animal can be ameliorated by the ingestion of sucrose. Similarly, feeding animals a diet high in fat content also impacts on the HPA axis. For example, feeding rats a diet containing 20% fat (control diet, 4% fat) increases both basal and stress-induced secretion of ACTH and corticosterone. In contrast, feeding rats a diet very high (60%) in fat can decrease hormonal responses to stress. Thus, in animals, as in humans, the macronutrient dietary content impacts the reactivity of the HPA axis.

Nonadrenal Production of Glucocorticoids in Obesity

Peripheral tissues, other than the adrenal cortex, are able to produce glucocorticoids. This occurs by conversion of inert 11-keto steroids (11-dehydrocorticosterone in rodents and cortisone in humans) to active

11-hydroxy steroids (corticosterone in rodents and cortisol in humans). This process is regulated by the 11 β -hydroxysteroid dehydrogenase (11 β -HSD) enzymes; 11 β -HSD-1 catalyzes the synthesis of active 11-hydroxy steroids, and 11 β -HSD-2 catalyzes the synthesis of the inactive 11-keto steroids. The 11 β -HSD-1 enzyme is predominantly expressed in white adipose tissue, the liver, and the central nervous system (particularly the paraventricular nucleus of the hypothalamus and the hippocampus). In genetic models of obesity such as the *ob/ob* mice (a naturally occurring mutation in the gene that encodes leptin) and the fatty Zucker rat (a mutation in the gene that encodes the leptin receptor), the expression of 11 β -HSD-1 mRNA and 11 β -HSD-1 enzymatic activity is increased within adipose tissue, whereas expression is reduced in the liver. Furthermore, overexpressing 11 β -HSD-1 mRNA within the adipose tissue of transgenic mice specifically increases local concentrations of corticosterone. Such animals have an obese phenotype, but this is not associated with increased plasma levels of corticosterone.

The Role of the Hypothalamic-Pituitary-Adrenal Axis in the Predisposition to Obesity

In rodents, genetic obesity (the *ob/ob* mouse and the fatty Zucker rat) causes elevated concentrations of corticosterone. Adrenalectomy attenuates (or completely reverses) the obese phenotype in these animals. Furthermore, in nonmutated rodents, adrenalectomy reduces food intake, and corticosterone replacement restores normal feeding behavior. The stimulatory effect of glucocorticoids on food intake, however, cannot be displayed in an adrenal-intact rodent.

In rodents, mild pain induced by tail pinch initially increases food intake, but this effect is abolished by repeated experience; chronic exposure to tail pinch reduces food intake and body weight. Similarly, in rats repeated restraint stress reduces food intake and body weight, and this effect can be mimicked by chronic treatment with corticosterone. This decrease, however, is in comparison to rats fed *ad libitum*. In contrast, animals pair-fed to a similar (and lower) plane of nutrition tend to have lower body weights and reduced fat deposition compared to animals exposed to either repeated restraint stress or glucocorticoid treatment.

Neuronal Systems Underlying Endocrine Perturbations Induced by Nutritional Stress

There are numerous neuropeptides and neurotransmitters that are involved in both the regulation of

appetite and neuroendocrine functions. Within the hypothalamus, this involves a number of interconnected nuclei. The arcuate nucleus is essential to the regulation of food intake and body weight and produces a number of neuropeptides, including the appetite stimulators neuropeptide Y (NPY) and agouti-related protein (AGRP) as well as the appetite inhibitors α -melanocyte stimulating hormone (α -MSH; derived from the proopiomelanocortin, POMC, gene) and urocortin 2. Other regions of the hypothalamus involved in regulating food intake and neuroendocrine function include the paraventricular nucleus, which produces the appetite inhibitors CRH, urocortin 1, urocortin 2 and urocortin 3; the lateral hypothalamic area, which contains the appetite stimulators melanin-concentrating hormone (MCH) and the orexin; and the ventromedial hypothalamic nucleus. Various appetite-regulating peptides impact on the HPA axis and these are summarized in Table 3.

In addition to the hypothalamus, several nuclei in the brain stem also regulate feeding behavior and the HPA axis. Brain stem nuclei can regulate HPA function by ascending projections to the forebrain or via descending spinal projections to the adrenal gland. For example, noradrenergic input from the

locus ceruleus to the paraventricular nucleus of the hypothalamus is involved in the regulation of HPA function as well as food intake. In sheep and rats, various studies document the effects of nutritional stress on the expression of genes that encode appetite-regulating peptides. For example, in sheep made lean by chronic food restriction, the expression of genes encoding NPY, AGRP, and MCH is increased, indicative of increased hunger drive; these animals, however, have normal basal cortisol secretion. On the other hand, in fasted rodents, the expression of the same genes (NPY, AGRP, and MCH) is increased and there is activation of the HPA axis. Species differences in neuronal circuitry regulating the HPA axis may be relevant to the way that alteration in body weight affects the stress axis.

Possible Hormonal Signals Involved in Mediating Nutritional Stress

Various hormonal signals relay information about nutritional state to the appetite-regulating and neuroendocrine centers of the brain. Three hormones that play an important role are leptin, ghrelin, and insulin.

Leptin

Leptin is primarily produced by white fat tissue, and its plasma levels reflect the level of adiposity. In rodents, there is a tight negative correlation between the plasma levels of leptin and corticosterone. In rodents and sheep, the fasting-induced activation of the HPA axis is ameliorated by leptin treatment. Leptin also attenuates the ACTH and corticosterone responses to a physical stressor (hind-limb restraint) in rats. This suggests that leptin may in fact influence the neuroendocrine stress axis under conditions not associated with either nutritional or metabolic perturbations. The mechanisms by which leptin exerts such an effect are most likely to be centrally mediated. *In vitro* studies show that leptin can influence CRH secretion from isolated rat hypothalami but not ACTH secretion from cultured rat pituitary cells. A subpopulation of CRH-containing cells in the hypothalamus expresses the leptin receptor, allowing for direct action at this level of the HPA axis. Leptin receptors are also found in many cell types in the brain that provide afferents to the paraventricular nucleus, so there is likely to be relayed effects as well. Such cell types include those that produce NPY, MCH, and POMC.

Ghrelin

Ghrelin is primarily secreted from the stomach and has been shown to potently stimulate the secretion of

Table 3 The effect of hypothalamic appetite-regulating peptides on the HPA axis^a

	Neuropeptide	Effect on food intake	Effect on HPA axis activity
Arcuate nucleus	NPY	↑	↑
	AGRP	↑	↑
	α -MSH	↓	↑ Basal ACTH and corticosterone ↓ IL-1-induced HPA axis activity
Paraventricular nucleus	CART	↓	↑
	Urocortin 2	↓	No effect
	CRH	↓	↑
	Urocortin 1	↓	↑
Lateral hypothalamic area	Urocortin 2	↓	No effect
	Urocortin 3	↓	No effect
	MCH	↑	↑
	Orexins	↑	↑
	Urocortin 1	↓	↑

^aACTH, adrenocorticotrophic hormone; AGRP, agouti-related protein; CART, cocaine and amphetamine regulated transcript; CRH, corticotropin releasing hormone; HPA, hypothalamic-pituitary-adrenal; IL, interleukin; MCH, melanin-concentrating hormone; MSH, melanocyte stimulating hormone; NPY, neuropeptide Y.

growth hormone. In some species, ghrelin also acts within the brain to stimulate food intake. Studies in sheep show that a preprandial rise in ghrelin levels stimulates a postprandial rise in growth hormone levels. Plasma levels of ghrelin fall postprandially and are increased by fasting or reduction in body weight; levels are lower in obese individuals. In addition to the stimulation of growth hormone, ghrelin activates the HPA axis, although not in all species. Ghrelin mRNA levels are increased after both physical (tail pinch) and nutritional (fasting) stress. Thus, ghrelin may be an additional factor involved in the regulation of the HPA axis under stressful stimuli. Ghrelin acts via the growth hormone secretagogue receptor, which is present in both the central nervous system and the pituitary gland. In rodents, the *in vivo* action of ghrelin on the HPA axis is exerted centrally, via cells such as those that produce NPY and AGRP in the arcuate nucleus. The orexin system in the lateral hypothalamus has also been linked to ghrelin-induced corticosterone secretion.

Insulin

The role of insulin in the regulation of food intake and neuroendocrine function is sometimes confused with the confounding effects of insulin-induced hypoglycemia. The latter increases food intake due to glucose deprivation, whereas in other circumstances insulin acts to decrease food intake. In rodents, diabetes causes increased basal concentrations of ACTH and corticosterone but reduced responsiveness to stress. Streptozotocin-induced diabetes impacts various neuronal systems, which in turn may influence the HPA axis. For example, streptozotocin-induced diabetes increases the concentration of norepinephrine within the paraventricular nucleus of the hypothalamus, and this parallels increasing levels of corticosterone. This effect is dependent on insulin concentrations because insulin replacement restores levels of norepinephrine within the paraventricular nucleus and the peripheral concentrations of corticosterone. Interestingly leptin treatment, in the absence of insulin

replacement, also corrects these streptozotocin-induced perturbations. In addition to effects on catecholaminergic systems, diabetes-associated changes in HPA axis function may be mediated via hypothalamic peptides such as NPY. In the rat, the arcuate nucleus contains insulin receptors and insulin decreases NPY mRNA expression in this region.

Further Reading

- Bronson, F. H. (1985). Mammalian reproduction: an ecological perspective. *Biology of Reproduction* **32**, 1–26.
- Cameron, J. L. (1996). Regulation of the reproductive hormone secretion in primates by short-term changes in nutrition. *Reviews of Reproduction* **1**, 117–126.
- Dallman, M. F., Pecoraro, N., Akana, S. F., et al. (2003). Chronic stress and obesity: a new view of “comfort food.” *Proceedings of the National Academy of Sciences USA* **100**, 11695–11701.
- Friedman, J. M. and Halaas, J. L. (1998). Leptin and the regulation of body weight in mammals. *Nature* **395**, 763–770.
- Henry, B. A. (2003). Links between appetite regulating systems and the neuroendocrine hypothalamus: lessons from the sheep. *Journal of Neuroendocrinology* **15**, 697–709.
- Horvath, T. L. and Diano, S. (2004). The floating blueprint of hypothalamic feeding circuits. *Nature Reviews Neuroscience* **5**, 662–667.
- Krysiak, R., Obuchowicz, E. and Herman, Z. S. (1999). Interactions between the neuropeptide Y system and the hypothalamic-pituitary-adrenal axis. *European Journal of Endocrinology* **140**, 130–136.
- Seckl, J. R. and Walker, B. R. (2004). 11 β -hydroxysteroid dehydrogenase type 1 as a modulator of glucocorticoid action: from metabolism to memory. *Trends in Endocrinology and Metabolism* **15**, 418–424.
- Smith, M. S. and Grove, K. L. (2002). Integration of the regulation of reproduction function and energy balance: lactation as a model. *Frontiers in Neuroendocrinology* **23**, 225–256.
- van der Lely, A. J., Tschop, M., Heiman, M. L., et al. (2004). Biological, physiological, pathophysiological and pharmacological aspects of ghrelin. *Endocrine Reviews* **25**, 426–457.

Food Shift Effect

F K Stephan

Florida State University, Tallahassee, FL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 167–169, © 2000, Elsevier Inc.

Endogenous Circadian Clocks
Physiological Correlates of Entrainment by Meals
Shift Work and Jet Lag
Conclusion

Glossary

<i>Circadian rhythm</i>	A cycle with a period of approximately 24 h.
<i>Entrainment</i>	A zeitgeber forcing its period and phase on a rhythm.
<i>Phase</i>	A point in the circadian cycle.
<i>Zeitgeber</i>	A periodic external cue that entrains circadian clocks.

Endogenous Circadian Clocks

In 1972, the neural substrate responsible for the generation and entrainment of circadian rhythms by light was identified. Lesions of this small structure, the suprachiasmatic nucleus (SCN) of the hypothalamus, appeared to abolish circadian rhythms in behavior and physiology. However, despite such lesions, rats were still able to anticipate one meal per day by increased activity, a rise in core temperature and serum corticosterone, and other physiological changes that occurred approximately 2–4 h prior to meal access. There is now compelling evidence that these functions are mediated by an endogenous circadian clock: (1) meal anticipation occurs only if the interval between meals is in the circadian range (~22–30 h), (2) anticipatory activity persists for several days when rats are food-deprived or the activity free runs (in some rats) if meals are presented outside the circadian range, and (3) phase shifts of meal-time result in transients, that is, gradual daily adjustments until the activity reentrains to the new meal time (see [Figure 1](#)). Despite many efforts, this feeding-entrained oscillator (FEO) has not yet been localized. Nevertheless, most mammals, as well as some inframammalian species, are now known to have the capacity to anticipate daily meals, and it would be surprising if humans lacked such a circadian

clock. It is well known that hunger sensations arise around the customary meal time. What has been difficult to explain from a purely homeostatic view is that the omission of a meal does not produce monotonically increasing feelings of hunger; instead, hunger subsides until the next customary meal time. However, this makes sense if hunger sensations are based in part on an internal timing signal and not entirely on an energy-deficit signal.

Physiological Correlates of Entrainment by Meals

Before addressing the potentially adverse effects of changes in meal time or meal omission, a brief summary of the main effects of feeding one meal per day on the rat's digestive physiology is in order. As mentioned earlier, serum corticosterone levels rise prior to expected meals. Corticosterone is both a metabolic and a stress hormone, and the brain has receptors for this steroid. Perhaps most surprising is the observation that disaccharidases (enzymes for digesting sugars) in the duodenum also increase before the meal is available. In addition, several aspects of liver function, cell division in the gastrointestinal tract, and intestinal motility become entrained to periodic meals. The increase in core temperature also may facilitate digestion. The general inference to be drawn from these observations is that the digestive physiology is not simply reactive to ingested nutrients but is capable of anticipating and preparing for meals, most likely under the influence of the circadian FEO. If the expected meal is displaced in time, these efforts are for naught and for several days these preparatory responses are out of synchrony with actual food consumption.

Shift Work and Jet Lag

Gastrointestinal malaise is a common complaint of shift workers and travelers across time zones. Along with impaired performance and other complaints, this has been attributed to the desynchronization of SCN-controlled rhythms (e.g., the sleep–wake cycle) from behavior. The discovery of a separate circadian clock raises the possibility that some of these symptoms are caused by forcing the FEO to phase shift to a new time. Unfortunately, the scientific literature regarding the role of meal timing and meal

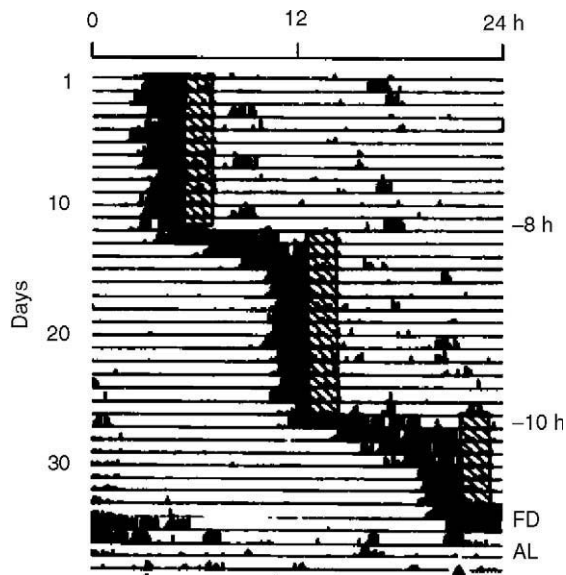


Figure 1 Anticipatory wheel running from a rat with a lesion of the suprachiasmatic nucleus (arrhythmic when fed *ad libitum*). Food is available for 2 h day⁻¹, indicated by the vertical cross-hatched bars. Most wheel running occurs approximately 3 h before meal time. When the meal time is delayed by 8 or 10 h, anticipatory activity takes 3 or 4 days to reentrain to the phase-shifted meal. The activity persists during food deprivation (FD), but is not observed during *ad libitum* (AL) feeding.

composition in producing jet-lag symptoms in humans is extremely sparse. Only very recently has an attempt been made to assess the consequences of limiting food intake to a single meal. A carbohydrate-rich meal in the morning for 3 consecutive days led to a phase advance of 1 h of the body temperature rhythm, whereas evening meals had no effect. A similar effect has been noted in mice in that the nocturnal activity rhythm is also phase advanced if they are limited to one meal per day. In rats, the FEO is easily phase shifted by a glucose meal, but seems less responsive to a fat meal. A number of other studies indicate that there are interactions between light-entrainable and food-entrainable circadian systems; the period and phase of one has an effect on the other. Thus, there is a possibility that particular meal schedules might facilitate adjustment to temporal shifts in daily routines, but this has not yet been specifically investigated. In an interesting study, patients in a vegetative state (which reduces or eliminates the effects of activity or other external factors) received intraduodenal infusions of a liquid diet continuously, only during the day or only during the night. No consistent temperature rhythms were observed with continuous infusions, whereas day infusions resulted in a peak temperature at 8 p.m. and night infusions resulted in a peak at 4 a.m. Because these peaks

occurred at different times relative to meal onset and end, they were probably not entirely due to the thermogenic action of food but may reflect the entrainment of some physiological processes by a circadian oscillator. When humans live in temporal isolation for more than 3 weeks, approximately one-half of the subjects display internal desynchronization. The core body temperature rhythm continues to free run with a period of approximately 25 h, but the sleep-wake cycle assumes a much longer period, between 30 and 48 h. When subjects were asked to eat only two meals per day (breakfast and lunch), the interval between meals was proportional to the wake time. Before internal desynchronization, the average interval between meals was approximately 7.4 h; after internal desynchronization (when the sleep-wake cycle had assumed a period of 35.4 h), the interval increased to 10.4 h. It should be noted that subjects were entirely unaware that their sleep-wake cycle had become abnormally long. It has long been suspected that internal desynchronization is evidence for two separate circadian clocks, but alternative explanations have also been offered. In particular, none of the studies establishes the existence of a separate feeding-entrained circadian clock in humans. Nevertheless, these studies at least suggest a circadian timing component to human hunger. However, in blind humans who have difficulties adhering to a 24-h schedule, regular meals and social routines have not been very effective in forcing entrainment of the sleep-wake cycle.

In other human studies regarding the effect of light or exercise on circadian rhythms, mealtime is either not reported or small meals are given at frequent intervals so that the effects of substantial meals on physiological measures cannot be assessed. Furthermore, temperature data (and other data) are often reported as group means or in other ways that make it impossible to determine even whether temperature or serum cortisol levels in humans rise in anticipation of meals. However, the evidence is quite compelling that such rhythms persist in constant conditions, that is, that they are controlled by one or more endogenous circadian oscillators.

A nearly universal observation is that circadian rhythms adjust more readily to phase delays (e.g., westward flights) than to phase advances (e.g., eastward flights). This asymmetry is even more pronounced in rats with SCN lesions that are entrained to one meal per day. As mentioned earlier, rats can anticipate meals at intervals as long as 30 h, but have difficulty anticipating meals presented at intervals shorter than 22 h, indicating that the FEO slows down more readily than it accelerates. When meals

are shifted to a later time, anticipatory activity shifts by $2\text{--}4\text{ h day}^{-1}$ until it reaches the new phase position (see [Figure 1](#)). Conversely, after phase advances of 6 h or more, many rats display delays in anticipatory activity, again of about 3 h day^{-1} , i.e., the oscillator delays by 16–18 h over 5–6 days to catch up with the phase-advanced meal time. It is interesting that some humans also show delaying transients after rotating to an earlier shift (e.g., from the night to afternoon shift).

Conclusion

During the past decades, it has become increasingly clear that the regulation of energy balance is largely proactive rather than reactive. Circadian clocks, such as the FEO, entrain to predictable external events and ready the system for the arrival of nutrients. Phase shifts of mealtime temporarily disrupt these functions and probably account, at least partially, for the symptoms of jet lag or shift-worker malaise. The release of adrenal corticosteroids before meals could be a warning signal to the brain to replenish energy reserves. If food is consumed, serum levels decline, but if the meal is not available, serum levels might remain elevated for many hours. This, as well as other wasted physiological changes that occur in anticipation of meals, is undoubtedly not beneficial. There is no direct evidence as yet (because it has not been studied) that serious harm results from phase shifts of meals; however, it is conceivable that prolonged erratic eating habits contribute to the stress experienced by hectic modern lifestyles.

See Also the Following Articles

Circadian Rhythms, Genetics of; Sleep Loss, Jet Lag, and Shift Work.

Further Reading

- Aschoff, J. (1998). Human perception of short and long time intervals: its correlation with body temperature and the duration of wake time. *Journal of Biological Rhythms* 13, 437–442.
- Czeisler, C. A. and Megan, E. J. (1990). Human circadian physiology: interaction of the behavioral rest-activity cycle with the output of the endogenous circadian pacemaker. In: Thorpy, M. J. (ed.) *Handbook of sleep disorders*, pp. 117–137. New York: Marcel Dekker.
- Krauchi, K., Werth, E., Cajochen, C., et al. (1998). Are carbohydrate-rich meals a zeitgeber in humans? *Society for Research on Biological Rhythms Abstract* 96.
- Mills, J. N., Minors, D. S. and Waterhouse, J. M. (1978). Adaptation to abrupt time shifts of the oscillator(s) controlling human circadian rhythms. *Journal of Physiology* 285, 455–470.
- Mistlberger, R. E. (1994). Circadian food-anticipatory activity: formal models and physiological mechanisms. *Neuroscience and Biobehavioral Reviews* 18, 171–195.
- Stephan, F. K. (1992). Resetting of a feeding-entrainable circadian clock in the rat. *Physiology & Behavior* 52, 985–995.
- Stephan, F. K. (1997). Calories affect zeitgeber properties of the feeding entrained circadian oscillator. *Physiology & Behavior* 62, 995–1002.
- Van Cauter, E., Sturis, J., Byrne, M., et al. (1994). Demonstration of rapid light-induced advances and delays of the human circadian clock using hormonal phase markers. *American Journal of Psychiatry* 266, E953–E963.

Freud, Sigmund

D J Lynn

Thomas Jefferson University, Philadelphia,
PA, USA

© 2007 Published by Elsevier Inc.

Introduction

Background and Personal Life

Freud the Theoretician and Proponent

Freud the Clinician

Freud the Organizer and Movement Leader

Freud and Stress

Freud in Perspective

Introduction

Sigmund Freud (1856–1939) is best known as the originator of psychoanalysis, which is at once a method of treatment for emotional disorders, a system of ideas concerning human emotional life (both healthy and disordered), and an intellectual movement. His historical significance stems from his activities in several frequently interlocking roles – theoretician,

proponent, clinician, organizer, and leader. Although practitioners of psychoanalysis have divided themselves into diverging groups adhering to different doctrines, many psychoanalysts continue to rely on Freud's theoretical conclusions and technical recommendations, and his writings are still cited in current American textbooks of psychiatry. In addition, some writers in the humanities, literary critics and historians in particular, have relied extensively on Freud's theories. Indeed, Freud's ideas, or corruptions of them, have become commonplace in Western popular culture.

Background and Personal Life

Sigmund Freud was the oldest of eight children born to a Jewish family in what was then Freiberg, in the Austro-Hungarian Empire, and is now Příbor, in the Czech Republic. His father, Jakob Freud, was a wool merchant; during hard times he moved the family to Vienna when Sigmund was 4 years old. A strong student in gymnasium and then at the University of Vienna, Freud had adolescent ambitions of becoming a leader in progressive politics but put them aside to take up the natural sciences. The geographical and social context was the Austro-Hungarian Empire, which was undergoing very rapid industrialization and modernization. Much has been written on the question of how much influence Freud's Jewish ancestry had upon his experiences, attitudes, and beliefs.

Freud's first scientific mentor was Ernst Wilhelm von Brücke, and his early laboratory studies involved comparative neuroanatomy. He earned his medical degree in 1881, but he remained a laboratory researcher for a time. He did not engage in clinical work until 1882, after he had met Martha Bernays, his future wife, and had decided that his earnings in research would never support married life. His work in neuroanatomy reached its peak in his microscopic studies of the medulla in the laboratory of Theodor Meynert. His papers on axon tracts and on his method for using gold solutions for staining them were published in both German and English in 1884, but disappointment followed when his staining method could not be replicated. Freud chose the clinical specialty of neurology – not psychiatry, which in the late nineteenth century was largely concerned with treating and studying psychotic patients in mental hospitals. As a neurologist in office practice, Freud encountered patients with symptoms that were then classified into such syndromic categories as neurasthenia and hysteria. Freud's neurological studies included important work on aphasia and paralyzes in children.

From October 1885 to February 1886, Freud studied in Paris with Jean-Martin Charcot, the premier

French neurologist of the era. Charcot was then immersed in the study of hysteria, using hypnosis to explore patients' histories and concluding that the paralyzes and seizures that he saw in his patients, and demonstrated so dramatically in his clinical presentations, had their roots in ideas. Freud's interest in these clinical problems had been stimulated by the descriptions given by his friend and senior colleague, Josef Breuer, of Breuer's treatment of the woman known in later publications as Anna O.

Freud married Martha Bernays in 1886. The couple had three sons and three daughters. In 1923, he was first given a diagnosis of oral carcinoma. This neoplasm, related to his constant cigar smoking, required several surgical procedures and produced intermittent pain over the ensuing years. Freud lived and worked in Vienna until the Nazi Anschluss in March of 1938 placed him and his family in great peril. Through the intercession of Ernest Jones and Freud's patients Princess Marie Bonaparte and William Bullitt, and diplomatic pressure by the United States, Freud was allowed to leave Austria with his wife and children in June. A condition imposed by the Germans was that Freud state that he had been treated with due respect. In response, Freud is reported to have declared "I can heartily recommend the Gestapo to anyone" – a fine example of his celebrated sense of humor. Freud spent the last year of his life in London, seeing some patients and continuing to write, despite his advanced age, until July 1939. Pain and debilitation caused by his spreading oral carcinoma then became overwhelming. On September 23, 1939, he died in what would today be called an assisted suicide.

Freud the Theoretician and Proponent

Freud generated theoretical proposals and conclusions that spanned a breathtaking range of subjects, including the technique of clinical psychoanalysis, the etiology of mental disturbances, human psychological functioning in health and disease, the pattern of development of the human personality, and the basic forces motivating human life.

Concerning the technique of clinical psychoanalysis, Freud prescribed an approach based upon anonymity, neutrality, and confidentiality and requiring from the patient free associations, meaning the vocalization of an uncensored flow of thoughts. Freud emphasized the importance of interpretations expressed by the analyst and giving meaning to dreams, feelings, and symptoms experienced by the patient. Invariably, patients encountered difficulties in reporting, or even experiencing, uncensored thoughts. Freud reported that these difficulties, termed resistances, could

themselves be understood and thus could serve to promote the exploratory process. The transference, the set of emotions experienced by the patient toward the analyst, initially seen as an obstacle to exploration, came to be seen as a crucial source of information for understanding the patient's development and reconstructing childhood relationships. The countertransference, defined as the psychoanalyst's attitudes and feelings toward the patient, subsequently received attention; Freud eventually prescribed that the analyst should undergo psychoanalysis in order to minimize its impact on the treatment.

Early in his psychoanalytic work, in 1896, Freud published his conclusion that the etiology of hysteria – the intermittent paralyses and other unexplained symptoms studied by Charcot and others – was invariably a delayed response by the patient to childhood experiences of seduction (sexual abuse or rape in today's terms) of the child by a trusted adult, typically the patient's father. Freud retracted this view within a few years and emphasized instead the significance of frustrated childhood wishes for sexual gratification by an adult.

Freud's initial psychoanalytic view of the functioning of the human mind involved what has been called a topographical model. In this conception, ideas, feelings, and wishes are processed in three different mental sectors. In the conscious sector, items that are not overly threatening could be remembered, considered, and acted upon. A preconscious sector would contain items that are not currently in awareness but that could be readily retrieved. An unconscious sector would contain extremely important memories, wishes, and feelings that would be too threatening to be recognized in consciousness. A postulated mental component known as the censor would perform the function of allowing, or preventing (through repression), items from reaching consciousness. Despite the functioning of the censor, unconscious ideation would be expressed in several interesting ways, through symptoms and symptomatic behavior, for example, or in dreams or parapraxes (slips of the tongue). In this view, an important function of psychoanalysis would be to strengthen the personality and resolve symptoms by interpreting these expressions and thereby making the unconscious conscious.

In later writings, Freud proposed a different model of mental functioning, although he never fully abandoned the topographic model. This tripartite model postulated three psychic structures. The id was described as a biologically rooted source of primitive urges. The superego, a rather rigid accumulation of prohibitions and rules imposed by parental and societal authorities, was thought to be formed in the resolution of the Oedipus complex, an emotional

situation culminating in the fourth or fifth year of life, in which the boy experiences wishes to possess his mother (emotionally and sexually) and conflicting fears of paternal retaliation and general disapproval. The ego was described as the component of the mind that can think, perceive, adapt, learn, compromise, and create; this component faces the task of reconciling demands from the other components and from the external world and enabling the person to conduct a life. In this view of mental functioning, important features of emotional life, including symptoms, choices, capabilities, and limitations, were seen as arising from conflicts between the psychic components. The task of psychoanalysis in this view would be to strengthen the ego through the development of insight into these intrapsychic conflicts.

A persistent feature of Freud's theoretical views throughout his psychoanalytic writing was the notion that no mental life would occur without energy that would have biological drives or instincts as its ultimate sources. Initially, libido or eros, a sexual drive, manifesting itself in different transformations, was seen as the sole underlying source. Later, an aggressive or destructive drive, a death instinct, was postulated in addition. This latter drive remained controversial, neither universally accepted by Freud's followers nor consistently invoked by Freud himself.

Freud had superb talents as a writer. In general, his psychoanalytic contributions were in the form of essays that were not constrained by a conventional scientific format. At times he could deploy a formidable array of rhetorical devices to convince the reader that his points were valid or even inescapable. In 1930, the city of Frankfurt honored Freud's creativity and literary accomplishments with the Goethe Prize. All of his psychoanalytic writings, with the exception of *Thomas Woodrow Wilson: A Psychological Study*, co-authored by his patient, William C. Bullitt, have been translated into English in the *Standard Edition of the Complete Psychological Works of Sigmund Freud*.

Freud the Clinician

In his earliest years as a physician and neurologist, Freud had used the standard methods of the era: a limited array of drugs, electrical currents (Faradic and Galvanic treatments), minor surgery, and hypnosis. By 1896 he was concentrating on what he called psychoanalysis. In the beginning, his analytic treatments were brief and tended to culminate in a single interpretation, but by 1896 he had begun to see at least a few patients in treatments that would last for years.

His clinical work was important for several reasons. First, since he did not directly observe the

psychoanalytic work of others, his own clinical work gave him his only opportunities to validate technical recommendations and to observe the clinical phenomena arising in psychoanalysis. Second, much of his analytic activity, especially after World War I, involved didactic psychoanalyses of trainees, and many of these trainees became productive, and useful, subordinates in the psychoanalytic movement. Third, even many of his patients who were not trainees and who never practiced psychoanalysis were prominent and wealthy people who retained a connection to Freud and his movement and proved to be helpful financially and in other ways.

Some of Freud's published discussions of individual cases, such as Schreber and Little Hans, do not refer to direct clinical psychoanalytic work conducted by Freud. His writings do contain extensive discussions of three of his own clinical cases. In the case of Dora, a young woman Freud treated in 1900, Freud reached a formulation involving an interpretation of her experience that was at variance with her own view. Freud concluded that the sexual advances from an older married man, which she had described as disturbing and unwanted, had been disturbing precisely because she had desired them, unconsciously. When Freud confronted her with this interpretation, she bolted in anger and terminated the treatment. He subsequently concluded that he had pressed this interpretation on her prematurely, but of course its validity, as well as its timing, is open to question.

Freud described the case of the Rat Man, a young army officer he treated in 1907, at length; in addition, his actual contemporaneous notes of the conduct of this analysis were discovered and published. Freud concluded that this patient suffered from an obsessional neurosis that Freud described as essentially cured by his interpretations. This case is usually seen as the most clinically successful of the ones that he described in his published works. Unfortunately, this man, later identified as Ernst Lanzer, died in 1914 at the onset of World War I, so that little longitudinal follow-up information is available.

The Wolf Man (later identified as Serge Pankejeff), a young Russian nobleman treated by Freud in 1909–1914 and later in 1919–1920, was Freud's best-studied case. Freud's published discussion focused on the Wolf Man's infantile neurosis, which Freud attempted to reconstruct. Central to Freud's formulation was his interpretation of the patient's wolf dream as symbolically referring to and describing a scene from this patient's infancy. This patient survived into the 1970s and was interviewed later in his life by two authors. Muriel Gardiner, a practicing psychoanalyst, described him as largely healed and contented. Karin Obholzer, a journalist, reported two

important findings. First, Pankejeff told her that he had received little or no clinical benefit from his psychoanalysis by Freud. Second, Pankejeff had never accepted Freud's interpretation of the wolf dream; he regarded Freud's reconstruction of a scene from his infancy as totally implausible.

Increasing quantities of information about many of Freud's psychoanalyses have gradually become available over the past half century in the form of Freud's letters, analysts' letters, memoirs and reports by analysts, and other sources. Recent revelations concerning his psychoanalytic treatments of some of his patients, such as Horace Frink, as described by Edmunds, or his daughter Anna, as described by Young-Bruehl, have sparked controversy. In a study of 43 of the cases he treated after 1907, Lynn and Vaillant found that in a majority of these cases, he deviated from the anonymity, neutrality, and confidentiality that he prescribed in his recommendations.

Freud the Organizer and Movement Leader

The development and maintenance of an organized group of psychoanalytic followers, forming a movement, always commanded Freud's attention, as shown especially in his correspondence with Ernest Jones and with Sandor Ferenczi. Freud was equally concerned with his followers' professional success, achievement of recognition, capacity to make intellectual contributions as members of his movement, and loyalty. Freud expressed his view of the optimal organization and functioning of the psychoanalytic movement in his essay *On the History of the Psycho-Analytic Movement*, first published in 1914. He envisioned a single leader with overwhelming authority. As the first leader, he wanted to find a successor who would be both capable and reliable. These considerations partly explain the pattern of his relations with a series of prominent male psychoanalysts. After apparently close collaboration, including personal friendship and mutual trust, there would be theoretical disagreement, a mutual sense of injury, competition for prominence, increasing estrangement, and repudiation. This series clearly included Carl Jung, Alfred Adler, Otto Rank, and (except for the overt repudiation), Sandor Ferenczi.

In 1912 Freud's most trusted followers, Sandor Ferenczi, Otto Rank, Karl Abraham, Ernest Jones, and Hans Sachs, formed a secret committee to oversee the movement, respond to criticisms, and assure Freud's continuing control. In his role of leader and organizer, Freud needed to reward loyalty and reliability rather than originality, creativity, or independent intellectual achievement. He generally saw

new members as converts; he referred to critics as opponents or even enemies.

Freud and Stress

Freud viewed anxiety as a signal that could refer to a number of different dangers. He was especially concerned with anxiety arising from intrapsychic threats, such as the impending emergence of repressed wishes or conflicts. In many of these situations, even the anxiety itself would be unconscious; its existence could only be inferred or deduced by the psychoanalyst. Castration anxiety, anxieties concerning unconscious homosexual wishes, and anxieties concerning unconscious incestuous wishes were frequently invoked in the formulation and interpretation of patients' experiences and symptoms. Freud's work did not direct attention to the biological aspects of states of anxiety, nor to the role played by overwhelming external stressors. Even in his brief discussions of war neuroses, Freud emphasized intrapsychic conflict as the essential element that either made these neuroses possible or promoted them. Freud's influence was reflected in the official diagnostic terminology of American psychiatry codified in the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-I) (1952) and retained in the DSM-II (1968). These listings explicitly described anxiety as an intrapsychic signal, included anxiety disorders as neuroses, and excluded any anxiety disorder specifically related to an external stressor (such as the current posttraumatic stress disorder).

Freud in Perspective

Sigmund Freud must be seen as one of the most influential intellectuals of the twentieth century. There is no indication that the influence of his system of ideas in such fields as literary criticism and historical analysis, or in the general popular culture, will recede.

Within the medical profession, Freud's ideas gained their greatest influence in American psychiatry in the middle of the twentieth century. As late as the 1970s, American psychiatry was divided into schools of thought, each of which had its own journals and took its own approach to the definition of problems to be studied, the terminology to be used, the clinical methods to be advocated, and the basic etiological assumptions to be adopted. An adherent of a particular school often felt no obligation to read or consider contributions from other schools. In this context, the extent of the influence that Freud's system of ideas

had upon psychiatrists was very uneven. At one extreme, for some psychiatrists in some departments, Freud's conclusions were seen as scientifically established discoveries that served as the basis for any treatment or any academic discussion. In other departments, Freud's conclusions were either explicitly rejected or, more frequently, ignored. In recent years, American psychiatry has taken important steps toward a systematically scientific approach to clinical issues. In general, training programs no longer explicitly adopt locally accepted assumptions, and advocates of any treatment approach are expected to produce data supporting their claims.

In this current environment, Freud's conclusions will retain influence to the extent that they can be validated with reproducible results. (Advocates of psychoanalysis cannot, in this context, adequately deal with skeptics by simply maintaining that nobody has conclusively demonstrated that the method does not work.) Freud's own psychoanalytic publications have shortcomings that limit the extent to which they can be used as validation: a lack of raw observational data, an inadequate description of Freud's actual methods, and a commingling of postulates, hypotheses, observations, conclusions, and speculations without adequate designation. The psychoanalytic situation contains unavoidable difficulties for any clinical researcher in the collection and verification of data, in the generation of adequate numbers of cases, and in the standardization of technique between analysts. Efforts by latter-day psychoanalysts to validate conclusions through systematic outcome research have not produced any conclusive results, and it is unclear how, and by whom, such studies would be carried out in the future.

See Also the Following Articles

Child Sexual Abuse; Emotional Inhibition; Hysteria; Psychoanalysis; Psychotherapy.

Further Reading

- Brabant, E., Falzeder, E. and Giampieri-Deutsch, P. (eds.) (1993). *The correspondence of Sigmund Freud and Sandor Ferenczi, vol. 1, 1908–1914*. Cambridge, MA: Belknap Press of Harvard University Press.
- Brabant, E., Falzeder, E. and Giampieri-Deutsch, P. (eds.) (1996). *The correspondence of Sigmund Freud and Sandor Ferenczi, vol. 2, 1914–1919*. Cambridge, MA: Belknap Press of Harvard University Press.
- Edmunds, L. (1988). His master's choice. *Johns Hopkins Magazine* 40, 40–49.
- Ferris, P. (1997). *Dr. Freud, a life*. Washington, D.C.: Counterpoint.

- Freud, S. (1957). *The standard edition of the complete psychological works of Sigmund Freud*. Strachey, J. (ed.). London: Hogarth Press.
- Freud, S. and Bullitt, W. C. (1966). *Thomas Woodrow Wilson: a psychological study*. Boston, MA: Houghton Mifflin.
- Gardiner, M. (1971). *The wolf-man by the wolf-man*. New York: Basic Books.
- Gay, P. (1988). *Freud: a life for our time*. New York: Norton.
- Gedo, J. (2002). The enduring scientific contributions of Sigmund Freud. *Perspectives in Biology and Medicine* 45, 200–211.
- Grosskurth, P. (1991). *The secret ring: Freud's inner circle and the politics of psychoanalysis*. Reading, MA: Addison-Wesley Publishing Co.
- Jones, E. (1953–57). *The life and work of Sigmund Freud*. (vols. 1–3). New York: Basic Books.
- Lynn, D. J. and Vaillant, G. E. (1998). Anonymity, neutrality, and confidentiality in the actual methods of Sigmund Freud: a review of 43 cases, 1907–1939. *American Journal of Psychiatry* 155, 163–171.
- Mahony, P. J. (1986). *Freud and the rat man*. New Haven, CT: Yale University Press.
- Obholzer, K. (1982). *The wolf man – sixty years later*. Shaw, M. (trans.). New York: Continuum Publishing Co.
- Paskauskas, R. A. (ed.) (1993). *The complete correspondence of Sigmund Freud and Ernest Jones 1908–1939*. Cambridge, MA: Belknap Press of Harvard University Press.
- Roazen, P. (1976). *Freud and his followers*. New York: Knopf.
- Shorter, E. (1997). *A history of psychiatry*. New York: John Wiley and Sons.
- Westen, D. (1999). The scientific status of unconscious processes: is Freud really dead? *Journal of the American Psychoanalytic Association* 47, 1061–1106.
- Young-Bruehl, E. (1988). *Anna Freud: a biography*. New York: Summit Books.



GABA (Gamma Aminobutyric Acid)

J D C Lambert

University of Aarhus, Aarhus, Denmark

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 177–190, © 2000, Elsevier Inc.

Introduction to the Role of GABA as a Neurotransmitter
Neuronal Receptors for GABA
GABAergic Neurons and Their Role in Neuronal Activity

Glossary

<i>Amino acid</i>	A small organic molecule consisting of a backbone to which is attached an amino group ($-\text{NH}_2$) and an acidic group ($-\text{COOH}$). Twenty naturally occurring amino acids are incorporated into proteins under direct genetic control.
<i>Benzo-diazepine</i>	The generic name for a class of chemical substances that are in wide clinical use for their anxiolytic, sedative, and muscle-relaxant properties. These actions are produced by enhancing inhibitory transmission at synapses that use GABA as a transmitter.
<i>Subunit</i>	The receptors for many neurotransmitters are composed of a number of (usually five) distinct proteins encoded by separate genes. When the receptor is assembled from identical subunits encoded by a single gene, the assembly is termed homomeric; when subunits encoded by different genes are incorporated, the assembly is termed heteromeric.
<i>Synapse</i>	The very small contact area between two nerve cells where chemical transmission occurs. The chemical messenger (the neurotransmitter) is stored in vesicles in the presynaptic terminal. Nervous activity causes these vesicles to release the neurotransmitter into the synaptic cleft, following which it activates specific

receptors on the postsynaptic membrane. These transduce the chemical message into an electrical signal, which can be either excitatory or inhibitory.

Gamma aminobutyric acid (GABA) is the principal inhibitory transmitter in the central nervous system (CNS). GABA acts on three distinct receptor complexes: GABA_A , GABA_B , and GABA_C receptors. GABA_A receptors are heteromeric structures assembled from 5 subunits selected from up to 16 different glycoprotein subunits. This heterogeneity confers discrete pharmacological and physiological properties on the assembled complex. The five subunits surround a central ion-selective channel, which conducts chloride ions into the neuron when GABA binds to the receptor. This inhibits the neuron by causing an increase in membrane potential and conductance. The GABA_A receptor complex also contains a number of distinct modulatory sites that are the target for many neuroactive and therapeutic drugs, including benzodiazepines, barbiturates, steroids, alcohol, and a variety of general anesthetics. The activation of these modulatory sites generally enhances neuronal inhibition at GABA_A synapses, which probably accounts for the anxiolytic, sedative, hypnotic, and anesthetic properties of these substances.

GABA_B receptors are formed by a single protein molecule that is coupled with intracellular enzymes and effector systems. These can activate K^+ -selective channels, which inhibits the neuron, or depress the activity of Ca^{2+} -selective channels, which decreases transmitter release from presynaptic terminals. GABA_C receptors have many structural and functional similarities to GABA_A receptors, but they exhibit sufficient pharmacological and mechanistic differences to merit classification as a separate receptor.

In view of its ubiquitous role in the integration of neuronal activity, it is hardly surprising that dysfunction of GABAergic transmission has been implicated in a wide variety of neurological and psychiatric

disorders, including epilepsy, anxiety states, alcoholism, nociception, Huntington's chorea, and other movement disorders.

Introduction to the Role of GABA as a Neurotransmitter

GABA was already shown to be present in the CNS in 1950. Using the technique of microiontophoresis to apply pharmacologically active substances to single

neurons in the CNS, it was subsequently shown that GABA caused a powerful inhibition of virtually every neuron to which it was applied. This, along with its simple structure and ubiquitous occurrence, meant that there was initially great skepticism as to whether GABA could fulfill the criteria required for a central inhibitory transmitter (see [Box 1](#)). These doubts were allayed by the use of two naturally occurring substances that potently and selectively block the actions of exogenously applied GABA, as well as synaptic

Box 1

Classification of GABA Receptors

A single neurotransmitter usually activates a number of distinct postsynaptic receptors. These are distinguished and identified on the basis of their responses to the actions of agonists and antagonists. Agonists are molecules that bear a chemical resemblance to the natural transmitter and, on binding to the receptors, are able to operate the effector system and thereby mimic the response to the transmitter itself. As the name implies, antagonists prevent the agonist from evoking a response. There are two main classes of antagonist: competitive and noncompetitive. Competitive antagonists bind to the receptor site, but are unable to operate the effector system. However, they impede access of the agonist (including the natural transmitter) to the receptor site and thereby block the response. Bicuculline is an example of a competitive antagonist at GABA_A receptors, and CGP 55845A competitively antagonizes responses at GABA_B receptors. Noncompetitive antagonists do not interact with the receptor itself but block the action of the effector mechanism. Picrotoxin, for example, binds

within the ionophore, which is gated by the GABA_A receptor, and thereby inhibits the response to GABA.

On the basis of such pharmacological studies, three main families of postsynaptic receptors for GABA have been identified. These have simply been named alphabetically in the chronological order of their discovery: GABA_A, GABA_B, and GABA_C receptors. The table shows the structural, pharmacological, and mechanistic features of these three receptors. GABA_A and GABA_C receptors are both pentameric structures surrounding a Cl[−]-selective ionophore. GABA_C receptors have been accorded their own classification based primarily on their lack of sensitivity to bicuculline and lack of modulatory sites. GABA_B receptors are coupled to intracellular enzymes and other mechanisms that mediate the response. Although GABA_A and GABA_B receptors obviously share the same natural transmitter, they otherwise have very little in common, either with respect to their molecular structure, pharmacology, or mechanism of action.

Specific properties of GABA receptors

	GABA _A	GABA _B	GABA _C
Receptor mechanism	Cl [−] -selective ionophore	G-protein-coupled metabotropic receptor	Cl [−] -selective ionophore
Receptor composition	Pentamer constructed from α_{1-6} , β_{1-4} , γ_{1-4} , δ_1 , and ϵ_1	Dimer constructed from GABA _B R1 + GABA _B R2	Pentamer constructed from ρ_{1-2}
Single channel conductance	≈30 pS	—	≈8 pS
Mean channel open time	≈25 ms	—	≈150 ms
Receptor desensitization	Yes	No	No
Potency of GABA (EC ₅₀) ^a	≈10 μ M	≈10 μ M	≈1 μ M
Selective agonist	THIP, ^b P4S ^c	Baclofen	CACA ^d
Selective antagonist	Bicuculline	Saclofen, CGP 55845A	3-APA ^e

^aMedian effective concentration.

^b4,5,6,7-Tetrahydroisoxazole[4,5-c]pyridin-3-ol.

^cPiperidine-4-sulfonic acid.

^dThere are no potent selective agonists for GABA_C receptors, but cis-4-amino crotonic acid is a partial agonist.

^e3-Aminopropylphosphonic acid.

potentials evoked by the stimulation of inhibitory nerves. These substances are the alkaloids bicuculline and picrotoxin, which have played a pivotal role in past and present research on GABAergic inhibition. The ability of both alkaloids to cause sustained seizures and convulsions was known long before their actions on GABA receptors were resolved. However, this provided the first clue as to how manipulation of the GABAergic system could alter the state of excitability in the nervous system.

In 1980, a bicuculline-insensitive action of GABA was discovered that could be mimicked by the antispastic agent baclofen. A pharmacologically distinct receptor for GABA had been identified, and it was decided to name receptors alphabetically in chronological order of their discovery: bicuculline-sensitive receptors were termed GABA_A, and bicuculline-insensitive receptors were termed GABA_B.

Early studies of inhibition mediated by GABA_A receptors at various sites in the CNS had disclosed discrete differences in the responses, which implied the existence of different isoforms of the receptor. Following its cloning in the late 1980s, it became apparent that the GABA_A receptor is assembled from five subunits and that the identity of these determines the properties of the assembled complex. These studies also led to the identification of pharmacologically distinct GABA_C receptors. The classification of the GABA receptors is given in [Box 1](#).

Neuronal Receptors for GABA

GABA_A Receptors

The GABA_A receptor is a heterooligomeric complex that is probably assembled from five subunits. These are grouped around a central ion-conducting channel (ionophore), which is opened when two molecules of GABA (or another agonist) bind to the receptor. The subunits are encoded by at least 16 genes, which are grouped into five families: α_{1-6} , β_{1-4} , γ_{1-4} , δ_1 , and ϵ_1 . The degree of sequence homology of amino acids within each family is around 70%, whereas it is less than 30% between the families. Furthermore, the mature mRNAs for α_6 , β_2 , β_4 , and γ_2 can be alternatively spliced to include or omit amino acid sequences, yielding long (L) or short (S) versions of these subunits, respectively. The subunits, which are composed of around 500 amino acids (molecular mass ranging from 48 to 64 kDa), have a similar topographical structure with the following features:

1. A large extracellular N-terminal domain containing the binding site (receptor) for GABA and a variety of modulatory ligands. This region also contains two to four glycosylation sites, which

could be important for the intracellular sorting and assembly of proteins.

2. Four hydrophobic transmembrane domains (TM1–4), which are linked by polypeptide chains. TM2 contains clusters of basic amino acids, which are believed to line the channel wall and determine its selectivity to anions.
3. A large intracellular loop connecting TM3 and TM4 that in some subunits contains sites that can be phosphorylated. Whether phosphorylation increases or decreases receptor function depends on the enzyme involved.

If there were no restrictions on how the subunits could assemble, an immense number of isoforms would exist. However, there are certain constraints on which combinations can form functional complexes. Thus, the presence of an α subunit is essential for producing a functional channel, and the GABA_A receptor itself is probably located at the interface between two dissimilar subunits. The overall stoichiometry of the pentamer is likely to be two α , two β , and one γ subunits, which are arranged alternately ([Figure 1](#)). The α subunits need not be identical and the γ subunit can be replaced by a δ or an ϵ subunit. *In situ* hybridization and immunoprecipitation studies have shown that approximately 20 isoforms account for the vast majority of GABA_A receptors, which are expressed in the CNS in a structure and cell-specific manner ([Table 1](#)). The most abundant combination is $\alpha_1 \beta_2 \gamma_2$, which accounts for almost half the receptors and is expressed in most brain regions. Different subunit combinations are often expressed within the same neuron.

This repertoire of 20 isoforms is sufficient to confer substantial functional and pharmacological diversity on the GABA_A receptors. Examples of this diversity include the sensitivity to GABA, the kinetics of operation of the ionophore, and the actions of modulators. An interesting example of how these properties are exploited concerns the expression of GABA_A receptors during development of the CNS.

GABA_A receptors are expressed in the membrane some time before specialized synaptic contacts develop. The α subunit is selected from α_2 , α_3 , or α_5 , which form receptors having a relatively high affinity for GABA and respond to ambient levels of GABA in the extracellular fluid. The intracellular Cl[−] concentration is relatively high early in development: E_{Cl^-} is positive to the resting membrane potential and the activation of the GABA_A receptors evokes an excitatory response (see [Box 2](#)). This depolarization activates voltage-dependent Ca²⁺ channels and the influx of Ca²⁺ activates a host of intracellular processes, which contribute to the neurotrophic response.

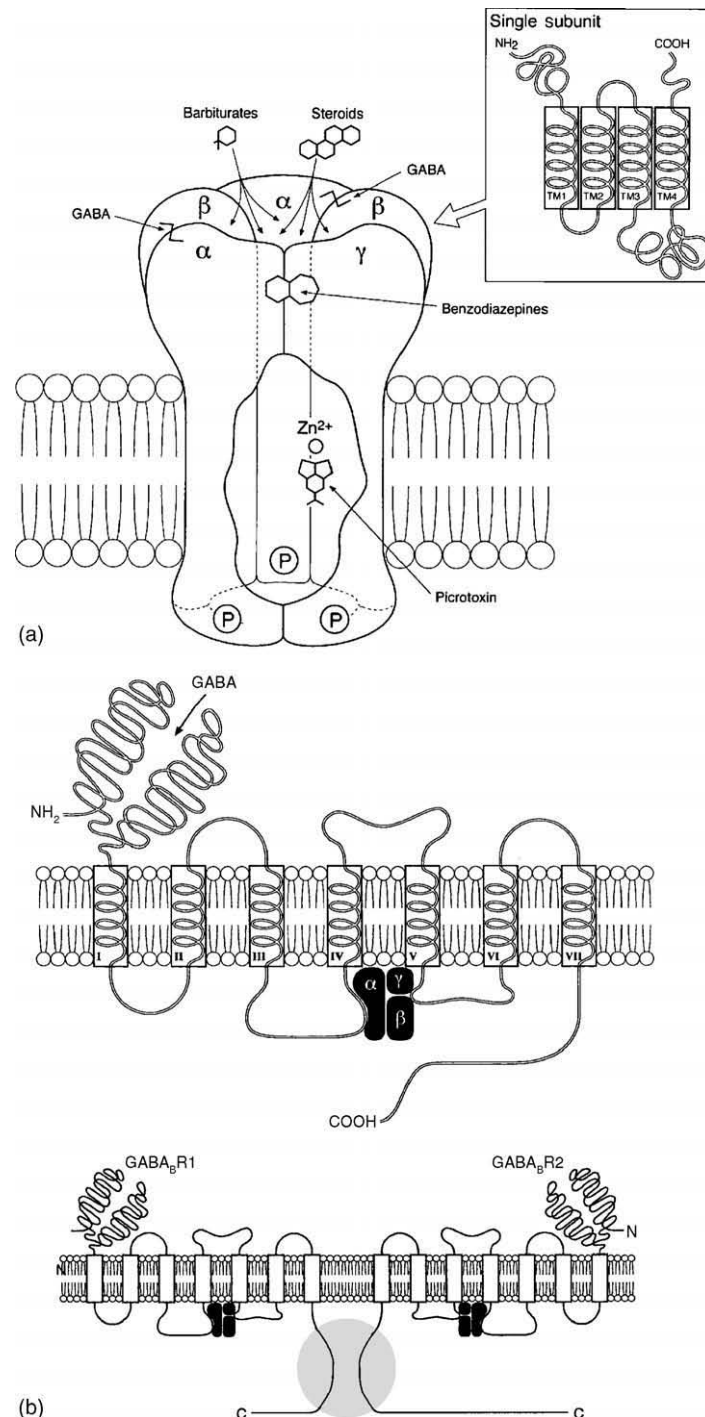


Figure 1 The molecular structure of GABA receptors. a, GABA_A receptor complex; b, GABA_B receptor. (a) The GABA_A receptor complex is assembled from five individual subunits (with an overall stoichiometry of 2 α , 2 β , and 1 γ) that surround a central chloride-selective channel (ionophore). The insert shows an individual subunit, which is composed of approximately 500 amino acids and has four transmembrane domains (TM1–4). Two GABA binding sites (receptors) are located between the α and β subunits. The GABA_A receptor also contains a number of modulatory sites. The benzodiazepine receptor is located at the interface between α and γ subunits, whereas the receptors for steroids and barbiturates show little subunit specificity (indicated by the multiple arrowheads). Anesthetics, ethanol, and the convulsant picrotoxin exert their actions at sites associated with the channel. The areas marked by an encircled P on the inside of the channel denote sites that can be phosphorylated and that modify channel function. (b) The GABA_B receptor has seven transmembrane domains (I–VII) with a large extracellular N-terminal region that forms the receptor for GABA. The complex is coupled to a G-protein complex on the inside of the membrane, which mediates the cellular effects following receptor activation. The lower diagram shows the functional GABA_B receptor. This is a dimer formed by two receptors, GABA_BR1 and GABA_BR2, which associate at the C-terminal region. Adapted from (a) McKernan and Whiting (1996); (b) Kaupmann et al. (1998); Kuner et al. (1990).

Table 1 Distribution of the major GABAA receptor subtypes in the rat brain^a

GABA _A receptor subtype	Relative abundance (%)	Location and putative function
$\alpha_1 \beta_2 \gamma_2$	43	Most brain areas Interneurons in hippocampus and cortex Cerebral Purkinje cells
$\alpha_2 \beta_{2/3} \gamma_2$	18	Spinal cord motoneurons Hippocampal pyramidal cells
$\alpha_3 \beta_n \gamma_2/\gamma_3$	17	Cholinergic and monoaminergic neurons
$\alpha_2 \beta_n \gamma_1$	8	Bergmann glia Nuclei of the limbic system
$\alpha_5 \beta_3 \gamma_2/\gamma_3$	4	Predominantly hippocampal pyramidal cells
$\alpha_6 \beta \gamma_2$	2	Cerebellar granule cells
$\alpha_6 \beta \delta$	2	Cerebellar granule cells
$\alpha_4 \beta \delta$	3	Thalamus and hippocampus dentate gyrus
Other minor subtypes	3	Present throughout the brain

^aAdapted from McKernan and Whiting (1996).

Synaptogenesis starts about 2 weeks after birth. The receptors cluster at the postsynaptic density and are exposed to a high concentration of GABA released from the presynaptic vesicles (see [Box 3](#)). These receptors predominantly contain the α_1 subunit, which has a relatively low affinity for GABA and is less susceptible to activation by extrasynaptic GABA. Meanwhile, an outwardly directed Cl^- transporter reduces the intracellular Cl^- concentration, and Cl^- becomes negative with respect to the resting membrane potential. The activation of the synaptic GABA_A receptors then produces the familiar inhibitory response.

Another example of the exploitation of receptor diversity is that α_1 subunits are expressed throughout the somatodendritic axis of hippocampal CA1 pyramidal neurons, whereas α_2 subunits are expressed on the initial segment of the axon ([Table 1](#)), where the output signal is generated (see [Box 2](#)).

Modulatory sites associated with GABA_A receptors

The receptors for a number of therapeutically important drugs are associated with the GABA_A receptor complex. Separate and distinct receptors have been identified for benzodiazepines, barbiturates, and steroids, whereas alcohol and some general anesthetics also act through the GABA_A system ([Figure 1](#)). The binding of ligands to these receptors does not generally operate the GABA_A ionophore directly. Rather, these act as modulatory sites that are allosterically coupled to GABA_A receptors. The occupation of a modulatory receptor increases the affinity of the GABA_A receptor and thereby increases the probability that the ionophore will open in response to GABA. Synaptic inhibition is prolonged, which depresses activity in neuronal networks. The therapeutic and behavioral effects of actions at modulatory sites are (in ascending order): anxiolytic → sedative → hypnotic → anesthetic.

Benzodiazepines The action of benzodiazepines shows subunit specificity. The minimum requirement for benzodiazepine sensitivity is an α subunit selected from α_1 , α_2 , α_3 , or α_5 , together with either a γ_2 or γ_3 subunit. The benzodiazepine receptor itself is probably located at the interface between the α and γ subunits. Benzodiazepines have no actions at receptors containing α_4 , α_6 , or γ_1 subunits, whereas the identity of the β subunit seems to have little influence on the effect of benzodiazepines.

The benzodiazepine receptor possesses the unique feature that the identity of the ligand determines whether it will interact positively or negatively with the GABA_A receptor. Agonists at the benzodiazepine receptor (such as diazepam and flunitrazepam) increase the likelihood that the ionophore will open in response to GABA and have therapeutic profiles as anxiolytic, anticonvulsant, sedative, and hypnotic. On the other hand, inverse agonists such as the β -carboline, β -CCM, and DMCM decrease the probability of channel opening, and their behavioral profile is correspondingly proconflict and proconvulsant. Benzodiazepine receptor antagonists block the actions of both agonists and inverse agonists.

In accordance with the intense efforts devoted to this therapeutically important locus, at least seven chemical classes of substances have been identified that interact with the benzodiazepine receptor. Investigations with representatives of these have shown further discrete subunit-dependent actions and have helped identify two main subtypes of benzodiazepine receptor. BZ₁ receptors have a high affinity for diazepam and β -CCM. They are characterized by containing an α_1 subunit and are widely expressed in the CNS (particularly the cerebellum). BZ₂ receptors have a high affinity for diazepam but a low affinity for β -CCM. BZ₂ receptors contain α_2 , α_3 , or α_5 subunits and are mainly expressed in a complementary pattern

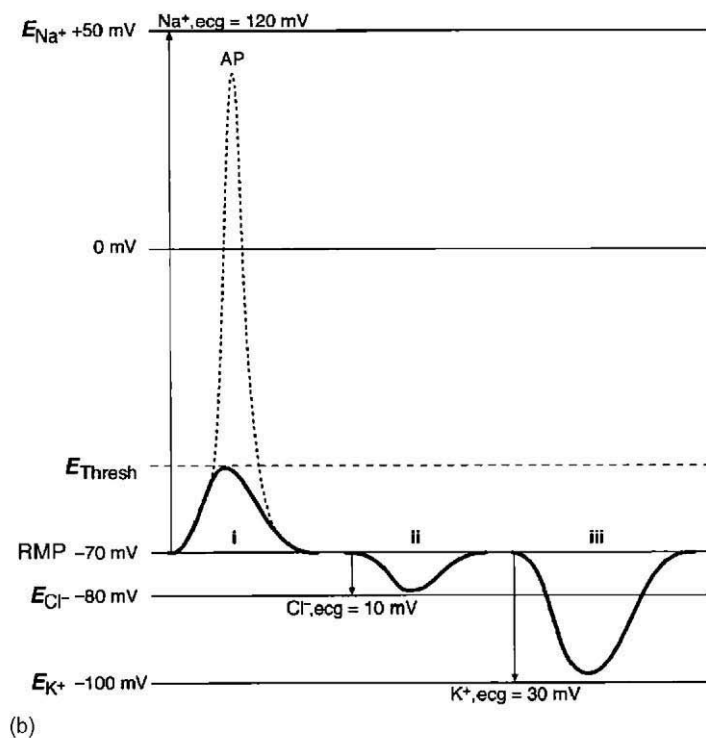
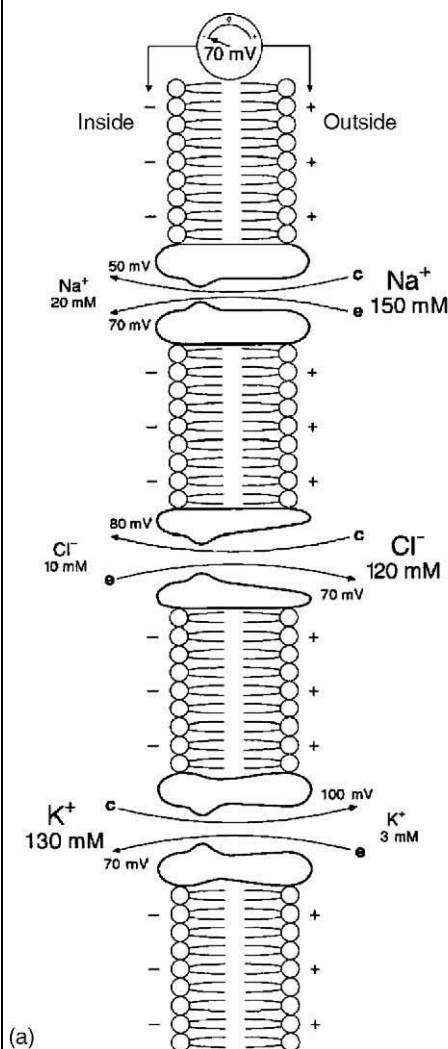
Box 2**Why GABA Is an Inhibitory Transmitter**

In order to understand why and how GABA acts as an inhibitory transmitter, it is necessary to consider the electrochemical basis of neuronal activity, which is illustrated in the accompanying diagrams.

(a) There is a potential difference across the membrane of the resting neuron, with the inside being approximately -70 mV with respect to the outside. This potential arises because of the asymmetrical distribution of the major monovalent ions (Na^+ , K^+ , and Cl^-) across the membrane. The diagram shows representative ionophores for these ions, along with their relative concentrations on both sides of the membrane. The resting membrane is relatively permeable to K^+ and Cl^- , but it is impermeable to Na^+ (indicated by the relative sizes of the channels through the ionophores). Each ionic species is under the influence of two gradients: a chemical gradient (**c**) favoring movement from high to low concentration and an electrical gradient (**e**) favoring movement toward the opposite polarity. The net driving force on a particular ion is given

by the sum (e.g., for Na^+) or difference (e.g., for K^+ and Cl^-) of these two gradients. In order to compare these gradients directly, the Nernst equation is used to convert the concentration gradient to an equivalent electrical gradient, the values of which are given on the arrows representing the chemical gradient and on horizontal lines in part (b). When the potential gradient exactly counterbalances the concentration gradient, equilibrium pertains and no net ionic movement occurs.

(b) A diagram of potential levels showing resting membrane potential (RMP), the equilibrium potentials for the ions (E_{Na^+} , E_{K^+} , E_{Cl^-}), and the threshold for action potential generation (E_{Thresh}). Zero potential is also shown for reference. A selective increase in membrane permeability to one of the ions causes it to move down its electrochemical gradient (ecg), and the membrane potential moves toward the equilibrium potential for that particular ion. Such a change in permeability can arise because the nerve cell membrane contains a multitude of

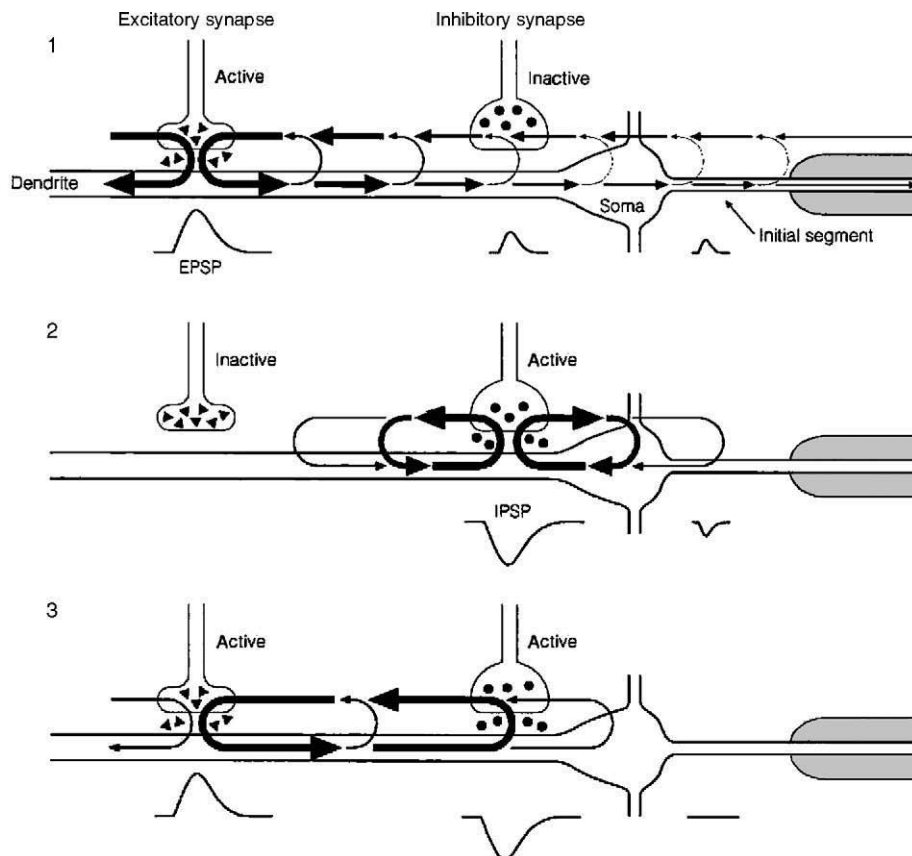


Box 2
Continued

different ion-selective channels that can be activated (gated) by changes in the transmembrane potential or by the binding of ligands and transmitters as to specific receptors (e.g., GABA). A selective increase in Na permeability (i) results in a depolarization and, if E_{Thresh} is reached, evokes a regenerative action potential (AP, stippled line). The AP is a positive-going potential of approximately 100 mV lasting for 1–3 ms and is the output signal of the neuron. Factors that move the membrane potential toward E_{Thresh} increase the likelihood of neuronal firing and are therefore termed excitatory. Conversely, increasing the membrane potential (hyperpolarization) makes it less likely that the neuron will fire an AP, and these influences are therefore inhibitory. GABA is an inhibitory transmitter because it acts on GABA_A and GABA_B receptors to cause a selective increase in permeability to Cl^- ions (ii) and K^+ ions (iii). Because both E_{Cl^-} and E_{K^+} are more negative than the resting membrane potential, action at both receptors causes a hyperpolarization, which inhibits the neuron. This hyperpolarization is, however, not the only reason why the neuron is inhibited.

(c) It is also necessary to consider the spatial interactions between excitatory and inhibitory synapses. A typical nerve cell consists of a cell body (soma)

from which a more or less extensive dendritic tree arises. This acts as the receiving and integrating area for excitatory and inhibitory synaptic contacts. The axon also arises from the soma and conducts the APs to the terminals where it causes transmitter release. The initial segment of the axon (or axon hillock) has the highest density of voltage-gated Na^+ channels and therefore the lowest threshold for AP generation. As to whether this threshold is reached depends on the ongoing integration of synaptic activity on the somatic and dendritic membranes. (1) A representative excitatory synapse is selectively activated. These synapses use glutamate as a transmitter, which gates ionophores that are selectively permeable to Na^+ . Positive current flows into the postsynaptic membrane at the synapse. Because the bilipid membrane has a high resistance, this current can flow relatively long distances within the neuron before exiting through the membrane and returning to the active synapse to complete the current loop. This spread of current allows the possibility of both spatial and temporal summation with activity at other synapses. For example, if a sufficient number of excitatory synapses are activated, the initial segment is depolarized to firing level and an AP arises. (2) Inhibitory GABA_A synapses are



(c)

Box 2
Continued

usually located close to the soma, where they can exert a correspondingly greater influence on the initial segment. The selective activation of the inhibitory synapses causes Cl^- ions to move through the ionophore (which is equivalent to an outward movement of positive current, as shown by the direction of the arrowhead), which results in an IPSP. (3) When excitatory and inhibitory synapses are

activated simultaneously, the GABA_A ionophores provide a low-resistance pathway through which the excitatory current will pass. This means that much less depolarizing current reaches the initial segment and the output of the neuron is inhibited. Thus, GABA_A synapses inhibit the neuron both by causing a hyperpolarization and by shunting excitatory synaptic activity.

to BZ_1 receptors (i.e., where the concentration of α_1 is low). Note that the aforementioned changes in receptor composition during development actually represented a shift from BZ_2 to BZ_1 receptors.

Barbiturates and steroids The GABA_A receptor is also modulated by barbiturates and certain metabolites of gonadal and adrenal hormones (such as progesterone and deoxycorticosterone, respectively). When synthesized in the brain, these metabolites are termed neurosteroids (to distinguish them from neuroactive steroids, which may be synthesized elsewhere) and are therefore an endogenous ligand for a modulatory site associated with the GABA_A receptor. Steroids were originally thought to interact with the barbiturate receptor until the separate identity of the two receptors was established.

The actions of barbiturates and steroids have much more in common with one another than they do with benzodiazepines. Barbiturates and steroids potentiate responses to GABA by prolonging the duration of channel open time, which contrasts to the increase in frequency of channel openings produced by benzodiazepine agonists. Furthermore, at concentrations greater than are required for the modulatory effect on GABA_A receptors, both barbiturates and steroids can gate the ionophore directly by an action that does not involve the GABA_A receptor itself. Such direct gating is never seen with benzodiazepines. Antagonists of the steroid receptor have been characterized (e.g., dehydroepiandrosterone), but so far no antagonists of the barbiturate receptor have been identified.

The actions of barbiturates and steroids do not show absolute subunit dependency, although their effects may be influenced more subtly by the actual composition of the receptor. The therapeutic and behavioral actions of barbiturates and steroids are generally much more pronounced than benzodiazepines. The former tend to cause heavy sedation and anesthesia, whereas benzodiazepines generally have milder effects. The lack of subunit specificity means barbiturates and steroids are likely to enhance GABA_A ergic

actions more globally and indiscriminately. For example, the GABA_A receptors expressed in the thalamus do not contain γ subunits (Table 1) and are not, therefore, responsive to benzodiazepines. The responses are, however, enhanced by barbiturates and steroids.

Other agents that modulate activity of the GABA_A receptor In addition to the barbiturate anesthetics, a wide range of other anesthetics appear to potentiate responses mediated by GABA_A receptors. These include the inhalation anesthetics halothane, isoflurane, and diethylether and the intravenous anesthetics propofol and etomidate. Augmentation of GABA_A ergic transmission is likely to make a marked contribution to the anesthetic action of these substances.

GABA_A ergic transmission is also modulated by ethanol. Interestingly, the effect of ethanol seems to be mediated by phosphorylation of the γ_{2L} subunit.

Side effects of agents that act at the GABA_A receptor Familiar problems are associated with long-term therapeutic treatments with many agents that act at the GABA_A modulatory sites. Prolonged use of benzodiazepines, barbiturates, and ethanol all cause tolerance, which is followed by dependence and addiction. There are also manifestly unpleasant symptoms on withdrawal of treatment. The increasing knowledge of the structure and function of the GABA_A receptor subtypes has intensified efforts to develop drugs that act at specific subunits (or combinations thereof). These would have a narrower clinical profile with fewer side effects and could be used for conditions such as anxiety, epilepsy, and insomnia.

GABA_B Receptors

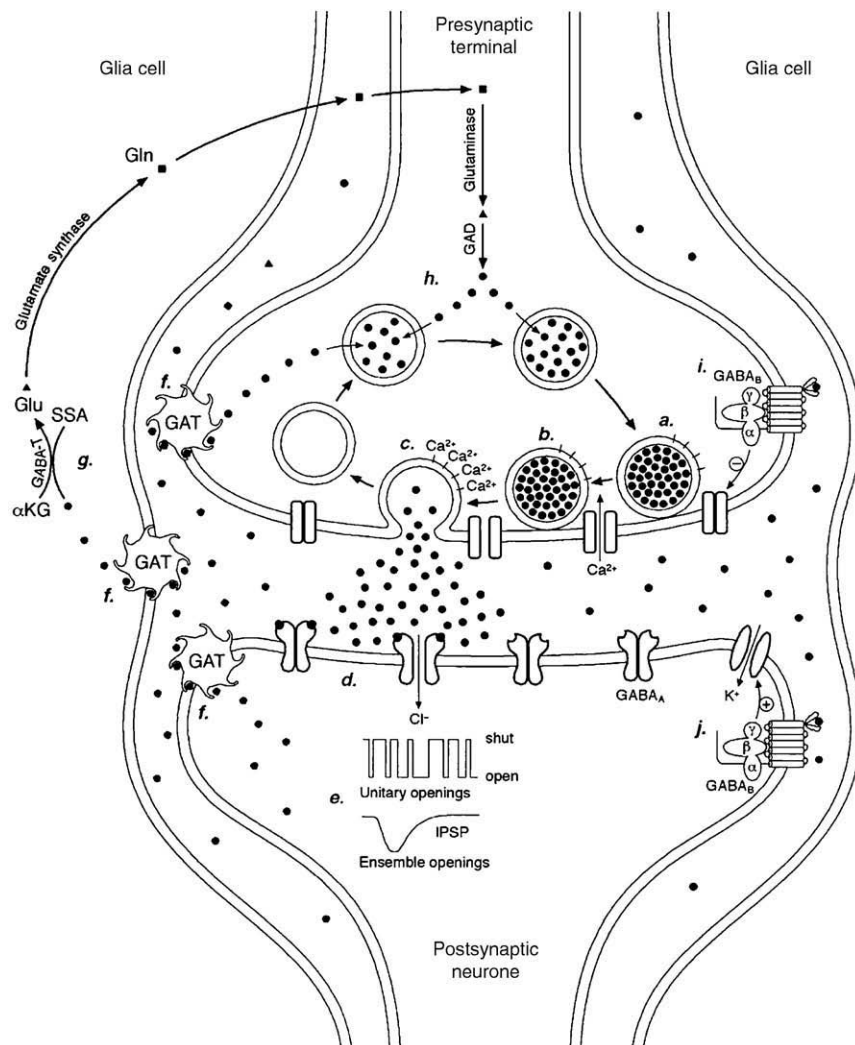
The GABA_B receptor is formed by a single protein containing seven hydrophobic transmembrane domains and a large N-terminal extracellular domain that forms the receptor for GABA (Figure 1). This structure is characteristic of metabotropic receptors, which are not coupled directly to an ionophore but exert their

Box 3Synaptic Transmission at a GABA_A Synapse

GABA fulfills the following criteria required for a neurotransmitter: it is present in the nerve terminal, it is released upon nerve stimulation, it activates specific receptors on the postsynaptic membrane, a specific uptake mechanism is available to remove it from the extracellular fluid, and its direct application to neurons mimics the effect of nerve stimulation. These aspects are illustrated in the accompanying diagram, which depicts the sequential steps (a–j) in synaptic transmission at a GABA_Aergic synapse.

The GABA_A synapse contains pre- and postsynaptic specializations at the area called the active zone. Four active zones are illustrated, each represented by a single presynaptic vesicle opposed by a single receptor complex in the postsynaptic membrane. The synaptic cleft is approximately 100 nm wide. (a.) Vesicles filled with GABA are docked at the presynaptic specialization in close proximity to a

voltage-sensitive Ca²⁺-selective channel. (b.) On arrival of an action potential at the nerve terminal, the Ca²⁺-selective channel opens transiently. (c.) Ca²⁺ ions pass into the terminal and bind to four Ca²⁺ sensors associated with the vesicle. This causes exocytosis of the vesicle and GABA is expelled into the synaptic cleft, where it quickly reaches a peak concentration of approximately 500 μM. The empty vesicle leaves the membrane and is recycled. (d.) The binding of two molecules of GABA to each receptor complex causes the subunits to tilt slightly. This opens the ionophores and allows Cl[−] ions to flow across the postsynaptic membrane. (e.) Recordings from single-receptor complexes using the patch-clamp technique have shown that the ionophore flickers between open and shut states in the presence of GABA. Current flowing through the whole ensemble of 20–30 ionophores clustered together at the active zone creates a smooth, but transient, potential



Box 3

Continued

change across the postsynaptic membrane. This is termed the inhibitory postsynaptic potential (IPSP), and lasts approximately 50 ms. The reasons why the IPSP depresses neuronal activity are given in [Box 2](#). Although 90% of the GABA released from the vesicle has diffused away from the active zone within 0.5 ms, some molecules still remain bound to the receptor complexes. However, these ionophores close of their own accord. This process is known as desensitization, and it is thought to protect the GABA_A receptors from overactivation. (f.) Synaptically released GABA is ultimately removed from the extracellular fluid by uptake transporters (GAT) located on both neurons and the supportive glia cells. (g.) The enzyme GABA-transaminase (GABA-T) catalyzes the conversion of GABA and α -ketoglutarate (KG; derived from glucose metabolism) to succinic semialdehyde (SSA) and glutamate (Glu). Glutamate, which cannot cross the extracellular space because it is neuroactive, is then

converted to glutamine (Gln) by the enzyme glutamate synthase. Glutamine diffuses across the membranes and into the presynaptic terminal, where it is reconverted to glutamate by glutaminase, which in turn is converted to GABA by glutamic acid decarboxylase (GAD). (h.) Specific transporters fill the recycling vesicles with GABA, where it is stored at a concentration of approximately 100 mM. These dock with the presynaptic membrane at the active zone, thereby completing the cycle. (i.) GABA_B autoreceptors located on the presynaptic terminal are activated by GABA present in the extracellular space. These are negatively coupled to the Ca²⁺ channels by G-proteins. The depression of Ca²⁺ entry through these channels decreases vesicle release in response to subsequent activation. (j.) Postsynaptic GABA_B receptors are positively coupled by G-proteins to K⁺-selective channels. This mediates a slow IPSP, thereby inhibiting the postsynaptic neuron.

actions via second messengers that alter the activity of intracellular enzymes and other mechanisms.

Two splice variants of the GABA_B receptor, GABA_BR1a and GABA_BR1b, were originally cloned and tested for their ability to mimic the effects of native GABA_B receptors. Although they showed a similar sensitivity to antagonists, their affinity for GABA was approximately 100 times less than the native receptors. Moreover, when expressed in mammalian cells, they showed poor translocation to the membrane and coupled only weakly to the effector mechanisms. The reasons for these anomalies have recently been resolved with the cloning and characterization of a second GABA_B receptor, GABA_BR2. This has 35% sequence homology with the GABA_BR1 receptor and has a similar structure. The two receptors associate together at the C-terminus. This is expressed in the cell membrane as a functional dimer, where it displays the characteristics of the native GABA_B receptor. The binding of GABA to this dimer causes the activation of GTP-binding proteins (G-proteins) on the inside of the membrane. Depending on the identity of the G-protein, this stimulates or inhibits effector enzymes (e.g., adenylyl cyclase and phospholipase C) and proteins (e.g., ionophores), which mediate the cellular response. The binding of a single molecule of GABA to the GABA_B receptor activates many molecules of G-protein, thereby amplifying the response greatly. However, due to the number of intermediate steps involved, G-protein-mediated transduction is characterized by a relatively

long latency (around 40 ms), followed by a time course of action lasting a second or more.

GABA_B receptors mediate two location-specific effects. On postsynaptic dendritic membranes, the G-proteins are positively coupled to a K⁺ selective ionophore. Because E_{K^+} is negative to the resting membrane potential (see [Box 2](#)), the activation of these channels gives rise to a relatively large inhibitory postsynaptic potential (IPSP), which lasts for 1–2 s. GABA_B receptors are also located on presynaptic terminals where they are negatively coupled to the voltage-dependent Ca²⁺ channels. The reduction of Ca²⁺ flux through these channels reduces the probability that a vesicle will be released in response to an action potential (AP) and thereby decreases synaptic transmission. These GABA_B receptors are located on the presynaptic terminals of a variety of neurons that use a range of transmitters. In the special case in which they are located on the terminals of GABAergic neurons themselves, they are termed autoreceptors. During repetitive activity, these autoreceptors serve to decrease the synaptic release of GABA, which will promote the expression of postsynaptic excitatory responses from other sources.

Although pre- and postsynaptic GABA_B receptors apparently show differences in their sensitivity to agonists and antagonists, it is presently unresolved as to whether these represent pharmacologically distinct receptors. Furthermore, it is not certain to what extent GABA_B receptors are located within the postsynaptic density (see [Box 1](#)). They are certainly

expressed at extrasynaptic sites, where they respond to ambient levels of GABA in the extracellular fluid (which partly arises as spillover from nearby GABA_A synapses). This notion that GABA_B receptors play more of a neuromodulatory role rather than a direct-transmitter role is supported by the fact that relatively strong stimulation of GABAergic neurons is required to activate pre- and postsynaptic GABA_B receptors.

GABA_B receptors show high concentrations in certain regions of the CNS, including the cerebellum, thalamus, and spinal cord. Although this pattern is distinct from that of GABA_A receptors, there are many regions where the two receptors overlap. It is likely that the majority of single neurons express both GABA_A and GABA_B receptors.

Effects mediated by GABA_B receptors have been implicated in the neuronal processes associated with memory (long-term potentiation), slow-wave sleep, muscle relaxation, and antinociception. The GABA_B agonist, baclofen, has long been in clinical use as an antispastic agent and also shows promise in the treatment of chronic pain and withdrawal symptoms from cocaine. GABA_B receptor antagonists do not presently have clinical applications, but have been shown to be effective in animal models of absence epilepsy and cognitive impairment.

GABA_C Receptors

GABA_C receptors are formed from homo- or heteromeric assemblies of ρ_1 or ρ_2 subunits and are structurally and functionally similar to GABA_A receptors. Although ρ subunits are probably the ancestral form of the GABA_A subunits, they do not apparently assemble together with other subgroups. GABA_C receptors are characterized by their lack of sensitivity to bicuculline and high sensitivity to GABA (Box 3). The GABA_C-operated ionophores have a substantially lower conductance than the GABA_A ionophores (probably because its diameter is slightly smaller), and they do not desensitize in the maintained presence of GABA. Finally, GABA_C receptors lack the modulatory sites associated with the GABA_A receptor. Functional GABA_C receptors have so far been demonstrated only in the retina. However, ρ subunits occur at many other sites in the CNS (e.g., the cerebellum), and mRNA for ρ_2 is present throughout the brain.

GABAergic Neurons and Their Role in Neuronal Activity

Because the only known function of GABA is to act as an inhibitory neurotransmitter, GABAergic neurons may be identified by histochemical demonstration

of the synthesizing enzyme, glutamic acid decarboxylase (GAD). The majority of GABAergic neurons are interspersed between the principal neurons of a given anatomical structure; then they are termed interneurons. However, they can also have long axons that project to other regions in the CNS. Although sharing the same transmitter, GABAergic neurons otherwise exhibit a wide range of morphological, histochemical, and functional characteristics. For example, at least a dozen distinct types of GABAergic neurons can be identified in the cerebral cortex.

Some of these have very narrow axonal fields confined to a cortical column, whereas others have extremely extensive axonal arborizations that can extend over large distances. GABAergic neurons themselves are under the control of other synaptic influences: they are excited by glutamatergic synaptic inputs and inhibited by other GABAergic neurons. The nature and degree of synaptic inhibition they produce on the neurons with which they form synaptic contacts depends on

1. Their intrinsic membrane properties and firing patterns. Many GABAergic neurons are able to fire APs at relatively high frequencies with little accommodation. This maximizes GABA release at the terminals (which in turn may be under the feedback control exerted by autoreceptors).
2. The number and location of the synaptic contacts on the postsynaptic membrane. Such contacts do not occur randomly but are placed in a strict topographical order on the postsynaptic membrane.
3. The identity of the postsynaptic receptors (e.g., GABA_A or GABA_B receptors and their possible subtypes).

By controlling the excitability of the postsynaptic neuron, GABAergic neurons participate in the regulation of network activity. Expressed in its simplest form, GABAergic inhibition curtails the activity of principal neurons in response to an excitatory synaptic input (feed-forward inhibition) and after an output signal has been generated (feedback inhibition) (Figure 2). These inhibitory influences act as a brake on the principal neuron, which could otherwise enter an unstable oscillatory state in which relatively modest excitation can result in an explosive response. Indeed, pharmacological blockade of GABAergic inhibition is often used in experimental situations to evoke activity that is reminiscent of epileptic discharges. It is becoming increasingly evident that GABAergic inhibition cannot be regarded simply as a counterweight to intrinsic and synaptic excitatory processes. GABAergic synapses participate in the spatiotemporal integration of activity in neuronal networks and thereby play a major role in the processing of information. GABAergic neurons are

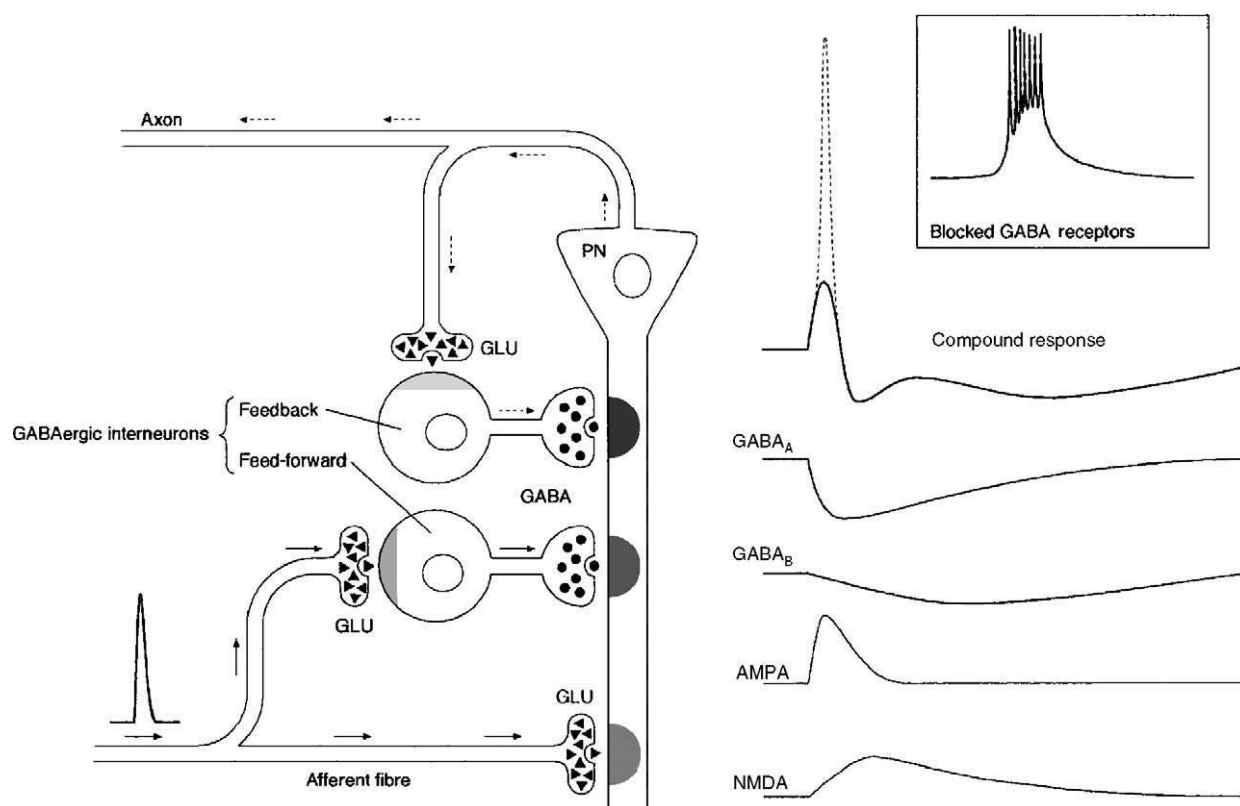


Figure 2 The role of GABA receptors in determining network excitability. The diagram shows how GABAergic inhibition controls the activity of principal neurons (PN). Only single synapses of each type are shown for clarity. Stimulation of the afferent input pathway (bottom left) evokes an action potential, which spreads in the nerve fiber (arrows). This ultimately results in a compound synaptic response in the PN (shown on the right), the components of which arise as follows. The initial response is a rapid depolarization by an excitatory postsynaptic potential (EPSP), which is mediated by two different types of excitatory receptors for glutamate (AMPA and NMDA). If sufficient excitatory synapses are activated, the threshold for the generation of an action potential (stippled line) may be reached. Meanwhile, collateral fibers also activate GABAergic interneurons. These release GABA on to the principal neuron after a short delay (feed-forward inhibition). The GABA activates both GABA_A and GABA_B receptors, which activate Cl⁻ and K⁺-selective ionophores, respectively. Both actions hyperpolarize the neuron, thereby ensuring that the membrane potential returns toward the resting level and prevents the overexcitation of the PN. If the PN fires an action potential (AP), this is transmitted down its axon (stippled arrows). Axon collaterals excite GABAergic interneurons, which feed back on to the PN and inhibit it (feedback inhibition). The insert shows the response to a single afferent stimulation following pharmacological blockade of the GABA receptors. The synaptic excitation is no longer curtailed and also recruits voltage-dependent ionic conductances in the membrane. This gives rise to a large depolarization surmounted by a number of action potentials, which is reminiscent of epileptic activity.

particularly prevalent in the cerebral cortex, where they serve to synchronize the firing of cortical pyramidal neurons into columnary activity.

See Also the Following Articles

Anxiolytics; Hippocampal Neurons.

Further Reading

- Biggio, G., Sanna, E. and Costa, E. (eds.) (1995). *GABA_A receptors and anxiety: from neurobiology to treatment*. New York: Rover.
- Costa, E. (1998). From GABA_A receptor diversity emerges a unified vision of GABAergic inhibition. *Annual Review of Pharmacology and Toxicology* 38, 321–350.

- Johnston, G. A. R. (1996). GABA_C receptors: relatively simple transmitter-gated ion channels. *Trends in Pharmacological Science* 17, 319–323.
- Kaupmann, et al. (1998). *Proceedings of the National Academy of Sciences* 95, 14991–14966.
- Kuner, et al. (1990). *Science* 283, 74–77.
- McKernan, R. M. and Whiting, P. J. (1996). Which GABA_A-receptor subtypes really occur in the brain? *Trends in Neuroscience* 19, 139–143.
- Miller, R. J. (1998). Presynaptic receptors. *Annual Review of Pharmacology and Toxicology* 38, 201–227.
- Sieghan, VIT. (1995). Structure and pharmacology of gamma-aminobutyric acid_A receptor subtypes. *Pharmacological Reviews* 47, 181–234.
- Smith, G. B. and Olsen, R. W. (1995). Functional domains of GABA_A receptors. *Trends in Pharmacological Science* 16, 162–168.

Gastrointestinal Effects

R Murison and A M Milde

University of Bergen, Bergen, Norway

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R Murison, volume 2, pp 191–195, © 2000, Elsevier Inc.

Brain–Gut Interactions

Gastrointestinal Disorders and the Influence of Stress

Conclusion

Glossary

<i>Amygdala</i>	A walnut-shaped brain structure that plays a key role in fear.
<i>Atropine</i>	An anticholinergic drug.
<i>Bombesin</i>	A peptide important in various activities both in brain and other organs.
<i>Brain-derived neurotrophic factor (BDNF)</i>	A protein important for maintaining function in various brain areas.
<i>Cholecystokinin (CCK)</i>	A hormone produced by the mucosa of the upper intestine that stimulates contraction of the gallbladder.
<i>Cholinergic</i>	Liberating, activated by, or resembling the physiologic action of acetylcholine.
<i>Functional dyspepsia</i>	A nonorganic disorder typified by stomach discomfort and pains that can not be explained by organic changes.
<i>Gastrin</i>	A polypeptide hormone secreted by the pyloric mucosa that stimulates the pancreas to release pancreatic fluid and the stomach to release gastric acid.
<i>Helicobacter pylori</i>	A family of gram-negative bacteria whose members are spiral shaped, showing corkscrewlike motility; they are found in the vast majority of patients with peptic ulcer.
<i>Hippocampus</i>	A ridge that extends over the floor of the descending horn of each lateral ventricle of the brain.
<i>Histamine</i>	An amine derivative of histidine that is widely distributed in human tissues.
<i>Hypophysectomy</i>	Surgical removal of the pituitary gland.
<i>Inflammatory bowel disease (IBD)</i>	A group of organic disorders of the bowel, including ulcerative colitis and Crohn's disease.
<i>Interferon</i>	A protein that is produced by intact animal cells when they are infected with viruses; it acts to inhibit viral reproduction and to induce resistance in host cells.

Interleukins

Any of a class of proteins that are secreted mostly by macrophages and T lymphocytes and induce growth and differentiation of lymphocytes and hematopoietic stem cells.

Irritable bowel syndrome (IBS)

A functional (nonorganic) disorder characterized by discomfort and pain in the bowel.

Ischemia

Localized anemia as a result of obstruction of the blood supply or to vasoconstriction.

Neuropeptide Y

A polypeptide released by axons at the synapse; it may act as a neurotransmitter and have a direct effect on synapse function or as a neuromodulator, having a long-term effect on postsynaptic neurons.

Neuroticism
Nonsteroidal anti-inflammatory drugs (NSAIDs)
Nonulcer dyspepsia (NUD)
Oxytocin

A neurotic condition, character, or trait.
Drugs that are widely used for the treatment of inflammatory disorders (such as rheumatoid arthritis) and other painful conditions.
Functional dyspepsia.

Peptic ulcer

A polypeptide hormone secreted by the neurohypophysis that stimulates contraction of the uterine muscles.

An ulcer involving the mucosa, submucosa, and muscular layers of the stomach or duodenum, due in part at least to the action of gastric acid–pepsin juice.

Rapid eye movement (REM) sleep

The part of the sleep cycle during which the eyes move rapidly, accompanied by a loss of muscle tone and a low-amplitude encephalogram recording.

Sacral nerves

Any of five pairs of spinal nerves in the sacral region that innervate the muscles and skin of the lower back, lower extremities, and perineum and branch to the hypogastric and pelvic plexuses.

Serotonin

A compound derived from tryptophan that functions as a local vasoconstrictor, plays a role in neurotransmission, and has pharmacological properties.

Sigmoid colon

The S-shaped portion of the colon between the descending colon and the rectum.

Somatostatin

A peptide secreted by the hypothalamus that acts primarily to inhibit the release of growth hormone from the anterior pituitary.

Vagotomy

An operative procedure involving cutting of the vagus nerve, previously used in treatment for ulcers.

Vagus nerve The tenth cranial nerve, forming an important part of the parasympathetic system and carrying information in both directions between the brain and the periphery.

The functions of the gut encompass ingestion, digestion, absorption, and transport of ingested materials through the gastrointestinal (GI) tract. Digestive processes rely on the presence of acid and enzymes to break down the nutrients into individual molecules that can be absorbed across the intestinal lining for passage into the bloodstream. At the same time, the mechanical and chemical means to digest food particles also represent potential threats (aggressive factors) to the integrity of the gut mucosa. If the stomach lining secretes more hydrochloric acid than needed, it has the potential of irritating the duodenal and gastric lining, thus enhancing their sensitivity. Other aggressive factors include nonbeneficial bacteria (in particular *Helicobacter pylori* in humans) and our tendency to self-treatment with NSAIDs. Such a self-digesting potential is prevented by protective factors: (1) the secretion of protective mucus from various mucus glands throughout the GI tract, (2) the secretion of bicarbonate in saliva to remove hydrogen ions, and (3) a high rate of cell regeneration. Endogenous prostaglandins are important for the proper function of these protective factors, as is also an adequate mucosal blood flow. Various stressors are able to modulate all of these functions. Because many effects appear to be stressor-specific and indeed species-specific, several generalizations must be made in what follows.

Brain-Gut Interactions

For stress to impact the GI system, pathways must be in place to permit a bidirectional communication between the central nervous system (CNS) and the gut. Such mechanisms include both nerve pathways and the endocrine system. Among the former are parasympathetic vagal connections terminating in the enteric nervous system, which comprises two plexuses, or networks of neurons, embedded in the wall of the GI tract; the outermost collection of neurons, called the myenteric (or Auerbach's) plexus; and the innermost group of neurons, called the submucosal (or Meissner's) plexus. Included in the latter are the sympathetic innervations via the spinal cord. Among the hormonal influences are epinephrine, adrenocorticotrophic hormone (ACTH), cortisol, and thyrotropin-releasing hormone (TRH). In addition, several immune-system substances (as interferon and interleukins) are now known to impinge on the GI tract.

Secretions

Gastric acid secretion is regulated by an interplay of several neural (cholinergic), hormonal (gastrin), and paracrine (histamine and somatostatin) mechanisms, where histamine is a potent inducer of acid secretion. TRH acts in the brain to stimulate gastric acid, pepsin, and serotonin secretion. TRH is also involved in mechanisms regulating mucosal blood flow, contractility, emptying, and activation of the parasympathetic outflow to the stomach, which may produce gastric ulcerations. Both vagotomy and atropine block this effect. It would appear that the main sites of TRH action in the brain for these effects lie in the dorsal vagal complex, nucleus tractus solitarius, nucleus ambiguus, and the dorsal motor nucleus of the vagus, which receive direct input from the amygdala and indirect input from the ventral hippocampus. Other centrally acting peptides shown to increase gastric secretion include CCK, gastrin, and oxytocin, whereas somatostatin inhibits the secretion of gastrin and histamine and appears to have a direct inhibitory effect on parietal cells.

Centrally acting corticotropin-releasing hormone (CRH), on the other hand, stimulates sympathetic activity while inhibiting parasympathetic activity. Injected into the paraventricular nucleus, CRH reduces acid secretion while increasing bicarbonate secretion. These central effects of CRH are independent of the hypothalamic-adrenocortical axis and vagal activity, but are prevented by adrenalectomy and noradrenergic blocking agents, suggesting that they are mediated by sympathetic outflow. CRH also inhibits the development of stress-induced gastric ulcerations. Apart from somatostatin, other peptides and proteins shown to inhibit gastric acid secretion include bombesin, neuropeptide Y, interleukin-1, platelet activating factor, and the opioid peptides.

Paradoxically, therefore, the immediate effect of central neuroendocrine stress-related peptides appears to be an inhibition of gastric acid secretion. Although the renowned Stewart Wolf and Harold Wolff studies from 1947 on the fistulated patient Tom suggested gastric hypersecretion in response to anxiety, resentment, and hostility, they also noted a transitory hyposecretion during situations associated with overwhelming despair and dejection. Results from animal studies indicate that acute stress is associated with a hyposecretion of acid followed by a hypersecretion (a rebound effect).

Mucus production is also inhibited by stress. The underlying mechanism for this effect appears to be related to the capacity of adrenal steroids to block the prostaglandin synthesis necessary for mucus production within the mucous epithelial cells, as well as

modulating both gastric vasodilatation necessary for adequate circulation and hydrogen ion (H^+) release from the parietal cells. Of the other protective factors, CRH and stress promote bicarbonate secretion but reduce blood flow, leaving the mucosa at risk for focal ischemia. Conversely, Filaretova et al. suggest that corticosteroids may act to maintain gastric blood flow during stress.

The overall picture therefore suggests an acute response to stress of closing down the upper GI system to allow redistribution of energy resources to the organs involved in fight-or-flight behavior. Both human and animal data, however, point to the possibility of local rebound processes in the wake of this acute response. Because the various response systems (e.g., acid secretion, mucus secretion, and blood vessel constriction) do not necessarily follow the exact same time course, a rebound-related increase in acid secretion without the simultaneous reinstatement of adequate blood supply and mucus secretion represents a potential danger to the mucosal integrity.

Gastrointestinal Motility

GI motility is controlled by the reflex receptive relaxation mediated by vagal inhibitory fibers and endocrine influences. Perceived emotional stressors such as fear generally result in acute decreases in gastric motility and gastric emptying, evoked dysrhythmic gastric myoelectrical activity, and increases in colonic activity. These effects are mimicked by CRH and blocked by its antagonists. Pathways for these effects are believed to include higher brain functions that provide descending information to the gut via vagal and sacral nerve pathways rather than via the hypothalamo-adrenocortical axis. Motility is under the influence of several transmitters, most attention being focused on serotonin. Selective serotonin reuptake inhibitors (SSRIs) have proven effective in treatment of motility disorders.

Stomach The majority of mammalian animal studies (rats, mice, dogs, and guinea pigs) indicate that acute stress of various types (shock, avoidance, noise, and handling) leads to a delay in gastric emptying associated with reduced contractile activity or a series of slow contractions outlasting the stress period. It is important to take into account the nutrient content and nature of the meal (acoustic stress delays emptying but only with nutritive meals bulk) and the time course of events (cold increases gastric motility in rats, but this is followed by a decrease). The effects of stress in rat stomach are mimicked by the central application of CRH, which can be blocked by both vagotomy and norepinephrine antagonist drugs.

Early human studies suggested both increases and decreases in gastric motility in response to anger and resentment. Later studies have more consistently shown reduced motility in response to a number of stressors (cold pain, painful transcutaneous stimulation, and noise).

Small intestine The stomach connects the esophagus to the small intestine, and the majority of food digestion takes place here. Stress may alter normal transit time, the gut contents taking longer to travel from the small to the large intestine, and more liquid may be extracted, forming the basis for constipation. This is a potent symptom of the constipated predominant IBS, a functional, nonorganic, pathological bowel disorder.

Large intestine The large intestine is divided into the cecum, colon, and rectum, the cecum making up the first 2 or 3 in. of the large intestine. The colon is subdivided into the ascending, transverse, descending, and sigmoid colon. The last portion of the large intestine is the rectum, which extends from the sigmoid colon to the anus (approximately 6 in.). The nerve supply to the large intestine contains both sympathetic and parasympathetic nerves; sympathetic stimulation inhibits activity and parasympathetic stimulation causes an increase in defecation reflexes. The colon contains the majority of microorganisms in the intestines; these microorganisms weigh more than 1 kg and consist of both friendly and unfriendly bacteria.

Animal studies are fairly consistent in showing increases of colonic motility and colonic transport in response to acute stress or central CRH. The effects of CRH are abolished by noradrenergic blockers but not by adrenalectomy or hypophysectomy. Exposure of rats to psychological stress (as in water-avoidance stress) induces abnormal colonic motility (increased fecal pellet output), enhanced colorectal sensitivity (induced distention), and increased colonic permeability (e.g., urine secretion of 51-chromium ethylenediaminetetraacetic acid).

Gastrointestinal Disorders and the Influence of Stress

Background

The majority of studies have focused on the immediate effects of acute stress on GI function. Clearly, chronic effects of acute stress or effects of chronic stress on GI structure and function will have consequences for illness and disease states.

For decades there has been a clear dissociation in the literature between the functional and organic GI

disorders. The former include functional dyspepsia (FD), (NUD), IBS, and constipation. Organic disorders include peptic ulcer, IBD, cancers, and acute stress ulcer. Traditionally, stress (psychosomatic factors) has been implicated in a number of GI disorders, including gastric and duodenal (peptic) ulcer, gastroesophageal reflux (GERD), and IBS. Apart from the obvious acute effects of stress on the gut, the clinical and epidemiological data support the notion that stress is associated with symptom worsening. A number of reports from war veterans and civilian women indicate high rates of GI symptoms associated with posttraumatic stress disorder (PTSD). However, the relationship between PTSD and the GI system has not received the same research interest as, for example, PTSD's relationship to the cardiovascular system. It is important, as ever, that we be cautious in attributing a disease state to self-reports of prior stress. The presence of abnormal psychological or psychiatric profiles among a patient group does not mean that there is a psychological etiology to the disorder; the psychological changes are at least as likely to be a consequence of the disease (as is suggested for inflammatory bowel disorder). Equally, an absence of psychological/psychiatric abnormalities in a patient group is not sufficient to discard the role of stress in either the onset of the disease process or vulnerability to developing disease symptoms during the life span.

Functional Gastrointestinal Disorders

Patients diagnosed with functional GI disorders (FGID) are frequently reported to have been earlier exposed to severe and/or chronic psychosocial stress and to suffer from emotional disturbances. In recent years, an interest has emerged for studying pain circuits, brain activation, and metabolism (using functional magnetic resonance imaging and positron emission tomography) and neuroendocrine responses to various stressors. There is a consensus that psychological stress plays an important role in patients with IBS in precipitating and/or worsening the symptoms. This has been supported by animal studies of, for example, the long-term consequences of neonatal stress and maternal separation. Induced learned helplessness and chronic mild stress influence subsymptoms such as sleep abnormalities (in REM sleep in particular), production of BDNF, and γ -aminobutyric acid (GABA) receptor activity, which are central to symptoms of depression or anxiety.

Functional dyspepsia FD is a chronic, upper-centered abdominal pain or discomfort characterized by the absence of organic changes that could explain the patient's predominant symptoms, which include nausea, early satiety, vomiting, abdominal distention,

bloating, and weight loss. Nonetheless, some patients may exhibit prepyloric erosive changes of the gastric mucosa that may be confused with acid-induced changes related to gastroesophageal reflux disease (GERD). The role of *Helicobacter pylori* in FD is still under discussion. Although it is not believed to be a cause, testing for and treatment of the infection is regular medical procedure. There is considerable symptom overlap among FD, IBS, peptic ulcer, and GERD. Therefore, a more extended medical examination is often required (e.g., upper intestinal endoscopy). The pathophysiological mechanisms commonly proposed for FD include lowered vagal tone, impaired duodeno-jejunal motility, delayed gastric emptying, impaired gastric meal accommodation, and increased visceral sensitivity in both the stomach and duodenum. The role of stress in the etiology of FD remains unclear and clinical studies show a lack of consistency. Both stress responders and nonstress responders show delayed gastric emptying and increased rates of gastric motility. In addition, psychological profiles and personality traits of FD patients show similar inconsistency although there is some reason to believe that these patients are more prone to seek medical assistance due to general health worries, which may be connected to negative life experiences in early life.

A majority of FD patients presenting to a physician have some psychological or psychiatric abnormality, usually anxiety, mood disorder or neuroticism, or combinations of these, and they also report more adverse life events such as physical and sexual abuse, more relatives with illnesses, and failure to thrive. Personality factors such as trait anger reactivity and neuroticism, coping styles, life stressors, and degree of emotional support are shown to have predictive value in distinguishing among the FGID subgroups. Females with depression, state anxiety, and trait anger temperament are shown to have more FD-like symptoms.

Irritable bowel syndrome Acute studies show that stress increases colonic activity. From the point of view of theories that explain IBS, we are more concerned with accounting for individual differences that can be traced back to earlier stress episodes (i.e., proactive effects). An interesting series of animal studies by Stam has adopted just this approach. A single short series of uncontrollable foot shocks failed to change basal colonic activity in rats but increased colonic activity in response to the presentation of an aversive novel stimulus (shock prod), despite there being no difference in the behavioral response. Such sensitization of the colonic response to a mild (and controllable) stressor after earlier experience with a more severe (and uncontrollable)

stressor provides an attractive model for IBS. Indeed, patients with IBS do report that stressful life events exacerbate their symptoms, and early life stress may be an important contributory factor in the etiology of the disorder. Patients are characterized by higher levels of anxiety, and it is noteworthy that the levels and numbers of shock used by Stam are of the same order of magnitude as those known to induce heightened anxiety in rats. There appears to be an emotional component involved, and there is a gender difference, in that women are affected more than men across all age groups. In addition, there are ethnic differences, in that whites and Jews seem to be more affected than other ethnic groups.

The data are highly suggestive of a link between sexual abuse and later development of IBS symptoms. Twice as many IBS patients report a history of sexual abuse than the normal population, suggesting a sensitization effect of earlier stress on motility function and visceral sensitivity. Closer analysis, however, has led to the conclusion that sexual abuse is more likely to be associated with health-care use in general rather than to any specific GI disorder. In connection with this, IBS patients exhibit visceral hypersensitivity, although their peripheral pain thresholds are normal or even elevated compared to healthy individuals.

Recent studies with brain imaging techniques have revealed that normal subjects respond to rectal distention with the activation of the arcuate nucleus but that IBS patients do not. On the other hand, IBS patients respond to the same stimulus (or anticipation of it) with the activation of the prefrontal cortex. IBS nonpatients do not differ from normals in terms of psychosocial features, but IBS patients do. This may indicate that stress is not a major factor in the etiology of the disorder but rather in general illness behavior.

Organic Gastrointestinal Disorders

Peptic ulcer From being one of the classic psychosomatic diseases, peptic ulcers (stomach or duodenal lining eroded by stomach acid) have to a large extent been removed from the realm of stress-related disorders. The impetus for this radical change was the discovery that the overwhelming majority of patients harbored *H. pylori* and that the elimination of this particular bacterium was almost 100% successful in curing the disorder. A number of issues remain unresolved, however. The incidence of *H. pylori* in the population (50–80% in the older population) is considerably higher than the incidence of peptic ulcer (2–10% lifetime prevalence). The association between the disease and the infection is weaker for gastric peptic ulcer than for duodenal ulcer, as is the efficacy of the eradication treatment. *H. pylori* alone is therefore not sufficient to cause the disease. Other

putative factors include smoking, heavy drinking, and use of NSAIDs. There is considerable evidence from prospective studies that stress is also involved in both the onset and development of the disease. Several studies of groups exposed to chronic stress (e.g., prisoners-of-war and air traffic controllers) also provide evidence for the role of stress. Given the presence of *H. pylori*, it has been suggested that stress acts through the neuroimmune system to modify the rate of colonization of the bacteria. In the absence of the bacteria, it is suggested that peptic ulcer may develop from acute stress ulcerations. The role of *H. pylori* should not be underestimated in also having a potential role in slowing the healing process.

Inflammatory bowel disease IBD is a collective term that refers to both ulcerative colitis (UC) and Crohn's disease (CD). UC involves the inflammation of the mucosal lining of the large intestine, whereas CD may cause the inflammation of the lining and wall throughout the GI tract. The cause of IBD is not known, but there are several theories that combine genetics, immunology, and bacterial and neonatal factors. There is general agreement that UC and CD are not related to stress in any etiological manner; instead, the psychological symptoms associated with these disorders are the result of the poor quality of life in these patients, and stress may induce or exacerbate symptoms of the already existent disease state. Treatment programs focus on assessing possible stressors that may influence healing. Studies show that patients with active symptoms are more likely to report stressful experiences than patients with a more quiescent disease state. Some individuals are especially prone to symptom flare-ups and rehospitalization. Flare-ups of inflammation are reported to be elevated when the affected patients experience unpredictable or uncontrollable stress. The susceptibility to reactivate intestinal inflammation appears to be related to CD4 lymphocyte activity.

There are a number of animal models developed for studying the underlying interactions between stress and UC. Experimentally induced colitis in rats has been shown to be aggravated and sustained by stress. Animal studies also show that neonatal maternal separation sensitizes adult rats to chemically induced colitislike symptoms in addition to altering the intestinal microflora and disrupting the permeability of the intestine.

Stress Ulcer

All species studied develop gastric mucosal abnormalities in response to severe or chronic stress (see **Ulceration, Gastric**). In humans, the acute gastric ulcer is generally observed after severe physical

trauma such as burns, surgery, sepsis, or injury to the CNS (Cushing's ulcer). The pathological background for these changes is most likely an exacerbation of superficial gastric erosions, caused by a combination of (1) restricted blood supply leading to local ischemia, (2) gastric acid, and (3) changes in gastric motility. There is no evidence that adrenocortical secretions play an aggressive role in this process; instead, the mediating mechanisms appear to be CRH-related sympathetic activation and vagally mediated gastric acid secretion.

Conclusion

Both animal and human studies provide clear evidence that stressors can and do impinge on the GI tract, most clearly at the gastric and intestinal levels. Although historically, many disorders were regarded as multifactorial, with stress as one of many contributory factors, the empirical evidence has been lacking due to methodological difficulties. There is no doubt that stress results from these disorders and that stress may exacerbate symptoms. In the case of the functional disorders, current medical attitudes have been influenced by the patients' visceral hypersensitivity. In the case of ulcer disease, the *Helicobacter* revolution has largely excluded a role for psychological stress factors. The basic physiological and psychobiological data together with comparative studies, however, clearly support the older notion that stress has an important contributory role in the etiology and maintenance of these multifactorial disease states.

See Also the Following Article

Ulceration, Gastric.

Further Reading

- Bennett, E. J., Piesse, C., Palmer, K., et al. (1998). Functional gastrointestinal disorders: psychological, social, and somatic features. *Gut* **42**, 414–420.
- Collins, S. M. (2001). Stress and the gastrointestinal tract IV. Modulation of intestinal inflammation by stress: basic mechanisms and clinical relevance. *American Journal of Physiology* **280**, G315–318.
- Drossman, D. A. (1998). Presidential address: gastrointestinal illness and the biopsychosocial model. *Psychosomatic Medicine* **60**, 258–267.
- Goebell, H., Holtmann, G. and Talley, N. J. (eds.) (1998). *Functional dyspepsia and irritable bowel syndrome: concepts and controversies*. Dordrecht: Kluwer Academic Publishers.
- Hernandez, D. E. and Glavin, G. B. (eds.) (1990). *Special issue on neurobiology of stress ulcers. Annals of the New York Academy of Sciences* **597**.
- Levenstein, S. and Kaplan, G. A. (1998). Socioeconomic status and ulcer: a prospective study of contributory risk factors. *Journal of Clinical Gastroenterology* **26**, 14–17.
- Mayer, E. A., Naliboff, B. and Munakata, J. (2000). The evolving neurobiology of gut feelings. *Progress in Brain Research* **122**, 195–206.
- Medalie, J. H., Stange, K. C., Zyzanski, S. J., et al. (1992). The importance of biopsychosocial factors in the development of duodenal ulcer in a cohort of middle-aged men. *American Journal of Epidemiology* **136**, 1280–1287.
- Milde, A. M., Enger, O. and Murison, R. (2004). The effects of postnatal maternal separation on stress responsivity and experimentally induced colitis in adult rats. *Physiology & Behavior* **81**, 71–84.
- Samuelson, L. C. and Hinkle, K. L. (2003). Insights into the regulation of gastric acid secretion through analysis of genetically engineered mice. *Annual Review of Physiology* **65**, 383–400.
- Sapolsky, R. M. (2004). *Why zebras don't get ulcers* (3rd edn.). New York: Henry Holt and Company.
- Sartor, R. B. and Sandborn, W. J. (2004). *Kirsner's inflammatory bowel diseases*. Edinburgh: Saunders.
- Söderholm, J. D., Yates, D. A., Gareau, M. G., et al. (2002). Neonatal maternal separation predisposes adult rats to colonic barrier dysfunction in response to mild stress. *American Journal of Physiology* **283**, G1257–1263.
- Stam, R., Ekkelenkamp, K., Frankhuijzen, A. C., et al. (2002). Long-lasting changes in central nervous system responsivity to colonic distention after stress in rats. *Gastroenterology* **123**, 1216–1225.
- Thompson, W. G. (1996). *The ulcer story*. New York: Plenum Press.

Gender and Stress

R J Handa and W C J Chung

Colorado State University, Fort Collins, CO, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R J Handa and R F McGivern, volume 2, pp 196–204, © 2000, Elsevier Inc.

Evidence Supporting Gender Differences in Stress Responses

Basic Mechanisms of Sexual Differentiation of Neural Function

Clinical Implications for Gender Differences in Stress Responses

Glossary

<i>5α-Dihydro-testosterone (DHT)</i>	Androgen responsible for the masculinization of the genitalia during development. It is metabolized from testosterone by the enzymatic actions of 5 α -reductase.
<i>α-Fetoprotein</i>	A protein produced by the liver. α -Fetoprotein can bind circulating estrogen and sequester it from the brain.
<i>Activational actions of sex steroids</i>	The actions of estrogens and androgens that occur transiently, in direct response to the presence of hormone. Activational effects can occur at any time during life.
<i>Aromatase</i>	An enzyme belonging to the P450 family of enzymes. Aromatase is primarily responsible for the metabolism of testosterone to estrogen.
<i>Critical period of sexual differentiation</i>	The period of time when the central nervous system is most sensitive to the organizing actions of sex steroids. The timing and length of the critical period can vary depending on the species, parameter, and hormone examined.
<i>Estrogen</i>	The main class of sex steroid hormones found in females. Estrogen is produced predominantly in the ovary; however, it can be produced <i>de novo</i> in other tissues and even in brain tissue in both women and men.
<i>Organizational actions of sex steroids</i>	The actions of estrogen and androgen that result in permanent changes in the structure or function of the brain. Organizational effects result from exposure during development or during a period of high sensitivity to the hormone in question.

Testosterone

The main class of sex steroid hormone found in males. Testosterone is produced predominantly by the testis; however, circulating testosterone is also found in females. Testosterone is metabolized to estrogenic as well as androgenic products.

Evidence Supporting Gender Differences in Stress Responses

Gender Differences in Adrenal Function

The early work of Hans Selye described adrenal gland function and its role in homeostasis. This work also provided the first compelling evidence for potential sex differences in stress responses. In supporting the observations of Selye, many studies showed that the adrenal gland is larger in females than in males, even when corrected for sex differences in body weight. Correspondingly, the rate of hormone secretion by the adrenal gland differs between males and females. Moreover, some studies in rodents reported that females show a greater diurnal fluctuation in corticosterone or cortisol, the main steroid hormone secreted by the adrenal cortex (corticosterone in the mouse and rat, cortisol in the hamster and human; [Figure 1](#)). Similarly, the adrenal response in secreting these glucocorticoid hormones following stressful, or perceived stressful, situations such as exposure to noxious fumes, exposure to a novel environment, immobilization, or a brief electric shock (see [Glucocorticoid Negative Feedback, Glucocorticoids, Effects of Stress on, Glucocorticoids – Adverse Effects on the Nervous System, Hippocampus, Overview](#)) resulted in peak elevations in adrenal glucocorticoid hormone levels that were 1.5–2 times greater in females than in males.

In adult animal models, sex differences in adrenal function appear to be relatively confined to glucocorticoids because gender differences in other adrenal cortex-derived steroid hormones, such as progesterone and androgen, have not been reported in the literature. Androgen and progesterone are secreted predominantly by the gonads at relatively high levels. Therefore, gonad-derived steroids may mask any sex differences in adrenal-derived steroids, which may underlie the absence of these types of studies in the literature.

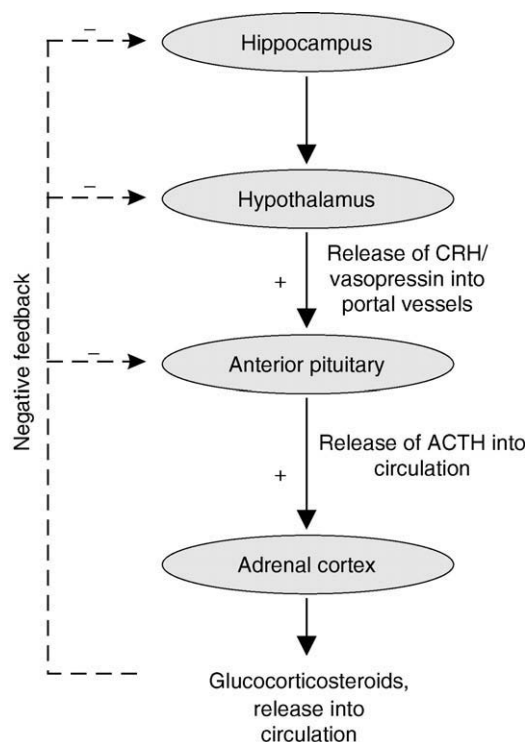


Figure 1 Schematic representation of the hypothalamic-pituitary-adrenal (HPA) axis. The positive feedback portion of the HPA starts at the level of the paraventricular nucleus of the hypothalamus. Neurons residing in this nucleus release corticotropin-releasing hormone (CRH) and vasopressin into the hypophyseal portal vessels to the anterior pituitary to stimulate the synthesis and release of ACTH. This in turn stimulates both synthesis and release of adrenal glucocorticoids into the general circulation to elicit an appropriate stress response. The presence of increased levels of glucocorticosteroids in turn induces an inhibitory effect (i.e., negative feedback) on the HPA axis, which is thought to be mainly mediated through glucocorticoid-responsive neurons in the hippocampus.

Sex Differences in Neuroendocrine Function

Pituitary hormone secretion The sex difference in adrenal function is in part due to sex differences in the neuroendocrine axis that regulates adrenal hormone secretion. This axis has been termed the hypothalamic-pituitary-adrenal (HPA) axis because of the interaction of each of these tissues in regulating one another's secretory activity (Figure 1). Similar to sex differences in the adrenal glucocorticoid release, adrenocorticotrophic hormone (ACTH) is also released in a sexually dimorphic fashion. ACTH is a hormone secreted by the anterior pituitary gland into the general circulation, which in turn stimulates adrenal synthesis and secretion of glucocorticoids. Animal studies also demonstrated that the increases in circulating ACTH levels following physical or psychological stressors are greater and more prolonged in females than in males. An enhanced ACTH response in female

rodents seems to depend on the hypothalamic peptide called corticotropin releasing hormone (CRH), suggesting that the sex difference in ACTH response may, in part, be the result of a sex difference in anterior pituitary responsiveness to CRH.

Hypothalamic function The secretion of ACTH from the anterior pituitary is controlled by hormones that are synthesized in neurons residing in the paraventricular nucleus of the hypothalamus (PVN). These hormones are released into the hypophyseal portal vasculature via the median eminence (Figure 1). The most important releasing factors for ACTH appear to be CRH and vasopressin. Both CRH and vasopressin are potent ACTH secretagogues, and vasopressin potentiates the ACTH-releasing activity of CRH several-fold. Animal studies have shown that the PVN in females contains more CRH neurons than in males. Moreover, females exhibited a more pronounced diurnal variation in hypothalamic CRH immunoreactivity than males. These data are consistent with physiological data demonstrating a greater activation of the HPA axis in females.

Studies have demonstrated sex differences in the release of vasopressin following a stressor. Indeed, vasopressin was released in female rats following immobilization stress, whereas this was not the case in male rats. Although, vasopressin immunoreactivity in the PVN is not sexually dimorphic, the presence of a sex difference in the functional sensitivity of vasopressin cells in the PVN to stressors may help explain sex differences in neuroendocrine responses.

Negative feedback regulation The stimulatory limb of the HPA axis (CRH, vasopressin, and ACTH) is kept in check by continuous feedback inhibition by circulating glucocorticoid hormones. Thus, it has been hypothesized that the intensity of this negative feedback action may be an underlying factor in the sexually dimorphic secretion of adrenal glucocorticoid hormones. The negative feedback glucocorticoids is mediated by two receptor types – mineralocorticoid receptor (MR) and glucocorticoid receptor (GR) – that reside in target neurons in the hypothalamus, hippocampus, and anterior pituitary gland, as well as other brain areas. Both receptors appear to be involved in differing aspects of the negative feedback regulation. Because of their high affinity for glucocorticoids, MRs are thought to regulate basal hormone secretion, whereas GRs, which exhibit a lower affinity for glucocorticoids, are thought to return stress-responsive elevations in glucocorticoids back to baseline. Assuming that the number of receptors accurately reflects the sensitivity of a tissue for a hormone, it is logical to conclude that sex differences in the HPA axis may

result from a sex difference in the number of receptors in a given cell. A greater negative-feedback sensitivity would be indicated by more receptors and consequently result in lower basal and stress-activated hormone secretion. On the other hand, lower levels of receptors would indicate weaker negative feedback and thus result in higher basal and stress-responsive hormone secretion. This hypothesis has been supported by sex differences in the concentration and function of MRs and GRs. Overall, females tend to have lower levels of MRs and GRs in many brain regions, including the hypothalamus and hippocampus, which is consistent with the pattern found in other tissues, such as the thymus and liver. Given similar levels of hormone, GR- or MR-mediated functions (e.g., autoregulation of the receptors) was reported to be less sensitive to glucocorticosteroid modulation in females than in males.

Sex Differences in Behavioral Responses to Stress

Sex differences in behavioral responses to stress have been well-documented in humans and animals. Laboratory animals have provided much of what we know regarding gender differences in response to stress. When placed in a given stressful situation, animals undergo a series of well-described behaviors that have been interpreted to indicate the degree of fear or anxiety that is being experienced. Freezing, defecation, and vocalization are a few good examples of these behaviors. Females appear to be less sensitive to stress-associated novel or aversive environments. For example, when placed into a novel environment or an elevated plus maze, female animals show less freezing and defecation than males. Females also enter a compartment in which they have been previously shocked more readily. In these behavioral paradigms, females also show more activity and exploration, behaviors that have been interpreted as signifying a lesser expression of fear and anxiety. However, when exposed to a variety of chronic stress paradigms, females are found to habituate significantly less than males.

In a well-studied animal model of chronic stress known as learned helplessness, animals are subjected to a repeated inescapable stressor, such as an intermittent tail shock. Subsequently, when allowed to escape, male rats are impaired in their escape performance, whereas female rats do not show this deficit to the same extent. However, when stressed repeatedly over a long period of time, it appears that males adapt more readily (eventually do not show reduced activity), whereas females fail to adapt over the same time period. It appears that this sex difference in adaptability may be dependent on levels of glucocorticoids because the reduction of the poststress glucocorticoid

levels in females to levels attained in males allows the adaptation to repeated restraint. However, these effects are not always observed across studies. One reason may be related findings from studies showing that both strain and stress exposure during early development are important determinants in the expression of stress-related sex differences in behavior in adulthood.

Basic Mechanisms of Sexual Differentiation of Neural Function

Sexual differentiation of the brain occurs as a result of the presence of gonadal steroid hormones acting during either development or adulthood ([Figure 2](#)). The actions of gonadal steroid hormones on the brain have been classified as being either organizational or activational. Organizational actions are permanent changes in brain structure or function that occur during development. Generally, these permanent effects of gonadal steroid hormones are greatest during a defined critical window of opportunity, usually during gestation or early postnatal life. The removal of gonadal steroid hormones after this critical period does not alter the organization or function of the brain appreciably. In contrast, the activational effects of gonadal steroid hormones are those that occur transiently, in direct response to the presence of a gonadal steroid hormone. Indeed, the removal of gonadal steroid hormones from the circulation, through castration or ovariectomy, results in a loss of effect. The activational effects of gonadal steroid hormones can occur throughout life, but are generally thought to have a maximal influence on gender differences during adulthood.

The literature regarding sexual differentiation of the brain suggests that the mammalian brain is an undifferentiated or perhaps neutral state and that, in response to normal circulating levels of testosterone (from the fetal male gonad), it develops toward the male direction. In contrast, if testosterone is low or absent, the brain develops toward the default mode, which is female (see [Figure 2](#)).

Although testosterone or its androgenic metabolite DHT may play a role in the masculinization of the brain, through their intracellular androgen receptor, in rodent models the functional gonadal steroid hormone for the organizational sexual differentiation of the brain is considered to be estrogen ([Figure 2](#)). This steroid hormone is produced intracellularly as a metabolite of testosterone due to the actions of the enzyme aromatase. The defeminizing effects of estrogen on the developing brain are mediated by estrogen receptors (ERs), although the exact role of each of the two forms of ERs, ER α and ER β , is not known.

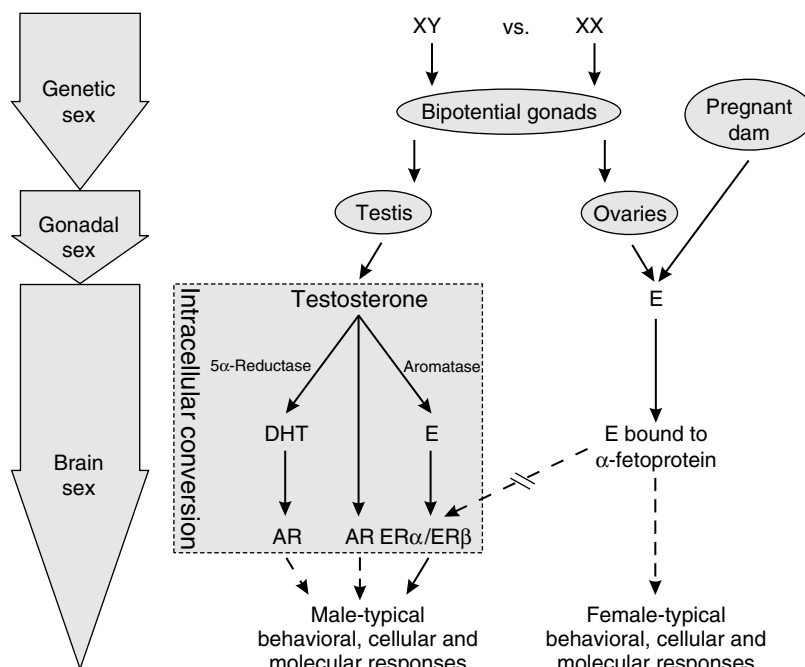


Figure 2 Schematic representation of the gonadal steroid hormones involved in the sexual differentiation of the vertebrate brain during development. According to this scheme, α -fetoprotein sequesters circulating estrogen, thereby preventing the defeminization (or masculinization) of the male or female brain. Therefore, the brain is thought to be initially programmed to develop in the female direction. However, α -fetoprotein is bypassed in the male brain because circulating testosterone is converted intraneuronally into estrogen by the enzyme aromatase, thus enabling estrogen to defeminize (or masculinize) the male brain, which is mediated through estrogen receptors. AR, androgen receptor; DHT, 5 α -dihydrotestosterone; E, estrogen; ER, estrogen receptor.

Circulating estrogen from the mother or the developing ovary of the female is bound to a liver-derived protein called α -fetoprotein, which is present in high levels during perinatal development. Consequently, circulating estrogens are sequestered and cannot defeminize (or masculinize) the developing male or female brain. In contrast, defeminization (or masculinization) by estrogen can occur in the male brain because the estrogen-sequestering abilities of α -fetoprotein are circumvented due to its intracellular origin through the aromatization of testosterone. In rodents, the high levels of α -fetoprotein decrease during the second and third weeks of life, which coincides with increasing levels of unbound estrogens in the circulating, thus allowing the organizational and activational effects of estrogens to occur at puberty and in adulthood.

Organizational Effects of Gonadal Steroid Hormones on Stress Responses

Gonadal steroid hormones may act during early development to permanently organize adult responses to stress. This is best illustrated by examining the effects of perinatal hormone removal and replacement on the development of the adult pattern of behavior in response to stress. An example is the earlier described sex difference in the performance

of the elevated plus maze and open-field test, in which female rats showed more activity and were less anxious than male rats. Indeed, the activity of an adult male rat that was castrated postnatally (a procedure that removes circulating levels of testosterone and its associated metabolites, such as estrogen) is similar to that of a female rat. The castration of male rats up to 21 days after birth prevents the development of the sex difference in activity. In contrast, neither the castration of adult males nor estrogen or testosterone replacement affects the sex difference in the level of activity. Similarly, the treatment of female rats with testosterone or estrogen up through 4 weeks after birth, but not in adulthood, results in a more malelike activity pattern. These studies indicate a critical period of sensitivity for the organizing actions of gonadal steroid hormones; moreover, they also show that sex differences in the activity level do not arise until after puberty.

Few studies have examined the occurrence of similar organizational effects of sex steroids on neuroendocrine responses to stress. The neonatal gonadectomy of male rats produces a more female-like hormonal response to stress in adulthood, but much of this may be due to the absence of hormones in adulthood because adult gonadectomy has a similar effect. On balance, it appears that the predominant effects of gonadal

steroid hormones on neuroendocrine stress responses are activational in nature, with organizational influences contributing a much more modest amount to the overall sex difference.

Activational Effects of Gonadal Steroid Hormones on Stress Responses

The activational effects of gonadal steroid hormones on stress responses can be demonstrated easily by the gonadectomy of males and females in adulthood. The castration of male rats results in an increase in stress-responsive ACTH and corticosterone secretion. Similarly, the ovariectomy of female rats results in a decrease in stress-responsive HPA axis activity. These are gonadal steroid hormone-specific effects because they were reversed by the replacement of gonadal steroid hormones to gonadectomized animals. Moreover, androgen and estrogen have opposing actions in regulating stress-responsive hormone secretion. The treatment of gonadectomized male and female rats with estrogen augments subsequent ACTH and corticosterone responses to stress, whereas treatment with testosterone, or the nonaromatizable androgen, DHT, effectively inhibits HPA axis responses to stress. These studies provide a good mechanistic understanding as to why the stress-responsive hormone secretion is greatest in female rats during proestrus, when estrogen levels are highest.

Testosterone and estrogen can limit or augment the ACTH and corticosterone response to stress at several sites. These include the adrenal gland, anterior pituitary, hypothalamus, and higher brain centers such as the hippocampus. Receptors for androgen and estrogen exist in all these tissues, and the effects of estrogen and androgen on cellular parameters relevant to HPA activation have been demonstrated in all tissues.

At the level of the adrenal gland, estrogen has been shown to increase the adrenal secretory response to ACTH administration. Similarly, some studies showed that in the presence of estrogen the anterior pituitary gland responds to CRH with a greater secretion of ACTH than in the absence of estrogen. Increases in CRH immunoreactivity and mRNA have also been demonstrated in the hypothalamus in response to estrogen treatment. Whether these are direct actions on CRH neurons or mediated transsynaptically is currently in question. However, studies have demonstrated that the promoter sequence of the human CRH gene contains sequences (estrogen response elements) that respond to occupied estrogen receptors by increasing gene transcription, suggesting a direct action on the CRH gene. More recently, studies in rodents and humans showed the presence of ERs in some populations of CRH neurons of the PVN.

Similarly, androgen also appears to act at several sites. Castration increases adrenal enzyme activity and the synthesis of corticosterone by the adrenal gland *in vitro*, suggesting that testosterone acts to inhibit adrenal activity. Although testosterone does not appear to alter the anterior pituitary ACTH response to CRH administration, it does decrease CRH immunoreactivity in the hypothalamus of castrated rats. A direct action of testosterone on CRH synthesis or secretion does not seem to be the case because CRH neurons do not contain androgen receptors; thus upstream sites containing androgen receptors and projecting to CRH cells in the hypothalamus have been implicated, including the medial preoptic area, septum, and bed nucleus of the stria terminalis.

Sex differences in circulating estrogen and testosterone are also responsible for the sex differences in brain GR and MR levels or function. These effects underlie the reported sex differences in the negative feedback modulation of the HPA axis. Estrogen decreases GR and MR levels and their respective mRNAs in some brain areas and interferes with the ability of GR to autoregulate gene expression following changes in circulating glucocorticoids. These effects of estrogen could arise secondarily as a result of estrogen-induced increases in corticosteroid-binding globulin (CBG). In response to estrogen exposure, increased levels of CBG can sequester circulating glucocorticoid hormones from the brain, resulting in a decrease in negative feedback. However, estrogen has also been reported to increase GR in some tissues, such as the pituitary gland, and decrease the efficacy of corticosteroid receptor agonists that are not bound by CBG. Thus, it is likely that the actions of estrogen on the negative feedback are not solely mediated by an increase in CBG.

At the molecular level, interactions between ERs and GRs have been demonstrated with each modulating the functional response of the other. Testosterone also appears to interact with glucocorticoids in the regulation of GRs, but not MRs; however, a clear role for androgen receptors in altering negative feedback mechanisms has yet to be demonstrated.

Clinical Implications for Gender Differences in Stress Responses

Hans Selye's work in the 1940s also introduced the concept that stress accelerates the onset of disease through pathophysiological changes in endocrine and autonomic regulatory systems. Stress alone does not cause disease, but it is hypothesized to increase an individual's susceptibility to disease. In the past quarter century, this concept has been supported by results from numerous clinical and epidemiological studies

related to cardiovascular disease, affective disorders, gastrointestinal disorders, diabetes, and physiological disorders associated with normal aging.

The mechanisms linking stress to disease states in humans are not fully understood, but hormonal changes, induced by the stress response, are recognized as a major influence. Glucocorticoids have a primary role but one that is modulated by gonadal steroid hormones. Androgen receptors, ERs, and GRs are found in many tissues that are stress-responsive. Some of these include the heart muscle, blood vessels, and immune tissues such as the spleen, thymus, and bone marrow. Gonadal steroid hormone receptors are also found in brain regions that mediate emotional and autonomic responses to stress, such as the hippocampus, hypothalamus, and amygdala. Thus gender differences in stress-related disease susceptibility appear to involve an interaction between glucocorticoid and gonadal hormones.

The role of stress in psychological and physiological disorders is complex. Animal studies have demonstrated that adult stress responsiveness is influenced developmentally by both genetics and the early environment. This bidirectional influence of both factors makes it difficult to predict the stress-related susceptibility to disease in individuals. However, physiological gender differences are still evident in humans regardless of early environment.

Many of these differences parallel sex differences in animals. For instance, the unstressed basal heart rate in women is greater than in men. Women also exhibit a larger exercise-induced increase in heart rate. However, in response to anger and psychological stress, men exhibit greater emotional and cardiovascular lability with respect to both heart rate and blood pressure. This lability, resulting in larger increases in systolic pressure, is thought to contribute significantly to the greater risk in men for cardiovascular morbidity. Protection from heart disease in women is also related to estrogen, which provides protection against arteriosclerosis until menopause. The anti-atherosclerotic effects of estrogen also appear to involve the presence of ERs in cardiovascular tissue.

Physical or psychologically stressful stimuli also produce a greater stimulation of the HPA axis in women than in men, similar to the stress effects observed in other mammals. Administering CRH to women induces a greater release of ACTH and a prolonged elevation of cortisol than administering it to men, indicating that these gender differences probably involve processes directly under the regulation of the central nervous system. However, unlike animals, gender differences in the human HPA axis responsiveness to psychological stressors appear to be

much more dependent on the type of psychological stress and the environmental circumstances in which it occurs. Studies suggest that psychological factors associated with social stress, such as taking care of an elderly parent, have more impact on HPA reactivity in women than in men, perhaps placing women's health at greater risk under such circumstances.

A primary factor underlying gender differences in the HPA response to stress appears to be higher estrogen levels in women. During the normal menstrual cycle, HPA activity is enhanced during the follicular phase, when estrogen levels peak, and is reduced during the luteal phase, when progesterone levels are elevated. The removal of the ovaries prior to menopause reduces the HPA sensitivity to CRH stimulation. The effect of estrogen on HPA reactivity is not gender-specific because normal men who were administered a short-term regimen of estrogen also exhibited increased HPA activity. One possible explanation for such effects of estrogen is provided by observations that estrogen therapy increases circulating levels of CBG in women. Although not yet shown to be the case, such increases in CBG, and the resulting decreases in free cortisol, could play a role in the elevated HPA activity in response to stress reported in women (due to a decrease in cortisol bioavailability). Similarly, it remains to be determined whether men and women are equally sensitive to the influence of estrogen on the HPA and whether organizational influences of gonadal steroids confer a differential sensitivity to the later hormonal modulation of HPA activity.

The incidence of autoimmune disease is also sexually dimorphic, and the severity and onset are exacerbated by life stress. The overall incidence in women is approximately 2.5-fold greater than in men. In contrast to the protective role estrogen plays in cardiovascular disease, it appears to be a major contributory factor to the gender difference observed in autoimmune diseases such as multiple sclerosis, rheumatoid arthritis, and lupus. Estrogen can regulate immune responses through ERs present on thymocytes, macrophages, and endothelial cells, all of which also express GRs. The interaction between estrogen and glucocorticoids on immunomodulation is a rapidly expanding area of research interest and one that is likely to underlie the gender difference in autoimmune disease susceptibility.

Numerous clinical studies have reported a greater incidence of depression in women. The reasons underlying this gender difference are controversial, but both hormonal status and stress may be involved. Women have been reported to be more susceptible to depression during periods of fluctuations in estrogen and progesterone levels. However, although such

changes have often been cited as causes of mood disorders in women, the literature does not provide strong support for a direct causal relationship between gonadal hormone status and depression or anxiety disorders. An additional mediating factor may be life stress. The onset of clinical depression is well known to be exacerbated by life stress, and data suggest that a link to a gender bias in mood disorders involves an increased susceptibility in some women to the combined effects of stress and hormonal status. Thus, fluctuations in estrogen and progesterone during the menstrual cycle have been proposed to confer a susceptibility to depression in women that can be exacerbated by life stress.

A gender difference is also observed in the incidence of posttraumatic stress disorder (PTSD), which is a stress-induced pathological condition that is clinically distinct from depression or other affective disorders. Long-term physiological changes are observed in PTSD patients that involve increased adrenergic tone and alterations in HPA function. Following exposure to prolonged psychological trauma, women exhibit PTSD symptoms twice as frequently as men, even when the type of trauma is equivalent. The organizational effects of hormones may also play a role in gender differences in the incidence of PTSD. This stems from findings showing that experiencing psychologically traumatizing events prior to puberty greatly increases the susceptibility to PTSD in adulthood in females compared to males. This increased susceptibility appears to relate specifically to experiencing childhood psychological trauma because accidents or injury during this period do not increase the risk for PTSD in either gender when later psychological trauma is experienced.

In summary, current research studying the interrelationships among gender, stress, and pathophysiology strongly implicate a role for gonadal hormones in predicting gender differences related to disease or psychopathology. However, a definitive understanding of the role of gender in the elicitation of stress-induced pathophysiology awaits a better understanding of the complex relationship among the environment, steroid hormones, and other stress-induced biochemical factors.

See Also the Following Articles

Adrenal Cortex; Androgen Action; Autoimmunity; Depression Models; Estrogen; Glucocorticoid Negative Feedback; Glucocorticoids, Effects of Stress on; Glucocorticoids – Adverse Effects on the Nervous System; Hippocampus, Overview; Learned Helplessness; Menopause and Stress; Selye, Hans.

Further Reading

- Beatty, W. W. (1979). Gonadal hormones and sex differences in nonreproductive behaviors in rodents: organizational and activational influences. *Hormones and Behavior* 12, 112–163.
- Breslau, N., Glenn, C., Andreski, P., et al. (1997). Sex differences in posttraumatic stress disorder. *Archives of General Psychiatry* 54, 1044–1048.
- Chrousos, G. P., Torpy, D. J. and Gold, P. W. (1998). Interactions between the hypothalamic-pituitary-adrenal axis and the female reproductive system: clinical implications. *Annals of Internal Medicine* 129, 229.
- Gorski, R. A. (1985). The 13th J. A. F. Stevenson memorial lecture: sexual differentiation of the brain: possible mechanisms and implications. *Canadian Journal of Physiology and Pharmacology* 63, 577–594.
- Handa, R. J., Burgess, L. H., Kerr, J. E., et al. (1994). Gonadal steroid hormone receptors and sex differences in the hypothalamo-pituitary-adrenal axis. *Hormones and Behavior* 28, 464–476.
- Handa, R. J., Price, R. H., Jr. and Wilson, M. E. (1998). Steroid hormone receptors and the assessment of feedback sensitivity. In: VandeKar, L. (ed.) *Methods in neuroendocrinology: the cellular and molecular neuropharmacology series*, pp. 49–68. Boca Raton, FL: CRC Press.
- Seeman, M. V. (1997). Psychopathology in women and men: focus on female hormones. *American Journal of Psychiatry* 154, 1641–1647.
- Turner, B. B. (1997). Influence of gonadal steroids on brain corticosteroid receptors: a minireview. *Neurochemical Research* 22, 1375–1385.
- Vogele, C., Jarvis, A. and Cheeseman, K. (1997). Anger suppression, reactivity and hypertension risk: gender makes a difference. *Annals of Behavioral Medicine* 19, 61–69.
- Wilder, R. L. (1995). Neuroendocrine-immune system interactions and autoimmunity. *Annual Review of Immunology* 13, 307–338.

Gene–Environment Interactions in Early Development

I S P Davis and R Plomin

Institute of Psychiatry, King's College London,
London, UK

© 2007 Elsevier Inc. All rights reserved.

Quantitative Genetics

Molecular Genetics

Behavioral Genomics and GE Interplay

Future QTL Research on GE Interplay

Glossary

<i>DNA pooling</i>	A molecular genetic technique that combines DNA from individuals in order to estimate allele frequencies for the group of individuals.
<i>Family chaos</i>	A measure of the family environment derived from the Confusion, Hubbub, and Order Scale.
<i>Genotype–environment correlation (rGE)</i>	Genetic influence on exposure to environments.
<i>Genotype–environment interaction (GxE)</i>	Genetic sensitivity to environments.
<i>Microarray</i>	A technology that allows the genotyping of hundreds of thousands of polymorphisms simultaneously.
<i>Quantitative genetics</i>	The statistical study of the genetic and environmental origins of individual differences in complex traits, using information about the relatedness of individuals; in contrast to molecular genetics, which attempts to identify specific genes.
<i>Quantitative trait locus (QTL)</i>	A genetic polymorphism that plays a role in the etiology of a trait influenced by many such loci.
<i>Single nucleotide polymorphism (SNP)</i>	A genetic polymorphism involving the substitution of a single letter (nucleotide base pair) of the four-letter genetic alphabet.

The interplay between nature and nurture has long been characterized as adversarial; the truth is much more complex, and more interesting. Although there are many ways of thinking about the issue, this article considers genotype–environment (GE) interaction and correlation from a quantitative genetic

perspective, where GE interaction refers to genetic sensitivity to environments and GE correlation refers to genetic influence on exposure to environments. Diathesis-stress theories predict that individuals' sensitivity to stressful events and access to social support networks depends on their genotype. Although evidence for such interactions is – as yet – inconclusive, a more convincing case can be made for genetic influence on environmental measures: gene–environment correlations. The ability to detect gene–environment interplay has traditionally relied on the use of specific environmental measures in genetically sensitive designs such as twin or adoption studies. This ability is greatly enhanced by incorporating genotyping for specific genes: new methods for doing this are likely to see future stress research making use of genetic measures as a matter of course.

Quantitative Genetics

Over the past two decades, there has been an increasing tendency to include measures of the environment in quantitative genetic research; without measures of the environment, quantitative genetic designs such as twin and adoption studies cannot contribute much to our understanding of GE interaction and correlation. This section begins with quantitative genetic research on GE interaction and correlation that incorporates measures of the environment and then turns to molecular genetic research that incorporates information from DNA. The effects of stressful environments on learning ability and disability are used as an example of how molecular genetics can contribute to our understanding of gene–environment interplay.

GE Interaction

The great advantage of research with animal models is the ability to manipulate both genes and environment experimentally so that main effects and interaction are assessed directly. That is, genetically different animals such as selected low and high lines or inbred strains can be studied in different environments. Animal models are powerful, but they have been hard-pushed to find GE interaction, especially for learning. The classic example in psychology textbooks is Cooper and Zubek's study of maze-bright and maze-dull rats reared in enriched or deprived environments. However, although it is highly cited, its conclusions are doubtful: Henderson studied GE

interaction systematically using inbred strains of mice and found few such effects.

Adoption designs are similar to animal designs, except that they are only quasi-experimental in their separation of genes and environment (Figure 1). Parents usually share both genes and environment with their offspring, but adoption separates nature and nurture by creating a distinction between genetic parents whose children are adopted away and environmental parents who adopt these children. Using the phenotype of biological parents as an indicator of genetic risk and the phenotype of adoptive parents as an indicator of environmental risk, we can find genetic effects without environmental effects (Figure 1a) and environmental effects without genetic effects (Figure 1b). Studies usually reveal both genetic and environmental effects (Figure 1c). The adoption design can also identify GE interaction (Figure 1d), in which genetic effects depend on environmental effects or vice versa. However, as with animal studies, it has been difficult to find GE interaction, especially for cognitive abilities. Twin studies are much more limited in terms of finding GE interaction because they are essentially constrained to detecting differences in heritability as a function of the environment.

For example, in the Twins Early Development Study (TEDS) of a large twin sample assessed at 2, 3, 4, and 7 years, GE interaction was looked for,

but not much evidence for it was found using quantitative genetic methods. Despite this, one promising example of GE interaction is that heritability of verbal ability is significantly higher (about 0.60) in high chaos families and lower (about 0.30) in low chaos families. However, the twin design lacks power to detect even this limited type of GE interaction: about 1000 pairs of each type of twin are needed to detect a heritability difference of 60% versus 40%. As discussed later, the lack of quantitative genetic evidence for GE interaction does not preclude the likelihood that molecular genetic research will find interactions between specific genes and specific environmental stressors.

GE Correlation

In spite of the current interest in GE interaction, it is likely that GE correlation, the nature of nurture, will eventually be seen as more important. Animal models are powerful for investigating GE interaction, but laboratory research has rarely considered GE correlation because it requires a menu of environments from which individuals can choose. The first step in addressing GE correlation is to investigate the extent to which there is genetic influence on environmental measures. This can be done by treating environmental measures as dependent measures in quantitative genetic analyses. Dozens of such twin and adoption

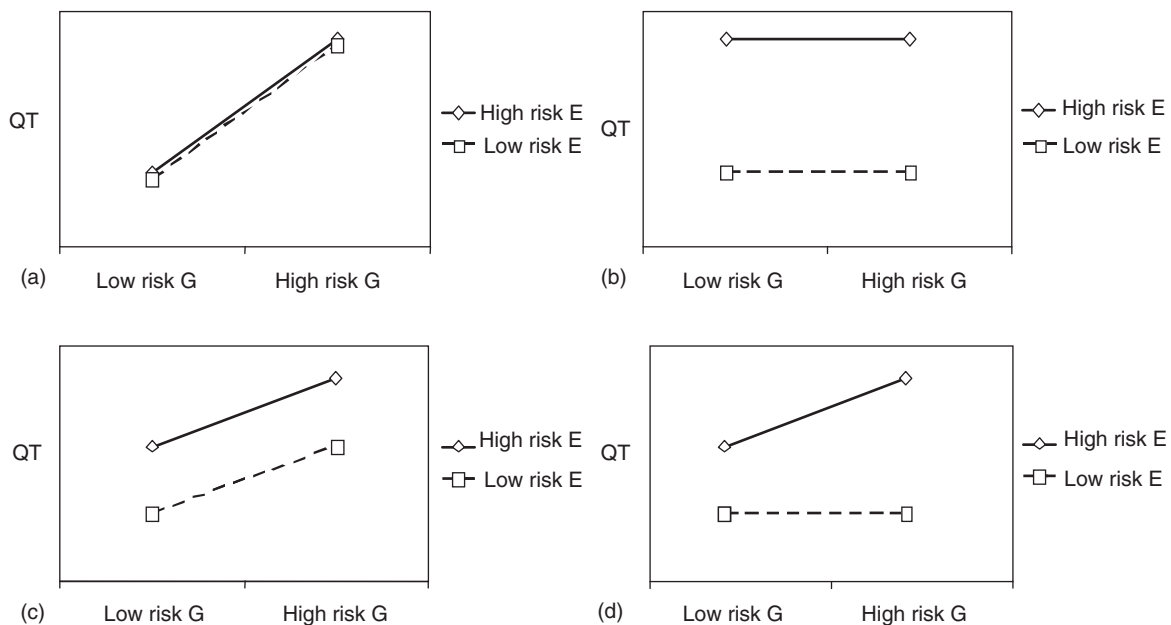


Figure 1 G, E, and GxE in an adoption study. G refers to low or high genetic risk, as indexed by the phenotype (quantitative trait [QT]) of the biological parents. E refers to the environment of the adoptive home. a, A main effect of G, but no effect of E on the phenotype. b, A main effect of E, but no effect of G. Very often, there is a main effect of both G and E, as illustrated in c. d, A main effect of both G and E, and a GxE interaction of the diathesis stress type, in which environmental risk (stress) has a disproportionately large effect on individuals at genetic risk (diathesis).

studies consistently find genetic influence on environmental measures: nearly all measures of family environment show genetic influence. Most studies that have addressed this question find genetic influence on environmental measures outside the family as well, such as peer groups, life events, and social support, which are of particular relevance to stress.

How can environmental stressors show genetic influence when environments do not have DNA? The answer is that our measures of stress are not independent of us – genetic influence on our behavior is reflected in our experiences. If genetics affects both environmental measures and behavioral measures, then the correlation between environmental measures and behavioral outcomes may be genetically mediated. Life events such as illness and injury have been shown to have a heritability of around 40%, implying that they do not happen capriciously to individuals: they happen, or are perceived to happen, more often to the genetically predisposed. As expected, those life events that the individual is involved in or has control over show more substantial heritability than those that happen to others. Recent research suggests that minor, chronic, daily stressors may be more important in determining outcomes than major life events. New studies are expected to show that these also show genetic influence. In a similar way, social support networks show heritability: studies have demonstrated that although the extensiveness of the network is not heritable, the quality of the relationships is.

The method underlying all these studies is to use multivariate genetic analysis to analyze genetic and environmental sources of covariance between environmental stressors and behavioral outcomes. Studies of this sort consistently show some genetic mediation of the association. In this sense, quantitative genetic research provides much more evidence for GE correlation than for GE interaction. GE correlation suggests an important model of environments that goes beyond the usual passive view in which the environment (and, by extension, stress) is what happens to us. Instead, environments are active: we select, modify, construct and, afterward, reconstruct our experiences partially based on our genetic propensities. From a molecular genetic point of view, GE correlation means that we ought to be able to find specific genes associated with environmental measures.

In summary, quantitative genetic research does not find much evidence for GE interaction, but it finds considerable evidence for GE correlation. This section has not gone into great detail because quantitative genetic results on GE interaction and correlation are not very relevant for future research in this area, which will use specific genes as well as specific

environments. The rest of this article considers molecular genetic approaches.

Molecular Genetics

So far, evidence of the existence of interplay between genes and the environment has been considered. Research on GE interaction and correlation will be much more powerful when we are able to include specific genes as well as specific environments, that is, when the G in GE refers to genes in molecular genetic analyses rather than genotypes in quantitative genetic analyses. There have been several replicated linkages and associations, but most people would agree that progress has been slower than expected in terms of finding genes for complex behavioral traits. The problem is that we are looking for needles in a haystack, but we have only had the power to detect very large needles. Very rarely, the effects of a single gene can account for all the variance in a trait. This is the case with phenylketonuria (PKU), in which a mutation in both copies of a single gene results in the inability to metabolize the amino acid phenylalanine. This causes a buildup of toxic products that damage the developing brain. The environmental intervention of a low phenylalanine diet has the effect of preventing mental impairment in children with PKU but has no effect on other children: a dramatic example of a GE interaction.

Despite the slow progress in identifying quantitative trait loci (QTLs) associated with common disorders and complex traits, some evidence has been found for GE interaction using candidate QTLs that affect genetic sensitivity to environments. For example, there are reports that the influence of maltreatment in early life on the development of adult antisocial problems is moderated by a functional polymorphism in the promoter region of the *MAOA* (monoamine oxidase A) gene. Likewise, the influence of stressful life events on depression may be moderated by a functional polymorphism in the promoter region of the *5-HTT* (5-hydroxytryptophan [serotonin] transporter) gene. These examples of GE interaction emerged despite the absence of a main effect for these two genes.

For the vast majority of complex traits, the largest QTL effect is unlikely to account for more than 1% of the variance. Very large samples are needed to attain power to detect such small effects, and even larger samples are needed to detect GE interaction. Rather than just examining a few candidate genes, the trend in molecular genetic research is toward systematic genome-wide scans that require hundreds of thousands of DNA markers. The costs for such genome-wide association scans on very large samples are substantial.

One new method that goes some way to solving this problem is genotyping of pooled DNA on microarrays, a method called SNP-MaP (single nucleotide polymorphism microarrays and pooling).

The first application of SNP-MaP involved mild mental impairment (MMI), which can be seen as the low end of general cognitive ability (g). g was assessed in 10 000 children at 7 years, and two independent SNP-MaP studies using a microarray that assessed 10 000 SNPs (the 10K GeneChip) compared high and low g groups. Five SNPs were identified that remained significantly associated with g when individually genotyped for 5000 children. On average, each SNP in the set accounts for just 0.2% of the variance in g . However, as the effects of the SNPs are uncorrelated and additive, it is possible to combine their effects to create a composite SNP set using an additive model in which each SNP is scored 0, 1, or 2, depending on the number of high- g -associated alleles. The SNP set accounts for nearly 1% of the variance of g scores at 7 years. As shown in Figure 2, SNP set scores are normally distributed and almost linearly related to g . The difference between those scoring low on the SNP set and those scoring high is about one-third of a standard deviation, or 5 IQ points.

Repeating the analysis using the newly available 500K microarray should find many more SNPs accounting for much more of the variance. It will then be possible to use SNP set scores as genetic risk indicators in research into environmental stressors and to select genotypically for extreme SNP set groups to use

in paradigms such as neuroimaging that cannot study large samples.

Behavioral Genomics and GE Interplay

Once genes or QTL sets associated with behavioral dimensions or disorders are found, the next phase of research will investigate the pathways between genes and behavior via the brain, which will represent a paradigm shift in research on GE interplay. The term functional genomics, which has been used to describe this direction for research, usually refers to bottom-up research that begins with understanding the gene product in a cell and works up toward the brain and behavior. However, other levels of analysis can also be useful in understanding pathways between genes and behavior: behavioral genomics refers to a complementary approach that begins with behavior of the individual and works down toward the brain. For example, we can investigate how genetic effects unfold during development or ask multivariate questions about heterogeneity and comorbidity. As the present concern is GE interaction and correlation, the focus will be on behavioral genomic analyses of GE interplay.

To return to the earlier example, the 5-SNP set associated with g at 7 years was used in behavioral genomic analyses of GE interaction and correlation. The environmental measures employed in these analyses included distal environmental measures such as parental occupation or education and proximal measures such as family chaos and discipline in early

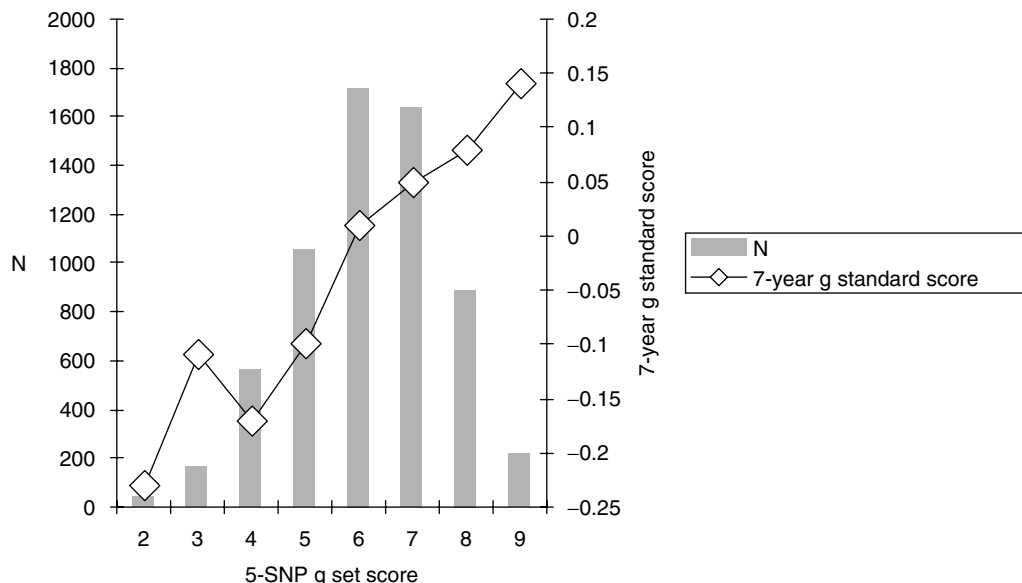


Figure 2 Distribution of 5-SNP g set and association between 5-SNP set and 7-year g . The distribution of scores for the SNP set (derived on the basis of higher scores for high- g -associated alleles) is normal (gray bars). The line represents the means of the 5-SNP set score plotted against means of the g composite score at 7 years.

childhood in those homes. Although GE interaction can be examined in several different ways, the extent to which environmental stressors moderate the SNP set association with *g* was investigated. Using family chaos as an example, the top panel (a) of Figure 3 shows a main effect for family chaos (higher *g* scores

for children in families with less chaos) but no statistically significant interaction using linear and nonlinear continuous analyses.

The other three environmental variables shown in Figure 3 yielded significant but modest GE interaction. For harsh parental discipline (Figure 3b),

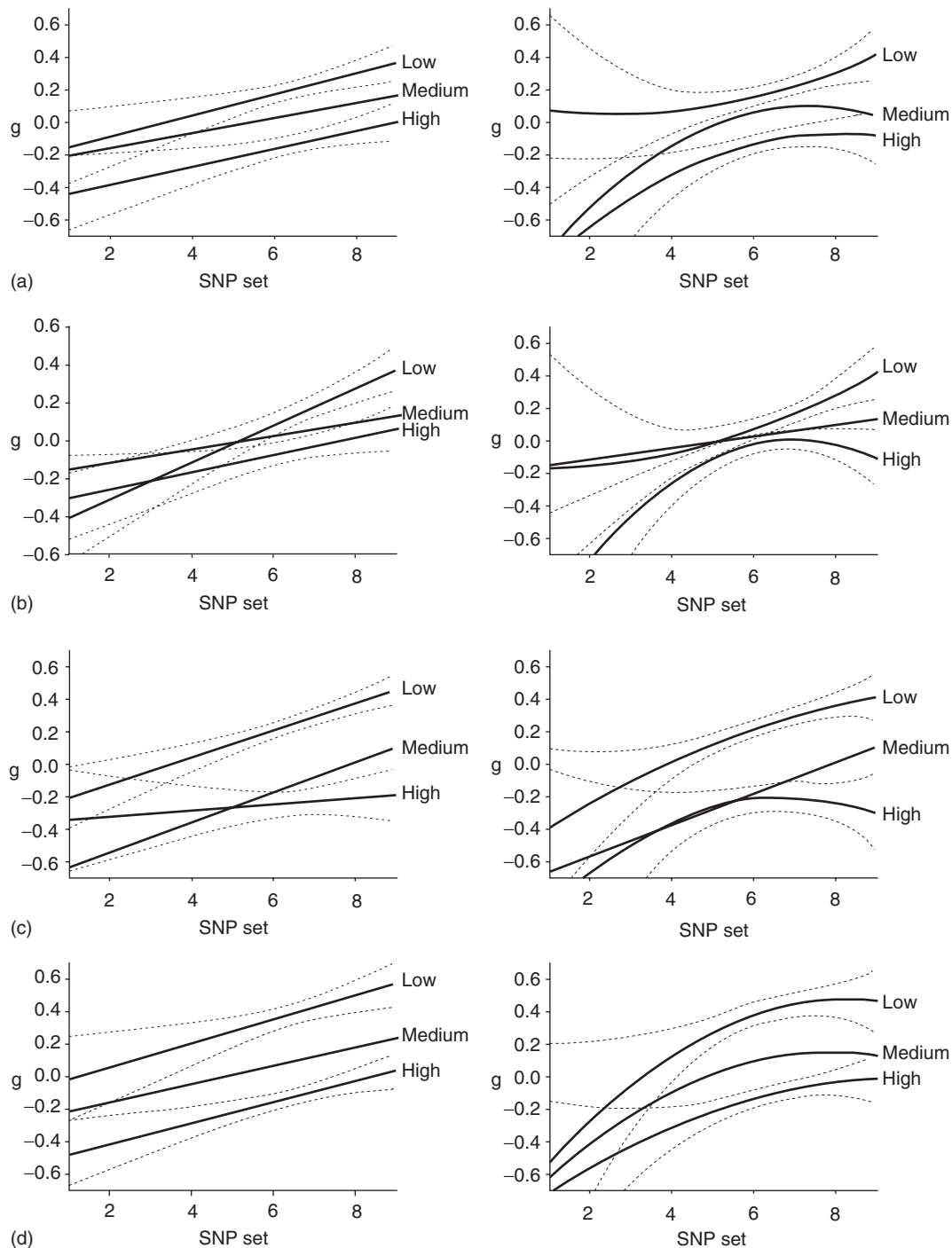


Figure 3 *g* scores as linear and nonlinear functions of SNP set scores and level (low, medium, and high) of environmental risk for chaos (a), harsh parental discipline (b), father's occupational class (c), and mother's highest qualification (d). Ninety-five percent confidence intervals (dotted lines) are shown for low and high levels of environmental risk.

environmental effects are greater at both the high and the low SNP set scores. For high SNP set scores, children with good environments (low harsh discipline) have higher than expected g scores, whereas children with low SNP set scores and harsh parental discipline have lower than expected g scores. In other words, this GE interaction suggests that children with a genetic propensity toward high g profit disproportionately from a good environment, and children with a genetic propensity toward low g suffer disproportionately from a bad environment. A similar sort of GE interaction was found for father occupation (Figure 3c): environmental effects were greater at both the high and the low genetic extremes.

We also found a significant GE interaction for the fourth environmental measure, maternal education (Figure 3d). However, this GE interaction represents a different sort of interaction, one that is compatible with recent findings from quantitative genetics suggesting that genetic effects on g are weaker for families with lower socioeconomic status. The result for maternal education is compatible with this finding: the association between SNP set and g is weakest for families with the least-educated mothers.

What about GE correlation? As noted earlier, GE correlation refers to the correlation between genes and environment. In contrast to GE interaction, which concerns environmental moderation of the association between the SNP set and g, GE correlation focuses on the extent to which the environment mediates the association between the SNP set and g. Such mediation requires that the environmental measure correlate with the SNP set and also correlate with g. Because the environmental measures were chosen for their correlation with g, the simplest GE correlation analysis was conducted, which was merely to correlate the environmental measures with SNP set scores, rather than conducting more formal mediation analyses. Both proximal measures of the family environment (harsh parental discipline and family chaos) were correlated significantly with the SNP set for g, suggesting GE correlation. It is interesting that the distal measures of family environment (father's occupation and mother's education) showed no GE correlation. Of the three types of GE correlation (passive, evocative, and active), finding GE correlation for distal environmental variables would be most easily characterized as passive GE correlation: the children's family environments would be correlated with their genetic propensities because their parents share the same genes. In contrast, these analyses suggest that GE correlation is stronger for proximal environments, implying evocative (reactive) GE correlation: children evoke reactions based on their genetic propensities.

In summary, there is evidence for both GE interaction and GE correlation in early development. One final caveat is that GE interactions may result from GE correlations under certain circumstances; true GE interaction should remain significant after controlling for GE correlation.

Future QTL Research on GE Interplay

In addition to using QTLs identified on the basis of their main effect in order to investigate GE interaction and correlation, future research on GE interplay will profit from identifying QTLs that contribute to GE interaction and correlation regardless of their main effect. In the previous example, SNPs were identified based on their association with g (regardless of environment) by comparing SNP allele frequencies for low versus high g groups. Figure 4 shows a simple two-by-two interaction design in which cases are selected based on low versus high environmental stress as well as on low versus high g. Using DNA pooled separately for the four groups, QTL associations representing the main effect of the QTL on g regardless of environment stressors can be detected by comparing groups 1 and 2 versus groups 3 and 4. The important advantage of this design is that GE correlation and GE interaction can also be detected independent of their contribution to the main effect. GE correlation can be assessed as the main effect of the environment, comparing groups 1 and 3 versus groups 2 and 4, thus identifying QTLs that are associated with low versus high environmental stress. Finally, GE interaction can be detected by comparing the diagonals (groups 1 and 4 versus groups 2 and 3).

The future of research on genotype-environment interaction and correlation looks bright, especially as the G in GE increasingly comes to represent genes in molecular genetic analyses rather than genotypes in quantitative genetic analyses.

Low g	1	2
	3	4
High g		
		Low E High E

Figure 4 Identifying QTLs on the basis of GE interaction and correlation. Using four, rather than two, DNA pools will facilitate these studies. For example, comparing SNP allele frequencies for pools 1 and 2 versus pools 3 and 4 tests for a main effect for g. Comparing pools 1 and 3 versus 2 and 4 reveals GE correlation, a QTL association with E. Comparing pools 1 and 4 versus 2 and 3 assesses GE interaction, a QTL association for g that depends on E.

See Also the Following Articles

Animal Models (Nonprimate) for Human Stress; Antisocial Disorders; Depression and Manic-Depressive Illness; Dopamine, Central; Economic Factors and Stress; Education Levels and Stress; Environmental Factors; Familial Patterns of Stress; Genetic Factors and Stress; Genetic Predispositions to Stressful Conditions; Learning and Memory, Effects of Stress on; Life Events Scale; Major Depressive Disorder; Serotonin; Social Status and Stress; Social Support.

Further Reading

- Asbury, K., Wachs, T. and Plomin, R. (2005). Environmental moderators of genetic influence on verbal and nonverbal abilities in early childhood. *Intelligence* 33, 643–661.
- Butcher, L. M., Meaburn, E., Liu, L., et al. (2004). Genotyping pooled DNA on microarrays: A systematic genome screen of thousands of SNPs in large samples to detect QTLs for complex traits. *Behavior Genetics* 34, 549–555.
- Butcher, L. M., Meaburn, E., Knight, J., et al. (2005). SNPs, microarrays, and pooled DNA: identification of four loci associated with mild mental impairment in a sample of 6000 children. *Human Molecular Genetics* 14, 1315–1325.
- Caspi, A., McClay, J., Moffitt, T. E., et al. (2002). Role of genotype in the cycle of violence in maltreated children. *Science* 297, 851–854.
- Caspi, A., Sugden, K., Moffitt, T. E., et al. (2003). Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science* 301, 386–389.
- Cooper, R. M. and Zubek, J. P. (1958). Effects of enriched and restricted early environments on the learning ability of bright and dull rats. *Canadian Journal of Psychology* 12, 159–164.
- Harlaar, N., Butcher, L., Meaburn, E., et al. (2005). A behavioural genomic analysis of DNA markers associated with general cognitive ability in 7-year-olds. *Journal of Child Psychology & Psychiatry* 46, 1097–1107.
- Henderson, N. D. (1972). Relative effects of early rearing environment on discrimination learning in housemice. *Journal of Comparative and Psychological Psychology* 72, 505–511.
- Meaburn, E., Butcher, L. M., Liu, L., et al. (2005). Genotyping DNA pools on microarrays: tackling the QTL problem of large samples and large numbers of SNPs. *BMC Genomics* 6, 52.
- Plomin, R. (1977). Genotype–environment interaction and correlation in the analysis of human behavior. *Behavior Genetics* 7, 83.
- Plomin, R. (1994). *Genetics and experience: the interplay between nature and nurture*. Thousand Oaks, CA: Sage Publications Inc.
- Plomin, R. and Crabbe, J. C. (2000). DNA. *Psychological Bulletin* 126, 806–828.
- Plomin, R., DeFries, J. C., McClearn, G. E., et al. (2001). *Behavioral genetics* (4th edn.). New York: Worth Publishers.
- Rowe, D. (2003). Assessing genotype–environment interactions and correlations in the postgenomic era. In: Plomin, R., DeFries, J. C., Craig, I. W. & McGuffin, P. (eds.) *Behavioral genetics in the postgenomic era*, pp. 71–86. Washington, D.C.: American Psychological Association.

General Adaption Syndrome (GAS) See: Alarm Phase and General Adaptation Syndrome; Selye, Hans.

Genetic Factors and Stress

C A Koch and C A Stratakis

National Institute of Child Health and Human Development, Bethesda, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 205–211, © 2000, Elsevier Inc.

Genes Regulating the Stress System
Stress Effects in Prokaryotes and Eukaryotes
Summary

Glossary

- Allele** Alternative forms of a gene at a given locus.
- Apoptosis** Programmed cell death inducible by DNA damage that exceeds the capacity of repair mechanisms.
- Chaperone** Any protein that binds to an unfolded or partially folded target protein, thus preventing misfolding, aggregation, and degradation of the target protein and facilitating its proper folding.

<i>DNA polymerase</i>	The enzyme responsible for replication of DNA, which is accomplished by using each complementary strand of the DNA double helix as a template for the synthesis of a new strand.
<i>Epistasis</i>	The situation in which expression of one gene obscures the phenotypic effects of another gene.
<i>Fitness</i>	A measure of fertility and therefore of the contribution to the gene pool of the succeeding generation; for a given genotype, the number of offspring surviving to reproductive age is relative to the number of wild-type genotypes.
<i>Genotype</i>	The genetic constitution of an individual or the alleles at specific genetic loci.
<i>Heavy chain-binding protein (bip)</i>	A protein discovered by its interaction with the heavy chain of IgG antibodies in the endoplasmic reticulum, keeping proteins unfolded and preventing misfolding.
<i>Homology</i>	Similarities between newly derived sequences and previously determined sequences.
<i>Intron</i>	A segment of a gene that is initially transcribed into RNA but is then removed from the primary transcript by splicing together the exon sequences on either side of it; intronic sequences are not found in mature mRNA.
<i>Phenotype</i>	The observed result of the interaction of the genotype with environmental factors.
<i>Pleiotropy</i>	The diverse effects of a single gene or gene pair on several organ systems and functions.
<i>RNA primer</i>	A short RNA sequence that is paired with one strand of DNA and provides a free 3' OH end at which a DNA polymerase starts the synthesis of a deoxyribonucleotide chain.

Stress is the response of the cell or the organism to challengers of homeostasis (a dynamic state of balance between the stress response and the stressors). Homeostasis is essential for the preservation of life. In humans and other higher organisms, components of the stress system include various neurotransmitters and hormones, which in most species constitute the effectors of the hypothalamic-pituitary-adrenal (HPA) axis. Thus, the cholinergic and catecholaminergic systems, corticotropin releasing hormone (CRH), arginine vasopressin, adrenocorticotrophic hormone (ACTH), and the glucocorticosteroids are all essential parts of the stress system in higher organisms.

At the cellular level, the major players are different but still analogous to the effectors of the HPA axis: the DNA repair system, the heat shock proteins (hsp), and the products of many housekeeping genes that

aim at maintaining homeostasis. In order to find genes that have a role in the regulation of the stress system in humans, it is essential to look at the effects of stress at all levels of development. This article reviews the effects of stress on various systems and notes the role of genetics in each one of them.

Genes Regulating the Stress System

Genetic loading can be described as follows. An individual with large innate susceptibility needs only a minor degree of internal or external stressors to reach the threshold for symptom expression and vice versa. The prevalence of psychosomatic symptoms among members of the same family suggests the existence of a genetic predisposition. When the fingerprints of 155 adult patients with psychosomatic abdominal pain since childhood were examined for congenital markers, digital arches were found in 64% of patients, compared to 10% of controls without abdominal pain. Genetic factors play an important role in the pathogenesis of psychiatric illnesses. The location of potential genes of interest for manic-depressive/bipolar disorder, a state of profound stress, has been summarized by Wahlstrom.

Role of the Corticotropin Releasing Hormone Gene and Other Components of the Hypothalamic-Pituitary-Adrenal Axis

CRH plays a key role in the regulation of the stress response with the CRH receptor type-1 (CRH-R1) as mediator. CRH-R1 is highly expressed in the anterior pituitary, neocortex, hippocampus, amygdala, and cerebellum. The activation of this receptor stimulates adenylate cyclase. CRH-R1-deficient mice show an impaired stress response as indicated by the absence of increased ACTH and corticosterone levels following stress.

Although acute and long-standing internal or external conflicts/stressors predispose to depression, a state associated with stress and elevated glucocorticosteroid levels, the CRH gene on human chromosomal region 8q13 is not linked to manic-depressive/bipolar disorder. Other components of the HPA axis have also been tested for linkage. On chromosome 5, the glucocorticoid receptor gene and the genes for β_2 - and α_1 -adrenergic receptors were excluded as candidates for manic-depressive disorder based on parametric lod score analyses. Mutation analysis in the gene for the adrenocorticotrophic receptor located on chromosome 18p11.2 was negative. In addition, parametric methods of analysis excluded the area of the proopiomelanocortin gene on chromosome 2 for linkage to bipolar disorder.

Role of the Tyrosine Hydroxylase Gene

Another component of the stress system is the catecholaminergic system. Tyrosine hydroxylase (TH) catalyzes the rate-limiting step in catecholamine synthesis. The TH gene is regulated by an array of biochemical mechanisms in response to environmental factors. In animals, a long-term increase of TH activity can be elicited by electrical stimulation and environmental stress or drugs such as reserpine. In humans, the involvement of the TH locus in manic-depressive disorder is supported by association studies. However, the precise genes with a causative role in bipolar disorder are not yet known. It is possible that a particular variant of the tetrarepeat located in the first intron of the TH gene plays a role. Studying mechanisms of regulation of TH gene expression in rats showed the existence of stressor specificity and stressor-specific pathways. In rats exposed to cold or insulin stress, adrenal denervation (splanchnectomy) prevented a rise in adrenal medullary TH mRNA levels, whereas immobilization-induced stress led to an increase in TH mRNA levels, suggesting that genetic changes in TH may indeed affect the stress response.

Role of Dopamine Receptors

The main effects of dopamine on the central nervous system include locomotion and behavior. Forebrain dopamine systems are involved in mediating the behavioral effects of chronic mild stress. At present, five subtypes of dopamine receptors are known and are classified as D1-like and D2-like, according to their similarity to the first and second cloned receptor. D1-like receptors are encoded by intronless genes and D2-like receptor genes are divided by intron sequences, suggesting that dopamine receptors have evolved from two main ancestral genes.

D1 receptors are predominant in the basal ganglia such as the nucleus accumbens and are involved in a prolactin stimulatory pathway in response to high environmental temperature stress. D1 receptor-deficient mice fail to thrive after weaning, with selective impairment of motivated behavior, including drinking and eating. These mice also demonstrated that the D1 receptor is important in drug addiction and reward mechanisms. The role of dopamine D1 receptor actions in reducing the function of the prefrontal cortex (a higher cognitive center) has been shown in various studies, including the infusion of a dopamine D1 receptor agonist in the prefrontal cortex, which induced cognitive deficits similar to those seen in stress. Increased catecholamine receptor stimulation in the prefrontal cortex leads to stress-induced working memory deficits, which can be ameliorated by agents that

prevent catecholamine release or that block dopamine receptors.

Mice lacking D2 receptors have an impaired reproductive function. Alterations of the D3 receptor function may play a role in behavioral disorders associated with hyperactivity, such as attention deficit hyperactivity disorder, a disease of prefrontal cortex dysfunction.

Stress Effects in Prokaryotes and Eukaryotes

Effects of Stress on Life History Traits in *Drosophila*

In natural populations, there may be significant genetic variation for stress resistance and life history traits. Different genotypes may respond differently to environmental stress, and the variation in these norms of reaction may be of great importance to the maintenance of genetic variation in metabolic traits. Each physiological response may be mediated by different genes in different individuals. Similarity between parents and progeny may be altered if the genes that contribute to variation in a trait in one environment do not contribute in another environment. Responses of life history traits to selection may depend on environmental conditions in different ways. Heritabilities for morphological traits are more constant than for life history traits.

Interestingly, selection for stress resistance can alter many traits. In selected lines of *D. melanogaster* undergoing stress, including starvation, heat, radiation, and ethanol and acetic acid exposure, a decrease in metabolic rate, early fecundity, and early behavioral activity, as well as an increase in male longevity, occurred. In *D. simulans*, desiccation resistance was associated with resistance to starvation and toxic ethanol. In *D. serrata*, selection for desiccation resistance also resulted in an increased resistance to starvation. When *D. melanogaster* and *D. simulans* were selected for increased cold intolerance, cross-generational effects were detected in all lines with marked effects of selection on early fecundity. Selected lines in both species showed reduced fecundities specifically associated with the selection response.

However, not all responses to stressful conditions are associated with correlated changes in life history traits. For instance, selection for increased knock-down resistance to heat in *D. melanogaster* led to lines that were two- to fivefold more resistant to this stress. This selection response did not alter life history traits and was associated with allelic changes in two heat shock genes, suggesting that such genes may

have specific effects on stress resistance and do not necessarily have life history costs.

These findings show that responses to stress can have many secondary effects unrelated to stress. Organisms under permanent stressful conditions may be expected to show life history changes that reduce their fitness under more favorable conditions. Many trait values seen in natural populations may reflect the effects of infrequent stressful conditions. The overall fitness of an individual that is unable to survive an infrequent stress will be zero, regardless of performance under normal conditions. The complexity of pleiotropic effects and epistatic interactions contribute to the likelihood that genotype and environment interactions play an important role in the maintenance of genetic variation.

Effects of Stress on Phenotype

In animals, most populations experience at least sporadically stressful conditions and change drastically in size when studied over long intervals. Environmental changes such as sudden climate changes can have many indirect effects, including competitive interactions and susceptibility to diseases. Animals developed many mechanisms for coping with sudden environmental shifts. For example, rapid changes at the cellular level involving the synthesis of enzymes and heat shock proteins occur in response to most stresses. Extreme conditions can influence the expression of phenotypic variability and the extent to which this variability is determined genetically. For instance, in *D. melanogaster* stress can produce abnormal morphological characteristics when the flies are exposed to stresses early in their development. However, the experimentally induced extreme conditions used to generate these mutant phenotypes probably do not exist under field conditions, and the resulting morphological peculiarities such as ultra-bithorax in *Drosophila* would probably not survive in nature.

Effects of Cellular Stress on Genotype

Genetic variants may result from the ability of cells to repair damaged DNA or from the limited accuracy of the cellular machinery for DNA replication. By means of the fluctuation test, Luria and Delbrück attempted to clarify whether bacterial mutants arise spontaneously or specifically in response to a particular kind of stress. They concluded that genetic variants do originate randomly and without selective pressure. However, the possibility that, in addition to spontaneous mutants, genetic variants might have been induced by selective pressure was not excluded

in that experiment; it was investigated and confirmed by Cairns in 1988. The degree of variation between parents and progeny differs enormously with different forms of life, depending on the stability of the living conditions.

Under nonstress conditions, variants predominantly result from replication errors. Eukaryotic organisms possess four (yeast, and probably other lower eukaryotes) or five (higher eukaryotes) different DNA polymerases. Cells do not need to assure a high accuracy of RNA primer synthesis because mistakes occurring during primer synthesis do not contribute to the mutation rate. Cells possess a battery of enzyme systems for DNA repair to avoid the loss of integrity of DNA genomes. The DNA repair system evolved to restore the original status of the damaged genome structure and to improve the ability of the cell to survive. However, some of the DNA repair processes are error-prone and are therefore important sources of genetic variants. In response to DNA breaks, recombination–repair mechanisms may construct a genome in an order that is different from the original. A cell will, naturally, rather survive as a variant than be killed as the original. Repairing DNA under the condition of prolonged stress may induce more genetic variants than absolutely necessary, as shown in starving bacterial cells that developed induced replication-independent mutants of selective advantage in the face of inhibited proliferation and threatened survival of the population. During severe stress, multiple mutations occur at high rates in cells that survived the hypermutable state with extensive DNA damage.

Effects of Stress on Heat Shock Gene Regulation and in Human Disease

Heat shock proteins: history Heat shock proteins were first described in 1962 when the chromosomes of the giant salivary gland of *D. busckii* were exposed to an elevated temperature that resulted in areas of swelling on the chromosome known as puffs. These puffs indicated transcriptionally active genes that were found to belong to the hsp family. The hsp are called molecular chaperones because they bind to unfolded proteins and thereby prevent the aggregation and misfolding of these proteins during cell stress.

Over the following years, stress protein research attracted the attention of basic and clinical investigators in many disciplines. Among other topics, studies on stress proteins include gene expression and the complex signal transduction pathways that control the regulation of stress genes; other studies focus on the role of stress proteins as molecular chaperones that regulate various aspects of protein folding and

transport and on the role of stress proteins in human disease.

Heat shock proteins: function Via the heat shock response as a homeostatic mechanism, cells are able to survive exposure to extreme environmental stress. Thermally induced stress proteins can also protect cells against other stresses that themselves induce stress protein synthesis. Such cross-protection is known as cross-tolerance.

The elevated synthesis of a few proteins following exposure to heat or other stressful stimuli, including ischemia; anoxia; hypoxia; cytokines; and certain drugs such as vasopressin, isoproterenol, and calcium channel blockers, occurs in all organisms ranging from prokaryotic bacteria such as *Escherichia coli* to mammals, including humans. After the genes encoding these stress proteins had been cloned, other genes encoding related stress proteins were identified based on their homology. Subsequently, many stress proteins were found to be encoded by multigene families encoding proteins with similar but distinct features. Major eukaryotic stress proteins are usually referred to as hsp with a number that reflects the molecular mass in kilodaltons, for example, hsp70 (Table 1).

Based on their molecular mass, the hsp are subdivided into two groups: large and small hsp. Small hsp possess a domain of homology to α A,B-crystallin proteins from the vertebrate eye and are less conserved than large hsp. Small hsp caught the attention of researchers because these proteins are able to induce protection against various toxic chemicals and interfere with inflammatory mediators. Cell death usually occurs by two distinct processes: necrosis and apoptosis. Small hsp protect cells against stress-induced necrosis and can also prevent differentiating cells

from apoptosis. Therefore, small hsp participate in essential physiological processes in unstressed cells. There are at least 4 major small hsp in *Drosophila*, 3 in mammalian cells and yeast, and more than 20 in plants. Although small hsp in plants are represented by at least two multigene families, the small hsp27 in humans is encoded by a single gene located on chromosome 7.

There seems to be a link between small hsp expression and tumorigenicity. Several cancer cell lines, including breast cancer, were found to show changes in the expression of mammalian hsp25/27. Small hsp expression is also considered a prognostic parameter of some human tumors. For instance, in patients with breast cancer or osteosarcoma the expression of hsp27 is associated with a poor outcome.

Many of these stress proteins are also expressed in normal, unstressed cells, indicating that the function of hsp is likely to be required in normal cells but more so in stressed cells. Because this function has been conserved in evolution, it seems to be very important in all cells ranging from bacteria to humans. A large body of evidence shows that the primary role of the majority of hsp is creating the proper folding of other proteins in order to enable these proteins to fulfill their appropriate functions. Stressed cells produce aberrant proteins that are folded improperly and sometimes too abnormal to be refolded properly. For such situations, some stress-inducible proteins, such as the hsp ubiquitin, are involved in protein degradation.

Among the first genes that are transcribed from the embryonic genome are genes encoding stress proteins such as hsp90, emphasizing the importance of the complex regulatory processes of the stress proteins in normal and stressed cells.

Table 1 Eukaryotic stress proteins

Family	Member	Functional role	Comment
Ubiquitin	Ubiquitin	Protein degradation	In the nucleus found conjugated to histone H2A
Hsp27	Hsp27, Hsp26, etc.	Actin-binding proteins, protein folding	In different organisms very variable in size (12–40 kDa) and number
Hsp47	Hsp47	Protein folding of collagen and possibly other proteins	Homology to protease inhibitors
Hsp56	Hsp56 (=FKBP59)	Protein folding, in the steroid receptor complex associated with hsp90 and hsp70	Peptidyl prolyl isomerase activity, target of immunosuppressive drugs
Hsp60	Hsp60	Protein folding and mycobacterial unfolding, organelle translocation	Major antigen of many bacteria and parasites that infect humans
Hsp70	Hsp70, Hsc70, Hsx70, Grp78 (=Bip)	Protein folding and unfolding, assembly of multiprotein complexes	Hsx70 only in primates
Hsp90	Hsp90, Grp94	Maintenance of proteins such as steroid receptors in an inactive form until appropriate	<i>Drosophila</i> and yeast proteins known as hsp83
Hsp100	Hp104, Hsp100	Protein turnover	ATPase activity, role in tolerance to extreme temperature

The specific transcription factor heat shock factor 1 (HSF1) induces stress protein genes after exposure to heat or other stresses and exists in an inactive (monomeric) form in unstressed cells. The regulation of these stress protein genes in response to nonstressful stimuli or in normal cells remains to be elucidated.

Stress proteins seem to be essential for the survival of cells and organisms. Thus far in mammals there is no knockout mouse lacking specific hsp genes. In *Drosophila* and in yeast, the inactivation of the hsp90 gene turned out to be lethal. However, it is possible that the inactivation of an individual stress protein gene is not lethal in all cases if genetic and functional redundancy are present. If, however, the function of a specific family of stress proteins is completely eliminated, survival is unlikely.

Heat shock proteins and chaperones show specific expression patterns during development and cell differentiation as well as during aging. hsp genes do not act as developmental genes to direct development along one pathway or another. Rather, as important partners in differentiation, cell division, and apoptosis, hsp act at the cellular level to regulate development.

The maintenance of homeostasis in mammals is achieved by many mechanisms, including the release of cytokines in response to stressful challenges. It is well known that, via interleukin-6, tumor necrosis factor α and interleukin-1 activate the HPA axis, including the CRH neuron. Cytokines are redundant and functionally pleiotropic. By activating specific transcription factors that bind and transactivate hsp genes, cytokines are also capable of inducing stress proteins/hsp, such as hsp90, hsp70, and hsp27 (Table 2).

Mild stress, sufficient to induce hsp, could protect cells against a subsequent, more severe stress. Transgenic animals overexpressing hsp70 showed an enhanced resistance to cardiac ischemia. In the clinical setting, studies are underway to identify drugs that could induce the synthesis of stress proteins in a

nonstressful manner and to find ways to deliver stress protein genes to damaged organs *in vivo* using gene therapy procedures.

On the other hand, the enhanced expression of stress proteins may play a role in many different human diseases. Several studies report that the synthesis of stress proteins is induced by stress caused by disease; this appears to be stress protein- and disease-specific. For instance, hsp90 is specifically overexpressed in a subset of patients with systemic lupus erythematosus, whereas the expression of other stress proteins is unaffected. General or specific overexpression of stress proteins can evoke an autoimmune response that leads to the production of antibodies or T cells that recognize these proteins and possibly cause a damaging immune response. Such immune responses to individual stress proteins, especially to hsp60, have been described in rheumatoid arthritis, multiple sclerosis, and diabetes mellitus type 1.

hsp90 is one of the most abundant proteins in eukaryotic cells. It makes up 2% of total cellular protein even under nonstress conditions and is highly evolutionary conserved with homologs in bacteria, yeast, *Drosophila*, and humans. hsp90 is needed for cell survival, as demonstrated by mutation analysis in yeast. hsp90 is important for the function of steroid receptors and perhaps certain viral oncogenic kinases. The hsp gp96, hsp90, and hsp70 isolated from tumors were able to immunize and elicit protective immunity specifically against the tumors from which the hsp were isolated. The hsp isolated from normal tissues were found to elicit immunity to any cancers tested. hsp96 molecules are not immunogenic per se but act as carriers of antigenic peptides. The hsp90 family may be a novel target for anticancer drug development such as vaccines, especially because cancer cells are very sensitive to the pharmacological disruption of hsp90 function.

Exciting work in yeast is centered around conformational alterations in prionlike proteins, which are believed to cause neurodegenerative diseases in humans, including Creutzfeldt-Jacob, mad cow disease, and possibly even Alzheimer's. The delineation of an evolutionary-conserved chaperone related to hsp70 whose expression changes in response to environmental stress has been found to control the inheritance of a yeast prionlike element (release factor Sup35) called PSI. Further studies of this complex non-Mendelian genetic system may elucidate how the cellular response to stress may cause disease.

Summary

Stress, a state of threatened homeostasis, leads to an individual stress response that is subserved by the

Table 2 Candidate genes

Transcription factors	Glucocorticoid receptor, mineralocorticoid receptor, androgen receptor, estrogen receptor, progesterone receptor, aryl hydrocarbon receptor, heat shock factor 1, mutated p53
Protein kinases	Tyrosine kinases (Src family, sevenless, weel, p185), serine/threonine kinase (Raf-1/v-Raf, casein, heme-regulated eIF-2 α , CDK4/CDK6, MEK)
Other proteins	Cytoskeletal proteins (tubulin, actin, intermediate filaments), calmodulin, proteasomes, G-protein $\beta\gamma$ subunit

stress system(s). Each physiological response may be mediated by different genes in different individuals, perhaps explaining why the same/similar internal and external stressors affect individuals in different ways. This is complicated by the complexity of pleiotropic effects and the interactions between genotype and environment that contribute to the maintenance of genetic variation. At the molecular and cellular level, stress evokes the synthesis of enzymes and over-expression of certain genes, many of which encode heat shock/stress proteins. These stress proteins are expressed not only in stressed cells but also in normal cells underlining their essential function. For instance, a functional ubiquitin pathway that helps eliminate damaged proteins in the cytosol, nucleus, and endoplasmatic reticulum is an essential feature of the stress response in eukaryotes as well as of unstressed cells. All cells have multiple ubiquitin genes, each of which encodes a fusion between ubiquitin and another protein. In line with Aristotle's *in medias res* (everything in moderation), it appears that individuals and, at the molecular and cellular level, cells exposed to too much stress or to no stress at all have a shorter longevity than cells and individuals that are coping with mild stress sufficient to induce protective factors such as hsp against subsequent more severe stresses.

See Also the Following Articles

Apoptosis; Chaperone Proteins and Chaperonopathies; Dopamine, Central; *Drosophila* Studies; Heat Shock Proteins: HSP60 Family Genes.

Further Reading

Accili, D., Drago, J. and Fuchs, S. (1998). Dopamine receptors in human disease: lessons from targeted mouse mutants. In: Spiegel, A. M. (ed.) *Contemporary endocrinology: G proteins, receptors, and disease*, pp. Totowa, NJ: Humana Press.

- Cairns, J., Overbaugh, J. and Miller, S. (1988). The origin of mutants. *Nature* 335, 142–145.
- Clark, A. G. and Fucito, C. D. (1998). Stress tolerance and metabolic response to stress in *Drosophila melanogaster*. *Heredity* 81, 514–527.
- Jenkins, N. L., Sgro, C. M. and Hoffmann, A. A. (1997). Environmental stress and the expression of genetic variation. *EXS* 83, 79–96.
- Kvetaňský, R., Pacak, K., Fukuhara, K., et al. (1995). Sympathoadrenal system in stress: interaction with the hypothalamic-pituitary-adreno-cortical system. *Annals of the New York Academy of Sciences* 771, 131–158.
- Latchman, D. S. (ed.) (1999). *Handbook of experimental pharmacology: stress proteins*. Berlin: Springer.
- Luria, S. and Delbrück, M. (1943). Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* 28, 491–511.
- Newnam, G. P., Wegrzyn, R. D., Lindquist, S. L., et al. (1999). Antagonistic interactions between yeast chaperones hsp104 and hsp70 in prion curing. *Molecular and Cellular Biology* 19, 1325–1333.
- Stratakis, C. A., Sarlis, N. J., Berrettini, W. H., et al. (1997). Lack of linkage between the corticotropin-releasing hormone gene and bipolar affective disorder. *Molecular Psychiatry* 2, 483–485.
- Thor, A., Benz, C., Moore, D., et al. (1991). Stress response protein (srp-27) determination in primary human breast carcinomas: clinical, histologic, and prognostic correlations. *Journal of the National Cancer Institute* 83, 170–178.
- Timpl, P., Spanagel, R., Sillaber, I., et al. (1998). Impaired stress response and reduced anxiety in mice lacking a functional corticotropin-releasing hormone receptor 1. *Nature Genetics* 19, 162–166.
- Uozaki, H., Oriuchi, H., Ishida, T., et al. (1997). Over-expression of resistance-related proteins (metallothioneins, glutathione-S-transferase, heat shock protein 27 and lung resistance-related protein) in osteosarcoma. *Cancer* 79, 2336–2344.
- Wahlstrom, J. (ed.) (1998). *Genetics and psychiatric disorders*. New York: Elsevier Science.
- Wintersberger, U. (1991). On the origins of genetic variants. *FEBS Letters* 285, 160–164.

Genetic Modifications of Stress Response

See: Genetic Factors and Stress; Genetic Predispositions to Stressful Conditions; Gene Environment Interactions in Early Development; Genetic Variation of HPA Axis Activity and Function in Farm Animals; Genetic Polymorphisms in Stress Response; Genetic Testing and Stress.

Genetic Polymorphisms in Stress Response

I W Craig

King's College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

Genes and the Stress Responses
 Polymorphisms of Genes in the Autonomic Stress
 Response Pathway
 Polymorphisms of Genes in the Neuroendocrine Stress
 Response Pathway
 Loci Whose Polymorphisms Interact with Stressful
 Influences in Mediating Behavioral Outcomes
 Conclusion

Glossary

<i>Allele</i>	One of two or more possible forms of a gene. An individual may possess two identical or two different alleles, except for the genes on the X-chromosome in males (they possess only a single copy of this chromosome).
<i>Body mass index (BMI)</i>	A measure of body fat based on height and weight that applies to both adult men and women.
<i>Case/control studies</i>	Comparisons between groups of individuals with particular features or disorders and groups apparently normal for these aspects; often employed to search for differences in the genotype or allele frequencies between the two groups.
<i>Genotype</i>	The configuration of alleles at a genetic locus. If both alleles are identical the individual is homozygous; if the two alleles differ the individual is heterozygous.
<i>Haplotype</i>	The pattern of alleles at two or more polymorphic sites represented on one of the chromosome homologs.
<i>Intragenic feature</i>	A feature located within the conventional boundaries of a gene.
<i>Kilobase (kb)</i>	One thousand base pairs.
<i>Linkage disequilibrium</i>	The existence of combinations of alleles at two or more polymorphic sites within a region at frequencies greater than expected by random assortment.
<i>Missense mutation</i>	A base change in the coding region of the gene that results in the substitution of one amino acid in the encoded protein for another.
<i>Polymorphism</i>	The existence at gene locus of more than one allele with the rarest at appreciable frequency (usually assumed to be >1%) in the population.

Promoter

The section of DNA that precedes a gene and controls its activity.

Single nucleotide polymorphism (SNP)

A polymorphism at a single base position in DNA.

Stressor

Stimulus input that creates stress and initiates a response.

Genes and the Stress Responses

The genetic basis for the stress response is complex; nevertheless, the genes involved can be divided into two broad categories: those encoding the autonomic reaction to stressful situations and those involved in the neuroendocrine stress response. The typical rapid reactions to perceived stress are well recognized to be associated with the fight-or-flight response and are mostly mediated by the sympathetic nervous system with the release of norepinephrine from widely distributed synapses and epinephrine from the adrenal medulla. Implicit in this classic description of the rapid stress response is the recognition of the potential violent and aggressive outcomes that stress can engender. A more delayed response is mediated by the release of cortisol, the major glucocorticoid in humans. This is initiated by neural signals to the hypothalamus that releases corticotropin releasing hormone (CRH), which in turn results in the release of adrenocorticotrophic hormone (ACTH) from the anterior pituitary, thereby stimulating cortisol production by the adrenal cortex. The production of cortisol assists in restoring homeostasis in response to stress; however, prolonged exposure can be potentially harmful.

Reactions to acute stress are also usually accompanied by perturbations of serotonergic activity, as indicated by altered levels of serotonin (5-HT) or its metabolites. Serotonin is a significant regulator of morphogenetic activities during early central nervous system development, including cell proliferation, migration, and differentiation.

Outcomes following exposure to stress may, therefore, be highly variable depending on a wide range of genetic factors controlling these various aspects. A simplified overview of some of the most important elements of stress response and relevant genes is shown in [Figure 1](#). The significance of genetic variation in the better-characterized components of the stress response pathways is considered next. Several case/control studies have examined the association between genes involved in the stress response and somatic

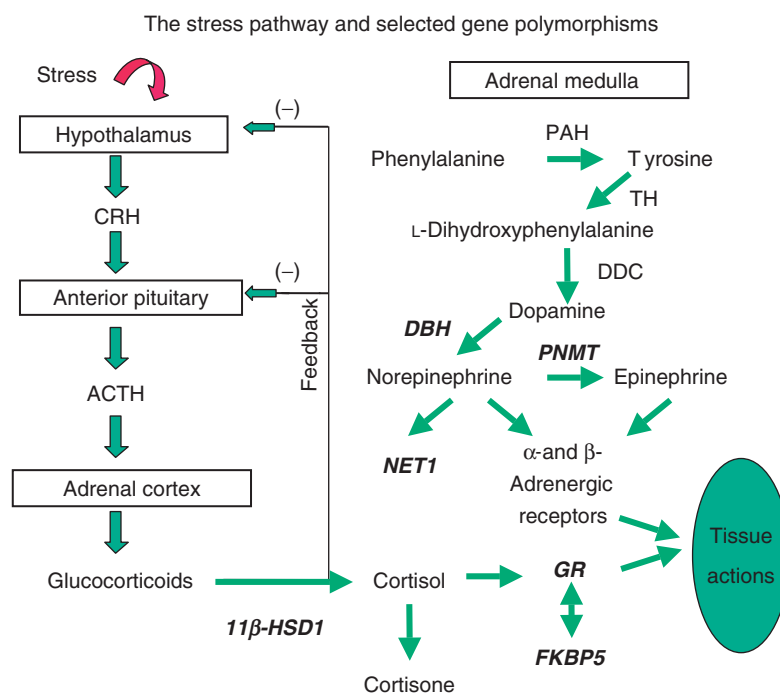


Figure 1 The stress pathway illustrating selected genes (in bold italics). 11 β -HSD1, 11- β -hydroxysteroid dehydrogenase enzymes; ACTH, Adrenocorticotrophic hormone; GR, glucocorticoid receptor; CRH, corticotropin releasing hormone; DBH, dopamine β -hydroxylase; DDC, DOPA decarboxylase; FKBP5, glucocorticoid receptor co-chaperone; NET1, norepinephrine transporter; PAH, phenylalanine hydroxylase; PNMT, phenylethanolamine *N*-methyl transferase; TH = tyrosine hydroxylase. α - and β -Adrenergic receptors include ADRA1A, ADRA1B, ADRA1D, ADRA2A, ADRA2B, ADRA2C, ADRB1, ADRB2, and ADRB3.

or behavioral traits. Some of these are noted; however, it is important to recognize that there are potential confounding factors to such studies and only consistent replication and supporting molecular evidence can establish the formal verification of such links.

Polymorphisms of Genes in the Autonomic Stress Response Pathway

The rapid stress response to stressful stimuli is mediated by the release of norepinephrine in the brain. This leads to the activation of the sympathetic nervous system and the adrenal medulla with release of epinephrine in the blood. Hence, the genes regulating the biogenesis and degradation of both epinephrine and norepinephrine and those encoding their cognate receptors are key players in the primary stress response. The epinephrine and norepinephrine metabolic pathways share components with those for other neurotransmitters. The steps in the synthesis of epinephrine from amino acid precursors are illustrated in Figure 1. The last two critical steps are undertaken by the enzymes dopamine β -hydroxylase (DBH) and phenylethanolamine *N*-methyl transferase (PNMT).

Dopamine β -Hydroxylase

The dopamine β -hydroxylase gene (*DBH*) is located on chromosome 9. A SNP in the presumed controlling

region of *DBH*, which results in the substitution of a C to T substitution 1021 bp before the transcription start, accounts for up to 50% of the variance in plasma levels of the enzyme. Other variants flanking, or within, the potential promoter region include a polymorphism that depends on the copy number of an intragenic dinucleotide repeat approximately 4 kb upstream of the transcription start and a 19-bp insertion/deletion located 118 bp upstream.

In addition several SNPs have been reported within the coding region. Of these, an arginine to cysteine substitution in exon 11 has been suggested to be the most likely functionally significant variant. Several studies have linked a variation in plasma levels and *DBH* genotypes to various physiological and behavioral disorders, including hypertension, psychosis, depression, migraine, attention deficit hyperactivity disorder, and alcoholism. These reported associations await replication and consolidation.

There are also reports of individuals with undetectable levels of norepinephrine and epinephrine, together with elevated levels of dopamine. Such individuals have low, or absent, circulating DBH enzyme caused by deleterious mutations affecting both copies of the *DBH* gene and manifest a range of symptoms, including orthostatic hypotension.

Although alterations to this key enzyme in the synthetic pathway for norepinephrine and epinephrine

are likely to have significant consequences for the stress response, no specific behavioral or physiological features relating to the reactions to stressors *per se* have been reported.

Phenylethanolamine *N*-Methyl Transferase

The gene for PNMT, a key enzyme, is located on chromosome 17. Two SNPs that may affect the regulation of the gene have been identified in the promoter. When tested in case/control studies, significant association was observed between these and early-onset, but not late-onset, versions of Alzheimer's disease, and there are data suggesting a possible association between these SNPs and hypertension in African American groups. In addition, there is evidence suggesting that a variant of this gene may be associated with blood pressure regulation abnormalities in the stroke-prone, spontaneously hypersensitive rat strain.

Norepinephrine and Epinephrine Receptors

The adrenergic receptors for epinephrine and norepinephrine are divided into two major classes, alpha (subdivided into α_1 and α_2) and beta (subdivided into β_1 , β_2 , and β_3). These receptors are targeted by a variety of drugs, including the so-called beta-blockers; hence genetic variation in them may be an important feature for pharmacotherapy. The activation of the β -adrenergic receptors stimulates adenylate cyclase.

α_1 -Adrenergic receptors The α_1 -adrenergic receptors are G-coupled trans-membrane receptors that are capable of binding both epinephrine and norepinephrine and thereby mediate some aspects of the sympathetic nervous system. The genes encoding them include *ADRA1A* (formerly known as *ADRA1C*, on chromosome 8), *ADRA1B* (formerly known as *ADRA1*, on chromosome 5) and *ADRA1D* (formerly known as *ADRA1A*, on chromosome 20). Of these, *ADRA1B* is of particular interest because fine mapping with highly polymorphic markers localizes it to a region that also contains the dopamine receptor gene *DRD1A*. Because this chromosome region has been linked to variability in systolic blood pressure, genetic variation in either *ADRA1B* or *DRD1A*, or both, may play an important role in influencing interindividual differences in this feature; however, specific functional variants in these have yet been described.

α_2 -Adrenergic receptors The release of norepinephrine is controlled by negative feedback from α_2 -adrenergic receptors, which are located on the presynaptic side of nerve terminals, including nerve muscle junctions. The targets of the released norepinephrine on heart muscle cells (myocytes) are

β_1 -adrenergic receptors. The individual roles of the three highly homologous α_2 -adrenergic receptors (*ADRA2A*, *ADRA2B*, and *ADRA2C*, encoded by genes located in humans on 10, 2, and 4, respectively) have been studied by gene disruption in mice. Both *ADRA2A* and *ADRA2C* subtypes were found to be required for the normal presynaptic control of transmitter release from sympathetic nerves in the heart and from central noradrenergic neurons. In addition, a small deletion (base pairs 322–325) in the *ADRA2C* gene in humans has been shown to decrease function and may act in concert with the beta-receptor gene *ADRB1* as a risk factor for heart attack.

A polymorphism leading to the deletion of three glutamic acid residues from a repeating motif in the third intracellular loop of *ADRA2B* has been associated with variations in basal metabolic rate.

β -Adrenergic receptors β_1 -Adrenergic receptors encoded by the gene *ADRB1* (also known as *ADRB1R*, located on chromosome 10) are present in heart, and their stimulation causes an increase in heart rate. Many beta-blockers act on these targets (and on β_2 receptors to a lesser extent). One variant of the β_1 -adrenergic receptor gene, which changes an arginine at position 389 in the protein to glycine, has been shown to possess increased function in tissue culture. There is also a polymorphic variant of a small deletion (base pairs 322–325) in the gene for the α_2C -adrenergic receptor, *ADRA2C*, that has decreased function. A marked increase in the risk of heart failure among black patients has been observed for those who are homozygous for this and the *ADRB1* variant. This suggests that genotyping of these two genes may be a useful approach for identifying people at risk for heart failure. Further support for an interaction between these two genes has been provided by recent studies indicating that *ADRA2A* and *ADRB1* proteins form heterodimers when expressed in tissue culture cells. In Caucasian women, it has been observed that the arg389 allele of *ADRB1* was associated with greater body weight and BMI.

β_2 -Adrenergic receptor Several polymorphic variants have been described, the most common and significant of which is a substitution causing a change from glycine to arginine at position 16. β_2 -Adrenergic receptor agonist drugs are widely employed in the treatment of asthma. It is therefore of interest that an excess of the glycine allele has been observed in nocturnal asthmatics. It appears that this substitution imparts enhanced agonist-promoted downregulation of the receptor.

A second polymorphism is centered on amino acid 27, where a glutamine to glutamic acid replacement

is observed. An association with obesity has been reported, with an excess of the glutamine version in nonobese individuals. Possible links of this variant with asthma have also been noted.

A third variant involving a threonine to isoleucine replacement at amino acid 64 exists at lower frequencies. This has a profound functional consequence that results in a fivefold reduction in sensitivity to β_2 -receptor agonist-mediated vasodilatation. This suggests a mechanistic explanation of decreased survival in patients carrying this variant.

β_3 -Adrenergic receptor The β_3 -adrenergic receptor, ADRB3, is located mainly in adipose tissue and for this reason has been studied in relation to obesity. A mutation resulting in the substitution of arginine for tryptophan at amino acid position 64 has been observed at a frequency of approximately 10%. Many studies have examined the relationship of the polymorphism to obesity, BMI, and noninsulin-dependent diabetes mellitus, with some concluding that the locus may have a role in predisposing these phenotypes; but other studies have failed to replicate the observations.

Norepinephrine Transporter

The norepinephrine transporter protein 1 (NET1) is also known as the norepinephrine transporter (NAT1) or the solute carrier family 6, member 2 (SLC6A2). The gene responsible is localized on chromosome 16. Approximately 80% of the reuptake of norepinephrine is mediated at nerve terminal synapses by the norepinephrine transporter. SLC6A2 is a target for the tricyclic antidepressants, including imipramine and desipramine and also for amphetamines and cocaine.

Following a report of SNP in the restriction site for the enzyme *TaqI*, a number of detailed investigations have revealed a range of rare mutations associated with orthostatic intolerance and more than 20 polymorphisms, some of which resulted in missense substitutions; however, no association with behavioral disorders were observed in the case/control studies. An insertion/deletion polymorphism of 4 bp in one of several AAGG repeat motifs in the promoter region at a position overlapping a transcription factor binding site has been reported, and there is some evidence supporting an association between the presence of the insertion and restricting anorexia nervosa. Rare genetic variants encoding a transporter with reduced affinity for norepinephrine have also been described. Carriers have been observed to manifest clinical symptoms (orthostatic intolerance) and metabolic evidence for disordered uptake of norepinephrine.

In summary, studies on the genes implicated in autonomic norepinephrine and epinephrine pathways

have demonstrated the potential for a great deal of interindividual diversity. The consequences of this have major impact on a wide variety of disorders, including several with direct or indirect links to stress. Other autonomic pathways are mediated through acetylcholine release and detection.

Polymorphisms of Genes in the Neuroendocrine Stress Response Pathway

Further sources of genetic diversity in individual reactions to stress are the genes implicated in the delayed response of the hypothalamus-pituitary-adrenal (HPA) axis, leading to the release of cortisol from the adrenal cortex initiated by neural signals to the hypothalamus. Other genes are involved in the subsequent mobilization of hypothalamic peptide hormones (including CRH), resulting in the secretion of ACTH from the anterior pituitary and the release of glucocorticoid hormones (cortisol) from the adrenal cortex. Glucocorticoids are produced by the adrenal cortex and restore homeostasis in response to stress; however, prolonged exposure resulting from chronic stressful occurrences is potentially harmful. Cortisol (the most prominent glucocorticoid in humans) or corticosterone (in rats) is largely bound to corticosteroid-binding globulin in the systemic system. It has a classic steroid hormone mode of action via binding to a cognate intracellular receptor (the glucocorticoid receptor) and subsequent migration to the nucleus, where the complex interacts with a specific gene transcription apparatus. The loss or profound diminishment of glucocorticoid secretion leads to a state of deranged metabolism and an inability to deal with stressors, which, if untreated, is fatal. Functional genetic polymorphisms in some important genes of this pathway are recognized.

The Glucocorticoid Receptor

Two specific corticoid receptors are recognized: the glucocorticoid receptor (GR) and the mineralocorticoid receptor (MR). Whereas the latter binds aldosterone and glucocorticoids with similar high affinity, GR is selective for the glucocorticoids. It appears that the main response to raised levels of steroids resulting from stress is mediated primarily by GRs. There have been several polymorphisms described for the GR gene, including one in the promoter region. Because there is strong linkage disequilibrium across the gene, there are relatively few different combinations of allelic patterns, with four major haplotypes recognized. Knowledge of the haplotype may help determine the response to glucocorticoid treatments. Variants

within the coding region include two polymorphisms at adjacent codons for the amino acids 22 and 23 in the protein. Whereas the base change at codon 22 does not change the amino acid (glutamic acid), the polymorphism at codon 23 results in a change from arginine to lysine. Individuals with a specific combination of alleles at the promoter site and the codon 22–23 sites have an associated relative resistance to glucocorticoids. A further polymorphism resulting in a change from asparagine to serine at codon 363 appears to be associated with hypersensitivity to glucocorticoids. A *BclI* restriction enzyme polymorphism has also been employed in case/control studies, and there is evidence for an association with hypersensitivity to glucocorticoids that may be related to lower BMI in older individuals.

A wide variety of membrane-bound GRs have been reported in a range of tissues including the liver, lymphocytes, adrenal gland, and pituitary. There is a paucity of information concerning any structural features distinguishing these various forms, and there are no data on how any genetic variation may affect them, other than through changes affecting the GR itself. Significantly, however, the GR function depends on interactions with several components of a large molecular complex. Five co-chaperones of GR are recognized as significant in regulating its activity, and they are encoded by the following genes:

- *BAG1* – BCL2-associated anthanogene 1
- *STUB1* – STIP 1 homologous and U box-containing protein 1
- *TEBP* – Thyroid transcription factor 1 (TITF1)
- *FKBP4* – FK-506 binding protein 4
- *FKBP5* – FK-506 binding protein 5

A recent analysis of SNPs in these genes found three within the *FKBP5* locus to be significantly associated with the increased recurrence of symptom episodes in depressed patients and a rapid response to prescribed antidepressant treatment. Evidence for this apparent association was strengthened by genotyping a range of additional SNPs within and flanking the gene. One marker in particular (rs1360780) seems to track the tightness of the control of glucocorticoid signaling and HPA axis reactivity mediated by the reported interaction of the GR and *FKBP5*.

11- β -Hydroxysteroid Dehydrogenase Enzymes

Glucocorticoids are metabolized by the 11 β -hydroxysteroid dehydrogenase enzymes (11 β -HSD1). There are at least two forms of 11 β -HSD1: type I (HSD11B1) and type II (HSD11B2). Both catalyze the conversion of cortisol to cortisone; however, the type I enzyme also catalyses the reverse reaction. Both

are therefore important in the stress response. The HSD11B2 enzyme inactivates 11-hydroxysteroids in the kidney, the circumventricular organ of the brain, and some other tissues, thus protecting the nonselective MR from occupation by glucocorticoids. A base change confers a restriction-fragment-length polymorphism (detected by the enzyme *AluI*) in exon 3 of this gene. This variant and a polymorphic microsatellite marker of the *HSD11B2* gene having 12 alleles have been employed to compare salt-sensitive (SS) and salt-resistant (SR) individuals in otherwise nonhypertensive white males. Homozygosity for one of the microsatellite alleles was higher in SS subjects than in SR subjects and the cortisol-to-cortisone ratios were higher in homozygous subjects than in the corresponding control subjects. It is possible that the decreased HSD11B2 activity observed in SS subjects reflects its involvement in the SS blood pressure response in humans and that variants of the gene contribute to different blood pressure responses to salt in humans.

Loci Whose Polymorphisms Interact with Stressful Influences in Mediating Behavioral Outcomes

The Serotonin Transporter

One particularly interesting connection between the perception of stress response and the serotonergic system is through the serotonin transporter, which removes the neurotransmitter from the synaptic cleft. The serotonin transporter (SERT, 5-HTT) plays a pivotal role in brain 5-HT homeostasis. It is also the initial target for both antidepressant drugs and drugs of abuse, some of which are potent neurotoxins. In humans, the gene for the transporter (variously designated as *SERT*, *HTT* or *SLC6A4*) is characterized by functional variants in the promoter region conferring high or low activity in expression of the locus. Several studies have linked variations of the two predominant alleles with anxiety response. Caspi and colleagues were the first to show that it appears that individuals with the low activity are more vulnerable to stressful life events and more likely to become clinically depressed as a result. Glatz and colleagues, in 2003, suggested a possible connection between stress and the serotonin system. They were able to demonstrate that a potent glucocorticosteroid hormone, dexamethasone, stimulates the expression of *SERT* in tissue culture at physiological concentrations. The hormone interacts with the promoter in an allele-specific manner, and this interaction may help to explain the differential vulnerability of individuals who are more susceptible to stressful experiences.

Monoamine Oxidase Type A

A similar picture emerges for variation of another gene involved in the serotonin system, the gene for monoamine oxidase type A, MAOA. Although this enzyme also has a role in metabolizing other neurotransmitters, including dopamine, drugs inhibiting its activity have long been employed in the treatment of behavioral disorders (particularly depression) and are thought to work through altering 5-HT levels. There are two major variants in the promoter of this gene that confer either high or low activity. Because the gene is X-linked, males exhibit either one or the other phenotype without complications from the presence of a second allele. There is a consensus view emerging that males who have the low-activity form of the gene are more sensitive to the effects of maltreatment and presumptive stressful environment manifested by their increased risk of antisocial behaviour and aggression. The high-activity form of the gene appears to provide protection against this outcome.

Conclusion

The response of individuals to stress varies considerably and depends on a wide range of environmental experiences, together with cognitive and genetic factors. Considerable progress has been made in characterizing some of the genes in the pathways relevant both to the fight-or-flight response and to those involved in the release and detection of glucocorticoid hormones. Many of the most significant changes to such genes are rare mutations. There are in addition, however, several variants of genes that are widely distributed in the population, and, although their impact may not be directly observable as an altered stress response, they may have important consequences with respect to a range of complex disorders including hypertension and allergic reactions. Furthermore, recent studies have shown common variants in the key enzymes involved in 5-HT metabolism (MAOA and SERT) can interact with stressful life experience to predispose individuals to manifesting antisocial behaviour or depression. Variants in several other related genes have subsequently also demonstrated an interaction with stress in predicting behavioral outcomes.

Further Reading

- Caspi, A., McClay, J., Moffitt, T. E., et al. (2002). Role of genotype in the cycle of violence in maltreated children. *Science* **297**, 851–854.
- Caspi, A., Sugden, K., Moffitt, T. E., et al. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* **301**, 386–389.
- Cubells, J. F. and Zabetian, C. P. (2004). Human genetics of plasma dopamine β -hydroxylase activity: applications to research in psychiatry and neurology. *Psychopharmacology* **174**, 463–476.
- Dionne, I. J., Garant, M. J., Nolan, A. A., et al. (2002). Association between obesity and a polymorphism in the beta-1-adrenoceptor gene (gly389arg ADRB1) in Caucasian women. *International Journal of Obesity* **26**, 633–639.
- Small, K. M., Wagoner, L. E., Levin, A. M., et al. (2002). Synergistic polymorphisms of beta-1- and alpha-2C-adrenergic receptors and the risk of congestive heart failure. *New England Journal of Medicine* **347**, 1135–1142.
- Stober, G., Nothen, M. M., Porzgen, P., et al. (1996). Systematic search for variation in the human norepinephrine transporter gene: identification of five naturally occurring missense mutations and a study of association with major psychiatric disorders. *American Journal of Medical Genetics: Neuropsychiatric Genetics* **67**, 523–532.
- van Rossum, E. F. C., Koper, J. W., van den Beld, A. W., et al. (2003). Identification of the *BclI* polymorphism in the glucocorticoid receptor gene: association with sensitivity to glucocorticoids *in vivo* and body mass index. *Clinical Endocrinology* **59**, 585–592.
- van Rossum, E. F. C., Roks, P. H. M., de Jong, F. H., et al. (2004). Characterization of a promoter polymorphism in the glucocorticoid receptor gene and its relationship to three other polymorphisms. *Clinical Endocrinology* **61**, 573–581.
- Wüst, S., van Rossum, E. F. C., Federenko, et al. (2004). Common polymorphisms in the glucocorticoid receptor gene are associated with adrenocortical responses to psychosocial stress. *Journal of Clinical Endocrinology and Metabolism* **89**, 565–573.

Relevant Websites

Online Mendelian inheritance in man and Entrez SNP database, <http://www.ncbi.nlm.nih.gov>.

Genetic Predispositions to Stressful Conditions

D Blackwood and H Knight

University of Edinburgh, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D Blackwood, volume 2, pp 212–217, © 2000, Elsevier Inc.

How Genes and Stressful Events Can Interact

Measuring Life Events

Twin Studies of Exposure to Stress

Molecular Genetic Approaches

Glossary

Genes

It is estimated the human genome contains approximately 30,000 genes distributed on the 23 pairs of human chromosomes, and about one-third of genes are expressed in the brain. Each gene is the specific DNA code for a protein.

Genetic association studies

Studies designed to detect differences in the frequencies of DNA polymorphisms in a group of individuals with a particular phenotype contrasted with a group of control subjects from the same population.

Genetic polymorphism

The occurrence of two or more variants (alleles) of DNA sequence in a population. There are many types of sequence variation in the human genome, including single nucleotide polymorphisms (SNPs), microsatellites, deletions, and insertions that have proved valuable for locating genes in human disease using linkage and association strategies. A well-known example of genetic polymorphisms is those giving rise to the proteins of the ABO blood groups.

Genotype

The complete genetic composition of an organism.

Life events

Events usually measured by a trained interviewer using standardized interview schedules to obtain a record of events over a stated time. The inventory is likely to cover common and major events, including financial and marital problems, bereavement, ill health, employment change and infringements of the law. Some investigators separate the effects of different classes of life events, for example, an independent event such as a close friend being involved in an accident is contrasted with the subject him- or herself being in an accident, an event

over which the subject may have some control or which directly threatened him or her.

Phenotype twin studies

The physical expression of the genotype. In genetic research, studies based on the fact that identical (monozygotic, or MZ) twins have exactly the same genes and nonidentical (dizygotic, or DZ) twins share on average 50% of their genes. Both types of twins share similar environments. In characteristics that are genetically determined, MZ twin pairs are more alike (concordant) than DZ pairs, whereas MZ and DZ concordance rates are similar in phenomena due solely to environmental influences.

It has often been taken for granted that anxiety and depression are the natural consequences of stressful events such as sudden bereavement, loss of a job, or marital disharmony and that stresses in childhood such as those related to poor parenting, physical or emotional abuse, or loss of a family member may have an impact on behavior well beyond childhood. However, many early studies of the effects of stressful situations on behavior have not taken into account the effects of genes. It is now abundantly clear that genetic factors play an important and sometimes major role in many human behaviors, personality dimensions, and the development of specific patterns of illness.

How Genes and Stressful Events Can Interact

Kendler and Eaves gave a detailed account of three models to explain how genes and the environment can interact to cause psychiatric illness. These models also help us understand how genetic influences and stressful events interact in everyday life not necessarily associated with psychiatric symptoms. They form a useful basis for discussing a number of recent empirical studies.

The first and simplest model hypothesizes that symptoms of anxiety and depression are the result of the addition of genetic and environmental contributions. This assumes that people, regardless of their genotype, will respond in the same way to an environmental stress, and the probability of being exposed to a stress is independent of genetic background. This model has been the basis of much research in social psychiatry and has, for example, highlighted the

importance of early loss and poor social support as risk factors for the development of depression during adulthood in women. Such a direct causal link between environmental challenges and behavior is illustrated by the normal physiological and behavioral response to a life-threatening situation that is often termed the fight-or-flight response. However, the assumption that environmental stresses and genotype act independently does not explain the empirical observation that some individuals appear to be more sensitive than others to stressful life events.

A second model proposes that genes influence the individual's response to stressful events in the environment. The experience of life events differs widely between individuals and is related to personal characteristics including self-esteem, social support, mood, and personality traits. An example might be that individuals with high scores on neuroticism (a trait that is in part genetically determined) respond to loss or personal threat differently than those with low scores on neuroticism. Animal studies provide an example of how gender differences influence response to stress. Shors and colleagues have reported that the effect of acute stress on memory in rats differs between males and females; the degree of retrograde amnesia or enhanced memory depends on the sex of the animal. This suggests that the genetic differences between the sexes may underlie the expression of different stress response strategies.

Environmental control of the expression of genes is illustrated by animal studies showing that the stress of maternal separation during early development influences the growth of specific brain structures, leading to changes in adult behaviors. Different strains of animals respond differently to early stress, confirming a genetic background to the response. Under this model, however, genotype and environment are not correlated. The model accounts for the influence of genetic factors on the response of an individual to stressful events, but it does not include an effect of genes on the probability that a person will be exposed to stressful environmental situations.

A third, and in many ways the most interesting, model states that a person's genotype alters the probability that he or she will be exposed to a stressful environment. Genes and environment do not act independently, but are correlated. In effect, individuals throughout their lives select the level of environmental stress to which they are exposed. Specific stressful events such as road traffic accidents and injuries at work relate to a number of factors, including personality, drug and alcohol intake, and psychiatric illness. A common but indirect example is a person who is alcoholic and exposed to the consequences of heavy drinking. Since alcoholism is partly a heritable

condition, as shown by family twin and adoption studies, one of the effects of the genes that increase risk of drinking behavior is also to increase exposure of the alcoholic to a range of stressful life events, including marital difficulties, assaults, road traffic accidents, and homelessness. Genes may also influence how a person recalls, perceives, and reports on their experiences of stressful life events.

Measuring Life Events

Over the past 30 years the concept of life events has been developed and refined as a means to study the effects of stressful conditions in the environment on human behavior and psychopathology. Scales such as the Social Readjustment Rating Scale of Holmes and Rahe were developed to be used as a measure of an individual's exposure to environmental stress and have become well-researched tools in social psychiatry. However, the reporting of life events may be influenced indirectly by genetic factors. One of the first studies to address the issue of how life event reporting is governed by social and other factors was carried out in New Zealand by Fergusson and Horwood, who obtained information on life events exposure including financial problems, marital problems, bereavement, ill health, and employment changes over a 6-year period for over 1000 women. The results made it clear that reporting stressful events was not random, and two major determinants of vulnerability to reporting life events were the level of social disadvantage of the woman and her level of trait neuroticism derived from the Eysenk Personality Inventory. Women of socially disadvantaged backgrounds and women with high neuroticism scores consistently reported high exposure to life events over the 6-year period.

Twin Studies of Exposure to Stress

Twin studies that include the comparison of monozygotic (identical) and dizygotic (nonidentical) twin pairs provide a powerful means to quantify the relative contributions to human behaviors of genetic and environmental factors. Rigorous statistical analysis including correlations between twin pairs permits calculation of variance due to genetic factors and shared and individual environments. If a twin experiences a life event that is entirely random and not influenced by his or her personal characteristics, we would not expect the co-twin to show increased likelihood of experiencing a similar event. The co-twin would have the same risk as the general population. If shared family environment was responsible for the life event, both MZ and DZ twins would share an

equally increased exposure. If genetic factors contribute to the event, MZ twins will show greater concordance than DZ twins for that variable because they have more genes in common.

Kendler, in a study of more than 2000 twin pairs selected from the population in Virginia, found clear evidence that reported life events were influenced by both genetic and environmental factors. There was a significant familial resemblance for reported total life events: the correlations for MZ twins (0.43) and DZ twins (0.31) both differed significantly from zero ($p < 0.001$), and the correlation for monozygotic twins consistently exceeded that for dizygotic twins. These results support the view that personality characteristics increase the probability of experiencing stressful life events. The genetic effect was observed only when life events were divided into those that affected the individual directly (personal events), for example, if the individual was in a car accident, as opposed to events involving someone else (network events), for example, if a close relative was in a car accident. Personal events included marital difficulties, being robbed or assaulted, interpersonal difficulties, financial problems, illness or injury, and legal problems, and these were all significantly influenced by genetic factors that explained nearly 20% of the variance. Shared environment had little effect on personal events. In contrast, network events affecting other people were not influenced by genetic factors, and the similarities between twins were entirely attributable to shared family environment. It is understandable that personal events should be under some degree of genetic control, since these events are dependent on the person's behavior whereas network events are not.

This twin study confirmed that the occurrence of stressful life events is more than just bad luck. A combination of familial and genetic influences may predispose the individual to create for themselves a high-risk environment leading to a complex relationship between life events and stress-related disorders. In a further analysis of the female twins from the Virginia Twin Registry, Kendler and Karkowski-Shuman demonstrated that the genetic liability for depression in these twins was associated with an increased risk for stressful life events that included assault, marital problems, job loss, serious illness, financial problems, and social difficulties with relatives and friends. In the same twin sample, the genetic liability to alcoholism was related to the risk of being robbed and having trouble with the police. One of the ways that genes contribute to psychiatric illness is probably by causing individuals to select themselves into high-risk environments. Silberg, in a study of juveniles from the Virginia Twin Registry, proposed a model whereby depression in girls is partly influenced

by genes that mediate adverse life events, and the influence of genes on life events is not stable but increases during adolescence.

Lyons specifically examined how genes influenced exposure to combat trauma in a study of more than 4000 male twin pairs who had served in the Vietnam War. A range of data was collected, including volunteering for service in Vietnam, actual service, a composite index of combat experiences, and the award of combat decorations. Correlations within twin pairs for actually serving in Vietnam were 0.41 (MZ) and 0.24 (DZ), and those for combat experiences were 0.53 (MZ) and 0.3 (DZ). These figures generated an estimate of heritability of 35 to 47%. The family environment had no significant effect on any of the variables, and the results are unlikely to be due to biases from self report because being awarded a combat medal gave similar correlations. The differences between MZ and DZ twin pairs show that exposure to combat is partly heritable and is not influenced by the socioeconomic status of the family of origin. A number of factors could influence combat experience, including the personality trait of novelty seeking, which is known to be genetically determined. Individuals scoring high on novelty seeking will be more likely to volunteer for active service. Similarly, genetic traits that increase the likelihood of a person suffering physical and psychiatric illnesses will act to decrease the likelihood that the person will seek out and experience exposure to trauma in combat.

A twin study by Stein extended these observations to noncombat situations in a study of 400 Canadian civilian twin pairs rated for exposure to trauma and symptoms of posttraumatic stress disorder (PTSD). Seventy-five percent of these twins reported experience of trauma that could be divided into assaultive (robbery or sexual assault) and nonassaultive (road traffic accident or natural disaster). Analysis of heritabilities showed that genetic effects contributed to risk of exposure to assaultive but not nonassaultive trauma, and there was a correlation between genetic effects and PTSD symptoms. It was concluded that many of the same genes that influence exposure to assaultive trauma appear to influence susceptibility to PTSD symptoms in their wake.

Studies of twins have provided important evidence that genetic predisposition has an important role in stress-related activity and the experience of stressful life events ranging from engagement in wartime combat to exposure to assault in civilian life. It is also clear that genetic factors influence how people respond to stress, explaining why some people and not others develop symptoms of PTSD. How these genetic influences are mediated is not known, but a variety of mechanisms have been proposed.

Molecular Genetic Approaches

We can now be fairly sure that a person's complement of genes strongly influences personality traits, but it is highly unlikely that in due course a gene for neuroticism or any other personality dimension will be uncovered. It is more probable that these traits are determined by a number of genes, and the inheritance of different variants of these genes will lead to the particular traits in a manner that has long been familiar to animal breeders.

Complex phenotypes including human personality characteristics are often termed quantitative traits because a number of different genes are thought to act additively in a quantitative manner to produce the phenotype. The genes contributing to the additive effect are termed quantitative trait loci (QTL). The terms polygenic (caused by many additive or interacting genes) and multifactorial (caused by many genes interacting with environmental factors) have also been used to describe the same situation in which a complex phenotype is under a variety of genetic influences. The completion of the sequencing of the human genome, followed by the Hapmap project, which catalogued hundreds of thousands of normal DNA sequence variants (polymorphisms) present at high-density across the genome, have given researchers unrivalled tools for detecting genes contributing to inherited human diseases and complex behavioral traits. Linkage studies use many markers on each of the 23 pairs of chromosomes to identify the small DNA region that is inherited along with a particular disease or trait within a particular family. Another approach for identifying genes contributing to monogenic disease or complex disorders is to study very large cohorts of subjects rated for a particular quantitative trait and compare the frequency of selected DNA polymorphisms between cases and a matched sample of controls recruited from the same population. Association studies are used to study candidate genes or small selected regions of a chromosome previously identified by family linkage analysis. Candidate genes coding for proteins in neurotransmitter pathways, including the serotonin (5-HT) and dopamine systems, have been the focus of particular attention because these neurotransmitters are known from animal and clinical studies to play an important role in regulating mood and a variety of behaviors relevant to traits of impulsiveness, sensation seeking, and addictive behavior. The neurotransmitter 5-HT is an important regulator of early brain development, and 5-HT neurons play known roles in integrating emotion, cognition, and motor function. Genes related to 5-HT are good candidates for a role in human personality, and the 5-HT transporter gene

apparently influences a constellation of traits related to anxiety and depression. Individuals who carry a common small deletion in the promoter region of the 5-HT transporter gene are more susceptible to depression. A possible role for the transporter gene as a mediator of the effects of environmental stress on moods was described by Caspi in a study of a New Zealand birth cohort of individuals. At the age of 26, these individuals were assessed for symptoms of depression and anxiety, and self-reported life events occurring in the previous 5 years were recorded. To address the question of why stressful experience leads to depression in some people and not in others, the 5-HT transporter gene was studied and a crucial difference was observed between individuals who became depressed after experiencing stress and those who were tolerant of stress. Those who became depressed were more likely than nondepressed individuals to carry a small deletion on the promoter region of the gene. This raises the possibility that different variants of the 5-HT promoter gene could moderate how a person is going to react to adverse life experiences.

The neurotransmitter dopamine, like 5-HT, is distributed widely in the nervous system and has a range of functions, including the modulation of mood. It is of considerable interest that a receptor in the dopamine family (DRD4) may influence novelty-seeking behavior. In a study described by Noble, a person who inherits one variety of the DRD4 receptor appears to be more likely to rate highly when assessed for novelty seeking. The same trait was also more common in boys with one particular set of variants in the DRD2 gene. However, the combined DRD4 and DRD2 polymorphisms contributed more markedly to novelty-seeking behavior than when these two gene variants were individually considered. Substantially more needs to be done to unravel the complex relations between stress and symptoms in populations, and it is likely that many more candidate genes will be examined by these methods in large populations to try to detect gene-environment interactions.

Neurotransmitter function is just one of many possible areas of gene-environment interaction providing a link between chronic stress and ill health. An entirely different approach was reported by Epel and colleagues, who investigated the effects on cellular aging of chronic psychological stress experienced by a group of mothers looking after a chronically ill child. They found a significant association between perceived stress and shorter telomere length in peripheral blood mononuclear cells. Telomeres are regions at the ends of each chromosome. It has been noted that telomere length becomes progressively shorter in cells as they age. In this study, the women who reported the highest levels of emotional stress had, on average,

shorter chromosome telomeres than the women who looked after children without disability. Thus, reduced telomere length due to increased oxidative stress in the cell, a determinant of longevity of cells, was associated with perceived psychological stress. It was striking that telomere shortening in women presumed to be under the most stress was shorter on average by an amount equivalent to 10 years of aging compared to low-stress women.

Another area of current interest is how genetic factors during prenatal development and childhood may exert an influence on the way the adult responds to stress several decades later. Welberg and Seckl have extensively reviewed the role of steroids and the hypothalamic-pituitary-adrenocortical (HPA) axis as a system involved in prenatal programming in the brain. In animal studies, prenatal stress leads to long-lasting effects on the HPA axis function in adult offspring, in general programming a persistently hyperactive system and producing clear alterations in the animals' behavior and response to stressful situations. Adult rats exposed to prenatal stress show an increase in behaviors reminiscent of human anxiety. Extrapolation from animal studies to human situations requires extreme caution. However, in a number of clinical studies poor maternal nutritional status in pregnancy and low birth weight were correlated with increased blood pressure, obesity, and glucose intolerance later in life. A possible link between such early life stress and altered behavior or psychopathology in the offspring once they reach adulthood remains an interesting conjecture, and steroids in the HPA axis are the focus of considerable research as possible mediators of such a link.

See Also the Following Articles

Gene Environment Interactions in Early Development; Genetic Factors and Stress; Life Events and Health.

Further Reading

Brown, G. V. and Harris, T. O. (1993). Aetiology of anxiety and depressive disorders in an inner-city population 1. Early adversity. *Psychological Medicine* 23(1), 143–154.

- Epei, E. S., Blackburn, E. H., Lin, J., et al. (2004). Accelerated telomere shortening in response to life stress. *Proceeding of the National Academy of Sciences USA* 101(49), 17312–17315.
- Gaspi, A., Sugden, K., Moffitt, T. E., et al. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 301(5631), 386–389.
- Kendler, K. S., Neale, M., Kessler, R., Heath, A. and Eaves, L. (1993). A twin study of recent life events and difficulties. *Archives of General Psychiatry* 50(10), 789–796.
- Kendler, K. S. and Karkowski-Shuman, L. (1997). Stressful life events and genetic liability to major depression: genetic control of exposure to the environment? *Psychological Medicine* 27(3), 539–547.
- Lvons, M. J., Goldberg, J., Eisen, S. A., et al. (1993). Do genes influence exposure to trauma? A twin study of combat. *American Journal of Medical Genetics* 48(4), 22–27.
- Mormede, P., Courvoisier, H., Ramos, A., et al. (2002). Molecular genetic approaches to investigate individual variations in behavioral and neuroendocrine stress responses. *Psychoneuroendocrinology* 27(5), 563–583.
- Noble, E. R., Ozkaragoz, T. Z. and Ritchie, T. L. (1998). D2 and D4 dopamine receptor polymorphisms and personality. *American Journal of Medical Genetics* 88(3), 257–267.
- Seedat, S., Niehaus, D. J. and Stein, D. J. (2001). The role of genes and family in trauma exposures and posttraumatic stress disorder. *Mol Psychiatry* 6(4), 360–362.
- Shors, T. J. (2004). Learning during stressful times. *Learning and Memory* 11(2), 137–144.
- Silberg, J., Pickles, A., Rutter, M., et al. (1999). The influence of genetic factors and life stress on depression among adolescent girls. *Archives of General Psychiatry* 56(3), 225–232.
- Stein, M. B., Jang, K. L., Taylor, S., Vernon, P. A. and Livesley, W. J. (2002). Genetic and environmental influences on trauma exposure and posttraumatic stress disorder symptoms, a twin study. *American Journal of Psychiatry* 159(10), 1675–1681.
- Veenema, A. H., Meijer, O. C., de Kloet, E. R. and Koolhaas, J. M. (2003). Genetic selection for coping style predicts stressor susceptibility. *Journal of Neuroendocrinology* 15(3), 256–267.
- Welberg, L. A. and Seckl, J. R. (2001). Prenatal stress, glucocorticoids and the programming of the brain. *Journal of Neuroendocrinology* 13(2), 113–128.

Genetic Testing and Stress

V Senior and M Cropley

University of Surrey, Guildford, UK

© 2007 Elsevier Inc. All rights reserved.

Overview of Genetic Testing

Stress and Uptake of Genetic Testing

Emotional Impact of Genetic Testing

Cognitive Appraisal of Genetic Testing

Glossary

<i>Autosomal recessive condition</i>	A genetic condition in which both copies of a gene need to be faulty in order for the specific condition to develop.
<i>Carrier testing</i>	A test to identify whether an individual has one copy of the faulty genes that contribute to an autosomal recessive condition.
<i>Human genome project</i>	An international collaboration to map the entire human genome.
<i>Mutation</i>	An inherited change in a sequence of DNA.
<i>Predictive genetic testing</i>	A test to identify whether an asymptomatic individual has a genetic mutation that predisposes him or her to developing a specific disease.

Overview of Genetic Testing

The knowledge generated by the Human Genome Project has already started to revolutionize our understanding of health and illness. It is now possible to identify people who have a genetic predisposition to an increasing number and variety of diseases. This knowledge offers great potential for disease prevention, developing more effective treatments, and, indeed, the hope of treating diseases that are currently incurable. This new understanding of the human genome also brings with it a whole host of ethical, social, and psychological dilemmas. Although genetic testing for certain conditions means that those at risk can reduce their chances of developing the disease, for other conditions genetic testing means that some people have to live with the knowledge that they are almost certain to develop an incurable disease. Genetic testing has implications for decisions about whether to have children and who to have them with, what to do if a fetus is known to have a genetic disease, and whether individuals will be discriminated against when seeking employment or insurance.

In this article, we focus on the role of stress in predictive genetic testing. This is because most of the well-conducted psychological evaluations conducted to date that assess stress using validated measures concerned predictive genetic testing. Also, in other forms of genetic testing, such as carrier testing, the psychological issues primarily concern decisions about pregnancy and childbirth. Carrier testing is concerned with identifying people who have a copy of a deleterious mutation for an autosomal recessive condition. These people do not develop the disease in question, but if they have children with someone who is also a carrier of the mutation, each child has a 25% chance of having the condition. Examples of these conditions are Tay-Sachs disease, sickle cell disease, and cystic fibrosis. Predictive genetic testing is concerned with identifying people who are at high risk of developing the disease themselves. All of the predictive genetic tests currently available raise different medical, psychological, and ethical issues.

One of the first genetic tests developed was for Huntington disease, an incurable neurodegenerative condition that typically develops in middle age or later life. People who carry the mutation definitely develop Huntington disease at some point, and those who do not carry the mutation definitely do not develop the disease. There is understandable concern that offering at-risk families genetic testing will constitute an enormous psychological burden for those who test positive. There are a number of exemplary longitudinal studies conducted worldwide to evaluate the consequences of genetic testing for these families. There are also an increasing number of psychological evaluations of genetic testing for a predisposition to cancer, particularly breast and ovarian cancer but also forms of colon cancer. In these cases, the psychological issues include the stress that may be associated with being found to be a carrier of a predisposing mutation and behavioral responses to genetic risk information. These behaviors include whether people adhere to recommendations about surveillance to detect possible cancers and also decisions about whether to have surgery to remove tissues likely to be affected prior to the development of any cancer (called prophylactic surgery). In the case of an inherited predisposition to cancer, being a carrier of a mutation increases the cancer risk considerably, typically making the likeliness of developing cancer 60–80%, but it does not make cancer inevitable. In addition, those found not to carry the genetic

mutation still have at least the same risk as the general population of developing cancer, which means it is important that these people do not believe they are not at risk. For other diseases, genetic testing is still in its infancy. For example, it is possible to offer predictive genetic testing for the predisposition to coronary heart disease and Alzheimer's disease. In the case of coronary heart disease, predictive genetic testing can be offered for familial hypercholesterolemia, an inherited predisposition to raised cholesterol levels. Because the treatment for this condition, in the form of cholesterol-lowering statin medication, is very effective, the psychological issues raised pertain particularly to preventative behavior, that is, whether those identified adhere to medication regimens and engage in other risk-reducing behaviors such as eating a low-fat diet and not smoking. Like Huntington disease, Alzheimer's disease is not curable at present. However, genetic testing for Alzheimer's indicates whether an individual has an increased risk for the disease rather than providing the information that an individual definitely will or will not go on to develop it. In this case, it is necessary for those undergoing testing to understand quite complicated probabilities.

Genetic testing programs differ in how they are organized. In some cases, everyone within a particular population may be offered genetic testing. Such population-based screening programs are most common for detecting carriers of recessive genes and for the prenatal and neonatal screening of infants. More common among predictive genetic testing programs is some form of cascade screening. These programs start with an individual who either has the disease or is at known risk (called the proband) and, if a deleterious mutation is identified in this individual, genetic testing is then offered to other biological relatives.

All these factors have an impact on the emotional state, risk perceptions, and expectations of those offered genetic testing and receiving the results of genetic tests. An interactional model of stress is widely accepted within the stress community, whereby stress is conceptualized as a process involving an interaction of individuals and their environment. One such example is the transactional model of stress and coping proposed by Lazarus and Folkman in the 1980s. Using this framework, we address three main questions with respect to genetic testing: does stress influence the decision to have a genetic test, is receiving the results of genetic testing stressful, and does genetic testing influence secondary appraisal processes and coping strategies? Each of these issues is now considered in turn.

Stress and Uptake of Genetic Testing

It is widely recognized that feelings of stress about health lead people to be more likely to engage in risk-reducing behaviors such as stopping smoking, taking exercise, and adopting a healthy diet. The picture with regard to disease-detecting behaviors, of which having a genetic test is one example, is less clear cut. Disease-detecting behaviors are actions that can inform the individual whether they have or are likely to develop a particular disease but that, in themselves, do not change the disease or risk status of the individual. Similar to models of stress and behavioral performance in nonhealth domains, it has been proposed that a certain amount of perceived stress motivates the performance of disease-detection behaviors. However, high levels of perceived stress are thought to lead to a greater likelihood of fear-control strategies and the subsequent avoidance of the behavior.

Only a proportion of those offered genetic testing will decide to have the test. There are a wide variety of factors influencing this decision, including the disease being tested for and whether treatment is available, sociodemographic factors, and the expectations and motivations of the individual. Whether greater perceived stress facilitates or inhibits the uptake of the offer of genetic testing has not been widely investigated. What is known is that people who believe they can cope with the results of genetic testing are the ones who have a greater interest in and uptake of genetic testing. In addition, people's perceiving themselves as being at greater risk for having the gene and developing the disease is typically associated with a greater uptake of genetic testing. In one of the early reports of genetic testing for Huntington disease, the Canadian Collaborative Study of Predictive Testing group found that participants who declined genetic testing (along with those who received uninformative test results) were the most distressed over a period of 12 months. In 1997 Caryn Lerman and colleagues explicitly tested the influence of distress on requesting the results of genetic testing for hereditary breast and ovarian cancer susceptibility. After controlling for sociodemographic factors and objective risk, those with higher levels of cancer-specific distress were nearly three times more likely to request the results of genetic testing. No evidence of a curvilinear (inverted-U) relationship between distress and genetic testing was found, although this may be because, on the whole, participants did not have high levels of distress. Although there are inherent difficulties in investigating whether stress predicts the uptake of genetic testing given that people who decline the offer of genetic testing are less likely to want to take

part in psychological evaluations, there is clearly a need to attempt to understand this relationship. As genetic testing becomes more ubiquitous, there is a need to find out whether those who are offered genetic testing and decline it constitute a particularly vulnerable group.

Emotional Impact of Genetic Testing

Given the justifiable concern that finding out one has a high risk of developing a serious genetic disease may prove to be a huge burden, it is unsurprising that the most widely investigated issue over the past 20 or so years has been the emotional impact of receiving genetic test results. The original studies continue to report the long-term emotional effects of testing, and more studies are needed as new genetic testing procedures for various diseases are implemented. As already discussed, having an uninformative genetic test result for Huntington disease (along with declining the genetic testing) was found to be more distressing than receiving a positive result. An uninformative genetic test result is one in which it is not possible to definitively state whether or not the individual is a carrier of a deleterious mutation. In this study, the higher distress of those receiving uninformative results was quite possibly due to the motivation for testing. Many people who seek predictive genetic testing do so hoping to resolve the uncertainty they have been living with, and uninformative results fail to help them achieve this goal. Most studies, however, report the impact of positive and negative genetic test results and these are considered next.

Negative Genetic Test Results

As might be expected, across different genetic tests, a consistent finding is that immediately after individuals learn they do not have the predisposing gene they experience psychological relief. Thus levels of anxiety, depression, distress, and illness-specific measures of distress typically drop to significantly lower levels than they were at the time when baseline assessments were made. Some studies find that these lower levels of psychological stress are maintained over the coming months and years, whereas other studies find that the level of stress slowly rises back to the baseline level. This latter finding may well be explained by the fact that the threat of disease is but one of the problems that individuals face. Despite the relief invoked by not being a carrier of the disease-causing gene, other problems remain unresolved and quickly start to impinge on the individuals' emotional state. In addition, these people come from families affected by serious diseases and still need to cope with this

situation and, possibly, the feelings of guilt that others are affected when they are not.

A consistent finding across studies is that distress prior to testing is the biggest predictor of distress after testing, regardless of the test result. A positive effect on the mean stress levels in individuals who receive a negative test result can hide the fact that some individuals have problems coping with testing negative for the gene. These people may have lived their lives and made important decisions (such as not having children) on the basis of their expectation of developing the disease in the future. Almqvist and colleagues recently reported that one person who tested negative for the gene for Huntington disease attempted suicide on the day he received the test result. Therefore, it is important that it not be assumed that all patients who receive negative genetic test results do not require additional support and counseling.

Positive Genetic Test Results

Some studies found that, immediately after receiving positive genetic results, people's levels of distress are no different than their baseline levels. Other studies found increased levels of distress in this group. An unexpected and important finding is that any increased distress is usually transitory. As early as 1 month and for at least 1 year after being given this bad news, the levels of distress are consistently no different to those reported for the baseline levels. In some studies, the levels of distress up to 1 year later are no different than the levels reported by those receiving negative test results. As with the finding for negative test results, a concentration on the mean scores can hide the fact that some individuals do experience great and long-term difficulties coming to terms with the information that they are at high risk of developing a serious disease. There is a need to ensure that these individuals, in particular, receive appropriate support in the months and years following the receipt of genetic test results. Because distress prior to testing predicts distress following testing, the process of identifying potentially vulnerable individuals can be put in place as soon as they come forward for testing.

Huntington disease was the first disease for which a predictive genetic test was developed, and these studies were also the first to report the results of long-term psychological evaluations. These are crucial because they show whether those who receive positive genetic tests start to experience greater distress when they approach the time when they can expect to develop symptoms. In 2003, Almqvist and colleagues reported that for those who carry the gene for Huntington disease the levels of distress,

which were reduced soon after testing and up to 2 years later, had increased at the 5-year assessment. However, this increase was no higher than the baseline levels of distress. Although some participants did report distress levels in the clinical range, the mean scores were below the clinical cut-off at all the times measured. Therefore, there are reasons to be optimistic that, for the majority of people, genetic testing does not result in psychological harm.

Many authors have noted that there are a variety of explanations to account for the finding that positive test results do not, on the whole, lead to long-term emotional difficulties. Most of the people undergoing predictive genetic testing have lived most of their lives with the knowledge that their family members are affected by this disease and that they are at risk. Thus, many enter genetic testing programs with the hope of resolving uncertainty about their own risk status and that of their relatives. In many contexts, living with uncertainty can be more stressful than certainty. In addition, only a proportion of those who are at risk decide to have genetic tests, and, as stated previously, these are typically those who feel they can cope with the results. Finally, many of the genetic testing programs evaluated to date have at their core the provision of excellent information and counseling, and this will make no small contribution to the psychological well-being of recipients of genetic testing. There is a need to ensure that, when genetic testing is expanded to other clinical environments and other diseases for which people may have less awareness of their own risk, it is still accompanied by the provision of high-quality information and support.

Cognitive Appraisal of Genetic Testing

There are rapid developments in the understanding of the genetic basis of multifactorial diseases such as cancer, cardiovascular disease, and diabetes, in which both genetic and behavioral factors can contribute to developing the disease. In this context, whether genetic testing influences the cognitive appraisal of the disease becomes a key issue. If having a genetic test leads people to appraise the disease differently, this may have implications for whether they will engage in behaviors that can reduce their risk.

In the general public, studies suggest that there is a common perception that the genetic risk of disease is less controllable than other risk factors. From a stress and coping perspective, the appraisal of a health threat as uncontrollable could initiate different coping procedures than its appraisal as controllable. Specifically, genetic risk information could invoke fatalism and could demotivate the individual from engaging in risk-reducing behaviors. This hypothesis

was tested – and was not supported – in a UK trial of genetic and nongenetic diagnoses of familial hypercholesterolemia. In this study, genetic testing appeared to influence the cognitive appraisal of the disease by reinforcing a biological model of the disease, such that the more genes were endorsed as contributing to coronary heart disease, the more medication was endorsed as an effective means of reducing risk. Thus, there is some preliminary evidence that genetic risks are indeed appraised differently and may lead to greater engagement in biological rather than behavioral ways of reducing risk. As with the majority of the studies conducted to date, this research was conducted with participants already highly aware of their increased risk of disease.

Some studies have found that genetic test results do have an impact on behavior, such as an increased adherence to surveillance behaviors in those with an inherited predisposition to cancer. Indeed, some people who are found not to have a genetic predisposition to cancer find it difficult to accept that their risk is at the same level as the general population and still desire regular screening to detect cancer. To date there is little evidence to suggest that such populations feel falsely reassured that a negative genetic test means they have no risk of developing the disease. Some people who do have an inherited predisposition to cancer opt to have prophylactic surgery (e.g., a mastectomy before any cancer has developed). Although this is clearly a difficult decision, the option to undergo surgery is an understandable coping strategy for those who have seen family members affected by cancer. Other studies, however, have not found any behavioral effects of genetic testing.

Given the equivocal findings across the few studies to date that have investigated behavioral outcomes, there is a need for improving our understanding of the mechanisms that may explain these differences. A number of authors are beginning to conceptualize the impact of genetic testing from a stress and coping and from a self-regulatory perspective. These conceptualizations should pave the way for developing an evidence base of the mechanisms by which cognitive and emotional appraisal and self-regulatory processes impact on behavior following genetic testing. This developing evidence base will lead to a greater certainty about how best to implement genetic testing for multifactorial diseases. In particular, as genetic testing for common diseases such as cancer and coronary heart disease becomes more widespread, there is a need to investigate the impact of genetic testing on cognitive appraisal and behavior in populations that have less prior awareness of their risk. These populations may have very different emotional responses and cognitive appraisals of their risk status than the

highly aware populations studied so far. Based on the evidence to date, however, there is cause for cautious optimism that if attention is paid to providing high-quality information and support to at-risk families the potential medical benefits of genetic testing can be realized without causing emotional or behavioral harm.

Further Reading

- Almqvist, E. W., Brinkman, R. R., Wiggins, S., et al. (2003). Psychological consequences and predictors of adverse events in the first 5 years after predictive testing for Huntington disease. *Clinical Genetics* **64**, 300–309.
- Bleiker, E. M. A., Hahn, D. E. E. and Aaronson, N. K. (2003). Psychosocial issues in cancer genetics: current status and future directions. *Acta Oncologica* **42**, 276–286.
- Broadstock, M., Michie, S. and Marteau, T. (2000). Psychological consequences of predictive genetic testing: a systematic review. *European Journal of Human Genetics* **8**, 731–738.
- Butow, P. N., Lobb, E. A., Meisser, B., et al. (2003). Psychological outcomes and risk perception after genetic testing and counselling in breast cancer: a systematic review. *Medical Journal of Australia* **178**, 77–81.
- Gooding, H. C., Organista, K., Burack, J., et al. (2006). Genetic susceptibility testing from a stress and coping perspective. *Social Science & Medicine* **62**, 1880–1890.
- Lerman, C., Croyle, R. T., Tercyak, K. P., et al. (2002). Genetic testing: psychological aspects and implications. *Journal of Consulting and Clinical Psychology* **70**, 784–797.
- Lerman, C., Schwartz, M. D., Narod, S., et al. (1997). The influence of psychological distress on use of genetic testing for cancer risk. *Journal of Consulting and Clinical Psychology* **65**, 414–420.
- Marteau, T. M. and Richards, M. (eds.) (1996). *The troubled helix: social and psychological implications of the new human genetics*. Cambridge, UK: Cambridge University Press.
- Marteau, T. M. and Weinman, J. (2006). Self-regulation and the behavioural response to DNA risk information: a theoretical analysis and framework for future research. *Social Science & Medicine* **62**, 1360–1368.
- Senior, V. and Marteau, T. M. (in press). Causal attributions for raised cholesterol and perceptions of effective risk-reduction: self-regulation strategies for an increased risk of coronary heart disease. *Psychology & Health*.

Genetic Variation of HPA Axis Activity and Function in Farm Animals

P Mormède

INRA – Université Victor Segalen Bordeaux 2,
Bordeaux, France

© 2007 Elsevier Inc. All rights reserved.

Evidence for Genetic Variability
Sources of Genetic Variability
Functional Consequences of Genetic Variability
Perspectives

Glossary

Corticosteroid-binding globulin (CBG) A glycoprotein specifically carrying glucocorticoid hormone in blood plasma.

Genotype The specific genetic makeup (the specific genome) of an individual. It codes for the phenotype of that individual.

Heritability The proportion of the observed variation in a particular phenotype and, in

Hypothalamic-pituitary-adrenal (HPA) axis
Phenotype

a particular study, that can be attributed to the contribution of genotype (inheritance).

The neuroendocrine system responsible for glucocorticoid hormone secretion.

A specific manifestation of a trait that varies between individuals. The phenotype is determined to some extent by genotype, or by the identity of the alleles that an individual carries at one or more positions on the chromosomes. Quantitative phenotypes are determined by multiple genes and are influenced by environmental factors.

Quantitative trait locus (QTL)

A polymorphic site on a chromosome containing alleles that differentially influence the expression of a quantitative trait.

Single nucleotide polymorphisms (SNPs)

Stable mutations consisting of a change at a single base in a DNA molecule. They are the most common type of genetic variation.

Evidence for Genetic Variability

A wide range of variability has been observed in hypothalamic-pituitary-adrenal (HPA) axis function, both basal and in response to stress. The contribution of genetic factors is shown by the differences among genetically defined stocks. Most available data have been obtained from pigs. Wild boars have the highest levels of cortisol, followed by Chinese breeds, including the most-studied Meishan. The highly selected lean breeds Large White and Landrace have the lowest levels, and Piétrain and Duroc pigs show intermediate levels. Few data are available in other species.

If a trait is influenced by genetic factors, it should be possible to increase or decrease the value of this trait by genetic selection. Indeed, divergent lines have been obtained by genetic selection in poultry (response to adrenocorticotropin [ACTH] or social stress), Japanese quail (response to immobilization), turkey (cold stress), and trout (confinement stress).

Sources of Genetic Variability

This article will not deal with major gene effects on HPA axis activity as seen in endocrine diseases induced by corticosteroid receptor or enzymatic defects. Traits such as circulating cortisol levels are normally distributed in genetically heterogeneous populations, suggesting that several genes may influence the phenotype, each gene with a small effect, and with complex gene $\times$ gene and gene $\times$ environment interactions. Several sources of genetic variability have been identified in farm animals.

Adrenal Sensitivity to ACTH

In the 1980s, D. P. Hennessy (Victoria, Australia) demonstrated in pigs that the increase of plasma cortisol after ACTH injection was variable among individuals but stable across time for a given animal. Similar differences in cortisol secretion were shown in response to corticotropin-releasing hormone (CRH), physical exercise, or insulin-induced hypoglycemia, although the release of ACTH was not different among individuals. Furthermore, the metabolic clearance of cortisol bears no relationship with the response to ACTH. Altogether, these data demonstrate that the key index of individual differences in HPA axis activity is the adrenal sensitivity to ACTH that was shown to be heritable in pigs. Several experimental results confirm genetic influences on the sensitivity of adrenal glands to ACTH. For instance, Meishan pigs (high cortisol levels) have a higher response to ACTH than Large Whites (low cortisol). In poultry, divergent lines could be selected on this trait. In quails,

selection for high corticosterone to immobilization increased the steroidogenic response of adrenocortical cells to ACTH, and in turkeys, selection for a high corticosterone response to cold stress also increased the sensitivity of the adrenal gland to ACTH. In trout, the divergent cortisol response to confinement stress mostly results from the selection of animals with divergent interrenal gland response to ACTH, together with a large difference in the expression level of genes involved in corticosteroid hormone synthesis. Taken together, these data show that the adrenal sensitivity to ACTH is an important mechanism by which genetic factors influence HPA axis response to stress.

Hormone Bioavailability

Quantitative trait locus (QTL) analysis is a strategy to search for gene polymorphisms influencing a quantitative trait. It is based on the linkage between phenotypic value and genetic polymorphisms in a segregating population, usually an F2 intercross or a backcross. A QTL analysis run on an F2 intercross between the Large White and Meishan breeds revealed the large influence of a locus located at the end of the q arm of chromosome 7 on basal and post-stress (novel environment exposure) levels of cortisol, explaining, respectively, 7.7% and 20.7% of the phenotypic variance, with the Meishan alleles increasing cortisol levels. Other loci on chromosome 1 and 17 were shown to influence poststress ACTH levels. The comparison of genetic maps from different species shows that the end of the long arm of chromosome 7 is homologous to the telomeric end of the long arm of human chromosome 14. In this region corticosteroid-binding globulin (CBG), a glycoprotein specifically carrying cortisol in blood plasma, has been mapped. CBG is an important factor influencing bioavailability of the hormone, as it has been shown in humans that differences in CBG levels are a major determinant of differences in plasma cortisol levels. Therefore, the *Cbg* gene is an interesting positional and functional candidate gene to explain QTL influences on plasma cortisol levels. Indeed, comparison of gene sequences in Large White and Meishan pigs revealed several polymorphisms influencing CBG levels and function.

Corticosteroid Hormone Receptors

Two receptors (mineralocorticoid receptor [MR] and glucocorticoid receptor [GR]) mediate the biological effects of corticosteroid hormones. They act as transcription factors to increase or decrease the expression of target genes. The efficiency of corticosteroid hormone receptors is a major site of genetic variation, but few data are available in farm animals.

Functional Consequences of Genetic Variability

Physiopathological consequences of genetic variations in HPA axis activity reflect the wide range of corticosteroid hormone functions. Most data are related to growth and carcass composition. Indeed, glucocorticoid hormones influence feeding behavior; they also exert catabolic effects in peripheral tissues and anabolic effects in the liver so that they reduce growth rate and promote the accumulation of fat at the expense of proteins. In pigs, breeds with a higher body fat content such as Duroc and Meishan also produce more cortisol than lean breeds such as Large White and Landrace. In a segregating population (F2) of Large White $\times$ Duroc pigs, the fat content of the carcass was positively correlated with urinary levels of cortisol measured in the bladder after slaughter. In other studies in pigs, the magnitude of the adrenal response to ACTH – the main mechanism of genetic variability in HPA axis activity – was shown to be negatively correlated with body weight and growth rate. In sheep and turkeys, selection for leanness is accompanied by a reduction of HPA axis reactivity. Conversely, selection for a reduced adrenal response to cold stress in turkeys also increased body weight and egg production in turkeys, and growth rate was also inversely related to cortisol levels in divergently selected lines of rainbow trout.

Glucocorticoid hormones are also potent modulators of the immune system, and data obtained in chickens illustrate the consequence of genetic variation in HPA axis activity/reactivity on resistance to parasites or infectious diseases.

Perspectives

Current research in molecular genetics aims at the identification of gene polymorphisms responsible for individual phenotypic variation. For instance, there are documented mutations in the *Cbg* gene modifying its functional properties and associated with variable carcass fat content and meat quality of pigs. A single nucleotide polymorphism in the *CRH* gene was found to influence rib-eye area and carcass weight in beef cattle. In the future, it should be possible to optimize the functioning of the HPA axis by genetic selection based on molecular polymorphisms. When considering growth rate and carcass fat content, a low level of cortisol production should be favorable in the context of modern breeding for

meat production. However, the picture is much less clear when considering the influence of HPA axis on other phenotypes. For instance, a lower cortisol response to social stress in poultry can either increase or decrease resistance depending on the infectious agent, and recently it was shown that plasma cortisol levels measured at birth are the best predictor of the genetic merit for piglet survival, with highest levels increasing survival. Further research will be necessary to better define the optimal activity of the HPA axis and to refine the molecular tools for farm animal selection.

See also the Following Articles

Adrenal Cortex; Comparative Anatomy and Physiology; Corticosteroid Receptors; Corticosteroid-Binding Globulin (Transcortin); Hypothalamic-Pituitary-Adrenal; Obesity, Stress and; Genetic Polymorphisms in Stress Response.

Further Reading

- DeRijk, R. and de Kloet, E. R. (2005). Corticosteroid receptor genetic polymorphisms and stress responsivity. *Endocrine* **28**, 263–270.
- Desautels, C., Sarrieau, A., Caritez, J. C., et al. (1999). Behavior and pituitary-adrenal function in large white and Meishan pigs. *Domestic Animal Endocrinology* **16**, 193–205.
- Gayraud, V., Alvinerie, M. and Toutain, P. L. (1996). Inter-species variations of corticosteroid-binding globulin parameters. *Domestic Animal Endocrinology* **13**, 35–45.
- Geverink, N., Foury, A., Plastow, G. S., et al. (2006). Cortisol-binding globulin and meat quality in five European lines of pigs. *Journal of Animal Science* **84**, 804–811.
- Kino, T. and Chrousos, G. P. (2004). Glucocorticoid and mineralocorticoid receptors and associated diseases. *Essays in Biochemistry* **40**, 137–155.
- Mormède, P., Courvoisier, H., Ramos, A., et al. (2002). Molecular genetic approaches to investigate individual variations in behavioral and neuroendocrine stress responses. *Psychoneuroendocrinology* **27**, 563–583.
- Ousova, O., Guyonnet-Duperat, V., Iannuccelli, N., et al. (2004). Corticosteroid binding globulin: a new target for cortisol-driven obesity. *Molecular Endocrinology* **18**, 1687–1696.
- Plastow, G. S., Carrion, D., Gil, M., et al. (2005). Quality pork genes and meat production: a review. *Meat Science* **70**, 409–421.
- Tempel, D. L. and Leibowitz, S. F. (1994). Adrenal steroid receptors: interactions with brain neuropeptide systems in relation to nutrient intake and metabolism. *Journal of Neuroendocrinology* **6**, 479–501.

Genetics of Stress See: Genetic Factors and Stress; Genetic Predispositions to Stressful Conditions; Gene Environment Interactions in Early Development; Genetic Variation of HPA Axis Activity and Function in Farm Animals; Genetic Polymorphisms in Stress Response; Genetic Testing and Stress.

Ghrelin and Stress Protection

**T Brzozowski, M Pawlik, D Drozdowicz,
Z Sliwowski, S J Konturek and W W Pawlik**
Jagiellonian University Medical College, Cracow, Poland
P C Konturek
University Erlangen-Nuremberg, Erlangen, Germany

© 2007 Elsevier Inc. All rights reserved.

Pathogenesis of Stress-Induced Gastric Ulcerations
Discovery, Receptors, and Physiological Functions
of Ghrelin

Role of Ghrelin in Maintenance of Gastric
Mucosal Integrity

Mechanism of Gastroprotection Induced by Ghrelin

Neural Aspects of Gastroprotective Activity of Ghrelin

Ghrelin Attenuates Inflammation and the Generation of
Reactive Oxygen Metabolites

*Nitric oxide
(NO)*

*Prostaglandins
(PG)*

Prostanoids

*Vascular endo-
thelial growth
factor (VEGF)*

synapses) and neurohormones (when released into the bloodstream)

Intracellular gaseous mediator released from vascular endothelium, sensory nerves, and gastrointestinal epithelial cells. It contributes to the mechanism of gastrointestinal mucosal integrity by increasing gastrointestinal microcirculation and protective mucus and bicarbonate secretion.

Metabolites of enzyme cyclo-oxygenase activity, the derivatives of arachidonic acid that exhibit cytoprotective activity.

A group of compounds derived from unsaturated 20-carbon fatty acids, primarily arachidonic acid, via the cyclo-oxygenase pathway. They are extremely potent mediators of a diverse group of physiological processes. These compounds play an essential role in the control of gastrointestinal circulation and gastrointestinal mucosal defense. A group of arachidonic acid metabolites which include mainly prostaglandins and thromboxanes.

One of the important growth factors, responsible for angiogenesis and acceleration of ulcer healing and process of gastrointestinal mucosal repair.

Glossary

D cells Endocrine cells belonging to the amine precursor uptake and decarboxylation (APUD) system known to release the hormone, somatostatin.

Endocrine cells (EC) Release the hormone motilin and the neurotransmitter serotonin.

Enterochromatoffin-like cells (ECL) Known to release histamine, a potent stimulant of gastric acid secretion in the stomach. They are the predominant endocrine cell-type of the oxyntic (acid-producing) mucosa of the stomach. Histamine acts as the positive paracrine stimulator of the release of hydrochloric acid from the parietal cell.

Gastric blood flow (GBF) Defined as the flow of the blood through submucosal vessels in the gastric mucosa.

Gr cells Ghrelin-producing cells identified in the gastric mucosa.

N^G-nitro-L-arginine (L-NAME) A nonselective inhibitor of the enzyme nitric oxide synthase. It has been used experimentally to induce hypertension.

Neuropeptides Small proteins (peptides) comprised of a chain of amino acids that are synthesized in, and secreted by, nerves and serve as neurotransmitters (when released at

Pathogenesis of Stress-Induced Gastric Ulcerations

Stress ulcerations are defined as acute gastric mucosal lesions occurring as complications in severely ill patients after burns, sepsis, major surgery, or trauma to the central nervous system. Critically ill patients are at increased risk of developing stress-related gastric mucosal lesions and gastrointestinal bleedings. Stress can induce acute gastric mucosal lesions by a complex of psychological factors, psychosis, or brain trauma, influencing individual vulnerability and specific brain pathways regulating autonomic functions. These stress-induced gastric microbleeding erosions are accompanied by a marked decrease in the gastric blood flow (GBF), a substantial decrease in the

gastric mucosal prostaglandin (PG) content and nitric oxide (NO) release, inhibition of gastric mucus and bicarbonate secretion, and the increased formation of free oxygen radicals in the gastric mucosa leading to gastric ischemia and inflammation. Among various stress models used in animals, the most reproducible results can be obtained by water immersion and restraint stress (WRS), which appear to act synergistically in the production of gastric ulcerations. Numerous studies have demonstrated the usefulness of cold-restraint stress and proved it to be a clinically relevant experimental model for studying of pathomechanism of acute gastric damage and protection. It has been shown that gastric acid plays an important role in the development of stress-induced gastric ulcer and is the most common endogenous factor responsible for the destruction of epithelial cells. Gastric mucosal lesions induced by different time-dependent exposures to WRS trigger time-course dependent changes of the ultrastructure of gastric parietal cells secreting gastric acid. Many physiological agents such as PG, NO donors, calcitonin gene-releasing factor (CGRP), and gastrointestinal hormones such as cholecystokinin (CCK) and leptin have been reported to attenuate the gastric erosions induced by stress of different origin and durations. The mechanism of the recovery of gastric mucosa from stress injury is not fully understood and the underlying healing of stress-induced gastric lesions appears to be multifactorial. It has been proposed that mucosal integrity and mucosal repair after stress damage, involve the enhancement of GBF and increased cell proliferation mediated by expression and subsequent release of growth factors such as epidermal growth factor (EGF) and transforming growth factor (TGF)- α .

Discovery, Receptors, and Physiological Functions of Ghrelin

Ghrelin is a recently described 28-amino acid peptide that has been discovered in rat and human gastrointestinal tracts, particularly in gastric mucosa, as an endogenous ligand for growth hormone (GH) secretagogue receptor (GHSR). Ghrelin stimulates food intake and body weight gain exerting a modulating effect on energy expenditure acting through afferent nerves and directly on hypothalamic feeding centers. This peptide was also shown to enhance the gastric motility and gastric secretion. Recent studies revealed that ghrelin is produced by special neuroendocrine cells located mainly in oxyntic mucosa but not in enterochromatoffin-like cells (ECLs), endocrine cells (ECs) or D cells. Ghrelin has no relevant homology with any known gastrointestinal peptides but

potently stimulates the release of GH and acts as a natural ligand for the GHSR which is the endogenous counter of the family of synthetic, peptidyl and non-peptidyl GH secretagogues. Ghrelin stimulates food intake and body weight gain exerting a modulating effect on energy expenditure. Administration of ghrelin causes weight gain via appetite stimulation and reduced fat oxidation. The release of ghrelin may be influenced by the status of fasting and nutrient feeding because central and peripheral administration to rats results in an increase in their feeding behavior. Previous studies in humans revealed that gastrectomy produced a dramatic fall in the plasma ghrelin levels, whereas fasting and anorexia nervosa were accompanied by elevated plasma ghrelin concentrations supporting the notion that the gastrointestinal tract, primarily the stomach, is a major source of circulating ghrelin, which could be considered as a starvation-related hormone. Interestingly, obestatin, which is encoded by the same gene as ghrelin has recently been reported to counteract physiological effects of ghrelin.

A recent study revealed that endocrine Gr cells of the stomach are a major source of circulating ghrelin acting via activation of two receptor subtypes. Two GHS-R subtypes are generated by alternative splicing of a single gene: the full-length type 1a receptor (GHS-R1a); and a carboxyl-terminally truncated GHS-R type 1b (GHS-R1b). The GHS-R1a is the functionally active, signal-transducing form of the GHS-R, while the GHS-R1b is devoid of high-affinity ligand binding and signal transduction activity. Ghrelin molecules, produced by endocrine cells of gastric glands exist in two major molecular forms, ghrelin and des-*n*-octanoyl ghrelin (des-acyl ghrelin). The acylation by *n*-octanoic acid of the hydroxyl group of their third residue, which is either serine or threonine, is essential for binding of ghrelin to GHS-R1a. The anorexia associated with cancer chemotherapy in rodents is reported to be alleviated by exogenous ghrelin and similar effects have been described in humans. This may be due to a reduction in activity of ghrelin-producing cells, present predominantly in the oxyntic mucosa.

Role of Ghrelin in Maintenance of Gastric Mucosal Integrity

Little is known about the factors that might affect ghrelin release in the stomach and whether this peptide can contribute the mechanism of gastric mucosa integrity. The role of ghrelin in the mechanism of gastric mucosal defense and gastroprotection has been investigated by revealing that peripheral administration of ghrelin reduces the formation of lesions induced by ethanol and cold stress ([Figure 1](#)).

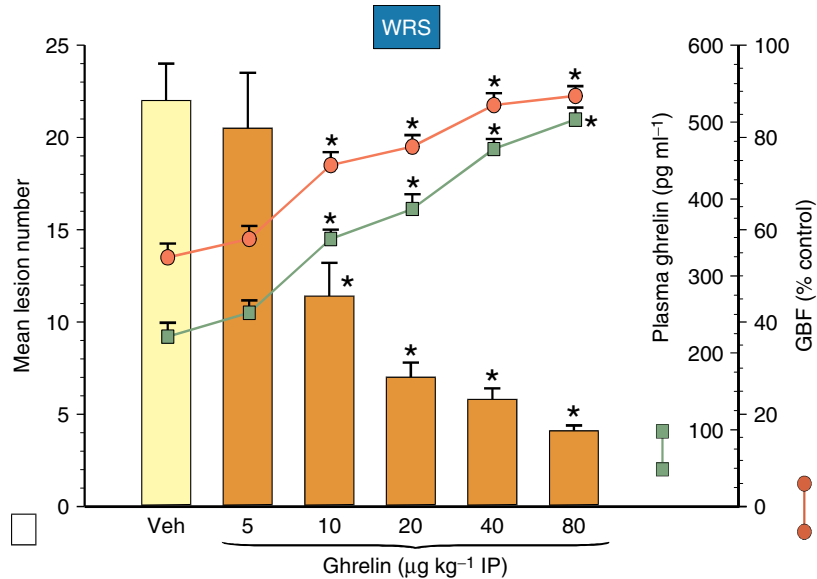


Figure 1 The mean number of water immersion and restraint stress (WRS)-induced gastric lesions, gastric blood flow (GBF), and plasma immunoreactivity of ghrelin in rats treated with vehicle (saline) or with various doses of ghrelin (5–80 µg kg⁻¹ intraperitoneal). Please note a significant and dose-dependent reduction in the number of gastric lesions induced by WRS accompanied by a significant rise in the GBF and plasma ghrelin levels. Means ± standard error of means of 6–8 rats. Asterisk indicates a significant change as compared to the vehicle control values. Reproduced from *Regulatory Peptides* 120, Exogenous and endogenous ghrelin in gastroprotection against stress-induced gastric damage. 39–51, 2004, with permission from Elsevier.

In agreement with the original hypothesis that prostanooids, NO, and sensory neuropeptides cooperate in the mechanism of maintenance of gastric integrity, it was proposed that NO and sensory neuropeptides may mediate these gastroprotective effects because the blockade of NO synthase (NOS) activity with N^G-nitro-L-arginine (L-NAME) and the functional ablation of sensory afferent nerves with capsaicin were both found to attenuate them. It is of interest that centrally applied ghrelin can influence gastric acid secretion, while attenuating lesions caused by stress and this effect is mediated by the vagal innervation and the expression of calcitonin gene related peptide (CGRP) messenger ribonucleic acid (mRNA), an important neuropeptide released from sensory nerve endings. Both these events are considered as important components of the brain-gastrointestinal tract axis involved in the gastroprotective effects of ghrelin especially against stress-induced gastric lesions.

In contrast, the *Helicobacter pylori* (*H. pylori*) infection of the human gastric mucosa which is now considered to play a role as the causal factor in the pathogenesis of gastritis and peptic ulcer was found to attenuate the mucosal expression and release of ghrelin and to reduce appetite. This was confirmed in *H. pylori*-infected gastric mucosa of Mongolian gerbils, which is now widely accepted as an appropriate model to study the mechanism of *H. pylori* infection under experimental conditions. It is of interest, that after successful cure of *H. pylori* in

healthy asymptomatic subjects the plasma ghrelin concentration increased, which in turn led to enhanced appetite and weight gain. This implies that increase in ghrelin in *H. pylori* cured gastric mucosa may contribute to the increasing obesity seen in Western populations, where the prevalence of *H. pylori* is low.

Mechanism of Gastroprotection Induced by Ghrelin

Recently, endogenous PGs have been implicated in the control of food intake and appetite but the possibility that these cytoprotective arachidonic acid metabolites could also play an important role in the gastroprotective effect of ghrelin has not been explored. Moreover, the question remains whether ghrelin contributes to gastroprotection against gastric lesions caused not only by artificial irritants such as ethanol, but also can protect against those caused by vascular disturbances resulting from stress or ischemia-reperfusion (I/R) that lead to severe microbleeding erosions and the fall in the microcirculation (Figure 2). Therefore it was important to confirm whether endogenous PGs as well as the expression of cyclo-oxygenase (COX)-1 and COX-2 are involved in the possible gastroprotective activity of ghrelin against stress-induced gastric erosions.

Our study supports the notion that ghrelin-induced protection and hyperemia involves co-activation of NO and sensory afferents and endogenous PG.

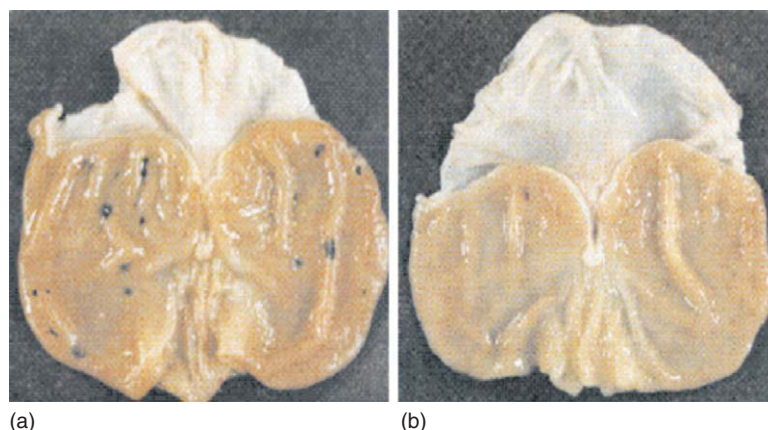


Figure 2 Representative photomicrograph showing gross appearance of gastric mucosa of rat pretreated 30 min before the exposure to 3.5 h of water immersion and restraint stress (WRS) with a, vehicle (1 ml of saline intraperitoneal (IP)) or b, ghrelin ($20 \mu\text{g kg}^{-1}$ IP). Please note that many bleeding erosions are observed macroscopically in vehicle-pretreated gastric mucosa as compared to that treated with ghrelin indicating a protective activity of this hormone against the formation of WRS-induced gastric lesions.

Arachidonic acid metabolites were believed to act as the classic mediators of cytoprotection but recent studies revealed that in contrast to NO, endogenous PG do not appear to contribute to the observed gastroprotection by peptides such as leptin and CCK against ethanol-induced gastric lesions. Questions arose whether the suppression of COX by nonselective COX inhibitor, indometacin, or the highly selective COX-2 inhibitor, rofecoxib, could influence the gastroprotective and hyperemic activity of ghrelin. It has been documented that ghrelin-induced protection and hyperemia is accompanied by the enhancement in the mucosal PGE_2 generation. Both, indometacin and rofecoxib greatly attenuated the protective and hyperemic effects of ghrelin, indicating that endogenous PG possibly derived from COX-1 and COX-2 PG pathway may mediate these beneficial effects of this hormone on the stomach (Figure 3).

Ghrelin applied centrally exhibits comparable gastroprotective activity as when administered peripherally against the mucosal damage induced by corrosive substance such as ethanol and nontopical ulcerogen such as stress. This suggests that exogenous ghrelin could be considered as an important protective factor for the gastric mucosa. This notion is supported by observation that ghrelin attenuated the gastric lesions induced by ethanol in various concentrations and those provoked by stress while producing an increase in gastric mucosal blood flow and the plasma level of this hormone. The results of secretory studies revealed that ghrelin applied by the intracerebroventricular (ICV) and intraperitoneal (IP) routes in doses that were gastroprotective against stress and ethanol injury significantly raised gastric acid secretion suggesting that the protective effect of this peptide occurred despite an increase of the gastric secretory function,

thus representing genuine gastroprotective activity as previously proposed for the phenomenon of cytoprotection. The ghrelin-induced protection after its central administration was accompanied by a significant and dose-dependent rise in the plasma ghrelin concentrations by WRS suggesting that ghrelin-evoked gastric hyperemia could be an important mechanism of the protective effect of this peptide in rat stomachs (Figure 1).

It has been shown that ghrelin, primarily produced in the stomach, but also in the hypothalamus, is an orexigenic peptide and affects both feeding at the levels of hypothalamus (arcuate nuclei) and GH secretion partly due to an activation of gastric vagal afferents transmitting visceral sensory signals to the brain. The observation that ghrelin stimulates gastric acid secretion is in keeping with the original observation that an increase in gastric secretion was achieved after parenteral but not topical administration of ghrelin. The mechanism by which ghrelin increases gastric acid secretion after peripheral as well as central administration could be related to elevation of plasma gastrin concentration, since a considerable increase in this hormone level was notified following ghrelin application. Gastrin could contribute, at least in part, to the secretory as well as protective and hyperemic effects of ghrelin but its effect on histamine release should also be considered. The notion that the protective activity of ghrelin depends, at least in part, on gastrin release is at variance with some studies claiming that ghrelin release does not operate under gastrin; however, these studies failed to directly address the effect of ghrelin on gastrin release. It is not excluded that ghrelin, which is known for the release of GH, acts via GH release to protect gastric mucosa against stress damage as GH was shown to

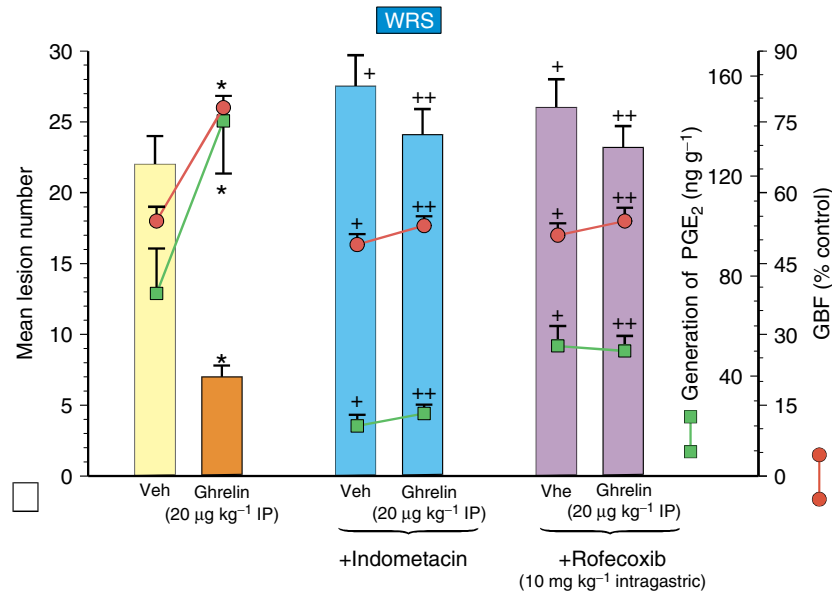


Figure 3 The mean number of gastric lesions induced by water immersion and restraint stress (WRS) and accompanying changes in gastric blood flow (GBF) in rats treated with vehicle (saline) and ghrelin ($20 \mu\text{g kg}^{-1}$ intraperitoneal (IP)) with or without pretreatment with nonselective COX inhibitor, indometacin (5 mg kg^{-1} IP) or COX-2-selective inhibitor, rofecoxib (10 mg kg^{-1} intragastric). Mean $\pm$ standard error of mean of 6–8 rats. Asterisk indicates a significant decrease as compared to the value obtained in vehicle-control animals. Cross indicates a significant change as compared to the respective value obtained in animals without pretreatment with COX inhibitors. Double cross indicates a significant change as compared to the respective value obtained in ghrelin-treated animals without pretreatment with COX inhibitors. Reproduced from *Regulatory Peptides* 120, Exogenous and endogenous ghrelin in gastroprotection against stress-induced gastric damage. 39–51, 2004, with permission from Elsevier.

enhance the healing of gastric ulcers, while increasing plasma gastrin level.

Neural Aspects of Gastroprotective Activity of Ghrelin

Ghrelin is an orexigenic peptide and affects both feeding at the levels of hypothalamus (arcuate nuclei) to stimulate neuropeptide Y (NPY) and Agouti-related peptide (AgRP)-related neurons and to release GH partly due to an activation of gastric vagal afferents transmitting visceral sensory signals to the brain. An experimental study in a stress model of gastric injury revealed that ghrelin mRNA was upregulated in the gastric mucosa exposed to WRS. This was followed by an increase in the plasma ghrelin level indicating that endogenous ghrelin might act as local integrity peptide to limit the extent of gastric damage provoked by WRS. This is similar to leptin, another gastric hormone involved in the control of food intake though acting in opposite direction than ghrelin. As expected, ghrelin-induced protection after its central administration was accompanied by a significant and dose-dependent rise in the plasma ghrelin concentrations and marked attenuation of the fall in the GBF caused by WRS suggesting that ghrelin-evoked gastric hyperemia could be an important

mechanism of the protective effect of this peptide in the rat stomach. This notion is based on the well-documented facts that an appropriate microcirculatory blood supply helps to orchestrate various lines of mucosal defense system in gastric mucosa. In addition, as mentioned before, ghrelin significantly increases the expression of mRNA for CGRP, which is known to increase the gastric mucosal blood flow. The mechanism by which ghrelin improves this gastric mucosal blood flow could be related to a direct effect of ghrelin on blood vessels due to a potent vasodilatory activity of this peptide. It is assumed that the direct protective effect of ghrelin can not be ruled out; however, no study has been yet undertaken to demonstrate that ghrelin exerts direct effects on vasculature.

Since the mechanism of gastric mucosal defense include NO that could be released from vascular endothelium, sensory nerves, or gastric epithelial cells, the hypothesis was tested that ghrelin-induced gastroprotection involving NO could originate from the activation of afferent sensory neurons by this peptide. The deactivation of rat primary afferent nerves, using a neurotoxic dose of capsaicin about 2 weeks before the stress experiment, aggravated WRS-induced gastric damage as compared to rats without capsaicin denervation and significantly reduced the GBF

when compared to that in animals with intact sensory nerves. Moreover, it has been demonstrated that such a capsaicin-induced deactivation of sensory nerves significantly attenuated the gastroprotective activity of central and peripheral ghrelin and completely abolished the ghrelin-induced rise in GBF. Moreover, the replacement therapy with exogenous CGRP, the major neuropeptide released from sensory afferent nerves, restored the protective and hyperemic activity of ghrelin against the WRS-induced gastric lesions (Figure 4). In addition, the expression of mRNA for CGRP, the major neuropeptide released from sensory afferent nerves, was enhanced in ghrelin-treated animals indicating that this major sensory nerve neuropeptide is essential for microcirculatory response and of crucial importance for the gastroprotective activity of ghrelin.

Involvement of sensory nerves in ghrelin-induced gastroprotection does not exclude the possibility that centrally applied ghrelin enhances the central parasympathetic outflow to the stomach and that also vagal efferent nerves are involved in gastroprotection afforded by centrally and peripherally administered ghrelin. It was proposed, for instance, that the inhibitory effects of ghrelin on gastric acid secretion require intact vagal pathways because vagotomy abolished the increase in gastric secretion induced by this

peptide. Indeed, vagotomy significantly attenuated the ghrelin-afforded gastroprotection and the accompanying rise in the GBF after its central and peripheral administration indicating that vagal pathway plays an important role in the mediation of the protective and hyperemic effects of this peptide against lesions evoked by stress. These data could be interpreted that ghrelin induced protection against damage induced by stress might be due to the stimulation of vagal cholinergic pathways that are involved in the recruitment of CGRP from afferent sensory nerves. Thus, evidence was provided that these protective and hyperemic effects of ghrelin may affect vagal nerves, the principal component of brain-gastrointestinal tract axis and involve cooperation between endogenous NO and sensory nerves releasing neuropeptides such as CGRP.

The finding that centrally applied ghrelin raised NO in stress model, is consistent with the observations that protective effects of ghrelin against ethanol can be attenuated by the pretreatment with L-NAME suggesting an involvement of NOS-NO pathway in the gastroprotective effect of ghrelin applied ICV. This excessive gastric NO production by ghrelin may originate from the upregulation of constitutive NOS (cNOS) rather than inducible NOS (iNOS) in the gastric mucosa. It seems likely that NO contributes to gastroprotection afforded by central and peripheral

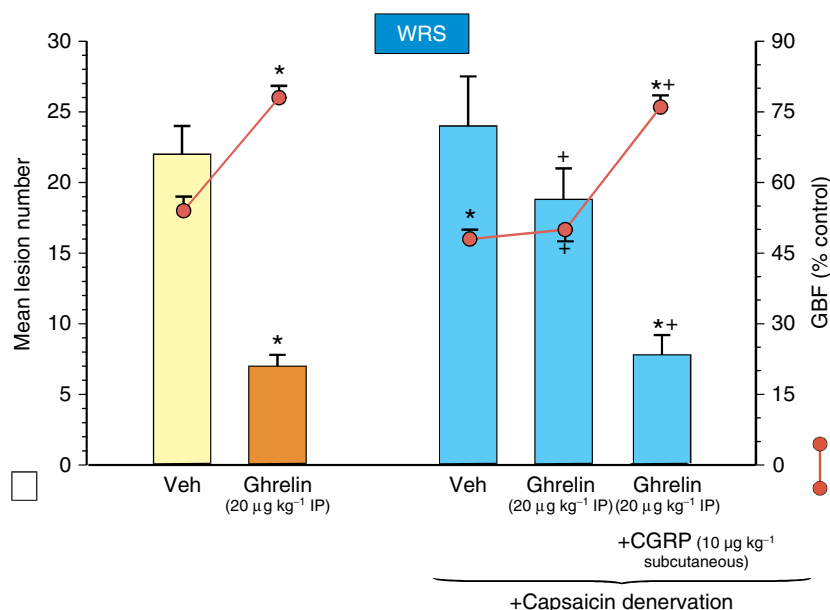


Figure 4 The mean number of WRS-induced gastric lesions and the alterations in the gastric blood flow (GBF) in rats with intact sensory afferent nerves and those with capsaicin denervation with or without pretreatment with vehicle (Veh; control) and ghrelin ($20 \mu\text{g kg}^{-1}$ intraperitoneal). Please note that ghrelin-induced protection against WRS lesions and the accompanying increase in the GBF are almost completely abolished by functional ablation of afferent sensory neurons by capsaicin. These effects are further restored by the concomitant treatment of ghrelin with exogenous CGRP ($10 \mu\text{g kg}^{-1}$ subcutaneous) in capsaicin-denervated animals. Mean $\pm$ standard error of mean of 6–8 rats. Asterisk indicates a significant change compared to the value in vehicle-treated rats. Cross indicates a significant change compared to the value obtained in rats without capsaicin denervation. Asterisk and cross indicate a significant decrease as compared to ghrelin-treated animals with capsaicin denervation of sensory afferents.

ghrelin as proposed earlier for CCK and leptin, both also implicated in control of appetite and gastroprotection.

Therefore, it is concluded that ghrelin, besides its recognized function in the control of appetite, energy homeostasis, fat, and carbohydrate metabolism and gastrointestinal motility, could be considered as an important gastroprotective factor, expressed locally in the gastric mucosa in response to mucosal injury. This notion is supported by recent observations that mRNA for ghrelin is upregulated in the gastric mucosa exposed to WRS. Furthermore, it has been demonstrated that beneficial effect of ghrelin involves an activation of the capsaicin-sensitive sensory nerves in the gastric mucosa that are responsible for the increased production of CGRP in gastric mucosa exposed to WRS. This notion is in keeping with the previous studies demonstrating the importance of the activation of capsaicin-sensitive sensory nerves in the protection of gastric mucosa against acute gastric injury. The contribution of CGRP to the ghrelin-induced gastroprotection is further supported by the observation that the functional ablation of sensory nerves by capsaicin attenuated significantly the protective activity of ghrelin and completely abolished the rise in GBF induced by this peptide. The addition of exogenous CGRP to ghrelin in rats with deactivated sensory nerves restored the protection and accompanying rise in the GBF with the extent similar to that observed in ghrelin-treated rats with intact sensory nerves. These results indicate that sensory nerves, and their neuropeptides such as CGRP, are essential for microcirculatory response and of crucial importance for the gastroprotective activity of ghrelin.

Previous studies revealed that CGRP released in response to acute gastric mucosal injury leads also to increased production of NO due to activation of constitutive endothelial NOS. The exposure of rats to WRS that produced gastric lesions was associated with a significant downregulation of cNOS mRNA expression as compared to vehicle-control rats. The treatment with ghrelin of such WRS exposed rats, dose-dependently increased the expression of cNOS which reached the level of this expression comparable to that observed in the vehicle-control gastric mucosa indicating that ghrelin is capable of preventing the decrease in expression of cNOS mRNA caused by the exposure of this mucosa to WRS. The implication of NO in gastroprotection induced by ghrelin was further determined by employing of an inhibitor of NOS, L-NAME injected IP without or with addition of L-arginine, the substrate of NOS. The gastroprotective and hyperemic effects induced by ghrelin were significantly attenuated by L-NAME, but reversed by

addition to L-NAME of L-arginine. All these observations support the notion that NO plays a crucial role in the mechanism of gastric protection against WRS damage afforded by ghrelin.

Ghrelin Attenuates Inflammation and the Generation of Reactive Oxygen Metabolites

The beneficial effects of ghrelin on the protection of gastric mucosa against injury induced by WRS could be related to the suppression of free oxygen radicals and anti-inflammatory properties of this peptide. Recent studies indicate that ghrelin could be considered as an important mediator in the regulation of inflammation due to its anti-inflammatory activity. Pretreatment with ghrelin just before exposure to WRS and I/R caused a significant reduction in the mRNA expression of proinflammatory cytokine, TNF- α , confirming that anti-inflammatory effect of ghrelin contributes to the gastroprotection and hyperemia exhibited by this peptide. This mechanism seems to be of importance since the upregulation of TNF- α expression represent a central pathophysiological event in the mechanism of acute mucosal injury induced by stress. Further support for the potent anti-inflammatory effects of ghrelin is the fact that ghrelin inhibited the activation of NF κ B, which is considered as a critical signaling molecule in inflammation. Previous studies revealed that transcription factor NF κ B activation is involved in acute injury of stomach and other organs. In its inactive form, NF κ B is sequestered in the cytoplasm and bound by members of I- κ B family of inhibitor proteins. Phosphorylation of I- κ B by an I- κ B kinase complex leads to the translocation of NF κ B-p65 into the nucleus. Activation of NF κ B induces gene programs leading to transcription of factors that promote inflammation. Therefore, it is assumed that a potential mechanism whereby ghrelin could modulate inflammatory response is blocking of activation of NF κ B. This notion was confirmed by observation that ghrelin inhibits an activation of NF κ B in human endothelial cells *in vitro*.

Ghrelin not only prevents the acute gastric injury induced by WRS, but also accelerates the healing of these lesions. At 6 h after the end of standard 3.5-h WRS, a significant acceleration of healing of the WRS-induced lesions was observed in rats treated with ghrelin as compared to those treated with vehicle. The mechanism by which ghrelin could stimulate the healing of WRS-induced lesions is poorly understood. Previous studies demonstrated that the growth factors stimulating angiogenesis play a crucial role in the mechanism of healing process. The acceleration of healing of acute WRS-induced gastric erosions was

accompanied by the increased expression of hypoxia-inducible factor-1 α (HIF-1 α). The increased expression of HIF-1 α in the ulcerated mucosa leads to increased expression and activation of vascular endothelial growth factor (VEGF). For instance, the concomitant rise in the expression of HIF-1 α and VEGF was observed at day 6 following the combination of ghrelin and I/R and this was accompanied by a marked decrease of the gastric mucosal lesions produced by I/R. At 24 h after the end of WRS and I/R, an increase in expression of HIF-1 α , while resulting in downregulation of the VEGF expression was observed in ghrelin-treated animals. This early reduction of VEGF expression despite increased expression of HIF-1 α at 24 h after I/R injury or following mucosal recovery from WRS lesions could be attributed to the increased generation of free oxygen radicals and increased lipid peroxidation which can inhibit the activation of VEGF. It is known that healing of WRS and I/R injury involves an activation of antioxidant enzymes resulting in the inhibition of free oxygen species generation and subsequent enhancement in VEGF mRNA expression. This could, at least in part, explain significant increase in the expression for

VEGF mRNA at 24 h following WRS damage and at day 6 after I/R-induced gastric injury.

In summary, the administration of exogenous ghrelin applied by the ICV or IP routes that is accompanied by a significant increment in its plasma levels, exhibits dose-dependent gastroprotection against the WRS-induced gastric lesions. Ghrelin is effective not only in gastroprotection against WRS damage but also accelerates the healing of these lesions. Existing evidence indicates that these protective and hyperemic actions of ghrelin may affect brain-gastrointestinal tract axis and involve cooperation between endogenous PG, NO, and sensory nerves releasing neuropeptides such as CGRP (Figure 5). Ghrelin exerts a potent anti-inflammatory activity and inhibits activation of NF κ B, an important transcription factor, which seems to be an essential step in the formation of WRS injury. As CGRP and cNOS/NO system are activated in gastric mucosa of ghrelin-treated animals, it is assumed that these two factors could trigger the gastroprotective mechanisms and contribute to the maintenance of gastric mucosal integrity in experimental animals and possibly also in humans exposed to noxious agents and to adverse conditions such as stress.

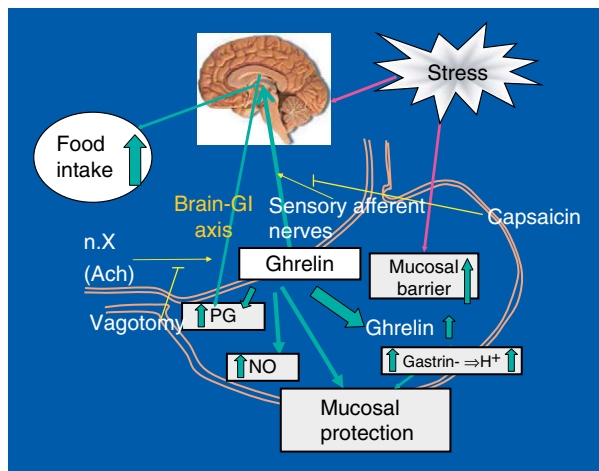


Figure 5 Scheme of the proposed mechanisms involved in gastroprotective action of ghrelin against the formation of mucosal damage induced by stress. Besides recognized involvement of ghrelin in the regulation of food intake and appetite behavior, this hormone exhibits gastroprotective and hyperemic activity. Ghrelin-induced protection may involve the activation of brain-gastrointestinal tract (GI) axis via vagal and sensory innervation and interactions of this hormone with potent gastroprotective factors including nitric oxide (NO), calcitonin gene releasing factor (CGRP) released from sensory afferent nerve endings and gastrin, known to exhibit gastroprotective and ulcer healing properties. The gastroprotective activity of ghrelin is impaired by suppression of sensory afferent and vagal efferent activities by capsaicin denervation and vagotomy, respectively, confirming the importance of neural aspects in the mechanism of gastroprotection induced by this hormone. n.X, vagal nerve; PG, prostaglandin.

Further Reading

- Ariyasu, H., Takaya, K., Tagami, T., et al. (2001). Stomach is a major source of circulating ghrelin, and feeding state determines plasma ghrelin-like immunoreactivity levels in humans. *Journal of Clinical Endocrinology and Metabolism* 86, 4753–4758.
- Brzozowski, T., Konturek, P. C., Konturek, S. J., et al. (2000). Central leptin and cholecystokinin in gastroprotection against ethanol-induced damage. *Digestion* 62, 126–142.
- Brzozowski, T., Konturek, P. C., Konturek, S. J., et al. (2000). Expression of cyclooxygenase (COX)-1 and COX-2 in adaptive cytoprotection induced by mild stress. *Journal of Physiology* 94, 83–91.
- Brzozowski, T., Konturek, P. C., Konturek, S. J., et al. (2001). Classic NSAID and selective cyclooxygenase (COX)-1 and COX-2 inhibitors in healing of chronic gastric ulcers. *Microscopy Research and Technique* 53, 343–353.
- Brzozowski, T., Konturek, P. C., Konturek, S. J., et al. (2004). Exogenous and endogenous ghrelin in gastroprotection against stress-induced gastric damage. *Regulatory Peptides* 120, 39–51.
- Brzozowski, T., Konturek, S. J., Śliwowski, Z., et al. (1996). Role of capsaicin-sensitive sensory nerves in gastroprotection against acid-independent and acid-dependent ulcerogens. *Digestion* 57, 424–432.
- Date, Y., Kojima, M., Hosoda, H., et al. (2000). Ghrelin, a novel growth hormone-releasing acylated peptide, is synthesized in a distinct endocrine cell type in the

- gastrointestinal tracts of rats and humans. *Endocrinology* **14**, 4255–4261.
- Kojima, M., Hosoda, H., Date, Y., et al. (1999). Ghrelin is a growth-hormone-releasing acylated peptide from stomach. *Nature* **402**, 656–660.
- Konturek, P. C., Brzozowski, T., Pajdo, R., et al. (2004). Ghrelin – a new gastroprotective factor in the gastric mucosa. *Journal of Physiology and Pharmacology* **55**, 325–336.
- Konturek, P. K., Brzozowski, T., Konturek, S. J., et al. (1990). Role of epidermal growth factor, prostaglandin and sulfhydryls in stress-induced gastric lesions. *Gastroenterology* **99**, 1607–1615.
- Konturek, S. J., Brzozowski, T., Majka, J., et al. (1992). Adaptation of the gastric mucosa to stress. Role of prostaglandin and epidermal growth factor. *Scandinavian Journal of Gastroenterology* **27**, 39–45.
- Masuda, Y., Tanaka, T., Inomata, N., et al. (2000). Ghrelin stimulates gastric acid secretion and motility in rats. *Biochemical and Biophysical Research Communications* **267**, 905–908.
- Robert, A. (1979). Cytoprotection by prostaglandins. *Gastroenterology* **77**, 761–767.
- Sibilia, V., Rindi, G., Pagani, F., et al. (2003). Ghrelin protects against ethanol-induced gastric ulcers in rats: studies on the mechanisms of action. *Endocrinology* **144**, 353–359.
- Whittle, B. J. R., Lopez-Belmonte, J. and Moncada, S. (1990). Regulation of gastric mucosal integrity by endogenous nitric oxide: interactions with prostanoids and sensory neuropeptides in the rat. *British Journal of Pharmacology* **99**, 607–611.

Glia or Neuroglia

G W Bennett and D E Ray

University of Nottingham Medical School,
Nottingham, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
G W Bennett, volume 2, pp 218–223,
© 2000, Elsevier Inc.

Introduction

Anatomy and Structure

Physiology and Pharmacology

Neuron–Glial Cell Interactions

Glossary

<i>GABA</i>	Gamma-aminobutyric acid, the major inhibitory neurotransmitter in the brain.
<i>CR3</i>	Complement receptor 3, a marker for immune cells.
<i>ED1</i>	Product of the ectodermal dysplasia 1 gene, a marker of immune activation.

Introduction

In the past, neuroscientists concentrated almost exclusively on the electroactive components of the nervous system (neurons) because of their direct relationship to function, and neglected the nonelectroactive silent cells (glia), which outnumber them by

at least 10 to 1 in much of the nervous system. Classically, glia were named for and relegated to the status of glue responsible for binding together the more important neurons, although Ramón y Cajal and others classified them and made useful proposals about their functions. This neglect has now been partly redressed, and their important roles in development, in maintaining function in the face of physiological stresses, and in the response to injury and infection are now recognized.

Subsequent studies confirmed that there are basically three types of glial cells, namely:

1. Macroglia:
2. Astrocytes: cells with many processes and possible function
3. Oligodendrocytes: cells with few processes that form the myelin of myelinated nerve fibers
4. Microglia:
5. Macrophages: cells in white matter specifically activated in injury and degeneration

In addition to these glial cells, the normal brain also contains pericytes and perivascular macrophages. Pericyte processes partially cover the basement membrane of brain blood vessels in a way similar to those of astrocytes. Pericytes, however, express smooth muscle actin and are contractile. Both pericytes and perivascular macrophages express immune markers and have sentinel phagocytic functions.

Astrocytes are the most abundant cells in the mammalian brain and consist of (1) those located in white

matter (nerve fibers), which were termed fibrous astrocytes due to the large numbers of fibrils present in the cells, and (2) those located in grey matter (nerve cell bodies), which have fewer fibrils and were termed protoplasmic astrocytes. Despite the survival of this early attempt at classification, it is evident that many other intermediary forms exist.

Functions of Glial Cells

Macroglial cells, astrocytes, and oligodendrocytes all possess complex roles during development and, in the adult, involve complex interactions with neurons. Suggested supportive roles include metabolic supply of energy and other substrates (though the extent of this remains to be substantiated) and the ability to mop up released neurotransmitters (e.g., Glu and GABA) and ions (e.g., K^+ ions). Astrocytes are often closely associated with blood vessels. They regulate endothelial cell function and the permeability of blood vessels as well as producing many factors present in extracellular space, including growth factors and steroids that have multiple regulatory roles. Oligodendrocytes specialize in myelination in the central nervous system (CNS), a role confined to the Schwann cell in peripheral nerves. Such functions are critical to the function of the normal, healthy brain, but microglial cells become activated in pathological states and share many functions of the macrophages of the immune system. Such cells become cytotoxic during inflammatory responses and produce and respond to specialized proteins called cytokines, which mediate the complex interactions between microglia, astrocytes, and neurons, as well as infiltrating immune cells resulting from disruption of the normal blood-brain barrier during these pathological processes.

Specific biochemical markers of glial cells have contributed enormously to our understanding of the roles of these cells. Immunohistochemical studies in the 1970s identified a glial fibrillary acidic protein (GFAP) as an associated marker of fibrous astrocytes, though many GFAP-negative cells predominate in gray matter and various other markers for these have emerged.

In Vitro Studies of Glial Cells

In the 1980s glial cell culture techniques were developed in which the cell-specific markers enabled cell identity to be confirmed. Such *in vitro* methods have been coupled with imaging techniques to record changes in intracellular Ca^{2+} and other ions and patch clamp methods to study individual cells and those in more intact brain slice preparations.

Applications of these techniques have demonstrated that glial cells have complex signaling processes similar

to neurons. Compared to neurons, glial cells generally have a more negative resting membrane potential due to a higher permeability to potassium. Glial cells therefore usually do not produce action potentials and remain passive (or silent) during intracellular recording, though more recent studies using patch clamping techniques have identified complex membrane properties. In addition, glial cells in culture, especially astrocytes, have been widely used to characterize membrane voltage-dependent channels (predominantly potassium channels and also low densities of sodium and calcium channels).

Many pharmacological and molecular studies indicate that glial cells possess a variety of transmitter, hormonal, and other receptors, though most glial cells appear to communicate with each other through specialized gap junctions rather than chemical synapses. This enables cytoplasmic exchange of many molecules, and such syncytial coupling may explain the lack of active membrane properties of glia cells *in situ*.

Anatomy and Structure

The staining techniques of Golgi (first described in 1885) provided the means for the earliest identification and morphology of neuroglial cells and neurons in the vertebrate nervous system.

Astrocytes

Astrocytes are the most abundant cells in the mammalian brain but are somewhat polymorphic. Those close to nerve fibers (white matter) are termed fibrous, since they are relatively hypertrophic with prominent intermediate filaments. The less prominent astrocytes associated with nerve cell bodies and synapses (gray matter) are termed protoplasmic, although intermediate forms exist, such as those associated with blood vessels.

All these share a common phenotype, with uniform expression marker proteins such as aldolase 3, but expression of others such as S-100 and GFAP may vary markedly. The most extreme instance of this phenotypic diversity within astrocytes is represented by the adult neuronal precursor cells in the subventricular zone, which have most of the characteristics of the astrocytes seen elsewhere yet retain the ability to dedifferentiate and form neurons by asymmetrical division.

Oligodendrocytes and Schwann Cells

Similarly, oligodendrocytes, which were found in early anatomical studies to produce the myelin sheath in the CNS, are analogous to the Schwann cells associated with peripheral neurons; they also exhibit

structural diversity and heterogeneity, including possible distinct perivascular as opposed to perineuronal oligodendrocytes. Although various subtypes have been identified, the growth and development of oligodendrocytes and their specific interaction with the neuron have been less well studied than those of the astrocyte.

Schwann cells function to protect peripheral neurons and nerve fibers from the surrounding tissues, are separated by a basal lamina, and are broadly of two types. The major group is ensheathing Schwann cells that are associated with unmyelinated nerve axons in parts of the autonomic and sympathetic nerves and dorsal root ganglia, where they offer strength and protection. Better known are the cells that form myelin sheaths around larger, fast-conducting nerve fibers (myelinating Schwann cells). The myelin sheath occurs in blocks with intervening nodes of Ranvier that possess high concentrations of potassium ion channels. This segmental arrangement, as well as providing important neuronal insulation, allows conduction of ionic currents within the enclosed fiber and facilitates saltatory (jumping) conduction along the fiber, resulting in a marked increase in conduction velocity of impulses. Recent research has strengthened the long-held view that Schwann cells are also metabolically active in facilitating regeneration processes and thereby enable peripheral nerve fiber recovery from injury, a capacity absent in central neural processes that lack Schwann cells. This prompted studies that have shown that in some situations it is feasible to experimentally transfer populations of Schwann cells into the CNS and thereby activate brain axonal regeneration.

A special type of CNS glial cell is the ependymogial (radial glia) cells, which line the ventricular spaces of brain. These are specially modified astrocytes of a characteristic bipolar structure with a ciliated apical pole protruding into the cerebrospinal fluid (CSF). In many vertebrate species, some of these cells include specialized sensory epithelia that develop mechanoreception (e.g., organ of Corti and vestibular organs). They also include tanycytes, which have a specialized role in the transportation of molecules between local blood vessels and brain circumventricular organ regions of the ventricular system; they are of particular importance in regions of the hypothalamus for the distribution of neuropeptide hormones. Other radial glial cells have important roles in development and demonstrate transitory existence.

Microglia

Microglial cells in the brain may outnumber nerve cells, and in the adult become particularly important

in pathological states. Those present in the normal (nonpathological) brain are termed resting or ramified (many branched appearance) microglia, while activated (or reactive) microglia cells may be nonphagocytic or become phagocytic microglia in fully activated pathology. Microglia derived from the systemic immune system invade during development and take up a quiescent resting morphology, their fine processes establishing nonoverlapping territories that cover the whole brain. A number of immune marker proteins such as CR3 can be used to identify resting microglia, and when activated they express a wide variety of markers characteristic of immune signaling and cytolytic functions such as ED1, interleukins, and nitric oxide synthase.

Physiology and Pharmacology

Ion Channels and Receptors

With their many processes surrounding synapses and cell bodies, astrocytes are well placed to dominate and regulate extracellular homeostasis, especially since they express a variety of ion channels, including potassium (K^+), sodium (Na^+), calcium (Ca^{2+}), and ligand-gated channels. The latter reflects the local pattern of neuronal neurotransmitter types, and all have largely been studied using primary astrocytes in culture and patch clamp techniques. Similarly, whole cell patch clamp methods have been applied to *in situ* preparations using hippocampal or other brain slices. These studies have identified many types of K^+ channels in astrocytes, indicating an important function of these cells in the regulation of extracellular levels of this ion, consistent with the notion that maintaining K^+ homeostasis is an important function of astrocytes.

Astrocytes, oligodendrocytes, and Schwann cells also possess voltage-gated Na^+ and Ca^{2+} ion channels, but at much lower densities than those in neurons. This causes some difficulties regarding understanding their precise function, since there are insufficient levels for action potentials to be generated. With respect to Schwann cells, some authors have suggested that the ion channels may be a local store of channels that can be passed to neighboring neural axons, since they share similar properties, but this remains to be directly demonstrated. Na^+ channels may also produce sufficient Na^+ exchange for the activation of astrocyte Na^+/K^+ -ATPase activity. Like with many other cells, modulation of glial intracellular Ca^{2+} may be an important signaling mechanism, since Ca^{2+} influx is activated by various stimuli, although the roles of calcium in astrocytes and other glia remain to be clarified.

Such ionic changes may be regulated by ligand-gated channel receptors, including GABA_A, that also

occur in the membranes of cultured astrocytes. Astrocytes also express high-affinity carrier systems for various neurotransmitter-like molecules, especially amino acids (e.g., Glu and GABA), and amino acid uptake into these cells may well contribute to the physiological regulation of these transmitter synapses as well as offer possible cytotoxic protection from such amino acids. Recent work suggests similar regulation of arginine, which is the precursor of nitric oxide, now considered a fundamental signaling molecule in neural tissues.

Astrocytes also contribute to extracellular ionic homeostasis by sequestering potentially neuronotoxic metals such as mercury and aluminum.

Metabotropic Receptors

Metabotropic (G-protein-coupled) receptors are also present in glial cell membranes, and there is good evidence for functional roles for glutamatergic (Glu), GABA_B, cholinergic (ACh), serotonergic (5-HT), adrenergic (NA), and various peptidergic (including angiotensin II, somatostatin, endothelium peptides, and substance P) receptors in cultured astrocytes, as shown by binding studies, the use of receptor-selective antibodies, and *in situ* hybridization methods. Similarly, biochemical analyses of second messenger systems, such as cyclic AMP, as well as cytosolic ionic changes, have demonstrated the pharmacological properties of such receptors. Similar receptors for hormones and growth factors are also present.

Gap Junctions

Gap junctions are predominant intercellular cytoplasmic communications between glial cells, though they also occur in other cells, including neurons. They consist of specialized membrane protein channels that enable the flow of inorganic and organic molecules between cells, usually consisting of six similar connexin molecules. Various types of connexin have been identified as products of a specific family of genes and have been associated with specific glial cells and neurons.

Intercellular communication via gap junctions provides a mechanism whereby the whole local population of astrocytes can make a coordinated contribution to maintaining ionic homeostasis in the face of physiological or pathological stress. The potential depolarizing effect of an accumulation of extracellular potassium during transiently high levels of neuronal activity can be minimized by its movement into and distribution between astrocytes within the active region until activity returns to normal. In addition, these astrocytic end feet that are in contact with capillaries show a high level of expression of

aquaporins, which may play a role in bulk movement of water within the brain.

Neuron–Glial Cell Interactions

Metabolism

The physical location of astrocytes between metabolic supply via the vasculature and metabolic demand in the neuron, combined with their unique possession of an energy reserve in the form of glycogen granules, is a clear indication of their potential role in energy metabolism. Furthermore, since astrocytes express glutamine synthetase, which is lacking in neurons, there is important metabolic cycling of glutamate and glutamine between neurons and glia. In addition to the considerable physiological importance of this cycle, astrocytic glutamine synthetase also plays an important role in defending the brain against hyperammonemia. Perivascular astrocytes are also the first casualties in hepatic encephalopathy.

Although there is no analogous compartmentation of the enzymes involved in glycolysis, it has been proposed that astrocytic conversion of glucose to lactate is followed by transfer of this lactate to neurons (the lactate shuttle hypothesis) and that this is the major pathway for glycolysis. There is recent evidence that the glycolytic response seen within 10 s of neuronal activation is specifically localized within astrocytes. Circulating lactate (elevated during fatiguing muscle activity) can also effectively substitute for glucose as an energy supply in the brain.

While astrocytes have been the primary focus of research, oligodendrocytes also respond metabolically to increased demand (as after traumatic myelin damage) in a complex manner with changes appropriate to the activation of sphingoglycolipids for myelination.

Neurotransmitter Inactivation

Astrocytes are closely associated with neuronal synapses in the CNS and are equipped with the means to degrade most neurotransmitters. Work in the late 1960s showed that various transmitters flowing out of the synaptic cleft are taken up and inactivated by glial elements. However, the question remains regarding the physiological importance of such processes during normal transmitter function, and it is possible that such processes function as a protective process during excess transmitter release and/or during pathological states. However, it is well established that astrocytes participate in the glutamate-glutamine cycle, in which metabolized transmitters can be taken up by the astrocyte and returned to neurons as a precursor.

Response to Injury: Microglia and Inflammatory Disease and Neuroglia and Stress

Response to local injury Although the brain does not participate in the normal systemic response to inflammation, the local presence of abnormal material (e.g., plasma proteins, cell debris, bacteria) or cytokines in the brain extracellular space activates both sentinel microglia and astrocytes. These recruit circulating neutrophils and monocytes, which arrive by vascular transcytosis, and all begin the process of cell killing and removal of debris. In addition to producing pro-inflammatory cytokines, activated microglia themselves become phagocytic. This local inflammation is normally self-limiting, and results in a glial scar that persists long after the initial damage has resolved. During inflammation, the normally tight blood–brain barrier becomes open to small molecules, under the influence of locally secreted cytokines and vascular endothelial growth factor. Normal vascular integrity is restored as the inflammatory process dies down. The later phase of the inflammatory process is associated with debris removal and replacement of nonneuronal elements, after which recruited macrophages leave the brain. However, the persistence of hypertrophic astrocytic processes within the scar severely limits the degree of axonal regrowth that is possible, although surviving axons passing through lesions can be remyelinated. In contrast, peripheral nervous system Schwann cells activate in a manner analogous to that seen during early development and facilitate axonal regeneration. In the future it may be possible to overcome the inhibitory influence of astrocytes and oligodendrocytes on axonal regrowth by blocking astrocyte-derived signaling molecules.

Local autoimmune reactions are seen in some disease states, most notably multiple sclerosis. In multiple sclerosis, apparently random, but most commonly periventricular, oligodendrocytes are targeted with primary demyelination and secondary neurodegeneration. The process of focal vascular infiltration and cytolysis progresses slowly, either continuously or, more commonly, at irregular intervals (the relapsing remitting form). The acute inflammatory process never extends beyond these focal areas, although magnetic resonance studies do show minor abnormalities in otherwise normal white matter. The precipitating factors that direct the immune system to what is to become a new focal area of damage are unknown, but may be minor vascular lesions that would be asymptomatic in a person without autoantibodies. A similar condition (relapsing experimental allergic encephalitis) can be produced in experimental animals autoimmunized with myelin basic protein.

Response to systemic injury and stress Glial cells express a variety of hormonal and growth factor receptors, which seem to be critical during development as well as during stress-induced injury or disease. Evidence indicates that thyroid hormones, especially T_3 , which are well known for their requirement in the maturation of neurons, also bind to stem cell astrocyte nuclei and regulate glial differentiation during development. A number of developmental neurotoxins are known to act by targeting glial proliferation, and maternal stress can have marked and irreversible effects on brain maturation.

More significantly, it is now well documented that astrocytes and oligodendrocytes synthesize many steroid hormones during development, while adult cultured cells have been shown to express receptors for several steroids, including estrogen, progesterone, glucocorticoids, mineralocorticoids, and androgens. Moreover, recent developments indicate that many such steroids elicit rapid nongenomic effects on both neuron and glial cells via membrane-bound receptors in addition to the established intracellular and nuclear genomic receptors known to activate RNA and protein synthesis.

The term neurosteroid describes the neuroactive steroids that are synthesized and released from glial cells and act as local neuromodulators in the brain. Such local steroids may be modulated in turn by the rise of hormones in the circulation during stress. For example, neurosteroids such as dehydroepiandrosterone (DHEA) alter in stress and during depression and may inhibit or reverse the stress effects of corticosteroids and other stress- or drug-induced behaviors. Furthermore, recent evidence suggests that cortisol-induced loss of brain hippocampal neurons may be attenuated by DHEA; an interesting possibility is that glial steroids may function as part of a neuroprotective mechanism.

The adult brain participates in systemic inflammation or trauma by producing sickness behavior (fever, loss of appetite, social withdrawal, and inactivity). This is thought to be mediated via the few small open barrier brain areas (notably the area postrema) that lack a blood–brain barrier and have resident macrophages that can respond directly to circulating cytokines and then signal to activate resident microglia elsewhere in the brain.

There is also some evidence that residual semi-activated states in resident microglia from prior brain damage or the accumulation of minor damage seen in normal aging may render otherwise normal brain areas sensitive to systemic cytokines. This has been suggested as a mechanism whereby systemic infections may precipitate a sudden deterioration in chronic neurodegenerative conditions such as Alzheimer's disease.

A number of studies have reported generalized increases in blood–brain barrier permeability in animals subjected to stress. Of those studies showing increased leakage into the brain parenchyma, several have not proved reproducible, and others did not allow sufficient time for markers to be cleared from the normal brain vasculature. The better methodologies have shown no evidence for blood–brain barrier compromise with stress, other than in extreme hypertension. However, in the case of dural blood vessels on the brain surface, there is some evidence that stress can lead to increased permeability. This stress effect, in rats and mice, is mediated by trigeminal activation of resident dural mast cells, and the effect may be relevant to the genesis of migraine headaches.

See Also the Following Articles

Blood–Brain Barrier, Stress and; Inflammation; Infection; Macrophages.

Further Reading

- Abbott, N. J. (2002). Astrocyte-endothelial interactions and blood-brain barrier permeability. *Journal of Anatomy* 200, 629–638.
- Kettermann, H. and Ransom, B. R. (1995). *Neuroglia*. Oxford: Oxford University Press.
- Kreutzberg, G. W. (1996). Microglia; a sensor for pathological events in the CNS. *Trends in Neurosciences* 19(8), 312–318.
- Martin, P. (1997). *The sickening mind*. London: Harper-Collins.
- Merrill, J. E. and Benveniste, E. N. (1996). Cytokines in inflammatory brain lesions: helpful and harmful. *Trends in Neurosciences* 19(8), 331–338.
- Perry, V. H. (2004). The influence of systemic inflammation on inflammation in the brain: implications for chronic neurodegenerative disease. *Brain, Behavior, and Immunity* 18, 407–413.
- Verkhratsky, A. and Kettermann, H. (1996). Calcium signalling in glial cells. *Trends in Neurosciences* 19(8), 346–352.

Glucocorticoid Effects on Memory: the Positive and the Negative

O T Wolf

University of Bielefeld, Bielefeld, Germany

© 2007 Elsevier Inc. All rights reserved.

Introduction

Effects of Stress on Different Memory Systems

Clinical Implications

Summary and Conclusion

*Procedural
memory*

Different forms of nonconscious (implicit) long-term memories that express themselves through changes in performance. These forms of memories are largely independent of the medial temporal lobe memory system.

*Working
memory*

A system for temporarily (short-term) storing and managing information. This system is dependent on regions of the prefrontal cortex.

Glossary

*Classical
conditioning
Declarative
memory*

A form of learning caused by the association (or pairing) of two stimuli.

The conscious (explicit) long-term memory for events or facts, mediated by the medial temporal lobe.

*Glucocorticoids
(GCs)*

Steroid hormones secreted from the adrenal cortex in response to stress. Cortisol and corticosterone are naturally occurring GCs in humans and animals.

Meta-analysis

A statistical procedure used to pool the results of multiple studies investigating a particular effect. This approach leads to a quantitative summary of a specific field of interest.

Introduction

Everyone has experienced instances of stress influencing his or her memory. There are situations in which we are unable to retrieve prior well-learned information, an example of how stress might interfere with our memory. In contrast, we remember very well specific embarrassing, shameful, or frightening events from the past; this is an example of how stress can enhance our memory. Finally, there are situations of chronic stress or stress associated psychiatric disorders that are characterized by hypo- or hyperamnesia. This article provides a brief and selective overview of what is known about the role of GCs in mediating

some of these effects. Advances in the field of psychoneuroendocrinology within the past few years have contributed substantially to a more differentiated and balanced view of the effects of stress hormones on memory in animals and humans.

Interaction with a stressor leads to a cascade of neuroendocrine stress responses designed to facilitate adaptation to the changing demands. The hypothalamic-pituitary-adrenal (HPA) axis, together with the sympathetic nervous system (SNS), is one of the most important systems in this respect. The hypothalamus activates the HPA axis in response to input from several other brain regions. Corticotropin releasing hormone (CRH) and vasopressin reach the pituitary through the portal blood system, which leads to the secretion of adrenocorticotrophic hormone (ACTH). In response to ACTH release from the pituitary, the adrenal secretes GCs. In most laboratory rodents (rats and mice) corticosterone dominates, whereas in humans cortisol is the main adrenal GC. GCs, as lipophilic steroid hormones, enter the brain relatively easily and exert multiple effects in regions throughout the brain. These effects are often mediated through the binding to the two receptors for the hormone: the mineralocorticoid receptor (MR) and the glucocorticoid receptor (GR). These two receptors differ in their affinity for cortisol (with the MR having a much higher affinity) and also in their localization in the brain. In addition, GCs can exert rapid nongenomic effects by influencing ion channels or neurotransmitter receptors at the membrane level. We emphasize that CRH, vasopressin, and ACTH can, on their own, influence cognition; the focus of this article, however, is on the effects of GCs on the memory of humans and animals.

To date, there is a substantial, almost overwhelming, literature on stress and memory. A PubMed search conducted with the two search terms stress and memory produced more than 3000 hits. Even a more focused search using glucocorticoids and memory as search terms still resulted in almost 400 hits. In order to provide a structured overview of this large and rapidly growing field, I use the concept of multiple memory systems in the brain to guide the reader through this jungle. Within each discussed memory domain, animal and human data are discussed. Acute stress effects are described, followed by chronic effects.

Effects of Stress on Different Memory Systems

There is a general agreement that there is not a single memory system in the brain but multiple interacting memory systems. Abundant literature exists on this topic, and it is not the goal of the current review to

provide an in-depth overview of different models and ideas; the important question of the comparability of different memory systems in different species is not addressed here either. I focus instead on declarative or explicit long-term memory because this domain has been, until now, investigated most intensely with respect to stress effects. In addition, I briefly outline the findings for procedural memory, working memory, and conditioning (or associative learning); these three memory systems have been sparsely investigated, at least in humans.

Declarative Memory

Acute effects Declarative memory is the explicit storage of facts and events, which can later be intentionally retrieved. It is a relational and flexible memory system. This type of memory is tested in the laboratory with word lists, paired associates, or short stories in humans. In animals, spatial maze tasks are used most often, the most famous one being the Morris water maze (MWM). Long-term memory can be further divided into different memory phases: acquisition (or initial learning), consolidation (or storage), and retrieval (or recall). For the successful completion of declarative tasks, an intact medial temporal lobe region (the hippocampus and surrounding cortical structures) is essential for the acquisition of the task. The role of the hippocampus in retrieval is under debate, and it might fulfill a time-limited role only. Especially in humans, it is the prefrontal cortex regions that are of great importance for (effortful) retrieval. The literature regarding the effects of stress on declarative memory is quite diverse and confusing, with groups reporting enhancing as well as impairing effects of GCs on this form of memory. However, it has become apparent that this is largely due to the fact that the different memory phases are modulated by GCs in an opposing fashion.

GCs enhance memory consolidation, and this aspect represents the adaptive and beneficial mode of GCs central nervous system action. [Figure 1a](#) depicts the chain of events underlying this phenomenon. This process has been conceptualized as the beneficial effects of stress within the learning context or intrinsic stress, emphasizing that a stressful episode is remembered better. This effect is mediated by the action of stress-released GCs on the hippocampal formation, and it is very well documented in animals. Studies by Roozendaal and McGaugh showed that adrenergic activation in the basolateral amygdala (BLA) appears to be a prerequisite for the modulating effects of GCs on other brain regions (e.g., the hippocampus). Lesions of the BLA, as well as β blockade, abolish the enhancing effects of posttraining GC administration.

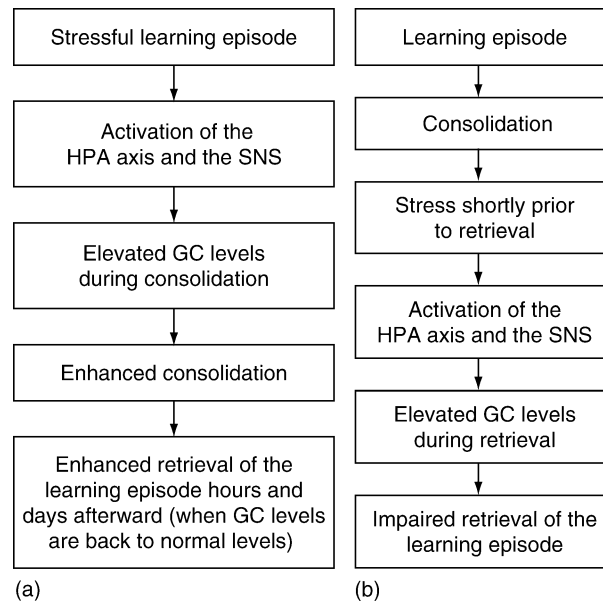


Figure 1 Schematic diagram of the cascade of events underlying the effects of GCs. a, Positive effects on long-term memory consolidation; b, negative effects on long-term memory retrieval. GC, glucocorticoid; HPA, hypothalamic-pituitary adrenal; SNS, sympathetic nervous system.

Although the enhanced memory consolidation is adaptive and beneficial, this process appears to occur at the cost of impaired retrieval (see [Figure 1b](#)). Using a 24-h delay interval, deQuervain and colleagues were the first to show that stress or GC treatment shortly before retrieval testing impairs memory retrieval in rats in the MWM. Further studies revealed that, again, an intact BLA and adrenergic activation appear to be necessary for the occurrence of this negative GC effect. Roozendaal summarized these findings as indicating that stress puts the brain into a consolidation mode, which is accompanied by impaired retrieval. Such a retrieval reduction might facilitate consolidation by reducing interference.

In humans, the literature on the acute effects of cortisol on declarative memory is somewhat divergent. Few studies reported beneficial effects, some reported negative effects, and several others found no acute effects. We recently conducted a meta-analysis in order to provide a quantitative summary of the current state of the field. The results supported the model derived from animal studies showing that cortisol impairs memory retrieval. Similar observations have been made in humans who were exposed to a psychosocial laboratory stressor shortly before retrieval testing took place.

To date, the beneficial effects on consolidation could not be clearly established using a meta-analysis, which in part might reflect a power problem. At least some studies, using emotionally arousing learning material and a sufficiently long retention delay in order to assure that GCs can influence consolidation

and in addition are back to baseline at the retrieval testing, observed enhanced memory consolidation for emotionally arousing material.

Interestingly, the beneficial effects on consolidation as well as the impairing effects on retrieval in humans are more pronounced for emotionally arousing material. This observation fits nicely with the observation in animals that GCs can exert effects on memory only in the presence of adrenergic activity in the amygdala. This arousal could be the result of specifics of the learning material and/or specifics of the testing conditions.

In the meta-analysis, time of day appeared as a second modulatory factor. Studies that administered cortisol before initial acquisition observed impairing effects on memory when they were conducted in the morning, a time of high endogenous cortisol levels in humans. In contrast, studies in the evening were more likely to observe beneficial effects. This supports an inverted U-shaped function between cortisol levels and memory in humans with both too low and too high levels at the time of acquisition being associated with impairments, especially when retrieval is tested at times when cortisol levels are still elevated.

In sum, studies in animals and humans converge on the idea that GCs acutely enhance memory consolidation of emotional arousing material while impairing memory retrieval. In addition, within this framework, emotional arousal and a nonlinear dose-response relationship are important modulatory variables.

Cortisol also influences declarative memory consolidation during sleep. Here the relationships appear

to be different than in the daytime. Low cortisol levels during the first half of the night are a prerequisite for a sleep-induced enhancement of declarative memory consolidation. Moreover, blocking the rise in cortisol levels, which typically occurs during the second part of the night, leads to enhanced emotional memory facilitation.

The interesting issue of individual differences in responsiveness to GCs has received little attention so far. Such differences might be able to account for some of the variance observed within as well as between studies. Individual differences could reflect genetic factors, differences in tissue GC sensitivity, or differences in lifetime cortisol exposure. Attempts to characterize subgroups with high versus low cortisol sensitivity could be one important venue for future research.

Chronic effects The picture is different for more chronic effects as a result of prolonged stress or GC treatment – here negative effects prevail. Chronic stress in animals has repeatedly been shown to result in impaired spatial-memory performance. The neurobiological underpinnings of those effects are still under debate. Stress-induced dendritic atrophy (or dendritic remodeling) in the hippocampus is one possible mechanism for the negative effects of chronic stress on spatial memory in rodents. However, several additional mediators such as reduced neurogenesis and dysregulation of several catecholamines have to be considered. Recent studies demonstrated substantial plasticity in the adult rodent brain. For example, stress-induced dendritic atrophy appears to be reversible.

Interestingly, the negative effect of chronic stress on spatial memory might occur only in male rats and not in females. In fact, similarly, acute stress appears to impair spatial memory only in males. The area of sex differences in response to stress is an important and still relatively neglected field. However, authors who went through the extra effort of studying male and female animals have often observed quite striking sex differences in the effects of stress on a wide range of target systems. These observations in combination with evidence for sex differences in emotional memory lateralization and sex differences in the effects of adrenergic manipulations on emotional memory emphasize that caution should be used when generalizing findings obtained in one sex to predict responses in the opposite sex.

In humans, the effects of chronic long-term GC treatment cannot, of course, be studied experimentally due to ethical considerations. Studies that administered GCs for several (3–10) days observed declarative memory impairments. Moreover, studies of patients receiving GC therapy and studies of hypercortisolemic

Morbus Cushing patients observed reduced declarative memory performance, which was associated with hippocampal atrophy. Some treatment studies in Cushing patients suggest that hippocampal atrophy in Cushing is reversible, which fits the observations in rodents previously mentioned.

Aging leads to alterations of HPA activity. Basal activity seems to increase and negative feedback is less efficient. Animal studies have suggested that enhanced HPA activity is associated with poorer memory. In older, otherwise healthy humans, several observational studies also reported associations between elevated or rising cortisol levels and declarative memory impairments. Whether these associations are specific for declarative memory or are, in fact, broader (also including working memory or attention) is currently under debate. Obviously all these human studies do not allow a clear cause and effect interpretation. Moreover, in most of these studies, the acute effects of cortisol cannot be differentiated from the chronic effects. In addition, the possible structural correlate of these hormone-performance associations remains to be firmly established. Hence, the possible association between rising cortisol levels and atrophy of the hippocampus is still not sufficiently understood, and the current empirical situation is somewhat mixed.

Finally, it has to be reiterated that other hormonal influences also might be important. With respect to aging, elevated cortisol levels could be part of the metabolic syndrome. Other aspects of this nonfavorable endocrine condition are impaired glucose tolerance and hypertension. These alterations have also been associated with memory impairment during aging, so future studies on this issue should consider taking a broader set of neuroendocrine measurements.

Procedural (Nondeclarative) Memory

This implicit and automatic form of (long-term) memory is not dependent on the medial temporal lobe. The brain areas involved depend on the specific content of the learning material. The neocortical areas, as well as the cerebellum and the basal ganglia, are involved in procedural memory. This is assessed in the laboratory with priming tasks (e.g., word-stem priming) or with skill-learning tasks (e.g., mirror tracing). In the few studies that used such tasks in the context of stress or GC research, a modulating effect of GCs was not observed, thus supporting the assumption that this memory form is not strongly influenced by GCs.

Working Memory

Working memory refers to the ability to store for the short term and manipulate material. This framework

has replaced the older notion of short-term memory and emphasizes the active processing capacity of this system. Neuroanatomically, this function has been linked to the prefrontal cortex (PFC). There are a large number of GRs in the PFC, but, in contrast to the effects of GCs on the hippocampus, the effects on the PFC are less well characterized. Acute impairing effects on working memory have been reported in animals and in some studies in humans, but the results have been somewhat inconsistent.

Regarding the chronic effects of GCs on working memory, impaired performance has been reported in animals and humans. In animals, chronic stress was associated with dendritic atrophy in prefrontal regions, which was reversible. In humans, elevated basal GC levels have been associated in some studies with smaller PFC volumes. The PFC is, of course, not only involved in working memory but also in executive functions. These domains have not been well investigated in humans with respect to GC effects and additional research efforts are still needed.

Classical Conditioning (Associative Learning)

Classical conditioning is one of the most basic forms of learning, and it can be demonstrated in snails, fruit flies, rodents, and humans. The brain regions involved depend on the specific stimuli used and the timing of the conditioned stimulus (CS)–unconditioned stimulus (UCS) association. For example, the cerebellum is involved in simple eyelid conditioning, whereas the amygdala is critical for fear conditioning and the hippocampus is important for contextual conditioning and trace conditioning. In rodents, the effects of stress on conditioning has been studied repeatedly. For eyelid conditioning, Shors and coworkers observed that stress enhanced conditioning in male rats. This was true for delay as well as trace conditioning. Interestingly, they observed the opposite effect in female animals – stress impaired conditioning. These findings are the opposite of the effects of acute stress on spatial memory, in which stress enhanced memory in female rats while impairing it in males. Although the enhancing effects of stress or GCs have been observed for fear conditioning, the issue of sex differences has not been addressed systematically. With respect to chronic stress, enhanced fear conditioning in the presence of hippocampal dentritic atrophy has been detected. In humans, classical conditioning has not been studied intensively in the context of neuroendocrine stress research; however, the available studies also observed sex differences. Because this form of learning and memory is of relevance for several psychiatric disorders, more research appears to be necessary.

Clinical Implications

The basic-science studies discussed thus far are also of substantial clinical relevance. For example, patients with posttraumatic stress disorder (PTSD) have been reported to show reduced basal cortisol levels, which is probably due to an enhanced negative feedback of the system. Recent small-scale clinical studies suggested that cortisol treatment shortly after the trauma might help to prevent PTSD. Moreover, even patients with chronic PTSD appear to benefit from a low-dose cortisol treatment. For such patients, the impaired emotional memory retrieval, in combination with an enhanced and more elaborate (place and time) reconsolidation, might help to reduce the core symptoms of the disorder, illustrating that basic science studies in the field of stress neuroendocrinology can help us to understand and possibly treat major psychiatric disorders.

Summary and Conclusion

Research over the past decade has substantially helped us gain a better understanding of the effects of GCs on memory. It is obvious that a more differentiated picture of stress effects on memory has evolved and that oversimplifications such as “stress impairs memory” are no longer supported by the existing literature. Increases in GCs induced by acute stress enhance memory consolidation but impair memory retrieval. These impairing effects of stress-induced GC secretion might hinder us in performing well in an important exam or at an important job interview. In addition, GCs also appear to reduce the amount of items that can be stored in working memory, but this is less well documented than the effects on declarative long-term memory. Interestingly, there is now convincing evidence that adrenergic activation in the amygdala is a prerequisite for several acute effects of GCs on memory. This activation can be induced either through the cognitive task itself, through the test situation (e.g., novel unfamiliar environment), or through the test material (e.g., emotional slides).

Substantial sex differences have been observed in several animal studies, but evidence in humans is still sparse. The direction of the effect appears to be task-specific. More knowledge about sex differences should, in the long run, help us understand sex differences in stress-associated disorders. Future studies need to investigate whether these sex differences are related to differences in the neuroendocrine stress response or differences in the response of the brain to endocrine stress messengers.

Chronic stress or chronic GC treatment has mostly negative effects on the brain and on the body. These observations are relevant to psychiatric disorders as well as to the aging process. However, it is encouraging that recent research has repeatedly observed evidence for preserved plasticity and structural remodeling once the stress has ceased or GCs return to normal levels.

Especially rewarding for all basic-science research that has been done on stress is that already several clinical applications have been implemented or are on the horizon.

Acknowledgments

The author's work was supported by grants from the German Research Foundation (DFG WO 733/6-1 and WO 733/7-1).

Further Reading

- Aerni, A., Traber, R., Hock, C., et al. (2004). Low-dose cortisol for symptoms of posttraumatic stress disorder. *American Journal of Psychiatry* **161**, 1488–1490.
- Buchanan, T. W. and Lovallo, W. R. (2001). Enhanced memory for emotional material following stress-level cortisol treatment in humans. *Psychoneuroendocrinology* **26**, 307–317.
- Cahill, L. (2003). Sex-related influences on the neurobiology of emotionally influenced memory. *Annals of the New York Academy of Sciences* **985**, 163–173.
- Conrad, C. D., Jackson, J. L., Wiczorek, L., et al. (2004). Acute stress impairs spatial memory in male but not female rats: influence of estrous cycle. *Pharmacology Biochemistry & Behavior* **78**, 569–579.
- Conrad, C. D., LeDoux, J. E., Magarinos, A. M., et al. (1999). Repeated restraint stress facilitates fear conditioning independently of causing hippocampal CA3 dendritic atrophy. *Behavioral Neuroscience* **113**, 902–913.
- Convit, A. (2005). Links between cognitive impairment in insulin resistance: an explanatory model. *Neurobiology of Aging* **26**(supplement1), 31–35.
- Croiset, G., Nijssen, M. J. and Kamphuis, P. J. (2000). Role of corticotropin-releasing factor, vasopressin and the autonomic nervous system in learning and memory. *European Journal of Pharmacology* **405**, 225–234.
- De Kloet, E. R., Joels, M. and Holsboer, F. (2005). Stress and the brain: from adaptation to disease. *Nature Reviews Neuroscience* **6**, 463–475.
- De Kloet, E. R., Oitzl, M. S. and Joels, M. (1999). Stress and cognition: are corticosteroids good or bad guys? *Trends in Neurosciences* **22**, 422–426.
- de Quervain, D. J., Roozendaal, B. and McGaugh, J. L. (1998). Stress and glucocorticoids impair retrieval of long-term spatial memory. *Nature* **394**, 787–790.
- Het, S., Ramlow, G. and Wolf, O. T. (2005). A meta-analytic review of the effects of acute cortisol administration on human memory. *Psychoneuroendocrinology* **30**, 771–784.
- Kuhlmann, S., Piel, M. and Wolf, O. T. (2005). Impaired memory retrieval after psychosocial stress in healthy young men. *Journal of Neuroscience* **25**, 2977–2982.
- LaBar, K. S. and Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience* **7**, 54–64.
- Luine, V. (2002). Sex differences in chronic stress effects on memory in rats. *Stress* **5**, 205–216.
- Lupien, S. J. and Lepage, M. (2001). Stress, memory, and the hippocampus: can't live with it, can't live without it. *Behavioural Brain Research* **127**, 137–158.
- Lupien, S. J., Schwartz, G., Ng, Y. K., et al. (2005). The Douglas Hospital Longitudinal Study of Normal and Pathological Aging: summary of findings. *Journal of Psychiatry & Neuroscience* **30**, 328–334.
- MacLulich, A. M., Deary, I. J., Starr, J. M., et al. (2005). Plasma cortisol levels, brain volumes and cognition in healthy elderly men. *Psychoneuroendocrinology* **30**, 505–515.
- McEwen, B. S. (2002). Sex, stress and the hippocampus: allostasis, allostatic load and the aging process. *Neurobiology of Aging* **23**, 921–939.
- Roozendaal, B. (2000). Glucocorticoids and the regulation of memory consolidation. *Psychoneuroendocrinology* **25**, 213–238.
- Roozendaal, B., Okuda, S., de Quervain, D. J., et al. (2006). Glucocorticoids interact with emotion-induced noradrenergic activation in influencing different memory functions. *Neuroscience* **138**(3), 901–910.
- Schelling, G., Roozendaal, B. and de Quervain, D. J. (2004). Can posttraumatic stress disorder be prevented with glucocorticoids? *Annals of the New York Academy of Sciences* **1032**, 158–166.
- Shors, T. J. (2004). Learning during stressful times. *Learning & Memory* **11**, 137–144.
- Squire, L. R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological Review* **99**, 195–231.
- Tulving, E. (2001). Episodic memory and common sense: how far apart? *Philosophical Transactions of the Royal Society of London B Biological Sciences* **356**, 1505–1515.
- Wagner, U., Degirmenci, M., Drosopoulos, S., et al. (2005). Effects of cortisol suppression on sleep-associated consolidation of neutral and emotional memory. *Biological Psychiatry* **58**, 885–893.
- Wolf, O. T. (2003). HPA axis and memory. *Best Practice & Research: Clinical Endocrinology & Metabolism* **17**, 287–299.
- Wolkowitz, O. M., Reus, V. I., Canick, J., et al. (1997). Glucocorticoid medication, memory and steroid psychosis in medical illness. *Annals of the New York Academy of Sciences* **823**, 81–96.
- Yehuda, R. (2002). Post-traumatic stress disorder. *New England Journal of Medicine* **346**, 108–114.

Glucocorticoid Negative Feedback

M F Dallman

University of California, San Francisco, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 224–228, © 2000, Elsevier Inc.

Characteristics of Feedback in the Hypothalamic-Pituitary-Adrenocortical Axis

Site(s) of Negative Feedback in the Hypothalamic-Pituitary-Adrenocortical Axis

Glossary

<i>Adrenocorticotrophic hormone (ACTH)</i>	A hormone made in corticotroph cells of the anterior pituitary; it stimulates the synthesis and secretion of corticosteroids from the adrenal cortex.
<i>Arginine vasopressin (AVP)</i>	A hormone that, when cosecreted with CRH, potentiates ACTH release.
<i>Corticotropin releasing hormone (CRH)</i>	A hormone synthesized in the hypothalamus; it stimulates ACTH synthesis and secretion from anterior pituitary corticotrophic cells.
<i>Dexamethasone</i>	A potent synthetic glucocorticoid that does not bind to transcortin (corticosteroid-binding globulin, CBG).
<i>Glucocorticoid receptor (GR)</i>	Low-affinity intracellular corticosteroid receptor that prefers dexamethasone, cortisol, and corticosterone.
<i>Limbic system</i>	A system of neural structures (amygdala, bed nuclei of the stria terminalis, hippocampus, septal nuclei, and hypothalamic and midbrain cell groups) that subserve the emotional and other associative memory, determining overall responsiveness to stressors.
<i>Mineralocorticoid receptor (MR)</i>	High-affinity intracellular corticosteroid receptor that selectively binds aldosterone, corticosterone, and cortisol.
<i>Paraventricular nucleus (PVN)</i>	Hypothalamic site of neuronal synthesis of CRH and AVP, which are secreted from axonal endings in the median eminence to act on pituitary corticotrophs.
<i>Stressors</i>	Stimuli that challenge the physical or psychological stability of the organism and cause the activation of the hypothalamic-pituitary-adrenal axis.
<i>Transcortin (CBG)</i>	A high-affinity binder for cortisol and corticosterone made by the liver; it reversibly binds corticosteroids in the circulation; also called corticosteroid-binding globulin.

Negative feedback inhibition by adrenal glucocorticoids is the means by which the brain is informed that the pituitary ACTH-secreting cells and adrenocortical fasciculata cells have done its bidding. However, a great deal of flexibility is required for an effective feedback system to serve its appropriate function. Thus, there are various time domains and mechanisms by which the glucocorticoids act to inhibit further brain-induced stimulation of the hypothalamic-pituitary-adrenocortical (HPA) axis, the different sites at which the steroids act, and the state-dependent modifications of the brain response to the feedback action of the steroids. The net result of this seeming complexity is a spectacularly functional feedback system that maintains glucocorticoid levels at values appropriate for the physiology of the organism at most times.

Characteristics of Feedback in the Hypothalamic-Pituitary-Adrenocortical Axis

The wide variation in total circulating corticosteroid levels, from the low nanomolar range during the daily circadian trough to the micromolar range during stress, provides a problem for feedback regulation by corticosteroids; the solution is to use two receptors of differing affinities to regulate the brain activation of the adrenocortical system. In addition, there are three temporal domains for the inhibitory effects of corticosteroids and the brain also manages to modify the apparent efficacy of the negative feedback effects of corticosteroids under persistently stressful conditions, thus maintaining an HPA axis that is responsive under conditions in which circulating corticosteroid levels are higher than the usual set point.

Two Receptors Mediate the Feedback Effects of Glucocorticoids

Figure 1 indicates the problem that must be solved by a corticosteroid feedback system in order to control accurately the activity of the HPA axis under circadian and stressful conditions. During the trough of the basal rhythm, total cortisol levels range between 50 and 100 nM. Because at low values ~99% of this represents cortisol bound to CBG, the free cortisol levels that can diffuse from the vascular compartment into the interstitial fluid to bathe cells range between ~0.5 and 1.0 nM. By contrast, at the peak of the circadian rhythm and after stress, total cortisol ranges up to micromolar concentrations and free

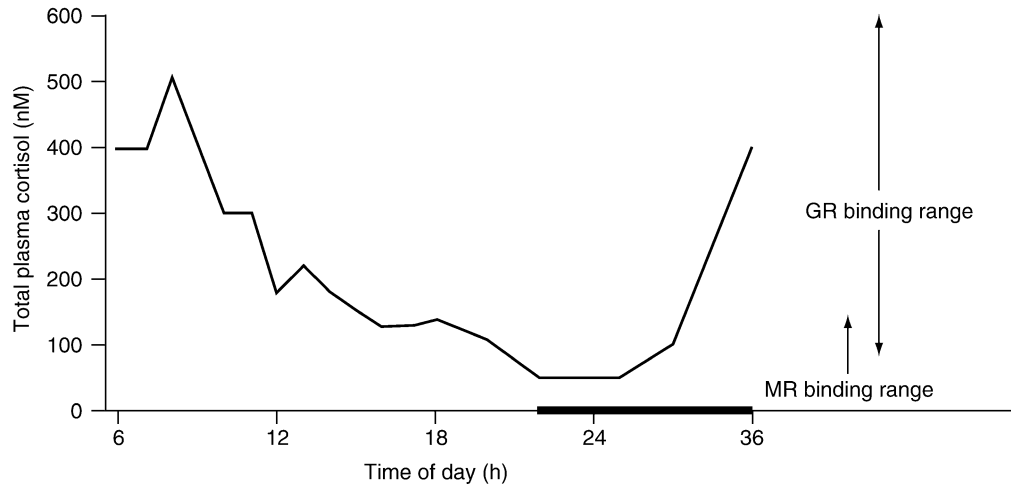


Figure 1 The two-receptor system that provides feedback information to the central HPA axis about adrenocortical performance. A hypothetical normal circadian rhythm in total plasma cortisol (CBG-bound and free) in humans. The dark bar on the abscissa represents the hours of darkness and sleep. The peak occurs just prior to awakening and the beginning of the daily activity cycle. The trough (MR occupancy) occurs early in the hours of sleep and inactivity before values rise again. Stressor-induced increases in cortisol can reach micromolar concentrations. Clearly, if GRs are occupied, MRs are as well. The estimated occupancy of MRs and GRs by free steroid at the various levels of total steroid is shown on the right.

cortisol levels approach 100–300 nM or even higher. Because receptors characteristically bind hormone over a 100-fold range, it is clear that a single receptor for cortisol cannot control the range of cortisol concentrations required during the normal physiological excursions of the HPA axis. The solution is the use of two receptors of differing affinities to inform the brain of adrenocortical activity.

1. The MR is a high-affinity receptor (K_d 0.5–1 nM) that has roughly equally high affinity for aldosterone, cortisol, and corticosterone. In cells that are selectively responsive to aldosterone, such as the principal cells of the renal nephron, there is an enzyme (11- β -hydroxysteroid dehydrogenase type II) that oxidizes the 11-hydroxyl groups of cortisol and corticosterone to the inactive 11-keto form. Aldosterone exists as an 11-18-hemiketal in solution and is not oxidized by this enzyme; therefore, it can exert its actions selectively in these target cells. Cells that do not contain the enzyme but do have MRs bind all three steroids; however, cortisol and corticosterone circulate in higher concentrations and gain entry to the brain much more easily than aldosterone; therefore, these steroids primarily occupy MRs in the brain. In the brain, MRs have a highly discrete distribution, primarily in neurons in entorhinal cortex, limbic structures, and motor output neurons. There is a very low density of MRs in the neurons of the hypothalamus.

2. The GR is a lower-affinity (K_d 10–30 nM) corticosteroid receptor that prefers dexamethasone $\gg$ cortisol, corticosterone $>$ aldosterone. The GRs are

distributed in every cell type in the organism. The distribution of GR in the brain is very widespread in both glial cells and neurons. CRH neurons in hypothalamus are well endowed with GRs, as are the neurons in the limbic system, the monoaminergic systems, and others.

Both occupied MRs and GRs dimerize in the cell nucleus to alter transcriptional rates of tissue-specific genes. In addition, MRs and GRs in the same cell can heterodimerize and act transcriptionally. Moreover, it is increasingly clear that the hormone receptor complexes can function in protein–protein interactions with other transcription factors to alter transcription rates. At this time, the specific effect of occupancy of MRs and GRs on transcription is impossible to predict. Nonetheless, occupancy of MRs and GRs by cortisol or corticosterone serves to inhibit activity in the HPA axis, generally by an action on brain.

Feedback regulation in the animal and its utility Using adrenalectomy and infusion of MR and GR agonists as well as treatment with specific antagonists in intact animals, it is clear that occupancy of the MRs is required to maintain ACTH secretion at low levels during the trough of the circadian rhythm. By contrast, at the circadian peak and during stress, the occupancy of both MR and GR restrains ACTH secretion, maintaining it within the normal limits observed in intact organisms. Because occupied GRs and MRs are powerful modulators of cell function, it is important to maintain their overall concentrations at

tightly regulated levels. Many corticosteroid-sensitive end points change their levels and activity when there is slight deviance on either side of the normal daily mean value. In studies of the daily mean of plasma cortisol in humans, values range from 6 to 8 $\mu\text{g dl}^{-1}$ when levels collected over 24 h are averaged. Similarly, in rats, the 24-h mean corticosterone concentration ranges between 4.5 and 7 $\mu\text{g dl}^{-1}$. Both species have relatively high circulating CBG levels. Above or below these ranges in unstressed humans and rats, untoward metabolic, behavioral, and immunological changes occur.

Temporally Discrete Effects of Corticosteroids on Hypothalamic-Pituitary-Adrenal Axis Function

Feedback inhibition of (CRH, AVP, and) ACTH secretion can be shown to exist in fast, intermediate, and slow time domains when corticosteroids are administered exogenously. Normally these segue smoothly from one to the other, but they can be distinguished mechanistically, and the degree of inhibition increases as the duration and amount of the corticosteroid exposure increases. In all time domains, the magnitude of the effect is strictly proportional to the dose of steroid, between the limits of threshold and saturating quantities.

1. The fast effects occur within milliseconds on neurons and must be exerted by an effect at the cell membrane. A membrane receptor has not yet been characterized. The iontophoresis of corticosteroids onto neurons or bathing tissue slices with corticosteroids shows effects within milliseconds to seconds on membrane function. Similarly, in animals the stimulation of pathways to CRH neurons does not evoke the normal response within minutes of a systemic injection of corticosteroids. At the pituitary corticotroph, CRH-stimulated ACTH secretion is inhibited within a few minutes of the application of corticosteroids; this inhibition is not altered by the inhibition of RNA or protein synthesis, showing that it is not mediated by the usual genomic effects of corticosteroids, although it may be mediated by interference with the generation of adenylyl cyclase, the intracellular messenger for CRH. The magnitude of the fast inhibition of ACTH secretion by corticosteroids acting on the corticotroph is small ($\sim 25\%$ of the maximal inhibition possible); thus, fast feedback leaves the organism responsive to further stressors and capable of secreting nearly normal amounts of ACTH.

2. Intermediate feedback has its onset at approximately 30 min after a pulse or continuous exposure to the steroid and lasts for a period of hours. The inhibition of protein synthesis with cyclohexamide blocks this feedback and, although it is powerful and can

completely inhibit CRH-induced ACTH secretion in the whole animal, ACTH secretion can usually be stimulated by further treatment with CRH or AVP. This temporal domain of corticosteroid feedback appears to dampen the excitability in the HPA axis without abolishing it (i.e., excitability is returned to prestress levels so that subsequent responsivity is not excessive). At the pituitary, AVP-stimulated ACTH secretion is less sensitive than CRH to corticosteroid inhibition.

3. Slow feedback is found in conditions in which the supraphysiological levels of exogenous corticosteroids have been provided for days or weeks. All components of the HPA axis are unresponsive to stimulation, both by intense stressors and, at the pituitary, by CRH and AVP. The CRH and AVP mRNA expression in the hypothalamus and proopiomelanocortin expression in the pituitary is markedly decreased. This condition is most frequently observed in humans with active Cushing disease or after prolonged treatment of disease with high levels of synthetic glucocorticoids; it takes months to recover from this inhibition when the high levels of steroid are removed.

Conditional Feedback in the Hypothalamic-Pituitary-Adrenal Axis

Up to this point, glucocorticoid feedback properties have been discussed with respect to the normal unstressed organism. Under conditions of persistent or chronic stress, the axis appears to change, and sensitivity to the feedback action of corticosteroids is slightly to markedly reduced. Thus, under chronic stress, mean daily plasma corticosteroid concentrations increase, frequently only because circadian trough levels are elevated but occasionally because there are major stressor-induced elevations throughout the circadian cycle. With chronic stress, the degree of the inhibition of stimulated ACTH secretion by exogenous (or endogenously secreted) glucocorticoids is decreased. In rats, chronic stress often stimulates increased expression of AVP in CRH-synthesizing neurons, thus providing a means for amplifying the corticotroph response to the CRH and AVP secreted. However, it also appears that neural sites upstream from the final hypothalamic CRH and AVP motor neurons of the HPA axis are less sensitive to corticosteroid feedback.

A decreased number of GRs has been reported, primarily in hippocampus, after many chronic stressors and after treatment with exogenous glucocorticoids, suggesting that one mechanism of reduced feedback sensitivity may reside in GR downregulation. This is not the only explanation for reduced feedback sensitivity during chronic stress, however, because there are several examples in which both MR and GR

numbers are normal throughout brain but glucocorticoid feedback insensitivity is found.

Because the strength of inputs to the CRH and AVP cells stimulates dose-related output, it is likely that, although corticosteroid dose-related functions are occurring normally under chronic stress conditions, the level of input to the CRH and AVP cells increases. Thus, normal responsivity in the HPA axis is maintained, whereas the apparent feedback efficacy of glucocorticoids is reduced. It does appear that new stimuli presented to chronically stressed animals causes facilitated responses in the central neural pathways to the CRH and AVP neurons.

The result of the apparent decrease in glucocorticoid feedback under conditions of chronic stimulation of the HPA axis is that there is dampening of the response to the ongoing stimulus but full or even exaggerated responsiveness to a new and different stimulus. Both effects are useful – the degree of hyperglucocorticoidism is limited by ongoing glucocorticoid feedback, but the organism is still capable of responding to a new stimulus that may demand a momentary shot of the hormones for optimal response.

Site(s) of Negative Feedback in the Hypothalamic-Pituitary-Adrenocortical Axis

Feedback is primarily exerted in the brain under normal conditions, with only moderate increases in mean corticosteroid levels or low-magnitude acute stressors. Although the CRH and AVP neurons in the PVN are plentifully supplied with GRs, this site does not appear to be inhibited directly by the naturally occurring corticosteroids. Instead, the corticosteroids appear to inhibit the activity of neuronal pathways that eventually impinge on CRH and AVP neurons. At high sustained corticosteroid levels, there is also clearly negative feedback on anterior pituitary corticotrophs.

Feedback in the Brain

Microinfusions of low amounts of MR and GR antagonists into the cerebral ventricles or into specific brain sites augment the magnitude of ACTH responses to a variety of stressors, strongly suggesting that stress-induced glucocorticoid secretion decreases the stimulatory input to the PVN during both the fast- and intermediate-feedback time domains. Small crystalline implants of corticosterone placed chronically into prefrontal cortex, the preoptic area, and the dorsal hippocampus have been shown to reduce stress or adrenalectomy-induced ACTH secretion. However, it appears that steroids implanted into brain must bathe those neurons that are activated by

the specific stressor used to excite the HPA axis. That is, there appear to be distributed sites at which steroids may inhibit HPA responses to stressors, not a single feedback site that globally inhibits PVN neuronal responses to stressors. Corticosteroids, acting through both MRs and GRs, have been shown to alter neuronal properties and activity in the hippocampus, and antagonists infused into both the amygdala and hippocampus have been shown to alter conditioned behaviors. However, it has not yet been directly demonstrated that implants in these sites alter the amount of ACTH secreted in response to a stressor. It may well be that corticosteroid receptors in these sites have little to do with modulating CRH and AVP secretion directly. In the brain, it appears that for feedback effects of steroid implants to be observed on ACTH secretion the steroid must act on sites that are also activated by stress; the feedback action is selective and specific.

Feedback at the Pituitary

Corticotrophs contain GRs but not MRs, and the direct inhibitory effects of corticosteroids are evident at the anterior pituitary. The concentrations of the naturally occurring steroids required to effect the inhibition of ACTH secretion are relatively high. This is probably because corticotrophs contain the high-affinity CBG corticosteroid binder apparently taken up from the circulation. The CBG in the corticotrophs sequesters the natural corticosteroids and high concentrations (10^{-8} M) are required to observe inhibition, although dexamethasone, which does not bind to CBG, is a more effective inhibitory agent. Moreover, corticotrophs tend to maintain sensitivity to AVP, although sensitivity to CRH decreases when corticosteroid levels are elevated. The relative insensitivity of the corticotroph to corticosteroid inhibition is useful because it allows the corticotroph to be the site of last resort, requiring the highest levels of signal, for the global inhibitory control of the HPA axis.

See Also the Following Articles

Adrenocorticotrophic Hormone (ACTH); Corticotropin Releasing Factor (CRF); Feedback Systems; Glucocorticoids, Effects of Stress on; Hypothalamic-Pituitary-Adrenal.

Further Reading

- Antoni, F. A. (1993). Vasopressinergic control of pituitary adrenocorticotropin comes of age. *Frontiers in Neuroendocrinology* 14, 76–122.
- Bamberger, C. M., Schulte, H. M. and Chrousos, G. P. Molecular determinants of glucocorticoid receptor function and tissue sensitivity to glucocorticoids. *Endocrine Review* 17, 245–261.

- Dallman, M. F., Akana, S. F., Cascio, C. S., Darlington, D. N., Jacobson, L. and Levin, N. (1987). Regulation of ACTH secretion: variations on a theme of B. In: Clark, J. H. (ed.) *Recent progress in hormone research*, pp. 113–173. Orlando, FL: Academic Press.
- Dallman, M. F., Akana, S. F., Scribner, K. A., et al. (1992). Stress, feedback and facilitation in the hypothalamo-pituitary-adrenal axis. *Journal of Neuroendocrinology* 4, 517–526.
- Dallman, M. F. and Bhatnagar, S. (1999). *Chronic stress and energy balance: role of the hypothalamo-pituitary-adrenal axis*. Washington, DC: American Physiological Society.
- De Kloet, E. R. (1991). Brain corticosteroid receptor balance and homeostatic control. *Frontiers in Neuroendocrinology* 12, 95–164.
- Joels, M. (1997). Steroid hormones and excitability in the mammalian brain. *Frontiers in Neuroendocrinology* 18, 2–48.
- Keller-Wood, M. E. and Dallman, M. F. (1984). Corticosteroid inhibition of ACTH secretion. *Endocrine Review* 5, 1–24.

Glucocorticoid Receptor Mutant Mice as Models for Stress-Induced Affective Disorders

P Gass

Central Institute of Mental Health (CIMA), University of Heidelberg, Heidelberg, Germany

© 2007 Elsevier Inc. All rights reserved.

Corticosteroids, Corticosteroid Receptors, and Depression
Animal Models of Depression
Mice with Targeted Mutations of GR
GR Mutant Mice Represent a Tool to Study Molecular
Correlates of Depression-like Behavior

Glossary

<i>Corticosteroids</i>	Hormones synthesized in the adrenal glands, secreted in a circadian rhythm and in response to stress. The main endogenous glucocorticoids are corticosterone in rodents and cortisol in humans.
<i>Corticosteroid receptors</i>	Intracellular receptor molecules for corticosteroid hormones. On hormone binding they function as transcription factors in the nucleus of the cell to regulate the expression of target genes. There are two types: mineralocorticoid receptors and glucocorticoid receptors.
<i>Depression</i>	Psychiatric disorder characterized by core symptoms such as depressed mood, anhedonia, and loss of interests. The syndrome also includes changes in cognition, psychomotor alterations, and vegetative symptoms (changes in sleeping, eating, and sexual behaviors).
<i>Hypothalamic-pituitary-adrenocortical (HPA) axis</i>	The brain-controlled activation pathway for the secretion of corticosteroid hormones after stress through a hormonal signaling cascade, involving release of

Learned helplessness

Stress

Transgenic mice

corticotropin-releasing hormone from the paraventricular nucleus of the hypothalamus in the brain and adrenocorticotrophic hormone from the pituitary gland.

A psychobiological concept that predicts that stressors are particularly apt to induce depressive states when they are unpredictable and uncontrollable.

A physiological and psychological state caused by a real or perceived challenge to homeostasis, which is accompanied by a hormonal stress response.

Animals with a gene knockout (loss of function) or a transgene expressing an additional gene product (gain of function), designed by targeted mutagenesis. With modern techniques, genes can be overexpressed or shut down in a brain regional or temporal-specific way.

Corticosteroids, Corticosteroid Receptors, and Depression

Dysregulations and dysfunctions of corticosteroid receptors have been implicated in the pathogenesis of stress-related psychiatric disorders such as depression and posttraumatic stress disorder. It is still unclear, however, whether these corticosteroid receptor disturbances are a cause or a consequence of affective disorders. Clinical studies have demonstrated a hyperactivity of the hypothalamic-pituitary-adrenocortical (HPA) system with elevated plasma cortisol levels in many patients with major depression. In these patients, diminished corticosteroid receptor expression or function has been postulated as a causative

factor for a deficient feedback of cortisol that may explain their increased HPA activity and stress sensitivity. An alternative hypothesis claims that the primary cause of the HPA dysregulation is an up-regulation of hypothalamic corticotropin-releasing hormone (CRH) levels, which secondarily leads to corticosteroid receptor downregulation in the limbic system and the hypothalamus that perpetuates the disease state. According to the latter concept, the changes of the HPA system may result from a primary disturbance of monoaminergic systems and their widespread connections to higher brain centers, including cortex, limbic system, and hypothalamus.

Corticosteroids exert their effects via two types of receptors: high-affinity receptors for cortisol called mineralocorticoid receptors (MRs) because of their ability to bind also mineralocorticoids (e.g., aldosterone), and low-affinity receptors called glucocorticoid receptors (GRs), which mainly bind cortisol or corticosterone. MRs are predominantly found in the limbic system, in particular in the hippocampus, whereas GRs are ubiquitously expressed throughout the central nervous system. Functionally, MRs are thought to be involved in the physiological maintenance of the HPA system, while GRs are supposed to control the recovery from stress. Corticosteroid receptors are intracellular proteins and function as transcription factors. Their ligands are lipophilic and can easily cross the cytoplasmic membrane. Inside the cell, activation of the receptors provokes their translocation to the nucleus, where they bind to specific sequences on the DNA (glucocorticoid response elements [GREs]) and enhance or repress the transcription rate of their target genes. GRs can also up- or downregulate gene transcription by direct interaction with other transcription factors such as CREB or AP-1.

Activation of corticosteroid receptors in neurons can influence diverse cellular processes such as energy metabolism, signal transduction, and even structural plasticity. Functional consequences include the control of excitability in limbic brain regions, in particular in the hippocampus. The corticosteroid receptor-mediated effects at the cellular level have consequences for processes in which the hippocampal formation plays an essential role, e.g., in the neuroendocrine regulation of the HPA system and also in behavior. Thus, apart from emotional behavior, corticosteroids also influence perception as well as learning and memory.

According to current concepts, GR-mediated functions in particular are disturbed in patients with severe depressive episodes. However, since studies of the GR in the human brain are not feasible *in vivo*, such evidence is mostly derived from pharmacological

challenge tests, such as the dexamethasone suppression test, which indirectly measures GR function in the pituitary and maybe in part also in the central nervous system. Thus, the analysis of direct effects of GR dysfunction and its potential role in emotional behavior is restricted to animal models. For this purpose, targeted mutagenesis in mice seems to be particularly promising.

Animal Models of Depression

Good animal models of human depression should fulfill the following criteria as best as possible: strong behavioral similarities (e.g., similar core symptoms), common etiology, similar pathophysiology, and common treatment. Such a model could then be used to elucidate molecular and biochemical mechanisms underlying the pathogenesis of depression. One of the core symptoms of severe depression in humans and mice is anhedonia, i.e., the lack of interest in pleasurable activities. In mice, anhedonia can be measured by a loss of preference for sweet solutions over tap water. Another core symptom is coping deficits in aversive but avoidable life situations. Confusion between a model and a test should be avoided. A model can be defined as an organism or a particular state of an organism that reproduces aspects of human pathology, providing a certain degree of predictive validity. A test is an end-point behavioral or physiological measure designed to assess the effect of a pharmacological or environmental manipulation. In this respect, models of depression are different from the tests used to monitor the effects of antidepressants. Behavioral despair paradigms such as the Porsolt forced swim test and the tail suspension test cannot be considered models of depression, even though their efficiency to screen for antidepressant-like activity is widely accepted. Nevertheless, increased despair behavior in these tests can be regarded as part of a depressive syndrome.

Two paradigms induce anhedonia in mice: the chronic mild stress model and the learned helplessness paradigm. In the chronic mild stress model, animals are subjected for several weeks to a series of mild stressors (intermittent food and water deprivation, overnight illumination, cage tilt, noise, etc.). This treatment induces a reduction of sucrose preference, decreased intracranial self-stimulation, altered sexual and aggressive behavior, loss of body weight, sleep disturbances, and overactivation of the HPA system. These symptoms are reversible upon administration of antidepressants. In the learned helplessness paradigm, animals are exposed to inescapable electric footshocks, which induce a number of deficits similar to those observed after chronic stress: failure in avoidance learning, lower preference for sucrose solution, decreased appetite, loss

of body weight, and overactivation of the HPA system. Again, these effects are reversed by the administration of antidepressants. Thus, the behavioral and somatic deficits observed in these two models mimic the human depressive syndrome.

A major drawback of the previous models is the difficulty of knowing whether the pathological changes observed are due to the animals' depressive state or rather reflect the manipulation by which this state was induced, e.g., stress or pharmacological treatment. An alternative animal model of affective disorders would be an endogenous, genetic model, independent of external factors, which mimics essential aspects of the human disorder and responds to standard regimens of therapy. Mouse mutants with altered HPA system activity are candidate strains for a murine depression-like syndrome, because HPA system imbalances are a key biological marker for a major depression in humans. In particular, alterations of corticosteroid receptors appear promising in modeling affective disorders. Different strains of mutant mice have been or can be obtained with over- or underexpression of glucocorticoid receptors. Current techniques even allow a conditional gene disruption in specific brain regions, in the best case inducible at a selective time point.

Mice with Targeted Mutations of GR

Several mouse strains with altered GR expression or function have been generated (Table 1). The mice produced earlier (before 2000) did not exhibit a behavioral depression-like phenotype. Thus, mice with a GR point mutation (GR^{dim}) that prevents GR dimerization and binding to its cognate element GRE have not revealed any changes in emotional behavior. Surprisingly, a GR antisense model with reduced expression in the brain as well as a strain with a

brain-specific GR knockout (GR^{NesCre}) exhibited reduced depression-like behaviors. However, these paradoxical findings may be explained by the fact that a complete lack of GR prevents all stress- or corticosterone-induced effects usually mediated by GRs and results in a stress-resistant phenotype. With respect to human genetics and psychopathology, the genetic defects generated in these early mouse strains are very unlikely to occur in humans.

To mimic certain quantitative or functional alterations of GRs – as postulated for depressive patients – mice were generated that have either a 50% reduced or a twofold increased GR gene dosage (Figures 1 and 2). Moreover, in these mice the genetic context controlling the GR gene was preserved as much as possible to maintain the regulatory principles governing the normal gene. Such mice have been generated by homologous recombination (GR^{+/-} mice) and by a transgenic approach (YGR mice). The latter strain contains two additional copies of the GR gene. This strain was generated by transfer of a yeast artificial chromosome of 290 kb in length, providing the transgenes with the natural chromosomal context of the mouse genome, resulting in expression of the GR transgenes in a copy number-dependent and position-independent manner. According to the molecular hypotheses of depression, one would expect that mice with reduced expression levels of GR (GR^{+/-}) have a predisposition to develop depression-like behavior. Current hypotheses would also predict that mice overexpressing GR (YGR) are more resistant to developing depression-like features than wild-type mice. To test these hypotheses, GR^{+/-} and YGR mice were subjected to a standardized battery of tests for emotional behavior, including tests for depression-like signs such as despair and helplessness. Furthermore, the HPA system was analyzed under baseline conditions, after stress, and following a Dex/CRH challenge. Moreover, the content of brain-derived neurotrophic factor (BDNF) was determined in the hippocampus, since BDNF has been postulated to play a critical role in pathogenesis and therapy of affective disorders.

GR^{+/-} Mice Exhibit a Depression-like Behavioral and Neuroendocrinological Phenotype

GR^{+/-} mice did not exhibit phenotypical changes under basal conditions in a test battery for locomotor, exploratory, anxiety-related, and emotional behaviors. In this context, it is of interest that a selective loss of GR function in hepatocytes did not affect gluconeogenesis under basal conditions, but caused a gluconeogenesis deficit under challenge conditions by fasting the animals. Similarly, GR^{+/-} mice exhibited a depression-like phenotype, i.e., increased

Table 1 Mouse strains with targeted GR mutations

Mouse strain	Reference
Transgenic mice with reduced GR expression in the brain	Pepin et al. (1992)
Mice with destructured GR dimerization (GR ^{dim})	Reichardt et al. (1998)
GR ^{NesCre} mice with CNS-specific GR knockout	Tronche et al. (1999)
Transgenic mice with increased GR expression (YGR mice)	Reichardt et al. (2000)
Mice with forebrain-specific GR overexpression	Wei et al. (2004)
GR ^{CamKIIαCre} mice with forebrain-specific GR knockout	Boyle et al. (2004)
GR ^{+/-} mice	Ridder et al. (2005)

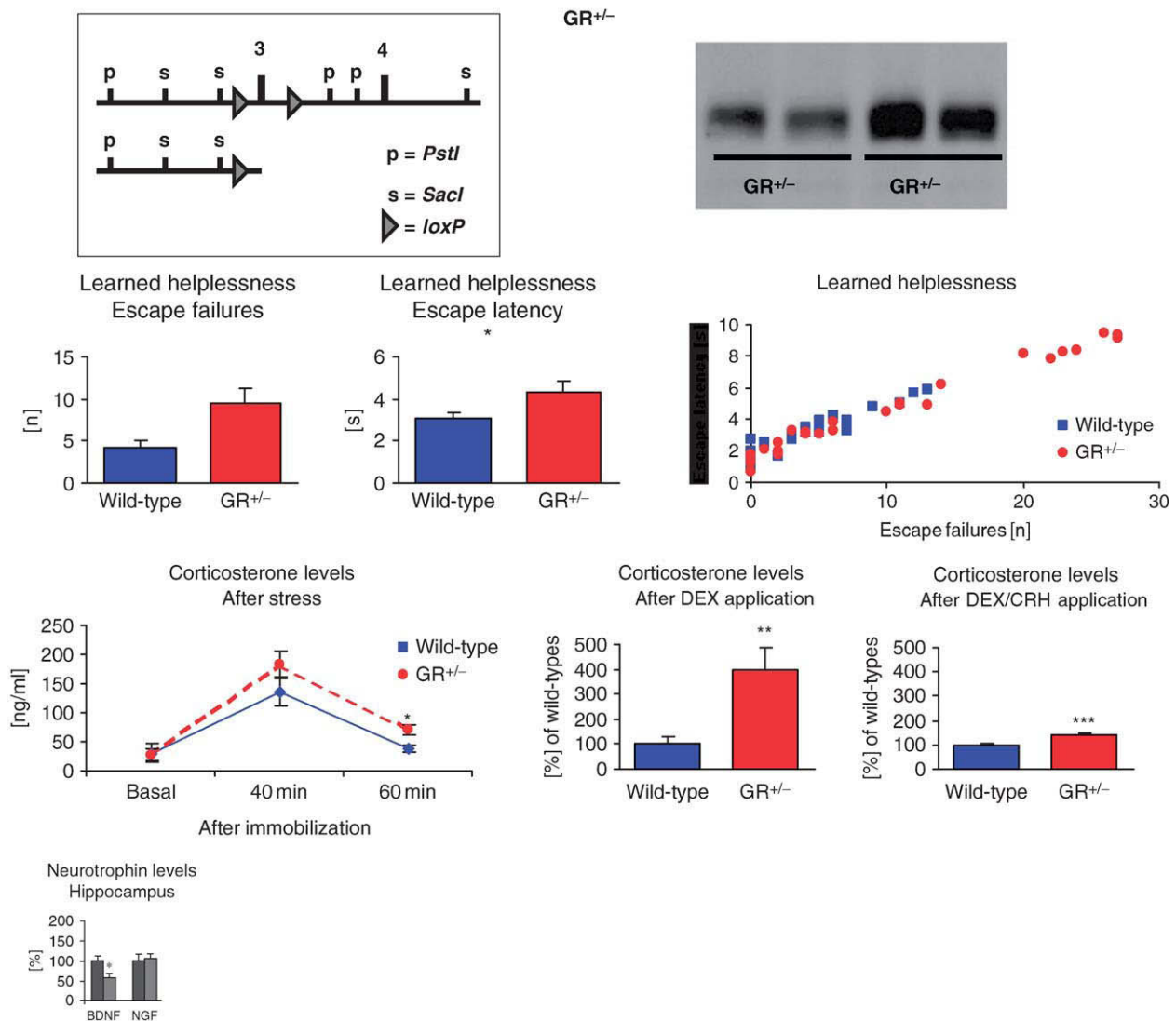


Figure 1 Genetic manipulation. The GR null allele was obtained by deleting the third exon (and a subsequent frame shift), leading to a truncated GR protein without functional domains. Heterozygous mice that underexpress GR exhibit normal baseline behaviors, but after stress exposure they demonstrate significantly increased learned helplessness. Similar to depressed patients, they show a disinhibition of the HPA system and a pathological Dex/CRH test. In agreement with the neurotrophin hypothesis of depression, they have reduced BDNF levels in the hippocampus. Thus, these mice represent a strain with a predisposition to develop depression-like features after stress.

helplessness, after being challenged by environmental stress (Figure 1). The learned helplessness model is a rodent depression model with good face and construct validity. It is based on the concept that depressed individuals show a loss of coping strategies in aversive environmental situations. GR^{+/-} mice demonstrated significantly prolonged escape latencies and increased failures in the active avoidance task of the learned helplessness paradigm after two sessions of footshock exposure. The impaired coping behavior strongly indicated the development of depression-like behavior in GR^{+/-} mice.

This altered behavioral reactivity to stress was accompanied by depression-like alterations of the

HPA system. Similar to the behavioral level, changes of the HPA axis were not present under basal conditions in GR^{+/-} mice, but a significant disinhibition occurred after stress. Furthermore, GR^{+/-} mice exhibited a pathological Dex/CRH test, currently the most relevant biological marker in patients for both florid depression and the risk of developing a depressive episode. This test combines the suppressive effect of dexamethasone with the stimulatory potency of CRH. In 70–80% of severely depressed patients the CRH-elicited adrenocorticotrophic hormone (ACTH) and cortisol response is blunted, and dexamethasone pretreatment does not elicit a suppressive effect and paradoxically induces enhanced cortisol levels

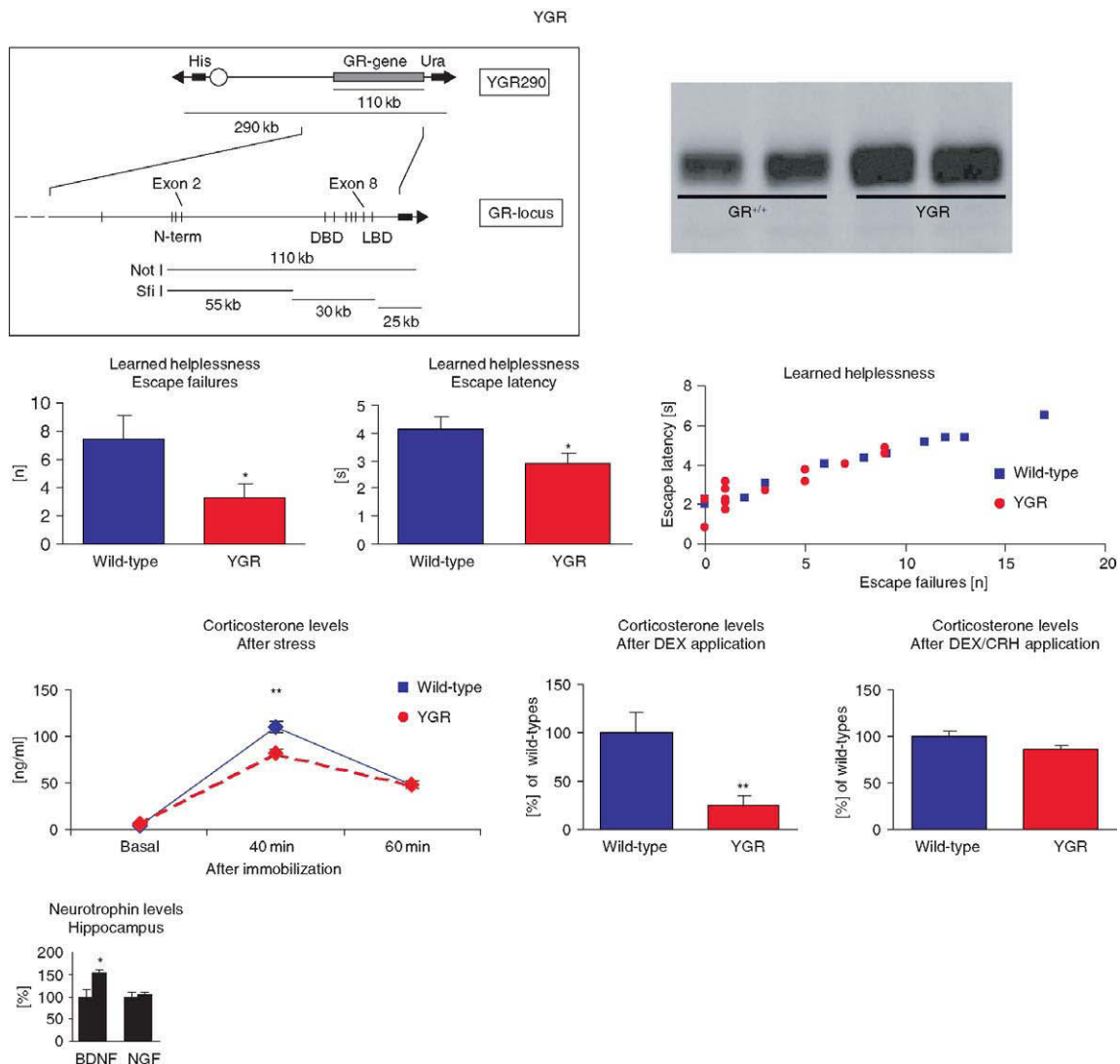


Figure 2 Genetic manipulation. YGR mice were generated by introducing two copies of the GR gene into the mouse genome with the help of the yeast artificial chromosome (YAC), which contains the entire GR locus of 110 kb with all regulatory sequences. Transgenic mice that overexpress GR in the brain are less sensitive to stress. Following immobilization, their HPA system returns to homeostasis faster than in wild-type. Behaviorally, they demonstrate better coping when exposed to stress than control littermates, even though baseline behavior is unaltered. Therefore, these mice represent a depression-resistant strain.

following a CRH challenge. In healthy individuals, in contrast, the Dex/CRH test creates the situation of a pharmacological (partial and transient) adrenalectomy in which hypothalamic CRH expression is increased due to a decrease in plasma cortisol. Similarly to depressed patients, GR^{+/-} mice display a characteristic nonsuppression following dexamethasone treatment and the paradoxical increase of corticosterone levels after combined Dex/CRH application.

Normal Baseline Behavior and Neuroendocrinology of GR^{+/-} Mice May Reflect Developmental Compensation

Using modern transgenic techniques, recently a mouse strain was generated with a forebrain-specific

complete GR knockout, induced by the calcium-calmodulin-dependent protein kinase II (CamKII) promoter. In contrast to GR^{+/-} mice, these mice exhibited at 3 weeks of age a depression-like phenotype under basal (unstressed) conditions. This indicated that GR underexpression from early development on does not lead to an overt phenotype under baseline conditions with low levels of stress. This could be due to a mere gene dosage effect. Furthermore, a complete GR knockout in some brain areas with preserved GR levels in other regions may have resulted in a different phenotype than a more subtle and homogenous GR reduction throughout the brain. However, the depressive phenotype of forebrain-specific GR knockout mice also indicated that crucial

brain regions responsible for the phenotype of GR^{+/-} mice may be located in the forebrain. On the other hand, the lack of a phenotype under baseline conditions of GR^{+/-} mice might also be caused by early ontogenetic compensatory or adaptive processes, possibly reflecting the condition of patients with a high genetic (familial) risk for depression who are initially inconspicuous but develop depressive episodes following stressful life events.

Since depression is a multifactorial disease, the moderate dysregulation of a single system may not be sufficient to induce depression-like alterations in mice under basal conditions. Therefore, provoking or destabilizing factors, such as stress or early childhood environmental factors, may be needed in addition to the existing genetic predisposition to effectively elicit a behavioral and also a neuroendocrinological phenotype. This concept may explain why the depression-like phenotype of GR^{+/-} mice manifests only after stress. The latter fact, however, can be regarded as a strength of the heterozygous model, because depression-vulnerable humans often develop a depressive episode only following external or internal stress factors.

YGR Mice Overexpressing GR Are Stress Resistant

A genetic approach to prove the specificity of the depression-like alterations of GR^{+/-} mice seemed to be the behavioral and neuroendocrinological examination of mice overexpressing GR. In these mice one would expect the opposite phenotype of GR^{+/-} mice. Indeed, YGR mice carrying four copies of the GR gene demonstrate a stress resistance on both the behavioral and the hormonal level. These animals exhibit less helplessness, lower corticosterone levels after stress, and oversuppression following dexamethasone treatment (Figure 2). This result was in good agreement with the predictions of the GR hypothesis of depression.

Recently it was also reported that mice overexpressing GR specifically in the forebrain displayed increased emotional lability in the Porsolt forced swim test, indicated by increased baseline immobility. Furthermore, these animals exhibited increased anxiety-like behavior in the elevated plus-maze and in the dark-light box. This was not observed in the YGR mice that express GR under control of its own regulatory sequences. These contradictory results could be explained by the late onset and forebrain restricted mutation induced by CamKII promoter used in the former study. Conflicting results also come from the human field. While enhanced glucocorticoid feedback has always been considered positive for better limiting the stress response, it is important to remember that overactive GR may transduce a stronger

glucocorticoid signal, i.e., it may produce effects of glucocorticoid excess. In a large elderly Dutch (Rotterdam) population cohort, two different polymorphisms in GR have been found, one that enhances sensitivity to dexamethasone and the other that confers resistance to dexamethasone. However, the group with increased sensitivity to dexamethasone (and lower basal cortisol) shows a glucocorticoid excess phenotype with greater incidence of dementia and reduced overall survival. Thus, our understanding of the animal models needs to examine additional possibilities of what reduced or increased glucocorticoid signaling entails beyond HPA feedback or learned helplessness models.

GR Mutant Mice Represent a Tool to Study Molecular Correlates of Depression-like Behavior

The results in GR^{+/-} mice resemble the human situation in depressive disorders, in which individuals at risk are predisposed to develop depressive episodes after stress. In this respect, GR^{+/-} mice differ from mice with a forebrain-specific complete GR knockout, which already demonstrate depression-like behavioral and neurochemical alterations under baseline conditions. However, the findings in both strains indicate that compromised GR function constitutes a crucial molecular risk factor in the pathophysiology of depression. This suggests that the mouse lines with altered GR expression are suitable tools for further physiological, biochemical, and pharmacological investigations of GR function with regard to depressive disorders. These analyses are also of clinical relevance, since in preliminary studies GR antagonists have been of value in the treatment of severe depressive episodes with psychotic features.

Recently, the so-called neurotrophin hypothesis of depression postulated that a downregulation of brain-derived neurotrophic factor (BDNF) is crucial for the pathogenesis of depression in humans and rodents. A stress- or corticosterone-induced BDNF downregulation in the hippocampus can be partially blocked by adrenalectomy or can be reversed by antidepressant therapy. In accordance with this hypothesis, GR^{+/-} mice – which have a predisposition to stress-induced depression-like behavior – show a significant downregulation of BDNF in the hippocampus, while YGR mice exhibit a significant BDNF upregulation. Indeed, this was the first experimental evidence that compromised GR function concurrently evokes a BDNF dysregulation and a predisposition to depressive behavior. Future experiments should investigate in these strains further mechanisms of neural plasticity thought to be involved in the molecular and

cellular pathogenesis of depression, e.g., monoaminergic systems, arginine vasopressin, and neurogenesis. Using modern molecular techniques such as transcriptomics and proteomics, completely new or even unknown gene products that are involved in the pathogenesis of depression may be detected. Such gene products may represent new targets for antidepressive treatment and could help to develop therapeutic strategies that are faster and more efficient than those currently in use.

See Also the Following Articles

Corticosteroid Receptors; Corticosteroids and Stress; Depression Models; Glucocorticoid Negative Feedback; Glucocorticoids, Effects of Stress on; Glucocorticoids – Adverse Effects on the Nervous System; Glucocorticoids, Overview; Glucocorticoids, Role in Stress; Mineralocorticoid Receptor Polymorphisms.

Further Reading

- Akil, H. (2005). Stressed and depressed. *Nature Medicine* 11, 116–118.
- Altar, C. A. (1999). Neurotrophins and depression. *Trends in Pharmacological Science* 20, 59–61.
- Beato, M., Herrlich, P. and Schütz, G. (1995). Steroid hormone receptors: many actors in search of a plot. *Cell* 83, 851–857.
- Belanoff, J. K., Rothschild, A. J., Cassidy, F., et al. (2002). An open label trial of C-1073 (mifepristone) for psychotic major depression. *Biological Psychiatry* 52, 386–392.
- Boyle, M. P., Brewer, J. A., Funatsu, M., et al. (2005). Acquired deficit of forebrain glucocorticoid receptor produces depression-like changes in adrenal axis regulation and behavior. *Proceedings of the National Academy of Sciences USA* 102, 473–478.
- Charney, D. S. (1998). Monoamine dysfunction and the pathophysiology and treatment of depression. *Journal of Clinical Psychiatry* 59(Supplement 14), 11–14.
- Cryan, J. F. and Mombereau, C. (2004). In search of a depressed mouse: utility of models for studying depression-related behavior in genetically modified mice. *Molecular Psychiatry* 9, 326–357.
- De Kloet, E. R., Vreugdenhil, E., Oitzl, M. S. and Joels, M. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Reviews* 19, 269–301.
- Heuser, I., Yassouridis, A. and Holsboer, F. (1994). The combined dexamethasone/CRH test: a refined laboratory test for psychiatric disorders. *Journal of Psychiatric Research* 28, 341–356.
- Holsboer, F. (2000). The corticosteroid receptor hypothesis of depression. *Neuropsychopharmacology* 23, 477–501.
- Pepin, M. C., Pothier, F. and Barden, N. (1992). Impaired type II glucocorticoid-receptor function in mice bearing antisense RNA transgene. *Nature* 355, 725–728.
- Reichardt, H. M., Umland, T., Bauer, A., Kretz, O. and Schutz, G. (2000). Mice with an increased glucocorticoid receptor gene dosage show enhanced resistance to stress and endotoxic shock. *Molecular and Cellular Biology* 20, 9009–9017.
- Reichardt, H. M., Kaestner, K. H., Tuckermann, J., et al. (1998). DNA binding of the glucocorticoid receptor is not essential for survival. *Cell* 93, 531–541.
- Ridder, S., Chourbaji, S., Hellweg, R., et al. (2005). Mice with genetically altered glucocorticoid receptor expression show vulnerability for stress-induced depressive reactions. *Journal of Neuroscience* 25, 6243–6250.
- Seckl, J. and Meaney, M. J. (2004). Glucocorticoid programming. *Annals of the New York Academy of Sciences* 1032, 63–84.
- Tronche, F., Kellendonk, C., Kretz, O., et al. (1999). Disruption of the glucocorticoid receptor gene in the nervous system results in reduced anxiety. *Nature Genetics* 23, 99–103.
- Wei, Q., Lu, X. Y., Liu, L., et al. (2004). Glucocorticoid receptor overexpression in forebrain: a mouse model of increased emotional lability. *Proceedings of the National Academy of Sciences USA* 101, 11851–11856.

Glucocorticoid Receptor Mutations and Polymorphisms

J W Koper

Erasmus MC, Rotterdam, The Netherlands

© 2007 Elsevier Inc. All rights reserved.

Glucocorticoid Receptor Mutations
Glucocorticoid Receptor Polymorphisms

Glossary

<i>Mutation</i>	Changes to the genetic material.
<i>Polymorphism</i>	The occurrence in the same habitat (DNA sequence) of two or more forms of a trait in such frequencies (usually >1%) that the rarer cannot be maintained by recurrent mutation alone.
<i>SNP</i>	A DNA sequence variation, occurring when a single nucleotide: A, T, C, or G – in the genome differs between members of the species.

Glucocorticoid Receptor Mutations

Some abnormalities in glucocorticoid (GC) response may be explained by mutations in the gene coding for the GC receptor (GR; NR3C1). Such mutations can occur either in the germline, causing systemic GC resistance or hypersensitivity, or they can occur somatically, resulting in localized changes in GC sensitivity, e.g., in tumors.

Germline Mutations, Glucocorticoid Resistance, and Glucocorticoid Hypersensitivity

Based on the results of experiments with knockout mice, it is clear that total absence of the GR is not compatible with life, and the germline mutations that have been identified in the GR gene have always been found to confer a relative GC resistance, which was compensated by increased activity of the hypothalamic-pituitary-adrenal (HPA) axis. Typically, the symptoms of most patients could be explained by the concurrent overproduction of adrenal androgens (hirsutism, fertility problems in women, male precocious puberty in a boy) and mineralocorticoids (hypokalemia, hypertension).

Worldwide, nine different cases of GC resistance caused by mutations in the GR gene have been described (Table 1). Seven of these nine mutations are located in the ligand-binding domain and resulted in reduced ligand affinity. One mutation resulted in the

(heterozygous) destruction of a splice site, and hence in a reduced amount of GR messenger ribonucleic acid (mRNA) and GR protein per cell. One mutation was located in the deoxyribonucleic acid (DNA)-binding domain and resulted in the ablation of trans-activating capacity. The low number of cases of GR-gene mutations stands in strong contrast to the situation with respect to the androgen receptor, where in patients with the partial- or complete androgen insensitivity syndromes hundreds of different mutations have been identified. Several explanations for this are possible: (1) the GR is more essential to (prenatal) survival than the androgen receptor, resulting in the occurrence of relatively mild mutations only, (2) the symptoms of GC resistance are relatively mild, resulting in an underestimation of the number of cases.

Several cases of GC hypersensitivity (typically, the presence of the symptoms of Cushing's syndrome combined with low serum cortisol levels) have been described; however, none of these have been explained by the presence of mutations in the GR gene. Due to the extremely low number of cases there are no data available on the effects of this type of mutations on the stress response.

Somatic Mutations in Tumors and Tumor Cell Lines

GCs play an important role in the treatment of a number of hematological malignancies and lack of response to this treatment might be associated with mutations in the GR gene. However, in the current literature only very limited numbers of such mutations has been reported. Similarly, involvement of GR-gene mutations in adrenocorticotrophic hormone

Table 1 Mutations in the glucocorticoid receptor gene associated with glucocorticoid resistance

Amino acid change ^a	Nucleotide change ^a
R477H	1562G > A
I559N	1808T > A
V571A	1844T > C
2013-2014del, IVSF+1-+2del	2013delGAGT
D641V	2054A > T
G679S	2168G > A
V729I	2317G > A
I747M	2373T > G
L773P	2450T > C

^aNumbering is based upon Genbank Accession number NM_000176.

Table 2 Polymorphisms in the glucocorticoid receptor gene associated with changes in glucocorticoid sensitivity

Amino acid change	Nucleotide change	rs number
NA	5'-UTR, 3447 bp upstream of NM_000176, C → T	rs10052957
ER22/23EK	GAG AGG → GAA AAG	rs6189 and rs6190
N363S	AAT → AGT	rs6195
NA	IVS-2,+646 C → G	NA "Bcl-I"
NA	3'-UTR exon 9β, A → G	rs6198

NA, not available.

(ACTH)-secreting pituitary adenomas has been shown in only a small number of studies.

Glucocorticoid Receptor Polymorphisms

The National Center for Biotechnology Information (NCBI) single nucleotide polymorphism database currently shows 461 SNPs in the human GR gene. Most of these are in noncoding sequences of the gene and have not (yet) been investigated with respect to functionality. Five SNPs in the GR gene have been investigated more extensively (Table 2). The main effects that have been observed for these polymorphisms are on body composition and metabolic profile (fat mass, fat distribution, lean mass, and serum lipids), but two polymorphisms have also been associated with changes in the response to psychosocial stress. Carriers of the N363S polymorphism showed an increased HPA-axis response when challenged with the Trier Social Stress Test, while carriers of the Bcl-I polymorphism showed a reduced response.

In general it can be said that the effects of these polymorphisms are relatively small. They can barely be measured in individuals, and their effects have been suggested to be due to lifelong exposure to small changes in GC sensitivity, rather than to changes in the acute HPA response.

Further Reading

- Bray, P. J. and Cotton, R. G. H. (2003). Variations of the human glucocorticoid receptor gene (NR3C1): pathological and *in vitro* mutations and polymorphisms. *Human Mutation* **21**, 557–568.
- Brufsky, A. M., Malchoff, D. M., Javier, E. C., et al. (1990). A glucocorticoid receptor mutation in a subject with primary cortisol resistance. *Transactions of the Association of American Physicians* **103**, 53–63.

- Charmandari, E., Kino, T. and Chrousos, G. P. (2004). Familial/sporadic glucocorticoid resistance: clinical phenotype and molecular mechanisms. *Annals of the New York Academy of Sciences* **1024**, 168–181.
- Charmandari, E., Raji, A., Kino, T., et al. (2005). A novel point mutation in the ligand-binding domain (LBD) of the human glucocorticoid receptor (hGR) causing generalized glucocorticoid resistance: the importance of the C terminus of hGR LBD in conferring transactivational activity. *Journal of Clinical Endocrinology and Metabolism* **90**, 3696–3705.
- DeRijk, R. H., Schaaf, M. and de Kloet, E. R. (2002). Glucocorticoid receptor variants: clinical implications. *Journal of Steroid Biochemistry and Molecular Biology* **81**, 103–122.
- DeRijk, R. H., Schaaf, M. J. M., Turner, G., et al. (2001). A human glucocorticoid receptor gene variant that increases the stability of the glucocorticoid receptor β-isoform mRNA is associated with rheumatoid arthritis. *Journal of Rheumatology* **28**, 2383–2388.
- Detera-Wadleigh, S. D., Encio, I. J., Rollins, D. Y., et al. (1991). A TthIII1 polymorphism on the 5' flanking region of the glucocorticoid receptor gene (GRL). *Nucleic Acids Research* **19**, 1960.
- Huizenga, N. A., Koper, J. W., De Lange, P., et al. (1998). A polymorphism in the glucocorticoid receptor gene may be associated with and increased sensitivity to glucocorticoids *in vivo*. *Journal of Clinical Endocrinology and Metabolism* **83**, 144–151.
- Hurley, D. M., Accili, D., Stratakis, C. A., et al. (1991). Point mutation causing a single amino acid substitution in the hormone binding domain of the glucocorticoid receptor in familial glucocorticoid resistance. *Journal of Clinical Investigation* **87**, 680–686.
- Karl, M., Lamberts, S. W., Detera-Wadleigh, S. D., et al. (1993). Familial glucocorticoid resistance caused by a splice site deletion in the human glucocorticoid receptor gene. *Journal of Clinical Endocrinology and Metabolism* **76**, 683–689.
- Karl, M., Lamberts, S. W., Koper, J. W., et al. (1996). Cushing's disease preceded by generalized glucocorticoid resistance: clinical consequences of a novel, dominant-negative glucocorticoid receptor mutation. *Proceedings of the Association of American Physicians* **108**, 296–307.
- Malchoff, C. D. and Malchoff, D. M. (2005). Glucocorticoid resistance and hypersensitivity. *Endocrinology and Metabolism Clinics of North America* **34**, 315–326.
- Mendonca, B. B., Leite, M. V., de Castro, M., et al. (2002). Female pseudohermaphroditism caused by a novel homozygous missense mutation of the GR gene. *Journal of Clinical Endocrinology and Metabolism* **87**, 1805–1809.
- Murray, J. C., Smith, R. F., Ardinger, H. A., et al. (1987). RFLP for the glucocorticoid receptor (GRL) located at 5q11–5q13. *Nucleic Acids Research* **15**, 6765.
- Ruiz, M., Lind, U., Gafvels, M., et al. (2001). Characterization of two novel mutations in the glucocorticoid receptor gene in patients with primary cortisol resistance. *Clinical Endocrinology* **55**, 363–371.

- Schaaf, M. J. and Cidlowski, J. A. (2002). Molecular mechanisms of glucocorticoid action and resistance. *Journal of Steroid Biochemistry and Molecular Biology* 83, 37–48.
- Seckl, J. R. (2004). Prenatal glucocorticoids and long-term programming. *European Journal of Endocrinology* 151, U49–U62.
- van Rossum, E. F. C. and Lamberts, S. W. J. (2004). Polymorphisms in the glucocorticoid receptor gene: associations with metabolic parameters and body composition. *Recent Progress in Hormone Research* 59, 333–357.
- van Rossum, E. F. C., Russcher, H. and Lamberts, S. W. J. (2005). Genetic polymorphisms and multifactorial diseases: facts and fallacies revealed by the glucocorticoid receptor gene. *Trends in Endocrinology and Metabolism* 16, 445–450.
- Vottero, A., Kino, T., Combe, H., et al. (2002). A novel, C-terminal dominant negative mutation of the GR causes familial glucocorticoid resistance through abnormal interactions with p160 steroid receptor coactivators. *Journal of Clinical Endocrinology and Metabolism* 87, 2658–2667.
- Wüst, S., van Rossum, E. F. C., Federenko, I. S., et al. (2004). Common polymorphisms in the glucocorticoid receptor gene are associated with adrenocortical responses to psychosocial stress. *Journal of Clinical Endocrinology and Metabolism* 89, 565–573.

Glucocorticoid Receptors See: Corticosteroid Receptors; Membrane Glucocorticoid Receptors.

Glucocorticoids – Adverse Effects on the Nervous System

R M Sapolsky

Stanford University, Stanford, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R M Sapolsky, volume 2, pp 238–243, © 2000, Elsevier Inc.

Background

GCs and the Function and Structure of Hippocampal Neurons

GCs and the Birth of Hippocampal Neurons

GCs and the Death of Hippocampal Neurons

GCs and the Health of the Human Hippocampus

Glossary

Dendritic atrophy The shrinking of the dendritic processes by which neurons receive excitatory or inhibitory inputs from neighboring neurons.

Glucocorticoids Steroid hormones released by the adrenal gland basally and in response to virtually all stressors.

Hippocampus A region of the brain that plays a central role in learning and memory and in the

regulation of the adrenocortical axis; it is atypically vulnerable to a number of neurological diseases and has an atypically high concentration of receptors for glucocorticoids.

Neuron death The death of brain cells that can be due to disease or as part of normal development (i.e., programmed cell death).

It has long been appreciated that prolonged stress or exposure to glucocorticoids (GCs) can have adverse effects throughout the body. In recent decades, it has become clear that these adverse effects can extend to the nervous system, particularly the hippocampus, a brain region critical to learning and memory and with an abundance of corticosteroid receptors. These effects of GCs and stress include (1) reversible atrophy of dendritic processes, (2) inhibition of neurogenesis in the adult hippocampus, (3) compromising of the ability to neurons to survive a variety of insults, and (4) overt neurotoxicity. This article considers the emerging insights as to the cellular and molecular mechanisms underlying these events, as well as the possible clinical implications (including in the realm of normative aging, treatment with exogenous GCs,

Cushing's disease, posttraumatic stress disorder, and depression).

Since the time of Walter Cannon, there has been recognition of the critical role of the stress response in maintaining allostatic balance in the face of insult and injury. However, since the time of Hans Selye, there has been the complementary recognition that excessive activation of the stress response can be deleterious, increasing the likelihood of a disease occurring or exacerbating some preexisting disease. The pages of this encyclopedia are filled with ample documentation of the adverse effects of overactivation of the stress response on metabolism, cardiovascular function, reproductive physiology, and so on, and these realms represent the cornerstones of stress pathophysiology. Only more recently has it become clear that excessive activation of the stress response can have deleterious effects on the nervous system.

Background

An array of physiological and psychological stressors causes secretion of catecholamines by the sympathetic nervous system and of glucocorticoids (GCs) by the adrenal cortex. While such peripheral catecholamines do not pass the blood–brain barrier, the steroidal GCs do so readily and appear to be the culprits in damaging the nervous system.

The deleterious effects of GC excess within the nervous system are centered in (but not exclusive to) the hippocampus, the limbic structure that plays a central role in learning and memory. The first hints of such deleterious actions came at a time when it was not yet obvious why the hippocampus should be particularly sensitive to GCs and thus were mostly ignored. This changed in the late 1960s, with the classic work of Bruce McEwen and colleagues, who demonstrated the extremely high levels of receptors for GCs (i.e., corticosteroid receptors) in the hippocampus. Since then, vast numbers of studies have documented the extensive neurochemical and electrophysiological sensitivity of the hippocampus to GCs, and it is basically impossible to think of stress, GCs, and the nervous system without thinking first about the hippocampus.

In the mid-1980s, E. Ron de Kloet and colleagues made the critical observation that there are two types of corticosteroid receptors in the nervous system, with the hippocampus having ample amounts of both. These are the high-affinity, low-capacity mineralocorticoid receptor (MR), and the low-affinity, high-capacity glucocorticoid receptor (GR). The existence of these two receptors, and their possessing such different affinities, clarify a key feature of GC action in the hippocampus. Specifically, mild

elevations of GCs, or mild transient stressors, enhance hippocampal-dependent cognition, as well as the synaptic plasticity that mediates such cognition (such long term potentiation [LTP]); this makes sense, in that mild, transient stressors are what we think of as stimulation, something that very much enhances cognition. In contrast, major and/or prolonged stressors or GC exposure has opposite, disruptive effects on cognition. This U-shaped curve of GC effects on hippocampal function (a concept popularized by David Diamond) was explained by the two-receptor system in the hippocampus. Specifically, the high-affinity MRs are mostly occupied basally and are then fully occupied by mild elevation of GC concentrations. Such MRs mediate the salutary effects of GCs upon cognition (i.e., the permissive effects of GCs basally, and the enhancing effects of mild stressors). In contrast, GRs, which are minimally occupied basally and are not heavily occupied until there is a major stressor, mediate the disruptive effects.

GCs and the Function and Structure of Hippocampal Neurons

As just discussed, mild stressors, via MR occupancy, enhance hippocampal-dependent cognition and LTP, whereas major stressors, via GR occupancy, disrupt both. The mechanisms underlying these actions are well understood.

Major stressors and prolonged GC exposure can also cause structural changes in hippocampal neurons. Neurons are highly arborized cells, with information flowing in from neighboring neurons via a dense tree of processes called dendrites, and with information flowing out through a major cable called the axon. The extent of dendritic arborization – the length and complexity of processes – is now recognized to be highly dynamic, being remodeled by a variety of local, systemic, and even experiential influences. Work initiated by McEwen and colleagues has shown that in the rat, sustained exposure to stress or to excessive GC levels over the course of weeks can cause atrophy of dendrites in hippocampal neurons. Reflecting the dynamic nature of dendritic remodeling, the atrophy appears to gradually reverse when the stress or GC overexposure stops. Subsequent work from this group and others has shown that the same atrophy occurs in the nonhuman primate hippocampus as well.

It is the GR receptor that appears to mediate the deleterious GC effects on the dendrites. Moreover, GCs appear to cause the atrophy through the glutamate neurotransmitter system. As evidence, atrophy can be prevented by blocking NMDA glutamate receptors in the hippocampus or by the use of drugs

that blunt glutamate release (such as the seizure suppressor phenytoin). Thus, the phenomena of GC-induced dendritic atrophy and of neuroendangerment appear to be linked mechanistically.

What are the consequences of this dendritic atrophy? There are at least three possibilities. First, part of the staggering complexity of the nervous system comes from the multiplicity of synaptic connections between neurons, and a typical neuron can form thousands of synapses with neighbors both near and far. The existence of such neuronal networks is obviously dependent upon an intact dendritic tree, and dendritic atrophy is likely to disrupt and simplify networks within the hippocampus. Stress or GC overexposure on the scale needed to cause dendritic atrophy will also disrupt cognition, and it seems plausible that the two are interrelated. Second, as will be reviewed shortly, GCs and stress can make hippocampal neurons less likely to survive coincident neurological insults. It has been suggested that the dendritic atrophy represents, in effect, a hibernation, of a neuron, an involution that protects it from a coincident insult. Third, in contrast, others have suggested that the atrophy represents a marker of the neuron being made more vulnerable, rather than less, to an insult. While it remains unclear whether the first two models are correct, most existing evidence supports the last.

GCs and the Birth of Hippocampal Neurons

For much of the twentieth century, a central dogma of neuroscience was that the adult mammalian brain did not make new neurons. The 1990s saw an overthrow of this view, with the recognition that a handful of generally ignored pioneering studies in the 1960s disproved this maxim, and the study of adult neurogenesis is now one of the most vibrant fields of neuroscience. Such neurogenesis occurs in only two regions of the brain (the hippocampus being one of them) and is stimulated by environmental enrichment, learning, voluntary exercise, and estrogen. An enormous amount of work has demonstrated that these are, indeed, neurons, that they can migrate some distance from their site of birth, and that they form functional synapses with preexisting neurons. Some of the current issues in the field include (1) whether such neurogenesis occurs in other brain regions as well (particularly the cortex); (2) how much neurogenesis occurs (with estimates ranging from rates that would completely replace certain hippocampal cell fields throughout the lifetime to rates that place the levels orders of magnitude lower); and (3) whether these new neurons are necessary for any

aspects of hippocampal function. These issues are quite contentious in this often fractious field.

One of the more consistent findings is that stress and GCs potently and rapidly (over the course of hours) inhibit adult hippocampal neurogenesis, a finding pioneered by Elizabeth Gould and colleagues. The mechanisms underlying this are unclear. As but one example, early reports suggested that neural progenitor cells in the adult hippocampus do not express MRs or GRs, and that the inhibitory effects of GCs had to be through neuron/progenitor interactions; more recent work, using more sensitive assays, suggests that progenitors do indeed possess GRs. Also unclear is how much the effects of GCs on neurogenesis contribute to any of the disruptive effects of GCs upon hippocampal function.

GCs and the Death of Hippocampal Neurons

GC Endangerment

A body of work initiated by Robert Sapolsky and colleagues in the 1980s demonstrated the ability of transient bursts of GC excess to endanger hippocampal neurons, compromising their ability to survive a coincident neurological insult. As such, the greater the GC levels at the time of an insult, the greater the extent of hippocampal damage. Such insults include epileptic seizures, hypoxia-ischemia, hypoglycemia and other metabolic insults, oxidative insults, the β -amyloid peptide (the fragment of the amyloid precursor protein that has been implicated in the neuron death of Alzheimer's disease), and gp120 (the coat protein of the AIDS virus that is thought to play a role in the neurodegeneration seen in AIDS-related dementia). This endangerment is of a considerable magnitude, with stress levels of GCs causing many-fold increases in insult-induced damage in some cases.

Features of this endangerment include the following: (1) it is mediated by GRs, rather than MRs; (2) while most dramatic in the hippocampus, the endangerment can also occur in the cortex and striatum; (3) it is a direct GC effect in the hippocampus (rather than secondary to some peripheral GC action), as the same phenomenon occurs in hippocampal cultures and tissue slices; and (4) it consists of GCs increasing the amount of insult-induced necrotic death, rather than an increase in the amount of apoptotic death (a form of programmed cell death).

In vitro models have made it possible to uncover some of the mechanisms underlying the endangerment. One realm concerns the disruptive effects of GCs upon neuronal energetics. GCs inhibit glucose transport in many peripheral tissues (as part of the

strategy to divert energy to exercising muscle during a stressful crisis); this effect is due to a reduction in the number of glucose transporter molecules in the cell membrane. GCs cause a similar inhibition in the hippocampus, leaving neurons energetically vulnerable, thereby worsening insult-induced declines in energy substrates (such as ATP). Such insults cause neurons to release damaging amounts of the excitatory neurotransmitter glutamate. This, in turn, causes a wave of calcium influx in the postsynaptic neuron, which causes an array of calcium-dependent degenerative events; of most significance is the generation of oxygen radicals. It is quite costly for a neuron to contain the excesses of glutamate, calcium, and oxygen radicals, and the mild energetic vulnerability caused by GCs makes each of these steps spiral a bit more out of control, increasing the likelihood of the neuron succumbing.

In addition to these energetic features, GC endangerment probably also arises from the demonstrated abilities of the hormone to increase calcium currents in hippocampal neurons, to decrease the activity of calcium-extrusion mechanisms, to suppress the levels or efficacy of neurotrophins (neuronal growth factors), and to disrupt some of the defenses normally mobilized by neurons during neurological insults.

At present, these data are derived exclusively from rodent and tissue culture studies. Should subsequent work demonstrate GC endangerment in the primate hippocampus, they have a number of disturbing implications. Numerous individuals are administered high-dose GCs to control autoimmune or inflammatory disorders, and this might well exacerbate the damaging effects of some neurological crisis. Plus, GCs are often administered after strokes in order to decrease brain edema (swelling). Most studies show that GCs are less effective at reducing such edema than nonsteroid anti-inflammatory compounds; the present studies suggest that the use of GCs might even worsen the neurological outcome. Finally, stroke, seizure, and cardiac arrest all cause massive GC secretion; such endogenous GC release might add to damage. In effect, what we have learned to view as the typical extent of neurological damage caused by these insults might well reflect the damaging effects of the maladaptive hypersecretion of GCs at that time.

GC Neurotoxicity

Independent of a coincident neurological insult, GCs can directly influence the life and death of a hippocampal neuron. In an unexpected finding, Robert Sloviter and colleagues reported that GCs are required for the survival of dentate gyrus neurons. In the complete absence of the hormone, such neurons undergo apoptosis within days, and

very low levels of GCs, via MR occupancy, prevent this effect.

A separate literature showed that truly prolonged exposure to elevated GC levels could directly kill hippocampal neurons. Studies initiated by Phil Landfield and colleagues showed that the loss was centered in the CA3 region of the hippocampus and was highly relevant to hippocampal aging. That region of the hippocampus loses neurons with age, and these studies suggested that it is the cumulative extent of GC exposure over the lifetime that acts as a major determinant of the extent of neuron loss (as well as of hippocampal-dependent memory problems) in old age. As such, stress has the capacity to accelerate hippocampal aging, whereas behavioral or social manipulations that reduce cumulative GC exposure can delay such aging. The finding of GC- and of stress-induced neurotoxicity was replicated by other investigators, including in the nonhuman primate hippocampus.

The mechanisms underlying such neurotoxicity are not well understood. The assumption by most in the field, however, is that it is on a continuum with GC neuroendangerment – in effect, if a day of GC overexposure can compromise the ability of a neuron to survive a massive neurological insult, months of overexposure can compromise its ability to survive the minor metabolic challenges of everyday life, especially during aging. The difficult studies required to test this idea remain to be carried out.

Despite the replication of the finding of GC neurotoxicity, the phenomenon is more controversial than the other adverse GC effects that have been discussed. The disagreements have focused on a number of issues. The basic finding of neuron loss during hippocampal aging has been called into question by some researchers counting neurons with the technique of stereology. Some investigators suggest that overt neurotoxicity can be caused only by massive amounts of ever-shifting stressors (to avoid habituation of the system) or by pharmacological GC exposure, questioning the relevance of the phenomenon to normative aging. GC hypersecretion and hippocampal neuron loss in old age are quite variable among rodents. There appear to be strains in which neither occurs to any dramatic extent. Moreover, dramatic differences in the extent of successful aging can occur within strains. Michael Meaney and colleagues, for example, have shown that remarkably subtle differences in the quality of early maternal care can change the entire trajectory of hippocampal and adrenocortical aging in a rat, in some cases completely avoiding the neuron loss or GC hypersecretion; thus, neither component of that cascade model is an inevitable part of aging. Finally, to the extent that

primates and humans hypersecrete GCs in old age, it is a nonlinear phenomenon, emerging only in extreme old age or when coincident with an insult (such as major depression or Alzheimer's disease).

GCs and the Health of the Human Hippocampus

Work in the past decade suggests that these findings regarding the deleterious effects of GCs and stress apply to the human hippocampus. These findings have been most striking in a number of realms.

Cushing's Syndrome

Cushing's syndrome can arise from any of a number of different types of tumors and results in vastly elevated levels of circulating GCs. Magnetic resonance imaging (MRI) studies by Monica Starkman and colleagues of Cushingoid patients have shown selective atrophy of the hippocampus. More strikingly, it is the same patients who have the most severe GC hypersecretion, the most severe hippocampal atrophy, and the most profound cognitive deficits. Post mortem studies to uncover whether this atrophy is due to neuronal loss, atrophy, or compression remain to be carried out. However, when the GC hypersecretion is corrected by removal of the tumor, the hippocampal gradually returns to normal size, in agreement with the model of reversible dendritic atrophy.

Exogenous GC Administration

Synthetic GCs are frequently administered in different realms of clinical medicine. Most cases involve transient exposure to fairly low concentrations, often through routes that do not result in much access of the steroids to the brain. However, there are hundreds of thousands of individuals on long-term, systemic high-dose GC treatment to control various autoimmune or inflammatory diseases. Difficult case-controlled work by Pamela Keenan and colleagues demonstrates that independent of the disease, treatment with such steroids increases hippocampal-dependent cognitive deficits. The mechanisms underlying this effect (e.g., whether it is due to dendritic atrophy, inhibition of neurogenesis, and so on) are unknown.

Clinical Depression

Major clinical depression is an archetypical example of a stress-related disease; as but three pieces of evidence, exogenous GC treatment increases the risk of depression, major stressors do as well, and about half of patients with major depression present with GC hypersecretion. Yvette Sheline and colleagues have

observed that a history of prolonged depression is associated with selective hippocampal atrophy on MRI. In this observation, since replicated, the more prolonged the depression, the more atrophy. While the necessary post mortem cell counting has not been carried out, the atrophy appears to be permanent in that it is demonstrable in individuals long after the depressions have resolved; whether this is due to loss of neurons or inhibition of neurons that would otherwise have been born is not clear. Some researchers have suggested a scenario of reverse causality, in which individuals who (for other unrelated reasons) have a smaller than average hippocampus are more at risk for clinical depression. However, such volume loss is not demonstrable shortly after diagnosis, and more prolonged depressions are associated with more volume loss; both findings argue against reverse causality.

Posttraumatic Stress Disorder (PTSD)

Probably the most contentious studies concerning GC effects on the human hippocampus concern PTSD. A large numbers of studies, with the first coming from J. Douglas Bremner and colleagues, demonstrate volume loss in the hippocampi of individuals with PTSD. Furthermore, the more volume loss, the more severe the cognitive impairments, and there is little evidence that either tends to reverse over time. The controversial issues are threefold.

First, are GCs the cause of the volume loss? This question is central to the second issue as well, but the first version of this question focuses on whether GC levels are elevated in PTSD. While it might seem intuitive that this is the case, many (but not all) studies suggest lower than normal levels and that such hyposecretion may even precede and predispose toward succumbing to PTSD. The studies showing hyposecretion, pioneered by Rachel Yehuda and colleagues, also suggest an enhanced sensitivity to GCs in these patients. This would then shift the issue from whether excessive GC levels play a role in the volume loss to whether excessive sensitivity to GCs plays a role.

The second controversy concerns when the volume loss occurs. There are three possibilities: (1) volume loss occurs in response to the trauma itself; (2) volume loss occurs in response to the biology of the prolonged posttraumatic period (i.e., the period of the PTSD); and (3) a small hippocampus (arising for reasons unrelated to stress or GCs) precedes and predisposes toward PTSD. This issue is not resolved, and, at present, there are some striking findings supporting, as well as arguing against, each of these possibilities. One reasonable solution is that from the

standpoint of hippocampal neurobiology, PTSD is a heterogeneous disorders, with all three occurring on occasion. Such heterogeneity is certainly seen in the clinical profile of the disorder.

The final issue is whether the volume loss is due to loss of neurons, blockage of birth of new neurons, atrophy of dendritic processes, loss of glia, and so on. As with the other realms of studies of the human hippocampus, resolving this issue requires difficult quantitative postmortem studies.

Obviously, many issues remain to be resolved and are likely to have some important consequences in neurology, psychiatry, and gerontology. Most broadly, the interrelated phenomena of GC neuroendangerment, inhibition of neurogenesis, dendritic atrophy, and neurotoxicity shift the realm of stress-related disease to the nervous system. For nearly 70 years, stress physiology has taught that cardiovascular function, metabolism, and immunity can be sensitive to the adverse effects of stress. This newer literature demonstrates that this vital and fragile portion of our bodies, the brain, can be sensitive to stress as well.

See Also the Following Articles

Alzheimer's Disease; Brain Trauma; Corticosteroid Receptors; Corticotropin Releasing Factor (CRF); Cushing's Syndrome, Medical Aspects; Cushing's Syndrome, Neuropsychiatric Aspects; Glucose Transport; Hippocampal Neurons; Hippocampus, Corticosteroid Effects on; Post-traumatic Stress Disorder, Neurobiology of.

Further Reading

- Gould, E. and Gross, C. (2002). Neurogenesis in adult mammals: some progress and problems. *Journal of Neuroscience* 22, 619–623.
- McEwen, B. (2002). Sex, stress and hippocampus: allostasis, allostatic load and the aging process. *Neurobiology of Aging* 23, 921.
- Sapolsky, R. (1999). Stress, glucocorticoids and their adverse neurological effects: relevance to aging. *Experimental Gerontology* 34, 721.
- Sapolsky, R. (2000). Glucocorticoids and hippocampal atrophy in neuropsychiatric disorders. *Archives of General Psychiatry* 57, 925.

Glucocorticoids, Effects of Stress on

J C Buckingham

Imperial College School of Medicine, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 229–237, © 2000, Elsevier Inc.

Mechanisms Controlling Glucocorticoid Secretion in Stress
Glucocorticoid Secretion in Repeated or Sustained Stress
Age- and Gender-Related Changes in the Stress Response

Glossary

Hypothalamo-hypophyseal portal vessels

A specialized portal blood system responsible for the transportation of neurohormones (termed releasing hormones) from the median eminence area of the hypothalamus to the anterior pituitary gland.

Paraventricular nuclei (PVNs)

Specialized areas of the hypothalamus lying on either side of the third ventricle; these nuclei include the cell bodies of neurons that project to the median eminence area of the hypothalamus, and they play a key role in initiating the pituitary-adrenal response to stress.

The glucocorticoids (cortisol and/or corticosterone, depending on the species) are steroid hormones produced by cells in the zona fasciculata of the adrenal cortex under the influence of the hypothalamic-pituitary complex (**Figure 1**). They exert diverse actions in the body and thereby play a key role in the regulation of crucial homeostatic mechanisms, such as those concerned with metabolic control and immune/inflammatory responses. The actions of the steroids are effected intracellularly, mainly via glucocorticoid receptors (GRs) but also via mineralocorticoid receptors (MRs). In normal circumstances, the rate of glucocorticoid secretion is maintained within narrow limits, varying according to a circadian rhythm that is coupled to the sleep-wake cycle; serum levels of the steroids are thus normally maximal at the end of the sleep phase and fall to a nadir some hours after waking. However, glucocorticoid secretion is increased markedly in conditions of physical and emotional stress; hence, the glucocorticoids are frequently termed stress hormones. The stress-induced changes in circulating glucocorticoids are superimposed on the existing circadian tone and vary in their rate of onset, magnitude, and duration according to the nature, duration, and intensity of the insult. Significant

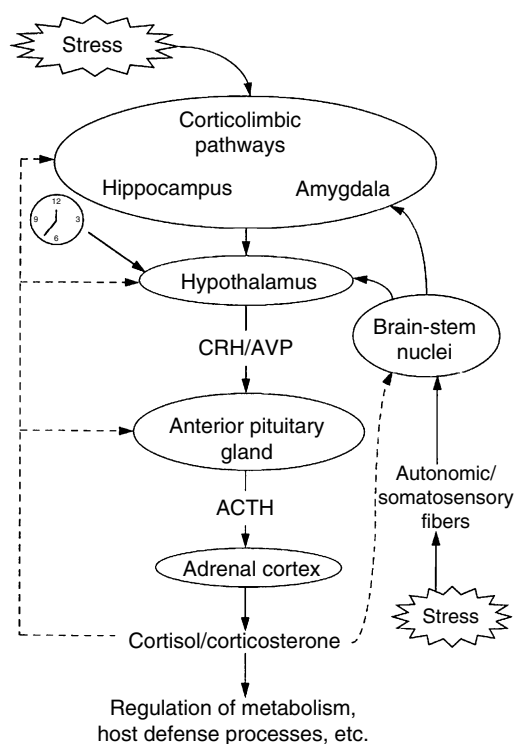


Figure 1 The hypothalamic-pituitary-adrenal axis. Dotted lines indicate inhibitory influences; intact lines indicate stimulatory influences; the clock indicates circadian influences. ACTH, adrenocorticotropic hormone; AVP, arginine vasopressin; CRH, corticotropin releasing hormone.

interspecies variations also occur; for example, humans are particularly sensitive to emotional trauma, whereas rats respond very readily to physical stimuli such as cold and pigs are relatively unresponsive to many stimuli (e.g., histamine and cold) but show a marked elevation in serum cortisol when deprived of food.

Mechanisms Controlling Glucocorticoid Secretion in Stress

The Hypothalamic-Pituitary-Adrenocortical Axis

The release of the glucocorticoids in stress is orchestrated mainly by the hypothalamus, which receives and integrates neural and humoral information from many sources and initiates the release of two neurohormones, termed corticotropin releasing hormone (CRH) and arginine vasopressin (AVP), into the hypothalamo-hypophyseal portal vessels. These peptides are transported to the anterior pituitary gland via the portal vessels, where they act synergistically to cause the release of adrenocorticotropic hormone (ACTH), which, in turn, stimulates cortisol/corticosterone secretion. The hypothalamus thus provides a

vital link between the perception of physical and emotional stress and the regulation of key homeostatic mechanisms. The responsivity of the hypothalamic-pituitary-adrenocortical (HPA) axis to incoming stimuli is modulated by a servo system through which the sequential release of CRH/AVP and ACTH from the hypothalamus and anterior pituitary gland, respectively, is regulated negatively by the glucocorticoids themselves (Figure 1); the magnitude of the HPA response to stress thus depends on the preexisting glucocorticoid tone.

Adrenocorticotropic Hormone and Steroidogenesis

Cortisol and corticosterone are not stored to any great degree in the adrenal cortex but are synthesized on demand and promptly pass by diffusion across the cell membrane into the systemic circulation. The primary trigger for the synthesis of the steroids is ACTH, a polypeptide hormone produced by a specialized subgroup of cells in the anterior pituitary gland, the corticotrophs. ACTH acts via specific membrane-bound receptors, termed type 2 melanocortin (MC2) receptors, to activate the rate-limiting enzyme in the steroidogenic pathway; it thus causes the conversion of cholesterol to pregnenolone, from which cortisol and corticosterone are formed rapidly. In addition, ACTH acts on the vasculature to increase adrenal blood flow and thereby aid the passage of the newly synthesized steroids out of the cells and into the general systemic circulation. Increases in circulating glucocorticoids may be apparent within 10 min of the onset of the stress, although, in some instances (e.g., following an immune insult) the response may be delayed. Approximately 95% of cortisol/corticosterone in blood is protein bound, mainly to a specific high-affinity corticosteroid-binding globulin (CBG) but also to albumin, which has a lower affinity for the steroids. In principle, only the free steroid has ready access to the corticosteroid receptors (GRs and MRs), which are located intracellularly in the target cells. However, local mechanisms that release the steroid from its carrier protein(s) appear to exist in at least some target tissues; these include (1) interactions of the carrier protein with cell surface CBG receptors and (2) enzymatic cleavage of CBG (e.g., by serine proteases in inflamed tissues), a process that lowers the affinity of the binding protein for the steroids. The access of the steroids to their receptors is further regulated within the target cells by 11β -hydroxysteroid dehydrogenase (11β -HSD) enzymes, which control the interconversion of cortisol and corticosterone to their inactive counterparts, cortisone and 11 -dehydrocorticosterone. The biochemical half-life of the steroids is approximately 90 min, but

the biological effects may be sustained for very much longer. Metabolism is hepatic and the resulting metabolites are conjugated and excreted in the bile and the urine.

Synthesis and Secretion of Adrenocorticotrophic Hormone

ACTH is synthesized in the corticotrophs by the post-translational cleavage of a large precursor molecule, proopiomelanocortin (POMC). POMC also gives rise to a number of other peptides that are coreleased with ACTH and are termed collectively ACTH-related or POMC peptides. These include an N-terminal peptide, N-POMC(1–76), of unknown function and β -lipotropin (β -LPH), which, in turn, may be partially degraded to β -endorphin (which is sometimes used as a marker of ACTH secretion) and γ -LPH. In some circumstances (e.g., adrenal insufficiency), more extensive N-terminal processing may occur, either within the pituitary gland or at the adrenal level, generating peptides that appear to augment the steroidogenic capacity of the adrenal cortex by potentiating the steroidogenic response to ACTH (γ -melanocyte-stimulating hormone) or enhancing adrenocortical mitogenesis (N-POMC(1–46)).

ACTH is released from the anterior pituitary gland into the bloodstream in an episodic manner under the direction of hypothalamic CRHs. The frequency of pulses is normally fairly constant over the 24-h period, and it is changes in the amplitude that give rise to the circadian profile of ACTH and, hence, glucocorticoid secretion. Increases in pulse amplitude also underpin the stress-induced excursions in pituitary-adrenocortical activity although, in some instances, changes in pulse frequency may also occur.

Hypothalamic Control of Adrenocorticotrophic Hormone Secretion

The concept that the secretion of ACTH is driven by a neurohormone, termed corticotropin-releasing factor (CRF), that is released from the hypothalamus and transported to the pituitary gland via the hypothalamic-hypophyseal portal vessels was first recognized in the late 1940s by Harris. However, some 35 years elapsed between Harris's pioneering work and the identification of the active principles. By the late 1970s it was becoming apparent that CRF activity of the hypothalamus could not be ascribed to a single factor, and the concept of a multifactorial CRF rapidly gained credence. Subsequent work identified a 41-amino-acid residue neuropeptide, termed CRH, as the principal component and also pointed to a key role for AVP.

Corticotropin releasing hormone The 41-amino-acid neuropeptide, CRH, was first identified in ovine hypothalami by Vale and colleagues; the human peptide was identified subsequently and shown to differ from the ovine peptide by seven amino acids but to be identical to the rat peptide. The CRH gene is well characterized in the rat, mouse, and human and is expressed at multiple sites within the central nervous system (e.g., the hypothalamus, amygdala, and brain stem) and also in the periphery. CRH destined for the anterior pituitary gland is synthesized in the cell bodies of parvocellular neurons, which originate in the paraventricular nucleus (PVN) of the hypothalamus and project to the external lamina of the median eminence. These neurons terminate in close apposition with the capillaries of the hypothalamo-hypophyseal portal complex into which they release their secretory products for transportation to the anterior pituitary gland. The concentration of CRH in hypophyseal portal blood is high (vs. peripheral blood) and alters in parallel with perturbations of pituitary-adrenal function (e.g., stress), as also does the expression of the peptide and its mRNA in the PVN–median eminence tract. The actions of CRH on the corticotrophs are effected by specific type 1 CRH receptors (CRH-R1s), and the blockade of these receptors with selective antagonists attenuates corticotropic responses to specific stressors and to exogenous CRH. The stress response is also impaired in mice in which genes encoding either CRH or CRH-R1 are ablated, although resting ACTH levels appear to be normal in mutant mice. CRH-R1s are coupled positively to adenyl cyclase via a G_s -protein. The stimulation of the receptor therefore causes a rise in intracellular cAMP with subsequent activation of protein kinase A, which phosphorylates Ca^{2+} channels, thereby permitting Ca^{2+} influx and the release of ACTH by exocytosis. The activation of CRH-R1s in the pituitary gland also augments POMC gene expression and hence ACTH synthesis and, in the longer term, causes the proliferation of the corticotrophs.

Arginine vasopressin The ACTH-releasing activity of AVP was first recognized in the 1950s, and evidence that AVP may be an important component of a multifactorial corticotropin releasing complex emerged originally from studies performed in rats congenitally lacking AVP (Brattleboro strain) and was subsequently supported by work using specific antisera and antagonists to the peptide. AVP *per se* shows only weak ACTH-releasing activity, acting via the V1b subclass of AVP receptors, which use the phosphatidylinositol signal transduction system.

However, it also potentiates the corticotropic responses to CRH markedly and thus plays a key role in determining the magnitude of the stress response. Surprisingly, the biochemical basis of this important synergistic response is poorly defined, although it is apparent that the activation of the V1b receptor potentiates the accumulation of 3'5' cAMP induced by CRH, probably by a mechanism involving protein kinase C-dependent protein phosphorylation. AVP destined for the corticotrophs is derived primarily from a population of CRH-containing parvocellular neurons projecting from the PVN to the median eminence; thus, as for CRH, the concentration of AVP in the portal system is relatively high (vs. peripheral blood) and is increased further by stress. In addition, some AVP may be directed to the corticotrophs from magnocellular neurons of the neurohypophyseal system (the primary source of AVP in peripheral blood) via the interconnecting short portal vessels. Within the parvocellular system, CRH and AVP are colocalized in the same secretory granules and may therefore be cosecreted. However, histological studies suggest that not all CRH-positive neurons express AVP in normal circumstances, thus permitting the ratio of AVP : CRH released to vary in a stimulus-specific manner; equally, not all corticotrophs express receptors for AVP, thereby permitting another level of control. Although AVP is generally considered to be the second CRH, this is not always the case. For example, in the sheep, AVP has precedence over CRH; similarly, in rodents and possibly in humans, AVP appears to be the dominant factor driving ACTH release in conditions of repeated or sustained stress.

Other hypothalamic factors Other hypothalamic factors may also contribute to the regulation of ACTH, including the pituitary adenylyl cyclase-activating peptide (PACAP), which exerts a stimulatory influence, and atrial natriuretic peptide (ANP) and substance P, both of which exert inhibitory influences.

Stress-Induced Activation of Corticotropin Releasing Hormone/Arginine vasopressin Neurons

The parvocellular CRH/AVP neurons in the hypothalamus are well positioned to monitor, integrate, and respond to emotional insults and physical changes in the internal or external environment that threaten homeostasis (i.e., stress). Significantly, they receive inputs from many ascending and descending nervous pathways; in addition, they are sensitive to a variety of substances that may be relayed by the bloodstream or the cerebrospinal fluid (e.g., glucose and steroids) and to local factors released from interneurons or

glial cells (e.g., eicosanoids and cytokines). Major roles have been identified for (1) inputs from the brain-stem nuclei, particularly the nucleus tractus solitarius (NTS), that relay sensory information directly to the PVN and also to other parts of the brain and (2) pathways projecting to the PVN from the prefrontal cortex and the limbic system (hippocampus; amygdala; lateral septum; and the bed nucleus of the stria terminalis, BNST), but other pathways, such as those arising from raphe nuclei, are also likely to contribute. Although the precise complement of neural pathways, humoral substances, and local mediators that orchestrates the release of CRH/AVP in stress almost certainly depends on the nature of the stimulus and the status of the individual at the time of the stress, some classification according to the type of stress is now possible. For example, HPA responses to visceral homeostatic stressors that lack a cognitive component (e.g., acute hypotension) are initiated by afferent fibers that signal via the brain-stem nuclei and the noradrenergic pathways that project from the brain stem to the PVN; some homeostatic stressors also signal via humoral mechanisms; for example, the responses to hypoglycemia and hypertonicity involve intrahypothalamic glucose- and osmotic-sensitive neurons, respectively. Visceral stressors with a significant cognitive component (e.g., restraint, foot shock, or cold) use somatosensory afferents and central pathways involving both the brain-stem nuclei and the corticolimbic system to activate the PVN. The corticolimbic system also plays a fundamental role in effecting the HPA responses to emotional stressors; the importance of the hippocampus, which is critical to cognitive function, particularly learning and memory function, is well documented in this regard, as is the role of the amygdala, a key area in coordinating the behavioral, autonomic, and neuroendocrine responses to fear or anxiety.

Challenges to the host defense system, such as those provoked by infection, injury, or inflammation, pose a major threat to homeostasis and thus constitute a severe stress. Not surprisingly, in experimental animals and in humans the HPA axis responds to such insults with a marked increase in its activity. This response is driven by the army of mediators (e.g., interleukins, eicosanoids, histamine, and phospholipase A2) released from the activated immune/inflammatory cells. Some of these agents may target the HPA axis directly, acting mainly at the hypothalamic level to trigger CRH/AVP release but also on the anterior pituitary gland and adrenal cortex to augment or sustain the final adrenocortical response. Others may act locally at the site of the infection or inflammatory lesion to stimulate sensory fibers and

thereby activate the central pathways that precipitate CRH/AVP release. In addition, many of the mediators provoke widespread pathological effects in the body (e.g., hypotension or hypoglycemia), which themselves stimulate the axis by the mechanisms outlined earlier. The HPA response to immune insults is thus a complex event potentially requiring the integration of a broad spectrum of stimuli, the complement of which reflects the nature and intensity of the insult. Much research has focused on the influence of the cytokines, interleukin (IL)-1 β and IL-6, which are particularly important in initiating the HPA response to bacterial and viral infections. Three main mechanisms have been proposed to explain their actions.

1. The cytokines act directly on the PVN to trigger CRH release via a mechanism dependent on prostanoid generation. This mechanism has been refuted on the basis that the cytokines, which are large polypeptides (17–26 kDa), are unlikely to penetrate the blood–brain barrier and thereby gain access to their targets. However, the possibility that cytokines enter the brain via fenestrated regions of the barrier (e.g., organum vasculosum lamina terminalis, OVLT), areas in which the permeability is increased by local inflammation, or the recently described transporter system cannot be discounted.

2. The cytokines act at the level of the blood–brain barrier itself, exerting prostanoid-dependent actions on the endothelium that trigger CRH/AVP release by mechanisms that may include the activation of neural pathways and/or induction of *de novo* synthesis of cytokines in the hypothalamus and elsewhere in the brain that are required to drive the neuroendocrine response. This concept is supported by the presence of IL-1 receptors in the endothelial cells and by increased IL-1 mRNA expression in the hypothalamus and other brain areas of rodents subjected to endotoxin challenge.

3. Hypothalamic responses to peripheral cytokines involve the sequential activation of sensory fibers, brain-stem nuclei (notably the NTS), and the ventral noradrenergic pathways (and possibly others) to the PVN. This view concurs with claims that the expression of IL-1 β mRNA in the hypothalamus and the hypersecretion of ACTH release induced by endotoxin are both dependent on the integrity of the vagus and that HPA responses to local inflammatory lesions induced by turpentine require C fibers. However, it is at odds with reports that cytokines activate the NTS, and hence connections between the NTS and the PVN, by mechanisms that are independent of the vagus but which may involve local prostanoid-dependent interactions with the blood–brain barrier.

Negative Feedback Control of the Hypothalamic-Pituitary-Adrenal Axis

Role of glucocorticoids Adrenocortical responses to incoming stimuli are effectively contained within appropriate limits by the glucocorticoids themselves, which exert powerful inhibitory influences on the secretion of ACTH and its hypothalamic-releasing hormones. Thus, elevations in circulating glucocorticoid levels brought about, for example, by the administration of exogenous steroids effectively suppress the HPA responses to stress and also the circadian rise in pituitary-adrenocortical activity. Conversely, the ACTH response to stress is exaggerated in subjects with adrenocortical insufficiency (e.g., Addison's disease) but corrected by glucocorticoid replacement therapy. The feedback actions of the steroids are highly complex because they are exerted at multiple sites within the axis, effected by multiple molecular mechanisms, and operate over three time domains: rapid or fast feedback, early delayed feedback, and late delayed feedback.

Phases of steroid feedback Rapid feedback develops within 2–3 min of an increase in glucocorticoid levels. It is of short duration (<10 min) and is followed by a silent period during which the stress response is intact, even though the circulating steroid levels may remain elevated. Rapid feedback is sensitive to the rate of change rather than the absolute concentration of steroid in the blood and may therefore provide an important means whereby the HPA response to stress is terminated; alternatively, it may serve to blunt the responses to a second stress incurred at a time when the steroid levels are still rising in response to the first stress. The early delayed phase of feedback develops approximately 0.5–2 h after an acute elevation in plasma steroid levels. It is normally maximal within 2–4 h and, depending on the intensity of the stimulus, may persist for up to 24 h. Both the circadian and the stress-induced excursions in ACTH secretion are suppressed with an intensity that is directly proportional to the concentration of steroid previously reached in the blood. Late delayed feedback, which has a latency of approximately 24 h, develops only after a very substantial rise in steroid levels and is frequently the consequence of repeated or continuous administration of high doses of corticosteroids. It may persist for days or weeks after the steroid treatment is withdrawn and is thus largely responsible for the potentially life-threatening suppression of the HPA axis associated with abrupt cessation of long-term treatment with high doses of steroids in, for example, rheumatoid arthritis.

Mechanisms of glucocorticoid feedback The rapid feedback actions of the glucocorticoids are exerted primarily at the hypothalamic level; weaker actions also occur at the pituitary level, but the contribution, if any, of extrahypothalamic sites in the central nervous system is ill defined. The underlying molecular mechanisms are poorly understood. Although there are some reports to the contrary, the majority of data suggests that these effects are nongenomic (and hence do not involve MRs or GRs) and that they may be mediated by receptors that are located close to or within cell membranes and which, when activated, inhibit exocytosis, possibly by blocking Ca^{2+} . In contrast, the early and late delayed phases of feedback inhibition are mediated by classic intracellular corticosteroid receptors located primarily in the anterior pituitary gland, the hypothalamus, and the hippocampus, but also elsewhere in the brain. Both the high-affinity MRs and the lower-affinity GRs are involved. Current evidence suggests that MRs (located predominantly in the hippocampus) may maintain the tonic inhibitory influence of steroids on HPA function in nonstress conditions because these receptors, which are not protected by $11\beta\text{-HSD}$, are normally almost fully occupied by glucocorticoids, even at the nadir of the circadian rhythm. GRs, however (found in particular abundance in the anterior pituitary gland and hypothalamus) are occupied extensively only when the glucocorticoid level is raised by stress, disease, or the administration of exogenous steroids; thus, unlike MRs, these receptors serve to detect phasic changes in steroid levels and to adjust the activity of the HPA axis accordingly. Early delayed feedback requires the *de novo* generation of protein second messengers that serve to inhibit the release of ACTH and its hypothalamic-releasing factors; several proteins are likely to be involved, including lipocortin 1 (LC1, annexin 1), a Ca^{2+} - and phospholipids-binding protein strongly implicated in the anti-inflammatory and antiproliferative actions of the steroids. The genes encoding POMC in the pituitary gland and CRH/AVP in the parvocellular PVN neurons may also be repressed during the early delayed phase of feedback; however, the consequent downregulation of ACTH, CRH, and AVP synthesis emerges relatively slowly and, with existing techniques, is not normally discernible until some hours after the onset of inhibition of peptide release. Late delayed feedback is characterized by the inhibition of the synthesis and the release of CRH/AVP and ACTH in the hypothalamus and anterior pituitary gland, respectively. The inhibition of synthesis again reflects the inhibitory influence of the steroids on the genes encoding each of the three peptides, but

the concomitant inhibition of peptide release cannot be explained fully in this way because peptide stores are rarely fully depleted.

Short loop and ultrashort loop feedback mechanisms

In addition to the corticosteroids, other hormones of the HPA axis participate in the feedback regulation of ACTH secretion. Of particular importance are the POMC-derived peptides, ACTH and β -endorphin, both of which have been shown to exert powerful inhibitory actions on the secretion of CRH/AVP by the hypothalamus. Because neither peptide penetrates the blood-brain barrier readily, their short loop feedback actions are probably exerted at the level of the median eminence, which lies outside the blood-brain barrier. Ultrashort loop feedback actions have been attributed to both CRH and AVP, but the physiological significance of such responses is unclear.

Glucocorticoid Secretion in Repeated or Sustained Stress

The bulk of studies on stress and HPA function have focused on acute stresses, and relatively little is known of either the characteristics or the mechanisms controlling the HPA responses to repeated or sustained stress. Such conditions are, of course, inherent to our lifestyle and environment, and a deeper understanding of their influence on neuroendocrine function is thus highly desirable. However, for both ethical and practical reasons, it is difficult to develop appropriate animal models and, as described later, the available data suggest that the profile of the HPA responses observed may vary considerably according to the nature of the stress.

In some cases, HPA responses to repeated (successive or intermittent) or sustained (chronic) stressful stimuli are attenuated. For example, tolerance develops to stresses such as cold, handling, saline injection, or water deprivation (induced by the replacement of drinking water with physiological saline). Some degree of cross-tolerance may occur between stresses that invoke similar pathways/mechanisms to increase CRH/AVP release. Thus, for example, repeated handling reduces the subsequent adrenocortical response to a saline injection; however, rats tolerant to the C fiber-dependent stress of exposure to a cold environment respond to the novel stress immobilization with a normal or even exaggerated rise in serum corticosterone. In other situations, however, tolerance does not develop; thus, HPA responses to intermittent stresses of electric foot shock, insulin hypoglycemia, or $\text{IL-1}\beta$ are maintained or even enhanced, as are those to the chronic stress of septicemia.

The sustained release of ACTH in repeated/chronic stress appears to be driven mainly by AVP. Thus, in rats subjected to repetitive foot shock, brain surgery, or injection of IL-1 β or lipopolysaccharide (LPS), AVP is synthesized in and released from parvocellular neurons in the PVN in preference to CRH; in addition, the sensitivity of the pituitary gland to AVP is well maintained, whereas that to CRH is reduced. Similarly, the expression of AVP, but not CRH, in the parvocellular PVN is increased in rats with experimentally induced arthritis; moreover, although the arthritic rats respond to LPS with a rise ACTH secretion, they are relatively unresponsive to a novel neurogenic stress normally driven mainly by CRH. The switch to AVP as the principal driving force to the corticotrophs in repetitive/chronic stress may reflect the disruption of the negative feedback mechanism for the expression of MRs in the hippocampus, and GRs in the PVN are reduced in conditions of chronic stress; moreover, the capacity of dexamethasone to block the HPA response to IL-1 β is impaired in rats pretreated 3 weeks with the cytokine. In the same vein, AVP and disrupted feedback have both been implicated in the etiology of the hypercortisolemia evident in depression, although others have advocated a role for CRH.

Age- and Gender-Related Changes in the Stress Response

Development

The functional development of the HPA axis begins in fetal life, and in the later stages of gestation the fetus secretes substantial quantities of cortisol/corticosterone. It also responds to stress with increases in adrenocortical activity, which are driven by the hypothalamic-pituitary axis and which are sensitive to the negative feedback actions of the glucocorticoids. Paradoxically, in the rat and several other species, HPA function regresses postnatally and a period ensues during which the ability of the axis to respond to stress is compromised. In the rat, this stress hyporesponsive phase (SHRP) normally persists from postnatal day 5 (P5) to P14, although depending on the stimulus a mature stress response may not emerge for up to 5–6 weeks. The refractoriness of the HPA axis to stress in the SHRP is not absolute and may be at least partially overridden in certain cases. For example, IL-1 β , hypoglycemia, glutamate, and histamine all elicit modest rises in serum ACTH and corticosterone throughout this period. Exceptionally, the hypersecretion of ACTH provoked by LPS in the SHRP matches that observed in the adult, although the

associated rise in serum corticosterone does not; moreover, paradoxically, both the ACTH and the corticosterone responses to kainic acid are more robust at P12 than P18. These and many other data suggest that the processes involved in the maturation of the stress response are complex and stimulus-dependent. They are therefore likely to involve a variety of factors that act in concert at different levels of the axis during the developmental period. These include immaturity of the neural and humoral mechanisms effecting the release of CRH/AVP in stress, enhanced sensitivity of the axis to the regulatory actions of the glucocorticoids, and impairment of the synergy between CRH and AVP.

Perinatal Programming of the Stress Response

Glucocorticoids play a key role in developmental programming, and a growing body of evidence suggests that inappropriate exposure of the developing fetus/neonate to the steroids or to factors that influence the expression/function of their receptors may trigger a plethora of pathologies (e.g., growth retardation or impairment of central nervous system development) and predispose the individual in later life to conditions as diverse as hypertension, insulin resistance, cognitive and other behavioral disorders, and allergic/inflammatory disease. Because the SHRP coincides with an important phase of development, it seems likely that it may serve to protect the organism from the potentially harmful effects of overexposure to glucocorticoids.

Despite its tight control, the developing HPA axis exhibits a high degree of plasticity, and disturbances in its activity at critical stages of development precipitate abnormalities in glucocorticoid secretion that persist into adult life and may increase the susceptibility of the individual to disease. Perinatal manipulations that alter the expression of GRs (but not MRs) in the brain are particularly hazardous in this regard because, by disrupting the negative feedback mechanisms, they alter the responsiveness of the HPA axis to stress throughout the lifetime of the individual. For example, exposure of neonatal rats to dexamethasone or LPS results in the adult in downregulation of GRs in the hippocampus and hypothalamus, decreased sensitivity of the HPA axis to the negative feedback actions of the glucocorticoids, and exaggerated HPA responses to stress. Similar long-term changes in GR expression occur in rats subjected to prolonged periods of maternal separation; the consequent impairment of the negative feedback system and changes in HPA function are associated with alterations in T-cell antibody responses and in the incidence

of age-related neuropathies. In contrast, neonatal handling, which increases hippocampal GR expression by mechanisms involving serotonin, reduces the magnitude and duration of the stress response in the adult; maternal care may therefore help program events that regulate responses to stress in the adult and thus reduce vulnerability to disease.

Sexual Dimorphism in Hypothalamic-Pituitary-Adrenal Function

In adults, distinct sexually dimorphic patterns of glucocorticoid secretion emerge. Serum glucocorticoid concentrations are consistently higher in the female than in the male, with further increases occurring toward the middle of the menstrual/estrous cycle just prior to ovulation and in the late stages of pregnancy. These changes are due partly to the positive effects of estrogen on the expression of CBG. In addition, estrogen exerts significant effects at the hypothalamic level, increasing the synthesis and release of CRH, a phenomenon that may be implicated in the etiology of the relatively high incidence of emotional disorders (e.g., depression and anxiety) in the female that are characterized by enhanced CRH secretion. Further modulation may be brought about by progesterone, which, when present in large amounts (e.g., in the premenstrual phase or pregnancy), binds readily to but is only weakly active at MRs in the hippocampus.

Stress Responses in the Aging

In both sexes, serum glucocorticoid concentrations tend to increase in aging individuals, particularly in disease states; in addition, although sometimes delayed, the HPA responses to stress are frequently exaggerated and prolonged. This aging process has been attributed in part to the downregulation of the corticosteroid receptors (GRs and MRs) in the hippocampus and hypothalamus and to the consequent emergence of a resistance of the HPA axis to the negative feedback actions of the glucocorticoids. The consequences are potentially important because hyperglucocorticoidemia is considered to be an important contributory factor in a number of age-related diseases, including neurodegenerative conditions, cognitive dysfunction, osteoporosis, noninsulin-dependent diabetes mellitus, and the decline in immunocompetence.

See Also the Following Articles

Adrenal Cortex; Adrenocorticotrophic Hormone (ACTH); Corticotropin Releasing Factor (CRF); Feedback Systems; Glucocorticoid Negative Feedback; Glucocorticoids, Role in Stress; Hypothalamic-Pituitary-Adrenal; Vasopressin.

Further Reading

- Buckingham, J. C. (1996). Stress and the neuroendocrine immune axis: the pivotal role of glucocorticoids and lipocortin 1. *British Journal of Pharmacology* **118**, 1–19.
- Buckingham, J. C., Gillies, G. E. and Cowell, A.-M. (1997). *Stress, stress hormones and the immune system*. Chichester, UK: John Wiley.
- de Kloet, E. R., Vreugdenhil, E., Citzl, M. S., et al. (1988). Brain corticosteroid receptor balance in health and disease. *Endocrine Review* **19**, 269–301.
- Ericsson, A., Arias, C. and Sawchenko, P. E. (1997). Evidence for an intramedullary prostaglandin-dependent mechanism in the activation of stress-related neuroendocrine circuitry by intravenous interleukin-1. *Journal of Neuroscience* **17**, 7166–7179.
- Gillies, G. E., Linton, E. A. and Lowry, P. J. (1982). Corticotropin releasing activity of the new CRH is potentiated several times by vasopressin. *Nature* **299**, 355–357.
- Harris, G. W. (1955). *Neural control of the pituitary*. London: Edward Arnold.
- Kellerwood, M. E. and Dallman, M. F. (1984). Corticosteroid inhibition of ACTH secretion. *Endocrine Review* **5**, 1–25.
- Li, H. Y. and Sawchenko, P. E. (1998). Hypothalamic effector neurons and extended circuitries activated in “neurogenic” stress: a comparison of footshock effects exerted acutely, chronically and in animals with controlled glucose levels. *Journal of Comparative Neurobiology* **393**, 244–266.
- Lowry, P. J. (1984). Pro-opioidin: the multiple hormone precursor. *Bioscience Reports* **4**, 467–482.
- Muglia, L., Jacobson, L. and Majzoub, J. A. (1996). Production of corticotropin-releasing hormone deficient mice by targeted mutation of embryonic stem cells. *Annals of the New York Academy of Sciences* **780**, 49–59.
- Rosenfeld, P., Surchecki, A. J. and Levins, S. (1992). *Neuroscience and Biobehavioral Reviews* **16**, 553–568.
- Seckl, J. R. (1997). 11β -hydroxysteroid dehydrogenase in the brain: a novel regulator of glucocorticoid action? *Frontiers in Neuroendocrinology* **18**, 49–99.
- Smith, G. W., Aubry, J. M., Dellu, F., et al. (1992). Corticotropin releasing factor receptor 1-deficient mice display increased anxiety, impaired stress response and observant neuroendocrine development. *Neuron* **20**, 1098–1102.
- Tilders, F. J. and Solimidi, E. D. (1998). Interleukin 1β -induced plasticity of hypothalamic CRH neurons and long-term hyperresponsivity. *Annals of the New York Academy of Sciences* **840**, 65–73.
- Vale, W., Spiess, J., Rivier, L., et al. (1981). Characterisation of a 41-residue ovine hypothalamic peptide that stimulates secretion of corticotropin and β -endorphin. *Science* **213**, 1394–1397.
- Vamvakopoulos, N. C. and Chrousos, G. P. (1994). Hormonal regulation of human corticotropin releasing hormone gene expression: implications for the sexual dimorphism of the stress response and immune inflammatory reaction. *Endocrine Review* **15**, 409–420.

Glucocorticoids, Overview

B E Pearson Murphy

McGill University Health Center, Montreal, PQ, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by B E P Murphy, volume 2, pp 244–260, © 2000, Elsevier Inc.

Introduction

Evolution

Binding Proteins

Biosynthesis

Actions

Metabolism

Absorption

Regulation of Levels of Glucocorticoids

Measurement

Antiglucocorticoids

Uses of Glucocorticoids and their Inhibitors as
Therapeutic Agents

Glossary

Glucocorticoids

(From Greek *γλυκος* meaning sweet and Latin *cortex* meaning bark.) Steroids produced by the adrenal cortex which regulate carbohydrate metabolism; however their actions are by no means confined to carbohydrate metabolism.

Steroids

Chemicals containing the cyclopentanoperhydrophenanthrene nucleus ([Figure 1](#)).

Introduction

Chemistry

Structurally glucocorticoids (GCs) are all modifications of corticosterone ([Table 1](#)). Modifications which enhance activity include the addition of a 17 α -hydroxy group (cortisol, [Figure 2](#)), a Δ 1-double bond (prednisolone), a 9 α -fluoro (fludrocortisol), a 16 α - or 16 β -methyl or a 16 α -hydroxy group. The 11-keto analogs of these compounds are often included because, although they themselves are biologically inactive, they are readily converted to the active 11-hydroxylated forms. Cortisone, the 11-keto derivative of cortisol, was identified before cortisol, so that for many years cortisol was known as hydrocortisone (i.e., cortisone), with two additional hydrogens giving the 11-hydroxyl group. Oral cortisone and prednisone, the 11-keto analog of prednisolone, are widely

used therapeutically since they are about 80% converted to the active forms as they pass through the liver after absorption in the gastrointestinal tract.

History

In the late 1930s and early 1940s many steroids were isolated from the adrenal cortex of animals and man. Glands from 20 000 cattle were collected to give 1000 kg of tissue from which 25 g of material was further processed to determine more than 30 different steroids, including the GCs and mineralocorticoids.

Evolution

Steroids are present throughout the whole of the biosphere of the Earth, and are thought to be ancient bioregulators, present throughout the plant and animal kingdoms, which evolved prior to the appearance of eukaryotes. They seem to be primordial biomolecules which are among the earliest chemical information transmitters and have had a crucial role in the evolution of life. Even bacteria, viruses, and blue-green algae contain steroids.

GCs are secreted in all vertebrates, but not in other life forms. In mammals and bony fish, the main GC is cortisol, except in some rodents where it is corticosterone. In birds, reptiles, amphibians, and cartilaginous fish, the main GC is corticosterone.

Binding Proteins

There are two types of binding proteins for steroids: one type which inactivates them, and the other which activates them.

Transport Proteins

Corticosteroid-binding globulin (CBG), present in blood, binds GCs noncovalently in such a way that their biological activity is suppressed; thus it is the

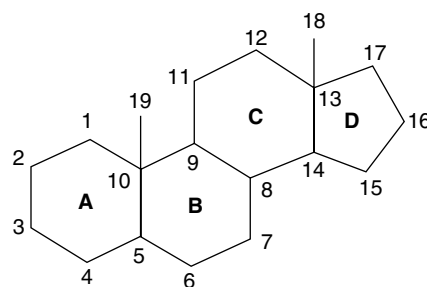


Figure 1 Cyclopentanoperhydrophenanthrene nucleus.

Table 1 Modifications of corticosterone to give other glucocorticoids, and their relative potencies in humans

Steroid	Modification of corticosterone	Anti-inflammatory potency	Salt-retaining potency	LHPA-axis suppression
Corticosterone	—	0.35	15	
Cortisol	17 α -OH	1	1	1
Prednisolone	Δ^1 -, 17 α -OH	4	0.75	4
Methylprednisolone	6 α -methyl-, Δ^1 -, 17 α -OH	6	0.5	4
Fludrocortisol	9 α -fluoro-, 17 α -OH	15	280	
Dexamethasone	16 α -methyl-, Δ^1 -, 9 α -fluoro-, 17 α -OH	26	<1	17
Betamethasone	16 β -methyl-, Δ^1 -, 9 α -fluoro-, 17 α -OH	25	<1	
Triamcinolone	16 α -hydroxyl-, Δ^1 -, 9 α -fluoro-, 17 α -OH	5	<1	4

LHPA, limbic-hypothalamic-pituitary-adrenal.

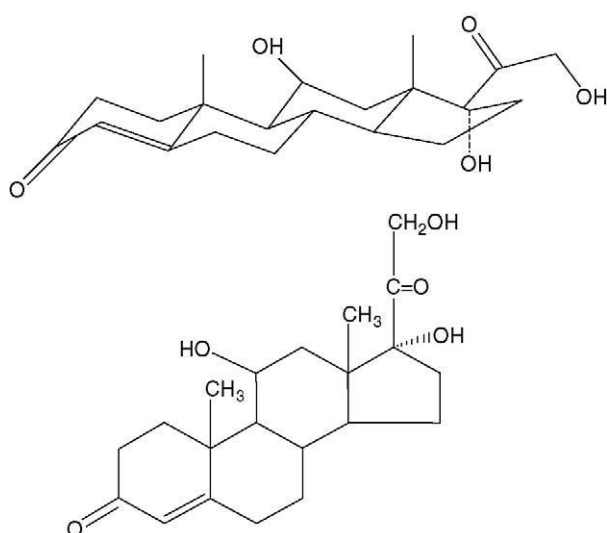


Figure 2 Structure of cortisol. Top: The three-dimensional structure. A dotted line denotes that the C17 OH is below the plane of the molecule, while a straight line, as for the C18- or C19-hydroxyl denotes that the CH₃ is above the plane. Asterisks indicate asymmetric carbon atoms. Bottom: Depicted in the customary fashion and showing numbering of the carbon atoms.

unbound or free hormone which is active. These proteins are found in virtually all vertebrates and their specificities for GCs vary; in some cases the binding is very highly specific and can be used to measure the steroid that is bound. They are saturable so that at high levels of GCs, proportionally less is bound. Such proteins are important in the transport of GCs since they prevent undue transfer of the steroid into tissues where it would be subject to enzymatic action and/or degradation. They thus provide a reservoir to damp down rapid changes. In conditions of severe stress, they are reduced. GCs are also bound reversibly and noncovalently to albumin, but this is a nonspecific and loose type of binding.

Glucocorticoid Receptors

Another kind of protein that binds GCs is the GC receptor (GR), which mediates the action of the

steroid and is found in or on the target cells for the hormones. This binding is of high affinity and specificity and is saturable. In the nucleus, the hormone-receptor complex is activated and binds to deoxyribonucleic acid (DNA), leading to transcription.

The actions of GCs are mediated by two types of receptors: GRs and mineralocorticoid receptors (MRs). GRs bind GCs but not aldosterone, while MRs bind both GCs and aldosterone, the selectivity of the MRs being dependent upon the presence of 11 β -dehydrogenase (particularly in the kidney) to inactivate the GCs. In the hippocampus where MRs are present in high concentrations, GCs serve to activate both.

Biosynthesis

The synthetic pathway appears to be much the same throughout the vertebrates: acetate $\rightarrow$ cholesterol $\rightarrow$ pregnenolone $\rightarrow$ corticosterone and/or cortisol (Figure 3). Each step is catalyzed by an enzyme. Some of these enzymes are known to occur also in plants and bacteria.

Steroidogenic tissues can take up low-density lipoprotein cholesterol from the plasma (about 80%), mobilize intracellular cholesterol ester pools, or synthesize cholesterol *de novo* from acetate. GCs are synthesized from cholesterol by the P450 group of enzymes (oxidative enzymes that have a characteristic 450-nm absorbance maximum when reduced with carbon monoxide). The first step is achieved by P450_{SCC}, the side-chain cleavage (SCC) enzyme, which cleaves the side chain from the C20 to give pregnenolone. This occurs in the inner mitochondrial membrane. 21-Hydroxylation occurs in the smooth endoplasmic reticulum while 11-hydroxylation occurs in the mitochondrion. In humans, the route to corticosterone and aldosterone production takes place in the adrenal zona glomerulosa while that to cortisol occurs in the zona fasciculata and the reticularis; the 11-hydroxylating enzymes involved differ slightly in the two regions, while the P450_{C17}-hydroxylase is

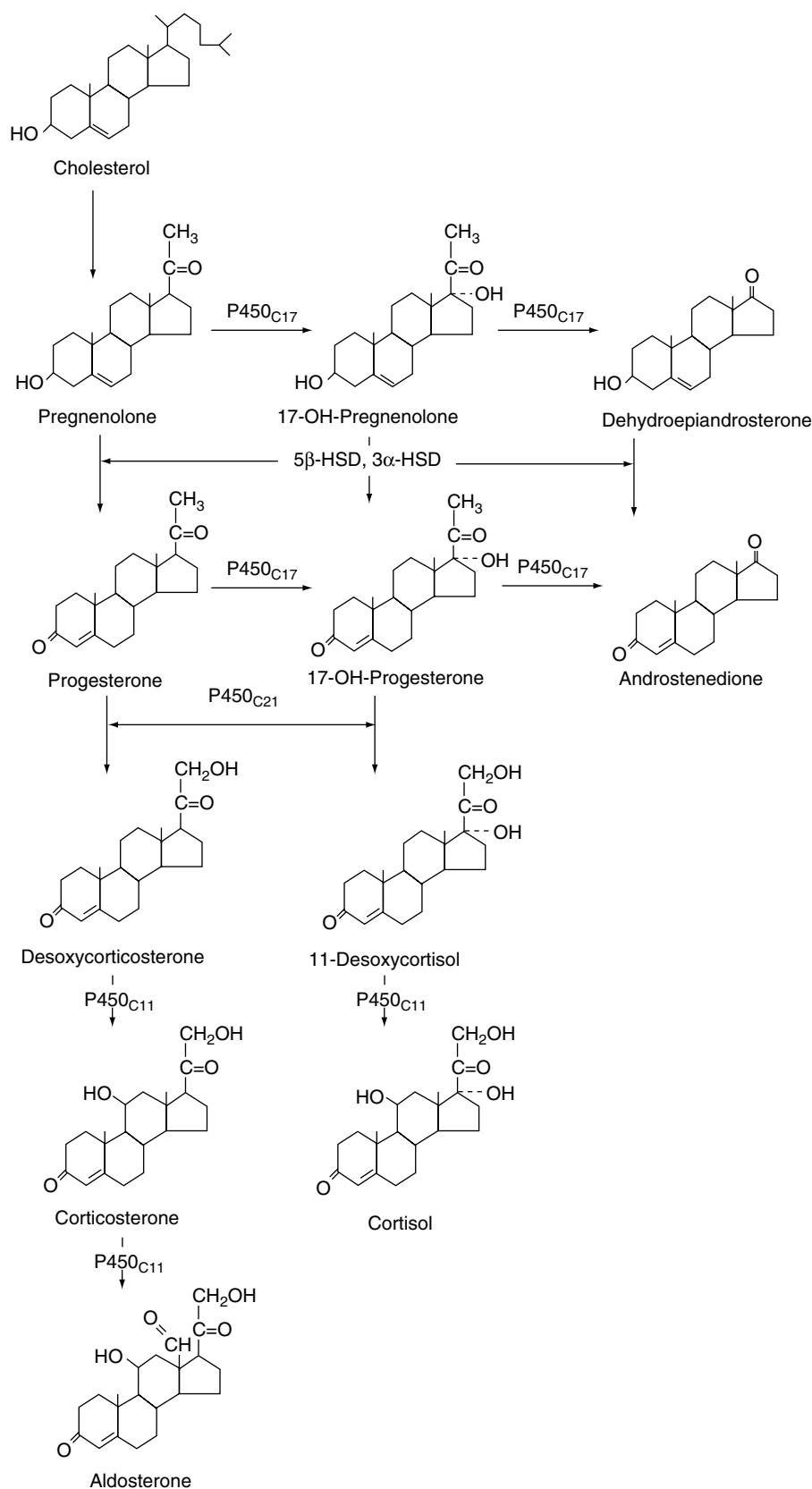


Figure 3 Biosynthesis of adrenal corticosteroids. 3 α -HSD, 3 α -hydroxysteroid dehydrogenase; 5 β -HSD, 5 β -hydroxysteroid dehydrogenase.

lacking in the zona glomerulosa. The P450_{C17} enzyme also cleaves the C20–C21 from C17, giving rise to dehydroepiandrosterone (DHEA) and the adrenal androgen pathway. If the GC pathway is blocked, as in congenital adrenal hyperplasia, this route becomes prominent. DHEA is sulfated in the reticularis to give DHEA sulfate (DHAS), which is quantitatively, though not physiologically, the most important product of the adult adrenal cortex.

In fetal primates, the adrenal is relatively much larger than in the adult, and is made up mainly of a fetal zone which is lost at about the time of birth. This zone lacks 3 β -hydroxysteroid dehydrogenase isomerase so that the major steroid produced by this zone is DHEA, mainly in the sulfated form. The function of this zone remains speculative.

Actions

GCs are necessary for life and affect every cell of the body. They are thought to enter the cell mainly by passive diffusion, and then bind to the GR. They also act nongenomically at membrane receptors.

Carbohydrate Metabolism

Adrenalectomized animals become hypoglycemic because they cannot maintain their hepatic glycogen stores. GCs activate glycogen synthase and inactivate the phosphorylase which mobilizes glycogen. Although glucagon and epinephrine are also gluconeogenic, they are ineffective in the absence of GCs which act permissively.

GCs enhance gluconeogenesis from protein, thereby raising blood glucose and repleting liver glycogen. This is accomplished by increasing release of glucogenic amino acids from peripheral tissues such as skeletal muscle. GCs also activate gluconeogenic enzymes, including glucose-6-phosphatase and phosphoenolpyruvate carboxykinase. Large amounts result in negative nitrogen balance, and if prolonged, depletion of body protein, observed clinically as thinning of the skin, muscle wasting, myopathy, osteoporosis, and in children, retarded growth.

Fat Metabolism

GCs, administered acutely, activate lipolysis in adipose tissue, so that plasma free fatty acid levels are reduced. However, when given chronically, increased availability of carbohydrate produces an antiketogenic effect, and clinically a redistribution of fat from the extremities to the face and trunk, especially in the supraclavicular and dorsocervical areas, the latter producing the characteristic buffalo hump seen in Cushing's syndrome.

Electrolyte and Water Metabolism

Although much less potent than the mineralocorticoids, GCs in high doses may produce sodium and water retention, hypokalemia, and hypertension. Conversely, in their absence, there is renal sodium wasting, with a concomitant loss of water, hemoconcentration, hyponatremia, and hypovolemia. Hypotension may occur, although this is usually avoided if sufficient mineralocorticoid is present. Vasopressin (VP) is increased and free water clearance is decreased in GC insufficiency, which may help to maintain normal blood pressure. GCs stimulate plasma atrial natriuretic peptide (ANP) production, and potentiate ANP action on the kidney.

Immunologic Function and Inflammation

GCs alter immune responses and inhibit inflammatory reactions. These actions can have both beneficial and deleterious results. While GCs are essential to maintain life in the face of stressors, excessive GCs may cause reactivation of latent infections such as tuberculosis. In contrast, they may benefit autoimmune diseases. A local inflammatory response involves the movement of cells and fluid to the site of inflammation. Histamine, a potent vasodilator, is released at the site, and prostaglandins are involved. Within 4 h after a rise in GCs in response to stress, peripheral lymphocyte and monocyte numbers fall, but granulocytes enter the circulation from lymphatic tissues and sites of inflammation. GCs inhibit synthesis of cytokines, proliferation of monocytes and their differentiation into macrophages; they also inhibit the phagocytic and cytotoxic functions of macrophages.

Effects on Bone and Connective Tissue

If GC excess is prolonged, osteopenia occurs because GCs inhibit osteoblast function – including down-regulation of insulin-like growth factor I (IGF-I) which promotes osteoblast proliferation – thereby decreasing new bone formation. Osteoclast numbers and activity are increased. Intestinal calcium absorption is decreased and parathyroid hormone levels are secondarily increased; these effects can be reversed by giving vitamin D and calcium. Renal calcium reabsorption is decreased resulting in an increase in calcium excretion, and which may cause the formation of renal stones.

Since GCs also suppress synthesis of the extracellular matrix components collagen and hyaluronidate, clinically excessive GCs can impair wound healing and make connective tissue more friable, resulting in easy bruisability.

Oral administration of GCs may be associated with symptoms of gastric or duodenal ulcer; possibly this is

only apparent, as ulcers caused by other factors, and which would heal if left alone, may become clinically troublesome if wound healing is impaired.

Neuropsychiatric Effects

GCs are released during learning and are necessary for establishment of an enduring memory. However excessive amounts impair spatial learning. Such processes involve the hippocampus.

Cortisol increases brain excitability, and emotional disturbances and even psychoses may occur in states either of deficiency or excess. In states of GC excess – either endogenous as in Cushing's syndrome or exogenous from therapeutic administration – altered behavior has been frequently noted.

Interestingly, in adrenal insufficiency, depression is also common, usually in association with apathy and lethargy; it improves with GC replacement.

The actions of GCs in the brain are mediated both genomically, whereby they enhance transcription, and nongenomically by direct effects on the excitability of neuronal networks, especially in the hippocampus, where GRs are present in the highest density. They can affect many neurotransmitter systems, including noradrenergic, serotonergic, cholinergic, dopamine, and GABAergic activity.

Development

Excessive concentrations of GCs, either endogenous or exogenous, inhibit skeletal growth in children. In the human fetus, the concentrations of circulating and tissue cortisol are kept low by the action of 11 β -hydroxysteroid dehydrogenase type 2, which converts cortisol to inactive cortisone in almost all tissues until close to birth.

However GCs are important in stimulating differentiation of many cell types. In lung, they induce surfactant production, which is necessary for infant mammals to adapt to air-breathing at birth. GCs cause differentiation of the neural crest precursor cells of the adrenal medulla to differentiate into chromaffin cells which produce catecholamines.

Metabolism

Steps Involved

The GCs are metabolized primarily in the liver, and are excreted mainly through the kidneys in a series of steps catalyzed by enzymes ([Figure 4](#) and [Table 2](#)). These include:

1. Oxidation of the C11-hydroxyl to a ketone – the only reaction which is readily reversible

2. Reduction of the double bond at C4–5 to give 5 α - and, mainly 5 β -dihydro derivatives – these are almost entirely further reduced at the 3-ketone
3. Reduction of the 3-ketone to give 3 α (mainly) and 3 β -tetrahydro derivatives, which are in part further reduced at the C20-ketone
4. Reduction of the C20-ketone to give 20 α - and (mainly) 20 β -derivatives of any of the above compounds, but mainly of the tetrahydro derivatives to give the cortols (retaining the 11 β -hydroxyl group) and cortolones (with the 11-ketone group)
5. Cleavage of the side chain of tetrahydrocortisol to give etiocholanolone
6. Oxidation of the 21-OH of the cortols to give the cortolic acids, and of the cortolones to give the cortolonic acids
7. Conjugation of the various metabolites with glucuronic acid (mainly at the 3 β -hydroxyl) or sulfate (mainly 3 α -hydroxyl), rendering the steroids more water-soluble
8. 6 β -Hydroxylation of cortisol to give 6 β -hydroxycortisol, also a much more water-soluble steroid

Factors Affecting Hepatic Metabolism

The following factors affect hepatic metabolism.

1. Thyroid hormones influence the rates of metabolism. In hyperthyroidism, metabolism of GCs is accelerated, while the opposite occurs in hypothyroidism. While these changes affect urinary excretion of metabolites, plasma corticoid levels are maintained, owing to the feedback mechanism which remains unaltered
2. In states of GC excess, hydroxylation to 6 β -hydroxycortisol increases
3. In old age, corticoid metabolism decreases to some extent, while clearance of glucuronides is decreased; plasma levels usually rise slightly
4. In cirrhosis of the liver, 5-reduction is decreased
5. In obesity, greater amounts of metabolites are excreted, but the plasma levels remain essentially unchanged
6. Various drugs affect metabolism, e.g., rifampin accelerates metabolism, particularly to 6 β -hydroxycortisol. Phenytoin, phenobarbital, and mitotane all divert metabolism toward 6 β -hydroxycortisol

Extrahepatic Metabolism of Glucocorticoids

The kidney is another important site of metabolism, since its 11 β -hydroxysteroid dehydrogenase (type 2) converts cortisol to cortisone, thereby limiting access of cortisol to the MR which binds both aldosterone

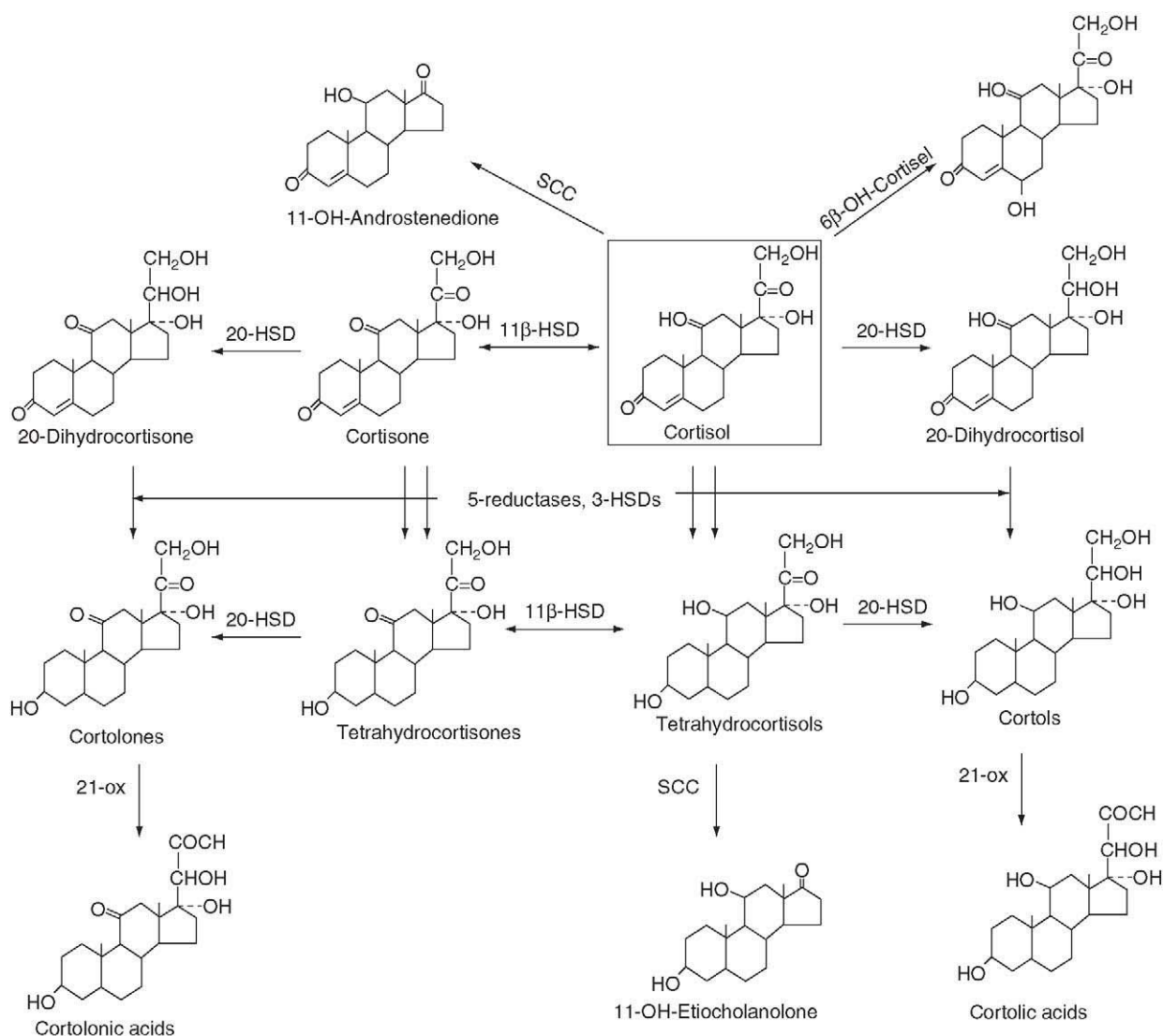


Figure 4 Metabolism of cortisol. HSD, hydroxysteroid dehydrogenase; ox, oxidase; SCC, side-chain cleavage.

Table 2 Cortisol and its metabolites in human urine^a

Steroid	Approximate % of total
Tetrahydrocortisols (5β > 5α; 3α > 3β)	20
Tetrahydrocortisones (5β > 5α; 3α > 3β)	20
Cortolones	20
Cortols	10
Cortolic and cortolonic acids	10
11-OH-etiocholanolone	5
6β-OH-cortisol	1
11-OH-androstenedione	1
Cortisone, conjugated	1
Cortisol, conjugated	1
Cortisone, unconjugated	0.3
Cortisol, unconjugated	0.1

^aWith the possible exception of 6β-hydroxycortisol, all the metabolites occur mainly as conjugates, glucuronides exceeding sulfates.

and cortisol equally. Since aldosterone is present in much smaller amounts than cortisol, this permits aldosterone to exert its physiological action on the receptor. Genetic or pharmacologic alterations in the enzyme may cause mineralocorticoid problems.

Absorption

While steroids are generally administered via the oral route, they can be given effectively in many other ways including percutaneous, vaginal, rectal, nasal, and pulmonary, as well as intravenously, subcutaneously, and intramuscularly.

Intravenously administered steroids are diluted throughout the whole blood volume within a few minutes; subcutaneously injected steroids are absorbed within minutes to a few hours; intramuscularly

injected steroids are more slowly absorbed, particularly if diluted in a medium which slows absorption, in which case absorption may be delayed for days to weeks. Absorption through skin (especially damaged skin) is rapid – usually detectable 30–100 min after application, and maximal by 16–24 h.

Regulation of Levels of Glucocorticoids

The Limbic-Hypothalamic-Pituitary-Adrenal Axis

Stimuli reaching the brain from extracerebral sites impact on the limbic system, where they are relayed to the hypothalamus, causing stimulation of corticotropin-releasing hormone (CRH) and VP. CRH and VP pass via the portal blood to the pituitary, causing pro-opio-melanocortin to be processed into adrenocorticotrophic hormone (ACTH), endorphin, and other peptides, which enter the systemic circulation. ACTH stimulates the adrenal to produce cortisol and/or corticosterone which in turn feed back on GRs in the limbic system (principally the hippocampus and amygdala), hypothalamus, and pituitary to quench the stimulation, thus maintaining an ever-changing but essentially constant equilibrium (Figure 5).

Adrenocorticotropin ACTH acts by increasing synthesis of GCs, since adrenal storage of corticoids is minimal. ACTH stimulates the processing of cholesterol to Δ^5 -pregnenolone, the rate-limiting step

in synthesis. It does this by binding to specific cell-surface receptors present on the adrenal cells, which activates adenylate cyclase thereby increasing cyclic adenosine monophosphate (cAMP), in turn activating cAMP-dependent protein kinase and phosphorylation of a number of proteins. If stimulation continues for more than a few minutes, increased synthesis of the various enzymes of the steroidogenic pathway are stimulated, as well as cell growth; if prolonged, the adrenal hypertrophies. Conversely, if ACTH levels fall for a prolonged period of weeks or months, the adrenal atrophies.

Other possible stimulators include direct effects of cytokines on the adrenal cortex and stimulation through neural connections.

Dexamethasone suppression test The dexamethasone-suppression test is often used to assess the suppressibility of the limbic-hypothalamic-pituitary-adrenal (LHPA) axis in humans; most commonly a 1 mg dose is given at bedtime, and blood samples are drawn once or more during the next day. The test is considered normal if the levels fall below 140 nmol l^{-1} ($5 \text{ } \mu\text{g dl}^{-1}$), or preferably below 60 nmol l^{-1} ($2.2 \text{ } \mu\text{g dl}^{-1}$).

Insulin tolerance test The insulin tolerance test is commonly used to evaluate the integrity of the LHPA axis. It measures the response to hypoglycemia induced by giving an injection of insulin, usually 1 μg ,

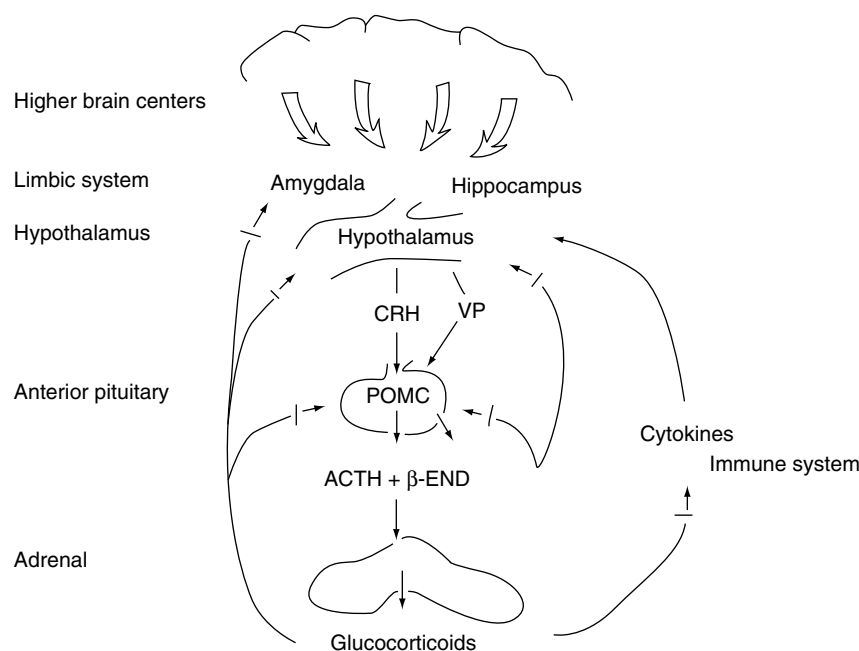


Figure 5 The limbic-hypothalamic-pituitary-adrenal (LHPA) axis. CRH, corticotropin-releasing hormone; VP, vasopressin; POMC, pro-opiomelanocortin; ACTH, adrenocorticotrophic hormone; β -END, β -endorphin.

and following glucose and GC levels at 30, 60, and 90 min after injection. With normal CBG levels, the cortisol level should reach at least 500 nmol l^{-1} ($18 \text{ } \mu\text{g dl}^{-1}$).

Plasma Levels of Glucocorticoids under Various Conditions

Plasma levels of GCs are generally in the range of 10–1000 nM. They vary diurnally, being highest shortly after awakening, and gradually falling during the period of activity (e.g., during the night in rodents, during the day in humans and monkeys). Plasma levels among species vary depending on the affinity and level of their CBG in plasma and of the GRs in the hippocampus, hypothalamus, and pituitary. GCs are also present in red blood cells in appreciable amounts.

In humans in the unstressed state, the plasma level of cortisol is $50\text{--}600 \text{ nmol l}^{-1}$ during the day, being highest at 6–7 a.m. and lowest at midnight; corticosterone levels are about 10% of these and cortisone levels about 20%. Levels of cortisol in squirrel monkeys are higher at about 3000 nmol l^{-1} although their CBG levels are virtually undetectable. In rats, cortisol is absent, and corticosterone ranges at $50\text{--}200 \text{ nmol l}^{-1}$. Levels of corticosterone in fish (e.g., $15\text{--}200 \text{ nmol l}^{-1}$ in trout) and reptiles are somewhat lower. In salmonids, and many animals, cortisol levels change seasonally, often in conjunction with CBG levels. In stressed animals, the levels of GCs rise within minutes to levels up to 10 or more times the resting levels.

Levels of Glucocorticoids in Other Tissues

GC levels in brain are as high or higher than those of plasma because brain is a fatty tissue and GCs are very fat soluble. GCs are found in all tissues but levels are usually lower than those in plasma and lymph. Levels are low in cerebrospinal fluid because the CBG level is low. They are very high in the adrenal cortex, the site of synthesis.

Cortisol–Cortisone Interconversion

Although in adult humans the level of cortisol greatly exceeds that of cortisone, the reverse is true in fetuses where oxidation by 11β -hydroxydehydrogenase (11β -HSD) greatly exceeds reduction. 11β -HSD consists of two enzymes: type 1 which predominates in adult liver and converts cortisone to cortisol; and type 2 which predominates in adult kidneys, allowing aldosterone to act in the presence of a 100-fold molar excess of circulating cortisol by converting cortisol

to cortisone. This type is also high in skin, salivary glands, and nonpregnant uterus.

In the fetus, type 2 predominates in all tissues except liver and intestine. It is particularly high in the placenta, converting maternal cortisol to cortisone, and corticosterone to 11-dehydrocorticosterone. This conversion, which inactivates cortisol and corticosterone, protects the fetus from excessive GCs which would interfere with growth. Type 2 11β -HSD is inhibited by licorice. The same enzyme converts prednisolone to prednisone but does not act on dexamethasone.

The human uterus, though oxidative in the non-pregnant state, becomes strongly reductive in late pregnancy, increasing the cortisol level eightfold so that the membranes are in contact with an environment high in cortisol. This increased cortisol concentration may play an anti-immune role in the uterine wall, the single maternal tissue, apart from maternal blood, in direct contact with fetal tissue.

A similar relationship between corticosterone and its 11-keto analog has also been demonstrated in rodents.

Conditions of Glucocorticoid Excess

Cushing's syndrome In human Cushing's syndrome due to pituitary ACTH-secreting tumors (Cushing's disease), the morning cortisol levels are usually only minimally elevated, but the normal diurnal rhythm is lost so that the 24-h urinary cortisol, which reflects the integrated unbound cortisol level, is elevated. However the suppression of ACTH production by dexamethasone, although attenuated, is usually maintained.

In cases of ectopic ACTH-producing tumors, generally occurring in lung, cortisol levels are usually higher, and are frequently unsuppressed by dexamethasone.

Adrenal tumors cause varying levels of GCs, which are usually unresponsive to dexamethasone suppression, and are often associated with increases of adrenal androgens and/or mineralocorticoids.

Depression is common in Cushing's syndrome, and other behavioral disturbances may occur. In some cases the psychiatric features dominate over the physical features, leading to misdiagnosis. If the condition is successfully treated, either surgically or medically, the psychiatric manifestations usually subside.

Acute stress When an acute stress (threat to homeostasis) is perceived by the brain, impulses to the hypothalamus stimulate the production of CRH and VP, which, through ACTH, stimulate the production of GCs. If the stimulus is prolonged, the axis usually

adapts to it over several days, so that ACTH and GCs return to normal.

In rats, it has been shown that the response of GCs to stress in later life can be modified by neonatal handling, which appears to increase the efficiency of the stress response, and to protect the animal from potentially damaging effects of GCs.

Chronic stress Should severe stress be continued for long periods, neurodegeneration or suppressed neurogenesis in the hippocampus occurs. Similar changes also occur after 3 weeks of treatment with a high dose of GCs.

Major depression Increased levels of GCs in major depression (seen in 60–70%) are the best-documented biochemical abnormality in psychiatric disease. There is also a decrease in the diurnal variation and, in about 50%, a resistance to suppression of cortisol by dexamethasone. Major depression is twice as common in women as in men. Although CRH levels are increased, ACTH levels are not increased, so that there appears to be an alteration in adrenal responsiveness which is not understood. These changes are reversed on remission.

Conditions of Glucocorticoid Deficiency

Adrenal insufficiency (Addison's disease) In adrenal insufficiency, first described by Addison in 1855, levels of GCs are low normal or low, and respond poorly or not at all to ACTH. Insufficiency may occur due to destruction of the adrenal cortex (primary), to deficient ACTH secretion (secondary), or to deficient hypothalamic secretion of ACTH secretagogues (tertiary).

Congenital adrenal hyperplasia (genetic defects in cortisol synthesis) Inherited enzymatic defects of GCs synthesis may become manifest in fetal life, the most common being due to lack of the various enzymes involved. In order to maintain adequate GCs levels, ACTH levels are increased to overcome the block, and hypertrophy and hyperplasia of the adrenal cortex occur. Also, due to increased ACTH, synthetic steps which are not blocked produce increased quantities of their products, particularly of precursors of GCs and of adrenal androgens. While males are usually asymptomatic, the excess androgens usually cause ambiguous genitalia in female fetuses. When initially recognized, this syndrome of masculinization of female fetuses associated with enlarged adrenal glands was called the adrenogenital syndrome.

21-Hydroxylase deficiency is the most common defect accounting for 95% of all cases of congenital

adrenal hyperplasia and is particularly prevalent among the Yupik Eskimos, where the incidence is 1 in 500 births and the heterozygote frequency is 1 in 11. Milder cases of 21-hydroxylase deficiency are masculinized while more severely affected infants have the salt-losing form, causing hyponatremia, hyperkalemia, acidosis, and dehydration, sometimes leading to vascular collapse. Very mild cases may not become manifest until puberty, when hirsutism becomes prominent. If suspected, this defect can be diagnosed prenatally and treated with small doses of dexamethasone, which crosses the placental barrier (unlike cortisol, which is converted to cortisone in the placenta).

Deficiency of the following enzymes may also occur but is rare: cholesterol desmolase, 3 β -hydroxysteroid dehydrogenase/isomerase, 17-hydroxylase, and 11 β -hydroxylase.

Measurement

Competitive Protein-Binding Assays

GCs are currently usually measured by competitive protein-binding (CPB, displacement analysis). In this type of assay, the steroid is measured according to its ability to displace a tracer form of itself from a protein which binds it with high affinity and specificity. Such proteins include endogenous plasma proteins (CBGs), endogenous GR proteins, enzymes, or most commonly, antibodies raised to GCs modified to make them antigenic. The sensitivity of such methods is very great (usually a few picograms) and depends on the affinity of the steroid for the protein (usually of the order of 10^9 M⁻¹ or greater); thus for plasma only a few microliters of sample is required for the determination of cortisol or corticosterone concentrations. The specificity varies with the binding protein used, with the method of separating bound and free moieties, and with the steroid milieu in which the GCs are being measured. Since no protein is absolutely specific for any single ligand, each method must be carefully validated for the circumstances under which it is being used. The sample to be measured must be freed of any binding proteins that may interfere.

Assays Using Corticosteroid-Binding Globulin

Assays employing CBG were the first such assays to be used for GCs, and were introduced in 1963. Human CBG was the first to be used, but is less specific for cortisol than are dog or horse CBG. Each CBG has its own spectrum of specificity and can be chosen according to the milieu in which the GC is to be measured. These assays have the advantages that the binding properties of the CBG are essentially invariant, so

that it is not necessary to determine the specificity or affinity of each batch of protein; also, no preparation beyond dilution is required.

Radioreceptorassays and Radioenzymaticassays

Receptor proteins can be used to measure their ligands, and enzymes can be used to measure their substrates in similar fashion, but are not usually used for GCs.

Radioimmunoassays

Immunoassays have the important advantage that a suitable protein can be raised to virtually any steroid, and since each antibody is slightly different, the most suitable one for the milieu in which the GC is to be measured can be chosen. A disadvantage is that each batch, even from the same animal, may vary with respect to affinity and specificity and must be characterized separately to ensure consistent results. Initially it was thought that the use of monoclonal antibodies might be advantageous but the sensitivity attained is usually less than with polyclonal antibodies.

Although radioactive isotopes have been the most common labels used (hence the term radioimmunoassay), other labels are now available (see below).

Fluorescence Immunoassays

Fluorescence immunoassays employ either a fluorophore label, or utilize enzyme-immunoassays in which substrates are converted to fluorescent endproducts.

Chemiluminescence Immunoassays

Chemiluminescence immunoassays use luminescent compounds which emit light during the course of a chemical reaction, usually enzymatic. Such assays have achieved the highest levels of sensitivity to date.

Determinations in Various Body Fluids

Plasma It is important to suit the method to the purpose. For example, a method which may be adequate for the measurement of cortisol in adult human blood may be, and usually is, entirely unsuitable for its measurement in neonatal or cord blood, which contain many steroids in high concentrations (e.g., progesterone and 17-hydroxyprogesterone) that are not present or are present in only minimal amounts in nonpregnant adult blood.

Urine Similarly, such a method is rarely, if ever, adequate to measure cortisol in urine, where there are very large amounts of competing steroids, so that extra steps such as extraction and chromatography must be performed first. Although the true mean value in human urine was established in 1980 to be

about 60 nmol day⁻¹, published values continue to exceed these by a factor of 2–10 throughout the literature, so must be carefully scrutinized.

Saliva In saliva, which is convenient for the taking of samples, and more closely approximates free GCs, it must be taken into consideration that the level of cortisone (at least in humans) may exceed that of cortisol, so may interfere even with an antibody that ostensibly binds cortisone poorly. While saliva does not contain CBG, the salivary glands do contain 11 β -hydroxysteroid hydroxylase type 2, which converts about half the cortisol reaching them to cortisone, so that salivary cortisol levels are about 50% of those of unbound plasma cortisol.

Indirect Methods

Prior to the development of CPB assays, GCs were measured by indirect methods. All of these are much less specific and less sensitive than CPB methods. They include fluorescence methods (steroids with an 11-hydroxyl group fluoresce under ultraviolet light in alkali or acid), the measurement of 17-hydroxycorticoids (Porter-Silber chromogens) as an approximation of GCs, the use of tetrazolium salts to measure compounds containing an α -ketol group, 17-ketogenic steroids minus 17-ketosteroids as an approximation of GCs, ultraviolet absorption after chromatography, and double isotope dilution derivative assay.

Antiglucocorticoids

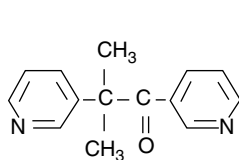
Inhibitors of steroidogenesis

A selection of inhibitors of steroidogenesis is given in [Figure 6](#).

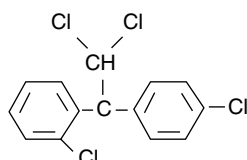
Metyrapone mainly inhibits P450_{C11}, preventing 11-desoxycortisol from being hydroxylated at C11 to give cortisol, and 11-desoxycorticosterone from being converted to corticosterone, thereby lowering cortisol, corticosterone, and aldosterone and resulting in excessive amounts of the precursors due to stimulation of ACTH secretion.

Mitotane, which is related to the insecticide DDT (4,4'-(2,2,2-trichloroethane-1,1-diyl)bis(chlorobenzene)), is toxic to adrenal tissue, causing necrosis and hemorrhage. It also affects peripheral metabolism of cortisol, greatly increasing its conversion to 6 β -hydroxycortisol.

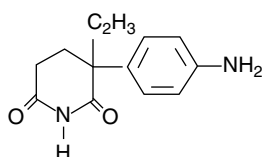
Aminoglutethimide affects GC synthesis in several ways: it inhibits P450_{sc}, thus decreasing the conversion of cholesterol to pregnenolone and the production of all the adrenal steroids. Although it may be useful in decreasing hypercortisolism, the blockade



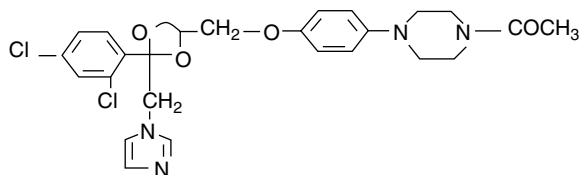
Metyrapone



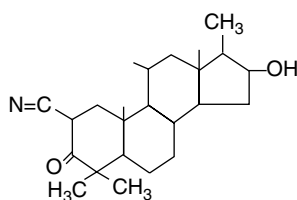
Mitotane (o,p'-DDD)



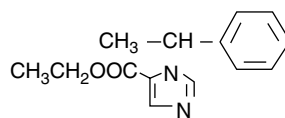
Aminoglutethimide



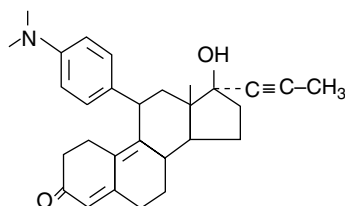
Ketoconazole



Cyanoketone



Etomidate



Mifepristone (RU 486)

Figure 6 Structures of selected antiglucocorticoids.

is often overridden by increased secretion of ACTH in response to the lowered cortisol levels.

Cyanoketone inhibits 3β -hydroxysteroid dehydrogenase but it is not currently used clinically.

Ketoconazole is an antifungal agent which inhibits $P450_{\text{scc}}$ and $P450_{\text{C11}}$.

Etomidate is an anesthetic agent and its main effect is on $P450_{\text{C11}}$, although it also affects $P450_{\text{scc}}$.

Inhibitor of Glucocorticoid Receptor

Mifepristone (RU 486) inhibits GCs production by binding to the GR, and has been used successfully in some cases of Cushing's syndrome. As with agents that directly interfere with synthesis, the blockade

caused by this receptor inhibitor tends to be overcome by increased production of CRH and ACTH.

Uses of Glucocorticoids and their Inhibitors as Therapeutic Agents

Therapeutic Uses of Glucocorticoids

The first therapeutic use of GCs was that of cortisone by PS Hench and his coworkers in 1949 to treat rheumatoid arthritis. Since then, a great many disorders have been treated using a variety of GCs, including allergic diseases such as asthma and dermatitis, autoimmune diseases such as systemic lupus

erythematosis and Graves' ophthalmopathy, and inflammatory disorders such as ulcerative colitis. They have also been used to decrease cerebral edema associated with intracerebral masses, to treat septic shock, and to suppress graft rejection after transplantation of organs.

When GCs are given therapeutically, a feeling of euphoria is commonly experienced, although prolonged use often results in depression, and less often mania, especially if higher doses are used. There is no typical presentation as a variety of mental symptoms is observed, including often irritability, insomnia, obsessive-compulsive behavior, and anxiety. Altered perception, sensory flooding, delusions and hallucinations, delirium, and even frank dementia may occur in about 5% of cases. The time required for mental symptoms to occur on treatment varies from almost immediate to several months, with a median onset of 11 days. These effects are usually largely reversible after reduction of dose or discontinuation of the steroid.

Cortisone itself is inactive but it is rapidly converted to cortisol in the liver. Although the natural GCs are very effective for many purposes, if given in large amounts they have many undesirable side effects, including bloating, hypertension, peptic ulcer, trunkal obesity, osteoporosis, and depression or mania.

Since GCs have so many different biological actions, much effort has been made to develop synthetic GCs that have a particular spectrum of effects deemed desirable but lack other effects deemed undesirable (Table 1). Examples are steroids which have strong anti-inflammatory effects but lack the mineralocorticoid effects. The modification of a double bond at C1–2 as in prednisolone (or its cortisone-like 11-keto analog prednisone) increases the GC activity while reducing the mineralocorticoid activity. Addition of a fluorine atom at C9 increases both GC (about 12-fold) and mineralocorticoid (about 25-fold) activity. Including both these modifications and adding a 16 α -methyl group gives dexamethasone which has about 30-fold greater GC activity but relatively little mineralocorticoid activity. Since the steroids themselves are relatively insoluble in water, they are often formulated as organic esters such as acetates, which are somewhat more soluble, or as salts such as sodium succinates which are freely soluble.

Although extremely useful, GCs must be used with considerable caution, and preferably for only limited periods, because of their important side effects that can occur with all forms, and include bloating, osteoporosis, trunkal obesity, and other signs and symptoms of GC excess.

One of the most dangerous effects of large doses of GCs, whether natural or synthetic, is suppression of the LHPA axis. This can occur after only a few weeks of high-dose GC therapy and may require several months for recovery to occur. For this reason, therapy is never stopped abruptly but only very slowly so that the organism has time to adjust.

Uses of Antiglucocorticoids as Therapeutic Agents

Inhibitors of steroidogenesis are used to treat conditions of GC excess, including the various forms of Cushing's syndrome, and more recently, major depression. These agents may be effective for short periods, but their effects are often overcome after a few weeks or months by the normal feedback mechanism. In depressed patients, this may be long enough to induce a remission, in which case the levels remain normal after stopping the medication.

See Also the Following Articles

11 β -Hydroxysteroid Dehydrogenases; Corticosteroid Receptors; Corticosteroid-Binding Globulin (Transcortin); Corticosteroids and Stress; Corticotropin Releasing Factor (CRF); Cushing's Syndrome, Medical Aspects; Cushing's Syndrome, Neuropsychiatric Aspects; Dexamethasone Suppression Test (DST); Glucocorticoid Negative Feedback; Glucocorticoids, Effects of Stress on; Glucocorticoids – Adverse Effects on the Nervous System; Hippocampus, Overview; Hypothalamic-Pituitary-Adrenal; Steroid Hydroxylases.

Further Reading

- DeKloet, R. E., Vreugdenhil, E., Oitzl, M. S., et al. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Reviews* **19**, 269–301.
- Dixon, P. F., Booth, M. and Butler, J. (1967). The corticosteroids. In: Gray, C. H. & Bacharach, A. L. (eds.) *Hormones in blood*, 2nd ed., vol. 2, London: Academic Press.
- Ehrhart-Bornstein, M., Hinson, J. P., Bornstein, S. R., et al. (1998). Intraadrenal interactions in the regulation of adrenocortical steroidogenesis. *Endocrine Reviews* **19**, 101–143.
- Falkenstein, E., Tillmann, H.-C., Christ, M., et al. (2000). Multiple actions of steroid hormones – a focus on rapid nongenomic effects. *Pharmacological Reviews* **52**, 513–555.
- Fieser, L. F. and Fieser, M. (1959). *Steroids*. New York: Reinhold Publishing Corporation.
- Meaney, M. J., Mitchell, J. B., Aitken, D. H., et al. (1991). The effects of neonatal handling on the development of the adrenocortical response to stress: implications for neuropathology and cognitive deficits in later life. *Psychoneuroendocrinology* **16**, 85–103.

- Murphy, B. E. P. (1967). Some studies of the protein-binding of steroids and their application to the routine micro and ultramicro measurement of various steroids in body fluids by competitive protein-binding radioassay. *Journal of Clinical Endocrinology and Metabolism* 27, 973–990.
- Murphy, B. E. P. and Branchaud, C. L. (1994). The fetal adrenal. In: Tulchinsky, D. & Little, B. (eds.) *Maternal-fetal endocrinology*, 2nd edn., pp. 275–295. Philadelphia: W. B. Saunders.
- Murphy, B. E. P., Ghadirian, A. M. and Dhar, V. (1998). Neuroendocrine responses to inhibitors of steroid biosynthesis in patients with major depression resistant to antidepressant therapy. *Canadian Journal of Psychiatry* 43, 279–286.
- Sandor, T. and Mehdi, A. Z. (1979). Steroids and evolution. In: Barrington, E. J. W. (ed.) *Hormones and evolution*, vol. 1, pp. 1–72. New York: Academic Press.
- Stewart, P. M. (2003). The adrenal cortex. In: Larsen, P. R., Kronenberg, H. M., Melmed, S., et al. (eds.) *Williams textbook of endocrinology* (10th edn., pp. 491–551). New York: W. B. Saunders.
- Westphal, U. (1986). *Steroid-protein interactions II*. New York: Springer-Verlag.
- Wild, D. (ed.) (2001). *The immunoassay handbook*, 2nd edn.). New York: Nature Publishing Group.
- Wolkowitz, O. M. and Reus, V. I. (1999). Treatment of depression with antiglucocorticoid drugs. *Psychosomatic Medicine* 61, 698–711.

Glucocorticoids, Role in Stress

J C Buckingham

Imperial College School of Medicine, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 261–269, © 2000, Elsevier Inc.

Secretion in Stress, Delivery to Target Tissues, and
Molecular Basis of Action
Functions of Glucocorticoids in Stress

Glossary

- Cell-mediated immunity* The process by which thymus-derived lymphocytes (T cells) respond to specific antigens, undergo clonal expansion, and mount a cytotoxic response.
- Humoral immunity* The process by which specific antigens activate bone marrow-derived lymphocytes (B cells), causing cell proliferation and the production and release of specific antibodies; also called antibody-mediated immunity.

Glucocorticoids are steroid hormones produced by the adrenal cortex. Their principal role is to restore homeostasis in conditions of stress. However, prolonged elevations in circulating corticosteroids, which may occur in conditions of chronic or repeated

stress, are potentially harmful and may contribute to the pathogenesis of a wide range of disorders.

Secretion in Stress, Delivery to Target Tissues, and Molecular Basis of Action

Secretion of Glucocorticoids in Stress

The glucocorticoids (GCs), cortisol and/or corticosterone, depending on the species, are produced by cells in the zona fasciculata of the adrenal cortex. In normal circumstances, serum GC concentrations are maintained within narrow limits with pronounced excursions occurring only in accord with the well-documented circadian rhythm or in response to physical or emotional stress. Circadian and stress-induced increases in GC secretion are driven by neural and humoral factors that impinge on the hypothalamus to cause the release into the hypothalamo-hypophyseal portal vessels of two neuropeptides, corticotropin releasing hormone (CRH) and arginine vasopressin (AVP); these peptides in turn act synergistically on the anterior pituitary gland to initiate the release of adrenocorticotrophic hormone (ACTH), the primary trigger of steroidogenesis. The secretion of cortisol/corticosterone is regulated further by a complex servo mechanism through which the steroids themselves negatively regulate the release of ACTH and its hypothalamic-releasing hormones. The hypothalamic-pituitary-adrenocortical (HPA) responses to repeated

or sustained stress are complex and depend on both the nature of the stress and the species. For example, in rats the adrenocortical response to cold stress is attenuated progressively with each exposure; that is, the axis becomes tolerant to the stimulus. In contrast, the HPA responses to repeated immune insults are well maintained and may become exaggerated, possibly because the feedback actions of the steroids are overcome.

Delivery of Glucocorticoids to Their Target Cells and Receptors

GCs in the systemic circulation are largely protein-bound (95%), mainly to a specific globulin, termed corticosteroid-binding globulin (CBG, or transcortin), but also to albumin. In principle, only the free steroid has ready access to the corticosteroid receptors, which are located intracellularly in the target cells. However, local mechanisms that release the steroid from its carrier protein(s) appear to exist in at least some target tissues: these include (1) interactions of the carrier protein with cell-surface CBG receptors and (2) enzymatic cleavage of CBG (e.g., by serine proteases in inflamed tissues), a process that lowers the affinity of the binding protein for the steroids. Access of the steroids to their receptors is regulated further within the target cells by 11β -hydroxysteroid dehydrogenase (11β -HSD) enzymes that control the interconversion of cortisol/corticosterone to their inactive counterparts, cortisone/ 11 -dehydrocorticosterone.

Molecular Basis of Glucocorticoid Action

The actions of the GCs are exerted mainly via the classic pathway of steroid action – the free steroid enters the target cells and binds to an intracellular receptor that is subsequently translocated to the nucleus, where it initiates specific changes in DNA transcription and, hence, in protein synthesis. In many instances, the activated steroid–receptor complex induces or represses transcription by binding directly to specific response elements in the promoter region of the gene; in others, however, it may interfere with the activity of other transcription factors, such as activator protein (AP)-1 and nuclear transcription factor (NF) κ B. In addition, the activated receptor may regulate posttranscriptional events, including mRNA transcript stability, RNA translation, and posttranslational processing of the protein.

In accord with the intracellular mode of action, the majority of biological responses to GCs are slow to emerge; some, however, are manifested more rapidly; for example, cortisol hyperpolarizes hippocampal neurons and coeliac ganglia within 1–2 min of

contact. The mechanisms underlying these rapid action are unknown, but increasing evidence favors a membrane action of the steroids, possibly via a cell surface receptor.

Corticosteroid Receptors

Two distinct subtypes of intracellular corticosteroid receptors have been identified: the mineralocorticoid receptor (MR) and the glucocorticoid receptor (GR). The MR has a high and approximately equal affinity for aldosterone and the endogenous GCs, corticosterone and cortisol ($K_d \sim 1$ – 2 nM). In contrast, the GR is of a lower affinity (K_d cortisol/corticosterone 10–20 nM), is glucocorticoid-selective, and does not bind aldosterone readily; a truncated form of this receptor, termed GR- β , which does not bind ligand but which has the capacity to bind DNA, also exists. MRs in the kidney are protected effectively from cortisol/corticosterone by type 2 11β -HSD (11β -HSD2), which acts primarily as a dehydrogenase and inactivates the steroids rapidly, thus permitting selective access of aldosterone to the receptors. MRs in the brain, however, are not always protected from GCs in this way and in many instances may be almost fully occupied at low physiological concentrations of cortisol/corticosterone. However, the delivery of GCs to GRs in organs such as the liver and white adipose tissues may be facilitated by type 1 11β -HSD (11β -HSD1), which acts primarily as a reductase and thereby regenerates the active steroids from their inactive 11 -dehydro metabolites. Although the MR undoubtedly mediates several facets of glucocorticoid activity in nonstress conditions, particularly in the central nervous system (CNS), the responses to the raised levels of the steroids observed in, for example, stress are effected primarily via GRs. These receptors are widely distributed in the body; cell responsiveness depends not only on the presence of GRs, but also the receptor concentration, which is known to fluctuate during development, the cell cycle, following disturbances in endocrine status, and on the activation of 11β -HSD1 enzymes.

Functions of Glucocorticoids in Stress

General Considerations

The concept that corticosteroids play a key role in the body's responses to stress emerged from the work of Selye on the alarm reaction and the general adaptation syndrome. It is now evident that cortisol/corticosterone displays two modes of activity. The first is a permissive or proactive role in which the steroids (1) maintain the basal activity of the HPA axis and

set its threshold for responding to stress and (2) normalize the body's response to stress by priming the body's defense mechanisms by, for example, facilitating the effects of catecholamines on lipid and carbohydrate metabolism; upregulating the expression of receptors for inflammatory mediators; and acting centrally to aid processes underlying selection attention, the integration of sensory information, and response selection. The second mode of cortisol/corticosterone action is suppressive or protective and enables the organism to cope with, adapt to, and recover from a stressful insult. Key to this mode are the powerful anti-inflammatory and immunosuppressive actions of the steroids that prevent the host defense mechanisms, which are activated by stress, from overshooting and damaging the organism. Other important protective actions include the ability of the steroids to redirect the metabolism to meet energy demands during stress, to exert important effects within the brain that promote memory processes, and to impair nonessential functions such as growth and reproductive function. The permissive actions of the steroids are evident at low physiological concentrations and are mediated at least partially, although probably not exclusively, by the high-affinity MRs. In contrast, the suppressive actions of the steroids emerge only when cortisol/corticosterone levels are raised by, for example, stress and are effected by the lower-affinity GRs.

The permissive and protective actions of GCs are complementary and enable the organism to mount an appropriate stress response and to maintain

homeostasis. Dysregulation of either by genetic or environmental factors is potentially harmful and may predispose the individual to a variety of diseases that, depending on the site of the lesion, may include depression, disorders of the host defense system, osteoporosis, obesity, diabetes mellitus, hypertension, and other cardiovascular diseases. Long-term elevations in GC production that may occur in conditions of chronic stress are particularly hazardous and may have severe deleterious effects on the body. The effects of acute and sustained elevations in serum GCs are compared in [Table 1](#).

Glucocorticoids and the Host Defense System

GCs exert multiple effects on the host defense system. In essence, they quench the early and late stages of the inflammatory response; that is, they suppress the initial vasodilation, the infiltration of leukocytes and pain, and the subsequent proliferative events associated with wound healing and tissue repair. They also oppose the changes in vascular permeability that occur in inflammation and thus reduce both edema formation and the increased permeability of the blood-brain barrier. In addition, GCs are powerful inhibitors of cell-mediated (i.e., T-cell-dependent) immune responses and may also modulate humoral (B-cell-dependent) responses. Chronic elevations in serum GCs thus induce immunosuppression and render the individual susceptible to infection.

The anti-inflammatory and immunosuppressive actions of the steroids are attributed to their powerful actions on the growth, differentiation, distribution,

Table 1 Response to acute and long-term elevations in serum glucocorticoids

<i>System</i>	<i>Acute^a</i>	<i>Long-term</i>
Host defense	Protection from potentially harmful inflammatory mediators	Immunosuppression and vulnerability to infection Poor tissue repair/wound healing
Metabolism	Mobilization of energy stores (↑ glycogen stores, ↑ gluconeogenesis, ↑ blood glucose, ↑ lipolysis, ↑ protein catabolism, ↓ peripheral glucose uptake/utilization)	Insulin-resistant steroid diabetes mellitus Centripetal obesity Protein depletion in muscle, connective, and other tissues
Musculoskeletal	Protein catabolism Altered Ca ²⁺ homeostasis	Raised serum lipids and cholesterol Impaired growth Muscle wasting Loss of connective tissue Osteoporosis and disturbed Ca ²⁺ homeostasis
Central nervous system	Improved cognitive function	Mood changes (depression and psychotic episodes) Neurodegeneration
Cardiovascular	Salt and water retention Inhibition of the production of vasoactive inflammatory mediators	Hypertension and other cardiovascular disease
Reproductive	Inhibition of hypothalamic-pituitary-gonadal function	Menstrual irregularities Infertility (male and female)
Gastrointestinal tract	Reduced bicarbonate and mucus production	Increased susceptibility to ulcers

^a↑, increased; ↓, decreased.

function, and life span of monocytes, macrophages, polymorphonuclear cells, and lymphocytes. At the cellular level, GCs reduce the number of circulating lymphocytes, monocytes, and eosinophils by causing the cells to be marginated to the lymphoid tissues, mainly the bone marrow. In contrast, blood neutrophil levels rise as the steroids prevent the cells migrating to inflamed tissues by altering the expression of adhesion proteins on their surface and on the endothelium and thereby impairing the adherence and attachment processes. The profile of circulating white cells is also influenced by the fact that GCs also induce the apoptosis (programmed cell death) of thymocytes, mature T cells, and eosinophils, although not of neutrophils.

Several processes contribute to the suppression of cell-mediated immunity invoked by the steroids. These include (1) direction of the development of undifferentiated T helper (Th) cells away from the Th1 phenotype, which is critical to cell-mediated responses, and toward the Th2 phenotype, which facilitates B-lymphocyte-dependent responses; (2) inhibition of the synthesis of interleukin (IL)-1 and IL-2 by antigen-presenting cells and T cells, respectively, and hence the impairment of antigen presentation and T-cell proliferation; and (3) induction of T-cell apoptosis. Very high levels of GCs may also depress the synthesis of the immunokines required for immunoglobulin production and, with time, decrease serum immunoglobulins. However, lesser amounts of the steroids stimulate antibody production, probably because of their positive effect on the differentiation of Th2 cells. Other immune cell functions influenced by GCs include phagocytosis and antigen processing. GCs decrease Fc receptor expression on macrophages and thus prevent the recognition of particulate antigens that are antibody-bound or optimized for subsequent clearance. They also affect mast cell and basophil function in a number of rodent models, inhibiting IgE-dependent degranulation and hence the release of histamine and leukotriene C4.

The diverse effects of GCs on immune/inflammatory cell function are explained at a biochemical level by the striking ability of the steroids to inhibit the synthesis, release, and/or activity of the army of mediators normally produced when these cells are activated. Early studies focused on the powerful inhibitory action of the steroids on the phospholipase A2 (PLA2)-dependent generation of the pro-inflammatory eicosanoids (prostanoids, leukotrienes, and epoxides) and platelet-activating factor (PAF). Several lines of evidence indicate that GCs reduce the expression of the PLA2 subtype (group II) associated with the pathogenesis of inflammation and upregulate the synthesis of

PLA2 inhibitory proteins, notably lipocortin 1 (also called annexin 1), which also influences other aspects of the inflammatory response. However, studies have shown that GCs also suppress prostanoid generation by repressing the expression of the inducible isoform of cyclooxygenase (COX II), which is normally expressed in abundance by activated monocytes, macrophages, fibroblasts, and endothelial cells.

GCs also exert profound inhibitory effects on the synthesis of many of the cytokines that are central to the manifestation of immune/inflammatory responses; these include IL-1, IL-2, IL-3, IL-6, IL-8, IL-12, granulocyte colony-stimulating factor (G-CSF), granulocyte-macrophage colony-stimulating factor (GM-CSF), interferon (IFN)- γ , tumor necrosis factor (TNF)- α , and monocyte chemotactic protein. Various mechanisms may contribute to these responses, including the downregulation of gene transcription and acceleration of mRNA degradation, together with effects at the translational and posttranslational levels. The biological actions of the cytokines may also be impaired; this may be explained in part by alterations in receptor expression and, for example, by the IL-2 receptor being downregulated by GCs. However, in many cases (e.g., IL-1, IL-6, and IFN- γ), the repression of cytokine expression is associated with the upregulation of the respective cytokine receptors ([Figure 1](#)); this phenomenon may help explain the permissive role of the steroids in priming the immune system for action as well as the ineffective responses mounted when GC levels are raised (e.g., by stress) or when adrenal function is impaired.

The release of many other inflammatory mediators may also be prevented by virtue of the powerful regulatory actions of the GCs on the enzymes responsible for their biosynthesis/degradation. For example, GCs suppress the inducible form of nitric oxide (NO) synthase and hence the generation of NO by macrophages. However, they promote the metabolism of nonlipid inflammatory mediators (e.g., bradykinin) by upregulating the enzymes responsible for their metabolism. Further important actions of the GCs include reductions in the serum complement components, in the release of mediators such as 5-hydroxytryptamine (5-HT), and in the production/activity of the proteolytic enzymes (e.g., elastase and collagenase) that are concerned with the late stages of inflammation.

Metabolic Actions of Glucocorticoids

GCs exert complex effects on carbohydrate, lipid, and protein metabolism; their principal actions are catabolic, and they thus mobilize energy stores and oppose the actions of insulin ([Figure 2](#)).

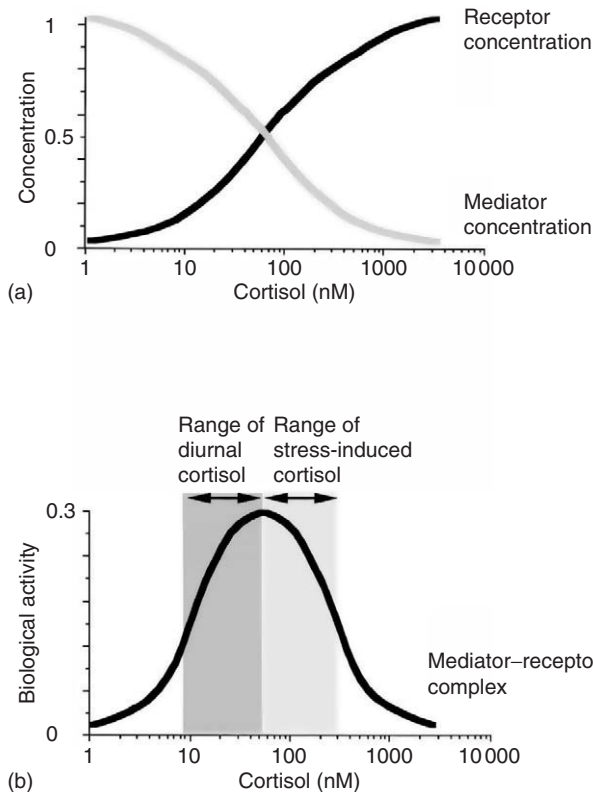


Figure 1 Mathematical model for the permissive and suppressive effects of glucocorticoids over a physiological concentration range. (a), Dose–response curves for a theoretical mediator and its specific receptors; (b), mathematical composite of mediator–receptor complex formation indicating biological activity and suggesting a mechanism by which low cortisol levels could be permissive and stress-induced cortisol surges suppressive in effect. From Munck, A. and Náraj-Fejes-Tóth, A. (1992), The ups and downs of glucocorticoid physiology: permissive and suppressive effects revisited, *Molecular and Cellular Endocrinology* **90**, C1–C4, copyright John Wiley & Sons Limited. Reproduced with permission.

Carbohydrate metabolism

The principal target is the liver, where GCs increase glycogen storage and gluconeogenesis. The effects of the steroids on glycogen storage reflect an increase in glycogen deposition due to the activation (dephosphorylation) of glycogen synthase and a concomitant decrease in glycogen breakdown due to inactivation of the glycogen-mobilizing enzyme glycogen phosphorylase. The increase in glucose synthesis is explained partly by the upregulation of key enzymes in the gluconeogenic pathway, such as phosphoenolpyruvate carboxykinase (PEPCK). In addition, the catabolic effects of the steroids on protein metabolism in peripheral tissues (e.g., muscle) and lipolysis in adipose tissue increase the availability of substrates (glucogenic amino acids and glycerol), while the fatty acids released with glycerol provide an important energy source for the process. The actions of GCs in the

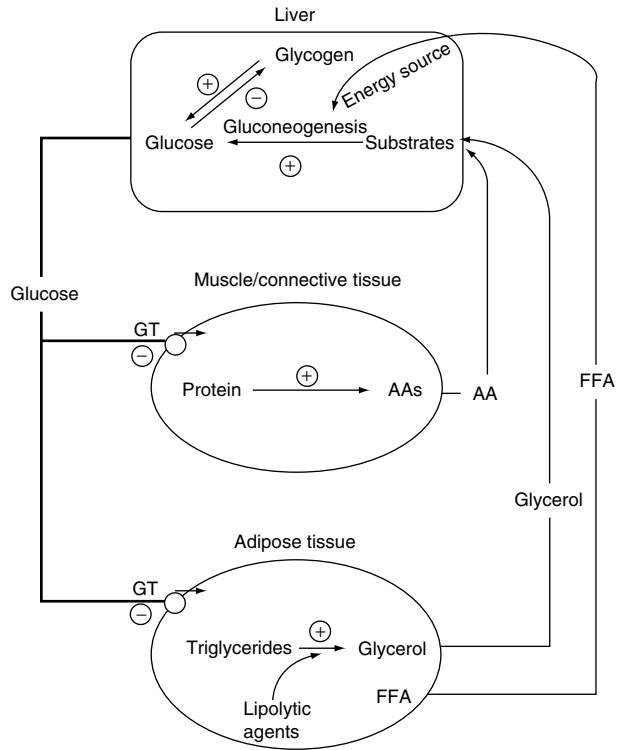


Figure 2 Schematic diagram illustrating the effects of glucocorticoids on carbohydrate, protein, and lipid metabolism. Note that the positive effects of the steroids on lipolysis involve interactions with other lipolytic agents (e.g., growth hormone and catecholamines). ⊕, stimulation; ⊖, inhibition; AA, amino acid; FFA, free fatty acid; GT, glucose transporter.

liver are effected by GR, and the delivery of the steroids to the receptors is facilitated by an abundance of 11 β -HSD1, which serves mainly as a reductase.

Although much of the glucose formed in the liver in response to GCs is stored locally as glycogen, some is released into the circulation. Blood sugar levels are enhanced further by the actions of GCs in the peripheral tissues, which lead to the downregulation of the glucose transporters and, hence, to decreased glucose uptake and use.

Lipid metabolism GCs exert a permissive influence on lipid metabolism, producing an acute lipolytic action that appears to reflect an increased responsiveness to catecholamines, growth hormone (GH), and possibly other lipid-mobilizing agents. In the long term, however, GCs produce a striking redistribution of fat, with enhanced deposition in the trunk (centripetal obesity), dorsocervical and subcapsular regions (buffalo hump), and face (moon face), together with a loss of fat in the upper and lower limbs, giving rise to the 'lemon-on-sticks' appearance that is exacerbated by the changes in protein metabolism. Blood lipid and cholesterol levels are also raised. The underlying mechanisms are poorly defined; they may

involve alterations in insulin secretion triggered by the changes in glucose homeostasis, and it is interesting to note that white adipose tissue is rich in GRs and 11 β -HSD1. Paradoxically, this pattern of fat redistribution is not seen in experimental animals.

Protein metabolism GCs have a powerfully catabolic effect on protein in muscle and other tissues with the consequent release of amino acids into the circulation, which are converted to glucose in the liver. Long-term elevations in serum GCs may cause severe muscle wasting.

Skeleton and Connective Tissue

Skeleton GCs exert direct and indirect effects on bone, and persistent elevations of serum GCs may lead to osteoporosis. The steroids decrease bone formation by inhibiting (1) the proliferation and differentiation of osteoclasts and (2) the production of type 1 collagen, osteocalcin, and other matrix components such as mucopolysaccharides and sulfate glucoaminoglycans while increasing the expression of collagen. In addition, they exert complex effects on the expression of growth factors and their receptors and reduce Ca^{2+} stores by decreasing intestinal Ca^{2+} absorption, increasing renal Ca^{2+} excretion, and raising serum parathyroid hormone (PTH) levels. GCs also promote bone resorption through mechanisms that are secondary to the rise in PTH and possibly also via direct actions on the bone.

Sustained elevations in serum GCs inhibit linear growth in children and experimental animals. The mechanisms responsible have yet to be defined but are likely to reflect the actions of the steroids on the bone, principally at the site of the epiphyses. Although there are reports to the contrary, the majority of data suggests that the serum levels of GH, insulin-like growth factor (IGF)-1, and IGF-binding proteins are normal in these subjects.

Connective tissue The actions of GCs on connective tissue lead to poor wound healing and tissue repair. The steroids decrease fibroblast proliferation and function; inhibit the production of the matrix proteins, collagen, and hyaluronidate; and augment the production of the matrix-degrading enzymes collagenase, elastase, and hyaluronidase. However, GCs augment the production of fibronectin, a matrix glycoprotein.

Cardiovascular

Elevations in serum GCs raise blood pressure. This may be at least partially explained by mineralocorticoid-like actions of the steroids, which cause increased

salt and water retention, and by the anti-inflammatory actions of the steroids, which prevent the release of local vasodilator substances such as prostacyclin and bradykinin. However, other mechanisms also come into play. These include (1) the upregulation of vascular α_1 -adrenoceptors and consequent increased sensitivity of the vessels to the vasoconstrictor actions of norepinephrine; (2) increased cardiac output, possibly as a result of increased β_1 -adrenoceptor expression in the myocardium; and (3) increased production of angiotensinogen and increased sensitivity of the blood vessels to the vasoconstrictor actions of angiotensin II. The endothelium is rich in GRs, and it is conceivable that the steroids also modulate the release/activity of other locally produced vasoactive substances. In addition, reports that GRs are abundant in the brain-stem nuclei raise the possibility that GCs may also act centrally to alter sympathetic outflow.

Mood and Behavior

GCs released in stress may exert significant effects on cognitive function, mood and behavior, and reception of sensory inputs. Subjects whose serum GCs are raised due to either Cushing's disease or exogenous steroids also show marked disturbances in sleep, with a reduction in time spent in the rapid eye movement phase.

Studies on laboratory animals have shown that GCs act in the neocortex and limbic system (hippocampus, septum, and amygdala) to modify learning and memory processes in a context-dependent manner. Thus, an acute rise in GCs may facilitate learning and memory processes but also eliminate learned behavior that is no longer of relevance. The complex actions of the steroids involve both MRs and GRs, with MRs effecting the appraisal of information and response selection and GRs promoting the processes underlying the consolidation of acquired information. The long-term effects of GCs on learning and memory are less well researched and the data available are conflicting, although, when given with neurotoxic agents, learned behavior is impaired by the steroids.

GC levels are raised in some patients with depression, but it remains to be determined whether this is cause or effect; indeed, the effects of GCs on mood appear variable. Thus, approximately one-half of patients with iatrogenic or endogenous Cushing's disease show psychological disturbances; depression is the most common but some subjects show euphoric or manic behavior and, in some cases, overt psychoses. Patients with adrenal insufficiency also frequently experience mood changes; paradoxically, the most common is also depression.

Increasing evidence suggests that sustained elevations in serum GCs, such as those that occur due to chronic stress, may contribute to the processes of neurodegeneration. The toxic actions of the steroids, which have been studied most widely in the hippocampus, are associated with altered glutaminergic transmission, a long-term enhancement of Ca^{2+} influx, and depletion of ATP stores, with a consequent impairment of essential energy-requiring processes such as glucose uptake. The vulnerability of the neurons may be enhanced further by a reduction in neurotrophic capacity because the elevations in serum GCs induced by chronic stress or the administration of exogenous steroids inhibit the expression of brain-derived neurotrophic factor mRNA and also tyrosine kinase, which is required for growth factor action.

Reproductive Function

GCs act at multiple sites within the hypothalamic-pituitary-gonadal axis to impair reproductive function; hence, chronic elevations in serum GCs lead to infertility and menstrual irregularities. An important site of action of the steroids is the hypothalamus. In addition, the steroids exert significant effects at the pituitary level to depress gonadotropin-releasing hormone (GnRH)-driven gonadotropin release and may also exert further regulatory actions within the gonads.

Development

GCs exert both positive and negative effects on cell growth. In addition, they fulfill complex organizational effects pre- and postnatally and thus play a key role in developmental programming. Not surprisingly, the HPA axis is tightly regulated during perinatal life and, in rodents, the stress response is attenuated during a critical phase postnatally. At this time, the developing HPA axis shows a high degree of plasticity, and a growing body of evidence suggests that disturbances in the GC milieu may (1) trigger a plethora of pathologies (e.g., growth retardation or impairment of CNS development) and (2) precipitate abnormalities in GC secretion that persist into adult life and may increase the susceptibility of the individual to conditions as diverse as hypertension, insulin resistance, cognitive and other behavioral disorders, and allergic/inflammatory disease. Perinatal manipulations that alter the expression of GRs (but not MRs) in the brain are particularly hazardous in this regard because, by disrupting the negative feedback mechanisms, they alter the responsiveness of the HPA axis to stress throughout the lifetime of the individual. For example, the exposure of neonatal rats to stimuli that raise serum GCs, such as

an immune insult (the injection of bacterial lipopolysaccharide) or the administration of exogenous GCs, results in the adult in the downregulation of GRs in the hippocampus and hypothalamus, decreased sensitivity of the HPA axis to the negative feedback actions of the GCs, and exaggerated HPA responses to stress. Similar long-term changes in GR expression occur in rats subjected to prolonged periods of maternal separation; the consequent impairment of the negative feedback system and changes in HPA function are associated with alterations in T-cell antibody responses and in the incidence of age-related neuropathies. In contrast, neonatal handling, which increases hippocampal GR expression by mechanisms involving serotonin, reduces the magnitude and duration of the stress response in the adult; maternal care may therefore help to program events that regulate responses to stress in the adult and thus reduce vulnerability to disease.

Other Actions

GCs inhibit the production of mucus and bicarbonate in the stomach and small intestine by virtue of their ability to block prostaglandin production; sustained elevations in GCs thus increase susceptibility to gastric and duodenal ulcer formation. The steroids also promote Na^+ absorption in the colon by mechanisms that involve MRs and GRs while acting via GRs to augment Na^+ excretion by stimulating the synthesis of atrial natriuretic peptide (ANP) and augmenting the actions of ANP on the kidney, thereby opposing the Na^+ -retaining influence exerted via MRs in the distal renal tubule. In addition, GCs may enhance free-water clearance by depressing the synthesis of vasopressin.

See Also the Following Articles

Adrenal Cortex; Adrenocorticotrophic Hormone (ACTH); Cytokines; Glucocorticoids, Effects of Stress on; Glucocorticoids – Adverse Effects on the Nervous System.

Further Reading

- Barnes, P. J. (1998). Anti-inflammatory actions of glucocorticoids: molecular mechanisms. *Clinical Science* **94**, 557–572.
- Brann, D. W. and Manesh, V. B. (1991). Role of corticosteroids in female reproduction. *FASEB Journal* **5**, 2691–2698.
- Buckingham, J. C., Gillies, G. E. and Cowell, A.-M. (1997). *Stress, stress hormones and the immune system*. Chichester, UK: John Wiley.
- de Kloet, E. R., Vreugdenhil, E., Oitzl, M. S., et al. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Review* **19**, 269–301.

- Lukert, B. P. and Raiz, L. G. (1994). Glucocorticoid-induced osteoporosis. *Rheumatic Disease Clinics of North America*.
- Munck, A., Guyre, P. M. and Holbrook, N. J. (1984). Physiological functions of glucocorticoids and their relation to pharmacological actions. *Endocrine Review* **51**, 25–44.
- Munck, A. and Náraj-Fejes-Tóth, A. (1992). The ups and downs of glucocorticoid physiology: permissive and suppressive effects revisited. *Molecular and Cellular Endocrinology* **90**, C1–C4.
- Nestler, J. E. and McClanahan, M. A. (1992). Diabetes and adrenal disease. *Baillière's Clinical Endocrinology and Metabolism* **6**, 829–851.
- Pilkis, S. J. and Granner, D. K. (1992). Molecular physiology of the regulation of hepatic gluconeogenesis and glycolysis. *Annual Review of Physiology* **54**, 885–909.
- Robinson, I., Gabrielsson, B., Klaus, G., et al. (1995). Glucocorticoids and growth problems. *Acta Paediatrica* **84**, 81–85.
- Selye, H. (1952). *The story of the adaptation syndrome*. Montreal: Acta.
- Tilders, F. J. and Schmidt, E. D. (1998). Interleukin 1 induced plasticity of hypothalamic CRH neurons and long-term stress hyperresponsiveness. *Annals of the New York Academy of Sciences* **840**, 65–73.
- Vinson, G. P. and Anderson, D. C. (1996). *Adrenal glands, vascular system and hypertension*. Bristol, UK: Society for Endocrinology.

Glucose Transport

A L McCall

University of Virginia Health System, Charlottesville, VA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A L McCall, volume 2, pp 270–275, © 2000, Elsevier Inc.

Introduction

Glucose Transport

Overview of Glucose Transport Regulation

Stress Hormones and Glucose Transport

GLUT1 as a Stress-Related Protein

Metabolic Stress and Glucose Transport

Overall Effects of Transport Regulation

Signaling Cascades and Glucose Transport

Summary

Glossary

<i>ATP/ADP</i>	Ratio of adenosine triphosphate to adenosine diphosphate, an index of cellular energy stores.
<i>Dinitrophenol</i>	A mitochondrial inhibitor associated with increased glucose transport <i>in vitro</i> .
<i>GLUT</i>	Glucose transport proteins enumerated in the order of their cloning, e.g., GLUT1, GLUT2.
<i>Mitogen-activated protein kinase</i>	A family of enzymes activated by hormones, cytokines, stress, and growth factors that result in a cascade of kinase activity and thereby influence cellular signaling.

Introduction

Glucose transport supplies fuel needed for energy metabolism by all mammalian cells. The supply of glucose is particularly important for cells, such as the neurons of the brain, that have a high metabolic rate supported by an obligate consumption of glucose as fuel. Transport is regulated by a variety of factors, including those associated with stress. The transport proteins that accomplish glucose transport are modulated in their expression, cellular distribution, synthesis, and half-lives by stress-related factors. Such factors include stress hormones, metabolic stress such as cellular energy demand or reduced supply of fuel or oxygen, and stress-related kinase signaling. The net effect of such regulation may be to ensure appropriate distribution of glucose fuel during stress to tissues that most require this particular fuel.

Glucose Transport

Simple diffusion of glucose across cellular phospholipid barriers is limited by its modest hydrophobicity. To overcome this, transfer of glucose across cell membranes uses a process known as facilitated diffusion, which is a carrier protein-mediated process. Two types of facilitated diffusion exist: energy dependent and energy independent. A few cells, such as the gut and kidney epithelium, are directly energy dependent and use a sodium-glucose cotransport. For most cells, however, energy-independent facilitated diffusion predominates. Facilitated diffusion via carrier proteins transports glucose down a concentration

gradient in a saturable manner and is the major mechanism for glucose entry into mammalian cells. Such transport is accomplished by different isoforms of a superfamily of hydrophobic, integral membrane proteins often referred to as GLUTs or glucose transporters (see [Figure 1](#)). The cloning of the first member of this family, GLUT1, has facilitated much of the physiological investigation of glucose transport and these proteins. The homology of these transporters has led to additional members of this family of proteins being

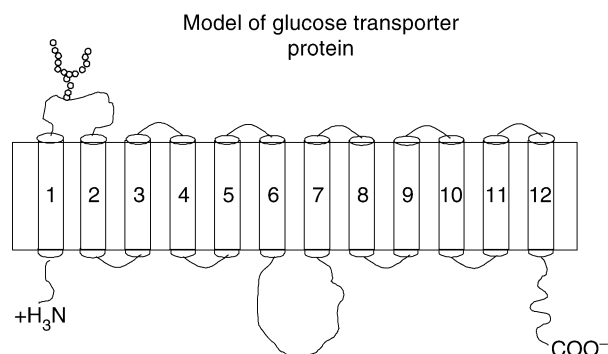


Figure 1 A model of glucose transporters. The diagram illustrates the putative 12 membrane-spanning domains of the glucose transporter's proposed structure, based on its cloning by Mueckler. This structure is reproduced for all class I GLUTs but is slightly different in placement of the carbohydrate moiety on the exofacial loop in other GLUT types. In this model of GLUT1, an N-linked asparagine residue on the exofacial (upper) surface of the cell membrane has an attached carbohydrate moiety that in certain isoforms may alter the efficiency of transport function. N- and C-termini of the proteins reside in intracellular sites.

identified. Their characterization physiologically in relation to stress signaling is as yet incomplete. [Table 1](#) lists the members of the GLUT family (excluding HMIT) of proteins and partially characterizes them. In brief, the GLUT family of transporters of hexoses and polyols now has 13 members, including 12 hexose transporters and one myoinositol transporter (HMIT). There are three classes within the transporter family: class I includes GLUT1–4, class II includes GLUTs 6, 8, 10, and 12, and class III includes 5, 7, 9, 11, and HMIT.

Overview of Glucose Transport Regulation

Glucose transport is a highly regulated process that varies considerably from one cell type to another. Some cells, such as red blood cells and brain neurons, have obligate consumption of glucose. For other cells, a facultative use of glucose exists, permitting other metabolic fuels, such as fatty acids, to supply the bulk of local energy requirements. Much evidence suggests that the regulation of glucose transport is also isoform specific. Regulation of transport occurs in response to altered energy requirements of tissues, so it is not surprising that one form of cellular stress, energy lack, is a potent regulator of glucose transport and transporter expression by different tissues. Regulation of GLUT proteins may occur by variation of the amounts of synthesis or degradation of the GLUT protein or mRNA. An increased transcription of GLUT mRNA or other regulatory effects on GLUTs

Table 1 The GLUT family (SLC2), with each member (except the related SMIT myoinositol transporter) listed and given its class, and some characteristics of tissue distribution and other features.

Common name	Class	Tissue distribution	Features/Comments
GLUT1	I	Brain microvessels, glia, erythrocytes, placenta, kidney	Widely expressed, synthesis induced by cell culture, growth factors, oncogenes, heat shock, cell stress, development
GLUT2	I	Liver, kidney, pancreatic β cells, small intestine	Low-affinity (high K_m) transporter, role in insulin release
GLUT3	I	Neurons, placenta, fetal skeletal muscle, platelets, leukocytes	High affinity (low K_m), expressed in glucose-dependent cells, regulated by development, metabolic stress
GLUT4	I	Skeletal muscle, heart, brain regions, vessels	Insulin responsive, translocates from intra-cellular compartment to plasma membrane stress hormone modulation
GLUT5	II	Testes, small intestine,	High affinity (low K_m) for fructose
GLUT6	II	Brain, spleen, leukocytes	Function ill defined, formerly GLUT9
GLUT7	III	Unknown	Homologous to GLUT5,? fructose transporter
GLUT8	II	Testes, brain, adrenal, liver spleen, fat, lung, blastocyst	High affinity for glucose, hormone/stress modulation
GLUT9	III	Kidney, liver	Transport activity unclear, GLUT5 homology
GLUT10	II	Heart, lung, brain, liver	High 2DG affinity
GLUT11	III	Liver, brain, lung, muscle	Low glucose affinity, multiple splice variants
GLUT12	II	Heart, muscle, fat, prostate	Substrate specificity not known

may occur as a result of stress hormones, such as glucocorticoids and epinephrine. In some tissues, these hormones have differing effects on expression of GLUTs or their transport activity, emphasizing their tissue-specific regulation. The effects of growth factors, physiological factors often involved in stress responses, increase GLUT1 transcription.

Stress Hormones and Glucose Transport

Stress hormone responses influence glucose transport in a tissue-specific and isoform-specific manner. Among the important influences on glucose transport are the stress hormones, particularly adrenal glucocorticoids. Another common stress hormone response that regulates glucose transport is the secretion of catecholamines, epinephrine and norepinephrine, by the adrenal medulla and the sympathetic nervous system. Glucocorticoid stress responses or exposure to synthetic glucocorticoid drugs, such as dexamethasone, regulates the expression of several glucose transporter isoforms. In cultured adipocytes and in skeletal muscle, which are both insulin-sensitive tissues, glucocorticoids downregulate glucose transport by decreasing the plasma membrane concentration of glucose transporters, particularly GLUT4, the insulin-sensitive glucose transporter. This was shown by several groups. This cellular redistribution of GLUT4 is opposite to the effects of insulin.

Kinase signaling may be important for glucose transport adaptations in response to stress hormones. In 3T3-L1 adipocytes, downregulation of GLUT4 activity has been found. The alteration is not from decreased PI3 kinase signaling, as might be expected, nor is it from downstream regulation in this insulin-responsive pathway. Instead, a decrease in p38 mitogen-activated protein (MAP) kinase phosphorylation occurs. This is caused by upregulation of MAP kinase phosphatases.

In adipocytes, glucocorticoids may stimulate the synthesis of GLUT1, however, consistent with tissue and isoform specificity of regulation. GLUT1 is also overexpressed in the small blood vessels of the brain (which comprise the blood-brain barrier) after *in vivo* dexamethasone treatment. β -adrenergic drugs and the stress hormone catecholamines may also reduce insulin-sensitive glucose transport via action of GLUT4 in tissues such as muscle and fat, although precise mechanisms are incompletely defined.

In the gut, GLUT2 may be downregulated as a result of stress. A complex effect of stress (restraint stress and uncontrolled diabetes have been studied) is seen with GLUT8 in the hippocampus. Short-term

stress had no effect alone but normalized the increased GLUT8 mRNA levels resulting from diabetes. The GLUT8 protein shows no change in total protein, but instead a redistribution induced by stress in GLUT8 protein to high-density microsomal fractions occurs. Intriguingly, both uncontrolled diabetes and restraint stress lower GLUT8 protein in high-density microsomes, and the two combined reduce levels to a greater degree than either alone. These examples of stress hormone-mediated effects on glucose transport suggest the importance of tissue, stress, and isoform specificity when understanding the effects of stress on glucose transport. The net effect of such regulation may redistribute glucose to tissues such as the brain that most need it as a fuel. Tissues, such as muscle, during stress may utilize other fuels, e.g., fatty acids. Some of these stress responses are thus thought to be adaptive, although given the biphasic nature of some, there may also be maladaptive or deleterious aspects, particularly of combined stresses.

GLUT1 as a Stress-Related Protein

GLUT1, the ubiquitously expressed glucose transporter, which is constitutively expressed at low levels by most, if not all, cells, is regulated by cellular stress. A variety of cell stressors, including glucose starvation, mercaptoethanol, calcium ionophores, tunicamycin, and heat shock, upregulated glucose transport and GLUT1 expression in L8 muscle cells and in 3T3L1 fibroblasts in an isoform-specific fashion. Another glucose transporter, GLUT4, failed to respond similarly. The GLUT1 stress responses were regulated in a manner very similar to that seen for another cell stress protein, glucose-regulated protein 78 (GRP78), in these studies. Subsequently, other studies confirmed the stress-related increase in GLUT1 expression in a variety of tissues and cell types. GLUT1 is thus the prototypical stress-responsive GLUT isoform. Nonetheless, many other GLUT isoforms are regulatable by factors associated with stress.

Metabolic Stress and Glucose Transport

Hypermetabolism

One aspect seen frequently with cellular stress or injury is a transient hypermetabolism. Perhaps this hypermetabolism is used to provide biochemical energy in the form of ATP or its high-energy phosphate bond equivalents for restoration of ionic cellular equilibrium and energy-dependent repair processes. The increased energy requirements after stress or

injury might exert regulatory effects on glucose transport and transporters through an energy intermediate in the metabolic pathways of glucose. A potential linkage to altered metabolism and glucose transport is the fuel sensor enzyme AMP kinase, which accelerates catabolic pathways such as glycolysis and fatty acid oxidation to restore the ATP/AMP ratio that may be depleted. Increased glucose transport can be seen in some tissues as a result of AMP kinase activation.

Mitochondrial Inhibitors

Metabolic stress may regulate glucose transport and transporters *in vivo* and *in vitro*. Inhibitors of mitochondrial oxidative phosphorylation have long been known to increase glucose transport *in vitro*, although the precise metabolic pathways and mechanisms are complex and only partly elucidated. This phenomenon was first described in cultured fibroblasts. Mitochondrial inhibitors may similarly affect GLUT proteins. Hypoxia and mitochondrial inhibitors or uncouplers of oxidative phosphorylation (dinitrophenol, antimycin, and azide) act similarly to hypoxia and increase both GLUT1 mRNA and protein expression *in vitro*. Increasing evidence supports the idea that GLUT3 may also respond to these stimuli, albeit through different mechanisms than GLUT1 or GLUT4.

Fuel Deprivation: Oxygen, Glucose, Ischemia

In a like manner, glucose transport is upregulated by fuel deprivation *in vitro* and *in vivo*. This deprivation may occur due to glucose lack, hypoxic conditions, or ischemia. Glucose deprivation induced upregulation of glucose transport has been shown in fibroblasts, in muscle cells, and in many other cell types. Some forms of feedback probably occur by metabolic intermediates in the metabolism of glucose and/or by feedback directly from the energy state of the cell, although the specific metabolic compound is not clear. Perhaps ATP levels or ATP/ADP ratios may be involved in mediating this increased glucose transport. Indeed, a substantial body of evidence suggests that the regulation of glucose transport in many tissues is influenced by the availability of the transport substrate itself, i.e., glucose. Altered blood flow, reduced oxygen, or glucose availability may increase transport. Hypoxia and hypoglycemia can exert their effects on glucose transport (GLUT1) at least partly through action of vascular endothelial growth factor (VEGF). In some situations this occurs through extension of the half-life of the mRNA for VEGF. Hypoxia and ischemia may also induce hypoxia-inducible factor 1 alpha (HIF-1 α), which in turn may be involved in

upregulation also through AMP kinase, the previously mentioned fuel sensor enzyme.

Overall Effects of Transport Regulation

The net effect of GLUT regulation in stress is to redistribute glucose fuel to tissues that have an obligate requirement for this energy substrate. Glucocorticoids may upregulate GLUT1 protein expression in tissues such as the brain, resulting in increased availability of fuel for that tissue. A complex aspect of that upregulation, however, is that endothelial cells may also have a relatively acute effect to decrease transport activity of glucose as well, leaving the net effect to vary depending upon the balance of transport-inhibiting effects that are probably not related to GLUT expression and enhancing effects that probably are related to GLUT expression. In contrast to effects in the brain of glucocorticoids, in insulin-dependent tissues, which employ GLUT4 as their primary glucose transporter, an intracellular sequestration of this transporter may prevent draining of glucose fuel by muscle during stress responses.

Although there is controversy about this point in some tissues, evidence suggests that local energy metabolism may be affected by such stress-mediated regulation of glucose transporters and glucose transport. For example, with *in vivo* hypoglycemia, an upregulation of brain endothelial transport of glucose occurs in animal studies and in some human studies. At blood glucose levels of about 3 mM (equal to 56 mg/dl), brain dysfunction normally begins to occur and presumably is related to a lack of adequate glucose entry into the brain to support the rather rigid requirement for glucose by brain metabolism. With an adaptation of glucose transport that occurs within several days of repeated hypoglycemia (in this case produced *in vivo* by overdoses of exogenous insulin), there is maintenance of glucose-6-phosphate and brain creatine phosphate and ATP levels. Presumably, the increased transport efficiency, which appears to involve both brain vascular (GLUT1) and neuronal glucose transporters (GLUT3), effectively operates to shuttle a greater amount of glucose across the brain endothelial and neuronal cellular barriers. This helps feed starving neurons. Interestingly, although such an adaptation would seem to be an unalloyed benefit, it has been suggested that a reduced awareness of hypoglycemia also occurs as a result. This reduced awareness may actually increase the risk of subsequent hypoglycemia in those with insulin-treated diabetes, which might risk brain injury from fuel deprivation. Glucose sensing in the brain may occur in discrete

neurons, which may be important in hypoglycemia responses and in feeding.

Signaling Cascades and Glucose Transport

Stress in cells and whole organisms is associated with a wide variety of hormonal and biochemical changes that induce cellular responses via many mechanisms. One mechanism for signal transduction in stress occurs via certain types of MAP kinase pathways, particularly the p38 and jun-kinase subtypes. The p38 kinases, mammalian homologs of osmotic stress kinases found in yeast, may help regulate both basal- and stress-activated expression of glucose transporters, including GLUT1 and GLUT3, according to work performed in cultured muscle cells. The fuel sensor kinase, AMP kinase, is able to stimulate GLUT expression in some tissues as well.

Summary

It is clear that stress can either activate or suppress glucose transport in many tissues. GLUT1, the widely expressed isoform, appears to be a prototype of the stress-activated glucose transporter isoform, but GLUT3 and GLUT4 may also increase or decrease in response to certain kinds of cellular stress. New data suggest that GLUT2 and GLUT8 show stress-related regulation as well. Less is known about the regulation of other GLUT isoforms. Hormones associated with stress response may be important for increasing or, in some instances, decreasing cellular glucose transport. The preservation of an adequate supply of glucose fuel to tissues is believed to be adaptive, but occasionally the redirection of glucose to certain tissues may also have adverse consequences. Adaptations of glucose transport and GLUT transporters are responsive to many signaling pathways activated by stress, including hormones, kinase signaling, fuel or oxygen or energy lack, and probably other pathways as yet incompletely delineated. Stress-mediated changes in glucose transporter amount or function probably act to redistribute this fuel to tissues such as the brain that have an obligate demand for it.

Acknowledgments

A.L.M. acknowledges support for laboratory research reported in part in this article from the National Institutes of Health (NINDS; NS22213, NS17493, NIDDK; DK 58006) and the Juvenile Diabetes Foundation International.

Further Reading

- Bazuine, M., Carlotti, F., Tafrechi, R. S., Hoeben, R. C. and Maassen, J. A. (2004). Mitogen-activated protein kinase (MAPK) phosphatase-1 and -4 attenuate p38 MAPK during dexamethasone-induced insulin resistance in 3T3-L1 adipocytes. *Molecular Endocrinology* 18(7), 1697–1707.
- Bazuine, M., Carlotti, F., Rabelink, M. J., Vellinga, J., Hoeben, R. C. and Maassen, J. A. (2005). The p38 mitogen-activated protein kinase inhibitor SB203580 reduces glucose turnover by the glucose transporter-4 of 3T3-L1 adipocytes in the insulin-stimulated state. *Endocrinology* 146(4), 1818–1824.
- Joost, H. G., Bell, G. I., Best, J. D., et al. (2002). Nomenclature of the GLUT/SLC2A family of sugar/polyol transport facilitators. *American Journal of Physiology Endocrinology and Metabolism* 282(4), E974–E976.
- Klip, A., Tsakiridis, T., Marette, A. and Ortiz, P. (1994). Regulation of expression of glucose transporters by glucose: a review of studies in vivo and in cell cultures. *FASEB Journal* 8, 43–53.
- Kumagai, A., Kang, Y., Boado, R. and Pardridge, W. (1995). Upregulation of blood-brain barrier GLUT1 glucose transporter protein and mRNA in experimental chronic hypoglycemia. *Diabetes* 44, 1399–1404.
- Lee, M., Hwang, J. T., Lee, H. J., et al. (2003). AMP-activated protein kinase activity is critical for hypoxia-inducible factor-1 transcriptional activity and its target gene expression under hypoxic conditions in DU145 cells. *Journal of Biological Chemistry* 278(41), 39653–39661.
- Levin, B. E., Routh, V. H., Kang, L., Sanders, N. M. and Dunn-Meynell, A. A. (2004). Neuronal glucosensing: what do we know after 50 years? *Diabetes* 53(10), 2521–2528.
- Nordal, R. A., Nagy, A., Pintilie, M. and Wong, C. S. (2004). Hypoxia and hypoxia-inducible factor-1 target genes in central nervous system radiation injury: a role for vascular endothelial growth factor. *Clinical Cancer Research* 10(10), 3342–3353.
- Pirolli, G. G., Grillo, C. A., Charron, M. J., McEwen, B. S. and Reagan, L. P. (2004). Biphasic effects of stress upon GLUT8 glucose transporter expression and trafficking in the diabetic rat hippocampus. *Brain Research* 1006(1), 28–35.
- Shepherd, E. J., Helliwell, P. A., Mace, O. J., Morgan, E. L., Patel, N. and Kellett, G. L. (2004). Stress and glucocorticoid inhibit apical GLUT2-trafficking and intestinal glucose absorption in rat small intestine. *Journal of Physiology* 560(Pt 1), 281–290.
- Simpson, I., Chundu, K., Davies-Hill, T., Honer, W. and Davies, P. (1994). Decreased concentrations of GLUT1 and GLUT3 glucose transporters in the brains of patients with Alzheimer's disease. *Annals of Neurology* 35, 546–561.
- Stein, I., Neeman, M., Shweiki, D., Itin, A. and Keshet, E. (1995). Stabilization of vascular endothelial growth factor mRNA by hypoxia and hypoglycemia and coregulation with other ischemia-induced genes. *Molecular and Cellular Biology* 15(10), 5363–5368.

- Taha, C., Tsakiridis, T., McCall, A. and Klip, A. (1997). Stress-activated pathways in the regulation of glucose transporters. *American Journal of Physiology Endocrinology and Metabolism* 273, 68–76.
- Uehara, Y., Nipper, V. and McCall, A. (1997). Chronic insulin hypoglycemia induces GLUT-3 protein in rat

- brain neurons. *American Journal of Physiology Endocrinology and Metabolism* 272, E716–E719.
- Weber, T., Joost, H., Kuroda, M., Cushman, S. and Simpson, I. (1991). Subcellular distribution and phosphorylation state of insulin receptors from insulin- and isoproterenol-treated rat adipose cells. *Cell Signaling* 3, 51–58.

Glutamate See: Glucocorticoids – Adverse Effects on the Nervous System.

Glycobiology of Stress

G Lauc and M Flögel

University of Zagreb, Zagreb, Croatia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G Lauc and M Flögel, volume 2, pp 276–282, © 2000, Elsevier Inc.

Glycoconjugates Are Important Mediators of Many Physiological Functions

Glycoproteins in Stress

Lectins in Stress

Stress Affects the Activity of Glycosyltransferases

Model of Glycobiology of Stress

Glossary

<i>Glyco-conjugate</i>	A compound composed of an oligosaccharide linked to a protein or a lipid.
<i>Glycolipid</i>	A compound containing both lipid and oligosaccharide moieties.
<i>Glycoprotein</i>	A protein molecule containing one or more oligosaccharide attachments.
<i>Glycosylation</i>	The posttranslational modification of a protein by the addition of a carbohydrate moiety; catalyzed by specific enzymes called glycosyltransferases.
<i>Glycosyl-transferase (GT)</i>	A group of enzymes that transfer monosaccharides to a growing oligosaccharide chain on glycoconjugates.
<i>Lectin</i>	A physiological receptor for carbohydrate structures attached to glycoconjugates.

Sialyltransferase

An enzyme that transfers sialic acid to a glycoconjugate.

The metabolic response to psychological stress is a very complex and demanding physiological process that involves numerous organs and organ systems. Although it is highly important for survival in an ever-changing environment, its excessive activation is associated with various detrimental effects. A number of epidemiological and experimental studies conducted during the past years have clearly demonstrated a link between stress and the development and course of many diseases, from simple virus infections and gastric ulcers to cardiovascular diseases and cancer. It is estimated that up to two-thirds of all visits to the physician's office are associated with stress.

Molecular mechanisms underlying the link between the response to stress and the development of disease appear to be exceedingly complex and are only partly understood. Although hormonal changes are key mediators of the physiological changes in stress, other factors appear to be decisive in the development of stress-associated disorders. Molecular response of a target cell appears to be defined not only by the incoming hormone but also by the current molecular situation within the cell itself. A good illustration of this principle is the way in which corticosteroids affect neurons in the hippocampus. If neurons are at rest, corticosteroids apparently have no effects. These effects become visible only when neurons are shifted from their basal condition by the action of neurotransmitters.

After more than 50 years of research, a lot is known about the endocrinology of the stress response, but the key molecular mechanism that explains how and why corticosteroids and other stress hormones cease to be beneficial and start to cause damage is still not known. One important link between stress and disease is the stress-induced decrease in the immune response, but the decrease in the immune response cannot explain all the effects of stress and other mechanisms have to be involved. Glycosylation appears to be involved, and the modification of glycoconjugate structures and their interaction with lectin receptors modulate at least some of these processes.

Glycoconjugates Are Important Mediators of Many Physiological Functions

Glycosylation is a very complex posttranslational modification that plays an important role in the integration of higher organisms. The surface of all cells is covered by oligosaccharide structures covalently attached to proteins and lipids (glycoproteins and glycolipids). Escalation in the extent and diversity of protein glycosylation correlates with the appearance of multicellular life, and apparently most of the interaction between cells and their surroundings include carbohydrate–protein recognition. An essential difference between proteins and the oligosaccharide structures attached to them is the fact that the carbohydrate part is not encoded in the DNA (Figure 1). Although genes unequivocally define the structure of each protein, there is no template for attached oligosaccharides. This makes them inherently sensitive to all changes within the cell, and glycan structures that are being produced at any individual moment actually mirror all relevant past events in the cell. In addition, the evolution of carbohydrate structures is much faster than the evolution of proteins, and often in different species we see significant differences in carbohydrate structures attached to nearly identical protein or lipid backbones.

Glycan parts of glycoconjugates are being synthesized by GTs that act in a sequence, and each

enzyme transfers specific monosaccharide to a specific acceptor contributing to the final glycan structure (Figure 2). It is estimated that over 300 specific enzymes (glycosidases and GTs) are involved in the synthesis of carbohydrate structures on glycoconjugates, and the whole process is very complex and energy expensive for the cell. Defects that disturb early phases of glycoconjugate synthesis are incompatible with multicellular life, whereas mutations in the end modifications of oligosaccharides result in specific physiological alterations. Because oligosaccharide structures are not encoded in genes, the final structure of a glycan is determined by the current cellular repertoire of expressed GTs (the expression, intracellular localization, and regulation of specific GTs that is commonly referred to as the glycosylation phenotype).

At the level of the individual molecule, from the structural point of view, there is no significant difference between the protein and carbohydrate parts. Glycoprotein is an entity composed of a protein and a carbohydrate part, and both parts have structural and functional roles that are specific to a given protein. Due to very high structural variability and the lack of adequate methods to analyze them, the carbohydrate parts of glycoconjugates were ignored by the most of the scientific community until very recently. However, thanks to the development of novel methods, in the last decade there has been an exponential increase in the knowledge about the physiological roles of carbohydrate structures. Now it is known that glycan part of glycoconjugates are crucial for many physiological processes from fertilization and development to regulation of hormonal activity and the formation of memory. Recent evidence has established differential glycan processing as a novel regulator of protein function, and the speculation that the invention of glycosylation was the key evolutionary step that enabled the development of multicellular organisms appears to be correct.



Figure 1 Central dogma of molecular biology amended with the biosynthesis of glycoproteins. Contrary to proteins, which all have their templates in the DNA, the carbohydrate structures attached to glycoproteins are not defined by genes. Their exact structures are determined by the regulation of expression, intracellular localization, and activity of various glycosyltransferases.

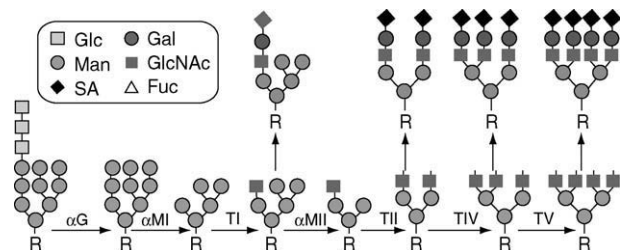


Figure 2 Schematic representation of processing of N-linked glycans. N-linked glycans are being synthesized by sequential action of numerous glycosyltransferases and glycosidases that add or remove specific monosaccharides to a growing oligosaccharide chain.

Glycoproteins in Stress

The influence of stress on glycosylation was first studied on gastric mucosa in rats. Mucosal cells of rats stressed by immersion in water were found to incorporate up to 50% less N-acetylgalactosamine than control animals. Subsequently, it was also shown that stress is associated with changes in the binding of lectins to gastric mucosa. Stress-induced alterations were found in the binding of several lectins, the most specific being changes detected with peanut agglutinin (PNA) lectin that specifically recognizes galactose $\beta(1,3)$ linked to N-acetylgalactosamine. These studies were not continued, but they undoubtedly indicated that stress does have some influence on glycosylation. *Helicobacter pylori* (causative agent of most peptic ulcers) attaches to the gastric mucosa through specific interactions with carbohydrate structures, and the stress-induced changes in the expression of these structures might be one of the molecular mechanisms that explains the effects of stress on the development of gastric ulcers.

An indirect confirmation of the hypothesis that structures of glycoconjugates might change in stress came from the studies of acute-phase proteins in depression. For some time, it is known that depression can significantly influence the immune system, especially some positive acute-phase proteins such as haptoglobin or α_1 -acid glycoprotein. Similar changes were detected in association with surgical stress, and recently it was shown that even a single episode of uncontrollable stress can activate acute-phase response in the experimental animals. Nearly all acute-phase proteins are glycosylated, and in addition to changes in the concentration of the whole proteins, depression is also associated with changes in the individual carbohydrate structures attached to acute-phase proteins. Specific changes in carbohydrate structures on glycoproteins also occur during inflammation and some diseases that are associated with stress. In their recent work, van Dijk and colleagues proposed that these changes might be associated with the regulation of the immune response.

The first indications that psychological stress can influence human glycoconjugates came from the studies of Flögel and colleagues in early 1990s. They used lectin-western blot to analyze changes in the glycosylation of proteins in sera of prisoners released from the Serbian concentration camps during the war in Croatia and Bosnia. Significant changes in glycosylation patterns were established with several different lectins, each recognizing specific segments of the oligosaccharide structure (Table 1).

Corticotropin releasing hormone (CRH) plays a crucial role in integrating the body's overall response

Table 1 Specificity of lectins^a

Lectin	Source	Specificity
GNA	<i>Galanthus nivalis</i>	Man- $\alpha(1,3)$, Man- $\alpha(1,6)$; Man- $\alpha(1,2)$ -Man
SNA	<i>Sambucus nigra</i>	Sia- $\alpha(2,6)$ -Gal
MAA	<i>Maackia amurensis</i>	Sia- $\alpha(2,3)$ -Gal
PNA	Peanut	Gal- $\beta(1,3)$ -GalNAc
DSA	<i>Datura stramonium</i>	Gal- $\beta(1,4)$ -GlcNAc

^aLectins are physiological receptors for carbohydrate structures in animals and plants. Plant lectins are easy to isolate and they are commercially available as tools for specific recognition of the exactly defined segments of carbohydrate structures.

to stress, including the integration of the response of the immune system to physiological, psychological, and immunological stressors. In addition to the hypothalamus, where it was initially identified, CRH, its receptors, and related ligands are present throughout the brain and periphery. By activating the glucocorticoid and catecholamine secretion, CRH from the central nervous system (CNS) mediates the suppressive effects of stress on the immune system. The action of CRH is exhibited through CRH receptors CRH1 and CRH2, which are differentially expressed on brain neurons located in neocortical, limbic, and brainstem regions of the CNS and on the pituitary corticotrophs. An intriguing aspect of both CRH receptors and CRH binding protein is their glycosylation pattern. CRH1 contains five potential N-glycosylation sites. Early studies showed that its glycosylation differs in different regions of the CNS, indicating potential functional roles for different glycoforms. The conversion of oligosaccharide chains to oligomannose structures by kifunensine (an inhibitor of mannosidase I that prevents the conversion of oligomannose to complex glycans) was reported not to change the ligand-binding properties of individual receptors, but mutation experiments have shown that the presence of at least three (out of five) oligosaccharide chains is required for normal CRH1 function.

Another system in which glycosylation might play an important role in the stress response is the cholinergic activation of the brain. It is generally accepted that in the CNS stress induces primarily cholinergic hyperactivation. Within several hours, the excess acetylcholinesterase (AChE) exerts a protective effect by retrieving cholinergic balance through enzymatic acetylcholine (ACh) hydrolysis. In addition to hydrolysis of ACh at the brain cholinergic synapses and neuromuscular junctions, AChE also affects cell proliferation, differentiation, and responses to various insults. Through alternative splicing of its 3' end, AChE is considered to be one of the general stress-responding proteins, both in the brain and in the periphery. AChE

is glycosylphosphatidylinositol (GPI)-anchored protein, and efficient glypiation is necessary for proper localization and secretion. For more than 10 years, it has been known that differences in glycosylation affect the stability of the enzyme, but the functional consequences of altered glycosylation are still not understood. Pharmacokinetic profiling of the AChE glycoforms demonstrated a correlation between circulatory longevity and the number of attached N-glycans. Glycosylation of AChE was shown to be altered in breast cancer, but this was not studied in detail. Abnormally glycosylated forms of the enzyme also accumulate in the cerebrospinal fluid in some neurological disorders, and changes in AChE glycosylation were reported to be a potential diagnostic marker for Alzheimer's disease.

Contraception, which is a form of hormonal stress for the body, also affects glycosylation. Oral estrogen treatment induces an increase in the degree of branching and a decrease in fucosylation and sialyl Lewis x expression on α_1 -acid glycoprotein compared to individuals receiving no estrogens or transdermal estrogen treatment. The effects of oral estrogens are identical in both males and females, and they can be reduced by the administration of progestagen. The effects of oral estrogens on glycosylation of α_1 -acid glycoprotein are the opposite of those induced by inflammation, indicating that estrogens can modulate the glycosylation-dependent inflammatory actions of α_1 -acid glycoprotein. Interestingly, estrogens that are applied transdermally do not exert this effect on hepatic glycosylation.

Lectins in Stress

Some glycans on glycoconjugates have only structural roles, but some also take part in specific recognition processes. One of the major mechanisms by which glycoconjugates perform their molecular functions is the interaction with their specific molecular receptors, lectins. Hundreds of endogenous lectins function as physiological receptors for oligosaccharides that interpret molecular information encoded in glycans. They take part in numerous physiological processes including folding, intracellular transport, fertilization, regulation of the inflammatory response, and brain plasticity.

The best-known example of lectin function is inflammation, in which the interaction between lectins called selectins and carbohydrate structures on glycoproteins represents the first decisive step leading to the adhesion of circulating lymphocytes to the endothelium at the site of inflammation. Carbohydrate-lectin interactions are very interesting because they can be inhibited with small oligosaccharides that

are generally nontoxic and, using modern carbohydrate cycling technologies, are relatively inexpensive to prepare for therapeutic purposes.

The first indications that lectins might change in stress came from a study of a phenomenon only remotely associated with stress. Bardosi and colleagues were analyzing the influence of prolonged anesthesia on mannose receptor and some other lectins on murine peripheral blood polymorphonuclear leukocytes. They demonstrated reduced expression of mannose receptor and changes in the expression of receptors for several other sugars, indicating that prolonged anesthesia affects the regulation of lectin expression.

Evidence that psychological stress can influence lectins came from studies of Lauc and colleagues on lectins in livers of rats exposed to immobilization stress. Immobilization stress was found to influence galectin-3, a galactose-specific lectin, so that it binds to another nuclear lectin CBP67 (67-kDa carbohydrate binding protein). The formation of the complex from CBP67 and galectin-3 resulted in the binding of the galactose-specific galectin-3 to the glucose-affinity column, for which it shows no affinity under normal conditions. This is a very interesting effect because galectin-3 is involved in mRNA splicing and changes in its functions might have profound effects on the whole cell. However, whether this binding was mediated by protein-protein interactions or through lectinlike binding of galectin-3 to galactose residues on carbohydrate structures attached to CBP67 (as well as other details of this interaction) is not known.

Galectin-3 is a versatile galactoside-binding lectin that has been implicated in numerous cellular functions. It also appears to be affected by stress and, interestingly, different types of stress have exactly opposite effects on its expression. As shown by Dumić and colleagues in 2000, whereas exposure to UV light or transfer to *in vitro* conditions induces galectin-3 in cultured cells, immobilization stress *in vivo* results in a decrease in galectin-3 in mouse spleen and liver. The regulation of galectin-3 expression involves transcription factor nuclear factor (NF)- κ B, which connects galectin-3 with corticosteroids (and CRH) and places its expression downstream from hormonal signals in the stress-response pathway.

A major problem in studying changes of lectin activity in different physiological processes is the lack of adequate methods for measuring lectin activity in complex biological samples. A new method (Figure 3) using photoaffinity glycoprobes labeled with digoxin was recently developed, and it is hoped that it will enable easier identification of changes in lectins in different diseases. Until now, only one new lectin was found to appear in stress. It was found

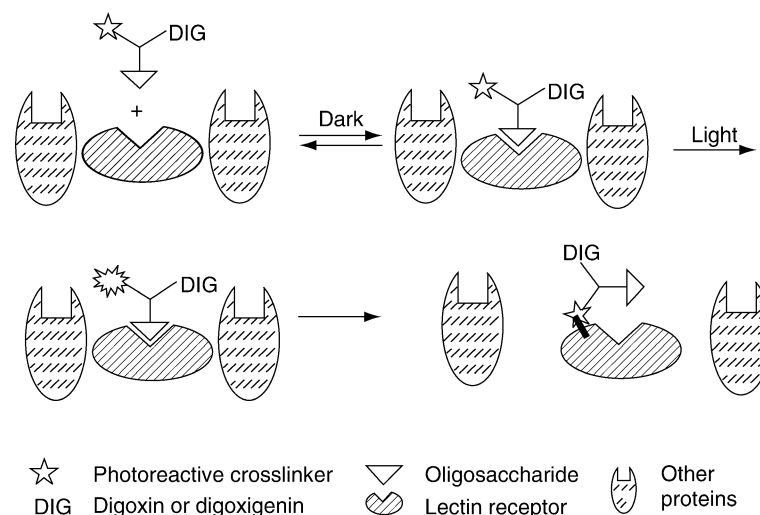


Figure 3 Photoaffinity method for the detection of lectins. Photoaffinity glycoprobes containing target carbohydrate structures are incubated with biological samples to allow the formation of noncovalent complexes between the probe and lectin receptors in the sample. Illumination activates the photoreactive cross-linker, and it forms covalent bonds with neighboring molecules, mostly lectin receptors. The result is a lectin with a covalently incorporated digoxin tag (DIG) that can be easily identified with labeled antibodies against digoxin.

to bind to glucose-containing glycoprobes, and according to the electrophoretic mobility of the protein isolated from rat liver named CBP33.

Stress Affects the Activity of Glycosyltransferases

The carbohydrate parts of glycoconjugates are synthesized by sequential action of numerous GTs, enzymes that are specific for both the structure of the glycoconjugate acceptor and the monosaccharide that is being added. The appearance of novel or altered glycoconjugate structures in stress indicates that stress should somehow affect the activity of GTs, but the only GTs whose relation with stress have been studied up to now are the sialyltransferases (enzymes that transfer sialic acids to glycoproteins).

For some time it has been known that corticosteroids can change the activity of sialyltransferases *in vitro*. Breen and colleagues have shown that corticosteroids and other hormones from the adrenal gland significantly influence the activity of sialyltransferases *in vivo*. Adrenalectomy and the subsequent administration of corticosterone and/or aldosterone significantly influence the activity of sialyltransferases in various rat tissues. Whereas sialyltransferases in some tissues such as the kidney are apparently not influenced by adrenalectomy or by the addition of steroid hormones, sialyltransferases in the liver are under the negative control of corticosteroids. Adrenalectomy results in the increased activity of sialyltransferases in the liver that cannot be reverted to normal values by the administration of dexamethasone.

Enzymes that perform the same function in the brain react to same hormonal signals in the exactly opposite way. Adrenalectomy and the consequential lack of circulating corticosteroids lead to the decrease of total sialyltransferase activity in the brain. The subsequent administration of exogenous corticosteroids exhibits regional specificity, with the enzyme activities in the cortex, cerebellum, and brain stem being stimulated by both dexamethasone and aldosterone and enzyme activity in the hippocampus being stimulated only by aldosterone. Total sialyltransferase activity in some tissue does not represent the activity of a single enzyme but the sum of the activities of all enzymes that transfer sialic acids. Breen and colleagues also studied the effects of corticosteroids on two individual enzymes, $\alpha(2,3)$ -sialyltransferase and $\alpha(2,6)$ -sialyltransferase. Both enzymes transfer sialic acids, but they link them to different carbon atoms on the preceding sugar in the carbohydrate structure.

As shown by Dabelic and colleagues in 2004, immobilization stress affects sialyltransferase activity in different rat tissues. Acute and chronic stress have different effects, but, even more interesting, the same type of stress has opposite effects on sialyltransferase activity in different tissues. In general, stress induces sialyltransferase activity in extraneural tissues and suppresses the activity of the same enzymes in most brain tissues. The fact that same enzymes respond differently to the same hormonal signals depending on their cellular environment exemplifies the fact that the molecular setup of the targeted cell, and not hormonal signal by itself, is the decisive player in

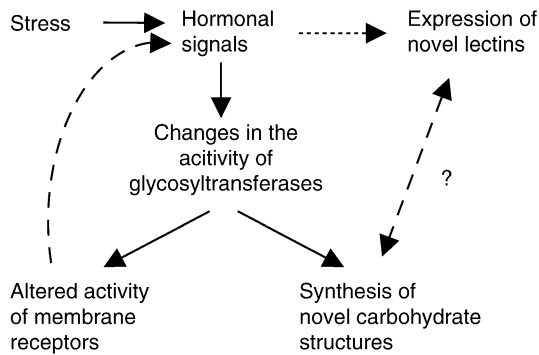


Figure 4 Hypothetical model of the glycobiology of stress. Hormones mediate numerous stress-associated changes, and among them alter the activity of different glycosyltransferases. Changed activity of glycosyltransferases results in the appearance of different carbohydrate structures on glycoproteins, which could either present a novel structures with potential to interact with specific endogenous lectins or modify the activity of various membrane receptors. Although this has yet to be proven, it could be one of the mechanisms explaining why prolonged stress has different effects than acute stress. In the same time, it is known that stress is associated with changes in composition and activity of endogenous lectins, but the exact mechanisms linking stress and changes in lectin activity are still not known.

determining the direction and consequences of the stress response.

Model of Glycobiology of Stress

At the moment only several fragments of the glyco-biological mechanisms involved in the physiological response to psychological stress are known, but the complete picture is slowly emerging (Figure 4). Stress causes numerous changes in the circulating hormones, and many molecular details of this process are known. It is also known that corticosteroids affect the activity of at least one GT both *in vitro* and *in vivo*. The altered activity of GTs results in different carbohydrate structures attached to glycoproteins, and these changes have been demonstrated both in humans and in experimental animals. A change in the carbohydrate structures attached to a glycoprotein is a well-established way to change its structural and functional properties, and recently this was shown

to be one of the mechanisms that control the activity of membrane receptors. Although this type of glycosylation-mediated receptor modulation in stress still has to be proven, it is a very interesting hypothesis. On the other hand, new glycoconjugate structures could also represent novel signals on the cell surface that could alter the interaction of the cell with neighboring cells in a process analogous to the selectin-mediated adhesion of lymphocytes. Stress is also known to be associated with the appearance of novel lectins, but the exact mechanism of this process is not known. These lectins could be receptors for either novel or normal glycoconjugate structures, translating their structures into molecular functions. Although most of this is still speculative, it is hoped that more will soon be known about the molecular role of glycoconjugates, their lectin receptors, and GTs in the physiological response to psychological stress.

Further Reading

- Assil, I. Q. and Abou-Samra, A. B. (2001). N-glycosylation of CRF receptor type 1 is important for its ligand-specific interaction. *American Journal of Physiology, Endocrinology and Metabolism* **281**, E1015–E1021.
- Axford, J. (2001). The impact of glycobiology on medicine. *Trends in Immunology* **22**, 237–239.
- Chitlaru, T., Kronman, C., Velan, A., et al. (2002). Overloading and removal of N-glycosylation targets on human acetylcholinesterase: effects on glycan composition and circulatory residence time. *Biochemical Journal* **363**, 619–631.
- Dabelic, S., Flogel, M., Maravic, G., et al. (2004). Stress causes tissue-specific changes in the sialyltransferase activity. *Zeitschrift für Naturforschung* **59**, 276–280.
- Lauc, G., Supraha, S., Lee, Y. C., et al. (2003). Digoxin derivatives as tools for glycobiology. *Methods in Enzymology* **362**, 29–37.
- Nalivaeva, N. N. and Turner, A. J. (2001). Post-translational modifications of proteins: acetylcholinesterase as a model system. *Proteomics* **1**, 735–747.
- Partridge, E. A., Le Roy, C., Di Guglielmo, G. M., et al. (2004). Regulation of cytokine receptors by golgi N-glycan processing and endocytosis. *Science* **306**, 120–124.

Gonadotropin Secretion, Effects of Stress on

M Ferin

Columbia University, New York NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M Ferin, volume 2, pp 283–288,

© 2000, Elsevier Inc.

Control of Normal Gonadotropin Secretion

The Gonadotropin Response to Stressors: Cross-Talk between the Hypothalamic-Pituitary-Adrenal and the Hypothalamic-Pituitary-Gonadal Axes

Central Pathways Mediating the Stress Response

Stressors and the Reproductive Cycle

Glossary

<i>Amenorrhea</i>	The absence of menstruation, denoting a silent hypothalamic-pituitary-ovarian endocrine axis.
<i>Arginine vasopressin (AVP)</i>	An adjunct neurohormone in control of the hypothalamic-pituitary-adrenal endocrine axis. It is colocated with and released together with CRH in response to stressors.
<i>Corticotropin releasing hormone (CRH)</i>	The main neurohormone in control of the hypothalamic-pituitary-adrenal endocrine axis. Its release from the hypothalamic paraventricular nucleus is activated by stressors. It stimulates the pituitary to secrete adrenocorticotrophic hormone.
<i>Estradiol</i>	The main sex steroid secreted by the ovarian follicle.
<i>Follicular phase</i>	The first stage of the menstrual cycle, during which the process of folliculogenesis is completed through the maturation of the soon-to-be-ovulated dominant follicle.
<i>Follicle stimulating hormone (FSH)</i>	A gonadotropin that promotes the recruitment of a new cohort of follicles at the start of each follicular phase.
<i>Gonadotropin releasing hormone (GnRH)</i>	The hypothalamic structure responsible for the pulsatile release of GnRH, in a process that is essential for the normal function of the reproductive process.
<i>pulse generator</i>	Two hormones, follicle stimulating hormone and luteinizing hormone, released by the anterior pituitary following stimulation by GnRH.
<i>Gonadotropin</i>	
<i>Luteal phase</i>	The second stage of the menstrual cycle, when the ovulated follicle is transformed into a new structure, the corpus luteum, which secretes both estradiol and progesterone.

Luteinizing hormone (LH)

A gonadotropin that stimulates the release of estradiol from the maturing follicle and of progesterone from the corpus luteum.

Progesterone

The main sex steroid secreted by the corpus luteum. It prepares the uterus for the implantation of the fertilized egg.

Proopiomelanocortin (POMC)

The common precursor molecule to β -endorphin (an endogenous opioid peptide) and to α -melanocyte stimulating hormone. Both peptides modulate the gonadotropin response to stressors.

Control of Normal Gonadotropin Secretion

Hypothalamic Control: The Gonadotropin Releasing Hormone Pulse Generator

The hypothalamic-pituitary-gonadal (HPG) axis, which governs the reproductive system, is driven by the brain. Gonadotropin (FSH and LH) secretion by the anterior pituitary gland is controlled by the hypothalamus through the release of GnRH. GnRH, a 10-amino-acid peptide, is synthesized within several hypothalamic nuclei and released into the hypophyseal portal circulation to stimulate the release of both gonadotropins. This direct vascular connection between the hypothalamus and the anterior pituitary allows for the rapid and undiluted transport of minute amounts of GnRH to its target, the pituitary gonadotroph. Gonadotropins, in turn, act on the gonads, the ovaries or testes, to induce specific morphological changes and to stimulate the release of gonadal steroids, such as testosterone, estradiol, and progesterone. Constant communication between the several levels of the HPG axis through the negative feedback loop is essential for normal function; estradiol or testosterone secreted by the ovaries or testes continuously feeds back to the hypothalamus and pituitary to adjust GnRH, LH, and FSH secretion. The interruption of this negative feedback loop, for instance, following castration, results in an increase of GnRH, FSH, and LH release.

Significantly, GnRH is released from the hypothalamus in a pulsatile fashion (the GnRH pulse generator), and thus tonic secretion of gonadotropins is pulsatile in nature. Each pulse of LH consists of an abrupt and brief release of the hormone followed by an exponential decrease according to its half-life. A pulsatile GnRH stimulus is required to increase the transcription of the gonadotropin subunit genes.

Pathological events that interfere with the function of the GnRH pulse generator disrupt the pituitary-gonadal axis. For example, the complete absence of GnRH pulsatility results in a dormant reproductive axis, whereas lesser abnormalities, such as a decrease in GnRH pulse frequency, may interfere with proper reproductive function.

Gonadotropins and the Reproductive Cycle

The reproductive cycle, estrous cycle (in rodent and sheep) or menstrual cycle (in nonhuman primate and human) involves a remarkable coordination of hormonal secretion and morphological changes within the HPG axis. In the primate, a rise in FSH at the start of the follicular phase promotes the recruitment of a cohort of antral follicles from which the dominant follicle destined for ovulation will be selected. Stimulated by the pulsatile release of the gonadotropins, the dominant follicle grows exponentially and releases estradiol in amounts proportional to its size. The maturity of the dominant follicle is marked by a high estradiol release, which when a threshold is reached activates the positive estradiol feedback loop; this induces the ovulatory LH surge and ovulation. After ovulation, the dominant follicle is transformed into the corpus luteum, which secretes both estradiol and progesterone when stimulated by LH. Substantial progesterone secretion promotes uterine differentiation in preparation for implantation of the fertilized egg.

A proper pattern of GnRH and gonadotropin pulsatile release is required for the normal reproductive cycle to occur. In the human, for example, an optimal GnRH-LH pulse frequency of approximately one pulse every 90 min is required for normal folliculogenesis to occur during the follicular phase. A pulse frequency that is too low may fail to bring the recruited follicle to a timely maturation, and a too rapid pulse frequency may result in pathology (such as polycystic ovary syndrome).

The Gonadotropin Response to Stressors: Cross-Talk between the Hypothalamic-Pituitary-Adrenal and the Hypothalamic-Pituitary-Gonadal Axes

Effects of Stressors on Tonic Gonadotropin Release

The activation of the hypothalamic-pituitary-adrenal (HPA) axis by several types of stressors, such as immune/inflammatory challenges as simulated by the administration of endotoxin or cytokines, temporary restraint, or insulin-induced hypoglycemia, results in an inhibition of tonic pulsatile LH release, mainly

in the form of a decrease in LH pulse frequency. In women, psychogenic stress has also been related to lower basal tonic levels of FSH, LH, and estradiol compared to normal cycling women; these patients have reduced integrated 24-h LH and FSH concentrations, reduced LH pulse frequency, and increased LH interpulse interval.

Overall evidence suggests that this decrease in tonic pulsatile LH release is primarily related to a central action of stressors exerted through an inhibitory action on the hypothalamic GnRH pulse generator; a decline in the generator's ability to properly stimulate the gonadotroph, mainly through a reduction in the frequency of pulsatile GnRH release, appears to be the main mechanism involved. In addition, there is experimental evidence that stressors may also subtly influence the gonadotroph's responsiveness to GnRH and thereby decrease the amplitude of the gonadotropin response. This effect may reflect a pituitary action of adrenal corticosteroids released in response to the activation of the HPA axis by the stressor.

Effects of Stressors on the Surge Mode of Luteinizing Hormone Release

Data in the rodent and ewe also suggest that an immune/inflammatory challenge or restraint (and possibly other stressors) can also suppress the estradiol-induced ovulatory LH surge. In the rodent, this effect appears to be exerted at the level of GnRH perikarya, where the percentage of GnRH neurons expressing the c-fos protein and the GnRH primary transcript are significantly reduced by these stressors. The administration of a GnRH agonist concomitantly with the stressor completely restores the LH surge. In the ewe, endotoxin was shown to prevent the GnRH surge into the hypophyseal portal circulation, also suggesting a central effect of the challenge. Significantly, this inhibitory action is seen only if the stressor acts quite ahead of the expected time of the LH surge, primarily at the initial time of activation of the estradiol positive feedback signal.

Stressors and an Early Short-Term Rise in Gonadotropins

In addition to the previously mentioned inhibitory action of stressors on gonadotropin release, investigators have also reported a very transient rise in LH in rodents within the first hour of the response to various stressors, such as handling; restraint; exposure to ether or to a novel environment; or visual, audio-genic, and thermal stressors. This rapid and transient stimulatory effect of stressors on tonic LH release has been reported mostly in the intact, but generally not in the castrated, male rat, and in men but not women.

The mechanisms behind this phenomenon and its role in the stress response remain to be elucidated.

Central Pathways Mediating the Stress Response

The Hypothalamic Paraventricular Nucleus

The paraventricular nucleus (PVN) plays a pivotal role in the response to stressors. The neuroendocrine activation of the HPA axis by stressors involves the release of two neuropeptides, corticotropin releasing hormone (CRH) and AVP, from discrete populations of hypophysiotrophic neurons in parvocellular areas of PVN. These two neuropeptides, which are colocalized within perikarya of PVN neurons and in secretory granules, play a complementary role in mediating the effects of stressors on the HPA axis (see [Figure 1](#)).

There is good experimental evidence for an active mediatory role by CRH in the inhibition of GnRH and gonadotropins that follow stressors. In the non-human primate, the administration of a CRH receptor antagonist or of a drug decreasing endogenous CRH activity prevents not only the cortisol increase but also the inhibition of pulsatile GnRH–LH release induced by stressors such as an immune/inflammatory challenge ([Figure 2a](#)) or insulin-induced hypoglycemia. Similar mediatory effects of a CRH antagonist or of CRH antiserum have been reported in the rodent after various stressors, except in the response to an inflammatory challenge. The latter is in a marked difference with the primate. The administration of

CRH also results in a rapid inhibition of pulsatile LH release and in a decrease in the frequency and duration of GnRH pulses. Data in the nonhuman primate suggest that this inhibitory action is not mediated by downstream activation of the pituitary–adrenal axis: indeed, the inhibitory effects of a CRH infusion on pulsatile LH release cannot be mimicked by adrenocorticotrophic hormone (ACTH) infusion, are not reversed by treatment with metyrapone (an inhibitor of adrenal steroidogenesis), and persist following adrenalectomy.

Data in CRH-deficient male and female mice indicate that stressors such as restraint or food withdrawal can still inhibit tonic LH secretion or the proestrus LH surge, suggesting that molecules other than CRH may also play a role. Although there fewer data, AVP has also been shown to mediate the inhibitory action of certain stressors on the HPG axis and conceivably may take over in the CRH-deficient animal. In the ovx monkey, the inhibitory effects of a central infusion of interleukin (IL)-1 on pulsatile LH release are prevented by the co-administration of an AVP antagonist, even though the antagonist is without effect on the cortisol increase ([Figure 2b](#)). Some, but not all, investigators also report that the same antagonist prevents hypoglycemia-induced LH suppression in ovx and in male monkeys.

The Endogenous Opioid Peptides

There is substantial evidence to indicate that the inhibitory action of CRH and of most stressors on pulsatile GnRH–LH release occurs at suprapituitary sites that not only involve the neuroendocrine HPA axis but also mediation by endogenous opioid peptidergic pathways. Endogenous β -endorphin is known to be a powerful inhibitor of the GnRH pulse generator; for instance, during the luteal phase of the menstrual cycle of the human and nonhuman primate, there is a marked decreased in LH pulse frequency that parallels elevated endogenous β -endorphin activity and that is prevented by the administration of naloxone, a μ -opioid receptor antagonist. Hypothalamic β -endorphin activity is stimulated by CRH and by stressors. Naloxone administration also prevents the CRH- or AVP-mediated suppression of GnRH–gonadotropin release and restores pulsatile LH secretion ([Figure 3a](#)). This opioid antagonist also prevents the LH inhibition that follows several types of stressors, except for insulin-induced LH inhibition.

β -Endorphin derives from POMC, a large glycoprotein precursor molecule synthesized primarily in the hypothalamus and pituitary gland. Posttranslational processing of POMC is tissue-specific and in the hypothalamus yields not only β -endorphin but also α -melanocyte stimulating hormone (α -MSH).

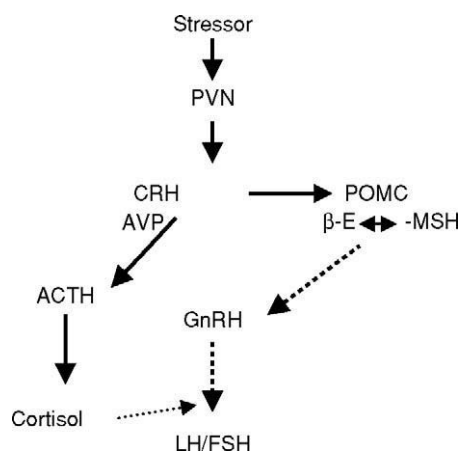


Figure 1 Main pathways mediating the inhibitory effects of stressors on pulsatile gonadotropin release. α -MSH, α -melanocyte stimulating hormone; β -E, β -endorphin; ACTH, adrenocorticotrophic hormone; AVP, arginine vasopressin; CRH, corticotropin releasing hormone; FSH, follicle stimulating hormone; GnRH, gonadotropin releasing hormone; LH, luteinizing hormone; POMC, proopiomelanocortin; PVN, paraventricular nucleus.

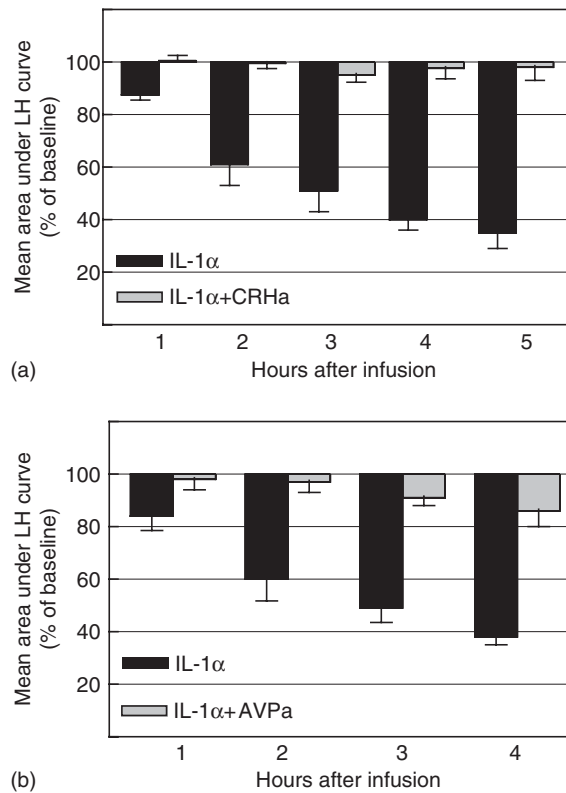


Figure 2 Reversal of the inhibitory effects of an intracerebroventricular infusion of interleukin (IL)-1 on luteinizing hormone (LH) release. a, By a CRH antagonist (CRHa); b, By a VP antagonist (AVPa). Reprinted with permission from (a) Feng, Y. J., Shalts, E., Xia, L., Rivier, J., Rivier, C., Vale, W., Ferin, M. (1991) An inhibitory effect of interleukin-1 on basal gonadotropin release in the ovariectomized rhesus monkey: reversal by a corticotropin-releasing factor antagonist, *Endocrinology* **128**, 2077–2082, Copyright 1991, The Endocrine Society; and (b) Shalts, E., Feng, Y. J., Ferin, M. (1992) Vasopressin mediates the interleukin-1 α -induced decrease in LH secretion in the ovariectomized rhesus monkey, *Endocrinology* **131**, 153–158, Copyright 1992, The Endocrine Society.

α -MSH has been shown to antagonize some of the actions of β -endorphin and may function as an endogenous opioid antagonist. In fact, α -MSH antagonizes the inhibitory action of exogenously administered β -endorphin on LH release and restores normal pulsatile LH release in ovx monkeys treated with CRH (Figure 3b) or subjected to an immune/inflammatory challenge. These data support the notion that post-translational processing of POMC may represent an important step in the response mechanisms to stressors and that the ratio of β -endorphin to α -MSH may be more significant than the absolute β -endorphin concentration in determining the ultimate effects of stressors on pulsatile LH release.

Stressors and the Reproductive Cycle

Stress and Amenorrhea

The reproductive cycle requires a remarkable coordination of hormonal secretion and morphological changes within all levels of the hypothalamic-pituitary-ovarian-uterine axis. Thus, it is not surprising

that stressors can disrupt the normal estrous or menstrual cycle. If sufficiently intense or chronic, stressors can completely interrupt the cycle. In women, psychogenic stress, usually the result of overt or latent psychological disturbances, is known to interfere with normal reproductive function and thought to be a common cause of amenorrhea in a syndrome known as functional hypothalamic chronic anovulation (FHCA). When fully established, the syndrome is characterized by ovarian quiescence, amenorrhea, and infertility. Several studies in the human underscore the association among FHCA, increased HPA activity in the form of an amplified circadian cortisol excursion, and low basal tonic levels of FSH and LH coupled with reduced LH pulse frequency.

Stress and Menstrual Cycle Defects

Although the long-term consequences of chronic stress may include amenorrhea and infertility and result in the full-blown FHCA syndrome, more recent data also demonstrate that stressors that initially are

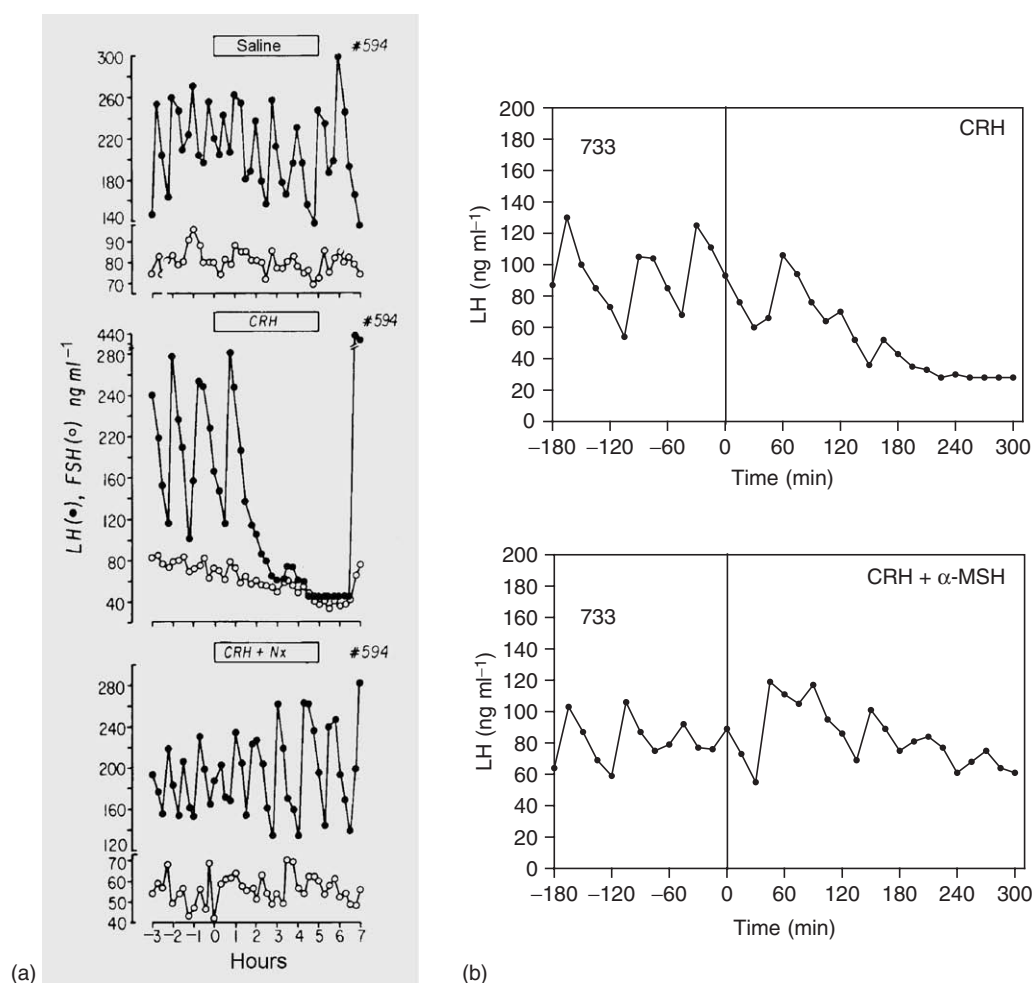


Figure 3 Reversal of the inhibitory effects of CRH on pulsatile LH release. a, By naloxone (Nx); b, By α -MSH. Reprinted with permission from (a) Gindoff, P. R., Ferin, M. (1987) Endogenous opioid peptides modulate the effect of corticotropin-releasing hormone (CRH) on gonadotropin release in the primate, *Endocrinology* **121**, 837–842, Copyright 1987, The Endocrine Society; and (b) Shalts, E., Feng, Y. J., Ferin, M., Wardlaw, S. L. (1992) α -Melanocyte-stimulating hormone antagonizes the neuroendocrine effects of corticotropin-releasing factor and interleukin-1a in the primate, *Endocrinology* **131**, 132–138, Copyright 1992, The Endocrine Society.

of an insufficient intensity to induce amenorrhea may nevertheless produce temporary cyclic defects.

Prolongation of the follicular phase As shown in the ewe and monkey, endotoxin administration, mimicking an inflammatory/immune challenge, or exercise during the follicular phase results in a lengthening of the follicular phase and in the interruption of the normal follicular estradiol profile. The stressor prevents the development of high-frequency LH pulses that are the essential stimulus for the preovulatory increase in estradiol secretion; without these, the estradiol rise is curtailed and the preovulatory LH surge is delayed. Significantly, in the monkey, the delay in the resumption of a normal follicular phase greatly exceeds the duration of the stress challenge itself.

Inadequate luteal phase In the monkey, a short-term inflammatory/immune challenge or a psychogenic-like stress, whether confined to the follicular or luteal phase of the menstrual cycle, invariably result in reduced luteal function, as demonstrated by overall diminished progesterone and inhibin A secretion during the luteal phase (Figure 4). In many animals, these short-term stress challenges continue to interfere with normal luteal function through the first or two subsequent stressor-free cycles. The lower progesterone levels are related to the decrease in LH induced by the stressor. Inadequate luteal phases have also been reported during exercise, suggesting that this syndrome may well represent a general initial response to stressors and/or energy-related demands in the normal cyclic female.

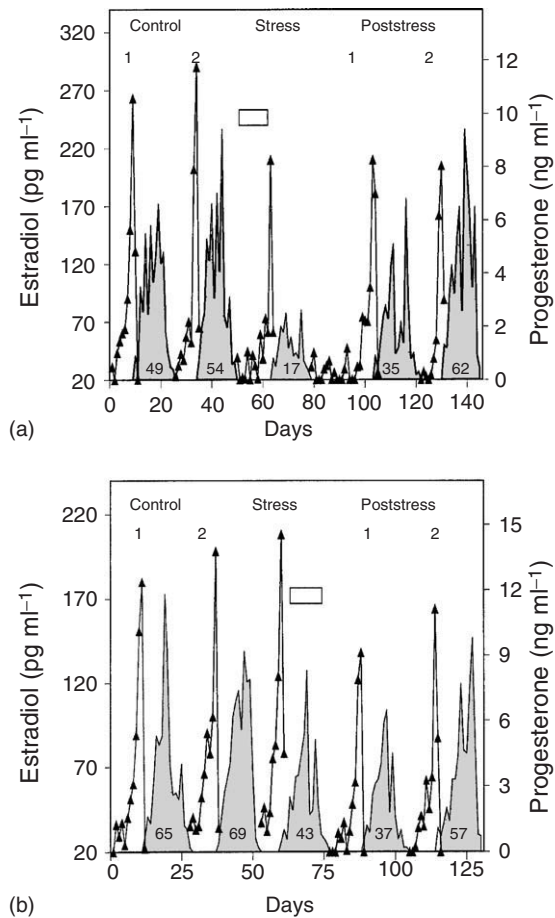


Figure 4 Inadequate luteal phase following a 10-day psychogenic-like stress. a, During the follicular phase; b, During the luteal phase. Illustrated are daily levels of estradiol during the follicular phase (triangles) and of progesterone during the luteal phase (shaded areas) throughout two control menstrual cycles, the stress cycle, and two poststress cycles. Luteal progesterone secretion is significantly decreased in the stress cycle and in the first of the two poststress cycles (numbers within the shaded areas are integrated luteal progesterone values). Reprinted with permission from Xiao, E., Xia-Zhang, L., Ferin, M. (2002) Inadequate luteal function is the initial clinical cyclic defect in a 12-day stress model that includes a psychogenic component in the Rhesus monkey, *Journal of Clinical Endocrinology & Metabolism* 87, 2232–2237, Copyright 2002, The Endocrine Society.

Premature ovulatory luteinizing hormone surge
Studies in the primate have shown that the activation of the HPA axis by endotoxin or IL-1 in the mid- to late follicular phase (but not at other stages of the cycle) results in a large release of LH. Data in the monkey suggest that progesterone, which is released in small amounts by the adrenals in response to the stressor, may be responsible for this acute release of LH because this effect is prevented by a progesterone or CRH antagonist. Smaller increases in LH have also been reported in women following exercise during the mid-follicular phase.

The precise physiopathological significance of these three cyclic defects in response to a stressor during the menstrual cycle remains to be further studied in patients. The data, however, suggest that the HPG axis is closely monitored by the brain and that a stress challenge, even one not sufficient or extended enough to interrupt the cycle and to result in amenorrhea, may well induce endocrine changes too subtle to be appreciated in the clinical environment but nonetheless substantial enough to temporarily impact on fertility or to provide a temporary suspension of pregnancy in suboptimal conditions. For instance, decreased luteal progesterone secretion in response to a stressor may interfere with the process of implantation of a fertilized egg and result in spontaneous abortion by delaying the endometrial changes required for proper implantation. Prematurely elevated LH concentrations in the follicular phase may damage the maturing follicle and/or oocyte and desynchronize the ovulatory signal from a timely follicular maturation process and thereby result in early pregnancy loss.

See Also the Following Article

Reproduction, Effects of Social Stress on.

Further Reading

- Battaglia, D. F., Brown, M. E., Krasa, H. B., et al. (1998). Systemic challenge with endotoxin stimulates corticotropin-releasing hormone and arginine vasopressin secretion into hypophyseal portal blood: coincidence with gonadotropin-releasing hormone suppression. *Endocrinology* 139, 4175–4181.
- Berga, S. L. (1996). Functional hypothalamic chronic anovulation. In: Adashi, E. Y., Rock, J. A. & Rosenwaks, Z. (eds.) *Reproductive endocrinology, surgery, and technology*, pp. 1061–1075. Philadelphia: Lippincott-Raven.
- Chrousos, G. P., Torpy, D. J. and Gold, P. W. (1998). Interactions between the hypothalamic-pituitary-adrenal axis and the female reproductive system: clinical implications. *Annals of Internal Medicine* 129, 229–240.
- Feng, Y. J., Shalts, E., Xia, L., et al. (1991). An inhibitory effect of interleukin-1 on basal gonadotropin release in the ovariectomized rhesus monkey: reversal by a corticotropin-releasing factor antagonist. *Endocrinology* 128, 2077–2082.
- Ferin, M. (1999). Clinical review 105: stress and the reproductive cycle. *Journal of Clinical Endocrinology & Metabolism* 84, 1768–1774.
- Ferin M. (2006). Stress and the reproductive system. In: Knobil, E. & Neill, J. (eds.) *The physiology of reproduction* (3rd edn., pp. 2627–2696). San Diego, CA: Elsevier.
- Gindoff, P. R. and Ferin, M. (1987). Endogenous opioid peptides modulate the effect of corticotropin-releasing hormone (CRH) on gonadotropin release in the primate. *Endocrinology* 121, 837–842.

- Karsch, F. J., Battaglia, D. F., Breen, K. M., et al. (2002). Mechanisms for ovarian cycle disruption by immune/inflammatory stress. *Stress* 5, 101–112.
- Rivest, S. and Rivier, C. (1995). The role of corticotropin-releasing factor and interleukin-1 in the regulation of neurons controlling reproductive functions. *Endocrine Review* 16, 177–199.
- Shalts, E., Feng, Y. J. and Ferin, M. (1992). Vasopressin mediates the interleukin-1 alpha-induced decrease in luteinizing hormone secretion in the ovariectomized rhesus monkey. *Endocrinology* 131, 153–158.
- Shalts, E., Feng, Y. J., Ferin, M., et al. (1992). α -Melanocyte-stimulating hormone antagonizes the neuroendocrine effects of corticotropin-releasing factor and interleukin-1 α in the primate. *Endocrinology* 131, 132–138.
- Xiao, E., Xia-Zhang, L., Barth, A., et al. (1998). Stress and the menstrual cycle: relevance of cycle quality in the short- and long-term response to a 5-day endotoxin challenge during the follicular phase in the rhesus monkey. *Journal of Clinical Endocrinology & Metabolism* 83, 2454–2460.
- Xiao, E., Xia-Zhang, L. and Ferin, M. (2002). Inadequate luteal function is the initial clinical cyclic defect in a 12-day stress model that includes a psychogenic component in the Rhesus monkey. *Journal of Clinical Endocrinology & Metabolism* 87, 2232–2237.
- Xiao, E., Xia-Zhang, L., Thornell, D., et al. (1996). Interleukin-1 stimulates luteinizing hormone release during the midfollicular phase in the rhesus monkey: a novel way in which stress may influence the menstrual cycle. *Journal of Clinical Endocrinology & Metabolism* 81, 2136–2141.

Graves' Disease (Thyrotoxicosis)

W M Wiersinga

University of Amsterdam, Amsterdam, Netherlands

© 2007 Elsevier Inc. All rights reserved.

Introduction

Historical Observations

Association between Stress and Graves' Disease

Stress and Graves' Disease: A Causal Relationship?

Glossary

<i>Autoimmunity</i>	Immune response directed at self antigens (belonging to the body of the individual), giving rise to antibodies against these antigens.
<i>Euthyroid</i>	Normal thyroid function.
<i>Graves' disease</i>	A multisystem autoimmune disease characterized by hyperthyroidism, enlarged thyroid gland (goiter), and frequently eye changes.
<i>Thyroid peroxidase (TPO)</i>	An enzyme in the thyroid gland involved in thyroid hormone synthesis.
<i>Thyrotoxicosis</i>	A disease caused by an excess of thyroid hormones; also called hyperthyroidism.
<i>Thyrotropin (TSH)</i>	The thyroid-stimulating hormone, which after release from the pituitary stimulates the thyroid gland to synthesize and release thyroid hormone.

Introduction

Thyrotoxicosis is a syndrome characterized by the clinical and biochemical manifestations of thyroid hormone excess in the target tissues of thyroid hormone. The most prominent clinical features of thyrotoxicosis are weight loss despite increased appetite, palpitations and tachycardia, fine tremor of the fingers, heat intolerance, perspiration, and a warm moist skin, all readily explained from an increased metabolic rate induced by thyroid hormone excess. Many patients also suffer from nervousness, irritation, anxiety, or mood disturbances, especially depression. Thyrotoxicosis is a very prevalent disease, occurring more frequently in women than in men. The annual incidence rate is approximately 0.4 per 1000 women and 0.1 per 1000 men.

Thyrotoxicosis can be caused by various diseases such as Graves' disease, toxic multinodular goiter, toxic adenoma, and destructive thyroiditis. Most prevalent, however, is Graves' disease, which is present in 75–90% of all thyrotoxic patients. Its relative frequency is dependent on ambient iodine intake, being lower in iodine-deficient areas. The peak incidence of Graves' disease is between ages 20 and 50 years. Graves' disease is an autoimmune disease characterized by antibodies directed against the thyrotropin receptor (TSH-R) on the surface of follicular epithelial cells in the thyroid gland; these antibodies

(called thyroid-stimulating immunoglobulins) after binding to the TSH-R stimulate the thyroid gland to synthesize and release thyroid hormone, thereby causing the clinical picture of thyrotoxicosis. The disease is referred to as Graves' hyperthyroidism (GH), a typical autoimmune thyroid disease. At the other end of the spectrum of autoimmune thyroid diseases is chronic lymphocytic thyroiditis, also called Hashimoto's disease, characterized by the presence of autoantibodies against thyroid peroxidase (TPO), an enzyme involved in thyroid hormone synthesis. Hashimoto's disease frequently results in hypothyroidism. TPO antibodies are also found in the majority of patients with Graves' hyperthyroidism.

Autoimmune thyroid disease frequently runs in families. The autoimmune reaction against thyroid antigens is thought to develop against a certain genetic background, provoked by environmental factors. Recent twin studies have indicated that genes are responsible for approximately 80% of the risk to get Graves' disease. Among these genes are HLA-DR3, CTLA-4, and PTPN-22, but their contribution is relatively small and most of the genes involved are still unidentified. It follows that 20% of the risk on Graves' disease is derived from the environment. Smoking is a clear risk factor for Graves' hyperthyroidism (especially for the eye changes in Graves' disease) but not for Hashimoto's hypothyroidism; smoking might even protect against the occurrence of TPO antibodies. Ever since the early descriptions of Graves' disease, stress has been implicated in its pathogenesis.

Historical Observations

In August 1803, Elizabeth S., a 21-year-old patient of Dr. Parry, became acutely frightened when she lost control of her wheelchair and careered downhill. Two weeks later she developed swelling of the thyroid gland, palpitations of the heart, nervous troubles, marked pulsation of the carotid arteries, and a small hard regular pulse of 96. Since Dr. Parry's description, the relation between stressful life events and the precipitation of clinical Graves' disease has been underlined by many other physicians, including Graves and von Basedow. In humans, a sudden fright may cause acute exophthalmos and thyroid enlargement with pulsating thrills (*Schreckbasedow*). The French refer to the thyroid as *la glande d'emotion*. An increase in the incidence of Graves' disease has been observed during every major war since the Franco-Prussian conflict, and the German term *Kriegsbasedow* reflects this association. Indeed, hospital admissions for Graves' disease increased fivefold in occupied Scandinavian countries during

World War II, returning to normal levels after liberation. Many textbooks in the second half of the twentieth century thus conclude that a period of severe emotional stress almost always precedes the onset of Graves' disease. The occurrence of Graves' hyperthyroidism in President George Bush, Sr., shortly after the Gulf War in 1991 apparently underlines this conclusion. The conclusion, however, is not supported by all studies. Environmental stress caused by the civil unrest in Northern Ireland was not associated with a higher incidence of thyrotoxicosis. Moreover, defining and quantifying stress is difficult. The death of a spouse is generally viewed as a stressful life event, highly distressing for the survivor after a long-standing marriage of mutual love and caring. But the death of an abusive wife-beating husband might come as a welcome relief to his widow. In addition, we may ask: which came first, the stress or Graves' disease? Increased anxiety and hyperreactivity to stressors might be an early manifestation of Graves' hyperthyroidism, preceding other symptoms and signs of the disease by months.

Association between Stress and Graves' Disease

Since the early 1990s there have been six case-control studies on the relation between stress and the onset of Graves' hyperthyroidism (see [Table 1](#)). The cases are well-documented patients with Graves' hyperthyroidism and were compared with controls who in general are reasonably well selected at random from a healthy population and matched for sex and age, although not always for marital and social status. Stress evaluation is standardized, applying either interviews or questionnaires, asking for stressful life events in the year prior to the diagnosis or prior to the first symptoms and signs of Graves' hyperthyroidism. The results have always been obtained in a retrospective manner, making all studies liable to recall bias. It has been argued that memory is differentially affected by mood, and, in particular, that negative events or memories are better recalled in the presence of a negative affect. Because hyperthyroid patients frequently have depressive symptoms, their negative affect could explain the increased reporting of negative life events in comparison with healthy controls. To avoid the effect of hyperthyroidism, some studies have chosen to perform the stress evaluation after the patient's thyroid function has been restored to normal levels; in doing so, the time interval for memorizing past events becomes longer, which by itself may increase recall bias. Evaluating stressful life events in the year prior to the first occurrence of hyperthyroid symptoms and signs is preferable to evaluating them

Table 1 Case-control studies on the relationship between stress and the onset of Graves' hyperthyroidism

Author (year)	Cases ^a		Controls ^b		Stress evaluation			Relation stress and Graves
	n	%	Matching criteria	Origin	Method ^c	Period	When done	
Winsa et al. (1991)	208	82	A, S	General population	1	1 year prior to diagnosis	Hyperthyroid; <2 wk after diagnosis	Yes
Sonino et al. (1993)	70	83	A, S, M, C	Advertisement, hospital staff	2	1 year prior to first signs	Euthyroid; 10 months (3–39) after first sign	Yes
Kung (1995)	95	84	A, S, M, C	Volunteers in community	1	1 year prior to diagnosis	Hyperthyroid; at referral	Yes
Radosavljevic et al. (1996)	100	93	A, S	Outpatients with other conditions	2	1 year prior to diagnosis	Not stated	Yes
Yoshiuchi et al. (1998)	228	80	A, S	Healthy controls from database	1	1 year prior to diagnosis	<1 month after diagnosis	Yes
Matos-Santos et al. (2001)	31	71	A, S, M, C	Preventive medicine appointments	1	1 year prior to first symptoms	Euthyroid after treatment	Yes

^aAll Graves'.^bA, age; C, social class; M, marital status; S, sex. Number of controls same as number of cases, except for Winsa (1991) study with 372.^c1 = Sarason's questionnaire Life-Event Survey, with modifications in Yoshiuchi study 5; 2 = Paykel's interview for Recent Life Events.

in the year prior to the time at which the diagnosis was made by the physician, but it cannot always be known with certainty when precisely the first symptoms and signs started. Nevertheless, all six studies observed more stressful life events in the patients with Graves' hyperthyroidism than in the controls.

Winsa et al. reported more negative stressful life events in patients than in controls (odds ratio, OR, 1.33; 95% confidence interval, CI, 1.20–1.48); the odds increased with increased scores attached to the negative life events, from OR 1.0 for a score of 0 points to OR 6.3 (95% CI 2.7–14.7) for a score of ≥ 5 points; the number of positive events was similar in both groups. The prevalence of divorce was higher and job satisfaction was lower in patients than in controls, but a score for mastery in life (apart from work) did not differ between the groups. Patients had type A behavior not more frequently than controls (28 vs. 18%, not significant). The effect of negative life events was similar in patients with or without familial occurrence of thyroid disease, implying different mechanisms for these two risk factors.

Sonino et al. also reported more life events in patients than in controls (1.51 vs. 0.54, $p < 0.001$); patients had more independent events (defined as not likely to be a consequence of the illness) and events with negative impact than controls, according to a blind rater.

Kung likewise observed a higher number of negative events in patients than in controls (0.94–1.54 vs. 0.38–0.88, $p = 0.003$) and higher negative event

scores (1.55–2.69 vs. 0.58–1.45, $p = 0.002$); the reporting of positive and neutral events was similar in both groups. Groups did not differ in coping ability, but Graves' patients experienced more daily hassles (12.11–17.05 vs. 6.36–11.00, $p = 0.001$) and higher hassle scores (17.55–25.34 vs. 8.92–15.87, $p = 0.001$).

Radosavljevic et al. reported a mean number of 3.18 stressful life events in patients versus 1.85 in controls ($p = 0.0001$). Life events that were significantly more prevalent in patients than in controls were change in time spent on work, unemployment, arguments at work, change in work conditions, increased arguments with spouse or fiancé, hospitalization of family member for serious illness, and moderate financial difficulties.

Yoshiuchi et al. found a lower score for problem-oriented coping skills and a higher prevalence of smoking but not of daily hassles in patients than in controls. After adjustment for these variables, stressful life event were significantly associated with the risk of Graves' disease in women but not in men; the relative risk was 7.7 (95% CI 2.2–27) for women with highest stress scores compared with lowest stress scores.

Matos-Santos et al. reported a significantly greater number and impact of negative stressful life events in patients with Graves' disease than in healthy controls.

Taken together, these six studies, despite their limitations, provide fair evidence for the association

between negative stressful life events and the onset of Graves' hyperthyroidism. The lack of evidence for such an association in a previous study by Gray and Hoffenberg might be explained by the inclusion of thyrotoxic patients without Graves' disease.

Stress and Graves' Disease: A Causal Relationship?

The case-control studies do not provide conclusive evidence that stress plays a causative role in the immunopathogenesis of Graves' hyperthyroidism because of their retrospective nature. It would be ideal to have prospective studies with regular stress evaluation in euthyroid subjects at risk for developing Graves' disease to see whether the stress load is higher in subjects who become hyperthyroid than in subjects who remain euthyroid; such studies have not been done. But there have been two prospective studies in patients with Graves' hyperthyroidism who were investigated when euthyroid on antithyroid drug treatment. In both studies, Yoshiuchi and Fukao observed higher scores of daily hassles in patients who experienced recurrent hyperthyroidism after the discontinuation of antithyroid drugs than in patients who remained euthyroid.

Matos-Santos et al. evaluated stressful life events in Graves' hyperthyroid patients and controls and also in 31 patients with toxic nodular goiter – a nonautoimmune disease causing thyrotoxicosis. The number and impact of negative stressful life events were higher in patients with Graves' disease than in patients with toxic nodular goiter or controls, whereas the difference between patients with toxic nodular goiter and controls was not significant. The data suggest that hyperthyroidism itself is not responsible for more stressful life events preceding the diagnosis of thyrotoxicosis; the data, rather, support the view that stress precipitates autoimmunity.

There are almost no data on a putative relationship between stress and TPO antibodies, an early marker of autoimmune thyroid disease, except a study by Strieder et al., who compared euthyroid women with or without TPO antibodies. No difference was found between both groups in the number of stressful life events, the number of daily hassles, or the scores of negative and positive affect scales. Stress apparently seems more related to thyroid autoimmunity against the TSH receptor, causing Graves' hyperthyroidism. This is underlined by a significant increase in the incidence of Graves' disease in eastern Serbia during the civil war in former Yugoslavia. The specificity of this finding is indicated by the fact that during the same time period the incidence of toxic nodular goiter did not increase. Chiovato et al. did

not find a single case of Graves' hyperthyroidism in patients with panic disorder; they hypothesize that failure to activate the hypothalamic-pituitary-thyroid axis by chronic endogenous stress due to panic disorder (as opposed to exogenous stress due to life events) might explain why panic disorder does not precipitate Graves' hyperthyroidism.

It can be concluded that the association between stressful life events and the onset of thyrotoxicosis is specific for Graves' hyperthyroidism and that stress probably is involved in the immunopathogenesis of Graves' disease. How stress contributes to the precipitation of the disease is presently poorly understood. The activation of the hypothalamus-pituitary-adrenal axis would be expected to cause immunosuppression rather than autoimmunity. Paradoxically, however, a surprising increase in Graves' disease was reported in patients with multiple sclerosis undergoing anti-T cell immunosuppressive therapy. Regulatory CD4⁺ CD25⁺ T-cells may be involved. These cells suppress autoimmune disease; if these cells are more sensitive to immunosuppression and slower to regenerate than other T cell subsets, this may explain the paradoxical observations.

Further Reading

- Chiovato, L., Marino, M., Perugi, G., et al. (1998). Chronic recurrent stress due to panic disorder does not precipitate Graves' disease. *Journal of Endocrinological Investigation* 21, 758–764.
- Dayan, C. M. (2001). Stressful life events and Graves' disease revisited. *Clinical Endocrinology* 55, 13–14.
- Fukao, A., Takamatsu, J., Murakami, Y., et al. (2003). The relationship of psychological factors to the prognosis of hyperthyroidism in antithyroid drug-treated patients with Graves' disease. *Clinical Endocrinology* 58, 550–555.
- Gray, J. and Hoffenberg, T. (1985). Thyrotoxicosis and stress. *Quarterly Journal of Medicine* 214, 153–160.
- Kung, A. W. C. (1995). Life events, daily stresses and coping in patients with Graves' disease. *Clinical Endocrinology* 42, 303–308.
- Matos-Santos, A., LacerdaNobre, E., Costa, J. G. E., et al. (2001). Relationship between the number and impact of stressful life events and the onset of Graves' disease and toxic nodular goitre. *Clinical Endocrinology* 55, 15–19.
- Mizokami, T., Li, A. W., El-Kaissi, S., et al. (2004). Stress and thyroid autoimmunity. *Thyroid* 14, 1047–1055.
- Paunkovic, N., Paunkovic, J., Pavlovic, O., et al. (1998). The significant increase in incidence of Graves' disease in Eastern Serbia during the civil war in the former Yugoslavia (1992 to 1995). *Thyroid* 8, 37–41.
- Radosavljevic, V. R., Jankovic, S. M. and Marinkovic, J. M. (1996). Stressful life events in the pathogenesis of Graves' disease. *European Journal of Endocrinology* 134, 699–701.

- Rosch, P. J. (1993). Stressful life events and Graves' disease. *Lancet* 342, 566–567.
- Sonino, N., Girelli, M. E. and Boscaro, M. (1993). Life events in the pathogenesis of Graves' disease: a controlled study. *Acta Endocrinologica* 128, 293–296.
- Strieder, T. G. A., Prummel, M. F., Tijssen, J. G. P., et al. (2005). Stress is not associated with thyroid peroxidase antibodies in euthyroid women. *Brain, Behavior, and Immunity* 19, 203–206.
- Winsa, B., Adami, H-O. and Bergström, R. (1991). Stressful life events and Graves' disease. *Lancet* 338, 1475–1479.
- Yoshiuchi, K., Kumano, H., Nomura, S., et al. (1998). Psychosocial factors influencing the short-term outcome of antithyroid drug therapy in Graves' disease. *Psychosomatic Medicine* 60, 592–596.
- Yoshiuchi, K., Kumano, H., Nomura, S., et al. (1998). Stressful life events and smoking were associated with Graves' disease in women, but not in men. *Psychosomatic Medicine* 60, 182–185.

Grieving

A Ray

Harvard Medical School and Dana-Farber Cancer Institute, Boston, MA, USA

H Prigerson

Dana-Farber Cancer Institute, Harvard Medical School, and Brigham and Women's Hospital, Boston, MA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S Noaghiul and H Prigerson, volume 2, pp 289–296, © 2000, Elsevier Inc.

Introduction

Diagnosis of Complicated Grief

Outcomes and Risk Factors of Complicated Grief

Treatment

Glossary

<i>Complicated grief</i>	The experience of debilitating grief symptoms 6 months postloss and beyond.
<i>Separation anxiety</i>	Fear and anxiety provoked by the threat of separation from a primary attachment figure.

Introduction

Grieving describes the emotional state associated with losing a person to whom one is attached. Bereavement is a natural part of life, and the normal process of grieving is self-limited and is most commonly not a cause for clinical concern. Despite this, research shows that bereavement is one of the most challenging life experiences and that bereaved people, as a group, have a higher risk of mortality and morbidity. This increased risk is largely a function of the significant minority for whom the process of grieving becomes overwhelming, causing distress and disruption for an unusually long period. In these cases of

complicated grief, there is cause for clinical concern and appropriate interventions are required.

Normal Grief

The vast majority of bereaved individuals experience normal or uncomplicated grief. Although normal grieving is painful and disruptive initially, eventually bereaved survivors are able to accept the loss as a reality and move on with life. They are able to carry out daily functions and gradually begin to enjoy leisure activities.

One of the most commonly cited models of normal grief resolution among professionals and laypersons is Kubler-Ross's stage theory. This theory proposes orderly progression through the following stages: disbelief or shock, yearning, anger and protest, depressed mood, and ultimate acceptance of the loss. Recent research attempting to test this theory suggests that multiple stages may be present simultaneously. Instead of clear-cut stages, on average the bereaved person's distress gradually declines as acceptance increases.

For the first few months following the loss, the symptoms of uncomplicated grief are difficult to distinguish from complicated grief. After 6 months, however, the majority of bereaved individuals have reached some level of acceptance, are able to engage in productive and enjoyable activities, and can remain connected with others while developing new relationships. Bereaved survivors who are experiencing adjustment difficulties 6 months postloss, such as suicidal thoughts and gestures, major depressive disorder (MDD), or posttraumatic stress disorder (PTSD), and those who are experiencing persistent symptoms of grief are a cause for clinical concern.

Resilient Responses to Bereavement

Some bereaved people exhibit little or no distress post-loss. Bereavement experts have debated the causes for this. Some have suggested that minor grief

symptoms may signify a more superficial attachment to the deceased. Others have suggested that low distress levels following a significant personal loss are the sign of a maladjusted, emotionally distant person. The traditional view has been to see the lack of grief as denial or an inhibition of the normal grieving process. Thus, these bereaved people have been counseled to assist in the resolution of their latent grief. Studies have shown, however, that some of these subjects may have a healthy resilience to loss that allows them to recover with little disruption in functionality. By this view, resilient subjects have little to benefit from grief counseling because they show little evidence of struggling internally with the loss. Still, using the absence of grief as a marker for resilience is somewhat controversial. Because grief intensity is in direct proportion to the attachment to the deceased, the absence of grief is truly a mark of experienced resilience only if the attachment to the deceased was very strong. This is often difficult to assess. However, there is much to be gained in further understanding the resilient response so that individuals struggling with grief can adopt some of these successful coping strategies.

Diagnosis of Complicated Grief

Descriptive Features

The term complicated grief has been used interchangeably with pathological grief, abnormal grief, atypical grief, and pathological mourning. At one time traumatic grief was also used, but since the attacks on September 11, 2001, this term has been reserved for the psychological reaction following a traumatic event and for grief following the loss of a person during such an event. The term complicated grief is now primarily used to describe the experience of losing a person to natural causes.

Complicated grief can be described as a state of chronic mourning. The bereaved survivor cannot accept the reality of the loss and as a result is reluctant to make any adaptations to life. The psychological distress is characterized by intense yearning for the absent loved one. His or her constant focus on the absence of the deceased creates a barrier to maintaining and forging relationships. This simply heightens the sense of isolation the bereaved person feels. This state of grieving has been shown to be a risk factor for suicidal thoughts and ideation, over and above any confounding influences of diagnoses such as MDD.

Diagnostic Criteria

The proposed set of diagnostic criteria for complicated grief from the *Diagnostic and Statistical Manual of Mental Disorders* (5th edn.; DSM-V) is shown in Table 1. This diagnosis requires that the bereaved person be persistently yearning for the deceased in a manner that disrupts his or her daily life. In addition, four of the following eight symptoms must be experienced to a distressing and disruptive degree (i.e., several times a day, interfering with social or occupational functioning): trouble accepting the death, inability to trust others since the loss, excessive bitterness related to the death, feeling uneasy about moving on, detachment from formerly close others, feeling life is meaningless without the deceased, feeling that the future holds no prospect for fulfillment without the deceased, and feeling agitated since the death. Also, as mentioned earlier, the severity and number of symptoms must persist for at least 6 months, regardless of when those 6 months occur in relation to the loss. This allows both chronic and delayed types of grief to be included in the diagnosis.

There is evidence to support the 6-month delay in making a complicated grief diagnosis. It is important

Table 1 Criteria for complicated grief proposed for DSM-V^a

Criterion A	Yearning, pining, and longing for the deceased; yearning must be experienced at least daily over the past month or to a distressing or disruptive degree
Criteria B	In the past month, the person must experience four of the following eight symptoms as marked, overwhelming, or extreme: - Trouble accepting the death - Inability to trust others since the death - Excessive bitterness or anger about the death - Feeling uneasy about moving on with his or her life (e.g., difficulty forming new relationships) - Feeling emotionally numb or detached from others since the death - Feeling life is empty or meaningless without the deceased - Feeling the future holds no meaning or prospect for fulfillment without the deceased - Feeling agitated, jumpy, or on edge since the death
Criterion C	The symptom disturbance causes marked dysfunction in social, occupational, or other important domains
Criterion D	The symptom disturbance lasts at least 6 months

^aFrom Prigerson, H. G., Vanderwerker, L. C. and Maciejewski, P. K. (in press). Chapter 8. In: Stroebe, M., Hansson, R., Schut, H. & Stroebe, W. (eds.) *Handbook of bereavement research and practice: 21st century perspectives*. New York: American Psychological Association Press.

to allow time for the normal grieving symptoms to subside, in order to detect those who are unusually distressed. This allows for a more conservative threshold for diagnosing complicated grief, and we are less likely to include false-positive patients whose symptoms will eventually naturally resolve. Also, patients still experiencing symptoms at 6 months were found to be less likely to experience natural grief resolution and more likely to experience negative outcomes 13–23 months postloss.

Differential Diagnosis

Bereavement may be complicated by a number of other psychiatric disorders. Major depression is the primary reported disorder following the death of a loved one. However, complicated grief symptoms form a symptom cluster distinct from the depressive or anxiety-related symptom clusters. Depressive symptoms include depressed mood, psychomotor retardation, and decreased self-esteem; complicated grief is specifically associated with yearning, disbelief, detachment, bitterness, and anger related to the death. PTSD is another complication associated with bereavement, especially if the death was associated with a traumatic incident. Studies have shown, however, that unlike PTSD, complicated grief in response to a natural death is not accompanied by fears of violent harm to self and others, hypervigilance, or avoidance of a fear-inducing stimulus. Complicated grief may, of course, occur in conjunction with PTSD in the context of the violent death of a loved one. In fact, complicated grief does frequently occur in addition to MDD and/or PTSD; however, many cases of complicated grief are currently missed due to the reliance on guidelines for psychiatric diagnoses provided in the DSM-IV.

Other factors that set complicated grief apart from MDD and PTSD are the risk factors, outcomes, and responses to treatment.

Outcomes and Risk Factors of Complicated Grief

Physical and Psychological Outcomes

Bereaved subjects suffering from complicated grief have been found to suffer myriad negative consequences. Long periods of distress and dysfunction may be responsible for this. Complicated grief has been associated with heightened risks of cancer, hypertension, cardiac events, and suicidal ideation. It has also been associated with disability, functional impairment (social, familial, and occupational dysfunctions), hospitalization, and reduced quality of life. Those suffering from complicated grief are also more likely to engage

in unhealthy behaviors such as alcohol and cigarette consumption. Psychologically, patients with complicated grief are more likely to develop comorbid conditions with depressive symptoms and major depressive episodes. The relative risks for negative health outcomes associated with complicated grief are summarized in Table 2.

Risk Factors

Given the negative outcomes of untreated complicated grief, the identification of at risk individuals should be a priority for clinicians. Several possible risk factors for the development of complicated grief have been identified in studies. One important risk factor is the closeness of the relationship between the bereaved and the deceased. Relationships that were dependent, confiding, and close led to the poorest bereavement adjustment. In fact, relationships that are conflicted at baseline result in lower rates of bereavement disorders. The belief that complicated grief is fundamentally an attachment disturbance is borne out by the identification of weak parental bonding in childhood as a risk factor for complicated grief. Bereaved people with a damaged sense of security due to childhood abuse or severe neglect have a higher risk of experiencing complicated grief as adults. Childhood separation anxiety was identified as a risk factor as well. Also, individuals who are generally averse to change of any kind (e.g. lifestyle changes) are more likely to experience complicated grief. Two factors that have been found to reduce the risk of complicated grief are advance preparation for the death and a strong support network.

Thus, in summary, patients who have attachment difficulties at baseline and who feel unprepared and unsupported during and after the death are at higher

Table 2 The relative risk of some negative health outcomes among bereaved individuals diagnosed with complicated grief at 6 months postloss^a

Health outcome (incidence)	Relative risk	Confidence interval (95%)
High systolic blood pressure at 13 months post loss	1.11	1.02–1.20
Depression at 13 months post loss	2.72	0.84–8.81
Changes in smoking at 13 months post loss	16.7	0.86–320.4
Changes in eating at 13 months post loss	7.02	1.62–30.6
Heart trouble at 25 months post loss	1.15	1.04–1.27
Problems with alcohol at 25 months post loss	1.25	0.98–1.63

^aFrom El-Jawahri, A. and Prigerson, H. G. (in review). Evidence-based guidelines for the diagnosis and treatment of complicated bereavement. *Journal of Palliative Medicine*.

risk of experiencing complicated grief. This indicates that interventions geared toward preparing caregivers for the impending death of a loved one, as well as promoting secure alternative engagements to others, would be helpful in the treatment of persistent grief symptoms.

Treatment

As previously mentioned, there is a strong connection between a lack of preparedness for the death of a loved one and the development of complicated grief postloss. Thus, it would seem that interventions administered preloss might prove promising at reducing bereaved survivors' risk of complicated grief. Clinicians who provide palliative care might extend their efforts to the caregivers of a dying patient, encouraging them to say goodbye and accept the impending loss. This would entail a greater comfort with prognosticating the end of life, an issue that is fraught with controversy and complexity.

Just as physician involvement and communication can increase the chances of caregivers being prepared for the death, similarly, hospice enrollment has been shown to reduce the risk of MDD. This is consistent because hospice enrollment usually requires acceptance that death will occur within 6 months, intimating caregiver preparedness. In the Netherlands, a study found that caregivers of cancer patients who died by euthanasia were less likely to experience disruptive grief symptoms than those whose loved ones died a natural death. This, too, is consistent, given that euthanasia implies a greater acceptance of the patient's prognosis on the part of the caregivers.

Concrete interventions to help caregivers prepare for the loss center around proactively assisting caregivers to "hope for the best, plan for the worst." Traditionally, getting one's affairs in order include making an effort to visit the patient and expressing appreciation, regrets, and reassurances before it is too late. In addition, health-care professionals should not hesitate to recommend hospice and palliative care services when their need becomes apparent. An earlier referral to a hospice has been shown to be beneficial in promoting a peaceful death and lessening distress for caregivers.

As noted earlier, after the loss, in most cases grief is manageable and resolves naturally, requiring no professional interventions. Bereavement interventions are more appropriate for those individuals who are particularly at risk or those who present with unremitting, severely disruptive grief symptoms six months postloss.

Research has shown that interventions such as nortriptyline and interpersonal psychotherapy, which

were effective for the treatment of depressive symptoms, were not effective in treating grief symptoms. To date, there have been no randomized controlled trials of pharmacotherapies for the reduction of grief-symptom severity. Open trials of selective serotonin reuptake inhibitors have been conducted and suggest that these and similar agents show promise for the reduction of complicated grief symptom severity, but randomized placebo-controlled trials are needed before more definitive conclusions can be drawn about their efficacy for the amelioration of complicated grief symptomatology.

Future pharmacotherapeutic options may also evolve from ongoing studies of functional neuroanatomy. In one study, grieving subjects were exposed to visual and verbal stimuli to remind them of the deceased. The stimuli led to activation in the posterior cingulate cortex, cerebellum, and medial/superior frontal gyrus. Thus, grief is mediated through a complex network involving affect processing, visual imagery, autonomic regulation, memory retrieval, processing of familiar faces, and modulation/coordination of these functions. New insights into the long-term health consequences of grief may follow, as well as therapies to modulate the neurochemical response.

As mentioned previously, a lack of social support is a significant risk factor for the development of complicated grief. Thus, it follows that bereaved individuals who lack a social network may be helped greatly by group therapy and support groups. Studies show that the perception of social support leads to a reduction of depression, whereas group therapy for complicated grief leads to an increase in perceived support.

One therapy specifically developed for complicated grief is known as complicated grief treatment. This combines a number of different approaches, including psychoeducation about normal grief and complicated grief, a dual focus on adjusting to the loss and restoration of a satisfying life, and motivational therapy to create appropriate life goals. This therapy was found to be more effective in a shorter time than interpersonal psychotherapy in a population of complicated grief patients. However, given the eclectic nature of the treatment, it is difficult to discern which specific component is providing the main benefit. Using a model in which grief is conceptualized as trauma, the therapy also includes desensitization through imagined conversations with the deceased, retelling of the death scene, and viewing photographs of the deceased. For nontraumatic deaths, however, this aspect of the treatment may be less important than dealing with issues of attachment, separation, loss, and reattachment.

Further research will yield insight into the most effective therapies for bereaved people suffering from

overwhelming grief symptoms. Most important, an understanding of complicated grief will encourage clinicians to help prepare caregivers of terminally ill patients for their loss. Providing appropriate resources before the loss will better facilitate normal grieving and grief resolution and ensure that survivors can more successfully adapt and resume a functional and joyful life.

Further Reading

- Bowlby, J. (1980). *Attachment and loss*. Vol. 3: *Loss, sadness and depression*. New York: Basic Books.
- Brown, G. W. and Harris, T. O. (1989). Depression. In: Brown, G. W. & Harris, T. O. (eds.) *Life events and illness*, pp. 49–94. New York: Guilford Press.
- Clayton, P. J. (1990). Bereavement and depression. *Journal of Clinical Psychiatry* 51, 34–38.
- El-Jawahri, A. and Prigerson, H. G. (in review). Evidence-based guidelines for the diagnosis and treatment of complicated bereavement. *Journal of Palliative Medicine*.
- Freud, S. (1953–1974). Mourning, melancholia (1917). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud*, (vol. 14), pp. 239–258. London: Hogarth Press and the Institute of Psychoanalysis.
- Jacobs, S. (1993). *Pathologic grief: maladaptation to loss*. Washington, DC: American Psychiatric Press.
- Kubler-Ross, E. (1969). *On death and dying*. New York: Macmillan.
- Parkes, C. M. and Weiss, R. S. (1983). *Recovery from bereavement*. New York: Basic Books.
- Prigerson, H. G., Bierhals, A. J., Kasl, S. V., et al. (1997). Traumatic grief as a risk factor for mental and physical morbidity. *American Journal of Psychiatry* 154, 616–623.
- Prigerson, H. G., Frank, E., Kasl, S. V., et al. (1995). Complicated grief and bereavement-related depression as distinct disorders: preliminary empirical validation in elderly bereaved spouses. *American Journal of Psychiatry* 152, 22–30.
- Prigerson, H. G. and Jacobs, S. C. (2001). Perspectives on care at the close of life. caring for bereaved patients: “all the doctors just suddenly go.” *Journal of the American Medical Association* 286, 1369–1376.
- Prigerson, H. G., Maciejewski, P. K., Reynolds, C. F., III, et al. (1995). Inventory of complicated grief: a scale to measure maladaptive symptoms of loss. *Psychiatry Research* 59, 65–79.
- Prigerson, H. G., Vanderwerker, L. C. and Maciejewski, P. K. (in press). Chapter 8. In: Stroebe, M., Hansson, R., Schut, H. & Stroebe, W. (eds.) *Handbook of bereavement research and practice: 21st century perspectives*. New York: American Psychological Association Press.
- Shear, M. K., Frank, E., Houck, P., et al. (2005). Treatment of complicated grief: a randomized controlled trial. *Journal of the American Medical Association* 293, 2601–2659.
- Zisook, S. and Shuchter, S. (1993). Uncomplicated bereavement. *Journal of Clinical Psychiatry* 54, 365–372.

Relevant Website

National Cancer Institute. <http://www.nci.nih.gov>.

Group Therapy

N Wong

Uniformed Services University of the Health Sciences,
Bethesda, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by N Wong, volume 2, pp 297–301, © 2000, Elsevier Inc.

Rationale for Group Therapy
Therapeutic Elements of Group Therapy
Applications
Conduct of Group Therapy

Glossary

Abreaction

A process by which repressed material, particularly a painful experience or conflict is brought back into consciousness. In the process, the person not only recalls but relives the emotional situation and experiences the appropriate emotional response.

Catharsis

The verbalization of ideas, thoughts, and expressed material that is accompanied by an emotional response and that produces a state of relief to the individual.

<i>Corrective recapitulation of the primary family group</i>	The phenomenon that patients in a therapy group relate to the group leaders and other members in ways reminiscent of how they once interacted with their parents, siblings, or other significant figures as they grew up; also called family reenactment.
<i>Counter-transference</i>	Reactions evoked in the therapist in response to the patient. These feelings, conscious or unconscious, may be realistic or arise from unresolved pathology. They are regarded as analogs of the patient's transferences.
<i>Group debriefing</i>	An intervention and stress management technique in which the major elements of a trauma are reviewed by participants in an attempt to prevent or reduce the immediate or long-term adverse psychological changes. Although it is therapeutic, is conducted in a group setting, and employs many principles and techniques of group therapy, group debriefing is not truly group therapy. It may proceed to group therapy if the members request that meetings continue because of persistent symptomatology and both the leader and members agree to further sessions.
<i>Group facilitator</i>	A support or educational group leader who assists the group in maintaining its boundaries, keeps the group on task, and provides group feedback by sharing observations and summarizing the process of the group with its members. The intent of the facilitator is not to attempt individual personality changes but to encourage self-learning, attitudinal shifts, and stress reduction.
<i>Group therapist/leader</i>	The most important person in the group; responsible for selecting and preparing patients, forming the group, and conducting the group in a manner that models ideal group behavior and creates therapeutic group norms to enable its members to attain their therapeutic goals, which may involve personality and behavior changes.
<i>Group therapy</i>	A widely accepted psychiatric treatment modality that uses therapeutic forces within the group, constructive interactions between members, and interventions of a trained leader to change the maladaptive behavior, thoughts, and feelings of people with a psychiatric disorder.
<i>Repression</i>	The unconscious act of controlling and inhibiting an unacceptable impulse, emotion, or thought.
<i>Suppression</i>	The conscious act of controlling and inhibiting an unacceptable impulse, emotion, or thought.

<i>Transference</i>	In the classic sense, all aspects of a patient's feelings and behavior toward a therapist, including rational and irrational aspects that arise from unconscious sources; usually classified into positive and negative reactions. In the general sense, it includes real aspects of the relationship and the unconscious influences of parental or other meaningful figures early in the individual's development.
<i>Universality</i>	The experience by the patient in a group setting that he or she shares many similarities and features in common with others, that "We are all in the same boat."

Rationale for Group Therapy

Some of the principles of group therapy have been applied with success in the fields of business and education in the form of training, sensitivity, and encounter groups. Individuals participating in such groups are generally not regarded as having a psychiatric disorder. Group therapy is also used in the prevention of psychiatric illness, and some of its tenets are applied in debriefing formats for acute stress reactions. Because of the numerous forms of group therapy, it is more accurate to speak of the many group therapies. However, for the sake of simplicity, it should be understood that the generic term group therapy in this article is used to describe a wide array of treatments in a group setting with their particular requirements and different objectives and goals.

Treatment in groups is a proven and effective therapeutic modality for stress-related conditions. Historically, the group has served as a protector and nurturer of the individual and a purveyor of culture. Although originating in the United States in 1905, group methods were initially used by English psychiatrists who were under pressure to treat the vast number of psychiatric casualties suffering from stress and trauma during World War II. In the safe environment of the group, patients or members have the opportunity to interact with others in a setting reflective of society at large; to gain mutual support; to be exposed to new learning situations; to express and benefit from altruism, which in itself can be beneficial; and to find opportunities for multiple healthy identifications and catharsis. In the group, a here-and-now focus and interpersonal relationships are stressed more than in individual therapy.

Therapeutic Elements of Group Therapy

Numerous conceptual models of group therapy exist, but they share in common certain therapeutic

elements or factors that have been studied and found to be curative for individuals in groups. More than 200 group process and therapeutic elements have been described in the literature to account for the curative changes that occur in group therapy. These can be classified into the broad categories of intellectual, emotional, and action factors. One of the most widely accepted classifications identifies 11 primary elements: the installation of hope, universality, the imparting of information, altruism, the corrective recapitulation of the primary family group, the development of socializing techniques, imitative behavior, interpersonal learning, group cohesiveness, catharsis, and existential factors. The therapeutic factors, are frequently interdependent and share similarities.

Not all of these curative or therapeutic factors need to be present for the group to be effective. Certain types of groups require the presence of particular elements. For example, in groups dealing with stress disorders arising from an acute or immediate traumatic event, emphasis is placed on imparting information, abreaction, catharsis, and universality. When the stress has been chronic and caused conditions of a chronic nature, there is a greater focus on interpersonal learning, the corrective recapitulation of the primary family group, and the development of socializing techniques, in addition to the aforementioned factors. In the latter setting, transference and countertransference are more likely to occur.

Types of Groups

Group therapy is used in inpatient and outpatient settings, institutional work, partial hospitalization units, halfway houses, community settings, and private practice. Open-ended groups exist for an indefinite time period, are open to new members, and pursue a variety of therapeutic goals. Closed-ended groups are time-limited, closed to new members, and focus on achieving a particular goal or purpose. Time-limited groups, or short-term groups, are commonly employed to treat people with acute stress-related disorders. The duration of treatment is ordinarily stated at the start of the group and rarely exceeds several months. Individuals exposed to a traumatic or catastrophic event may be seen for only one group session.

Placement in a long-term group is usually indicated for patients who have experienced chronic stress from underlying medical conditions or needing to reduce chronic stress that may contribute to the formation of physical or emotional illnesses. Long-term groups typically run for several months to several years and convene at regular and frequent intervals such as once or twice weekly.

Group Size

The size of the group varies considerably in the treatment of stress-related conditions. When a catastrophic or traumatic event involves a large number of people, debriefing groups may have as many as 100 or more people, whereas a therapy group for posttraumatic stress disorder (PTSD) patients usually consists of 5–10 members. Generally, smaller groups are more appropriate for an intensive and prolonged group therapy experience.

Group Composition

Individuals placed in a group setting for the treatment or prevention of psychiatric symptoms or illness in which stress plays a prominent role are admitted on the basis of their primary symptom (such as phobias, anxiety, or depression), an illness or diagnosis (such as PTSD; cancer; or human immunodeficiency virus, HIV), or an event (such as rape, surviving the Holocaust or a concentration camp, or a recent disaster). Although secondary considerations of age, gender, socioeconomic, and ethnic factors are of lesser importance in stress-related groups than in traditional group therapy, a compatible matching of these elements may ensure greater success for the individual and the group.

Applications

Group therapy is useful for people who have experienced a vast number of situations or medical conditions in which stress plays a major role. Some examples include PTSD patients, sexual or physically abused victims, concentration camp survivors and their families, bereaved people, cancer patients, and family members and caretakers of patients with chronic mental or physical illnesses. There are established groups for patients with HIV and acquired immunodeficiency syndrome (AIDS), with chronic fatigue syndrome, with ulcerative colitis, with heart disease, with diabetes, with obesity, with fibromyalgia, and on long-term dialysis. The list is varied and growing.

Conduct of Group Therapy

The basic principles of group therapy are observed in most therapy groups, although the approach varies with the type of group that has been formed in response to the specific situation. A few of the common approaches and indications for group treatment are described next.

Group Debriefing

Perhaps the most widely known and used intervention to treat individuals exposed to acute traumatic situations, the aims of group debriefing are to manage the immediate emotional distress and to prevent adverse long-term psychological changes. There are several debriefing protocols; one of the best known and most widely practiced is the critical incident stress format developed by J. Mitchell in 1981. The protocol consists of consecutive phases in which various aspects of a traumatic exposure are explored and worked through by the participants. The focus remains on the individual and not on the group. Members are encouraged to recall their personal experiences of the traumatic event to allow for ventilation and to attempt to understand and master the confusion and trauma. Education and the mobilization of social support are key elements in this approach, and symptomatic individuals are identified and referred for further treatment. Some of the goals of debriefing are to assist individuals to regain emotional and cognitive controls; reduce cognitive dysfunction; diminish traumatic conditioning; and reduce group scapegoating, projection, and nihilistic and antisocial attitudes that may arise in response to the trauma and to correct improper information about the event. Debriefing is usually conducted in naturally occurring groups, is considered short term, lasting at most only a few sessions. Because this approach has the potential to produce negative outcomes as well as positive results, it is essential that the leader or leaders be well trained in group work.

Cancer Groups and Groups for the Terminally Ill

Depending on the nature of the illness and prognosis, cancer and terminally ill patient support groups may be short or long term and open- or closed-ended. Breast cancer group therapy has been the best researched of the group approaches and in many studies it has been shown to increase the length and quality of patients' lives. Patients suffering from terminal cancer experience significant stress in attempting to deal with their anguish, loss of control over what is happening to them, physical pain and disability, and sense of uselessness while awaiting death. Through the group experience, they are afforded the opportunity to find meaning outside themselves and provided a sense of purpose as they support one another and share their life experiences.

Patients in the terminal stage of cancer frequently harbor irrational and rational anger toward their families, the helping professions, and society in general. The group experience often assists them to transcend their immediate plight to find a deeper meaning for their existence and to accept death with greater

equanimity. Similarly, the caretakers of these patients who are members of cancer support groups derive emotional support and gain helpful information from others to reduce their stress levels and to deal with the resentment, guilt, and anger that may develop in the course of caring for terminally ill relatives.

The hypothetical 5-year waiting period during which no evidence of recurrence has been found and a cure is likely is a stressful time for many post-treatment cancer patients. Group therapy for these recovering patients may be long term and supportive in nature as individuals share medical information, personal setbacks, and coping strategies. For many members, it is a time for personal growth and maturation as they struggle with their illness and the impact it has had on their lives.

The main therapeutic elements present in successful cancer groups appear to be group cohesion, social support, enhancing communication skills with family and health-care workers, relaxation training, and the opportunity for emotional expressiveness.

Support Groups for Other Medical Conditions

Group therapy has been useful for a wide array of medical illnesses that are partially caused or aggravated by stress. In turn, stress may arise as a consequence of suffering with a medical illness. Most of these groups are open-ended, long term, and meet on a regular basis. Many of the therapeutic elements found in cancer support groups are also present in these groups. Major emphasis is placed on imparting information, learning coping skills, social support, the installation of hope, and dealing with here-and-now situations. Unlike those in cancer support groups, the patients in these groups may not face imminent death despite the chronicity and severity of their illnesses.

Posttraumatic Stress Disorder Groups

Group therapy is one of the most widely used treatments for patients with PTSD. Individuals diagnosed with acute PTSD, in which the symptoms have been present for less than 3 months, may be treated successfully in a time-limited or short-term group. Individuals showing symptoms of a chronic PTSD generally require long-term treatment.

In the various group therapies used for patients with these disorders, the goals and objectives are to contain their severe distress and suffering and to improve the quality of their lives. To attain these ends, a variety of group approaches are employed, ranging from a cognitive behavioral focus to the use of psychodynamically oriented psychotherapy. The behavioral model emphasizes the reduction of symptoms, the acquisition and maintenance of coping skills, and

the systematic examination of the assumptions and misattributions that govern the patient's behavior, expectations, and appraisal of events (as described in the works of Foa and colleagues in the 1990s). Prominent in these groups are the mechanisms of group cohesiveness, imparting of information, imitative behavior, interpersonal learning, and installation of hope.

Since the 1970s, Horowitz and his colleagues have researched and firmly established that a psychodynamically oriented approach is successful in the treatment of PTSD. The goal for the traumatized patient is to reconcile the occurrence of the trauma with his or her personal concept and the world. The principal aim of this therapy is to assist individuals to integrate their traumatic experiences through the therapeutic reexperiencing of the events in a supportive environment. Attempts are made to gain insight into the conscious and unconscious meaning of the symptoms to enable patients to master the trauma and restore functioning. This approach acknowledges the influence of dispositional variables on the patients' response to the trauma and therapy and emphasizes the use of the corrective recapitulation of the primary family group, interpersonal learning, insight, and catharsis more than other types of group treatments for stress.

Staff and Caretaker Support Groups

Because of a heightened awareness of dementias, especially Alzheimer's disease, the increasing longevity of the world population, and the large numbers of the chronically mentally and physically ill, participation and interest in staff and caretaker groups have grown significantly. Typically, health-care providers and caretakers operate under high and sustained daily levels of stress, and they benefit from participation in a group. The emphasis on these groups is to enable individuals to provide better patient care and smoother functioning in their surroundings through improved staff well-being. The groups are usually conducted by a facilitator instead of a group therapist or leader, and the goals and objectives are to assist members in reducing stress rather than to explore personal conflicts or to attempt personality change. Group meetings may occur weekly, on an ad hoc basis, or for a set number of sessions in defined modules over the span of a year.

Benefits gained from participating in the group experience include decreased anxiety, reduced sense of isolation, more open communication between caregivers, increased self-esteem, improved adaptation in working with patients or ill family members, and greater group cohesion and teamwork.

Scapegoating and irrationality diminish, patient-care goals become more realistic, and staff members are better able to tolerate and accept pain and suffering in their patients without feeling excessive personal distress.

Concentration Camp and Torture Survivor Groups

A large number of Holocaust survivors exist in the world, with Israel alone housing over 250,000. In addition to Holocaust survivors, there are concentration camp survivors from the Bosnia-Herzegovina war and survivors of torture from Central and South America and nations in Africa. Many elderly Holocaust survivors manifest the survivor syndrome, typically, reporting anxiety, sleep disorders, affective disturbances, and survivor guilt. As these individuals age, their symptoms intensify because of economic and health problems and because of the constant reminder of their previous plight due to the unstable world situation with reports of internment camps, torture, terrorism, and serious religious and ethnic conflicts. Exiled survivors of torture must deal with the acute trauma, loss of family or close ones, often forced migration, and adaptation to a new culture.

Group therapy is supportive in nature and attempts to help survivors cope with the anxieties and painful memories that are still fresh or have been rekindled by disturbing world events, work through the mourning for lost relatives and their own personal losses, improve their mental equilibrium, and enhance their adaptation to their inner and outer worlds. A major objective is to create a sense of safety and well-being in the groups so patients can express themselves freely. Attempts are made early in therapy to create supportive group interpersonal relationships with active leader interventions to help members integrate their painful past with current joys and successes in the present. Many of these groups are open-ended both for membership and in duration because rarely are these individuals able to completely set aside what has happened to them and to successfully proceed in life without a periodic resurgence of disturbing symptoms.

Rape and Sexual Abuse Victims Groups

Victims of rape and sexual abuse typically experience effects on body image, identity, self-regulation, and interpersonal functioning. If the abuse has occurred in childhood and is ongoing, these areas may be affected profoundly. Victims experience both acute and chronic stress when reminded of the trauma, and they are frequently conflicted in their social and sexual relationships. Group therapy helps victims through the reenactment of the traumatic situation and abuse

in the confines of a safe environment, increases their capacity for self-soothing, deals therapeutically with the shame, and teaches patients how to end the cycle of repeated victimization.

The therapeutic or curative factors of universality, interpersonal learning, catharsis, developing socializing techniques, and imparting information are prominent in the group processes. The groups may be short or long term, depending on the acuity of the situation. These groups are often open-ended to accommodate newcomers and to allow previous members to return for periodic reinforcement and support of their newly acquired knowledge and behaviors. Psychoeducation is provided in these groups to instruct patients about abuse effects, in developing problem-solving skills, and in becoming active in supportive topic-focused discussions.

In summary, group therapy or a group therapeutic approach is an effective modality in the treatment of acute and chronic stressful conditions. Indications for group therapy mirror those for individual approaches; but the group has as an advantage the presence of group therapeutic or curative factors not present in individual therapy. The group becomes a safe haven and laboratory where members learn from one another, practice new behaviors in a supportive setting, and are provided opportunities to show altruism, which in itself is a constructive healing force. The group serves as a microcosm of society from which members can venture forth again into the larger society with new knowledge and personal gains from the group experience.

Some individuals find it difficult, if not impossible, to develop sufficient trust and the sense of confidentiality in a group setting, and they may require an introductory period of individual care before introducing them to group therapy. In addition, individuals may benefit from a combination of concurrent individual and group approaches to test out new behaviors and learning.

Perpetrator Groups

Perpetrator groups are gaining in popularity with the increase in public awareness about sexual predators and about emotional and physical abuse in relationships. Common in institutional and court-directed settings, many are now present in both community and private locations. Individuals are afforded a safe environment for catharsis, to sense their impact on others with similar backgrounds, to gain understanding into their behavior, and to learn techniques to control and modify their destructive behavior. These groups tend to be long term and open-ended, and they may employ treatment principles found in alcohol

and substance abuse groups. Anger management techniques and psychoeducation also play important roles.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Cancer Treatment; Holocaust Survivors, Experiences of; Posttraumatic Therapy; Reenactment Techniques.

Further Reading

- Abramowitz, I. A. and Coursey, R. D. (1989). Impact of an educational support group on family participants who take care of their schizophrenic relatives. *Journal of Consulting and Clinical Psychology* 57, 232–236.
- Allan, R. and Scheidt, S. (1998). Group psychotherapy for patients with coronary heart disease. *International Journal of Group Psychotherapy* 48, 187–214.
- Baider, L. and De-Nour, A. K. (1989). Group therapy with adolescent cancer patients. *Journal of Adolescent Health Care* 10, 35–38.
- Classen, C., Butler, L. D., Koopman, C., et al. (2001). Supportive-expressive group therapy and distress in patients with metastatic breast cancer: a randomized clinical trial. *Archives of General Psychiatry* 58, 494–501.
- Forester, B., Kornfeld, D. S., Fleiss, J. L., et al. (1993). Group psychotherapy during radiotherapy: effects on emotional and physical distress. *American Journal of Psychiatry* 150, 1700–1706.
- Foy, D. W., Ruzek, J. I., Glynn, S. M., et al. (2002). Trauma focused group therapy for combat-related PTSD: an update. *Journal of Clinical Psychology* 58, 907–918.
- Freehill, K. M. (1992). Critical incident stress debriefing in health care. *Critical Care Clinics* 8, 491–500.
- Gregurek, R. (1999). Countertransference problems in the treatment of a mixed group of war veterans and female partners of war veterans. *Croatian Medical Journal* 40, 493–497.
- Kelly, J. A. (1998). Group psychotherapy for persons with HIV and AIDS-related illnesses. *International Journal of Group Psychotherapy* 48, 143–162.
- Lansen, J. and Cels, J. P. (1992). Psycho-educative group psychotherapy for Jewish child-survivors of the Holocaust and non-Jewish child-survivors of Japanese concentration camps. *Israeli Journal of Psychiatry & Related Sciences* 29, 22–32.
- Wallis, D. A. (2000). Reduction of trauma symptoms following group therapy. *Australian and New Zealand Journal of Psychiatry* 36, 67–74.
- Wong, N. (2005). Group psychotherapy and combined individual and group psychotherapy. In: Sadock, B. J. (ed.) *Kaplan and Sadock's comprehensive textbook of psychiatry* (8th edn., pp. 2568–2583). Philadelphia: Lippincott Williams and Wilkins.
- Yalom, I. D. (1995). *The theory and practice of group psychotherapy*. New York: Basic Books.

Gulf War Syndrome, Psychological and Chemical Stressors

H Soreq

Hebrew University of Jerusalem, Jerusalem, Israel

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by H Soreq, volume 2, pp 302–312, © 2000, Elsevier Inc.

Acute Psychological Stress and Exposure to
Anticholinesterases Cause Similar Delayed Effects
The Blood–Brain Barrier Is Disrupted under Stress
Early Immediate Response Genes Modulate Acute
Stress Signals
Hypothalamic–Pituitary–Adrenal Involvement in Acute
Stress Responses
Poststress Changes Are Long-Lasting
Acetylcholinesterase Accumulation Confers Protection
against Anticholinesterases
Acetylcholinesterase Gene Expression Associates with State
and Trait Anxiety
Delayed Nervous System Deterioration under Persistent
Acetylcholinesterase Overexpression
Retrospective and Prospective Implications

Glossary

Acetylcholinesterase (AChE) inhibitors Chemical agents capable of preventing the enzyme acetylcholinesterase from performing its catalytic function, namely, the hydrolysis of acetylcholine. These inhibitors may vary in their chemical structure and include natural compounds, drugs, commonly used insecticides, and chemical warfare agents.

Alternative splicing A molecular process through which a single gene may produce more than one protein with distinct properties.

Blood–brain barrier (BBB) The complex structures and processes that prevent the free penetration of compounds into the brain (and free exit from it). It includes the epithelial cells lining the walls of blood vessels in the brain and the end foot of astrocytes surrounding these blood vessels. The tight junctions between endothelial cells in brain blood vessels are major players in this function, as are several proteins and active pumps controlling the move of compounds in and out of the brain.

Psychological stress An emotional state of anxiety and hyperexcitation that may have various origins

and causes and that varies in intensity. It probably is important to ensure rapid, alert functioning under external insults; however, it may also cause delayed adverse consequences at different levels and functions.

Transgenic overexpression

The excessive production of a foreign protein in the cells and tissues of an organism that has been made transgenic by the introduction of foreign DNA into its genome. The transgene included in this DNA directs the production of its protein product in the host organism, leading to overexpression because it is produced in addition to the normal functioning of the host gene.

Acute Psychological Stress and Exposure to Anticholinesterases Cause Similar Delayed Effects

Posttraumatic stress disorder (PTSD) involves delayed cognitive deterioration, depression, irritability, and persistent deficits in short-term memory, as summarized by Robert Sapolsky in 2003. Stress-induced changes in nervous system structure and function appear in many diverse species, suggesting that they reflect an evolutionarily conserved adaptation to environmental insults. For example, magnetic resonance imaging (MRI) demonstrates reduced hippocampal volume in PTSD patients.

Surprisingly similar effects are often reported years after acute or chronic exposures to cholinesterase inhibitors. The use of the carbamate anticholinesterase (anti-AChE) pyridostigmine during the Gulf War for prophylactic protection under threat of chemical warfare emphasizes the importance of this topic, as does the increasing use of anti-AChE drugs for the treatment of Alzheimer's disease. [Table 1](#) summarizes these reports.

Neurophysiological stress events are initiated by millisecond-long bursts of neurotransmitter release, whereas the long-term, persistent brain changes occur at a very slow pace, sometimes over years. Moreover, stress responses primarily elicit hormonal modulations in peripheral tissues, and ACh controls inflammatory reactions. Yet, it is not clear how this imprints on brain structure and function in a long-delayed manner. Finally, the link between structural and functional properties in the brain itself calls for an

Table 1 Common outcomes of psychological stress and anticholinesterase exposure^a

	<i>Organism</i>	<i>Inducer</i>
<i>Neuropathological phenomena</i>		
Excessive sprouting and retraction of neuronal processes	Mollusk	Variety of stressors
Decreased density of dendritic spines, where synaptic connections take place	Transgenic mice overexpressing AChE	Cage life
Atrophy of apical dendrite in the hippocampus region CA3	Tree shrews (primate)	Restraint stress
Reduced hippocampal volume	Humans (war veterans)	Acute traumatic war experience
Degenerative changes in the spinal cord	Hens	Low-level exposure to the organophosphate anti-AChE leptophos
Degeneration of the hippocampus	Rats	Low-level exposure to the organophosphate anti-AChE fenthion
Arrest of neuronal proliferation	Primates (rhesus)	Sociological stress
<i>Impaired functioning</i>		
Enhanced long-term depression of hippocampal activity	Rats	Mild behavioral stress
Enhanced long-term depression and suppressed long-term potentiation of hippocampal activity	Rats	Restraint stress and electric tail shock
Increased permeability of the blood–brain barrier	Mice	Forced swimming stress, transgenic AChE overexpression
Behavioral changes	Primates (<i>Callithrix jacchus jacchus</i>)	Psychological stress
Deficits in short-term memory, long-term psychopathologies	Humans	Combat-related stress, childhood abuse
Impaired firing of CA1 neurons in the hippocampus	Rabbits	Acute exposure to the organophosphate anti-AChE parathion
Suppressed long-term potentiation in the hippocampus	Rats	Acute exposure to the organophosphate anti-AChE soman
Memory impairments	Rats	Low-level exposure to the organophosphate anti-AChE DFP or phosphorodithioate
Increased permeability of the blood–brain barrier	Rats	Soman intoxication
Deficient cognitive performance	Rhesus monkeys	Soman intoxication
Progressive deterioration of memory and general cognitive decline	Humans	Organophosphate pesticide poisoning
Intensified traumatic memories	Mice	Immobilization stress

^aLong-term consequences of psychological insults and anticholinesterase exposures in various evolutionarily remote species. Note the similarities in the nature of the changes observed in both the structural and functional properties of neurons. AChE, acetylcholinesterase; DFP, diisopropyl-fluorophosphonate.

in-depth study of the causal relationships leading to these changes.

The Blood–Brain Barrier Is Disrupted under Stress

Several studies have demonstrated enhanced penetration of the brain by anti-AChEs under psychological stress. This stress-induced disruption of the BBB can explain some of the nervous system-associated sequelae reported by Gulf War veterans. Soldiers at the Gulf were exposed to unknown doses and combinations of various harmful xenobiotics. For prophylactic protection from the anticipated chemical warfare agents, which would have irreversibly blocked acetylcholinesterase (AChE), Gulf War soldiers were administered pyridostigmine, a reversible carbamate cholinesterase inhibitor. Pyridostigmine includes a quaternary ammonium group that under normal circumstances prevents its crossing the BBB. Thanks to this property, it is used routinely to treat peripheral neuromuscular

junction deficiencies in patients with the autoimmune disease myasthenia gravis. During peacetime clinical tests that involved healthy volunteers, pyridostigmine caused mild, primarily peripheral side effects. In contrast, during the Gulf War, pyridostigmine in the same dosage caused a significant increase in reported central nervous system (CNS) symptoms. Friedman and colleagues showed, in animal experiments, that the dose of pyridostigmine required to block 50% of brain AChE was 100-fold lower in stressed mice.

The sudden need for energy in the stressed brain suggests that the disruption of the BBB may be beneficial when neurons are overactivated. Nevertheless, BBB disruption may also allow toxic agents and unneeded proteins to penetrate the brain, which may impair brain functioning under stress. However, the linkage between neuronal excitability and the nature of those impairments is still largely obscure. AChE inhibitors, which increase cholinergic signaling by preventing acetylcholine (ACh) hydrolysis, were reported to cause transient permeabilization of the

BBB. Revealing the molecular mechanisms underlying the links between cholinergic and other elements in BBB-associated cells should therefore contribute significantly toward our understanding of the mechanism(s) leading to BBB disruption under psychological stress. Because cholinergic neurotransmission is imbalanced following traumatic impacts, this question was addressed using transgenic mice overexpressing synaptic AChE. Diffusion-weighted MRI showed impaired water diffusion and BBB functioning, and DNA microarray analyses demonstrated the enhanced production of multiple ion channels and the water channel aquaporin 4. Thus, persistent cholinergic imbalance may by itself cause BBB malfunctioning.

Early Immediate Response Genes Modulate Acute Stress Signals

The similar symptoms reported for patients with PTSD and those exposed to anti-AChEs pointed to ACh as the common inducer of these delayed responses. Several pieces of evidence support this conclusion.

1. Robust elevations of ACh levels occur under both PTSD and anti-AChE intoxication.
2. ACh interaction with an ACh receptor(s) elevates intracellular Ca^{2+} levels, inducing the transcription of genes that have promoters with Ca^{2+} response elements (CRE motifs).
3. One of these Ca^{2+} -induced genes encodes the early immediate transcription factor *c-Fos*. Interestingly, *c-Fos* expression is elevated drastically within minutes under stress.

Both the choline acetyltransferase gene, *ChAT*, and the AChE gene, *ACHE*, are subject to stress-induced regulation. The first intron in the *ChAT* gene includes the coding sequence for the vesicular ACh transporter VACHT. Together they form the cholinergic locus. Therefore, these two mRNA transcripts are coregulated, and the early immediate gene (IEG) *c-Fos* can control both the initiation and the termination of cholinergic neurotransmission. This operates through regulating ACh production by *ChAT*, its vesicle packaging by VACHT, and its hydrolysis by AChE. Measurements of mRNA levels in conjunction with neurophysiology tests have indeed demonstrated the suppression within 3 h of the initial cholinergic hyperexcitation, most likely through the suppression of ACh synthesis (through decreases in *ChAT* and VACHT mRNA levels) and the enhancement of its hydrolysis (through increases in AChE mRNA levels) within 30 min postinsult. However, the association of stress processes with the bimodal suppression of the initial increase in released ACh does not tell us how

specific *c-Fos* dimers sort out their respective promoters to induce the observed opposing effects. Moreover, it is not clear how much of the increased intracellular Ca^{2+} is due to influx through specific channels in the plasma membrane and how much of it results from Ca^{2+} release from intracellular stores. In addition, other IEGs may also contribute to the observed phenomenon. For example, the activation of muscarinic cholinergic receptors induces multiple IEGs. Also, the time scale of the Ca^{2+} -induced transcriptional response should be refined. Most important, the initial phase of the feedback response probably leads to delayed cascades of the transcription of other relevant genes, and both the scope and the nature of these secondary phenomena are still unknown.

Hypothalamic-Pituitary-Adrenal Involvement in Acute Stress Responses

Stress-induced activation of the hypothalamic-pituitary-adrenal (HPA) axis enhances the release of adrenocorticotrophic hormone (ACTH) from the hypothalamus and of corticosterone from the adrenals. Corticosterone, in turn, stimulates the production of several cytokines, among them interleukin (IL)-1, tumor necrosis factor (TNF), and IL-6. Tracey and colleagues have found that ACh controls the production of pro-inflammatory cytokines by tissue macrophages. This suggests that AChE overproduction and the resultant reduction in ACh levels induce an increase in the inflammatory reactions, as summarized in [Figure 1](#). Thus, the capacity for AChE overproduction under stress affects both peripheral and brain-mediated responses. When produced in the brain, IL-1 interacts with IL-1 type I and type II receptors. This induces *c-Fos* overproduction, which further potentiates the peripheral production of steroid hormones. Therefore, many responses evoked by stress depend on the activation of the HPA axis, are prevented by adrenalectomy, and can be mimicked by exogenous corticosterone (or cortisol in humans). Glucocorticoids induce increases in intracellular Ca^{2+} within the hippocampal neurons, inducing neurotoxic effects in the brain, mainly atrophy and degeneration of hippocampal neurons that are enriched in glucocorticoid receptors. Brain endothelial cells also express IL-1 receptors, suggesting a link between this cytokine, the BBB, and the HPA axis.

Not all stress responses are totally dependent on adrenal functions. An example is the cholinergic feedback response to stress and anti-AChEs described earlier, which apparently includes an HPA-independent component. This is supported by the following:

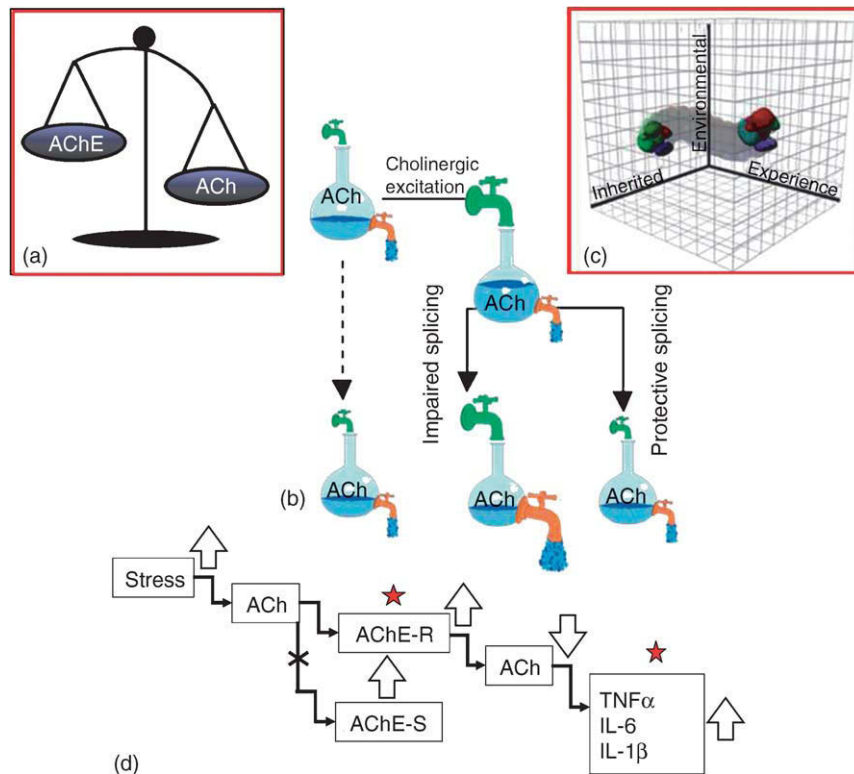


Figure 1 Cholinergic reactions to stress. a, Acetylcholinesterase (AChE) and acetylcholine (ACh) balance determines the cholinergic signaling status; b, cholinergic excitation modifies the alternative splicing of AChE-mRNA transcripts; c, origins of risk factors affecting stress reactions; d, links among stress, cholinergic signaling, and inflammatory reactions. In b, cholinergic synapses are presented as glass beakers, with ACh inflow shown as the top taps and ACh outflow as water streaming out. Impaired splicing (i.e., accumulated synaptic AChE-S) potentiates the outflow, whereas protective splicing (production of AChE-R) maintains balanced flow. When mimicked in transgenic mice, impaired splicing exacerbates stress neuropathology, whereas protective splicing ablates it. In c, a schematic graph shows inherited, experience-derived (e.g., increased body mass index), and environmental measures (e.g., insecticides exposure) affecting the level of risk for stress-induced damages. In d, stress induces ACh release, which induces a shift from synaptic AChE-S to the stress-induced AChE-R splice variant. The excess AChE reduces ACh, which relieves the control over production by tissue-residing macrophages of pro-inflammatory cytokines. Therefore, AChE-R levels are associated with those of pro-inflammatory cytokines (stars). AChE-R, readthrough variant of acetylcholinesterase; AChE-S, synaptic form of acetylcholinesterase; IL, interleukin; TNF α , tumor necrosis factor α .

1. Acute ACh release and its delayed suppression occurred in adrenalectomized rats.
2. Stress-induced overproduction of *c-Fos* mRNA occurs in adrenalectomized mice.
3. The transcriptional feedback response occurs in brain slices.
4. Transgenic mice with a cholinergic, but not HPA, imbalance show impaired BBB and stress neuropathology.

Altogether, this information calls for further research into the specific contributions of the HPA and internal brain processes to the cholinergic feedback response.

Poststress Changes Are Long-Lasting

Alternative splicing produces three AChE isoforms. Of these, the readthrough variant (AChE-R) encodes

a hydrophilic enzyme form secreted as a globular monomer following stress events. The amphiphilic soluble properties of the stress-induced AChE isoform may provide easy accessibility to extracellular spaces flooded under stress with excess ACh. This form of the enzyme is therefore best suited for suppressing the activation of ACh receptors that is associated with stress or anti-AChE exposures.

Biochemical analyses of catalytically active AChE demonstrated that globular monomers are produced following treatment with diisopropylfluorophosphate (DFP; a potent AChE inhibitor) and in the cerebrospinal fluid (CSF) of Alzheimer's disease patients.

In the nonstressed brain, AChE mRNA is positioned around the nucleus of cortical neuronal somata. Following stress or under anti-AChEs, AChE-mRNA appears throughout the cell body, extending to prolonged dendrites. Intracellular changes in the

distribution of mRNAs are often associated with modified functions of their protein products. Therefore, Meshorer and coworkers postulated in 2002 that this altered localization may be physiologically significant.

The poststress accumulation of AChE takes place primarily in cholinceptive brain regions involved in learning and memory, such as the cortex and the hippocampus. The accumulation of the enzyme in the hippocampus is relatively short-lived (a few hours). This suggests the cessation of the poststress AChE-mRNA transcription (or the enhanced degradation of AChE-mRNA or protein) in brain regions that extend cholinergic afferents to the hippocampus. In contrast, cortical AChE levels remain significantly elevated over 3 days poststress. This may reflect continuous local production in cortical cholinergic interneurons. The initiation of CNS stress responses can thus lead, within milliseconds and continuing for several days at least, to AChE excess in the mammalian cortex. Interaction of the stress-induced AChE with signaling kinases can explain the changes induced in neuronal functioning.

Acetylcholinesterase Accumulation Confers Protection against Anticholinesterases

The accumulation of brain AChE increases the concentration of binding sites for anti-AChEs. For

example, Kaufer et al. demonstrated in 1998 that pyridostigmine doses that reduce body temperature in normal, untreated FVB/N mice have no effect when administered 24 h after stress treatment. Similarly, transgenic mice overexpressing AChE in their brain neurons also display resistance to the hypothermic response to pyridostigmine. This supports the conclusion that the stress-induced resistance to anti-AChEs is due to AChE accumulation because elevated AChE levels persist in the brains of these mice. Pyridostigmine protects against the lethal effects of chemical warfare agents that are anti-AChEs. The protection conferred by pyridostigmine may therefore be related, at least in part, to AChE overproduction that is induced by pyridostigmine during the first few hours following treatment.

Acetylcholinesterase Gene Expression Associates with State and Trait Anxiety

Recent genotype–phenotype interaction studies support the notion of AChE interactions with stress responses. In the U.S. population, volunteers with polymorphisms debilitating the functioning of the *ACHE* gene or the adjacent paraoxonase *PON1* gene on the long arm of chromosome 7 were found to display increased trait anxiety or the tendency for state anxiety, as shown in Figure 2.

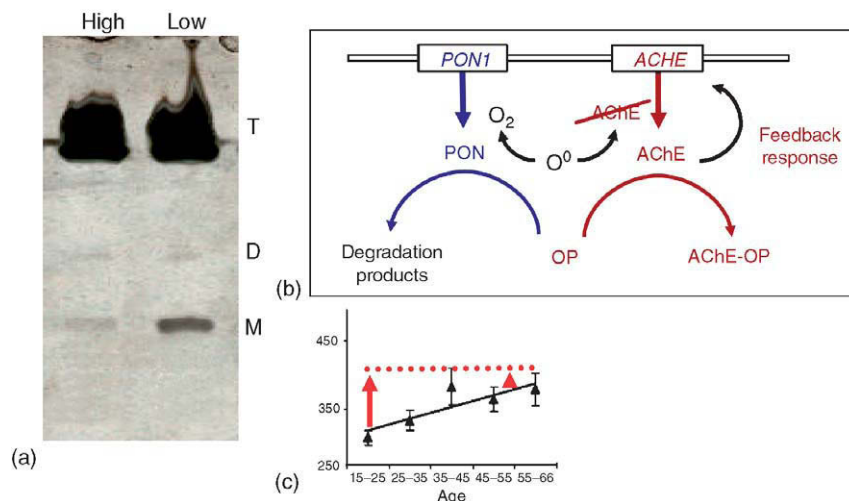


Figure 2 Association of AChE and PON variations with state and trait anxiety. a, Gel electrophoresis of serum samples from volunteers with high or low trait anxiety; b, the putative genomic link; c, age-dependent changes in the capacity for AChE overproduction ($R^2 = 0.73$). In a, note that subjects with low anxiety show more AChE monomers (M) than those with high anxiety, although tetramers (T) and dimers (D) remain unmodified ($N = 472$ volunteers). Thus, the alternative splicing process yielding AChE-R monomers links to trait anxiety. In b, the *ACHE* and *PON1* genes are positioned close to one another on the long arm of chromosome 7. Exposure of AChE to organophosphate inhibitors (OP) inhibits its activity by forming irreversible AChE–OP complexes. This induces a feedback response leading to AChE overproduction. However, paraoxonase (PON) degrades OPs, thus protecting the AChE protein. In addition, PON functions as a peroxidase, protecting blood proteins (AChE included) from oxidative stress. In c, blood AChE levels increase by over 25% between the ages 15 and 66, suggesting a progressive limitation in the capacity for stress-induced AChE overproduction (arrows).

Delayed Nervous System Deterioration under Persistent Acetylcholinesterase Overexpression

Apart from its capacity to hydrolyze ACh, the AChE protein possesses other, noncatalytic activities. These are apparently related to neurite extension in cholinergic as well as dopaminergic neurons and are fully sustained in genetically inactivated AChE. This suggests that the structural effects of AChE might also operate when the AChE protein is inactivated catalytically by anti-AChEs. The long-term outcome of neuronal AChE excess was studied in patients with the autoimmune disease myasthenia gravis and in rats with experimental autoimmune myasthenia. In both cases, the chronic overproduction of AChE-R was observed. Antisense oligonucleotide suppression of this excess enzyme improved neuromuscular functioning in treated rats. Parallel outcomes in clinical trials was reported at the U.S. Society of Neurology in 2004. In the neuromuscular junctions of AChE transgenic mice (the synapses connecting motor neurons with the innervated muscle), impaired

neuromuscular transmission was observed. Similar phenomena were reported for patients with neurodegenerative or neuromotor syndromes associated with cholinergic malfunction, such as Down syndrome, Alzheimer's disease, and myasthenia gravis. Indeed, Farchi and co-workers reported the deterioration of the physiological responses of AChE transgenic muscle, reminiscent of those of myasthenia patients.

Exposure to agricultural insecticides, which serves as a case study for the risks involved in exposure to chemical warfare agents, also increases the risk of Parkinson's disease. It was known for several decades that this risk also involves predisposing genotype-phenotype elements. In a study performed in Israel, this increased risk was found to be associated with debilitating polymorphisms in the *ACHE* and the *PON1* genes. Exposed patients showed an increased incidence of such debilitating polymorphisms, suggesting that an inherited interactive weakness of AChE and *PON1* was the cause. These findings, reported by BenMoyal-Segal and co-workers, are schematically presented in Figure 3.

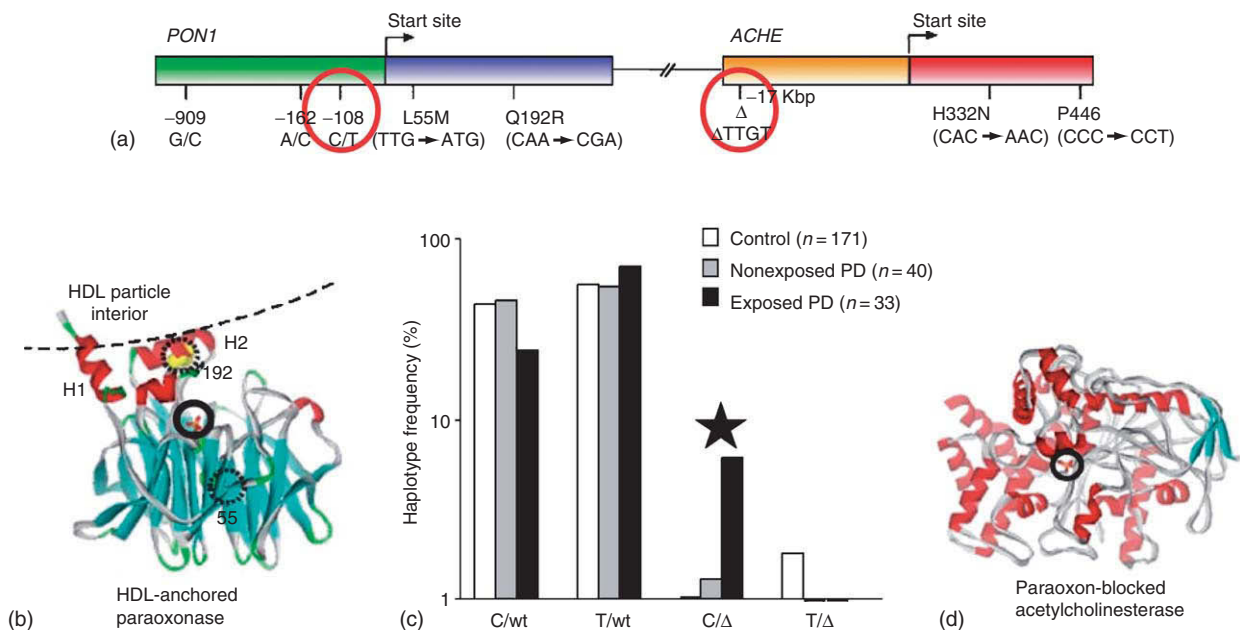


Figure 3 Risk-associated *ACHE/PON1* polymorphisms are strongly overrepresented in exposed Parkinson's disease (PD) patients. a, The *ACHE/PON1* locus; b, the *PON1* protein and frequent polymorphisms; c, graphs of increased incidence of debilitating *ACHE/PON1* haplotype in insecticide-exposed Parkinson's patients; d, the AChE protein and its interaction with paraoxon. In a, a diagram of the adjacent *ACHE* and *PON1* genes is shown, with the tested nucleotide polymorphisms marked below. The two debilitating polymorphisms found frequently in patients with Parkinson's disease who were continuously exposed to agricultural insecticides are circled. In b, the crystal structure of paraoxonase is shown with the two frequent changes in it, in positions 55 and 192. Note that *PON1* interacts with blood lipids (HDLs). The paraoxon binding site is circled. In d, the crystal structure of the AChE protein is shown. Paraoxon is circled in its deep-buried binding site, where its binding blocks AChE functioning irreversibly. Star indicates a statistically significant difference in the incidence of the debilitating *ACHE/PON1* polymorphisms within insecticides-exposed Parkinson's disease patients. C/Δ, carriers of the robust *PON1* and the debilitated AChE variant; C/wt, carriers of the two robust variants; HDL, high-density lipoprotein; T/Δ, carriers of the two debilitated variants; T/wt, carriers of debilitated *PON1* and robust AChE.

Retrospective and Prospective Implications

The feedback response to acute stress and anti-AChEs apparently varies in its intensity with age, body mass index, gender, and ethnic origin. Homozygous carriers of mutations in the *ACHE* homologous *BCHE* gene, encoding serum butyrylcholinesterase (BCHE; e.g., the atypical variant), have anomalous responses to cholinergic drugs. In such individuals, relatively low doses of anti-AChEs may penetrate the brain quite effectively, which results in higher effective doses than intended. This creates a genetic predisposition to adverse responses to anti-AChEs. This inherited sensitivity, and the findings already summarized, raise a concern with regard to the long-term dangers of chronic exposure to household and agricultural insecticides because these are frequently organophosphate or carbamate anti-AChEs. Moreover, the partial symptomatic relief offered by the currently approved Alzheimer's disease anti-AChE drugs appears to be limited in its duration, suggesting that the feedback response to stressor anti-AChEs marks the initiation of progressive damage caused by the production of excess AChE. AChE, an important constituent of cholinergic neurotransmission, may thus cause neuropathological damage when it accumulates in the mammalian brain over long periods. Like many other essential proteins, AChE should be there at the right time and in the right amount or else it can initiate a vicious cycle of long-term harmful responses.

Acknowledgments

Writing this article would not have been possible without the contribution of my co-workers. Special thanks are due to Daniela Kaufer, Alon Friedman, Ella Sklan, and Noa Farchi for research contributions and to Eyal Soreq for artwork. Research was supported by grants from the Israel Science Fund, The Israel Ministry of Defense, and Ester Neuroscience, Ltd.

See Also the Following Articles

Combat Stress Reaction; Persian Gulf War, Stress Effects of; War-Related Posttraumatic Stress Disorder, Treatment of.

Further Reading

Benmoyal-Segal, L., Vander, T., Shifman, S., et al. (2005). Acetylcholinesterase/paraoxonase interactions increase the

risk of insecticide-induced Parkinson's disease. *FASEB Journal* 19(3), 452–454.

Birikh, K., Sklan, E., Shoham, S., et al. (2003). Interaction of "readthrough" acetylcholinesterase with RACK1 and PKC β II correlates with intensified fear-induced conflict behavior. *Proceedings of the National Academy of Sciences USA* 100, 283–288.

Brenner, T., Hamra-Amitay, Y., Evron, T., et al. (2003). The role of readthrough acetylcholinesterase in the pathophysiology of myasthenia gravis. *FASEB Journal* 17(2), 214–222.

Darreh-Shori, T., Hellstrom-Lindahl, E., Flores-Flores, C., et al. (2004). Long-lasting acetylcholinesterase splice variations in anticholinesterase-treated Alzheimer's disease patients. *Journal of Neurochemistry* 88(5), 1102–1113.

Farchi, N., Soreq, H. and Hochner, B. (2003). Chronic acetylcholinesterase overexpression induces multilevelled aberrations in mouse neuromuscular physiology. *Journal of Physiology* 546(1), 165–173.

Friedman, A., Kaufer, D., Shemer, J., et al. (1996). Pyridostigmine brain penetration under stress enhances neuronal excitability and induces early immediate transcriptional response. *Nature Medicine* 2, 1382–1385.

Kaufer, D., Friedman, A., Seidman, S., et al. (1998). Acute stress facilitates long-lasting changes in cholinergic gene expression. *Nature* 393, 373–377.

Meshorer, E., Erb, C., Gazit, R., et al. (2002). Alternative splicing and neuritic mRNA translocation under long-term neuronal hypersensitivity. *Science* 295(5554), 508–512.

Meshorer, E., Biton, I., Ben-Shaul, Y., et al. (2005). Chronic cholinergic imbalances promote brain diffusion and transport abnormalities. *FASEB Journal* 19(8), 910–922.

Nijholt, I., Farchi, N., Kye, M., et al. (2004). Stress-induced alternative splicing of acetylcholinesterase results in enhanced fear memory and long-term potentiation. *Molecular Psychiatry* 9(2), 174–183.

Pollack, Y., Gilboa, A., Ben-Menachem, O., et al. (2005). Acetylcholinesterase inhibitors reduce brain and blood interleukin-1 β production. *Annals of Neurology* 57(5), 741–745.

Sapolsky, R. (2003). Bugs in the brain. *Scientific American* 288(3), 94–97.

Sklan, E. H., Lowenthal, A., Korner, M., et al. (2004). Acetylcholinesterase/paraoxonase genotype and expression predict anxiety scores in Health, Risk Factors, Exercise Training, and Genetics study. *Proceedings of the National Academy of Sciences USA* 101(15), 5512–5517.

Soreq, H. and Seidman, S. (2001). Acetylcholinesterase – new roles for an old actor. *Nature Reviews Neuroscience* 2, 294–302.

Tracey, K. J. (2002). The inflammatory reflex. *Nature* 420(6917), 853–859.

Weinberger, D. R. (2001). Anxiety at the frontier of molecular medicine. *New England Journal of Medicine* 344(16), 1247–1249.

Headache See: Migraine.

Health and Socioeconomic Status

T Chandola and M Marmot

University College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M Marmot and A Feeney, volume 2, pp 313–321, © 2000, Elsevier Inc.

Socioeconomic Gradient in Health

Model

Explanations

Glossary

<i>Psychosocial factors</i>	Measurements of psychological phenomena that relate to specific social environments.
<i>Social exclusion</i>	People or areas that suffer from a combination of linked problems such as poor health, unemployment, inadequate skills, low incomes, poor housing, high crime rates, lack of educational opportunities, and family breakdown.
<i>Socioeconomic gradient</i>	The positive association between greater socioeconomic disadvantage and the amount of death, disability, and disease.
<i>Socioeconomic status</i>	An individual's rank or status in a social hierarchy, usually evaluated with reference to a person's access to and consumption of goods, services, and knowledge in a society.

Socioeconomic Gradient in Health

The association between health and socioeconomic status results in the socioeconomic gradient in health. In rich countries, people who are more socioeconomically disadvantaged have poorer health and higher

levels of chronic and infectious diseases. In poor countries, there is a similar socioeconomic gradient in infectious diseases and an emerging socioeconomic gradient in chronic diseases such as obesity.

These observed socioeconomic gradients are not simply a question of poverty. [Figure 1](#) shows the infant mortality rate by levels of wealth in some of the poorest countries in the world. There is a strong socioeconomic gradient in infant mortality, with richer people, in the poorest countries of the world, having much lower rates of infant mortality compared to poorer people. It is not just the poor who have high infant mortality rates. Those who are not poor, but also not rich, have higher infant mortality rates compared to the richest groups in poor countries.

Furthermore, within rich countries, there is a socioeconomic gradient in health even among relatively well-off sections of society. One of the most famous examples of this comes from the Whitehall and Whitehall II studies of British civil servants. In the original Whitehall study, which started in 1960, the initial finding after 5 years of follow-up was that coronary heart disease mortality was higher in the lower employment grades of the British civil service. After 25 years of follow-up, there was a similar finding of an inverse gradient in mortality: the lower the grade, the higher the mortality. Data from the more recent Whitehall II study (1985–2004) shows a similar inverse gradient in all-cause mortality. In [Figure 2](#), as we proceed lower down the civil service hierarchy of jobs, the risks of mortality increase for all-cause, cardiovascular, and coronary heart disease deaths. It is not just the poor who suffer from higher levels of mortality and disease. Those who cannot be described as being in poverty, such as clerical British civil servants, have poorer health than the people at the top of the socioeconomic hierarchy.

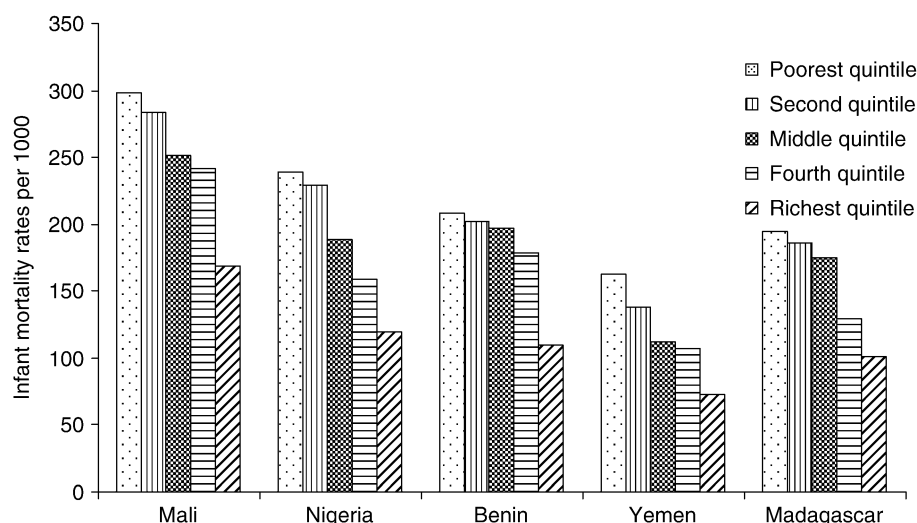


Figure 1 Infant mortality rates by wealth quintiles in Mali, Nigeria, Benin, Yemen and Madagascar. Data from Gwatkin et al. (2000).

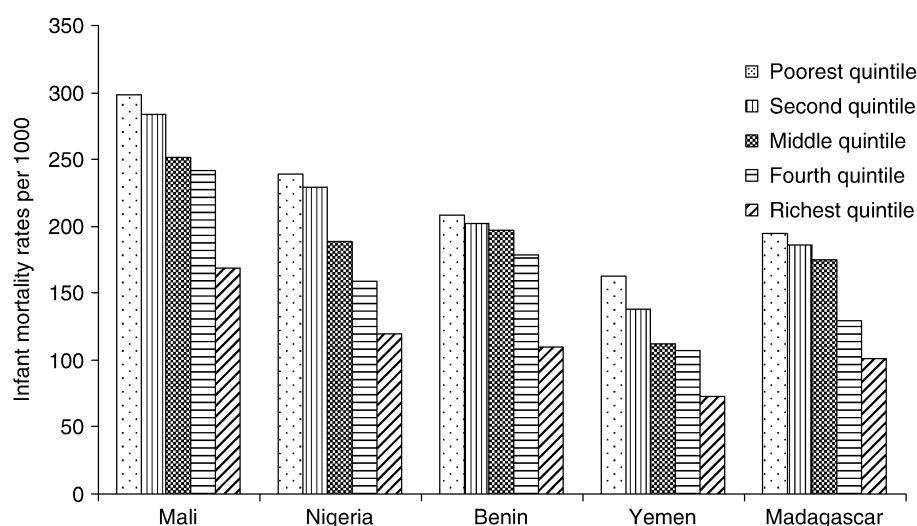


Figure 2 Percentage distribution of deaths by employment grade, adjusted for age and sex: Whitehall II cohort followed up for 19 years.

Socioeconomic status (SES) refers to an individual's rank or status in a social hierarchy, usually evaluated with reference to a person's access to and consumption of goods, services, and knowledge in a society. A number of different measures have been used as proxy indicators of SES, including income, educational level, and occupational prestige. Although they all display similar socioeconomic gradients in health and disease, it is important to distinguish between different measures of SES, as they may affect health through different mechanisms and pathways. For example, education is likely to have indirect effects on health through positive effects on occupation and income as well as on other factors influencing health, including health-related behaviors such as smoking and diet.

Furthermore, measurements of SES need to specify which period of life they are meant to represent. Socioeconomic disadvantage can accumulate over a lifetime, from early childhood, to school level education, to work experiences and living conditions, and on to retirement. Each of these periods of socioeconomic disadvantage can have independent effects on health later on in life. Because being disadvantaged at an earlier period increases the risk of being disadvantaged later in life, measures of SES across the life course are usually correlated with each other. However, a measure of SES in mid-life may not adequately capture the specific risks for poor health that arise from living in disadvantaged socioeconomic circumstances in childhood.

Any explanation of the association between SES and health has to take account of the socioeconomic gradient. This is not just a question about a divide between those living in poverty and others. Explanations need to take account of the socioeconomic gradient in risk factors for disease and mortality and their distribution across the population. Furthermore, the challenge for explanations is not only to explain the socioeconomic gradient, but also to suggest how it is possible for everyone in the population to have the good health enjoyed by those at the top of the socioeconomic ranks. In order to progress toward eliminating socioeconomic gradients in health, we need to understand the causes of this gradient in health. For this, we need to have an appropriate model to describe how SES can get under our skin and affect health.

Model

It is not difficult to understand how poverty in the form of material deprivation – dirty water, poor nutrition, lack of medical care – can account for the

poorer health and higher mortality of poor people. However, poverty alone cannot explain the socioeconomic gradient in health. First, dirty water, lack of calories, and poor medical care cannot account for why British civil servants who work in clerical jobs are over twice as likely to die compared to those in the top ranks of the civil service. Furthermore, the poverty-only model fails to take into account the fact that relief of such material deprivation is not simply a technical matter of providing clean water or better medical care – those resources are socially determined. Even among poorer sections of society, those who are relatively better off are more likely to take advantage of policy interventions to provide better material living conditions, thereby potentially perpetuating and increasing the socioeconomic gradient in health.

In order to reduce inequalities in health, we need to understand the mechanisms and pathways that link the socioeconomic circumstances to health. [Figure 3](#) presents a simplified model of how different causal factors link the social structure to cardiovascular disease outcomes. This helps us to understand the different mediators and pathways linking SES to health.

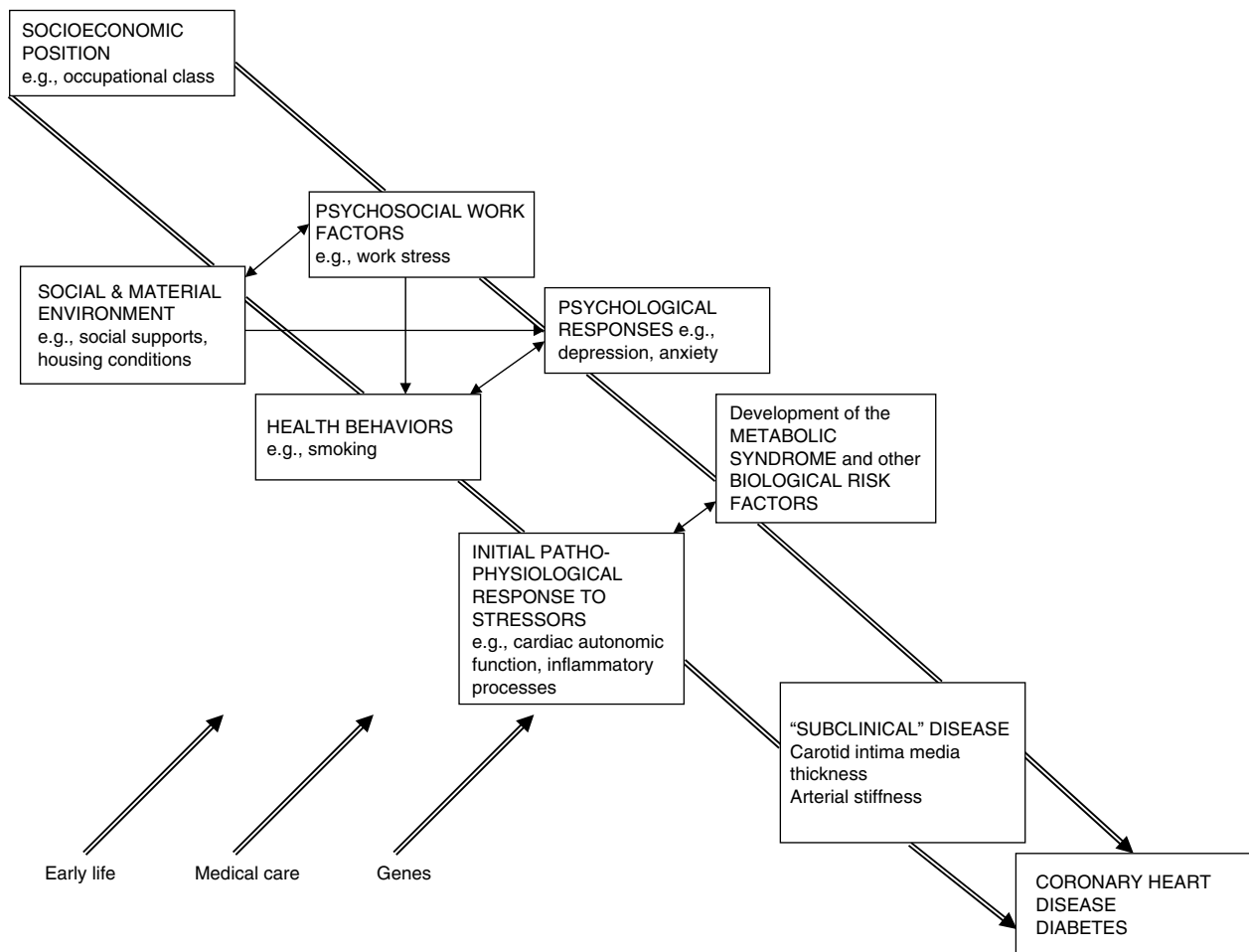


Figure 3 Model of social determinants of health.

This is not meant to be a complete model but a simplified representation. There are further causal factors that may be added to the model, and more causal arrows that may be drawn. This model, in particular, draws our attention to how potential causal factors in the work, social, living, and material environments relate to health behaviors or to more direct psychological processes that influence the psycho-neuro-endocrine-immune systems of the body.

This model helps us to understand how different causal factors interrelate, and it also provides a guide to potential points of intervention for reducing inequalities in health. For example, the metabolic syndrome is a risk factor for heart disease. Treatment of the metabolic syndrome through medical care may reduce heart disease, but may not reduce the socioeconomic gradient in heart disease if poorer people are less likely to have access to or get treated for the metabolic syndrome. Another potential point of intervention could be to change the health behaviors that influence the development of the metabolic syndrome. Getting more people to quit smoking may reduce the development of the syndrome, but telling people to quit smoking alone is not enough. The model suggests that there is a socioeconomic gradient in smoking that alerts us to the socioeconomic factors that underlie smoking behaviors. In order to reduce the socioeconomic gradient in the metabolic syndrome and consequently reduce heart disease, we need to understand the causes of the causes: the living, working, social, and material environmental contexts in which these health behaviors originate and operate. So if stressful work circumstances make it hard to give up smoking, we need to identify interventions that reduce stress at work. If being poor makes it hard to afford quitting smoking aids, then these aids need to be made freely available.

Explanations

The concept of multiple causal factors operating in conjunction with each other, as outlined in [Figure 3](#), is not unique to the development of chronic diseases. Although the tubercle bacillus is the specific cause of tuberculosis, most people infected with the tubercle bacillus do not have clinical tuberculosis. Other factors, such as those associated with poverty and nutrition, influence susceptibility to the disease. In the case of tuberculosis, exposure to an infectious agent is not enough to cause a disease. The exposure, in conjunction with the socioeconomic circumstances, affects the development of the disease.

Similarly, we do not believe that the material, behavioral, psychosocial, and environmental causal

factors outlined in [Figure 3](#) operate in isolation. The division of explanations of the socioeconomic gradient in health into these categories often dismisses the interrelationships and interactions between all of them. So instead, the following sections highlight some of the broad groups of factors related to the social determinants of health, all of which encompass different dimensions and combinations of material, behavioral, psychosocial, and environmental factors.

Early Life

Research from observational and intervention studies shows that the foundations of adult health are laid in early childhood and even before birth. Slow growth, poor emotional support, and socioeconomic deprivation in childhood increase the risk of poor physical, emotional, and cognitive functioning in adulthood. There is a strong patterning of these childhood risk factors for poor adult health by SES.

A number of factors associated with pregnancy, such as nutritional deficiencies during pregnancy, maternal stress, maternal smoking, misuse of drugs and alcohol, insufficient exercise, and inadequate prenatal care, can lead to less than optimal fetal development, low birth weight, and subsequently to an increased risk of heart disease in adulthood. A mother living in poorer socioeconomic circumstances is much more likely to experience some of these risk factors for poor fetal development compared to a mother living in better circumstances.

Infancy is a particularly important period of time for the development of integrated biological systems with long-term consequences for adult health. Research shows that biological systems in infancy are highly malleable through cognitive, emotional, and sensory inputs into the brain. Insecure emotional attachment and poor stimulation in infancy can lead to problematic behaviors in childhood, and adolescence as well as low educational attainment. This in turn increases the risk of social marginalization in adulthood and the ability to obtain employment or steady jobs. Furthermore, health-promoting behaviors in adulthood, such as exercise, healthy diet, and not smoking, are influenced by such parental behaviors in childhood.

It has been argued that parental poverty starts a chain of social and biological risk. Parental poverty increases the risk of poor fetal development, stimulation, and emotional and cognitive inputs in infancy. This leads to slow growth in childhood, poor behavior and educational attainment at school, leading to a raised risk of unemployment, low-status jobs, and lower SES. All of these links in the chain have consequences for physical and mental health in adulthood.

and old age – some of them indirectly, through influencing the risks further down the chain, others more directly influencing adult health. However, it is also possible to break the links between the chains at critical periods of social and biological development. This is the area where interventions on reducing the socioeconomic gradient in health need to concentrate.

Social Exclusion

Absolute poverty, defined in terms of a lack of basic material necessities of life, continues to exist, even in the richest countries. The long-term unemployed, many ethnic minority groups, guest workers, disabled people, refugees, and homeless people are at risk of living in absolute poverty in rich countries. Homeless people living on the streets have the highest rates of premature death. In poorer countries, the proportion of people living in absolute poverty is much larger than in richer countries, and this contributes to the poorer health in these countries relative to richer countries.

However, absolute poverty alone does not explain all of the socioeconomic gradient in health. Relative poverty, often defined as living on less than 60% of the national median income, affects a greater proportion of people in a population than absolute poverty. It could deny people access to decent housing and jobs, education, transport, and other vital factors for full participation in social life.

The stresses of living in absolute and relative poverty are particularly harmful during pregnancy, to babies, and to the elderly, when people are particularly vulnerable. People move in and out of poverty during their lives, so the number of people who experience poverty and social exclusion during their lifetime is far higher than the number of socially excluded people at any one time.

Furthermore, poverty and social exclusion increase the risk of social isolation, influencing divorce and separation, disability, illness, and addiction. Social isolation affects individuals and communities. Communities and neighborhoods that experience high levels of unemployment, deprivation, poor housing, and limited access to services often have far poorer health outcomes than would be expected by aggregating the number of poor people living in those communities. Social isolation in such communities is hard to escape from, as social exclusion and social isolation are intertwined, often forming a vicious cycle.

Unemployment

Unemployed people and their families have substantially increased risks of premature death and poorer physical and mental health compared to those with

stable employment. People who are long-term unemployed, in particular, have much poorer health than the employed. The health effects of unemployment are linked to both its psychological consequences and the financial problems it brings.

The health effects start when people first feel that their jobs are threatened, even before they become unemployed and before any change in their financial circumstances. Job insecurity has been shown to increase the risk of anxiety and depression, sickness absence, smoking, lack of exercise, and heart disease. Unsatisfactory jobs can be just as harmful to health as unemployment; merely having a job will not always protect physical and mental health. As job insecurity grows, it can act as a chronic stressor whose pernicious effects on health increase with the length of exposure.

Disability and ill health increase the risk of unemployment. So the association between unemployment and health could be a result of prior ill health (or selection) or the effects of unemployment on health. There is evidence for both causal models. Any interventions aimed at reducing the association between unemployment and health need to take account of the risk of unemployment through ill health as well as the risk of ill health through unemployment.

Addiction

Alcohol dependence, illicit drug use, and cigarette smoking are all closely associated with socioeconomic disadvantage. Addictions to these substances may help alleviate some of the stressors associated with socioeconomic disadvantage, but in turn, they can generate further problems as well as have long-term consequences on biological systems.

Some of the transitional economies of Central and Eastern Europe experienced great social upheaval following the end of Communist governments at the end of the twentieth century. In these countries, deaths linked to alcohol use – such as accidents, violence, poisoning, injury, and suicide – rose sharply in the 1990s. Some have argued that this sudden increase in alcohol-related deaths was due to the changed socioeconomic circumstances in these countries. In the Whitehall II cohort, a stressful work environment, measured by an imbalance between effort and rewards at work, was found to be a risk factor for alcohol dependence in men.

Smoking is a well-established and important cause of ill health and premature death. Socioeconomic disadvantage is associated with high rates of smoking and very low rates of quitting. Furthermore, smoking is a major drain on poor people's incomes, thereby contributing to further financial deprivation as well

as increasing risks of ill health. Simply increasing the price of cigarettes, to decrease the rates of smoking, may actually increase the socioeconomic gradient in smoking. Tackling addictions requires attention to the social determinants of these addictions.

Food

A good diet and adequate food supply are central for promoting health and well-being. Shortage of food and lack of variety cause malnutrition and deficiency diseases, which in turn are causes of other infectious and chronic diseases. While food deficiency occurs in poorer countries, in richer countries, excess food intake contributes to cardiovascular diseases, diabetes, cancer, and obesity. Furthermore, food poverty can exist side by side with food abundance. Some developing countries are experiencing the double burden of obesity and malnutrition-related diseases.

Economic growth and improvements in housing and sanitation resulted in an epidemiological transition from infectious to chronic diseases, as well as a nutritional transition. Diets have changed to overconsumption of energy-dense fats and sugars, producing more obesity. In rich countries, as well as some developing poor countries, obesity is becoming more common among the poor than the rich. The main dietary difference between socioeconomic groups is the sources of nutrients. The poor tend to substitute cheaper processed foods for fresh food, with much higher levels of energy-dense fats and sugars. People on low incomes are least able to obtain key nutrients from their diets.

The important public health issue is the availability and cost of healthful, nutritious food. Telling people they should eat healthfully is of little use when they do not have access to healthful food, either because they cannot afford it or because they live in food deserts where shops sell processed foods that are inexpensive but unhealthful.

Exercise

As mechanization has reduced the exercise involved in jobs, housework, and commuting, people need to find new ways of building physical activity and exercise into their lives. Regular exercise protects against heart disease and diabetes by limiting obesity. It also promotes a sense of well-being and can protect older people from depression.

There is a strong socioeconomic gradient in leisure-time physical activity – people from lower socioeconomic positions exercise less. This may be because they have less time to spend on exercise, they do not have access to leisure sports facilities, or they live in an environment that does not encourage walking or cycling.

Social Support

Supportive relationships are important emotional and practical resources that contribute to health. Social support may buffer the effects of negative events in a person's life, helping them to cope better with the effects of unemployment or disability. Alternatively, social support may directly contribute to a person's health by making him or her feel cared for, esteemed, and valued. Belonging to a social network and having mutual obligations have a powerful protective effect on health.

Social isolation, or the absence of social networks, is associated with increased rates of premature death and poor chances of survival after a heart attack. People who get less social and emotional support from others are more likely to have depression, poorer well-being, and higher levels of disability from chronic diseases. In addition, negative aspects of support from close relationships can lead to poor mental and physical health. Poverty can contribute to social isolation and negative social support.

Social cohesion, defined as the quality of social relationships and the existence of trust, mutual obligations, and respect in communities, helps to protect people and their health. Some argue that social inequalities tend to corrode good social relations and destroy social cohesion. Societies with high levels of income inequality tend to have less social cohesion and more violent crime. The breakdown of social relations, which is linked to greater inequality, reduces trust and increases levels of violence. Declines in social cohesion in some communities have also been linked to increases in rates of heart disease.

Stress

Disadvantaged social and psychological circumstances can cause long-term stress. Continuing anxiety, insecurity, low self-esteem, social isolation, and lack of control over work and home life have powerful effects on health. Such psychosocial risks accumulate during life and increase the chances of poor mental health and premature death. People in lower socioeconomic groups are more likely to experience these psychosocial stressors.

Psychosocial and other stressors activate hormones and the nervous system by triggering the fight-or-flight response: raising the heart rate, mobilizing stored energy, diverting blood to muscles, and increasing alertness. This stress response diverts energy and resources away from many physiological processes important to long-term health maintenance. While biological stress responses to acute stressors are healthy, both the cardiovascular and immune systems may be affected if there are repeated exposures

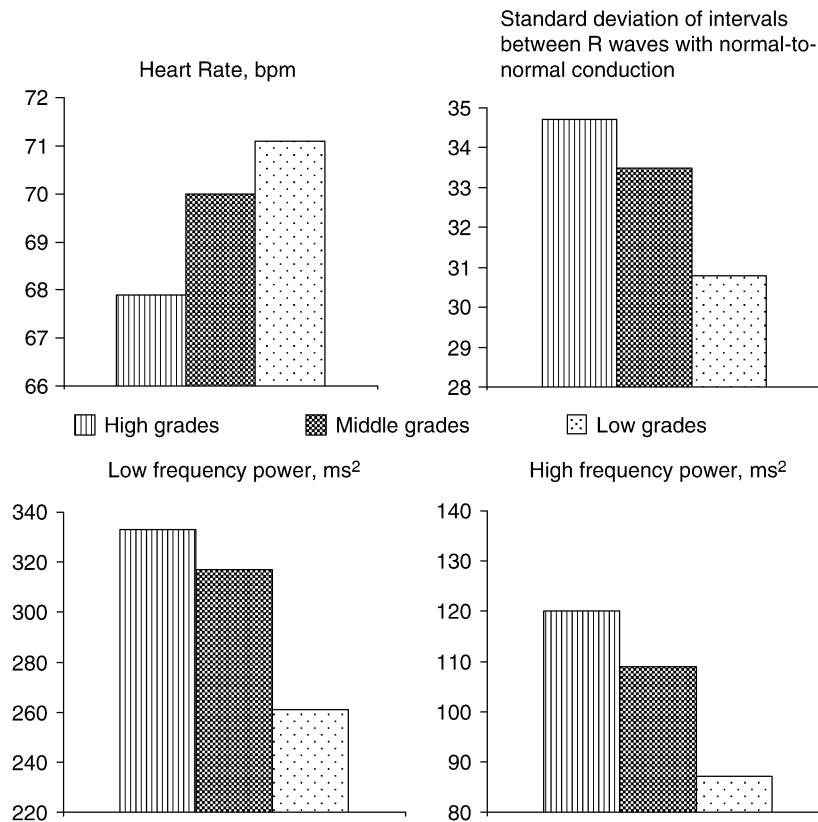


Figure 4 Heart rate variability by employment grade. Data from Hemingway et al. (2005).

to psychosocial stressors or if it goes on for too long. In the Whitehall II cohort, men from lower civil service grades had higher heart rates and lower heart rate variability, indicating poorer autonomic function (see Figure 4). In addition, those with low job control, an indicator of stress at work, were more likely to have impaired autonomic function.

Policy

Health policy was once thought to be about the provision and funding of medical care. This is now changing. While medical care can prolong survival and improve prognosis after some serious diseases, the socioeconomic conditions that make people ill in the first place are more important for the health of the population.

A review of policies in European countries identified several that took action on the social determinants of health, including policies on taxation, old age pensions, sickness benefits, maternity or child benefits, unemployment benefits, housing policies, labor markets, and community facilities. These policies may not be directly related to health, but the evidence suggests that they form part of the broader causes that indirectly influence health. The World Health Organization has set up an independent Commission on the Social

Determinants of Health, with the mission to influence public policy, both national and global, to take into account the evidence on the social determinants of health and interventions and policies to address them.

Policies at all levels of government, and between national governments, need to take account of the evidence on the social determinants of health. Improving access to medical care is important for reducing the socioeconomic gradient in health. However, policies relating to poverty reduction, education, work stress, transport, unemployment, addiction, and social exclusion – policies that cut across the whole business of government – are also needed.

The evidence suggests that telling individuals what they should do to improve their health does not reduce socioeconomic gradients in health. Understanding the causes of the causes implies emphasizing the social determinants of these health-promoting behaviors and recommending social, political, and economic interventions leading to healthier behaviors and healthier outcomes.

See Also the Following Articles

Economic Factors and Stress; Psychosocial Factors and Stress; Social Networks and Social Isolation.

Further Reading

- Adler, N., Marmot, M. and McEwen, B. (eds.) (1999). *Socioeconomic status and health in industrial nations: social, psychological and biological pathways* (vol. 896). New York: New York Academy of Sciences.
- Bartley, M. (2004). *Health inequality*. Cambridge: Polity Press.
- Berkman, L. F. and Kawachi, I. (eds.) (2000). *Social epidemiology*. New York: Oxford University Press.
- Brunner, E. J. (1997). Socioeconomic determinants of health: stress and the biology of inequality. *British Medical Journal* 314, 1472–1476.
- Gwatkin, D. R., Rustein, S., Johnson, K., Pande, R. and Wagstaff, A. (2000). *Socioeconomic differences in health, nutrition and population: health, nutrition and population discussion paper*. Washington, D.C.: World Bank.
- Hemingway, H., Shipley, M., Brunner, E., Britton, A., Malik, M. and Marmot, M. G. (2005). Does autonomic function link social position to coronary risk? The Whitehall II Study. *Circulation* 111, 3071–3077.
- Krieger, N. (2001). A glossary for social epidemiology. *Journal of Epidemiology & Community Health* 55, 693–700.
- Kuh, D. and Ben-Shlomo, Y. (1997). *A life course approach to chronic disease epidemiology*. Oxford, UK: Oxford University Press.
- Marmot, M. (2004). *The status syndrome*. New York: Henry Holt.
- Marmot, M. (2005). Social determinants of health inequalities. *The Lancet* 365(9464), 1099–1104.
- Marmot, M. and Wilkinson, R. (eds.) (2006). *Social determinants of health* (2nd edn.). Oxford: Oxford University Press.
- Wilkinson, R. and Marmot, M. (eds.) (2003). *Social determinants of health: the solid facts* (2nd edn.). Copenhagen: World Health Organisation.

Health Behavior and Stress

A Steptoe

University College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A Steptoe, volume 2, pp 322–325, © 2000, Elsevier Inc.

Definition of Health Behavior

Importance of Links with Stress

Different Behavior–Stress Relationships

Health Behaviors as Coping Responses

Study Designs in the Investigation of Health Behavior

Conclusions

Glossary

Health risk behavior

An action carried out by people with a frequency or intensity that increases risk of disease or injury, whether or not the person is aware of the link between the activity and risk of disease or injury.

Positive health behavior

An action that may help prevent disease, detect disease and disability at an early stage, promote and enhance health, or protect from risk of injury.

Definition of Health Behavior

There are two broad classes of health-related behaviors: risk behaviors and positive health behaviors. A health risk behavior can be defined as an activity carried out by people with a frequency or intensity that increases risk of disease or injury, whether or not the person is aware of the link between the activity and risk of disease or injury. Common health risk behaviors include smoking, drinking and driving, drug abuse, and certain sexual practices. The definition of risk behavior recognizes that “dose” is important, as some activities carry no risk in moderation but are problematic when carried out at a high level (e.g., alcohol consumption).

Positive health behaviors are activities that may help prevent disease, detect disease and disability at an early stage, promote and enhance health, and protect from risk of injury. Such behaviors are defined by medical, scientific, and social consensus and are not immutable. As knowledge evolves and the nature of environmental risks is modified, different actions may come to be defined as positive health behaviors. Typical positive health behaviors are regular physical exercise, avoidance of dietary fat, consumption of fruit and vegetables, use of sunscreen protection,

breast self-examination, regular dental care, and use of vehicle seat belts.

Importance of Links with Stress

Not all health behaviors are related to stress. Even when behaviors are linked with stress, there are typically many other influences on the frequency and intensity of the activity. Health behaviors are embedded within a complex sociocultural matrix, with factors such as tradition, legislation, taxation, income, and systems of provision all influencing lifestyle and health behavior. Thus, alcohol use is not only modified by stress, but also is related to religious and social traditions, costs relative to disposable income, family experiences and peer group influences, and individual attitudes and habits. Leisure-time physical activity can be used as a way of coping with stress, but is also affected by other motives such as interest in sport and concern about appearance and by availability of a safe place to exercise, cost, time constraints, health status, and cultural attitudes.

Nevertheless, stress does have a significant role in many health behaviors, and this can be important for understanding associations between stress and health. The reason is that if stress does increase the occurrence of a health risk behavior or decreases a positive health behavior, then these behaviors may mediate the influence of stress on health risk. Indeed, changes in behavior sometimes play a more important role in the stress–health relationship than neuroendocrine, autonomic, and immunological responses. For instance, one study evaluated the impact of work stress on the development of hypertension (high blood pressure) over a 5 year period among male air traffic controllers. Some 20% of the air traffic controllers who initially had normal blood pressure developed hypertension during the study. This might be assumed to have been the result of sustained neuroendocrine and autonomic activation elicited by working in this responsible and stressful job. However, it was found that alcohol consumption also rose in individuals who became hypertensive, and it was this factor that accounted for the increase in blood pressure. Thus, the association between work stress and cardiovascular disease was mediated by a health-related behavior rather than by activation of biological stress pathways.

Other examples can be found in the epidemiological literature on psychosocial factors and health risk. A study reported by Everson and colleagues of a large sample of middle-aged men in Finland explored associations between hostility and cardiovascular disease mortality over a 9 year period. It was found that high

levels of hostility measured at the beginning of the study predicted cardiovascular mortality independently of known risk factors such as cholesterol and socioeconomic status. However, when smoking, alcohol consumption, body mass index, and physical activity were taken into account, the association between hostility and mortality was diminished substantially. This suggests that the influence of hostility was mediated not by neuroendocrine pathways, but rather by differences in health behaviors. The more hostile individuals were also those who smoked, drank alcohol excessively, and were overweight and physically inactive, which accounted for the increase in cardiovascular risk.

Health behaviors also influence links between stress and health outcomes in patients with medical conditions. One important health behavior in many illness states is adherence to medication and health advice. Low adherence can result in ineffective treatment and risk to future health. Poor adherence is affected by life stress and by states of subjective distress and depression. This might explain why, for instance, cardiac patients who suffer depression and distress following a myocardial infarction are at increased risk for recurrent cardiac events.

Thus, the relationship between health behavior and stress is of interest not only in itself, but also for its role in mediating health effects.

Different Behavior–Stress Relationships

The association between stress and health behavior can take several forms, which complicates systematic investigation, particularly in humans. The different possibilities are outlined in [Figure 1](#). The first model ([Figure 1a](#)) shows that behavior change is one aspect of the psychobiological stress response. Just as stressful episodes elicit distress and physiological adjustments, they also stimulate changes in behavior. Stressful life events, for instance, may lead to increases in cigarette smoking or dietary fat intake, or an increased likelihood of drinking and driving, in the same way that they elicit upsetting emotions. The different elements of psychobiological stress responses are typically coupled only loosely. Hence, although changes in health behavior would be expected to correlate with an increased distress in response to stressful episodes, the association may be relatively weak.

The second possibility is that changes in health behavior are stimulated by the distress that people experience during a stressful episode in their lives ([Figure 1b](#)). For example, increased alcohol consumption may occur in response to family conflict in

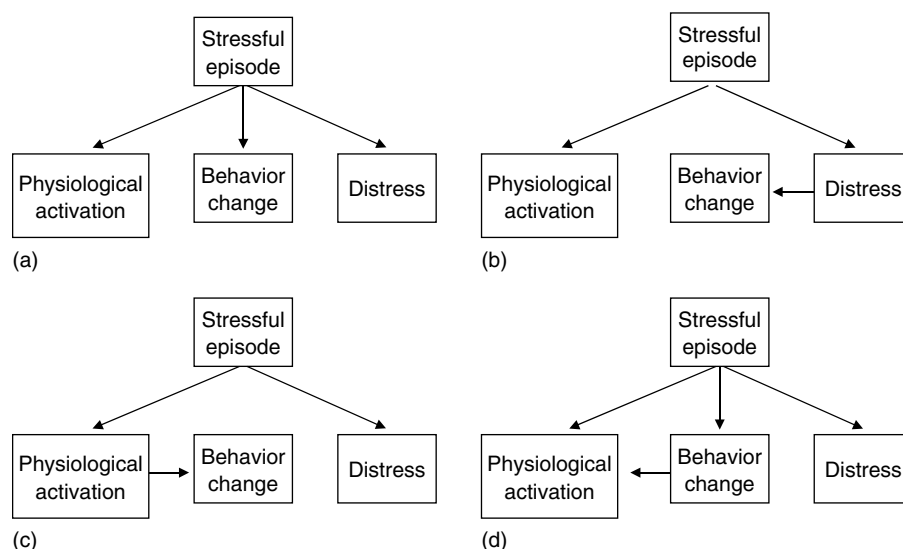


Figure 1 Stress and health behaviors.

individuals who become very depressed or anxious. Under these circumstances, the changes in health behavior arise not as a direct consequence of exposure to stressors, but indirectly as a result of mood change. If a person experiencing a threatening event does not become distressed or copes well with the event, there may be no change in behavior. A threshold of distress may exist below which no alteration in behavior will be measurable.

Third, the physiological adjustments that take place as part of the stress response may provoke changes in health behavior (Figure 1c). It has been proposed, for example, that elevated glucocorticoids in response to stress modulate the control of feeding and preference for foods high in fat and sucrose, and that these serve in turn to dampen adrenocortical and sympathetic nervous system responses. This pathway may contribute to the effects of stress on long-term fat consumption and hence on cardiovascular and metabolic disease risk. When such a process operates, the behavior change will take place only secondarily and not directly as part of the stress response.

Health behaviors can, in turn, influence the magnitude of the physiological adjustments that contribute to stress responses (Figure 1d). Alcohol consumption, smoking, and diet all have effects on immune function. If changes in these behaviors take place during stressful episodes, they may modulate the extent of immunosuppression during stress. Similarly, regular physical exercise has been shown in some studies to reduce the magnitude of cardiovascular and affective reactions to stress, whereas functional foods may alter physiological adjustments by increasing the bioavailability of neurotransmitter substrates.

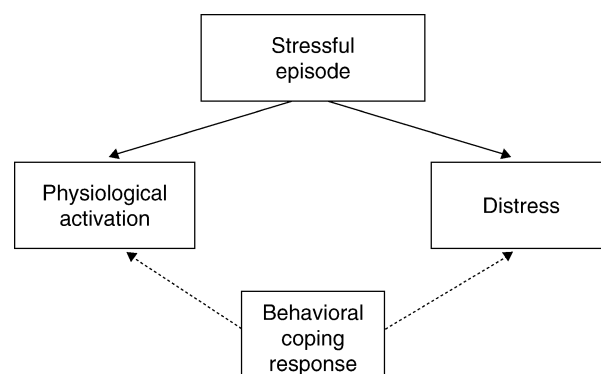


Figure 2 Health behavior change as a coping response.

Health Behaviors as Coping Responses

Thus far, changes in health behavior have been described as direct or indirect components of the stress response. A further possibility is that people change their health behavior in an effort to cope with distress and ameliorate their negative emotional states (Figure 2). Health behavior change may therefore not be part of the stress response, but may occur as one element of the way in which people attempt to manage stress responses. There are several familiar examples of health behavior changes as coping responses. Some people turn to comfort foods such as chocolate when upset, whereas others engage in vigorous physical exercise in order to improve their mood or distract themselves from the troubling situation. These health behaviors may be effective in managing stress. Thus, it has been found that vigorous

physical activity in smokers who are trying to quit can help ameliorate sensations of craving. But in other cases, the beneficial effects of the behavior as a coping response may be more apparent than real. For instance, a large proportion of smokers believe that smoking helps relieve tension and manage stress. However, studies indicate that smoking is associated with poor well-being rather than positive states and that giving up smoking leads to reductions in perceived stress. While smoking may aid mental concentration by increasing alertness in the short term, there is little evidence that it is an effective coping behavior in most circumstances. In Western society, alcohol is sometimes used as a palliative; people drink to cope and disengage from their troubles. Unfortunately, this palliative also has adverse effects and may ultimately become a source of stress in itself.

These different associations between health behavior and stress are not mutually exclusive. An increase in alcohol consumption may, for example, be a direct response to stress (Figure 1a), as well as being a coping behavior (Figure 2). Links between stress and health behavior are therefore not straightforward, which in turn has implications for research designs.

Study Designs in the Investigation of Health Behavior

Experimental studies in animals and humans can be used to tease out associations between stress and health behavior. They allow antecedent conditions to be controlled carefully so that the impact of defined stimuli can be explored systematically. However, because much interest is focused on mental and physical health outcomes, observational studies of naturally occurring stressors are also essential. A major problem that has limited progress in the investigation of stress and health behavior has been the use of cross-sectional designs in naturalistic studies. These prevent conclusions from being drawn about causality. For example, a study of drug dependency might show that drug-dependent individuals frequently live in deprived conditions, experience a high frequency of adverse life events such as family disruption, divorce, and job loss, and have poor social supports. There is therefore an association between stress and drug dependency, but are these adverse social conditions the effects of drug dependence or did they stimulate drug use in the first place? Without longitudinal or prospective study designs, cause and effect cannot be distinguished.

Another design problem in some studies of health behavior is sample selection. A study that focuses on individuals with problems related to a particular behavior may produce results that are not broadly

representative. Many studies of stress and alcohol use, for instance, have been carried out with samples of people suffering from alcohol-related problems. There may be other people in the population, however, who drink as much alcohol as the people in these samples without suffering from health problems. Studies of individuals with health problems will be biased toward observing stress-related effects that would not manifest with a population-based sampling frame.

Conclusions

Health risk and positive health behaviors show complex relationships with stress, as they can be both elements of the stress response and coping behaviors. A combination of research designs is needed to delineate these associations, and the inferences that can be drawn from cross-sectional studies are very limited. Health behaviors require detailed investigation in the context of stress research, as they contribute to many of the associations between stressful phenomena and health risk identified in the literature.

See Also the Following Articles

Alcohol and Stress: Social and Psychological Aspects; Diet and Stress, Non-Psychiatric; Exercise; Psychosocial Factors and Stress; Smoking and Stress.

Further Reading

- Dallman, M. F., Pecoraro, N. C. and la Fleur, S. E. (2005). Chronic stress and comfort foods: self-medication and abdominal obesity. *Brain, Behavior and Immunity* **19**, 275–280.
- DeFrank, R. S., Jenkins, C. D. and Rose, R. M. (1987). A longitudinal investigation of the relationships among alcohol consumption, psychosocial factors, and blood pressure. *Psychosomatic Medicine* **49**, 236–249.
- Everson, S. A., Kauhanen, J., Kaplan, G. A., et al. (1997). Hostility and increased risk of mortality and acute myocardial infarction: the mediating role of behavioral risk factors. *American Journal of Epidemiology* **146**, 142–152.
- Greeno, C. G. and Wing, R. R. (1994). Stress-induced eating. *Psychological Bulletin* **115**, 444–464.
- Kassel, J. D. and Hankin, B. L. (in press) Smoking and depression. In: Steptoe, A. (ed.) *Depression and physical illness*. Cambridge, UK: Cambridge University Press.
- Steptoe, A. and Ayers, S. (2004). Stress, health and illness. In: Sutton, S., Baum, A. & Johnston, M. (eds.) *Sage handbook of health psychology*, pp. 169–196, London: Sage.
- Steptoe, A. and Wardle, J. (2004). Health behaviour: prevalence and links with disease. In: Kaptein, A. & Weinman, J. (eds.) *Health psychology*, pp. 21–51, Oxford: Blackwell.

- Stewart, S. H. (1996). Alcohol abuse in individuals exposed to trauma: a critical review. *Psychological Bulletin* 120, 83–114.
- Ussher, M., West, R., McEwen, A., Taylor, A. and Steptoe, A. (2003). Efficacy of exercise counselling as an aid for smoking cessation: a randomized controlled trial. *Addiction* 98, 523–532.

- Ziegelstein, R. C., Fauerbach, J. A., Stevens, S. S., Romanelli, J., Richter, D. P. and Bush, D. E. (2000). Patients with depression are less likely to follow recommendations to reduce cardiac risk during recovery from a myocardial infarction. *Archives of Internal Medicine* 160, 1818–1823.

Heart Disease/Attack

G J Baker, S Suchday and D S Krantz

Uniformed Services University of the Health Sciences,
Bethesda, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition article
by G J Baker, S Suchday and D S Krantz, volume 2, pp 326–332,
© 2000, Elsevier Inc.

Introduction

Overview of Pathophysiology

Psychosocial Risk Factors for Coronary Heart Disease

Individual Coronary Heart Disease Risk Factors

Cardiac Stress Management Interventions

Conclusion

Glossary

<i>Atherosclerosis</i>	A condition in which fatty plaques build up in the coronary arteries.
<i>Cardiovascular reactivity</i>	An excessive hemodynamic response to stressors.
<i>Coronary heart disease</i>	A usually symptomatic condition in which cardiac function is impaired (thought to result from atherosclerosis).
<i>Myocardial infarction (MI)</i>	A heart attack.
<i>Myocardial ischemia</i>	The inadequate supply of blood to cardiac tissue.

Introduction

Chronic cardiovascular diseases (CVD), including coronary heart disease (CHD), high blood pressure, and stroke, is the leading cause of death in developed countries. CVD is caused by the interaction of physiological, environmental, and behavioral variables, many of which are a function of individuals' lifestyles

and can, to some extent, be controlled and modified by the individual.

Despite the decline in cardiovascular mortality, CVD continues to be the leading cause of death in the United States, and an increasing body of evidence suggests that exposure to chronic and acute psychological stress may play a role in the development and/or triggering of cardiovascular pathology. This article provides a selective overview of current knowledge regarding the role of acute and chronic stress as risk factors for coronary heart disease.

Overview of Pathophysiology

Coronary atherosclerosis, which is a precursor to symptomatic coronary heart disease, results from the formation of plaques (fatty deposits) on the arterial walls that obstruct blood flow to the heart. Myocardial ischemia results from this inadequate blood supply and is often accompanied by chest pain, also known as angina pectoris. The occurrence of frequent and severe myocardial ischemia increases the risk for cardiac-rhythm disturbances, which can result in sudden cardiac death. Prolonged or severe ischemia, or plaque rupture in the coronary arteries, which may result in complete blockage, can lead to more serious manifestations of CHD. These include death of cardiac tissue (MI), also known as a heart attack, and sudden cardiac death.

Current models of the effects of acute physical and mental stress on cardiac pathophysiology suggest that the activation of the autonomic nervous system induced by behavioral factors may increase the risk of cardiovascular events. **Figure 1** illustrates pathways through which acute mental stress is linked to myocardial ischemia and clinical coronary events such as heart attack or sudden death. These links are thought to exist in individuals with preexisting coronary artery disease or dysfunction, prior MI, or poor

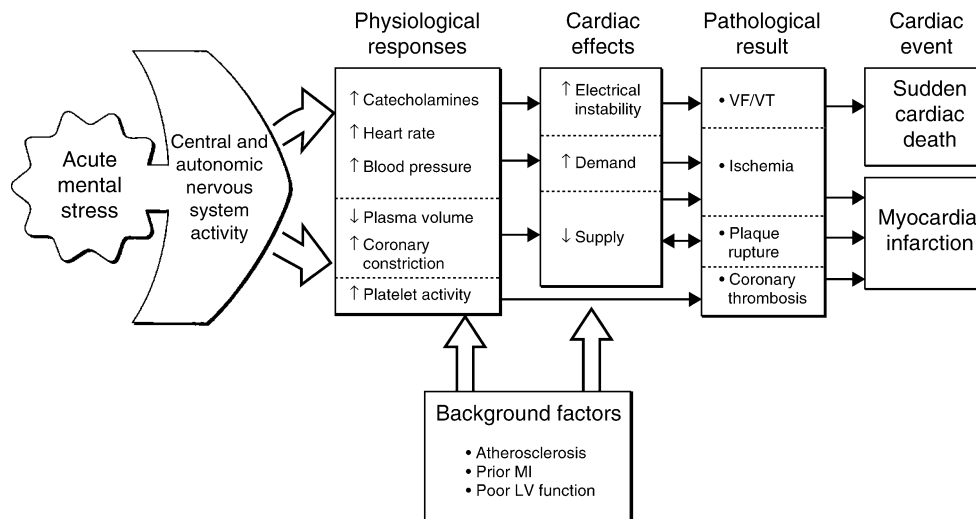


Figure 1 Pathophysiological model of the actions of acute stress as a trigger of myocardial infarction and sudden death in vulnerable individuals. Via its actions on the central and autonomic nervous system in patients with coronary artery disease, stress can produce a cascade of physiological responses that may lead to myocardial ischemia, ventricular fibrillation/tachycardia (VF/VT), plaque rupture, or coronary thrombosis (MI, myocardial infarction; LV, left ventricular). Reprinted from *Cardiology Clinics*, Volume 14(2), May 1996 with permission.

cardiac function, who are therefore at higher risk for these clinical events.

Psychosocial Risk Factors for Coronary Heart Disease

Psychosocial factors that have been implicated as risk factors for developing heart disease include individuals' accessibility to social and economic resources, the work environment, and stressful life circumstances. Animal model studies in primates also suggest that conditions involving psychological stress can promote the development of atherosclerosis. Each of these factors is discussed in the following sections.

Animal Model Studies

A series of studies involving cynomolgus monkeys has been conducted to examine the effects of mental stress on coronary disease. Monkeys not only have similar coronary disease pathology to humans, but they also exhibit many of the same behavioral characteristics that have been cited as potential risk factors, such as competition, aggression, and the formation of social hierarchies. Stressful situations were created for these monkeys by rearranging their social groups periodically, known as the unstable condition. Monkeys were then labeled as either dominant or subordinate based on their behavior patterns in the groups.

In one study, more coronary atherosclerosis developed in dominant male monkeys in the unstable social condition than in their subordinate counterparts. The differences in disease progression between

dominants and subordinates in the stable social condition were less pronounced. Another study, using female monkeys, found more coronary artery disease in subordinate animals than in dominants. Reproductive function was also disrupted in many of the subordinate monkeys, which may have decreased the animals' resistance to coronary disease.

Low Social Support as a Coronary Heart Disease Risk Factor

Lack of social resources, particularly inadequate social support and social isolation, can increase the risk of CVD. Social resources can be categorized as social networks or as social support derived from those networks. Social support is defined as the instrumental, emotional, and informational aid obtained from an individual's social ties and community resources, and it exerts a mediating or moderating impact on health. A lack of supportive networks also impacts health by interfering with practical aid. Research indicates that people with access to strong social networks are at a relatively lower risk of developing CHD. Social networks also exert a beneficial impact on recovery after incidence of disease.

Industrial countries demonstrate a graded, inverse relationship between socioeconomic status (SES) and morbidity and mortality, with higher disease rates appearing in poorer populations. SES has typically been defined using standard measures, including education, income, occupation, or combinations of these three factors. Although it has been shown that the prevalence of standard CHD risk factors (e.g., high

blood pressure and smoking) diminishes with increasing SES, that does not fully explain the increase in coronary mortality attributable to low SES. Regardless of other prognostic factors, social isolation and low household incomes predict an increased risk of cardiovascular death over a 5-year period. Presumably, psychological reactions to MI may be exacerbated by diminished social and economic resources.

Occupational Stress

Karasek and Theorell attempted to define specific characteristics of stressful occupations that promote risk of coronary disease. Specifically, high job strain (defined as conditions of high demand and few opportunities to control outcomes) predicted job stress-related illnesses, especially CHD. The Karasek job demand/control hypothesis explained cardiovascular disease and mortality in studies of Swedish workers and in studies of U.S. men and women.

A study of occupational stress and CHD in women also suggests that the relationship between conditions in the workplace and cardiovascular disease, at least in part, depends on family and work demands on the individual and the amount of control the individual has over these situations. The Framingham Heart Study, a major epidemiologic study of heart disease conducted by the National Institutes of Health, examined these issues. Haynes and Feinleib analyzed data from this study to address the question of whether the increase of women working outside the home has affected their cardiovascular health adversely. An 8-year follow-up of middle-age women in the mid-1960s examined the development of CHD in this population. Women who had been employed outside the home for more than half of their adult lives were compared with housewives and with men. The results indicated that generally there was not a significantly greater risk of subsequent heart disease in working women. However, there was a greater risk of CHD development in clerical workers, working women with children, and women whose bosses were unsupportive. It is thought that the increased risk in these groups is due to low job control, high family demands, or low job satisfaction. It is interesting to note that the risk of CHD increased linearly with the number of children for working women but not for housewives.

Acute Stress as a Trigger of Cardiovascular Symptoms

Research has also focused on the role of acute stress and of emotions as triggers of coronary disease symptomatology in individuals with preexisting disease. One early study observed that 40 out of 100 sudden cardiac death patients had experienced a stressful life

event (i.e., death of a spouse) within the 24 h preceding their deaths and that loss events (e.g., death of a loved one) were more common in sudden death victims than in controls. These studies support the findings of previous research that observed increased mortality of widowers from cardiovascular events.

Research has also shown an increase in cardiovascular deaths and increased rates of MI following disasters and personal traumas. The Iraqi missile attacks on Israel during the initial days of the Gulf War have been the subject of studies that examined the effect of these attacks on fatal and nonfatal cardiac events among the population living close to Tel Aviv. The intensive care unit of a Tel Aviv medical center reported an increase in acute MI cases during the week following the missile attacks (January 17–25, 1991) compared with the week prior to the attacks and with a control period consisting of the same week a year earlier.

An increase in cardiac mortality has been attributed to acute psychological stress following earthquakes and other natural disasters. For example, during the 1994 Los Angeles earthquake, the incidence of critical-care admissions for acute MI reportedly increased (odds ratio 2.4). Unstable angina, however, did not increase in the week following the earthquake compared to the week prior to the disaster, especially for hospitals within 15 miles of the earthquake epicenter.

Psychologically stressful events have also been implicated as environmental triggers occurring prior to the onset of MI. In a study of patients with acute MI, nearly half reported one or more possible triggers, the most common of which was emotional distress. Mittleman and colleagues used a more sophisticated epidemiological methodology in which each patient's activities just prior to MI were compared to a control period defined as his or her own usual levels of activity in order to examine the most common physical and mental triggers of MI. This study of MI patients interviewed a median of 4 days after the MI revealed that 39 out of 1623 patients (2.4%) reported experiencing anger within the 2 h prior to MI onset. The relative risk of MI following an anger episode was more than twice the risk during nonanger periods.

Mental Stress as a Trigger of Ischemia

Although the most important end points for studies of cardiovascular disease are MI and sudden cardiac death, research has also focused on myocardial ischemia as an intermediate measure of the effects of stress on CHD. Studies of myocardial ischemia also have the benefit of being unconfounded by reporting biases

because ischemia often occurs without symptoms that are known to the patient.

Using a structured diary method, Barry and colleagues observed that most of the episodes of daily life ischemia occurred during moderate or low levels of physical activity (e.g., talking on the phone, speaking with a friend, or clerical work). There was a graded relationship between the likelihood of ischemia and the intensity of both mental and physical activities, suggesting that as levels of physical and mental activity become more intense, they are increasingly likely to evoke transient daily life ischemia.

Gabbay and coworkers asked 63 subjects who were undergoing ambulatory electrocardiographic monitoring to fill out a more sophisticated diary. Results showed that ischemia occurred most often during moderately intense physical and mental activities. The amount of ischemic time (5%) was approximately equal for high-intensity physical and mental activities, as compared to only 0.2% of ischemic time in low-intensity activities. In summary, ambulatory electrocardiographic monitoring studies illustrate that various physical and mental activities are associated with the occurrence of ischemia during daily life. Both physical and mental activities can act as potent triggers. Thus, it seems that many features of daily life ischemia may be due to behavioral influences, including daily life stress.

Individual Coronary Heart Disease Risk Factors

Type A Behavior

The type A behavior pattern was originally described by cardiologists Friedman and Rosenman in the 1950s as a behavior pattern characterized by agitation, hostility, rapid speech, and an extremely competitive nature. The contrasting type B behavior pattern consists of a more laid-back style and a lack of the type A characteristics mentioned previously. A structured interview was developed to measure type A behavior based on behaviors such as speech characteristics and subjects' responses to various questions.

In the 1960s and 1970s, many studies were conducted to probe the relationship between type A behavior and heart disease. Most of the studies revealed a correlation between type A behavior and coronary heart disease in both men and women, which is comparable and independent of the effects of smoking and hypertension. For example, two major studies obtained results supporting the findings that type A behavior is a risk factor for heart disease. The Western Collaborative Group Study (WCGS) followed

initially healthy men for 8.5 years. The men were given questionnaires and took part in interviews to determine their type A status at the outset of the study. Those individuals who were identified as type A were more likely to have developed heart disease over the course of the 8.5 years of the study than the type B group. The Framingham Heart Study also showed that type A behavior was a predictor of CHD among white-collar men and women who worked outside of the home.

Since the 1980s, however, most studies have not corroborated the relationship between type A behavior and heart disease. The Multiple Risk Factor Intervention Trial (MRFIT) assessed whether interventions to reduce coronary risk factors, such as high blood pressure, smoking, or high cholesterol levels, decreased the potential for coronary disease in high-risk men and women. After 7 years, the results did not show a relationship between these measures of type A behavior and the incidence of the first heart attack. Further, there were reports from research that indicated that, after a heart attack, type B patients were more likely to die than type A patients. This finding could be explained as the result of healthier type A patients being the ones who initially survived their first heart attack. Regardless, however, this result certainly presented doubts about type A behavior as a coronary risk factor.

It is unclear why these studies resulted in such inconsistent findings. Some have suggested that type A behavior may be a risk factor for younger individuals rather than for older or high-risk individuals, such as the ones tested in the MRFIT study. Still, it seems that there are certain aspects of type A behavior, particularly anger and hostility, that remain correlated with coronary disease, even in studies in which overall type A behavior was not related to CHD.

Hostility and Anger

As noted earlier, it is possible that some particular characteristics of the multifaceted type A behavior are more pertinent risk factors for CHD than others, and perhaps the type A concept as a whole does not constitute increased risk. Factors that have been shown consistently to correlate with CHD are hostility, anger, certain speech characteristics drawn from the structured interview, and not expressing anger or irritation. For example, an analysis of WCGS data showed that the strongest predictors of CHD were, in fact, the potential for hostility, vigorous speech, and reports of frequent anger or irritation. The MRFIT study, which did not show a correlation between type A behavior and coronary disease, also demonstrated a relationship between hostility and an increased risk of CHD.

The Cook and Medley hostility inventory, a scale that measures certain attitudes such as cynicism and mistrust of others, has been shown to relate hostility to the occurrence of heart disease. In one study, a 25-year follow-up of physicians who completed the MMPI while attending medical school revealed that high Cook-Medley scores predicted not only the incidence of CHD but also mortality from all causes. This relationship was independent of the individual effects of other risk factor such as smoking, age, and the presence of high blood pressure. Evidence also suggests that low hostility scores are correlated with a decrease in death rates during the 20-year follow-up of the almost 1900 participants in the Western Electric Study. Research also illustrates that hostility scores are higher in low socioeconomic groups, men, and non-whites in the United States, as well as being correlated positively with smoking prevalence. Therefore, it is possible to suggest that hostility may account for some of the socioeconomic and gender differences in death rates from cardiovascular diseases.

Depression and Exhaustion

Clinical depression, which includes symptoms such as fatigue, exhaustion, and dysphoric mood, is prevalent among patients with CHD and is associated with poor health outcomes in CHD patients. Studies have reported that nearly one in five cardiac patients can be diagnosed as clinically depressed. Depression among post-MI patients, however, goes largely untreated, despite its prevalence. In several studies, the presence of a major depressive episode in CHD patients is associated with poor psychosocial rehabilitation and increased medical morbidity.

Appels and colleagues identified a syndrome of fatigue or exhaustion called vital exhaustion, which overlaps with some of the symptoms of depression. Research indicates that vital exhaustion is predictive of the development or worsening of cardiovascular disease. Further, the predictive value of fatigue or exhaustion cannot be explained by the effects of illness on mood or energy levels. Studies are currently underway to assess whether risk of death and recurrence in post-MI patients can be lessened by treating the symptoms of depression.

Responsiveness (Reactivity) to Acute Stress

It has been known for some time that individual reactions to stress vary widely, and research suggests that physiological responses to stress may be implicated in the development of CHD and/or high blood pressure. Cardiovascular reactivity is measured by assessing changes in heart rate, blood pressure, or other cardiac measures in response to stress, as opposed to the

sole measurement of resting levels of cardiovascular function. The magnitude of these responses varies widely from one individual to another. Some people are so-called hot reactors and display large increases to the challenging task, whereas others show very little or no increase at all. For example, evidence shows that behaviors present in type A individuals are often accompanied by the same types of responses that are believed to link psychosocial stress to CHD.

It may be possible that high levels of stress reactivity is a risk factor for CHD or a marker for other stress-related risk processes. Keys and coworkers followed a group of initially healthy men for 23 years and found that the magnitude of their diastolic blood pressure reactions in a cold pressor test (immersing the hand in very cold water) was a good predictor of subsequent heart disease; in fact, it was a better predictor than many of the standard risk factors. However, a later attempt to replicate these findings was unsuccessful.

Cardiac Stress Management Interventions

Evidence suggests that psychosocial interventions, when combined with the standard medical treatments for these patients and stress management, may be one way to reduce cardiac morbidity and mortality. For example, the Recurrent Coronary Prevention Project (RCPP) assessed the effects of modifying type A characteristics, such as anger, impatience, and irritability, on the recurrence of heart attacks and CHD-related deaths. Patients were divided into three groups: a cardiology counseling treatment group, a combined cardiology counseling and type A-behavior modification group, or a no treatment control group. After 4.5 years of follow-up, the type A-behavior modification group displayed significantly lower rates of heart attack recurrence than the cardiology counseling and control groups.

The Ischemic Heart Disease Life Stress Monitoring Program assigned heart attack patients to either a treatment or a control group. The treatment group was exposed to life stress monitoring and intervention, whereas the control group received only standard medical follow-ups. The results at the 1-year follow-up showed that cardiac deaths decreased by 50% and that this lower death rate persisted for 6 months beyond the end of the study.

Ornish and colleagues conducted an intervention study on 28 patients with documented coronary artery disease. Patients were assigned to either a usual care control group or a treatment group that was given a low-fat vegetarian diet, yoga-based stress

management training, smoking cessation program, and aerobic exercise. After 1 year, the results demonstrated that the treatment group showed slightly less coronary artery blockage, whereas the controls showed a progression in disease. Further, the more compliant patients in the treatment group exhibited greater disease reduction and the least compliant patients showed less change.

Finally, Blumenthal and coworkers examined the effects of an exercise regimen and stress management on CHD. The Myocardial Ischemia Intervention Trial assigned patients to receive 4 months of stress management, 4 months exercise training, or simply their standard medical follow-up care. The patients in the stress management group showed a lower risk of cardiac events in the 38 months of follow-up than the control group. There was also a lower risk associated with the exercise training group, but this did not reach significant levels.

Conclusion

There are many physiological, environmental, and behavioral influences that affect the development of cardiovascular disease. Many of the standard CHD risk factors have behavioral components, and increasingly, evidence points to important psychosocial risk factors for CHD as well, including occupational stress, hostility, and physiological reactivity to stress. In cardiac patients, acute stress, low levels of social support, and depression also seem to have an impact on the development and progression of CHD. However, several studies suggest that stress management programs, when added to usual care medical treatments, can significantly lessen cardiac morbidity in cardiac patients. Thus, it is important to consider more than simply the physiological aspects of coronary disease when discussing treatment and prevention. Greater public awareness of the aspects of coronary disease that are under individuals' control may aid in the struggle to reduce cardiac morbidity and mortality in the United States; however, continued research is needed to examine and explain these phenomena further.

See Also the Following Articles

Anger; Atherosclerosis; Captivity, Recovery from; Eating Disorders and Stress; Sleep, Sleep Disorders, and Stress; Social Stress, Animal Models of; Trauma and Memory.

Further Reading

Barefoot, J. C. and Lipkus, I. M. (1994). The assessment of anger and hostility. In: Siegman, A. W. & Smith, T. W. (eds.)

Anger, hostility, and the heart, pp. 43–66. Hillsdale, NJ: Lawrence Erlbaum.

Barry, J., Selwyn, A. P., Nabel, E. G., et al. (1988). Frequency of ST segment depression produced by mental stress in stable angina pectoris from coronary artery disease. *American Journal of Cardiology* 61, 989–993.

Blumenthal, J. A., Wei, J., Babyak, M., et al. (1997). Stress management and exercise training in cardiac patients with myocardial ischemia: effects on prognosis and on markers of myocardial ischemia. *Archives of Internal Medicine* 157, 2213–2223.

Carney, R. M., Freedland, K. C., Rich, M. W., et al. (1995). Depression as a risk factor for cardiac events in established coronary heart disease: a review of possible mechanisms. *Annals of Behavioral Medicine* 17, 142–149.

Case, R. B., Moss, A. J., Case, N., et al. (1992). Living alone after myocardial infarction. *Journal of the American Medical Association* 267, 515–519.

Cohen, S. and Wills, T. A. (1985). Social support and the buffering hypothesis. *Psychological Bulletin* 98, 310–357.

Gabbay, F. H., Krantz, D. S., Kop, W. J., et al. (1995). Triggers of myocardial ischemia during daily life in patients with coronary artery disease: physical and mental activities, anger, and smoking. *Journal of the American College of Cardiology* 27, 585–592.

Frasure-Smith, N. and Prince, R. (1987). The ischemic heart disease life stress monitoring program. *Psychosomatic Medicine* 5, 485–513.

Friedman, M. and Rosenman, R. H. (1974). *Type A behavior and your heart*. New York: Knopf.

Haynes, S. B. and Feinleib, M. (1980). Women, work, and coronary disease: prospective findings from the Framingham Heart Study. *American Journal of Public Health* 70, 133–141.

Kaplan, G. A. and Keil, J. E. (1993). Socioeconomic factors and cardiovascular disease: a review of the literature. *Circulation* 88, 1973–1998.

Kaplan, J. R., Adams, M. R., Clarkson, T. B., et al. (1984). Psychosocial influences on female “protection” among cynomolgus macaques. *Atherosclerosis* 53, 283–295.

Karasek, R. A. and Theorell, T. G. (1990). *Health, work, stress, productivity, and the reconstruction of working life*. New York: Basic Books.

Keys, A., Taylor, H. L., Blackburn, H., et al. (1971). Mortality and coronary heart disease among men studied for 23 years. *Archives of Internal Medicine* 128, 201–214.

Kop, W. J., Appels, A. P. W. M., de Leon, C. M., et al. (1994). Vital exhaustion predicts new cardiac events after successful coronary angioplasty. *Psychosomatic Medicine* 56, 281–287.

Krantz, D. S., Kop, W. J., Gabbay, F. H., et al. (1996). Circadian variation in ambulatory myocardial ischemia: triggering by daily activities and evidence for an endogenous circadian component. *Circulation* 93, 1364–1371.

Krantz, D. S. and Manuck, S. B. (1984). Acute psychophysiological reactivity and ride of cardiovascular disease: a review and methodological critique. *Psychological Bulletin* 96, 435–464.

- Matthews, K. A. and Haynes, S. G. (1986). Type A behavior pattern and coronary disease risk: update and critical evaluation. *American Journal of Epidemiology* **123**, 923–960.
- Mittleman, M. A., Maclure, M., Sherwood, J. B., et al. (1995). Triggering of acute myocardial infarction onset by episodes of anger. *Circulation* **92**, 1720–1725.
- Muller, J. E., Tofler, G. H. and Stone, P. H. (1989). Circadian variation and triggers of onset of acute cardiovascular disease. *Circulation* **79**, 733–743.
- Ornish, D., Brown, S. E., Scherwitz, L. W., et al. (1990). Can lifestyle changes reverse coronary heart disease? *Lancet* **336**, 129–133.
- Shumaker, S. A. and Cjowski, S. M. (eds.) (1994). *Social support and cardiovascular disease*. New York: Plenum.
- Tofler, G. H., Stone, P. H., Maclure, M., et al. (1990). Analysis of possible triggers of acute myocardial infarction (the MILIS study). *American Journal of Cardiology* **66**, 22–27.
- Williams, R. B., Barefoot, J. C., Califf, R. M., et al. (1992). Prognostic importance of social and economic resources among medically treated patients with angiographically documented coronary artery disease. *Journal of the American Medical Association* **267**, 559–560.

Heart Failure, Stress Effects

M Alevizaki

Evgenidion Hospital and Alexandra Hospital, Athens University School of Medicine, Athens, Greece

G P Chrousos

Evgenidion Hospital and Athens University School of Medicine, Athens, Greece

© 2007 Elsevier Inc. All rights reserved.

Pheochromocytoma

Tumor arising from cells of neural crest origin of the medulla of the adrenal gland.

Renin-angiotensin-aldosterone system

Participates in the regulation of the extracellular fluid volume and the long-term arterial pressure, as well as in the maintenance of Na and K blood concentrations.

Glossary

<i>Apoptosis</i>	(From the Greek words <i>apo</i> = from and <i>ptosis</i> = falling); one of the main types of programmed cell death.
<i>Atrial natriuretic peptide (ANP)</i>	Released by atrial cardiomyocytes and has a direct diuretic, natriuretic and vasodilator effect.
<i>Cardiomyocytes</i>	Striated muscle cells of the heart, deriving from cardiac myoblasts; they participate in the contraction of the heart.
<i>Catecholamines</i>	Biogenic amines produced mainly from the adrenal medulla and the postganglionic neurons of the sympathetic nervous system; they participate in the stress system, influencing heart rate, blood pressure and various metabolic parameters.
<i>Cytokines</i>	Polypeptides produced in a variety of tissues; mediators of the inflammatory and immune responses.
<i>Nitric oxide synthase (NOS)</i>	Enzyme responsible for the synthesis of nitric oxide (NO).
<i>Norepinephrine (noradrenaline)</i>	A catecholamine, released both from noradrenergic neurons of the sympathetic nervous system and from the medulla of adrenal glands.

Heart failure is the end result of declining pumping capacity of the heart; it follows some initial injury to the heart, which may be of either an acute or progressive nature. In recent years, neurohormonal mechanisms that may be involved in the progression of heart failure have been proposed. These represent compensatory physiological responses activated to assist the heart and to slow down the deterioration of its pumping function. The increased secretion of catecholamines; release of various peptides with vasodilatory effects, such as the atrial natriuretic peptides; and production of the gas hormone nitric oxide all contribute in the assistance of the heart. Several cells from within the heart, however, produce other hormones, such as endothelin, with vasoconstrictive effects, which, when present at increased levels or when administered in excess, are able to partially reproduce the syndrome of heart failure in experimental animal models. These effects are generally reversible.

Stress is classically linked to elevated adrenergic activity. In the failing heart the original increase in adrenergic activity is beneficial because it facilitates an increase in the heart rate and contractility.

However, in the progress of heart failure, further increases in circulating norepinephrine occur, which may be detrimental to the myocardium. Elevated norepinephrine concentrations are positively correlated with the severity of heart failure, with the more elevated levels linked to a worse prognosis and higher mortality. Chronically elevated catecholamines may be cardiotoxic, and heart failure associated with pheochromocytoma is a well-recognized entity. Prolonged and excessive β -adrenergic receptor stimulation may result in intracellular calcium overload, which may promote cardiomyocyte apoptosis. Therefore, the effects of increased adrenergic stimulation are paradoxical and change as the disease progresses, starting as beneficial and ending as detrimental. Slowly, the failing heart may lose its ability to respond to catecholamines.

The use of antagonists of neurohormonal factors may reverse heart failure progression to some extent; thus, the use of both β -adrenergic blockade and of angiotensin-converting enzyme inhibitors partially reverse the progression of heart failure.

Stress is also associated with the production of pro-inflammatory cytokines. Recently, an important role of pro-inflammatory cytokines in the pathophysiology and progression of heart failure has been recognized. These cytokines may be produced under the influence of stress by cardiac myocytes independently of the activation of the immune system and have direct adverse effects on the cardiovascular system. Studies have specifically shown that tumor necrosis factor (TNF)- α , interleukin (IL)-2, and IL-6 exert immediate negative inotropic effects on the heart and, thus, aggravate heart failure. Increased circulating TNF- α levels are associated with more severe degrees of heart failure. It also appears that these molecules may have adverse effects on the remodeling of the cardiac cells and left ventricular structure through effects on apoptosis, as well as on the endothelium, by influencing the transcription of the NO synthase.

This knowledge has led to clinical trials studying the effects of TNF- α blockade on the progression of heart failure; these studies, however, have not shown encouraging results. Interestingly, the blockade of nuclear factor (NF)- κ B activity, which is a ubiquitous transcription factor involved in the transcription of the pro-inflammatory cytokine genes, has also been associated with an improvement in heart failure. It has thus been suggested that NF- κ B might be a promising target for future drug development for the treatment of heart failure. The current general view is that there is an interplay among catecholamines, the renin-angiotensin system, and pro-inflammatory cytokines that is involved in the pathophysiology of

heart failure. Thus, many conventional therapeutic modalities may be acting through their effects on the pro-inflammatory cytokines. One example is that angiotensin type I receptor antagonists reduce the level of circulating inflammatory mediators in patients with heart failure.

Another interesting observation linking hormones that help the body adapt to stress with heart failure was recently published. It was shown that corticotropin releasing hormone receptor 2 (CRH-R2)- and urocortin 2-deficiencies were associated with heart failure in an animal model and that the reversal of this abnormality was associated with improvement in cardiovascular function.

Finally, it was recently suggested that certain cases of cardiovascular disease may be linked to posttraumatic stress disorder (PTSD). This is not surprising because these conditions are characterized by increased sympathetic activity and elevated concentrations of circulating cytokines such as IL-6. PTSD has, indeed, been associated with heart-rate abnormalities and increased incidence of coronary artery disease; thus, although direct links of PTSD with heart failure have not been shown, the association of this disorder with coronary artery disease points in that direction.

See Also the Following Articles

Catecholamines; Depression and Coronary Heart Disease; Cardiovascular Disease, Stress and; Primate Models, Cardiovascular Disease; Stress Management and Cardiovascular Disease; Stress, NPY and Cardiovascular Diseases.

Further Reading

- Bale, T. L., Hoshijima, M., Gu, Y., et al. (2004). The cardiovascular physiologic actions of urocortin II: acute effects in murine heart failure. *Proceedings of the National Academy of Sciences USA* 101, 2697–3702.
- Bozkurt, B., Kribbs, S., Clubb, F. J., Jr., et al. (1998). Pathophysiologically relevant concentrations of TNF α promote progressive left ventricular dysfunction and remodeling in rats. *Circulation* 97, 1382–1391.
- Cotter, G., Milo-Cotter, O., Rubinstein, D., et al. (2006). Posttraumatic stress disorder: a missed link between psychiatric and cardiovascular morbidity? *CNS Spectrums* 11, 129–136.
- Deswal, A., Petersen, N. J., Feldman, A. M., et al. (2001). Cytokines and cytokine receptors in advanced heart failure: an analysis of the cytokine database from the versarinone trial (VEST). *Circulation* 103, 2055–2059.
- Gurlek, A., Kilickap, M., Dincer, I., et al. (2001). Effect of losartan on circulating TNF alpha levels and left ventricular systolic performance in patients with heart failure. *Journal of Cardiovascular Risk* 8, 279–282.

- Kapadia, S., Dibbs, Z., Kurrelmeyer, K., et al. (1998). The role of cytokines in the failing human heart. *Cardiology Clinics* 16, 645–656.
- Levine, B., Kalman, J., Mayer, L., et al. (1990). Elevated circulating levels of TNF in severe chronic heart failure. *New England Journal of Medicine* 223, 236–241.
- Mann, D. L. (2004). Activation of inflammatory mediators in heart failure. In: Mann, U. (ed.) *Heart failure: a companion to Brownwald's Heart Disease*, pp. 159–180. Philadelphia: Saunders.
- Port, J. D., Linseman, J. V. and Bristow, M. R. (2004). Adrenergic receptor signaling in chronic heart failure. In: Mann, U. (ed.) *Heart failure: a companion to Brownwald's Heart Disease*, pp. 145–158. Philadelphia: Saunders.

Heart Rate

J R Jennings

University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J R Jennings, volume 2, pp 333–337, © 2000, Elsevier Inc.

Forms of Heart Rate Measurement

Heart Rate Assessment in Behavioral Medicine

Heart Rate Assessment in Psychological Science

Summary

Glossary

<i>Baseline</i>	A condition to which responses are compared. This can be a resting condition with no stimulation or a condition that has all the elements of the task or stressor condition except for the critical task or stress element.
<i>Electrocardiogram</i>	A changing voltage over time collected from the chest of a person using surface electrodes. The electrocardiogram's changing voltages show atrial contraction (p wave) followed by ventricular contraction (qrs wave) and then ventricular recovery (t wave). Because of its relatively large amplitude, the qrs wave is typically used as the marker for the computation of heart rate or interbeat interval, although the heartbeat is initiated by the atrial contraction (p wave).
<i>Heart rate</i>	The number of contractions (beats) of the heart expressed per unit time, typically the number of heartbeats per minute.
<i>Interbeat interval</i>	The time, typically expressed in milliseconds, between contractions of the heart.

Phasic response

A transient change in the rate of the heart (or other index) induced by an environmental or physiological stimulus. Phasic responses are typically compared to a baseline and last for a second or more but less than a minute.

Stressor/task

Any event that poses a challenge to an individual. Conditions that threaten to induce harm or pain or frustrate achievement of positive rewards are termed stressors. Conditions that pose a requirement to mentally or physically manipulate objects are typically termed tasks; these may or may not also be stressors.

Tonic response

A sustained change in the rate of the heart (or other index) induced by an environmental or physiological condition. Tonic responses are typically compared to a baseline and last for a minute or more.

Forms of Heart Rate Measurement

Clinical

The pulse, obtained by palpitation over an artery, is the basic means of obtaining heart rate in the medical clinic. The contractions of the heart are counted based on the mechanical sensation in the thumb or finger of the observer induced by the expansion of the artery due to cardiac contraction. The experience of the pulse during stress probably arises from a similar mechanism: mechanical sensations from the viscera induced by cardiac contraction – contractions that increase in force during stress. Collection of heart rate via palpitation requires the presence of another person and physical contact. The person must accurately sense the heartbeats and time the interval over which the counts are made. Use of palpated heart rate in studies of stress has been reduced due to the

possibility of error in judgments of timing and beat occurrence as well as the likelihood that the presence of a person may alter any laboratory-induced stress. Heart rate derived by this technique is typically expressed as beats per minute.

Physiological

Heart rate in physiological studies is typically derived from measurements taken from the electrocardiogram. The number of qrs waves per unit time are counted or the time between these waves is measured. Time between waves (interbeat interval) can be translated to the rate of the heart for any pair of beats. For example, if the time between heartbeats is 1000 ms (1 s), then the corresponding heart rate would be 60 beats/min; that is, 60 of these intervals of 1000 ms would equal 60 000 ms, 60 s, or 1 min. Such detailed measurements permit the tracking of how the heart is reacting beat by beat (or second by second) to environmental and physiological stimuli.

Heart rate arises from the integration of an individual's intrinsic rate with the influences of the sympathetic and parasympathetic branch of the nervous system. The heartbeat is initiated in the sinus node of the heart, which is a patch of spontaneously active neural tissue located in the cardiac atrium. The sinus node discharge rate depends on a basic intrinsic or constitutional rate for that person as modulated by the speeding influences of sympathetic nervous system firing and the slowing influences of parasympathetic system (vagus nerve) firing upon the sinus node. Thus, a second-by-second recording of heart rate provides a second-by-second record of a balance between sympathetic and parasympathetic nervous system influences on the heart.

Psychological

The sensitivity of heart rate to nervous system activation provides the justification for its use as an index of psychological stress. Stress is typically assumed to activate the sympathetic nervous system and deactivate the parasympathetic nervous system. Direct influences of the sympathetic nervous system on the sinus node as well as indirect influences via circulating catecholamines (derived primarily from the endocrine gland, the adrenal medulla) will activate β -adrenergic receptors, resulting in a speeding of the heart rate and an increase in cardiac contractility. Both nervous and endocrine reactions are transient responses to stressors. Some individuals report being in stressful circumstances lasting a year or more; such circumstances are unlikely to be directly reflected in heightened heart rate or even adrenomedullary reactions because these are reactions that return to baseline values after a brief

period. Specific stressors arising from stressful circumstances can induce tonic reactions or phasic reactions of heart rate. Reactions to stressor lasting minutes to hours can be detected by examining tonic reactions in which the average heart rate over periods of minutes can be demonstrated to be elevated relative to a baseline period. Acute reactions to specific stressors may induce transient nervous system responses lasting seconds rather than minutes. Such phasic responses are typically assessed by comparing heart rate (or interbeat interval) for the seconds after the stressor to that of the seconds just before the stressor. The rapidity with which heart rate returns to baseline can also be informative. Increased vagal activation or decreased sympathetic activation will return heart rate to baseline values. Psychological factors, such as rumination over the stressful incident, may prolong heart rate speeding in some individuals. Concurrently, physiological factors may play a role in the recovery of heart rate, e.g., some individuals may have rapid vagal activation after a stressor with consequently brief heart rate reactions to stress. In short, the measurement of heart rate as an index of stress must be gauged to the type of stress, most particularly its temporal duration.

Heart Rate Assessment in Behavioral Medicine

Cardiovascular Reactivity

Because of its influence on the cardiovascular system, stress may influence the course of cardiovascular disease. A general hypothesis, the reactivity hypothesis, states that the cardiovascular reactivity resulting from repeated stressful events might contribute to the development of hypertension, myocardial infarction, and stroke. Presumably certain individuals are predisposed to show consistent and large cardiovascular reactions to stress, or are placed in an environment in which stressors are frequent and repetitive, or both. This hypothesis has been widely investigated using both laboratory challenges to assess reactivity and the ambulatory monitoring of heart rate and blood pressure as individuals perform their usual activities. The consistency within an individual over time of heart rate responses to laboratory stress has been established. Heart rate also changes in response to daily events in ambulatory recordings.

The number and size of heart rate responses do not seem, however, to predict cardiovascular disease onset and course; blood pressure increases in response to stressors have proven more predictive. The extensive parasympathetic influence on heart rate may explain this difference in the predictive utility of heart rate and blood pressure.

Heart Rate Variability

Derived measures of heart rate related to nervous system function have been found to correlate positively with cardiovascular health. Parasympathetic influences on heart rate are exerted via the vagus nerve, which innervates the heart (along with innervation by the cardiac sympathetic nerves). The vagus is more important than the sympathetic nerves for most transient changes in heart rate. Vagal firing influencing the heart waxes and wanes with respiration; this results in heart rate increasing during inspiration and decreasing during expiration. The rhythmic changes in heart rate over time, heart rate variability, can then be used to measure vagal influence on the heart. Heart rate has been found to show rhythmic changes with lower frequencies than respiration as well. For example, a frequency of 0.1 Hz has been related in part to sympathetic activity related to blood pressure, and heart rate shows a 24-h rhythm related to the sleep–wake cycle. Variability at the respiratory rate has been shown to be due almost completely to vagal/parasympathetic activity. Variability at other frequencies is not as well defined. Relatively greater variability has in general, however, been associated with cardiovascular health. A related measure positively associated with cardiovascular health is baroreceptor sensitivity. The baroreceptors are sense organs in major arteries that sense pressure change and transmit this to the brain, which then alters sympathetic and parasympathetic control to return blood pressure to normal. Pressure increases induce a vagal activation reducing heart rate. Greater heart rate change per unit of heart rate change (baroreceptor sensitivity) is correlated with better cardiovascular health. The associations of cardiac variability and baroreceptor sensitivity with cardiovascular health have led to the hypothesis that cardiac variability, most specifically parasympathetically induced variability, may act to buffer the influence of stress on the cardiovascular system.

Heart Rate Assessment in Psychological Science

Motivation/Affect

A psychological analysis of stress typically suggests that the threat or stressor induces an affective reaction as well as influences a person's motivational state. Subsequently, coping with the stressor may require other psychological processes. Behavioral medicine investigations have not typically attempted to separate the different processes influenced by stress, but the psychological analysis raises the question of whether heart rate reactions are controlled by one or

more of the different psychological processes. The inference of a psychological process that is not directly observable is difficult, and this difficulty is compounded when the processes are as complex and conceptually overlapping as categories of affect and motivation. Studies do show that heart rate increases with the induction of most emotions, with the exception of disgust. Similarly, cues for reward or punishment can induce heart rate increases, perhaps with stronger increases related to cues for reward. Interpretation is further complicated by the possibility that coping with the affect or acting upon the motivational cue may elicit other processes controlling heart rate. For example, the induction of affect with an imaged scenario that contains an action element appears to induce greater heart rate change than induction with a more passive scenario.

Attention/Cognition

Coping refers to a set of action and mental processes that permit the person to redefine the threat and/or develop a means of neutralizing the threat. Thus, a multitude of cognitions and actions may be elicited as coping. If heart rate responds to processes such as attention and memory processes, then the involvement of these processes in coping, or even in the initial evaluation of the stressor, may be the source of phasic heart rate changes after a stressor. Muscle activation is a strong metabolic stimulus that induces rapid increases in heart rate. Similarly, change in breathing patterns, e.g., breath holding, have rapid, large effects on heart rate. These changes can readily be demonstrated in the absence of a stressor, while it is difficult to ensure that reactions to a stressor do not include muscle tension changes or changes in breathing pattern. Of greater psychological interest is attentive waiting, which induces clear phasic changes in heart rate. Heart rate decreases during the anticipation of any significant task or task element. Conceptually, heart rate seems to decrease whenever an action must be chosen at a specific point in time and potentially competing actions must be inhibited. The heart slows reliably prior to a signal to respond, even when that signal is accompanied by an electrical shock. This slowing of the heart rate does not appear to be due to respiratory changes or the inhibition of the peripheral musculature. The relevance of this observation to behavioral medicine studies of reactivity is clear. Many studies have used as their stressor/challenge computer games with periods of anticipatory attention.

Interpretive Issues

Heart rate can serve as an index of stress, but its utility and validity can be increased by a few simple

precautions. First, stress and stress reactions should be conceptualized. Should heart rate in response to an increase in muscle tension be considered a stress reaction or a muscle artifact? Second, the stressor and overall conditions of eliciting the stress reaction should be controlled or assessed so it is possible to separate heart rate changes due to motor activity, respiration, transient attention, and affect/motivation. If a tonic response is of greatest interest, than scoring of the heart rate should occur after transient changes related to attention, breath holding, and muscle tensing. Fortunately, many of these phasic changes disappear after the initial few seconds of a task or stimulus presentation. Care must be taken in interpretation to recognize all the possible sources – other than those defined as stress – of the observed change in heart rate.

Summary

Heart rate increases rapidly during stress. The response may be initiated by motivational threat, affect, or the thoughts and actions induced by the stressor. Separation of the cause of the heart rate change is difficult in both the laboratory and real life. The degree of heart rate increase in response to stress or psychological challenge has been less predictive of medically significant cardiovascular disease than degree of change of blood pressure. The degree of variability of heart rate over time has, however, been associated with relatively better cardiovascular health.

Acknowledgment

The author acknowledges the assistance in writing this article from the support of NIH grant HL57529.

See Also the Following Articles

Blood Pressure; Cardiovascular Disease, Stress and; Sympathetic Nervous System.

Further Reading

- Bernston, G. G. and Cacioppo, J. T. (2000). Homeostatic to allodynamic regulation. In: Cacioppo, J. T., Tassinary, L. G. & Bernston, G. G. (eds.) *Handbook of psychophysiology* (2nd edn., pp. 459–490). Cambridge, UK: Cambridge University Press.
- Dekker, J. M., Crow, R. S., Folsom, A. R., et al. (2000). Low heart rate variability in a 2-minute rhythm strip predicts risk of coronary heart disease and mortality from several causes: the ARIC Study. Atherosclerosis risk in communities. *Circulation* 102, 1239–1244.
- Huikuri, H. V., Jokinen, V., Syvanne, M., et al. (1999). Heart rate variability and progression of coronary atherosclerosis. *Arteriosclerosis, Thrombosis, and Vascular Biology* 19(8), 1979–1985.
- Jennings, J. R., Berg, W. K., Hutcheson, J. S., et al. (1981). Publication guidelines for the heart rate studies in man. *Psychophysiology* 19, 226–231.
- Jennings, J. R., van der Molen, M. W., Somsen, R. J., et al. (2002). Vagal function in health and disease: studies in Pittsburgh. *Physiology and Behavior* 77(4–5), 693–698.
- Jennings, J. R., Kamarck, T. W., Everson-Rose, S. A., et al. (2004). Exaggerated blood pressure responses during mental stress are prospectively related to enhanced carotid atherosclerosis in middle-aged Finnish men. *Circulation* 110(15), 2198–2203.
- Kamarck, T. W., Jennings, J. R., Debski, T. T., et al. (1992). Reliable measures of behaviorally-evoked cardiovascular reactivity from a PC-based test battery: results from student and community samples. *Psychophysiology* 29, 17–28.
- Krantz, D. S. and Manuck, S. B. (1984). Acute psychophysiological reactivity and risk of cardiovascular disease: a review and methodologic critique. *Psychological Bulletin* 96, 435–464.
- Lovallo, W. R. (2005). *Stress and health: biological and psychological interactions* (2nd edn.). Thousand Oaks, CA: Sage.
- Noble, D. (1979). *The initiation of the heart beat*. Oxford, UK: Clarendon Press.
- Task Force (1996). Heart rate variability: standards of measurement, physiological interpretation, and clinical use. *Circulation* 93, 1043–1065.

Heat Resistance

A J L Macario and E Conway de Macario

New York State Department of Health and
The University at Albany (SUNY),
Albany, NY, USA

Published by Elsevier Inc.

Diversity of Life Forms and Habitats

Definition

Chaperones and Chaperonins

Thermoprotectants

Intrinsic Molecular Features

Cell Envelopes and Multicellular Structures

Spores

Applications

Glossary

Biofilm

A flat multicellular structure, one or very few cells thick, that consists of a single or several (consortium) species. The single-species variety appears during the normal life cycle of an organism or in response to stress. Most biofilms must adhere to a surface as a support for growth and are, like granules, more heat-resistant than are their individual components living outside the structure. Other biofilms, such as the lamina formed by the archaeon *Methanoarcina mazei* S-6, do not need to adhere to a surface for growth.

Granule

A multicellular structure or consortium composed of two or more species of organisms; it is more resistant to stressors such as heat than are the free organisms outside it.

Heat resistance

The capacity to withstand temperatures higher than 37°C without a significant loss of function; also called thermotolerance. More specifically, a cell condition characterized by an increase in the capacity to withstand temperatures above the optimal for that cell's growth, whether 37°C, higher, or lower. The cell condition involves an active process with a differential regulation of gene expression and protein levels induced by the elevation of temperature above the optimal one (heat shock); also called acquired or induced heat resistance.

Hyper-thermophile Packet

An organism with an optimal temperature for growth of 80°C or higher.

A globular multicellular structure, composed of only one species, formed by certain microbes (e.g., the archaeon *Methanosarcina mazei* S-6) as part of their life cycles or in response to environmental stressors. The packet is more resistant to stressors, including heat shock, than are the individual cells living outside it.

Thermophile

An organism with an optimal temperature for growth of 50–70°C; it prefers temperatures higher than 37°C and has constitutive heat resistance, namely an inborn capacity to carry out all of its physiological functions at temperatures that are not optimal for (or will damage or kill) human cells.

Thermo-protectant

A nonprotein molecule that plays an active role in thermotolerance by protecting proteins, and perhaps other cellular components, from the damaging effects of elevated temperatures, whether during normal growth (thermophiles and hyperthermophiles) or on a sudden increase above the optimal temperature (heat shock). There are several thermo-protectants that have rather simple structures. They are different from molecular chaperones and chaperonins, which are proteins.

Diversity of Life Forms and Habitats

In nature, living cells and organisms are diverse, encompassing a wide range of species that thrive under very varied conditions. Temperature is one of the major conditions that significantly influence the life of cells. As a reference point, bear in mind that the normal body temperature of adult humans is 36.8 (±0.4)°C and that human cells grow well in laboratory cultures at 37°C. There are, however, very many organisms, both single- and multicelled, on our planet that have optimal temperatures for growth (OTGs) that differ from 37°C, either lower or higher and sometimes radically.

Organisms are classified, according to their OTGs, into the following groups: psychrophiles (OTG 15°C or lower), mesophiles (OTG 35–40°C), thermophiles (OTG 50–70°C), and hyperthermophiles (OTG 80°C or higher). Within the mesophiles, those that can still grow at 0–4°C are termed psychrotolerant; within the

hyperthermophiles, those that have OTGs above 95°C are termed extreme thermophiles.

Although organisms grow best at their OTGs, they can also grow and multiply, albeit more slowly, within temperature ranges on either side of their optimum. For example, the mesophilic organism *Methanosarcina mazei* S-6 displays its fastest growth rate at 37–40°C, but it is capable of growing and multiplying within the 25–45°C range. Likewise, the extreme thermophile *Pyrolobus fumarii* has an OTG of 106°C, but it can grow at temperatures between 90 and 113°C. Determinations of this sort have shown that the maximum temperature at which any given organism can grow is closer to its OTG than to the minimum that still allows its survival. This suggests that, as a general rule, organisms are more sensitive to overheating than to cooling. Excess heat may kill (lethal temperature), whereas a comparable excess of cold may only slow down the functions of the cell.

Cells exposed to an elevation of temperature above their OTGs are stressed, or heat shocked, and they mount a stress or heat shock response. This response is characterized by the activation of stress (heat shock) genes along with the slowing down or shutting off of many other genes. The stress genes produce stress (heat shock) proteins, some of which are molecular chaperones; the latter play a role in cell survival during stress and recovery after stress.

Definition

Heat resistance, also called thermotolerance, is a condition that makes the cell capable of withstanding a temperature excess. Because of this capability, the cell can survive a heat shock that would otherwise be lethal.

If we use as a reference point the OTG for human cells, 37°C, we can recognize two main types of thermotolerance: constitutive and acquired. The former pertains to organisms that normally grow at temperatures considered stressful or lethal for human cells, as exemplified by the thermophiles and hyperthermophiles.

Acquired thermotolerance, in contrast, is an experiential tolerance that an organism gains after a heat shock. The following experiment illustrates a typical case of acquired thermotolerance. Cells from a hyperthermophilic organism with an OTG of 80°C, but that can grow within the 60–86°C range, was heat-shocked when exposed to 88°C (sublethal temperature for this organism) for 1 h, but it was killed by exposure to 92°C (lethal temperature). Remarkably, cells that had been heat shocked at 88°C did not die when they were subsequently exposed to the lethal temperature. The heat-shocked cells had acquired

thermotolerance; the cells were able to survive a lethal temperature because they had been prepared (preconditioned) by preexposure to a sublethal, heat-shocking temperature.

Cross-Tolerance

Temperature excess is a cell stressor, but it is not the only one known. There are many cell stressors of mechanical (e.g., compression and shearing), physical (e.g., heat and ultraviolet radiation), or chemical (e.g., alcohol, hydrogen peroxide, high salinity, and lack of oxygen) nature that elicit a stress (heat shock) response in cells. Interestingly, stress tolerance induced by one stressor may also result in efficient resistance (i.e., cross-tolerance) to other stressors. Thus, thermotolerance induced by heat shock may also increase the resistance to a chemical stressor and vice versa. Cross-tolerance to stressors is of extraordinary potential value to fields such as medicine and surgery.

Mechanism

Heat resistance or thermotolerance is mediated by a variety of factors and mechanisms: stress proteins, molecular chaperones, chaperonins, thermoprotective compounds, innate characteristics of proteins that make them thermostable, cell-wall composition and architecture, formation of multicellular structures, and formation of spores.

One of the key harmful events in heat stress is protein denaturation. If the stress is too drastic (a heat-shocking temperature far above the OTG or very prolonged heat shock), the cellular proteins denature and aggregate. Cell death ensues. An illustrative example is the hard-boiled egg. There is no known way to restore a hard-boiled egg to a viable form so that the denatured, aggregated proteins disaggregate and regain their functional, native conformation. There are, however, many instances in which protein denaturation does not reach such extreme, irreversible levels. Milder stresses cause only partial denaturation, and the damage is reversible; the proteins can be restored to their functional form. The recovery is mediated by several mechanisms, one of which involves the molecular chaperones. These chaperones are also involved in preventing protein damage and in eliminating protein molecules damaged beyond repair.

Thermotolerance is based on various mechanisms and a key one is the presence and use of all the cellular molecules and mechanisms that prevent and cure protein denaturation and also those that eliminate irreversibly damaged and aggregated proteins before these reach levels that will clog the intracellular machinery.

Chaperones and Chaperonins

Chaperones and chaperonins belong to the large group of stress (heat shock) proteins, although not all of them are the product of stress-inducible genes. Their functions are to assist nascent polypeptides in folding correctly so that each can reach a functional configuration (its native conformation), to protect the polypeptides from aggregation during synthesis and during stress, and to refold them after partial unfolding (reversible denaturation) caused by stressors such as heat.

Thermoprotectants

Thermoprotectants are simple cellular components that confer resistance to other molecules, by mechanisms not yet fully elucidated. They are considered integral to thermotolerance because they have been found in hyperthermophilic organisms and because laboratory tests *in vitro* have indicated that they stabilize protein molecules, rendering them heat resistant.

Two thermoprotectants have been studied in some detail: di-*myo*-inositol phosphate (DIP), and cyclic diphosphoglycerate (cDPG). These molecules are believed to be important for the physiology of organisms living in high-temperature ecosystems (e.g., hot springs, hot pools and mud pots, solfataras and geysers, and hydrothermal vents at the bottom of the ocean) and for their survival when stressed by the variations in temperature, pH, and hydrostatic pressure that are characteristic of these extreme environments. Trehalose, a disaccharide, has also been implicated in the mechanism of heat resistance, either alone or in association with chaperones.

Intrinsic Molecular Features

Many proteins, including enzymes, of thermophilic and hyperthermophilic organisms withstand heat stress better than do their counterparts from mesophiles, even in the absence of thermoprotectants. They are believed to have built-in structural features that make them thermostable, but the nature of these features has not been determined. It has been proposed that innumerable subtle interactions occur among various points of the molecule that confer great stability, even in the face of stressors, and enhanced resistance to the denaturation induced by the latter. The identification and measurement of these interactions have proven elusive up to the present time.

Cell Envelopes and Multicellular Structures

Many organisms, particularly those living in extreme environments, are protected from harsh environmental factors by specialized cell envelopes that are resistant to the stressors (e.g., high temperature or salinity) that are inherent to their ecosystems. In addition, many species form multicellular structures that are more resistant to stressors than are single cells. In these structures, cells are held together either by an intercellular connective material, consisting of a flat sheet with a thickness of one or very few cells or in a globular conglomerate resembling a sphere (Figure 1). The former, flat structure is called a biofilm, and it is usually, but not always, attached to a supporting surface (e.g., rock in natural habitats or sand particles in methanogenic bioreactors).

The globular structures are of two types; they are formed either by a single species or by an association of two or more species. The multispecies type is called a granule, and it constitutes a microbial consortium that can live in both natural and manufactured ecosystems. Examples of manufactured ecosystems are the anaerobic methanogenic bioreactors used in waste-processing plants. In the granular consortium, various species of microorganisms cooperate with one another in a food web that bioconverts organic molecules while being shielded from the shearing and other stressors typical of bioreactors, such as temperature and pH variations. Indeed, the consortium is very resistant to these stressors, and it survives stresses that would kill its microbial components outside the granule.

The single-species type of globular multicellular structure is formed by a single species of organism as part of its life cycle or in response to stressors. An example of a single-species multicellular structure is that built by the methanogenic microorganism *Methanosarcina mazei* S-6 in culture and in bioreactors (Figure 1). The globular morphotype, termed a packet, is considerably more resistant to mechanical, chemical, and thermal stressors than is the single-celled form of this archaeon.

Spores

Some microbes subjected to harsh environmental conditions, such as scarcity of nutrients, desiccation, and exposure to toxic substances, undergo sporulation. They form a very stress-resistant structure called a spore or endospore. The microbe becomes dehydrated, and its genetic material is enclosed in a thick coat with a special composition that is, most likely,

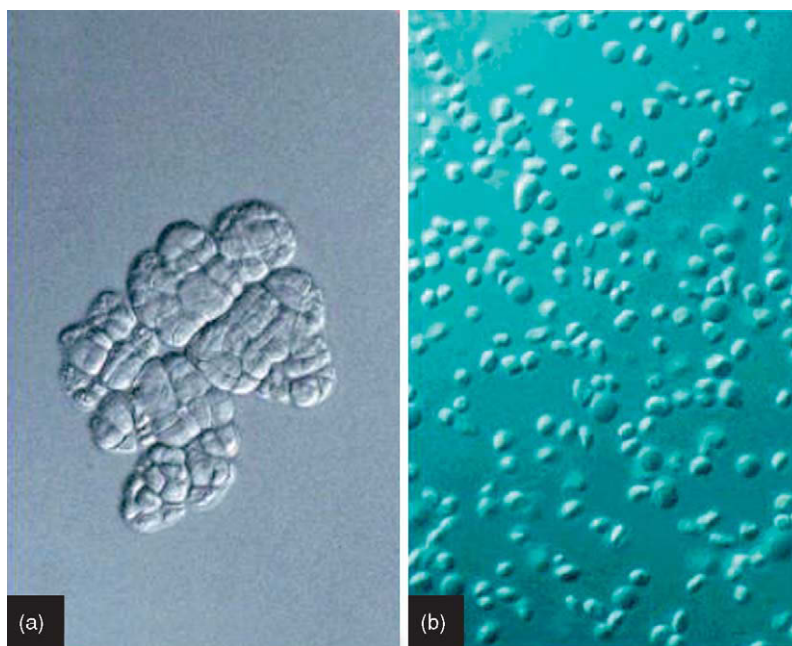


Figure 1 Archaeal microbe *Methanosarcina mazei* S-6. (a) Cells growing in a packet, a globular, heat-resistant multicellular structure; (b) cells growing individually. The packet is composed of many cells joined together by an intercellular connective material that also enwraps the whole structure. From A. J. L. Macario and E. Conway de Macario (2001), *Frontiers in Bioscience* **6**, d262–d283, [online].

responsible for the extreme resistance of spores to stressors, including heat.

Because spores are produced by microbes that can cause disease in humans and animals (e.g., *Clostridium tetani* and *C. botulinum*), they are of special interest to the field of medicine and to the food industry, which must resort to sterilization and pasteurization (both involving heating) to prevent contamination and infection.

Applications

Stress resistance is a desirable attribute from the organism's perspective because confrontation with stressors, such as heat shock, is a frequent occurrence in the life of any cell or organism. However, as pointed out in the previous section, microbial stress resistance may be undesirable from the human perspective, when the organisms is pathogenic and must be destroyed.

In other circumstances, when stress resistance is a convenient attribute from the medical or industrial perspective, attempts at making certain cell types more resistant to stressors are not only justified but also are an attractive scientific and biotechnologic endeavor. Means are being developed at the present time to enhance cell survival during stress and to accelerate cell recovery after stress. Briefly, the idea is to provide a cell with those elements now known to

play (or that are suspected of playing) a key role in maintaining its normal array of proteins in working condition. These elements can help in restoring the functional conformation of damaged proteins or in degrading those damaged beyond repair during and after stress. Genetic engineering and preconditioning are two approaches to achieving this; they can be applied independently of one another or in combination. Essentially, in the genetic engineering approach, an extra set of genes encoding the proteins of interest (e.g., molecular chaperones or proteases) are grafted into the genome of the target cell whose stress resistance is to be increased. Alternatively, the extra set of genes is incorporated into the cytoplasm of the cell via a self-replicating extrachromosomal element such as a plasmid. In both instances, the expression of the foreign genes increases the levels of the stress proteins inside the cell, which in turn enhances the folding of newly formed, much needed polypeptides and also the refolding of partially denatured molecules. Of course, it is a possibility that secondary effects may also occur that are negative from the standpoint of the cell, tissue, or organism we are trying to protect from the consequences of stress. Because of this possibility, many tests with experimental models are necessary to certify the benefits of this type of anti-stress treatment before it is applied to valuable cells in bioreactors or to humans in a therapeutic context.

The preconditioning approach involves the use of mild stressors, for example, chemical compounds known to induce stress genes, to generate cross-tolerance to potent stressors or the use of a potent stressor at low doses to induce a stress response of low intensity. Both procedures confer a sort of cellular memory. When the cell is challenged with a potentially very damaging or lethal stressor against which a cross-tolerance was previously induced, the cell has increased chances of survival. Likewise, a cell preconditioned with low doses of a potent stressor also has increased chances of survival when challenged with the same stressor at doses that would be lethal for a nonpreconditioned cell.

These strategies and procedures are being applied and have a promising future in the biotechnology industry and in medicine and surgery. In the biotechnology field, bioreactors (also called fermentors) are used for a variety of purposes ranging from the large-scale production of useful molecules (e.g., hormones and antibiotics) to the bioconversion of wastes. However, a drawback is that the cells in bioreactors are susceptible to stressors generated by the process itself or received with the influent (e.g., wastes). Oscillations or abrupt changes in temperature, pH, and content of toxic chemicals are common events in bioconversion technology. These stressors cause frequent bioreactor malfunctions or failures because the bioprocessing cells suffer stress, malfunction, and even die. Obviously, the generation of resistant cells, capable of withstanding the stressors, will improve the efficiency and economic value of this technology.

Another obvious need for cells equipped with strong antistress mechanisms exists in surgery. Poor oxygenation, resulting in cell stress, occurs in many tissues; most important, it occurs in the brain and heart during large-scale surgical operations, particularly those necessitating blood exchange. It is, therefore, crucial to prevent brain- and heart-cell damage and to ensure the preservation of these vital tissues. To achieve this goal, a promising strategy is to induce stress tolerance. Stress proteins, specifically molecular chaperones, are being tested in experiments whose aim is to invigorate the cells. A stress protein or its gene is introduced into the tissue to confer resistance in a manner similar to that described for microbes in bioreactors that process wastes.

An alternative approach takes advantage of cross-tolerance. Heat resistance is induced in the cells of tissues targeted for protection from ischemia hypoxia (e.g., brain and heart tissues) by mild heat shock. This preparatory mild stress, or preconditioning, is relatively harmless, yet it is sufficient to elicit a heat shock response and heat resistance, which are efficient in

counteracting the harmful effects of ischemia hypoxia during a surgical procedure.

Last, replacement therapy, genetic or otherwise, may be indicated in chaperonopathies – diseases in which deficient chaperones are the primary etiology of or significant contributors to the pathogenesis. In these instances, chaperone genes may be used in genetic engineering approaches, as described earlier, or purified molecular chaperones may be administered directly, as antibodies are sometimes provided to treat infections.

On the negative side, microbial heat resistance is an obstacle to the sterilization of surgical and laboratory instruments, plasticware and glassware, sera, solutions, culture media, and any other object or substance that must be freed from pathogenic organisms while avoiding a significant loss of biological or medicinal properties. The heat resistance of microbes also presents a problem for the food industry whenever it is necessary to improve the safety and preservation of edible products. Some very serious food-borne pathogens, such as *Clostridium botulinum*, form spores and can be in a dormant state for years until favorable environmental conditions induce their return to a vegetative, pathogenic status. The food industry must apply heat and other means to get rid of the pathogens, but only to the extent that the nutritional and sensory qualities of the product are not impaired significantly. The achievement of such a balance between killing or disarming the pathogens and preserving food quality is a major challenge for science and technology.

See Also the Following Articles

Chaperone Proteins and Chaperonopathies; Heat Resistance; Heat Shock Genes, Human; Heat Shock Response, Overview; Viral Virulence and Stress; Chaperonopathies; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles.

Further Reading

- Attfield, P. V. (1997). Stress tolerance: the key to effective strains of industrial baker's yeast. *Nature Biotechnology* 15, 1351–1357.
- Beaucamp, N., Harding, T. C., Geddes, B. J., et al. (1998). Overexpression of *hsp70i* facilitates reactivation of intracellular proteins in neurons and protects them from denaturing stress. *FEBS Letters* 441, 215–219.
- Bloch, E., Rachel, R., Burggraf, S., et al. (1997). *Pyrolobus fumarii*, gen. and sp. nov., represents a novel group of archaea, extending the upper temperature limit for life to 113 °C. *Extremophiles* 1, 14–21.
- Conway de Macario, E. and Macario, A. J. L. (2003). Molecular biology of stress genes in methanogens:

- potential for bioreactor technology. *Advances in Biochemical Engineering/Biotechnology* 81, 95–150.
- da Silva, R., Lucchinetti, E., Pasch, T., et al. (2004). Ischemic but not pharmacological preconditioning elicits a gene expression profile similar to unprotected myocardium. *Physiological Genomics* 20, 117–130.
- Dinh, H-K. B., Zhao, B., Schuschereba, S. T., et al. (2001). Gene expression profiling of the response to thermal injury in human cells. *Physiological Genomics* 7, 3–13.
- Gross, C. and Watson, K. (1997). Transcriptional and translational regulation of major heat shock proteins and patterns of trehalose mobilization during hyperthermic recovery in repressed and derepressed *Saccharomyces cerevisiae*. *Canadian Journal of Microbiology* 44, 341–350.
- Holden, J. F., Adams, M. W. W. and Baross, J. A. (2000). Heat-shock response in hyperthermophilic microorganisms. In: Bell, C. R., Brylinsky, M. & Johnson Green, P. (eds.) *Microbial biosystems: new frontiers* (vol. 2), pp. 663–670. Halifax, Canada: Atlantic Canada Society for Microbial Ecology and Acadia University Press.
- Juneja, V. K., Klein, P. G. and Marmer, B. S. (1998). Heat shock and thermotolerance of *Escherichia coli* O157:H7 I in a model beef gravy system and ground beef. *Journal of Applied Microbiology* 84, 677–684.
- Kregel, K. C. (2002). Heat shock proteins: modifying factors in physiological stress responses and acquired thermotolerance. *Journal of Applied Physiology* 92, 2177–2186.
- Macario, A. J. L. and Conway de Macario, E. (2001). The molecular chaperone system and other anti-stress mechanisms in Archaea. *Frontiers in Bioscience* 6, d262–d283. To see tables and figures go to: <http://www.bioscience.org/2001/v6/d/macario/fulltext.htm>
- Madigan, M. T., Martinko, J. M. and Parker, J. (2003). *Brock biology of microorganisms* (10th edn.). Upper Saddle River, NJ: Prentice Hall.
- Mestril, R. (2005). The use of transgenic mice to study cytoprotection by the stress proteins. *Methods* 35, 165–169.
- Mohsenzadeh, S., Saupe-Thies, W., Steier, G., et al. (1998). Temperature adaptation of house keeping and heat shock gene expression in *Neurospora crassa*. *Fungal Genetics and Biology* 25, 31–43.
- Nollen, E. A. A., Brunsting, J. F., Roelofsen, H., et al. (1999). In vivo chaperone activity of heat shock protein 70 and thermotolerance. *Molecular and Cellular Biology* 19, 2069–2079.
- Pedone, E., Bartolucci, S. and Fiorentino, G. (2004). Sensing and adapting to environmental stress: the archaeal tactic. *Frontiers in Bioscience* 9, 2909–2926. To see tables and figures go to: <http://www.bioscience.org/2004/v9/af/1447/fulltext.htm>
- Pespeni, M., Hodnett, M. and Pittet, J-F. (2005). In vivo stress preconditioning. *Methods* 35, 158–164.
- Piper, P. (1998). Differential role of Hsps and trehalose in stress tolerance. *Trends in Microbiology* 6, 43–44.
- Roberts, M. F. (2004). Osmoadaptation and osmoregulation in archaea: update 2004. *Frontiers in Bioscience* 9, 1999–2019. To see tables and figures go to: <http://www.bioscience.org/2004/v9/af/1366/fulltext.htm>
- Sabehat, A., Weiss, D. and Lurie, S. (1998). The heat-shock proteins and cross-tolerance in plants. *Physiologia Plantarum* 103, 437–441.
- Scandurra, R., Consalvi, V., Chiaraluce, R., et al. (2000). Protein stability in extremophilic archaea. *Frontiers in Bioscience* 5, d787–d795. To see the complete article go to: <http://www.bioscience.org/2000/v5/d/scan/scan.pdf>
- Sharp, F. R., Massa, S. M. and Swanson, R. A. (1999). Heat-shock protein protection. *Trends in Neuroscience* 22, 97–99.
- Takai, K., Nunoura, T., Sako, Y., et al. (1998). Acquired thermotolerance and temperature induced protein accumulation in the extremely thermophilic bacterium. *Rhodothermus obamensis*. *Journal of Bacteriology* 180, 2770–2774.
- Tolson, J. K. and Roberts, S. M. (2005). Manipulating heat shock protein expression in laboratory animals. *Methods* 35, 149–157.
- Weibezahn, J., Tessarz, P., Schlieker, C., et al. (2004). Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. *Cell* 119, 653–665.
- Westerheide, S. D., Bosman, J. D., Mbadugha, B. N. A., et al. (2004). Celastrols as inducers of the heat shock response and cytoprotection. *Journal of Biological Chemistry* 279, 56053–56060.

Heat Shock Genes, Human

A J L Macario and E Conway de Macario

New York State Department of Health and
the University at Albany (SUNY),
Albany, NY, USA

Published by Elsevier Inc.

Introduction

Molecular Phylogeny

Origins of the Human Cell

Microanatomic Components of the Human Cell

Stress Proteins in the Various Compartments of the Human Cell

Stress Genes

The Human Genome

Perspectives

Glossary

<i>Endoplasmic reticulum (ER)</i>	A eukaryotic cell component consisting of a labyrinthine system of flat cavities delimited by a membrane that is continuous with the outer membrane bounding the nucleus. The ER is involved in the synthesis and transport of proteins and other molecules.
<i>Genome</i>	The complete set of genetic information carried by a cell (or organism), representing its entire hereditary complement.
<i>Genomics</i>	The cataloguing and large-scale investigation of the sequence, structure, and function of whole genomes, or of whole sets of genes and their interactions and expression patterns.
<i>Mitochondrion</i>	A eukaryotic cell organelle, evolutionarily related to gram-negative bacteria; it houses the cellular respiration apparatus and produces energy.
<i>Molecular phylogeny</i>	The science of classifying organisms through comparisons of molecular sequences. This type of classification reflects the evolution of individual molecules and shows evolutionary relationships between molecules and organisms.
<i>Nucleolus</i>	A relatively small region within the nucleus that contains the ribosomal RNA (rRNA) genes; this is where rRNA synthesis occurs.
<i>Nucleus</i>	The distinctive organelle of the eukaryotic cell (it is not present in prokaryotes), consisting mainly of packed genetic material (DNA) encased in a double membrane. Gene expression begins in

Proteomics

this organelle with transcription, which generates the messenger RNA (mRNA). The cataloguing and study of the full set of proteins encoded by a genome; also, the analysis of the set of proteins produced by a genome or a large group of genes (expression profile) under various circumstances, e.g., conditions of cell stress vs. basal conditions (no stress).

Introduction

Heat shock genes are also called stress genes. The latter is a more appropriate name, since these genes are induced not only by heat shock but also by other stressors. Hence, the term stress genes is given preference in this article.

A better understanding of human stress genes can be gained by examining them from the evolutionary standpoint rather than by studying them in isolation, without connection with the same genes in other species or with their evolutionary history. Modern molecular phylogeny has provided the means to look at genes and their proteins while taking into consideration their structural and evolutionary relations, at a level of resolution that was not attainable in the past. Furthermore, genome sequencing has now become a reality, expanding the possibilities for gene analysis and comparison beyond previously imaginable limits.

Molecular Phylogeny

Molecular phylogeny is a relatively new scientific discipline that involves the comparative analysis of the nucleotide sequences of genes and the amino acid sequences and structural features of proteins from which evolutionary histories and relationships, and in some cases also functions, can be inferred.

The basic classification of organisms furnished by molecular phylogeny is based upon the comparison of 16S/18S (small subunit) ribosomal RNA (rRNA) sequences. Thus, all living cells are divided into three main evolutionary branches or phylogenetic lines of descent termed domains: bacteria (eubacteria), archaea (archaeobacteria), and eucarya (eukaryotes).

In addition to the small ribosomal subunit rRNA, other molecules such as stress (heat shock) proteins have been used for phylogenetic analyses. One of the most convenient targets for these analyses is the stress protein Hsp70(DnaK), for the following reasons. (1) It is long enough (about 600 amino acids on average,

of which nearly 500 are conserved) to permit the occurrence of mutations at an observable level; (2) it is highly conserved, being present in virtually all organisms and cells with few exceptions, with an amino acid sequence that is very similar in all organisms; (3) it has segments that are more highly conserved than others, which allows for reliable alignments for comparative analyses of matching structural domains; and (4) it has distinctive structural features called motifs or signatures, some of which have a known function, that facilitate comparisons among molecules from different organisms, including functional comparisons without actually performing functional assays in the laboratory.

Origins of the Human Cell

Comparative analyses of sequences and sequence signatures have provided information on the evolution of organisms and consequently of their relationships (what or who is derived from what or whom). These analyses have also clarified the origins of the eukaryotic cell (including the human cell) as a whole, and those of its components.

With the preceding in mind, we can now add that a better understanding of human stress genes also requires familiarity with the components of the eukaryotic cell and their origins, as revealed by molecular phylogenetic analyses and studies of full genome sequences.

Microanatomic Components of the Human Cell

Human cells are composed of the cytoplasmic membrane that encloses the cytoplasm or cytosol in which the organelles reside. Examples of the latter are the nucleus with the nucleolus, mitochondria, Golgi apparatus, and endoplasmic reticulum (ER). Plant cells also have these various organelles, and in addition they have chloroplasts, in which photosynthesis takes place. Stress proteins are present in all the cellular compartments, whereas the genes that code for them are in the DNA of the nucleus.

Molecular phylogenetic analyses of nucleotide and amino acid sequences of various types of molecules, including stress genes and proteins, have revealed that the genes in the nuclear DNA are a mixed assemblage of bacterial and archaeal components side by side with typically eukaryotic constituents. Furthermore, it has been shown that mitochondria are derived from ancestors of modern gram-negative bacteria and that chloroplasts are descendants of the ancestors of today's cyanobacteria.

It has been concluded that the human cell is an assemblage of bacterial, archaeal, and eukaryotic

components that coexist in a functional unit. This is the result of billions of years of genetic changes, exchanges, gains, and losses during evolution from the most primitive ancestral forms to today's *Homo sapiens*.

Stress Proteins in the Various Compartments of the Human Cell

A sample of stress proteins that have been studied in human cells and whose corresponding genes have been identified and characterized, at least partially, is given in [Table 1](#), which classifies stress proteins based on molecular mass.

Comparative sequence analyses of the type mentioned in the previous sections have provided insights into the evolutionary origins of the stress proteins residing in the cytosol, nucleus, mitochondria, and ER. For instance, mitochondrial Hsp60 is derived from ancestors of modern gram-negative bacteria (it is equivalent to the GroEL chaperonin of extant bacteria), whereas the cytosolic Hsp60 is the homolog of the chaperonin subunits that compose the thermosome in modern archaea. The Hsp70(DnaK) in all compartments of the human cell is related to a gram-negative bacterial ancestor and is similar to the homologs of extant gram-negative bacteria but distinct from those of modern gram positives.

Stress Genes

Human stress genes and their proteins have been studied only to a limited extent compared to those from other eukaryotic organisms. A sample of the human cell types, cell lines, and tissues in which expression of the stress genes has been assessed is displayed in [Table 2](#).

In general, constitutive (basal) gene expression was measured and compared with expression induced by a stressor, for instance, a temperature higher than optimal for cell growth (heat shock) or some other agent of medical significance, such as a heavy metal. Whenever possible, the levels of expression in pathological cells were compared with those in normal counterparts. Overall, the results showed alterations in the levels of expression in cells from diseased tissues, increases more often than decreases, as compared with normal controls. However, the physiopathological or pathogenic significance of these alterations could not be established with certainty in most instances. This deficiency is chiefly attributable to the difficulties in designing properly controlled experiments, including direct manipulation of cells or tissues, when performing research involving human subjects and clinical samples. Hence, many more data are available from experiments done with simple

Table 1 Examples of human stress (heat shock) and molecular chaperone proteins whose genes have been at least partially identified and characterized

Family		Locale of residence in the cell ^a		
Name(s)	Mass (kDa)	Cytosol/Nucleus	Mitochondria	Endoplasmic reticulum (ER)
Heavy; high molecular weight; Hsp100	100 or higher	Hsp110; Hsp105alpha; Hsp105beta		
Hsp90	81–99	Hsp90alpha (Hsp90; Hsp84) Hsp90beta (Hsc90; Hsp86)		Hsp90alpha (Hsp90; Hsp84); Hsp90beta; (Hsc90; Hsp86); Grp94
Hsp70; chaperones	65–80	Hsp70–1; Hsp70–2 (Hsc70); Hsp-Hom		Grp78 (BiP)
Hsp60	55–64	Hsp60	Cpn60(Hsp60)	
chaperonins group I		CCT (TriC; TCP-1) 8–9 subunits		
chaperonins group II				
Hsp40; DnaJ	35–54	Hsp40; HDJ-1; HSJ-1; HSJ-2; HDJ-2; Hsc54		HDJ-1 (Hsp40); Hsp47
Other	34 or lower			
small Hsp; sHsp		Hsp27; alpha-crystallin family (HspB1–10)		
Miscellanea		PDIase; PPIase	Hsp23(GrpE); Cpn10(Hsp10)	PDIase

^aThe corresponding genes are in the nuclear DNA.**Table 2** Some human cell types, cell lines, and tissues in which expression of stress (heat shock) genes has been investigated^a

Cells/lines/tissues	Gene
Cells	
Fibroblasts	<i>hsp70i</i> ; <i>hsc70(hsp73)</i> ; HSP70B _p ; <i>hsp60</i>
Melanocytes	<i>Hsp70</i>
Lymphoid tumor cells	<i>hsp72(hsp70i)</i>
Macrophages	<i>hsp70</i>
Mononuclear cells from peripheral blood	<i>hsp90</i> ; <i>hsp70</i>
Hepatocytes	<i>hsp70–1</i> ; <i>hsc70</i> ; <i>hsp90</i>
Unbilical vein endothelial cells	<i>grp78</i>
Wilms tumor cells	<i>hsp70i</i>
Cell line	
Breast cancer T47-D	<i>grp94</i> ; <i>grp78</i> ; <i>hsp72</i>
Embryonic kidney 293	<i>hsp47</i>
Erythroleukemia K562	<i>hsp70</i> ; <i>hsp70–1</i> ; <i>hsp70–2</i> ; <i>hsp70B'</i>
RAJI	<i>hsp70–1</i> ; <i>hsp70–2</i> ; <i>hsp70B'</i>
HL60	<i>hsp70–1</i> ; <i>hsp70–2</i> ; <i>hsp70B'</i>
HeLa	<i>hsp105</i> ; <i>grp78</i> ; <i>hsp70i(hsp72)</i> ; <i>hsc70(hsp73)</i>
Human lung tumor HS24	<i>hsc70-like</i> or HSP24/p52
Melanoma	<i>grp94</i> ; <i>hsp90α</i> ; <i>hsp90β</i> ; HSP70–1; HSP70–2; <i>hsp27</i>
Tissue	
Arterial intima	<i>hsp70</i> ; <i>hsp60</i>
Jejunal epithelium	<i>hsp60</i>
Liver	<i>hsp60</i>
Lung	<i>hsp73</i> ; <i>hsp72</i> ; <i>hsp60</i>
Skeletal muscle	<i>hsp90</i> ; <i>hsp70</i>
Stomach	<i>hsp60</i>
Retina	<i>alpha-crystallin</i>

^aConstitutive (basal) and stressor (usually heat shock)-induced expression was studied in many cases, with pathologic cells or tissues compared to their normal counterparts whenever possible.

organisms (e.g., the bacterium *Escherichia coli*, the archaeon *Methanosarcina mazei*, and the unicellular fungus *Saccharomyces cerevisiae*) and animal models (e.g., mice, rats, and gerbils) than are available from research carried out directly in human cells or tissues.

In animal models, it has been shown, for example, that the levels of stress proteins, especially Hsp70 (DnaK), are elevated in tissues damaged by ischemia-hypoxia. Research has focused on determining the presence of Hsp70 (DnaK) in brain and heart tissues afflicted by ischemia-hypoxia. Around the dead, central zone of an infarct in the brain (stroke) or in the heart (cardiac attack), caused by occlusion of an artery and blood flow interruption, there are zones in which the brain or heart cells are not dead, but only injured. These zones are called penumbra in the brain and areas of stunned myocytes in the heart. Cells in such zones are damaged mainly due to protein denaturation. The latter may be reversible, particularly with the help of stress proteins that, like Hsp70 (DnaK), are molecular chaperones.

The Human Genome

The human genome has been sequenced, and the information is now being processed, with the data available at various websites (Table 3). It will take considerable effort over the coming years to process all the data and to complete the annotation, based on thorough analyses of each gene (genomics) and gene product (proteomics). A hint at the enormous complexity of the situation is provided, for instance, by the fact that at present there are more than 30 candidate genes that encode Hsp40 (DnaJ)-related proteins and more than 20 that encode Hsp70 (DnaK)-related

proteins. A similar situation holds true for all the other heat shock proteins and molecular chaperones.

Perspectives

A full view of the human species' stress gene candidates should become available when the sequence of the human genome has been more fully analyzed. In addition to identification of the genes through comparisons of their sequences – and those of the proteins encoded in them – with homologs of known functions from other organisms, experiments will be required to establish whether these genes are transcribed and translated and also to determine these genes' responses to various stressors. Based on the results, it should become possible to devise strategies for using the genes, their regulatory components and factors, and the proteins that they produce to both prevent and treat stress-induced lesions and diseases.

Also, a detailed phylogenetic analysis of each heat shock protein and molecular chaperone will be necessary if we are to understand their evolution and if we are to gain insight into their biological roles in organisms from the three phylogenetic domains. These analyses will also provide valuable information on the evolution and function of each protein's structural-functional domain and motif present in human heat shock proteins and molecular chaperones. In turn, this information will be instrumental for elucidating the nature of genes and gene products that are abnormal in genetic and acquired chaperonopathies; currently, only preliminary information is available for these disorders, suggesting that the affected molecules are chaperones.

See Also the Following Articles

Chaperone Proteins and Chaperonopathies; Chaperonopathies; Heat Resistance; Heat Shock Response, Overview; Viral Virulence and Stress; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles.

Further Reading

- Andley, U. P., Mathur, S., Griest, T. A. and Petrash, J. M. (1996). Cloning, expression, and chaperone-like activity of human alpha-crystallin. *Journal of Biological Chemistry* 271, 31973–31980.
- Bukach, O. V., Seit-Nebi, A. S., Marston, S. B. and Gusev, N. B. (2004). Some properties of human small heat-shock protein Hsp20 (HspB6). *European Journal of Biochemistry* 271, 291–302.
- Ellgaard, L. and Ruddock, L. W. (2005). The human protein disulphide isomerase family: Substrate interactions and functional properties. *EMBO Reports* 6, 28–32.

Table 3 Human genome websites

Website	Chromosome
http://www.gdb.org	1–22, X, Y
http://www.ensembl.org	1–22, X, Y
http://www.ncbi.nlm.nih.gov/genome/guide/human/	122, X, Y
http://www.jgi.doe.gov	5, 16, 19
http://www.sanger.ac.uk/HGP/	1, 6, 9, 10, 11, 13, 20, 22, X
http://genome.ucsc.edu	1–22, X, Y
http://www.hgsc.bcm.tmc.edu	1–22, X, Y
http://www.ncbi.nlm.nih.gov/Sitemap/index.html#HumanGenome	1–22, X, Y
http://www.shgc.stanford.edu/finishing/human.html	5, 16, 19
http://www.nature.com/cgi-taf/DynaPage.taf?file=nature/journal/v434/n7034/full/nature03466_fs.html	2, 4

- Hansen, J. J., Bross, P., Westergaard, M., et al. (2003). Genomic structure of the human mitochondrial chaperonin genes: HSP60 and HSP10 are localised head to head on chromosome 2 separated by a bidirectional promoter. *Human Genetics* **112**, 71–77.
- Hinton, A., Gatti, D. L. and Ackerman, S. H. (2004). The molecular chaperone, Atp12, from *Homo sapiens*. *Journal of Biological Chemistry* **279**, 9016–9022.
- Ishihara, K., Yasuda, K. and Hatayama, T. (1999). Molecular cloning, expression and localization of human 105 kDa heat shock protein, hsp105. *Biochimica et Biophysica Acta* **1444**, 138–142.
- Kappe, G., Franck, E., Verschuure, P., Boelens, W. C., Leunissen, J. A. M. and de Jong, W. W. (2003). The human genome encodes 10 alpha-crystallin-related small heat shock proteins: HspB1–10. *Cell Stress and Chaperones* **8**, 53–61.
- Kleizen, B. and Braakman, I. (2004). Protein folding and quality control in the endoplasmic reticulum. *Current Opinion in Cell Biology* **16**, 343–349.
- Leung, T. K. C., Rajendran, M. Y., Monfries, C., Hall, C. and Lim, L. (1990). The human heat shock protein family. *Biochemical Journal* **267**, 132–152.
- Macario, A. J. L. (1995). Heat-shock proteins and molecular chaperones: implications for pathogenesis, diagnostics, and therapeutics. *International Journal of Clinical Research* **25**, 59–70.
- Macario, A. J. L. and Conway de Macario, E. (1999). The archaeal molecular chaperone machine: Peculiarities and paradoxes. *Genetics* **152**, 1277–1283.
- Macario, A. J. L. and Conway de Macario, E. (2002). Sick chaperones and ageing: a perspective. *Ageing Research Reviews* **1**, 295–311.
- Macario, A. J. L. and Conway de Macario, E. (2004). The pathology of anti-stress mechanisms: a new frontier. *Stress* **7**, 243–249.
- Macario, A. J. L., Grippo, T. M. and Conway de Macario, E. (2005). Genetic disorders involving molecular-chaperone genes: a perspective. *Genetics in Medicine* **7**, 3–12.
- Meacham, G. C., Lu, Z., King, S., Sorscher, E., Tousson, A. and Cyr, D. M. (1999). The Hdj-2/Hsc70 chaperone pair facilitates early steps in CFTR biogenesis. *EMBO Journal* **18**, 1492–1505.
- Nagai, N., Tetuya, Y., Hosokawa, N. and Nagata, K. (1999). The human genome has only one functional *hsp47* gene (CBP2) and a pseudogene (*pshsp47*). *Gene* **227**, 241–248.
- Orchard, S., Hermjakob, H. and Apweiler, R. (2005). Annotating the human proteome. *Molecular and Cellular Proteomics* **4**, 435–440.
- Snoeckx, L. H. E. H., Cornelussen, R. N., van Nieuwenhoven, F. A., Reneman, R. S. and van der Vusse, G. J. (2001). Heat shock proteins and cardiovascular pathophysiology. *Physiology Reviews* **81**, 1461–1497.
- Won, K.-A., Schumacher, R. J., Farr, G. W., Horwich, A. L. and Reed, S. I. (1998). Maturation of human cyclin E requires the function of eukaryotic chaperonin CCT. *Molecular and Cellular Biology* **18**, 7584–7589.

Heat Shock Proteins See: Chaperone Proteins and Chaperonopathies; Heat Shock Response, Overview.

Heat Shock Proteins: HSP60 Family Genes

H Kubota

Institute for Frontier Medical Sciences, Kyoto University,
Kyoto, Japan

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by H Kubota, volume 2, pp 347–349, © 2000, Elsevier Inc.

Introduction

Members of the HSP60 Family

Response to Heat Shock and Related Stress

Function in Protein Folding

Medical Aspects

Glossary

Chaperonin
containing
TCP-1 (CCT)

The HSP60 family member protein distributed in the eukaryotic cytosol and nucleus. CCT is also called TRiC or c-cpn.

Chaperonins

Alternative name for the HSP60 family, preferentially used by molecular biologists and biochemists. This term is often

	confused with chaperones, which refers to all molecular chaperones. The term chaperonin is also used for the 10-kDa co-chaperone of the HSP60 family. In this case, the HSP60 family and 10-kDa co-chaperone are called chaperonin 60 (Cpn60) and 10 (Cpn10), respectively.
<i>GroEL</i>	The HSP60 family member protein contained in <i>Escherichia coli</i> and related bacterial species.
<i>Group II archaeobacterial chaperonins</i>	The HSP60 family member proteins contained in archaeobacterial species.
<i>Heat shock proteins (HSPs)</i>	Proteins induced by heat shock and related stress.
<i>HSP60 (60-kDa heat shock protein)</i>	The HSP60 family member protein distributed in mitochondria.
<i>Molecular chaperone</i>	A group of proteins that associate with unfolded proteins and assist in their folding.
<i>Protein folding</i>	The process by which newly synthesized polypeptides or denatured proteins produce functional conformations.
<i>Rubisco subunit binding protein</i>	The HSP60 family member protein distributed in plastids.

Introduction

Heat shock proteins (HSPs) are a group of proteins induced by heat shock and related stress including heavy metals and amino acid analogs. Most HSPs were named according to their molecular mass in kilodaltons with the abbreviation HSP (such as HSP90, HSP70, and HSP60) in mammals. Proteins in the HSP60 family (also called the chaperonin family) form large complexes composed of subunits of approximately 60 kDa. Each complex contains 14–18

subunits and shows a double doughnut-like structure. HSP60 family proteins act as molecular chaperones, assisting in refolding of proteins denatured by heat and related stress and in folding of newly synthesized polypeptides. These proteins were found to be distributed in all organisms investigated to date and in most compartments of eukaryotic cells, including the cytosol, nucleus, mitochondria, and chloroplast. The ubiquitous distribution and chaperone activity of the HSP60 family proteins indicate that the HSP60 family proteins play fundamental roles in the survival of organisms by mediating production and recovery of functional conformations of proteins in the cell.

Members of the HSP60 Family

The HSP60 family includes GroEL of eubacteria, HSP60 of mitochondria, Rubisco subunit binding protein of plastids, the chaperonin containing TCP-1 (CCT, also called TRiC or c-cpn) of eukaryotic cytosol and nucleus, and archaeobacterial chaperonins (e.g., TF55/56, thermosome). These proteins can be divided into two groups based on the amino acid sequence homology and structural similarities (Table 1). GroEL, HSP60, and Rubisco subunit binding protein are classified as group I, while CCT and archaeobacterial chaperonins are classified as group II. GroEL is a major heat-inducible protein of *Escherichia coli* and plays a fundamental role in the stress response by assisting in the refolding of denatured proteins. GroEL is also important for the folding of newly synthesized proteins. HSP60 assists in refolding of mitochondrial proteins imported from the cytosol across the mitochondrial membrane, or after denaturation by heat and related stress. A considerable level of HSP60 is also found in the cytosol, although roles of the cytosolic HSP60 have not been clarified. Rubisco subunit binding protein mediates folding of ribulose-1,5-bisphosphate carboxylase/oxygenase

Table 1 Characteristics of HSP60 family proteins^a

	Organism	Localization	Amino acid identity with		Rotational symmetry	Subunit species	Inducibility by stress
			GroEL	CCT			
Group I							
GroEL	Eubacteria	Soluble	–	<20%	7	1	+
HSP60	Eukaryotes	Mitochondria	60%	<20%	7	1	+
RSBP ^b	Plants	Plastids	60%	<20%	7	2	+
Group II							
CCT (TRiC, c-cpn)	Eukaryotes	Cytosol and nucleus	<20%	–	8	8–9	+
TF55/56	Archaeobacteria	Soluble	<20%	40%	9	3	+
Thermosome	Archaeobacteria	Soluble	<20%	40%	8	2	+

^aMKKS and BBSIO show very weak homology to HSP60 family proteins.

^bRubisco subunit binding protein.

(Rubisco) in plastids. CCT is known to assist in the folding of newly synthesized actin, tubulin, and some other proteins in eukaryotic cytosol and is composed of eight different subunit species encoded by independent genes. Archaeobacterial chaperonins are the major heat-inducible proteins and bind denatured proteins, although they do not always show refolding activity. Recently, a gene encoding novel mammalian chaperonin-like protein was discovered by chromosome mapping for the McKusick–Kaufman syndrome (MKKS) gene. The MKKS gene was also found to be the sixth locus for Bardet–Biedl syndrome (BBS6). The amino acid sequence of MKKS shows a weak identity to those of group II subfamily members. Although this protein distributes in the cytosol with a highly concentrated distribution in the centrosome when overexpressed in mammalian cells, its function and stress inducibility have not been determined. Very recently another BBS gene product (BBSIO) was found to show very weak identity to the HSP family, although no experimental information is currently available for the BBSIO protein.

Response to Heat Shock and Related Stress

GroEL is required during normal and stress conditions for the survival of *E. coli*. Expression of GroEL in *E. coli* is strongly induced by a variety of stressors such as heat shock, ethanol shock, heavy metals, and amino acid analogs. These treatments cause denaturation of proteins in the bacteria, and the denatured proteins require help from molecular chaperones to recover functional conformations. The heat shock response of GroEL is mediated by a transcription factor called σ_{32} , which is the product of the *RpoH* gene. The translation of σ_{32} becomes more efficient at higher temperatures by structural changes of its mRNA, and the translated protein transiently becomes more stable after heat shock. The activity of σ_{32} is inhibited by GroEL but not by substrate-occupied GroEL, suggesting the presence of a transcription factor–chaperone network for sensing and controlling protein folding environment *in vivo*.

Mammalian HSP60 is constitutively expressed under normal conditions and upregulated under stress conditions. Heat shock and related stress induce HSP60 expression, and this response is regulated by a family of transcription factors called heat shock factors (HSFs) through the heat shock elements (HSEs) in the HSP60 gene promoter. In vertebrates, four members of this family (HSF1–4) are known, and HSF1 plays the most important role in controlling the heat shock response. HSF1 is activated by heat shock and related stress, and the HSF1 activity is controlled by

HSP70 and HSP90 by monitoring the protein folding environment at least in part.

CCT is expressed at a basal level under normal growth-arrested conditions such as terminally differentiated cells *in vivo*. CCT expression is upregulated by growth stimulation. For example, CCT is highly expressed in embryos, tumors, and rapidly growing cultured cells. Under the rapidly growing conditions, CCT expression is upregulated by transcription factors whose expression and activity are highly dependent on growth stimulation. In addition, CCT is highly expressed in tissues that are abundant in substrate proteins. CCT expression is particularly high in testis to assist in the folding of tubulin, a major component of sperm tail. CCT is further induced by chemical stress including arsenite and proline analog treatments. As CCT gene promoters contain HSE-like sequences that are recognizable by HSF1 and HSF2, these factors may play a role in the chemical stress-induced expression of CCT. In contrast, CCT was not induced by heat shock as far as investigated. Thus, the mechanism of stress-induced regulation of CCT expression is probably somewhat different from that of typical HSPs.

In several archaeobacterial species, chaperonins become the most major protein under stress conditions. This is considered to be due to severe environmental conditions. For example, thermophilic species grow at temperatures above 80°C.

Function in Protein Folding

The major contribution of HSP60 family proteins in the stress response is to assist in refolding denatured proteins. All the members bind denatured proteins but do not bind functional proteins that are completely folded. Most members of this family show ATPase activity and use it to obtain conformational changes that allow these proteins to release target proteins. Activity of group I members is modulated by co-chaperones of approximately 10 kDa, such as GroES and cpn-10, although homologs of the co-chaperone for group II members are not known. All HSP60 family members have back-to-back stacked double-ring-like structure composed of 14–18 subunits, although rotational symmetry differs between group I and II: group I members show sevenfold symmetry whereas group II members show eight- or ninefold symmetry. Denatured substrate protein is trapped in the central cavity of these cylinder-like molecules near their opening in the absence of ATP. In the case of GroEL, ATP and GroES are bound to the cylinder following trapping substrate. GroES acts as a lid of the GroEL cylinder, and the substrate is encapsulated in the GroEL–GroES cage. In this protected environment, the single substrate

protein is allowed to fold properly without forming nonproductive aggregation. After ATP hydrolysis, the GroES lid and folded substrate are released from the GroEL cylinder. In contrast, group II members including CCT and archaeobacterial chaperonins are able to produce cage-like structures for encapsulating substrate proteins in the absence of GroES-like co-chaperones.

The HSP60 family proteins are considered to cooperate with other molecular chaperones such as HSP70 and HSP40 (DnaJ homolog) family proteins. The HSP60 family proteins also play important roles in the folding of newly synthesized proteins with other molecular chaperones and co-chaperones. In this view, *E. coli* strains overexpressing GroEL can be used to produce some recombinant proteins efficiently.

Medical Aspects

It is known that HSP60 homologs in mycobacteria and related species are highly antigenic in mammals. Immune responses to these proteins and/or HSP60 itself are considered to be involved in several diseases mediated by the immune system, including reactive arthritis, juvenile chronic arthritis, rheumatoid arthritis, Lyme disease, Bechet's disease, systemic lupus erythematosus and insulin-dependent diabetes mellitus.

Recently, SPG13, a dominant human gene responsible for hereditary spastic paraplegia, was mapped to the gene encoding mitochondrial HSP60. Hereditary spastic paraplegia is a genetically heterogeneous neurodegenerative disorder characterized by progressive spasticity and weakness of lower limbs. HSP60 carrying the V72I mutation that is found in the patient group cannot compensate for the function of GroEL when expressed in *E. coli*, whereas wild-type HSP60 can. These observations suggest that the chaperone function of HSP60 plays an important role in preventing diseases caused by protein misfolding, including neurodegenerative disease.

Although exact function of MKKS protein is unknown, its weak homology to the HSP60 family suggests that a particular function may be shared between MKKS and the HSP60 family proteins. Mutations in the MKKS gene cause human developmental anomalies including vaginal atresia with hydrometrocolpos, polydactyly, and congenital heart defects. MKKS also causes BBS, and provides more severe symptoms when mutated together with other BBS genes. BBS is a genetically heterogeneous disorder characterized by obesity, retinal dystrophy, polydactyly, mental retardation, renal malformations, and hypogenitalism. MKKS knockout mice show BBS-like phenotypes including retinal degeneration, failure of sperm flagella formation, photoreceptor

degeneration, obesity, and deficits in olfaction and social dominance. As MKKS and some BBS proteins are highly concentrated in the centrosome and/or basal body of cilia, MKKS protein function may be related to the activity of these organelles and may work together for this activity. It is also possible that MKKS protein functions in protein transport, as suggested for other BBS proteins. As recently identified BBS10 also shows weak homology to the HSP60 family proteins, the function of this BBS protein may be related to that of MKKS.

See Also the Following Articles

Chaperone Proteins and Chaperonopathies; Genetic Factors and Stress; Heat Shock Genes, Human; Heat Shock Response, Overview.

Further Reading

- Barral, J. M., Broadley, S. A., Schaffar, G. and Hartl, F. U. (2004). Roles of molecular chaperones in protein misfolding diseases. *Seminars in Cell and Developmental Biology* 15, 17–29.
- Beales, P. L. (2005). Lifting the lid on Pandora's box: the Bardet-Biedl syndrome. *Current Opinion in Genetics & Development* 15, 315–323.
- Bukau, B. (ed.) (1999). *Molecular chaperones and folding catalysts: regulation, cellular function and mechanisms*. New York: Harwood Academic Publishers.
- Ellis, R. J. (ed.) (1996). *The chaperonins*. San Diego, CA: Academic Press.
- Fares, M. A., Moya, A. and Barrio, E. (2004). GroEL and the maintenance of bacterial endosymbiosis. *Trends in Genetics* 20, 413–416.
- Fenton, W. A. and Horwich, A. L. (2003). Chaperonin-mediated protein folding: fate of substrate polypeptide. *Quarterly Review of Biophysics* 36, 229–256.
- Gething, M.-J. (ed.) (1997). *Guidebook to molecular chaperones and protein-folding catalysts*. Oxford, UK: Oxford University Press.
- Gomez-Puertas, P., Martin-Benito, J., Carrascosa, J. L., Willison, K. R. and Valpuesta, J. M. (2004). The substrate recognition mechanisms in chaperonins. *Journal of Molecular Recognition* 17, 85–94.
- Kubota, H. (2002). Function and regulation of cytosolic molecular chaperone CCT. *Vitamins and Hormones* 65, 313–331.
- Latchman, D. S. (ed.) (1999). *Stress proteins, handbook of experimental pharmacology* (Vol. 136). Berlin: Springer-Verlag.
- Morimoto, R. I., Tissieres, A. and Georgopoulos, C. (eds.) (1994). *The biology of heat shock proteins and molecular chaperones*. New York: Cold Spring Harbor Laboratory Press.
- Prakken, B. J., Roord, S., Ronaghy, A., Wauben, M., Albani, S. and van Eden, W. (2003). Heat shock

protein 60 and adjuvant arthritis: a model for T cell regulation in human arthritis. *Springer Seminars in Immunopathology* 25, 47–63.

Spieß, C., Meyer, A. S., Reissmann, S. and Frydman, J. (2004). Mechanism of the eukaryotic chaperonin: protein folding in the chamber of secrets. *Trends in Cell Biology* 14, 598–604.

Thomas, J. G., Ayling, A. and Baneyx, F. (1997). Molecular chaperones, folding catalysts, and the recovery of

active recombinant proteins from *E. coli*: to fold or to refold. *Applied Biochemistry and Biotechnology* 66, 197–238.

van Eden, W. and Young, D. (eds.) (1996). *Stress proteins in medicine*. New York: Dekker.

Young, J. C., Agashe, V. R., Siegers, K. and Hartl, F. U. (2004). Pathways of chaperone-mediated protein folding in the cytosol. *Nature Reviews Molecular and Cell Biology* 5, 781–791.

Heat Shock Response, Overview

A J L Macario and E Conway de Macario
New York State Department of Health and the
University at Albany (SUNY), Albany, NY, USA

Published by Elsevier Inc.

Heat Shock and Stress

Definition

Scope

Events

Characteristics

Mechanisms

Clearing the Rubble

Milestones

New Developments to 2005

Perspectives

Glossary

Heat shock gene

Microarray

A gene inducible by heat shock or any other stressor; also called stress gene.

A miniaturized collection of samples (e.g., cDNA made from the mRNAs present in a cell under specific experimental conditions), anchored in a predetermined spatial order to many microareas delimited on a support (e.g., microchip, glass slide, filter, or silicon wafer); this makes the samples available for reaction (e.g., cDNA–DNA hybridization) in a high-throughput mode with known, labeled probes (e.g., DNA with sequences of known genes).

Messenger ribonucleic acid (mRNA)

A molecule composed of bases in a sequence complementary to that of the coding region of a gene; it has the message to synthesize the protein encoded in the gene.

Protein

A molecule formed by amino acids in a sequence dictated by the sequence of bases in the corresponding mRNA.

Ribosome

A large, multimolecular structure present in the cytoplasm of all cells at which protein synthesis takes place according to the message encoded in the mRNA.

Stressor

An agent, which can mechanical, physical, psychological, or chemical, capable of causing cell stress and inducing a stress (heat shock) response. Although the term heat shock response may be used as synonym of stress response, it is more properly applied only to stress responses caused by the stressor heat.

Systems biology

Large-scale, computer-assisted analyses of genomic and proteomics data to elucidate intra- and intercellular gene and protein interactions as they might occur *in vivo*. The purpose of such analyses is ultimately to provide mathematical models and simulations of cellular pathways and networks and of their dynamics. An example is the pathways followed by each chaperoning system and the networks formed by two or more of these systems between them and/or between them and protein-degrading systems so as to maintain intracellular protein homeostasis.

Transcription

The process of synthesis of the mRNA, marking the initial main step of gene expression.

Transcriptome

The whole suite of mRNAs that can be transcribed from a genome; also, more specifically, the subset of mRNAs that are transcribed from a genome of a particular cell (or organism or tissue), at any given time, under specific conditions, for example, under conditions of cell stress

Translation

or under basal, constitutive conditions (no stress).

The process of synthesis of a protein molecule on the ribosome, in which amino acids are linked in a linear non-branching chain that follows the order encoded in the corresponding mRNA.

Heat Shock and Stress

A heat shock response consists of the series of events occurring in a cell that are induced by heat stress (i.e., a sudden temperature elevation) and that reflect the cell's perturbed status compared with the status prior to the stressor's impact. Among all such events, only a few are considered in this article, whose focus is intracellular proteins produced by stress-inducible genes and the phenomena mediated by these proteins.

Many stressors other than a sudden increase in temperature can cause virtually the same series of events as does heat shock. Such cellular reactions are also called heat shock responses for historical reasons. To avoid confusion, the term stress response should be used regardless of the stressor, and the stressor should be mentioned in all cases. Examples of common cell stressors are listed in [Table 1](#). In this article, the expressions stress response, stress gene, and stress protein are preferred over heat shock response, heat shock gene, and heat shock protein, respectively.

Definition

The shock (or stress) produced on an object by an abrupt temperature elevation is called heat shock or

heat stress. Shock and stress are considered synonyms within the context of this article and may be defined as the impact – and its consequences – of an agent, such as heat, on an object. Stress is a perturbed status of the object generated by the impacting agent. The object may be inanimate (e.g., a piece of metal) or living (e.g., a cell or an organism made up of one or more cells). This article deals chiefly with cell stress.

Scope

Stress may be defined more specifically from several different standpoints, depending on the natures of the object and the stressor and on the particular field of knowledge being considered, materials, physical, mechanical, psychological, physiological, or medical sciences, just to name some representative examples.

This article deals with cellular stress caused by heat shock, or heat stress, and by other stressors that produce very similar effects, pertaining to biology in its broadest sense – that of all living organisms, whether uni- or multicelled. If the organism is unicelled, cellular and organismal stresses are synonymous. An illustrative list of organisms that have been used to study the stress response is in [Table 2](#).

It must be borne in mind that although the effects of stressors on the cell are, on the whole, not specific, there are details of the cell's response to stress that vary according to the stressor.

Events

The stress response manifests itself in diverse cellular changes that affect many of the cell's components and properties, from gross morphology to the most

Table 1 Examples of cell stressors^a

Type/group	Name/description
Physical	Heat (including fever), cold, irradiation (e.g., ultraviolet rays), magnetic fields
Mechanical	Compression, shearing, stretching
Nutritional	Starvation: multiple, specific (carbon, glucose, nitrogen, phosphate, nitrate)
pH	Alkalosis, acidosis, pH shift
Osmotic	Changes in the concentration of salt, sugars, other osmolytes (hyper- and hypoosmotic shocks)
Antibiotics	Puromycin, tetracycline, nalidix acid
Alcohols	Ethanol, methanol, butanol, propanol, octanol
Oxygen	Oxygen-derived free radicals (ROS), hydrogen peroxide, anaerobiosis to aerobiosis shift (e.g., reperfusion), hypoxia/anoxia (ischemia)
Biological	Infection, inflammation, fever
Psychological	Emotions, conflicts; hormonal imbalance (HPA axis, autonomic nervous system)
Metals	Cadmium, copper, chromium, zinc, tin, aluminum, mercury, lead, nickel
Other	Desiccation, benzene and derivatives, phenol and derivatives, teratogens, carcinogens, mutagens, arsenite, arsenate, amino acid analogs, nicotine, anesthetics, insecticides, pesticides

^aThese agents cause stress not only in human and other (e.g., mouse and rat) eukaryotic cells but also in prokaryotic (e.g., bacterial) cells. HPA, hypothalamic-pituitary-adrenal; ROS, reactive oxygen species. Modified from Macario, A. J. L. and Conway de Macario, E. (2000), Stress and molecular chaperones in disease, *International Journal of Clinical and Laboratory Research* **30**, 49–66, with permission from the copyright owner.

Table 2 Examples of organisms whose cells and tissues have been used as experimental models to study cell stress and the stress (heat shock) response

Type	Name
Prokaryote	
Bacterium	<i>Agrobacterium tumefaciens</i> , <i>Anabaena</i> sp., <i>Bacillus subtilis</i> , <i>Borrelia burgdorferi</i> , <i>Bradyrhizobium japonicum</i> , <i>Campylobacter jejuni</i> , <i>Clostridium acetobutylicum</i> , <i>Escherichia coli</i> , <i>Lactococcus lactis</i> , <i>Rhodothermus obamensis</i> , <i>Shewanella oneidensis</i> , <i>Staphylococcus aureus</i> , <i>Streptomyces coelicolor</i> , <i>Syneccoccus</i> sp.
Archaeon	
Methanogen	<i>Methanosarcina mazei</i> , <i>Methanosarcina thermophila</i> , <i>Methanococcus voltae</i>
Extreme halophile	<i>Halobacterium halobium</i> , <i>Haloferax volcanii</i>
Thermophile	<i>Thermoplasma acidophilum</i>
Hyperthermophile	<i>Archaeoglobus fulgidus</i> , <i>Pyrodictium occultum</i> , <i>Sulfolobus shibatae</i>
Eukaryote	
Unicelled	<i>Arxiozyma telluris</i> , <i>Mrakia frigida</i> , <i>Saccharomyces cerevisiae</i> (yeasts) <i>Tetrahymena thermophila</i> (protozoan) <i>Euglena gracilis</i> (alga)
Multicelled	<i>Neurospora crassa</i> (fungus) <i>Schistosoma mansoni</i> , <i>Caenorhabditis elegans</i> (worms) <i>Crassostrea gigas</i> (oyster) <i>Drosophila melanogaster</i> (insect) <i>Salmo gairdnerii</i> [rainbow trout], <i>Danio rerio</i> [zebrafish] (fishes) <i>Arabidopsis thaliana</i> [wall cress], <i>Lycopersicon esculentum</i> [tomato] (plants) <i>Gallus gallus</i> [chicken] (bird) <i>Mus musculus</i> [mouse], <i>Rattus norvegicus</i> [rat], <i>Meriones unguiculatus</i> [gerbil], <i>Homo sapiens</i> [human] (mammals)

detailed biochemical and molecular mechanisms. A list of cellular components and properties affected by stressors such as heat shock and the methods used to study the cell's response or reaction to stressors are shown in [Table 3](#).

Characteristics

The stress response is characterized by a decrease in the levels of many proteins and a concurrent increase in the levels of others. In addition, a series of proteins appear that were not detectable prior to the stressor's impact. The proteins that increase or appear as a consequence of stress are the stress proteins. The subset of these proteins that are the product of genes induced by heat shock are called heat shock proteins (Hsps). However, for historical reasons the same name Hsp is also used to designate proteins encoded in genes that are inducible by any stressor. An illustrative sample of stress proteins, including Hsps, is presented in [Table 4](#). In this table, some molecules (thermoprotectants) that are not proteins but that, like the stress proteins, function to counteract the effects of stressors are also mentioned. These nonproteinic antistress compounds are mentioned to convey the idea that the cell possesses a variety of antistress mechanisms beyond the stress proteins and also to draw attention to

the fact that all the proteins (e.g., enzymes) that participate in the production and trafficking of these antistress compounds must work during the stress response and that, therefore, they are part of the antistress mechanisms, too.

Mechanisms

Many genes are downregulated or shut off as a consequence of stress. If the affected genes are ones that produce proteins, their protein products decrease or disappear. In contrast, other genes are upregulated or induced *de novo* by stress. These are the heat shock or stress genes *sensu lato*. Among them, there is a group that belongs to the category of protein producers. Their products are the heat shock or stress proteins, some of which are important for cell survival during stress and for cell recovery after stress. A subset of heat shock proteins are molecular chaperones; they assist other proteins to fold correctly during synthesis and to refold if they have been partially denatured by stress. One of the best-known molecular chaperones is the stress protein Hsp70(DnaK), which interacts with at least two others, Hsp40 (DnaJ) and GrpE in prokaryotes (or BAG-1, or HspPB1 in eukaryotes), to bring about protein folding or refolding.

Table 3 Cellular components and properties affected by heat shock and methods used to study the stress (heat shock) response of the cell

<i>Component / property</i>	<i>Method</i>
Viability, growth rate	Cell counts
Morphology (size; shape; formation of multicellular structures or spores)	Various types of light and electron microscopies Biochemical and immunochemical analyses of cell envelopes and intercellular connective materials Histochemistry Immunohistochemistry
Proteins	
General contents	One- and two-dimensional electrophoresis
Synthesis, general	Incorporation of radioactively labeled amino acids followed by quantification of radioactive fractions resolved by electrophoresis or other fractionation methods (e.g., chromatography) DNA microarray-proteomics (two-dimensional electrophoresis plus other methods, e.g., western blotting, mass spectrometry, microsequencing) to determine the exact nature of the proteins synthesized in response to stress Two-hybrid method to detect protein-protein interactions followed by bioinformatics (systems biology, neural networks)
Synthesis, specific	Separation and/or quantification of target stress proteins using specific antibodies (immunoprecipitation, western blotting) Identification of protein complexes by electron microscopy combined with immunochemistry and functional assays (e.g., determination of ATPase activity and capacity to assist protein folding or refolding) Microsequencing
Genes ^a	Measurement of transcript (mRNA) for stress proteins: quantity, size (length), transcription initiation and termination sites (northern blotting, primer extension, S-1 nuclease protection, reverse PCR) DNA microarray-transcriptomics and proteomics (two-dimensional electrophoresis plus other methods, e.g., northern and western blottings, mass spectrometry, microsequencing) to determine the exact nature of the proteins synthesized in response to stress and, thus, identify genes involved Systems biology, neural networks

^aPresence or absence of genes is determined by genome sequencing; cloning and sequencing; and Southern, northern, and western blottings. PCR, polymerase chain reaction.

Table 4 Some stress (heat shock) proteins proven or believed to play a role during the stress (heat shock) response and in cell recovery after stress^a

<i>Group</i>	<i>Examples</i>
1	Molecular chaperones and chaperonins, Hsp70(DnaK), Hsp40(DnaJ), GrpE, GroEL, GroES, Hsp90, Hsp104, sHsp
2	Regulators of stress (heat shock) genes: sigma factors (e.g., σ^{32} , σ^E , σ^B), positive regulators or activators (e.g., HSF), negative regulators (e.g., HrcA, HspR)
3	Proteins that constitute the intercellular connective material of multicellular structures formed as a response to stress
4	Transport proteins: e.g., archaeal TrkA
5	Proteases: Clp family, Lon/La, HtrA, proteasome components
6	Enzymes and cofactors involved in the synthesis of other stress proteins (listed in the preceding groups 1–5) and other molecules, such as thermoprotectants (trehalose; di- <i>myo</i> -inositol phosphate; cyclic diphosphoglycerate)
7	All proteins, enzymes, cofactors, and gene-regulatory factors involved in sporulation

^aHSF, heat shock factor; Hsp, heat shock protein. Modified from Macario, A. J. L. and Conway de Macario, E. (2000), Stress and molecular chaperones in disease, *International Journal of Clinical and Laboratory Research* **30**, 49–66, with permission from the copyright owner.

A balanced set of proteins, each in its native functional conformation and at its physiological concentration, is key to cell survival. Heat shock and other stresses alter the cellular proteins. These tend to lose their functional shapes; they denature. Protein

denaturation is a central harmful factor for the cell during stress. Thus, the mechanisms for preventing protein denaturation, and for reversing it when it is still reversible, are major survival tools for the cell. Among these tools, the stress proteins, including the

molecular chaperones, are prominent. Therefore, it is not surprising that the levels of these proteins are maintained or augmented in response to stress, whereas many other proteins decrease or become undetectable.

There are several ways for the cell to readjust the concentration of stress proteins when it has been affected by a stressor. One way is to upregulate the genes that produce the useful proteins, and another is to diminish the degradation of the latter. All cellular proteins undergo a cycle of birth, maturation, function, and degradation (equivalent to death). By postponing the degradation of stress proteins, the cell enhances the chances of its survival in the face of stress.

Another mechanism for increasing the levels of useful proteins during stress is to facilitate their synthesis from their mRNAs. This may be achieved in various ways. One way is to increase the life span of the mRNAs that encode stress proteins (mRNAs are usually short-lived). If the stability of the mRNA increases, the same mRNA molecules can reach the ribosome and be translated more than just once, avoiding delays due to the unavailability of newly synthesized mRNA molecules. Furthermore, translation itself may be enhanced, so mRNA is decoded more rapidly and more mRNA molecules are translated per ribosome.

In summary, gene activation, involving both an initiation of the transcription of genes inactive in the absence of stress and an increase in the rate of transcription of genes already active before stress, is not the only mechanism that brings about the increase of heat shock proteins. Prolongation of the life span of mRNAs and of the stress proteins themselves also plays a role in building up the cellular defenses against the ravages of stress.

Clearing the Rubble

The rescue of proteins partially damaged by stress is a key mechanism for cell survival, as explained earlier. However, many molecules may be damaged irreversibly and cannot be rescued. These denatured proteins tend to aggregate and to form relatively large precipitates. These may be toxic, and they may also obstruct other cellular components. This is the reason why the cell has a battery of enzymes that digest seriously damaged, useless proteins. These enzymes are called proteases, and some of them are considered stress proteins because they handle the key task of removing stress-damaged proteins to clear the way for the trafficking and functioning of other, viable molecules essential for survival. Furthermore, the degradation products generated by the action of proteases on their substrates serve as building materials

for the synthesis of other molecules necessary for mounting a stress response.

In summary, a series of mechanisms are set in motion by stress. The main objective of the cell is to maintain a range of functionally competent proteins at the concentrations required for survival. The production of nonessential proteins is stopped, the synthesis of required proteins is started or increased, and damaged or useless molecules are eliminated.

Milestones

The heat shock response has been the focus of research efforts for at least 4 decades. If we bear in mind the current division of living organisms into three major lines of descent or phylogenetic domains – Eucarya (eukaryotes), Bacteria (eubacteria), and Archaea – it is possible to see that the chronology of specific contributions proceeded from the Eucarya to Bacteria, and only considerably later reached the Archaea. Currently, organisms from the three domains are used as experimental models, but those from the Archaea are still in the minority.

In 1962 it was reported that a temperature elevation induced distinct puffing patterns in the chromosomes of the salivary gland of the fly *Drosophila buschii*. These puffs were transient but reproducible, and they always affected the same chromosomal regions (bands). Puffing occurred when the fly larvae were moved out of their optimal temperature for growth (i.e., 25°C) and into a temperature of 30°C and kept there for 30 min. The same results were obtained with another fly species, *D. melanogaster*, which is much used for research.

At the time, these results were interpreted to represent gene activation induced by a change in the environmental temperature, which as we know now is what constitutes a heat shock response. Similar effects were obtained with 2–4 dinitrophenol and salicylate, two compounds that today we would call chemical cell stressors (see [Table 1](#)). These observations established that heat shock activates some genes, always the same ones in a defined location in the chromosome, and that similar gene activation can be induced by agents other than heat. The conclusion was that heat shock and other stressors induce gene transcription with production of new mRNAs.

That these mRNAs were actually the precursors of new proteins, whose synthesis was set in motion by heat shock, was demonstrated in 1974, once again in the salivary glands of *D. melanogaster* and also in other tissues. The results showed a correlation between the typical heat shock-induced puffing patterns, mRNA synthesis, and the appearance of several new proteins in the tissues of the fly.

The new proteins were visualized by one-dimensional gel electrophoresis. At this time, it became clear that heat shock induces the synthesis of proteins that are not present at detectable levels under normal physiological conditions. It also became clear that gene activation and protein synthesis are not restricted to a single tissue or organ (e.g., the salivary glands) but are more generalized.

The main characteristics of the heat shock response were in fact defined by these early studies with the fruit fly. These characteristics are the transient activation of a restricted, defined group of genes by an abrupt temperature upshift with the synthesis of a set of mRNAs and proteins, occurring in all types of tissues and cells. Similar effects can be produced by other agents or stressors.

Thus, the field of heat shock or stress response was opened for further investigations at the biochemical, molecular, and genetic levels by these pioneering studies on a eukaryotic organism, an insect. Several years passed before knowledge of the stress response of prokaryotes began to emerge. The prokaryotic organism most used for research in the field has been the bacterium *Escherichia coli*. Investigations with this model system were reported from the beginning of the 1980s. A number of stress genes and proteins were identified, and many of them have been characterized thoroughly, including detailed structure–function determinations by very sophisticated techniques such as crystallography, electron microscopy, and computer-assisted imaging analysis. Also, considerable progress has been made toward the elucidation of the mechanisms that regulate stress genes, addressing the following questions. How are these genes instructed to start transcription, and to stop it, according to the cell's needs? How does the cell sense the temperature or other environmental changes (stressors) around it and transmit from its surface the orders for activity or inactivity to the genes?

Over the years, since the time of the initial studies with *E. coli*, other bacterial species have been added to the list of model systems used for studying the stress response. These species include representatives from many of the known bacterial groups, such as gram positives and gram negatives; mesophiles, thermophiles, and hyperthermophiles; and pathogens and nonpathogens (see Table 2). In parallel, eukaryotes other than *Drosophila* have also been used to study the stress response and the stress genes and proteins. Examples are the single-celled fungus *Saccharomyces cerevisiae*, known also as baker's yeast, and several other organisms representing the major groups of eukaryotes (see Table 2).

These studies with members of the domains Bacteria and Eucarya proliferated in the 1980s and 1990s,

not only because they helped to advance our understanding of important cellular functions and mechanisms under physiological conditions and during and after stress (see Table 3) but also because some of the organisms are pathogens for humans and for animals valued by humans. Hence, it was pertinent to examine the role of the stress response in health and disease. In addition, other model organisms have existing or potential industrial applications and are, therefore, of commercial interest.

While the research activities mentioned so far were ongoing, the identification and characterization of the third evolutionary domain, the Archaea, proceeded steadily since the late 1970s. However, it was only in 1991 that a heat shock gene, *hsp70(dnaK)*, was first identified by cloning and sequencing in an organism of the domain Archaea, the mesophilic methanogen *Methanosarcina mazei*. In the same year, the chaperonin complex called the thermosome was visualized in another member of the Archaea, a hyperthermophile. Both findings opened the doors to exciting new territories for exploration. Research with archaeal organisms has progressed, driven by the same motivations that propelled research on organisms of the other two domains. Archaeal species are attractive to the scientist because they possess a series of characteristics that are unique to them and that were unknown before these organisms were discovered. Some of these characteristics are extreme compared with what is known for the more familiar, and better studied, eukaryotes and bacteria. In addition, there are archaeal species and molecules from Archaea that have potential in the biotechnology industry, including the development of clean and efficient means for waste treatment with the generation of useful or harmless by-products. In a variety of these applications, the stress response and the stress genes and proteins play a key role. Consequently, their manipulation is a promising way to improve extant technology and develop better alternatives than those currently available.

Currently, research continues on organisms of the three domains. The goals are to discover other stress genes, to characterize more thoroughly the mechanisms that bring about the stress response and the mechanisms involved in the regulation of the stress genes, and to design strategies and procedures to manipulate these genes and their protein products to prevent or cure lesions caused by stressors in plants and animals (see Table 4).

New Developments to 2005

Recent developments pertain to (1) stress genes in the Archaea; (2) chaperones and disease; (3) genomics

and automated methods by which sequenced genomes can be studied and by which large amounts of information about expressed genes (including stress genes), their regulatory networks, and the interactions of their protein products in the cell can be obtained and analyzed; (4) high-throughput technology to measure gene transcription at the level of the entire genome and to determine the full range of proteins present in a cell or tissue at any given time; and (5) techniques to measure the presence of stressors in environmental samples such as wastewaters, soil, food, paints, and construction materials.

Stress genes in Archaea proved to have a distribution unexpected from what was known from bacteria and eukaryotes. For example, the molecular chaperone machine Hsp70(DnaK)-Hsp40(DnaJ)-nucleotide exchange factor, which is present in all bacteria and eukaryotes tested, was not found in several archaeal species. Furthermore, the chaperonin system group I, which had been considered to be exclusive to bacteria and bacterial-related eukaryotic organelles, was discovered in some species of Archaea. Two new chaperonin group II subunits were discovered in an archaeal species. This finding elevated the total number of subunits in the Archaea to five (the maximum known before was three), approaching a total number of eight, which is characteristic of eukaryotes.

The concept that defective molecular chaperones can contribute to pathogenesis not only as auxiliary factors but also as primary etiological agents was developed, and a number of diseases were identified as examples of this newly recognized pathological entity, chaperonopathy. The negative impact of chaperonopathies on stress responses, on survival in the face of protracted stress, and on recovery after stress remains to be evaluated. This will most likely be an area of intensive research in the first decades of the twenty-first century.

High-throughput automatic technologies with relatively low turnaround times were developed for sequencing genomes, which increased considerably the information available in databases. In parallel, similar, automated methods were standardized for mining genome sequences for stress genes and their regulatory genes and DNA elements, and to obtain clues on the multiple interactions of stress proteins and molecular chaperones with other proteins and cell components (systems biology and cellular neural networks) and with foreign compounds. The results of these genomic, computer-assisted searches will be instrumental in guiding laboratory experiments, focusing on a few selected genes and proteins for elucidating their functions, regulation, and interaction *in vitro* and *in vivo*.

Also, microarray technology has radically changed the analysis of gene expression profiles at the whole genome level. Microarray analysis provides enormous quantities of data for the comparison of different tissues from healthy and sick organisms or of cells under basal conditions (constitutive gene expression) and cells affected by stressors (stress-induced gene expression). This approach has shown that the number of genes induced by stressors is relatively large considering the total number of genes in any given genome, encompassing not only the genes that encode the classic Hsps and chaperones but also many others whose products have a variety of other roles. Thus, it has become clear beyond doubt that there are many stress genes and proteins but that only a fraction of them belong to the Hsp and chaperone families. Likewise, it is clear that not all molecular-chaperone genes are induced by stressors and, therefore, not all chaperones can be classified as Hsps.

The increasing concerns about the contamination of the environment and food with toxic compounds (i.e., cell stressors) have given impetus to the development of novel assays for measuring the potency of these compounds for causing cell stress.

Perspectives

In addition to the stressors discussed in the preceding sections, psychological stress due to professional, financial, social, and family pressures and obligations also cause cell stress via the hormonal and biochemical imbalances inherent in this unpleasant condition. The list of stressors becomes longer and longer as new compounds are isolated from natural sources or are synthesized in the laboratory and released into the environment (added to water, foods, fuels, paints, building materials, clothing, etc.). The hole in the ozone layer of the atmosphere and the greenhouse effect add another dose of stressors (radiation and overheating) on top of all the others already alluded to. No doubt plants and animals – including humans – will have to face stress as the scourge of our time. The more that is known about the heat shock (stress) response, the easier it will be to find ways to counteract the effects of stressors. Research in this field promises to be vigorous in the near future and will most likely attract many scientists. The topic should also be of interest to virtually everybody, regardless of educational level or profession. The public at large ought to know, at least in a general way, what can cause cell stress and how to avoid stressors and abate their deleterious effects. For these purposes, the usual means of using good hygiene and common sense should be helpful, along with efforts to promote legislation aimed at reducing the number and quantities of stressors.

The malfunctioning of antistress mechanisms has recently come into prominence as a new field of medicine. For example, abnormalities in chaperone molecules have an impact on the chaperoning systems throughout the organism and have been implicated in various genetic and acquired pathological disorders. This new medical field will no doubt attract the attention of more and more scientists and practitioners to improve diagnostics, therapeutics, and prevention.

See Also the Following Articles

Chaperone Proteins and Chaperonopathies; Heat Resistance; Heat Shock Genes, Human; Viral Virulence and Stress; Chaperonopathies; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles.

Further Reading

- Abdollahi, M., Ranjbar, A., Shadnia, S., et al. (2004). Pesticides and oxidative stress: a review. *Medical Science Monitor* 10, RA141–RA147.
- Bomholt, S. F., Harbuz, M. S., Blackburn-Munro, G., et al. (2004). Involvement and role of the hypothalamo-pituitary-adrenal (HPA) stress axis in animal models of chronic pain and inflammation. *Stress* 7, 1–14.
- Conway de Macario, E., Maeder, D. L. and Macario, A. J. L. (2003). Breaking the mould: Archaea with all four chaperoning systems. *Biochemical and Biophysical Research Communications* 301, 811–812.
- Gao, H., Wang, Y., Liu, X., et al. (2004). Global transcriptome analysis of the heat shock response of *Shewanella oneidensis*. *Journal of Bacteriology* 186, 7796–7803.
- Gash, A. P., Spellman, P. T., Kao, C. M., et al. (2000). Genomic expression programs in the response of yeast cells to environmental changes. *Molecular Biology of the Cell* 11, 4241–4257.
- Gibney, E., Gault, J. and Williams, J. (2004). A novel immunocytochemistry technique to measure Hsp70 induction in L929 cells exposed to cadmium chloride. *Biomarkers* 9, 353–363.
- Helman, J. D., Wu, M. F., Kobel, P. A., et al. (2001). Global transcriptional response of *Bacillus subtilis* to heat shock. *Journal of Bacteriology* 183, 7318–7328.
- Hickey, A. J., Conway de Macario, E. and Macario, A. J. L. (2002). Transcription in the Archaea: basal factors, regulation, and stress-gene expression. *Critical Reviews in Biochemistry and Molecular Biology* 37, 537–599.
- Holden, J. F., Adams, M. W. W. and Baross, J. A. (2000). Heat-shock response in hyperthermophilic microorganisms. In: Bell, C. R., Brylinsky, M. & Johnson Green, P. (eds.) *Microbial biosystems: new frontiers* (vol. 2), pp. 663–670. Halifax, Canada: Atlantic Canada Society for Microbial Ecology and Acadia University Press.
- Krone, P. H., Blechinger, S. R., Evans, T. G., et al. (2005). Use of fish liver PLHC-1 cells and zebrafish embryos in cytotoxicity assays. *Methods* 35, 176–187.
- Macario, A. J. L. and Conway de Macario, E. (2000). Stress and molecular chaperones in disease. *International Journal of Clinical and Laboratory Research* 30, 49–66.
- Macario, A. J. L. and Conway de Macario, E. (2004). The pathology of anti-stress mechanisms: a new frontier. *Stress* 7, 243–249.
- Macario, A. J. L., Lange, M., Ahring, B. K., et al. (1999). Stress genes and proteins in Archaea. *Microbiology and Molecular Biology Reviews* 63, 923–967.
- Macario, A. J. L., Malz, M. and Conway de Macario, E. (2004). Evolution of assisted protein folding: the distribution of the main chaperoning systems within the phylogenetic domain Archaea. *Frontiers in Bioscience* 9, 1318–1332.
- Maeder, D. L., Macario, A. J. L. and Conway de Macario, E. (2005). Novel chaperonins in a prokaryote. *Journal of Molecular Evolution* 60, 409–416.
- Schoffl, F., Prandl, R. and Reindl, A. (1998). Regulation of the heat-shock response. *Plant Physiology* 117, 1135–1141.
- Stintzi, A. (2003). Gene expression profile of *Campylobacter jejuni* in response to growth temperature variation. *Journal of Bacteriology* 185, 2009–2016.
- Thakurta, D. G., Palomar, L., Stormo, G. D., et al. (2002). Identification of a novel cis-regulatory element involved in the heat shock response in *Caenorhabditis elegans* using microarray gene expression and computational methods. *Genome Research* 12, 701–712.
- Traustadottir, T., Bosch, P. R. and Matt, K. S. (2005). The HPA axis response to stress in women: effects of aging and fitness. *Psychoneuroendocrinology* 30, 392–402.
- Voellmy, R. (2005). Dominant-positive and dominant-negative heat shock factors. *Methods* 35, 199–207.
- Weber, H., Polen, T., Heuveling, J., et al. (2005). Genome-wide analysis of the general stress response network in *Escherichia coli*: sigma S-dependent genes, promoters, and sigma factor selectivity. *Journal of Bacteriology* 187, 1591–1603.
- Zhao, R., Davey, M., Hsu, Y.-C., et al. (2005). Navigating the chaperone network: an integrative map of physical and genetic interactions mediated by the Hsp90 chaperone. *Cell* 120, 715–727.

Relevant Websites

<http://www.bioscience.org/current/special/macario.htm>

Hemostasis and Stress

R von Känel

University Hospital Berne, Berne, Switzerland

© 2007 Elsevier Inc. All rights reserved.

Historical Perspective

Hemostasis Pathways

Hemostasis Factors and Cardiovascular Disease

Acute Stress

Chronic Stress

Conclusions and Future Directions

Glossary

<i>Blood coagulation</i>	Physiological process as part of hemostasis in which liquid blood is changed into a blood clot of platelet aggregates and fibrin.
<i>Endogenous anticoagulants</i>	Proteins produced by the body to regulate thrombin activity during blood coagulation.
<i>Fibrinolysis</i>	Physiological process of hemostasis in which the fibrin clot is broken down.
<i>Hemostasis</i>	Physiological process in which bleeding is halted in a damaged blood vessel requiring the combined regulated activity of vascular, platelet, and plasma factors.
<i>Hypercoagulability Platelets</i>	A shift in the hemostatic balance favoring formation and maintenance of a clot. Small anuclear blood cells that adhere to sites of blood vessel injury, where they become activated and aggregate with each other.
<i>Thrombosis</i>	Pathological process in which a blood clot forms in the lumen of a blood vessel.

Historical Perspective

The possible influence of stress and stress hormones on blood coagulation and fibrinolysis has attracted researchers for more than a century. As early as 1903, Vosburgh and Richards observed that epinephrine added to the surface of the pancreas and injected into the abdominal cavity of dogs hastened the coagulation time of whole blood, in some cases to as much as four-fifths the coagulation time of control animals. In 1911, von den Velden reported a similar decrease in coagulation time in humans after subcutaneous injection of epinephrine. Only a few years later, Walter Cannon and his co-workers demonstrated in a series

of carefully conducted experiments that stimulation of the splanchnic nerve, pain, fear, and enrage ment all shortened blood clotting in the cat. They noted that rapid coagulation did not occur if adrenals had been removed or were exhausted because of previous excitement, when cats had been caged near dogs. These observations prompted Cannon to state that “rapid coagulation may reasonably be considered as an instance of adaptive reaction serviceable to the organism in the injury which may follow the struggle that fear or rage may occasion.”

Patients with hemophilia A lack substantial activity of clotting factor VIII (FVIII:C) such that severe bleeding occurs with minor traumas or even spontaneously. In search of treatment options for this bleeding disorder, systematic investigations in the 1960s and 1970s found that parenteral administration of epinephrine evoked a dose-dependent increase in plasma levels of antihemophilic FVIII:C with concomitant shortening of the clotting time of whole blood. Electric brain stimulation evoked an increase in plasma FVIII:C and a corresponding decrease in the clotting time in dogs. Plasma FVIII:C also rose after air inflation in the skull during a diagnostic procedure in humans. In terms of fibrinolysis, anecdotal reports from the 1940s observed fibrinolytic activation by epinephrine and by a variety of acute naturalistic stressors eliciting a fear response, e.g., planned surgery, air raid warning, and exams.

These earlier studies suggested what has become consensus today, namely, that in a healthy organism, acute stress concomitantly activates the coagulation and fibrinolysis part of hemostasis to result in net hypercoagulability. Clearly, this supports Cannon, who argued that thickening of blood in response to an acute and potentially life-threatening stressor is actually physiological and provided our ancestors with an evolutionary benefit. Indeed, we can readily imagine that hunters who were injured and women who gave birth increased their odds of surviving if hypercoagulability constrained excessive bleeding. Equally exciting, some of this early research can be viewed as first evidence for a regulatory mechanism of hemostasis by the brain involving the sympathomedullary axis and catecholamines.

Hemostasis Pathways

The scheme in [Figure 1](#) simplifies complex and closely intertwined hemostasis pathways. Activation of hemostasis results in a sequential interaction between

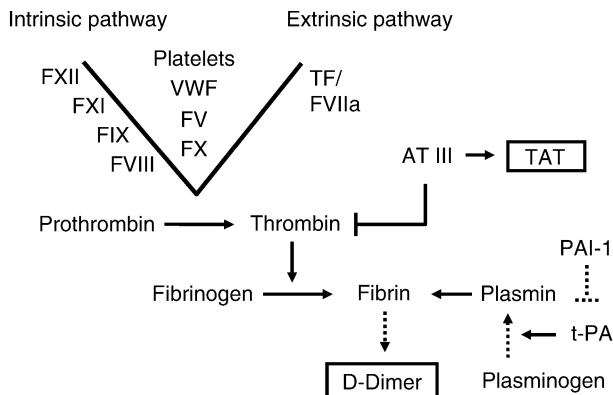


Figure 1 Hemostasis pathways. The figure shows coagulation steps in solid lines and fibrinolysis steps in dashed lines with the symbol $\rightarrow$ meaning activation of a step and the symbol $\mid$ meaning inhibition of a step. Hypercoagulability markers thrombin-antithrombin (AT) III complex (TAT) and clotting factor V (FV) and X (FX) are involved in both coagulation pathways. FXII, FXI, FIX, and FVIII are components of the intrinsic pathway, and tissue factor (TF) and FVII belong to the extrinsic pathway of blood coagulation. PAI-1, plasminogen activator inhibitor-1; t-PA, tissue-type plasminogen activator. Cf. text for details of activation steps.

enzymes and inhibitors of the coagulation and fibrinolysis cascades, of platelets, and of the von Willebrand factor (VWF). Two pathways of blood coagulation are commonly distinguished. Whereas the intrinsic pathway is initiated by FXII, the extrinsic pathway is triggered when tissue factor becomes exposed to the bloodstream at sites of tissue damage, e.g., on material of a ruptured atherosclerotic plaque. Tissue factor instantly binds to circulating FVII. Both pathways result in activation of several clotting factors in a sort of cascade, and eventually converge in a common activator complex. This complex converts prothrombin to thrombin, which, in turn, converts fibrinogen to fibrin. Together with platelet aggregates, fibrin forms a blood clot that, e.g., may critically occlude a coronary artery to result in myocardial infarction. Circulating VWF mediates adherence of platelets to sites of blood vessel injury where platelets are activated and aggregate with each other through fibrinogen bridges. Negatively charged phospholipids on activated platelets catalyze activation of circulating clotting factors.

Termination of clot formation involves several anticoagulant steps, e.g., binding of antithrombin III to thrombin. This step neutralizes thrombin, a potent activator of several clotting factors and platelets, in a thrombin/antithrombin III (TAT) complex. Fibrinolysis is initiated by tissue-type plasminogen activator (t-PA) converting plasminogen into fibrin-cleaving plasmin. The breakdown of cross-linked fibrin chains yields soluble fibrin fragments such as D-dimer. The antifibrinolytic enzyme plasminogen activator

inhibitor-1 (PAI-1) is the main inhibitor of t-PA, neutralizing t-PA in a t-PA/PAI-1-complex.

TAT and D-dimer are viewed as coagulation activation markers because they combine several coagulation steps leading to thrombin and fibrin formation, respectively. In addition, TAT and D-dimer are hypercoagulability markers indicating exaggerated coagulation activity. It is important to note that, unlike individual clotting and fibrinolysis factors, D-dimer indicates activation of the entire hemostatic system, i.e., fibrin formation by the coagulation system and fibrin degradation by the fibrinolytic system.

Hemostasis Factors and Cardiovascular Disease

Since the mid-1980s, numerous large-scale and prospectively designed epidemiological studies have been published suggesting that numerous components of the hemostatic system may predict cardiovascular disease. Most of this research has investigated the predictive value of hemostasis factor for coronary artery disease (CAD) and acute myocardial infarction, while a minor proportion of previous work has addressed the role of hemostasis factors in cerebrovascular and peripheral arterial disease. Meta-analyses have shown CAD risk factor status for fibrinogen, D-dimer, and VWF, whereas FVII:C, FVIII:C, TAT, PAI-1, t-PA, and platelet activity were risk factors of CAD and myocardial infarction across a series of individual investigations.

Remarkably, hemostasis factors not only predicted recurrent cardiac events in patients with cardiovascular diseases but also prospectively predicted the cardiovascular disease risk in individuals who were in apparently good health at study entry. Moreover, many studies showed that the prospective relationship between hemostasis factors and coronary events was independent of sociodemographic factors, lifestyle factors, and the established cardiovascular risk factors of smoking, diabetes, hypertension, and obesity, all of which may affect hemostasis. Altogether, this abundant epidemiological research challenges the view of some researchers that hemostasis factors are mere risk markers of cardiovascular disease and do not actively contribute to atherosclerosis.

It is assumed that a procoagulant milieu, as reflected by increased activity of clotting factors, coagulation activation markers, platelets, and the VWF on the one hand and impaired fibrinolysis on the other, will gradually contribute to atherosclerosis progression over many decades by promoting fibrin deposits in the atherosclerotic vessel wall and inflammation. In addition, upon rupture of a vulnerable atherosclerotic plaque, a procoagulant milieu may accelerate

and enhance intravascular thrombus growth and critical occlusion of a coronary artery. Chronic low-grade systemic activation of coagulation processes with chronic stress and repetitive hypercoagulability in response to daily hassles, for example, thus provide one psychobiological pathway through which stress might accelerate atherosclerosis progression. Moreover, the stress procoagulant response is exaggerated in atherosclerotic vessels because the endothelium has become dysfunctional and partially deprived of its anticoagulant properties. Therefore, an exaggerated stress procoagulant response could enhance thrombus formation at times of plaque rupture triggered by acute stress-induced blood pressure peaks.

Acute Stress

Summary of Findings

Table 1 summarizes findings from well-conducted and standardized laboratory studies on acute stress effects on hemostasis factors in apparently healthy subjects reaching beyond individual findings. Different types of speech tasks, mental arithmetic, the mirror star tracing task, and the Stroop color-word conflict test were regularly applied as mental stressors in these studies. This research suggests that acute stress elicits a significant increase both in procoagulant factors (i.e., FXII:C, FVII:C, FVIII:C, fibrinogen, VWF, platelet activity) and in profibrinolytic t-PA. However, the finding of an increase in coagulation activation markers TAT and D-dimer suggests that the concomitant activation of virtually all hemostasis components with acute stress results in net hypercoagulability compatible with the evolution paradigm of Cannon. Unlike changes in blood pressure and cortisol, changes in the coagulation system do not adapt in response to the same stressor (i.e., combined speech task and mental arithmetic) applied three times on healthy subjects within an interval of

1 week. Apparently, evolution empowered humans to mount the same hypercoagulability any time we might potentially be injured independently of whether or not we perceive the situation as threatening.

Modulators of the Procoagulant Stress Response

Certain diseases, psychological states, and life circumstances have been shown to exaggerate the acute stress procoagulant response. In particular, it has been shown that D-dimer formation is increased and fibrinolytic activation is attenuated in acutely stressed patients with cardiovascular disease. Also, recovery of platelet activity after stress was delayed in patients with CAD as compared to those without. Anxiety, depression, and poor coping strategies (i.e., low problem solving and low personal mastery) have been shown to be associated with greater D-dimer formation and antifibrinolytic activity with acute stress in elderly subjects who were caregivers for a spouse with Alzheimer's disease. These caregivers also experienced greater acute procoagulability to stress when they experienced relatively more negative life events in the preceding month and if their spouse showed relatively more severe dementia. In accordance with the stress reactivity paradigm, these findings suggest that the stress procoagulant response, because of being hyperactive, might embed potential cardiovascular harm in certain individuals with certain diseases and at certain times of their life.

However, hemostatic effects of negative affect are not straightforward and may differ between elderly individuals, many of whom have cardiovascular disease, and healthy individuals. In the latter, an inverse relationship was found between levels of different domains of negative affect (e.g., anxiety and depression) and the magnitude of coagulation changes with acute stress. We may assume that endothelial dysfunction in the elderly might interact with negative affect to promote acute hypercoagulability. In contrast, greater hypercoagulability with less negative affect might still reflect normal physiology in healthy individuals preparing the organism for fight or flight.

Table 1 Changes in hemostasis factors with stress

Hemostasis factor	Acute stress	Chronic stress
FVII:C	↑	↑
FXII:C	↑	
FVIII:C	↑	
Fibrinogen	↑	↑
Thrombin/antithrombin III complex	↑	—
D-dimer	↑	↑
von Willebrand factor	↑	↑ / —
Platelet activity markers	↑	
Plasminogen activator inhibitor-1	—	
Tissue-type plasminogen activator activity	↑	↓

↑, increases with stress; ↓, decreases with stress; —, no response to stress.

Physiological Mechanisms

Molecular studies in *in vitro* and *in vivo* studies in animals and humans convincingly demonstrate that the sympathetic nervous system (SNS) regulates hemostatic activity. Much of this knowledge derives from human studies in which adrenergic compounds, particularly the stress hormones epinephrine and norepinephrine, had been infused together, with or without previous blockade of adrenergic receptors by various drugs. **Table 2** summarizes adrenergic receptors and mechanisms assumed to be involved

Table 2 Adrenergic receptors and hemostasis regulation

<i>Hemostasis factor</i>	<i>Adrenergic receptors (ARs)</i>	<i>Mechanisms</i>
FVIII	Endothelial β 2-AR	Release from endothelial cells and the liver
VWF	Endothelial β 2-AR	Release from endothelial cells
t-PA	Endothelial β 2-AR	Release from endothelial cells
	(β 1- and) β 2-adrenergic stimulation	Increases t-PA clearance
	α 1-adrenergic stimulation	Reduces t-PA clearance
Platelets	Platelet α 2-AR	Platelet activation
	Platelet β 2-AR	Inhibition of platelet activation
PAI-1	β 1- and/or β 2-adrenergic stimulation	PAI-1 mRNA expression

FVIII, clotting factor VIII, VWF, von Willebrand factor, t-PA, tissue-type plasminogen activator; PAI-1, plasminogen activator inhibitor-1.

in hemostasis regulation. Via stimulation of vascular endothelial β 2-adrenergic receptors, catecholamine surge releases preformed FVIII, VWF, and t-PA from endothelial storage pools into the circulation. This phenomenon takes place within minutes and shows dose dependency. Catecholamines also stimulate release of FVIII from the liver and affect hepatic clearance of t-PA. In addition, platelets are activated by stimulation of their α 2-adrenergic receptors, with platelet β 2-adrenergic receptors likely exerting an inhibitory effect on platelet activation by catecholamines. Greater sensitivity of the β 2-adrenergic receptor was associated with greater thrombin formation in response to acute stress.

Stress-induced hemoconcentration was not associated with acute stress-induced changes in coagulation factors and activities. This suggests that coagulation activation during stress is an active phenomenon and does not merely reflect a passive increase of factors due to stress-induced plasma volume contraction. To date, there is little evidence for a significant effect of the hypothalamic-pituitary adrenal (HPA) axis and cortisol activity on the acute stress procoagulant response. Conceivably, evolution wanted the blood to clot instantly after injury as mediated by the SNS. In contrast, the HPA axis peaks its activity only 15 to 30 min after stress and, therefore, is a system likely too slow to critically limit bleeding.

Chronic Stress

Summary of Findings

Table 1 shows that, as with acute stress, chronic stress leads to an increase in procoagulant factors, namely, FVII:C, fibrinogen, and D-dimer, while data for VWF are conflicting. Notably, effects of chronic stress on fibrinolytic activity are distinctly different from those observed with acute stress. Chronic stress impairs the fibrinolytic activity such that the balance between procoagulant factors and profibrinolytic factors is shifted toward hypercoagulability to an extent that

exceeds the physiological net hypercoagulable changes seen with acute stress in healthy individuals.

These findings derive from investigations on hemostasis factors in subjects with high job strain, with low socioeconomic status, and who are caregivers of a demented spouse. In these studies, job strain was mainly defined as a mismatch between job demands and decision latitude and an imbalance between effort spent and reward obtained, moderated by a personality trait of overcommitment to work. Although relationships between various job strain variables and elevated plasma levels of FVII:C, fibrinogen, and PAI-1 on the one hand and decreased t-PA activity on the other were independent, most of the association between low socioeconomic status and hypercoagulability became insignificant when statistics controlled for health habits.

The number of negative life events in the preceding month correlated positively with plasma D-dimer in Alzheimer caregivers. In another study, chronically stressed dementia caregivers had higher D-dimer levels than age- and gender-matched noncaregiving controls, even when controlling for cardiovascular diseases and risk factors, lifestyle, and medication. Increased D-dimer was also observed in elderly hypertensive subjects in the weeks after they had experienced catastrophic stress. A fairly consistent line of research suggests that a state of excessive exhaustion ensuing from long-term psychological strain is associated with elevated PAI-1. Therefore, an antifibrinolytic effect could help explain the observation that profound feelings of fatigue may precede an acute myocardial infarction.

Physiological Mechanisms

The pathophysiological underpinnings of the link between chronic stress and hemostatic changes are far from clear, though gene regulation, the SNS, and the HPA axis could all be involved. This seems particularly true in terms of PAI-1. In accordance with the observed decrease in PAI-1 with chronic stress,

restraint stress and injection of nonselective β -adrenergic substances in animals (i.e., epinephrine, isoprenaline) both resulted in upregulation of PAI-1 mRNA (Table 2) with a concomitant increase of PAI-1 activity in tissue homogenates. Moreover, homozygote carriers of the 5G allele of the PAI-1 4G/5G gene polymorphism who were exhausted had higher PAI-1 levels than homozygote 5G allele carriers who were not exhausted. In contrast, homozygote 4G carriers and heterozygote 4G/5G carriers did not differ in their PAI-1 levels in relation to their state of exhaustion.

When controlling for a whole range of possible confounders, greater overnight catecholamine secretion predicted higher morning levels of fibrinogen and D-dimer, while overnight cortisol excretion was not independently related to morning PAI-1 levels in a working population. However, research on the role of the HPA axis and endogenous cortisol activity on hemostasis in situations of chronic stress is too scarce to make a definite statement. Notably, patients with Cushing's disease, which is characterized by high production of endogenous cortisol, have an increased risk of thrombosis. In addition, subjects who received exogenous glucocorticoids either experimentally or as treatment manifested a hypercoagulable state.

Conclusions and Future Directions

More than a century of research on stress effects on hemostasis strongly suggests that, in response to acute stress, healthy individuals show a hypercoagulable state that is physiological and that has served the human species to enhance the odds of surviving fight-and-flight situations. However, this evolutionary benefit could be harmful today in an environment characterized by comparably longer lasting psychological stressors than those encountered in ancient times when humans either were killed or could escape. Daily hassles experienced in a rapidly changing society, years of marital stress, and job strain, as well as becoming a caregiver for a loved one, may elicit repetitive and sustained hypercoagulability that has lost its protective nature, thereby contributing to atherosclerosis and eventually myocardial infarction.

There is good evidence to assume that, of the stress-mediating systems, it is the sympathomedullary system that plays a major role in hemostasis regulation. However, data on the role of the HPA axis are rare, and the effect of the parasympathetic nervous system on hemostasis has not been explored. Although there is preliminary evidence that the brain transforms perceived stress into a physiological stress response resulting in rapid hemostatic activation further downstream in the peripheral blood and gene

expression of hemostasis molecules, much research needs to be done to better elucidate these pathways. Also, we lack prospective studies investigating whether increased stress procoagulant reactivity has any impact on hard cardiovascular endpoints later in life. Often in the same large cohort, researchers reported that both hemostasis factors and stress predicted cardiovascular endpoints. However, none of these studies has investigated an additive or even multiplicative effect of stress and hemostasis on endpoints. If stress interacts with hypercoagulability in predicting cardiovascular disease, carefully conducted intervention studies on effects of stress management in its broadest sense on hemostatic function seem warranted. We could learn from these studies whether individuals with hypercoagulability exceeding its physiological magnitude could profit from, e.g., cognitive-behavioral stress management for better cardiovascular health.

See Also the Following Articles

Acute Stress Response: Experimental; Atherosclerosis; Beta-Adrenergic Blockers; Cardiovascular System and Stress; Caregivers, Stress and; Catecholamines; Corticosteroids and Stress; Fibrinogen and Clotting Factors; Genetic Factors and Stress; Heart Disease/Attack; Sympathetic Nervous System; Cardiovascular Disease, Stress and.

Further Reading

- Brydon, L., Magid, K. and Steptoe, A. (2006). Platelets, coronary heart disease, and stress. *Brain, Behavior, and Immunity* 20, 113–119.
- Cacioppo, J. T., Berntson, G. G., Malarkey, W. B., et al. (1998). Autonomic, neuroendocrine, and immune responses to psychological stress: the reactivity hypothesis. *Annals of the New York Academy of Science* 840, 664–673.
- Camacho, A. and Dimsdale, J. E. (2000). Platelets and psychiatry: lessons learned from old and new studies. *Psychosomatic Medicine* 62, 326–336.
- Cannon, W. B. and Mendenhall, W. L. (1914). Factors affecting the coagulation time of blood IV. The hastening of coagulation in pain and emotional excitement. *American Journal of Physiology* 34, 251–261.
- Cannon, W. B. (1953). *Bodily changes in pain, hunger, fear, and rage. An account of recent researches into the function of emotional excitement* (2nd edn.). Boston, MA: Charles T. Branford Company.
- Folsom, A. R. (2001). Hemostatic risk factors for atherothrombotic disease: an epidemiologic view. *Thrombosis and Haemostasis* 86, 366–373.
- Lee, K. W. and Lip, G. Y. (2003). Effects of lifestyle on hemostasis, fibrinolysis, and platelet reactivity: a systematic review. *Archives of Internal Medicine* 163, 2368–2392.

- Loscalzo, J. and Schafer, A. I. (eds.) (2003). *Thrombosis and haemorrhage* (3rd edn.) Philadelphia, PA: Lippincott Williams & Wilkins.
- Markovitz, J. H. and Matthews, K. A. (1991). Platelets and coronary heart disease: potential psychophysiologic mechanisms. *Psychosomatic Medicine* **53**, 643–668.
- Preckel, D. and von Känel, R. (2004). Regulation of hemostasis by the sympathetic nervous system – any contribution to coronary artery disease? *Heart Drug* **4**, 123–130.
- von Känel, R. and Dimsdale, J. E. (2000). Effects of sympathetic activation by adrenergic infusions on hemostasis in vivo. *European Journal of Haematology* **65**, 357–369.
- von Känel, R. and Dimsdale, J. E. (2003). Hemostatic alterations in patients with obstructive sleep apnea and the implications for cardiovascular disease. *Chest* **124**, 1956–1967.
- von Känel, R. and Dimsdale, J. E. (2003). Fibrin D-dimer: a marker of psychosocial distress and its implications for research in stress-related coronary artery disease. *Clinical Cardiology* **26**, 164–168.
- von Känel, R., Mills, P. J., Fainman, C. and Dimsdale, J. E. (2001). Effects of psychological stress and psychiatric disorders on blood coagulation and fibrinolysis: a bio-behavioral pathway to coronary artery disease? *Psychosomatic Medicine* **63**, 531–544.
- von Känel, R., Preckel, D., Zraggen, L., et al. (2004). The effect of natural habituation on coagulation responses to acute mental stress and recovery in men. *Thrombosis and Haemostasis* **92**, 1327–1335.
- von Känel, R., Kudielka, B. M., Preckel, D., et al. (2005). Opposite effect of negative and positive affect on stress procoagulant reactivity. *Physiology & Behavior* **86**, 61–68.

Herpes Simplex See: Herpesviruses.

Herpesviruses

D A Padgett, M T Bailey and J F Sheridan
The Ohio State University, Columbus, OH, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D A Padgett and J F Sheridan, volume 2, pp 357–364, © 2000, Elsevier Inc.

Introduction
Virology
Immunobiology and Recurrent Disease
Stress and Herpesvirus Reactivation
Conclusion

Glossary

<i>Capsid</i>	The protein coat or shell that contains the nucleic acid genome of the virus particle.
<i>Capsomeres</i>	The individual morphological units (or sides) on the surface of the capsid.
<i>Envelope</i>	The lipid-containing outer membrane that surrounds the entire virus particle. The envelope arises from budding of

Icosohedron

Latent infection

Nucleocapsid

mature herpesvirus particles through the nuclear membrane.

Geometric shape with 20 equal sides or capsomeres (similar to a soccer ball). The herpesvirus capsid is an icosohedron.

When the virus persists in a quiescent (occult or cryptic) form most of the time with very little viral replicative activity. Herpesviruses are characterized by intermittent flare-ups of clinical disease.

Includes the capsid and the enclosed nucleic acid.

Introduction

The Herpesviridae represent a large group of viruses that are pathogenic for many vertebrates. To date, there are more than 130 identified herpesviruses, nine of which are classified as human herpesviruses. These include herpes simplex virus type 1 (HSV-1), HSV-2, human cytomegalovirus (HCMV), varicella-zoster virus (VZV), Epstein–Barr virus (EBV), and human herpesviruses 6A, 6B, 7, and 8. The name herpesvirus is derived from the Greek word *herpein*, which means

to creep. Such a name is quite appropriate, as this family of viruses is characterized by chronic, latent, and recurrent infections that appear to creep or crawl across the skin.

The pattern of infection with a herpesvirus depends on the portal of entry, and the site of viral replication is restricted by the surface proteins expressed by each species of virus. The herpesvirus family members cause a range of diseases, including infectious mononucleosis, chicken pox, shingles, gingivostomatitis (i.e., fever blisters and cold sores), genital herpes, and Kaposi's sarcoma. The severity of disease is chiefly dependent upon the immune status of the host. For example, in newborns or in immunocompromised adults, infection by HSV can be severe, resulting in encephalitis and possibly death. In contrast, healthy adults resolve the primary infection, and the virus becomes latent. In fact, several members of the herpesvirus family cause lifelong infections in humans that are characterized by periods of quiescence interrupted by reactivation and expression of clinical symptoms.

The type of cell in which a herpesvirus can become latent is specific to each herpesvirus. For HSV, the development of latency means that virions reside quietly in neural tissues without causing disease. In contrast, EBV establishes latency in B lymphocytes. In each case, periodic reactivation of the latent virus is frequently accompanied by clinical symptoms, production of infectious virus, and the development of herpetic lesions.

Virology

Characteristic Structure

Herpesviruses are large DNA viruses that share similar architectural details; they are morphologically indistinguishable from one another. Each member has a double-stranded DNA genome arranged in what has been described as a toroid core; in such a core, the DNA is spooled around a cylinder of proteins. The herpesvirus's toroid core is encapsulated within an icosahedral capsid. The capsid is composed of 162 capsomeres and is itself packaged within a lipid envelope. This envelope is formed from the nuclear membrane of an infected cell from which progeny viruses bud. Within this nuclear-lipid envelope are virus-encoded glycoproteins, or spikes, about 8 nm long. An additional feature of the herpesvirus family is a tegument, which is an amorphous, asymmetrical structure located between the capsid and envelope. The entire virion measures 120–200 nm, and the naked capsid is approximately 100 nm in diameter.

Characteristic Genome

The genome of herpesviruses is double-stranded linear DNA ranging from 120 to 250 kilobase pairs. The variation in size of the DNA genome among herpesvirus family members is due to a characteristic arrangement of their DNA sequences. Their genomes have terminal and internal repeat sequences that are bounded by a unique long (UL) and a unique short (US) region. These repeats allow some members of the family to undergo genomic rearrangement of their unique regions, giving rise to different genomic isomers. In addition, because of the high incidence of rearrangement, spontaneous deletions occur, and defective viral particles are common. Within the herpesvirus family, there is little DNA homology among the different species, with a few notable exceptions to this generality. Specifically, HSV types 1 and 2 share 50% sequence homology and herpesvirus 6 and 7 share 30–50% homology. However, because of the low homology among the herpesviruses, digestion of the genome with restriction endonucleases will yield a characteristic fragment pattern specific for each family member. This allows for rapid epidemiologic tracing of a given strain.

The genome of the typical herpesvirus is large and encodes for at least 100 different proteins, many of which (>35) are structural in nature. Additionally, all herpesviruses encode for a number of virus-specific enzymes that are involved in nucleic acid metabolism. These can include DNA polymerase, thymidine kinase, dUTPase, ribonucleotide reductase, helicase, and primase. It is important to note that none of these enzymes are packaged within the mature virus particle. They are expressed only during replication within the host cell.

Subclassification

More than 130 different herpesviruses have been identified, including nine human isolates. The herpesviruses are subdivided into three subclasses without regard to DNA sequence homology. Assignment to one of the three subfamilies, the alpha-, beta- and gammaherpesviruses, is based on biological properties, with emphasis given to host-range restriction and cellular tropism.

Alphaherpesviruses Alphaherpesviruses have a broad host range; they replicate rapidly in many types of cells. Although these are highly lytic viruses, members of this subclass can establish latency in neurons and glial cells. This group includes human HSV-1 and HSV-2, human VZV, and bovine herpesvirus type 1 (BHV-1).

Betaherpesviruses Betaherpesviruses have a restricted host range and grow slowly in tissue culture. Infection typically causes cells to enlarge, which is termed cytomegalia. These members can establish latency in hematopoietic progenitor cells and can cause persistent infections of epithelial and glandular cells. This group includes human and murine cytomegaloviruses (CMV), human herpesvirus 6, and human herpesvirus 7.

Gammaherpesviruses Gammaherpesviruses have a restricted host range. These viruses can replicate in epithelial cells and establish long-term latency in lymphocytes. In addition, some of these viruses can transform or immortalize their host cell. This group includes EBV and Kaposi's sarcoma-associated human herpesvirus (also called human herpesvirus type 8).

Virus Replication

Viruses are obligate intracellular parasites; as such, members of the Herpesviridae family must enter a host cell to undergo replication. Entry is a multistep process beginning with weak association of the virus envelope with the sugar moieties of cell membrane-associated proteoglycans (i.e., heparan sulfate and chondroitin sulfate). The second step of entry involves specific binding of the virus to a surface receptor. For example, EBV recognizes complement receptor 2 (CD21 or CR2) and HLA class II molecules. Human cytomegalovirus recognizes β_2 -microglobulin, while HSV-1 binds to a member of the tumor necrosis factor/nerve growth factor receptor family. In addition to these common receptors, other glycoproteins have also been implicated in herpesvirus recognition of host target cells, and the group of receptors has been termed herpes virus entry mediators (HVEMs). Distribution of HVEMs determines the tropism of herpesvirus for specific cell types. After binding, the virus enters the host cell by fusion with the cell membrane. After entry, the viral envelope remains part of the cell's plasma membrane, but the capsid is transported through the cytoplasm to a nuclear pore. At the pore, the DNA is released from the capsid through the pore and into the nucleus. Once inside the nucleus, the viral DNA immediately circularizes.

Expression of the viral genome is tightly regulated and sequentially ordered in a cascade fashion. Immediate-early genes are expressed, yielding alpha proteins (alpha proteins are those that are produced in the absence of prior virus gene transcription). The alpha genes contain promoters that are recognized by the host cell's own transcription factors and are, therefore, turned on as soon as the viral DNA enters the nucleus of the cell. During a productive herpesvirus infection, the host cell's enzymes are hijacked

for the good of the virus. For example, herpesviruses are dependent upon the host cell's RNA polymerase II for all viral gene transcription. Even though the cell's own enzymes are used by the virus during replication/infection, normal cellular DNA and protein synthesis virtually stop as viral replication begins. Synthesis of the immediate-early alpha proteins promotes the expression of the next set of genes, or early genes, which are translated into beta proteins.

Beta proteins are those that are expressed before DNA replication begins but require alpha gene expression. Synthesis of the beta proteins represents the telltale sign that viral DNA synthesis has begun. It is believed that the beta proteins include the enzymes that are required to replicate the herpesvirus genome (i.e., DNA polymerase, thymidine kinase, dUTPase, ribonucleotide reductase, helicase, and primase). Because it circularizes when it enters the nucleus, herpesvirus DNA is synthesized by a rolling circle or concatametric mechanism that appears seemingly endless. The length of the replication cycle varies from about 18 h for herpes simplex virus to over 70 h for cytomegalovirus.

As viral DNA replication begins, late transcripts are produced that give rise to gamma proteins (gamma proteins are those that are expressed after DNA synthesis). While many of the alpha and beta proteins are enzymes or DNA-binding proteins, the gamma proteins are structural in nature and are the proteins of the tegument, toroid core, capsid, and envelope. In other words, synthesis of the gamma proteins is necessary for assembly of infectious virus progeny. Of interest, the virus's capsid is assembled before the newly synthesized viral DNA is inserted. This occurs in the cell nucleus. After encapsidation (i.e., DNA insertion into the capsid), maturation of infectious progeny occurs by budding of nucleocapsids through the inner nuclear membrane. Enveloped viral particles are then released from the cell through the plasma membrane by a poorly defined mechanism. The host cell that produces infectious progeny does not survive.

Overview of Diseases Caused by the Herpesviridae

As there are over 130 different types of herpesviruses, it is not surprising that they cause a wide variety of diseases. As their subclassification into alpha-, beta-, and gammaherpesviruses suggests, primary infection and possible latency may involve different cell types and present a widely different set of symptoms and clinical disease.

Herpes simplex viruses Herpes simplex viruses (HSV) are extremely widespread in the human population and exhibit an ability to replicate in many host

tissues. They also infect other animals. There are two main forms of the herpes simplex viruses, human herpes viruses (HHV)-1 and -2, which are also known as HSV-1 and HSV-2. Both viruses grow rapidly and are extremely cytolytic. Typically, these viruses infect epithelial cells found in mucosal tissue and establish latent infections in peripheral neurons. HSV-1 has been classically associated with oropharyngeal lesions characterized by recurrent fever blisters. HSV-1 is spread by contact, usually involving infected saliva. HSV-2 has been associated with recurrent genital herpes and is usually transmitted sexually or from mother to neonate during vaginal birth. Both viruses can cause neurologic disease, with HSV-1 being the leading cause of sporadic encephalitis in the United States. Both HSV-1 and HSV-2 can cause severe primary infections in newborns. Latent infection followed by recurrent disease is the hallmark of these viruses. HSV-1 and HSV-2 reside in latently infected ganglia in a nonreplicating state with only a few viral genes being expressed; infectious virus cannot be isolated during latency. Persistence of the viral genome within a latently infected cell lasts for the lifetime of the host. Following a reactivating stimulus, the virus follows axons back to the peripheral site and replication proceeds in the epithelial cells. Spontaneous reactivation occurs despite HSV-specific immunity. More than 80% of the human population harbors HSV-1 in a latent form, but only a small percentage experience reactivation. Neither the cellular nor the molecular basis for reactivation is completely understood.

Varicella-zoster virus Varicella-zoster virus (VZV or human herpesvirus 3) is highly contagious and is a very common virus. Like the HSVs, VZV is an alphaherpesvirus. The typical route of infection for VZV is the mucosa of the upper respiratory tract. After primary replication in the mucosal epithelium, the virus enters the circulation and eventually localizes to the skin, where it causes two distinct syndromes. Upon primary infection, usually in children, it causes chicken pox (also called varicella), which is characterized clinically by a generalized vesicular eruption of the skin and mucous membranes. However, VZV subsequently establishes a latent infection, and upon reactivation, the virus causes shingles (also called zoster). Shingles is a sporadic, incapacitating disease of adults or immunocompromised individuals that is clinically characterized by a rash limited in distribution to the skin innervated by a single sensory ganglion. The lesions are similar to varicella.

Epstein-Barr virus Epstein-Barr virus (EBV or human herpesvirus 4) causes infectious mononucleosis

and is latent in most adults. EBV is commonly transmitted by infected saliva (hence the nickname the kissing disease). Primary infection involves epithelial cells of the oropharynx and parotid gland. Viral shedding occurs for weeks to months after infection. Following replication in epithelial cells, EBV infects B cells and quickly becomes latent. Latent EBV is a factor in the development of nasopharyngeal carcinoma, Burkitt's lymphoma, and other lymphoproliferative disorders in immunocompromised individuals. As it has been known for years that cells in tissue culture could be immortalized by EBV, we now realize that EBV is the prototypic gammaherpesvirus and is oncogenic in humans.

Cytomegalovirus Cytomegalovirus (CMV or human herpesvirus 5) infects respiratory epithelial cells or epithelial cells within salivary glands or the kidneys. CMV is believed to infect a majority of the human population worldwide, often without symptoms of overt disease. However, CMV infection causes congenital defects and mental retardation in more than 5000 infants in the United States per year. CMV is the largest herpesvirus, with a genome of approximately 240–250 kilobase pairs. Infection is characterized by massive enlargement of infected cells. Human CMV is the prototypical betaherpesvirus. Primary CMV infections in immunocompromised hosts are more severe than in normal hosts. Therefore, organ transplant recipients on immunosuppressive therapy and individuals with AIDS are predisposed to life-threatening infections by CMV.

Human herpesvirus 6 and human herpesvirus 7 Human herpesvirus 6 (HHV-6) and human herpesvirus 7 (HHV-7) were first recognized in 1986 and 1990, respectively. Both viruses were isolated from cultures of peripheral blood cells from patients with lymphoproliferative disorders involving T cells. Both have been classified as betaherpesviruses, as infection results in the formation of giant cells. HHV-6 and HHV-7 share significant sequence homology; however, they are antigenically and immunologically distinct. HHV-6 has been further subclassified into HHV-6A and HHV-6B based on distinct cellular tropism. Although HHV-6B is associated with roseola or exanthema subitum, HHV-6A and HHV-7 have not been associated with any known human disease.

Human herpesvirus 8 Human herpesvirus 8 is also known as Kaposi's sarcoma-associated herpesvirus; it was originally isolated in the mid-1990s from an AIDS patient with Kaposi's sarcoma. It is present in all forms of Kaposi's sarcoma, in primary B-cell effusion lymphomas, and in some forms of Castleman's

disease, each of which is more prevalent in immunocompromised patients. As each of these diseases is a form of cancer, HHV-8 has been characterized as a human tumor virus or gammaherpesvirus. HHV-8 expresses a number of cellular regulatory genes that the virus appears to have acquired from mammalian cells as a way of defeating host defense mechanisms.

B virus B virus is highly pathogenic for humans, although transmissibility of the virus from its normal host to humans is limited. Typically, B virus infects Old World monkeys and is enzootic in rhesus, cynomolgus, and other macaque monkeys. B virus infections in the natural host seldom cause overt disease, but latency is established and B virus is easily reactivated under stressful conditions. Infection of humans usually results from monkey bites and is associated with a high rate of mortality due to an acute, ascending myelitis and encephalomyelitis.

Immunobiology and Recurrent Disease

The course or natural history of a herpes viral infection is influenced by the site of viral entry and host defenses. Primary infection with a herpesvirus induces both innate and adaptive immune responses, which result in the termination of viral replication and cessation of clinical symptoms. However, unlike acute viral infections, in which a successful immune response leads to elimination of the virus, infection with herpesviruses frequently leads to latency and thus the opportunity for episodic recurrent disease.

Innate Immunity

All mammals are born with a genetically inherited resistance to infectious microbes termed innate or natural resistance. Expression of natural resistance does not require previous exposure to the viral antigen (as does the adaptive immune response). Recent evidence suggests that toll-like receptor (TLR)-2 and TLR-9, which are pattern recognition receptors, are the first to detect that the host has been infected with herpesvirus. Activation of the TLR pathway initiates the innate immune response, which is aimed at restricting the early spread and replication of the virus without regard to the type of virus. Macrophages, natural killer (NK) cells, and killer cells mediating antibody-dependent cellular cytotoxicity (ADCC) all have roles in restricting the early spread of the virus. Infection by members of the Herpesviridae family is also accompanied by the expression of genes encoding the type I interferons that impart resistance to uninfected cells surrounding an infected cell.

Adaptive Immunity

Most infections with members of the Herpesviridae result in the induction of virus-specific humoral (antibody) and cell-mediated immune responses. In contrast to innate immunity, it takes several days for these herpesvirus-specific adaptive immune responses to be generated. A primary infection with HSV, whether asymptomatic or symptomatic, results in the production of virus-specific antibodies (both binding and neutralizing). These antibodies play multiple roles in mediating resolution of the primary infection through direct viral neutralization and through lysis of the infected cells by an ADCC mechanism. Virus-specific T cells are also very important during infection; individuals with cell-mediated immunodeficiency suffer severe herpetic disease. Clearance of virus correlates with the presence of CD8⁺ T cells with virus-specific cytolytic activity. Immunological memory in both the B cell and T cell compartments also follows primary infection. However, neither the successful termination of virus replication during primary infection nor the establishment of immunological memory prevents reactivation of latent virus and the development of recurrent lesions.

Recurrent Disease

Whether the time to reactivation of latent virus is weeks to months (as is the case for HSV) or years to decades (as is the case for VZV), it is apparent that immune memory, which results from the resolution of a primary infection with these viruses, offers little protection against recurrent disease. The development of recrudescence lesions is a two-step process. The first step occurs at the cellular/molecular level and involves the regulation of viral and host transcription factors influencing the state of expression of the latent viral genome. The balance between positive and negative transcription factors determines whether the genome will be reactivated, and this process is mediated by numerous hormonal and growth factor signals. These signals may result from many different stimuli. Psychological stress, physical trauma, UV irradiation from a sunburn, and hormonal fluctuations of the menstrual cycle have all been associated with reactivation of latent virus.

The period of time in which the reactivated virus gets to replicate, and thus cause clinical symptoms, is immunologically limited (and represents the second step in the process). Healthy adults generally limit clinical symptoms during a recurrence to a couple of days or less, whereas individuals with diminished cell-mediated immunity, as seen in AIDS, often develop severe disease that lasts much longer.

Stress and Herpesvirus Reactivation

The capacity of latent herpesviruses to reactivate and replicate is essential for completion of the viral life cycle. In the case of HSV-1 latency, reactivation of herpesvirus from ganglia results in the appearance of infectious virus at the site of the initial primary infection – typically the lips or gingiva. During the latent phase of any herpesvirus infection, viral gene expression is highly restricted; latency-associated transcripts (LATs) are among the few viral genes routinely detected during latency. It is believed that LAT gene expression establishes and maintains latency as mutations in LAT or its promoter limit herpesvirus latency.

The factors that lead to reactivation are even less clear. The limited knowledge of the pathogenesis of herpesvirus reactivation has been generated from studies in laboratory animals including mice, guinea pigs, and rabbits. In these models, reactivating stimuli range from mechanical and pharmacological to immunological alterations of cells and surrounding tissues. It is known that the first viral genes that are expressed during reactivation are the herpesvirus alpha, or immediate-early, genes whose promoter regions contain sequences that are similar to the host's cells own promoters. Reactivating stimuli may drive transcription of cellular genes whose promoter regions share homology with the promoters of the herpesvirus alpha genes. Thus, it is thought that the alpha genes are transcribed by the host cell's own transcription factors. The virus is then triggered for reactivation.

It is commonly thought that physical and psychological stress may be contributing factors to reactivation of herpesvirus. The impact of psychological stress on immune function is well documented in the literature and supports the hypothesis that psychological stress can reactivate latent virus. Several studies have shown that reactivation of latent HSV infections is more frequent in individuals who experience traumatic life events such as the death of a family member, the stress of interpersonal problems, or work-related difficulties. Similar relationships between psychosocial stress and reactivation of other herpesviruses, including EBV and VZV, have also been reported. It is thought that the physiologic alterations that ensue during stress alter the cellular and molecular microenvironment and serve as reactivators of latent herpes infections.

In mammals, stressors typically activate a specific set of core stress responses, including the hypothalamic-pituitary-adrenal (HPA) axis and sympathetic nervous system (SNS). Activation of the HPA axis results in increases in glucocorticoids such as cortisol, while activation of the SNS results in increases in circulating and tissue levels of catecholamines such as epinephrine and norepinephrine. Each of these

alterations in host physiology has been linked to herpes reactivation. For example, a model for *in vivo* reactivation has been established in the rabbit eye. Epinephrine induces reactivation and shedding of mature infectious virus in the rabbit's tears. Glucocorticoids have also been implicated in reactivation. For example, cortisol can drive viral replication in primary cell cultures established from latently infected mouse trigeminal ganglia. Other experimental stressors (e.g., UV irradiation and transient hyperthermia), in a variety of animal models, can also cause recrudescence. However, the events that occur between stress and production of infectious herpesvirus are incompletely understood *in vivo*. Therefore, important biological knowledge remains unknown concerning reactivation.

Conclusion

The herpesviruses represent a diverse group of viruses that have evolved unique relationships with their hosts. Members of this viral group have been associated with diverse disease manifestations ranging from malignant transformation to latency and reactivation. Herpes infections occur from the cradle (neonatal HSV encephalitis) to the grave (shingles caused by reactivation of VZV in the later decades of life). They affect healthy individuals (e.g., CMV and mononucleosis) and immunodeficient individuals (e.g., human herpes virus type 8 which causes Kaposi's sarcoma in AIDS patients). Reactivation of the latent members of the herpesvirus family has been associated with multiple forms of stress (from physical stressors such as UV irradiation to psychological stress such as social conflict). Unfortunately, detailed knowledge of the molecular steps that trigger reactivation is lacking. Additional research is needed to gain a better understanding of these mechanisms and to help develop new prevention and treatment strategies that limit the replication, spread, and reactivation of the Herpesviridae.

See Also the Following Articles

Disease, Stress Induced; HIV Infection/AIDS.

Further Reading

- Brooks, G. F., Butel, J. S. and Ornston, L. N. (2004). Herpesviruses. In: Jawetz, E., Melnick, J. L. & Adelberg, E. A. (eds.) *Medical Microbiology* (23rd edn., pp. 429–453). McGraw-Hill Companies.
- Enright, A. M. and Prober, C. G. (2004). Herpesviridae infections in newborns: varicella zoster virus, herpes

- simplex virus, and cytomegalovirus. *Pediatric Clinics of North America* 51, 889–908.
- Knipe, D. M. and Howley, P. M. (eds.) (2001). *Fields virology* (4th edn.). Philadelphia, PA: Lippincott Williams & Wilkins.
- Mitchell, B. M., Bloom, D. C., Cohrs, R. J., Gilden, D. H. and Kennedy, P. G. (2003). Herpes simplex virus-1 and varicella-zoster virus latency in ganglia. *Journal of Neurovirology* 9, 194–204.
- Moore, P. S. and Chang, Y. (2003). Kaposi's sarcoma-associated herpesvirus immunoevasion and tumorigenesis: two sides of the same coin? *Annual Review of Microbiology* 57, 609–639.

Hippocampal Neurons

R L Spencer and S T Bland

University of Colorado, Boulder, CO, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R D Brinton and T W Berger, volume 2, pp 364–371, © 2000, Elsevier Inc.

Overview

Hippocampal Formation

Laminar Organization of Dentate Gyrus and Hippocampus

Lammelar Organization (Trisynaptic Circuit) of Dentate Gyrus and Hippocampus

Principal Neurons

Interneurons

Intrinsic and Extrinsic Neural Connections

Neurochemistry

Neuroplasticity

Hippocampal Function

Glossary

<i>Adenylyl cyclase</i>	A protein enzyme that when activated, for example by a G-protein, catalyzes the formation of an intracellular second-messenger molecule, cyclic adenosine monophosphate (cAMP).
<i>Cation channel</i>	A protein complex embedded in a cell's membrane that allows positively charged ions (e.g., Na ⁺ , K ⁺ , or Ca ²⁺) to pass through the membrane.
<i>G-proteins</i>	Guanine nucleotide (guanosine triphosphate, GTP, and guanosine diphosphate, GDP) binding proteins that associate with the intracellular portion of metabotropic receptors. G-protein alterations are the first step in a cascade of intracellular molecular changes triggered by neurotransmitter or hormone (ligand)

binding of metabotropic receptors at the membrane surface.

Neuropil

Brain region that consists primarily of neuronal processes (dendrites and axons) rather than neuronal cell bodies.

Spatial learning

Learning tasks in which optimal performance requires that the organism has learned the relative spatial relationship between objects in the environment. Spatial learning can be demonstrated by tasks in which the organism is able to select the most direct route to a location that it has been to before or, alternatively, by tasks in which the organism avoids revisiting places that it has been to before.

Overview

The hippocampal formation consists of the hippocampus and the closely associated dentate gyrus and subiculum. The hippocampal formation first attracted attention as a primitive region of the cortex that had an architecturally simple and orderly cellular organization. Strikingly detailed drawings of the microscopic organization of the hippocampal formation were provided by Ramón y Cajal in the late 1800s (**Figure 1**). This structure is dominated by the orderly alignment of two principal neuronal cell types: pyramidal cells and granule cells. Moreover, the principal cells of the hippocampal formation were found to be interconnected in such a way as to produce a simple intrinsic neurocircuit (the trisynaptic circuit). Interest in the hippocampal formation intensified when, based on human amnesic cases, it was found to be essential for the formation of new episodic memories and the acquisition of declarative facts. The intense study of the hippocampus has yielded the first characterized cellular functional response (long-term potentiation) believed to reflect the fundamental neuroplasticity underlying memory formation.

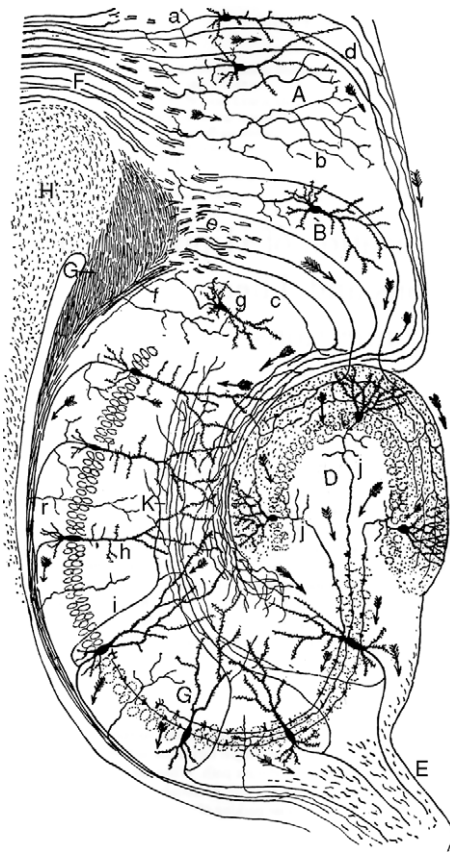


Figure 1 Drawing of a transverse section of the hippocampal formation highlighting the principal cells (e.g. B, D, G, h), their interconnections and presumed flow of neural impulses (arrows). Drawing by S. Ramón y Cajal, first published in 1911 (Ramón Y Cajal, *Histologie du système nerveux de l'Homme et des Vertébrés*. Paris: A. Maloine, 1911).

Subsequent study of the hippocampus has revealed a range of neuroplasticity phenomena that include dynamic changes in neuronal arborization, spine density, and synapse formation. Investigations in recent years have determined that neurogenesis is a routine feature of the adult dentate gyrus portion of the hippocampal formation. The characterization of these forms of hippocampal neuroplasticity has revolutionized recent neuroscience with the discovery that these neuroplastic changes are not restricted to critical developmental periods or recovery from brain injury but are a normal daily occurrence that persists throughout the life span.

Hippocampal Formation

Each of the components of the hippocampal formation is characterized by a single principal cell body layer sandwiched between two zones of neuropil, yielding a simplified version of cortex with only

three distinct layers. Extensively associated with the hippocampal formation, in terms of interconnections, is the adjacent six-layer entorhinal cortex. The hippocampus and its interleaved partner, the dentate gyrus, are two phylogenetically primitive cortical structures that are buried underneath the cortical mantle. In the human brain, these two enfolded sheets of rudimentary cortex are located within the temporal lobe (Figure 2). In the rodent brain, these structures extend more dorsal-anteriorly. The dorsal portion of the hippocampus is located immediately beneath the corpus callosum and has an anterior extension that approaches the septal region (Figure 3). The most caudal extent of the hippocampus is located in the ventral half of the caudal portion of rodent forebrain.

Dissected out of a rat brain, the combined dentate gyrus and hippocampus are somewhat banana-shaped. The longitudinal axis of this structure is often referred to as the septal-temporal axis of the hippocampus and dentate gyrus. The rostral approximately two-thirds of this structure is known as the dorsal hippocampus, and the caudal third is known as the ventral hippocampus. This structure is usually studied and depicted in the transverse/cross-sectional plane (Figure 3). The cellular organization of the hippocampus and dentate gyrus in mammals is remarkably consistent throughout its longitudinal extent. Recent anatomical studies, however, have demonstrated important differences in intrinsic and extrinsic connections within the dorsal and ventral hippocampus. These differences may have important functional significance. For example, there is some evidence that the dorsal hippocampus is more important for spatial learning and memory and that the ventral hippocampus selectively contributes to fear and other emotion-dependent behaviors and learning.

Laminar Organization of Dentate Gyrus and Hippocampus

The principal cell body layer of the hippocampal formation consists primarily of pyramidal cell bodies (hippocampus and subiculum) or granule cell bodies (dentate gyrus). The surrounding neuropil zones primarily comprise the proximal principal cells' dendrites and axons and the nerve terminals of extrinsic and intrinsic afferents. This laminar organization can be easily visualized when examining microscopically a transverse section of dentate gyrus and hippocampus (Figure 4a). Anatomists have subdivided the neuropil zones into different strata, based primarily on microscopic differences in the principal cell processes and the afferents and efferents present within each strata (Figure 4b). Note that each of these cortical

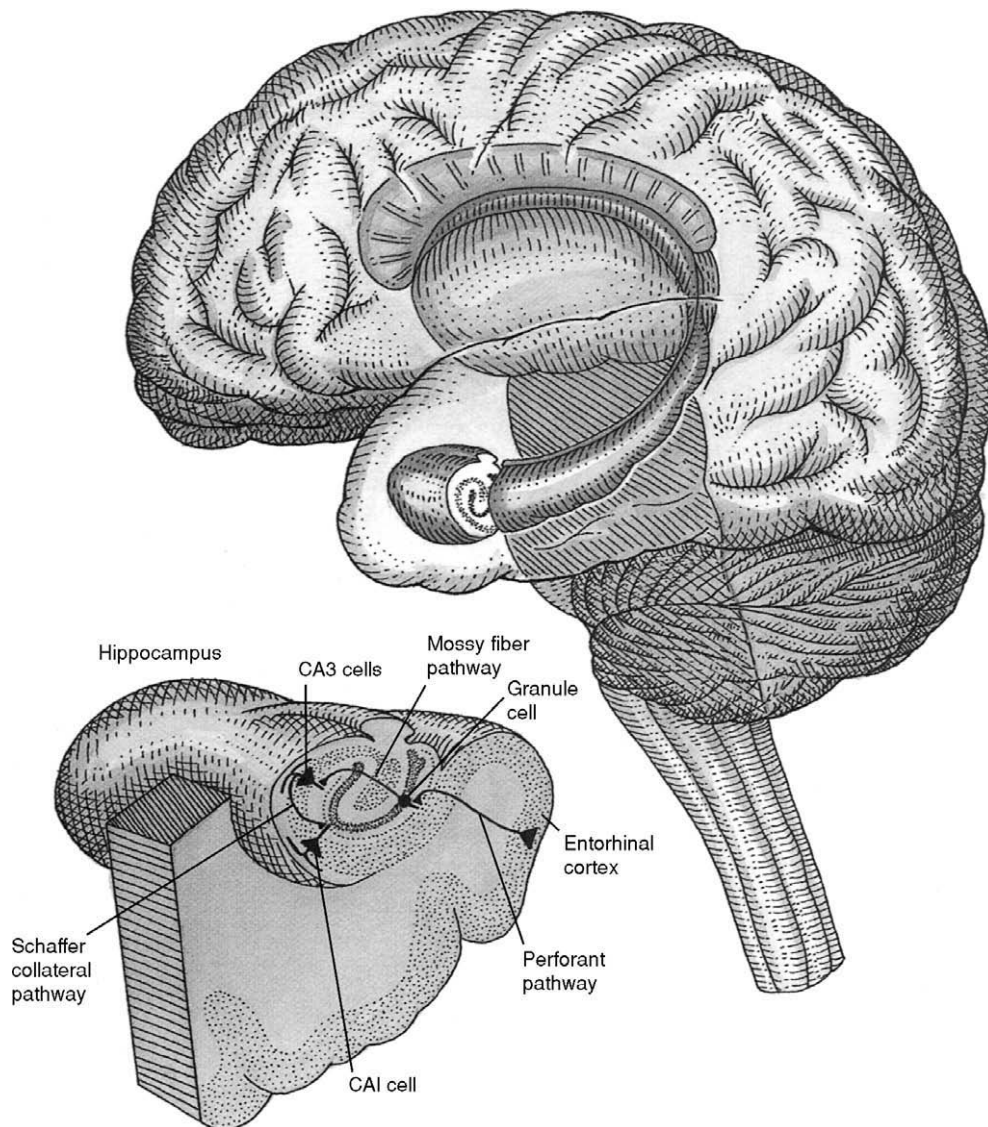


Figure 2 Relative location of hippocampal formation in human brain. The hippocampus and its interleaved partner, the dentate gyrus, reside beneath the cortical surface in the temporal lobe. An enlarged transverse slice of the hippocampus and dentate gyrus is depicted in the lower left portion of the figure. From Squire, L. R. and Kandel, E. R. (2000), *Memory, from mind to molecules*, New York: Scientific American Library.

sheets has been folded in half along its longitudinal axis. Therefore, when we view a cross section of this structure in the rat brain, the orderly aligned principal cell bodies form a C-shaped (hippocampus) or V-shaped (dentate gyrus) formation. As is evident when we view a horizontal cross section of ventral hippocampus, one edge of the hippocampal sheet is adjacent to the subiculum and entorhinal cortex (Figure 4c–d). The other edge of the hippocampal sheet is tucked inside the two folds of the dentate gyrus. The C-shaped formation of pyramidal cell bodies in the hippocampus is subdivided into three subfields based on morphological and interconnection distinctions. These subfields are designated

CA1, CA2, and CA3 (drawing on the Latin name for the hippocampus, *cornus Ammonus*, Ammon's horn). The CA3 region is adjacent to the dentate gyrus, and the CA1 region is adjacent to the subiculum. The smaller CA2 region has morphological, phenotypical, and connectional features that overlap with those of the CA3 and CA1 subregions, but in combination makes for a clearly distinct functional zone of cells. The granule cells in the dentate gyrus are relatively uniform throughout the V-shaped formation in terms of morphology, innervation, and projections. The two folds of dentate gyrus giving rise to the V-shaped formation are commonly referred to as the enclosed and free blades, or the

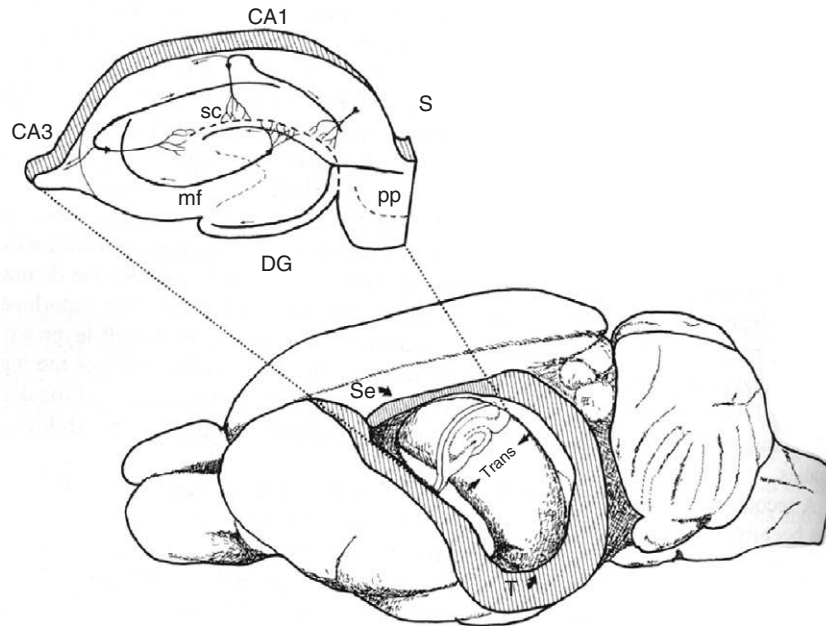


Figure 3 Relative location of hippocampal formation in rat brain. The hippocampus (including CA1 and CA3 subregions) and dentate gyrus (DG) form a banana-like structure with one end of the longitudinal axis located dorsal-rostrally near the septal (Se) region and the other end located ventral-caudally near the temporal (T) pole of the rat forebrain. An enlarged transverse slice is depicted in the upper portion of the figure. mf, mossy fiber; pp, perforant path; sc, Schaffer collateral; S, subiculum. From Amaral, D.G. and Witter, M.P. (1989) The three-dimensional organization of the hippocampal formation: a review of the anatomical data. *Neuroscience* 31, 571–591.

suprpyramidal and infrapyramidal blades, of the dentate gyrus. Recent studies indicate that there are some clear differences in the relative afferent connections to the two blades of the dentate gyrus as well as discriminable tendencies in their projection to CA3 subregions. The intervening space between the two granule cell body-forming blades of the dentate gyrus is known as the hilus or polymorphic region of the dentate gyrus.

Lammellar Organization (Trisynaptic Circuit) of Dentate Gyrus and Hippocampus

Significantly, many of the intrahippocampal neural connections are maintained within a narrow cross-sectional area (lammellae). A simple unidirectional trisynaptic connection among the components of the hippocampal formation was deduced by Ramón y Cajal based on microscopic architecture (Figure 1). Around 1970, Andersen and colleagues provided electrophysiological evidence to support a predominantly lammellar organization of hippocampal function. This trisynaptic circuit (Figure 5a) begins with pyramidal neurons predominantly localized in layer II of entorhinal cortex. These cells project (perforant pathway) to both blades of the dentate gyrus

and form synapses on granule neuron dendrites (synapse 1). The granule neurons project (mossy fiber pathway) to CA3 and form synapses on the apical dendrite of CA3 pyramidal neurons within the stratum lucidum (synapse 2). A collateral axonal branch of the CA3 pyramidal neurons projects (Schaffer collateral) to CA1 and forms synapses on basilar and apical dendrites of CA1 pyramidal neurons (synapse 3). The primary output from this circuit, and from the hippocampus overall, proceeds from CA1 neurons to the adjacent subiculum and entorhinal cortex as well as to subcortical structures.

More recent investigations have determined that there is considerable divergence and convergence within the longitudinal plane of connections between the entorhinal cortex and the dentate gyrus and between CA3 pyramidal neurons and other CA3 and CA1 cells. Only the mossy fiber connections between the dentate gyrus and CA3 are predominantly restricted to a cross-sectional lammellae. Nevertheless, a fairly thin cross section of the dentate gyrus and hippocampus contains a number of intact cells of each trisynaptic circuit component. Therefore, the hippocampal formation is ideally suited for *ex vivo* study, in which a 250- to 500- μ m cross-section of this structure contains an electrophysiologically functional trisynaptic circuit. Consequently, with appropriate bath conditions for

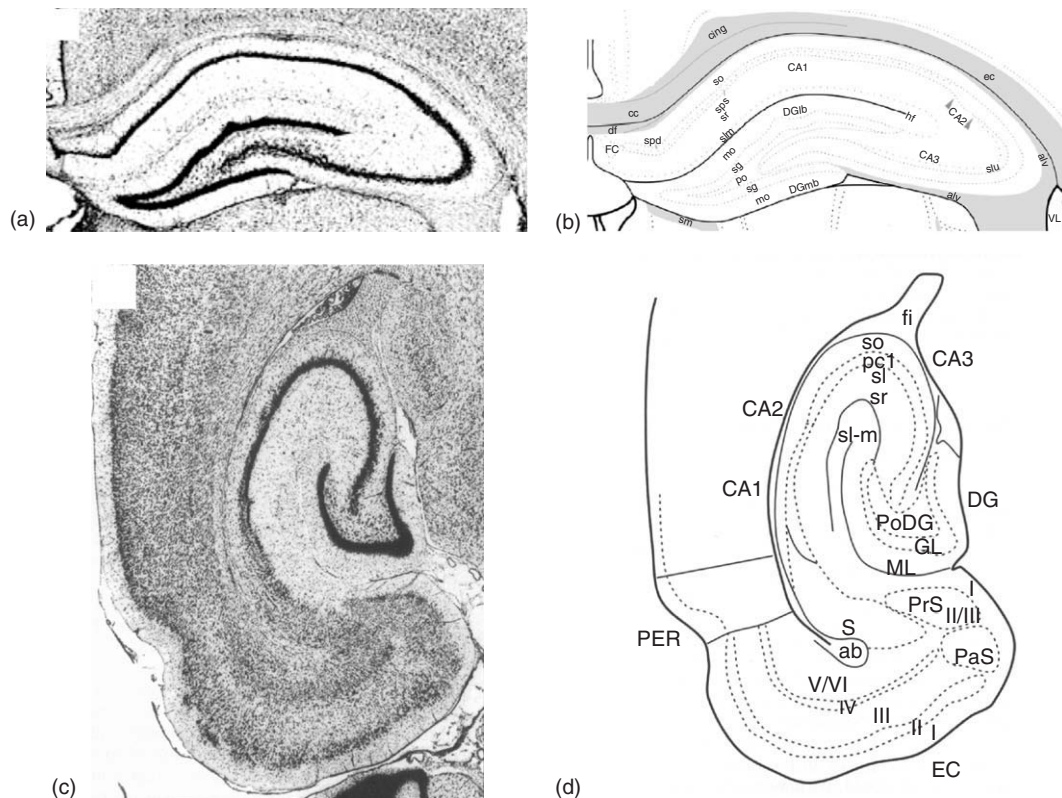


Figure 4 Laminar organization of rat hippocampal formation. a, Photomicrograph of histologically stained tissue of transverse section of the dorsal hippocampus (coronal plane); b, corresponding line drawing; c, photomicrograph of histologically stained tissue of transverse section of the ventral hippocampus (horizontal plane); d, corresponding line drawing. The strata of the hippocampus shown are stratum oriens (so), pyramidal cell body layer (spd, sps, or pcl), stratum lucidum (sl), stratum radiatum (sr), and stratum lacunosum-moleculare (slm). The strata of the dentate gyrus shown are the molecular layer (mo), granule cell body layer (gl or sg), and the hilus or polymorphic region (po). The tightly packed and orderly aligned pyramidal and granule cell bodies result in a darkly stained C- or V-shaped formation in the hippocampus or dentate gyrus, respectively. ab, angular bundle; alv, alveus; cc, corpus callosum; cing, cingulum; DGlb, dentate gyrus lateral or enclosed blade; Dgmb, dentate gyrus medial or free blade; EC, entorhinal cortex (layers I–VI); fi, fimbria of fornix; hf, hippocampal fissure; PaS, parasubiculum; PER, perirhinal cortex; PrS, presubiculum (layers I–III); S, subiculum; VL, lateral ventricle. Panels a and b from Swanson, L. W. (2004), *Brain maps: structure of the rat brain* (3rd edn.), Amsterdam: Elsevier Academic Press.; panels c and d from Witter, M. P. and Amaral, D. G. (2004), Hippocampal formation, In: Paxinos, G. (ed.) *The rat nervous system* (3rd edn.), pp. 635–704. Amsterdam: Elsevier Academic Press.

neuronal survival, a cross-sectional brain slice of the hippocampus and dentate gyrus maintains many functional features of an intact hippocampal circuit. The *ex vivo* electrophysiological study of hippocampal slices has been instrumental in deciphering the cellular and molecular correlates of hippocampal-dependent learning (Figure 5b).

Principal Neurons

All the principal cells of the hippocampal formation (pyramidal and granule cells) are believed to be glutamatergic. Pyramidal cells have a distinctly pyramidal or tear-drop shape. Pyramidal cells in CA3 and CA2 have larger cell bodies (approximately 25 μm in diameter) than those in CA1 (approximately 15 μm in diameter) (Figure 6). Separate multibranching dendritic trees emerge from the apex and base of

each cell. The base of the cell is defined by the surface from which the axon emerges, and the separate dendritic trees are referred to as apical and basilar dendrites, with each cell having one or two apical dendrites and several basilar dendrites. All hippocampal pyramidal cells are aligned in the same orientation, with the apical dendritic trees radiating toward the center of the C-shaped semicircle of cells (filling the stratum radiatum). The basilar dendrites extend in the opposite direction (filling the stratum oriens). Hippocampal pyramidal cell axons penetrate the stratum oriens and travel as part of the alveus to distal hippocampal or subcortical targets.

Granule cells of the dentate gyrus have small spherical or ovoid-shaped cell bodies (approximately 8–15 μm in diameter). One moderately complex dendritic tree (filling the stratum moleculare) is attached to the pole of the cell opposite from where

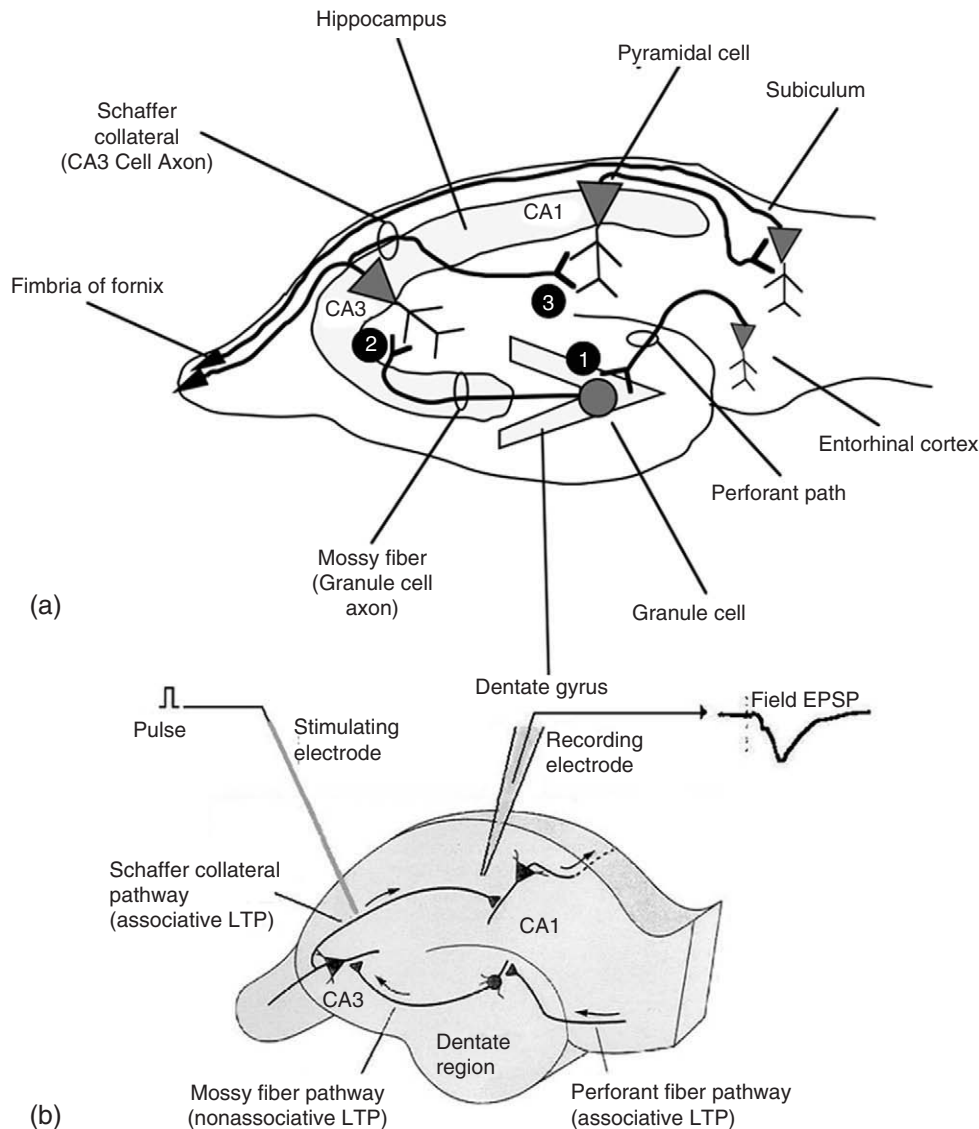


Figure 5 Hippocampal-formation trisynaptic neural circuit. a, An intact unidirectional trisynaptic neural circuit in a 250- to 500- μ m transverse section of rat hippocampal formation (lamellae); b, *ex vivo* stimulation of a transverse slice of hippocampal formation placed in a chamber containing a solution of artificial cerebrospinal fluid to assess the electrophysiological activity in one component of the trisynaptic circuit. In (a) representative pyramidal cells (present in the entorhinal cortex, CA3, CA1, and subiculum) and a granule cell (in the dentate gyrus) that participate in the circuit are depicted. The three featured synapses of the circuit are numbered in the temporal order in which they are engaged by neural activity, starting with pyramidal neurons in the entorhinal cortex. In (b) a stimulating and recording electrode is applied to the tissue slice in order to assess electrophysiological activity within one component of the trisynaptic circuit. EPSP, excitatory postsynaptic potential; LTP, long-term potentiation. Adapted from Squire, L. R. and Kandel, E. R. (2000), *Memory, from mind to molecules*, New York: Scientific American Library.

the axon emerges and projects. There is also a population of large-diameter cells in the hilus region of the dentate gyrus called mossy cells that also appear to be glutamatergic. These cells receive input from the mossy fibers and project back to the dentate gyrus granule cells. Interestingly, mossy cells preferentially project bilaterally to granule cells at other levels of the longitudinal axis than that in which they reside.

The dendritic trees of hippocampal formation principal cells are densely covered with bulbous spines

(Figure 6). The head of these spines form the postsynaptic site for much of the excitatory input to these cells. There are especially large prominent spines on CA3 neurons and mossy cells called thorny excrescences, which form the postsynaptic site for the mossy fiber input from the dentate gyrus. Inhibitory synapses, largely arising from γ -aminobutyric acid (GABA)ergic interneurons, are mostly localized to the soma, dendritic shafts, and axonal initial segments of the principal cells.

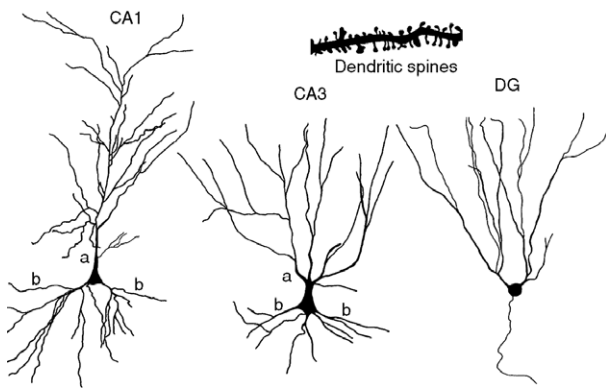


Figure 6 Hippocampal formation principal cells. Camera lucida tracings of representative CA1 and CA3 pyramidal cells in the hippocampus and a granule cell in the dentate gyrus (DG) from rat brain are shown. All three representative cells are oriented such that the axon emerges from the bottom of the cell body (an axon is clearly evident for the granule neuron). Multibranched apical (a) and basilar (b) dendritic trees are attached to the pyramidal cell body. The dendritic trees for the granule cell are attached to the upper portion of the cell body. A camera lucida tracing of an enlarged portion of a CA1 apical dendrite is also included. This tracing illustrates the dense presence of spines that cover the dendrites of these principal cells. Adapted from Gould, E., Woolley, C.S., Frankfurt, M., et al. (1990), *Gonadal steroids regulate dendritic spine density in hippocampal pyramidal cells in adulthood*. *Journal of Neuroscience* 10,1286–1291. © 1990 by the Society for Neuroscience.

Interneurons

There are various types of interneuron cell bodies scattered throughout all strata of the dentate gyrus and hippocampus. Most of these interneurons are presumed to be GABAergic, and they have been further categorized by the differential expression of various neuropeptides and calcium-binding proteins. Many of these interneurons have extensive axon collaterals and arborization that extend fairly long distances within both the longitudinal and transverse plane and form numerous synaptic contacts with principal cells. Although there are substantially fewer hippocampal interneurons than principal cells (probably totaling less than 1% of all hippocampal neurons), this pattern of connectivity allows them to have substantial effects on overall hippocampal function. Two of the better-characterized types of inhibitory interneurons are the basket cells and the chandelier or axoaxonic cells. The basket cells are so named because their multibranched axons form an extensive interwoven contact with the cell body of target pyramidal cells. The chandelier cells form strong synaptic contacts on the initial axonal segment of pyramidal cells. It should also be noted that there are many synaptic contacts between inhibitory interneurons as well as axodendritic contacts from principal cells. Thus, the

interneurons are interconnected in such a way as to mediate complex feedforward and feedback inhibition and disinhibition within the hippocampal formation.

Intrinsic and Extrinsic Neural Connections

A major source of neural input to the dentate gyrus, as previously described, is from the entorhinal cortex and constitutes the first component of the trisynaptic circuit. There are considerable reciprocal connections of the entorhinal cortex with other regions of neocortex, most notably the adjacent perirhinal cortex and the nearby postrhinal and piriform cortex. There are also substantial direct interconnections with the medial prefrontal, insular, and retrosplenial cortex. Thus, the entorhinal cortex is believed to provide the dentate gyrus and hippocampus with the rich multimodal sensory content of ongoing experience. In addition to the extensive input to the dentate gyrus, the entorhinal cortex also provides (via the temporal-ammonic pathway) direct monosynaptic parallel input to the CA3 region of the hippocampus, as well as direct input to CA1 pyramidal cells. Thus, both CA3 and CA1 neurons are in the position to integrate or compare direct and processed information from the entorhinal cortex. The projection of CA1 neurons back to the entorhinal cortex, either directly or via the subiculum, makes the entorhinal cortex the principal bidirectional interface between the hippocampal formation and various neocortical regions.

In addition to the trisynaptic circuit, there are extensive commissural and longitudinal associational connections within the hippocampus. The major source of these additional associational connections are provided by the non-Schaffer collateral branch of CA3 pyramidal cell axons, with additional contribution from CA2 pyramidal cells and mossy cells of the hilus.

There are other important neural inputs to the dentate gyrus and hippocampus from various populations of biogenic amine-containing neurons that may primarily modulate the general excitatory tone and neural synchronization within this structure. In addition, the CA1 region receives direct excitatory input from basal nuclei of the amygdala and midline nuclei of the thalamus. Interestingly, the amygdala input to the hippocampus is largely restricted to the ventral hippocampus.

The major output neurons of the hippocampal formation are pyramidal cells in CA1, subiculum, and deep layers of the entorhinal cortex. CA1 and subiculum pyramidal cells provide a large subcortical projection via the fornix to the septum and hypothalamus, with minor contribution from CA3 pyramidal cells. There are also strong direct connections of the

CA1 and subiculum with the medial prefrontal cortex. However, as previously indicated, there are more extensive indirect connections of the CA1 and subiculum with the prefrontal and other cortical areas via the entorhinal cortex.

Neurochemistry

Amino Acid Neurotransmitters

The excitatory neurotransmitter, glutamate, is released from hippocampal-formation principal cells and is therefore the neurotransmitter used by the perforant path, mossy fibers, and Schaffer collaterals. Glutamate receptor subtypes include the ionotropic receptors (iGluRs), which directly gate cation channels and are named after the ligands that preferentially bind to them: *N*-methyl-D-aspartate (NMDA), α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA), and kainate. All the iGluRs are present in the excitatory pathways of the hippocampus. AMPA and kainate receptors mediate fast excitatory postsynaptic potentials (EPSPs), and NMDA receptors mediate slower EPSPs. All the iGluRs open Na^+ and K^+ channels. NMDA as well as some AMPA and kainate receptors are also permeable to Ca^{2+} . The NMDA receptor is unique in that channel permeability is both ligand- and voltage-dependent.

There is also a family of metabotropic glutamate receptors (mGluRs) that are coupled to guanine nucleotide-binding proteins (G-proteins). mGluRs have both pre- and postsynaptic localization and can coexist in the postsynaptic density with iGluRs. Postsynaptically, mGluR regulate synaptic transmission by stimulating specific second-messenger cascades and generating slow synaptic responses. When expressed presynaptically, mGluR activation can modulate transmitter release.

The inhibitory neurotransmitter GABA can also act on both ionotropic and metabotropic receptors. GABA_A receptors are ionotropic and gate ion channels that are permeable to the anion Cl^- . GABA also binds a metabotropic receptor, GABA_B, which is coupled to a G-protein that, when activated, opens K^+ channels. When expressed postsynaptically, GABA_B receptor activation leads to a slow hyperpolarization. When expressed presynaptically, GABA_B receptor activation reduces neurotransmitter release.

Biogenic Amines

The major cholinergic input to the hippocampus is from the medial septal nucleus. Acetylcholine (ACh), similar to glutamate and GABA, can act on both ionotropic (nicotinic receptors) and metabotropic (muscarinic) receptors, both of which are expressed

in the hippocampus. Nicotinic ACh receptors are ligand-gated ion channels that mediate fast excitatory transmission via an influx of Na^+ and Ca^{2+} . Nicotinic receptors are expressed on both excitatory and inhibitory interneurons in the hippocampus and can be expressed pre- and postsynaptically. Muscarinic receptors are very densely expressed in the hippocampus, and different subtypes are differentially expressed on specific populations of neurons: M_1/M_3 on principal cells and M_2/M_4 on interneurons. The actions of muscarinic receptor activation can be excitatory or inhibitory, depending on the type of G-protein to which the receptor is coupled: M_1/M_3 are coupled to G_q (which stimulates phospholipase C), and M_2/M_4 are coupled to $\text{G}_{i/o}$ (which can inhibit adenylyl cyclase or activate a K^+ channel). Muscarinic receptor activation potentiates NMDA receptor activity, probably by the activation of the M_1 subtype. M_2 can be expressed as autoreceptors on cholinergic terminals. Acetylcholine is also known to regulate the release of dopamine, serotonin, and neuropeptides by the activation of presynaptic nicotinic and muscarinic receptors.

Cholinergic neurons that project to the hippocampus from the medial septal nucleus and the nucleus of the diagonal band of Broca contribute to the synchronous discharge of principal cells reflected by the prominent theta frequency hippocampal electroencephalograph (EEG). Theta is observed both during voluntary movement (Type 1, 8–12 Hz) and during alert immobility (Type 2, 4–6 Hz), although only Type 2 theta is sensitive to cholinergic manipulations.

Dopamine afferent terminals are observed in the ventral subiculum, in all strata of CA1, in the stratum oriens of CA3, and in the hilus of the dentate gyrus. The source of hippocampal dopamine comes from the midbrain ventral tegmentum and substantia nigra and there is some evidence for different populations of dopaminergic cells innervating dorsal and ventral hippocampal formations. Dopamine receptors comprise two types, D1-like (D1 and D5) and D2-like (D2, D3, and D4), and all are metabotropic. D5 is the most prominent D1-like receptor in the hippocampus and is coupled to G_s (which stimulates adenylyl cyclase). D5 receptors are expressed on pyramidal cells and can enhance NMDA- and AMPA-mediated currents. Dopamine D4 is the most prominent D2-like receptor in the hippocampus. D4 receptors are densely expressed on pyramidal neurons in CA1, on granule cells in the dentate gyrus, and on GABA interneurons. D4 receptors are coupled to $\text{G}_{i/o}$ and decrease NMDA-mediated excitatory transmission.

Norepinephrine innervation of the hippocampus arises primarily from the locus ceruleus, and the hippocampus contains one of the highest densities

of norepinephrine-containing terminals in the brain. Receptors for norepinephrine (adrenergic receptors) have been divided into two classes: α and β , which are further subdivided based on the specific G-proteins to which they are coupled. β -adrenergic receptors are coupled to G_s , while α_1 types are coupled to G_q and α_2 types to G_i . Electrical stimulation of the locus ceruleus produces an inhibition of hippocampal pyramidal neuron firing, followed by a transient increase in firing. The inhibition appears to be mediated by α_2 and the excitation by β receptors.

Serotonin (5HT) input to the hippocampus derives almost exclusively from the median raphe nucleus. Serotonin receptors in the hippocampus are preferentially expressed on calbindin-positive GABA interneurons, and the specific subtypes expressed include 5HT_{1A}, which is coupled to $G_{i/o}$; 5HT_{2A}, which is coupled to G_q ; and 5HT₃. The 5HT₃ receptor is an ionotropic receptor, the activation of which opens a nonspecific cation channel, resulting in fast depolarization. The activation of 5HT₃ receptors can thus inhibit the firing of pyramidal neurons by a GABA interneuron-mediated mechanism. There are also autoreceptors of the 5HT_{1B/D} subtype on serotonergic terminals that inhibit the release of serotonin. Intracellular recording studies show that the most prominent effect of serotonin application to the hippocampus is hyperpolarization of pyramidal neurons caused by an increase in K^+ conductance.

Histamine projections arise from the tuberomammillary nucleus of the hypothalamus. Histamine H₂ receptors are dense in CA1 and CA3 and are primarily located on pyramidal cell dendrites, but are also found on granule cells in the dentate gyrus. H₂ receptors are coupled positively to adenylyl cyclase, and the activation of these receptors causes strong excitation. Histamine H₁ receptors are mostly found in CA3, where their activation causes hyperpolarization. The H₁ receptor is coupled to stimulation of phospholipase C via $G_{q/11}$. H₃ receptors are also inhibitory (coupled to $G_{i/o}$) and are located on perforant path terminals in the dentate gyrus, where their activation can suppress glutamate secretion.

Other Neuromodulators

The hippocampus is rich in other neuromodulators, including neuropeptides, growth factors, cytokines, and endocannabinoids as well as their receptors. Neuropeptides are typically coreleased with classical neurotransmitters. Neuropeptides expressed in hippocampal neurons include cholecystokinin, somatostatin, neuropeptide Y, leptin, galanin, corticotropin releasing hormone, and the opioid peptides β -endorphin, enkephalin, and dynorphin. Receptors for the

gonadal and adrenal steroid hormones and thyroid hormone are also expressed in the hippocampus. Growth factors that are abundant in the hippocampus include brain-derived growth factor (BDGF), nerve growth factor (NGF) and fibroblast growth factor (FGF). Growth factors perform neurotrophic and neuroprotective functions and are critically involved in development and neuroplasticity. Cytokines regulate intercellular communication, including glial-to-neuronal and immune-to-neuronal communication. Cytokine ligands and their receptors found in the hippocampus include interleukin-1, tumor necrosis factor, and fractalkine. Eicosanoids are metabolites of arachidonic acid, which is derived from dietary linoleic acid. Several of the eicosanoids, including anandamide (arachidonylethanolamine) and arachidonylglycerol, are known as endocannabinoids because they act on the cannabinoid receptor CB1. High levels of CB1 expression are found throughout the hippocampus.

Neuroplasticity

Long-term potentiation (LTP) reflects a cellular model of learning and memory, first described in rabbit hippocampus in 1973. LTP is induced by applying a brief high-frequency electrical stimulus to an excitatory synaptic pathway, such as the hippocampal Schaffer collateral pathway, the perforant pathway, or the mossy fiber pathway (Figure 5b). For periods of hours (*in vitro*) to weeks (*in vivo*) after this conditioning stimulus, a test stimulus can subsequently provoke an increased synaptic response that is 50–200% above baseline. In the Schaffer collateral–CA1 synapse and the perforant path–dentate gyrus synapse, LTP is dependent on the activation of the NMDA receptor and subsequent Ca^{2+} influx (associative LTP), whereas in the mossy fiber–CA3 synapse, LTP appears to be dependent on kainate receptors (nonassociative LTP).

LTP has four basic characteristics that resemble features of learning and memory:

1. Temporal specificity: the presynaptic cell must fire before the postsynaptic cell.
2. Cooperativity: many synapses must be active to induce LTP.
3. Associativity: a strong input can induce potentiation at a weakly activated, adjacent synapse on the same postsynaptic cell.
4. Input specificity: potentiation is only induced at the synapses that received the conditioning stimulation.

In contrast to the potentiation caused by brief high-frequency stimulation, prolonged low-frequency stimulation of a pathway induces a long-term depression

(LTD) of the synaptic response to a test stimulus. Like the typical LTP found in most pathways (but not the mossy fiber pathway), LTD is also dependent on the activation of the NMDA receptor. It is thought that different patterns of Ca^{2+} influx produced during LTP and LTD induction determine the differential responses to these stimuli; high-frequency stimulation results in a large Ca^{2+} elevation and activation of the Ca^{2+} /calmodulin-dependent protein kinase (CaMK) cascade, whereas the smaller but more prolonged Ca^{2+} elevation produced by low-frequency stimulation leads to activation of a phosphatase cascade.

Hippocampal Function

The hippocampal formation appears to play a uniquely pivotal role in the brain's ability to form rapid and long-lasting associations between environmental stimuli in a way that allows for the learning of new concrete and abstract factual information (declarative memory) and detailed recall and recognition of events and places (episodic memory). The hippocampus in rodents appears to have an especially important role in learning the spatial configuration of places. Many CA1 and CA3 pyramidal cells behave as place cells; on initial exposure to a new environment, these cells acquire within several minutes the ability to increase their neuronal firing rate whenever the rat returns to a particular place in that environment.

The role of the hippocampus in stress response is less clear. As part of the limbic system, the involvement of the hippocampus in the generation of a stress state has often been presumed, but its anatomical neighbor, the amygdala, appears to play a much more important role than the hippocampus in this process. The cells of the hippocampus, however, appear to be especially sensitive to the effects of various stressors, perhaps in part due to their high expression of adrenal steroid receptors. Although the hippocampus may not be directly involved in the generation of a stress state, in the rat its more ventral regions contribute regulatory influences on hypothalamic-pituitary-adrenal (HPA) axis activity. Whether the primate hippocampus also has direct modulatory effects on the HPA axis remains to be determined.

See Also the Following Articles

Glucocorticoids – Adverse Effects on the Nervous System; Hippocampus, Corticosteroid Effects on; Hippocampus, Overview; Learning and Memory, Effects of Stress on; Memory and Stress; Neurogenesis; Steroid Hormone Receptors; Glucocorticoid Effects on Memory: the Positive and Negative.

Further Reading

- Amaral, D. G. and Witter, M. P. (1989). The three-dimensional organization of the hippocampal formation: a review of anatomical data. *Neuroscience* 31, 571–591.
- Andersen, P. (1975). Organization of hippocampal neurons and their interconnections. In: Isaacson, R. L. & Pribram, K. H. (eds.) *The hippocampus Vol. 1: Structure and development*. New York: Plenum Press.
- Andersen, P., Bliss, T. V. P. and Skrede, K. (1971). Lamellar organization of hippocampal excitatory pathways. *Experimental Brain Research* 13, 222–238.
- Bland, B. H. and Oddie, S. D. (2001). Theta band oscillation and synchrony in the hippocampal formation and associated structures: the case for its role in sensorimotor integration. *Behavioral Brain Research* 127, 119–136.
- Bliss, T. V. and Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *Journal of Physiology* (London) 232, 331–356.
- Eichenbaum, H. and Otto, T. (1992). The hippocampus – what does it do? *Behavioral and Neural Biology* 57, 2–36.
- Freund, T. F. and Buzsaki, G. (1996). Interneurons of the hippocampus. *Hippocampus* 6, 347–470.
- Joels, M. (2001). Corticosteroid actions in the hippocampus. *Journal of Neuroendocrinology* 13, 657–669.
- Gould, E., Woolley, C. S., Frankfurt, M., et al. (1990). Gonadal steroids regulate dendritic spine density in hippocampal pyramidal cells in adulthood. *Journal of Neuroscience* 10, 1286–1291.
- Johnston, D. and Amaral, D. G. (2004). Hippocampus. In: Shepherd, G. M. (ed.) *The synaptic organization of the brain* (5th edn., pp. 455–498). Oxford: Oxford University Press.
- Malenka, R. C. and Bear, M. F. (2004). LTP and LTD: an embarrassment of riches. *Neuron* 44, 5–21.
- O'Keefe, J. and Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Oxford University Press.
- O'Reilly, R. C. and Rudy, J. W. (2001). Conjunctive representations in learning and memory: principles of cortical and hippocampal function. *Psychological Review* 108, 311–345.
- Ramón y Cajal, S. (1968). The structure of Ammon's horn. Original publication in Spanish, 1893. Kraft, L. M. (trans.). Springfield, IL: Charles C. Thomas.
- Risold, P. Y. and Swanson, L. W. (1996). Structural evidence for functional domains in the rat hippocampus. *Science* 272, 1484–1486.
- Squire, L. R. and Kandel, E. R. (2000). *Memory, from mind to molecules*. New York: Scientific American Library.
- Swanson, L. W. (2004). *Brain maps: structure of the rat brain* (3rd edn.). Amsterdam: Elsevier Academic Press.
- Witter, M. P. and Amaral, D. G. (2004). Hippocampal formation. In: Paxinos, G. (ed.) *The rat nervous system* (3rd edn., pp. 635–704). Amsterdam: Elsevier Academic Press.

Hippocampus, Corticosteroid Effects on

M Joëls and H Karst

University of Amsterdam, Amsterdam, Netherlands

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M Joëls and H Karst, volume 2, pp 372–378, © 2000, Elsevier inc.

Patch clamp recording

electrical activity, (i.e., the neurotransmitter response).

A method in which a very tight seal between an electrode and the cell membrane is established, allowing the recording of currents through single ion channels. If the membrane patch under the electrode is disrupted, the recording of potentials or currents over the whole cell is possible.

Mechanism of Corticosteroid Action

Targets for Corticosteroid Action

Changes in Network Properties

Cellular Effects after Chronic Hypo- or Hypercorticism

Glossary

<i>Extracellular recording</i>	A method for recording of electrical activity on the outside of neurons; often applied <i>in vivo</i> when recording the collective response of a group of neurons to the stimulation of afferents (i.e., the field response).
<i>Hippocampal slice</i>	A transversal slice of the hippocampal formation (~350 μm thick) kept alive <i>ex vivo</i> , in which the main synaptic pathways are still intact.
<i>Intracellular recording</i>	A method for recording of electrical activity in which the cell is impaled with a sharp electrode; generally used to record the potential or resistance over the entire membrane and changes thereof in the presence of neurotransmitters.
<i>Ion currents</i>	Ion fluxes over the neuronal membrane through specific ion channels. The activation of many of these channels depends on the membrane potential. In addition, some channels become inactive in a voltage-dependent manner.
<i>Long-term potentiation (LTP)</i>	Prolonged strengthening of synaptic contacts by patterned stimulation of afferent fibers; thought to play an important role in memory processes.
<i>Long-term depression (LTD)</i>	Prolonged weakening of synaptic contacts by patterned stimulation of afferent fibers; thought to play an important role in memory processes.
<i>Neurotransmitter response</i>	Neurotransmitters change the electrical properties of neurons either through activation of ligand-gated ion channels or G-protein-coupled receptors, which, via the intermediate effects of G-proteins and/or second messengers, act on a specific type of ion channel. Electrophysiological methods record only the final change in

Corticosteroid hormones affect cellular and network properties in the hippocampus in a slow but persistent manner. The nature of the effects depends on the relative participation of mineralocorticoid and glucocorticoid receptors.

Mechanism of Corticosteroid Action

Corticosteroid hormones, which are secreted from the adrenal gland in particularly high amounts after stress, enter the brain and bind to specific receptors. Two main types of receptors have been recognized: the high-affinity mineralocorticoid receptor (MR) and the lower-affinity glucocorticoid receptor (GR). Both receptor types are expressed abundantly in the hippocampal formation. In fact, most principal neurons in the hippocampal subfields ([Figure 1a](#)) coexpress the two receptor types. This has raised the question of how the activation of the two corticosteroid receptor types affects cellular properties in the hippocampus. One of the premises of initial studies was that the activation of MRs may affect hippocampal cell properties in a different way than GR activation. This has proven to be correct.

For neurons that coexpress MRs and GRs, the relative participation of the two receptor types in cellular effects depends on several factors. First, the relative amounts of the two receptor types are important. Even within the hippocampus, the amounts vary among the different subfields. For instance, pyramidal neurons in the CA1 region and granule cells in the dentate gyrus (see [Figure 1a](#)) express high amounts of both MRs and GRs. Pyramidal neurons in the CA3 region, however, have relatively low amounts of the GR. Second, the affinity of the receptor for the natural ligand, in combination with the availability of the ligand, plays an important role. The bio-availability in the hippocampus is not only determined by adrenal release of corticosterone, but also by factors such as the passage over membranes in which p-glycoproteins play a major role as well as the presence of converting

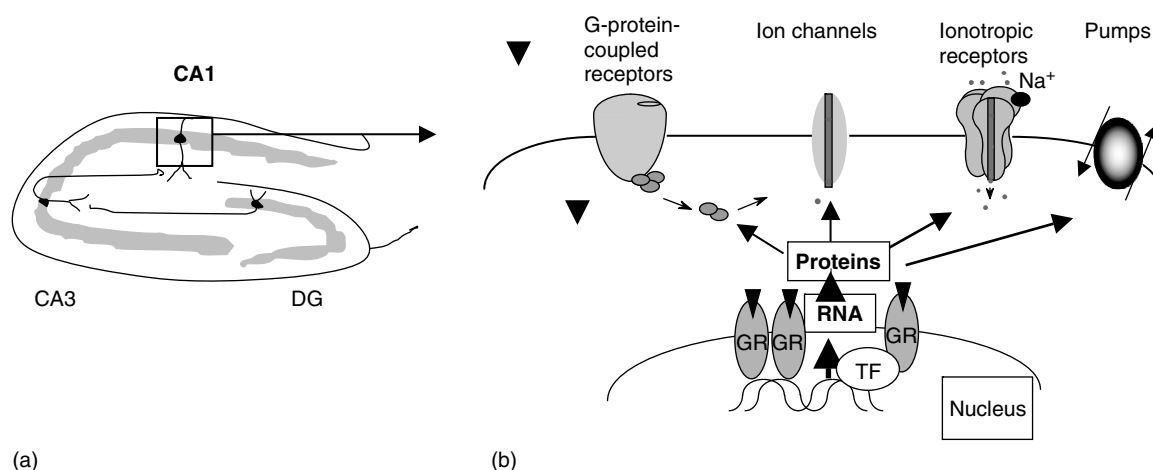


Figure 1 Corticosteroid effects on hippocampal cells. a, Representation of a hippocampal slice showing the main cellular subfields and the synaptic contacts between these subfields; b, Gene-mediated actions of corticosteroid hormones affecting the properties of hippocampal cells, aiming at various targets: ion channels, ionotropic receptors, G-protein-coupled receptors, and ion pumps. Intracellular corticosteroid receptors affect gene transcription either as homodimers binding directly to specific sequences in responsive genes or as monomers interacting with other transcription factors. CA1, cornus ammoni 1; CA3, cornus ammoni 3; DG, dentate gyrus; TF, transcription factor.

enzymes such as 11β -hydroxysteroid dehydrogenase. Low levels of the ligand activate MRs extensively due to their high affinity, whereas the lower-affinity GRs are only partly occupied. GRs are fully activated only when corticosteroid levels are very high. These two factors are not static entities but are prone to regulation, for example, during prolonged periods of exposure to stress.

Once corticosteroids are bound to their receptor, the steroid-receptor complex dissociates from attached proteins and translocates to the nuclear compartment. Rapid cellular effects of corticosteroids that are not gene-mediated have recently been described in hippocampus, but are not discussed here. Classically, it was thought that all gene-mediated corticosteroid actions, including those in the brain, involve binding of receptor homodimers to hormone response elements in the DNA (Figure 1b). Indeed, all functional effects studied so far in hippocampal cells appear to require DNA-binding of GR homodimers. However, molecular investigations – including microarray surveys in the hippocampus – indicate that effects can also take place via the interactions of steroid receptor monomers with other transcription factors. A third possible mode of action, via the heterodimerization of MR and GR molecules, has been suggested based on studies in transfected cells in which (among other things) protein–protein interactions were examined. It is still undecided whether MR–GR heterodimers play a major role in the development of corticosteroid actions in the hippocampus.

Over the past years, much more has become known about the genes that form a target for corticosteroid action in the hippocampus, as surveyed with serial

analysis of gene expression, microarrays, and single-cell RNA amplification of material from identified hippocampal cells. Thus, MR and GR activation each leads to altered expression of some 70–100 genes, of which roughly 50% was found to be upregulated and one-third were responsive to both receptor types. The hormone coordinates the expression of genes underlying aspects of energy metabolism, cell adhesion, synaptic transmission, oxidative stress, neurogenesis, cell cycle and structural plasticity. These transcriptional effects are likely to underlie cellular changes that have been observed, although the entire pathway has not yet been resolved for any of these. The functional data are in line with the molecular observations and indicate too that steroid-responsive genes may be involved in very diverse aspects of cellular function, such as cell structure and growth, energy metabolism, and membrane properties involved in signal transduction. Part of these cellular effects are summarized next.

Targets for Corticosteroid Action

Ion Channels

Under resting conditions, (passive) membrane properties of hippocampal CA1 and CA3 pyramidal cells are not affected by corticosteroid hormones. However, when the membrane potential of the cells is shifted from its resting level toward more depolarized or hyperpolarized levels (e.g., in response to a neurotransmitter), clear corticosteroid effects can be observed. These effects are generally slow in their onset and very persistent, which is in line with a gene-mediated mechanism of action. Patch clamp

and intracellular studies have shown that some, although not all, of the ionic conductances in hippocampal cells are modulated by MR or GR activation. This underlines that the hormone actions are pleiotropic but nevertheless show a distinct degree of selectivity.

Three types of ion conductances have been investigated with regard to corticosteroid sensitivity: potassium (K^+), sodium (Na^+), and calcium (Ca^{2+}). Of these, voltage-dependent Ca^{2+} conductances are clearly most sensitive to corticosteroid receptor activation. It was found that low levels of corticosteroids resulting in a predominant MR activation are associated with small Ca^{2+} currents. If activation of GRs occurs in addition to MRs (e.g., after stress exposure) this results in an increased amplitude of Ca^{2+} currents compared to the situation in which mostly MRs are occupied. Several types of Ca^{2+} channels contribute to the overall Ca^{2+} currents in hippocampal cells. The steroids affect at least L-type Ca^{2+} channels, which cause the sustained current seen with large depolarizations of the cell, such as happen during an action potential. Only the amplitude of the currents and not the kinetic properties or voltage dependency is altered by corticosteroid treatment. Interestingly, in the absence of corticosterone, Ca^{2+} current amplitudes are also large, pointing to a U-shaped corticosteroid dose-dependency (see Figure 2). These data indicate that MR activation is necessary to retain Ca^{2+} currents within certain limits (a situation that is preferable for a neuron), whereas GR activation in addition to MR activation increases the Ca^{2+} influx into the hippocampal cells.

Steroid modulation of Ca^{2+} influx indirectly affects the slow Ca^{2+} -dependent K^+ conductance, which is extremely sensitive to the intracellular Ca^{2+} concentration. This conductance is slowly activated when neurons are depolarized for a considerable period of time (>100 ms) and causes accommodation of the firing frequency during depolarization; because the channels are also slowly deactivated, a lingering after-hyperpolarization (AHP) of the membrane is seen after the depolarizing influence is terminated. Both accommodation and AHP suppress the transmission of excitatory input to the hippocampal cells. In agreement with the small Ca^{2+} influx seen with predominant activation of MRs, it was found that accommodation and AHP under those conditions are relatively small, which promotes the transmission of excitatory input. With additional GR activation, accommodation and AHP are increased, leading to a temporary suppression of excitatory signals. Thus, hippocampal CA1 neurons in which mostly MRs are activated effectively transmit excitatory signals without being exposed to large intracellular Ca^{2+} levels. The large Ca^{2+} influx seen with additional

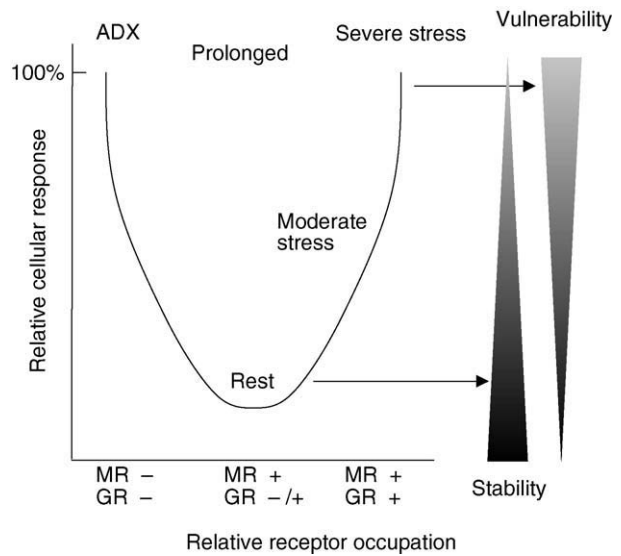


Figure 2 Many hippocampal cell properties, such as calcium currents, are large in the absence of adrenal steroids (adrenalectomy, ADX) or when both corticosteroid receptor types are extensively activated and are small with predominant MR activation. This results in a U-shaped corticosteroid dose-dependency, similar to what has been found for the potency to induce long-term depression versus long-term potentiation in the hippocampus. The relative MR and GR activation is indicated below the dose-response curve. Cellular responses associated with predominant MR occupation induce a situation of stability and resistance to degeneration. Additional GR activation (e.g., after a moderate stress) generally results in a temporary suppression of hippocampal cell activity, in agreement with the adaptive and normalizing role of this receptor in the stress response. The instability linked to GR activation can become deleterious when GR activation occurs concurrently with other challenging situations for the tissue, such as ischemia. Under these conditions, cellular vulnerability is increased if the relative GR to MR activation is high.

GR activation is normally restrained because Ca -dependent K^+ conductances prevent large-scale depolarization. However, when GR activation is associated with extensive depolarization of neurons, such as occurs during epileptic or ischemic insults, cells are at high risk of being exposed to very high Ca^{2+} levels. This may partly explain why hippocampal cells show a more profound loss of function and integrity after hypoxia-ischemia or epileptic seizures when corticosteroid levels are persistently elevated.

The K^+ and Na^+ conductances that were investigated appeared to be far less sensitive to corticosteroid receptor activation. Relatively minor effects were observed on the properties of the inwardly rectifying K^+ conductance I_h and on the fast, tetrodotoxin (TTX)-sensitive Na^+ conductance. Other K^+ conductances, such as the delayed rectifier or transient A current, showed no steroid dependency at all in the hippocampus. Although the effects on voltage-dependent Na^+ and K^+ conductances are

thus relatively small, they could still result in an altered shape of the action potential.

Neurotransmitter Responses

Based on chemical structure, three classes of neurotransmitters are recognized that affect the cellular function of hippocampal cells: amino acids, biogenic amines, and neuropeptides. Of these, the responses to biogenic amines appear to be most sensitive to corticosteroid receptor activation. Corticosteroid effects on amino acid-mediated transmission seem to be diverse and only partly gene-mediated. Neuropeptide actions have largely not been investigated in this respect.

The responses to serotonin (5-hydroxytryptamine, 5-HT) are strongly modulated by corticosteroid receptor activation. Serotonin can bind to several types of receptors on CA1 pyramidal neurons, including the 5-HT_{1A} receptor. The activation of this receptor leads, through G-protein-coupling, to the opening of K⁺ channels, which results in a hyperpolarization of the membrane. With intracellular recording, it was found that the 5-HT_{1A} receptor-mediated hyperpolarization is small under conditions of predominant MR activation. This may be due to a MR-dependent downregulation of 5-HT_{1A} receptor mRNA expression, as was observed in the dentate gyrus and in some studies also in the CA1 region. The activation of GRs in addition to MRs (e.g., after exposure to a physical stressor) induces large 5-HT_{1A} receptor-mediated responses. The molecular mechanism underlying these large responses is unclear because 5-HT_{1A} receptor expression and binding are not increased after GR activation. Potential targets for steroid modulation downstream of the 5-HT_{1A} receptor are probably also not modulated because steroid actions on the G-protein function or the K⁺ channel opened by the G-proteins cannot explain the changes in 5-HT_{1A} response. In addition to modulation of 5-HT_{1A} receptor-mediated responses, corticosteroids also affect responses or expression of 5-HT₄ and 5-HT₇ receptors.

Depolarizing responses mediated via muscarinic acetylcholine receptors appear to have a similar steroid dependency. Because both serotonergic and cholinergic responses are relatively large in the absence of corticosteroid hormones, there seems to be a U-shaped dose-dependency for the transmitter responses, as in the case of Ca²⁺ conductances (see [Figure 2](#)). This may signify the existence of a common denominator underlying all these steroid effects. It is possible that corticosteroid receptor activation results in the altered activity of kinases or phosphatases, which affect several membrane components in a comparable way through posttranslational modifications. This does not exclude the development of steroid

actions aiming at specific genes. This is supported by the observation that responses to norepinephrine, mediated by β 1-adrenoceptors, show an inverse linear steroid dose-dependency rather than the U-shape.

Whereas the steroid modulation of biogenic amine responses is slow in onset, persistent, and requires *de novo* protein synthesis, steroid effects on amino acid-mediated responses seem to have different features. This was mostly investigated by examining the steroid modulation of synaptically evoked responses in the CA1 hippocampal region. The stimulation of Schaffer/commissural fibers in this region leads to a subsequent activation of glutamatergic AMPA and NMDA receptors. It was found that the administration of very high amounts of corticosterone, resulting in concomitant MR and GR occupancy, suppresses field responses mediated by glutamate receptors. With these very high doses of corticosterone a rapid onset of effects is seen and the effects are very prominent when extracellular Ca²⁺ concentrations are high. Moreover, the steroid effects are particularly evident after the repeated stimulation of the afferent pathway. The data suggest that under these circumstances energy-dependent processes in the signal transduction and cellular integrity may be impaired. When glutamate transmission is recorded under more physiological conditions (i.e., comparable to the design used to study effects on ion currents and aminergic responses), a slow GR-mediated enhancement rather than impairment of the responsiveness to synaptic stimulation becomes apparent.

Inhibitory responses mediated by GABA_A receptors are affected only by extreme (and probably physiologically not relevant) doses of corticosterone. GABA_B receptor-mediated responses were found to be stable with predominant MR activation, but declined rapidly with repeated stimulation both in the absence of steroids and in the presence of very high amounts of corticosterone.

Changes in Network Properties

The corticosteroid effects on the properties of individual hippocampal cells or small groups of cells clearly alter the function of hippocampal networks in which these cells participate. In addition, corticosteroid hormones could change network properties that depend critically on the connections and associative contacts between cells in a network.

Examples of processes that depend both on the characteristics of individual hippocampal cells and on the connections between cells are LTP and LTD. The former refers to the situation in which the high-frequency stimulation of afferents results in a prolonged strengthening of synaptic contacts. The fact

that LTP is induced rapidly in hippocampal subfields, persists for hours to days, and involves associativity of inputs has led to the hypothesis that LTP may be the neurobiological substrate for learning and memory, although a direct link is still disputed. Stimulating hippocampal pathways at a lower frequency can lead to LTD, the weakening of synaptic contacts. The processes of LTP and LTD each require specific intracellular signaling cascades, phosphorylation targets, and Ca^{2+} levels. In nonstressed animals, LTP can be induced easily with high-frequency stimulation, whereas homosynaptic LTD is relatively hard to obtain.

Many studies over the past decade have demonstrated that MR and GR activation largely affect the likelihood to evoke LTP and LTD. Both in the dentate gyrus and in the CA1 region, conditions that result in the near-complete activation of MRs and only partial activation of GRs are associated with a very efficient induction and maintenance of LTP. When GRs are activated extensively, it becomes hard to evoke LTP, whereas LTD induction is promoted. In the absence of corticosteroid hormones, LTP induction is also suppressed. The degree of potentiation therefore shows an inverted U-shaped steroid dose-dependency, which points to an underlying process that may be common to and perhaps evolve from the steroid actions on Ca^{2+} influx and responses to neurotransmitters such as serotonin.

These observations, generally obtained by administering exogenous corticosterone or selective agonists or antagonists for the corticosteroid receptors, can be extended to behaviorally relevant situations. Thus, several studies have shown that the exposure of animals to stressors that are unpredictable and out of the context of a learning experience suppresses the likelihood of inducing LTP and instead allows the induction of LTD. Even less extreme situations, such as the subjection of animals to a novel environment out of the context of a learning experience, led to a similar suppression of LTP and promotion of LTD. Stress effects on LTP and LTD are sometimes quite rapid in onset and not always associated with extensive changes in GR activation. This indicates that a shift in the likelihood of evoking LTP depends not only on the actual MR–GR occupation but also on other factors, such as the recent history of the postsynaptic cells, which may be influenced by other stress hormones as well, such as norepinephrine and corticotropin releasing hormone (CRH). In the intact organism, the effects of corticosteroids on synaptic plasticity are also determined by the degree to which multiple brain regions, such as the prefrontal cortex and amygdala nuclei, are involved, which in turn depends on the nature of the stressor.

Cellular Effects after Chronic Hypo- or Hypercorticism

The corticosteroid effects on cellular properties and network function described previously, are generally observed over a period of several hours after a change in MR/GR occupation. It is therefore very likely (and this has indeed been demonstrated in several cases) that these effects on hippocampal function take place shortly after exposure to an acute stressful situation. Increased corticosterone levels after stress, resulting in GR activation on top of already activated MRs, evokes an array of effects, most of which temporarily suppress the transmittal of excitatory signals (see Table 1). These GR-mediated events may help restore hippocampal function following stress, which is in accordance with the normalizing and adaptive mode by which GRs participate in the stress response.

Although large variations in MR activation probably do not take place under physiological conditions, data nevertheless show that MRs play an important role in hippocampal function. First, MRs seem to be important for the full development of GR-mediated events (and vice versa). Second, MR activation is essential for maintaining a basal level of hippocampal activity (see Table 1). And MR activation is even a prerequisite for neuronal viability. This is most evident in granule cells of the dentate gyrus. Dentate granule cells proliferate from progenitor cells even in adulthood and after some time die through apoptosis. There is a continuous balance between the incorporation of newly formed cells and the removal of old cells. When rats are adrenalectomized, this balance is disturbed, so that 3 or more days after adrenalectomy proliferation is increased but also more cells

Table 1 Cellular effects observed with predominant MR activation or with concurrent MR and GR activation^a

	Excitability ↑	Excitability ↓
MR activation	Accommodation/AHP ↓ 5-HT _{1A} receptor response ↓ Glutamatergic response ↑ LTP ↑	Carbachol responses ↓
MR + GR activation	GABA _B response ↓ Glutamate release (?) ↑ Carbachol responses ↑	Accommodation/AHP ↑ 5-HT _{1A} receptor response ↑ β-Adrenergic response ↓ Glutamate response ↓ LTD ↑, LTP ↓

^aAHP, after-hyperpolarization; GR, glucocorticoid receptor; LTD, long-term depression; LTP, long-term potentiation; MR, mineralocorticoid receptor.

undergo apoptosis. In adrenalectomized rats undergoing apoptosis, the synaptic transmission as well as the plasticity between the entorhinal cortex and the dentate gyrus is impaired. The functional impairment most likely occurs prior to and independent of the change in cell turnover and may in fact form a trigger for an overall synaptic reorganization in the dentate. The impaired synaptic function, as well as the changes in cell turnover, can be prevented by treating animals with MR ligands. The fate of an individual cell also appears to be determined by its specific gene expression pattern. The data indicate that relatively young dentate cells can increase the expression of proteins that will help them survive, whereas older cells fail to do so, explaining why destabilizing factors such as enhanced calcium influx or impaired synaptic communication, seen in the absence of corticosterone, can become deleterious. Collectively, normal MR function in the dentate gyrus seems to be important in preserving the delicate balance between viability and cell death.

Although prolonged deprivation of corticosteroids is a threat for dentate granule cell survival, chronic overexposure of the hippocampus to corticosteroid hormones is also potentially harmful. Exposure of rats to 3 weeks of restraint stress causes atrophy of apical dendrites in CA3 pyramidal neurons. Under these conditions, both the presynaptic nerve terminals and the postsynaptic spine density in the CA3 area are altered. Chronic stress also decreases proliferation in the dentate gyrus and in fact may slow down the entire cell cycle for a considerable period of time, possibly by changing the expression of endogenous factors involved in the cell cycle. In addition to (or perhaps related to) the morphological effects, other hippocampal properties are also altered after chronic stress. Glutamate transmission is enhanced both in the CA3 area and dentate gyrus. Conversely, LTP is markedly impaired in all hippocampal subareas. Interestingly, 5-HT_{1A} receptor-mediated responses become gradually attenuated after chronic stress or prolonged overexposure to corticosterone. This may be one explanation for the increased risk of mood disturbances after prolonged hypercortisolemia.

The mechanisms underlying the changes in hippocampal morphology and cell turnover after chronic stress exposure are largely unknown. The observed enhancement in glutamate transmission may be a causative factor, but clearly sequential regulation of many gene targets, including genes encoding for corticosteroid receptors and chaperones, enzymes, structural proteins, cell cycle factors, and growth factors, in concert determines the outcome of the process. Understanding this sequence of events will

be one of the challenges for the future. It is already clear that the cellular actions induced by repeated overexposure to corticosterone are not a simple extrapolation of the effects seen after an acute rise in corticosteroid level. Also, in the hippocampus there is no evidence for steroid resistance – generally effects are strongest when chronic stress is combined with an acute rise in the level of corticosterone. Insight into these fundamental processes as well as the potential to intervene with pharmacological treatment (e.g., by applying antiglucocorticoids) will be essential for novel treatment strategies of stress-related disorders involving hippocampal dysfunction.

See Also the Following Articles

Corticosteroid Receptors; Hippocampal Neurons; Hippocampus, Overview; Neurogenesis.

Further Reading

- Datson, N. A., van der Perk, J., de Kloet, E. R., et al. (2001). Identification of corticosteroid-responsive genes in rat hippocampus using serial analysis of gene expression. *European Journal of Neuroscience* **14**, 675–689.
- De Kloet, E. R., Joëls, M. and Holsboer, F. (2005). Stress and the brain: from adaptation to disease. *Nature Reviews Neuroscience* **6**, 463–475.
- Gould, E. and Tanapat, P. (1999). Stress and hippocampal neurogenesis. *Biological Psychiatry* **46**, 1472–1479.
- Karst, H., Karten, Y. J., Reichardt, H. M., et al. (2000). Corticosteroid actions in hippocampus require DNA binding of glucocorticoid receptor homodimers. *Nature Neuroscience* **3**, 977–978.
- Kim, J. J. and Diamond, D. M. (2002). The stressed hippocampus, synaptic plasticity and lost memories. *Nature Reviews Neuroscience* **3**, 453–462.
- Kole, M. H., Swan, L. and Fuchs, E. (2002). The antidepressant tianeptine persistently modulates glutamate receptor currents of the hippocampal CA3 commissural associational synapse in chronically stressed rats. *European Journal of Neuroscience* **16**, 807–816.
- Korz, V. and Frey, J. U. (2003). Stress-related modulation of hippocampal long-term potentiation in rats: involvement of adrenal steroid receptors. *Journal of Neuroscience* **23**, 7281–7287.
- Maroun, M. and Richter-Levin, G. (2003). Exposure to acute stress blocks the induction of long-term potentiation of the amygdala-prefrontal cortex pathway *in vivo*. *Journal of Neuroscience* **23**, 4406–4409.
- McEwen, B. S. (1999). Stress and hippocampal plasticity. *Annual Review of Neuroscience* **22**, 105–122.
- Shakesby, A. C., Anwyl, R. and Rowan, M. J. (2002). Overcoming the effects of stress on synaptic plasticity in the intact hippocampus: rapid actions of serotonergic and antidepressant agents. *Journal of Neuroscience* **22**, 3638–3644.

Hippocampus, Overview

M J Meaney and S J Lupien

McGill University, Montreal, Quebec, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition article by M J Meaney and S J Lupien, volume 2, pp 379–385, © 2000, Elsevier Inc.

Hippocampal Regulation of Hypothalamic-Pituitary-Adrenal (HPA) Activity

Glucocorticoid Regulation of Hippocampal Activity

Chronic Stress and Hippocampal Function

Glucocorticoids and Hippocampal Function

Conclusions

Glossary

Feedback inhibition

A process whereby the end product of the system is able to inhibit the activity of higher order, stimulating events. In the case of the hypothalamic-pituitary-adrenal axis, the adrenal glucocorticoids inhibit corticotropin-releasing factor synthesis in hypothalamic neurons.

Plasticity

The ability of neurons to establish new contacts with other neurons through the formation of synaptic connections. The process can also refer to the strengthening of established connections.

Hippocampal Regulation of Hypothalamic-Pituitary-Adrenal (HPA) Activity

Under most conditions, the hypothalamic-pituitary-adrenal (HPA) axis lies under the dominion of specific releasing factors secreted by neurons located in paraventricular nucleus (PVN) of the hypothalamus, notably corticotropin-releasing factor (CRF) and arginine vasopressin (AVP). Neurons of the PVN project extensively to the portal capillary zone of the median eminence, furnishing a potent excitatory signal for the synthesis and release of adrenocorticotrophic hormone (ACTH) from the anterior pituitary, and thus the means by which neural activity associated with the stressor is transduced into endocrine signals. Elevated ACTH levels, in turn, increase the synthesis and release of glucocorticoids from the adrenal (in primates the primary glucocorticoid is cortisol, whereas in rodents it takes the form of corticosterone).

The responsivity of the HPA axis to stress is in part determined by the ability of the glucocorticoids to regulate ACTH release (i.e., glucocorticoid negative feedback). Circulating glucocorticoids act on pituitary and specific brain regions to inhibit the synthesis and secretion of releasing factors from hypothalamic neurons and pituitary ACTH. In addition to pituitary and hypothalamic sites, there is now considerable evidence of the importance of the limbic system, the hippocampus, and the front cortex in the regulation of HPA activity.

Circulating glucocorticoids bind with high affinity to two corticosteroid receptor subtypes: the mineralocorticoid receptors (MRs) and the glucocorticoid receptors (GRs). Both receptor subtypes have been implicated in mediating glucocorticoid feedback effects. Interestingly, ligand-activated MRs and GRs act as a common DNA-binding site (the corticosteroid receptor response element), but the receptors differ in their ability to interact with intracellular proteins (i.e., protein-protein actions).

Lesions to the hippocampus in the rat or the primate result in elevated corticosterone levels under basal and poststress conditions, whereas glucocorticoid implants into the hippocampus normalize ACTH levels in adrenalectomized rats. Hippocampectomized animals show a reduced suppression of HPA activity following exogenous glucocorticoid administration. Finally, hippocampectomy results in increased CRH mRNA levels in the parvocellular region of the PVN (a region that contains neurons whose axons terminate in the median eminence and increased portal concentrations of ACTH secretagogues). The negative feedback effects of glucocorticoids at the level of the hippocampus appear to involve both corticosteroid receptor subtypes. Implants of RU28318, an antagonist selective for the mineralocorticoid sites, and/or RU38486, an antagonist selective for the glucocorticoid site, increase basal and poststress HPA activity (i.e., antagonize endogenous glucocorticoid negative feedback). Levels of GR occupancy in the hippocampus and hypothalamus are correlated negatively with portal concentrations of CRH and AVP. Finally, both MR and GR agonists with selective antagonists underscore the state-dependent influence of these receptors. MR antagonists implanted into the hippocampus elevate basal as well as stress-induced HPA activity, whereas GR antagonists affect stress-induced activity. In humans, both mineralocorticoid and GR antagonists have been found to increase plasma ACTH levels, and, as in the rat, the effects of MR and GR blockade

depend on the time of day. These roles are consistent with corticosteroid receptor pharmacology, as modest basal steroid levels serve to occupy largely MRs. Dynamic variations in GR levels occur only under conditions of elevated steroid levels, such as with the circadian peak or during periods of stress.

Naturally occurring and experimentally induced variations in hippocampal corticosteroid receptor levels produce effects of HPA activity that are consistent with this formulation. In the rat, postnatal handling increases hippocampal GR but not MR expression, and the effect is largely unique to the hippocampus. Predictably, animals handled in early life show dampened stress-induced HPA activity as adults, whereas there is no effect on basal activity. Likewise, increased maternal licking/grooming during early life increases hippocampal GR expression, resulting in enhanced glucocorticoid negative feedback sensitivity, decreased hypothalamic CRF mRNA expression, and more modest HPA responses to stress. Lewis rats exhibit increased hippocampal MR levels and show lower basal and stress-induced activity in comparison to the parent Wistar strain. Thus, both MR and GR activation can potentially alter both basal and stress-induced HPA activity; under normal circumstances, the low corticosteroid levels prevailing under basal conditions selectively activate MRs. GR activation normally occurs during the peak in the circadian cycle or during periods of stress, and it is at this time that the GR influence on HPA activity is apparent.

Glucocorticoid Regulation of Hippocampal Activity

Glucocorticoids have been known for some time to regulate the excitability of hippocampal neurons, and more recent studies using models of long-term potentiation (LTP) have helped clarify the relevant mechanisms of action. LTP can be induced by the brief tetanic stimulation of glutaminergic afferents to either the CA1 or dentate gyrus region of the hippocampus and is defined by the enhanced synaptic response to subsequent stimulation. This LTP is thought to model the processes of synaptic plasticity that mediate learning. The magnitude of the hippocampal LTP response in the rat is decreased with adrenalectomy, and this effect is reversed with low levels of corticosterone replacement that affect largely MR occupancy. However, higher levels of corticosterone, which serve to occupy both MR and GR sites, reverse this effect, and with increasing GR activation, there is an even further suppression of LTP well beyond the effect of ADX, resulting in an inverted U-shaped relationship between circulating glucocorticoid levels and

hippocampal LTP. The mechanism underlying this effect involves the corticosteroid regulation of intracellular Ca concentrations. Following a prolonged sequence of stimulation, hippocampal neurons exhibit a slow, enduring afterhyperpolarization (AHP) that is mediated by a slow Ca-dependent K conductance, serving to decrease hippocampal excitability. Ca influx through voltage-gated channels, and thus the strength of AHP, is modest under conditions of MR occupancy and is increased substantially with the occupation of GRs. Elevated corticosterone levels increase mRNA levels for several Ca channels, whereas MR activation has the opposite effect.

There are also relevant corticosteroid effects on neurotransmitter systems. Glutaminergic transmission drives hippocampal LTP and is enhanced under conditions of MR activation and suppressed with increasing GR activation. Norepinephrine acts via cAMP-coupled β -adrenoreceptors to attenuate the magnitude of the AHP and thus enhance hippocampal excitability, an effect that is diminished under conditions of elevated corticosterone levels, which activate GRs. Serotonin (5-HT) acts via 5-HT_{1A} receptors to produce a hyperpolarization and thus a decrease in hippocampal excitability. This effect is enhanced through GR activation. Finally, GR activation can also serve to metabolically compromise the function of hippocampal neurons through effects on glucose transport.

These findings are consistent with the idea that hippocampal excitability and LTP formation are facilitated under conditions of low basal corticosterone levels that result in MR activation. As corticosterone levels rise, such as under conditions of stress, there is greater GR activation, which serves to increase AHP, dampen glutaminergic transmission, attenuate noradrenergic facilitation of excitability, and increase serotonergic inhibition. Taken together, these findings are consistent with the inverted U relationship between circulating corticosterone levels and hippocampal LTP. Importantly, steroid pharmacology is also consistent with the findings that stress impairs hippocampal LTP formation.

Chronic Stress and Hippocampal Function

Exposure to chronically elevated glucocorticoid levels further compromises hippocampal function to the point of endangering neuronal integrity and survival. The story again focuses, at least in part, on corticosteroid effects on Ca homeostasis. Chronic stress in rodents and primates produces a reversible regression of apical dendrites. The effect is mimicked by high

doses of corticosterone and is blocked by corticosterone synthesis inhibition or by Ca channel blockers. Interestingly, these effects parallel changes in the expression of a hippocampal-derived neurotrophic factor, brain-derived neurotrophic factor (BDNF). Chronic stress or chronically elevated corticosterone levels decrease BDNF mRNA expression as well as that for *trkB*, the BDNF receptor.

Elevated glucocorticoid levels compromise not only the maintenance of the existing synaptic form, but also synaptic sprouting. This finding is consistent with the GR-mediated effects on LTP, which appear to involve synaptic plasticity. Partial deafferentation in selected regions of the adult brain results in a sequence of compensatory events involving both the remaining afferent fibers and denervated dendrites, culminating in a reorganization of the circuitry in the denervated area. These events include sprouting of the remaining afferent fibers, restoration of spine density and length, and replacement of vacated synaptic contacts. One of the best examples of such reactive synaptogenesis occurs in the dentate gyrus of the hippocampus in animals following entorhinal cortex lesions. Entorhinal cortex lesions disrupt the perforant path, resulting in nearly 60% loss of synaptic input to the granule cells of the dentate gyrus. However, beginning as early as a few days after the lesion, new synapses are formed, and within 2 months virtually all the lost inputs are replaced. Following fimbria-fornix lesions, synaptic replacement occurs as a result of the growth of fibers from the adrenergic neurons in the superior cervical ganglion into the denervated region. In the case of entorhinal damage, the new synapses originate from the cholinergic cells of the medial septum, from glutaminergic commissural-associational pyramidal cells, and, to a lesser extent, from neurons of the contralateral entorhinal cortex.

Glucocorticoids suppress hippocampal sprouting in young animals in a dose-related manner. High glucocorticoid levels reduce hippocampal sprouting significantly following lesions of either the fimbria-fornix or the entorhinal cortex. This effect appears to be mediated by glucocorticoid effects on the expression of genes involved in synaptic repair, such as cellular receptor systems involved in cholesterol scavenging and microtubule-associated proteins.

The degree of glucocorticoid exposure appears to be a determinant of hippocampal pathology after neurological insult, as well as during the aging process. In humans, poststroke damage is correlated positively with circulating cortisol levels. In brief, Sapolsky's studies have shown that glucocorticoids endanger hippocampal neurons by (1) limiting the availability of energy substrates, thus rendering neurons metabolically compromised, and (2) increasing

extracellular glutamate concentrations by reducing glutamate uptake by glial cells. Both corticosteroid actions are GR mediated. These effects appear to be most apparent under conditions of insult associated with hyperexcitation and metabolic challenge. In contrast, MR activation appears to be neuroprotective under these same conditions.

Data from the laboratories of Sloviter and McEwen have illustrated another dimension to the corticosteroid regulation of neuronal survival in the hippocampus. In these studies, long-term adrenalectomized animals showed a profound decrease in neuron density in the dentate gyrus, a hippocampal subregion. In this case, the absence of glucocorticoid stimulation in adult animals resulted in neuron loss. Interestingly, the effects were observed uniquely in the granule cells of the dentate gyrus. Pyramidal cell density in Ammon's horn (CA1–CA4) was completely unaffected. In contrast, chronic exposure to elevated glucocorticoid levels results in the loss of pyramidal neurons in Ammon's horn, whereas the dentate is unaffected. Landfield's group reported that long-term adrenalectomized animals maintained under conditions of low corticosterone replacement showed no changes in neuron density in the dentate gyrus. The key here is the low level of corticosterone replacement. McEwen's group found that low levels of corticosterone or aldosterone replacement are sufficient to block the effects of adrenalectomy on neuron density in the dentate gyrus; both treatments selectively target the hippocampal MR. In contrast, Sapolsky's group showed that the loss of pyramidal neurons in Ammon's horn in the presence of elevated glucocorticoid levels is mediated by the GR. Thus, glucocorticoid potentiation of the neurotoxic effects of excitatory amino acids on hippocampal neurons is mimicked by RU28362, a GR agonist, and blocked by RU38486. In an interesting parallel, Meaney and Lupien showed that during the early postnatal period, an active period of hippocampal development, there are high, adult-like levels of MRs and very low levels of GRs. In essence, low levels of corticosteroid stimulation are necessary for the maintenance of hippocampal granule neurons, and chronically elevated glucocorticoid levels are hazardous for the integrity and survival of pyramidal neurons. The ideal state is one of tight titration around low basal glucocorticoid levels, an important argument for efficient glucocorticoid negative feedback.

Glucocorticoids and Hippocampal Function

Corticosteroid treatment has been known for some time to induce a reversible psychotic condition

(so-called steroid psychosis) in certain individuals. This condition also includes logical thinking. A dementia-like syndrome (including attentional deficits and memory impairments) occurs frequently in a group of patients given high doses of corticosteroids for disorders not related to the central nervous system and who did not show evidence of psychosis. These findings are consistent with studies of Cushing's patients, in which cortisol levels are elevated profoundly. Cushing's patients consistently show attentional deficits and memory impairments, and the magnitude of the impairments correlates with plasma cortisol levels. Interestingly, the older subjects (>45 years) were affected more seriously; these subjects showed significantly greater memory impairments than younger subjects. Importantly, these deficits were reversible; surgical intervention to correct the hypercortisolemic state resulted in normalized cortisol levels and improved cognitive performance.

In addition, cognitive impairments, including attentional deficits and verbal and visual memory impairments, are often associated with clinical depression and are correlated with increased cortisol levels. Again, both the increase in plasma cortisol and the magnitude of the cognitive disturbances are more severe in elderly patients, often resulting in a condition of so-called pseudodementia. Importantly, in hypercortisolemic patients (especially dexamethasone-resistant patients) the cognitive disturbances are more pronounced. Results showed that subjects who received hydrocortisone treatment presented an impaired performance in the declarative memory task but not in the nondeclarative memory task, thus suggesting that cortisol interacts with hippocampal neurons to induce cognitive deficits. Another study reported comparable deficits in declarative memory in response to higher cortisol levels associated with major stress.

Depression is also associated with hippocampal atrophy, and is commonly associated with increased cerebrospinal levels of CRF and adrenal hypertrophy, suggesting a degree of glucocorticoid feedback insensitivity. Considering the evidence for hippocampal atrophy in this propulsion, it is tempting to speculate on the potential involvement of the hippocampus dysfunction as a source of this feedback insensitivity and the hypercorticoid state. The problem is that the hippocampus can be seen as both victim and villain in this conspiracy. It is both a source of feedback regulation over HPA activity and a primary target for elevated glucocorticoid levels. It also appears to be a target for therapeutic intervention. Antidepressant drugs increase hippocampal expression of both MR and GR, enhance glucocorticoid negative feedback sensitivity, and reduce hypothalamic CRF expression.

Among aged rats, increased hippocampal GR levels are associated with enhanced feedback inhibition of HPA activity and reduced evidence of hippocampal aging.

While the results of the clinical studies have been largely correlative, there is considerable evidence for the idea that the exogenous administration of corticosteroids impairs cognitive function. The oral administration of 10 mg of hydrocortisone leads to a significant decrease in memory performance as tested 60 min after hydrocortisone intake. In this study, declarative and nondeclarative memory performance was measured in order to assess the influence of corticosteroids on the hippocampal formation process. The logic for the inclusion of this amnesic dissociation is due to the fact that studies report that the hippocampus is essential for a specific kind of memory, notably declarative, while it is not essential for nondeclarative memory. Declarative memory refers to the conscious or voluntary recollection of previous information, whereas nondeclarative memory refers to the fact that experience changes the facility for the recollection of previous information without affording conscious access to it (priming). Thus, this somewhat specialized role of the hippocampus serves as the basis for specific hypotheses regarding the effects of acute administration of corticosteroids on human cognition. The results showed that subjects who received hydrocortisone treatment presented an impaired performance in the declarative memory task but not in the nondeclarative memory task, thus suggesting that cortisol interacts with hippocampal neurons to induce cognitive deficits.

Although most of the effects of exogenous administration of glucocorticoids on human psychophysical and cognitive processing have been found using hydrocortisone as a compound, some other studies have used other compounds and reached different conclusions. Poor performance (errors of commission; incorrectly identifying distractors as targets) on verbal memory tasks occurs in normal volunteers following the administration of prednisone (80 mg/day for 5 days). The general cognitive deficit described in this study involved the relative inability to discriminate previously presented relevant information (target) from irrelevant information (distractors) in a test of recognition memory. The authors concluded that exogenous administration of corticosteroids may diminish the encoding of meaningful stimuli and impair selective attention (i.e., reducing the ability to discriminate relevant from irrelevant information).

It appears that medium-high levels of circulating glucocorticoids may first affect the process of selective attention, thus impairing further explicit acquisition of information, whereas high-very high levels

of circulating corticosteroids may affect explicit memory. This suggestion has been confirmed by a dose-response hydrocortisone infusion study in normal controls in which a 300 µg/kg/h infusion of hydrocortisone impaired selective attention capacity, whereas a 600 µg/kg/h infusion impaired both selective attention and explicit memory.

The results of these studies suggest that in humans, as in rodents, elevated glucocorticoids serve to compromise hippocampal function and promote cognitive impairments. This hypothesis was tested in a population of elderly, generally healthy human subjects in a study that examined cognitive function in relation to individual differences in HPA activity and hippocampal integrity. Basal cortisol levels were correlated inversely with impairments, including deficits in spatial learning and memory, and were associated with the degree of hippocampal loss. These findings are consistent with those using rodent populations. Prolonged exposure to elevated glucocorticoid levels can serve to compromise hippocampal integrity and certain forms of learning and memory. There is an impressive level of functional equivalence between the rodent and human model of hippocampal aging and the apparent role of corticosteroids in promotion degeneration.

Conclusions

The assumption drawn from earlier rodent studies is that all of this is entirely irreversible. However, more recent studies have emphasized the role of dendritic atrophy as a major component of the loss of hippocampal volume associated with a prolonged exposure to elevated corticosteroid levels. Neuron loss may be a more minor component of the earlier phases of hippocampal degeneration than was originally thought. Second, more recent studies have underscored the very dynamic effects of elevated circulating glucocorticoid levels on cognitive process, suggesting that attentional impairments that culminate in learning and memory deficits need not necessarily be seen as arising only under circumstances in which prolonged corticosteroid exposure has led to rampant structural damage. Rather, the processes by which glucocorticoids impair cognition can be dynamic.

The central question for human studies, then, is whether in at least some cases the functional impairments associated with elevated glucocorticoid exposure are reversible. Rodent studies suggest that stress-induced hippocampal atrophy is reversible. The critical question is whether it can be reversible in an elderly human population. As has so often been the case, the seminal experiment on this issue comes from the Landfield laboratory, which showed that the

effects of chronic stress on hippocampal function were more serious and enduring in older than in younger animals. It might well be that this scenario emerges from a conspiracy between the endangering effects of glucocorticoids against a background of aging, which involves a more limited level of trophic support and thus a decreased ability to repair the damage associated with the stressor. In our minds, the question of reversibility is critical.

See Also the Following Articles

Cushing's Syndrome, Medical Aspects; Cushing's Syndrome, Neuropsychiatric Aspects; Hippocampal Neurons; Hippocampus, Corticosteroid Effects on; Neurogenesis.

Further Reading

- Bodnoff, S. R., Humphreys, A., Lehman, J., Diamond, D. M., Rose, G. M. and Meaney, M. J. (1995). Enduring effects of elevated glucocorticoid levels on spatial memory deficits, dampened long-term potentiation, and hippocampal neuron loss in mid-aged rats. *Journal of Neuroscience* 15, 61–69.
- Dallman, M. F., Akana, S. F., Bradbury, M. J., Strack, A. M., Hanson, E. S. and Scribner, K. A. (1994). Regulation of the hypothalamo-pituitary-adrenal axis during stress: feedback, facilitation and feeding. *Seminars in Neuroscience* 6, 205.
- de Kloet, E. R. (1991). Brain corticosteroid receptor balance and homeostatic control. *Frontiers in Neuroendocrinology* 12, 95.
- DeKosky, S., Scheff, S. and Cotman, C. (1984). Elevated corticosterone levels: A possible cause of reduced axon sprouting in aged animals. *Neuroendocrinology* 38, 33.
- Duman, R. S., Heninger, G. and Nestler, E. (1997). A molecular and cellular theory of depression. *Archives of General Psychiatry* 54, 597–606.
- Francis, D. and Meaney, M. J. (1999). Maternal care and the development of stress responses. *Current Opinion in Neurobiology* 9, 128–134.
- Issa, A., Gauthier, S. and Meaney, M. J. (1990). Hypothalamic-pituitary-adrenal activity in aged cognitively impaired and cognitively unimpaired aged rats. *Journal of Neuroscience* 10, 3247.
- Joels, M. and de Kloet, E. R. (1989). Effect of glucocorticoids and norepinephrine on excitability in the hippocampus. *Science* 245, 1502.
- Kerr, D. S., Campbell, L. W., Hao, S.-Y. and Landfield, P. W. (1989). Corticosteroid modulation of hippocampal potentials: Increased effect with aging. *Science* 245, 1505.
- Kirschbaum, C., Wolf, O. T., May, M., Wippich, W. and Hellhammer, D. H. (1996). Stress and drug-induced elevation of cortisol levels impair explicit memory in healthy adults. *Life Science* 58, 1475.
- Landfield, P., Baskin, R. K. and Pitler, T. A. (1981). Brain-aging correlates: Retardation by hormonal-pharmacological treatments. *Science* 214, 581.

- Lupien, S., DeLeon, M., DeSanti, S., et al. (1998). Longitudinal increase in cortisol during human aging predicts hippocampal atrophy and memory deficits. *Nature (Neuroscience)* 1, 59–65.
- Martignoni, E., Costa, A., Sinforiani, E., et al. (1992). The brain as a target for adrenocortical steroids: cognitive implications. *Psychoneuroendocrinology* 17, 343–354.
- McEwen, B. S. (1999). Stress and hippocampal plasticity. *Annual Review of Neuroscience* 22, 105–122.
- Meaney, M. J., Aitken, D. H., Bhatnagar, S., van Berkel, C. and Sapolsky, R. M. (1988). Postnatal handling attenuates neuroendocrine, anatomical, and cognitive impairments related to the aged hippocampus. *Science* 239, 766.
- Newcomer, J. W., Craft, S., Hershey, T., Askins, K. and Bardgett, M. E. (1994). Glucocorticoid-induced impairment in declarative memory performance in adult humans. *Journal of Neuroscience* 14, 2047.
- Newcomer, J. W., Selke, G., Melson, A. K., et al. (1999). Decreased memory performance in healthy humans induced by stress-level cortisol treatment. *Archives of General Psychiatry* 56, 527.
- Poirer, J., May, P. C., Osterburg, H. H., Geddes, J., Cotman, C. and Fince, C. E. (1989). Selective alterations of RNA in rat hippocampus after entorhinal cortex lesioning. *Proceedings of the National Academy of Sciences USA* 87, 303.
- Reul, J. M. H. M. and deKloet, E. R. (1985). Two receptor systems for corticosterone in rat brain; microdistribution and differential occupation. *Endocrinology* 117, 2505.
- Roy-Byrne, P. P., Weingartner, H., Bierer, L. M., Thompson, K. and Post, R. M. (1986). Effortful and automatic cognitive processes in depression. *Archives of General Psychiatry* 43, 265.
- Rubinow, D., Post, R., Savard, R. and Gold, P. (1984). Cortisol hypersecretion and cognitive impairment in depression. *Archives of General Psychiatry* 41, 279.
- Sapolsky, R. M. (1992). *Stress, the aging brain, and the mechanisms of neuron death*. Cambridge, MA: MIT Press.
- Sloviter, R., Valiquette, G., Abrams, G., Ronk, E., Sollas, P., Paul, L. and Neubort, F. (1989). Selective loss of hippocampal granule cells in the mature rat brain after adrenalectomy. *Science* 243, 535.
- Smith, M. A., Makino, S., Kvetnansky, R. and Post, R. M. (1995). Stress and glucocorticoids affect the expression of brain-derived neurotrophic factor and neurotrophin-mRNAs in the hippocampus. *Journal of Neuroscience* 15, 1768–1777.
- Starkman, M. N., Gebarski, S. S., Berent, S. and Scheingart, D. E. (1992). Hippocampal formation volume, memory dysfunction, and cortisol levels in patients with Cushing's syndrome. *Biological Psychiatry* 32, 756.
- Varney, N. R., Alexander, B. and MacIndoe, J. H. (1984). Reversible steroid dementia in patients without steroid psychosis. *American Journal of Psychiatry* 141, 369.
- Wolkowitz, O. M., Weingartner, H., Rubinow, D. R., et al. (1993). Steroid modulation of human memory: biochemical correlates. *Biological Psychiatry* 33, 744.

Hiroshima Bombing, Stress Effects of

R J Lifton

Harvard Medical School and Cambridge Health Alliance,
Cambridge MA, USA

Published by Elsevier Inc.

I was recently told that a prominent Zen Buddhist leader in this country has made a vow to himself that, in any conversation he enters into, he must bring up the subject of nuclear threat. That includes – so the story goes – the most casual social conversations, talks with taxi drivers, brief exchanges with strangers, whatever. The vow did not sound strange to me. Without being quite aware of it, I had, in effect, made a similar, if much more modest, vow to myself: that in any public statement I make on nuclear threat, I will bring up the subject of Hiroshima.

My strong impulse is to convey to everyone Hiroshima's indelible impact upon me. And I also

have come to sense, as both psychiatrist and antinuclear activist, the special power of specific Hiroshima or Nagasaki images for those able to receive them. Images that extreme are best described in relatively quiet, understated tones that encourage reflection on our contemporary situation. The themes put forward in this article (originally from a chapter of mine in a book co-authored with Richard Falk, *Indefensible Weapons*) are meant to portray Hiroshima not as an event of the past, but as a source of necessary knowledge for the present and the future.

Those Hiroshima images continue their insistent claim on me: the searing details of survivors' experiences; the moving scenes, collective pain, and ultimate inadequacy of the 6 August day of commemoration; our own bittersweet family celebration, with a few Hiroshima friends, of our son's first birthday. Put most simply, the 6 months spent in Hiroshima in

1962 have formed in me a special constellation of truth that, when I am wise enough to draw upon it, informs everything I have to say about nuclear threat and much else.

I arrived in Hiroshima in the early spring of 1962. I intended no more than a brief visit. But very quickly I made a discovery that I found almost incomprehensible. It had been 17 years since the dropping of the first atomic weapon on an inhabited city – surely one of the tragic turning points in human history – and no one had studied the impact of that event. There had of course been research on the physical aftereffects of the bomb, and there had been brief commentaries here and there on the behavior of some of the survivors at the time of the bomb and afterward. But there had been no systematic examination of what had taken place in people's lives, of the psychological and social consequences of the bomb.

I came to a terrible, but I believe essentially accurate, rule of thumb: the more significant an event, the less likely it is to be studied. Again, there are reasons. One reason, certainly relevant to Hiroshima, has to do with the fear and pain the event arouses – the unacceptable images to which one must, as an investigator, expose oneself. To this anxiety and pain I can certainly attest.

But another source of avoidance is the threat posed to our traditional assumptions and conventional ways of going about our studies. We would rather avoid looking at events that, by their very nature, must change us and our relation to the world. We prefer to hold on to our presuppositions and habits of personal and professional function. And we may well sense that seriously studying such an event means being haunted by it from then on, taking on a lifelong burden of responsibility to it.

I was able to stay in Hiroshima and conduct interview research with people there over a 6-month period. The best way I know to describe a few of my findings that might be of use to us now is to look at the Hiroshima experience as taking place in four stages.

The first stage was the immersion in the sea of dead and near-dead at the time the bomb fell. This was the beginning of what I have called a permanent encounter with death. But it was not just death: it was grotesque and absurd death, which had no relationship to the life cycle as such. There was a sudden and absolute shift from normal existence to this overwhelming immersion in death.

Survivors recalled not only feeling that they themselves would soon die but experiencing the sense that *the whole world was dying*. For instance, a science professor who had been covered by falling debris and temporarily blinded remembered: "My body seemed all black. Everything seemed dark, dark all over.

Then I thought, 'The world is ending.'" And a Protestant minister, responding to scenes of mutilation and destruction he saw everywhere, told me: "The feeling I had was that everyone was dead. The whole city was destroyed. . . . I thought all of my family must be dead. It doesn't matter if I die. . . . I thought this was the end of Hiroshima, of Japan, of humankind." And a writer later recorded her impressions:

I just could not understand why our surroundings changed so greatly in one instant. . . . I thought it must have been something which had nothing to do with the war, the collapse of the earth, which was said to take place at the end of the world, which I had read about as a child. . . . There was a fearful silence, which made me feel that all people . . . were dead. (Lifton, 1982 [1967] pp. 22–23.)

As psychiatrists, we are accustomed to look upon imagery of the end of the world as a symptom of mental illness, usually paranoid psychosis. But here it may be said that this imagery is a more or less appropriate response to an extraordinary external event.

In referring to themselves and others at the time, survivors described themselves as "walking ghosts", or, as one man said of himself: "I was not really alive." People were literally uncertain about whether they were dead or alive, which was why I came to call my study of the event *Death in Life* (1968).

Indicative of the nature of the event is the extraordinary disparity in estimates of the number of people killed by the bomb. These vary from less than 70,000 to more than 250,000, with the City of Hiroshima estimating 200,000. These estimates depend on whom one counts and how one goes about counting, and can be subject at either end to various ideological and emotional influences. But the simple truth is that nobody really knows how many people have been killed by the Hiroshima bomb, and such was the confusion at the time that nobody will ever know.

The second stage was associated with what I call "invisible contamination". Within hours or days or weeks after the bomb fell, people – even some who had appeared to be untouched by the bomb – began to experience grotesque symptoms: severe diarrhea and weakness, ulceration of the mouth and gums with bleeding, bleeding from all of the body orifices and into the skin, high fever; extremely low white blood cell counts when these could be taken; and later, loss of scalp and body hair – the condition often following a progressive course until death. These were symptoms of acute radiation effects. People did not know that at the time, of course; and even surviving doctors thought it was some kind of strange epidemic. Ordinary people spoke of a mysterious "poison."

But the kind of terror experienced by survivors can be understood from the rumors that quickly spread among them. One rumor simply held that everyone in Hiroshima would be dead within a few months or a few years. The symbolic message here was: none can escape the poison; the epidemic is total – all shall die. But there was a second rumor, reported to me even more frequently and with greater emotion: the belief that trees, grass, and flowers would never again grow in Hiroshima; that from that day on, the city would be unable to sustain vegetation of any kind. The meaning here was that nature was drying up altogether. Life was being extinguished at its source – an ultimate form of desolation that not only encompassed human death but went beyond it.

These early symptoms were the first large-scale manifestation of the invisible contamination stemming from the atomic particles. The symptoms also gave rise to a special image in the minds of the people of Hiroshima – an image of a force that not only kills and destroys on a colossal scale but also leaves behind in the bodies of those exposed to it deadly influences that may emerge at any time and strike down their victims. That image has also made its way to the rest of us, however we have resisted it.

The third stage of Hiroshima survivors' encounter with death occurred not weeks or months but years after the bomb fell, with the discovery (beginning in 1948 and 1949) that various forms of leukemia were increasing in incidence among survivors sufficiently exposed to irradiation. That fatal malignancy of the blood-forming organs became the model for the relatively loose but highly significant term "A-bomb disease." Then, over decades, there have been increases in various forms of cancer – first thyroid cancer, and then cancer of the breast, lung, stomach, bone marrow, and other areas. Since the latent period for radiation-induced cancer can be quite long, and since for many forms it is still not known, the results are by no means in. Researchers are still learning about increases in different cancers, and the truth is that the incidence of virtually *any* form of cancer can be increased by exposure to radiation.

An additional array of harmful bodily influences have been either demonstrated, or are suspected, to be caused by radiation exposure – including impaired growth and development, premature aging, various blood diseases, endocrine and skin disorders, damage to the central nervous system, and a vague but persistently reported borderline condition of general weakness and debilitation. Again, the returns are not in. But on a chronic level of bodily concern, survivors have the feeling that the bomb can do anything, and that anything it does is likely to be fatal. Moreover,

there are endless situations in which neither survivors themselves nor the most astute physicians can say with any certainty where physical radiation effects end and psychological manifestations begin. There is always a "nagging doubt". For instance, I retain a vivid memory of a talk I had in Hiroshima with a distinguished physician who, despite injuries and radiation effects of his own, had at the time of the bomb courageously attempted to care for patients around him. He spoke in philosophical terms of the problem of radiation effects as one that "man cannot solve"; but when I asked him about general anxieties he smiled uneasily and spoke in a way that gave me the strong sense that a raw nerve had been exposed:

Yes, of course, people are anxious. Take my own case. If I am shaving in the morning and I should happen to cut myself very slightly, I dab the blood with a piece of paper – and then, when I notice that it has stopped flowing, I think to myself, 'Well, I guess I am all right.' (Lifton, 1982[1967] p. 54).

Nor does the matter end with one's own body or life. There is the fear that this invisible contamination will manifest itself in the next generation, because it is scientifically known that such abnormalities *can* be caused by radiation. There is medical controversy here about whether genetic abnormalities have occurred: they have not been convincingly demonstrated in studies on comparative populations, but abnormalities in the chromosomes of exposed survivors have been demonstrated. People, of course, retain profound anxiety about the possibility of transmitting this deadly taint to subsequent generations. For instance, when I revisited Hiroshima in 1980, people said to me: "Well, maybe the next generation is okay after all, but what about the third generation?" The fact is that, scientifically speaking, no one can assure them with certainty that subsequent generations will not be affected. Again, nobody knows. So there is no end point for possible damage, or for anxiety.

No wonder, then, that a number of survivors told me that they considered the dropping of the bomb to be a "big experiment" by the United States. It was a new weapon; its effects were unknown; American authorities wanted to see what those effects would be. Unfortunately, there is more than a kernel of truth in that claim, at least in its suggestion of one among several motivations. More important for us now is the idea that any use of nuclear warheads would still be, in a related sense, "experimental."

The fourth stage of the Hiroshima experience is its culmination in a lifelong identification with the dead – so extreme in many cases as to cause survivors to feel "as if dead" and to take on what I spoke of as

an “identity of the dead.” Hiroshima and Nagasaki survivors became, in their own eyes as well as in those of others, a tainted group, one whose collective identity was formed around precisely the continuous death immersion and the invisible contamination I have been discussing. The identity can include what we may think of as paradoxical guilt – the tendency of survivors to berate themselves inwardly for having remained alive while others died, and for not having been able to do more to save others or to combat the general evil at the time of the bomb. In connection with the latter, the sense of “failed enactment” (Lifton, 1983 [1979] p. 174) can have little to do with what was possible at the time or with what one actually did or did not do.

More than that, survivors underwent what can be called a second victimization in the form of significant discrimination in two fundamental areas of life: marriage and work. The “logic” of the discrimination was the awareness of potential marriage partners (or families and go-betweens involved in making marriage arrangements) and prospective employers that survivors are susceptible to aftereffects of the bomb, making them poor bets for marriage (and healthy children) and employment. But the deeper, often unconscious feeling about atomic bomb survivors was that they were death-tainted, that they were reminders of a fearful event people did not want to be reminded of, that they were “carriers,” so to speak, of the dreaded “A-bomb disease.”

At the end of my study of these events, I spoke of Hiroshima, together with Nagasaki, as a last chance, a nuclear catastrophe from which one could still learn. The bombs had been dropped, there was an “end of the world” in ways I have described, yet the world still exists. And precisely in this end-of-the-world quality of Hiroshima lie both its threat and its potential wisdom.

Is Hiroshima, then, our text? Certainly as our *only* text, it would be quite inadequate. We know well that what happened there could not really represent what would happen to people if our contemporary nuclear warheads were used. When the Hiroshima and Nagasaki bombs were dropped, they were the only two functional atomic bombs in the world. Now there are approximately 50,000 nuclear warheads, most of them having many times – some a 100 or a 1000 or more times – the destructive and contaminating (through radiation) power of those first “tiny” bombs. While those early bombs initiated a revolution in killing power, we may speak of another subsequent technological revolution of even greater dimensions in its magnification of that killing power. The scale of Hiroshima was difficult enough

to grasp; now the scale is again so radically altered that holding literally to Hiroshima images misleads us in the direction of extreme understatement.

Yet despite all that, Hiroshima and Nagasaki hold out important nuclear-age truths for us. The first of these is the *totality of destruction*. It has been pointed out that Tokyo and Dresden were decimated no less than was Hiroshima. But in Hiroshima it was one plane, one bomb, one city destroyed. And the result of that single bomb was incalculable death and suffering.

A second Hiroshima truth for us is that of the weapon’s *unending lethal influence*. Radiation effects were (and are) such that the experience has had no cutoff point. Survivors have the possibility of experiencing delayed but deadly radiation effects for the rest of their lives. That possibility extends to their children, to their children’s children, indefinitely into the future – over how many generations no one knows. And we have seen how the physical and psychological blend in relation to these continuing effects.

A third truth, really derived from the other two, has to do with Hiroshima and Nagasaki survivors’ identification of themselves as *victims of an ultimate weapon* – of a force that threatens to exterminate the species. This sense had considerable impact on Hiroshima survivors, sometimes creating in them an expectation of future nuclear destruction of all of humankind and most of the earth.

And there is still something more to be said about Hiroshima and Nagasaki regarding our perceptions of nuclear danger. The two cities convey to us a sense of *nuclear actuality*. The bombs were really used there. We can read, view, and, if we will allow ourselves, *feel* what happened to people in them. In the process we experience emotions such as awe, dread, and wonder (at the extent and nature of killing, maiming, and destruction) – emotions surely appropriate to our current nuclear threat. Such emotions can transform our intellectual and moral efforts against nuclear killing into a personal mission – one with profound ethical, spiritual, and sometimes religious overtones. Hiroshima, then, is indeed our text, even if in miniature.

The argument is sometimes extended to the point of claiming that this sense of nuclear actuality has prevented full-scale nuclear war; that, in the absence of the restraining influence of Hiroshima and Nagasaki, the United States and the Soviet Union would have by now embarked upon nuclear annihilation. The claim is difficult to evaluate; and while I feel some of its persuasiveness, I do not quite accept it. In any case, one must raise a countervailing

argument having to do with another dimension of Hiroshima and Nagasaki's nuclear actuality: namely, the legitimization of a nation's using atomic bombs on human populations under certain conditions (in this case, wartime). Once a thing has been done, it is psychologically and in a sense morally easier for it to be done again. That legitimization can then combine with an argument minimizing the effects of the Hiroshima bomb: the claim that one has unfortunately heard more than once from American leaders that Hiroshima's having been rebuilt as a city is evidence that one can fight and recover from a limited nuclear war.

Here I may say that part of Hiroshima's value as a text is in its contrasts with our current situation. One crucial contrast has to do with the existence of an outside world to help. Hiroshima could slowly recover from the bomb because there were intact people who came in from the outside and brought healing energies to the city. Help was erratic and slow in arriving, but it did become available: from nearby areas (including a few medical teams); from Japanese returning from former overseas possessions; and, to some extent, from the American Occupation. The groups converging on Hiroshima in many cases contributed more to the recovery of the city as such than to that of individual survivors (physically, mentally, or economically). But they made possible the city's revitalization and repopulation.

In Hiroshima there was a total breakdown of the social and communal structure – of the web of institutions and arrangements necessary to the function of any human group. But because of the existence and intervention of an intact outside world, that social breakdown could be temporary.

Given the number and power of our current nuclear warheads, can one reasonably assume that there will be an intact outside world to help? I do not think so.

Like any powerful text, Hiroshima must be read, absorbed, and recreated by each generation searching for its own truths.

Further Reading

- Lifton, R. J. (1982 [1967]). *Death in life: the survivors of Hiroshima*. New York: Basic Books.
- Lifton, R. J. and Mitchell, G. (1995). *Hiroshima in America: fifty years of denial*. New York: Grosset/ Putnam Books.
- Lifton R. J. and Falk, R. (1982). *Indefensible weapons: the political and psychological case against nuclearism*. New York: Basic Books.
- Lifton, R. J. (1987). *The future of immortality and other essays for a nuclear age*. New York: Basic Books.
- Lifton, R. J. (1983 [1979]). *The broken connection: on death and the continuity of life*. New York: Basic Books.
- Lindqvist, S. (2001). *A history of bombing*. New York: The New Press.

HIV Infection/AIDS

B W Dixon

University of Pittsburgh School of Medicine and Allegheny County Health Department, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition article by B W Dixon, volume 2, pp 386–389, © 2000, Elsevier Inc.

Glossary

Acquired immuno-deficiency syndrome (AIDS)

A clinical state of severe immune compromise associated with HIV infection, characterized by opportunistic infections, malignant tumors (especially Kaposi sarcoma), or a CD4 lymphocyte count of less than 200.

CD4 cell

A specialized cell of the T portion of the immune system, also referred to as a helper T lymphocyte. This cell is responsible for maintaining immune competence and is one of the cells infected selectively with the HIV virus and destroyed as a consequence of viral replication. As the body's ability to

Stress Associated with Initial Testing for HIV

Stress Associated with the Presence of HIV Disease

Stress for Family and Friends of HIV-Infected Individuals

Stress for the Health-Care Team

Human immunodeficiency virus (HIV)

replenish CD4 cells fails, the CD4 cell count number falls. This is one predictor of worsening disease.

One of a group of RNA-containing retroviruses, a specific causative agent of AIDS. Similar but distinct retroviruses appear to be fairly widely disseminated in nature and are responsible for such diverse illnesses as feline leukemia, a simian AIDS-like illness, and, rarely in humans, a T-cell lymphoma.

Perhaps no disease in history – ancient or modern – has evoked such strong emotional and societal reaction as has HIV infection. Certainly the major infectious pandemics – plague of the Middle Ages, cholera of Victorian England, and influenza of 1918 – have had higher attack rates and greater mortality. Those diseases, however, were limited geographically and communication was primitive, so news of their occurrence was not disseminated readily. In contrast, HIV is of worldwide occurrence, and its end result, AIDS, which was first described in the early 1980s, was initially associated with a very high mortality rate. HIV infection in an asymptomatic state may persist for years and may respond well to newer treatments, resulting in a chronic rather than acutely fatal disease. Because, however, of the earlier picture of AIDS, many people still equate HIV infection with a rapidly progressive fatal disease. Those perceptions result in stress for the HIV-infected individual, for his or her family and friends, and for noninfected individuals who have no behaviors to put them at risk but who harbor unrealistic fears concerning the transmission of HIV. Although fear seems to be the biggest psychological stress associated with HIV, the reactions to fear are varied and include anxiety, guilt, anger and frustration, denial, and obsessive disorders. Whereas most individuals are able to handle their stress in an appropriate way, a number of people have psychological, psychosomatic, or physiological symptoms as a result of the stress associated with HIV.

Stress Associated with Initial Testing for HIV

Many emotional factors go into the decision to obtain a test for HIV. The need to acknowledge behaviors that participants have engaged in, either episodically or consistently, but with which they are uncomfortable, influences the decision. Fear of rejection by family and peers, should they be infected, appears to be important. Incomplete knowledge of the significance of a positive test for HIV antibodies and feelings of hopelessness at being able to deal with the

consequences, such as loss of future, loss of employment, isolation, and pain, have been reported. Fears of discrimination in housing, in the workplace, and in society are not uncommon. Proponents of anonymous testing point out that anonymous testing has a tendency to allay those fears. It is clear that some individuals, having mustered their courage to get a test, do not return for results. Some members of the health community who feel an obligation to ensure that individuals, particularly those who test positive, are counseled appropriately argue against anonymous testing. The role of anonymous testing in reducing stress is thus not clear.

Some home-based HIV antibody test kits have been marketed. They have a certain appeal to individuals who are concerned with privacy to a degree that they will not otherwise be tested. If the test is positive, however, individuals may not seek or may delay appropriate medical care and avoid the necessity of dealing with the stresses that result from a positive test or deal with them in inappropriate ways.

The test that is performed most commonly involves sequential testing, first by enzyme-linked immunosorbent assay (ELISA) fluorescent antibody testing and a confirmatory chromatographic (western blot) test. Although the ELISA test may be performed quite rapidly (results usually available within a matter of 24 h), the western blot test takes somewhat longer and, in most laboratories, is not run with great frequency. Because it is important to have a confirmatory western blot test in every presumptively positive ELISA test and because there may be a considerable delay between obtaining blood and receiving results, there is tremendous anxiety during the interval between obtaining blood and obtaining the results. Although some practitioners report negative ELISA test results within 24 h to reduce that anxiety, the consequence is that people who do not have their test returned in a short period of time assume (although it is not necessarily true) that their test is positive. For that reason, most testing centers have an interval of 5–7 days between obtaining blood and giving results to all individuals. At the initial encounter, it is important that the counselor or phlebotomist take extra precautions, particularly in individuals at risk for HIV, to explain the consequences of a positive test and to explain the testing procedure to reduce fears about being tested.

Stress Associated with the Presence of HIV Disease

Individuals newly diagnosed with HIV infection go through the entire spectrum of behaviors from denial to bargaining to acceptance, as occurs in many

chronic and serious diseases. These reactions have been well documented by Kubler-Ross and are not covered here. The initial denial of infection may be a barrier to obtaining care because, on initially being told of the presence of HIV infection, many people do not hear or integrate further information at that session. Thus, many health practitioners bring patients back repeatedly over a short period of time and provide easy access to someone skilled in dealing with the emotional responses to a serious life-threatening disease. Many individuals rationalize that, although they are HIV infected, at least they do not have AIDS. This is a beneficial coping mechanism during early infection.

Soon after the detection of HIV infection, analyses of viral load and CD4 count are often obtained, and this information is used in predicting not only the course of the disease in an individual but also in arriving at a plan for therapeutic intervention. Individuals who have a normal CD4 count and low viral load, in most instances, progress very slowly over a period of years. This allows sufficient time to deal with the potential psychological issues that arise. Despite the best intentions of the practitioner, a small number of people have committed self-destructive acts, including suicide, on learning of HIV infection. It is important that practitioners be aware of these potential adverse affects and allow for continuing support. For individuals with more advanced disease who meet the case definition of AIDS and for people continuing to show deterioration of immune function, modern therapies have been able to reverse or at least stabilize the disease, and numerous authors have pointed out that a positive outlook is associated with longer survival and less frequent adverse consequences, including opportunistic infections.

Unfortunately, as the disease progresses, the HIV infection itself may be associated with symptoms that can be confused with psychosomatic illness. Paresthesia in the limbs, abdominal pain, dysphasia, and changes in mental outlook and functioning may all be a result of HIV infection, its therapy or secondary infections, or neoplastic disease. Similar symptoms, however, may result on a purely psychosomatic basis without any organic cause. The challenge to the practitioner is to discern which symptoms are organic and which are psychosomatic and to deal with them by appropriate intervention. Continuity of care from a multifaceted health-care team, including physicians, social workers, advocates, and individuals who are themselves HIV-infected, appears to be important in reducing the psychological stress associated with HIV.

Whereas the stress associated with HIV infection does not appear to vary with specific racial, ethnic, or behavioral factors, individuals who acquire this

infection through intravenous drug use have a separate set of stressors because of the drug use. This must be dealt with, either prior to or concomitantly with the evaluation and treatment of the HIV infection, to maximize outcome. It is clear that treating individuals who remain active drug users with potent antiviral medications is, more often than not, fraught with failure and the development of antiviral-resistant organisms. A small but well-documented group of individuals have expressed anger at finding they are HIV-infected and have purposely tried to infect other members of society with HIV. An even larger group of individuals do not understand the infectiousness of a nonsymptomatic transmissible disease and continue to expose others by persistent risk-taking behavior. On learning that they have been the source of infection for a loved one – a sexual partner or other close societal member – they experience significant guilt.

In addition to the psychological and psychosomatic stresses exhibited by individuals who are HIV-infected, physiological changes occur as a result of HIV infection. It is well documented that stress factors of a nonspecific type have a depressant effect on the immune system and may hasten the onset of symptomatic HIV disease and reduce life expectancy. Diurnal and stress-related changes in CD4 cell count and immune function are well known. There seems to be general agreement that a fall in natural killer cell numbers and CD8⁺ lymphocytes, both of which are present in high numbers with heightened immune activity, may occur in the presence of stress. Their mechanisms of falling are not clear but may be mediated by glucocorticoids and endorphins. A similar relationship of stress to the recurrence of herpes simplex viral infections is well described, hence the term cold sore or fever blister, the symptom that occurs when a concomitant stressful situation, including infection, causes the herpes virus to reactivate. Individuals addicted to crack or cocaine in any form, which seems to have a direct immunosuppressive effect, have a shortened life expectancy.

Stress for Family and Friends of HIV-Infected Individuals

The stresses imposed on the families of people who are HIV-infected are enormous. In many instances, because of the behaviors that resulted in the individuals becoming HIV-infected (particularly sex with members of the same gender and drug use), adult individuals who have moved some distance from the family will return home when the diagnosis of HIV is made or when symptomatic AIDS occurs. The family is often unaware of the behaviors that resulted in the individuals becoming HIV-infected,

and the family members are left with the consequences of dealing with or accepting behaviors with which they are not comfortable, as well as the grief from the potential fatal outcome of AIDS. Fears concerning acceptance by relatives, neighbors, and friends add to family stress.

In the case of parents, many are aged and unable to understand adult children and the behaviors that have led to HIV. The physical stress of moving individuals or caring for them when they become weakened from disease is physically and emotionally taxing. For many parents, dealing with the potential death of a child prior to their own demise is stressful, and they may require professional help. In areas where there is a high prevalence of HIV, children may be born HIV-infected as a result of pre- or perinatal transmission, and there are many well-documented instances of parents abandoning a child on learning that he or she is HIV-infected. This may add grief and guilt to the psychological stress of the family.

Stress for the Health-Care Team

As HIV infection becomes a much more chronic disease, a variety of stresses have become more apparent in the health-care team. The difficulties of dealing with a generally young population that is losing weight and becoming physically impaired and dependent have been severe. The short natural history of the infection in the past allowed people to cope with this deterioration in health. As the disease becomes chronic, health practitioners, particularly volunteers

and people who themselves may be at risk for HIV, have become increasingly frustrated with their inability to deal with the disease over the long term. In some instances, anger has replaced sympathy, leading to depression and other stress-related symptoms in the health-care team.

See Also the Following Articles

AIDS; Herpesviruses; Immune Suppression.

Further Reading

- Antoni, M. H., August, S., LaPerriere, A., et al. (1990). Psychological and neuroendocrine measures related to functional immune changes in anticipation of HIV-1 serostatus notification. *Psychosomatic Medicine* 52, 496–510.
- Burack, J. H., Barrett, D. C., Stall, R. D., et al. (1993). Depressive symptoms and CD4 lymphocyte decline among HIV-infected men. *Journal of the American Medical Association* 270, 2568–2573.
- Glaser, R. and Kiecolt-Glaser, J. (1987). Stress-associated depression in cellular immunity: implications for acquired immune deficiency syndrome (AIDS). *Brain, Behavior, & Immunity* 1, 107–112.
- Hedge, B. (1991). Counselling the HIV-positive patient. *Practitioner* 235, 436–438.
- Kubler-Ross, E. (1970). *On death and dying*. New York: Macmillan.
- Lyketsos, C. G., Hoover, D. R., Guccione, M., et al. (1993). Depressive symptoms as predictors of medical outcomes in HIV infection. *Journal of the American Medical Association* 270, 2563–2569.

Holocaust Survivors, Experiences of

P Valent

Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Valent, volume 2, pp 396–398, © 2000, Elsevier Inc.

Cultural Background to Holocaust Experiences
Overview of Stress and Trauma Experiences in Different
Traumatic Situations
Stress and Trauma Experiences on the Moral and
Spiritual Levels
Mitigating Factors and the Truth

Glossary

<i>Holocaust</i>	The genocidal extermination of Jews during World War II.
<i>Stressors</i>	Noxious events that actually or potentially interfere with survival or the fulfillment of life's goals.
<i>Survivor</i>	A person who has lived through traumas, in this case the Holocaust.
<i>Trauma</i>	A state of irretrievable loss of a life promoting equilibrium.
<i>Traumatic stressors</i>	Situations that lead to trauma.

The experiences of the Holocaust have become a benchmark of suffering in stress and trauma situations. This was the result of the vast Nazi machinery of the day having directed all political, military, industrial, psychological, and bureaucratic resources to the unrelenting persecution and extinguishing of the culture, dignity, spirituality, and physical lives of all Jewish people. The extent and severity of the persecution resulted in every possible stress and trauma consequence occurring in large numbers and sharp relief, and, unlike in other genocides, they have been extensively recorded from both the perpetrators' and victims' standpoints. Further, the Holocaust has been a moral watershed that has not only tortured survivors but challenged our civilization. For all these reasons, the Holocaust is worth studying in its own right, in addition to being used as a source that can inform other traumatic situations.

The word experiences ambiguously subsumes events as well as subjective responses to them. This article concentrates on descriptions of stress and trauma inducing events (technically, stressors or traumatic stressors) of the Holocaust. (For the experiential consequences of the events, see **Holocaust, Stress Effects of**.) In what follows, first the cultural backdrop and the traumatic situations of the Holocaust are overviewed. Next, attacks on the morality and spirituality of the victims, including ultimate perversities and evil, are examined. Last, mitigating experiences are examined.

Cultural Background to Holocaust Experiences

The rise of Nazism ruptured an uneasy anti-Semitic equilibrium that had been present to varying degrees over the centuries in Europe, including Germany. The Nazis drew both on this anti-Semitism and on primitive tribal, religious, and ideological myths to demonize, dehumanize, and scapegoat Jews. Thus, Jews were portrayed variably as predatory monsters, subhuman parasites sucking German blood, and superhuman world conspirators.

To the Jews, many of whom had fought for their various countries, including Germany, and who were frequently at the forefront of local culture, the authorities' portrayals of them were an incredible betrayal. Persecution based on the false propaganda shattered their assumptions about a moral and just world. The total attacks on the deeper humanity and spiritual fiber of the victims made the physical attacks on their survival much harder to bear. The following two sections deal with the events that threatened physical and spiritual survival.

Overview of Stress and Trauma Experiences in Different Traumatic Situations

Wartime Situations

Some traumatic events, such as bombing and evacuation ahead of invading armies, were shared with the general population. Such events were felt to be far less noxious than the additional intentional sufferings for which Jews were earmarked.

The precursors to genocide were the following. In Germany and occupied territories, Jews had to display yellow stars on their clothes. Their identification cards indicated their Jewishness. Laws forbade their entry into public areas, segregated them in ghettos, and appropriated their means of livelihood and property. Jewish children were not allowed to attend schools.

In Germany, Krystallnacht in November 1938 broke the taboo on violence and killing. At government instigation, Jewish synagogues and property were smashed and burned, and Jews were killed. According to the new model of justice, Jews were blamed and collectively fined for the events.

As genocidal plans were ever more intensely implemented, segregation, robbery, and killing were accelerated. In occupied countries, within hours of entering villages and towns, Jews were rounded up and killed by firing squads, burned in synagogues, or beaten to death. Alternately, they were slowly killed by starvation and disease in ghettos.

Extermination camps with gas chambers and crematoria were established to increase the efficiency of killing great numbers of people. Entering such places was like entering hell. The only respite from being killed was temporary slavery, with harsh conditions and cruelty.

Some Jews escaped the Nazis to other countries in the early years. Others hid in forests, cellars, and holes. Some pretended to be non-Jews, which was called open hiding. The tension of hiding could be worse than being caught. Many were either caught or denounced and handed over.

Totality of attacks It should be emphasized that these attacks on Jews were total, relentless, and without mercy. Because the criterion for punishment was having 25% or more Jewish genes, there was no way out, such as conversion. Neither the elderly nor children were spared anything. Indeed, children were a prime target as the carriers of future genes. One and a half million children (90%) in Nazi-occupied Europe were murdered.

For the Nazis, physical and genetic extinction were not enough. Schools, culture (e.g., anything written

by Jews), and even cemeteries were destroyed. No evidence or memory of Jews was to survive.

Postwar

For those who survived the Holocaust, many postwar experiences were extremely stressful. Survivors had to come to terms with having lost most or all their relatives, friends, communities, property, and way of life. Dreams of reunion, taking up life where it had been disrupted, hopes that had sustained victims during their worst times, were now shattered. Worse, some local populations rejected returning Jews, withheld returning their property, expressed anti-Semitic views, and in some cases even killed the survivors.

In such circumstances, survivors had to hide their wounds and cope with the new survival challenges of oppression, political persecution, uprootings, and emigrations. The children who had survived had to cope with all these adult stressor events, as well as with their changed and stressed parents.

Stress and Trauma Experiences on the Moral and Spiritual Levels

In order to put the huge Nazi-inspired killing machine into action, it was essential to deprive Jews of human status. If they were deemed to not be human, then the principles of justice and law would not apply to them any more than they did to cattle or vermin and they could be killed without guilt like such animals. The Nazi machine therefore consciously used every means to first kill their victims' humanity in preparation for their physical killing.

Unrestricted and unabashed propaganda was used to portray the monstrous, parasitic, and demonic conspiracy nature of Jews. Their religion, culture, and values were attacked through constant lies. Personal humiliation and stripping of dignity were important conscious precedents and justifications to personal robbery, damage, and killing. Examples of public humiliations included cutting off old men's beards, public mocking, beating, and being made to perform menial tasks. Laws equated Jews with dogs. "No Jews or dogs allowed" signs appeared in public places. Jews were transported in cattle cars like cattle to their places of slaughter. Degradation during transportation was effected through lack of food and drink, forced lack of privacy, and having to excrete in crowded conditions. On arrival in the concentration camps, the dehumanization continued. Prisoners were herded by dogs and whips, were cursed and beaten, and were branded with numbers. Those not slaughtered immediately had all control taken away. They

had to work like animals, had to obey trivial orders and routine on the pain of death, and had their most intimate human functions such as of ingestion and excretion governed by guards. Even their deaths were undignified. Prisoners were stripped naked, beaten, and herded into communal impersonal killing machines. In other words, in addition to the constant threats to their lives, Jews had to contend with forces that were meant to strip them of their humanity, identity, and everything they cherished and loved, their very spiritual and existential meanings. Their souls were attacked as much as their bodies.

Some attacks were particularly cruel and perverse. They included the forced separation of families. When attempts were made to keep family members together, such as when mothers chose to stay with their babies, this love was punished by killing both the mothers and children. Victims were forced to make perverse choices, sacrificing some in order for all not to be killed. For instance, communities were forced to give up their old and young so that all would not be killed. Compounding this perversity, prisoners were forced to take part in the killing and the body-looting process. Worse, victims had to suppress their emotions in order to survive.

Thus, in addition to being the targets of physical sadism and abuse of unlimited power, victims were forced to shed civilized values, relinquish what they held dear, and even turn against what had made their lives meaningful. Mengele, who like the Angel of Death chose with a motion of his finger who was to live and who to die, became a symbol of the victims' powerlessness. The meaninglessness of the slaughter was highlighted by the banality of the industrialization and bureaucratization of the killing process, which used and parodied all the achievements of modern civilization, including the use of psychological knowledge to inure the killers to continuous killing.

Mitigating Factors and the Truth

In the midst of genocide, there was no lack of devotion and courage among Holocaust victims, who risked and helped one another whenever possible. Sometimes kindness and bravery were shown by outsiders at the risk of their lives. The truth of Holocaust experiences can often be absorbed against the background of survival, courage, and victory of goodness, as in Steven Spielberg's film *Schindler's List*. But the truth is that redemptive experiences were far outweighed by the suffering and pain, ineffectiveness of bystanders, meaningless deaths of millions,

and long-term moral and existential wounds of the survivors.

See Also the Following Article

Holocaust, Stress Effects of.

Further Reading

- Friedrich, O. (1982). *The kingdom of Auschwitz*. New York: HarperCollins.
 Gilbert, M. (1986). *The Holocaust; the Jewish tragedy*. London: Collins.
 Levi, P. (1987). *If this is a man*. London: Abacus.

Holocaust, Stress Effects of

P Valent

Melbourne, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Valent, volume 2, pp 390–395, © 2000, Elsevier Inc.

Stress Effects during the Holocaust

Stress Effects Soon after the Holocaust

Stress Effects over the Years

Stress Effects in Recent Times

Lessons and Challenges Regarding Stress Effects from the Holocaust

Glossary

<i>Holocaust</i>	The genocidal extermination of Jews during World War II.
<i>Stress effects</i>	Biological, psychological, and social responses to stressors; they can form the constituents of traumas and illnesses.
<i>Stressors</i>	Noxious events that actually or potentially threaten life and its fulfillment.
<i>Survivor</i>	A person who has lived through traumas, in this case, the Holocaust.
<i>Trauma</i>	A state of irretrievable loss of a life-promoting equilibrium.
<i>Traumatic stressors</i>	Stressors that lead to trauma.
<i>Traumatic situations</i>	Situations that often lead to trauma.

Just as the experiences of Holocaust survivors (see **Holocaust Survivors, Experiences of**) can serve as a benchmark for the stressor nature of Holocaust traumatic situations, so the range and depth of the consequent stress effects of Holocaust experiences can serve as a standard against which stress effects from other traumatic situations can be compared. Further, because Holocaust stress effects have been well documented for

over 60 years, the progression of stress effects – biological, psychological, and social – can be examined over an individual's life cycle, as well as down the generations.

Stress Effects during the Holocaust

Adults

Although, understandably, not scientifically researched in an extensive manner, reports indicate a wide variety of biological, psychological, and social stress responses associated with roundups and deportation, segregation into ghettos, concentration camps, and being in hiding.

Roundups and deportation During roundups and deportations, severe anxiety and reactive depressions were ubiquitous and were perceived as understandable rather than pathological. However, panic attacks, severe depressions, suicide attempts and completed suicides, and the new development of angina and hypertension in these situations, although still understandable, were indistinguishable from physical and psychiatric illnesses.

The sense in many of the neurotic and psychotic symptoms that also occurred can easily be discerned. For instance, agoraphobia developed in response to the danger of going outside. Hysterical symptoms to avoid deportation and paranoid delusions, such as of the gas man, are also quite understandable. The abeyance of prior neuroses and the development of new ones seemed to be dictated by survival needs.

Ghettos and hiding Anecdotal evidence indicates that stress effects in ghettos overlapped those of the roundups, deportations, and concentration camps. Hiding in some ways was more stressful than being caught because people had to cope with constant isolation in hostile circumstances and the fear of being caught could be worse than actually being caught.

Concentration camps On entering the camps, psychic shock (see **Disaster Syndrome**) was often described subjectively as apathy, exhaustion, and emotional blunting. When apathy and total exhaustion continued, they contributed to the 20–50% who just died without obvious disease soon after entering concentration camps. Older and middle-class people were more prone to such shock. Many of them also committed suicide.

In the psychiatric ward in the concentration camp Teresienstadt, all psychiatric diagnostic categories were reported to occur well above expected rates. The diagnoses included schizophrenia and manic-depression. Reports from other concentration camps also described not infrequent cases of depressions, phobias, paranoias (such as of being poisoned), hallucinations, and psychotic reactions. Most psychiatrically ill prisoners died or were shot if their behavior did not conform to strict rules. Some psychoneuroses and psychosomatic illnesses disappeared, and new ones appeared. In some cases, psychosomatic illnesses that went into abeyance during internment reappeared after liberation.

Most prisoners in concentration camps who were not actively murdered, succumbed to the just-as-lethal extermination process, which included starvation, cold, physical assaults, and disease. Physical symptoms such as stomach cramps, diarrhea, dizziness, and headaches due to a variety of diseases were frequent, although some may have been psychophysiological.

Coping and defenses Most people did what they could to help others and themselves. Mostly nothing could be done except to obey orders, resign oneself to circumstances, and maintain hope deep inside.

The most potent factor in survival was luck. But, in addition, to survive victims had to have a strong desire to live no matter what. Other factors that helped them to survive were the suppression of feelings, dissociation and detachment, and living in small stretches of time. Some were helped by intellectual curiosity, others by magical beliefs, humor, and art. Socially, younger people, those who paired to help one another, and those who had strong beliefs had better chances of survival. Religious beliefs helped some but hindered survival for others. However, hope was essential for everyone. The sudden loss of hope was deadening and could lead to a fast demise through infections such as typhus.

Children

Stressors were larger for children, and their responses to stress were frequently insufficient, especially among the youngest children. Children were often the first to succumb to starvation, cold, and disease,

and they had the highest mortality rates – 90% of Jewish children died in Nazi-occupied Europe. The surviving children's responses were shaped by their developmental phases.

Younger children The younger children's worlds fragmented easily. Stress effects frequently included somatic symptoms such as stomachaches, diarrhea and asthma and uncontained emotions and behavior. Younger children tended also to have relatively more vivid, concrete, and phantastic interpretations of events. For instance, human persecutors were admixed with images of monsters, and separations from parents were interpreted as punishment for being bad.

Older children The older children, especially those beyond the age of seven, experienced dread, fear, grief, despair, anger, and guilt akin to the adults. Their capacities and ingenuity sometimes saved adults.

Few children survived the concentration camps; however, in a relatively privileged part of Bergen-Belsen, children were noted to suffer night anxieties, enuresis, and phobias; to be irritable and aggressive; and to form gangs akin to civilian wayward youth.

Coping and defenses Especially when children felt themselves to be under their parents' protective shields, they could interpret the frightening events in part according to their normal developmental phases. For instance, they could see deportation as an adventure.

From about the age of four on, children's psychological defenses resembled increasingly those of adults in their abilities to dissociate and freeze their emotions and the meanings of events. In addition they had amazing capacities for obedience and discerning that life was at stake. For instance, they separated silently from their parents, were docile with strange caretakers, and kept silent when hidden, even when alone and in the dark for long periods. They assumed and lived out a series of false identities and arranged their psyches as desired.

Children also owed their survival to luck, an inner drive to live, suppressing feelings, and maintaining hope. The latter was often through a tenacious clinging to objects and memories representing loving parents. Being appealing, evoking caring impulses, and being plastic when making various adjustments also helped their survival.

Stress Effects Soon after the Holocaust

Adults

At liberation The joy of liberation was mitigated by continued physical debilitation, which in some

cases led to death. Unfortunately, too, many died of eating food that their frail digestive systems could not handle.

Mental debility from camp life could continue after liberation. For instance, in Bergen-Belsen concentration camp, over 40 schizophrenic and several hysterical survivors maintained the psychiatric symptoms that had helped them to survive hunger, death, and torture during their incarceration. Many of the survivors continued to suffer dulled and blunted affect and decreased memory. Others developed reactive despair and paranoid delusions only after liberation.

Postwar Many postwar stressors were worse than the wartime ones. The reality of murdered families, empty communities, and hostile countrymen shattered the dreams that had kept survivors going. Many let go now and took ill. Some died. The rates for suicide and schizophrenic illnesses rose. Some became chronically depressed. Others developed psychosomatic pains, dizziness, and weight-gain problems, all of which often became chronic. Amenorrhea, dysmenorrhea, abortions, and premature babies were not infrequent.

Coping and defenses Most victims resumed their survivor modes. They cut off emotions and memories, labored for today, and hoped for the future. Many married quickly and had children. Many started to achieve financial security, imbued with new hopes in their children and new countries.

Children

Children also enjoyed their freedom, but came to despair when their parents did not materialize or had changed from those in their dreams. Others grieved at having to leave their foster families. Many children were now treated for tuberculosis and other neglected illnesses, and this could mean further separation from their parents. Many exhibited neurotic symptoms and nonsocial behaviors. On the whole, the children coped, like the adults, by suppressing their emotions and memories and by concentrating on establishing their school and social lives.

Stress Effects over the Years

Adults: The Survivor Syndrome

The wide-ranging and marked biological, psychological, and social effects of the Holocaust came to be recognized only slowly and reluctantly. Psychiatry had to invent diagnostic labels to designate the salient image of the shuffling corpse, and the pervasively psychologically scarred survivor with numerous

symptoms. Survivor syndrome became the most commonly accepted label for these symptoms. Its features, described next, subsumed the many biological, psychological, and social symptoms and illnesses, which occurred in many fluctuating combinations and permutations.

Physical sequelae Research on ex-concentration camp prisoners (the same probably applies to other Holocaust survivors) found that they suffered excessive mortality and morbidity from a great variety of illnesses for decades compared to control populations. Illnesses included tuberculosis and other infections, cancer, hypertension, coronary heart disease, stomach ulcers, digestive disorders, hyperthyroidism, and accidents. Premature aging and arteriosclerosis were also noted.

Survivors also suffered excessive autonomic nervous system psychophysiological symptoms. Common sympathetic (fight-or-flight) symptoms were tension, muscle pains, and digestive and cardiovascular symptoms, parasympathetic symptoms included tiredness, dizziness, lassitude, and exhaustion. In addition survivors relived symptoms such as stomach cramps, diarrheas, and headaches along with other wartime memories or (in somatization) as symbolic substitutes for them. Finally, any of these symptoms could be used hypochondriacally to signal anxieties directly or indirectly related to Holocaust fears.

Psychosocial and psychiatric sequelae Psychosocial sequelae include mainly negative emotions, cognitions, defenses, psychiatric illnesses, and spiritual and existential dilemmas. They, in turn, affected relationships and worldviews.

- Emotions. Anxiety, fear, and terror were the most common emotions, followed by depression. The latter was often associated with suppressed grief and was often complicated by survivor guilt (see **Survivor Guilt**).

- Cognitions. Holocaust events could present as hypermnasias or amnasias. In the hypermnasias, the past events were relived vividly in dreams, nightmares, flashbacks, thoughts, and fantasies. The relived events might not be distinguished from the original ones. They could be especially set off by cues reminiscent of the initial Holocaust events; such cues could act as phobias, which survivors might avoid strenuously. The amnasias, on the other hand, were the consequences of avoidance and a variety of defenses such as dissociation, repression, regression, splitting, and projection. The tension between reliving and avoidance (the two main features of posttraumatic stress disorder, PTSD) was

noted to lead to restlessness, decreased concentration, sleeplessness, memory disturbances, moodiness, emotional lability, irritability, and physiological lability.

- **Psychiatric sequelae.** As in the physical arena, Holocaust stresses led to a wide variety of psychiatric symptoms and illnesses. They could be present from early on or could develop after a latent period (even decades later). Most commonly, anxiety and depression intensified into a range of anxiety and depressive disorders, the latter again associated with unresolved mourning. Dissociative defenses were associated with dissociative symptoms such as depersonalization, projection, and displacement, with suspiciousness, persecutory delusions, nighttime hallucinations, schizophreniform psychoses, and schizophrenia. All of these constellations could emerge years later and usually had Holocaust-laden content.

- **Spiritual and existential sequelae.** Survivors were ubiquitously plagued with moral, identity, spiritual, and existential problems, often causing more distress than the physical and psychological problems. Survivor guilt torments included “Why am I alive, when my family/such good people died?” and “Why did I acquiesce to go left, when my mother was sent right to her death?” Other questions included “How could such injustices occur?” “Where was the world/God?” “What does it all mean?” “Who am I now?” and “What purpose can my life have now?” Survivors could become cynical, lacking confidence in the world and themselves, always ready for further wounds.

Coping and defenses Many survivors did well financially due to their single-minded efforts, and their wives helped in their businesses and ran spotless homes. These were scotomized areas of fulfillment because they were intermingled with insecurity.

Survivors achieved meaning by having children to defeat Hitler’s genocide, bear witness, and execute missions of remembrance. Israel was the political answer to vulnerability. Many survivors found purpose by trying to prevent suffering in others. Thus, survivors provided more than their share of moral leadership, philanthropy and cultural accomplishments.

Children

Like their parents, child survivors also suffered excessive mortality and morbidity rates from all types of illnesses, including psychoses and suicides. Child survivors, unlike their parents, often could not remember the events that still plagued them. This was because of their young age at the time and the strong desire of adults that the children not remember. Thus,

they lived their unhappiness as a saturated solution of unacknowledged turmoil, out of which could crystallize clearer symptoms and illnesses. Child survivors coped, like their parents, by concentrating on survival, the future, their own children, and ameliorating suffering in the world.

Only in the 1980s, in their forties and fifties, were child survivors recognized, and they recognized themselves. They explored their identities and memories beyond the molds that had been arranged by others.

Second-Generation Holocaust Survivors

Instead of making sure that their post-Holocaust children’s experiences were linked to a normal external world, Holocaust survivors often linked their children’s experiences to the Holocaust that still dominated them. The children then became unconsciously attuned to their parents’ Holocaust experiences. For instance, a child was spanked for hopping happily in the park because to the parent this was an affront to all the dead; this child grew up to be a sad adult, pervaded by death, without knowing why. Children also absorbed unconscious roles, such as substituting for dead parents and prior children.

Second-generation children often felt overburdened, overprotected, and frustrated with parents who did not recognize them, yet felt unable to be angry with their fragile parents who themselves craved care and understanding. The children developed both symptoms due to identifying with their parents and symptoms resulting from their own attachment problems and unintegrated identities. These children often copied their parents’ means of coping, such as being devoted to their own children and being prominent in humanist and healing occupations.

Stress Effects in Recent Times

Adult Survivors

As Holocaust survivors reach the ends of their lives, it can be seen that the effects of stress and trauma are lifelong. Retirement, the illnesses of old age, diminished brain function, and helplessness can intensify memories, and, for instance, doctors and nurses may be experienced by survivors as Nazis whose injections kill the sick.

On the other hand, survivors enjoy seeing the generations flourish, and this can reconnect them with their pre-Holocaust memories. It is never too late to resolve guilt and integrate frozen meanings. For instance, many survivors have recently given testimonies to the Spielberg Shoah Foundation and have broken their silence with their children.

Child Survivors

Child survivors, having been recognized and having discovered their own identities, have emerged from hiding, which in a sense they had been in since the Holocaust. Although still suffering, some with increased symptoms like their parents, for many it is more important to retrieve their memories clearly, integrate their Holocaust traumas, and make meaning of their lives. To these retrieve memories, many have visited the places of their traumas and have broken the conspiracy of silence. Once they have regained their identity and memories, they trauma pains can be integrated and new meanings can be forged.

Child survivors have formed local, national, and international groups. They have also given testimonies and feel the special privilege and responsibilities associated with being the last witnesses of the Holocaust.

Second and Subsequent Generations

The integration of the Holocaust by second-generation children has been harder because their Holocaust-derived culture has been pervasive and ego-syntonic. Nevertheless, many have come to realize the sources of their problems and have attempted to resolve them with their parents, including going with them to the places of their suffering. Alternately, they have attempted to work out in psychotherapy who they were and why.

Lessons and Challenges Regarding Stress Effects from the Holocaust

If Holocaust stress and trauma effects are seen a kind of benchmark, what principles can be abstracted from them?

1. Biopsychosocial nature of stress effects. Holocaust stress effects indicate that a wide variety of biological, psychological, and social stress effects may occur in many combinations and are subsumed in a great variety of symptoms and illnesses.

2. Diagnostic categories arising from stress effects. This great variety of symptoms and illnesses challenges us to adopt a nonlinear model in which trauma is like the big bang at the center and the symptoms and illnesses freeze out at different points on different ripples emanating from such a center. Survivor syndrome and PTSD may then be conceptualized as generic entities that subsume stress-derived symptoms and illnesses.

3. Multidimensional nature of stress effects. In addition to the rippling of stresses and traumas into

symptoms and illnesses, the Holocaust, especially, indicates that trauma disrupts the human dimension, including morality, meaning, spirituality, identity, and purpose. The Holocaust further alerts us that stress effects take place in individuals, families, communities, and nations; in children and adults; and across generations.

4. Adaptive as well as maladaptive stress effects. During the Holocaust, many responses were adaptive and later could be elaborated into adaptive concerns that benefited large sections of the population. This must be remembered but not idealized – the suffering far outweighed any beneficial results of the traumatic events.

In conclusion, the stress effects of the Holocaust are very varied, deep, and intense. The sense of the overwhelming nature of the effects and the despair about being able to do anything about them parallels our general avoidance of trauma and despair about being able to treat it. Yet the fear is excessive, and helpful recognition of even the greatest traumas can be very beneficial and rewarding to both survivors and to those in contact with them.

See Also the Following Articles

Concentration Camp Survivors; Disaster Syndrome; Holocaust Survivors, Experiences of; Survivor Guilt; War-Related Posttraumatic Stress Disorder, Treatment of.

Further Reading

- Bergman, M. S. and Jucovy, M. E. (eds.) (1982). *Generations of the Holocaust*. New York: Columbia University Press.
- Dwork, D. (1991). *Children with a star; Jewish youth in Nazi Europe*. New Haven, CT: Yale University Press.
- Krell, R. and Sherman, M. I. (eds.) (1997). *Medical and psychological effects of concentration camps on Holocaust survivors; genocide: a critical bibliographic review*. New Brunswick, NJ, and London: Transaction Publishers.
- Krystal, H. (ed.) (1968). *Massive psychic trauma*. New York: International Universities Press.
- Marcus, P. and Rosenberg, A. (eds.) (1989). *Healing their wounds; psychotherapy with Holocaust survivors and their families*. New York: Praeger.
- Niederland, W. G. (1968). Clinical observations on the "survivor syndrome." *International Journal of Psychoanalysis* 49, 313–315.
- Valent, P. (1998). *From survival to fulfillment; a framework for the life-trauma dialectic*. Philadelphia: Brunner/Mazel.
- Valent, P. (2002). *Child survivors of the Holocaust*. New York and London: Brunner-Routledge.

Homeostasis

B S McEwen

Rockefeller University, New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by B S McEwen, volume 2, pp 399–400, © 2000, Elsevier Inc.

Characteristics of Homeostatic Systems

Problems with the Concept of Homeostasis and
Alternative Terminology

Glossary

Homeostasis The ability of an organism to maintain the internal environment of the body within limits that allow it to survive; self-regulating processes that return critical systems of the body to a set point, or range of operation, consistent with survival of the organism.

Characteristics of Homeostatic Systems

Homeostasis is highly developed in warm-blooded animals living on land, which must maintain body temperature, fluid balance, blood pH, and oxygen tension within rather narrow limits while at the same time obtaining nutrition to provide the energy to maintain homeostasis. This is because maintaining homeostasis requires the expenditure of energy. Energy is used for locomotion as the animal seeks and consumes food and water, for maintaining body temperature via the controlled release of calories from the metabolism of food or fat stores, and for sustaining cell membrane function as it resorbs electrolytes in the kidney and intestine and maintains neutral blood pH.

Homeostasis also refers to the body's defensive mechanisms. These include protective reflexes against inhaling matter into the lungs, the vomiting reflex as a protection to expel toxic materials from the esophagus or stomach, the eye blink reflex, and the withdrawal response to hot or otherwise painful skin sensations. Also included is the defense against pathogens through innate and acquired immunity.

According to Cannon, "Bodily homeostasis . . . results in liberating those functions of the nervous systems that adapt the organism to new situations, from the necessity of paying routine attention to the management of the details of bare existence. Without homeostatic devices we should be in constant danger of disaster, unless we were always on the alert to correct

voluntarily what normally is corrected automatically. With homeostatic devices, however, that keep essential body processes steady, we as individuals are free from such slavery – free to enter into agreeable relations with our fellows, free to enjoy beautiful things, to explore and understand the wonders of the world about us, to develop new ideas and interests, and to work and play, untrammelled by the anxieties concerning our bodily affairs."

The body has a considerable redundancy or margin of safety when it comes to many of these mechanisms. There are two kidneys and adrenal glands in case one fails, and the digestive tract is able to function adequately for food absorption even if intestinal length is shortened. Moreover, individuals are able to survive with one lung. And the cardiovascular system has a number of devices to maintain blood pressure after considerable blood loss. Finally, as Cannon notes, the body displays considerable resilience in the face of injury and after surgery to compensate for lost function.

Problems with the Concept of Homeostasis and Alternative Terminology

Yet there are paradoxes and problems with the notion of a set point that is always maintained. In discussing obesity, Cannon states, "What leads to the storage of fat in large amounts – sometimes in grotesquely large amounts – in some persons, and in slight amounts in others, is not well-understood." This and other aspects of pathophysiology, particularly in response to stress, has forced a reevaluation of the processes that reestablish homeostasis in the face of a challenge and that, under some circumstances, exacerbate disease (see **Allostasis and Allostatic Load**).

Another problem with homeostasis is that, taken literally, it is a rather static concept, akin to the notion of equilibrium in classical thermodynamics; in real life, an organism experiences changing set points around which it maintains stability over a finite period of time. Moreover, unlike the closed systems of classical thermodynamics, a living organism is an open system in which energy and matter flow in and out, with equilibrium being replaced by a steady state. Cannon recognized this and used the term steady state; he even suggested that homeodynamics might be a better term than homeostasis.

The same problem was also recognized by a number of other authors. Nicolaidis introduced the term homeorheusis in 1977; here stasis is replaced by rheusis, meaning "something flowing." Mrozovsky introduced the term rheostasis in 1990, "a condition

or state in which, at any one instant, homeostatic defenses are still present but over a span of time there is a change in the regulated level, or set point of a system.” He considered this concept advantageous in explaining how organisms adjust to changing environments such as the seasons of the year by storing body fat or changing reproductive physiology, and he believed that rheostasis allows for the kinds of physiological plasticity that is involved in evolution.

See Also the Following Articles

Allostasis and Allostatic Load; Feedback Systems.

Further Reading

- Cannon, W. B. (1929). Organization for physiological homeostasis. *Physiological Reviews* 9, 399–431.
- Cannon, W. B. (1939). *The wisdom of the body*. New York: W. W. Norton.
- Mrosovsky, N. (1990). *Rheostasis: the physiology of change*. New York: Oxford University Press.
- Nicolaides, S. (1977). Physiologie du comportement alimentaire. In: Meyer, P. (ed.) *Physiologie humaine*, pp. 908–922. Paris: Flammarion Medecine-Sciences.

Homosexuality, Stress and

J Drescher

New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

Heterosexism
Gender Beliefs and Gender Stress
Homophobia
Moral Condemnations of Homosexuality
Antigay Violence
The Stress of Being in the Closet

Glossary

<i>Being in the closet</i>	Psychological and behavioral activities intended to keep one's homosexuality a secret.
<i>Coming out</i>	Admitting one's homosexuality to oneself or revealing it to others.
<i>Dissociation</i>	A psychological mechanism that allows individuals to keep unwanted or anxiety-producing information out of their awareness.
<i>External homophobia</i>	The irrational fear and hatred that heterosexuals feel toward gay people.
<i>Gay bashing</i>	Physical violence directed at people because they are gay.
<i>Gender beliefs</i>	Cultural ideas about what constitutes the essential qualities of men and women.
<i>Gender identity</i>	An individual's psychological sense of being either a man or a woman.
<i>Heterosexism</i>	A belief system that naturalizes and idealizes heterosexuality and either dismisses or ignores a gay subjectivity.

Internal/ internalized homophobia
Moral condemnations of homosexuality
Outing

The self-loathing gay people feel for themselves.

Beliefs that regard homosexual acts as intrinsically harmful to the individual, to the individual's spirit, and to the social fabric.

The unwanted revelation of an individual's homosexuality to others by a third party.

Sexual orientation

A person's attraction to members of either the same sex (homosexual), the other sex (heterosexual), or to both sexes (bisexual).

Sodomy laws

Legislation that criminalizes homosexual behavior, ruled unconstitutional by the United States in 2003 by the Supreme Court.

Throughout the life cycle, gay, lesbian, and bisexual (GLB) individuals experience significant stressors that are a consequence of society's antihomosexual attitudes. Some of those attitudes and the stresses they engender are discussed here.

Heterosexism

Those with GLB identities are routinely stressed by cultural expectations that each individual have a heterosexual identity. This is known as heterosexism, a belief system that naturalizes and idealizes heterosexuality and either dismisses or ignores a gay subjectivity.

Heterosexism does not necessarily imply malice or ill will toward GLB individuals; idealizing oneself

does not necessarily require denigrating those who are different. However, heterosexist ideology typically frames its antihomosexual arguments in terms of the normal, the natural, or the traditional. Such arguments, sometimes implicitly and at other times explicitly, tend to equate identities that are not conventionally heterosexual with immorality and can rationalize condemnations of same-sex feelings and behaviors. Heterosexist attitudes are pervasive in many cultures, and they inevitably place gay people in the awkward and often stressful position of either (1) keeping their homosexuality a secret (being in the closet) or (2) having to reveal they are gay (coming out) to disprove cultural myths and assumptions about homosexuality. The most dramatic example of this stressful dilemma can be seen in the U.S. military's Don't Ask, Don't Tell policy, under which either revealing oneself to be gay or being found out to be gay can lead to a discharge.

In Western culture, most books, plays, movies, and print and television ads contain naturalized depictions of heterosexuality. "Boy meets girl, boy loses girl, boy gets girl" are the common elements of a heterosexist script. Heterosexism renders gay lives so invisible that at the time of writing this chapter, there is enormous media attention and controversy over *Brokeback Mountain*, a film in which the two male leads fall in love. Many heterosexual people have openly stated either their reluctance or opposition to seeing two men kiss on the silver screen.

Common expressions of heterosexism include valuing heterosexual relationships above homosexual ones, unquestioned opposition to gay parenting regardless of a gay parent's actual qualities as a parent, and the belief that lesbians and gay men are unsuited to teach small children, to serve in government, or to practice medicine. Heterosexist assumptions lead to separating men and women from one another in locker rooms. Yet gay teenagers may be stressed by having to hide any possible arousal they may experience on seeing naked peers in the locker room's ostensibly asexual environment. Similarly, a heterosexist world may make it difficult for elderly gay couples to openly live together in either retirement communities or assisted living programs.

A decade ago, marriage was implicitly assumed to be a ritual reserved for heterosexual couples. As a result of lawsuits initiated by gay couples in the 1990s and following the 2003 decision to allow same-sex marriage in Massachusetts, many other states rewrote their constitutions to specifically state that marriage should only be between a man and a woman. The stated impetus for these constitutional changes is often termed the "defense of marriage." Yet laws ostensibly written to preserve heterosexual relations

inevitably create stress for gay couples and their families. The U.S. General Accounting Office has computed over 1000 rights, privileges, and benefits accrued through marriage. Consequently, the heterosexist denial of legal marriage to long-term, committed gay couples adds to the latter's stress by denying them and their families basic needs, including clearly defined inheritance rights, pension benefits, survivor death benefits, custody rights in cases of divorce, and health benefits, just to name a few. A growing literature also indicates increased psychological resilience in those who marry; denying legal marriage to same-sex couples denies them marriage's psychological benefits as well.

Gender Beliefs and Gender Stress

Any discussion of antihomosexual stressors necessitates some discussion of both scientific and popular heterosexist beliefs about gender and sexuality. Modern sexology distinguishes a person's gender identity, or the sense that one is a man or a woman, from the person's sexual orientation, whether one is attracted to members of the same sex (homosexual), the other sex (heterosexual), or to both sexes (bisexual). However, the general culture uses conventional heterosexuality as an organizing frame of reference and frequently conflates the categories of gender identity and sexual orientation. These common but mistaken gender beliefs include the notion that an attraction to men is a female trait or that gay men feel and act like girls. Although such mundane gender beliefs can be stressful because they misrepresent GLB people, GLB individuals are not alone in having to contend with them. Gender beliefs also include ideas about the kind of shoes men should wear (construction boots rather than spike heels) and the kind of career a woman should choose (becoming a nurse rather than a doctor). In other words, gender beliefs are not confined to the realm of sexuality; they concern themselves with many aspects of day-to-day life.

Gender beliefs play an important role in the development of gender identity. One may be born with an anatomical gender, but one's psychological gender is a later acquisition that has both cognitive and affective elements. How one's gender identity actually develops is not precisely known, although empirical data suggest it is a long and complex process. Clinical work shows that a person's sense of being a man or a woman can involve processes that are entirely dependent on attitudes from the external environment over extended periods of time. Children must learn a psychological construct of gender that is based not solely on anatomy but also on myriad cultural and familial clues (e.g., boys like to play ball or girls are

made of sugar and spice). Consequently, the meanings of, for example, aggression in girls or a lack of athletic interests in boys can be internalized along a family or cultural model that codes these attributes as gender-specific. From this developmental perspective, a gender identity can be thought of as the accretion of experiences or interactions with the external and internalized gender-coding environment. That is to say, a person's gender is an activity occurring in a relational matrix.

For children who grow up to be heterosexuals, conformity to conventional gender beliefs and norms can sometimes offer a sense of comfort. However, gender beliefs can be stressful to heterosexuals as well. Real men and real women are powerful cultural ideals against which all men and women may unfavorably compare themselves. However, GLB individuals must contend with the dissonance created by early same-sex attractions and their learned cultural beliefs about the relationship between homosexuality and gender. Consequently, in some children who grow up to be gay, their developing gender beliefs about the meanings of their homoerotic feelings may lead them to question their own gender identities.

Seen from this perspective, children who grow up to be gay may be gender stressed. This is a process that can occur over a protracted period of time as early homoerotic feelings become linked to other gender-nonconforming interests and as these children come to realize that they are failing to meet the cultural and social expectations of their assigned gender. This is commonly seen in gay men who were effeminate boys, often denigrated for being sissies. However, even gay men who were more conventionally masculine as children recall gender stresses in their attempts to integrate their same-sex feelings into a male identity. Sometimes gender stress can lead to gender confusion. Clinically, gender confusion can be expressed as "If I am attracted to boys, I must be like a girl." But it might also include responses such as "I am depraved," "I will be alone," or "I have sinned." These are common and stressful gender beliefs associated with transgressing gender boundaries.

Homophobia

George Weinberg is generally credited with coining the term homophobia. He defines external homophobia as the irrational fear and hatred that heterosexuals feel toward gay people and internal (or internalized) homophobia as the self-loathing gay people feel for themselves. Weinberg suggests several origins of homophobia: the religious motive, the secret fear of being homosexual, repressed envy, and the threat to values. In fact, plethysmographic studies

demonstrate that in some men, homophobic attitudes may defend against unwanted homoerotic feelings.

Weinberg's early work has been followed by an ever-expanding literature devoted to studying intolerance of homosexuality. Other factors thought to contribute to homophobic attitudes include (1) rare contact with GLB people, (2) a lack of homosexual experiences of one's own, (3) perceiving that one's own community does not accept homosexuality, (4) living in communities where homosexuality is unacceptable, (5) being older, (6) being less educated, (7) identifying oneself as religious or as belonging to a conservative religion, (8) being less permissive about sexuality in general, and (9) expressing high levels of authoritarianism.

In 1954, the U.S. Supreme Court ruled in *Brown v. The Board of Education* that racial segregation was unconstitutional. In support of that decision, the justices cited the psychological literature of the time that demonstrated how stressful and harmful racism was to the self-esteem of black children. Just as one cannot grow up black in America without internalizing racism, one cannot grow up gay without internalizing homophobic and other antihomosexual attitudes. Self-hatred is extremely stressful. In the clinical setting, internalized homophobia can present as low self-esteem, difficulty in sustaining relationships, estrangement from one's family of origin, work difficulties, or self-medication with drugs or alcohol to reduce stress-induced anxieties.

Homophobia is often expressed in the association of male homosexuality with effeminacy. Sissies of all ages are regularly targets of homophobia. The low social status accorded to sissies leads some parents, fearful that their children will grow up to be gay, to seek professional treatment to change the behaviors of their effeminate sons. However, when mental health professionals and parents join the social chorus of derision, presumably for a child's own good, being labeled effeminate at an early age can entail great stress and suffering.

Although some gay men are effeminate, assuming that all gay men are effeminate falls under the rubric of stereotyping. In addition to disdain for effeminacy, homophobia can be expressed in other stereotypes, clichés, or caricatures of sexual identities. Common stereotypes of gay men describe them as being sexually promiscuous or compulsive, having a talent for the arts, being mama's boys, or lacking in courage. Common stereotypes of gay women label them as being man-haters, tomboys, overly aggressive, or physically unattractive. Both gay men and women have been stereotyped as sexual predators or being fated to lead single lives of loneliness. Clinical experience has demonstrated that it is extremely stressful

for GLB individuals to have to constantly prove to heterosexual families, friends, employers, and co-religionists – and sometimes to themselves – that these are just stereotypes.

Moral Condemnations of Homosexuality

People who morally condemn homosexuality regard homosexual acts as intrinsically harmful to the individual, to the individual's spirit, and to the social fabric. Those who morally condemn same-sex activities believe that the tradition of philosophical, legal, and religious opposition to homosexuality is, in and of itself, sufficient reason to forbid the open expression of homosexuality in the modern world. Although critics of these condemnations attribute them to homophobia, antihomosexuality stemming from moral beliefs is often more complex. For example, religious and secular social conservatives who morally condemn homosexuality claim they do so out of love or concern for something else. In today's culture wars, such individuals and groups actively and vocally oppose gay and lesbian civil rights at the municipal, state, and federal level. Their political activities, which arise from their moral condemnations, are an ongoing source of stress for the GLB community.

The most severe forms of moral condemnation are embodied in laws that criminalize homosexual behavior, once referred to as sodomy laws. In the mid-twentieth century, homosexuality was illegal in all 50 states. By 2003, only 13 states still had such laws. Although rarely enforced, they were often a source of stress for GLB populations. Selective enforcement of sodomy laws allowed antihomosexual law enforcement communities to harass and discriminate against gay citizens. Furthermore, it was routinely argued in states with sodomy laws that gay couples, as potential criminals, should not be allowed to adopt their foster children. Obviously, criminalizing the sexual behaviors felt to be part of one's normal sexuality is extremely stressful to those so charged. In 2003, however, the U.S. Supreme Court ruled in the case of *Lawrence and Garner v. Texas* that all state sodomy laws were unconstitutional.

Moral condemnations of homosexuality hampered the early public health response to the AIDS epidemic and, in the process, added significantly to the losses and stress of the GLB community. Initially – and incorrectly – perceiving AIDS to be primarily a disease of gay men, public health systems were unprepared to talk frankly about either homosexuality or the kind of sexual practices that transmitted the virus. In 1987, 6 years after the initial cases were identified, President Reagan first mentioned the word "AIDS." Although teaching safe sex practices to sexually

active individuals was important in slowing the spread of the epidemic, moral disapproval of homosexuality made federal funding of such educational efforts illegal. Even though most global AIDS cases have been transmitted heterosexually, many still argue, usually in a condemnatory way, that homosexuality causes AIDS rather than unsafe sexual practices that spread the virus. After experiencing thousands of preventable deaths, it is not unsurprising that many in today's GLB community feel their own government does not value their lives.

In their more subtle forms, moral condemnations of homosexuality include an attitude known as love the sinner, hate the sin. In earlier historical eras, religious authorities condemned sodomites to eternal damnation, to prison, and sometimes even physical torture. However, science and medicine's redefinition of homosexuality, as well as the growing number of openly gay and lesbian voices, had an enormous impact on religious attitudes. Consequently, religious groups were forced to integrate the accumulating scientific data regarding the presumptive etiologies and naturalness of homosexuality.

Today's religious authorities find themselves having to contend with the spiritual needs of a growing number of gay people, as well as those of their families. Some religious denominations (i.e., Episcopalians and the United Church of Christ) have adopted an accepting model of homosexuality and openly embrace lesbian and gay worshippers. Some denominations even admit openly gay men and women as priests, rabbis, and ministers. However, acceptance has not been the majority approach.

Many religious authorities have instead chosen to embrace the homosexual but not the homosexuality. Contrary to mainstream scientific findings, antihomosexual religious authorities believe one chooses to be a homosexual despite the prohibitions against it – "choice" being a charged word in moral debates about homosexuality. In recent years, antihomosexual groups have been publicizing so-called sexual conversion or reparative therapies intended to change homosexuals. Rather than excluding gay people, they invite them back into the religious fold if they change their ways. However, the overstated claims surrounding conversion therapies have generated much interpersonal stress in many religious GLB individuals, particularly when family members pressure them to undergo the process. The public relations effort promoting these treatments is worrisome because (1) their efficacy is questionable, (2) the etiological theories on which they are based are not supported by current scientific research, (3) the practitioners of these treatments incorrectly label homosexuality as a mental disorder, (4) there is sparse documentation

of the cure rates and when the rates are documented they are quite low (27–35%), (5) long-term results are rarely validated, (6) there are an increasing number of anecdotal reports that these treatments may cause harm, and (7) the false hopes raised in offering such treatments often delay individuals or their family members from accepting them as gay. In fact, the American Psychiatric Association has issued a position statement asking ethical practitioners to refrain from such practices and to do no harm.

Religious beliefs that disapprove of homosexuality constitute substantive, ongoing stressors in the lives of gay people, particularly for those who had strict, fundamentalist, or orthodox religious upbringings. Because they may no longer be participating in religious activities, they are often estranged from their communities where family, social, and religious activities are deeply intertwined. The social isolation that ensues from gay people's fear of moral condemnation can also be stressful.

Antigay Violence

In 1990 the federal government passed a law ordering a study of hate crimes, including attacks on gay men and women, the first time a federal civil rights law covered sexual orientation. In one common presentation of gay-bashing, gangs of young men descend on neighborhoods where gay people meet. Armed with bats, clubs, or other weapons, they attack anyone they believe to be gay.

Since the much-publicized beating to death of Matthew Shepard in 1998, there has been increased awareness and documentation of antigay violence and gay-bashing incidents. Publicized cases have also led to the open articulation of previously unspoken cultural beliefs, for example, the belief that violence is an excusable response to an unwanted sexual advance from a gay man. Termed the "gay panic defense," no court has yet accepted it as justification for antigay violence. Its underlying rationale is that gay victims are responsible for provoking the violence against them. Others, however, argue that antigay violence is the inevitable outgrowth of unexamined and unstated beliefs implicit in homophobia, heterosexism, and moral condemnation. They argue that not everyone can distinguish the sinner from the sin and that those who marginalize homosexuality are either directly or indirectly responsible for the brutality against gay men and women.

Leaving the unanswered question of antigay violence's origins aside, most GLB people have legitimate concerns about it. Holding one's partners hand in public or a chaste kiss may provoke a physical attack.

Furthermore, antigay violence is believed to be an underreported phenomenon because its victims are often reluctant to come forward. Victims often blame themselves for being in the wrong place at the wrong time. Even gay people who have never been assaulted are sensitive to the ways in which members of the dominant culture, both in positions of authority or on its margins, can make threats and do as they please.

The Stress of Being in the Closet

In their developmental histories, gay men and lesbians often report periods of difficulty in acknowledging their homosexuality, both to themselves and to others. Children who grow up to be gay rarely receive family support in dealing with antihomosexual prejudices. On the contrary, beginning in childhood – and distinguishing them from racial and ethnic minorities – gay people are often subjected to the antihomosexual attitudes of their own families and communities. Consequently, gay men and lesbians may spend long periods of their lives unable to acknowledge their homosexuality, either to themselves or to others. Throughout childhood and the life cycle, even the suspicion of being gay can lead to teasing, ridicule, family censure, or even violence. Consequently, many gay people come to regard their homosexuality as an unpleasant fact they would rather not know about themselves, let alone admit to others; their homosexuality must be kept out of their awareness and separated from their public persona. The psychological means by which they accomplish this separation is called dissociation.

Dissociation, of course, is not limited to gay men and lesbians. Almost everyone is capable of pushing unwanted knowledge out of one's awareness. However, given some of the consequences of being openly gay – estrangement from family, loss of employment, loss of home, loss of child custody, loss of opportunity, loss of status and even blackmail – dissociation may seem a viable option for survival. Some closeted gay people can reflexively speak without revealing the gender of the person being discussed or without providing any gendered details of his or her personal life. Some closeted gay people marry and, on the surface, live their lives as if they were typical heterosexuals. Although some closeted gay men and women may not act on their homosexual feelings, others may develop secret sexual lives. In fact, through dissociation, a whole double life can be lived and yet never be acknowledged.

This psychological effort to solve one problem, however, creates others. It is stressful to continuously hide significant aspects of the self or to vigilantly

separate aspects of the self from others. Constantly hiding creates difficulties in accurately assessing other people's perceptions of oneself as well as recognizing one's own strengths. Dissociation's impact on self-esteem can also make it difficult to feel one's actual accomplishments as reflections of one's own abilities. Hiding takes its toll, particularly because the psychological effort needed to maintain a double life often leads to errors in judgment that further add to one's stress. Maintaining a heterosexual identity while engaging in secret homosexual activity may inevitably lead a closeted individual into compromising situations.

Those who find this psychological split untenable may come out as either gay, lesbian, or bisexual. Coming out involves putting into words the feelings and ideas that previously had no acceptable form of open expression. Calling oneself gay or lesbian is an effort to reach some measure of self-acceptance. The process is not just about revealing oneself to others. It is a realization that previously unacceptable homosexual feelings or desires are actually part of one's self. Coming out is fraught with danger because of the social stigma attached to homosexuality due to antihomosexual attitudes in the culture. Given the difficulties associated with revealing a gay identity, it may seem a wonder that anyone comes out at all.

For some, coming out may be against their will, precipitated by an act of malice. In such cases of outing, a hidden homosexual identity is exposed by those seeking revenge, political advantage or financial gain. Fear of being outed is a major stressor in the lives of many closeted GLB individuals. However, those who come out of their own accord describe the experience positively, as "a switch being turned on," "coming home," or "discovering who I really am." For them, as stressful as it may be to come out, it may be even more stressful to stay in the closet.

Further Reading

- Adams, H. E. (1996). Is homophobia associated with homosexual arousal? *Journal of Abnormal Psychology* 105(3), 440–445.
- Cabaj, R. P. and Stein, T. S. (eds.) (1996). *Textbook of homosexuality and mental health*. Washington, DC: American Psychiatric Press.
- Drescher, J. (1998). *Psychoanalytic therapy and the gay man*. Hillsdale, NJ: The Analytic Press.
- Drescher, J., Stein, T. S. and Byne, W. (2005). Homosexuality, gay and lesbian identities, and homosexual behavior. In: Sadock, B. & Sadock, V. (eds.) *Kaplan and Sadock's comprehensive textbook of psychiatry* (8th edn., pp. 1936–1965). Baltimore, MD: Williams and Wilkins.
- Drescher, J. and Zucker, K. J. (2006). *Ex-gay research: analyzing the Spitzer study and its relation to science, religion, politics, and culture*. New York: The Haworth Press.
- Herek, G. (1984). Beyond homophobia: a social psychological perspective on the attitudes toward lesbians and gay men. *Journal of Homosexuality* 10, 1–21. (Reprinted in DeCecco J. (ed.) (1985). *Bashers, baiters & bigots: homophobia in American society*, pp. 1–21. New York: Harrington Park Press.)
- Hoffman, L. G., Hevesi, A. G., Lynch, P. E., et al. (2000). Homophobia: analysis of a "permissible" prejudice; a public forum of the American Psychoanalytic Association and the American Psychoanalytic Foundation. *Journal of Gay & Lesbian Psychotherapy* 4(1), 5–53.
- Kohlberg, L. (1966). A cognitive-developmental analysis of children's sex role concepts and attitudes. In: Maccoby, E. (ed.) *The development of sex differences*, pp. 82–172. Stanford, CA: Stanford University Press.
- Shilts, R. (1987). *And the band played on: politics, people and the AIDS epidemic*. New York: St. Martin's Press.
- Sullivan, A. (ed.) (1997). *Same-sex marriage: pro and con*. New York: Vintage Books.
- Weinberg, G. (1972). *Society and the healthy homosexual*. New York: Anchor Books.

Hostility

L H Powell and K Williams

Rush University Medical Center, Chicago, IL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by L H Powell, volume 2, pp 407–412,

© 2000, Elsevier Inc.

History of the Concept

Conceptualization

Assessment

Link to Physical Disease in Longitudinal Studies

Mechanisms of Association

Intervention

Glossary

<i>Anger</i>	An unpleasant emotion, ranging in intensity from irritation or annoyance to fury or rage, that is an emotional expression of hostility.
<i>Cynicism</i>	A component of hostility characterized by the belief that others are motivated by selfish concerns.
<i>Hostile interactional style</i>	A style of interaction characterized by the tendency to challenge, evade questions, and become easily irritated.
<i>Hostility</i>	A general cognitive personality trait including a devaluation of others' worth and motives, an expectation of wrong-doing in others, an oppositional approach to others, and a desire to see others harmed.
<i>Mistrust</i>	A component of hostility characterized by the belief that others are likely to be provoking and hurtful.
<i>Type A behavior pattern</i>	The early conceptualization of coronary-prone behavior that featured two key components: excessive time urgency and free-floating hostility.

History of the Concept

In the 1960s, the concept of the type A behavior pattern was introduced by cardiologists Meyer Friedman and Ray Rosenman to describe individuals who possessed excessive time urgency and free-floating hostility and who, by virtue of this behavior pattern, were believed to be coronary-prone. This conceptualization fostered hundreds of investigations in the 1970s and 1980s aimed at replicating early associations with coronary disease, refining its measurement, and understanding its physiological underpinnings.

In 1980, a classic paper was published by Williams and his colleagues that suggested that the hostility component of the type A behavior pattern was its toxic core ([Figure 1](#)). Angiography patients were divided by gender, type A behavior, and hostility, and these classifications were related to occlusive disease. For both males and females, hostility was a better predictor of $\geq 75\%$ occlusion than type A behavior. This seminal investigation was subsequently replicated in a large number of studies using a variety of subjects and study designs and resulted in a shift in thinking away from type A behavior toward hostility as a key coronary-prone behavior.

Conceptualization

Hostility is a negative attitude toward others consisting of enmity, denigration, and ill will. It is a stable personality trait with stability coefficients of 0.85 over 1 or 4 years, 0.67 over 5 years, and 0.40 over 24 years. Hostility appears to be a multidimensional trait with from two to four factor analyzed components. There is general agreement that the two key components are cynicism, or the belief that others are motivated by selfish concerns, and mistrust, or the belief that others are likely to be provoking and hurtful. Other components include such things as a hostile attributional style, which is the tendency to construe the actions of others as involving aggressive intent, and a hostile interactional style characterized by the tendency to challenge others, evade questions, and become easily irritated. Smith has argued for a transactional view of hostility, which includes aspects of both the person and the environment. When both hostility and social support are considered jointly, three distinct components emerge: (1) nonhostile and high socially supported healthy individuals, (2) highly hostile and low socially supported submissive individuals, and (3) highly hostile and high socially supported aggressive individuals.

Although hostility is often linked to the risk factors of anger and aggression, it is distinct from them in that it refers to a cognitive personality trait. Anger is an emotion that includes reactions ranging from mild irritation to intense anger and can be a transitory feeling state or an enduring predisposition. Aggression is a behavior defined as attacking or hurtful actions toward others, whether the harm is physical or verbal. Correlations among these characteristics are moderate, suggesting that they are related, but nonetheless distinct.

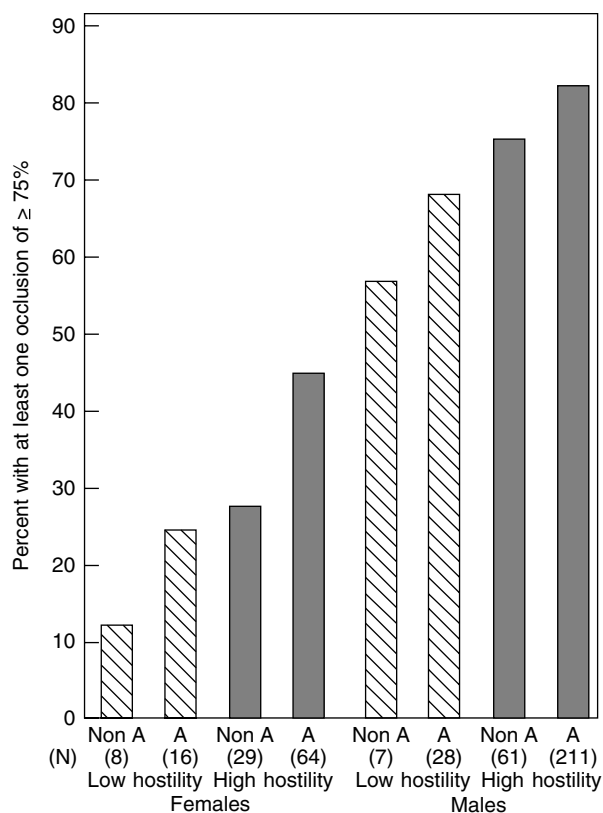


Figure 1 Relation of type A behavior pattern, hostility, and gender to presence of significant coronary occlusions. From Williams, R. B., et al. (1980). *Psychosomatic Medicine* 42, 539–549, with permission.

Studies that have examined the sociodemographic correlates of hostility have found that it increases with age, is higher in men than in women, is higher in minorities than in Caucasians, and is higher in individuals of lower than in those of higher socioeconomic status. Moreover, there is consistent evidence that hostility is associated with smoking, elevated body mass index, and heavy alcohol consumption. Since these correlates are, in themselves, risk factors for disease, they can serve as confounders for any association between hostility and disease. Confounding in this case would be making the mistake that the true association is between hostility and coronary disease, when in fact the true association is actually between one or more of these established coronary risk factors (which are linked to hostility) and coronary disease. It is important to control for these potential confounders in the design or analysis of studies assessing the health impact of hostility.

Assessment

There are three popular assessment measures of hostility, all of which have subscales. Two of these are self-report questionnaires and the third is a behavioral rating. [Table 1](#) presents selected items from these scales.

The Cook-Medley Hostility Scale

The Cook-Medley Hostility Scale is a subscale of the Minnesota Multiphasic Personality Inventory

Table 1 Selected items from the most popular hostility assessment scales

SELF-REPORT QUESTIONNAIRES

Cook-Medley Hostility Scale: Mark each item as “true” or “false”

Cynicism	I think most people would lie to get ahead
	Most people are honest chiefly through fear of being caught
Mistrust/hostile attributions	It is safer to trust nobody
	I tend to be on guard with people who are somewhat more friendly than I had expected
Hostile affect	Some of my family have habits that bother and annoy me very much
Aggressive responding	I have at times had to be rough with people who were rude or annoying

Buss-Durkee Hostility Inventory: Mark each item as “true” or “false”

Expressed hostility	I lose my temper easily but get over it
	When people yell, I yell back
	I raise my voice when arguing
Experiential hostility	I don't seem to get what's coming to me (resentment)
	I know that people tend to talk about me behind my back (suspicion)
	My motto is "Never trust strangers" (mistrust)

BEHAVIORAL RATINGS

Interpersonal Hostility Assessment Technique: Behavioral ratings from a structured interview, based upon style of interaction

Direct challenges	Direct or explicit challenge of the question or the interviewer
Indirect challenges	Implication that the answer was obvious or the interviewer was stupid for having asked it
Hostile withholding or evasion	Avoidance or refusal to answer a question, associated hostile tone and intent not to answer
Irritation	Irritated tone, impatience, exasperation with the interview or interviewer, aroused reliving of negative events, condescension or snide remarks, harsh generalizations, punched words with angry emphasis

(MMPI) that is composed of 52 true–false items. Several factor analyses of this scale have produced slightly different solutions ranging from two to four factors. The two-factor solution includes the components of cynicism and mistrust/hostile attributions. The four-factor solution includes the components of cynicism, mistrust/hostile attributions, hostile affect, and aggressive responding. Studies of hostility as a risk factor have used the entire scale, the cynicism subscale, or the Abbreviated Cook-Medley (ACM) with the four subscales. The major limitation of the Cook-Medley scale is that it correlates consistently with depression and anxiety, characteristics that may be outside of the conceptual domain of hostility.

The Buss-Durkee Hostility Inventory

The Buss-Durkee Hostility Inventory is composed of 66 true–false items that are grouped into seven subscales: assault, indirect hostility, irritability, negativism, resentment, suspicion, and verbal hostility. In factor analytic studies, it has been consistently shown to assess two correlated dimensions, expressive hostility and experiential hostility. Expressive hostility assesses verbal and physical aggressiveness and is more highly correlated with antagonism than with neuroticism. Experiential hostility assesses the subjective experience of resentment, suspicion, mistrust, and irritation and is more highly correlated with neuroticism than with antagonism.

The Interpersonal Hostility Assessment Technique (IHAT)

The IHAT is an evolution of the structured interview for type A behavior in which the focus is on behavioral ratings of a hostile interpersonal style. It provides more objective criteria for assessing hostility than the subjective judgments that have characterized past interview ratings. Four types of hostile behaviors are rated from type A interview scripts, including direct challenges to the interviewer, indirect challenges to the interviewer, irritation, and hostile withholding of information or evasion. IHAT ratings have been correlated ($r = 0.32$) with potential for hostility ratings from the structured interview for type A behavior and have been shown to be stable over a 4-year period ($r = 0.69$).

Link to Physical Disease in Longitudinal Studies

Figure 2 summarizes the longitudinal studies that have examined a link between hostility and physical disease. Approximately half of all longitudinal studies have found the expected association.

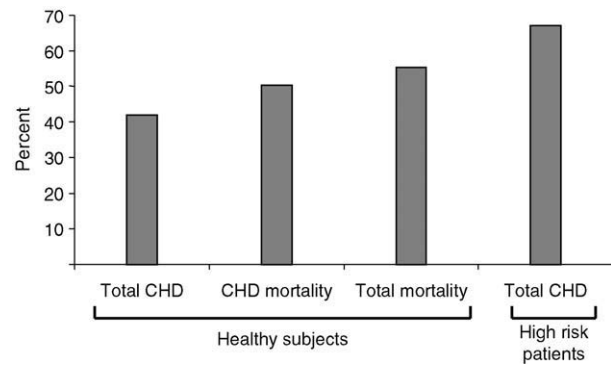


Figure 2 Percentage of prospective studies that have found a positive link between hostility and physical health.

Coronary Heart Disease in Healthy Subjects

To date there have been 12 longitudinal studies that have examined the role of hostility on the incidence of coronary heart disease (coronary mortality or myocardial infarction), six longitudinal studies that have examined the role of hostility on coronary heart disease (CHD) mortality, and two longitudinal studies that have examined the role of hostility on subclinical cardiovascular disease. Positive associations were found in five (41.7%) of the CHD studies, three (50%) of the CHD mortality studies, and two (100%) of the subclinical cardiovascular disease studies. Evidence exists that hostility is equally predictive in men and women, older and younger subjects, African American and Caucasian subjects, and the total scale and the abbreviated subscales.

Since hostility is embedded in a behavioral complex of other risk factors for coronary disease that includes lower education, higher body mass index, current smoking, and heavy alcohol consumption, these risk factors must be controlled for in multivariate analyses to avoid the problem of confounding. When the prospective studies that showed significant associations were examined for the adequacy of their control for these confounders, none of the studies controlled for all of these factors and most controlled only one or two of them. In one study, direct adjustment for these potential confounders eliminated a prior positive association. This suggests that any link between hostility and CHD may not be independent of standard risk factors but instead may be mediated by them.

Total Mortality in Healthy Subjects

There have been nine prospective studies of the link between hostility and total mortality. Of these, five (55.6%) found a positive association. The positive findings included both men and women and included both the total Cook-Medley scale and an abbreviated

version. None of these studies controlled for education, and very few controlled for body mass index. Thus, the bulk of the evidence suggests that an association exists. It is not known, however, whether this association is independent of, or mediated by, established risk factors.

Coronary Heart Disease in High-Risk Subjects

There have been nine prospective studies of hostility and clinical CHD outcomes in patients with evidence of existing coronary disease. Of these, six (66.5%) found a positive association. These studies are characterized by inadequate control for confounding, with one exception. In a study of postmenopausal women with existing CHD from the Heart and Estrogen/Progestin Replacement Study (HERS) cohort, both the total Cook-Medley scale and the cynicism subscale predicted total CHD events after excellent controls for demographic and CHD risk factor confounders in multivariate analyses. Thus, the evidence is slightly stronger that an association, which is independent of the risk incurred by established risk factors, exists between hostility and CHD in high-risk patients than in their healthy counterparts.

Mechanisms of Association

There are three key mechanisms by which hostility may be associated with physical disease. The first is a direct mechanism whereby hostility triggers neuroendocrine pathogenic mechanisms for physical disease. The second is an indirect association whereby hostility fosters elevations in established risk factors and these risk factors increase risk for physical disease. The third suggests that hostility is embedded in a constellation of psychosocial risk factors, which together increase the psychosocial burden and, concomitantly, the risk of physical disease. Literature on potential mechanisms is growing rapidly, particularly evidence for a direct physiological link between hostility and disease.

Physiological

There is consistent evidence that hostility exerts a direct effect on CHD and total mortality through the mechanisms of amplified cardiovascular reactivity and neuroendocrine responses to stressors. Evidence exists for an association between hostility and such pathophysiological mechanisms as heart rate and blood pressure reactivity, increased platelet aggregation, vasoconstriction of atherosclerotic arteries, silent ischemia, decreased ejection fraction, decreased threshold for ventricular fibrillation, and decreased heart rate

variability. There also is evidence that hostility activates the sympathetic branch, and/or suppresses the parasympathetic branch, of the autonomic nervous system, mediated through epinephrine, norepinephrine, testosterone, and serotonin. Recent research suggests other possible physiological mechanisms involving the metabolic syndrome and inflammatory processes. Hostility has been associated with the metabolic syndrome of high levels of plasma lipids, insulin resistance, visceral body fat, and elevated blood pressure. Hostility has also been associated with inflammatory processes, specifically pro-inflammatory cytokines, tumor necrosis factor- α , and interleukin-6.

Health Behaviors

Evidence is accumulating that hostility is associated with smoking, heavy alcohol consumption, a sedentary lifestyle, obesity and overweight, elevated blood pressure, and lowered HDL cholesterol. Moreover, hostile individuals may be more likely to ignore health warnings and fail to comply with medical regimens. Two studies exist that have directly tested this model by examining the association between hostility and CHD before and after adjusting for established risk factors. One found that the association was maintained after adjustment; the other found that the association was eliminated. Thus, there is evidence from one study to support this mechanism.

Psychosocial Vulnerability

Hostile individuals not only may be hostile but also may have other psychosocial characteristics that have been observed to be coronary prone, including depression, anxiety, low social support, and/or general distress. Correlations between measures of hostility and these other psychosocial risk factors are between 0.25 and 0.30. Thus, it may not be the hostility alone but the overall adverse psychosocial risk factor profile that is associated with physical disease. For example, one study found that general distress was a better predictor than hostility of 6-month recurrence in patients with myocardial infarction. Another study found that although hostility is related to lower social support, when it is available, it provides fewer physiological benefits to hostile persons.

Intervention

Intervention for hostility can take one of two forms. Pharmacological interventions focus on blocking the impact of key physiological mechanisms on cardiac targets. Behavioral interventions focus on altering the cognitive, behavioral, and emotional concomitants of hostility as a way to reduce the frequency and

intensity with which the physiological pathways are mobilized.

Pharmacological

β -adrenergic blocking drugs block the impact of increased sympathetic tone on the heart, triggered by hostility, and thereby reduce the hemodynamic response that may lead to plaque disruption. With this therapy, hostile individuals may construe the stressor as threatening, but the concomitant sympathetic arousal has little, if any, impact on heart rate or blood pressure. Aspirin has antiplatelet properties and thus may prevent sympathetically induced platelet aggregation that leads to thrombus formation. Estrogen therapy in women with existing coronary disease does not weaken the ability of hostility to predict recurrent coronary events.

Behavioral

Effective behavioral interventions for hostility must begin with the development of sensitivity to signs and symptoms of hostility in the individual him- or herself. This is a difficult task because hostility tends to be well rationalized, based upon real or perceived insults from the social environment. Interventions are most effective when they focus on the development of skills and competencies in all of the domains of hostility. Relaxation therapies help the individual to develop the skill to switch from sympathetic arousal to parasympathetic relaxation. Assertiveness training can help individuals to replace hostile interactional styles with less arousing and more effective alternatives. Cognitive therapy can help individuals to identify cognitions, such as blaming others, that are associated with arousal and to replace them with alternatives that are more rational and less arousing. The development of efficacy in controlling emotional arousal in response to environmental challenges has the effect of improving self-esteem, improving interpersonal effectiveness, and producing more positive feedback from the social environment. Progressively, as skills develop, the enhanced interpersonal effectiveness can challenge core cynical and mistrusting beliefs.

See Also the Following Articles

Aggression; Anger; Heart Disease/Attack; Type A Personality, Type B Personality.

Further Reading

- Barefoot, J. C., Dodge, K. A., Peterson, B. L., Dahlstrom, W. G. and Williams, R. B., Jr. (1989). The Cook-Medley hostility scale: item content and ability to predict survival. *Psychosomatic Medicine* 51, 46–57.
- Buss, A. H. and Durkee, A. (1957). An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology* 21, 343–349.
- Chaput, L. A., Adams, S. H., Simon, J. A., Blumenthal, R. S., Vittinghoff, E., Lin, F., et al. (2002). Hostility predicts recurrent events among postmenopausal women with coronary heart disease. *American Journal of Epidemiology* 156, 1092–1099.
- Cook, W. W. and Medley, D. M. (1954). Proposed hostility and pharisaic virtue scales for the MMPI. *Journal of Applied Psychology* 38, 414–418.
- Haney, T. L., Maynard, K. E., Houseworth, S. J., et al. (1996). Interpersonal hostility assessment technique: description and validation against the criterion of coronary artery disease. *Journal of Personal Assessment* 66, 386–401.
- McEwen, B. S. and Stellar, E. (1993). Stress and the individual. Mechanisms leading to disease. *Archives of Internal Medicine* 153, 2093–2101.
- Miller, T. Q., Smith, T. W., Turner, C. W., Guijarro, M. L. and Hallet, A. J. (1996). A meta-analytic review of research on hostility and physical health. *Psychiatric Bulletin* 119, 322–348.
- Siegmán, A. W. and Smith, T. W. (1994). *Anger, hostility, and the heart*. Hillsdale, NJ: Lawrence Erlbaum.
- Smith, T. W., Glazer, K., Ruiz, J. M. and Gallo, L. C. (2004). Hostility, anger, aggressiveness, and coronary heart disease: an interpersonal perspective on personality, emotion, and health. *Journal of Personality* 72, 1217–1270.
- Suarez, E. C. (2003). Joint effect of hostility and severity of depressive symptoms on plasma interleukin-6 concentration. *Psychosomatic Medicine* 65, 523–527.
- Suls, J. and Wan, C. K. (1993). The relationship between trait hostility and cardiovascular reactivity: a quantitative review and analysis. *Psychophysiology* 30, 615–626.
- Williams, R. and Williams, V. (1993). *Anger kills. Seventeen strategies for controlling the hostility that can harm your health*. New York: Times Books, Random House.

HPA Alterations in PTSD

R Yehuda

Mount Sinai School of Medicine, New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction

Basal Cortisol Levels in Posttraumatic Stress Disorder

Cortisol Levels in Response to Stress

Basal Corticotropin Releasing Hormone and
Adrenocorticotrophic Hormone Levels in Posttraumatic
Stress Disorder

Glucocorticoid Receptors in Posttraumatic Stress Disorder

The Dexamethasone Suppression Test in Posttraumatic
Stress Disorder

The Metyrapone Stimulation Test

The Corticotropin Releasing Hormone Challenge Test and
Adrenocorticotrophic Hormone Stimulation Test in
Posttraumatic Stress Disorder

Findings of Cortisol in the Acute Aftermath of Trauma

Conclusion

Glossary

*Hypothalamic-
pituitary-
adrenal (HPA)*

The HPA system is one of the major two neuroendocrine stress response systems of the body. Stress stimulates the release of the hypothalamic neurohormonal peptides, corticotropin releasing hormone and arginine vasopressin, that stimulate secretion of the pituitary adrenocorticotrophic hormone (ACTH), which in turn stimulates the secretion of adrenal corticosteroids. In addition to its role in the stress response, the HPA system plays a role in and serves as a biomarker of mental disorders and especially depression and posttraumatic stress disorder (PTSD).

Introduction

The study of the neuroendocrinology of posttraumatic stress disorder (PTSD) has presented an interesting paradox – some of the alterations described have not historically been associated with pathological processes. In particular, initial observations of low cortisol levels in a disorder precipitated by extreme stress directly contradicted the emerging and popular formulation of hormonal responses to stress, the glucocorticoid cascade hypothesis, which posited that stress-related psychopathology involves hypercortisolism as either a cause or consequence of disorder. It is now clear that insufficient glucocorticoid signaling

may produce detrimental consequences as robust as those associated with glucocorticoid toxicity.

With respect to PTSD, the majority of studies demonstrate alterations consistent with an enhanced negative feedback inhibition of cortisol on the pituitary and/or an overall hyperreactivity of other target tissues (adrenal gland and hypothalamus). Findings of low cortisol and increased reactivity of the pituitary in PTSD are also consistent with reduced adrenal output, although it is more difficult to fully explain cortisol elevations in response to stress and neuroendocrine challenge with this model.

Insofar as there have been discrepancies in the literature with regard to basal hormone levels, investigators more recently theorized that models of enhanced negative feedback, increased hypothalamic-pituitary-adrenal (HPA) reactivity, and reduced adrenal capacity may explain different facets of the neuroendocrinology of PTSD, including preexisting risk factors that may be related to certain types of early experiences, at least in certain people who develop PTSD. Notwithstanding that some aspects of HPA alterations may predate exposure to a traumatic event that precipitates PTSD, alterations associated with enhanced negative feedback inhibition also appear to develop over time in response to the complex biological demands of extreme trauma and its aftermath. Moreover, the findings of increased HPA reactivity may sometimes reflect a more nonspecific response to ongoing environmental challenges associated with having chronic PTSD. Certainly the absence of cortisol alterations in some studies imply that the alterations associated with low cortisol and enhanced negative feedback are present only in a biological subtype of PTSD. The observations in the aggregate, and the alternative models of pathology or adaptation suggested by them, must be clearly understood when using neuroendocrine data in PTSD to identify targets for drug development.

Basal Cortisol Levels in Posttraumatic Stress Disorder

The majority of the evidence supports the conclusion that cortisol levels in PTSD are different from those observed in acute and chronic stress and in major depression; that is, they are lower, rather than higher, compared to normal. However, it should be noted that there have been many reports that failed to observe low cortisol levels and (although fewer) that have reported cortisol increases in PTSD. This discrepancy may be due to the fact that basal alterations of cortisol in PTSD are subtle. That is, even

when cortisol is reported to be significantly lower in PTSD, levels fall within the normal endocrinological range. There are numerous other sources of potential variability in such studies related to the selection of subjects and comparison groups, adequate sample size, and inclusion/exclusion criteria, as well as considerations that are specific to the methods of collecting and assaying cortisol levels, that can explain the discrepancies in the findings. However, the simplest explanation for the disparate observations is that cortisol levels may not represent stable trait markers and are likely to fluctuate, making it difficult to consistently observe group differences.

There is better agreement in studies of circadian rhythmicity of cortisol in PTSD, although these are fewer in numbers. The advantage of these studies is that plasma samples can be obtained continuously throughout the 24-h cycle under controlled conditions. Thus, not only is it possible to confirm overall mean cortisol release but also to determine the times of day at which alterations are most likely to be present and note changes in circadian rhythm parameters. Indeed, it appears that the major difference between PTSD and non-PTSD groups is that cortisol levels are lower in the late night and remain lower for a longer period of time in PTSD during hours when subjects are normally sleeping.

Even in cases in which there is failure to find group differences in cortisol levels, there are often correlations within the PTSD group with indices of PTSD symptom severity or trauma exposure. Most often, investigators have reported negative correlations between cortisol levels (particularly integrated mean 24-h cortisol levels) and PTSD symptoms in combat veterans. Low cortisol levels have also been negatively correlated with duration since trauma, implying that low cortisol is associated with early traumatization. In fact, in some recent studies, cortisol levels were inversely related to childhood traumatic experiences, including emotional abuse, even in the absence of PTSD.

Cortisol levels have also been correlated with findings from brain imaging studies in PTSD. In one report, there was a positive relationship between cortisol levels and hippocampal acetylaspartate (NAA), a marker of cell atrophy presumed to reflect changes in neuronal density or metabolism, in subjects with PTSD, suggesting that, rather than having neurotoxic effects, cortisol levels in PTSD may have a trophic effect on the hippocampus. Similarly, cortisol levels in PTSD were negatively correlated with medial temporal lobe perfusion, whereas anterior cingulate perfusion and cortisol levels were positively correlated in PTSD but negatively correlated in trauma survivors without PTSD. These inverse correlations may result from an augmented negative hippocampal effect secondary to increased sensitivity of brain

glucocorticoid receptors, which would account for the inverse correlation in PTSD despite equal cortisol levels in both the PTSD and non-PTSD groups. On the other hand, the positive correlation between regional cerebral blood flow in the frontocingulate transitional cortex and cortisol levels in PTSD may reflect unsuccessful attempts of the frontocingulate transitional cortex to terminate the stress response, which has also been linked with low cortisol.

Cortisol may be related to specific or state-dependent features of the disorder, such as comorbid depression or the time course of the disorder. There is a long and rich tradition in psychosomatic medicine underscoring the importance of examining intrapsychic correlates of individual differences in cortisol levels in PTSD, reflecting emotional arousal and opposing antiarousal disengagement defense mechanisms or other coping styles. Finally, recent evidence also suggests that cortisol levels further decrease over time and in old age in trauma survivors.

Cortisol Levels in Response to Stress

Although cortisol levels may be lower at baseline, they may actually be elevated in PTSD relative to comparison subjects when challenged by physiological or psychological stressors. In particular, subjects with PTSD appear to show more anticipatory anxiety to laboratory stressors, which is reflected by increased cortisol levels. These studies provide at least some evidence that the adrenal is capable of producing ample cortisol levels at baseline and in response to stress. In fact, transient increases in cortisol levels are consistent with the notion of a more generalized HPA axis reactivity in PTSD, as reflected by enhanced negative feedback inhibition.

Basal Corticotropin Releasing Hormone and Adrenocorticotrophic Hormone Levels in Posttraumatic Stress Disorder

Interestingly, all published reports examining the concentration of corticotropin releasing hormone (CRH) in cerebrospinal fluid (CSF) in PTSD have concluded that levels of this peptide are increased in PTSD. The assessment of CSF CRH does not necessarily provide a good estimate of hypothalamic CRH release but, rather, an estimate of both the hypothalamic and extra-hypothalamic release of this neuropeptide. Nonetheless, that there is increased CRH release is somewhat of a paradox considering the lower cortisol levels. With respect to adrenocorticotrophic hormone (ACTH), the majority of studies have reported no detectable differences in ACTH levels between PTSD and comparison subjects even when cortisol levels obtained from the same sample were found to be significantly lower.

Lower cortisol levels in the face of normal ACTH levels can reflect a relatively decreased adrenal output. Yet, under circumstances of classic adrenal insufficiency, there is usually increased ACTH release compared to normal levels. Thus, in PTSD, there may be an additional component of feedback on the pituitary acting to depress ACTH levels so that they appear normal rather than elevated. Indeed, elevations in ACTH are expected not only from a reduced adrenal output but also from increased CRH stimulation. On the other hand, the adrenal output in PTSD may be relatively decreased but not enough to substantially affect ACTH levels. In any event, the normal ACTH levels in PTSD in the context of the other findings suggest a more complex model of the regulatory influences on the pituitary in this disorder than reduced adrenal insufficiency.

Glucocorticoid Receptors in Posttraumatic Stress Disorder

Type II glucocorticoid receptors (GRs) are expressed in ACTH- and CRH-producing neurons of the pituitary, hypothalamus, and hippocampus and mediate most systemic glucocorticoid effects, particularly those related to stress responsiveness. Low circulating levels of a hormone or neurotransmitter can result in increased numbers of available GRs to improve response capacity and facilitate homeostasis. However, alterations in the number and sensitivity of both type I (mineralocorticoid) and type II GRs can also significantly influence HPA axis activity and, in particular, can regulate hormone levels by mediating the strength of negative feedback. In PTSD, there is evidence for both an increased number of GRs (measured in lymphocytes) and increased responsiveness of the GR.

Observations regarding the cellular immune response in PTSD are also consistent with enhanced GR responsiveness in the periphery, as indicated by increased beclomethasone-induced vasoconstriction and enhanced delayed-type hypersensitivity of skin test responses. Because immune responses, like endocrine ones, can be multiply regulated, these studies provide only indirect evidence of GR responsiveness. However, when considered in the context of the observation that PTSD patients showed an increased expression of the receptors in all lymphocyte subpopulations, despite a relatively lower quantity of intracellular GRs, as determined by flow cytometry, and in the face of lower ambient cortisol levels, the findings more convincingly support an enhanced sensitivity of the GRs to glucocorticoids. Furthermore, there appears to be an absence of alterations of the mineralocorticoid receptor in PTSD, as investigated by examining the cortisol and ACTH response to spiro-lactone following CRH stimulation.

With respect to increased glucocorticoid responsiveness, this was recently determined using an *in vitro* paradigm in which mononuclear leukocytes were incubated with a series of concentrations of dexamethasone (DEX) to determine the rate of inhibition of lysozyme activity. Subjects with PTSD showed evidence of a greater sensitivity to glucocorticoids as reflected by a significantly lower mean lysozyme median inhibitory concentration of DEX (IC_{50-DEX} ; in nanomoles per liter). The lysozyme IC_{50-DEX} was significantly correlated with age at exposure to the first traumatic event in subjects with PTSD. The number of cytosolic GRs was correlated with age at exposure to the focal traumatic event.

The Dexamethasone Suppression Test in Posttraumatic Stress Disorder

In contrast to observations regarding ambient cortisol levels, results using the dexamethasone suppression test (DST) present a more consistent view of reduced cortisol suppression in response to DEX administration. The DST provides a direct test of the effects of GR activation in the pituitary on ACTH secretion, and cortisol levels following DEX administration are thus interpreted as an estimate of the strength of negative feedback inhibition, provided that the adrenal response to ACTH is not altered. Using 0.50 mg DEX, most investigators have found a more dramatic reduction in plasma cortisol levels. This exaggerated cortisol suppression in response to DEX has been observed in PTSD related to people with combat exposure, survivors of childhood sexual abuse, Holocaust survivors, children exposed to disaster, women exposed to domestic violence, and mixed groups of trauma survivors.

The Metyrapone Stimulation Test

Metyrapone prevents adrenal steroidogenesis by blocking the conversion of 11-deoxycortisol to cortisol, thereby unmasking the pituitary gland from the influences of negative feedback inhibition. If a sufficiently high dose of metyrapone is used, such that an almost complete suppression of cortisol is achieved, this allows a direct examination of pituitary release of ACTH without the potentially confounding effects of differing ambient cortisol levels. When metyrapone is administered in the morning when HPA axis activity is relatively high, maximal pituitary activity can be achieved, facilitating an evaluation of group differences in pituitary capability. The administration of 2.5 mg metyrapone in the morning resulted in a similar and almost complete reduction in cortisol levels in both PTSD and normal subjects (i.e., removal of negative feedback inhibition) but in a higher

increase in ACTH and 11-deoxycortisol in combat Vietnam veterans with PTSD, compared to nonexposed subjects. In the context of low cortisol levels and increased CSF CRH levels, the findings supported the hypothesis of a stronger negative feedback inhibition in PTSD. Both pituitary and adrenal insufficiency are not likely to result in an increased ACTH response to the removal of negative feedback inhibition because the former is associated with an attenuated ACTH response and reduced adrenal output would not necessarily affect the ACTH response. To the extent that ambient cortisol levels are lower than normal, an increased ACTH response following the removal of negative feedback inhibition implies that, when negative feedback is intact, it is strong enough to inhibit ACTH and cortisol. The increased ACTH response is most easily explained by increased suprapituitary activation; however, a sufficiently strong negative feedback inhibition would account for the augmented ACTH response even in the absence of hypothalamic CRH hypersecretion.

Evidence of increased ACTH was not observed when a lower dose of metyrapone was used, administered over a 3-h period at night (750 mg at 7:00 a.m. and 10:00 a.m.). However, interestingly, the manipulation produced a more robust suppression of cortisol in comparison subjects. Using a similar methodology, the effects of metyrapone was used to evaluate CRH effects on sleep. ACTH and cortisol levels were increased in the PTSD group relative to the controls the next morning, suggesting that the same dose of metyrapone did not produce the same degree of adrenal suppression of cortisol synthesis. Under these conditions, it is difficult to evaluate the true effect on ACTH and 11-deoxycortisol, which depends on achieving complete cortisol suppression, or at least the same degree of cortisol suppression, in the two groups. The endocrine response to metyrapone in this study do not support the model of reduced adrenal capacity because this would have yielded a large ratio of ACTH-to-cortisol release, yet the mean ACTH/cortisol ratio prior to metyrapone was no different in the PTSD subjects and the controls. On the other hand, the mean ACTH/cortisol ratio postmetyrapone was lower, although not significantly, suggesting, if anything, an exaggerated negative feedback rather than reduced adrenal capacity.

The Corticotropin Releasing Hormone Challenge Test and Adrenocorticotrophic Hormone Stimulation Test in Posttraumatic Stress Disorder

The infusion of exogenous CRH increases ACTH levels and provides a test of pituitary sensitivity. A blunted ACTH is classically thought to reflect a

reduced sensitivity of the pituitary to CRH, reflecting a downregulation of pituitary CRH receptors secondary to CRH hypersecretion. The results of studies in PTSD subjects using this challenge have also been mixed. Some studies have demonstrated a blunted ACTH response to PTSD; however, insofar as this blunting did not occur in the presence of hypercortisolism, as in depression, the attenuated ACTH response may reflect an increased negative feedback inhibition of the pituitary secondary to increased GR number or sensitivity. This explanation supports the idea of CRH hypersecretion in PTSD and explains the pituitary desensitization and resultant lack of hypercortisolism as arising from a stronger negative feedback inhibition. An augmented ACTH response to CRH has also been reported, which the authors interpreted as reflecting increased reactivity of the pituitary, although here too there was no evidence for increased cortisol levels.

Findings of Cortisol in the Acute Aftermath of Trauma

Recent data provided some support for the idea that low cortisol levels may be an early predictor of PTSD rather than a consequence of this condition. Low cortisol levels in the immediate aftermath of a motor vehicle accident predicted the development of PTSD in a group of accident victims consecutively presenting in an emergency room in two studies. In a sample of people who survived a natural disaster, cortisol levels were similarly found to be lowest in those with highest PTSD scores at 1 month posttrauma; however cortisol levels were not predictive of symptoms at 1 year. In the acute aftermath of rape, low cortisol levels were associated with prior rape or assault, themselves risk factors for PTSD, but not with the development of PTSD *per se*.

These findings imply that cortisol levels might have been lower in trauma survivors who subsequently developed PTSD even before their exposure to trauma and might therefore represent a preexisting risk factor. Consistent with this, low 24-h urinary cortisol levels in adult children of Holocaust survivors were specifically associated with the risk factor of parental PTSD. Interestingly, the risk factor of parental PTSD in offspring of Holocaust survivors was also associated with an increased incidence of traumatic childhood antecedents. In this study, both the presence of subject-rated parental PTSD and scores reflecting childhood emotional abuse were associated with low cortisol levels in offspring. Thus, it may be that low cortisol levels occur in those who have experienced an adverse event early in life and then remain different from those not exposed to early adversity. Although there might reasonably be HPA axis

fluctuations in the aftermath of stress, and even differences in the magnitude of such responses compared to those not exposed to trauma early in life, HPA parameters would subsequently recover to their (abnormal) prestress baseline.

Low cortisol levels may impede the process of biological recovery from stress, resulting in a cascade of alterations that lead to intrusive recollections of the event, avoidance of reminders of the event, and symptoms of hyperarousal. This failure may represent an alternative trajectory to the normal process of adaptation and recovery after a traumatic event.

One of the most compelling lines of evidence supporting the hypothesis that lower cortisol levels may be an important pathway to the development of PTSD symptoms are the results of studies showing that the administration of stress doses of hydrocortisone prevents the development of PTSD and traumatic memories and attenuate these symptoms once present. These findings support the idea that low cortisol levels may facilitate the development of PTSD in response to an overwhelming biological demand, at least in some circumstances.

Conclusion

Cortisol levels are most often found to be lower than normal in PTSD subjects, but can also be similar to or greater than those in comparison subjects. Findings of changes in circadian rhythm suggest that there may be regulatory influences that result in a greater dynamic range of cortisol release over the diurnal cycle in PTSD. Thus, although cortisol levels may be generally lower, the adrenal gland is certainly capable of producing adequate amounts of cortisol in response to challenge.

The model of enhanced negative feedback inhibition is compatible with the idea that there may be transient elevations in cortisol, but suggests that, when present, these increases are shorter-lived due to a more efficient containment of ACTH release as a result of enhanced GR activation. An increased negative feedback inhibition results in reduced cortisol levels under ambient conditions. In contrast to other models of endocrinopathy, which identify specific and usually singular primary alterations in endocrine organs and/or regulation, the model of enhanced negative feedback inhibition in PTSD is in large part descriptive and currently offers little explanation of why some individuals show such alterations of the HPA axis following exposure to traumatic experiences and others do not.

The wide range of observations observed in the neuroendocrinology of PTSD underscores the important observation that HPA response patterns in PTSD are fundamentally in the normal range and do not

reflect endocrinopathy. In endocrinological disorders, in which there is usually a lesion in one or more target tissue or biosynthetic pathway, endocrine methods can usually isolate the problem with the appropriate tests and then obtain rather consistent results. In psychiatric disorders, neuroendocrine alterations may be subtle, and therefore, when using standard endocrine tools to examine these alterations, there is a high probability of failing to observe all the alterations consistent with a neuroendocrine explanation of the pathology in tandem or of obtaining disparate results within the same patient group owing to a stronger compensation or reregulation of the HPA axis following challenge.

The next generation studies should aim to apply more rigorous tests of neuroendocrinology of PTSD based on the appropriate developmental issues and in consideration of the longitudinal course of the disorder and the individual differences that affect these processes. No doubt such studies will require a closer examination of a wide range of biological responses, including the cellular and molecular mechanisms involved in adaptation to stress, and an understanding of the relationship between the endocrine findings and other identified biological alterations in PTSD.

Acknowledgments

This work was supported by MH 49555, MH 55-7531, and MERIT review funding.

See Also the Following Articles

Depression and Manic-Depressive Illness; Depression Models; Depression, Immunological Aspects; Dexamethasone Suppression Test (DST); Hypocortisolism and Stress; Neuroendocrine Systems.

Further Reading

- Baker, D. G., West, S. A., Nicholson, W. E., et al. (1999). Serial CSF corticotropin-releasing hormone levels and adrenocortical activity in combat veterans with posttraumatic stress disorder. *American Journal of Psychiatry* 156, 585–588[errata] 986.
- Bremner, J. D., Vythilingma, M., Vermetten, E., et al. (2003). Cortisol response to a cognitive stress challenge in posttraumatic stress disorder related to childhood abuse. *Psychoneuroendocrinology* 28, 733–750.
- Boscarino, J. A. (1996). Posttraumatic stress disorder, exposure to combat, and lower plasma cortisol among Vietnam veterans: findings and clinical implications. *Journal of Consulting and Clinical Psychology* 64, 191–201.
- Delahanty, D. L., Raimonde, A. J. and Spoonster, E. (2000). Initial posttraumatic urinary cortisol levels predict subsequent PTSD symptoms in motor vehicle accident victims [in process citation]. *Biological Psychiatry* 48, 940–947.
- Holsboer, F. (2000). The corticosteroid receptor hypothesis of depression. *Neuropsychopharmacology* 23, 477–501.

- Raison, C. L. and Miller, A. H. (2003). When not enough is too much: the role of insufficient glucocorticoid signaling in the pathophysiology of stress related disorders. *American Journal of Psychiatry* **160**, 1554–1565.
- Schelling, G., Briegel, J., Roozendaal, B., et al. (2001). The effect of stress doses of hydrocortisone during septic shock on posttraumatic stress disorder in survivors. *Biological Psychiatry* **50**, 978–985.
- Yehuda, R. (2002). Status of cortisol findings in PTSD. *Psychiatric Clinics of North America* **25**, 341–368.
- Yehuda, R., Boissoneau, D., Lowy, M. T., et al. (1995). Dose-response changes in plasma cortisol and lymphocyte glucocorticoid receptors following dexamethasone administration in combat veterans with and without posttraumatic stress disorder. *Archives of General Psychiatry* **52**, 583–593.
- Yehuda, R., Halligan, S. L. and Grossman, R. (2001). Childhood trauma and risk for PTSD: relationship to intergenerational effects of trauma, parental PTSD and cortisol excretion. *Development and Psychopathology* **2001**, 733–753.
- Yehuda, R., Levengood, R. A., Schmeidler, J., et al. (1996). Increased pituitary activation following metyrapone administration in post-traumatic stress disorder. *Psychoneuroendocrinology* **21**, 1–16.
- Yehuda, R., Shalev, A. Y. and McFarlane, A. C. (1998). Predicting the development of posttraumatic stress disorder from the acute response to a traumatic event. *Biological Psychiatry* **44**, 1305–1313.
- Yehuda, R., Teicher, M. H., Trestman, R. L., et al. (1996). Cortisol regulation in posttraumatic stress disorder and major depression: a chronobiological analysis. *Biological Psychiatry* **40**, 79–88.

HPA Role in Depression See: HPA Alterations in PTSD; Depression and Manic-Depressive Illness; Depression Models; Dexamethasone Suppression Test (DST).

HSF (Heat Shock Transcription Factor) See: Chaperone Proteins and Chaperonopathies; Heat Shock Response, Overview.

Hurricane Katrina Disaster, Stress Effects of

C Piotrowski

University of West Florida, Pensacola, FL USA

© 2007 Elsevier Inc. All rights reserved.

Calamity

Hurricane Katrina

Methodological Issues

Impact on Human Populations

Glossary

Chaos theory A scientific paradigm that suggests that chaos and complexity are inherent to explaining phenomena; it aids in the understanding of how organizations respond, adapt, change, function, and renew.

Direct-contact victims

Emergency and rescue workers who searched for survivors and dead bodies, as well as morgue personnel.

Disaster mental health

A competency-based intervention model, based on disaster-related psychopathology, that addresses a community's mental health needs.

First-responders

All emergency personnel that aid in the immediate aftermath of a disaster, such as police, firefighters, rescue workers, medical staff, and special response units. The costliest hurricane to hit the United States prior to Hurricane Katrina; it caused \$44 billion in damage in 1992.

Hurricane Andrew

Mass-casualty event

An event that involves multidimensional stressors that may result in differential patterns of posttraumatic reactions, both short and long term.

<i>Northeast quadrant</i>	The storm quadrant with the strongest winds and rain (because hurricanes circulate counterclockwise). In Hurricane Katrina, this meant that the areas east of New Orleans, for 160 miles, received the brunt of the storm's wrath.
<i>Self-efficacy</i>	An individual's belief that his or her actions and behavior will produce the desired outcomes; coined by personality theorist Albert Bandura.

Calamity

Without a doubt, Hurricane Katrina (August 29, 2005) was the most destructive and costly natural disaster in U.S. history. Months after the storm, the total destruction, including the human toll, is still being calculated. Perhaps as a key indicator of the enormity of the devastation, 6 months after the disaster news accounts on every aspect of the tragedy still appear each night on the national news networks, the number of deaths are provided as an estimate, and over 100 of the deceased victims have not yet been identified. The numbers, to date, allow an appreciation of the impact:

- Deaths: ~1500, either during or immediately after the storm.
- Homeless residents: 800,000.
- Victims left unemployed: ~550,000.
- Displaced people: 2 million.
- Homes destroyed: 350,000.
- Penetration of the storm surge along the Louisiana, Mississippi, and Alabama coast was measured in miles not feet.
- Property losses: ~\$35 billion.
- Estimates for reclamation: 2 years to clear 22 million tons of waste and debris.
- Small businesses not in operation: 50%.
- Cars submerged under water: 200,000.
- Residents of the greater New Orleans area who may have to relocate permanently: ~75%.
- Area of the city of New Orleans that was flooded: ~70%.
- Evacuees separated from their pets: ~80%.
- Estimate of total economic costs and damage: varies widely, but will probably reach \$200 billion.

Hurricane Katrina

Hurricane Katrina was both an immensely powerful and a large storm. Initial forecasts estimated winds of 145 mph at landfall, but revised figures from the National Hurricane Center in Miami reported maximum sustained winds of approximately 125 mph; officially, the storm was a high category 3 hurricane.

Hurricane-force winds (>74 mph) extended 160 miles out from the eye of the storm. Unbelievably, the city of Waveland, Mississippi, was struck by a 28-ft storm surge. Several hours after the eye of the storm passed New Orleans, the levee system that protected the city was either breached or undermined in three strategic areas; within hours, approximately two-thirds of the city was submerged under up to 12 ft of water. Many residents drowned, and thousands sought higher ground.

The initial governmental response, particularly that of the Federal Emergency Management Agency (FEMA), can be regarded as a complete debacle; however, many regular citizens, volunteers, and health-care personnel extended themselves to save lives. Unfortunately, most residents were left to fend for themselves for almost a week before basic necessities arrived. The result was untold human suffering. Small coastal cities along the gulf coast from Port Sulphur, Louisiana, to Mobile, Alabama, were hard hit by high winds, torrential rain, and a destructive storm surge; some small towns on the Mississippi and Alabama coast were literally wiped off the map.

Outside of the impact area, the country felt the residual impact of the calamity: anger over government malaise, a sense of national shame, and high energy prices. Undoubtedly, the social consciousness of the public was severely tested, particularly with regard to the racial issues and poverty that were brought to the forefront in the aftermath of the storm. In this regard, Hurricane Katrina will always be considered both a natural and human-made disaster.

Methodological Issues

Although the United States suffered several devastating natural disasters in its history (e.g., the San Francisco earthquake, the Galveston hurricane, and Mississippi River floods), the calamity of Hurricane Katrina has no precedent with regard to human suffering, social displacement, structural damage, and financial cost. Summarily, a continuous stretch of 200 miles of the U.S. coast was adversely impacted, including a population of approximately 3.6 million residents. At the time of this writing (6 months post-disaster), there is limited empirical research on the storm, although many articles have appeared in general media publications. Because I am a disaster researcher, I can attest to the fact that there are a plethora of research studies underway, particularly at universities in the New Orleans area.

Researchers should be cautioned about several unique features of investigating this disaster. First and foremost, Hurricane Katrina was not a unitary impact event; that is, from a methodological

perspective, not all regions (and populations) were impacted in the same manner or experienced the same level or type of disaster. For example, New Orleans basically experienced a flood event and was not greatly impacted by hurricane winds (the city lay west of the eye of the storm). However, smaller coastal communities to the east of the eye (coastal Louisiana, Mississippi, and Alabama) felt the full brunt of hurricane-force winds and devastating storm surge. Some residents had recoverable property damage, whereas others were completely wiped out. Some residents evacuated to safety, whereas others rode out the storm. Many victims evacuated temporarily, whereas others were displaced to cities throughout the United States and many in this latter group will probably not return to their home areas. Moreover, there was a great disparity in the amount of governmental support and financial assistance that individuals received, commensurate with the amount of damage, need for shelter, and sustenance needs. Therefore, researchers need to be cognizant of these factors when studying select populations and be aware that not everyone experienced Hurricane Katrina in the same way. Green reviewed some of the difficulties encountered in designing sound disaster research studies, and several investigators have considered the heterogeneous nature of studying the effects of hurricanes on different cohort groups. Furthermore, due to the overwhelming nature of the disaster and its aftermath, Hurricane Katrina, for many victims, was a multiple traumatic loss: that is, a loss of loved ones, pets, residence, job, security, and the future.

Impact on Human Populations

Flooding

Much of New Orleans was severely flooded due to the failure of levee structures. Both natural and human-made disasters can result in flooding, for example, the Buffalo Creek dam collapse in 1972 and the Great Midwest Mississippi floods in 1993. The effects of extensive flooding were studied in the aftermath of Hurricane Floyd, which struck North Carolina in 1999. Emotional exhaustion, grief reactions and depression, posttraumatic stress disorder (PTSD) symptomatology, and somatic complaints were identified as the major sequela in several studies. The aged seemed to be the most affected.

Health Issues

The health consequences of natural disasters have been a major research area. For example, in the 1983 South Australian bush fires, Japan's 1995

Hanshin-Awaji earthquake, and California's 1994 Northridge earthquake, impacted populations were found to be vulnerable to a host of medical maladies, particularly stress-related conditions such as hypertension, gastrointestinal disorders, diabetes, and mental illness. Non-stress-related conditions were typically not affected. Most conditions were found to improve within 15 months postdisaster. Investigators have reported on the impact of immunological changes after natural disasters and found that adaptive coping to stress effects mediated immunological responses in affected populations.

Mental Health Symptoms

Psychiatric morbidity, particularly stress-related conditions, have been extensively studied in both natural and human-made disasters. Specifically, PTSD, depression, and substance abuse disorders have been reported to significantly increase following large-scale disasters. Researchers caution that acute stress disorder is prognostically important as a marker for the development of moderate or full PTSD symptoms in the long term. One study on mental health symptomatology following Hurricane Andrew reported that 36% of the sample had PTSD, 30% experienced major depression, and 20% showed signs of anxiety disorders. In fact, 56% had significant symptoms persisting beyond 6 months posthurricane. Due to the high prevalence of mental health problems experienced in the wake of hurricanes, researchers point to the critical importance of coping self-efficacy perceptions as a mediator between acute stress response and long-term distress following natural disasters.

Because geographical relocation was a major aspect of the Katrina disaster, it should be an area of particular research focus. Prior studies in this area concluded that displaced individuals in natural disasters exhibit more intense levels of psychopathology, particularly depression. Perhaps, this is the result of unresolved grief on the part of victims who were not only relocated but also lost their homes and needed to start over elsewhere. Moreover, in a study of family life following Hurricane Hugo in 1989, researchers found the divorce rate increased. In this regard, there have been recent media ad campaigns across the impacted Gulf Coast areas that address post-Hurricane Katrina family stress issues such as child and wife abuse.

Recovery Issues

There are a number of governmental recovery plans in place across the Gulf Coast, although their implementation seems to be progressing slowly. Research following Hurricane Andrew in 1992

indicated that disruption during rebuilding phases contributed to the exacerbation of mental health symptoms. In recent years, social scientists have advocated disaster mental health-planning initiatives via national organizations such as the American Psychological Association to address preparedness planning, mitigation, agency coordination, and community resource development. Particularly in the context of the complexity of disasters, interagency cooperation and coordination are crucial to meeting the social and psychological needs of victims posthurricane.

By all accounts, the governmental response in the immediate aftermath of Katrina and the subsequent lethargic recovery efforts can only be described as appalling. This unfortunate state of affairs can only impede individual and collective efforts at achieving some sense of normality among the affected population. Instead, much of the goodwill and recovery assistance to the affected population came from local church groups and organizations such as the Red Cross. Indeed, this is also what occurred after the double hit of Hurricanes Erin and Opal in Pensacola, Florida.

Special Populations

It is incumbent that researchers pay particular attention to population groups such as the aged, health-delivery workers, children, the disabled, first-responders, and impoverished citizens. Unique populations cope with stressors in diverse ways, and stress reactions tend to be heterogeneous. This was the finding of a study on children following the Oklahoma City bombing. Furthermore, the critical function of denial in both child and adult populations has received theoretical support in prior research studies. Last, there has been much media attention given to race and poverty issues in the aftermath of Katrina; this will be an important factor in ongoing research, although recent research findings on this issue are equivocal.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Earthquakes, Stress Effects of; Floods, Stress Effects of; Refugees, Stress in; Life Events and Health; Neurosis; Posttraumatic Stress Disorder – Clinical.

Further Reading

Benight, C. and Harper, M. (2002). Coping self-efficacy perceptions as a mediator between acute stress response

and long-term distress following natural disasters. *Journal of Traumatic Stress* 15, 177–186.

David, D., Mellman, T. A., Mendoza, L. M., et al. (1996). Psychiatric morbidity following Hurricane Andrew. *Journal of Traumatic Stress* 9, 607–612.

Elliot, J.R. and Pais, J. (2006). Race, class, and Hurricane Katrina: social differences in human responses to disaster. *Social Science Research* 35, 295–321.

Green, B. L. (1991). Evaluating the effects of disasters. *Psychological Assessment* 3, 538–546.

Jones, R. T., Frary, R., Cunningham, P., et al. (2001). Psychological effects of Hurricane Andrew on ethnic minority and Caucasian children and adolescents: a case study. *Cultural Diversity & Ethnic Minority Psychology* 7, 103–108.

McMillen, C., North, C., Mosley, M., et al. (2002). Untangling the psychiatric comorbidity of posttraumatic stress disorder in a sample of flood survivors. *Comprehensive Psychiatry* 43, 478–485.

Myers, D. and Wee, D. F. (2005). *Disaster mental health services*. New York: Brunner-Routledge.

Najarian, L. M., Goenjian, A., Pelcovitz, D., et al. (2001). The effect of relocation after a natural disaster. *Journal of Traumatic Stress* 14, 511–526.

Norris, F., Friedman, M. J. and Watson, P. J. (2002). 60,000 disaster victims speak. II: Summary and implications of the disaster mental health research. *Psychiatry* 65, 240–260.

Norris, F., Perilla, J. L. and Murphy, A. D. (2001). Postdisaster stress in the United States and Mexico: a cross-cultural test of the conceptual model of posttraumatic stress disorder. *Journal of Abnormal Psychology* 110, 553–563.

Piotrowski, C. (2006). Hurricane Katrina and organization development: part 1, implications of chaos theory. *Organization Development Journal* 24(3), 10–19.

Piotrowski, C. and Armstrong, T. (1998). Satisfaction with relief agencies during Hurricanes Erin and Opal. *Psychological Reports* 82, 413–414.

Piotrowski, C., Armstrong, T. and Stopp, H. (1997). Stress factors in the aftermath of Hurricanes Erin and Opal: data from small business owners. *Psychological Reports* 80, 1387–1391.

Piotrowski, C. and Dunham, F. (1983). Locus of control orientation and perception of “Hurricane” in fifth graders. *Journal of General Psychology* 109, 119–127.

Rotton, J., Dubitsky, S., Milov, A., et al. (1997). Distress, elevated cortisol, cognitive deficits, and illness following a natural disaster. *Journal of Environmental Psychology* 17, 85–98.

Solomon, G. F., Segerstrom, S., Grohr, P., et al. (1997). Shaking up immunity: psychological and immunologic changes after a natural disaster. *Psychosomatic Medicine* 59, 142–143.

Hydrocortisone See: Corticosteroids and Stress; Glucocorticoids, Overview.

11 β -Hydroxysteroid Dehydrogenases

J R Seckl

Edinburgh University, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J R Seckl, volume 2, pp 413–420, © 2000, Elsevier Inc.

Definition: 11 β -Hydroxysteroid Dehydrogenase

The Mineralocorticoid Receptor Paradox

11 β -Hydroxysteroid Dehydrogenase Type 2 in the Adult Central Nervous System

11 β -Hydroxysteroid Dehydrogenase Type 2 in the Developing Central Nervous System

11 β -Hydroxysteroid Dehydrogenase Type 2 and Fetal Programming

Programming the Hypothalamic-Pituitary-Adrenal Axis and Behavior

11 β -Hydroxysteroid Dehydrogenase Type 1 in the Brain

11 β -Hydroxysteroid Dehydrogenase Type 1 and the Hypothalamic-Pituitary-Adrenal Axis

11 β -Hydroxysteroid Dehydrogenase Type 1 and Brain Aging

Conclusion

Glossary

11 β -Hydroxysteroid dehydrogenase (11 β -HSD) An enzyme that catalyzes the interconversion of active cortisol (corticosterone in rats and mice) and inert cortisone (11-dehydrocorticosterone).

11 β -HSD type 1 (11 β -HSD1) An isozyme that predominantly catalyzes the reduction of inert cortisone to active cortisol in intact cells and organs.

11 β -HSD type 2 (11 β -HSD2) An isozyme that catalyzes the rapid dehydrogenation of active cortisol to inert cortisone.

Aromatase The enzyme that catalyzes the conversion of androgens to estrogens.

Carbenoxolone A nonselective 11 β -HSD inhibitor.

Cushing's syndrome The generic description of the phenotype observed with circulating glucocorticoid excess, due either to endogenous overproduction of cortisol or adrenocorticotrophic hormone (ACTH) or to exogenous pharmacotherapy.

Glucocorticoid receptor (GR) Intracellular receptor for active cortisol.

Intracrine

The process whereby intracellular enzymic transformations of steroids alter their access to nuclear receptors.

Metabolic syndrome

The co-association of risk factors for atheromatous disease (insulin resistance, type 2 diabetes, dyslipidemia, and hypertension) and abdominal (visceral) obesity.

Mineralocorticoid receptor (MR)

Intracellular receptor for active cortisol, selectivity for aldosterone *in vivo* generated by 11 β -HSD2.

Syndrome of apparent mineralocorticoid excess

A congenital syndrome of hypertension, hypernatremia, and hypokalemia due to inappropriate MR activation by glucocorticoids because of deleterious mutations of the 11 β -HSD type 2 gene and hence lack of the enzyme in the kidney.

Definition: 11 β -Hydroxysteroid Dehydrogenase

Elevated plasma cortisol or glucocorticoid pharmacotherapy has myriad adverse effects ranging from obesity, diabetes, and hypertension to depression and memory impairments. Glucocorticoid action is classically believed to be determined by the product of circulating glucocorticoid levels and the tissue density of corticosteroid receptors (GR and MR). However, recent discoveries suggest a key role for local enzymic tissue metabolism of glucocorticoids in determining receptor activation. The main enzymes involved are the 11 β -hydroxysteroid dehydrogenases (11 β -HSDs). These catalyze the interconversion of active cortisol (in humans and most mammals) and corticosterone (in rats and mice) and inert 11-keto forms (cortisone, 11-dehydrocorticosterone) ([Figure 1](#)). The roles of these enzymes in determining glucocorticoid actions in the central nervous system (CNS) and periphery are now becoming apparent.

The Mineralocorticoid Receptor Paradox

Purified MR binds aldosterone, cortisol, and corticosterone with equal and high affinity *in vitro*, but *in vivo*

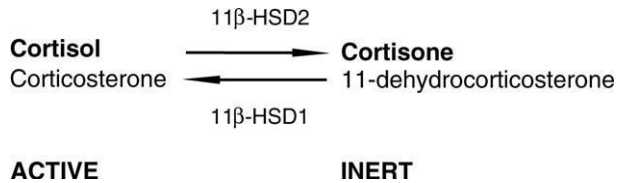


Figure 1 11 β -Hydroxysteroid dehydrogenases. Type 1 acts as a ketoreductase, regenerating active cortisol (corticosterone), whereas type 2 acts as a dehydrogenase, rapidly inactivating these physiological glucocorticoids.

only aldosterone binds to MR in the kidney, although there is much more glucocorticoid in the circulation. In resolution of this paradox, in 1988 groups in Edinburgh and Melbourne reported that 11 β -HSD was a prereceptor enzymic mechanism that bestowed aldosterone selectivity on the intrinsically nonselective MR *in vivo*. The 11 β -HSD activity associated with MR in the distal nephron acted as a dehydrogenase, rapidly inactivating physiological glucocorticoids (cortisol, corticosterone) to their inert 11-keto forms. Such 11 β -HSD activity forms an intracellular barrier to glucocorticoids, preventing the occupation of MR. When the enzyme is congenitally deficient in the syndrome of apparent mineralocorticoid excess or is inhibited by liquorice, glucocorticoids inappropriately access MR, causing sodium retention, hypertension, and hypokalemia. Although an NADP(H)-associated 11 β -HSD activity (now called 11 β -HSD1) and its encoding cDNA were isolated from rat liver by Monder and colleagues in the mid-1980s, this was shown not to be responsible for renal MR selectivity: its encoding gene HSD11B1 was intact in patients with the syndrome. Subsequently a second 11 β -HSD isozyme (11 β -HSD2) was cloned from the kidney and placenta in a variety of species. This isozyme is a high-affinity 11 β -dehydrogenase, which uses NAD as cofactor and is highly expressed in the distal nephron and placenta. Patients with the syndrome of apparent mineralocorticoid excess were homozygous or compound heterozygous for deleterious mutations in the HSD11B2 gene, and 11 β -HSD2 knockout mice faithfully recapitulated the syndrome. This work underlined the fundamental principle that steroid access to nuclear receptors *in vivo* is regulated by prereceptor metabolism. Other examples include 5- and 5'-monodeiodinases (thyroid hormone receptors), 5 α -reductase (androgen receptor), and aromatase (estrogen receptors).

11 β -Hydroxysteroid Dehydrogenase Type 2 in the Adult Central Nervous System

Although the vast majority of MR in the CNS are nonselective and bind glucocorticoids (cortisol,

corticosterone) *in vivo*, it is anticipated that any selective effects of aldosterone in the CNS will require MR gated by local 11 β -HSD2. In this regard, aldosterone mediates the central regulation of blood pressure and salt appetite. Thus, aldosterone infused intracerebroventricularly (ICV) elevates blood pressure in the rat, whereas the same dose given peripherally is ineffective, demonstrating a specific central action. Corticosterone ICV does not reproduce this action, and indeed antagonizes aldosterone hypertension. Similarly, aldosterone exerts central actions on salt appetite that cannot be reproduced by corticosterone. 11 β -HSD2 action in the CNS is therefore inferred. This is supported by ICV infusion of the 11 β -HSD inhibitor carbenoxolone, which increases blood pressure in the rat, an effect reversed by an MR antagonist, suggesting central 11 β -HSD2 prevents access of glucocorticoids to central MR. Although 11 β -HSD2 mRNA is only found in a few discrete sites in the adult rat brain, these loci also express MR and are believed to underlie the selective central effects of aldosterone to increase blood pressure (nucleus tractus solitarius, NTS), and modulate salt appetite (ventromedial hypothalamus, subcommissural organ, and central nucleus of the amygdala). In the mouse there is an even more limited distribution of 11 β -HSD2, with expression confined to the NTS. 11 β -HSD2 was not found in the human forebrain, although the brain stem has not been investigated. Thus, there is very little 11 β -HSD2 in the adult brain, located appropriately for the few CNS actions specifically ascribed to this hormone.

11 β -Hydroxysteroid Dehydrogenase Type 2 in the Developing Central Nervous System

Prenatal glucocorticoid administration reduces brain weight at birth; delays the maturation of neurons, glia, and cerebral vasculature; and retards CNS myelination. Given such widespread effects, it is unsurprising that GR and MR are highly expressed in the developing brain. However, whether these receptors are occupied by endogenous glucocorticoids until late gestation is uncertain because there is also plentiful 11 β -HSD2 in the CNS at midgestation in rodents and humans. This presumably functions to protect vulnerable developing cells from premature glucocorticoid action. Strikingly, 11 β -HSD2 expression is dramatically switched off in a CNS at the end of midgestation, coinciding with the terminal stage of brain nucleus development. This appears to be an exquisitely timed system of cellular protection from circulating glucocorticoids.

11 β -Hydroxysteroid Dehydrogenase Type 2 and Fetal Programming

Human epidemiological studies show that low birth weight at term is associated with an increased later occurrence of cardiovascular (hypertension and ischemic heart disease) and metabolic disorders (type 2 diabetes) in adult life. In rodents, sheep, and humans, excessive exposure to glucocorticoids during gestation, either bypassing of fetoplacental 11 β -HSD2 with nonsubstrate glucocorticoid such as dexamethasone or using 11 β -HSD inhibitors such as carbenoxolone, reduces birth weight. Although in such models, the weight deficit is regained soon after birth, the adult offspring show increased blood pressure and serum insulin and glucose levels. These effects are underpinned by an increased tissue sensitivity to glucocorticoids via increased GR density in key metabolic loci such as the liver and visceral adipose tissue depots.

Programming the Hypothalamic-Pituitary-Adrenal Axis and Behavior

Prenatal 11 β -HSD2 inhibition or bypass with the nonsubstrate dexamethasone permanently increases basal plasma corticosterone levels in the offspring in rats. The density of GR and MR in the hippocampus are reduced in this model, which is anticipated to reduce hypothalamic-pituitary-adrenal (HPA) axis feedback sensitivity. Moreover the glucocorticoid excess may drive, at least in part, the hypertension and hyperglycemia observed in such prenatal programming models. Maternal undernutrition also lowers placental 11 β -HSD2 and alters adult offspring metabolic, cardiovascular, and HPA axis function, suggesting that HPA programming may be a common outcome of prenatal environmental challenge. Overexposure to glucocorticoids *in utero* also leads to alterations in adult behavior to produce a phenotype reminiscent of anxiety. Prenatal glucocorticoid exposure also affects the developing dopaminergic system with clear implications for proposed developmental contributions to schizo-affective, attention-deficit hyperactivity and extrapyramidal disorders. Mechanistically, prenatal dexamethasone or 11 β -HSD inhibition increases corticotropin-releasing hormone (CRH) mRNA levels specifically in the central nucleus of the amygdala, a key locus for the effects of the neuropeptide on the expression of fear and anxiety. Intriguingly, fetoplacental 11 β -HSD inhibition also permanently increases GR and MR in the offspring amygdala. A direct relationship between brain GR levels and anxiety behaviors is supported by transgenic mice with selective loss of GR in the forebrain, which exhibit reduced fearfulness.

Intriguingly, low-birth-weight humans also exhibit HPA hyperactivity in adult life.

11 β -Hydroxysteroid Dehydrogenase Type 1 in the Brain

So what was the role of Monder's original 11 β -HSD (type 1), highly expressed in liver, adipose tissue and the brain? Studies in intact cells from liver, fat, lung, and other sites have shown that 11 β -HSD1, although bidirectional in tissue homogenates, is a predominant 11 β -ketoreductase in intact cells. This reaction regenerates active glucocorticoids from inert 11-keto forms. A similar directional predominance is seen in the intact liver *ex vivo* and in rodents and humans *in vivo*. In the brain, 11 β -HSD1-like immunoreactivity and mRNA are widespread. Fetal rat hippocampal cells in culture express 11 β -HSD1 but no 11 β -HSD2. Intact hippocampal cells show only 11 β -reductase activity. This reaction direction increases intracellular glucocorticoid levels. In searching for functions, it was noted that licorice-derived 11 β -HSD inhibitors given to rats alter hypothalamic secretagogues in pituitary portal blood *in vivo* and activate a series of regions of the CNS, including the hippocampus, a key glucocorticoid target with high GR and MR density where glucocorticoids act to suppress HPA activity and to modulate cognitive functions.

11 β -Hydroxysteroid Dehydrogenase Type 1 and the Hypothalamic-Pituitary-Adrenal Axis

11 β -HSD1 is expressed in the hippocampus, paraventricular nucleus of the hypothalamus (PVN) and pituitary, key sites of glucocorticoid negative feedback. 11 β -HSD1 knockout (11 β -HSD1^{-/-}) mice have hypertrophied adrenals and elevated basal adrenocorticotrophic hormone (ACTH) and corticosterone levels at the nadir of the rhythm. These effects are compatible with the requirement for increased glucocorticoid production from the adrenal cortex in this mouse secondary to increased metabolic clearance of corticosterone because of the lack of glucocorticoid regeneration, notably in the liver. Unexpectedly, basal plasma corticosterone levels are also elevated in 11 β -HSD1^{-/-} mice, suggesting reduced glucocorticoid feedback on the HPA axis and increased forward drive on the HPA system. Altered central drive is rather unlikely because CRH mRNA levels in the PVN are unaltered. In contrast, some findings suggest that 11 β -HSD1^{-/-} mice have attenuated negative feedback because they show less suppression of

corticosterone to exogenous cortisol 2 h prior to restraint stress. Interestingly, recent data show that obese Zucker rats have reduced 11 β -HSD1 in the hippocampus, which may explain why this strain has reduced HPA feedback sensitivity and hence has increased corticosterone levels. Altered brain 11 β -HSD1 has been reported in pregnancy, when the HPA axis becomes refractory to stressful stimulation. Interestingly, 11 β -HSD1 activity is selectively increased in the PVN in pregnancy, so locally increased 11 β -HSD1 may further amplify glucocorticoid feedback and hence attenuate HPA activity.

11 β -Hydroxysteroid Dehydrogenase Type 1 and Brain Aging

A subgroup of aged rodents and humans shows the association of cognitive decline with a chronic elevation of plasma corticosterone levels. If glucocorticoids are maintained at low levels in rats, by adrenalectomy and low-dose corticosterone replacement, age-related impairments are reduced. The hippocampus requires glucocorticoids for neuronal function and survival but is especially vulnerable to the adverse effects of chronic glucocorticoid excess, which causes atrophy of dendrites and cognitive dysfunction. Therefore, chronic glucocorticoid overexposure has been implicated in the pathogenesis of age-related decline in hippocampus-associated cognitive functions.

In primary cultures of hippocampal cells, 11 β -HSD1 potentiates kainic acid-induced neurotoxicity, not only by corticosterone but equally by intrinsically inert 11-dehydrocorticosterone, an effect prevented by carbenoxolone. 11-Dehydrocorticosterone increases kainic acid hippocampal toxicity *in vivo*, an effect also blocked by carbenoxolone. Crucially, old 11 β -HSD1^{-/-} mice learn as well as young mice in hippocampus tasks and avoid the cognitive decline seen in the majority of aged wild-type mice. This cognitive protection in aged 11 β -HSD1^{-/-} mice is associated with reduced intrahippocampal corticosterone levels, indicating the potency of intracellular metabolism by 11 β -HSDs in determining effective glucocorticoid action on target receptors. *In situ* hybridization studies in postmortem human brain material show high expression of 11 β -HSD1 mRNA in the hippocampus, prefrontal cortex, and cerebellum, mirroring rodent findings. In two small randomized double-blind placebo-controlled cross-over studies, the administration of the 11 β -HSD inhibitor carbenoxolone improved verbal fluency after 4 weeks in ten healthy elderly men and improved verbal memory after 6 weeks in 12 elderly patients with type 2 diabetes. Plasma cortisol was unaltered, suggesting that 11 β -HSD1 does not modulate HPA axis feedback

in humans, at least at the morning peak. Intriguingly, 11 β -HSD1 may play a role in pathogenesis because rare haplotypes of the HSD11B1 gene are associated with a risk of Alzheimer's disease.

Conclusion

11 β -HSDs interconvert active glucocorticoids and inert 11-keto forms. This seemingly obscure reaction is important in peripheral tissues and the CNS, determining glucocorticoid access to nuclear receptors and hence biological effects in health and disease. 11 β -HSD1 is an 11 β -ketoreductase that amplifies intracellular glucocorticoid levels. Studies in 11 β -HSD1^{-/-} mice show a role in HPA feedback. Deficiency of 11 β -HSD1 reduces intracerebral glucocorticoid levels and prevents the emergence of cognitive deficits with aging. 11 β -HSD2 is a potent 11 β -dehydrogenase, which is little expressed in the adult CNS but which plays an emerging role in preventing premature glucocorticoid actions on the developing nervous system. The maintenance of fetoplacental 11 β -HSD2 and the inhibition of 11 β -HSD1 have therapeutic potential.

Further Reading

- De Kloet, E. R. (2004). Hormones and the stressed brain. Stress: current neuroendocrine and genetic approaches. *Annals of the New York Academy of Sciences* **1018**, 1–15.
- Diaz, R., Brown, R. W. and Seckl, J. R. (1998). Ontogeny of mRNAs encoding glucocorticoid and mineralocorticoid receptors and 11 β -hydroxysteroid dehydrogenases in prenatal rat brain development reveal complex control of glucocorticoid action. *Journal of Neuroscience* **18**, 2570–2580.
- Edwards, C. R. W., Benediktsson, R., and Lindsay, R. (1993). Dysfunction of the placental glucocorticoid barrier: a link between the foetal environment and adult hypertension? *Lancet* **341**, 355–357.
- Harris, H. J., Kotelevtsev, Y. and Mullins, J. J. 2001. 11 β -hydroxysteroid dehydrogenase type 1 null mice have altered hypothalamic-pituitary-adrenal axis activity: a novel control of glucocorticoid feedback. *Endocrinology* **142**, 114–120.
- Kotelevtsev, Y., Holmes, M. C. and Burchell, A. (1997). 11 β -hydroxysteroid dehydrogenase type 1 knockout mice show attenuated glucocorticoid inducible responses and resist hyperglycaemia on obesity or stress. *Proceedings of the National Academy of Sciences USA* **94**, 14924–14929.
- Kotelevtsev, Y., Brown, R. W. and Fleming, S. (1999). Hypertension in mice lacking 11 β -hydroxysteroid dehydrogenase type 2. *Journal of Clinical Investigation* **103**, 683–689.
- Liu, D., Dioria, J. and Tannenbaum, B. (1997). Maternal care, hippocampal glucocorticoid receptors, and

- hypothalamic-pituitary-adrenal responses to stress. *Science* 277, 1659–1662.
- Matthews, S. G. (2002). Early programming of the hypothalamo-pituitary-adrenal axis. *Trends in Endocrinology and Metabolism* 130, 373–380.
- McEwen, B. S. (1999). Stress and hippocampal plasticity. *Annual Review of Neuroscience* 22, 105–122.
- Moisan, M.-P., Seckl, J. R. and Edwards, C. R. W. (1990). 11 β -hydroxysteroid dehydrogenase bioactivity and messenger RNA expression in rat forebrain: localization in hypothalamus, hippocampus and cortex. *Endocrinology* 127, 1450–1455.
- Mune, T., Rogerson, F. M. and Nikkilä, H. (1995). Human hypertension caused by mutations in the kidney isozyme of 11 β -hydroxysteroid dehydrogenase. *Nature Genetics* 10, 394–399.
- Phillips, D. I. W., Walker, B. R. and Reynolds, R. M. (2000). Low birth weight predicts elevated plasma cortisol concentrations in adults from 3 populations. *Hypertension* 35(6), 1301–1306.
- Rajan, V., Edwards, C. R. W. and Seckl, J. R. (1996). 11 β -hydroxysteroid dehydrogenase in cultured hippocampal cells reactivates inert 11-dehydrocorticosterone, potentiating neurotoxicity. *Journal of Neuroscience* 16, 65–70.
- Roland, B. L., Li, K. X. Z. and Funder, J. W. (1995). Hybridization histochemical localization of 11 β -hydroxysteroid dehydrogenase type 2 in rat brain. *Endocrinology* 136, 4697–4700.
- Sandeep, T., Yau, J., MacCullich, A., et al. (2004). Effects of 11 β -hydroxysteroid dehydrogenase inhibition on cognitive function in healthy elderly men and patients with type 2 diabetes. *Proceedings of the National Academy of Sciences USA* 101, 6734–6739.
- Seckl, J. R. (2004). 11 β -hydroxysteroid dehydrogenases: changing glucocorticoid action. *Current Opinions in Pharmacology* 4, 597–602.
- Seckl, J. R. (2004). Prenatal glucocorticoids and long-term programming. *European Journal of Endocrinology* 151, U49–U62.
- Seckl, J. R. and Walker, B. R. (2004). 11 β -hydroxysteroid dehydrogenase type 1 as a modulator of glucocorticoid action: from metabolism to memory. *Trends in Endocrinology and Metabolism* 15, 418–424.
- Stewart, P. M. and Krozowski, Z. S. (1999). 11 β -hydroxysteroid dehydrogenase. *Vitamins and Hormones* 57, 249–324.
- Weaver, I., Cervoni, N., Champagne, F., et al. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience* 7, 847–854.
- Welberg, L. A. M., Seckl, J. R. and Holmes, M. C. (2000). Inhibition of 11 β -hydroxysteroid dehydrogenase, the fetoplacental barrier to maternal glucocorticoids, permanently programs amygdala glucocorticoid receptor mRNA expression and anxiety-like behavior in the offspring. *European Journal of Neuroscience* 12, 1047–1054.
- Yau, J. L. W., Noble, J., Kenyon, C. J., et al. (2001). Lack of tissue glucocorticoid reactivation in 11 β -hydroxysteroid dehydrogenase type 1 knockout mice ameliorates age-related learning impairments. *Proceedings of the National Academy of Sciences USA* 98, 4716–4721.

Hyperreactivity (Cardiovascular)

A Georgiades

Duke University Medical Center, Durham, NC, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A Georgiades and M Fredrikson, volume 2, pp 421–424, © 2000, Elsevier Inc.

Cardiovascular Reactivity and Disease
Physiological Parameters
Assessments
Stress Tasks
Stability
Generalizability
Risk Factors and Hyperreactivity
Conclusion

Glossary

- Blood pressure (BP)** The pressure of the circulating blood on the walls of the arteries and veins. BP is usually measured with a sphygmomanometer in millimeters of mercury (mmHg).
- Coronary heart disease (CHD)** A disease associated with insufficient blood flow to the heart muscle, including myocardial infarction and angina pectoris.
- Diastolic blood pressure (DBP)** The lowest BP within a cardiac cycle.
- Hypertension** Elevated levels of arterial BP. According to Joint National Committee criteria of 2003, hypertension is systolic blood pressure levels over 140 and/or DBP levels over 90. Individuals in the prehypertensive range of 130–139 and/or 80–89 are at twice the risk of developing

	hypertension compared to those with lower values.
<i>Mental stress testing</i>	A method used to elicit cardiovascular stress reactivity in a laboratory setting through the administration of engaging, challenging, or aversive stimuli (stressors).
<i>Risk factor</i>	An attribute of individuals that is associated with an increased likelihood for developing a specific disease compared to individuals without the attribute.
<i>Systolic blood pressure (SBP)</i>	The peak blood pressure within a cardiac cycle.

Cardiovascular reactivity is defined as an individual's propensity to display cardiovascular reactions of greater or lesser magnitude compared to a baseline value when encountering stimuli or situations experienced as engaging, challenging, or aversive.

Cardiovascular Reactivity and Disease

It has been suggested that exaggerated cardiovascular reactivity plays an etiological role in cardiovascular disorders such as hypertension and CHD. Folkow suggested that the repeated elicitation of the defense reaction, which is associated with enhanced BP and heart rate (HR), translates into sustained hypertension by causing vascular hypertrophy in genetically susceptible or hyperreactive individuals. Because the normal outlet of the defense reaction (physical work) is inhibited during psychological stress, disruption of bodily homeostasis occurs. Through the repeated surges in arterial pressure and overperfusion of inactive muscles, lasting changes of the peripheral vascular resistance and a resetting of the baroreceptor set point may occur. Stress may further hamper adequate BP regulation by increasing renal vascular resistance, causing antinatriuresis with sodium retention, a decrease in the glomerular filtration rate, and the stimulation of renin release. Thus, it has been proposed that the development of high BP is stress-related.

Cardiovascular hyperreactivity has also been linked to the occurrence of CHD events. It has been hypothesized that exaggerated BP responses can cause endothelial injury, leading to accumulating atheroma, increasing the risk for cardiac arrhythmia. The adverse effects of stress are believed to be mediated by the autonomic nervous system, mainly enhanced sympathetic activity and subsequent physiological responses. For example, high plasma levels of norepinephrine and epinephrine that are associated with hyperreactivity may, in combination with cortisol, be toxic to the coronary vasculature, constituting

an additional mechanism mediating the adverse effects of stress on the cardiovascular system.

A number of cross-sectional studies show that cardiovascular hyperreactivity is more prominent among individuals at an increased risk of developing hypertension and CHD. In addition, mental stress is a potent trigger of myocardial ischemia both in the laboratory and in the ambulatory state among cardiac patients. Also, results from prospective studies show that cardiovascular reactivity may be a risk factor in cardiovascular disease progression. Although the evidence is not conclusive, studies have found that cardiovascular responses to standardized mental stress testing can predict the development of hypertension, left ventricular mass, and atherosclerosis. Excessive BP increases during exercise have also been suggested as a marker of future hypertension and to be associated with an increased risk of cardiovascular mortality.

Because growing evidence suggests that exaggerated cardiovascular responses are a risk factor or marker for the development of cardiovascular disease, research on cardiovascular stress reactivity has focused on characterizing the determinants of hyperreactivity and identifying possible ways of modifying it.

Physiological Parameters

The most commonly measured variables during mental stress reactivity testing are BP and HR. In general, noninvasive BP measurements during laboratory testing require multiple measurements made at frequent intervals (e.g., every 1–2 min). Automatic or semiautomatic devices are preferred, using a variety of available methods, such as Korotkoff sounds, oscillometry, and ultrasound.

HR recordings are usually derived from interbeat interval information available from recording the electrocardiogram (EKG) via electrodes adhered to the chest. Automated BP monitors often include measurements of HR from either EKG signals or arterial pulse detection. Holter monitoring records the EKG waveform over 24 h and involves several algorithms to describe patterns associated with the underperfusion of the heart muscle.

For continuous noninvasive BP monitoring, spectral analysis can be used, allowing the computation of short-term BP variability and baroreflex sensitivity. Newer noninvasive techniques such as impedance cardiography make it possible to document a more complete hemodynamic pattern of the stress response, including measurements of cardiac contractility, cardiac output, and vascular resistance. These measures add important new information to the area of cardiovascular reactivity research.

Assessments

Reactivity testing in the laboratory typically consists of an initial rest (baseline) period followed by a period during which the subject is exposed to a stressor. During baseline, the individual should be stimulated minimally by external events, that is, be in a relaxed state. Reactivity is quantified as a change score, calculated as the difference between resting baseline levels and mean or peak levels during stress. To increase reliability, two or more measurements should be taken and averaged for each condition. Because reactivity scores are sometimes correlated with baseline levels, residualized change scores may be calculated to eliminate baseline effects on the calculated change scores.

Stress Tasks

Stimuli (stressors) used in laboratory stress testing have different characteristics. One distinction is made between psychological stressors and physical stressors. Some widely used and well-characterized psychological stressors with emotional and/or cognitive challenging components are mental arithmetic, mirror tracing, anger interview, public speech, video games, and the Stroop test. Physical stressors commonly used are isometric handgrip, dynamic exercise, and the cold pressor test. Another distinction has been made between passive and active tasks, with the cold pressor being an example of the former and mental arithmetic an example of the latter. In addition, various stressors elicit different hemodynamic patterns, reflecting β - or α -adrenergic receptor activity. For example, cold pressor predominantly activates α receptors, whereas mental arithmetic is considered to be a β -adrenergic task.

Stability

Because there are profound individual differences in both the magnitude and the patterning of cardiovascular responses to stress, it is of importance to establish the stability of the response magnitude and patterning to a given stressor. Numerous studies suggest the BP and HR reactivity show acceptable test-retest reproducibility measured over both the short and long term. In general, the highest stability coefficients have been found for SBP and HR levels and for SBP and HR reactivity scores during standardized mental stress testing, with a somewhat lower stability for DBP levels and reactivity scores. Studies also indicate that the higher test reliability can be achieved by aggregating measurements across multiple stimuli and by aggregating over multiple tasks administered on multiple occasions. Thus, evidence suggests that

BP and HR reactivity in the laboratory constitute a stable individual trait.

Generalizability

The reactivity definition discussed so far refers to individual differences in the response magnitude to a given stimulus on a particular occasion. The underlying assumption is that hyperreactivity to laboratory stressors is a generalized individual propensity and that the individual response pattern elicited by laboratory-induced stress can be generalized to responses in the natural setting during encounters with everyday stress. Noninvasive ambulatory blood pressure (ABP) monitoring provides the opportunity to study the relationship between cardiovascular responses in the laboratory and in natural settings, and several studies have examined the relationship between cardiovascular measurements in the laboratory and during everyday life using automated ABP recordings. Most studies have found a strong positive relationship between ABP levels and laboratory BP levels, whereas the relationship between laboratory-induced BP reactivity and ambulatory levels or variability seems less consistent.

Risk Factors and Hyperreactivity

A number of cross-sectional data show that cardiovascular hyperreactivity is more prominent among individuals at an increased risk of developing hypertension and CHD. Individual differences in cardiovascular reactivity to stress are, in part, mediated genetically. Also, male gender, family history of CHD, borderline hypertension, African American ethnicity, and type A behavior and depression have all been associated with an increased risk for CHD, and these high-risk group have independently shown exaggerated cardiovascular reactivity in several studies. In addition, smoking status, serum cholesterol levels, fasting insulin levels, left ventricular hypertrophy, and carotid artery atherosclerosis are all important independent risk factors for cardiovascular disease. A number of studies have investigated the relationship between these risk factors and cardiovascular hyperreactivity.

1. Gender. Sex differences exist in resting baseline and HR levels, with women having higher HR and men, in general, having higher BP levels (until the age of 60). In general, men are more SBP reactive than women, whereas there are small or no differences in DBP reactivity. Studies on HR reactivity are inconsistent, with men and women having an equal likelihood of displaying augmented HR reactivity.

2. Hypertension. In a meta-analysis of published reactivity studies related to hypertension, it was concluded that the majority of studies have shown that individuals with a positive family history of hypertension and individuals with borderline hypertension display enhanced BP reactivity compared to individuals without a family history of hypertension and to normotensive controls.

3. Ethnicity. African Americans have enhanced cardiovascular hyperreactivity compared to Caucasians. Studies also show that African Americans display increased total peripheral resistance during mental stress testing in the laboratory compared to Caucasians.

4. Type A behavior. Individuals with type A behavior (i.e., showing enhanced hostility, time urge, and competitiveness) have increased sympathetically mediated reactivity compared to type B individuals. It has been suggested that it is the hostility component in the type A behavior pattern that is most closely related to enhanced cardiovascular reactivity.

5. Depression. Accumulating evidence show that depressive symptoms are related to cardiovascular morbidity and mortality. Studies evaluating the association between depression and hyperreactivity have so far only shown moderate support for a link between depressive symptoms and enhanced cardiovascular reactivity.

6. Smoking. Evidence is accumulating that smoking, in combination with psychological stress, may be particularly harmful, both because smoking increases in stressful situations and because the cardiovascular effects of smoking combined with those of psychological stress may be synergistic. Nicotine itself has acute effects on the activity of the cardiovascular system, leading to an increase of both BP and HR, but smoking status per se (being a smoker or a non-smoker) does not seem to affect the magnitude of the stress response. Nicotine administration has been found to enhance HR reactivity, but there is no conclusive evidence that nicotine enhances cardiovascular reactivity in general.

7. Lipids. Several studies have reported an association between serum lipid levels and cardiovascular reactivity to laboratory stressors, but findings are equivocal. Associations have been found to be lower or completely absent in youngsters but evident among older adults, which suggests an age-dependent association. Some evidence suggests that an enhanced cardiac responsiveness is associated positively with low-density lipoprotein levels but is related inversely with high-density lipoproteins.

8. Insulin resistance. Insulin resistance is part of the metabolic syndrome associated with hypertension and cardiovascular alternations. One mechanism

through which insulin may be linked to hypertension is through the stimulation of sympathetic nervous activity. Thus, insulin concentrations may be correlated with indices of sympathetic activity such as hyperreactivity even before it affects resting BP. Fasting insulin concentrations have been associated with cardiovascular reactivity to exercise. There are also studies indicating that insulin resistance may be more closely related to BP levels during mental stress testing than during rest.

9. Left ventricular mass. Left ventricular hypertrophy is an important target-organ consequence of hypertension and an independent risk factor for cardiovascular morbidity and mortality. Several studies have shown that left ventricular mass is poorly related to casual clinic BP levels, whereas some studies have shown stronger associations to daytime ambulatory BP levels and variability and to BP levels during mental stress testing. Left ventricular mass (index) has also been found to be associated with a vasoconstrictive response pattern among children and adolescents. A few studies have found cardiovascular reactivity to predict the development of left ventricular mass over time. However, further evidence is needed to establish the association between stress-induced cardiovascular reactivity and left ventricular mass.

10. Atherosclerosis. Carotic intima-media thickness is an important preclinical measure of atherosclerotic vascular disease and has been strongly associated with the development of CHD. BP responses to mental stress have been found to be associated to intima-media thickness as well as plak height. Although studies have also shown reactivity to mental stress to be prospectively related to intima-media thickness, further evidence is required to establish the relationship between cardiovascular hyperreactivity and atherosclerosis.

11. Inflammation. Atherosclerosis is considered an inflammatory disease. In the atherosclerotic process, proinflammatory cytokines, derived from excess adipose tissue, increase the expression of adhesion molecules that promote the sticking of blood leukocytes to the inner surface of the arterial wall. Recent studies report that acute psychological stress in humans significantly increases the serum concentrations of proinflammatory cytokines such as interleukin-6 (IL-6) and the IL-1 receptor antagonist (IL-1Ra), suggesting that stress reactivity may affect the atherosclerotic process by eliciting an activation of the inflammatory response system.

Conclusion

Cardiovascular hyperreactivity is associated with several risk factors for cardiovascular morbidity and

mortality, and, although the mechanisms are not fully established, there is evidence suggesting that heightened cardiovascular reactivity may independently contribute to the development of CHD and hypertension.

See Also the Following Articles

Ambulatory Blood Pressure; Beta-Adrenergic Blockers; Cardiovascular System and Stress; Heart Disease/Attack; Heart Rate; Hypertension; Type A Personality, Type B Personality.

Further Reading

- Folkow, B. S. (1982). Physiological aspects of primary hypertension. *Physiological Reviews* **62**, 374–504.
- Fredrikson, M. and Matthews, K. A. (1990). Cardiovascular responses to behavioral stress and hypertension: a meta analytic review. *Annals of Behavioral Medicine* **12**, 17–29.
- Kamarck, T. W. and Lovallo, W. R. (2003). Cardiovascular reactivity to psychological challenge: conceptual and

- measurement considerations. *Psychosomatic Medicine* **65**, 9–21.
- Linden, W., Gerin, W. and Davidson, K. (2003). Cardiovascular reactivity: status quo and a research agenda for the new millennium. *Psychosomatic Medicine* **65**, 5–8.
- Lovallo, W. R. and Gerin, W. (2003). Psychophysiological reactivity: mechanisms and pathways to cardiovascular disease. *Psychosomatic Medicine* **65**, 36–45.
- Schneiderman, N., Weiss, S. M. and Kaufmann, P. G. (1989). *Handbook of research methods in cardiovascular behavioral medicine*. New York: Plenum Press.
- Schwartz, A. R., Gerin, W., Davidson, K. W., et al. (2003). Toward a causal model of cardiovascular responses to stress and the development of cardiovascular disease. *Psychosomatic Medicine* **65**, 22–35.
- Steptoe, A., Ruddle, H. and Neus, H. (1985). *Clinical and methodological issues in cardiovascular psychophysiology*. New York: Springer-Verlags.
- Treiber, F. A., Kamarck, T., Schneiderman, N., et al. (2003). Cardiovascular reactivity and development of preclinical and clinical disease states. *Psychosomatic Medicine* **65**, 46–62.
- Turner, J. R. (1994). *Cardiovascular reactivity and stress: patterns of physiological response*. New York: Plenum Press.

Hypertension

A Steptoe

University College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A Steptoe, volume 2, pp 425–430, © 2000, Elsevier Inc.

Left ventricular hypertrophy

The enlargement of the left ventricle of the heart, one of the adverse effects of hypertension.

Normotension

Blood pressure below the criteria for hypertension.

Prevalence

Etiological Factors

Experimental Evidence

Psychological Characteristics

Lifestyle and Adverse Life Experience

Stress Management and Quality of Life

Glossary

White-coat hypertension

Blood pressure that is elevated when measured in the clinic by a physician or nurse but normal when measured in other situations.

Hypertension

Sustained elevated blood pressure.

Prevalence

Hypertension is the condition of sustained elevated arterial blood pressure. The criteria for defining hypertension are somewhat arbitrary because the distribution of blood pressure level in the population is continuous. The current definition of hypertension is a systolic blood pressure ≥ 140 mmHg and diastolic pressure ≥ 90 mmHg. Most authorities regard a blood pressure in the range 130–139/85–89 mmHg as high-normal, although the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure has defined levels of 120–139/80–89 mmHg as prehypertensive. Risk of coronary heart disease and stroke is directly related to blood

pressure level, so hypertension defines those individuals who are at an especially elevated risk.

The prevalence of hypertension in the United States is approximately 29% in men and 30.5% in women. Recent figures from the United Kingdom indicate that approximately 37% of men and 34% of women are hypertensive. However, blood pressure tends to rise with age, so these overall figures disguise the substantial increase in hypertension in middle age. Thus, in the United States, the prevalence rises from 8.1% in men and 2.7% in women ages 20–34 years to 44.9% and 53.9% in those ages 55–64 years. There are also striking ethnic variations, with high rates in African Americans and in African Caribbean people in the United Kingdom.

Between one-third and one-half of hypertensives are not diagnosed and are unaware of their condition. Because hypertension is not consistently associated with symptoms, it is frequently identified only during routine screening, so it may continue for many years before detection. The number of diagnosed hypertensives who are not treated is substantial, whereas as many as 50% of patients who are prescribed antihypertensive medication do not have their blood pressure controlled adequately. The problem of hypertension is therefore considerable.

Etiological Factors

The etiology of hypertension is complicated. In a minority of cases, high blood pressure is associated with a specific pathological abnormality such as pheochromocytoma or a renovascular disorder. However, in most cases there is no single identifiable cause, and the condition is regarded as multifactorial. This condition was once known as essential hypertension because an elevated pressure was considered necessary to drive blood through partially blocked arteries. Now, cases without a single cause are sometimes called primary hypertensives to distinguish them from individuals whose hypertension is secondary to another disease. Hypertension typically develops over many years, with blood pressure gradually increasing until it reaches threatening levels. The most consistent predictors of future hypertension are a positive family history, a blood pressure in childhood or early adult life that is above average, and elevated body weight. High salt and alcohol intake also contribute.

The primary reason for considering stress to be an etiologic factor relates to the influence of the autonomic nervous system and neuroendocrine factors in hypertension. There is clear evidence of increased sympathetic nervous system activity in hypertension, as documented through raised levels of plasma norepinephrine. Microneurographic studies have shown

elevated sympathetic nervous traffic to muscle. Hypothalamic activation may be responsible for these sympathetic responses, which in turn lead to inhibition of baroreceptor reflex sensitivity, peripheral vasoconstriction, and alterations of renal blood flow. Reduced parasympathetic activity may also be associated with the etiology of hypertension, together with the impairment of opioid-mediated counter-regulatory processes. Abnormalities of adrenocortical function have been described in some cases of hypertension, and pro-inflammatory processes may contribute as well.

The hypothesis underlying much research on this topic is that the stress-induced activation of sympathetic and other pathways occurs repeatedly over months and years in susceptible individuals, leading to hemodynamic adjustments that increase hypertension risk. Evidence relevant to this hypothesis derives from several sources, including animal and human experimental studies and investigations of the impact of adverse life experiences, chronic stressors, and psychological characteristics.

Experimental Evidence

Animal Studies

Evidence that permanently elevated blood pressure can be produced with behavioral stress in animals is controversial. Although it is clear that acute hypertensive episodes can be induced, the development of sustained hypertension is more elusive. Aversive conditions such as restraint stress, unpredictable shock, and approach-avoidance conflict produce chronic hypertension in some studies but not others. Effects vary across species and are more prominent in the presence of risk factors such as a genetic predisposition and heightened salt sensitivity. Some of the most compelling evidence comes from studies carried out by Henry and co-workers on the effects of social conflict in CBA mice. Male mice raised in isolation were placed in complex colonies that encouraged conflict. There were immediate blood pressure increases that reverted to control levels after their removal from the colony. However, irreversible hypertension developed with prolonged exposure to conflict, accompanied by alterations in catecholamine synthesis and myocardial fibrosis and degeneration. Species differences are apparent in primates as well, although hypertension can be induced by prolonged shock avoidance in rhesus monkeys.

Cardiovascular Reactivity to Acute Behavioral Stress in Humans

Blood pressure increases acutely in response to a variety of behavioral stressors in humans, ranging

from problem-solving tasks and emotionally laden interviews to painful stimuli such as the cold pressor test. The increase in blood pressure is typically accompanied by tachycardia, inotropic stimulation of the heart, vasodilatation in skeletal muscle and adipose tissue, and reduced blood flow to the kidneys, skin, and viscera.

The issue of whether these responses are involved in hypertension has been the subject of much research, stimulated in the 1950s by Jan Brod. Compared with normotensives, hypertensives typically show enhanced blood pressure responses to acute behavioral stress, accompanied by greater renal vasoconstriction and increased norepinephrine turnover. However, it is possible that this exaggerated cardiovascular responsivity is the result of the complex physiological adjustments that occur in hypertension and is not primary. Attention has therefore shifted to people with normal blood pressure who are at increased risk, for example, because of genetic factors. Results have again been variable, but a meta-analysis of this literature suggests that a proportion of people with a positive family history of hypertension are more reactive to tasks involving active behavioral responses (e.g., difficult problem solving) than are those without family histories. Young people with positive family histories may also have elevated blood pressure during their everyday lives, as assessed using automated ambulatory monitoring techniques.

Longitudinal studies have also been reported, testing the hypothesis that high reactivity to behavioral stress predicts future rises in blood pressure over 5–10 years. Most trials have found evidence in favor of this hypothesis. Impaired poststress recovery in blood pressure has also been shown to predict future increases. But such effects are not always observed, probably for two reasons. First, the investigation of acute responses to behavioral stress is complex and requires the careful standardization of procedures. Different types of stimuli elicit different patterns of physiological response. For instance, some tasks such as mental arithmetic elicit blood pressure responses that are primarily sustained through increased myocardial contractility and cardiac output, whereas others predominantly elicit increases in peripheral vascular resistance. These response profiles may not all be equally relevant to hypertension. Second, the nature of the hypothesized relationship between stress responsivity and hypertension should be considered. [Figure 1](#) outlines the presumed sequence of events. Heightened cardiovascular responses to acute stress *per se* will not increase the risk of hypertension. Such responses are relevant only if they are representative of the individual's experience during daily life. Stress-related factors are most likely to contribute to hypertension if reactive individuals are exposed over months or years to environments that elicit damaging patterns of intense response. This interaction has only

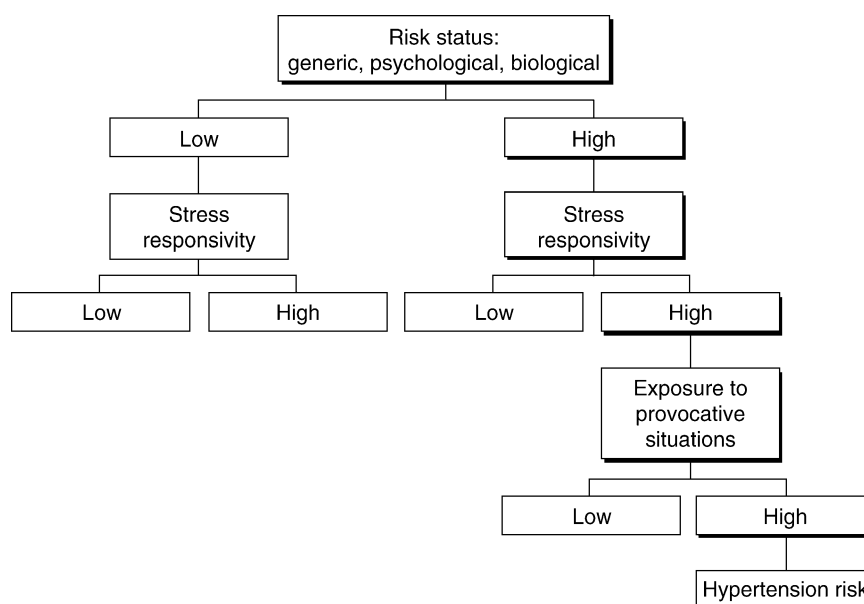


Figure 1 Schematic outline of the way in which acute cardiovascular and neuroendocrine stress responsivity might contribute to hypertension risk. Highlighted boxes indicate the highest risk group. The combination of high risk status, high stress responsiveness, and prolonged exposure to situations that provoke appropriate physiological responses is necessary.

been evaluated in a few studies to date. It should also be kept in mind that stress is only one of many factors contributing to hypertension and does not affect all those at risk in the same fashion.

Psychological Characteristics

The many comparisons that have been made between hypertensive patients and controls in terms of personality typically suffer from two limitations. First, the diagnostic label of hypertension may cause psychological distress and lead to changes on psychological tests. Second, because much hypertension is not diagnosed, any differences observed may be restricted to those who seek medical attention. Population-based studies involving measurements of blood pressure and psychological factors are therefore required.

Some population studies have shown prospective associations between tension or anxiety and future hypertension, but effects are small and inconsistent. A more persuasive body of evidence has linked hypertension with the ways that people cope with angry and hostile feelings. This notion was stimulated in the mid-1970s by research on hypertension and urban stress in Detroit. It was found that hypertension was more prevalent and that blood pressure was generally higher among black men living in high-stress areas (defined by population density, crime rate, and family instability) who also inhibited anger expression. Subsequent research has supported the link between anger inhibition and high blood pressure in a number of populations. However, it has also been found that the expression of anger is related to blood pressure, so perhaps the appropriateness of anger display and hostility management is more critical.

Hypertension is associated with mild deficits of cognitive function, as assessed by measures of memory, attention, and abstract reasoning. These effects vary with age, education, and whether blood pressure is being treated. The extent to which these impairments affect the performance of daily activities is not known. They may be the result of central nervous system neurophysiological sequelae of high blood pressure.

Lifestyle and Adverse Life Experience

What types of real-life stressful experience are likely to be associated with hypertension? If the inferences from laboratory studies are correct, they are experiences that are repeated or sustained over years and elicit active efforts to cope and maintain control. Acute traumatic events would not be expected to stimulate appropriate long-term cardiovascular and neuroendocrine responses, and indeed there is little

evidence that acute life events are associated with the development of hypertension. In contrast, some forms of chronic stress are associated with sustained increases in blood pressure. Experiences such as caring for demented relatives, marital conflict, and the experience of racism in ethnic minorities have been related to increased blood pressure. Interesting results have emerged from studies of migrants from traditional, isolated, rural cultures to modernized, developing and developed societies. For example, increases in blood pressure have been recorded among young adults from the Luo tribe in western Kenya who migrated to Nairobi. The blood pressure rise was associated with a higher salt intake and greater body weight, but stress-related increases in the sympathetic nervous stimulation may also have been involved. Similarly, studies of migrants from the isolated Pacific Tokelau islands to New Zealand have shown that blood pressure increases are greater among those who establish extensive contact with non-Tokelauans compared with individuals who maintain traditional ethnic affiliations.

Blood pressure has been shown in several studies to be associated inversely with the size of social networks and the extent of social support. A striking illustration of the protective effects of a tranquil lifestyle and supportive environment is illustrated in [Figure 2](#). This shows results from a long-term study of Italian nuns in a secluded order compared with nonsmoking, age-matched women living in local communities. It can be seen that the typical age-related increases in blood pressure were attenuated among the nuns, which was coupled with a significantly lower incidence of fatal and nonfatal cardiac events.

Work and Hypertension

Some types of work might be expected to have the characteristics that would lead to acute stress-induced increases in blood pressure and sympathetic activation, thereby heightening the risk of hypertension. An important early study by Cobb and Rose of air traffic controllers demonstrated an increased incidence and earlier onset of hypertension among men working in this highly demanding occupation compared with controls. Subsequently, an automated ambulatory recording apparatus has been used to measure blood pressure repeatedly as people go about their everyday lives. Blood pressure is typically higher at work than at home, and it is elevated in people who work in high-demand, low-control jobs. For instance, in the Work Site Blood Pressure Study in New York City, job strain (high demand, low control) was an independent predictor of hypertension after controlling for age, body mass index, type A behavior, 24-h sodium

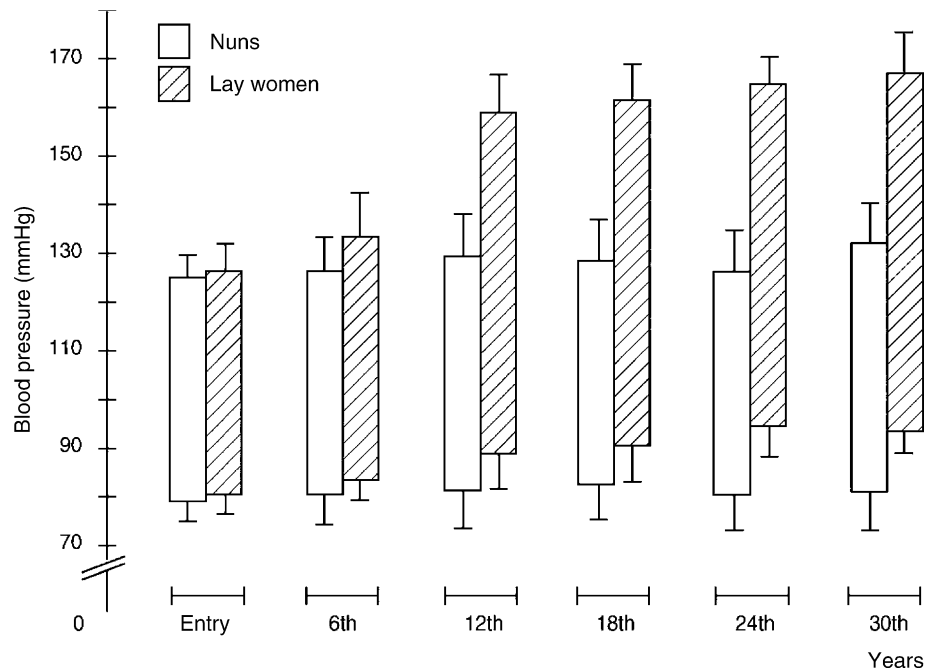


Figure 2 Results from a 30-year longitudinal study of 144 nuns from a secluded Catholic order and 138 age-matched nonsmoking women from the local community. The absence of age-related increases in blood pressure in the nuns is clear. Reproduced with permission from Timio et al. (1997). Blood pressure trend and cardiovascular events in nuns in a secluded order: a 30-year follow-up study. *Blood Pressure* 6, 81–87. © Scandinavian University Press.

excretion, physical activity at work, education, smoking, and alcohol consumption. Longitudinally, individuals exposed to more years of high job strain had higher blood pressure. This is consistent with the evidence that lack of control over challenging situations is associated with heightened physiological activation.

White-Coat Hypertension

White-coat hypertension is the phenomenon of transiently raised blood pressure during measurement in the clinic by a physician or nurse compared with blood pressure as recorded at home or using an ambulatory recording apparatus. The definition of white-coat hypertension is somewhat arbitrary, but one study found that 20% of patients referred with a diagnosis of hypertension, whose diastolic pressure remained in the range of 90–104 mmHg on several occasions in the clinic, were normotensive during ambulatory monitoring. The phenomenon is clearly stress-related, in that it is dependent on physiological responses to the threatening clinical situation. People who exhibit white-coat effects do not have unusual psychological profiles, but do show heightened blood pressure responses to stressful behavioral tasks. There is also limited evidence that white-coat hypertension is associated with the secondary effects of hypertension such as left ventricular hypertrophy. However,

the boundaries and the significance of the problem remain poorly understood.

Stress Management and Quality of Life

Despite research over the past quarter of a century, the use of stress management in hypertension remains controversial. A number of well-controlled trials have shown that stress management involving relaxation training (sometimes accompanied by biofeedback) can reduce the blood pressure and medication needs of hypertensives. Problems have arisen, however, in establishing genuine pretreatment levels of blood pressure against which to measure the effects of stress management. Blood pressure may fall progressively with repeated measurement due to adaptation and the habituation of white-coat effects, and these phenomena may be confused with the effects of stress management. The evidence for specific techniques such as transcendental meditation reducing the blood pressure of hypertensives is inconclusive.

The pharmacological treatment of hypertension is a complex and continuously developing field. Older treatments, such as β -adrenergic blockade and diuretics, have been joined by calcium antagonists and angiotensin-converting enzyme inhibitors. Some hypertensive medications have side effects that lead to impairments in the health-related quality of life. Stress

management may play a role in particular types of cases. One is the individual whose blood pressure is only mildly elevated (e.g., 170/100 mmHg), who might be able to reduce blood pressure without recourse to lifetime medication. The second is the established hypertensive who is already medicated but who wishes to try alternative methods to reduce drug intake or because their blood pressure is poorly controlled. Other behavior changes also benefit hypertensives, including weight loss, increased physical exercise, and restriction of salt and alcohol intake. Stress management should therefore be applied in the context of these other aspects of behavior modification.

See Also the Following Articles

Ambulatory Blood Pressure; Control and Stress; Demand-Control Model; Hypotension, Hypovolemia, and Septic Shock; Metabolic Syndrome and Stress; Stress Management and Cardiovascular Disease; War Stress in the Former Yugoslavia.

Further Reading

Chobanian, A. V., Bakris, G. L., Black, H. R., et al. (2003). The seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *Journal of the American Medical Association* 289, 2560–2572.

- Eisenberg, D. M., Delbanco, T. L., Berkey, C. S., et al. (1993). Cognitive behavioral techniques for hypertension: are they effective? *Annals of Internal Medicine* 118, 964–972.
- Henry, J. P. and Stephens, P. M. (1977). *Stress, health, and the social environment*. New York: Springer-Verlag.
- Muldoon, M. F., Terrell, D. F., Bunker, C. H., et al. (1993). Family history studies in hypertension research: review of the literature. *American Journal of Hypertension* 6, 76–88.
- Pickering, T. G., James, G. D., Boddie, C., et al. (1988). How common is white coat hypertension? *Journal of the American Medical Association* 259, 225–228.
- Rutledge, T. and Hogan, B. E. (2002). A quantitative review of prospective evidence linking psychological factors with hypertension development. *Psychosomatic Medicine* 64, 758–766.
- Steptoe, A. (1997). Behavior and blood pressure: implications for hypertension. In: Zanchetti, A. & Mancia, G. (eds.) *Handbook of hypertension – pathophysiology of hypertension*, pp. 674–708. Amsterdam: Elsevier Science.
- Timio, M., Lippi, G., Venanzi, S., et al. (1997). Blood pressure trend and cardiovascular events in nuns in a secluded order: a 30-year follow-up study. *Blood Pressure* 6, 81–87.
- Waldstein, S. R. and Elias, M. F. (2001). *Neuropsychology of cardiovascular disease*. Mahwah, NJ: Lawrence Erlbaum.
- Williams, D. R., Neighbors, H. W. and Jackson, J. S. (2003). Racial/ethnic discrimination and health: findings from community studies. *American Journal of Public Health* 93, 200–208.

Hyperthermia

J Roth

University of Giessen, Giessen, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J Roth, volume 2, pp 431–438, © 2000, Elsevier Inc.

Thermal Balance and Body Temperature
Physiological Responses to Heat Stress
Brain Function in Hyperthermia
Adaptation to Heat
Malignant Hyperthermia

Glossary

Adaptation A change that reduces the physiological strain produced by stressful components

Conductive heat transfer

Convective heat transfer

Evaporative heat transfer

Metabolic heat production

of the total environment. This change may occur within the lifetime of an organism (phenotypic) or be the result of genetic selection in a species (genotypic). The net rate of heat transfer by conduction down a thermal gradient, within an organism or between an organism and its external environment.

The net rate of heat transfer by convection between an organism and its external environment or the transfer of heat with the bloodstream from the thermal core through the thermal shell to the heat-dissipating body surface.

The loss of heat energy from the body to the ambience by evaporation of water from the skin and the surface of the respiratory tract.

The rate of transformation of chemical energy into heat in an organism.

<i>Radiant heat exchange</i>	The net rate of heat exchange by radiation between an organism and its environment.
<i>Temperature regulation</i>	Maintenance of the temperature of a body within a restricted range under conditions involving variable internal or external heat loads.
<i>Thermal stress</i>	Any change in the thermal relation between a temperature regulator and its environment that, if uncompensated by temperature regulation, would result in hyper- or hypothermia.
<i>Thermo-effector</i>	An organ and its function that affect heat balance in a controlled manner as part of the process of temperature regulation.

Thermal Balance and Body Temperature

If body temperature is to be maintained at a constant level, there has to be a thermal equilibrium. This means that heat production or heat gain and heat loss must be equal. Under resting conditions, basal metabolic heat production, which occurs due to muscle contractions, to biosynthesis of molecules within the body, and to the activity of ion pumps, is predominantly counteracted by so-called dry heat loss. Dry heat loss consists of radiation, convection, and conduction of heat from the body to the surroundings and is dependent on and proportional to the temperature difference between body and surroundings. This means that a dry heat loss would be equal to zero at an ambient temperature of 37°C for a human subject; at lower temperatures it would increase and at higher temperatures the body would even gain dry heat from the surroundings. However, dry heat loss from the body also depends on the conduction and convection of heat within the body and thus on peripheral blood flow. Vasoconstriction of skin vessels prevents heat flow from the body core to the surface, and the activation and recruitment of additional (to the basal) heat production can be delayed to a lower ambient temperature. As a consequence, there is a so-called thermoneutral zone, a range of ambient temperatures in which heat loss can be matched to resting heat production by vasomotor adjustments. For humans, the range of the thermoneutral zone is between ambient temperatures of 28 and 32°C in nonmoving air. Beyond the lower limit of the thermoneutral zone, body temperature can be kept constant only if regulatory mechanisms increase thermogenesis in proportion to heat loss. In humans, heat production, predominantly due to the activation of shivering, can be increased as much as three to five times the basal metabolic rate. At an ambient temperature, which requires more than the maximum of

available heat production to keep the body core temperature constant, the lower limit of the thermoregulatory range is reached. When this limit is passed, hypothermia develops, which results in cold death. Beyond the upper limit of the thermoneutral zone, thermal balance is achieved by an additional heat loss mechanism, the evaporation of sweat (wet heat loss in contrast to the dry heat loss mentioned earlier). To a small degree of about 20% of the total heat loss, evaporative heat transfer also contributes to the heat loss under thermoneutral conditions. This is achieved by evaporation of water that has diffused either to the skin or from membranes of the respiratory tract; this is called insensible perspiration. Evaporation of water from the skin depends on the difference of vapor pressure between skin and surrounding air. This means that evaporative heat loss is difficult under conditions of high relative humidity. At an ambient temperature higher than 37°C and saturated vapor pressure of the surrounding air (e.g., the situation after a tropical rain on a very hot day), evaporative heat loss can become impossible.

At the upper limit of the thermoneutral zone, additional (to the insensitive) evaporative heat loss is achieved by activation of glandular water loss via sweat glands. When the ambient temperature exceeds that of the body, evaporation of sweat remains the only mechanism of the body to keep the temperature constant. The high effectiveness of thermoregulatory sweat secretion is based on the high amount of heat that is lost by evaporated water (2400 kJ/l). In humans, sweating rates of more than 2 l of water per hour can be observed under thermal stress. This means that an amount of heat can be dissipated by the evaporation of sweat that corresponds to seven times the basic metabolic heat production. At an ambient temperature that requires more than the maximum of available heat dissipation to keep the body temperature constant, the upper limit of the thermoregulatory range is reached. When this limit is passed, hyperthermia develops, which can result in heat death. A schematic diagram illustrating thermal balance within and beyond the zone of normothermy is shown in [Figure 1](#).

Hyperthermia develops at ambient temperatures higher than the upper limit of the thermoregulatory range (see [Figure 1](#)). Also, during exercise, the rate of heat production increases above its level at rest mainly due to skeletal muscle contractions. The exercise-induced increase in heat production occurs immediately at the onset of exercise so that the rate of heat production exceeds the rate of heat dissipation. Therefore, body temperature is already rising during the early stage of exercise. When core temperature rises, heat-dissipating mechanisms (predominantly evaporation of

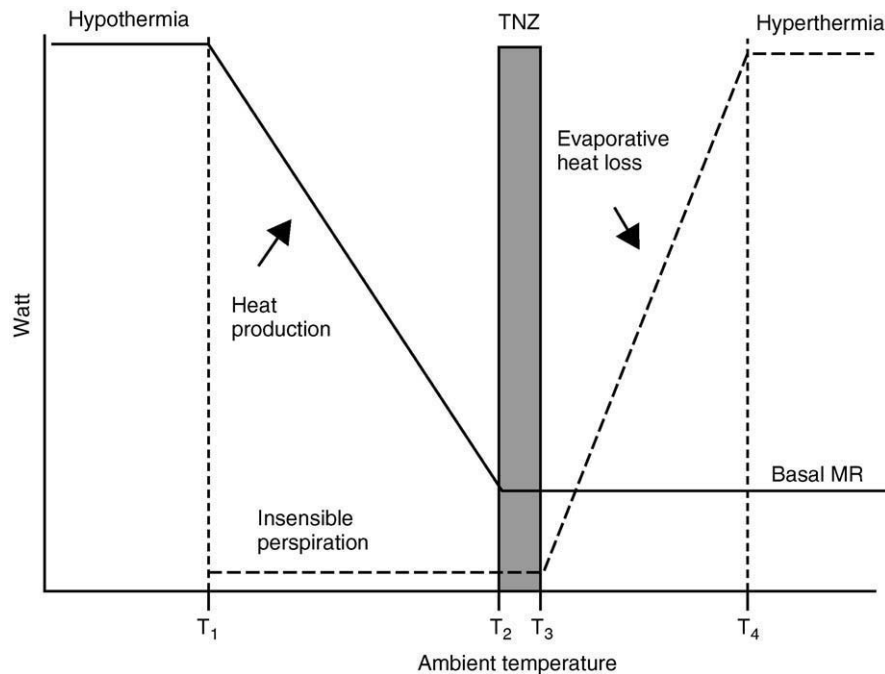


Figure 1 Schematic diagram illustrating thermal balance. Within the range of ambient temperatures between T_1 and T_4 there is an equilibrium between heat production and heat loss. Below T_1 , more heat is lost than can be produced (hypothermia develops). Above T_4 , production and influx of heat exceed the capacity for evaporative heat loss (hyperthermia develops). Between T_2 and T_3 (thermoneutral zone), heat loss can be matched to resting heat production by vasomotor adjustments. At T_2 there is a maximum of peripheral vasoconstriction, whereas at T_3 there is a maximum of peripheral vasodilation. MR, metabolic rate; TNZ, thermoneutral zone. Modified from Brück (1983), with permission from Springer-Verlag.

sweat) are activated, and the core temperature rises further with a smaller slope. Especially under conditions of high ambient temperature and/or high relative humidity, exercise can cause severe hyperthermia. During prolonged, exhausting exercise such as marathon running, rectal temperatures of higher than 41°C have been measured.

Physiological Responses to Heat Stress

During exposure to external heat or during production of additional internal heat, a number of vital homeostatic functions are influenced or even impaired. Important changes in cardiovascular responses, body fluid balance, endocrine responses, and responses of the immune system are described. In part, these adjustments serve to prevent or limit a developing hyperthermia.

Cardiovascular Responses

When the body has to counteract an increased heat load, it must direct enhanced blood flow to the skin, first to transport heat from the body core to the surface for dry heat loss (radiation, convection, conduction) and, second, to supply the water and electrolytes that are necessary for sweat formation.

Cardiovascular adjustments to whole body heating in humans can be summarized as follows. During an aggressive heating of the whole body surface with water-perfused suits, the skin temperature rises from 35 to 40.5°C and the blood temperature increases from 36.7 to 39.1°C . These body temperature changes are accompanied by a continuous rise of cardiac output from 6.4 to 13 l/min. The doubling of cardiac output is, to a large degree, driven by an increased heart rate (rise of about 30 beats per minute per degree increase of core temperature), whereas stroke volume changes little. Increased cardiac output during heat stress is accompanied by a regional redistribution of blood. Despite the doubled cardiac output, splanchnic, renal, and muscle blood flows are reduced by 1.2 l/min during the heating period. Thus, skin blood flow can approach about 8 l/min during external heat stress (6.6 l/min additional cardiac output plus 1.2 l/min redistributed blood). The dramatic increase in skin blood flow during heat stress occurs over the entire body surface and is distributed between two general vessel types: capillaries and arteriovenous anastomoses (AVAs). AVAs are shunt vessels between arterioles and venules. Opening AVAs accelerates blood flow and thereby heat flux substantially. AVAs are concentrated on the so-called acral sites of the skin (tips of extremities, nose, ears, tongue, and lips); there

is a lack of AVAs in nonacral sites (torso, forearms, upper arms, and legs). The major mechanisms by which skin blood flow is increased during heat stress are (1) a reduced activity in sympathetic vasoconstrictor nerves, (2) an increased activity of a sympathetic active vasodilator system, and (3) direct local vasodilative effects of elevated skin temperature. Cardiac output and distribution/redistribution of blood under resting conditions and during heat stress are summarized in [Figure 2](#).

It has to be noted that the increase of thermoregulatory skin blood flow is limited by the demands of the working muscles during hyperthermia induced by physical exercise.

Effects on Body Fluid Balance

Maintenance of body fluid volume and composition is strongly impaired by heat stress. Exposure to heat species specifically results in sweating, panting, or saliva spreading to increase evaporative heat transfer to the environment. As a consequence, the constancy of total body water is disturbed by its considerable decrease. In humans sweating heavily, water loss from the body can exceed 2 l/h. All body fluid compartments and intracellular fluid, as well as extravascular and intravascular extracellular fluid, contribute



Figure 2 Distribution of human cardiac output between the skin and other major organs. The left column illustrates the conditions at rest in normothermia and the right column shows the conditions during severe hyperthermia when core temperature is higher than 39°C. From Rowell (1983), with permission from the American Physiological Society.

to the water loss in heat stress. Without fluid replacement, a thermal dehydration results in hypovolemia (reduced intravascular compartment size) and hyperosmolality (increased plasma osmolality). Hypovolemia and hyperosmolality have adverse effects on thermoregulatory responses to heat stress. The rate of evaporative water loss is reduced in dehydrated humans and animals. As a consequence, heat stress results in faster onset and a stronger increase of body temperature in dehydrated subjects. Under conditions of restricted fluid supply, conservation of water is essential for survival. Therefore, the dehydrated organism tolerates an elevation of body temperature under conditions of heat stress to avoid a further loss of water by evaporation. This phenomenon is demonstrated impressively in the camel living at an ambient temperature of 34°C in the morning and 41°C in the afternoon. In the fully hydrated camel, which is watered every day, body temperature varies by about 2°C, between 36 and 38°C. When the camel is deprived of drinking water, however, daily body temperature fluctuations may increase to as much as 7°C, between 34 and 41°C. For a dehydrated camel, the amount of heat that is stored by this mechanism during the day equals a saving of 5 l of water. These interactions between water balance and body temperature control under the conditions of high external heat load are shown in [Figure 3](#).

Corresponding observations are made under conditions of high internal heat load (exercise). At ambient temperatures higher than 30°C, an exercising man who is dehydrated responds with a strong increase of core temperature, while water replacement allows

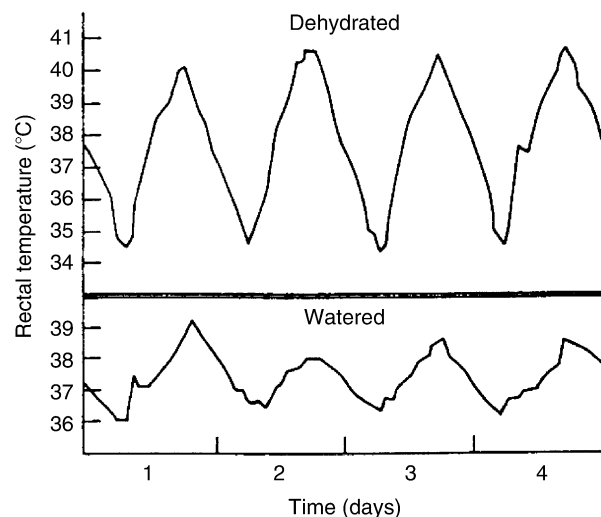


Figure 3 Daily body temperature fluctuations in a well-watered camel (about 2°C, bottom) and in a dehydrated camel deprived of drinking water (about 7°C, top). Modified from Schmidt-Nielsen et al. (1957), with permission from The American Physiological Society.

sustained sweating under the same external conditions so that the core temperature is maintained at a constant level. Sweating is thus inhibited by dehydration. In other words, the development of hyperthermia due to external or internal heat stress is accelerated if there is a strong demand for water conservation to avoid a pronounced deviation in the body fluid balance. Hypovolemia and hyperosmolality inhibit evaporative water loss in many species. It thus seems to be an evolutionary strategy to try to maintain body fluid balance during thermal dehydration with the consequence of a temporally developing hyperthermia.

Endocrinological Responses

Circulating levels of numerous hormones or amounts of hormones excreted with the urine change significantly under conditions of heat stress. As pointed out earlier, there are close interactions between heat defense strategies and the conservation of body fluid in hyperthermic conditions. Even a moderate level of dehydration by heat exposure or exercise affects the circulating levels of hormones that are involved in fluid and electrolyte homeostasis. Increasing dehydration is correlated with increments in arginine vasopressin, plasma renin activity, and aldosterone, hormones involved in the retention of water and electrolytes. In contrast, the secretion of atrial natriuretic peptide, which stimulates the excretion of sodium, is decreased by dehydration. Circulating levels of prolactin and adrenocorticotrophic hormone, which both belong to the so-called stress hormones, are elevated during heat stress. Apparently, whenever thermal stress (heat or cold) induces marked discomfort or physiological strain, the hypothalamic-pituitary-adrenal axis is activated significantly.

Acute heat stress reduces serum thyroid-stimulating hormone. Further, decreased thyroid activity during chronic exposure to heat is indicated by attenuated circulating levels of thyroxine (T_4) and triiodothyronine (T_3). Finally, activity of the sympathoadrenomedullary system is also altered under hyperthermic conditions. The effects of acute and chronic heat exposure have to be differentiated. In most reports dealing with this question, an acute exposure to heat results in a rise of circulating epinephrine (E), whereas increases, no changes, or even decreases of norepinephrine (NE) can be observed depending on the experimental conditions, especially the intensity of the imposed heat stress. Long-term acclimation to heat is accompanied more constantly by a reduced excretion of NE, as long as the acclimation temperature is not excessively high. In summary, if the physiological strain evoked by acute heat stress is sufficient, an elevation of circulating or excreted A and NE is observed. Long-term exposure to heat may be accompanied by an attenuation of sympathetic activity. The most important endocrinological responses to heat stress are summarized in [Table 1](#).

Immunological Responses

The thermal effects of fever, an evolutionarily conserved component of the acute phase response, has been associated with better survival and a shorter duration of disease in cases of infection. A possible explanation for this classical observation derived from a number of studies in which a whole body hyperthermia was induced in humans or experimental animals and various components of the innate and the specific immune defense mechanisms were analyzed under hyperthermic conditions. The benefits of heat for the immune system include enhanced neutrophil

Table 1 Endocrinological responses to heat stress^a

<i>Hormone</i>	<i>Stimulus</i>	<i>Direction of change</i>	<i>Species</i>
AVP ACTH PRA ANG II ALD	Acute heat exposure or severe heat stress in a sauna bath or exercise	Increase in plasma	Human
AVP	Acute heat exposure	Increase in plasma	Rat
TSH	Acute heat exposure	Decrease in serum	Rat
T_3 T_4	Long-term heat acclimation	Decrease in plasma	Rat, chicken, guinea pig
E	Acute heat exposure or exercise	Increase in plasma	Human, dog, rat
NE	Long-term heat acclimation	Decrease in plasma or urine	Human, guinea pig

^aData compiled from Zeisberger and Roth (1988) and Fregly and Blatteis (1996). AVP, arginine vasopressin; ACTH, adrenocorticotrophic hormone; PRA, plasma renin activity; ANG II, angiotensin II; ALD, aldosterone; TSH, thyroid-stimulating hormone; T_3 , triiodothyronine; T_4 , thyroxine; E, epinephrine; NE, norepinephrine.

and monocyte motility and emigration, enhanced phagocytosis, increased production and activity of interferon, increased T helper cell activation and antibody production, and induction of cytoprotective heat shock proteins in host cells. These effects of hyperthermia are thought to improve the pathogen-eliminating capacity by hindering the spread of local infection. The ability to warn the immune system distant from the site of primary infection suggests that hyperthermia or infectious fever has a messenger role for the immune system.

Brain Function in Hyperthermia

The Heat Stroke Syndrome

During prolonged hyperthermia, with a body temperature of 41°C or even higher, the brain suffers severe damage, which leads frequently to death. Alterations in microvascular permeability cause the development of cerebral edema. Postmortem findings further show microhemorrhages, tissue softening, and destruction of neurons. The victim exhibits disorientation, delirium, and convulsions. This syndrome is referred to popularly as heat stroke. Heat stroke can develop in response to severe external heat exposure. The heat illness syndrome can also be caused by exercise or work in hot environments. Perhaps the most famous example of heat stroke is the traditional 7-day pilgrimage to Mecca (Saudi Arabia), the so-called Makka Hajj. During the Hajj pilgrimage of 1982, high ambient temperatures (35–50°C), in addition to the large radiant heat load of the thermal environment resulting from the presence of a large number of human bodies in a very restricted area, caused more than 1500 cases of heat stroke. About 1100 people were treated successfully, whereas several hundred died. The use of molecular techniques has provided the basis for a number of pathophysiological changes in the heat-stressed brain.

Molecular Changes in the Heat-Stressed Brain

The expression of a number of molecules is upregulated in the brain under conditions of hyperthermia. Representative examples of these molecular changes are introduced briefly in the following sections. The role of these molecules in the hyperthermic brain, especially in the manifestation of detrimental effects, is the subject of current investigation. Knowledge about mechanisms by which all these molecules, which are induced in the hyperthermic brain, contribute to the heat stroke syndrome is still incomplete. Research is focusing on the roles of these molecules in the disturbance of microvascular permeability,

formation of edema, and cell injury within the hyperthermic brain.

Heat shock proteins Exposure of cells and tissues to stressful stimuli results in a rapid induction and synthesis of a specific set of proteins. These proteins are called heat shock proteins (HSPs) because their induction was originally discovered under hyperthermic conditions. However, because these proteins are also synthesized under the influence of nonthermal noxious stimuli, they are also called stress proteins. In the hyperthermic brain, the induction of HSPs is closely related to damaged brain areas and can thus be used as markers of cell injury. However, after induction of HSPs by a moderate hyperthermia, the brain becomes more resistant to a subsequent severe hyperthermia. HSPs may therefore have some neuroprotective properties. Experimental studies demonstrated that there is a correlation between the induction of HSPs and the breakdown of the blood–brain barrier in hyperthermia.

Glial fibrillary acidic protein Hyperthermic brain injury is accompanied by an activation of glial cells as indicated by the expression of the glia cell marker protein glial fibrillary acidic protein (GFAP). Under normal conditions, this protein is predominantly expressed in astrocytes. Under pathophysiological conditions, such as heat stress, increased induction of this protein is associated with a breakdown of the blood–brain barrier and vasogenic edema. The increased synthesis of GFAP is a part of the reactive response to almost any type of central nervous system (CNS) injury. Following injury, glia cells are activated to accomplish metabolic and structural roles in the CNS.

Cytokines Activation of the immune system during infection or inflammation is characterized by the expression of cytokines in the brain. Interleukin-1 (IL-1) is regarded as the most important cytokine to cause modifications of brain-controlled functions during infection. Evidence shows that the increased expression of IL-1 within the CNS is associated with brain injury. Heat stroke-induced cerebral ischemia and neuronal damage are accompanied by an increased production of IL-1 in the damaged brain. Because the survival time of rats after experimentally induced heat stroke can be prolonged by treatment with an IL-1 receptor antagonist, a role of IL-1 (and possibly other cytokines) in the manifestation of the heat stroke syndrome is indicated.

Prostaglandins A role of prostaglandins in fever is well established. Inhibition of cyclooxygenases, the

enzymes responsible for prostaglandin formation, suppresses fever. Also, the hyperthermic response to heat stress is reduced slightly by treatment with cyclooxygenase inhibitors. In addition, the hyperthermia-induced formation of edema and cellular damage in the brain are reduced by inhibitors of prostaglandin synthesis. Thus, prostaglandins seem to contribute to the detrimental effects that are induced by hyperthermia in the brain.

Nitric oxide and carbon monoxide Nitric oxide (NO) and carbon monoxide (CO) are novel chemical messenger molecules that are involved in physiological and pathophysiological processes within the brain. NO and CO are produced by specific enzymes: NO synthases and heme oxygenases. Both of these enzymes are upregulated in the brain under conditions of heat stress. In the brain, the products of these enzymes seem to contribute to the disturbance of microvascular permeability in hyperthermia.

Adaptation to Heat

Adaptation has been defined as a set of modifications within organisms that take place during prolonged or repeated actions of an environmental stimulus, a so-called stressor. The process of adaptation improves the ability to cope with stressor-induced disturbances that interfere with the homeostasis of an organism. Such disturbances (i.e., repeated or continuous exposure to heat) evoke specific and nonspecific responses and can finally lead to functional and morphological modifications that counteract the disturbance. As a nonspecific response to various stressors, including heat, central nervous circuits are activated that lead to the stimulation of the sympathoadrenal system and the hypothalamic-pituitary-adrenal axis. This nonspecific response or emergency reaction usually decreases in its intensity in response to repeated or prolonged actions of the same stressor. Specific adaptive responses to the stressor heat may lead to changes in the capacity of the effector systems that are predominantly activated in hyperthermia and to modifications of the regulatory characteristics of the control system being responsible for the thermal homeostasis under conditions of heat stress. The most profound heat-adaptive response of the thermoeffector system in humans is the increase of the sweat rate capacity, which nearly doubles during heat acclimation. This kind of adaptation to heat develops in response to repeated exposures to heat and to repeated severe exercise. Using this mechanism, a heat stress of the same intensity leads to much smaller hyperthermia at the end of the process of heat adaptation. The second component of the stressor-specific adaptive response

to heat concerns the modifications of the regulatory characteristics. Usually, sweating is activated when the mean body temperature rises to a certain value, the so-called threshold temperature for the onset of sweating. This sweating threshold is shifted to a lower mean body temperature after repeated exposures to external heat or to repeated exercise, which includes heat load such as marathon running. This means that the heat-adapted subject starts to sweat earlier and thereby delays or even prevents the developing hyperthermia. A different type of heat adaptation is observed in residents of areas with tropical climate. Here, the onset of sweating occurs at a higher body temperature than in controls. This kind of heat adaptation can be called tolerance adaptation, meaning that the subject tolerates an increase of body temperature before sweating is activated. This observation corresponds to the physiological response of the dehydrated camel (see [Figure 3](#)) and may be regarded as an avoidance of sweating (i.e., a loss of water) when living continuously in a hot climate.

Malignant Hyperthermia

Malignant hyperthermia is caused by a rare, genetically fixed defect in the transport of calcium within the sarcoplasmic reticulum. This defect has been identified as a mutation of the ryanodine receptor (calcium release channel) in the membrane of the sarcoplasmic reticulum. Some inhalation anesthetics such as halothane or isoflurane cause excessive release of calcium from the sarcoplasmic reticulum in genetically susceptible subjects. Also, the muscle relaxant succinylcholine can act as a triggering agent for malignant hyperthermia. As a consequence, uncoordinated muscle contractions with a tremendous rise in oxygen consumption and metabolic rate induce a severe hyperthermia that is accompanied by acidosis, tachycardia, or cardiac arrhythmias. After its initiation by one of the triggering substances, malignant hyperthermia leads frequently to death.

See Also the Following Articles

Critical Thermal Limits; Heat Resistance; Hypothermia; Prostaglandins; Thermal Stress; Thermotolerance, Thermoresistance, and Thermosensitivity.

Further Reading

- Adolph, E. F. (1956). General and specific characteristics of physiological adaptations. *American Journal of Physiology* **184**, 18–28.
- Blatteis, C. M. (1997). Thermoregulation: tenth international symposium on the pharmacology of thermoregulation.

- Annals of the New York Academy of Sciences* **813**, 111–116.
- Blatteis, C. M. (2004). Fever: pathological or physiological, injurious or beneficial? *Journal of Thermal Biology* **28**, 1–13.
- Brück, K. (1983). Heat balance and regulation of body temperature. In: Schmidt, R. F. & Thews, G. (eds.) *Human physiology*, pp. 531–547. Berlin: Springer.
- Brück, K. and Zeisberger, E. (1990). Adaptive changes in thermoregulation and their neuropharmacological basis. In: Schönbaum, E. & Lomax, P. (eds.) *Thermoregulation: physiology and biochemistry*, pp. 255–307. New York: Pergamon Press.
- Fregly, M. J. and Blatteis, C. M. (1996). *Handbook of physiology* (vol. 1, sect. 4). New York: Oxford University Press.
- IUPS Thermal Commission. (2001). Glossary of terms for thermal physiology. *Japanese Journal of Physiology* **51**, 245–280.
- Khogali, M. and Hales, J. R. S. (1983). *Heat stroke and temperature regulation*. Sydney: Academic Press.
- Rothwell, N. J. and Relton, J. K. (1993). Involvement of cytokines in acute neurodegeneration in the CNS. *Neuroscience and Biobehavioral Reviews* **17**, 217–227.
- Rowell, L. B. (1983). Cardiovascular adjustments to thermal stress. In: Shepard, J. T. & Abboud, F. M. (eds.) *Handbook of physiology* (vol. III), pp. 967–1023. Baltimore, MD: Waverly Press.
- Schmidt-Nielsen, K., Schmidt-Nielsen, B., Jarnum, S. A. and Houpt, T. P. (1957). Body temperature of the camel and its relation to water economy. *American Journal of Physiology* **188**, 103–112.
- Selye, H. (1950). *The physiology and pathology of exposure to stress*. Montreal: Acta Inc. Medical.
- Sharma, H. S. and Westman, J. (1998). *Progress in brain research: brain function in hot environments* (vol. 115). Amsterdam: Elsevier.
- Zeisberger, E. and Roth, J. (1989). Changes in peripheral and central release of hormones during thermal adaptation in the guinea pig. In: Malan, A. & Canguilhem, B. (eds.) *Living in the cold*, pp. 435–444. London: John Libbey Eurotext Ltd.
- Zeisberger, E. and Roth, J. (1996). Central regulation of adaptive responses to heat and cold. In: Blatteis, C. M. & Fregly, M. J. (eds.) *Handbook of physiology* (vol. I), pp. 579–595. New York: Oxford University Press.

Hyperthyroidism

I M Lesser and D L Flores

Harbor-UCLA Medical Center, Torrance, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by I M Lesser, volume 2, pp 439–440,

© 2000, Elsevier Inc.

Hypothalamic-pituitary-thyroid (HPT) axis

T4 and T3

Regulatory feedback loop between the brain (hypothalamus and pituitary gland) and the thyroid gland.

Circulating thyroid hormones (thyroxine and triiodothyronine, respectively) that have widespread effects throughout the body.

Causes of Hyperthyroidism

Behavioral and Neuropsychiatric Aspects of Hyperthyroidism

Association between Psychological Stress and Hyperthyroidism

Proposed Mechanism of Action Linking Stress and Hyperthyroidism

Glossary

Graves' disease A type of hyperthyroidism thought to be an autoimmune disease.

Hyperthyroidism Excessive functional activity of the thyroid gland leading to increased metabolism and symptoms such as weight loss, fatigue, anxiety, tremor, and increased heart rate.

Causes of Hyperthyroidism

There are multiple causes of hyperthyroidism, all leading to a state of hypermetabolic activity. As a final common pathway, all have an increase in circulating thyroid hormones (free thyroxine [T4] and free triiodothyronine [T3]), which in turn lead to a variety of symptoms: increased heart rate, weight loss, tremulousness, fatigue, anxiety, weakness, and increased tolerance to cold. The most commonly occurring type of hyperthyroidism is Graves' disease, first described in the 1830s and now known to be an autoimmune disorder. Other causes of hyperthyroidism include toxic nodular goiter, thyroid adenomas, and several types of thyroiditis (inflammatory conditions).

Each may have a different clinical presentation, but they share many of the same symptoms as well. Diagnosis is made by clinical history, laboratory determinations, and imaging studies of the thyroid gland.

Behavioral and Neuropsychiatric Aspects of Hyperthyroidism

Behavioral and psychological symptoms may be the presenting and most prominent aspects of hyperthyroidism. Patients complain of being anxious, restless, agitated, irritable, dysphoric, emotionally labile, and angry; having poor concentration and at times trouble thinking; feeling weak and fatigued; and having insomnia. More serious, and more rarely, paranoid thinking or frank psychosis can occur. The most common group of psychiatric disorders, in which the presenting symptoms may be similar, are the anxiety disorders, specifically, generalized anxiety disorder and panic disorder. Less often, there may be some symptoms overlapping with depression, but the full depressive syndrome usually is not present in the hyperthyroid individual. Even more rarely, manic symptoms may occur. Because of these symptoms, it is not uncommon for patients with Graves' disease to be prescribed psychotropic medication prior to recognition and work-up of their thyroid illness.

Association between Psychological Stress and Hyperthyroidism

The association of stress with hyperthyroidism dates back to the original case reports of the early 1800s. In 1825, Parry described the onset of a hyperthyroid state in a young woman who became symptomatic after falling down the stairs in a wheelchair. In 1835, Graves described the disease, which now bears his name, and implied that there was an innate mental disposition to the disorder. Since then, the association between stress and hyperthyroidism has been recognized, although the universality of this finding has been questioned. The majority of studies (but not all) have shown that stressful events, both negative and positive, in the year prior to the onset of the illness are more commonly seen in patients with Graves' disease than in patients with other forms of thyroid disease and in control subjects without thyroid illness. There also have been reports of an increased prevalence of Graves' disease in people under severe stress, such as refugees from German concentration camps and from the civil war in the former Yugoslavia, where there was a fivefold increase in disease incidence.

Proposed Mechanism of Action Linking Stress and Hyperthyroidism

The mechanism by which stress may lead to hyperthyroidism or, more specifically, to Graves' disease is not clear. Because Graves' disease is an autoimmune disorder, hypotheses have been made regarding the link between stress and the immune system. It has been shown that stress can affect a number of neurohormones that modulate the immune response, specifically decreasing the efficacy of this response. Stress affects the function of the hypothalamic-pituitary-thyroid (HPT) axis by suppressing the secretion of pituitary thyroid-stimulating hormone (TSH) and blunting its response to thyrotropin-releasing hormone. Furthermore, the HPT axis has feedback loops with the hypothalamic-pituitary-adrenocortical (HPA) axis, which also is very involved in response to stress. It is possible that individuals with a genetic predisposition to thyroid disease could, if their immune responsiveness is compromised, develop Graves' disease. Stress, and its effects upon immune function, could be one of the intervening variables involved in the cascade of events leading to this disorder. This theory has not been proven with any certainty, but remains a testable hypothesis.

See Also the Following Articles

Autoimmunity; Graves' Disease (Thyrotoxicosis); Hypothyroidism; Thyroid Hormones.

Further Reading

- Braverman, L. E. and Utiger, R. D. (eds.) (1996). *Werner and Ingbar's the thyroid* (7th edn.) Philadelphia, PA: Lippincott-Raven.
- Dayan, C. M. (2001). Stressful life events and Graves' disease revisited. *Clinical Endocrinology* 55, 13–14.
- Graves, R. J. (1835). Newly observed affection of the thyroid gland in females. *London Medical Surgery Journal* 7, 516–517.
- Harris, T., Creed, F. and Brugha, T. S. (1992). Stressful life events and Graves' disease. *British Journal of Psychiatry* 161, 535–541.
- Kung, A. W. (1995). Life events, daily stresses and coping in patients with Graves' disease. *Clinical Endocrinology* 42, 303–308.
- Lesser, I. M., Rubin, R. T., Lydiard, R. B., et al. (1987). Past and current thyroid function in subjects with panic disorder. *Journal of Clinical Psychiatry* 48, 473–476.
- Mizokami, T., Li, A. W., El-Kaissi, S., et al. (2004). Stress and thyroid autoimmunity. *Thyroid* 14, 1047–1055.
- Parry, C. H. (1825). *Collections from the unpublished writing of the late C. H. Parry* (vol. 2). London: Underwoods.

- Paunkovic, N., Paunkovic, J., Pavlovic, O., et al. (1998). The significant increase in incidence of Graves' disease in eastern Serbia during the civil war in the former Yugoslavia (1992–1995). *Thyroid* 8, 37–41.
- Sonino, H., Girelli, M., Boscaro, M., et al. (1993). Life events in the pathogenesis of Graves' disease: a controlled study. *Acta Endocrinology* 128, 293–296.
- Stern, R. A., Robinson, B., Thorner, A. R., et al. (1996). A survey study of neuropsychiatric complaints in patients with Graves' disease. *Journal of Neuropsychiatry and Clinical Neurosciences* 8, 181–185.
- Weisman, S. A. (1958). Incidence of thyrotoxicosis among refugees from Nazi prison camps. *Annals of Internal Medicine* 48, 747–752.
- Winsa, B., Adami, H., Bergstrom, R., et al. (1991). Stressful life events and Graves' disease. *Lancet* 338, 1475–1479.
- Whybrow, P. C. (1996). Behavioral and psychiatric aspects of thyrotoxicosis. In: Braverman, L. & Utiger, R. D. (eds.) *Werner and Ingbar's the thyroid* (7th edn., pp. 696–700). Philadelphia, PA: Lippincott-Raven.

Hyperventilation

C Bass

John Radcliffe Hospital, Oxford, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by C Bass, volume 2, pp 441–445, © 2000, Elsevier Inc.

Physiology

Causes

Clinical Features

The Hyperventilation Syndrome

Establishing the Diagnosis of Hyperventilation

Hyperventilation and Its Relation to Anxiety and Panic

Chronic Hyperventilation

Hyperventilation, Panic, and Asthma

Management of Hyperventilation

Breathing Exercises and Asthma

Glossary

<i>Alveolar PCO₂</i> (PACO ₂)	The partial pressure of carbon dioxide in alveolar space.
<i>Arterial PCO₂</i> (PaCO ₂)	The partial pressure of carbon dioxide in arterial blood.
<i>Dysfunctional breathing</i>	An abnormal pattern of breathing causing breathlessness, chest tightness, and other nonrespiratory symptoms that may not be associated with hyperventilation but that may respond to breathing retraining.
<i>End-tidal PCO₂</i> (PetCO ₂)	An uninvase measure of CO ₂ that is very close to arterial CO ₂ .
<i>Hyperventilation provocation test</i> (HVPT)	The patient is asked to hyperventilate for up to 3 min in an attempt to reproduce symptoms.

Hypocapnia

The reduction of CO₂ in the body that is the result of hyperventilation.

Respiratory alkalosis

A drop in PCO₂ with an arterial pH above 7.40 caused by hyperventilation.

Tetany

An abnormal increase in the excitability of nerves and muscles resulting in muscular spasms.

Physiology

Hyperventilation (HV) is a physiological term implying breathing in excess of metabolic requirements, thus producing arterial hypocapnia. The approach to HV requires a clear understanding of the physiological background of respiratory control. If carbon dioxide production remains constant and ventilation increases, there is an increased washout of carbon dioxide (i.e., hypocapnia) from the alveolar space with a consequent fall of alveolar PCO₂ (PACO₂) and arterial PCO₂ (PaCO₂). Stated another way, PaCO₂ below the normal range is invariably associated with a raised alveolar ventilation and excessive respiratory drive (see **Figure 1**). HV cannot be present without hypocapnia. Increases in ventilation without hypocapnia, respiratory tics, and other disorders of respiratory pattern are not HV. Following 15–20 min of voluntary hyperventilation to below 20 mmHg, hypocapnia can be maintained with only an occasional large breath, but even chronic persistent HV is associated with at least a 10% increase in ventilation. The cause or causes of this increased drive to breathe must always be sought in any individual with a HV-related disorder.

Hypocapnia and respiratory alkalosis develop rapidly after the onset of HV. If hypocapnia is sustained, renal compensation occurs through increased renal excretion of bicarbonate, but this is probably unimportant in humans. Tetany is often thought to be due

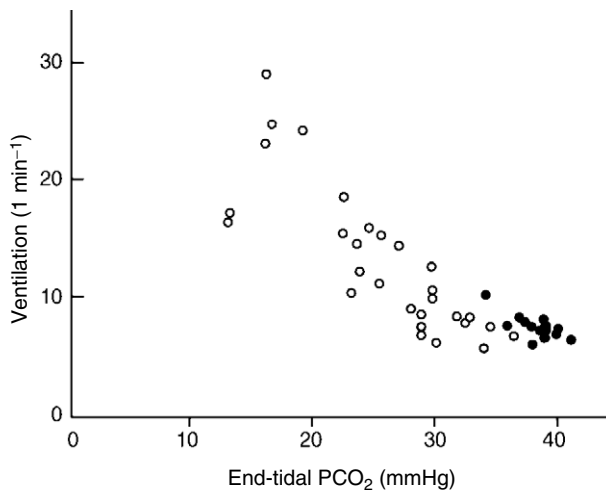


Figure 1 Relation between minute ventilation and PetCO_2 during resting breathing in eight normal subjects (closed symbols) and 16 patients with chronic hyperventilation (open symbols; two runs per subject). Each symbol is the mean of 40–100 breaths. From Gardner, W. N. (1996), *The pathophysiology of hyperventilation disorders*, *Chest* **109**, 516–534.

to a fall in serum calcium, but this has never been confirmed. There is, however, a rapid and consistent fall in the serum phosphorus level, and this provides a more likely explanation for the increase in neuronal excitability. Hypocapnic alkalosis is associated with a reduction in cerebral blood flow, which is probably due to a change of pH rather than PCO_2 , and is associated with significant cerebral hypoxia. Cerebral blood flow is linearly related to PaCO_2 , and there is a 2% reduction of arterial blood flow for each 1 mmHg decline in PaCO_2 . This means that severe HV leading to profound hypocapnia of 15–20 mmHg can reduce the cerebral circulation by approximately 50%. Hypocapnia can also cause abnormal esophageal motility and increases in tone in the muscles in the colon. Patients with known coronary vasospasm with or without significant fixed coronary artery obstructions can develop coronary vasospasm and reduced coronary perfusion following hypocapnia; this may be associated with the development of angina and ischemic electrocardiographic changes. Electroencephalographic recordings during HV often reveal paroxysmal or continuous high-voltage slow-wave activity of approximately $2\text{--}5 \text{ cycles s}^{-1}$.

Causes

HV can be caused by a number of factors, which often overlap and interact. Organic causes include a variety of factors such as drug effects (aspirin overdose), the early stages of alcohol withdrawal, respiratory disease such as asthma and pulmonary embolus, central nervous system disorders, and

chronic pain. Physiological causes include altitude acclimation, heat, and exercise. HV is also induced by speech, and progesterone in women reduces PaCO_2 by up to 8 mmHg in the second half of the menstrual cycle and even lower during pregnancy. Psychological factors include anxiety, depression, and panic attacks, but the relationship between anxiety and breathlessness is not simple.

Clinical Features

HV is clinically significant only if it results in physical complaints that are distressing or disabling. The way in which these symptoms are produced and interpreted is important in understanding why HV becomes a problem for some patients. A knowledge of the physiology suggests that the clinical manifestations of HV may be diverse and can involve almost any system in the body. For this reason, HV is said to have replaced syphilis as the great mimic in medical practice. No symptoms are truly specific for HV, but certain patterns are discernible. Common neurological complaints include dizziness, syncope, and tingling of the hands and feet (paresthesia), which may be unilateral and mimic a stroke; respiratory complaints include sensations of breathlessness or the inability to take a satisfactory breath (air hunger); cardiovascular manifestations are commonly chest pains often mimicking angina, palpitations, and racing pulse; gastrointestinal complaints include air swallowing (aerophagy) resulting in a bloated sensation, flatulence, and heartburn; musculoskeletal symptoms include muscle aches (a consequence of increased muscle tone with muscular tightness) and tetany (extremely rare); and general symptoms may include chronic and easy fatigability, weakness, the sensation of feeling cold, and poor concentration.

Although symptoms can be described in terms of organs and systems, it is probably more useful to divide them into those due to an increase in neuronal excitability; those due to a constriction of the blood vessels, especially in the cerebral, coronary, and skin circulations; and those of uncertain cause.

The Hyperventilation Syndrome

The term hyperventilation syndrome, first proposed in the 1930s, came under attack in the 1990s, and it was recommended that it be abolished. The reasons for this proposal were that (1) there were no established valid criteria, (2) the HV provocation test (designed to reproduce symptoms in patients with the syndrome) was found to be unreliable, (3) it is often impossible to establish the diagnosis in the absence of demonstrated hypocapnia, and (4) the link

between panic attacks and HV was found to be tenuous. Because most symptoms reported by patients did not require a fall in PaCO_2 , other terms such as dysfunctional breathing were suggested as an alternative. This umbrella term allows the inclusion of patients with and without HV, and it moves the focus of attention from physiological hypocapnia to a more pragmatic clinical approach.

Patients with dysfunctional breathing have an unsteadiness of breathing in response to stimuli such as exercise and a period of voluntary overbreathing, increased respiratory rate, abnormal orthostatic increases of respiratory gas exchange, predominantly intercostal respiratory effort, and frequent sighing. But not all patients with dysfunctional breathing hyperventilate or have hypocapnia.

The diagnosis of dysfunctional breathing can be difficult because the overbreathing aspect of the symptom complex may be episodic and difficult to demonstrate. The characteristic symptoms are common to other diseases, and there is no standard diagnostic test. This may lead to the underrecognition of the effects of abnormal breathing patterns, and symptoms may be wrongly attributed to other causes, resulting in inappropriate investigations and ineffective treatment, specially in patients with coexisting lung disease.

Establishing the Diagnosis of Hyperventilation

HV can coexist with other disorders, and it is important to decide which, if any, of the patient's symptoms could be due to HV. In particular, HV should be considered in any patient with a long history of unexplained physical symptoms involving a number of different organ systems.

Acute HV may be obvious from the fast rate and increased depth of breathing accompanied by the constellation of symptoms of hypocapnia. Because the disorder may be intermittent, however, a low PaCO_2 is not diagnostic; spot arterial or end-tidal samples may be normal. Conversely, making the measurement can induce overbreathing. Ambulatory transcutaneous measurement of PCO_2 for continuous monitoring is not yet responsive enough to be clinically useful, but it has led to important findings in recent research studies of *in vivo* panic attacks. End-tidal PCO_2 is sometimes measured uninvvasively by capnograph sampling from a nostril as part of lung function testing; the reduction of PCO_2 at rest, or during or after various provocations such as exercise, may suggest that symptoms are related to HV.

A commonly employed test is the HV provocation test (HVPT); the patient is asked to hyperventilate for

3 min (with concomitant monitoring of end-tidal PCO_2) and to keep PCO_2 below 20 mmHg for at least 1 min. After 3 min, the patient completes a checklist of symptoms experienced during overbreathing. Although this test has limitations (as previously mentioned), it is often used as a diagnostic aid and has some clinical utility, especially when the physician is attempting to help the patient reattribute physical symptoms to HV (and not some imagined sinister disease). Because there is no widely accepted and reliable diagnostic test, the incidence of HV has probably been underestimated.

Hyperventilation and Its Relation to Anxiety and Panic

Similarities have long been noted between the symptoms of panic attacks and HV. It has recently been established that CO_2 inhalation (and to a lesser extent HV) causes more anxiety and panic in patients with panic disorder than in normal controls. One theory proposed to explain this phenomenon is that the respiratory drive in panic disorder patients is controlled by abnormally sensitive CO_2 receptors in the central nervous system. Klein suggested that panic disorder patients possess an abnormal suffocation detector that processes asphyxia-related causes. According to Klein, HV among panic disorder sufferers is a result of the body's adaptive attempts to reduce panicogenic CO_2 . That is to say, chronic HV and sighing may adaptively keep PCO_2 below a depressed suffocation alarm threshold. The nature of the specific respiratory abnormality in panic disorder remains controversial, however, and investigators disagree on the interpretation of recent experimental data. Whereas some authors have interpreted biological challenge data as supporting a biological etiology for panic disorder, other investigators have argued that psychological factors (e.g., perceived safety, perceived control, or proximity to a safe person in the laboratory) mediate patients' responses to these challenges and that cognitive theories of panic disorder can account for these findings.

Recent work with ambulatory monitoring of CO_2 has failed to confirm the view that HV is a necessary factor for panic to occur. Although hypocapnia has been demonstrated to occur during a spontaneous panic attack in the laboratory, in a recent study there was a decrease in PCO_2 in only 1 of the 24 registered panic attacks that lasted 3 min, and even during this attack the degree of HV was not impressive. It is not possible, however, to discount a role for HV in the production of symptoms experienced during panic attacks. The crucial link between HV (when it occurs) and panic may be cognitive; panic attacks

occur when the bodily and mental sensations accompanying HV are interpreted as signs of impending personal catastrophe, such as madness or heart attack (see [Figure 2](#)).

Chronic Hyperventilation

An initial attack of HV caused by a combination of physiological, psychological, and possibly pathological factors could create a vicious circle and lead to increasing anxiety and further HV. The patient may misattribute the first attack of HV to serious disease and then perpetuate the disorder by becoming more anxious and breathing more deeply to relieve chest discomfort. CO₂ homeostasis is then reset at a lower level. Such a sequence could rapidly lead to invalidism and chronic HV. In these patients, PCO₂ is either chronically low at rest and throughout a number of provocations or only mildly reduced at rest but precipitated into prolonged reduction by either exercise or voluntary overbreathing.

Hyperventilation, Panic, and Asthma

The prevalence of anxiety disorders (including panic) is above expected levels in patients with asthma, and asthma is more common in people with psychiatric illness. The relationship appears strongest in those

with severe asthma and those with severe anxiety. Panic and anxiety can directly exacerbate asthma symptoms through HV.

In a recent longitudinal study, it was found that active asthma predicted subsequent panic disorder and that panic disorder predicted subsequent active asthma, this effect being stronger in women and in subjects younger than 30 years old. This relationship between physiological and psychological factors in asthma offers opportunities for research.

Management of Hyperventilation

The main aim of treatment is to reassure patients that they do not have a serious life-threatening illness and then to treat the underlying cause of the HV. If no cause is found or the cause is untreatable, it is often possible to teach patients to understand and adjust to the symptoms.

First, organic causes for the presenting symptoms (e.g., asthma or pulmonary embolus) should be excluded. Two of the main treatments used in the management of patients with HV-related symptoms in recent years have been breathing retraining and cognitive reattribution of physical symptoms to HV. Breathing retraining involves encouraging the patient to breathe in a slow regular fashion using the diaphragm; reattribution is aimed at helping the patient

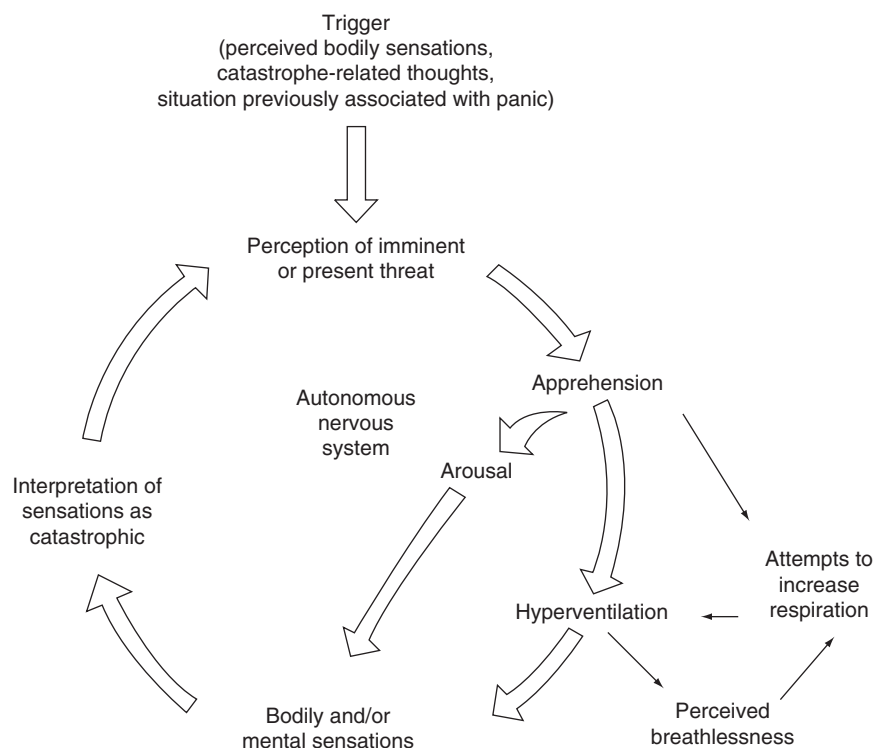


Figure 2 Principal components of the cognitive model of panic attacks in which hyperventilation is involved in the production of symptoms.

to ascribe the physical symptoms to anxiety and at explaining the relationship between stress and HV. In practice, however, it is very difficult to determine the relative contributions of these two therapeutic components to outcome.

Recent research supported the view that training in controlled breathing is not of specific benefit in patients identified as hyperventilators by the provocation test. Indeed, breathing retraining has been referred to as a rational placebo that is effective not by restoring levels of CO₂ but by inducing relaxation, providing a structured task and distraction, and giving the patient a sense of control and mastery. This may distract the patient from feelings of anxiety and promote self-reliance.

Breathing Exercises and Asthma

It has been reported that up to 30% of adults with asthma in the community may have symptoms suggestive of functional breathing disorders, and there is some evidence of a role for breathing exercises in the treatment of asthma complicated by these abnormal breathing patterns. A review of psychotherapeutic interventions for adults with asthma found conflicting evidence, however, with some positive studies but with insufficient data to draw firm conclusions.

Further Reading

- Bass, C. (1997). Hyperventilation syndrome: a chimera? *Journal of Psychosomatic Research* 42, 421–426.
- Gardner, W. N. (1996). The pathophysiology of hyperventilation disorders. *Chest* 109, 516–534.
- Garssen, B., de Ruiter, C. and van Dyck, R. (1992). Breathing retraining: a rational placebo? *Clinical Psychology Review* 12, 141–153.
- Hasler, G., Gergen, P., Kleinbaum, D., et al. (2005). Asthma and panic in young adults; a 20 year prospective community study. *American Journal of Respiratory and Critical Care Medicine* 171, 1224–1230.
- Klein, D. F. (1993). False suffocation alarms, spontaneous panics, and related conditions. *Archives of General Psychiatry* 50, 306–317.
- Salkovskis, P. (1988). Hyperventilation and anxiety. *Current Opinion in Psychiatry* 1, 76–82.
- Thomas, M. (2003). Breathing exercises and asthma. *Thorax* 58, 649–650.
- Thomas, M., McKinley, R., Freeman, E., et al. (2001). Prevalence of dysfunctional breathing in patients treated for asthma in primary care: cross sectional survey. *British Medical Journal* 322, 1098–1101.
- Thomas, M., McKinley, R., Freeman, E., et al. (2003). Breathing retraining for dysfunctional breathing in asthma: a randomized controlled trial. *Thorax* 58, 110–115.

Hypnosis

W G Whitehouse, E C Orne and M T Orne

University of Pennsylvania School of Medicine,
Philadelphia, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M T Orne and W G Whitehouse, volume 2, pp 446–452, © 2000, Elsevier Inc.

Historical Background

Induction of Hypnosis

Hypnotizability Assessment and the Phenomena of Hypnosis

Hypnotic Interventions for Stress Management

Conclusion

Glossary

Hypnotic induction

A series of suggestions that focus attention and bring about the transition from ordinary waking experience to hypnosis.

Hypnotizability

The potential to experience hypnosis; a relatively stable trait that varies from person to person and that can be assessed by standardized test procedures.

Nonspecific effect

A therapeutic outcome (e.g., symptom relief) that follows a hypnotic procedure but that occurs regardless of the patient's ability to experience deep hypnosis.

Historical Background

Special healing properties have been ascribed to trancelike conditions throughout civilizations and

time. Some of the earliest examples can be traced back 4000 years to ancient Egypt and by the 5th century BC to the Greek empire. The priests of these times established healing sanctuaries or sleep temples, where people suffering from mental or physical problems could seek relief. In exchange for some sacrifice or atonement, the priest or priestess performed incantations and rituals designed to heighten suggestibility and promote a trancelike sleep, during which the individual's dreams could be affected by healing communications from the gods. Similarly, witchcraft practiced during the Middle Ages, exorcisms to alleviate demonic possession, Afro-Caribbean voodoo rituals, and faith-healing in relatively modern times each appear to involve elements of what is today subsumed by the term hypnosis.

Historically, explanations for the effectiveness of many of these practices appealed to supernatural or metaphysical causes. More modern perspectives arose from controversies surrounding the work of the Austrian physician, Franz Anton Mesmer (1734–1815) and his followers. Working in France at the time of the American Revolutionary War, Mesmer observed that some patients derived benefit from the passing of magnets over their bodies. In due time, Mesmer concluded that he, himself, possessed the critical animal magnetism, a putative fluid within the body that, in combination with certain accoutrements, could be transferred to others as needed to heal them. This claim expressly identified the source of the healing power as a property of the magnetizer. Indeed, one of Mesmer's students, the Marquis de Puysegur, claimed to have successfully magnetized a tree on his estate, where some of his peasant workers obtained relief from their ailments. Eventually, de Puysegur found that he could influence the behavior of mesmerized patients merely by talking with them, much the same way that hypnosis is used today. The fluid theory of animal magnetism was soon discredited, however, in a series of investigations conducted by such scientific luminaries as Benjamin Franklin and Antoine-Laurent Lavoisier (1784). Nevertheless, the therapeutic effects – at first dismissed as the products of imagination and suggestion – were too striking to be ignored for very long.

By the 1840s, the current term hypnosis had been coined by James Braid, a surgeon from Edinburgh, Scotland, who conceptualized the process as one of focusing attention, leading ultimately to a state of artificial sleep. Subsequently, demonstrations of its analgesic uses in surgery – many by James Esdaile, a physician employed by the East India Company in Calcutta – were being documented in British medical periodicals. It was not long before hypnosis was adopted for use in the treatment of psychiatric and

neurological disorders following a number of celebrated cases by such renowned academics and physicians as Hippolyte Bernheim, Jean-Martin Charcot, Sigmund Freud, Pierre Janet, and Morton Prince. Today, hypnosis is commonly employed in clinical practice as an adjunctive psychological technique for the management of a variety of conditions, both acute and chronic, such as pain, anxiety, addiction, mood disturbance, dissociative disorders, and stress-related disorders.

Induction of Hypnosis

A typical hypnotic induction begins with the establishment of rapport between the hypnotist (sometimes referred to as the operator) and the to-be-hypnotized person (the subject). This serves not only to assure the subject of the hypnotist's competence and trustworthiness, but also to create a favorable expectation for a positive response to hypnosis (i.e., by explaining how hypnosis will be used with the subject and correcting any misconceptions about the procedure the individual may have). After this, the subject is invited to relax and be comfortable and to pick some target on which to fix his or her gaze. The hypnotist then exhorts the subject to experience a growing feeling of relaxation and proceeds to direct attention to simple, suggested perceptual changes (e.g., "Your eyelids are becoming heavier and heavier. . . they will soon shut of their own accord. . ."). The specific suggestions are often designed to take advantage of normal bodily responses to the situation, such as the strain and fatigue associated with prolonged visual fixation. Because most individuals readily share such experiences and respond successfully, judiciously selected suggestions of this sort may also serve to enhance expectations for further responsiveness to hypnosis. Once eye closure has been accomplished, the hypnotist will often test whether the subject is truly beginning to experience the condition of hypnosis or is merely complying. Thus, the hypnotist may suggest that it is becoming increasingly difficult, perhaps impossible, for the subject to lift his or her hand. The majority of subjects who pass this challenge (i.e., fail to lift their hand) are indeed usually capable of positive responses to a number of ensuing hypnotic suggestions, provided they are administered in a generally graded sequence of increasing difficulty.

Often early suggestions attempt to promote subjective alterations involving sensory and motor responses, such as the hand growing lighter and lighter and beginning to float upward from the armrest of the subject's chair. Coupled with complementary remarks that as the hand rises the subject will sink deeper and deeper into hypnosis, such a relatively easily experienced suggestion is but one of several deepening techniques

intended to shift the subject's awareness from the subject's ambient surroundings to a nearly exclusive focus on experiencing the hypnotist's suggestions. At this point, a suitable subject may experience marked distortions of perception, memory, and mood in response to carefully and methodically linked suggestions that are appropriate to the individual's level of hypnotic skill.

Hypnotizability Assessment and the Phenomena of Hypnosis

Contrary to the view, espoused by the eighteenth-century mesmerists and many modern-day stage hypnotists, that the power of hypnosis resides in the influence of the hypnotist, the scientific evidence is compelling that only a subset of people are capable of experiencing the full range of phenomena that constitute the domain of hypnosis. Thus, it is clear that the ability to respond to hypnosis is determined, in large part, by the individual about to undergo the procedure.

Formal attempts to measure hypnotic ability and to quantify aspects of hypnotic experience began in the United States in the 1930s. The most popular hypnotizability scales in use worldwide today were derived from efforts by Andre Weitzenhoffer and Ernest Hilgard of Stanford University in 1959. A tape-recorded adaptation of their Stanford Hypnotic Susceptibility Scales (SHSS), suitable for use with groups of individuals, was published in 1962 by Ronald Shor and Emily Carota Orne of Harvard University and is known as the Harvard Group Scale (HGS). The general design of these instruments includes a standardized induction and order of administration of hypnotic suggestions that is characterized by increasing difficulty over the course of 12 items. A person's hypnotizability is determined by the number of suggestions (which can range from 0 to 12) that evoke positive behavioral responses as judged by the hypnotist (SHSS) or by the subject (HGS). With the development of standardized scales for hypnotizability assessment, researchers have been able to identify three distinct clusters of hypnotic phenomena.

1. **Ideomotor responses.** One type of hypnotic response that nearly all individuals can experience to some degree is the tendency for vividly imagined bodily movement to produce corresponding physical movement on a seemingly involuntary basis. Such is the case, for example, when a subject responds to suggestions that his or her "forearm is beginning to feel lighter . . . as if a large helium-filled balloon is attached to the wrist and is gently lifting the arm . . . allowing it to float effortlessly upward and upward. . . ." Response

to ideomotor suggestions can also occur in the absence of prior hypnotic induction, indicating that some degree of suggestibility may also operate under normal waking conditions. However, although waking suggestibility is moderately correlated with hypnotic responsivity, evidence shows that the hypnotic induction procedure contributes to hypnotic outcomes. Thus, it is exceedingly rare for an individual to respond to suggestions in the waking condition and fail to respond following induction; on the other hand, many individuals' responsiveness to suggestion is substantially increased following an appropriate hypnotic induction.

2. **Challenge suggestions.** Challenge items are intended to create a contradictory subjective experience involving the inability to carry out one's own will. For instance, the hypnotized subject might be told to "interlock your fingers tightly . . . so tightly that they will be impossible to separate. In a moment, you may wish to try to separate your fingers, but you will find that you are unable to do so. Go ahead . . . try . . . just try to pull your fingers apart." If successfully passed (e.g., in this case, the fingers do not separate), challenge suggestions provide a compelling impression of external control over the individual's behavior – an impression that is often shared by the hypnotized person and onlookers.

3. **Cognitive/perceptual alterations.** Among the more difficult hypnotic suggestions are those designed to induce some kind of perceptual hallucination and those intended to alter the ability to remember personal experiences from the immediate or remote past. Hallucinatory suggestions can be either positive, in which a nonexistent stimulus is introduced by the hypnotist (e.g., it is suggested that a mosquito is buzzing annoyingly around the subject's face), or negative, in which an actual physical stimulus loses its perceptibility (e.g., the subject is told that he or she will be unable to smell anything, after which an open bottle of ammonia is passed under the nose). Suggested amnesia is an example of a commonly administered cognitive suggestion, whereby the subject is instructed that, on awakening, he or she will have no recollection of the events that transpired during the hypnotic proceedings. Similarly, posthypnotic suggestions are sometimes given during hypnosis with the intent that, following termination of hypnosis, the subject will perform a specific behavior in response to a prearranged cue but will not remember being told to do so. When properly used with suitable patients, amnesia and posthypnotic suggestions can be important clinical tools that may extend therapeutic gains outside the therapist's office.

In addition to delineating the general factor structure of hypnosis into ideomotor, challenge, and

cognitive domains, the use of standardized assessment scales has resulted in a great deal of information about hypnotizability in general. One important feature, from both clinical and research perspectives, is the manner in which hypnotizability is distributed in the general population. Hilgard and his associates found that hypnotizability, like many other human skills, follows a normal or bell-shaped pattern, in which the majority of people have the capacity to experience most hypnotic phenomena to some degree, 15–20% are generally unresponsive, and an approximately equal percentage of people are highly responsive. Recognition of such individual differences in the ability to respond to hypnotic procedures is an important consideration for its use clinically, particularly when treatment is predicated on a successful response to difficult cognitive suggestions, such as to experience analgesia (a negative hallucination) for either acute (e.g., dental) or chronic (terminal-stage cancer) pain.

Although scores on standardized assessment scales prove useful in predicting the types of hypnotic suggestions that individuals are most likely to experience, they are not absolute prognosticators of hypnotic potential. This is because, for individuals with scores in the moderately hypnotizable range, various combinations of passed items from different realms of difficulty are possible. The hypnotizability score is determined merely by the sum of items that are passed, not by their classification as ideomotor, challenge, or cognitive (which corresponds roughly to their difficulty level). Thus, although there is virtual assurance that a subject with a score of 2 (out of a possible 12 points) will be unable to experience a positive hallucination, the same cannot be said of a subject with a score of 5 or 6. Modifications of the original SHSS have been published to identify and characterize individual hypnotic profiles, but their use has been limited primarily to research purposes.

A final yet particularly important observation made possible by the availability of standardized scales of hypnotic ability is the general stability of an individual's hypnotic potential over time. This has been confirmed in numerous laboratories throughout the world using correlations among scores obtained for the same participants at varying test–retest intervals, in some cases spanning several decades. Although some research suggests that environmental circumstances can occasionally increase a subject's responsiveness to hypnotic suggestion (e.g., situations involving extreme sensory deprivation or explicit training in the skills necessary to respond positively to specific suggestions), the overall consensus of the evidence is that hypnotizability is one of several relatively enduring characteristics of an individual's personality.

Hypnotic Interventions for Stress Management

Although hypnosis is widely regarded as an effective cognitive-behavioral method for alleviating stress, there is surprisingly little formal research on this issue. There are, however, numerous clinical and experimental studies concerned with specific problems that are often regarded as stressors. From this kindred literature, we can extrapolate conclusions pertaining to the effectiveness of hypnotic techniques in reducing stress. Before doing so, however, it seems prudent to characterize the type of stress to which we refer.

Conceptualizations of stress vary widely. Some investigators have identified stress primarily as a stimulus in which some aspect of the physical or psychological environment exacts a toll on individuals and/or society; others view stress as the psychological or physiological response to such hostile environmental provocation; and still others embrace a more dynamic view whereby the individual's appraisal of the specific environmental pressures, his or her own coping resources, and available options all converge to determine the nature and extent of the stressful experience. From the perspective of scientists and practitioners concerned with the use of hypnotic interventions for stress management, each of these variant emphases holds viability. This is because hypnosis is one of the few psychological techniques that, given an appropriately responsive individual, can be introduced at various points to modify the experience of stress. Thus, it has the potential to mitigate the stimulus impact of a stressor by altering perceptual experience. It can also be enlisted to modify cognitive and affective factors that influence a subject's appraisal of the stressor and his or her ability to cope. In addition, hypnosis may provide specific relief of symptoms precipitated or maintained by exposure to stress.

Pain Control

The use of hypnosis to suppress pain provides dramatic evidence of the technique's value for stress management. Although the widespread availability and safety of chemical anesthetic agents today obviates the use of hypnosis as a general method of surgical anesthesia, there are numerous documented cases in which hypnotic techniques were effectively employed as the sole anesthetic for surgery (e.g., limb amputations, temporal lobectomy, cardiac surgery, tooth extractions, appendectomies, and cesarean sections, among others). An important caveat, however, is that the effectiveness of hypnosis in blocking instances of acute pain is often correlated with the patient's ability to be hypnotized. This has been demonstrated in both clinical samples of patients undergoing painful

medical procedures (e.g., bone marrow aspirations, lumbar punctures, and debridement and dressing of burn wounds) and experimental studies involving painful laboratory stressors (e.g., prolonged immersion of the hand in ice water). Another feature that may contribute to the efficacy of hypnosis in the control of acute pain in hypnotically responsive people is the use of direct suggestions delivered in a heterohypnotic or self-hypnotic context.

Hypnosis is decidedly less helpful in the management of certain forms of chronic pain, particularly those with a psychogenic component. Two issues are relevant here. The first concerns the impracticality of heterohypnotic treatments based on direct suggestion in blocking the experience of persistent chronic pain, most of which occurs outside the therapist's office. The alternative strategy, training in self-hypnosis, often requires regular contact with the therapist to reinforce its continued use by the patient in his or her daily life. The more formidable problem, however, concerns the powerful reinforcing value that pain behavior can acquire among family, friends, and other important members of the community. Thus, the sympathy and support provided by others – the so-called secondary gains – may serve to maintain the individual's pain and undermine the incentive to practice hypnotic techniques that otherwise might be helpful in bringing about relief. The most promising prognosis for patients suffering from functional pain of this nature is achieved by a combination of behavior therapeutic principles aimed at extinguishing pain behavior and the implementation of hypnotic or self-hypnotic approaches to minimize pain-related discomfort.

For other forms of chronic pain that do not involve a substantial functional component (e.g., pain from shingles, trigeminal neuralgia, and many types of cancer), hypnosis and self-hypnosis have proven useful as an adjunct to long-term pain management programs. Hypnosis, even self-hypnosis, derives much of its benefit from the interpersonal relationship between patient and therapist, which, among other consequences, tends to shape expectations for therapeutic improvement. In an appropriately supportive context, therefore, hypnotic techniques may prove beneficial without particular regard to the patient's overall ability to respond to hypnosis. For example, we have observed significant improvement among children and adolescents during the course of an 18-month intervention involving regular group training in self-hypnosis to control unpredictable episodes of pain associated with sickle-cell disease. Prior to the intervention, many of the children experienced frequent absences from school, occasionally for weeks

at a time, and the only treatments available included analgesics, narcotic medications, and hospitalization. Although the participants were enrolled in the study based on their expressed wish to help identify an effective nonpharmacological treatment for the disease and without regard to hypnotic ability, our findings indicated that, in this case, the ability to respond to hypnotic suggestion was not an overriding factor in determining how often self-hypnosis was practiced nor how beneficial it proved to be. Apparently, self-hypnosis training, along with consistent reinforcement of the technique during the regular group sessions, provided a much-needed coping skill; the group treatment sessions themselves were critical to maintaining motivation and providing a forum for the patients to share their common fears, misunderstandings, and concerns, which were not being adequately addressed by conventional medical management systems. For the group as a whole, the combination of these therapeutic approaches led to increased school attendance, improved nighttime sleep, less dependence on medication, and fewer bouts of pain, although the more severe episodes of pain still required nonhypnotic supplements for pain management.

Anxiety and Stress Disorders

Stress is inextricably linked to anxiety states, whether they are circumscribed by some anticipated source of realistic concern (e.g., impending open-heart surgery), are more persistent but specific fears (e.g., of elevators or spiders), or are unrelenting and pervasive hindrances to normal functioning (e.g., agoraphobia or social phobia). Hypnotic techniques have proven extremely effective in treating anxiety conditions. Available evidence suggests that there may be two distinct mechanisms responsible for the high success rate. The first, which is nonspecific to hypnosis, is the tendency for hypnotic induction to promote a profound state of relaxation and comfort in many individuals, regardless of their hypnotic capacity. As Joseph Wolpe demonstrated in his development of a behavior therapeutic approach to the treatment of anxiety disorders, relaxation is incompatible with fear. Thus, the cultivation of extreme relaxation, whether accomplished with hypnosis, self-hypnosis, or some other cognitive-behavioral technique (e.g., autogenic training, meditation, or progressive relaxation), will tend to inhibit symptoms of anxiety. The second mechanism is directly related to the ability to experience hypnotic-like conditions and applies to some of the more severe forms of anxiety and trauma-related disorders. Fred Frankel (Harvard Medical School) has compiled extensive data from case studies, which suggest that high hypnotizability may be a

risk factor for the development of phobic anxiety. The same positive association involving hypnotizability has been found for acute stress disorder, posttraumatic stress disorder, and dissociative disorders following traumatic experiences. Although such evidence appears to suggest that highly hypnotizable people may be more vulnerable to developing severe anxiety and stress-related conditions pursuant to adverse life events, the picture is far from clear. To date, the majority of studies that report such a relationship have employed cross-sectional designs, meaning that hypnotizability is measured only at the time (or shortly after) the patient is diagnosed with a particular disorder. Thus, it is unknown whether high hypnotizability constitutes a predisposition to anxiety and dissociative symptoms in the aftermath of trauma or whether it is one of several consequences of the stressful experience. In support of the latter possibility, several studies have shown that hypnotizability may be enhanced under adverse conditions, such as sensory deprivation or restricted environmental stimulation. Only studies that adopt a prospective, longitudinal design, in which hypnotizability assessments occur both before and after a stressful life event, can determine the role of hypnotizability in the causal sequence leading to stress-related psychopathology. Nevertheless, and most important, the same highly hypnotizable people who may appear to exhibit a propensity to develop anxiety and stress-related disorders also benefit markedly from psychotherapy that involves the use of hypnosis and training in self-hypnosis as a coping strategy. Similarly, hypnotic methods have been successful in treating survivors of trauma whose symptoms of hyperarousal, emotional numbing, and avoidance are maintained by uncontrollable intrusive flashbacks and recurrent nightmares. In the treatment of posttraumatic stress disorders, hypnosis is most effectively employed as an adjunctive technique, secondary to pharmacotherapy and/or supportive psychotherapy. On a cautionary note, however, the heightened suggestibility and vividness of the imaginings that hypnotized individuals often experience have been associated with a corruption in the accuracy of recollections obtained during hypnotic interviews, whereas their confidence that such memories are veridical tends to be increased. Thus, current evidence warrants that clinical professionals carefully evaluate how hypnosis should be used in the treatment of victims who claim full or partial amnesia for details of past traumatic experiences. As a relaxation technique, hypnosis and self-hypnosis can be highly beneficial; as a tool for recovering memories of past life events, hypnosis elevates the risk that believed-in, false memories may emerge.

Psychophysiological Disorders and Psychoneuroimmunology

Stressors trigger specific neuroendocrine and sympathetic nervous system responses, which, if prolonged, can lead to a number of clinical manifestations linked to parasympathetic inhibition, including dysregulation of respiratory, cardiovascular, digestive, eliminative, and sexual functions. An abundance of case studies attests to the value of hypnotic and self-hypnotic procedures in the treatment of such conditions as asthma, tension headache, irritable bowel syndrome, Crohn's disease, and insomnia. However, there is a paucity of controlled studies in which the distinctive contributions of hypnotic responsiveness and placebo factors to clinical improvement can be evaluated. The available evidence suggests that hypnosis may be as effective as related methods (e.g., biofeedback and progressive relaxation) aimed at countering sympathetic arousal in the treatment of psychophysiological disorders. The mechanism of action appears to involve the induction of a profoundly relaxed condition, which can be achieved in a majority of individuals regardless of their capacity to respond to hypnotic suggestion.

Recent molecular and pharmacological studies have identified intricate patterns of communication between the immune system and the central nervous system, which gives substance to the long-held belief that stress can precipitate illness. Hypnosis is a versatile psychological approach that can alter the perception, evaluation, or symptoms of stress and that might therefore prove effective in the management of infectious or malignant disease. Many studies have found that hypnosis is beneficial in reducing stress symptomatology, which may result in a corresponding improvement in immune function, but the benefit is often only weakly correlated with subjects' hypnotizability, suggesting that any immune enhancement related to hypnosis may be due to nonspecific effects. However, in one study conducted by our laboratory, we were able to confirm a relation between ratings of the depth of relaxation and increases in both the number and cytotoxic capacity of natural killer cells from whole-blood samples from medical students trained to manage stress using self-hypnosis. Other investigators have reported findings that suggest a relationship between the practice of hypnotic methods and indices of humoral and cellular immunity. In addition, recent studies by Gruzelier and colleagues in Great Britain have identified a positive association between hypnotizability and several indices of immunocompetence, which was most evident when hypnotic procedures included guided imagery that specifically addressed immune system dynamics.

At present, we can conclude only that hypnotic techniques appear to be effective in promoting well-being and reducing reactivity to stress, which might otherwise impair immunity to disease. The development of optimal strategies for the use of hypnosis to counteract stress-related immune compromise continues to be informed by further investigation.

Conclusion

As with any clinical intervention, hypnosis can lay claim to the efficacy of both specific and nonspecific components. Because the benefit from hypnotic therapy is a function of the treatment itself, as well as the patient's ability to experience hypnosis, several processes may operate independently or synergistically to determine therapeutic outcome. For individuals who are able to profoundly experience hypnotic suggestion, the stimulus, evaluation, or symptom features of a stressor can be directly addressed in treatment, and the same individuals may be able to avail themselves of the additional benefits that accrue from generalized relaxation and positive expectancies. For individuals of lesser hypnotic capacity, the non-cognitive aspects of hypnosis, particularly relaxation, clearly provide an effective technique for inhibiting sympathetic arousal. Finally, because hypnosis requires the establishment of a strong and trusting therapeutic alliance between the therapist and patient, mutual expectations regarding milestones for improvement, coupled with regular contact, may serve to enhance motivation sufficiently to assure adherence to treatment protocols. For some instances of stress, the ability to experience hypnosis may be paramount to effective coping, whereas, for other

stressors, simply the subject's belief in his or her ability to benefit from the exercise may prove helpful. It must be emphasized, however, that individuals seeking treatment with hypnosis should consult with a qualified physician, psychologist, or dentist who, in addition to advanced training and practice in one of these health-care specialties, has skills in the adjunctive therapeutic use of hypnotic techniques. When properly applied, hypnotic methods can have a beneficial role in the prevention and treatment of stress-related disorders.

See Also the Following Articles

Stress Management and Cardiovascular Disease; Stress Management, CAM Approach.

Further Reading

- Bowers, K. S. (1983). *Hypnosis for the seriously curious*. New York: Norton.
- Crasileck, H. B. and Hall, J. A. (1985). *Clinical hypnosis: principles and applications*. Orlando, FL: Grune & Stratton.
- Frankel, F. H. (1979). *Hypnosis: trance as a coping mechanism*. New York: Plenum.
- Gruzelier, J. H. (2002). A review of the impact of hypnosis, relaxation, guided imagery and individual differences on aspects of immunity and health. *Stress* 5, 147–163.
- Hilgard, E. R. (1965). *Hypnotic susceptibility*. New York: Harcourt, Brace & World.
- Orne, M. T. and Dinges, D. F. (1989). Hypnosis. In: Wall, P. D. & Melzack, R. (eds.) *Textbook of pain* (2nd edn., pp. 1021–1031). London: Churchill Livingstone.
- Patterson, D. R. and Jensen, M. P. (2003). Hypnosis and clinical pain. *Psychological Bulletin* 129, 495–521.

Hypocortisolism and Stress

C M Heim and U M Nater

Emory University School of Medicine, Atlanta, GA USA

© 2007 Elsevier Inc. All rights reserved.

What is Hypocortisolism?

Clinical Findings of Hypocortisolism

Stress and Hypocortisolism

Potential Mechanisms of Hypocortisolism

Effects of Hypocortisolism on Disease Vulnerability

Clinical Implications

Glossary

Cortisol

Glucocorticoid hormone synthesized and secreted from the adrenal cortex in primates. Cortisol is secreted in a diurnal rhythm, and its secretion is increased in response to acute stress. Cortisol controls hypothalamic-pituitary-adrenal axis activity through negative feedback inhibition and exerts metabolic, immune-regulatory, and behavioral effects that adapt the organism to stress.

<i>Functional somatic syndromes</i>	The presence of multiple physical symptoms that manifest without identifiable physical signs or medical abnormalities. Specific symptoms or symptom patterns are used to define syndromal disorders, such as chronic fatigue syndrome, fibromyalgia, irritable bowel syndrome, and chronic pelvic pain.	the person or others and a response of intense fear, helplessness, or horror.
<i>Glucocorticoid receptors</i>	Intracellular receptor molecules through which cortisol exerts its effects on target cells. There are two types of receptors, the mineralocorticoid receptors (MRs) and the glucocorticoid receptors (GRs). MRs are mainly located in the hippocampus, whereas GRs are widely distributed throughout the brain and periphery. During stress, MRs and GRs in concert control hypothalamic-pituitary-adrenal axis activity.	
<i>Hypothalamic-pituitary-adrenal (HPA) axis</i>	Neuroendocrine cascade that controls the secretion of glucocorticoids during stress. Corticotropin releasing hormone from the paraventricular nucleus of the hypothalamus stimulates the secretion of adrenocorticotrophic hormone from the anterior pituitary, which in turn stimulates the production and secretion of glucocorticoids (corticosterone in rodents, cortisol in primates) from the adrenal cortex.	
<i>Negative feedback inhibition</i>	A decrease in HPA axis activation in response to increasing levels of circulating glucocorticoids, and vice versa, mediated through glucocorticoid receptors at the hippocampal, hypothalamic, and pituitary levels. The feedback sensitivity of the pituitary can be determined by measuring cortisol suppression in response to the synthetic glucocorticoid dexamethasone. Standard doses of dexamethasone detect nonsuppressors (feedback resistance), whereas low doses detect hypersuppressors (increased feedback sensitivity).	
<i>Posttraumatic stress disorder (PTSD)</i>	An anxiety disorder that manifests secondary to the experience of a traumatic event and that is characterized by symptoms of reexperiencing of the event, avoidance of reminders of the event, and hyperarousal. Symptoms must last 1 month or longer and must be associated with significant impairment in work, family, or social functioning.	
<i>Stress</i>	A state caused by real or perceived challenge to homeostasis, which is accompanied by a hormonal response.	
<i>Trauma</i>	A stressful experience that involves actual or threatened death or serious injury or a threat to the physical integrity of	

Since the seminal studies by Hans Selye in 1936, stress has been associated with the activation of the hypothalamic-pituitary-adrenal (HPA) axis, ultimately resulting in the increased secretion of cortisol from the adrenal glands. The physiological effects of cortisol help the organism to maintain homeostasis under conditions of stress. The association between stress and increased cortisol secretion has been consolidated over the past decades to such an extent that stress and increased cortisol secretion nearly have become synonyms in the literature and, moreover, the presence of cortisol hypersecretion has been used to define states of stress. However, representing a major challenge for current concepts in stress research, an increasing number of studies have now provided convincing evidence that the adrenal gland can be hypoactive in some stress-related states. The landmark studies in the 1990s by Rachel Yehuda and her group on neuroendocrine dysregulation in posttraumatic stress disorder (PTSD) have catapulted the phenomenon of hypocortisolism into the center of neuroendocrine stress research. The phenomenon of hypocortisolism has since received growing attention. In fact, hypocortisolism does not merely represent a specific correlate of PTSD but has also been observed in response to severe acute, chronic, and developmental stress in both animal models and studies of human subjects. In addition, hypocortisolism has been observed in patients with chronic fatigue syndrome, fibromyalgia, and chronic pelvic pain, among other functional somatic syndromes. All these disorders are arguably related to stress and frequently coincide with mood and anxiety disorders, including PTSD. Although hypocortisolism appears to be a frequent and widespread phenomenon, its precise definition and epidemiology as well as its underlying mechanisms and physiological relevance remain speculative. It is also unknown whether there are independent or interactive genetic contributions to the manifestation of hypocortisolism in stress and illness. It has been suggested that hypocortisolism is causally involved in the pathophysiology of stress-related disorders; the persistent lack of the regulatory effects of cortisol may promote an increased vulnerability for the development of a spectrum of stress-related disorders characterized by fatigue, pain, stress sensibility, and anxiety. This model has important implications for the prevention and treatment of disorders related to hypocortisolism. This article summarizes current knowledge on the phenomenon of hypocortisolism in stress research.

What is Hypocortisolism?

There is no consensus on the definition of relative hypocortisolism in the published literature. In clinical endocrinology, the normal physiological range of urinary cortisol excretion is 20–90 $\mu\text{g day}^{-1}$. Thus, medical disorders of adrenal insufficiency, such as Addison's disease, are defined as an abnormal cortisol excretion below the normal physiological range, with endocrine challenge testing performed to elucidate the source of the insufficiency. Interestingly, adrenal insufficiency is related to fatigue, weakness, and pain. However, the stress-related phenomenon of hypocortisolism that is the focus of this article does not necessarily encompass cortisol secretion below the normal physiological range. Rather, reports of hypocortisolism in the stress literature refer to relative decreases of mean cortisol secretion in a defined study group compared to controls, with mean cortisol levels usually in the physiologically normal range. Thus, relative hypocortisolism as observed in stress-related conditions reflects subtle changes in HPA axis regulation that are uncovered when comparing groups. Consequently, to date, there are no diagnostic methods or normative values that allow us to decide whether an individual exhibits such relative hypocortisolism, independent of comparison subjects. When defining relative hypocortisolism, we must decide whether decreased cortisol levels should be measurable under basal conditions, during pharmacological dynamic testing, during stress challenge, or in all of these.

The most stringent definition of hypocortisolism might refer to decreased 24-h urinary cortisol excretion. On the other hand, certain tests, such as the adrenocorticotrophic hormone (ACTH)_{1–24} stimulation test, the corticotropin releasing hormone (CRH) test, or the low-dose dexamethasone suppression test, have the potential to be more sensitive to detect relative hypocortisolism by directly targeting the potential mechanisms of dysregulation, such as adrenocortical sensitivity and enhanced negative feedback sensitivity. However, blunted cortisol responses in a CRH stimulation test might not be readily interpreted as hypocortisolism, given that pituitary responses might be low due to elevated circulating glucocorticoid feedback and/or CRH receptor downregulation. Therefore, the identification of relative hypocortisolism usually requires the consideration of a pattern of neuroendocrine results across various conditions. The development of economic, reliable, and valid methods for detecting hypocortisolism as well as normative data is an important task for future research.

In sum, hypocortisolism may reflect many facets of neuroendocrine dysregulation, which may differ within and across populations, ultimately causing a

lack of cortisol effects in the organism. For the purposes of the following literature overview, hypocortisolism refers to a deficiency of cortisol, including (1) reduced basal cortisol secretion, (2) reduced adrenocortical sensitivity or capacity to challenge, or (3) enhanced negative feedback inhibition of the HPA axis. Reduced glucocorticoid signaling may also occur due to increased clearance or binding of free cortisol, as well as a reduced glucocorticoid sensitivity of target cells as a consequence of receptor abnormalities.

Clinical Findings of Hypocortisolism

Hypocortisolism has mainly been described in patients suffering from PTSD, chronic fatigue syndrome, fibromyalgia syndrome, and other functional somatic syndromes. Hypocortisolism also occurs in subsyndromal states of exhaustion and chronic pain.

Posttraumatic Stress Disorder

Because the development of PTSD by definition is related to extreme stress experiences, changes in neurobiological systems that participate in the mediation of the stress responses have been extensively studied in these patients. Paradoxically, initial studies demonstrated that basal cortisol concentrations, measured in urine or blood, were markedly reduced in combat veterans with PTSD compared to healthy controls. This finding has been replicated for Holocaust survivors with PTSD as well as for patients with abuse-related PTSD. Low basal cortisol secretion has been hypothesized to reflect hyperregulation of the HPA axis in the context of increased negative feedback sensitivity, as uncovered using low-dose dexamethasone suppression and metyrapone testing, although mild adrenal insufficiency might also play a role. In addition, findings of increased glucocorticoid receptor binding and function support the assumption of the increased negative feedback sensitivity of the HPA axis in PTSD. In the central nervous system (CNS), marked and sustained elevations of cerebrospinal fluid (CSF) concentrations of CRH have been reported in PTSD patients. Significantly, a circuit connecting the amygdala and the hypothalamus with the locus coeruleus, in which CRH and norepinephrine (NE) are the major neurotransmitters, participates in the regulation of vigilance, anxiety, and fear and in the integration of endocrine and autonomic responses to stress. Patients with PTSD exhibit markedly elevated CSF NE concentrations. Cortisol is known to inhibit the NE–CRH feedforward loop. Therefore, we might speculate that reduced cortisol availability leads to sensitization of the central stress responses,

with increased negative feedback acting at the pituitary level. Enhanced CRH and NE activation probably contribute to core symptoms of PTSD, such as sleep disruption, increased autonomic arousal, imprinted emotional memories, and enhanced startle reactivity. It must be noted that the finding of hypocortisolism has not been replicated in all PTSD studies, with both normal and elevated cortisol levels reported. Differences in type and timing of the trauma, symptom patterns, and comorbidity, among other factors, might contribute to this inconsistency. However, even normal cortisol output reported in some studies may be inappropriately low in the face of markedly elevated CRH-NE concentrations.

Chronic Fatigue Syndrome and Burnout

A large number of studies have investigated HPA axis alterations in chronic fatigue syndrome (CFS). Several independent studies reported decreased 24-h urinary free-cortisol excretion in patients with CFS. Although some studies failed to detect decreased cortisol secretion in CFS patients, a recent meta-analysis confirmed a moderate overall effect size for reduced 24-h urinary cortisol excretion in CFS. Studies of serial basal cortisol measures in saliva or blood over time have provided inconsistent findings. However, several chronobiological studies demonstrated lower amplitudes of cortisol in CFS patients when controlling for diurnal rhythm changes. CFS patients also exhibit decreased salivary cortisol awakening responses. In addition, decreased adrenal gland volume was reported for hypocortisolemic patients with CFS. In order to identify the origins of decreased adrenocortical activity in CFS, several studies have evaluated alterations in the sensitivity of the adrenal cortex to low doses of ACTH. Results are conflicting, with increased, normal, and decreased sensitivity reported. Inconsistent findings have also been obtained using pituitary or centrally acting challenge tests, such as the CRH challenge test and the insulin tolerance test. The only study that evaluated integrated HPA axis response to a psychosocial stressor found lower ACTH responses in CFS patients. In addition, there is one study suggesting the increased suppression of the HPA axis by a low dose of dexamethasone, although the effect was not confirmed at the pituitary level. Glucocorticoid receptor studies in CFS patients suggest decreased functionality. Taken together, although there is relatively solid evidence for relative hypocortisolism in CFS, results regarding the contributing mechanisms are inconclusive. Based on the concatenation of available findings, it has been suggested that hypocortisolism in CFS originates from the CNS rather than reflecting a primary adrenal dysfunction.

Regardless of its origin, hypocortisolism in CFS has been proposed to be directly linked to aberrations in the control of immune regulators, possibly leading to the characteristic symptom complex of fatigue, pain, cognitive disturbance, and sleep disruption.

Findings of hypocortisolism were also reported in nurses with burnout syndrome. In addition, a study of teachers with burnout syndrome, low morning cortisol, and hypersuppression of cortisol in response to a low-dose dexamethasone suppression test was reported. Such relative hypocortisolism was associated with physical complaints.

Fibromyalgia and Other Chronic Pain Syndromes

A series of studies have provided evidence for hypocortisolism in fibromyalgia syndrome (FMS). Three independent studies report reduced 24-h urinary cortisol excretion in patients with FMS relative to controls. Neuroendocrine challenge studies provided evidence for a primary adrenocortical impairment in FMS, although findings are conflicting. Patients with FMS also demonstrate reduced adrenocortical reactivity in the CRH stimulation test, with ACTH responses being either normal or increased. Several investigators have applied a standard dexamethasone suppression test to patients with FMS. Although two studies reported increased rates of nonsuppressors among FMS patients, several other studies revealed strikingly low rates of nonsuppressors in this patient population, rates that are even lower than nonsuppressor rates in healthy individuals. Nonsuppression seems to occur only in patients with comorbid depression. Based on these findings, we may expect that FMS is related to increased negative feedback sensitivity, similar to PTSD, although the hypothesis has not yet been directly tested. Despite the pronounced suppression of cortisol by dexamethasone, reduced glucocorticoid receptor affinity has recently been measured in the lymphocytes of patients with FMS, suggesting reduced regulatory effects of already low cortisol on target cells in the immune system, which might contribute to the increased perception of pain.

Findings of hypocortisolism have also been reported in women with chronic pelvic pain (CPP). In a series of studies of HPA axis function, women with CPP with no identified organic correlate demonstrated normal to low diurnal salivary cortisol levels. In response to CRH stimulation, normal plasma ACTH but reduced salivary cortisol concentrations were observed. These patients also exhibited an enhanced suppression of salivary cortisol to a low dose of dexamethasone, suggesting enhanced negative feedback sensitivity of the HPA axis. These same women had an increased rate of sexual and physical abuse experiences in addition to

PTSD and high levels of chronic stress. Interestingly, similar findings for women with CPP and verified pelvic adhesions were obtained. Evidence for relative hypocortisolism was also reported in other groups with chronic pain syndromes, such as patients with chronic headaches and children with recurrent abdominal pain. In addition, patients who had chronic sciatic pain 8 weeks after disk surgery exhibited decreased cortisol secretion after awakening, which was negatively correlated with pro-inflammatory cytokine activity. Significantly, a recent prospective study suggests that preoperative hypocortisolism predicts the development of persistent pain after disk surgery, as well as decreased pain threshold, increased NE activity, and pro-inflammatory cytokine production, in failed back syndrome (FBS) patients compared to patients who did not develop persistent pain and to healthy controls.

Other Somatic Disorders Related to Hypocortisolism

Relative hypocortisolism has also been reported for several other somatic disorders. Children and adults with asthma or atopic dermatitis exhibit decreased cortisol responses to psychosocial stress. Patients with rheumatoid arthritis have decreased basal cortisol levels, concordant with the assumption of hyperimmune activation in hypocortisolism. Basal hypocortisolism has been reported for atypical depression, which is characterized by lack of energy and predominantly somatic symptoms, closely resembling CFS. Finally, there is some evidence for low basal cortisol levels with blunted ACTH and cortisol responses to CRH challenge in patients with functional gastrointestinal disorders, such as irritable bowel syndrome (IBS). In IBS, the lack of cortisol effects might lead to the increased activation of CRH-NE systems, which in turn exert direct effects on gastrointestinal function, potentially leading to bowel symptoms.

Is There a Spectrum of Hypocortisolism-Related Disorders?

Hypocortisolism has been identified in diverse stress-related bodily disorders, although the potential mechanisms of hypocortisolism and its causal relationship in the pathophysiology of these disorders have not been sufficiently studied. Nevertheless, given the accumulating evidence for hypocortisolism in many patients suffering from diverse stress-related disorders, we may speculate that these disorders represent a spectrum of related disorders with common pathophysiological pathways. Interestingly, there is symptom overlap and high rates of comorbidity

among all the disorders related to hypocortisolism. There are also reports that cases with PTSD are at elevated risk for developing fatigue and pain over time. In addition, all these disorders share common risk factors, such as certain personality styles, chronic stress, and early adverse experience. Interestingly, the symptoms of all the disorders discussed so far are frequently triggered by acute or chronic stresses. Moreover, childhood abuse has been identified as a risk factor for the development of PTSD (after additional trauma in adulthood) as well as for CFS, FMS, CPP and functional gastrointestinal disorders. Although chronic stress and developmental stress have marked influences on the HPA axis, there are virtually no integrative studies assessing stress history and psychopathology together with HPA axis function in patients with bodily disorders. It is likely that stress experiences contribute to hypocortisolism in bodily disorders and might be linked to comorbid emotional problems. We next discuss whether there is preclinical evidence for an association between stress and hypocortisolism.

Stress and Hypocortisolism

Hypocortisolism does not seem to be an exclusive correlate of stress-related pathology; it has also been reported for healthy subjects living under ongoing stress as well as for healthy people with early developmental stress. In addition, some animal models of chronic and early-life stress have reported findings of decreased adrenocortical activity, particularly in nonhuman primates.

Studies in Humans

A small number of studies in humans suggest reduced adrenocortical activity or reactivity in states of chronic stress; some of these studies date back to the 1960s and 1970s. With increasing evidence for hypocortisolism in stress-related disorders, these early findings have regained attention in stress research. For example, decreased urinary excretion of the cortisol metabolite 17-hydroxycorticosterone (17-OHCS) was measured over several months in the parents of fatally ill children. Many of these parents demonstrated decreased 17-OHCS excretion below their initial baseline, even in phases of acute medical complications in their children. Lower than normal 17-OHCS excretion was also observed in soldiers of a special team in Vietnam who had been warned to expect an enemy attack; interestingly, 17-OHCS levels dropped even further on the day when the attack was anticipated. In another study, the same authors measured 17-OHCS excretion in helicopter medics in Vietnam. These

medics demonstrated a stable pattern of decreased 17-OHCS excretion whether they were flying or not. On flying days, some medics even demonstrated lower 17-OHCS levels than on their days off. Similar findings were obtained in civilian paramedics who demonstrated lower cortisol levels on work days compared to their days off. More recently, decreased plasma cortisol concentrations were measured in Bosnian prisoners of war, in white-collar employees with high work loads, and in chronically stressed teachers. Finally, evidence for marked hypocortisolism under basal conditions has been reported in severely neglected children, such as Romanian orphans. Maltreated children exhibit flat cortisol cycles throughout the day, particularly on high-conflict days. Interestingly, greater numbers of work hours for employed mothers and greater numbers of children in the household are also associated with lower morning cortisol levels and a flat cortisol cycle across the day in children. Decreased basal cortisol secretion and mild adrenocortical insufficiency in response to maximal ACTH stimulation has also been reported in adult women with histories of childhood sexual or physical abuse, particularly in the absence of depression. There are also findings of hypocortisolism in healthy subjects with early loss experiences. Taken together, these findings suggest that hypocortisolism does occur in humans as a function of severe chronic or early-life stresses. The physiological meaning of hypocortisolism is unclear; it might reflect a physiological state of preparedness in threatening situations, a state of exhaustion after chronic stress, and/or a state of adaptive counterregulation to hyperactive central stress response systems.

Animal Studies

The increasing evidence for the phenomenon of hypocortisolism in clinical and preclinical human research is not paralleled by an equal focus on hypocortisolism in animal models. Indeed, there exist only very few published studies that found low glucocorticoid secretion in animal models of stress. The most conclusive reports to date are studies in nonhuman primates. For example, when tested as adults, nonhuman primates exposed to prolonged periods of maternal or social deprivation exhibit marked behavioral, physiological, and neurobiological changes, including fearfulness and anxiety, social dysfunction, anhedonia, and changes in immune and autonomic function. In terms of HPA axis function, basal and stress-induced cortisol secretion has been reported to be either increased or reduced. Squirrel monkeys exposed to repeated social separations demonstrate markedly reduced cortisol responses to later separation stress and

increased glucocorticoid feedback sensitivity. Similarly, decreased morning cortisol excretion was recently reported for marmoset monkeys exposed to repeated maternal separation. Another paradigm in nonhuman primates involves the exposure of bonnet macaque mothers with infants to unpredictable conditions with respect to food access for 3 months, resulting in a diminished perception of security and a reduction of maternal care for the infants. When tested as adults, bonnet macaques reared under these conditions demonstrated elevated CSF CRH concentrations, sensitization of the NE system, and behavioral sensitization to fear stimuli. However, their cortisol levels were markedly decreased. Thus, early-life stress appears to result in decreased adrenocortical activity in several nonhuman primate models, closely paralleling findings in PTSD patients.

Although there now are a number of reports of hypocortisolism in nonhuman primates, very few studies on hypocortisolism in rodent models of stress have been published. In one study, social isolation resulted in a reduction of plasma corticosterone levels, both under basal and stress-stimulated conditions. Prenatal exposure to dexamethasone also resulted in decreased HPA axis activity in rat offspring. In rats, a single session of prolonged severe foot shock stress induced enhanced negative feedback sensitivity compared to stress-naïve animals. However, when the rats were stressed again 2 weeks later, no differences in stress responses were observed. Interestingly, one study showed that chronically stressed rats that initially demonstrated lower corticosterone levels developed higher acoustic startle responses and gastrointestinal damage than rats with initially higher levels, suggesting a causal role of relative hypocortisolism for promoting high central stress/fear responses and physical symptoms.

It should be noted that there are numerous reports that adult rats that were subjected to early handling (3–15 min daily) exhibited lower HPA responses to stress and generally better adaptation and health than nonhandled rats. However, it was found that the handling procedure induces increased maternal care after the reunion, probably reflecting a nurturing early environment, whereas the nonhandled rats were more deprived and therefore had relatively increased stress reactivity. This hypothesis was confirmed when directly tested in rats with high and low maternal care.

Potential Mechanisms of Hypocortisolism

As already mentioned, relative hypocortisolism has been described in a multitude of stress-related

conditions that share common symptoms and frequently manifest in comorbidity. Different mechanisms may contribute to decreased cortisol output under various conditions, and different facets of hypocortisolism have been emphasized for different study groups. For example, increased negative feedback inhibition has been emphasized for PTSD, adrenocortical dysfunction has been suggested for FMS, and central origins of hypocortisolism have been suggested for CFS. Clearly, to date, all of the possible contributing mechanisms have not been sufficiently studied in all of the disorders that appear to be related to hypocortisolism. It might well be that there is heterogeneity within diagnostic groups that might reflect different pathways to illness. Future studies should systematically scrutinize each of these disorders in terms of multiple potential mechanisms. Potential mechanisms include (1) reduced biosynthesis or depletion at several levels of the HPA axis (CRH, ACTH, and cortisol); (2) CRH hypersecretion and adaptive downregulation of pituitary CRH receptors, resulting in lower than normal ACTH and cortisol secretion once a stressor stops and CRH secretion returns to normal levels; (3) increased feedback sensitivity of the HPA axis; and (4) morphological changes (i.e., decreased adrenal gland volume). Lack of cortisol effects may also result from relative glucocorticoid resistance of target cells due to receptor changes; decreased availability of free, biologically active cortisol due to alterations in binding proteins; changes in enzyme activity (i.e., 11- β -hydroxysteroid dehydrogenase) that metabolizes glucocorticoid hormones on entry into the cell; and changes in the multidrug resistance pump, which extrudes cortisol but not corticosterone from the target cell.

It should be noted that superimposed factors, such as the nature or timing of the stressor, coping styles, personality, sex, and genetic variations probably contribute to hypocortisolism and might influence the specific mechanisms of its manifestation in relation to stress. For example, the timing of stress during development influences the direction of HPA axis changes in nonhuman primates. Passive-avoidant coping and alexithymic personality appear to be linked to hypocortisolism. Women generally have lower free-cortisol responses to psychosocial stress than men. Basal cortisol levels and HPA axis responses to stress have been shown to be influenced by genetic factors in twin studies. Finally, polymorphisms in genes involved in the stress response, such as the glucocorticoid receptor (GR) gene, have been shown to markedly moderate neuroendocrine responses to stress. In addition to these factors, it has been suggested that hypocortisolism might reflect

different symptom constellations or phases during the course of illness. Taken together, integrative longitudinal research approaches are needed to address the mechanisms of hypocortisolism, together with contributing factors and illness characteristics, in the future.

Effects of Hypocortisolism on Disease Vulnerability

There are several avenues through which the lack of cortisol might lead to characteristic symptoms of pain, fatigue, and anxiety/stress sensitivity. First, glucocorticoids secreted during stress induce gluconeogenesis, mobilize free fatty acids, and reduce the use of amino acids for protein synthesis, thereby increasing the organism's energy supplies. Altered metabolism during stress due to hypocortisolism might lead to exhaustion and fatigue.

Second, glucocorticoids exert inhibitory effects on the secretion of pro-inflammatory cytokines, such as interleukin (IL)-6 and tumor necrosis factor (TNF)- α , during stress and help return these cytokines to baseline levels after stress. These cytokines regulate cellular immunity and mediate pain. Increased or prolonged cytokine activity in the brain leads to impaired cognitive function and mood and anxiety symptoms through the stimulation of CRH-NE systems. Thus, the lack of glucocorticoid effects probably promotes the disinhibition of pro-inflammatory cytokines during stress, resulting in increased vulnerability to developing chronic fatigue, chronic pain, allergies, asthma, and emotional disorders. Interestingly, elevated pro-inflammatory cytokine levels have been reported for PTSD and several functional somatic syndromes, although the findings are inconsistent. Patients with adrenal insufficiency also have elevated levels of pro-inflammatory cytokines, such as IL-6, as well as symptoms of fatigue and pain.

Third, cortisol exerts direct effects on the brain and behavior. Thus, glucocorticoids exert negative feedback effects on the CRH-induced activation of the locus ceruleus NE cells. Relatively decreased cortisol output may have permissive effects toward an increased and sustained activation of the feed-forward cascade between the CRH and NE systems, perhaps in concert with elevated cytokine responses. CRH and NE are implicated in enhanced encoding of traumatic emotional memories, facilitation of fear conditioning, and increased vigilance and arousal, including increased startle responses, leading toward the three symptom clusters of PTSD. Enhanced CRH-NE activation might further stimulate the immune system, promoting fatigue and pain, and

might promote gastrointestinal disturbance. This model suggests that hypocortisolism indeed might occur as a preexisting (genetic or stress-related) risk factor that promotes the development of a spectrum of emotional and somatic disorders. In support of this hypothesis, increased autonomic activity and low cortisol secretion directly after the trauma are significant predictors of the development of PTSD. Low cortisol levels presurgery predict the development of ongoing pain and immune-NE hyperactivation after surgery in FBS patients.

Clinical Implications

If hypocortisolism indeed is causally involved in the development of certain stress-related disorders, reflecting a preexisting risk factor, this has seminal clinical implications. Once normative values are established, hypocortisolism could be measured in healthy subjects to predict the risk of developing PTSD or functional somatic disorders should severe or ongoing stress occur. Consequently, the substitution of hydrocortisone should have preventive effects. In fact, the administration of hydrocortisone during trauma has been shown to effectively prevent PTSD in humans in several studies. In addition, it was recently demonstrated that hydrocortisone treatment, simulating normal circadian cortisol rhythm, is effective in the treatment of PTSD. Randomized controlled clinical trials, furthermore, showed that some (but not all) patients with CFS benefit from hydrocortisone treatment.

In conclusion, given the high prevalence of disorders related to hypocortisolism, the development of studies aiming to describe the mechanisms of hypocortisolism across and within disorders, to scrutinize the effects of hypocortisolism in target cells of the immune system and CNS, and to reverse these effects surely is a promising and exciting area of future research.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Adrenal Cortex; Adrenal Insufficiency; Adrenocorticotrophic Hormone (ACTH); Allostasis and Allostatic Load; Animal Models (Nonprimate) for Human Stress; Arthritis; Asthma; Burnout; Child Abuse; Childhood Stress; Chronic Fatigue Syndrome; Corticosteroids and Stress; Corticotropin Releasing Factor (CRF); Cytokines; Dexamethasone Suppression Test (DST); Disease, Stress Induced; Endocrine Systems; Fatigue and Stress; Gastrointestinal Effects; Glucocorticoid Negative Feedback; Glucocorticoids, Effects of Stress on; Glucocorticoids,

Overview; Glucocorticoids, Role in Stress; Hypothalamic-Pituitary-Adrenal; Immune Response; Immune Suppression; Neuroendocrine Systems; Pain; Pharmacological Treatments of Stress; Posttraumatic Stress Disorder, Neurobiology of; Salivary Cortisol; Somatic Disorders; Stress Effects, Overview; Stress, Definitions and Concepts of.

Further Reading

- Adler, G. K. and Geenen, R. (2005). Hypothalamic-pituitary-adrenal and autonomic nervous system functioning in fibromyalgia. *Rheumatic Disease Clinics of North America* 31, 187–202.
- Clauw, D. J. and Chrousos, G. P. (1997). Chronic pain and fatigue syndromes: overlapping clinical and neuroendocrine features and potential pathogenic mechanisms. *Neuroimmunomodulation* 4(3), 134–153.
- Chrousos, G. P. and Gold, P. W. (1992). The concepts of stress and stress system disorders: overview of physical and behavioral homeostasis. *Journal of the American Medical Association* 267, 1244–1252.
- Cleare, A. J. (2004). The HPA axis and the genesis of chronic fatigue syndrome. *Trends in Endocrinology and Metabolism* 15(2), 55–59.
- Fries, E., Hesse, J., Hellhammer, J., et al. (2005). A new view on hypocortisolism. *Psychoneuroendocrinology* 30(10), 1010–1016.
- Geiss, A., Rohleder, N., Kirschbaum, C., et al. (2005). Predicting the failure of disc surgery by a hypofunctional HPA axis: evidence from a prospective study on patients undergoing disc surgery. *Pain* 114, 104–117.
- Gunnar, M. R. and Vazquez, D. M. (2001). Low cortisol and a flattening of expected daytime rhythm: potential indices of risk in human development. *Developmental Psychopathology* 13(3), 515–538.
- Heim, C., Ehler, U. and Hellhammer, D. H. (2000). The potential role of hypocortisolism in the pathophysiology of stress-related bodily disorders. *Psychoneuroendocrinology* 25(1), 1–35.
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* 338(3), 171–179.
- Raison, C. L. and Miller, A. H. (2003). When not enough is too much: the role of insufficient glucocorticoid signaling in the pathophysiology of stress-related disorders. *American Journal of Psychiatry* 160, 1554–1565.
- Sapolsky, R. M., Romero, L. M. and Munck, A. U. (2000). How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrine Reviews* 21(1), 55–89.
- Sternberg, E. M. (2001). Neuroendocrine regulation of autoimmune/inflammatory disease. *Journal of Endocrinology* 169(3), 429–435.
- Yehuda, R. (2002). Current status of cortisol findings in post-traumatic stress disorder. *Psychiatric Clinics of North America* 25(2), vii, 341–368.

Hypoglycemia

B M Frier

Royal Infirmary of Edinburgh, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by B M Frier, volume 2, pp 453–459, © 2000, Elsevier Inc.

Physiological Responses to Hypoglycemia
Effects on Cerebral Function
Causes and Investigations

Glossary

<i>Autonomic reaction</i>	Physiological manifestations of acute activation of the sympathetic (sympathoadrenal) division of the autonomic nervous system, through the autonomic neural stimulation of end organs and secretion of catecholamines.
<i>Awareness of hypoglycemia</i>	The subjective perception of the onset of symptoms generated by a decline in blood glucose.
<i>Cognitive dysfunction</i>	The impairment of intellectual and mental activities as a consequence of interference with cerebral function.
<i>Counter-regulation</i>	The process that promotes the secretion of hormones that oppose the actions of insulin, provoking metabolic changes that elevate blood glucose.
<i>Gluconeogenesis</i>	The production of new glucose from three carbon precursors, principally in the liver and kidney.
<i>Glucose homeostasis</i>	The metabolic maintenance of blood glucose within a narrow range by regulatory mechanisms. The main factors controlling this are the intake of nutrients, the rates of gastric emptying and intestinal absorption, the hepatic production of glucose, and the secretion of insulin and counter-regulatory hormones.
<i>Glycemic threshold</i>	The blood glucose concentration at which a specific event is triggered during the development of hypoglycemia (e.g., onset of symptoms).
<i>Glycogenolysis</i>	The breakdown of hepatic glycogen into glucose, which is released into the circulation.
<i>Neuroglycopenia</i>	The direct effect of glucose deprivation on neurons in the central nervous system.

Hypoglycemia (low blood glucose) is a biochemical abnormality that can be arbitrarily defined by a low blood glucose concentration. In a person with diabetes,

a blood glucose below 3.5 mmol l^{-1} (61 mg dl^{-1}) is considered to represent hypoglycemia; in a nondiabetic person, a blood glucose below 2.2 mmol l^{-1} (38 mg dl^{-1}) is almost always pathological. Hypoglycemia *per se* is not a clinical diagnosis, and a cause for the low blood glucose should be sought.

Physiological Responses to Hypoglycemia

The human brain requires a continuous supply of glucose as its principal source of energy and malfunctions rapidly if deprived of this substrate. To protect the brain and prevent a dangerous fall in blood glucose, multiple counterregulatory mechanisms have evolved to maintain glucose homeostasis. These include the secretion of counterregulatory hormones including glucagon, epinephrine (adrenaline), cortisol, and growth hormone; the centrally activated release of various neurotransmitters; and simultaneous inhibition of endogenous insulin secretion. Hepatic autoregulation can release glucose during profound hypoglycemia. The secretion of pancreatic glucagon occurs independently of the activation of counterregulatory mechanisms within the brain. In contrast, epinephrine is secreted secondary to the stimulation of the sympathetic division of the autonomic nervous system following the activation of autonomic centers within the hypothalamus. The other counterregulatory hormones are released through the activation of the hypothalamic-pituitary-adrenal axis. These hormones promote the breakdown of glycogen to glucose within the liver (hepatic glycogenolysis) and the production of three carbon precursors from proteolysis and glycogenolysis in the skeletal muscle that are used for glucose formation (gluconeogenesis) within the liver and kidney. The oxidation of free fatty acids, produced by lipolysis, provides energy for gluconeogenesis.

Hypoglycemia generates symptoms (1) as a direct consequence of glucose deprivation of neurons in the brain (neuroglycopenia) and (2) secondary to acute autonomic stimulation. Symptoms are idiosyncratic and age-specific, varying considerably among individuals, both in nature and intensity. Common autonomic symptoms are sweating, trembling, pounding heart, hunger, and anxiety; neuroglycopenic symptoms include inability to concentrate, drowsiness, confusion, difficulty with speech, and incoordination. Common nonspecific symptoms include nausea, tiredness, and headache. The recognition of the early onset of the symptoms of hypoglycemia is essential for people

with diabetes who are treated with insulin so that a correction of the low blood glucose can be made.

Symptomatic hypoglycemia is usually associated with the activation of the autonomic nervous system and a rise in plasma catecholamines of similar magnitude to that observed during major vascular events such as acute myocardial infarction and subarachnoid hemorrhage. Hypoglycemia can therefore represent a severe form of stress, which provokes major pathophysiological consequences. Catecholamines provoke hemodynamic changes through their direct effects on the cardiovascular system, causing changes in regional blood flow, and the effects on various

organs induce physiological changes that are identified as autonomic symptoms (Figure 1). This is sometimes described as an autonomic reaction. A transient rise in heart rate is accompanied by an elevation in systolic blood pressure and a decline in diastolic blood pressure, with increased myocardial contractility and cardiac output. Blood flow is increased to the brain, heart, liver, gut, retina, and skeletal muscles, but is decreased to organs such as the kidney and spleen. Hypoglycemia also provokes significant hemostatic and hemorrheological changes (mainly through hormonal action), which can affect the microcirculation.

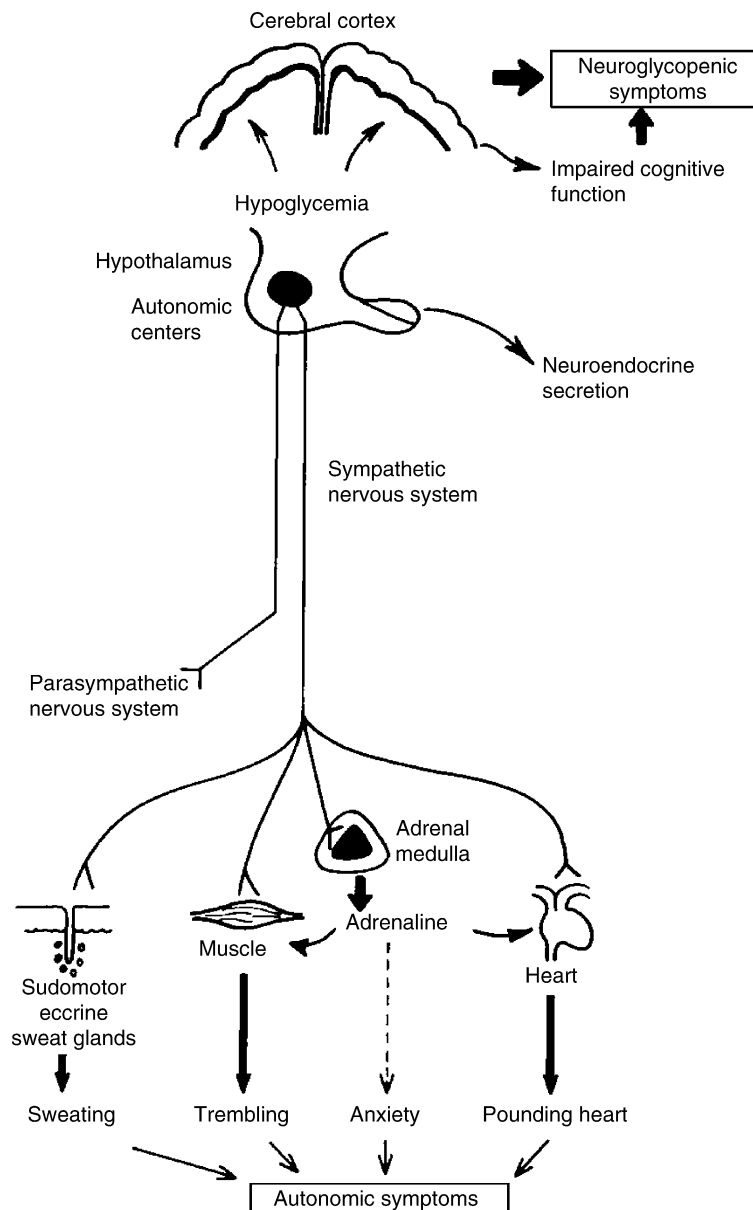


Figure 1 Autonomic neural pathways involved in the generation of the peripheral physiological manifestations and symptoms during insulin-induced hypoglycemia.

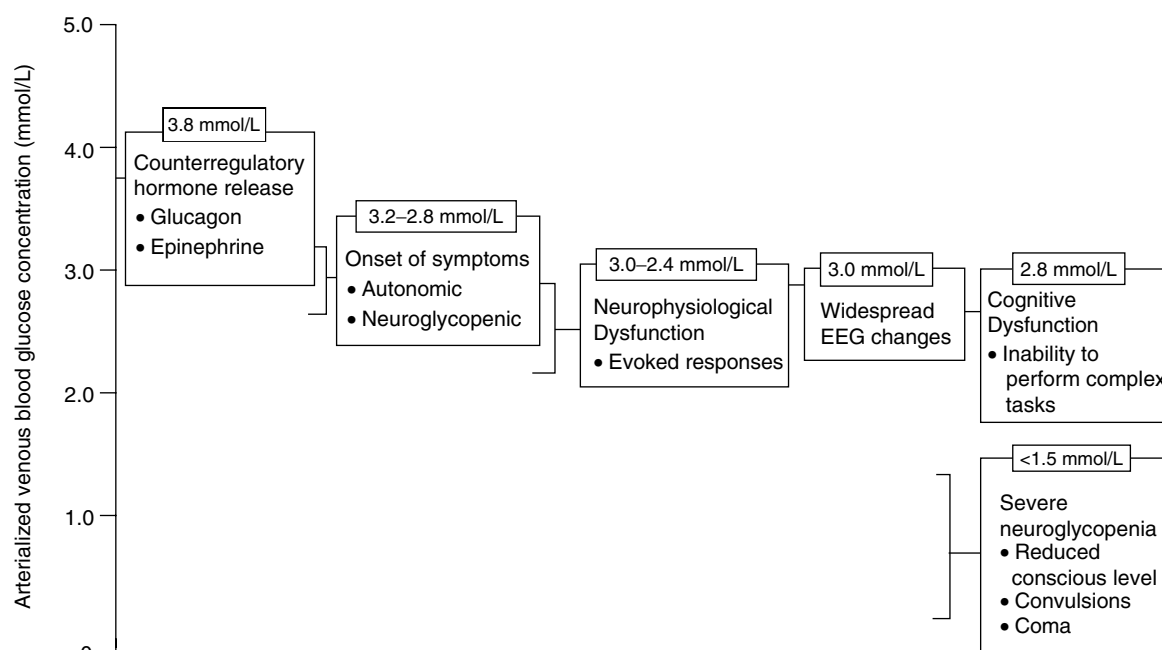


Figure 2 Glycemic thresholds for the secretion of counterregulatory hormones and the onset of physiological, symptomatic, and cognitive changes in response to acute hypoglycemia in nondiabetic adults. Glycemic thresholds are shown in relation to glucose values in arterialized venous blood. Reproduced from Frier, B. M. and Fisher, B. M. (eds.) (1999), *Hypoglycaemia and clinical diabetes*, Chichester, UK: John Wiley. Copyright 1999 John Wiley & Sons Limited. Reproduced with permission.

Effects on Cerebral Function

A progressive decline in blood glucose triggers a hierarchy of events that occur at individual glycemic thresholds (Figure 2). Counterregulation is stimulated when arterial blood glucose falls below 3.8 mmol l^{-1} , and impairment of cognitive function commences below 3.0 mmol l^{-1} . Early changes are subtle, and progressively greater cognitive impairment occurs around 2.6 mmol l^{-1} , with deteriorating performance evident in tests of several cognitive domains. Some of the neuroglycopenic symptoms of hypoglycemia are subjective manifestations of impaired cognitive function. Hypoglycemia also induces noncognitive changes in mood and behavior, including feelings of tense tiredness, unhappiness, fear, despondency, and, in some individuals, anger. Autonomic symptoms commence around 3.0 mmol l^{-1} , and the effects on neurophysiological function (sensory-evoked potentials and electroencephalographic abnormalities) become more prominent as blood glucose falls further. If the condition is untreated, confusion and drowsiness will progress to loss of consciousness and coma. The comatose state may be complicated by convulsions. Significant brain damage is rare and occurs only if the neuroglycopenia is protracted, leading to brain death. Thus, hypoglycemia can vary significantly in severity, ranging from asymptomatic biochemical hypoglycemia to profound coma. In people with insulin-treated

diabetes, severe hypoglycemia is defined by the need of the individual for external assistance for recovery, in which situation acute neuroglycopenia is interfering with the individual's capability for self-treatment. Severity is not determined by the intensity of the symptomatic response or the development of coma.

Glycemic thresholds are not fixed, can vary among individuals, and, in people who have diabetes, can be influenced by recurrent exposure to hypoglycemia. In some people with insulin-treated diabetes, their symptom profile alters with time, the perception of the onset of symptoms becomes diminished, and they may develop a degree of impaired awareness of hypoglycemia. This is a major risk factor for severe hypoglycemia. With the increasing duration of insulin-treated diabetes, deficiencies of counterregulatory hormonal secretion often develop as an acquired defect, and this also increases the risk of severe hypoglycemia. This problem co-segregates with impaired awareness of hypoglycemia. In affected individuals, the glycemic thresholds for the onset of autonomic symptoms and for counterregulatory hormonal secretion are reset at a lower blood glucose than usual, allowing neuroglycopenia to supervene and interfere with the patient's ability to self-treat hypoglycemia. These acquired abnormalities are associated with strict glycemic control (particularly when glycated hemoglobin is within the nondiabetic range) and with repeated

exposure to episodic (antecedent) hypoglycemia, and they occur in other conditions that cause chronic or recurrent hypoglycemia, such as insulin-secreting tumors (e.g., insulinoma). These changes represent an adaptation by the brain to allow it to function during glucose deprivation.

Although total cerebral blood flow is increased during hypoglycemia, the regional changes in blood flow within the brain vary considerably, being increased to frontal areas (including the cerebral cortex) and reduced in posterior regions of the brain. Localized changes in cerebral blood flow may provoke some of the neurological manifestations of hypoglycemia such as transient hemiplegia. However, the rate-limiting factor in protecting the brain against neuroglycopenia is the transport of glucose across the blood–brain barrier. Glucose transporters can be upregulated in response to chronic hypoglycemia, but this process is relatively slow and it takes time (hours to days) for the human brain to adapt and be able to function at low concentrations of blood glucose. Such a modification is essentially maladaptive because the warning symptoms of impending neuroglycopenia are obtunded, reducing the time available to detect the falling blood glucose and increasing the risk of progressing to severe hypoglycemia.

Hypoglycemic coma is associated with the release of excitatory amino acids, including glutamate and aspartate, which activate subtypes of excitatory amino acid receptors including N-methyl-D-aspartate receptors. Membrane depolarization occurs and leads to neuronal necrosis. The brain is affected in a rostro-caudal direction, with the cerebral cortex and hippocampus being most sensitive to neuroglycopenia; the brain stem and spinal cord are much more resistant. In fatal cases of hypoglycemic coma, the neuropathology of the brain is variable, but the laminar necrosis of neurons occurs in the third to fifth layers of the cerebral cortex. Survivors of severe protracted hypoglycemic coma develop cortical and hippocampal atrophy, with ventricular enlargement, often associated with a chronic vegetative state. The cumulative effects of recurrent severe hypoglycemia may cause intellectual impairment in the developing brain of infants and young children, but in adults (principally those with insulin-treated diabetes) the effect on cognitive impairment appears to be modest, with occasional anecdotal exceptions. Progression to severe intellectual impairment and dementia is rare, although a few cases have been described, particularly in patients exposed to chronic neuroglycopenia of long duration, for example in association with an undiagnosed insulin-secreting tumor. Although repeated exposure of people with insulin-treated diabetes to severe hypoglycemia may induce modest cognitive

impairment, this may primarily be caused by other co-existing morbidities, such as hypertension, chronic hyperglycemia (poor glycemic control), and cerebrovascular disease, and underlies the concept of a diabetic encephalopathy.

Causes and Investigations

Multiple causes of hypoglycemia are recognized in humans, but most are rare. In the nondiabetic population, hypoglycemia is most commonly associated with excessive alcohol consumption. However, by far the most common cause is iatrogenic, resulting from the inadvertent use of insulin or drugs that promote insulin secretion (mostly sulfonylureas), which are used for the treatment of diabetes. Occasionally, hypoglycemia is the presenting feature of an illness or occurs as an important epiphenomenon, and it is then considered to be spontaneous. The classification of hypoglycemia can be based on etiology (see [Table 1](#)) or, alternatively, on whether insulin is detectable. Subdivision into “fasting” and “reactive” types of hypoglycemia is no longer used. The spectrum of causes of hypoglycemia varies between hospital and community practice. In hospitalized patients, hypoglycemia is associated with diabetic therapies, chronic renal failure, infections, liver disease, shock, pregnancy, neoplasia, and burns. In the community, excessive alcohol consumption and therapies for treating diabetes predominate as the main causes.

Hypoglycemia is usually episodic, but the symptoms that are commonly experienced with acute insulin-induced hypoglycemia may be absent in people who present with spontaneous hypoglycemia. A low blood glucose level should be suspected in people who present as an acute medical emergency with collapse, altered consciousness, convulsions, abnormal neurological findings, or confusion. A suspicion of underlying hypoglycemia should be considered even when an obvious cause of cerebral abnormality is apparent, such as hemiplegic stroke, alcoholic intoxication, or cerebral malaria. An accurate and reliable history may be difficult to obtain from the patient, and an objective account from witnesses may be valuable. Where subacute neuroglycopenia presents with deteriorating cognitive function over months or years, psychiatric features may be prominent, and such patients may have been referred to a neurologist or psychiatrist for investigation of funny turns, feeling faint, odd behavior, paresthesiae, or convulsions. In some cases, the only indication that hypoglycemia is responsible is the finding of a low blood glucose level on routine biochemical screening.

Tumor-induced hypoglycemia may occur as a pre-terminal feature of an advanced malignancy,

Table 1 Causes of hypoglycemia in adults

Type	Causes
Toxic hypoglycemia	Therapeutic hypoglycemic agents: insulin, sulfonylureas Alcohol Drugs, e.g. quinine, salicylates, propranolol, pentamidine Poisons, e.g. certain types of mushroom
Pancreatic	Insulinoma: benign, malignant, multiple and microadenomatosis Islet hyperplasia with functional hyperinsulinism
Extrapancreatic neoplasms	Mesenchymal tumors Hemangiopericytoma Primary hepatic carcinoma Various other tumors
Liver and kidney disease	Hepatocellular disease; cirrhosis Advanced renal failure Congestive cardiac failure
Endocrine disease	Pluriglandular syndrome Pituitary insufficiency Adrenocortical insufficiency Selective hypothalamic insufficiency
Autoimmune hypoglycemia	Autoimmune insulin syndrome Anti-insulin receptor antibodies
Inborn errors of metabolism	Adult glycogenolysis (liver glycogen disease) Hereditary fructose intolerance
Miscellaneous	Sepsis; burns Prolonged starvation; anorexia nervosa Dialysis Excessive, strenuous exercise Diseases of the nervous system Falciparum malaria

particularly in patients with extensive hepatic metastases and also with some very large tumors such as fibrosarcoma. Some tumors secrete substances with insulin-like activity; malignancies involving the liver disrupt hepatic gluconeogenesis and the maintenance of glucose homeostasis. Insulinomas are pancreatic islet cell tumors that secrete insulin and usually cause recurrent fasting hypoglycemia. They are rare (one to two cases per million population per year) but are one of the commonest causes of spontaneous hypoglycemia in nondiabetic individuals. Symptoms may occur before meals or during exercise and are relieved by the consumption of food or a glucose drink. The tumors are usually small (0.5–5 mm in diameter) and may occur in any part of the pancreas. Approximately 10% are malignant.

A firm diagnosis of hypoglycemia is usually based on Whipple's triad:

- Symptoms are associated with fasting or exercise.
- Symptoms are relieved by glucose.
- Low blood glucose is demonstrated by biochemical analysis.

In the investigation of suspected hypoglycemia, low blood glucose should be confirmed by laboratory analysis of venous or capillary blood glucose. Hypoglycemia is often associated with inappropriately elevated plasma insulin concentrations. Blood should be collected during the emergency presentation, and the plasma frozen and stored, both for immediate confirmation of hypoglycemia and for additional analyses (e.g., for hormones, alcohol, and drugs), which may identify a cause. This can avoid subsequent unnecessary investigations and is of forensic and medical-legal importance in cases of attempted suicide or deliberate poisoning.

The nature and timing of blood samples for the measurement of glucose are important in the evaluation of the causes of hypoglycemia. The possible development of a large arteriovenous difference after glucose absorption precludes the use of glucose tolerance tests that require an oral glucose load because the arterial plasma glucose concentration is 15% higher than that estimated from whole venous blood. Provocative tests for investigating suspected hypoglycemia include prolonged fasting with or without exercise and the infusion of intravenous insulin to induce hypoglycemia to see whether C-peptide (an index of endogenous insulin secretion) is suppressed. Plasma concentrations of insulin, C-peptide, pro-insulin, and other hormones such as insulin-like growth factor (IGF)-1 may also be measured during confirmed hypoglycemia. Characteristic patterns of plasma glucose and insulin are associated with different causes of hypoglycemia, and a diagnosis of endogenous hyperinsulinism depends on evidence of inappropriate insulin and/or C-peptide secretion in the presence of hypoglycemia. With an insulinoma, a mesenchymal tumor, and a sulfonylurea overdose, the fasting plasma insulin is usually inappropriately high for the prevailing blood glucose and endogenous insulin secretion is not suppressed. This can be demonstrated by elevated plasma C-peptide during insulin-induced hypoglycemia. Factitious hypoglycemia is characteristically associated with elevated plasma insulin and absent plasma C-peptide in nondiabetic individuals, indicating that an overdose of exogenous insulin has been administered.

True reactive hypoglycemia may occur in patients who have had previous gastric surgery or who have gastrointestinal motility disorders. However, some individuals describe symptoms suggestive of, but without biochemical evidence of, hypoglycemia that occur 1–2 h after food. This is called the postprandial syndrome, and such patients should be evaluated with ambulatory blood glucose measurements or by serial measurements of plasma glucose and insulin

after a standard mixed meal, plus recording of symptoms and dietetic assessment.

The emergency treatment of acute hypoglycemia should not be withheld to await a confirmatory biochemical diagnosis. In diabetic patients, self-treatment of mild hypoglycemia requires the ingestion of a glucose drink or food. More severe hypoglycemia can be treated with intravenous glucose or intramuscular glucagon. In noniatrogenic causes of hypoglycemia, treatment depends on the specific diagnosis. Tumor-induced hypoglycemia may require prolonged therapy if surgical resection is not possible or the tumor cannot be located. Treatment with diet (regular intake of oral carbohydrate) and drugs that inhibit insulin secretion (e.g., diazoxide), thiazide diuretics, and somatostatin analogs (e.g., octreotide) may be beneficial. Patients with true reactive hypoglycemia usually respond to dietary measures such as the exclusion of highly refined sugars and the consumption of frequent regular meals that are high in fiber content; the α -glucosidase inhibitor acarbose may be beneficial.

Further Reading

- Cryer, P. E. (2002). Hypoglycaemia: the limiting factor in the glycaemic management of type I and type II diabetes. *Diabetologia* 45, 937–948.
- Cryer, P. E. and Frier, B. M. (2004). Hypoglycaemia. In: De Fronzo, R. A., Ferrannini, E., Keen, H. & Zimmet, P. (eds.) *International textbook of diabetes mellitus* (3rd edn., pp. 1079–1100). Chichester, UK: John Wiley.
- Frier, B. M. and Fisher, B. M. (eds.) (1993). *Hypoglycaemia and diabetes: clinical and physiological aspects*. London: Edward Arnold.
- Frier, B. M. and Fisher, B. M. (eds.) (1999). *Hypoglycaemia and clinical diabetes*. Chichester, UK: John Wiley.
- Marks, V. (2001). Spontaneous hypoglycaemia. *CPD Clinical Biochemistry* 3, 67–71.
- Marks, V. and Teale, J. D. (1998). Tumours producing hypoglycaemia. *Endocrine-Related Cancer* 5, 111–129.
- Service, F. J. (1995). Hypoglycemic disorders. *New England Journal of Medicine* 332, 1144–1152.
- Service, F. J. (1999). Hypoglycemic disorders. *Endocrinology and Metabolism Clinics of North America* 28, 467–661.

Hypotension, Hypovolemia, and Septic Shock

A Beishuizen and A B Johan Groeneveld

VU University Medical Center, Amsterdam,
The Netherlands

I Vermes

Medical Spectrum Twente, Enschede, The Netherlands

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A Beishuizen, I Vermes and C Haanen, volume 2, pp 460–467, © 2000, Elsevier Inc.

Introduction

Neuroendocrine Regulation in Hypotension and Hypovolemia
Septic Shock

Glossary

<i>Aldosterone</i>	A mineralocorticoid hormone produced by the zona glomerulosa of the adrenal cortex and a key regulator of potassium balance where it induces sodium and water retention.
<i>Angiotensin II</i>	An octapeptide which acts to increase peripheral arterial resistance, raising

Atrial natriuretic peptide (ANP)

Circulatory shock

Cytokines

Hypotension

Hypovolemia

Multiple organ dysfunction syndrome

blood pressure, and stimulates aldosterone biosynthesis.

A polypeptide hormone mainly synthesized in the cardiac atrium, released following atrial dilatation or increased intravascular volume, and involved in natriuresis and the regulation of blood pressure and volume homeostasis.

A generalized inadequacy of blood flow throughout the body, to the extent that the tissues are damaged because of too little flow, especially too little delivery of oxygen to tissue cells.

Polypeptide cell regulators, mainly produced by cells of the immune system, which all affect in some way the function of the immune system and have the ability to act as communication signals.

A systolic blood pressure of <90 mmHg or a reduction of ≥ 40 mmHg from baseline.

An abnormal decrease in blood (plasma) volume.

The presence of altered organ dysfunction in an acutely ill patient such that homeostasis cannot be maintained without intervention.

<i>Renin</i>	A glycoprotein hormone and proteolytic enzyme produced by the juxtaglomerular cells of the kidney in response to hypovolemia and hypotension and it cleaves its substrate angiotensinogen into angiotensin I.
<i>Sepsis</i>	The systemic response to infection.
<i>Septic shock</i>	Sepsis-induced hypotension, despite adequate fluid resuscitation along with the presence of perfusion abnormalities which may include, but are not limited to, lactic acidosis, oliguria, or an acute alteration in mental status.
<i>Sympathoadrenal system</i>	An anatomical and physiological unit of the adrenal medulla and the sympathetic nervous system.
<i>Vasopressin</i>	A polypeptide hormone produced in the hypothalamus and stored in the posterior pituitary from which it is secreted upon osmotic and hemodynamic stimuli. Once released, it induces vasoconstriction and promotes free water clearance. Also known as antidiuretic hormone.

present. In severe shock, the perfusion of heart and brain is inadequate with severe acidosis and an altered mental state. Finally, the cardiac output diminishes, which further compromises the blood pressure, and respiratory failure will develop.

Circulatory shock is characterized by a failure to maintain adequate tissue perfusion and an impaired oxygen delivery which will ultimately lead to anaerobic metabolism and adenosine triphosphate (ATP) depletion and mitochondrial dysfunction. In general, three stages of shock can be distinguished: early (non-progressive), progressive, and irreversible. The most frequent causes of shock can be divided in cardiogenic, distributive (such as sepsis or anaphylaxis), hypovolemic, and obstructive shock ([Table 1](#)).

Septic shock is more life threatening than hypovolemic or cardiogenic shock, because it is always accompanied by microcirculatory dysfunction and cellular membrane injury, which can lead to multiple organ dysfunctions. This severe stress state involves cardiac dysfunction, acute renal failure, adult respiratory distress syndrome, and hepatic failure.

Introduction

The survival of multicellular organisms is exquisitely dependent on the continuous provision of all body cells with fluid, oxygen, minerals, nutrients, and the uninterrupted elimination of metabolic waste products in order to provide the cells and tissues with the essential circumstances necessary for life. To ensure these vital conditions an intact vascular and microvascular blood flow is a primary prerequisite. Any serious failure of the circulatory system to maintain the cellular perfusion and a fluid homeostasis causes a syndrome of clinical symptoms due to abnormal cellular metabolism, membrane dysfunction, and cell death in tissues and organs throughout the body, eventually ending up in death of the whole organism.

The main symptoms of circulatory failure include hypotension and tachycardia and the patient may show profuse perspiration, the extremities are cool and mottled, and peripheral pulses are weak or impalpable. Due to inadequate perfusion of the brain, a shock state can be demonstrated by restlessness, anxiety, confusion, somnolence, and unconsciousness up to deep coma. In cases of mild shock, there is a decreased perfusion of nonvital organs, such as skin, adipose tissue, skeletal muscle, and bones. The mental state is unimpaired; the urine output may be slightly decreased. In moderate shock, the perfusion of liver, gastrointestinal tract, and kidneys is decreased, but that of the heart and brain is still intact. Urine production is diminished and metabolic acidosis is

Table 1 Classification and etiology of circulatory shock

Hypovolemic shock
Decreased preload and subsequently decreased cardiac output
Hemorrhage (external or internal)
Fluid loss (diarrhea, vomiting, heat stroke, inadequate repletion of insensible losses, burns, and "third spacing" (which can occur in intestinal obstruction, pancreatitis, and cirrhosis)
Cardiogenic shock
Pump failure (infarction, cardiomyopathy, myocarditis)
Arrhythmias (both atrial and ventricular)
Mechanical abnormalities (acute valvular insufficiency, rupture ventricular wall)
Distributive (vasodilatory) shock
Severe decrease in systemic vascular resistance, often associated with an increased cardiac output
Septic shock
Activation of the systemic inflammatory response (pancreatitis, burns, or multiple traumatic injuries)
Toxic shock syndrome
Anaphylaxis and anaphylactoid reactions
Addisonian crisis
Myxedema coma
Neurogenic shock
Post-cardiopulmonary bypass
Obstructive shock
Massive pulmonary embolism
Tension pneumothorax
Severe constrictive pericarditis
Pericardial tamponade
Severe pulmonary hypertension producing Eisenmenger's physiology

Neuroendocrine Regulation in Hypotension and Hypovolemia

Because hypotension and/or hypovolemia compromise the distribution of essential nutrients to body tissue, they represent a threat to homeostasis. Homeostatic systems such as blood and tissue oxygen, acid-base status, and body temperature must be maintained within narrow ranges. Accordingly, hypovolemia- and/or hypotension-induced hormonal responses are part of the well-orchestrated defense mechanisms of the three major central systems of the individual: the nervous, endocrine, and immune system. Acute changes in the hormonal functions are primarily adaptive and provide optimal conditions of the organism for the fight, including optimal intravascular volume and perfusion pressure. Under physiological conditions the acute reaction of the body to these challenges is twofold, turning on defense mechanisms that initiate a complex adaptive pathway (adaptive phase), and then shutting off this response when the threat is past (recovery phase). However, there are limits to the ability of the hormonal system to compensate adequately when the inappropriate hormonal response can contribute to the breakdown of the body tissue (catabolic phase). Hormonal changes during or following hypotension and/or hypovolemia are summarized in [Table 2](#). Salt and water balance is largely determined by factors regulating fluid and electrolyte intake and, more importantly, the ability of the kidney to regulate sodium and water excretion. Vasopressin, the renin-angiotensin-aldosterone (RAA) axis and the sympatho-adrenal system are the key players in the regulation of volume and water during stress ([Figure 1](#)).

Critical illness such as shock is characterized by striking alterations in the hypothalamic-anterior-pituitary-peripheral hormone axes, the severity of which is associated with a high risk of morbidity and mortality. There appears to be a biphasic neuroendocrine response to critical illness. The acute phase is characterized by an actively secreting pituitary, but most peripheral hormones are decreased, partly due to the development of target-organ resistance. This is considered to be beneficial, at least for short-term survival. In prolonged illness, uniform (hypothalamic) suppression of the neuroendocrine axes contributes to low levels of the target-organ hormones. These alterations are considered to be maladaptive and contribute to the catabolism and wasting.

Hypothalamo-Posterior Pituitary System

Vasopressin is a peptide hormone that is synthesized in the hypothalamic nuclei and transported to the posterior pituitary where it is stored. It is released

Table 2 Neuroendocrine changes during hypotension and hypovolemia

<i>Hormone</i>	<i>Acute phase</i>	<i>Chronic phase</i>	<i>Recovery phase</i>
HPA axis			
CRH	↑	↑	↓
ACTH	↑	↓	N
Cortisol	↑↑	↑	↑/N
DHEA(S)	↓↓	↓↓	N
CBG	↓↓	↓	N
RAA system			
Renin	↑	↑	↓/N
Angiotensin II	↑	↑	↓/N
Aldosterone	↑	↓	↓/N
Vasopressin	↑/N	↑	N
ANP	↑	↑↑	N
Sympatho-adrenergic system			
Epinephrine	↑↑	↑/N	N
Norepinephrine	↑	↑/N	N
Prolactin	↑	?	N
Insulin	↓	↑	↓
Glucagon	↑	↑	↑
GH	↑	↓	↑/N
IGF-I	↓	↓	↑/N
IGFBP3	↓	↓	N
Pituitary-thyroid axis			
TSH	N	↓	↓/N
TT4	N	↓	N
TT3	↓	↓	N
rT3	↑	↑	N
FT4	N	↓	N
FT3	↓	↓	↓/N
TBG	?	?	?

↓, decrease; ↑, increase; ?, unknown; ACTH, adrenocorticotrophic hormone; ANP, atrial natriuretic peptide; CBG, cortisol-binding globulin; CRH, corticotropin-releasing hormone; FT3, free triiodothyronine; FT4, free thyroxine; GH, growth hormone; HPA, hypothalamic-pituitary-adrenal; IGF-I, insulin-like growth factor I; IGFBP3, insulin-like growth factor binding protein-3; N, number; RAA, renin-angiotensin-aldosterone; rT3, reverse triiodothyronine; TBG, thyroxine-binding globulin; TSH, thyrotropin; TT3, total triiodothyronine; TT4, total thyroxine.

into the circulation upon stimulation by increased plasma osmolality or as a baroreflex response to decreases in blood volume or blood pressure. Vasopressin exerts antidiuretic, hemostatic, and vasoconstrictive effects via stimulation of V₁- and V₂-receptors. Furthermore, vasopressin serves as a key secretagogue in the body's complex endocrinologic orchestra ([Figure 1](#)). The physiological effect of vasopressin is to promote free water clearance by altering the permeability of the renal collecting duct to water. This effect is mediated by renal V₂-receptors that are coupled to adenylyl cyclase and the generation of cyclic adenosine monophosphate (cAMP). Vasopressin is also capable of causing vasoconstriction and increasing blood pressure. In fact, vasopressin is a more potent vasoconstrictor than norepinephrine and is capable of

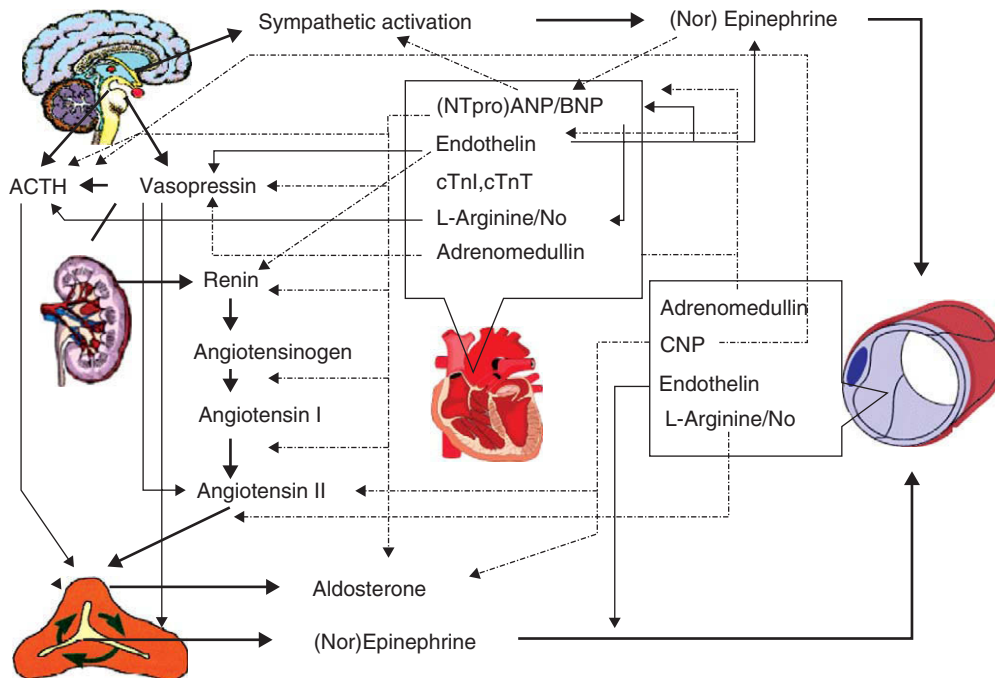


Figure 1 Neuroendocrine regulation of shock and hypotension. Thick lines represent positive feedback mechanisms. Thin lines represent less pronounced positive feedback regulation. Interrupted lines stand for negative feedback. ACTH, adrenocorticotrophic hormone; ANP, atrial natriuretic peptide; BNP, brain (B type) natriuretic peptide; CNP, C-type natriuretic peptide; cTnI, troponin I; cTnT, cardiac troponin T; NO, nitric oxide; NTpro, N-terminal pro.

increasing systemic vascular resistance in doses less than those required to produce antidiuresis. Conversely, vasopressin is a relatively weak pressor agent because it resets the cardiac baroreflex to a lower pressure. These actions are mediated via V_1 -receptors which are coupled to phospholipase C and increased intracellular Ca^{2+} concentration. Aside from the corticotropin-releasing hormone (CRH), vasopressin is known to be the most potent stimulator of adrenocorticotrophic hormone (ACTH) and, consequently, cortisol release. Additionally, vasopressin exerts stimulatory effects on prolactin and endothelin-1 release, and it modulates atrial natriuretic peptide (ANP) and angiotensin II secretion.

Increased vasopressin release is, therefore, a key part of the hypotension- and/or hypovolemia-induced neuroendocrine stress adaptation. Low plasma vasopressin concentration during stress is a sign of the inappropriate autonomic adaptation which has been described in patients with septic or hypovolemic shock. However, it is unclear if this relative vasopressin deficiency has clinical relevance or therapeutic consequences.

Renin-Angiotensin-Aldosterone System

The RAA system plays a pivotal role in the control of cardiovascular homeostasis (**Figure 1**). Renin is a glycoprotein hormone produced by the renal

juxtaglomerular cells in response to hypovolemia and hypotension, but is not a vasoactive substance. Its primary action is to cleave angiotensinogen into the decapeptide, angiotensin I. In the pulmonary circulation, angiotensin-converting enzyme (ACE) removes two more amino acids from angiotensin I to produce angiotensin II, which is the active hormone with a direct strong vasoconstrictor effect (40 times more potent than norepinephrine) and an effect on the adrenal cortex to stimulate aldosterone production. Aldosterone is the key regulator of the potassium balance and induces sodium and water retention. The appropriate response of the RAA axis to hypotension and hypovolemia is renin release, which, via angiotensin II, induces aldosterone secretion by the adrenal cortex, resulting in hyperreninemic hyperaldosteronism, which is a state of secondary hyperaldosteronism often caused by depletion of extracellular fluid and reflected by metabolic alkalosis and hypokalemia. Stress-induced ACTH and vasopressin release also induces aldosterone secretion independently of volume status. An elevated plasma-renin concentration accompanied by inappropriately low aldosterone level is present in some seriously ill patients with hypovolemia and hypotension, and is a sign of the insufficient adaptation of the RAA axis (hyperreninemic hypoaldosteronism, a transient and reversible syndrome characterized by a dissociation of

plasma renin activity and aldosterone production (20–30% of critically ill patients) possibly due to a transient diffuse impairment of the zona glomerulosa of the kidney. Activation of the RAA system occurs during any type of pump failure or shock.

The family of natriuretic peptides consists of ANP, brain (B type) natriuretic peptide (BNP) and C-type natriuretic peptide. ANP is a 28-amino acid polypeptide hormone synthesized in the heart and released into the circulation in response to stretch of the heart chambers, fluid loading, and congestive heart failure. ANP works to oppose the function of the RAA axis via inhibiting the secretion and effects of renin, the effects of angiotensin II, and the adrenal secretion of aldosterone (Figure 1). ANP acts on the kidney to increase sodium excretion and glomerular filtration rate to antagonize renal vasoconstriction. In the cardiovascular system, ANP antagonizes vasoconstriction and shifts fluid from the intravascular to the interstitial compartment. The diuretic and blood pressure lowering effect of ANP may be partially due to adrenomedullin.

BNP and N-terminal pro-brain natriuretic peptide (NTproBNP) are mainly produced in the human ventricle and are cleaved from their inactive precursor proBNP. Both ANP and BNP levels are elevated in a variety of conditions in the critically ill, such as sepsis, trauma, and shock and have prognostic value.

Sympatho-Adrenal System

The sympatho-adrenal response to severe stress is marked, rapid, and transient and plays an important role in the early metabolic changes during the acute phase of a stress response (Figure 1). Epinephrine is secreted from the adrenal medulla in direct response to increased sympathetic tone; however, the majority of the circulating norepinephrine is released from the sympathetic nerve endings. Both catecholamines support blood pressure, heart rate, myocardial contractility, cardiac output, respiration, and bronchial tone. Epinephrine is an important catabolic hormone, inducing lipolysis, glycogenolysis, and gluconeogenesis, and stimulates glucagon release and inhibits insulin release. A high level of plasma catecholamines has been found during the acute phase of a stress response, but the measurements of plasma catecholamine concentrations have limited value as an index of the sympathetic nervous system activity because of their extremely short half-life.

Hypothalamo-Pituitary-Adrenal Axis

The hypothalamic-pituitary-adrenal (HPA) axis does not follow a simple activation response but rather demonstrates a biphasic process during hypotension

and/or hypovolemia. There is a prompt, dramatic, and sustained increase in both cortisol and ACTH concentration in blood. This activation is accompanied by a loss of circadian rhythm, ACTH pulsatility, and CRH sensitivity of the pituitary. During the second catabolic phase the high plasma cortisol level is accompanied by paradoxically low ACTH levels. Vasoactive peptides such as vasopressin, endothelin, ANP, or calcitonin gene-related peptide (CGRP) are implicated in the adrenal activation in this stage. The degree of cortisol elevation correlates with the degree of homeostatic disturbances.

Neuropeptides

In addition to these classical stress hormones, several other hormones are augmented in response to hypotension and/or hypovolemia. There is no doubt that prolactin and growth hormone (GH) are released during this stress response but the biological meaning is much less clear. However, there are experimental observations that neuropeptides may play a causal role in the hypotension and/or hypovolemia-induced cardiovascular responses. There is now appreciable evidence that one such group of neuropeptides, the endogenous opioids, exert a major influence on cardiovascular responses as a counter-stressor system. In canine hemorrhagic shock models, administration of the opioid antagonist naloxone significantly improved mean arterial pressure without change in heart rate and total peripheral resistance; increased cardiac contractility and cardiac output; and reduced mortality. In patients with hemorrhagic shock, infusion of opiate antagonist resulted in improvement in hemodynamic status via a significant fall in heart rate together with improvement in stroke volume without reduction in cardiac output, and reduction in vasopressor requirements. Based on these observations, it has been suggested that although hypovolemic shock is best treated in the human by volume repletion, in an emergency situation, an opioid antagonist might stabilize the patient until the institution of conventional therapy. In addition to endogenous opiates, the endogenous release of the potent vasoconstrictor neuropeptide Y (NPY) is increased in hypovolemic shock. NPY co-exists and is co-released with norepinephrine, and causes direct vasoconstriction and potentiates norepinephrine-induced vasoconstriction in a complex manner. NPY plays a relatively minor role in the minute-to-minute maintenance of cardiovascular homeostasis; its role being one of a second line of defense that comes into play in situations of greater or more prolonged stress. However, the role of CGRP, the most potent vasodilator and hypotensive agent yet tested, has been also suggested as a

causal factor in hypotension-induced stress response. Plasma CGRP levels are related to hemodynamic changes and there is a correlation between the initial CGRP plasma level and the severity of the shock. It is possible that neuropeptides, especially endogenous opiates, represent a physiological buffer of the cardiovascular response to stress, namely as a major counter-stressor system exerting significant influence on cardiovascular responses to both physiological and pathophysiological stresses in animals and humans.

Septic Shock

Septic shock is the most common cause of death in intensive care units. The terms sepsis, severe sepsis, and septic shock are used to identify the continuum of the clinical response to infection. However, the definition of sepsis appears to be very complicated because of the nonhomogenous nature of the patient population studied and the underlying conditions related to sepsis. Septic shock is hypoperfusion with organ dysfunction in a septic patient. Sepsis represents a severe threat to homeostasis and it is related to a systemic inflammatory host response to an inciting event. In general terms, sepsis is a severe stress stimulus which disturbs the milieu interne and induces homeostatic responses specific to the stimulus and generalized responses when the disturbances are severe. Regulation of fluid and volume status is of critical importance to survival. Multiple mechanisms exist to maintain normovolemia and adequate blood pressure in sepsis, and when they fail septic shock arises. Septic shock is often characterized by massive systemic vasodilatation with low vascular resistance, and increased cardiac output. The relative lack of effective pharmacologic interventions highlights the complex pathophysiologic events involved in sepsis. In this paragraph we focus on the pathogenesis, and neuroendocrine-immune interactions (Figure 2).

Pathogenesis

The clinical manifestations of septic shock are the result of an excessive host response to an (usually bacterial) infection. Although host defense mechanisms are usually beneficial and designed to localize and neutralize invading micro-organisms, remove dead and damaged cells, and repair tissue damage, excessive activation may be detrimental. The inflammatory response to infection or injury is a highly conserved and regulated reaction of the organism. Immediately, the organism produces several soluble and lipid pro-inflammatory molecules that activate cellular defenses and then produces similar anti-inflammatory molecules to attenuate and halt the pro-inflammatory

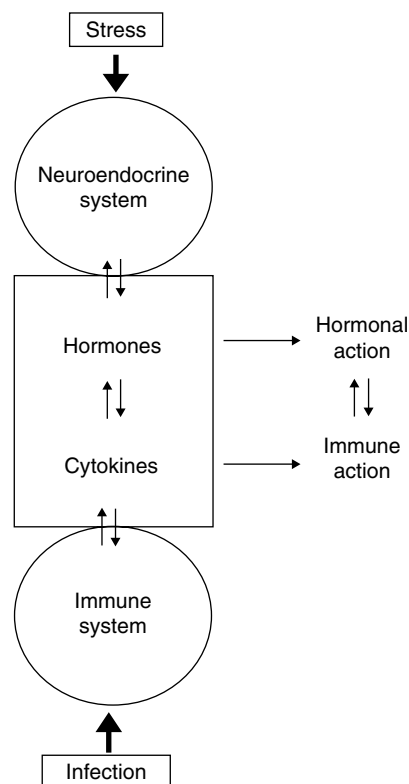


Figure 2 Bidirectional interactions between the neuroendocrine and immune systems during critical illness.

response. The complex interaction of these stress mediators can be seen as a cascade. The initial inflammatory response is kept in check by downregulating production and counteracting the effects of cytokines already produced. These mediators initiate overlapping processes that influence the endothelium, cardiovascular, hemodynamic, and coagulation systems directly. The release of many of these vasoregulators is often local, as part of the total septic-response picture. The overproduction of nitric oxide (NO), a potent vasodilator, has been held responsible for the hemodynamic and metabolic consequences of sepsis and endotoxemia. High levels of nitrite and nitrate, the stable end-products of NO metabolism, are found in patients with septic shock and these levels correlate with vasodilatation and the degree of shock.

The duration of illness also may alter the mix of mediators, leading to a state of metabolic disorders in which the body has no control over its own inflammatory response. If the balance cannot be established and homeostasis is not restored, a massive pro-inflammatory reaction (sepsis inflammatory response syndrome; SIRS) or a compensatory anti-inflammatory reaction (compensatory anti-inflammatory response syndrome, CARS) will ensue.

Bone and colleagues have proposed five stages in the development of septic shock and ultimately multiple organ dysfunction:

1. The local reaction at the site of injury or infection: the initial response is to induce a pro-inflammatory state in which mediators (TNF, IL-1, IL-6, interferon- γ , platelet-activating factor) have multiple overlapping effects designed to limit new damage and to ameliorate whatever damage has already occurred. They destroy damaged tissue, promote the growth of new tissue, and combat pathogenic organisms. A compensatory anti-inflammatory response soon ensures that the effects of these pro-inflammatory mediators do not become destructive. IL-4, IL-10, IL-11, IL-13, soluble TNF receptors, IL-1 receptor antagonists, and others work to diminish monocytic major histocompatibility complex class II expression, impair antigen presenting activity, and reduce the ability of cells to produce inflammatory cytokines. Local levels of these mediators can be much higher than are later found systemically.
2. The initial systemic response: if the original insult is severe enough, first pro-inflammatory and later anti-inflammatory mediators will appear in the systemic circulation. The pro-inflammatory mediators help recruit neutrophils, T cells and B cells, platelets, and coagulation factors to the site of injury or infection. Few significant clinical signs are produced.
3. Massive systemic inflammation: the loss of regulation of the pro-inflammatory response results in a massive reaction manifested as the systemic inflammatory response with (a) progressive endothelial dysfunction, (b) platelet sludging blocking the microcirculation, (c) activation of the coagulation system and (d) profound vasodilatation, fluid transition, and maldistribution of blood flow may result in severe shock.
4. Excessive immunosuppression: this 'immune paralysis' or 'compensatory anti-inflammatory response syndrome' is a result of an inappropriate anti-inflammatory reaction.
5. Immunologic dissonance: this is the final stage of MODS, an inappropriate, out-of-balance response of the immunomodulatory system. Sometimes, monocyte deactivation has been shown.

Basic and clinical science continues to reveal the molecular picture for the etiology and pathophysiology of sepsis, however, sepsis continues to puzzle and challenge physicians. Recent discoveries have shown that not only microbial products but also endogenous molecules can trigger Toll-like receptors (TLR),

which are held responsible for the stimulation of widespread inflammation. TLR-4, the endotoxin receptor, is constitutively suppressed and the first step in sepsis could be the release of TLR-4 from suppression.

Neuro-Immuno-Endocrine Interactions

The endocrine and immune systems are interrelated via a bidirectional network through which hormones and neuropeptides affect immune function and, in turn, immune responses are reflected in neuroendocrine changes (Figure 2). Apparently, there is constant interaction between neuroendocrine and internal immunoregulatory mechanisms that assures the fine tuning of both the neuroendocrine and the immune system so that both are able to preserve homeostasis during severe and life-threatening illnesses. Evidence shows that inflammatory cytokines are produced in the brain and endocrine glands and that these organs have cytokine receptors. This suggests that paracrine mechanisms mediated by locally produced cytokines play certain roles in causing endocrine changes in inflammation or sepsis. Similarly, various neuropeptides, including GH and PRL, are produced in immunocompetent cells and neuropeptide receptors are present on these cells.

Many cytokines, such as IL-1 and TNF, can activate the HPA axis, when given by systemic routes. These cytokines seem to act through the hypothalamus via release of CRH, which is the most important regulator of ACTH. IL-1, IL-2, IL-6, and TNF act directly on the adrenal cortex to enhance biosynthesis of glucocorticoids, possibly contributing to the persisting hypercortisolism, present in critical illness. Injection of bacterial lipopolysaccharide (LPS) into healthy volunteers increases plasma TNF within 90 min followed by increases of plasma ACTH and cortisol levels. Plasma IL-6 also increases 2–3 h after LPS injection, but IL-1 and IFN- γ remain unchanged. IL-6, synthesized through increases of IL-1 and TNF plays a pivotal role in activating the HPA axis. Interestingly, IL-6 suppresses IL-1 production indicating a negative feedback regulation within the cytokine cascades. Cytokines are also produced locally in a variety of tissues, including in the brain. Both IL-1 and IL-6 are involved in activation of the HPA-axis. Various stimuli, for example LPS, lead to the production of cytokines in the anterior pituitary gland, acting through a paracrine or autocrine mechanism to modulate ACTH secretion and to regulate growth of pituitary cells. Cytokines can also be produced in the adrenocortical or medullary cells, with a modulating function in steroid hormone production.

Recent studies have led to the discovery of a mediator that acts as an endogenous counter-regulator of glucocorticoid action within the immune system: macrophage migration inhibitory factor (MIF) produced by the anterior pituitary gland, macrophages and T cells in response to inflammatory stimuli and upon incubation with low concentrations of glucocorticoids. Once secreted, MIF overrides the anti-inflammatory and immunosuppressive effects of steroids on macrophage and T cell cytokine production. Patients with sepsis show elevated MIF levels, which are correlated with poor prognosis. Anti-MIF antibodies protect animals from septic shock.

In conclusion, it appears that the acute response of the HPA axis is mediated by cytokines in inflammatory foci. Since cytokine levels in the brain or endocrine organs under unstimulated conditions are very low, the action of cytokines locally produced by LPS requires a certain latent period. Therefore, the paracrine mechanisms may become significant under prolonged stimulation by infectious agents. Metabolic alterations induced by TNF are considered to be dependant on tissue TNF levels, but not on systemic levels. This suggests the importance of locally produced cytokines. As a result of the activation of the HPA axis, glucocorticoids are finally produced and act on the immune system to inhibit the overshoot of cytokine production and to enhance the defense mechanism of the organism. Thus, there exists a feedback mechanism over the immune and endocrine systems; glucocorticoids inhibit cytokine-induced CRH release and also production of certain cytokines in endocrine organs. Another feedback mechanism may exist in cytokine cascades too. Both IL-1 and IL-1 receptor antagonist are produced by the same cells, probably with a different time course, to modulate overstimulation by IL-1. Complex interactions seem to exist also among inflammatory cytokines. These feedback mechanisms are involved in maintaining homeostasis in inflammatory processes and sepsis.

Sepsis may not only be attributable to hyperactivation of the immune system (uncontrolled hyperinflammation) but systemic anti-inflammatory responses might dominate the body's stress response. The immune system may be severely compromised and unable to eradicate pathogens. Future therapy may be directed at enhancing or inhibiting the patient's immune response.

Further Reading

- Annane, D., Bellissant, E. and Cavaillon, J. M. (2005). Septic shock. *Lancet* **365**, 63–78.
- Beishuizen, A., Hartemink, K. J., Vermes, I., et al. (2005). Circulating and cardiovascular markers and mediators in the critically ill: an update. *Clinica Chimica Acta* **354**, 21–34.
- Beishuizen, A. and Thijs, L. G. (2004). The immunoneuro-endocrine axis in critical illness: beneficial adaptation or neuroendocrine exhaustion? *Current Opinion in Critical Care* **10**, 461–467.
- Besedovsky, H. O. and Del Rey, A. (1996). Immune-neuroendocrine interactions: facts and hypotheses. *Endocrine Reviews* **17**, 64–102.
- Bone, R. C., Grodzin, C. J. and Balk, R. A. (1997). Sepsis: a new hypothesis for pathogenesis of the disease process. *Chest* **112**, 235–243.
- Buckingham, J. C. (1998). Stress and the hypothalamo-pituitary-immune axis. *International Journal of Tissue Reactions – Experimental and Clinical Aspects* **20**, 23–34.
- Calandra, T. and Bucala, R. (1997). Macrophage migration inhibitory factor (MIF): a glucocorticoid counter-regulator within the immune system. *Critical Reviews in Immunology* **17**, 77–88.
- Chrousos, G. P. and Gold, P. W. (1992). The concept of stress and stress system disorders. Overview of physical and behavioral homeostasis. *JAMA* **267**, 1244–1252.
- Espiner, E. A. (1987). The effects of stress on salt and water balance. *Baillière's Clinical Endocrinology and Metabolism* **2**, 375–390.
- Little, R. A., Kirkman, E. and Ohnishi, M. (1998). Opioids and the cardiovascular response to haemorrhage and injury. *Intensive Care Medicine* **24**, 405–414.
- Hotchkiss, R. S. and Karl, I. E. (2003). The pathophysiology and treatment of sepsis. *New England Journal of Medicine* **348**, 138–150.
- Riedemann, N. C., Guo, R-F. and Ward, P. A. (2003). The enigma of sepsis. *Journal of Clinical Investigation* **112**, 460–467.
- Van den Berghe, G., de Zegher, F. and Bouillon, R. (1998). Acute and prolonged critical illness as different neuroendocrine paradigms. *Journal of Clinical Endocrinology and Metabolism* **83**, 1827–1834.
- Vanhorebeek, I., Langouche, L., and Van den Berghe, G. Endocrine aspects of acute and prolonged illness. *Nature Clinical Practice* **2**, 20–31.
- Vermes, I., Beishuizen, A., Hampsink, R. M., et al. (1995). Dissociation of plasma adrenocorticotropin and cortisol levels in critically ill patients: possible role of endothelin and atrial natriuretic hormone. *Journal of Clinical Endocrinology and Metabolism* **80**, 1238–1242.

Hypothalamic-Pituitary-Adrenal Axis

M F Dallman, S Bhatnagar and V Viau

University of California, San Francisco, San Francisco, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition article by M F Dallman, S Bhatnagar and V Viau, volume 2, pp 468–476, © 2000, Elsevier Inc.

The Hypothalamic-Pituitary-Adrenal Axis Is One
Neuroendocrine Arm of the Central Corticotropin
Releasing Hormone Stress System
Characteristics of the Hypothalamic-Pituitary-Adrenal
Axis
Target Sequelae of Corticosteroid Responses to Stress

Glossary

<i>Adrenocorticotrophic hormone (ACTH)</i>	A proopiomelanocortin-derived peptide of 39 amino acids that is released from the corticotroph and acts on MC-3 receptors on adrenocortical cells to stimulate the synthesis and secretion of adrenal corticosteroids.
<i>Arginine vasopressin (AVP)</i>	A peptide that, in addition to its behavioral and antidiuretic roles, potentiates the action of corticotropin-releasing factor to cause ACTH secretion from the corticotroph.
<i>Corticotropin releasing hormone (CRH)</i>	A 41-amino-acid peptide that acts centrally as a neurotransmitter and releasing factor for ACTH.
<i>Corticosterone (B)</i>	The major corticosteroid synthesized and secreted by rat adrenals; also called Kendall's compound B.
<i>Cortisol (F)</i>	The major corticosteroid synthesized and secreted by human adrenals; also called Kendall's compound.
<i>Glucocorticoid receptor (GR)</i>	A receptor found in cells throughout the body that mediates most of the systemic effects of corticosteroids secreted by the adrenal.
<i>Heterotypic stress</i>	The application of a novel stimulus to animals that have been previously or chronically stimulated by another stressor.
<i>Homotypic stress</i>	The reapplication of the same stressor that has been applied to animals previously.
<i>Paraventricular nucleus (PVN)</i>	The hypothalamic site of CRH and AVP neurons, which subserve both neuroendocrine and autonomic functions.
<i>Proopiomelanocortin (POMC)</i>	The 13-kDa precursor protein of ACTH that is synthesized in the anterior pituitary corticotrophs in response to CRH.

The hypothalamic-pituitary-adrenal (HPA) axis, which ultimately controls the secretion of glucocorticoids from the adrenal cortex, is simply one arm of a central CRH system that appears to be responsible for coordinated behavioral, neuroendocrine, autonomic, and immune responses to alterations in homeostasis. There are marked circadian rhythms in the HPA axis, well integrated with other circadian activity to optimize the daily release of CRH and AVP, and its primary function is to stimulate steroidogenesis by the adrenal cortex. In addition, the HPA axis is responsive to stimuli (stressors) that threaten homeostasis, either psychologically or physically. Responses to acute stressors are essential for life, and they turn on and off rapidly. Chronic stressors appear to change the response characteristics in the HPA axis, possibly through recruiting central neural pathways that normally contribute in only a minor fashion to the acute stress response. During chronic stress, both the CRH- and the glucocorticoid-regulatory components of the HPA axis serve to inhibit activity in other, energy-expending hormonal systems of the body, thus conserving energy stores for immediate use.

The Hypothalamic-Pituitary-Adrenal Axis Is One Neuroendocrine Arm of the Central Corticotropin Releasing Hormone Stress System

CRH in the PVN is the essential driver of the anterior pituitary corticotroph, causing the synthesis and secretion of ACTH, which then activates adrenocortical synthesis and the secretion of the corticosteroids: cortisol (F) in humans and corticosterone (B) in rats. ACTH appears to be a slave to hypothalamic release of CRH and AVP, and its primary function is to stimulate steroidogenesis by the adrenal cortex. Both F and B have important effects on central CRH synthesis, behavior, the immune system, and energy balance.

Neuroendocrine Corticotropin Releasing Hormone Neurons

CRH neurons, primarily in the medial parvocellular portion of the PVN, are known as neuroendocrine neurons. Their axons end at the hypothalamic median eminence, and CRH secreted from these diffuses into the hypophyseal-portal blood supply to excite, through its action on CRH receptor 1 (CRH-R1), the synthesis of POMC and secretion of ACTH into the general circulation. ACTH secreted from the pituitary

appears to have the single peripheral function of exciting adrenocortical MC-3 receptors to cause the synthesis and secretion of aldosterone from the zona glomerulosa and F (in humans) or B (in rats), or combinations of the two in other species from the zona fasciculata. The neuroendocrine CRH-synthesizing neurons in PVN also cosynthesize and secrete AVP, which greatly potentiates the effect of CRH on corticotroph ACTH secretion. The amount of AVP synthesized varies with the state of the organism. After adrenalectomy and the removal of F or B negative feedback or after some chronic stressors, AVP synthesis increases markedly in CRH cells. The increase in CRH neuronal AVP allows persistent responsiveness in the HPA axis under conditions when a high corticosteroid feedback signal would otherwise damp the system.

Autonomic Corticotropin Releasing Hormone Neurons in the Paraventricular Nucleus

There are also CRH cells in the dorsal, lateral, and medial parvocellular regions of PVN that innervate the dorsal vagal complex and interomedial lateral portions of the spinal cord, suggesting that they affect neural outflow to the peripheral autonomic nervous system.

Other Corticotropin Releasing Hormone Neuronal Groups

Within the brain, there are clusters of other CRH-synthesizing neurons, in addition to those in the PVN, that also appear to play key roles in behavior, autonomic activity, and energy balance. Some of these CRH-expressing cell groups are located in the central nucleus of the amygdala (CeA), bed nucleus of the stria terminalis (BNST), lateral hypothalamic area, parabrachial nuclei, and dorsal motor nucleus of the vagus. Evidence from lesion studies and central injections of CRH and its antagonists suggests strongly that the central CRH system regulates not only the neuroendocrine and autonomic but also behavioral and possibly the neural recruitment responses to applied stressors.

Several studies have shown that CRH expression in the PVN is decreased by corticosteroid treatment, whereas CRH in the CeA and BNST is increased, suggesting that the corticosteroids exert opposite effects on CRH at different sites in brain. However, recent evidence suggests strongly that neurons in the PVN inhibit neuronal expression of CRH in the amygdala; that is, it is not a direct effect of glucocorticoids on extra-PVN cells that causes increased CRH expression but rather neural signals from the PVN to the amygdala that are altered by decreased or increased corticosteroids. Thus, it appears that at least some

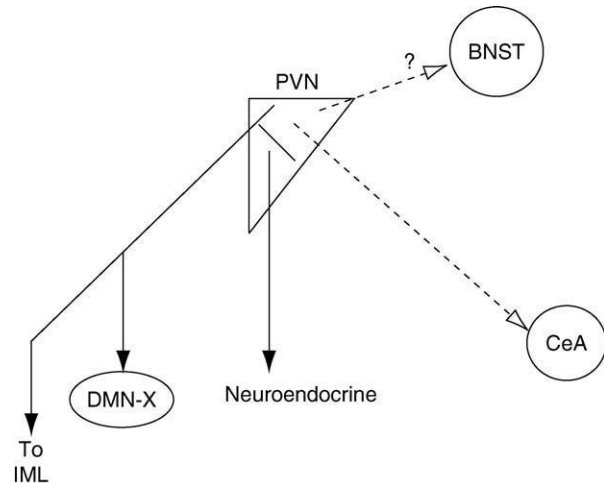


Figure 1 The PVN is central to the stress responsive systems of brain. CRH neurons in the PVN regulate activity in the HPA axis. Neurons in the PVN also regulate autonomic nervous system activity through the innervation of the dorsal motor nucleus of the vagus (DMN-X) and the interomedial lateral cell column of the spinal cord (IML) containing sympathetic preganglionic neurons. Recent results from Palkovits et al. show that pathways from the PVN also regulate CRH expression in the central nucleus of the amygdala (CeA). Although not documented, a similar pathway may exist for CRH neurons in the bed nucleus of the stria terminalis (BNST). Solid and dashed arrows indicate stimulatory and inhibitory connections, respectively.

central CRH pathways are also affected by activity in the PVN (Figure 1). It now seems possible that signals from the PVN serve to orchestrate the activity of the rest of the central CRH system.

Characteristics of the Hypothalamic-Pituitary-Adrenal Axis

The HPA axis exhibits three prominent characteristics that interact to alter the regulation of ACTH and corticosteroid secretion: a circadian rhythm in basal activity, sensitivity to corticosteroid (F and B) feedback, and responsiveness to stressors. The response to a single stress differs considerably from responses to either homotypic or heterotypic stressors on the background of chronic stress, suggesting major regulatory changes occur in the central neural pathways that activate the HPA axis.

Circadian Rhythm of Glucocorticoid Secretion

Vertebrates exhibit peak circulating corticosteroid levels at the onset of their daily activity cycle and trough levels during the early period of inactivity. The circadian rhythm is driven by the biological clock in the suprachiasmatic nuclei in mammals and fish and in the pineal in birds. In humans and rodents, F and B levels change over a wide range between the peak and trough of the rhythm, and both peak and

trough values are regulated by negative feedback effects of F or B on hypothalamic secretion of CRH. During the trough of the rhythm in rats, it appears that there is no CRH/AVP drive from the hypothalamus because neither lesions nor passive immunization with CRH antisera changes ACTH or B secretion; by contrast, CRH(AVP) secretion is essential to drive the circadian increase in B, as shown in the same experiments. Similarly, either adrenalectomy or reduction in normal B levels by use of adrenal B synthesis inhibitors elevates ACTH at both times of day. Both ACTH and direct neural regulation appear to account for the marked circadian rhythm in the amplitude of adrenal corticosteroid production.

Corticosteroid Feedback

Feedback regulation of stress-induced ACTH secretion is markedly affected by the circadian rhythm in HPA activity. There is high feedback sensitivity at trough times and relatively low sensitivity during peak times. In rats with fixed B levels, this is independent of the corticosteroid levels at the time of stressor application, but appears to correlate with the circadian drive to the CRH neurons.

Stressors

Acute signals that alter the balance in the organism stimulate acute HPA responses. These stimuli are most often termed stressors; however, there are also HPA responses associated with stimuli to hypothalamically regulated functions, such as feeding and copulation.

Acute stressors Acute, punctate stressors such as a 30-min psychological stress in humans or a similar period of restraint in rats provoke acute HPA responses that are usually completed within 2 h (Figure 2). The intensity and duration of the acute stimulus are related directly to the magnitude of the HPA response. These acute HPA responses are also modified to an extent by glucocorticoid feedback acting in the fast and intermediate time domains on mineralocorticoid receptors (MRs) and GRs. In rats, there is a clear behavioral role of the acute rise in B acting at GR-containing cells in the amygdala to consolidate memory of aversive events.

Persistent or chronic stressors markedly alter regulation in the hypothalamic-pituitary-adrenal axis Under conditions of chronic stress, there are changes in the amplitude of circadian rhythms, with increasing intensity of stress resulting in decreases in the amplitude of the rhythm through the elevation of trough values. Moreover, the magnitude and duration of acute stress responses under conditions of chronic stress depend on familiarity with the stressor. Chronic

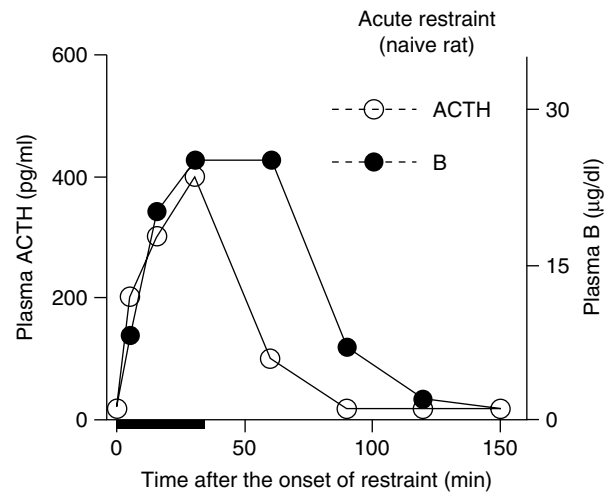


Figure 2 An example of typical ACTH and corticosterone (B) responses to acute restraint in naive rats not previously exposed to stressors. The duration of restraint was 30 min; thereafter rats were returned to their home cages. The ACTH response ceases at the end of restraint; because ACTH levels are high, the B response is prolonged for 30 min beyond the period of restraint.

or prior stress appears to recruit neural pathways that alter regulation of acute HPA responses.

Circadian rhythm Under resting conditions, without additional acute stress, the signs of mild chronic stress are increased trough corticosteroid levels, suggesting strongly that under chronically stressful conditions there is hypothalamic stimulation of ACTH secretion at all times of day. With increasing intensity of chronic stressors, ACTH and corticosteroid secretion is elevated above normal around the clock (Figure 3). The elevation of the trough corticosteroid values is biologically significant and has been shown to enhance glucocorticoid-dependent target responses.

Altered responses to homotypic and heterotypic stressors When acute stress is presented to chronically stressed rats, it becomes clear that marked stress-induced alterations in control of the HPA axis have occurred. In rats, it has been shown that if the acutely applied stimulus is a repetition of the chronic stimulus (homotypic stress), the duration and magnitudes of ACTH and B responses are markedly inhibited compared to the responses of the naive rat (Figure 4a; cf. Figure 2). By contrast, if the acutely applied stimulus is novel and different from the chronic stimulus (heterotypic stress), the duration and magnitudes of ACTH and B responses may be markedly facilitated compared to the responses of the naive rat (Figure 4b; cf. Figure 2). These results described for restraint as a final stimulus appear to be generalizable to other types of chronic stress followed by either homotypic stress or heterotypic stress.

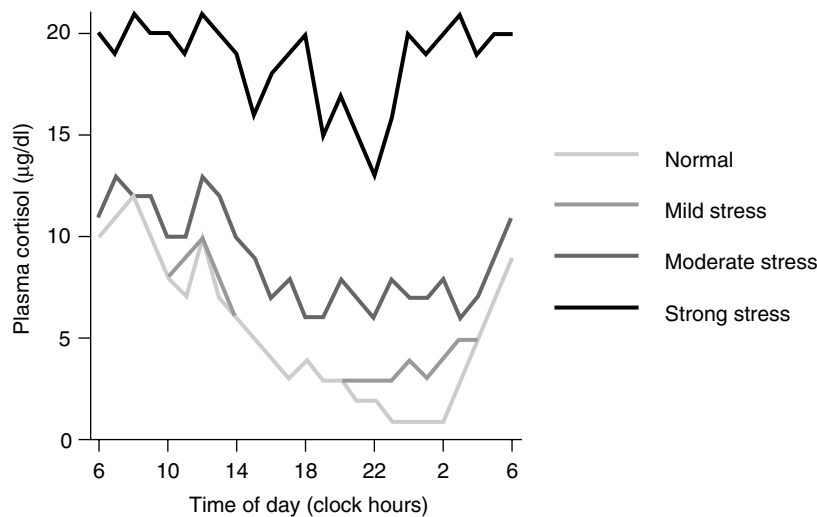


Figure 3 Examples of the effects of chronic stress of differing intensity on basal cortisol concentrations in humans. These examples are collected from a number of clinical papers and show stressors ranging from aging (which some may not perceive as a stress but which bears the hallmarks of chronic stress) or brief fasting at the low end to anorexia nervosa or coronary infarction at the high end. Note that under conditions of mild stress, only the trough levels of the basal rhythm are affected; as the stimulus intensity increases, mean F levels increase until, at highest intensity, the normal circadian rhythm disappears.

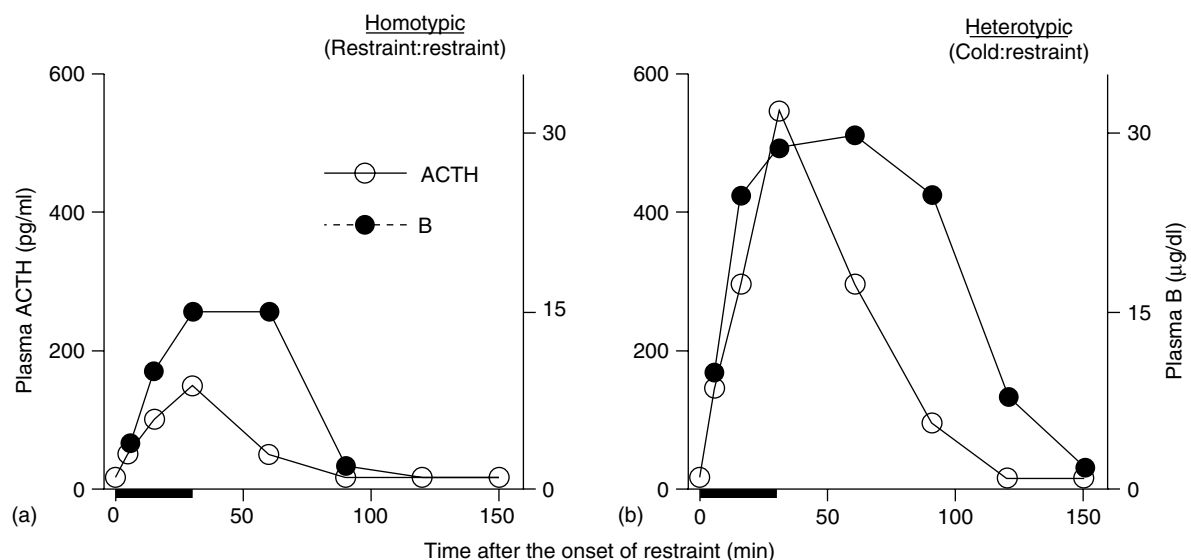


Figure 4 Typical HPA responses to restraint. a, As a homotypic stressor (restraint : restraint); b, As a heterotypic stressor (cold : restraint). Clearly, homotypic HPA responses are diminished (adaptation) and heterotypic responses are exaggerated (facilitation) in chronically stressed rats compared to naïve rats exposed only to the acute stress or restraint (see Figure 2). These are representative results (relative to results in concurrently run naïve rats) from separate sets of experiments of (a) Viau and (b) Bhatnagar.

If we were dealing only with repetitions of the same stimulus, it would be appealing to suppose that the final ACTH and B responses were reduced because of the prior and repeated stress-induced increases in B and its feedback effects on brain; however, because heterotypic stimuli elicit either normal or larger than normal responses, it is clear that prior stress-induced increases in B cannot explain the effect of chronic stress on regulation of the HPA axis.

Neural pathways mediating acute hypothalamic-pituitary-adrenal responses in chronically stressed animals There must be, as well, alterations in the neural pathways that respond to homotypic and heterotypic stimuli in chronically stressed rats.

In recent years, either *in situ* hybridization or immunocytochemical techniques have been used to measure changes in immediate early gene (most usually Fos) expression with time after stressors applied

to stress-naïve rats in order to determine which neural structures are excited by specific stimuli. These studies have shown that most stimuli activate a common set of structures; among these are the nucleus of the tractus solitarius (NTS), ventrolateral medullary and pontine catecholaminergic cell groups, lateral parabrachial nuclei, raphe nuclei, locus coeruleus (LC), PVN, thalamic paraventricular nuclei (PVTh), amygdala, BNST, lateral septum, and prefrontal cortex. Recently, these techniques have been used in chronically stressed rats exposed to either a homotypic or heterotypic stressor. In both sets of experiments, the numbers of Fos immunoreactivity (Fos-ir) cells in a given structure were compared in naïve and chronically stressed rats after a final stress (foot shock : foot shock and cold : restraint). In both reports, more neurons in PVTh stained with Fos in the chronically stressed rats. In marked contrast to the increased number of Fos-staining cells in the PVTh, after homotypic stress the numbers of Fos-expressing cells in the amygdala, PVN, raphe nuclei, and LC were decreased compared to naïve rats experiencing the stressor for the first time; the reduction in Fos-ir cells in the PVN probably reflected a reduction in ACTH that generally occurs after the *n*th application of a homotypic stressor. After heterotypic stress, the opposite was observed; the numbers of Fos-positive cells in the amygdala, PVN, raphe nuclei, and LC were increased in the chronically stressed rats compared to naïve rats experiencing restraint. In this case, the increase in PVN Fos-ir cells was shown to reflect elevated ACTH responses that occurred after the heterotypic stressor.

The role of the PVTh on acute restraint-induced ACTH secretion was studied in naïve and chronically cold-stressed rats by lesions of the structure. Lesions of the PVTh augmented restraint-induced ACTH secretion in chronically stressed rats, although they had no effect on the amplitude of the ACTH response to restraint in stress-naïve rats. Thus, it appears that the PVTh normally acts to inhibit HPA activity in chronically stressed, but not naïve, rats; perhaps, therefore, the increased number of activated cells in this structure represents a memory of prior stress. Because the number of Fos-positive cells in the PVTh increases after both homo- and heterotypic stressors, it does not seem likely that this structure exhibits opposite functional effects after the two types of stressors; it is more likely that the output of the PVTh on the HPA response to acute stress in chronically stressed rats is generally inhibitory and independent of the type of acute stress.

The output of the posterior PVTh, where increased numbers of Fos-ir neurons were found after heterotypic restraint stress, is quite discrete; this portion of the nucleus innervates the basolateral nucleus of the

amygdala (BLA), basomedial nucleus of the amygdala (BMA), and CeA – also sites where decreased numbers of Fos-ir cells were found after homotypic shock stress and where increased numbers of Fos-ir cells were found after heterotypic restraint stress in cold-exposed rats. Based on Fos staining in rats after the application of homotypic and heterotypic stressors and some functional results, the diagrammatic circuit in [Figure 5](#) shows that some pathways to the activation of the HPA axis may be recruited in the chronically stressed individual.

[Figure 5](#) clearly emphasizes the major role played by the amygdala in responding to acute stress under conditions of chronic or prior stress. Only a small amount of the known anatomic connectivity is shown in the figure. For instance, both the NTS and lateral parabrachial nucleus directly innervate all the other cell groups as well as the PVTh; the LC and raphe nuclei also innervate all the other cell groups shown in the diagram; and the PVN also innervates the amygdala, LC, raphe nuclei, lateral parabrachial nucleus, and NTS. Thus, the proposed circuit is massively parallel in nature. Afferent input probably ascends primarily through parallel inputs from the NTS (and ventral medullary aminergic pathways) and the parabrachial nuclei to more anterior brain sites. It is assumed that, under conditions of both homo- and heterotypic stress, the PVTh inhibits activity in the amygdala that leads to the activation of the HPA axis as well as behavioral and autonomic activity. The raphe nuclei, through secretion of serotonin at the lateral nucleus of the amygdala, may inhibit amygdalar activity, possibly mediating HPA adaptation to homotypic stress.

Although the circuit in [Figure 5](#) is based primarily on changes in the number of Fos-positive neurons after a final application of the same or a novel stress in chronically or repeatedly stressed rats, to date there has been no attempt to identify the phenotype of the increased Fos-positive cells after the application of novel stimuli. Validation of the model shown in [Figure 5](#) will require this identification and will also require the identification of the terminal fields of the neurons whose activity is altered specifically under conditions of novel stress on a background of chronic stress.

It is clear from this discussion that chronic stress alters the central regulation of stress-induced HPA function. In addition, the altered neural emphasis also effects changes in behavioral and autonomic outputs. The altered amygdalar output that proceeds through the CeA is compatible with changes in the activity of the amygdalar CRH system, which innervates all the indicated output structures. Thus, the major central stress mediator, CRH, may be in large part responsible

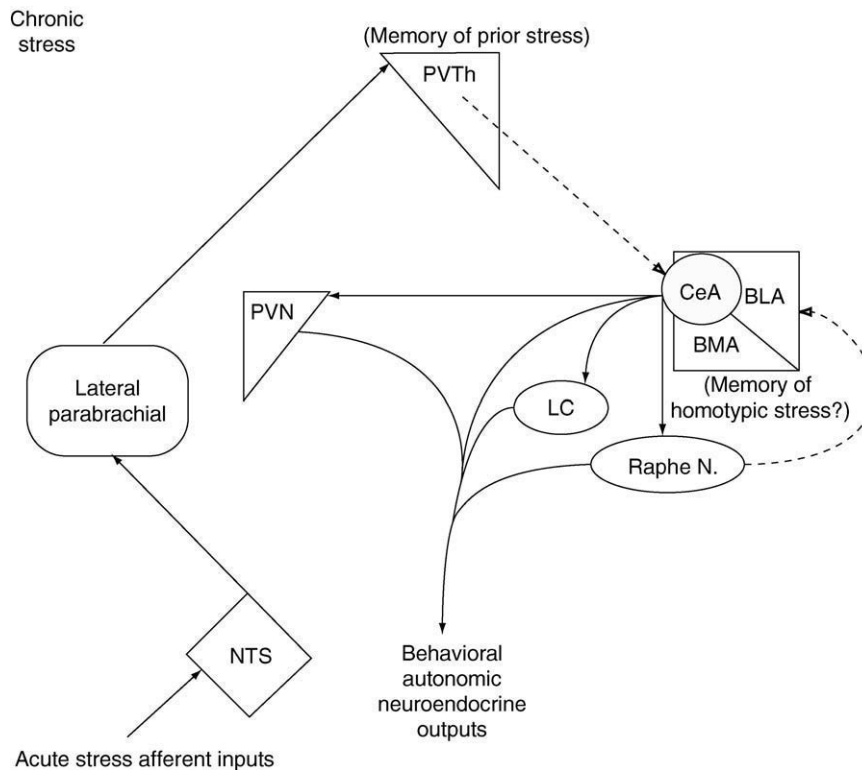


Figure 5 Model of pathways that are emphasized in HPA responses of chronically stressed rats to acute homotypic or heterotypic stressors. We assume that afferent inputs excited by stress excite the brain through activating neurons in the nucleus of the tractus solitarius (NTS) and, in turn, the lateral parabrachial nucleus. Based on results of Fos immunoreactivity (Fos-ir), the number of excited neurons in the thalamic paraventricular nuclei (PVTh) is increased after both types of chronic stressors. By contrast, compared to controls, the number of Fos-ir neurons is decreased after homotypic stress and increased after heterotypic stress in amygdala (CeA, BMA, and BLA), the PVN, the locus ceruleus (LC), and the raphe nuclei. Solid and dashed arrows indicate stimulatory and inhibitory connections, respectively. Only a selected portion of the interconnectivity is shown; both NTS and the parabrachial nucleus innervate all the other structures shown and, in turn, are innervated by the other structures. Moreover, the LC and raphe nuclei innervate all other structures. Thus, this is a massively parallel system.

for the changes in HPA responsivity observed after chronic stress.

Target Sequelae of Corticosteroid Responses to Stress

Because all cell types contain GRs, corticosteroid secretion above normal levels affects function in most tissues of the body. The alterations include changes in behavior and in endocrine and immune responses as well as in intrinsic activity of individual organs. Acute HPA responses to a single application of a stressor probably have considerable utility in that the elevated corticosteroid levels ensure an adequate substrate for increased metabolic need and help to sustain blood pressure and to depress immune function. However, chronic elevations in glucocorticoid levels induced by chronic stress can be deleterious to the organism. In addition to altering the expression of many gene products in the central nervous system, the

elevation of glucocorticoid levels inhibits central catecholaminergic systems and stimulates the expression of neuropeptide Y (NPY). The diminution of enzymatic activity in catecholaminergic systems has been correlated with reduced systemic catecholaminergic function. The stimulation of NPY in the hypothalamic arcuate nucleus has been correlated with increased food intake. In addition, glucocorticoids have been shown to be required at the amygdala for the consolidation of the memory of conditioned stimuli as well as for the inhibitory action of serotonin on glutamate-induced stimulation of lateral amygdalar cell activity. Slight elevations of glucocorticoid secretion above normal inhibit activity in the thyroid, growth hormone, and gonadal axes through actions both at the central nervous system and at the pituitary and peripheral endocrine targets.

At the tissue level, glucocorticoids inhibit immune responses and protect the organism through a sequenced feedback effect from the consequences of

overactivity in this potent system. However, because the glucocorticoids also, through central nervous system effects, alter peripheral catecholaminergic activity, the effect of stress on immune system function depends markedly on when the stress, or glucocorticoid treatment, is provided in temporal relationship to the immune challenge. Thus, the literature on this subject is confusing, showing that stress, or glucocorticoids, can enhance or inhibit immune system activity. The effects of glucocorticoids on immune cells are legion. The direct effects of glucocorticoids on target tissue activity may seem somewhat simpler. By themselves, glucocorticoids are catabolic hormones; they mobilize free fatty acids and amino acids from adipocytes and muscle and cause hyperglycemia through increasing hepatic gluconeogenesis. However, glucocorticoids also enhance insulin secretion. Normally, stress decreases food intake and insulin secretion so that the catabolic effects of the glucocorticoids predominate under stressful conditions; nonetheless, the stimulation of insulin by stress-induced glucocorticoid secretion occurs. This effect is perhaps through antagonizing the hepatic effects of insulin, and it tends to remodel energy stores so that more fat and less muscle are stored under conditions of high glucocorticoid levels.

Bone loss has become a hallmark of chronic stress, depression, anorexia nervosa, or glucocorticoid treatment. Again, these effects of the glucocorticoids are dual. They act directly on bone to decrease matrix deposition and osteoblast activity and at the hypothalamic-pituitary-gonadal axis to decrease gonadal steroid production. The gonadal steroids have positive effects on bone formation, and excess glucocorticoids provide a double hit on bone by their dual actions.

In summary, the HPA axis is one arm of the central stress system defined by CRH. Chronic stressors alter the regulation of the HPA axis, possibly by increasing the emphasis on activity in CRH pathways. Although acute glucocorticoid responses to stressors can be taken as important for the maintenance of homeostasis, chronic elevations in glucocorticoid secretion appear, on the whole, to be deleterious for the organism.

See Also the Following Articles

Adrenocorticotrophic Hormone (ACTH); Corticotropin Releasing Factor (CRF); Glucocorticoids, Role in Stress; Paraventricular Nucleus; Vasopressin.

Further Reading

Antoni, F. A. (1986). Hypothalamic control of adrenocorticotropin secretion: advances since the discovery of

41-residue corticotropin-releasing factor. *Endocrine Review* 7, 351–378.

Bhatnagar, S. and Dallman, M. F. (1998). Neuroanatomical basis for facilitation of hypothalamic-pituitary-adrenal responses to a novel stressor after chronic stress. *Neuroscience* 84, 1025–1039.

Buijs, R. M., Wortel, J., Van Heerikhulze, J. J., et al. (1997). Novel environment induced inhibition of corticosterone secretion: physiological evidence for a suprachiasmatic nucleus mediated neuronal hypothalamic-adrenal cortex pathway. *Brain Research* 758, 229–236.

Cullinan, W. E., Herman, J. P., Battaglia, D. F., et al. (1995). Pattern and time course of immediate early gene expression in rat brain following acute stress. *Neuroscience* 64, 477–505.

Dallman, M. F., Akana, S. F., Scribner, K. A., et al. (1992). Stress, feedback and facilitation in the hypothalamo-pituitary-adrenal axis. *Journal of Neuroendocrinology* 5, 517–526.

Dallman, M. F. and Bhatnagar, S. (1999). Chronic stress and energy balance: role of the hypothalamo-pituitary-adrenal axis. In: McEwen, B. S. (ed.) *Handbook of physiology: environmental physiology* (vol. 7), chap. 10. Washington, DC: American Physiology Society.

Dunn, A. J. and Berridge, C. W. (1990). Physiological and behavioral responses to corticotropin-releasing factor: is CRF a mediator of anxiety or stress responses? *Brain Research Reviews* 15, 71–100.

Herman, J. P., Prewitt, C. M.-F. and Cullinan, W. E. (1996). Neuronal circuit regulation of the hypothalamo-pituitary-adrenocortical stress axis. *Critical Reviews in Neurobiology* 10, 371–394.

Li, H.-Y. and Sawchenko, P. E. (1998). Hypothalamic effector neurons and extended circuitries activated in “neurogenic” stress: a comparison of footshock effects exerted acutely, chronically, and in animals with controlled glucocorticoid levels. *Journal of Comparative Neurology* 393, 244–266.

McGaugh, J. L., Cahill, L. and Roozendaal, B. (1996). Involvement of the amygdala in memory storage: interaction with other brain systems. *Proceedings of the National Academy of Sciences USA* 93, 13508–13514.

Munck, A., Guyre, P. M. and Holbrook, N. J. (1984). Physiological functions of glucocorticoids in stress and their relations to pharmacological actions. *Endocrine Review* 5, 25–52.

Owens, M. J. and Nemeroff, C. B. (1991). Physiology and pharmacology of corticotropin-releasing factor. *Pharmacological Reviews* 43, 425–473.

Palkovits, M., Young, W. S., III, Kovacs, K., et al. (1998). Alterations in corticotropin-releasing hormone gene expression of central amygdaloid neurons following long-term paraventricular lesions and adrenalectomy. *Neuroscience* 85, 135–147.

Stutzmann, G. E., McEwen, B. S. and LeDoux, J. E. (1998). Serotonin modulation of sensory input to the lateral amygdala: dependency on corticosterone. *Journal of Neuroscience* 18, 9529–9538.

Swanson, L. W., Sawchenko, P. E., Rivier, J., et al. (1983). Organization of ovine corticotropin-releasing factor immunoreactive cells and fibers in the rat brain: an immunohistochemical study. *Neuroendocrinology* 36, 165–186.

Taché, Y., Martinez, V., Million, M., et al. (1999). Corticotropin-releasing factor and the brain-gut motor response to stress. *Canadian Journal of Gastroenterology* 13(supplement A), 18A–25A.

Hypothalamus See: Hypothalamic-Pituitary-Adrenal; Neuroendocrine Systems.

Hypothermia

M J Taylor

Organ Recovery Systems, Inc., Charleston, SC, and Drexel University and Carnegie Mellon University, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M J Taylor, volume 2, pp 484–495, © 2000, Elsevier Inc.

Definitions

Background to Hypothermia as a Stress

The Extent of the Problem of Death due to Hypothermia

Pathophysiology of Hypothermia

Glossary

<i>Cold-induced fibrillation</i>	Very rapid, irregular contractions of the muscle fibers of the heart resulting in a lack of synchronization between heart-beat and pulse, with a resultant lack of cardiac output.
<i>Cold narcosis</i>	A state of stupor, unconsciousness, or arrested activity produced by cooling.
<i>Homeotherm</i>	A warm-blooded animal such as humans that regulates its core temperature at a constant temperature (usually 37°C, or 98.6°C) irrespective of environmental temperature.
<i>Ischemia</i>	Localized deprivation of oxygen (hypoxia) and nutrients to a tissue due to interruption of blood supply.
<i>Lipid phase transition</i>	A temperature-induced change of state of the lipid component of biological membranes involving a change from a gel phase to a liquid crystal phase. These physical changes of membrane fluidity

often result in functional alterations of membrane properties.

Thermogenesis

The production of heat in the body, usually by metabolic oxidation of food substrates.

Definitions

The prefix hypo means below or under, and therm means temperature, so hypothermia literally means below normal temperature. In humans and most other mammals, normal temperature is regarded as 37°C (98.6°F), but this single value belies the fact that normal temperature in healthy individuals varies within a degree or so, depending on factors such as the site of measurement (core or exterior) and time of day (circadian rhythm). Body temperature can be measured in a variety of ways. Exterior temperatures include sublingual (under the tongue), skin, or axillary (under the armpit) measurements. Core or interior temperatures reflect the mean temperatures of the vital organs and are approximated by esophageal, rectal, tympanic membranes, and pulmonary artery measurements. In humans, normal body temperature has traditionally been set at 37°C based upon 1 million axillary measurements from 25 000 individuals carried out by the nineteenth century investigator Carl Wunderlich in 1868. Modern-day corroboration of these data has been provided by a study of 148 healthy men and women that showed a mean oral temperature of 36.8°C, with 37.7°C as the upper limit of normal.

Humans are homeothermic mammals (warm-blooded animals), which means that the core body temperature is precisely regulated and kept constant at about 37°C, irrespective of the ambient

temperature. Even though core temperature is well maintained, the temperature of the peripheral shell (skin, fat, and muscle) ranges from 31°C to 35°C, with skin temperature at 28°C to 32°C. Thermoregulation impairment, prolonged exposure to an extreme environment, or both may lead to a decrease in core temperature and the onset of a pathological state. Nevertheless, under controlled conditions in a hospital setting, hypothermia can actually be used for therapeutic advantage in a variety of medical applications since cold can also be a protective modality. Hypothermia is defined in homeotherms as the reduction of core body temperature below 35°C (95°F), and clinically has been arbitrarily graded for convenience into degrees of hypothermia, as shown in [Table 1](#).

Background to Hypothermia as a Stress

Hypothermia is a double-edged sword in that cold has been both a friend and a foe to humanity. Historically, we know that cold can be a devastating stress, as illustrated by the nineteenth century records that Napoleon watched helplessly as his army was annihilated by the bitter Russian winter. Hypothermia remains a modern-day problem, and much of what we know has been gleaned from data collected from victims of accidental hypothermia. These include victims from outdoor recreational activities in which otherwise healthy individuals succumb to injury, and even death, as a result of overexposure to the cold weather conditions, as well as victims of urban problems such as old people living in substandard accommodations without adequate heating in the winter or homeless people sleeping on the street. The latter cases introduce some complications to the interpretation of hypothermic exposure due to confounding issues such as disease, drug or alcohol abuse, or age. In sharp contrast, deep hypothermia has saved some individuals from inevitable death due to prolonged hypoxic ischemia. There are remarkable accounts of children who have been rescued and have made complete recoveries after lengthy submersion in icy water that reduced their core body temperature to below 25°C (77°F) and caused them to stop breathing for periods as long as 40 min.

Under controlled conditions, deliberate hypothermia, while still a stress, can be used as a highly effective therapeutic modality to control, or counteract, the effects of hypoxia or ischemia associated with various medical and surgical treatments. From the very earliest medical manuscripts in the third to fifth centuries BC, we know that cold was recommended as a treatment for wounds, infections, sprains, fractures, and gouty swellings and to suppress hemorrhage (uncontrolled bleeding) and head injuries. Even total body hypothermia was recommended for tetanus and convulsions. Since the 1950s, a vast literature has emerged describing the value of hypothermia in widely varying diseases, including head injury, cerebral hemorrhage, cardiac arrest, carbon monoxide poisoning, neonatal asphyxia, and eclampsia. Today, the elective use of moderate and deep hypothermia has found numerous clinical applications centered principally upon protection of the heart and brain to facilitate techniques in cardiothoracic surgery and neurosurgery. Moreover, the robust cerebroprotective effect of mild hypothermia demonstrated in animals is currently undergoing clinical trials in the management of traumatic brain injury and may in the future be applied as an adjunctive technique in other critical situations such as victims of stroke and hemorrhagic shock.

The Extent of the Problem of Death due to Hypothermia

Accidental hypothermia is regarded as an escalating problem, but accurate statistics are not readily available. This is largely due to the fact that records often do not recognize or specify hypothermia as the primary cause of morbidity or mortality in victims of accidental hypothermia. Official health statistics based on death certificates tend to underestimate the actual incidence of hypothermia. This is partly due to the use of a single-cause coding system; for example, if a death is recorded as bronchopneumonia due to hypothermia, the cause of death is recorded as pneumonia and no statistical record of hypothermia is made. The issue is further complicated by the fact that hypothermia is often not diagnosed or recognized by physicians. Nevertheless, predictions have been made that significant numbers of people die from hypothermia-related causes. For example, based upon the research of Taylor and Fox in Britain, it has been predicted that 20 000 hypothermia-related deaths occur per year, predominantly in the elderly. Similar estimates in North America predicted 25 000 hypothermia-related fatalities per year in the United States and 8 000 in Canada. There is a growing awareness of exposure hypothermia associated with

Table 1 Classification of hypothermia

<i>Classification</i>	<i>Temperature range (°C)</i>	<i>Temperature range (°F)</i>
Mild	32–35	90–95
Moderate	27–32	81–90
Deep or profound	10–27	50–81
Ultraprofound	<10	<50

recreation, cold water immersion, and to some extent socioeconomic problems among people with predisposing factors such as drugs, alcohol intoxication, and age. However, attention also needs to be given to secondary hypothermia, such as that induced during shock resulting from trauma.

Pathophysiology of Hypothermia

Systemic Effects of Hypothermia

Surviving hypothermic exposure in humans is generally associated with core temperatures above 25°C (77°F); however, surviving deeper exposure is possible if appropriate treatment is employed. Physiological responses to cooling and warming depend on how heat exchange is mediated and whether the patient is conscious or anesthetized. Heat exchange may be either by surface phenomena (conduction, convection, and radiation) or in the clinical setting by core perfusion techniques. The pathophysiology and treatment of accidental hypothermia victims are complex issues often compounded by the nonhomogeneity of the patient population. For example, a drowning person may become seriously cooled in a matter of minutes, while elderly people living in substandard accommodations may undergo an insidious cooling process that develops over several weeks. Under these circumstances it has proved difficult to compare different methods of treatments. However, in the comparatively controlled environment of the hospital operating room, the physiological responses to hypothermia are well characterized and documented.

Effect of hypothermia on thermoregulation Hypothermia triggers an ordered progression of autonomic thermoregulatory responses. Initially, cutaneous vasoconstriction is initiated in the arteriovenous shunts of the extremities, mediated by noradrenergic sympathetic nerves. With moderate cooling, this reflex vasoconstriction of the skin is assisted by direct constrictor action of cold on the arterioles. Aided by reflex noradrenergic drive to the heart, the vasoconstriction can result in large increases in arterial pressure. So, the body tightly controls its core temperature, using its peripheral shell as a thermal buffer with the environment. As heat is lost from the periphery, the temperature gradient from the core to the periphery increases, and heat flows out of the core tissues into the peripheral shell. Eventually vasoconstriction can no longer compensate for the continuing heat loss of the periphery by sacrificing peripheral temperature, and shivering is necessary to maintain core temperature. With progressive cold exposure, reflex increases in muscle tone and shivering are

brought about by somatic motor nerves, and this can increase heat production by as much as fivefold for many hours at a time. Interestingly, newborn infants do not shiver, but utilize a nonshivering thermogenesis that relies upon metabolic heat production from the oxidation of brown fat. In general, when hypothermia reaches moderate degrees (<33°C), shivering gradually decreases and is followed by stiffness of muscles and joints.

Effect of hypothermia on the heart and brain Hypothermia has significant effects on nearly every organ system, but disturbances of cerebral and cardiac function as the temperature of the brain and heart fall below 35°C (95°F) are of most concern. One of the earliest signs of hypothermia is a change in personality. The individual may become uncooperative and show signs of listlessness, incoordination, and confusion, often with subsequent amnesia for events at the time of low body temperature. Loss of consciousness is highly variable and may occur at any point within the range 26–33°C. After an initial rise in cardiac output associated with shivering, cardiac function declines with slowed heart rate due principally to the direct effects of cold on the cardiac muscle. At temperatures between 17 and 26°C (63 and 79°F), the cardiac output is not sufficient to supply even the reduced oxygen requirements of the cold tissues. At 25°C (77°F), cardiac function is reduced to about 30% of normothermic levels. The most important and life-threatening complication of hypothermia is the associated cardiac irritability. Progressive cooling eventually causes cardiac arrhythmias, with atrial fibrillation occurring earlier (~30°C [~86°F]) than ventricular fibrillation (threshold at 28°C [82°F]). However, the latter is highly variable and is influenced by several factors, including acid–base balance. Rapid cooling of large, aged, or diseased hearts can predispose to early fibrillation, while small, immature hearts may not fibrillate at all during cooling but pass directly from bradycardia (slow heart rate) to diastolic arrest. Ventricular fibrillation may occur in the temperature range 20–27°C (68–81°F) and is the most dangerous and feared aspect of hypothermia, since death invariably ensues unless rewarming and defibrillation procedures are initiated.

The brain is a highly metabolic tissue requiring a constant supply of oxygen and, like the heart, is exquisitely sensitive to hypoxia and ischemic insult. Metabolism is markedly suppressed by a fall in temperature, and in general metabolic rate is decreased by 5–7% per degree Celsius. Oxygen consumption is a reasonable measure of metabolic activity, since for practical purposes, tissue and cellular stores of oxygen do not exist, and cells rely upon the circulation to

deliver oxygen in quantities determined by the rate of O₂ consumption. Cerebral blood flow (CBF) and cerebral metabolic rate of oxygen consumption also decrease by about 50% for every 10-degree drop in temperature. This correlates with declining electrical activity (EEG) as the brain cools with electrocerebral silence (no measurable electrical activity) at ~20°C. This profound effect of hypothermia on diminished oxygen demand is actually one of the most important protective aspects of clinical hypothermia. However, despite the unequivocal fact that hypothermia protects the brain from the effects of hypoxia, prolonged survival in hypothermia almost totally depends on having sufficient cardiovascular function for adequate perfusion of important organ systems for supply of substrates in addition to oxygen. Hypothermia also causes a shift of the oxygen-hemoglobin dissociation curve to the left, thus reducing tissue oxygen availability. This may contribute to the uncompensated metabolic acidosis that is commonly observed with hypothermia due to reduced tissue perfusion with accumulation of lactate and other acid metabolites.

Effect of hypothermia on respiration The rate of respiration decreases progressively during cooling to 30°C, but the partial pressure of oxygen remains essentially normal. CO₂ production is significantly elevated initially due to intense shivering, but as the core temperature falls, CO₂ production is diminished. Coupled with increased solubility of CO₂ in plasma at lower temperatures, this tends to decrease the partial pressure of CO₂ despite diminished respiratory rate. The net effect of these changes is a mild respiratory alkalosis in the early phase of acute hypothermia. This alkalosis is, however, gradually replaced by metabolic acidosis as cooling progresses, due to poor tissue perfusion and oxygenation, as discussed earlier. So, in general, respiration is depressed during hypothermia, but the rate is almost in proportion to metabolic needs until respiration becomes shallow and erratic and eventually stops in the region of 25–30°C.

Effect of hypothermia on the blood Hypothermia also results in marked hematological perturbations such as increased viscosity, leukopenia (abnormally low number of circulating leukocytes), and thrombocytopenia (reduced numbers of platelets leading to bleeding tendencies). Increasing blood viscosity, which changes by as much as 5% per degree Celsius, also contributes to an apparent rise in vascular resistance (decreased efficiency of vascular perfusion). Deep hypothermia also hinders clotting by slowing the enzyme-mediated coagulation cascade and by promoting fibrinolysis. Abnormalities in clotting

mechanisms occur in 68–100% of patients subjected to clinical hypothermia, making bleeding tendencies a major concern in people subjected to prolonged hypothermia.

Physiological signs and symptoms of cold exposure Hypothermia is now recognized as the number one killer of outdoor recreationalists, and along with this has come the need for increasing awareness of the problem. As a consequence, many outdoor organizations are taking active measures to educate their members about hypothermia as an environmental hazard and are teaching emergency care for hypothermia. A critical part of this process is the ability to recognize the signs and symptoms of hypothermic exposure. [Table 2](#) illustrates some of the characteristic signs and symptoms at various stages of hypothermia.

Because survival after cold exposure is critically dependent on early intervention to reverse the effects of hypothermia, emergency personnel have become increasingly skilled over the past several years at managing and treating victims of hypothermia. It is beyond the scope of this article to describe the recommended treatments for victims of accidental hypothermia, but [Figure 1](#) illustrates an algorithm of recommended actions. It is noteworthy that some early interventions can be initiated in the field, while more aggressive treatments can not be considered until the patient is transferred to an adequately equipped emergency vehicle or hospital.

Enhancing tolerance to cold exposure The risk of accidental hypothermia is ever present for people who live at high altitudes, and efforts to protect against cold exposure have focused on methods of decreasing heat loss by improving insulation. Nevertheless, some research has been devoted to finding ways of improving cold tolerance by modifying the thermoregulatory response to cold. The principal strategies employed have been thermal acclimation, physical exercise, dietary enhancement of thermogenesis, pharmacological enhancement of thermogenesis, and manipulation of the thermoregulatory set point. While none of these approaches have been accepted routinely to improve cold tolerance, it can be concluded from some studies that the pharmacological enhancement of cold thermogenesis using ephedrine in combination with methylxanthine represents the most promising method for delaying the onset of hypothermia in humans.

Contrasting outcome between victims of accidental hypothermia and patients experiencing clinical hypothermia In considering the chances of survival after the stress of hypothermia, an important

Table 2 Characteristics of hypothermic stress at different stages of cooling

Stage	Core temperature	Signs and symptoms
Mild hypothermia	37–35°C (98–95°F)	Normal, shivering can begin, cold sensation, goose bumps, unable to perform complex tasks with hands, shiver can be mild to severe, hands numb
Mild–moderate hypothermia	35–33°C (95–93°F)	Shivering/intense, muscle incoordination becomes apparent, movements slow and labored, stumbling pace, mild confusion, may appear alert. Use sobriety test; if unable to walk a 30-foot straight line, the person is hypothermic
	33–32°C (93–90°F)	Violent shivering persists, difficulty speaking, sluggish thinking, amnesia starts to appear, gross muscle movements sluggish, unable to use hands, stumbles frequently, difficulty speaking, signs of depression, withdrawn
Moderate hypothermia	32–30°C (90–86°F)	Shivering stops, exposed skin blue or puffy, muscle coordination very poor, inability to walk, confusion, incoherent/irrational behavior, but may be able to maintain posture and appearance of awareness
	30–27°C (86–82°F)	Muscle rigidity, semiconscious, stupor, loss of awareness of others, pulse and respiration rate decrease, possible heart fibrillation
Deep hypothermia	27–25°C (82–78°F)	Unconscious, heart beat and respiration erratic, pulse may not be palpable
	<25°C (<77°F)	Pulmonary edema, cardiac and respiratory failure, death (death may occur before this temperature is reached)

Adapted from the Princeton University Outdoor Action Guide, with permission.

distinction has to be made between accidental hypothermia and clinical hypothermia. Even though the final core temperature of the individual may drop to the same level, the survival rate from clinically induced hypothermia is much higher than that of accidental hypothermia. In the latter case, homeotherms respond to hypothermic stress with immediate activation of heat production as well as inhibition of heat loss. The timing and intensity of the thermogenic response, as well as the duration over which the process can be maintained, depend on factors such as sex, age, and health of the individual. Hypothermia ensues when maximum thermogenesis fails to balance minimum heat loss and heat deficit results. As hypothermia progresses, cold narcosis of the central nervous system supervenes and respiratory support declines. Heart rate also slows with decreasing temperature, but the heart may continue to beat for some time depending upon the cooling rate. Eventually, when the core temperature drops to a specific level (in the range 15–25°C [60–77°F]), the heart either goes into ventricular fibrillation or arrests. In the battle to maintain core temperature, the subject has exhausted itself.

In contrast to the preceding scenario, clinically induced hypothermia is a sedate, relatively nonstressful, and intensively monitored condition. The patient is initially anesthetized and rapidly cooled, often with the aid of cardiopulmonary bypass support, with little or no metabolic resistance. Moreover, the patient is monitored continuously throughout so that presenting complications can be corrected expeditiously or rewarming employed before they become lethal. It is important to explain why hypothermia is sometimes induced electively in patients for clinical

benefit. As described in the next section, it has long been known that hypothermia is a highly effective modality for protecting cells against the deleterious effects of hypoxia and ischemia. In general, the extent of protection is directly related to the degree of hypothermia, with colder temperatures offering greater protection. This is based on the fundamental effect of cooling on biophysical and biochemical processes as molecular activity and mobility is slowed. In the fields of cardiovascular surgery, neurosurgery, and more recently in trauma medicine, there is often a need in complicated cases for the surgeon to implement a temporary cessation of blood flow to the heart or the brain. Adjunctive hypothermia is often used to protect the organs against hypoxia and ischemia. In recent years, there has been a renewed interest in the application of mild to moderate hypothermia as a potent neuroprotective modality, and consideration is being given to a potential therapeutic benefit of hypothermia to avert the complications of traumatic brain injury and stroke.

Cellular Responses to Hypothermic Stress

A complete understanding of the effects of hypothermia on the body calls for examination of the events at the cellular level, since this is where the pathophysiology of hypothermic stress is ultimately mediated. A great deal is known about the effects of cold on cells, since cooling has proved to be the foundation of nearly all effective methods of protecting and preserving cells and tissues for applications such as transplantation. Transplantation science calls for effective methods of preservation since donor cells, tissues, and organs are required to withstand a period of ischemia and hypoxia as part of any transplantation

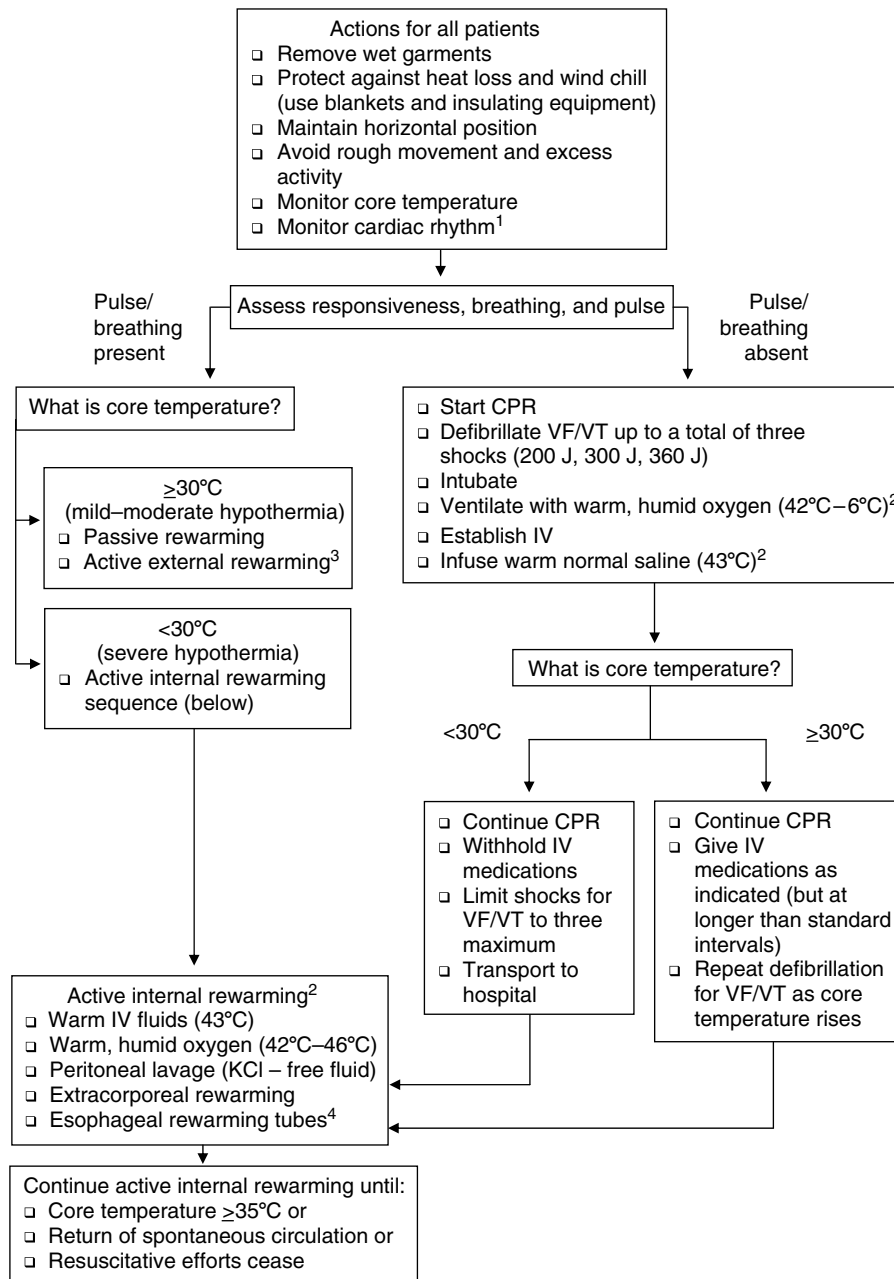


Figure 1 Algorithm for treatment of hypothermia. Notes: (1) This may require needle electrodes through the skin. (2) Many experts think that these interventions should be done only in hospital, though practices vary. (3) Methods include electric or charcoal warming devices, hot water bottles, heating pads, radiant heat sources, and warming beds. (4) Esophageal rewarming tubes are widely used internationally and should become available in the United States. Abbreviations: VF, ventricular fibrillation; VT, ventricular tachycardia; CPR, cardiopulmonary resuscitation; IV, intravenous. This algorithm is based on the Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiac Care published by the American Medical Association (*JAMA*, **268**, 2171, Copyright 1992, American Medical Association, all rights reserved) and is adapted from the article 'JAMA protocol for hypothermia' at <http://hypothermia.org/jama.htm>

procedure when the blood supply is temporarily interrupted. The basis of this hypothermic protection is that cooling can help to combat the deleterious effects of ischemia, but the consequences of cooling are not exclusively beneficial, such that hypothermic storage is a compromise between the benefits and detriments of

cooling. In the following, the detrimental effects only of hypothermic stress on cells is outlined.

General suppression of reaction rates The fundamental basis of all biological and chemical processes is molecular activity and mobility, which are

governed by thermal energy such that as temperature is lowered, molecular motion is slowed. The removal of heat from a system slows down both physical and chemical processes in proportion to the loss of heat and therefore to the fall in temperature. Since the processes of deterioration associated with ischemia and anoxia are mediated by chemical reactions, it has proved well founded to attempt to prevent or attenuate these changes by cooling. Biochemical processes involve molecular interactions that are invariably catalyzed by enzymes in reactions that require energy input from cellular stores such as ATP or creatine phosphate. Cooling can affect all components of these reactions, including the energy status of the substrate molecules, the stability of the enzyme protein, and the capacity of the cell to supply biological energy. The rate of biophysical processes such as diffusion of ions and osmosis declines linearly with temperature by approximately 3% per 10°C. It is apparent, therefore, that biophysical events are relatively little affected by the comparatively modest temperature changes imposed during hypothermic storage of tissues for transplantation. It is only at much lower temperatures that the rates of biophysical processes become significantly important, especially at sub-zero temperatures, when phase changes lead to both ice formation and solute concentration changes.

By comparison, the rate of chemical reactions, including the biochemical processes that constitute metabolic activity, is slowed significantly more by a given degree of hypothermic exposure. It is well established that within the temperature range 0–42°C, oxygen consumption in tissues decreases by at least 50% for each 10°C fall in temperature. The quantitative relationship between energy requirements for biochemical processes and temperature changes has been expressed mathematically in different ways:

1. The Arrhenius relationship: Biochemical processes, in common with all chemical reactions, occur only between activated molecules, the proportion of which in a given system is given by the Boltzmann expression, $\exp(-E/RT)$, where E is the activation energy, R is the gas constant, and T is the absolute temperature. According to the Arrhenius relationship, the logarithm of the reaction rate, k , is inversely proportional to the reciprocal of the absolute temperature:

$$-\log k = A(-E/2.3RT).$$

A graphical plot of $\log k$ against $1/T$ will yield a straight line with a slope of $E/2.3R$ if the relationship represents a single rate-limiting step.

2. The Van't Hoff rule relates the logarithm of a chemical reaction rate directly to temperature and is commonly expressed in the form of the respiratory quotient or temperature coefficient, Q_{10} , where Q_{10} is the ratio of reaction rates at two temperatures separated by 10°C. Accordingly,

$$Q_{10} = (K_2/K_1)10^{(T_2-T_1)}$$

For most reactions of biological interest, Q_{10} has a value between 2 and 3, but some complex, energy-dependent reactions have a Q_{10} between 4 and 6 and are more likely to stop completely at low temperatures. Both Q_{10} and Arrhenius plots have been used to quantitate changes in metabolic processes occurring in biological systems, whether they are enzyme reactions in single cells or the oxygen consumption of the entire human body. The Q_{10} for whole-body oxygen consumption is approximately 2, indicating that, in general, metabolic rate is halved for each 10°C drop in temperature.

Metabolic uncoupling While it is clear that cooling has a profound effect on biochemical reaction rates and that this in turn can slow degradative processes and reduce the rate of substrate and energy depletion, it is important to realize that not all reaction rates are affected to the same degree, or even in the same manner, by cooling. One hypothesis of chilling injury states that at a certain critical temperature within the chilling injury range, the membrane lipids undergo a transition from a liquid-crystalline to a solid gel state. The two main consequences of the transition thought to eventually result in cell injury are an increase in membrane permeability and an increase in the activation energy of membrane-bound enzymes such as those controlling respiration in mitochondria. Thus, the result would be the accumulation or depletion of metabolites at the point of entry into mitochondria. Hence, the membrane phase transitions in subcellular membranes could cause metabolic imbalance and provide one component of injury sustained by homeothermic cells during cold exposure.

Figure 2 shows schematically the large number of different, integrated chemical reactions that constitute common metabolic pathways within a cell. Each of these reactions will be affected by cooling in a different manner and to a different degree such that the possibility exists for uncoupling reaction pathways and producing harmful consequences. In preservation, cooling prolongs *in vitro* survival because it slows metabolism, reduces the demand for oxygen and other metabolites, and conserves chemical

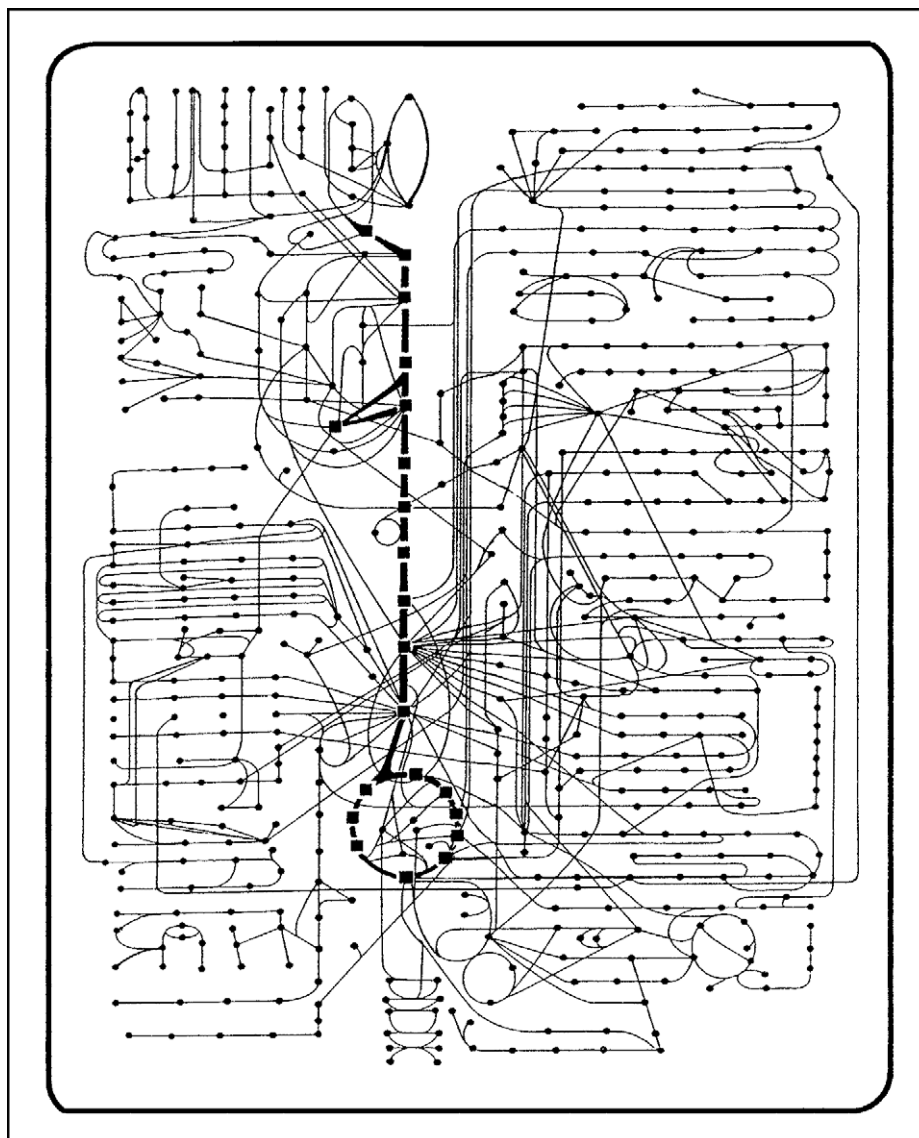


Figure 2 Schematic diagram illustrating the complexity of integrated biochemical pathways responsible for the structural and functional integrity of a typical cell. A healthy cell depends upon the synthesis and degradation of a wide variety of molecules. These processes are controlled by a complex series of biochemical reactions that produce and consume energy in a highly regulated manner. The diagram portrays schematically the complex integrated pathways that constitute the biochemical machinery of a typical cell. About 500 common metabolic reactions are shown by the connecting lines, with each metabolite represented as a filled circle. The centrally placed reactions of the main energy-producing pathways (glycolysis and tricarboxylic acid cycle) are shown as bold. A typical mammalian cell synthesizes more than 10 000 different proteins, a major proportion of which is enzymes. Each of these pathways can be affected differently by reduced temperatures such that the possibility exists for disruption of the normal metabolic processes, with harmful consequences as described in the text. Adapted from Alberts et al. (1994). *Molecular biology of the cell* (3rd edn.), Garland Publishing, with permission.

energy. However, it does not affect all reactions to the same extent, and the net result of cooling on an integrated, metabolizing system is complex, not entirely predictable, and not completely understood.

Studies of the effects of pure hypothermia on cell viability in the absence of any prior hypoxia have shown that the inactivation (killing) rates of cells exposed to reduced temperatures change slope at approximately 7–8°C, implying that there are distinct

mechanisms of hypothermic inactivation above and below this transition temperature. These values have been interpreted as suggesting that unbalanced metabolism is probably the rate-limiting step for hypothermic inactivation in the higher temperature range (8–25°C), and membrane lipid phase transition or cold denaturation of a critical protein is likely to be responsible for the strong temperature dependence in the lower range (<8°C).

Effect upon energy metabolism Under normal circumstances the supply of energy-rich compounds to fuel a cell's requirement for homeostatic control is continuously replenished by oxidative phosphorylation in the mitochondria. During cooling, however, there is a progressive exhaustion of chemical energy reserves in a cell despite the general suppressive effect of cooling on metabolism. The effect of cooling on metabolism is complex and should not be regarded as causing a simple uniform retardation of all biochemical reactions. For example, the complexity of low temperature effects on mitochondrial respiration is not limited to the impairment of translocase enzymes. It has been shown that other enzymes that control reactions of the tricarboxylic acid (TCA) cycle and the electron transport chain are affected differently by cold storage.

Effect upon ion transport and cell swelling Intracellular ionic composition and volume regulation of a cell are maintained by a pump-leak mechanism in which membrane-bound enzymes transport various ions and solutes to counter the passive diffusion driven by chemical potential gradients. These active pumps are inhibited by hypothermia both by its direct effect on enzyme activity and by the depletion of high-energy reserves as mitochondrial energy transduction fails. Nevertheless, the resultant passive redistribution of ions and water across the cell membrane and the concomitant change in the membrane potential have been demonstrated to be rapidly and fully reversible in the short term and may not be a source of permanent injury. Cells are known to have interrelated cation transport systems depending on energy supply and are thereby affected by reduced temperatures. Moreover, it is recognized that changes in the distribution of divalent cations (Ca^{2+} , Mg^{2+}) as a consequence of ischemia, hypoxia, and even cooling are especially important in cellular injury. The belief that calcium, in particular, is a mediator of cell death is based upon evidence accumulated over several decades from observations in a wide variety of tissues and pertains to cell death from a variety of causes. Cytosolic calcium concentration is 10 000-fold lower than extracellular concentrations, and it is now postulated that enhanced or unbalanced calcium influx across the plasma membrane represents a final common pathway in cell death mediated by various conditions that share the tendency to induce an abnormal membrane permeability for calcium. In recent years, the recognition of the central role of calcium-mediated effects in the death of cells of ischemically sensitive tissues such as heart and brain has led to intense scrutiny of the effects of perturbations in divalent cation homeostasis. Calcium functions as a

second messenger in the regulation of numerous biochemical and physiological processes, often by activating regulatory proteins, which, in turn, can activate many intracellular enzymes, including protein kinases. It has recently been proposed that the significant protective effect of mild hypothermia against ischemic brain injury might be mediated in part by inhibition of calcium-induced effects and the prevention of inactivation of important Ca-dependent enzyme systems such as the kinases. However, it is generally recognized that cooling has no beneficial effect in preventing, and may even accentuate, inhibition of the (Ca^{2+} , Mg^{2+})-ATPase pump mechanism, which in general is the most important factor in controlling intracellular Ca^{2+} concentrations. The effects of cooling on calcium homeostasis result not only from the inhibition of transmembrane pumping, but also from the inhibition of calcium sequestration by intracellular organelles, particularly the endoplasmic reticulum and mitochondria.

Proton activity changes The principal events of the ischemic cascade show that elevation of the concentration of protons, i.e., increasing acidity, is regarded as a contributory central event in the process of cellular injury ensuing from O_2 deprivation and energy depletion. Moreover, reduced temperatures are also known to influence pH regulation, which is another important homeostatic mechanism for cell survival. It has frequently been reported that hydrogen ion concentration increases in a variety of mammalian cells during hypothermic storage such that tissue pH has been recorded to fall to 6.5–6.8 within a few hours of cold. Acidity is widely recognized as a hazard for cells, with the accumulation of protons contributing to a variety of deleterious processes, including metabolic block of glycolysis and structural damage. Destabilization of lysosomes releasing harmful proteases and catalysis of oxidative stress by mobilization of free heavy metals have been implicated as mechanisms of cellular tissue injury during acidosis.

Effect of hypothermia on the generation of oxygen free radicals The emerging role of oxygen-derived free radicals (ODFRs) in tissue injury and their participation in reperfusion injury is well established. In essence, it is recognized that cooling increases the susceptibility of cells to produce free radicals and attenuates the natural defense mechanisms by which cells normally deal with the low-level free radical production in metabolism. Due to a lower activation energy, free radical reactions are depressed less by temperature reduction than the enzymatic processes used to scavenge them. It is the balance between production and removal of free radicals that is crucial

to the cell such that a combination of excessive radical generation and hindrance of normal defense mechanisms can redirect the processes in favor of injury. There is evidence that some of the natural defense mechanisms against free radicals become depleted during cold storage and that the addition of natural pharmacological scavengers including superoxide dismutase (SOD), catalase, or mannitol to cold storage media may improve the viability of hypothermically stored organs. It is clear, therefore, that in tissues exposed to cold there is potential for unbalanced free radical reactivity and elevated intracellular free calcium levels, and subsequent warming and reperfusion is likely to greatly potentiate these adverse events.

Structural changes The interrelationship between structure and function is fundamentally important for cellular homeostasis, which is governed by an intricate array of biochemical, physiological, and biophysical processes. Moreover, these processes are compartmentalized within the cell such that the structure of biological membranes and the cytoskeleton are integral components of cell viability and vitality.

The sensitivity of biological structures to temperature change is well known, and the thermal denaturation of proteins in particular is well documented. Thus, most proteins, as well as nucleic acids and many polysaccharides, are able to exist in their biologically active states only within a limited temperature range that is characteristic of the macromolecule and its environmental conditions such as ionic strength and pH. While a great deal more is known about heat denaturation of proteins at elevated temperatures, there are well-described examples of cold denaturation involving spontaneously unfolding of proteins or dissociation of the multisubunit structure into biologically inactive species that may or may not reassemble on rewarming to normal temperatures.

It has been postulated on the basis of measurements in both model systems and intact cells that a phase separation occurs within the plane of the membrane as cooling proceeds. Such cold-induced changes in the degree of membrane fluidity render it thermodynamically unfavorable for membrane proteins to remain in the gel phase such that they may be redistributed laterally into regions of low melting point lipids that remain in the liquid crystalline phase. This process is thought to result in packing faults due to the development of lipid-rich and protein-rich microdomains in a membrane undergoing phase transitions. One consequence of this could be a change in membrane permeability and alteration in the solute barrier function of the membrane. Phase separation of mammalian membranes has been demonstrated at temperatures

in the region of 10°C and below, and this correlates with the trend toward increased permeabilities to both ions and even large molecules such as proteins during cooling. In addition to phase separations, other forms of cold-induced membrane damage include the actual loss of membrane phospholipids, which is intuitively more deleterious than phase changes, which may in large part be reversible.

Cold-induced changes of cellular structural proteins that constitute cytoskeletal components such as microtubules have been known since the mid 1970s. This cold sensitivity appears to be mediated by depolymerization of component polypeptide units and in general is readily reversible upon rewarming. Nevertheless, the possibility exists for reassembly of the microtubules in a way that results in cellular abnormalities.

Thermal shock It has been discovered in some cells that the rate of cooling is also a determinant of injury; the phenomenon, which is not well understood, is referred to as thermal shock (cold shock) or chilling injury. The effects are thought to be related to the thermotropic properties of cell membranes involving phase transitions, as discussed briefly earlier. It has been proposed that rapid cooling in the absence of freezing can cause mechanical stresses on membranes induced by differential thermal contraction.

Cold shock proteins Changes in the structure of proteins are a common response to stress such that cells accumulate incorrectly folded proteins as a consequence of stresses such as hypoxia or temperature change. Proteins that experience unfolding or conformational changes do not remain soluble and are transformed into denatured proteins. It is now well established that one of the defense mechanisms adopted by cells to counteract such effects of stress is to synthesize new proteins commonly referred to as stress proteins or heat shock proteins (HSPs) because of their increased synthesis by many cell types after exposure to elevated temperatures. The general role of this group of highly conserved stress proteins is believed to be homeostatic in that they protect the cell against the harmful consequences of traumas and help promote a quick return to normal cellular activities once the shock has terminated.

The induction of stress proteins in response to cold (cold shock proteins) has been identified in invertebrate species such as bacteria and yeast, but it was not known until recently that similar families of cold shock proteins are also induced in mammalian cells.

Apoptosis vs. necrosis in cold-induced cell death It is now known that there are two distinct ways in

which cells may die. Necrosis is caused by a general failure of cellular homeostatic regulation following injury induced by a variety of deleterious stimuli, including hypoxia, toxins, radiation, and temperature changes. Apoptosis (programmed cell death) by contrast, is a regulated process that has been distinguished from necrosis by numerous morphological, biochemical and physiological differences. While many of the diverse stresses known to cause necrotic cell death have also been reported to induce apoptosis in a variety of cells, the role of low temperatures as a possible stimulus of programmed cell death has yet to be investigated in depth. Nevertheless, it has been shown that cultured mammalian fibroblast cells at the transition from logarithmic to stationary growth were killed by brief exposures to 0°C and rewarming at 37°C. Moreover, the kinetics of cell death served to distinguish cold-induced apoptosis from the lethal effects of longer-term cold storage (hours or days), and also showed characteristics inconsistent with direct chilling injury (cold shock). It is possible, therefore, that apoptosis may represent another manifestation of cold injury. Susceptibility to cold-induced apoptosis seems to depend critically upon cell growth cycle and the temperature of exposure. Recent studies in cultured hepatocytes and liver endothelial cells have indicated that reactive oxygen species (free radicals) might play a key role in mediating cold-induced apoptosis.

Finally, it is important to emphasize that both the severity and the reversibility of the catalog of effects outlined here are highly dependent upon both the duration and the extent of hypothermic exposure. Many of the consequences of hypothermic stress manifest in cells subjected to prolonged exposure during cold storage do not develop during brief exposure to mild or moderate hypothermia.

See Also the Following Articles

Heat Resistance; Heat Shock Proteins: HSP60 Family Genes; Hyperthermia.

Further Reading

Bristow, G. K. (1985). Clinical aspects of accidental hypothermia. In: Heller, H. C., Musacchia, X. J. &

- Wang, L. C. H. (eds.) *Living in the cold: physiological and biochemical adaptations*, pp. 513. New York: Elsevier.
- Collins, K. J. (1983). *Hypothermia: the facts*. New York: Oxford University Press.
- Fuller, B. J. and Grout, B. W. W. (eds.) (1991). *Clinical applications of cryobiology*. London: CRC Press.
- Kruuv, J. (1997). Survival of mammalian cells exposed to pure hypothermia in culture. In: Willis, J. (ed.) *Advances in molecular and cell biology* (vol. 19), pp. 143–192. Greenwich, CT: JAI Press.
- Lee-Chiong, T. L., Jr. and Stitt, J. T. (1996). Accidental hypothermia: when thermoregulation is overwhelmed. *Postgraduate Medicine* 99(1), 77–80, 83–84, 87–88.
- Lloyd, E. L. (1996). Accidental hypothermia. *Resuscitation* 32(2), 111–124.
- Lønning, P. E., Skulberg, A. and Åbyholm, F. (1986). Accidental hypothermia: review of the literature. *Acta Anaesthesiologica Scandinavica* 30, 601–613.
- Maher, J. and Hachinski, V. (1993). Hypothermia as a potential treatment for cerebral ischemia. *Cerebrovascular and Brain Metabolism Reviews* 5, 277–300.
- Maier, C. M. and Steinberg, G. K. (2003). *Hypothermia and cerebral ischemia: mechanisms and clinical applications*. Totowa, NJ: The Humana Press.
- Mercer, J. B. (1995). Enhancing tolerance to cold exposure – how successful have we been? *Arctic Medical Research* 54 (Supplement 2), 70–75.
- Pozos, R. and Born, D. (1982). *Hypothermia: causes, effects, and prevention*. Piscataway, NJ: New Century Publisher.
- Swan, H. (1973). Clinical hypothermia: a lady with a past and some promise for the future. *Surgery* 73, 736–758.
- Taylor, M. J. (1987). Physico-chemical principles of low temperature biology. In: Grout, B. W. W. & Morris, G. J. (eds.) *The effects of low temperatures on biological systems*, pp. 3–71. London: Edward Arnold.
- Taylor, M. J. (2006). Biology of cell survival in the cold: the basis for biopreservation of tissues and organs. In: Baust, J. G. & Baust, J. M. (eds.) *Advances in biopreservation* (chap. 2), pp. 15–62. Boca Raton, FL: CRC Press.
- Taylor, M. J., Elrifai, A. M. and Bailes, J. E. (1996). Hypothermia in relation to the acceptable limits of ischemia for bloodless surgery. In: Steponkus, P. K. (ed.) *Advances in low temperature biology* (vol. 3), pp. 1–64. London: JAI Press.
- Wilkerson, J. A., Bangs, C. C. and Hayward, J. S. (eds.) (1993). *Hypothermia, frostbite, and other cold injuries: prevention, recognition and pre-hospital treatment*. Seattle, Washington: Mountaineers Books.

Hypothyroidism

R T Joffe

McMaster University, Hamilton, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R T Joffe, volume 2, pp 496–499, © 2000, Elsevier Inc.

Introduction

Thyroid and Psychiatric Symptoms

Types of Hypothyroidism

Clinical Hypothyroidism

Subclinical Hypothyroidism

Thyroid in Psychiatric Illness

Glossary

Clinical hypothyroidism A disorder characterized by decreased production of thyroxine and triiodothyronine by the thyroid gland. It leads to a well-recognized clinical syndrome, of which psychiatric symptomatology is a prominent component.

Depressive symptoms A group of mood, cognitive, and physical symptoms that commonly occur, but the number of symptoms are insufficient to qualify for the diagnosis of major depression.

Major depression A characteristic syndrome that consists of mood, physical, and cognitive symptoms. This is one of the commonest psychiatric illnesses and is amenable to treatment with antidepressants and specific psychotherapies.

Subclinical hypothyroidism A condition characterized by normal levels of circulating thyroxine and triiodothyronine but elevated levels of thyrotropin (TSH). There is also an enhanced TSH response to thyrotropin-releasing hormone (TRH). A well-described clinical syndrome is usually not evident.

Thyrotropin (TSH) A hormone secreted by the pituitary that acts on thyroid cells to stimulate the production of thyroxine and triiodothyronine; also called thyroid-stimulating hormone.

Thyrotropin-releasing hormone (TRH) A tripeptide hormone released by the hypothalamus that stimulates the release of TSH from the pituitary. TRH and TSH play an important role in the regulation of thyroid hormone secretion, both through a stimulatory effect on the thyroid and by negative feedback regulation of circulating thyroid hormone

levels on the hypothalamus and pituitary. In this manner, appropriate thyroid hormone levels are maintained in the circulation.

Thyroxine (T_4)

One of two main circulating thyroid hormones produced by the thyroid gland. Although it is the major secretory product of the thyroid gland, this hormone is thought to exert its physiological action by conversion to triiodothyronine.

TRH test

A diagnostic test that involves the intravenous administration of TRH and the measurement of TSH at 15-min intervals for 90 min. In hypothyroidism, the TSH response to TRH is exaggerated, consistent with the high output of both TRH by the hypothalamus and TSH by the pituitary in response to a failing thyroid and decreased circulating T_4 and triiodothyronine levels. With the advent of new highly sensitive tests for the measurement of TSH, the TRH test is becoming obsolete.

Triiodothyronine (T_3)

One of two main circulating thyroid hormones produced by the thyroid gland. Only approximately 15–20% of circulating T_3 is secreted directly by the thyroid; the remainder comes from conversion from T_4 . T_3 is approximately four times more metabolically active than T_4 .

Introduction

Thyroid diseases, both hyperthyroidism and hypothyroidism, are among the commonest endocrine disorders seen in clinical practice. Excess or diminished production of thyroid hormone leads to the clinical syndromes of hyperthyroidism and hypothyroidism, respectively. A systematic evaluation of thyroid function tests helps verify the role of thyroid disease in the production of this symptomatology. This article reviews the clinical and biochemical features of hypothyroidism and, in particular, its relationship to psychiatric symptomatology.

Thyroid and Psychiatric Symptoms

It has long been known that there is an association between alterations of thyroid function and psychiatric symptomatology. Clinical thyroid diseases are commonly associated with psychiatric symptomatology, particularly alterations of mood and

cognition. In particular, the association between hypothyroidism and depression has been documented consistently in the medical literature for more than 100 years.

Both hyperthyroidism and hypothyroidism may present with psychiatric symptomatology. In hyperthyroidism, a broad range of psychiatric symptomatology may present, but in hypothyroidism, depression and alterations of cognition are observed most commonly. Symptoms of anxiety and even psychosis are much less frequent. The observation of this very strong relationship between thyroid dysfunction and psychiatric symptomatology has led to the conclusion that thyroid hormones or the thyroid axis may be important in the pathophysiology of primary psychiatric disorders, especially major depression and the dementia disorders. As a result, it is common clinical practice to do a routine thyroid screening in patients who present with psychiatric symptoms, particularly disorders of mood or cognition. The best evidence is that most patients with depressive or dementia illnesses are euthyroid, and it has even been questioned whether routine thyroid screening of psychiatric patients is cost effective. However, this practice prevails because it does present an important opportunity, however uncommon, to identify an easily reversible major psychiatric disorder associated with hyper- or hypothyroidism.

Types of Hypothyroidism

Hypothyroidism is categorized according to the degree of thyroid failure. Since the mid-1970s, the definition of various degrees of hypothyroidism has become standardized. Grade I, or clinical, hypothyroidism is characterized by clinical symptoms of thyroid underactivity and is associated with abnormally decreased levels of thyroxine (T_4) and triiodothyronine (T_3), elevated basal thyrotropin (TSH) levels, and an enhanced TSH response to thyrotropin-releasing hormone (TRH). In grade I hypothyroidism, clinical symptoms of thyroid hormone deficiency are clearly evident. In grade II hypothyroidism, peripheral T_4 and T_3 levels are within normal limits, but both basal and TRH-stimulated TSH levels, using the TRH test, are elevated abnormally. Patients with this type of hypothyroidism do not have the classic symptoms of clinical hypothyroidism. Grade III hypothyroidism is characterized by no clinical features of hypothyroidism, normal circulating levels of T_4 , T_3 , and basal TSH, but an abnormally elevated TSH response to TRH. Each of these subtypes of hypothyroidism is considered separately here.

Clinical Hypothyroidism

Clinical hypothyroidism is a relatively uncommon disorder affecting approximately 1% of the population. However, its prevalence is influenced strongly by both gender and age; it is substantially more common in women and also increases in prevalence with advancing age. Therefore, the prevalence of clinical hypothyroidism in women over the age of 60 is approximately 4%. The features of clinical hypothyroidism are often vague and nonspecific. They include symptoms such as facial puffiness, dry skin, hair loss, myalgia, paresthesias, reduced tendon reflexes, cold intolerance, constipation, and fatigue. However, psychiatric symptoms occur in the majority of patients and are reasonably consistent for diagnosis. They include depressive symptoms or even a major depressive syndrome and, less commonly, evidence of cognitive deterioration. Anxiety and psychotic symptoms are much less common. The consistency of the psychiatric presentation in clinical hyperthyroidism is remarkable and highlights a need to exclude clinical hypothyroidism as a primary diagnosis in all patients presenting with depression or dementia, particularly in the elderly. These psychiatric syndromes associated with clinical hypothyroidism are generally not amenable to treatment with psychotropic drugs such as antidepressants and require correction of the underlying thyroid disease. However, in a minority of patients, psychiatric symptoms may persist even with correction of the underlying endocrine disorder, and further intervention with psychotropics may be necessary.

The diagnosis of clinical hypothyroidism is established on the basis of the clinical history as well as the measurement of T_4 , T_3 , and TSH. The specific treatment is dependent on the underlying pathology. However, thyroid replacement therapy is necessary to restore normal thyroid function.

Subclinical Hypothyroidism

This refers to both grade II and grade III hypothyroidism, as defined earlier. These abnormalities of thyroid function are usually detected on routine thyroid screening and are not associated with any particular clinical syndrome. Approximately 5–15% of depressed patients have evidence of subclinical hypothyroidism. These cases of subclinical hypothyroidism may be associated with a variety of causes, particularly iodine deficiency, autoimmune thyroid disease, and, less commonly, ablation treatment of clinical hyperthyroidism. The medical consequences of subclinical hypothyroidism are uncertain. Most studies have focused on grade II hypothyroidism, and these

suggest that patients may have alterations in some physiological parameters such as myocardial contractility that may be reversed by thyroid hormone replacement therapy. Other studies have suggested that grade II hypothyroidism may be a risk factor for coronary artery disease due to alterations in serum lipoproteins. Although the clinical implications of grade II hypothyroidism require further clarification, it is known that between 5 and 10% per year of patients with subclinical hypothyroidism progress to clinical hypothyroidism.

The psychiatric consequences of subclinical hypothyroidism also remain to be clarified. As noted previously, approximately 5–15% of depressed patients have grade II or III subclinical hypothyroidism, although it is unclear whether this differs from the prevalence in the general population. Regardless, there is increasing evidence that subclinical hypothyroidism represents an increase in the vulnerability to depression and may also make patients more refractory to standard antidepressant therapies. Furthermore, if subclinical hypothyroidism is identified in a patient with depression, it is still a matter of controversy whether the depression in fact represents a clinical manifestation of the endocrine abnormality and, therefore, necessitates replacement therapy with thyroid hormone. Although the use of thyroid replacement therapy in subclinical hypothyroidism requires further study, the presence of treatment refractoriness may justify the use of thyroid hormones in order to enhance the response to antidepressants and, therefore, reduce the risk of chronicity and/or recurrence that are commonly associated with a major depressive illness.

Thyroid in Psychiatric Illness

As noted previously, the vast majority of patients with depression are euthyroid and have no evidence of thyroid dysfunction. Clinical hypothyroidism is extremely rare in major depressive illness, although 5–10% of patients with major depression have evidence of subclinical hypothyroidism. In fact, in the majority of patients with primary major depression, there is evidence of slight thyroid axis overactivity. It is suggested that primary major depression may be associated with a stimulation of TRH, which in turn leads to increased release of TSH and, therefore, stimulates the thyroid gland to produce more T_4 and T_3 . The most common alterations of thyroid function seen in primary major depression are slight elevations in measures of T_4 , normal circulating levels of T_3 , and blunted TSH responses to TRH. These alterations of thyroid function tests are suggestive of TRH stimulation. Although these alterations of thyroid function

tests are seen when groups of depressed patients are compared to either psychiatric or healthy controls, thyroid hormone levels in depressed patients are usually within normal limits.

Despite the fact that thyroid function is rarely decreased in depressed patients and in fact may be relatively increased, there is a long-standing practice of using thyroid hormones to treat depressive illness. Thyroid hormones used alone have not shown particular efficacy in treating depressive disorders, but T_3 in particular has been shown to augment therapeutic response when added to the antidepressant regimen in patients who have had no or only a partial response to treatment. The limited evidence to date suggests that T_3 may be more effective than T_4 in this regard. Although this remains to be clarified in future studies, the convention has been to use T_3 instead of T_4 in such cases, in contrast to thyroid replacement therapy for clinical hypothyroidism in which T_4 instead of T_3 is the usual treatment. There is insufficient evidence for the efficacy of either TRH or TSH in the treatment of depression. Although T_3 is gaining increasingly widespread acceptance as a treatment for resistant depression, its mechanism of action is unknown. However, it appears clear that it is not used to correct either clinical or subclinical hypothyroidism in the majority of cases.

See Also the Following Articles

Hyperthyroidism; Major Depressive Disorder; Thyroid Hormones.

Further Reading

- Barzel, U. S. (1995). Hypothyroidism: diagnosis and management. *Clinical Geriatric Medicine* 11, 239–249.
- Bilezikian, J. P. and Loeb, J. N. (1983). The influence of hyperthyroidism and hypothyroidism on alpha- and beta-adrenergic receptor systems and adrenergic responsiveness. *Endocrine Review* 4, 378–388.
- Bunevicius, R., Kazanavicius, G., Zalinkevicius, R., et al. (1999). Effects of thyroxine as compared with thyroxine plus triiodothyronine in patients with hypothyroidism. *New England Journal of Medicine* 340, 424–429.
- Gold, M. S., Pottash, A. L. C. and Extin, I. (1981). Hypothyroidism and depression: evidence from complete thyroid function evaluation. *Journal of the American Medical Association* 245, 1919–1922.
- Haggerty, J. J., Jr. and Prange, A. J., Jr. (1995). Borderline hypothyroidism and depression. *Annual Review of Medicine* 46, 37–46.
- Jackson, I. (1998). The thyroid axis and depression. *Thyroid* 8, 951–959.
- Joffe, R. T. and Levitt, A. J. (1993). *The thyroid axis and psychiatric illness*. Washington, DC: American Psychiatric Press.

Hypoxia, Role in Stress <i>See: Pressures, Effects of Extreme High and Low.</i>
--

Hysteria

K Pajer

The Ohio State University, Columbus, OH, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by K Pajer, volume 2, pp 500–502, © 2000, Elsevier Inc.

Definition and Historical Background
Treatment of Hysteria-Related Syndromes

Glossary

<i>Conversion disorder</i>	A psychiatric syndrome characterized by disturbances of bodily functions that do not correspond to known pathophysiological mechanisms.
<i>Dissociation</i>	The splitting of consciousness in an effort to repress painful memories; may result in either physical symptoms (somatization or conversion disorder), unconscious creation of alternative identities (multiple personality disorder), or inability to remain focused on present reality (borderline personality disorder).
<i>Free association</i>	A technique designed to uncover unconscious feelings. The patient lies down or sits comfortably and says anything that comes into his or her mind. Neither the patient nor the therapist makes any attempt to censor the thoughts, but the therapist listens and then interprets the associations.
<i>Repression</i>	Forgetting painful or frightening experiences. Repression is not a conscious process but, rather, an unconscious one; that is, the patient does not know that he or she has forgotten an event.
<i>Somatization disorder</i>	A psychiatric disorder in which the patient has many somatic complaints, but no underlying medical problem can be identified.

Definition and Historical Background

Clinical Vignettes

Patient 1 S.M. was 21 years old and in good health until she was out with her fiancé and he was hit by a swerving cab as he tried to push S.M. to safety. He died 3 days later. S.M. awakened on the day of the funeral to discover that she could not walk. She reported that she was paralyzed and could not walk or sit up. Her family tried to help her out of bed, and she simply collapsed. Her parents were quite upset, but she seemed relatively calm and told them it did not matter. She urged them to go to the funeral for her because she could not move. They consulted a prominent neurologist who could not find any abnormalities.

Patient 2 G.H. was 27 years old and sought treatment from a urologist for pain on urination and difficulty initiating urination. The urologist examined her and noted a urethral stricture that he treated with good success. Six weeks later, however, the patient called and said she was having pain in her pelvic area. He reexamined her, but could find nothing. The urologist suggested she consult her gynecologist, which she did. The gynecologist could not find any problem, but gave the patient some exercises to do to strengthen her pelvic muscles. The pain gradually subsided. However, the patient called 3 months later because of difficulties with her “bowels and stomach.” The gynecologist referred her to a gastroenterologist who did a complete workup, but could not find any abnormalities. During the evaluation, G.H. did tell the gastroenterologist that she had recently begun her first sexual relationship and was having a great deal of difficulty with this. The gastroenterologist referred her to a psychiatrist, who worked with G.H. for several months before she finally elicited a history of severe and repeated sexual abuse that had occurred during childhood. This abuse had produced

the urethral stricture and was now causing sexual and psychological difficulties for the patient as she tried to function in a normal adult romantic relationship.

Historical Changes in the Meaning of Hysteria

The word hysteria is familiar to most clinicians and lay people, but is elusive in definition. It is no longer used in psychiatric or psychological nomenclature, but, surprisingly, it was once the cornerstone of modern psychiatry. Several scholars of the history of psychiatry have proposed that the epidemic of hysteria afflicting American and British women in the late nineteenth and early twentieth centuries spawned the field of psychoanalysis. However, the disorder has now virtually disappeared. Furthermore, the term hysteria, once the focus of intensive treatment efforts and research, can scarcely be found in modern mainstream psychiatric literature.

The symptoms of hysteria can be found as early as 1900 BC in the Egyptian papyri (the actual word hysteria is derived from *hystera*, Greek for uterus). In both Egyptian and Greek writings, this multisymptom complex described in women was thought to be caused by a wandering uterus. Treatments were directed at coaxing the uterus to return to the pelvis and keeping it there once it returned. Assuming that sexual intercourse and childbirth would keep the uterus from moving, the most common prescription was for marriage and coitus.

Subsequent centuries brought changes in the definition and hypothesized etiology of hysteria. The definition became quite broad, referring to nearly any symptom of disease that was found more often in women than in men. Etiological explanations continued to focus on uterine pathology but also included causes such as satanic possession.

The nineteenth century brought a dramatic increase in the incidence of hysteria, although it is not clear why this happened. Treatments continued to focus on the uterus and newly discovered ovaries. The cauterization of the cervix, leeches applied to the cervix and vagina, and the removal of the ovaries were all common treatments that were only moderately successful. In the late nineteenth century, however, hypnosis was discovered to be useful for curing some symptoms of hysteria, such as the paralysis described earlier in S.M. Janet, Freud, and others believed that the mechanism of this cure was that hypnosis allowed the therapist, and then later the patient, to gain access to repressed feelings. In exploring these hidden feelings, buried conflicts could be resolved and the symptoms would disappear. The use of hypnosis diminished as Freud developed the technique of free association with interpretation by

the analyst. This technique became the hallmark of psychoanalysis and was eventually used for many psychiatric disorders.

The *Diagnostic and Statistical Manual* 4th edn. (DSM-IV) has reclassified hysteria into two other disorders in the somatoform disorders category: conversion disorder (CD) and somatization disorder (SD). CD is a psychiatric syndrome characterized by disturbances of bodily functions (e.g., blindness, hearing loss, paralysis, and lack of sensation) that do not correspond to known pathophysiological mechanisms. SD is a psychiatric disorder in which the patient has many different somatic complaints but no underlying medical problem can be identified.

As with hysteria, these disorders are more common in women. However, we now know from clinical and research data on combat and natural disasters that men may also present with unusual physical symptoms that do not have medical explanations. The primary risk factor for both disorders in men and women is trauma or extremely stressful events over which the patient has no control. The trauma may be acute, as in the case of S.M. or exposure to battle, or chronic and repressed, as in the case of G.H., who had been the victim of repeated sexual abuse. She had remained asymptomatic until her first sexual relationship reawakened bodily sensations associated with a time of great stress earlier in her life.

Overwhelming stress in the context of helplessness seems to be the key feature of situations that produce SD or CD. However, studies of the function of either the sympathoadrenal or hypothalamic-pituitary-adrenal (HPA) axis components of the stress-response system have not demonstrated consistent abnormalities. It does appear that patient with SD are more subjectively sensitive to pain, with greater stress-response system reactivity. The multiple illnesses in SD may be due to a sensitization of the brain cytokine system by past adverse experiences such that it responds with a sickness pattern to a lower threshold of stress than is usually required in other people.

Treatment of Hysteria-Related Syndromes

Conversion Disorder

A complete medical and psychiatric workup is essential to rule out any medical pathology that merits treatment. In some cases, even an acute trauma is not known to the physician or psychologist (e.g., rape or an abortion of a hidden pregnancy). In this case, repressed trauma can sometimes be discovered with an interview that takes place when the patient is

under the influence of amobarbital or lorazepam. On occasion, hypnosis is still used. The symptoms of a CD remit spontaneously in the majority of cases. The establishment of a nonjudgmental therapeutic relationship seems crucial for the treatment of patients whose symptoms do not subside quickly. With this type of a relationship, the exploration of latent trauma can be accomplished and healthier ways to cope with stress can be taught.

Somatization Disorder

As with CD, patients with suspected SD should receive complete medical workups. The primary danger in treating this disorder is that the patient can receive many invasive evaluations that are both expensive and disruptive to her life before someone notices the pattern of multiple, changing somatic complaints. There are three strategies for treating this disorder. The first is to control the frequent doctor shopping. This is accomplished most often by establishing a single physician to manage all aspects of care. This person keeps track of all medical care and schedules frequent contact with the patient. Second, some patients have had a positive response to antidepressant medication. The final treatment strategy is psychotherapy, focusing on any early trauma and the effect it may have on the patient's current ability to cope with stress. The prognosis for SD is not good unless the patient responds to an antidepressant or is motivated to embark on a course of psychotherapy.

See Also the Following Articles

Autonomic Nervous System; Combat Reaction, Chronic; Combat, Acute Reactions to; Corticosteroids and Stress; Posttraumatic Stress Disorder, Neurobiology of; Psychosomatic Medicine; Somatic Disorders; Trauma and Memory.

Further Reading

- Dantzer, R. (2005). Somatization: a psychoneuroimmune perspective. *Psychoneuroendocrinology* 30, 947–952.
- Janet, P. (1920). *The major symptoms of hysteria*. New York: Macmillan.
- Micale, M. (1989). Hysteria and historiography: a review of past and present writings, I–II. *History of Science* 27, 223–261, 319–351.
- Rief, W. and Barsky, A. J. (2005). Psychobiological perspectives on somatoform disorders. *Psychoneuroendocrinology* 30, 996–1002.
- Rief, W., Shaw, R. and Fichter, M. M. (1998). Elevated levels of psychophysiological arousal and cortisol in patients with somatization syndrome. *Psychosomatic Medicine* 60, 198–203.
- Scaer, R. C. (2001). The neurophysiology of dissociation and chronic disease. *Applied Psychophysiology and Biofeedback* 26, 73–91.
- Stone, J., Smyth, R., Carson, A., et al. (2005). Systematic review of misdiagnosis of conversion symptoms and “hysteria.” *British Medical Journal* 331(7523), 989–995.

Immobilization Stress

R Kvetnansky

Slovak Academy of Sciences, Slovak Republic

R McCarty

Vanderbilt University, Nashville, TN, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R Kvetnanský and R McCarty, volume 2, pp 503–506, © 2000, Elsevier Inc.

Variations in Immobilization
Advantages and Disadvantages
Acute versus Chronic Immobilization
Immobilization-Induced Gene Expression
Inventory of Neurohumoral and Genetic Responses to
Acute Immobilization
Summary and Conclusion

Glossary

Immobilization A model of stress that has been employed in studies of laboratory rats and mice. The procedure involves taping the four limbs of a rat or mouse to mounts secured to a metal frame using hypoallergenic tape. A pair of metal loops attached to the frame limits the range of motion of the animal's head. The frame, loops, and mounts are fabricated of stainless steel. The duration of a single episode of immobilization usually varies from 5 to 120 min or more. In addition, animals in chronic stress protocols may be immobilized each day for many weeks even months.

Restraint A model of stress whereby laboratory rats or mice are restricted in their range of movement. Each animal is secured in a plastic tube or a wire-mesh container that is small enough to prevent the animal from moving about or turning. The

frequency and duration of restraint stress vary depending on the questions being addressed by a given study, but a permanent restraint could last for many weeks.

Immobilization was introduced as a model of stress in animal experimentation in the mid-1960s. Since that time, Kvetnansky and colleagues have developed a standard protocol for immobilizing laboratory rats that involves securing the four limbs of a rat to mounts on a metal frame with hypoallergenic tape. The range of motion of an animal's head is limited by rings that are fixed to the metal frame. An acute episode of immobilization typically lasts from 5 to 120 min. In addition, animals can be immobilized daily (usually for 120 min) for many weeks to study the biological effects of adaptation to repeated immobilization stress.

During the first exposure to immobilization stress, laboratory rats or mice typically struggle vigorously to free themselves for 20–30 min. After that time, they will often cease struggling and remain motionless until a disturbance occurs. When one animal begins to struggle or if there is a sudden loud noise or some other form of sensory stimulation, many of the animals in the group that is being immobilized will begin to struggle again. Animals that have been immobilized daily for many days display signs of habituation to the stressor and calm down much sooner than naïve animals being immobilized for the first time.

Immobilization is a complex stressor and includes physical as well as psychological dimensions. The struggling and associated muscular exertion that occur during immobilization represent an intense form of physical exercise. After the first immobilization exposure, animals are so exhausted that they hardly move. In addition, the limited range of movement and exposure in an open area are significant sources of psychological stress to laboratory animals. It is also

quite common for repeatedly immobilized adult rats to lose a significant amount of body weight over a 1- to 4-week period as a result of an increase in physical training and a decrease in food intake. After that period, repeatedly immobilized animals begin to gain weight at a rate that is similar to unstressed controls. However, the absolute body weights of stressed rats remain below those of control rats.

Variations in Immobilization

Since its introduction as a laboratory stressor, immobilization has been modified and adapted for the unique needs of different laboratories. Unfortunately, it is all too easy to assume that all models of stress involving immobilization are the same. In fact, it is very difficult to compare studies unless the technique for immobilization is comparable. For example, many investigators have used restraint as a model of stress. Restraint may involve placing the test animal into a plastic tube or a wire-mesh container that is well ventilated. Although the animal's range of movement is severely limited, the limbs are not secured and the animal remains within an enclosed area. Many have argued that restraint stress is less intense a stressor than immobilization, and comparative studies using neural and endocrine measures have supported this assertion.

Immobilization may also be combined with other stressors, such as cold temperatures or even placing the immobilization board in shallow water such that the animal is partially submerged. Caution should be exercised in comparing such studies with others because the addition of a second stressor affects an animal's physiological and behavioral responses.

Immobilization has also been employed in the development of an animal model of long-term weightlessness. This model involves the continuous restriction of movement of laboratory rats for as long as 60 days to simulate some of the physiological changes that attend long-term space flight. The hypokinesia apparatus is modified such that the animal can gain access to food and water. It bears some similarity to studies with humans that have included long-term bed rest as a model of weightlessness associated with space flight.

Advantages and Disadvantages

Without doubt, immobilization has evolved into one of the most frequently employed models of stress in studies with laboratory rats and, recently, mice. There are several reasons for this. No special equipment is required other than the immobilization boards, which can be fabricated of stainless steel in a departmental

or institute shop and used for many years (the authors have used the same boards for more than 30 years). Once the animals have been immobilized, the investigator is free to do other things in the laboratory until the end of the stress session for that day. In addition, the stressor is sufficiently intense such that most stress-responsive biological systems are altered and their regulatory processes may be studied ([Tables 1 and 2](#)). Immobilization is well tolerated by laboratory rats and mice and can be used in studies of chronic intermittent stress over weeks or even months. Finally, there is extensive literature on biological responses to immobilization that may be consulted when an investigator is planning an experiment that involves using an acute or a repeated stressor.

The immobilization model of stress does have its limitations, however. The intensity of the stressor cannot be altered, as we can do with foot-shock stress or cold stress, for example. In addition, measures of behavioral responses to immobilization stress are not usually an element of the experimental protocols. This limits the use of this model of stress when investigators have a specific interest in relating physiological and behavioral responses to the same stressful stimulus.

Acute versus Chronic Immobilization

Immobilization stress has provided an especially useful model for exploring the molecular and physiological adaptations that occur when animals are exposed to a regimen of chronic daily stress. Although immobilization is an intense stressor, considerable evidence shows that many neural and endocrine systems do exhibit decreases in the magnitude of their responses following repeated daily bouts of immobilization stress. These habituated responses are also characteristic of other laboratory stressors (e.g., restraint, foot shock, swim stress, handling, and open field testing) that are less intense than immobilization.

In laboratory rats that are exposed to repeated daily episodes of immobilization stress, there is an enhanced response of many neural and endocrine systems if the same animals are exposed to a novel stressor. The proper control group to demonstrate this sensitization effect is a group of naïve control rats with no prior history of experimenter-controlled stress that is exposed to the novel stressor for the first time.

Immobilization-Induced Gene Expression

In studies of transcriptional regulation of genes involved in neurotransmitter and hormone biosynthesis, immobilization has served as a valuable stress

Table 1 Selected neurochemical and neurohumoral responses of laboratory rats to acute immobilization stress^a

Variable	Baseline	Stress	Peak time (min)
<i>Norepinephrine</i>			
Adrenal content (μg/adrenal)	3.2 ± 0.2	2.0 ± 0.3	120
Urine levels (μg/kg/24 h)	1.8 ± 0.2	4.9 ± 0.3	30
Plasma levels (pmol/ml)	1.25 ± 0.3	7.3 ± 1.2	20
<i>Epinephrine</i>			
Adrenal content (μg/adrenal)	17.5 ± 0.6	14.6 ± 0.4	120
Urine levels (μg/kg/24 h)	0.8 ± 0.06	2.6 ± 0.4	120
Plasma levels (pmol/ml)	0.20 ± 0.03	15.5 ± 2.7	5
<i>Plasma levels of:</i>			
DOPA (pmol/ml)	3.3 ± 0.2	3.9 ± 0.4	5
Dopamine (pmol/ml)	0.8 ± 0.09	1.25 ± 0.4	20
Dihydroxyphenylglycol (pmol/ml)	3.2 ± 0.3	9.3 ± 1.0	20
Methoxyhydroxyphenylglycol (pmol/ml)	15 ± 2.0	62 ± 15	60
Dihydroxyphenylacetic acid (pmol/ml)	5.2 ± 0.6	13.5 ± 0.2	20
Homovanillic acid (pmol/ml)	22 ± 4	55 ± 6	20
Plasma acetylcholinesterase (nmol/min/ml)	71 ± 5	112 ± 8	120
Plasma renin activity (ng/min/h)	4.4 ± 0.9	29.3 ± 4.7	10
Plasma aldosterone (pg/ml)	970 ± 90	3190 ± 460	120
Vasopressin (fmol/ml)	1.9 ± 0.2	6.1 ± 2.3	2
ACTH (pg/ml)	44 ± 4	790 ± 225	30
Corticosterone (μg/100 ml)	3.9 ± 0.2	24.5 ± 2.7	120
Prolactin (ng/ml)	7.2 ± 0.5	71 ± 17	20
<i>Microdialysate norepinephrine (pg/ml)</i>			
Paraventricular nucleus	178 ± 17	475 ± 37	30
Bed nucleus of the stria terminalis	92 ± 12	360 ± 38	30
Central nucleus of the amygdala	151 ± 8	425 ± 53	30

^aValues are means ± SEM. All data on plasma levels of various substances were obtained from chronically catheterized animals. All data on microdialysate norepinephrine levels were obtained from animals that were surgically prepared with chronic microdialysis probes. All values are estimates of data presented in figures of published papers that employed immobilization stress in studies using laboratory rats.

Table 2 Quantitative changes in gene expression of catecholamine biosynthetic enzymes during acute immobilization stress^a

Variable	Baseline	Stress	Peak time (h)
Adrenal medulla			
TH mRNA (amol/ng RNA)	0.46 ± 0.06	3.10 ± 0.56	3 after IMMO
DBH mRNA (amol/ng RNA)	6.00 ± 0.84	15.20 ± 1.57	3 after IMMO
PNMT mRNA (amol/ng RNA)	28.00 ± 2.36	141.0 ± 20.8	3 after IMMO
Stellate ganglia			
TH mRNA (amol/ng RNA)	0.017 ± 0.002	0.038 ± 0.005	24 after IMMO
DBH mRNA (amol/ng RNA)	2.60 ± 0.20	6.45 ± 0.61	24 after IMMO
PNMT mRNA	Present; not enough for quantification		
Brain stem			
TH mRNA (amol/μgRNA)			
Locus coeruleus	75.0 ± 8.7	268.0 ± 34.0	24 after IMMO
NTS (A2)	4.6 ± 1.0	9.6 ± 1.1	24 after IMMO
A1 area	17.0 ± 2.0	109.0 ± 11.0	3 after IMMO
Hypothalamus			
TH mRNA (amol/μg RNA)			
Paraventricular n.	5.2 ± 0.6	14.0 ± 1.8	3 after IMMO
Supraoptic n.	3.1 ± 0.9	10.4 ± 2.0	24 after IMMO
Dorsomedial n.	0.8 ± 0.1	5.5 ± 1.3	24 after IMMO
Arcuate n., lateral	8.0 ± 1.4	12.0 ± 0.6	24 after IMMO
Arcuate n., medial	11.0 ± 2.3	18.0 ± 2.5	3 after IMMO

^aValues are means ± SEM. Values are from data presented in figures of published papers that employed immobilization stress in rats. DBH, dopamine-B-hydroxylase; IMMO, immobilization; n., nucleus; PNMT, phenylethanolamine N-methyltransferase; TH, tyrosine hydroxylase.

model. Kvetnansky, Sabban, Wong, and co-workers have examined changes in the transcription rates of genes for enzymes involved in catecholamine biosynthesis in central and peripheral tissues. These include genes for tyrosine hydroxylase (TH), dopamine- β -hydroxylase (DBH), and phenylethanolamine-N-methyltransferase (PNMT) (Table 2). For example, their work has shown that adrenal medullary TH mRNA concentration (expressed in amol ng⁻¹ total RNA) increase approximately sixfold above control levels following a single 2-h period of immobilization. This increase is transient and levels of TH mRNA return to control levels within 24 h. However, if rats are immobilized for 2 h per day for 2–7 consecutive days, TH mRNA levels increase as much as sixfold above basal levels and these elevated levels are maintained up to 24 h later. The concentrations of DBH and PNMT mRNA in the adrenal medulla are 15–70 times higher, but are further elevated during immobilization stress. These increases are transient and within 24 h return to control levels. TH and DBH mRNA concentrations in the sympathetic ganglia show lower levels than those in the adrenal medulla. During immobilization, the ganglionic mRNA levels are significantly elevated; however, they stay increased even 24 h after the end of immobilization.

PNMT mRNA is also present in sympathetic ganglia in a very low concentration. PNMT gene expression has also been detected in other peripheral tissues such as the heart atria and ventricles, spleen, and thymus, and immobilization stress induced a many-fold increase of PNMT mRNA levels in these tissues.

TH mRNA levels also increase significantly within catecholaminergic cell bodies of the brain-stem areas, for example, the nucleus locus coeruleus, nucleus tractus solitarius, and A1 area, and in some hypothalamic nuclei following immobilization stress (Table 2). In these hypothalamic areas, catecholaminergic cells were not originally described.

The expression of genes that encode catecholamine biosynthetic enzymes is mediated by transcriptional mechanisms. Recent work identified various transcription factors that are associated with the brief or intermediate duration of a single or repeated immobilization stress exposure. It has been shown that the genes coding for the catecholamine biosynthetic enzymes contain several regulatory genomic elements in their promoters, for example, the cAMP response element (CRE), glucocorticoid response element (GRE), activating enhancer binding protein (AP)-1, AP-2, POU/Oct, and SP1 sites. The signaling molecules that may influence the transcription rates of the genes for TH, DBH, and PNMT, include P-cAMP response element binding protein (CREB), Jun N-terminal kinase (JNK), c-fos, Egr-1, Fra-1, and Fra-2.

With repeated daily stress exposure, Fra-2 was a major AP-1 factor induced in the adrenal medulla, but Egr-1 levels were also markedly elevated.

For PNMT gene expression, two transcription factors are important especially: glucocorticoid receptor and Egr-1. The crucial role of glucocorticoids in PNMT gene regulation has also been confirmed in corticotropin releasing hormone (CRH) (corticoliberin) knockout mice exposed to immobilization.

The signaling pathways triggered by immobilization differ in different tissues. In the adrenal medulla, immobilization triggered activation of the mitogen-activated protein kinase (MAPK), JNK, and induction of AP-1 factors, Egr-1 and phosphorylation of CREB. In contrast, in the sympathetic ganglia of immobilized animals were not observed changes in JNK and only little binding to the AP-1 motif of the TH promoter. However, there was an increase in CREB binding to the CRE site of the TH promoter. A single episode of immobilization also triggered the elevation of c-fos in the brain areas, especially in the locus coeruleus. In this area, increased phosphorylation of CREB was also observed after both single and repeated immobilization; however, the levels of Fra-2 and Egr-1 were not changed (in the adrenal medulla these transcription factors were highly activated). These results reveal differential mechanisms of regulation of gene expression of catecholamine biosynthetic enzymes by immobilization stress in two components of the sympathoadrenal system (adrenal medulla and sympathetic ganglia) as well as in the brain areas.

These findings related to the gene expression of catecholamine biosynthetic enzymes provide an excellent example of how immobilization has been useful in tracking the regulation of gene transcription during acute and chronic stress. In addition, these findings on gene expression may be compared with similar studies relating to the regulation of other genes because immobilization stress has been adopted as a model in a host of different laboratories.

Inventory of Neurohumoral and Genetic Responses to Acute Immobilization

Tables 1 and 2 present an abbreviated summary of biological responses of laboratory rats to a single 2-h period of immobilization stress. Table 1 emphasizes the responses of the sympathetic nervous system, the adrenal medulla, and the hypothalamic-pituitary-adrenocortical system because of the importance of these systems in contemporary stress research. Table 2 emphasizes changes in gene expression of catecholamine biosynthetic enzymes during the exposure of rats to an acute immobilization stress. This archival information was obtained from many

journal articles (see Further Reading for secondary reference sources). This information provides a useful inventory with which we may compare the responses of the same biological systems to other types of laboratory stressors.

Summary and Conclusion

Since the mid-1960s, immobilization has become one of the most frequently employed laboratory models of stress. The protocols for acute and chronic immobilization have been described clearly, and many laboratories have incorporated this procedure into their studies of stress. A rich and varied literature is available on biological responses to immobilization, and these findings have provided important insights into the molecular, genetic, endocrine, and neural adaptations that occur in laboratory rats following acute versus chronic stress.

See Also the Following Articles

Adrenal Medulla; Adrenocorticotrophic Hormone (ACTH); Animal Models (Nonprimate) for Human Stress; Corticosteroids and Stress; Genetic Factors and Stress; Hypothalamic-Pituitary-Adrenal.

Further Reading

Chrousos, G. P., McCarty, R., Pacak, K., et al. (eds.) (1995). *Special issue on stress: basic mechanisms and clinical implications. Annals of the New York Academy of Sciences* (771).

- Kvetnansky, R., McCarty, R. and Axelrod, J. (eds.) (1992). *Stress: neuroendocrine and molecular approaches*. New York: Gordon and Breach.
- Kvetnansky, R. and Mikulaj, L. (1970). Adrenal and urinary catecholamines in rats during adaptation to repeated immobilization stress. *Endocrinology* 87, 738–743.
- McCarty, R., Aguilera, G., Sabban, E. L. and Kvetnansky, R. (eds.) (1996). *Stress: molecular genetic and neurobiological advances*. Amsterdam: Harwood Academic.
- McCarty, R., Aguilera, G., Sabban, E. L. and Kvetnansky, R. (eds.) (2002). *Stress: neural, endocrine and molecular studies*. London: Taylor and Francis.
- Pacak, K., Aguilera, G., Sabban, E. L. and Kvetnansky, R. (eds.) (2004). *Special issue on stress: current neuroendocrine and genetic approaches. Annals of the New York Academy of Sciences* (1018).
- Sabban, E. L. and Kvetnansky, R. (2001). Stress-triggered activation of gene expression in catecholaminergic systems: dynamics of transcriptional events. *Trends in Neurosciences* 24, 91–98.
- Usdin, E., Kvetnansky, R. and Axelrod, J. (eds.) (1984). *Stress: the role of catecholamines and other neurotransmitters*. New York: Gordon and Breach.
- Usdin, E., Kvetnansky, R. and Kopin, I. J. (eds.) (1976). *Catecholamines and stress*. Oxford: Pergamon Press.
- Usdin, E., Kvetnansky, R. and Kopin, I. J. (eds.) (1980). *Catecholamines and stress: recent advances*. New York: Elsevier North Holland.
- Van Loon, G. R., Kvetnansky, R., McCarty, R. and Axelrod, J. (eds.) (1989). *Stress: neurochemical and humoral mechanisms*. New York: Gordon and Breach.
- Yehuda, R. and McEwen, B. (eds.) (2004). *Special issue on biobehavioral stress response: protective and damaging effects. Annals of the New York Academy of Sciences* (1032).

Immune Cell Distribution, Effects of Stress on

F S Dhabhar

Stanford University School of Medicine, Stanford, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by F S Dhabhar, volume 2, pp 507–514, © 2000, Elsevier Inc.

A Stress-Induced Decrease in Blood Leukocyte Numbers Represents a Redistribution Rather than a Destruction or Net Loss of Blood Leukocytes
Target Organs of a Stress-Induced Redistribution of Blood Leukocytes
Stress-Induced Redistribution of Blood Leukocytes
Conclusion

Glossary

B lymphocytes The antibody-producing cells of the body, originating in the bone marrow. B lymphocytes are also identified by the surface expression of specific molecules such as CD19.

Introduction

Stress

The Immune System

Stress-Induced Changes in Blood Leukocytes

Hormones Mediating a Stress-Induced Decrease in Blood Leukocytes

<i>Fluorescence-activated cell sorter (FACS)</i>	An instrument used to rapidly quantify the relative percentages of different cell types in a cell suspension. The identification of different cells is achieved based on the size, granularity, and fluorescence properties of the cell. Fluorescence is derived from fluorescently conjugated antibodies that bind specific markers on the cell surface.
<i>Immune response</i>	A reaction mounted by various components of the immune system, which consists of leukocytes, antibodies, cytokines and chemokines (messenger molecules), prostaglandins (inflammatory mediators), and accessory cells.
<i>Leukocytes</i>	Cells that constitute the immune system. These include T and B lymphocytes, natural killer (NK) cells, monocytes/macrophages, and granulocytes.
<i>Lymphocytes</i>	Leukocytes that specifically recognize and respond to foreign antigen.
<i>Monocytes/macrophages</i>	Cells that phagocytose foreign particles and self tissues that are injured (monocytes generally flow through the blood and mature into macrophages after they are activated and extravasate into tissues). They produce cytokines, which recruit other leukocytes to the site of an immune response, and they perform an important function known as antigen presentation, which is crucial to the initiation of specific immune responses by lymphocytes. Also known as mononuclear phagocytes.
<i>Natural killer (NK) cells</i>	Large granular lymphocytes that lyse virally infected cells and tumors in the absence of specific antigenic stimulation. They are identified by surface molecules such as CD16 and CD56.
<i>Neutrophils</i>	Cells that perform functions similar to mononuclear phagocytes and that form an important component of an acute inflammatory response; also known as polymorphonuclear leukocytes.
<i>T lymphocytes</i>	Leukocytes that are mediators of cellular immunity. Like all leukocytes, they are formed in the bone marrow but mature in the thymus. Helper T lymphocytes (Ths) orchestrate immune responses by secreting cytokines and recruiting other leukocytes to the site of an immune reaction and activating them at that site; most Th cells express a surface protein called CD4, which is used to identify them. Cytolytic T lymphocytes (CTLs) lyse cells that are infected by intracellular viruses or bacteria; most CTLs express a surface protein called CD8, which is used to identify them.

Introduction

Effective immunoprotection requires the rapid recruitment of leukocytes into sites of surgery, wounding, infection, or vaccination. Immune cells or leukocytes circulate continuously on surveillance pathways taking them from the blood, into various organs, and back into the blood. This circulation is essential for the maintenance of an effective immune defense network. The numbers and proportions of leukocytes in the blood provide an important representation of the state of distribution of leukocytes in the body and of the state of activation of the immune system. Numerous studies have shown that stress and stress hormones induce significant changes in absolute numbers and relative proportions of leukocytes in the blood.

Dhabhar et al. were the first to propose that a stress-induced decrease in blood leukocyte numbers may represent an adaptive response. These investigators suggested that acute stress-induced changes in blood leukocyte numbers represent a redistribution of leukocytes from the blood to other organs, such as the skin and lining of the gastrointestinal and urinary-genital tracts and their draining lymph nodes. They also suggested that such a leukocyte redistribution may enhance immune function in those compartments to which leukocytes traffic during stress. In agreement with this hypothesis, it has been demonstrated that a stress-induced redistribution of leukocytes from the blood to the skin is accompanied by a significant enhancement of a skin immune response. Here we discuss stress- and stress hormone-induced changes in blood leukocyte distribution and investigate the functional consequences of a stress-induced redistribution of blood leukocytes.

Stress

Although the word stress generally has negative connotations, stress is a familiar aspect of life, being a stimulant for some but a burden for others. Numerous definitions have been proposed for the word stress. Each definition focuses on aspects of an internal or external challenge, disturbance, or stimulus; on the perception of a stimulus by an organism; or on a physiological response of the organism to the stimulus. Physical stressors have been defined as external challenges to homeostasis, and psychological stressors have been defined as the "anticipation justified or not, that a challenge to homeostasis looms." (Sapolsky, 2005, p. 648). An integrated definition states that stress is a constellation of events, consisting of a stimulus (stressor) that precipitates a reaction in the brain (stress perception), which activates

physiological fight-or-flight systems in the body (stress response). The physiological stress response results in the release of neurotransmitters and hormones that serve as the brain's messengers to the rest of the body. This article examines the role played by stress and stress hormones in enhancing different aspects of an immune response. It is often overlooked that a stress response has salubrious adaptive effects in the short run, although stress can be harmful when it lasts a long time. An important distinguishing characteristic of stress is its duration and intensity. Thus, we define acute stress as stress that lasts for a period of a few minutes to a few hours and chronic stress as stress that persists for several hours per day for weeks or months. The intensity of stress may be gauged by the peak levels of stress hormones, neurotransmitters, by other physiological changes such as increases in heart rate and blood pressure, and by the amount of time for which these changes persist during stress and following the cessation of stress.

The Immune System

Immune Cells and Organs

The immune system consists of organs (bone marrow, spleen, thymus, and lymph nodes), cells (T cells; B cells; natural killer, NK, cells; monocytes/macrophages; and neutrophils), and messenger molecules (cytokines, chemokines, and prostaglandins). These components of the immune system are spread throughout the body and are in constant communication with one another. The appropriate distribution of immune cells in the body is crucial to the performance of the different functions of the immune system.

Functions of the Immune System

The functions of the immune system may be classified into six major categories:

1. Constant surveillance of the body in preparation for potential immunological challenges (such as those described in the other five categories).
2. Detection and elimination of infectious agents such as bacteria, viruses, fungi, and other microorganisms and parasites from the body.
3. Detection and elimination of noninfectious foreign matter from the body.
4. Wound repair and healing.
5. Clearance of debris that may result from noninjurious events such as programmed cell death (apoptosis) in host tissues.
6. Detection and elimination of tumors and neoplastic tissue.

Thus, the immune system may be regarded as the body's army with soldiers (immune cells) moving from their barracks (bone marrow, lymph nodes, spleen, and thymus) through highways (blood vessels and lymphatic ducts) and patrolling almost all organs within the body, especially those organs (skin, lining of gastrointestinal and urinary-genital tracts, lung, liver, lymph nodes, and spleen) that may serve as potential battle stations should the body's defenses be breached. In order to perform their functions, immune cells communicate with one another and with the rest of the body through messenger molecules such as cytokines and chemokines.

Immune Cell Trafficking

The appropriate distribution of immune cells between different tissues in the body is crucial to the performance of the immune functions. Leukocyte trafficking is the process by which leukocytes circulate between immune and nonimmune organs as they patrol the body and discharge their functions. Leukocyte trafficking is accomplished through interactions between adhesion molecules on leukocytes and endothelial cells that ultimately determine the location of immune cells within the body. The three major families of adhesion molecules are the selectins, the immunoglobulin superfamily, and the integrins. Adhesion molecules mediate three important steps (rolling, anchoring, and extravasation) that are involved in capturing leukocytes from the circulation. For example, a T cell flowing through the bloodstream may be selectively retained within the vasculature of the skin if a particular adhesion molecule is expressed on the T cell and its corresponding ligand is expressed on skin endothelial cells. Thus, hormones, cytokines, and chemokines may influence the location of leukocytes in the body by regulating the conformation or expression of adhesion molecules and/or their ligands on leukocytes and endothelial cells.

The Immune Response

It is important to note that, although immune reactions are classified into different categories (e.g., innate, acquired, cell-mediated, or humoral reactions), an immune response is often a combination of different immune reactions. The immune response protects the body from infections, wounds, and tumors. Although mechanisms for self- and counterregulation exist between various immune components, mechanisms also exist by which immune responses may be modulated or influenced by mediators such as hormones and neurotransmitters. Stress, and its consequent release of hormones and neurotransmitters, has

been shown to influence specific immune parameters such as leukocyte trafficking, NK cell activity, lymphocyte proliferation, antibody production, effector cell function, and cell-mediated immune reactions.

Stress-Induced Changes in Blood Leukocytes

The phenomenon of acute stress-induced changes in blood leukocyte numbers is well known. In fact, decreases in blood leukocyte numbers were used as an indirect measure for increases in plasma corticosterone before methods were available to directly assay the hormone. Stress-induced changes in blood leukocyte numbers have been reported in fish, mice, rats, rabbits, horses, nonhuman primates, and humans. This suggests that the phenomenon of stress-induced leukocyte distribution has a long evolutionary lineage and that perhaps it has functional significance.

Many of these studies have shown that stress-induced increases in plasma corticosterone are accompanied by a significant decrease in numbers and percentages of lymphocytes and by an increase in numbers and percentages of neutrophils. FACS analyses have revealed that absolute numbers of peripheral blood T cells, B cells, NK cells, and monocytes all show a rapid and significant decrease (40–70% lower than baseline) during stress. These stress-induced changes in leukocyte numbers are rapidly reversed on the cessation of stress.

Hormones Mediating a Stress-Induced Decrease in Blood Leukocytes

Several studies have examined stress hormone-induced changes in blood leukocyte numbers. It has been shown in rats that both adrenalectomy (which eliminates the corticosterone and epinephrine stress response) and cyanoketone treatment (which eliminates only the corticosterone stress response) virtually eliminate the stress-induced redistribution of blood leukocytes. Studies have also shown that adrenal steroid and catecholamine hormones mediate the stress-induced changes in blood leukocyte numbers and that glucocorticoid treatment induces changes in leukocyte distribution in mice, guinea pigs, rats, and humans. These studies clearly show that glucocorticoid hormones induce a significant decrease in blood lymphocyte numbers when administered under acute as well as chronic conditions. Similarly studies have delineated the importance of catecholamine hormones in mediating stress-induced changes in blood leukocyte distribution in rodents and humans.

A Stress-Induced Decrease in Blood Leukocyte Numbers Represents a Redistribution Rather than a Destruction or Net Loss of Blood Leukocytes

From the discussion so far, it is clear that stress and glucocorticoid hormones induce a rapid and significant decrease in blood lymphocyte, monocyte, and NK cell numbers. This decrease in blood leukocyte numbers may be interpreted in two possible ways. The decrease in cell numbers could reflect a large-scale destruction of circulating leukocytes. Alternatively, it could reflect a redistribution of leukocytes from the blood to other organs in the body. Experiments were conducted to test the hypothesis that acute stress induces a redistribution of leukocytes from the blood to other compartments in the body.

The first series of experiments examined the kinetics of recovery of the stress-induced reduction in blood leukocyte numbers. It was hypothesized that if the observed effects of stress represented a redistribution rather than a destruction of leukocytes, we would see a relatively rapid return of leukocyte numbers back to baseline on the cessation of stress. Results showed that all leukocyte subpopulations that showed a decrease in absolute numbers during stress showed a complete recovery, with their numbers reaching pre-stress baseline levels within 3 h after the cessation of stress. Plasma levels of lactate dehydrogenase (LDH), which is a marker for cell damage, were also monitored in the same experiment. If the stress-induced decrease in leukocyte numbers were the result of a destruction of leukocytes, we would expect to observe an increase in plasma levels of LDH during or following stress. However, no significant changes in plasma LDH were observed, further suggesting that a redistribution rather than a destruction of leukocytes was primarily responsible for the stress-induced decrease in blood leukocyte numbers.

It is important to bear in mind that, although glucocorticoids are known to induce leukocyte apoptosis under certain conditions, they have also been shown to induce changes in various immune parameters and in immune cell distribution in the absence of cell death. It has been suggested that some species may be steroid-resistant and that others may be steroid-sensitive and that glucocorticoid-induced changes in blood leukocyte numbers represent changes in leukocyte redistribution in steroid-resistant species (humans and guinea pigs) and leukocyte lysis in steroid-sensitive species (mouse and rat). However, it is now accepted that, even in species previously thought to be steroid-sensitive, changes in adrenal

steroids similar to those described here produce changes in leukocyte distribution rather than an increase in leukocyte destruction.

Target Organs of a Stress-Induced Redistribution of Blood Leukocytes

Based on this discussion, the obvious question we might ask is: Where do blood leukocytes go during stress? Numerous studies using stress or stress hormone treatments have investigated this issue. Using gamma imaging to follow the distribution of adoptively transferred radiolabeled leukocytes in whole animals, experiments showed that stress induces a redistribution of leukocytes from the blood to the mesenteric lymph nodes. It has been reported that anesthesia stress, as well as the infusion of adrenocorticotrophic hormone (ACTH) and prednisolone in rats results in decreased numbers of labeled lymphocytes in the thoracic duct, whereas the cessation of drug infusion results in the normal circulation of labeled lymphocytes. This suggests that ACTH and prednisolone (which produce hormonal changes similar to those observed during stress) may cause the retention of circulating lymphocytes in different body compartments, thus resulting in a decrease in lymphocyte numbers in the thoracic duct and a concomitant decrease in numbers in the peripheral blood. It has also been reported that a single injection of hydrocortisone, prednisolone, or ACTH results in increased numbers of lymphocytes in the bone marrow of mice, guinea pigs, and rats. Fauci et al. suggested that glucocorticoid-induced decreases in blood leukocyte numbers in humans may also reflect a redistribution of immune cells to other organs in the body. Finally, corticosteroids have been shown to induce the accumulation of lymphocytes in mucosal sites, and the skin has been identified as a target organ to which leukocytes traffic during stress. *In vitro* catecholamine treatment has also been shown to direct leukocyte traffic to the spleen and lymph nodes.

It is important to note that in these studies a return to basal glucocorticoid levels was followed by a return to basal blood lymphocyte numbers, further supporting the hypothesis that the decrease in blood leukocyte numbers is the result of a glucocorticoid-induced redistribution rather than a glucocorticoid-induced destruction of blood leukocytes. In view of this, Dhabhar et al. proposed that a transient stress-induced decrease in blood leukocyte numbers reflects a redistribution or redeployment of leukocytes from the blood to other organs (e.g., lymph nodes, bone marrow, and skin) in the body.

Stress-Induced Redistribution of Blood Leukocytes

Molecular Mechanisms

It is likely that the observed changes in leukocyte distribution are mediated by changes in either the expression or affinity of adhesion molecules on leukocytes and/or endothelial cells. It has been suggested that following stress or glucocorticoid-treatment specific leukocyte subpopulations (being transported by blood and lymph through various body compartments) may be selectively retained in those compartments in which they encounter a stress- or glucocorticoid-induced adhesion match. As a result of this selective retention, the proportion of some leukocyte subpopulations would decrease in the blood and increase in the organ in which they are retained (e.g., the skin).

Support for this hypothesis comes from studies that show that acute psychological stressors such as public speaking can induce significant changes in leukocyte adhesion molecules. Prednisolone has also been shown to induce the retention of circulating lymphocytes within the bone marrow, spleen, and some lymph nodes, thus resulting in a decrease in lymphocyte numbers in the thoracic duct and a concomitant decrease in numbers in the peripheral blood. Moreover, glucocorticoid hormones also influence the production of cytokines and lipocortins that, in turn, can affect the surface adhesion properties of leukocytes and endothelial cells. Further investigation of the effects of endogenous glucocorticoids (administered in physiological doses, and examined under physiological kinetic conditions) on changes in the expression/activity of cell surface adhesion molecules and on leukocyte-endothelial cell adhesion is necessary.

Functional Consequences

It had been proposed that a stress-induced decrease in blood leukocyte numbers may represent an adaptive response, that is, a redistribution of leukocytes from the blood to other organs such as the skin, lining of gastrointestinal and urinary-genital tracts, lung, liver, and lymph nodes that may serve as potential battle stations should the body's defenses be breached. It has been suggested that such a leukocyte redistribution may enhance the immune function in those compartments to which leukocytes traffic during stress.

Thus, an acute stress response may direct the body's soldiers (leukocytes) to exit their barracks (the spleen and bone marrow), travel the boulevards (blood vessels), and take up position at potential battle stations (the skin, lining of gastrointestinal

and urinary-genital tracts, lung, liver, and lymph nodes) in preparation for an immune challenge. In addition to sending leukocytes to potential battle stations, stress hormones may also better equip them for battle by enhancing processes such as antigen presentation, phagocytosis, and antibody production. Thus, a hormonal alarm signal released by the brain on detecting a stressor, may prepare the immune system for potential challenges (wounding or infection) that may arise due to the actions of the stress-inducing agent (e.g., a predator or attacker).

When interpreting data showing stress-induced changes in functional assays such as lymphocyte proliferation or NK activity, it may be important to bear in mind the effects of stress on the leukocyte composition of the compartment in which an immune parameter is being measured. For example, it has been shown that acute stress induces a redistribution of leukocytes from the blood to the skin and that this redistribution is accompanied by a significant enhancement of a skin cell-mediated immune (CMI) response. In what might at first glance appear to be contradicting results, acute stress has been shown to suppress splenic and peripheral blood responses to T-cell mitogens and splenic IgM production. However, it is important to note that in contrast to the skin, which is enriched in leukocytes during acute stress, the peripheral blood and spleen are relatively depleted of leukocytes during acute stress. This stress-induced decrease in blood and spleen leukocyte numbers may contribute to the acute stress-induced suppression of immune function in these compartments.

Moreover, in contrast to acute stress, chronic stress has been shown to suppress skin CMI and a chronic stress-induced suppression of blood leukocyte redistribution is thought to be one of the factors mediating the immunosuppressive effect of chronic stress. Again, in what might appear to be contradicting results, chronic stress has been shown to enhance mitogen-induced proliferation of splenocytes and splenic IgM production. However, the spleen is relatively enriched with T cells during chronic glucocorticoid administration, suggesting that it may also be relatively enriched with T cells during chronic stress, and this increase in spleen leukocyte numbers may contribute to the chronic stress-induced enhancement of immune parameters measured in the spleen.

It is also important to bear in mind that the heterogeneity of the stress-induced changes in leukocyte distribution suggests that using equal numbers of leukocytes in a functional assay may not account for stress-induced changes in relative percentages of different leukocyte subpopulations in the cell suspension

being assayed. For example, samples that have been equalized for absolute numbers of total blood leukocytes from control versus stressed animals may still contain different numbers of specific leukocyte subpopulations (e.g., T cells, B cells, or NK cells). Such changes in leukocyte composition may mediate the effects of stress even in functional assays using equalized numbers of leukocytes from different treatment groups. This possibility needs to be taken into account before concluding that a given treatment changes an immune parameter on a per-cell rather than a per-population basis.

Conclusion

An important function of endocrine mediators released under conditions of acute stress may be to ensure that appropriate leukocytes are present in the right place and at the right time to respond to an immune challenge that might be initiated by the stress-inducing agent (e.g., attack by a predator or invasion by a pathogen). The modulation of immune cell distribution by acute stress may be an adaptive response designed to enhance immune surveillance and increase the capacity of the immune system to respond to challenge in immune compartments (such as the skin, the epithelia of the lung, and the gastrointestinal and urinary-genital tracts) that serve as major defense barriers for the body. Thus, neurotransmitters and hormones released during stress may increase immune surveillance and help enhance immune preparedness for potential (or ongoing) immune challenge.

See Also the Following Articles

Immune Response; Immune Function, Stress-Induced Enhancement; Immunity.

Further Reading

- Benschop, R. J., Rodriguez-Feuerhahn, M. and Schedlowski, M. (1996). Catecholamine-induced leukocytosis: early observations, current research, and future directions. *Brain, Behavior, & Immunity* 10, 77–91.
- Cox, J. H. and Ford, W. L. (1982). The migration of lymphocytes across specialized vascular endothelium. IV: Prednisolone acts at several points on the recirculation pathway of lymphocytes. *Cellular Immunology* 66, 407–422.
- Dhabhar, F. S. and McEwen, B. S. (1996). Stress-induced enhancement of antigen-specific cell-mediated immunity. *Journal of Immunology* 156, 2608–2615.
- Dhabhar, F. S. and McEwen, B. S. (1999). Enhancing versus suppressive effects of stress hormones on skin immune function. *Proceedings of the National Academy of Sciences USA* 96, 1059–1064.

- Dhabhar, F. S. and McEwen, B. S. (2001). Bidirectional effects of stress & glucocorticoid hormones on immune function: possible explanations for paradoxical observations. In: Ader, R., Felten, D. L. & Cohen, N. (eds.) *Psychoneuroimmunology* (3rd edn., pp. 301–338). San Diego, CA: Academic Press.
- Dhabhar, F. S., Miller, A. H., McEwen, B. S., et al. (1995). Effects of stress on immune cell distribution – dynamics and hormonal mechanisms. *Journal of Immunology* **154**, 5511–5527.
- Dhabhar, F. S., Miller, A. H., McEwen, B. S., et al. (1996). Stress-induced changes in blood leukocyte distribution – role of adrenal steroid hormones. *Journal of Immunology* **157**, 1638–1644.
- Dhabhar, F. S. and Viswanathan, K. (2005). Short-term stress experienced at the time of immunization induces a long-lasting increase in immunological memory. *American Journal of Physiology: Regulatory, Integrative & Comparative Physiology* **289**, R738–R744.
- Dougherty, R. F. and White, A. (1945). Functional alterations in lymphoid tissue induced by adrenal cortical secretion. *American Journal of Anatomy* **77**, 81–116.
- Fauci, A. S. (1975). Mechanisms of corticosteroid action on lymphocyte subpopulations. I: Redistribution of circulating T and B lymphocytes to the bone marrow. *Immunology* **28**, 669–680.
- Goebel, M. U. and Mills, P. J. (2000). Acute psychological stress and exercise and changes in peripheral leukocyte adhesion molecule expression and density. *Psychosomatic Medicine* **62**, 664–670.
- Mills, P. J. and Dimsdale, J. E. (1996). The effects of acute psychologic stress on cellular adhesion molecules. *Journal of Psychosomatic Research* **41**, 49–53.
- Sapolsky, R. M. (2005). The hierarchy on primate health. *Science* **308**, 648–652.
- Stefanski, V., Solomon, G. F., Kling, A. S., et al. (1996). Impact of social confrontation on rat CD4 T cells bearing different CD45R isoforms. *Brain, Behavior & Immunity* **10**, 364–379.
- Viswanathan, K. and Dhabhar, F. S. (2005). Stress-induced enhancement of leukocyte trafficking into sites of surgery or immune activation. *Proceedings of the National Academy of Sciences USA*, **102**, 5808–5812.
- Zalcman, S. and Anisman, H. (1993). Acute and chronic stressor effects on the antibody response to sheep red blood cells. *Pharmacology, Biochemistry & Behavior* **46**, 445–452.

Immune Function See: Immune Response; Immune Function, Stress-Induced Enhancement; Immune Suppression; Immune System, Aging; Immunity; Immune Surveillance – Cancer, Effects of Stress on.

Immune Function, Stress-Induced Enhancement

F S Dhabhar

Stanford University School of Medicine, Stanford, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by F S Dhabhar, volume 2, pp 515–522, © 2000, Elsevier Inc.

Stress

The Immune System

General Assumption: Stress Suppresses Immune Function and is Detrimental to Health – Reasons for Modifying This General Assumption

Studies Showing Stress-Induced Enhancement of Immune Function

Functional Implications

The Stress Spectrum Hypothesis

Conclusion

Glossary

- Antigen** A foreign substance that induces an immune response.
- Cytokines** Protein hormones that mediate various functions of immune cells.
- Immune response** A reaction mounted by various components of the immune system, which consists of leukocytes, antibodies, cytokines and chemokines (messenger molecules), prostaglandins (inflammatory mediators), and accessory cells.
- Leukocytes** Cells that constitute the immune system. These include T and B lymphocytes, natural killer (NK) cells, monocytes/macrophages, and granulocytes.

Stress

Although the word stress generally has negative connotations, stress is a familiar aspect of life, being a stimulant for some but a burden for others. Numerous definitions have been proposed for the word stress. Each definition focuses on aspects of an internal or external challenge, disturbance, or stimulus; on the perception of a stimulus by an organism; or on a physiological response of the organism to the stimulus. Physical stressors have been defined as external challenges to homeostasis, and psychological stressors have been defined as the “anticipation justified or not, that a challenge to homeostasis looms.” (Sapolsky, 2005, p. 648). An integrated definition states that stress is a constellation of events, consisting of a stimulus (stressor) that precipitates a reaction in the brain (stress perception), which activates physiological fight-or-flight systems in the body (stress response). The physiological stress response results in the release of neurotransmitters and hormones that serve as the brain’s messengers to the rest of the body. This article examines the role played by stress and stress hormones in enhancing different aspects of the immune response. It is often overlooked that a stress response has salubrious adaptive effects in the short run, although stress can be harmful when it lasts a long time. An important distinguishing characteristic of stress is its duration and intensity. Thus, we define acute stress as stress that lasts for a period of a few minutes to a few hours and chronic stress as stress that persists for several hours per day for weeks or months. The intensity of stress may be gauged by the peak levels of stress hormones, neurotransmitters, by other physiological changes such as increases in heart rate and blood pressure, and by the amount of time for which these changes persist during stress and following the cessation of stress.

The Immune System

Immune Cells and Organs

The immune system consists of organs (bone marrow, spleen, thymus, and lymph nodes), cells (T cells; B cells; and natural killer, NK, cells; monocytes/macrophages; and granulocytes), and messenger molecules (cytokines, chemokines, and prostaglandins). These components of the immune system are spread throughout the body and are in constant communication with one another. The appropriate distribution of immune cells in the body is crucial to the performance of the different functions of the immune system.

Functions of the Immune System

The functions of the immune system may be classified into six major categories:

1. Constant surveillance of the body in preparation for potential immunological challenges (such as those described in the other five categories).
2. Detection and elimination of infectious agents such as bacteria, viruses, fungi, and other microorganisms and parasites from the body.
3. Detection and elimination of noninfectious foreign matter from the body.
4. Wound repair and healing.
5. Clearance of debris that may result from non-injurious events such as programmed cell death (apoptosis) in host tissues.
6. Detection and elimination of tumors and neoplastic tissue.

Thus, it has been suggested the immune system may be regarded as the body’s army with soldiers (immune cells) moving from their barracks (bone marrow, lymph nodes, spleen, thymus) through highways (blood vessels and lymphatic ducts) and patrolling almost all organs within the body, especially those organs (the skin, lining of gastrointestinal and urinary-genital tracts, lung, liver, lymph nodes, and spleen) that may serve as potential battle stations should the body’s defenses be breached. While performing their various functions, immune cells communicate with one another and with the rest of the body through messenger molecules such as the cytokines and chemokines.

Types of Immune Responses

Natural or innate immunity Natural or innate immune responses are initially mounted when an organism encounters an antigen or pathogen. These responses are rapidly activated and generally regarded as the first line of defense. Epithelial barriers, immune cells, the complement system, circulating proteins such as C-reactive protein, cytokines such as the α and β interferons (IFNs), tumor necrosis factor (TNF), interleukin (IL)-12, and IL-15 are involved in mediating innate immune reactions. Macrophages, neutrophils, and NK cells are the primary cellular mediators. Cells of the innate immune system recognize pathogens through surface receptors such as the toll-like receptors (TLRs), which recognize pathogen-associated molecular patterns (PAMPs). Innate immune responses are not enhanced by repeated exposure to antigens.

Acquired or specific immunity Acquired or specific immune responses are mounted following prior

exposure to an antigen; they are mediated by lymphocytes (T and B cells) and antibodies and are primarily regulated by lymphocyte-derived cytokines such as IL-2, IL-4, and IFN- γ . Acquired immune responses are directed against specific antigens and are often enhanced following repeated exposure to antigens. Acquired immunity can be divided into two broad categories: cell-mediated immunity (CMI) and humoral immunity.

Cell-mediated immunity CMI can be adoptively transferred from an immunized organism to naïve organism by the transfer of T cells and primarily involves cell-mediated clearing mechanisms. Intracellular infectious agents such as certain bacteria (e.g., *Listeria monocytogenes* or *Mycobacterium tuberculosis*) and viruses require a CMI response for their elimination. This response involves T helper (Th) cells which recognize infected cells with the help of specialized antigen-presenting cells (APC). On antigen recognition, Th cells release cytokines and chemokines, and they orchestrate a series of reactions involving cytolytic T cells (CTL), macrophages, and NK cells. These cells move to the site of infection and lyse or phagocytose infected cells and destroy them. In addition to their protective effects, CMI responses are also involved in harmful reactions such as organ transplant rejection, graft versus host disease, and certain autoimmune diseases. Delayed type hypersensitivity (DTH) is a form of CMI in which the macrophage is the major effector cell. The tuberculin skin test is a classic example of a CMI response. A characteristic CMI response peaks at 24–48 h following exposure to the antigen. Depending on the nature of the antigen, CMI reactions mediate beneficial (resistance to viruses, bacteria, and fungi) or harmful (allergic dermatitis, autoimmunity) aspects of the immune function.

Humoral immunity A humoral immune response is adoptively transferred from an immunized organism to a naïve organism by the transfer of cell-free plasma or serum and primarily involves antibody-mediated clearing mechanisms. Extracellular infectious agents such as bacteria (e.g., *Staphylococcus*, *Streptococcus*, or *Escherichia coli*), protozoan, and helminthic parasites require antibody-mediated responses for their elimination. Antibodies are proteins produced by B cells that neutralize extracellular infectious agents, which are then destroyed by macrophages, neutrophils, or complement proteins.

The integrated immune response It is important to note that, although immune reactions are classified

into different categories, an immune response is often a combination of innate as well as acquired immune reactions. *In vivo* immune responses necessarily involve intricate interactions between both the innate and adaptive arms of the immune system. Different components of the immune system work in synchrony to protect the body from infections, wounds, and tumors. Although mechanisms of self- and counter-regulation exist between various immune components, mechanisms also exist by which immune responses may be modulated or influenced by mediators such as hormones and neurotransmitters. Stress, and its consequent release of hormones and neurotransmitters, has been shown to influence specific immune parameters such as blood leukocyte numbers, NK cell activity, lymphocyte proliferation, antibody production, effector cell function, and CMI reactions. The immunoenhancing effects of stress are the focus of the rest of this article.

General Assumption: Stress Suppresses Immune Function and is Detrimental to Health – Reasons for Modifying This General Assumption

It is generally assumed that stress suppresses immune function and hence is detrimental to health. However, the suppression of immune function under all stress conditions should not be evolutionarily adaptive. It seems paradoxical that organisms should have evolved to suppress immune function at a time when an active immune response may be critical for survival, for example, under conditions of stress when an organism may be injured or infected by the actions of the stress-inducing agent (e.g., a predator). Stress is an intrinsic part of life for most organisms, and dealing successfully with stressors is what enables survival. Environmental challenges and most selection pressures, the chisels of evolution, are stressors that may be psychological (fear and anxiety), physical (wounding and infection), or physiological (food or water deprivation). One of the primary functions of the brain is to perceive stress and to warn and enable an organism to deal with its consequences. For example, when a gazelle sees a charging lion, the gazelle's brain detects a threat and orchestrates a physiological response that first prepares and then enables the gazelle to flee.

It has been suggested that under short-duration stress conditions, just as the stress response prepares the nervous, cardiovascular, musculoskeletal, and neuroendocrine systems for fight or flight, it may also prepare the immune system for challenges (e.g., wounding or infection) that may be imposed by the

stressor. In contrast to this hypothesis, it had previously been suggested that a stress-induced suppression of immune function may be evolutionarily adaptive because immunosuppression may conserve energy that is required to deal with the immediate demands imposed by the stressor. However, immunosuppression does not necessarily conserve energy, and some of the proposed mechanisms for immunosuppression, such as apoptosis of leukocytes, may expend rather than conserve energy. Moreover, the immune system may often be critically needed to respond immediately to the actions of the stress-inducing agent (e.g., wounding by a predator). Thus, whereas ovulation, copulation, and digestion can wait for the cessation of stress, the immune response cannot be assumed to be dispensable in terms of meeting the immediate demands of stress. Moreover, the time course for many of the proposed mechanisms for stress-induced immunosuppression, such as inhibition of prostaglandin synthesis, cytokine production, or leukocyte proliferation, is significantly longer than that seen during acute stress conditions. Thus, although energy conservation may play a role in stress-induced immunosuppression under some conditions, it is not clear that it does so under all stress conditions.

Another well-known paradoxical observation is that, on the one hand, stress is thought to suppress immunity and increase susceptibility to infections and cancer and, on the other, it is thought to exacerbate inflammatory (e.g., cardiovascular disease and gingivitis) and autoimmune (e.g., psoriasis, arthritis, and multiple sclerosis) diseases that should be ameliorated by the suppression of immune function. If stress were solely immunosuppressive, we would expect stress to ameliorate, not exacerbate, these diseases. A stress-induced enhancement of immune function would help explain the stress-induced exacerbation of diseases mediated by immunopathological reactions.

Keeping these considerations in mind, and based on their initial observations of the effects of stress and of the circadian corticosterone rhythm on leukocyte redistribution in the body, Dhabhar and McEwen demonstrated that stress has bidirectional effects on immune function such that acute stress is immunoenhancing and chronic stress is immunosuppressive. This article focuses on the immunoenhancing aspects of stress.

Studies Showing Stress-Induced Enhancement of Immune Function

A majority of studies in the field of psychoneuro-immunology have focused on the immunosuppressive

effects of stress, but several studies have also revealed that stress can be immunoenhancing. A stress-induced enhancement of immune function may be an important evolutionarily adaptive response that prepares an organism for potential immunological challenges for which stress perception by the brain, and subsequent stress hormone and neurotransmitter release, may serve as an early warning. Studies showing stress-induced enhancements in immune parameters *in vivo* and *in vitro* are briefly reviewed next.

Skin Cell-Mediated Immunity

There exist several mechanisms by which stress may enhance immune function. Studies showed that acute stress (2-h restraint) results in a significant redistribution of leukocytes from the blood to other organs (the skin, lymph nodes, and bone marrow) in the body and that adrenal stress hormones are the major mediators of this leukocyte redistribution. Because the skin was one of the targets to which leukocytes trafficked during stress, it was hypothesized that such a leukocyte redistribution may increase immune surveillance and consequently enhance immune function should the skin be exposed to an antigen following acute stress.

To test this hypothesis, researchers examined the effects of acute stress on skin immunity, using rodent models of CMI. In order to induce CMI, researcher initially sensitized animals to 2,4-dinitro-1-fluorobenzene (DNFB) by administering the chemical antigen to the skin of the dorsum. The sensitization phase of a CMI reaction is one in which the organism develops an immunological memory (through the generation of memory T cells) for the antigen with which it is immunized. Following sensitization, the ability of the animals to mount a CMI response against DNFB was examined by administering DNFB to the dorsal aspect of the right pinna. The CMI response was subsequently measured as an increase in pinna thickness that was proportional to the intensity of the ongoing immune reaction. This phase, also known as the challenge phase, involves the recruitment of memory T cells and effector cells such as macrophages, which mount an immune response against the antigen to which the animal was previously sensitized. Acute restraint stress administered immediately before the introduction of the antigen resulted in a large and long-lasting enhancement of skin CMI. Histological analysis identified significantly larger numbers of leukocytes in the skin of stressed animals both before and after exposure to antigen, and this suggested that a stress-induced redistribution of leukocytes was one of the factors mediating the stress-induced enhancement of skin

immunity. Acute stress has similarly been shown to enhance skin CMI in mice and, under certain conditions, may also enhance the CMI response in human subjects.

It has also been shown that the acute stress-induced enhancement of skin CMI is mediated by adrenal hormones. Thus, adrenalectomy, which eliminates the glucocorticoid and epinephrine stress response, eliminated the stress-induced enhancement of skin CMI. Low-dose corticosterone or epinephrine administration significantly enhanced skin CMI and produced a significant increase in T-cell numbers in the lymph nodes draining the site of the CMI reaction. These results suggest a novel role for adrenal stress hormones as endogenous immunoenhancing agents. They also show that hormones released during an acute stress response may help prepare the immune system for potential challenges (e.g., wounding or infection) for which stress perception by the brain may serve as an early warning signal.

Humoral Immunity

Wood et al. showed that exposure to foot shock stress enhanced humoral as well as CMI to keyhole limpet hemocyanin (KLH). These investigators administered a single acute foot shock session on days -1, 0, 1, and 3, relative to sensitization (day 0). They observed that, compared to nonstressed controls, animals stressed on days 0 or 1 showed enhanced immune responses with respect to serum anti-KLH IgG levels, splenocyte proliferation, and skin CMI to the KLH antigen. Similar results were almost simultaneously reported by researchers who found that the timing of stressor exposure was critical for determining its immunoenhancing effects. Several other studies have reported stress-induced enhancement of humoral immune responses. For example, stress has been shown to accelerate antigen removal in mice, rats, and pigs. It has also been proposed that glucocorticoids may shift the balance of an ongoing immune response in favor of humoral immunity, and physiological doses of glucocorticoids have been shown to enhance immunoglobulin production by mitogen- and IL-4-stimulated human lymphocytes in culture.

Cytokine Function

Several lines of evidence indicate that glucocorticoid stress hormones may act synergistically with cytokines to enhance specific immune reactions. Thus, although glucocorticoids inhibit the synthesis of cytokines under some conditions, they have also been shown to induce the release and potentiate the actions of cytokines under other conditions. For example, stress has been shown to induce increases in plasma

levels of IL-1, IL-6, and glucocorticoids to induce the production of migration inhibitory factor (MIF), which is, in turn, proposed to counter the actions of glucocorticoids. Moreover, glucocorticoids synergistically enhance the induction of acute phase proteins by IL-1 and IL-6. Glucocorticoids similarly enhance the biological responses of other cytokines such as IL-2, IFN- γ , granulocyte colony-stimulating factor (G-CSF), granulocyte macrophage colony-stimulating factor (GM-CSF), and oncostatin M. It has been suggested that synergistic interactions between glucocorticoids and cytokines may be mediated by the glucocorticoid-induced upregulation of cytokine receptors on target cells.

Proliferative Responses and Phagocytosis

Acute stressors such as handling and foot shock have been shown to enhance the mitogen-induced proliferation of blood lymphocytes. Acute stress has also been shown to enhance macrophage phagocytic function and blood NK cell activity. Several studies have shown that the *in vitro* administration of stress hormones to immune cells can enhance different aspects of the immune function. Low physiological concentrations of corticosterone have been shown to stimulate mitogen- and anti-T-cell receptor-induced proliferation of splenocytes. Similarly, glucocorticoid hormones have been shown to enhance nitric oxide and IL-1 β secretion by macrophages.

Functional Implications

The studies described here show that, under certain conditions, stress can enhance immune responses. These studies may help reconcile the apparent contradiction between reports showing that, on the one hand, stress suppresses immunity and increases susceptibility to infections and cancer but, on the other, exacerbates autoimmune diseases that should be ameliorated by the suppression of immune function. For example, under natural conditions, acute stress may serve a protective role by enhancing an immune response directed toward a wound or infection. However, a stress-induced enhancement of immune function could also be detrimental if the immune response is directed against an innocuous (poison ivy, nickel in jewelry, latex, etc.) or autoimmunogenic antigen, and this could explain the well-known stress-induced exacerbations of autoimmune diseases. It is also known that chronic stress suppresses immune function. This may explain stress-induced exacerbations of infections and cancer and the stress-induced suppression of wound healing. It has also been reported that high doses or prolonged administration of corticosterone or low doses

of dexamethasone are potently immunosuppressive, in keeping with their clinically well-known antisuppressive effects.

It is important to note that humans as well as animals experience acute or chronic stress as an intrinsic part of life. They also experience stress in conjunction with many standard diagnostic, clinical, and experimental manipulations. Unintended stress may significantly affect diagnostic measures and overall health outcomes. Thus, it may be important to account for the effects of acute versus chronic stress when conducting clinical, diagnostic, or experimental manipulations.

The Stress Spectrum Hypothesis

In view of the data presented here, and in view of the widely demonstrated immunosuppressive effects of stress, Dhabhar and McEwen proposed that a stress response and its consequent effects on immune function may be viewed in the context of a stress spectrum. One region of this spectrum is characterized by acute stress or eustress (i.e., conditions of short-duration stress that may result in immunopreparatory or immunoenhancing physiological conditions). An important characteristic of acute stress is a rapid physiological stress response mounted in the presence of the stressor, followed by a rapid shut-down of the response on cessation of the stressor. The other region of the stress spectrum is characterized by chronic stress or distress (i.e., repeated or prolonged stress that may result in dysregulation or suppression of immune function). An important characteristic of chronic stress is that the physiological response either persists long after the stressor has ceased or is activated repeatedly, which results in an overall integrated increase in the exposure of the organism to stress hormones. The concept of allostatic load has been proposed to define the psychophysiological wear and tear that takes place while different biological systems work to stay within a range of equilibrium (allostasis) in response to the demands placed by internal or external chronic stressors. We suggest that conditions of high allostatic load result in the dysregulation or suppression of immune function. Significantly, a disruption of the circadian corticosterone/cortisol rhythm may be an indicator and/or mediator of distress or high allostatic load.

The stress spectrum also proposes that acute or chronic stress is generally superimposed on a psychophysiological health maintenance equilibrium. The extent and efficiency with which an organism returns to its health maintenance equilibrium after stress depends on resilience, which we define as the reserve capacity of psychophysiological systems to recover

from challenging conditions. Factors such as coping mechanisms, sense of control, optimism, social support, early life experiences, learning, genetics, and sleep may be important mediators of psychological resilience. Factors such as neuroendocrine reactivity, genetics, environment, nutrition, and sleep may be important mediators of physiological resilience. The psychophysiological basis of resilience and reserve capacity are underinvestigated and provide an important opportunity for future research.

The stress spectrum, taken together with the preceding discussion, shows that the duration, intensity/concentration, and timing of exposure to stressor-induced physiological activation (neurotransmitters, hormones, and their molecular, cellular, organ-level, and systemic effects) are critical for determining whether stress enhances or suppresses/dysregulates immune function. The model shows that the stressor itself can be acute or chronic. Stress perception and processing by the brain and mechanisms mediating psychological and physiological resilience are critical for determining the duration and magnitude of the physiological stress response. Psychological resilience mechanisms are especially important in humans because they can limit the duration and magnitude of chronic stress responses. Psychogenic stressors are also very important in human subjects because they can generate stress responses long after stressor exposure (e.g., posttraumatic stress disorder following a severe traumatic experience or, in a milder form, lingering anger/mood disturbance following a social altercation) or even in the absence of a physical stressor or salient threat (e.g., worrying about whether one's romantic feelings will be reciprocated). Therefore, following stressor exposure and its processing by the brain, there ensues a physiological stress response. This response may consist of acute or chronic physiological activation (neurotransmitters; hormones; and their molecular, cellular, organ-level and systemic effects) that results in psychophysiological states that have different effects on overall health and immune function.

Conclusion

Stress has long been suspected to play a role in the etiology of many diseases, and numerous studies have shown that stress can be immunosuppressive and hence may be detrimental to health. Moreover, glucocorticoid stress hormones are widely regarded as being immunosuppressive and are used clinically as anti-inflammatory agents. Due to a host of psychosociopolitical factors, stress has unfortunately become an increasing and inevitable part of people's lives. Stress is also a major factor during the

diagnosis, treatment, and follow-up of most diseases. However, as this discussion shows, some stress conditions (i.e., acute or short-duration stressors) produce immunoenhancement. A determination of the physiological mechanisms through which stress and stress hormones enhance or suppress immune responses may help our understanding and treatment of diseases thought to be affected by stress. Thus, future studies should aim at developing biomedical treatments that harness an individual's physiology to selectively enhance (during vaccination, wounding, infections, or cancer) or suppress (during autoimmune or inflammatory disorders) the immune response, depending on what is most beneficial for the patient.

Further Reading

- Altemus, M., Cloitre, M. and Dhabhar, F. S. (2003). Cellular immune response in adult women with PTSD related to childhood abuse. *American Journal of Psychiatry* **160**, 1705–1707.
- Dhabhar, F. S. and McEwen, B. S. (1996). Stress-induced enhancement of antigen-specific cell-mediated immunity. *Journal of Immunology* **156**, 2608–2615.
- Dhabhar, F. S. and McEwen, B. S. (1997). Acute stress enhances while chronic stress suppresses immune function in vivo: a potential role for leukocyte trafficking. *Brain, Behavior & Immunity* **11**, 286–306.
- Dhabhar, F. S. and McEwen, B. S. (1999). Enhancing versus suppressive effects of stress hormones on skin immune function. *Proceedings of the National Academy of Sciences USA* **96**, 1059–1064.
- Dhabhar, F. S. and McEwen, B. S. (2001). Bidirectional effects of stress & glucocorticoid hormones on immune function: possible explanations for paradoxical observations. In: Ader, R., Felten, D. L. & Cohen, N. (eds.) *Psychoneuroimmunology* (3rd edn., pp. 301–338). San Diego, CA: Academic Press.
- Dhabhar, F. S., Satoskar, A. R., Bluethmann, H., et al. (2000). Stress-induced enhancement of skin immune function: a role for IFN γ . *Proceedings of the National Academy of Sciences USA* **97**, 2846–2851.
- Dhabhar, F. S. and Viswanathan, K. (2005). Short-term stress experienced at the time of immunization induces a long-lasting increase in immunological memory. *American Journal of Physiology: Regulatory, Integrative & Comparative Physiology* **289**, R738–R744.
- Glaser, R. and Kiecolt-Glaser, J. K. (2005). Stress-induced immune dysfunction: implications for health. *Nature Reviews Immunology* **5**, 243–251.
- Pruett, S. B. and Fan, R. (2001). Quantitative modeling of suppression of IgG1, IgG2a, IL-2, and IL-4 responses to antigen in mice treated with exogenous corticosterone or restraint stress. *Journal of Toxicology and Environmental Health A* **62**(3), 175–189.
- Sapolsky, R. M. (2004). *Why zebras don't get ulcers* (3rd edn.). New York: W. H. Freeman.
- Sapolsky, R. M. (2005). The influence of social hierarchy on primate health. *Science* **308**, 648–652.
- Saul, A. N., Oberyzy, T. M., Daugherty, C., et al. (2005). Chronic stress and susceptibility to skin cancer. *Journal of the National Cancer Institute* **97**, 1760–1767.
- Sephton, S. and Spiegel, D. (2003). Circadian disruption in cancer: a neuroendocrine-immune pathway from stress to disease? *Brain, Behavior & Immunity* **17**, 321–328.
- Straub, R. H., Dhabhar, F. S., Bijlsma, J. W. J., et al. (2005). How psychological stress via hormones and nerve fibers may exacerbate rheumatoid arthritis. *Arthritis & Rheumatism* **52**, 16–26.
- Viswanathan, K. and Dhabhar, F. S. (2005). Stress-induced enhancement of leukocyte trafficking into sites of surgery or immune activation. *Proceedings of the National Academy of Sciences USA* **102**, 5808–5813.
- Wood, P. G., Karol, M. H., Kusnecov, A. W., et al. (1993). Enhancement of antigen-specific humoral and cell-mediated immunity by electric footshock stress in rats. *Brain, Behavior & Immunity* **7**, 121–134.
- Zalcman, S. and Anisman, H. (1993). Acute and chronic stressor effects on the antibody response to sheep red blood cells. *Pharmacology, Biochemistry & Behavior* **46**, 445–452.

Immune Response

P J Delves

University College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

Innate and Adaptive Immune Responses
 Recognition of Pathogens and Other Antigens
 Generation of the Immune Response
 The Anatomy of the Immune Response
 Protective Immune Response
 Avoidance of Harmful Immune Responses
 Neuroendocrine–Immune Interactions

Glossary

<i>Adaptive (acquired) immune response</i>	Immunity mediated by lymphocytes and characterized by antigen-specificity and memory.
<i>Complement</i>	A group of serum proteins, some of which act in an enzymatic cascade, producing effector molecules involved in inflammation (C3a, C5a), phagocytosis (C3b), and cell lysis (C5b–C9).
<i>Cytokines</i>	Low-molecular-weight proteins that stimulate or inhibit the differentiation, proliferation, or function of immune cells.
<i>Dendritic cell</i>	Refers to an interdigitating dendritic cell which is major histocompatibility complex (MHC) class II-positive, Fc receptor-negative, and presents processed antigens to T cells in the T cell areas of secondary lymphoid tissues.
<i>Granulocyte</i>	Myeloid cells containing cytoplasmic granules (i.e., neutrophils, eosinophils, and basophils).
<i>Innate immune response</i>	Immunity which is not intrinsically affected by prior contact with antigen, i.e., all aspects of immunity not directly mediated by lymphocytes.
<i>Primary immune response</i>	The relatively weak immune response which occurs upon the first encounter of naive lymphocytes with their specific antigen.
<i>Secondary immune response</i>	The much stronger immune response which occurs upon the second encounter of primed lymphocytes with their specific antigen.
<i>Toll-like receptors (TLRs)</i>	A family of pattern recognition receptors involved in the detection of structures associated with pathogens or damaged host tissues.

The immune system has evolved to protect the body from foreign pathogens. It does so using quite a number of different cell types, most of which arise from pluripotent hematopoietic stem cells in the bone marrow and are collectively referred to as leukocytes (white blood cells). These stem cells can enter one of two differentiation pathways, the myeloid pathway which gives rise to monocytes/macrophages, dendritic cells, granulocytes (neutrophils, eosinophils, basophils, and mast cells), erythrocytes, and platelets, and the lymphoid pathway which gives rise to lymphocytes and natural killer (NK) cells ([Figure 1](#)). The bone marrow is referred to as a primary lymphoid organ and is where many of these cell types, including the B lymphocytes (B cells), reach maturity. However, the precursors of the T lymphocytes (T cells) need to leave the bone marrow and travel to the thymus in order to become fully developed. Given that a thymus is obligatory for the generation of T cells it is also referred to as a primary lymphoid organ. Upon maturation the lymphocytes, together with specialized antigen-presenting dendritic cells, primarily locate to secondary lymphoid organs (mucosa-associated lymphoid tissue (MALT), lymph nodes, and spleen) where the local architecture permits optimal stimulation of the immune response.

In order to become activated, the cells of the immune system need to detect the presence of a pathogen. Structures that can potentially be recognized by the immune system are broadly referred to as antigens and comprise an almost unlimited variety of protein, carbohydrate, lipid, and small molecular moieties. Antigenic structures are therefore associated not only with foreign organisms (viruses, bacteria, fungi, and parasites) and molecules (ingested, inhaled, or absorbed substances) but also the body's own (self) molecules and cells. Despite this capacity of the immune system to recognize an almost infinite variety of antigens, it must avoid making potentially damaging responses against self-antigens and against beneficial commensal organisms, food antigens, and so on. Thus, the cells of the immune system normally respond when they detect structures associated with pathogenic organisms.

Innate and Adaptive Immune Responses

Most cells of the immune system react with similar vigor however many times they encounter the same pathogen. This type of response is referred to as innate immunity ([Figure 2](#)) and involves phagocytic cells and

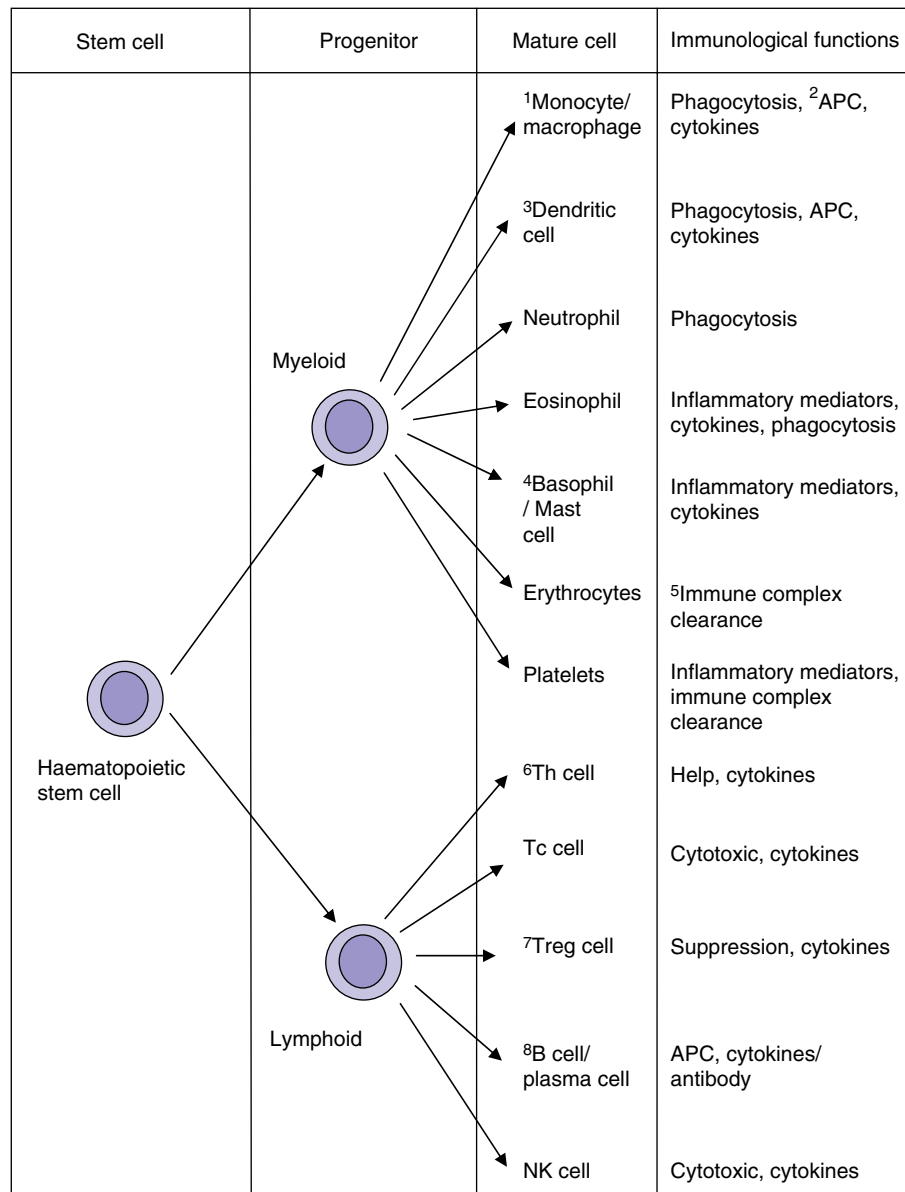


Figure 1 The major players in the immune response derived from hematopoietic stem cells in the bone marrow. Myeloid and lymphoid progenitors differentiate into a number of different cell types that, upon activation, mediate various immunological functions (examples of which are indicated). ¹Monocytes in the blood have functions that generally become more pronounced when these cells enter the tissues and differentiate into macrophages. ²APC, professional antigen-presenting cell that uses MHC class II molecules to present processed antigen to CD4 (usually helper) T cells. ³Dendritic cells (DC) come in a variety of forms. In addition to the myeloid dendritic cells, lymphoid dendritic cells (not shown) arise from lymphoid progenitors and have similar immunological functions to those of myeloid DC. Note that DCs are the only cell type that can act as APCs for naive T cells (i.e., T cells that have not previously encountered an antigen). There is also an entirely separate population of cells with a dendritic morphology that are present in secondary lymphoid organs and are referred to as follicular dendritic cells (FDC). FDC are nonphagocytic and lack the MHC class II required of professional APC, but instead present nonprocessed antigen (held on their surface in the form of immune complexes) directly to B cells. ⁴Blood basophils and tissue mast cells share functional properties but it is thought that they are independently derived rather than blood basophils giving rise to tissue mast cells. ⁵Immune complexes are composed of antigen bound to antibody and also often including complement components. ⁶Th cells assist other cells in the immune system partly by secreting immunostimulatory cytokines but also using cell-contact dependent mechanisms. ⁷Treg cells are heterogeneous, some inhibit other cells in the immune system by secreting cytokines with immunosuppressive properties while others use cell-contact dependent mechanisms to suppress immune responses. ⁸B cells can act as professional APCs for antigen-experienced Th cells, and when activated can themselves secrete cytokines. When B cells differentiate into plasma cells their role is to secrete large amounts of antibody.

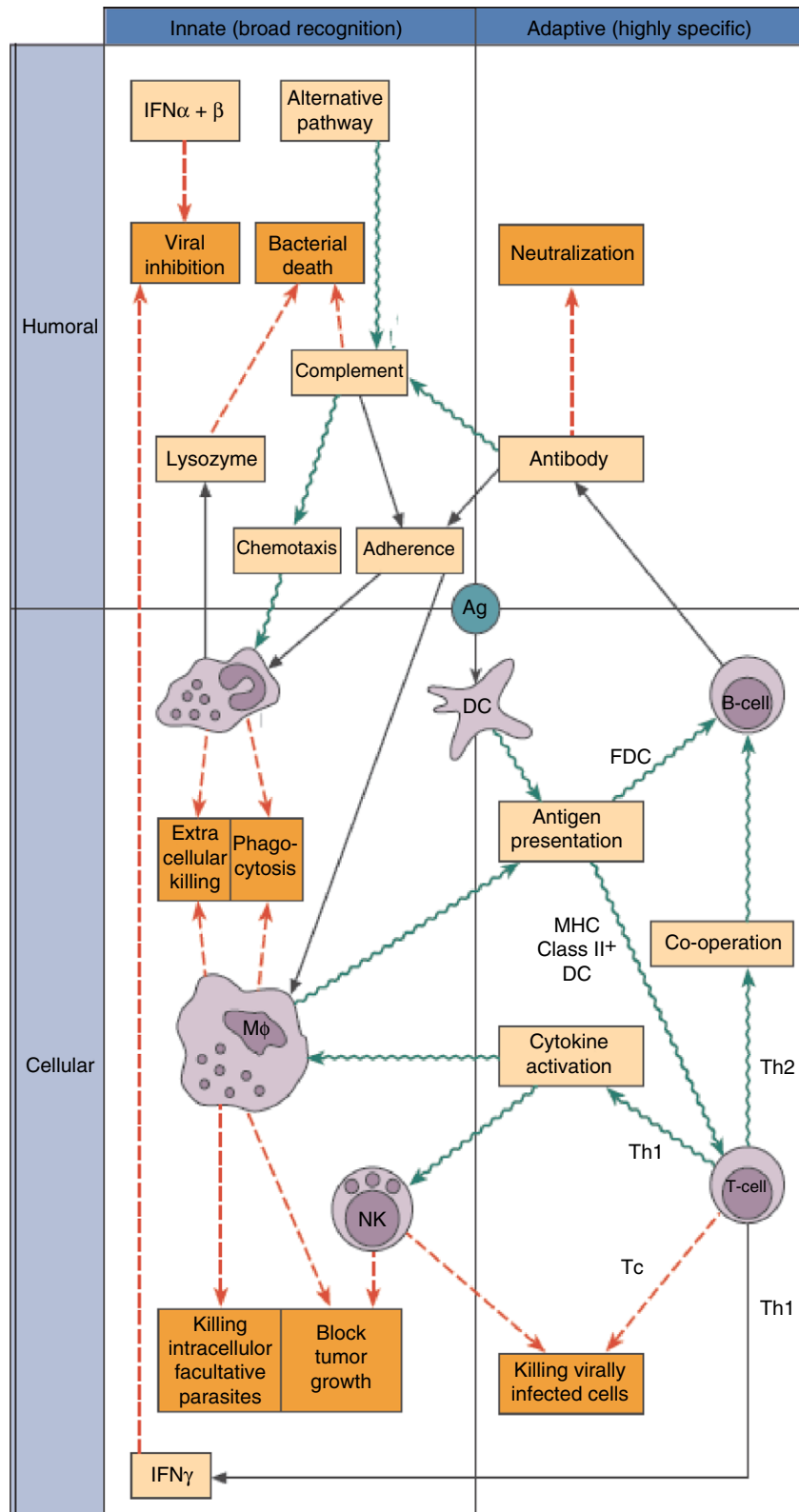


Figure 2 Some innate and adaptive immune responses. Humoral responses are those that are mediated by soluble molecules such as complement and cytokines in the innate response and antibody in the adaptive response. MHC class II⁺ dendritic cells act as professional APC to present processed antigen to T cells, while the entirely different follicular dendritic cells (FDC) can present native antigen to B cells. Ag, antigen; IFN γ , gamma interferon; M ϕ , macrophage. Green wavy lines indicate processes that activate, whereas red dotted lines indicate processes that inhibit or kill. (Modified with permission from Delves PJ, Martin SJ, Burton, DR, and Roitt IM. (2006). *Roitt's essential immunology*, 11th edn. Oxford: Blackwell Publishers).

NK cells, as well as basophils and mast cells. The phagocytic cells comprise short-lived neutrophils present in both the circulation and in the tissues, blood monocytes which differentiate into the more actively phagocytic macrophage upon entering the tissues, and dendritic cells in the tissues. Eosinophils are also weakly phagocytic but they are more adept at killing organisms that are too large to be engulfed and onto which they release toxic molecules. Basophils in the blood, and the functionally related mast cell in the tissues, release inflammatory mediators such as histamine, prostaglandins, and leukotrienes. The NK cells, present in both the circulation and the tissues, kill infected or malignant cells. Thus, the innate immune response can call upon a number of fundamentally different cell types. In contrast, the adaptive (also referred to as acquired) immune response is mediated by lymphocytes. These cells are exquisitely specific for individual antigens and when they encounter 'their' antigen for the first time they mount a primary immune response that produces not only cells and molecules to deal with the threat at hand but also memory cells. The latter constitute a population of lymphocytes that are expanded in number and have the capacity to generate the qualitatively and quantitatively superior secondary immune response. Of the two main populations of lymphocytes, B lymphocytes (B cells) and T lymphocytes (T cells), it is the B cells that are responsible for producing antibody molecules. T cells can be divided into three main subpopulations. Helper T cells (Th) assist, and regulatory T cells (Treg) can suppress, other cells in the immune response. Cytotoxic T cells (Tc), also referred to as cytotoxic T lymphocytes (CTL), kill infected or foreign eukaryotic cells. Most Th and Treg express a molecule called CD4 on their cell surface, whereas most Tc express CD8 instead.

Recognition of Pathogens and Other Antigens

In order to mount an immune response the cells of the immune system need to detect the presence of a threat. They do this using cell surface receptors which can be grouped into two main types: broadly reactive receptors and antigen-specific receptors.

Broadly Reactive Receptors

The cells involved in innate immune responses have a large number of receptors which are each encoded by entirely separate genes and which recognize structures that are common to many different organisms. These types of receptors are referred to as pattern recognition receptors (PRR) and their ligands

designated pathogen-associated molecular patterns (PAMPs). Recently PRRs have been found to also be present on the cells of the adaptive immune response (i.e., lymphocytes), and some PRRs have been shown to recognize certain endogenous (self) ligands as well as recognizing PAMPs. Although the receptors are highly specific for the structures they recognize, their ligands are molecules that are widely distributed amongst many different species of pathogen. Examples include the toll-like receptors (TLRs; such as TLR3 which recognizes viral double-stranded ribonucleic acid (RNA) and TLR5 which detects bacterial flagellin) and the mannose receptor. The latter binds to the terminal mannose structure characteristic of prokaryotes but which is always covered by other sugars in glycoproteins of eukaryotic origin.

Antigen-Specific Receptors

Lymphocytes possess highly antigen-specific receptors on their cell surface. The B-cell receptor for antigen is a transmembrane version of the antibody molecule, while the analogous receptor on T lymphocytes consists of an $\alpha\beta$ or $\gamma\delta$ T-cell receptor (TCR) heterodimer. These receptors contain variable regions which, as their name suggests, vary from one receptor to another and which confer antigen specificity. All the antigen receptors on an individual lymphocyte will possess the same variable region sequence, but those on a different lymphocyte will bear a different variable region and hence specificity for a different antigen. The receptors are generated by a unique gene recombination mechanism in which the nucleotide sequence encoding the variable region of the protein is constructed from more than one gene. Thus, one gene sequence is picked in a fairly random fashion from each of two (for immunoglobulin (Ig) light chains and for TCR α or γ chains) or three (for Ig heavy chains and for TCR β or δ chains) gene pools referred to as variable (V), diversity (D, if utilizing all three gene pools), and joining (J) gene segments and these sequences spliced next to each other ([Figure 3](#)). The Ig genes undergo recombination in the precursors of B lymphocytes while TCR genes undergo recombination in early T cells. In each lymphocyte a different recombination occurs and, together with a variety of additional molecular mechanisms which further increase diversity, leads to an almost infinite variety of antigen-specific receptors being generated. Therefore one lymphocyte will have antigen receptors that are specific for a particular component of streptococcus, another lymphocyte will be specific for an antigen from staphylococcus, another for a part of the influenza virus, and so on. There will only be a handful of lymphocytes specific for each antigen but collectively

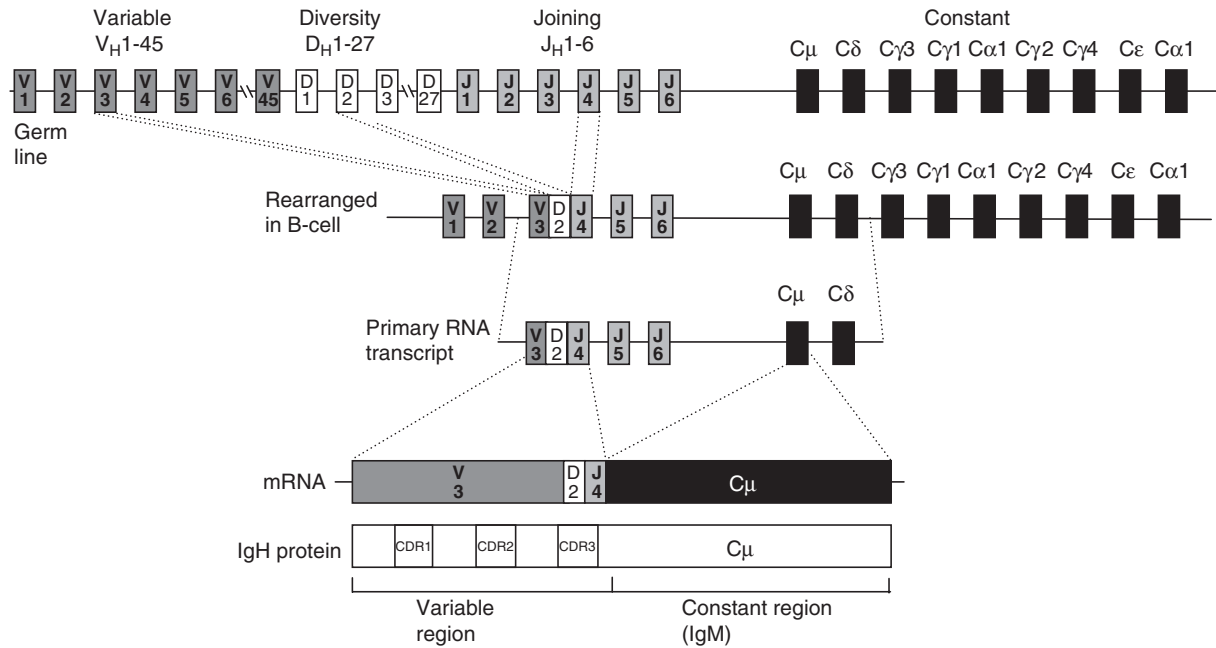


Figure 3 Antigen receptor gene recombination. In this example, immunoglobulin heavy chain (IgH) gene recombination is illustrated, but immunoglobulin light chain genes and the TCR genes undergo similar recombination processes in order to create a large number of different variable region sequences from a relatively small gene pool. Here the B cell has recombined the V3, D2, and J4 immunoglobulin heavy chain gene segments (not drawn to scale) and is utilizing the C_μ constant region gene segment from a primary RNA transcript that also includes C_δ . Alternative splicing of the primary RNA transcript is used to also produce IgD with the same antigen specificity. A different B cell might recombine V42, D16, and J6, etc., and would therefore have a different heavy chain variable region sequence. Recombination to provide variability for the immunoglobulin light chain variable region (not shown) is brought about by similar mechanisms using separate gene loci containing only V, J, and C segments. The heavy chain constant region gene transcription can include transmembrane exons. The heavy and light chain proteins polymerize into the 4 chain ($2H + 2L$) immunoglobulin molecule, which then sits in the cell surface membrane to function as the antigen-specific B-cell receptor. In plasma cells the transmembrane exons are not utilized and a secreted version of the antibody, with identical antigen-specificity, is produced. In the case of IgM the antibody polymerizes into pentamers prior to secretion. Class switching occurs later in the immune response with C_μ and C_δ usage being replaced with, for example, $C_\gamma1$ which specifies the IgG1 subclass of antibody but utilizing the same variable region and thus with the same antigen specificity. The areas of the variable region in both the heavy and the light chain which make contact with the antigen are referred to as complementarity determining regions (CDR).

all the lymphocytes will provide cover against virtually all antigens that might be encountered.

Generation of the Immune Response

Innate responses are initiated when the PRRs on the surface of phagocytic and other cells detect the presence of PAMPs. Additionally, foreign antigens will very often become coated with proteins that are generated by the complement system, a cascade of molecules activated during immune responses. Later on in the immune response the antigen will additionally be coated with antibody. Among the many different types of molecules on the surface membrane of immune system cells are receptors for complement molecules and for the nonantigen-binding end (Fc) of the antibody molecule. Microorganisms coated with antibody and/or complement are said to be opsonized for phagocytosis, the antibody and complement linking the organism to the phagocyte. Complement carries out other functions; it acts

as a chemotactic factor for phagocytes, it stimulates the release of inflammatory mediators from mast cells, and can mediate lysis of certain strains of bacteria.

The response of lymphocytes to an encounter with an antigen is somewhat different to that of the cells of the innate immune response. Whereas the latter go straight ahead with the task required, be it phagocytosis or the release of inflammatory mediators, the lymphocyte undergoes rapid cell division. The clone of lymphocytes with the appropriate antigen-specific receptor thus become greatly amplified in number during the adaptive immune response. T cells (at least the main population of T cells which bear an $\alpha\beta$ TCR) cannot recognize antigens directly, but instead see antigen that has been processed and presented. With respect to Th cells which have not previously encountered their specific antigen, i.e., naive T cells, this involves dendritic cells which phagocytose or endocytose antigen, then demolish its protein components using proteases which generate short peptides

of around 15 amino acids in length within the endosomal compartment of the cell. While still inside the endosomes these antigenic peptide fragments are combined with major histocompatibility complex (MHC) molecules of the class II type, and the complex of antigen peptide plus MHC class II is transported to the cell surface of the dendritic cell. It is the combination of peptide–MHC that is recognized by the TCR on the T cell, with the dendritic cell also expressing other cell-surface molecules such as CD80 and CD86 which supply the potent co-stimulation obligatory for the activation of naive T cells. Both macrophages and B cells also express peptide-bearing MHC class II molecules on their cell surface but due to their less profound ability to provide high levels of co-stimulation they only activate antigen-experienced T cells. Dendritic cells, macrophages, and B cells are collectively referred to as professional antigen-presenting cells (APC), a euphemism for cells which present antigen to CD4 T cells using MHC class II molecules. However, virtually all cells in the body are able to present peptide antigens using MHC class I molecules. In this case the peptides are usually generated from antigens present within the cytosol of a cell, including foreign antigens such as viruses. Cytosolic proteolysis for subsequent antigen presentation utilizes a modified version of the proteasome enzyme complex that carries out the normal cellular degradative processes. Peptide–MHC class I complexes expressed on the surface of infected cells are recognized by CD8+ Tc cells. Thus CD4+ T cells use their $\alpha\beta$ TCR to recognize peptides presented by MHC class II molecules on professional APC whereas CD8+ T cells use their $\alpha\beta$ TCR to recognize peptides presented by MHC class I molecules on infected cells.

The MHC is given a different name in different species, in humans it is referred to as human leukocyte antigen (HLA; although present not only on leukocytes) while in mice it is called H-2. In humans there are three main MHC class I gene loci, HLA-A, HLA-B, and HLA-C, and three class II loci, HLA-DP, HLA-DQ, and HLA-DR. In mice the main class I loci are H-2K and H-2D while the class II loci are referred to as I-A and I-E. The MHC molecules are encoded by the most polymorphic of all of the genes in the genome, and thus can vary enormously from individual to individual. It is this MHC polymorphism that is largely responsible for the immune system recognizing and subsequently rejecting organs or tissues from other members of the same species (allografts). The MHC genes are co-dominantly expressed and therefore each individual expresses two polymorphic variants of HLA-A, two of HLA-B, and so on. Certain variants are more common in the population than others and with common variants such as HLA-A2

in Caucasians it is quite frequently the case that an individual will be homozygous at that particular genetic locus.

The Anatomy of the Immune Response

The cells of the innate response such as phagocytic cells can act individually at the site of an infection. They are attracted to such sites by chemotactic factors, and their passage from the circulation out into the tissues is facilitated by the upregulated expression of adhesion molecules on vascular endothelial cells, ligands for which are expressed on the leukocyte cell surface.

As discussed previously, the Th cells of the adaptive response need to interact with professional APC in order to be activated. This requires cell–cell contact as the TCR on the T cell must bind to the peptide–MHC on the APC. The Th cells are also required to be in close proximity with the cells that will receive their help, which is provided by a combination of soluble molecules such as cytokines and by interactions between cell surface molecules on the Th cell and on the cell being helped. There are different subsets of cytokine-producing Th cells (e.g., Th1 cells, Th2 cells) and depending upon which subset is involved, the cytokine-mediated help that is given may facilitate antibody production by B cells, activation of Tc to kill infected cells, or enhancement of macrophage function ([Figure 2](#)). Given that Th, Tc, and B cells are all antigen-specific, that there are only a handful of cells specific for a given antigen, and that T cells require antigen to be presented to them, there is a need for a meeting place where the relevant cells can get together. This is the role of the secondary lymphoid organs; the spleen, lymph nodes, and MALT. These are the locations in which adaptive immune responses are initiated. Because the majority of pathogens will enter the body by ingestion or inhalation the MALT collectively provides a critical first line of defense against infection and acts as a major effector site for protective immune responses.

Protective Immune Response

The immune response is aimed at destroying foreign pathogens. Engulfment of microorganisms by phagocytes leads to their elimination when a wide variety of microbicidal compounds including toxic metabolites of oxygen such as hydroxyl radicals and enzymes such as lysozyme kill the invader. Eosinophils release toxins to achieve extracellular killing. Mast cells and basophils produce a plethora of molecules which collectively facilitate an acute inflammatory response with recruitment of leukocytes to the

site of the infection. Infected cells, and some tumor cells, are dealt with by Tc and NK cells which can both induce apoptotic cell death in the target cell. Many cells in the body are capable of releasing cytokines but Th cells are particularly adept at doing so. In addition to providing help for other cells in the immune system, some cytokines can be directly cytotoxic (e.g., tumor necrosis factor (TNF)). B cells produce antibody molecules which are collectively found in their secreted form at high levels ($10\text{--}15\text{ mg ml}^{-1}$) in the circulation and, with the exception of the physically large pentameric IgM, also readily diffuse into the tissues. Antibodies are able to directly neutralize bacterial toxins, prevent viruses binding to cells, agglutinate microorganisms and thereby immobilize them, activate complement, opsonize microorganisms for phagocytosis, and mediate antibody-dependent cellular cytotoxicity (ADCC) in which leukocytes of various types become linked by antibody to infected cells and then kill them.

Avoidance of Harmful Immune Responses

The immune system can potentially react against the body's own components, resulting in immunity to self (autoimmunity). This only becomes a problem if the autoimmune reactions cause pathology, in which case the individual develops an autoimmune disease. Approximately 3–5% of individuals develop autoimmune disease at some time in their life. There are an extremely large number of different diseases in which autoimmune responses are thought to play a role, with the most common including psoriasis, thyroid autoimmune diseases (primarily Hashimoto's thyroiditis and Graves' disease), rheumatoid arthritis, vitiligo, systemic lupus erythematosus, Sjögren's syndrome, celiac disease, Crohn's disease, pernicious anemia, ankylosing spondylitis, type I (insulin-dependent) diabetes, and multiple sclerosis. The fact that the vast majority of individuals do not develop pathogenic autoimmune responses is due to the phenomenon of immunological tolerance. Encounter with self-antigen early on in the development of a lymphocyte leads to its clonal deletion by apoptosis. This occurs in the thymus for T cells and in the bone marrow for B cells. The T cells undergo thymic education involving positive and negative selection events in the thymus. Cells are positively selected if the random rearrangement of their TCR genes generates a receptor capable of recognizing peptides presented by that individual's polymorphic variants of MHC (self-MHC). However, if the TCR recognizes peptides derived from self antigens then a subsequent negative selection step will induce apoptosis of the autoimmune T cell. Thus, the only T cells which leave

the thymus are ones that recognize foreign peptides presented by self-MHC. A somewhat similar process of negative selection occurs for B cells in the bone marrow, although their antigen receptor (the antibody molecule) recognizes antigen directly and therefore there is no requirement for positive selection for recognition of self-MHC molecules. The clonal deletion of T and B cells is not foolproof and some of the cells that leave the primary lymphoid organs are self-reactive. Therefore there are backup mechanisms to prevent harmful autoimmunity arising. Pathogenic foreign antigens are normally seen in the context of co-stimulatory activity whereas self-antigens do not provide co-stimulatory signals. When the lymphocyte receives a signal only through its antigen receptor in the absence of co-stimulation the result is not cell activation but rather cell inactivation, a process referred to as anergy. Unlike clonal deletion where the cell is eliminated, this leaves anergic cells which potentially can be re-activated. A further level of control rests with Treg cells that act to suppress any potentially harmful self-reactive cells. There are a number of different types of Treg cells, some of which arise naturally in the thymus and exert their effect via poorly understood cell contact-dependent mechanisms. Other Treg cells are induced to secrete cytokines that can mediate immunosuppressive activity, such as interleukin (IL)-10 and transforming-growth factor- β (TGF β).

While the avoidance of harmful immune responses to self-antigens is something that operates with a fair degree of success, the capacity of the body to prevent allergic responses to what should be innocuous foreign antigens would appear to be rather less effective. A substantial minority of individuals (estimates vary, but perhaps even as many as 45% of children and 30% of adults) suffer from allergies to varying degrees. This frequently involves atopy, a capacity to produce raised levels of the IgE class of antibody against the allergen. The class of antibody is of paramount importance here because the IgE becomes bound to the surface of mast cells and basophils by virtue of the fact that these cells possess high-affinity receptors for the Fc region of IgE (Fc ϵ R). If the IgE has specificity for environmental antigens (allergens) these cross-link the IgE molecules. As a consequence, cross-linking of the Fc ϵ Rs occurs leading to degranulation of the cells with the release of inflammatory mediators such as histamine.

Neuroendocrine-Immune Interactions

Although immune responses are driven by encounter with antigen and subject to regulation by a variety of cells and molecules which are clearly components of

the immune system, it is also known that factors such as nutrition, hormones, and neuropeptides can modify the immune response. Stressors such as exam anxiety, restraint, and noise are known to influence various immune functions including lymphocyte proliferation, IgA secretion, phagocytosis, and the activity of NK cells. Receptors for endorphins, enkephalins, acetylcholine, β -adrenoceptor agents, corticosteroids, insulin, growth hormone, estradiol, testosterone, and prolactin have all been described on various types of immune cell and their expression presumably infers some functional significance. To give an example, the cytokines IL-1, IL-6, and TNF act via the hypothalamic-pituitary-adrenal (HPA) axis to stimulate the release of glucocorticoids which are then capable of downregulating the activity of both Th1 cells and macrophages. Androgens and glucocorticoids generally suppress, while growth hormone, thyroxine, and insulin generally stimulate, immune responses. Lymphoid tissues, including the thymus, spleen, and lymph nodes, are richly innervated by norepinephrine-releasing sympathetic neurons. There is evidence from studies in knockout mice that in the absence of this neurotransmitter there is impaired production of the Th1-derived cytokines interferon (IFN) γ and TNF in response to challenge with antigen. Dendritic cells bear β -adrenoceptors and it has also been suggested that sympathetic nerves may be involved in controlling the migration of these APCs from the sites of inflammation to the local draining lymph nodes.

See Also the Following Articles

Apoptosis; Cytokines; Cytotoxic Lymphocytes; Glucocorticoids, Overview; Immune Cell Distribution, Effects of Stress on; Immune Function, Stress-Induced Enhancement; Immune Suppression; Immune System, Aging; Immunity; Infection; Lymph Nodes; Leukocyte Trafficking

and Stress; Lymphocytes; Macrophages; Natural Killer (NK) Cells; Neuroendocrine Systems; Neuroimmunomodulation; Primate Models, Behavioral-Immunological Interactions; Prostaglandins; Serotonin; Thymus; Vasoactive Peptides; Inflammation; Organ Transplantation, Stress of; Vaccination.

Further Reading

- Abbas, A. K. and Lichtman, A. H. (2006). *Basic immunology: functions and disorders of the immune system* (2nd edn.). Philadelphia, PA: Saunders.
- Ader, R. (ed.) (2006). *Psychoneuroimmunology* (4th edn.). San Diego: Academic Press.
- Daruna, J. H. (2004). *Introduction to psychoneuroimmunology*. San Diego: Academic Press.
- Delves, P. J., Martin, S. J., Burton, D. R., et al. (2006). *Essential immunology* (11th edn.). Oxford: Blackwell Science.
- Delves, P. J. (2006). Biology of the immune system. In: Beers, M. H., Porter, R. S., Jones, T. V., et al. (eds.) *The Merck manual* (18th edn, pp. 1320–1331). Whitehouse Station, NJ, USA: Merck.
- Goldsby, R. A., Kindt, T. J., Osborne, B. A., et al. (2003). *Immunology* (5th edn.). New York: WH Freeman.
- Padgett, D. A. and Glaser, R. (2003). How stress influences the immune response. *Trends in Immunology* **24**, 444–448.
- Reiche, E. M., Nunes, S. O. and Morimoto, H. K. (2004). Stress, depression, the immune system, and cancer. *Lancet Oncology* **5**, 617–625.
- Rhen, T. and Cidlowski, J. A. (2005). Mechanisms of disease: antiinflammatory action of glucocorticoids – new mechanisms for old drugs. *New England Journal of Medicine* **353**, 1711–1723.
- Shepherd, A. J., Downing, J. E. and Miyan, J. A. (2005). Without nerves, immunology remains incomplete – *in vivo* veritas. *Immunology* **116**, 145–163.
- Sternberg, E. M. (2006). Neural regulation of innate immunity: a coordinated nonspecific host response to pathogens. *Nature Reviews Immunology* **6**, 318–328.
- Vedhara, K. and Irwin, M. (2005). *Human psychoneuroimmunology*. Oxford: Oxford University Press.

Immune Suppression

P Prolo and F Chiappelli

UCLA School of Dentistry, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by F Chiappelli and D Hodgson, volume 2, pp 531–535, © 2000, Elsevier Inc.

Allostasis

Type I and Type II Allostatic Responses

The Allostatic Modulation of Immune Suppression

T Cell Anergy and Apoptosis

A New Strategy: Molecular Cartography

Glossary

<i>Allostasis</i>	Psychobiological process that brings about stability through change of state consequential to stress. The allostatic response encompasses a range of psychobiological outcomes that reflect physiologically important changes in the balance between the psychoneuroendocrine and the immune systems. Allostasis overload leads to statistically significant and clinically relevant damage to the organism due to excessive energy expenditure in the response to stress.
<i>Anergy</i>	A state of immune unresponsiveness, induced as soon as the antigen receptor of the T cell is stimulated, which stops T cell responses waiting for second signal from the antigen-presenting cell.
<i>Apoptosis</i>	The process of programmed cell death needed to eliminate cells that represent a threat to the integrity of the organism.
<i>Circadian</i>	A term derived from the Latin <i>circa diem</i> meaning about a day, refers to biological rhythms that have a cycle of about 24 hours.
<i>Molecular cartography</i>	The science of recognizing and identifying multifaceted and intricate array of gene interaction and gene products that characterize the function of each individual cell in the context of cell–cell interaction, tissue, and organ function.

The renowned XIX century French physiologist, Claude Bernard (1813–78) first proposed that defense of the internal milieu (*le milieu intérieur*, 1856) is fundamental to physiological regulation, from when the phrase homeostasis originated. Walter Cannon (1871–1945) proposed that organisms engage in a

dynamic process of adjustment of the physiological balance of the internal milieu in response to changing environmental conditions. Hans Selye (1907–82) established the cardinal points of the generalized stress response in his demonstration of concerted physiological responses to stressful challenges.

The immune response is finely regulated at the molecular, cellular, and physiological levels. Bacteriokines (i.e., bacterial cytokine inducers), produced by infectious pathogens, contribute to the control of the pathological inflammatory response. This article focuses on the homeostasis of the immune response from the perspective of the psychoneuroendocrine and cellular events that contribute to the suppression of immunity in response to stressors.

Allostasis

Allostasis refers to the psychobiological process that brings about stability through change of state consequential to stress. Psychoemotional stress can be defined as a perceived lack of fit of one's perceived abilities and the demands of the environment (i.e., person/environment fit). The perception of stress leads to anxiety, irritability and anger, tension, depressed moods, fatigue, a compendium of symptoms referred to as the sickness behavior (e.g., lethargy, nausea, and fever), and to a set of other bodily manifestations, including perspiration, blushing or blanching of the face, increased heart rate or decreased blood pressure, and intestinal cramps and discomfort, and suppressed immune resistance to antigenic challenges.

Sterling and Eyer (1988) defined allostasis as the set of mind–body systemic events aimed at regulating the recovery from stress. Heterostasis came to describe the situation where the demands upon the organism exceed its physiological capacity.

The process of allostatic regulation signifies the recovery and maintenance of internal balance and viability amidst changing circumstances consequential to stress. It encompasses a range of behavioral and physiological functions that direct the adaptive function of regulating homeostatic systems in response to challenges. The allostatic load refers to the set of physiological, psychological, and pathological side effects of the allostatic process, whereas allostatic overload signifies psychobiological collapse, and is associated with a variety of health outcomes systemically.

Type I and Type II Allostatic Responses

The allostatic response encompasses a range of psychobiological outcomes that reflect changes in the

balance between the psychoneuroendocrine and the immune systems. Allostasis overload leads to statistically significant and clinically relevant damage to the organism due to excessive energy expenditure in the response to stress.

Type I

Type I allostasis utilizes stress responses as a mean of self-preservation by means of developing and establishing temporary or permanent adaptation skills. The organism aims at surviving the perturbation in the best condition possible, and at normalizing the normal lifecycle. An example of Type I allostatic response is recurrent aphthous stomatitis (RAS, canker sores), a stress-associated T cell-dependent pathology of the gingival mucosa. RAS lesions are characterized with increased Th1 patterns of cytokine *in situ* (e.g., interleukin (IL)-2). Ongoing work confirms a statistically significant correlation between the clinical incidence of RAS, based on stringent clinical and diagnostic criteria, and scores on a version of the short form-36 health survey questionnaire (SF-36) adapted for dentistry.

Other examples of Type I allostatic response manifested in the oral cavity include teeth grinding: tooth pain that cannot be attributed to carries, cracked tooth syndrome, pulpitis or other circumferential inflammatory process (e.g., gum disease), temporomandibular disorder and pain, and pain at palpation of the masticatory muscles, altered dental anatomy at the coronal surfaces. When the subjects are followed up prospectively, changes in vertical distance of occlusion and the angle of Spee that refers to the curved line that follows the occlusal surface of the posterior teeth of the mandible and is viewed from a sagittal plane are also noted.

Examples of systemic Type I allostatic response include the physiological responses to electromagnetic field. In a paired double-blind study, healthy young adults were exposed to electromagnetic field during controlled sleep. A sharp reduction in secreted melatonin was observed, which was transient and not correlated with any alterations of note in immune responses. The observed twofold increase in IL-2 serum levels failed to alter significantly natural killer cell activity.

Type II

In Type II allostatic load, the stressful challenge is excessive, sustained, or continued, and drives allostasis chronically. An escape response cannot be found. The psychobiological and pathological responses to stress become chronic (e.g., depression, anxiety, hypercortisolemia, and glucocorticoid resistance), with consequential erosion in health and immune surveillance mechanisms.

Oral Lichen Planus

Lichen planus of the oral mucosa (oral lichen planus; OLP) is one example of a Type II allostatic response with evident manifestations of immune suppression in the oral cavity. This form of oral lichenoid lymphocytic mucosal inflammation involves basal cell lysis, and lymphocyte transmigration into the epithelial compartment, and brings about a statistically significant increase in incidence of oral squamous cell carcinoma, which implicates OLP as a premalignant lesion. This chronic inflammatory disorder afflicts 2–3% of the population, and may be in part attributed to aging-associated decrease in the subject's ability to recover from the psychoneuroendocrine-immune changes consequential to stress since OLP is prevalent in individual over 50 years of age. Men are at risk at an earlier age than women (40–49 years with men versus 50–59 years with women). OLP lesions are exacerbated with nonsteroidal anti-inflammatory drugs, antirheumatic drugs, angiotensin-converting enzyme inhibitors, antihypertensive drugs, antibiotics, or β -blockers. These pharmacologic treatments share the commonality of suppressing one or another arm of cellular immunity, suggesting a pathological parallel between OLP lesions and mucositis. The nervous system may play an important role *in vivo* in regulating cellular immune responses in OLP lesions, as suggested by the suggestive alignment of autonomic innervation with lymphocytes (Figure 1). Immune suppression is modulated by sympathetic innervation, and by the cholinergic anti-inflammatory pathway, which inhibits proinflammatory cytokine production and secretion. The relationship between OLP and immune suppression is confirmed by the observation that it appears to result from an abnormal immune response to an unidentified causative agent. Over 75% of invading T cells express the memory phenotype (CD45RO+), the β 1 integrin, CD29+, which renders them essentially sessile. A large proportion of the invading cells also appear phenotypically to be activated Treg (CD25+CD69+). Altered antigenicity of cell-surface epithelial cells causes them to be recognized as immunogenic, in a manner akin to an autoimmune disorder. Human hepatitis C virus (HCV) seems to be a predisposing factor for OLP, and comorbidity of HCV infection with OLP is noted in association with the HLA-Dr6 allele. The genetic propensity for OLP is emphasized by the observation that expression of HLA-Bw57 favors OLP, whereas HLA-Dq1 expression provides resistance against it.

Chronic Pain

A systemic example of Type II allostasis is chronic pain. Patients with fibromyalgia demonstrate altered

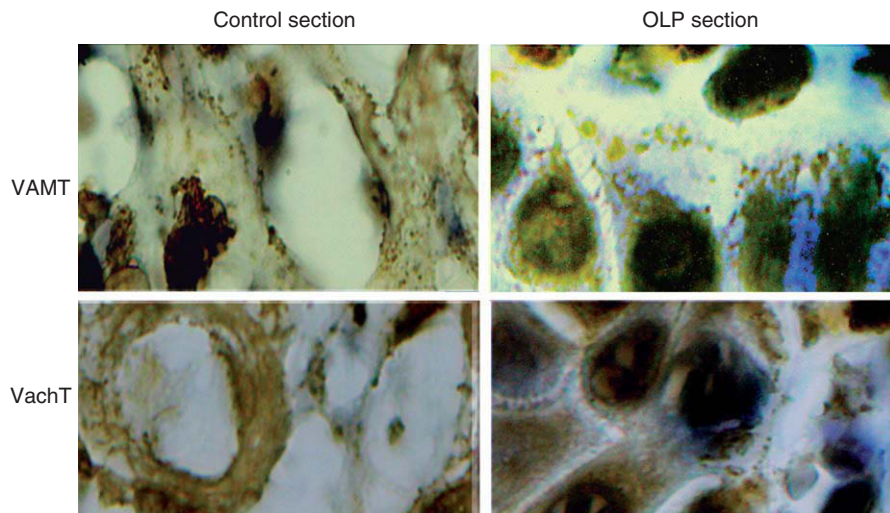


Figure 1 Double-label immunohistochemical staining of erosive OLP lesions with CD3 for T cells, and vesicles of acetylcholine transport (VachT; lower panels) and vesicles for monoamine transport (VAMT; upper panels) for autonomic innervation (magnification $\times 100$, oil immersion).

function of β -adrenoceptors. Activation of G-protein coupled receptors, such as the adrenoceptors, leads to an increase in the intracellular level of cAMP. Levels of cAMP serve to assess, albeit indirectly, the functional state of these receptors in immune cells. Patients with fibromyalgia were compared to sex- and age-matched control subjects, and basal and stimulated intracellular cAMP levels were measured, using the β -agonist, isoproterenol (10^{-3} M– 10^{-5} M, following a 5 min preincubation). Basal cAMP levels were shown to be elevated in patients with fibromyalgia (3.02 ± 0.44 pmol/ 10^6 cells, $P = 0.124$), compared to control subjects (2.26 ± 0.39 pmol/ 10^6 cells). Stimulation with isoproterenol at 10^{-3} M significantly increased intracellular cAMP in both groups; but isoproterenol at 10^{-5} M failed to produce this outcome in the patient group ($P = 0.74$), compared to the stimulation by isoproterenol obtained in the control group ($P = 0.012$). Taken together, these data show that patients with fibromyalgia harbor diminished adrenoceptor function in circulating immune cells.

Mucositis

Another example of Type 2 allostasis may be the clinical condition of mucositis. Oropharyngeal mucositis is a common and treatment-limiting side effect of cancer therapy. Oral and gastrointestinal mucositis can affect up to 100% of patients undergoing high-dose chemotherapy and hematopoietic stem cell transplantation, 80% of patients with malignancies of the head and neck receiving radiotherapy, and over 50% of patients receiving chemotherapy. Oral mucositis can be so severe as to lead to the need to

interrupt or discontinue the therapy protocol, and to hinder cure of the primary disease. Mucositis leads to immune suppression, and to an increased risk of local and systemic infection, opportunistic infections, and mortality due to sepsis. It is a critical condition since in 2004 alone, close to 1.5 million Americans were diagnosed with cancer and received some combination therapies for this disease, including surgery, radiation, chemotherapy, immunotherapy, and/or cell transplantation, and mucositis is the principal oral complication from cancer therapies.

The Allostatic Modulation of Immune Suppression

One important regulatory mechanism in allostasis is the interactive neuroendocrine-immune loop. Stress alters the regulation of both the sympathetic and the parasympathetic branches of the autonomic nervous system, and modulates immune suppression. Autonomic activation, and the elevation of hormones, including those produced by the hypothalamic-pituitary-adrenal axis, play a pivotal role in regulating cell-mediated immune surveillance mechanisms, including the production of cytokines that control inflammatory and healing events.

The experience of stress and of traumatic events leads to the release of glucocorticoids, which regulate cellular immune activity *in vivo* systemically and locally: they block the production of proinflammatory cytokines (e.g., IL- 1β , IL-6) and of Th1 cytokines (e.g., IL-2) at the molecular level *in vitro* and *in vivo*, but have little effects upon Th2 cytokines (e.g., IL-4). The activity of the proinflammatory cytokine IL- 1β

is blocked by IL-1 receptor antagonists (IL-1Ra, sIL-1RII) in combination with the soluble IL-1 accessory protein (IL-1R AcP). This results in a long-term beneficial effect in chronic inflammatory diseases. The net effect of challenging immune cells with glucocorticoids is to suppress immune T cell activation and proliferation, while maintaining antibody production. The secretion of glucocorticoids by the adrenal cortex is under the control of the anterior pituitary adrenocorticotropin hormone (ACTH). Immune challenges release proinflammatory cytokines (e.g., IL-1 β , IL-6), which induce hypothalamic secretion of the ACTH inducing corticotropin releasing hormone (CRH). It follows that the consequences of stress are not uniform. The immune-physiological impact of stress may be significantly greater in certain people, compared to others. The impact of stress is dynamic and multifaceted, and the same person may exhibit a variety of manifestations of immune suppression in response to stress with varying degrees of severity at different times. The outcome of stress can be multivalent, and individuals may position themselves along a spectrum of allostatic regulation, somewhere between Type I and Type II allostasis.

Several experimental protocols have been characterized for testing allostasis in animal and human research. Human subjects under stress (e.g., caregivers of elderly patients with dementia and spouses of military personnel in active combat duty) can be subjected to cognitive, sensory, or physiological challenges aimed at studying the allostatic process of immune-physiological adaptation to the challenge in addition to basal states of the stress response.

The Stroop Test

The Stroop color-word interference test is an example of a cognitive challenge that was originally developed in 1935 to examine patterns of cognitive decline. The Stroop challenge is measured as reaction time, and is high in patients with mild to moderate aging-related dementia of the Alzheimer's type (DAT). Patients with more severe dementia show less Stroop interference effect, when the reaction time is adjusted for color naming performance. These observations confirm the need and the importance of partitioning out underlying deficits for the understanding of complex cognitive processes in aging. Principal-components analyses show that disruption of semantic knowledge and speeded verbal processing in certain patients (e.g., patients with neuro-AIDS, patients with DAT) is a major contributor to the incongruent trial. The Stroop challenge is an effective modality for testing the glucocorticoids-cytokine response in healthy

young men and women. Two groups of responders can be distinguished based on salivary and plasma glucocorticoids response to the challenge. The bottom 40% of the responders for cortisol responses to the challenge characteristically manifested significantly higher salivary and plasma levels of IL-6, but lower subjective experience of stress, compared to the high-glucocorticoids responders (top 40 percentile). This confirms that individual variations exist in the response of the glucocorticoids-cytokine axis to mild stress, and that these variations can have potentially adverse health outcomes.

An ongoing study is comparing patients with major depression (*Diagnostic and Statistical Manual of Mental Disorders, 4th edition* (DSM-IV) criteria) with age-, sex-, and ethnicity-matched control subjects for the response on the Stroop test. Preliminary data indicate that the median Raw Color-Word score is 20% higher in the control, compared to the depressed subjects. The median predicted interference score based on Color-Word scores is indistinguishable among healthy and depressed subjects. These data confirm the literature in the substantial differences in percentile position of the median T scores: 66% percentile (depressed patients) versus 42% percentile (healthy control subjects). An important caveat of the Stroop challenge is that color confusion can be inconsistent in certain subgroups of subjects, especially as subjects advanced in age. The Stroop test should not be used for clinical diagnostic purposes, but rather limited as an experimental tool.

The Cold Pressor Test

Another widely used stress paradigm for testing allostasis in human subjects is the cold pressor challenge. The physiological mechanism underlying the response to this challenge, which consists in the immersion of the nondominant hand in an ice bath, involves a response temperature as constriction to blood flow of both superficial and deep tissues of the hand, and associated strong perception of pain. This challenge test is an adequate model to study psychobiological responses to stressful discomfort because cold-induced vasospasms are frequently present in syndromes of chronic pain, but it may have adverse effects in subjects with cardiovascular dysfunctions. Pain latency (e.g., length of time of hand immersion) is significantly shorter in certain populations of patients, such as patients with fibromyalgia and secondary concomitant fibromyalgia, compared to healthy control subjects.

This challenge alters the glucocorticoids-cytokine axis responsivity in healthy young adult men and women, and leads to time-dependent patterns of change in the circadian patterns of salivary IL-6.

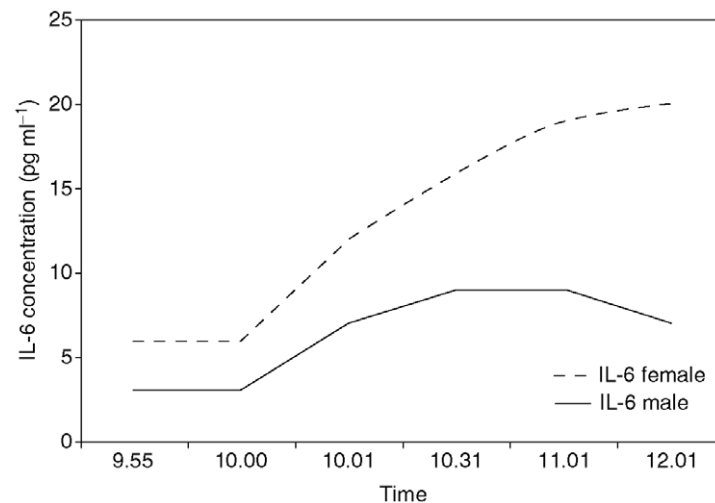


Figure 2 Salivary IL-6 response after cold pressor test applied at 10.00 a.m. on healthy volunteers. IL-6 response to stress was more pronounced in female (dashed line) compared to male subjects (solid line) ($P < 0.05$).

The cold pressor test significantly alters salivary IL-6 in control subjects. IL-6 levels were measured at times $t = -5$ min, $t = 0$ min (baseline levels), and $t = 1, 31, 61$, and 91 min postchallenge. Salivary IL-6 rose two- to threefold during the 90 min after cold pressor test. This effect was more pronounced in female compared to male subjects ($P < 0.05$; Figure 2).

The Dexamethasone-Suppression Test

The dexamethasone-suppression test (DST) is a widely used physiological challenge to study the functional response of the glucocorticoids-cytokine axis. Administration of 1–5 mg of dexamethasone (DEX) to healthy subjects in the evening (23.00 h) results in a flattened cortisol plasma level the following morning and day. The cortisol suppression response is an indication of DEX-sensitivity and normal hypothalamic-pituitary-adrenocortical (HPA) response. Significant changes in the migratory properties of circulating CD4⁺ T cells occur in DEX-sensitive control subjects. The use of the DST challenge is in monitoring the glucocorticoids-cytokine response in a variety of patient populations. In DEX-resistant (i.e., no cortisol suppression following DST) subjects, the migratory properties of circulating CD4⁺ T cells are not altered. An intimate relationship is evident between the physiological regulation of the neuroendocrine and cellular immune systems *in vivo* in healthy subjects, compared to patient populations. Stress states can lead to enhanced suppression of cortisol by DEX following the DST challenge, although discrepancies exist among reports. Neither low basal cortisol nor enhanced suppression of cortisol are consistent markers of a posttraumatic stress disorder, for

instance, major depression, or DAT. The lack of reliability, validity, and specificity of the DST as a diagnostic instrument, however, has not reduced the usefulness of this challenge modality to better understand neuroendocrine-immune interactions and the physiological regulation of immune suppression. Another caveat of this paradigm is that it provides a static challenge to the neuroendocrine-immune system: it leads to a shut-off, albeit temporarily, of glucocorticoids flux by acting on its feedback system. In that respect it presents a reflection of the regulatory response, rather than of the glucocorticoid-cytokine axis potential *per se*, as, for example the CRH stimulation test does. Therefore, the DST challenge is a static challenge of glucocorticoids-cytokine responses.

Alcohol and Substance Abuse

Subjects who abuse alcohol or other substances (e.g., cocaine) characteristically exhibit a deregulated pituitary-adrenal axis, elevated cortisol levels, a blunted T cell activation, and consequential immune suppression. Associated is a deregulated distribution of CD4⁺CD62L⁺ cells, which upon withdrawal of alcohol or cocaine abuse, normalizes together with plasma cortisol. Indeed, nicotine and ethanol, at concentrations approximating those found in the plasma of smokers and alcohol abusers, decrease immune surveillance of oral carcinoma by interfering with immune cell-mediated clearance of the tumor target. Significant ($P < 0.05$) diminution of migration, binding, activation, and cytotoxic activity by nicotine when coupled with ethanol contributes to the observed immune suppression.

Experimental Findings

Circulating memory CD4⁺CD45R0⁺ cells in patients with OLP comprise 66% of the circulating leukocyte population, compared to 49% in healthy control subjects, suggesting an imbalance between naïve and memory T cells. Moreover, serum levels of tumor necrosis factor (TNF)- α and IL-6 are significantly elevated in patients with OLP. These patients also exhibit significant alterations in pituitary-adrenal/cellular immune regulation, as do close to 50% of patients with DAT. In these patients, whole saliva can be used for the reliable assessment of neuroendocrine products and cytokines.

Serum, saliva, and cervicular fluid samples were collected for the determination of IL-6 in healthy age-matched control subjects, in patients with DAT, and in their caregivers. Subjects were assessed by a multidimensional evaluation that included collection of demographic data, cognitive function, level of impairment, and assessment of mood states by means of the profile of moods states. Mean IL-6 levels in both plasma and saliva were higher in the group of patients with DAT, compared to the caregiver and control subjects. Crevicular fluid determinations were unreliable due to the small volume collected. Saliva and plasma IL-6 levels were well correlated. Profiles of moods states scores were significantly higher in the caregiver group (54.80 ± 65.45 ; -16 – 166), compared to the patient group with DAT (39.6 ± 50.5 ; 6 – 129), and the control subjects (9.4 ± 12.1 ; -26 – 2) (Wilcoxon, $P < 0.05$). Caregivers showed lower emotional well-being and more emotional stress than patients and healthy noncaregiver controls.

T Cell Anergy and Apoptosis

T cell activation can lead to the proliferative response. Alternative pathways can attenuate or prevent the response of T cells, and lead to anergy (i.e., absence of response) or apoptosis (i.e., spontaneous programmed cell death), and consequential immune suppression.

T cell stimulation by triggering through the T cell receptor (TcR) in the absence of CD28-mediated costimulatory signals induces unresponsiveness to further stimulation in T cells, a phenomenon known as anergy. Anergic T cells do not produce IL-2 when challenged, and are defective in their ability to upregulate protein binding and transactivation at two critical IL-2 DNA enhancer elements: NF-AT (nuclear factor of activated T cells; a sequence that binds a heterotrimeric NFATp, fos, and jun protein complex), activator protein-1 (AP-1; that binds fos and jun heterodimers), the Janus tyrosine kinases (Jak), and the signal transducers and activators of transcription

(Stat) act as key regulatory components. The impaired DNA-protein interactions in anergic T cells is related to poor expression of the inducible AP-1 family members c-fos, fosB, and junB, attributable, at least in part, to a selective block in the expression of the AP-1 Fos and Jun family members in anergic T cells. Ligation of CD28 results in its phosphorylation and subsequent recruitment and activation of phosphatidylinositol (PI)3-kinase. Pharmacological and mutational analysis data show that disruption of PI3-kinase association with CD28 is neither necessary nor sufficient for costimulation of IL-2 production. T cell anergy also involves transcription regulators.

Calmodulin

Calmodulin (CaM) is an intracellular Ca²⁺-binding protein that mediates many of the effects of Ca²⁺. CaM binds to Fas directly and identifies the CaM-binding site on the cytoplasmic death domain (DD) of Fas. Fas (APO-1/CD95) is a cell-surface receptor that initiates apoptotic pathways, and its cytoplasmic domain interacts with various molecules suggesting that Fas signaling is complex and regulated by multiple proteins. The processes of T cell anergy and apoptosis are intertwined and act together to play critical role in immune suppression.

The Apoptosis Process

The process of apoptosis is a complex network, with different nodes connected by physical interactions of critical domains, including the caspase recruitment domains (CARDs); death effector domains (DEDs); DDs; BIR (baculovirus inhibitor of apoptosis proteins (IAP) repeat) domains of IAPs; Bcl-2 family proteins; caspase protease domains; and endonuclease-associated CIDE (cell death-inducing DNA fragmentation factor (DFF)45like effector) domains. Apoptotic cell death is thus a multifaceted process, whose initial signal is triggered in part by members of the TNF receptor superfamily, including Fas. When Fas interacts with the Fas-ligand, transmembrane signals are initiated through the associated FADD/MORT-1 motif, which interacts with caspase-8. Up to 10 caspases have been identified to date, which have been numbered based upon the chronological order of publication of their respective cDNAs. Active caspases are released from cleavage from the proenzyme into a larger peptide (17–22 kD) and a smaller peptide (10–12 kD) that associate to form the active enzyme in cells undergoing apoptosis. Caspase-8 possesses two motifs near the 5'-terminal that associate with FADD, referred to as the death

effector domain. Activated caspase-8 in turn activates other caspases, and acts directly on mitochondrial permeability and metabolism.

Mitochondria

Mitochondria are involved in apoptosis through release of soluble molecules, such as bcl-2 that is normally situated in the outer mitochondrial membrane near the inner–outer membrane contact sites where permeability pores form. The action of bcl-2 is to protect against apoptosis by preventing the opening of these pores and release of a 50-kD protease and cytochrome C. Additional events involve changes in the plasma membrane and translocation of phosphatidyl-serine from the inner to the outer aspect of the membrane, which can be reliably detected by its high affinity to the Ca^{2+} -dependent phospholipid-binding protein, Annexin-5 (35–36 kD). Signals reach the nucleus for activation of nucleases and degradation of chromosomes into nucleosomal subunits. These metabolites are detected either as fragments of increasing length on an agarose gel or as cytoplasmic content of nucleosomal proteins.

β -Amyloid

Case in point, the peptide found in the neural plaques that characterize Alzheimer's disease, β -amyloid, significantly modulates T cell-mediated immunity.

The 1–42 β -amyloid peptide leads to T cell anergy, and impairs the generation of CD4+CD25+ and CD8+CD25+ Treg by up to 25%. This effect is attributable to the 28–35 portion of the peptide, the same that is responsible for impairing the maturation of T cell *in vitro*. The peptide also induces apoptotic cell death of healthy human T cells, as monitored by augmentation of cytoplasmic oligonucleosomes following treatment with β -amyloid, and a over three-fold induction of caspase 3 (Figure 3).

A New Strategy: Molecular Cartography

Molecular cartography is the science of recognizing and identifying the multifaceted and intricate array of interacting genes and gene products that characterize the function and specialization of each individual cell in the context of cell–cell interaction, tissue and organ function. The relevance of molecular cartography pertains to every domain of physiology, including immunity.

Molecular cartography provides the fundamental knowledge and understanding of the genomic, the proteomic, and interactomic processes that regulate the emergence, stability, and function of immune cell populations, and the mechanisms by which the psychoneuroendocrine system regulates immunity. It produces new fundamental knowledge with respect to the processes that regulate immune suppression.

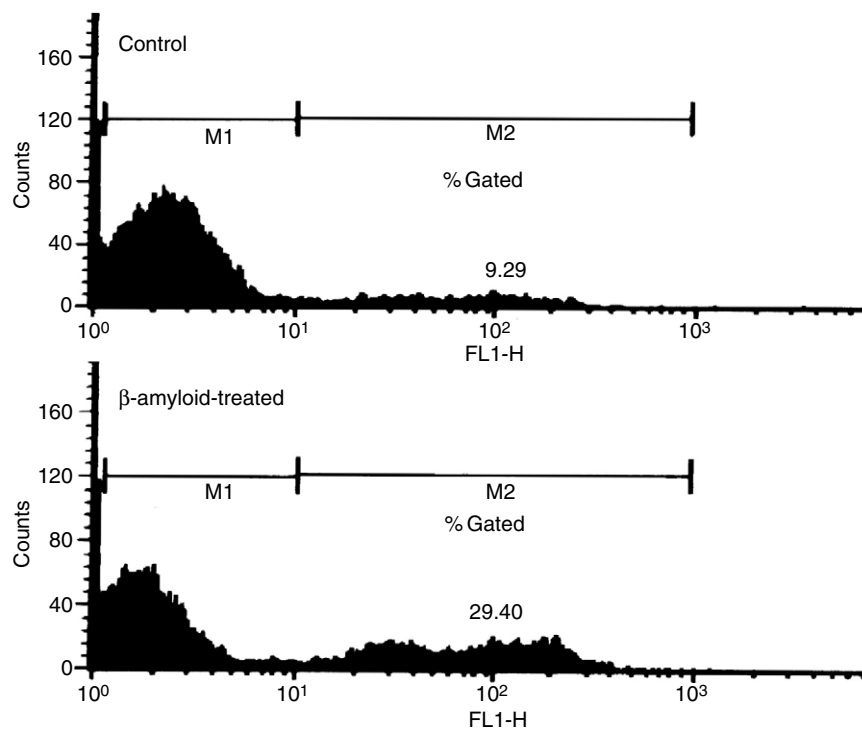


Figure 3 Caspase 3 induction in healthy human T cells exposed to β -amyloid peptide (1–42) for 48 h. Upper panel shows untreated (control) cultures. Lower panel shows β -amyloid-treated cultures. β -amyloid stimulated over threefold induction of caspase 3.

Further Reading

- Carli, G., Suman, A. L., Biasi, G., et al. (2002). Reactivity to superficial and deep stimuli in patients with chronic musculoskeletal pain. *Pain* 100, 259–269.
- Chiappelli, F., Kung, M. A., Nguyen, P., et al. (1997). Cellular immune correlates of clinical severity in oral lichen planus: preliminary association with mood states. *Oral Diseases* 3, 64–70.
- Chiappelli, F. (2005). The molecular immunology of mucositis: implications for evidence-based research in alternative and complementary palliative treatments. *Evidence-Based Complementary and Alternative Medicine* 2, 489–494.
- Chiappelli, F. and Cajulis, O. S. (2004). Psychobiological views on “stress-related oral ulcers” *Quintessence International* 35, 223–227.
- Chiappelli, F. and Hodgson, D. (2000). Immune suppression. In: Fink, G. (ed.) *Encyclopedia of stress*, (vol. 2), pp. 531–536. San Diego: Academic Press.
- Chiappelli, F., Prolo, P., Fiala, M., et al. (2005). Allostasis in HIV-1 infection and AIDS. In: Minagar, P. A. & Shapshak, P. (eds.) *Neuro-AIDS*. Hauppauge, NY: Nova Science Publisher (in press).
- Chiappelli, F., Prolo, P., Iribarren, J., et al. (2006). Alzheimer’s disease: new frontiers for the XXI century. In: Welsh, E. M. (ed.) *Trends in Alzheimer’s disease research*, pp. 233–258. Hauppauge, NY: Nova Science Publishers.
- Czura, C. J., Friedman, S. G. and Tracey, K. J. (2003). Neural inhibition of inflammation: the cholinergic anti-inflammatory pathway. *Journal of Endotoxin Research* 9, 409–413.
- Kunz-Ebrecht, S. R., Mohamed-Ali, V., Feldman, P. J., et al. (2003). Cortisol responses to mild psychological stress are inversely associated with proinflammatory cytokines. *Brain Behavior and Immunology* 17, 373–383.
- Maekawa, K., Twu, C., Lotaif, A., et al. (2003). Function of β -adrenergic receptor on mononuclear cells in female fibromyalgia patient. *Journal of Rheumatology* 30, 364–368.
- McEwen, B. and Wingfield, J. C. (2003). The concept of allostasis in biology and biomedicine. *Hormones and Behavior* 43, 2–15.
- Sterling, P. and Eyer, J. (1988). Allostasis: a new paradigm to explain arousal pathology. In: Fisher, S. & Reason, J. (eds.) *Handbook of life stress, cognition, and health*, pp. 629–645. New York: Wiley.
- van der PouwKraan, T. C., Kasperkovitz, P. V., Verbeet, N., et al. (2004). Genomics in the immune system. *Clinical Immunology* 111, 175–185.

Immune Surveillance – Cancer, Effects of Stress on

D Spiegel and F S Dhabhar

Stanford University School of Medicine,
Stanford, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Stress Reactivity and Disease Progression
Stress Hormones and Cancer Progression
Stress and Tumor Immunosurveillance
Psychosocial-Biobehavioral Intervention and Cancer

Glossary

<i>B cells</i>	Class of white blood cells that mount the humoral immune response, attacking foreign antigens such as those found on bacteria.
<i>Chemokines</i>	Proteins that attract the activity of various white blood cells.
<i>Cytokines</i>	Proteins that regulate immune activity and also affect the functioning of the central nervous system.
<i>DNA ploidy</i>	The relative disorganization of DNA replication in tumor cells.

*Helper and
cytolytic T cells*
*Hypothalamic-
pituitary-
adrenal
(HPA) axis*
*Immuno-
suppression*
*Natural Killer
(NK) cells*

*Neuroimmune
effects*
*Proopiomelano-
cortin (POMC)*

Class of white blood cells that kill virally infected or tumor cells.

A major hormonal stress response system in the body that regulates the secretion of cortisol, which mobilizes glucose into the bloodstream.

The reduction of the activity of components of the immune system.

A class of white blood cells that attacks transformed or dying cells, such as tumor cells.

Interacting effects of the central nervous system on immune function.

A precursor protein that is the source of adrenocorticotrophic hormone, which triggers cortisol release, and of endogenous opiates

Cancer can be understood as, among other things, a series of stressors: the diagnosis, arduous treatments, fears of death, and changes in social and physical environment. For example, as many as 80% of breast cancer patients report significant distress during initial treatment. Estimates of the prevalence of

psychiatric disorders among newly diagnosed cancer patients has ranged from 30–44%, and 20–45% exhibit emotional morbidity 1–2 years later. Although the majority of cancer patients do not have a psychiatric diagnosis, for those who do depression is the most common one, and the vast majority experience the diagnosis of cancer as a major stressor. Up to 10% have severe maladjustment as long as 6 years later. Thus, cancer is a disease that causes considerable psychological stress.

Cancer is largely a disease of aging; it is rare below the age of 50 and increasingly common with advancing age. To the extent that abnormalities of the endocrine stress response system, the hypothalamic-pituitary-adrenal axis (HPA), increase with age and affect cancer progression, their impact should increase with advancing age due to a greater likelihood of abnormality and a higher prevalence of disease. The bulk of cancer research has productively focused on the pathophysiology of the disease, emphasizing tumor biology, especially tumor characteristics such as DNA ploidy and estrogen/progesterone receptor status as predictors of disease outcome at the expense of studying the body's psychophysiological reactions to tumor invasion. These reactions are mediated by brain-body mechanisms, including the endocrine, neuroimmune, and autonomic nervous systems. Although a large portion of the variance in any disease outcome is accounted for by the specific local pathophysiology of that disease, some variability must also be explained by host resistance factors, which include the manner of response to the stress of the illness. For example, in a series of classic experiments in animals, Riley showed that crowding accelerated the rate of tumor growth and mortality. In a recent authoritative review of human stress literature, McEwen documented the adverse health effects of cumulative stressors and the body's failure to adapt the stress response to them. The activation of the HPA axis is an adaptive response to acute stress, but over time, in response to cumulative stress, the system's signal-to-noise ratio can be degraded so that it is partially on all the time, leading to adverse physiological consequences, including abnormalities of glucose metabolism, hippocampal damage, accumulation of abdominal fat, and depression. Abnormalities of HPA axis function, including glucocorticoid receptor hypersensitivity, have also been found to be associated with posttraumatic stress disorder. Thus, adverse emotional events, ranging from traumatic stressors to cumulative minor ones, are associated with HPA axis dysregulation. Persistently elevated or relatively invariant levels of cortisol may, in turn, stimulate tumor proliferation via differential gluconeogenesis in normal and tumor tissue, the activation of hormone receptors in tumor,

or immunosuppression. Glucocorticoid dysregulation may be associated with other stress-related endocrine and immune dysfunction that could have consequences for host resistance to cancer progression.

Stress Reactivity and Disease Progression

Reactivity to stressors involves not only the initial response but the system's ability to reset itself after stress. Older organisms seem to reset themselves after a stressor less readily, leaving stress-responsive systems more tonically upregulated for longer periods of time after the cessation of the stressor. Cancer affects older people particularly, and chronic disease-related stressors are perhaps likely to further decrease the flexibility of stress response mechanisms. Sapolsky and colleagues found that stressed older animals had not only persistently elevated cortisol levels after the stress was over but also much more rapid growth of implanted tumors. This finding was confirmed in a study that also identified the suppressive effects of stress hormones on natural killer (NK) cells as one potential mediator of the effect. It may not be so much the magnitude of stress-induced elevation of corticosteroids that is important in cancer outcomes as the duration or persistence of the elevation. These findings are particularly relevant because there is evidence that persistent elevations of cortisol, as measured by slower rates of cortisol decline throughout the day, have been associated with both shorter cancer survival and the suppression of NK cells. These lymphocytes, identified with CD56 cell surface markers, do not attack specific antigens but rather destroy transformed or dying cells, such as cancer cells. The number and killing activity of these NK cells have been found to be associated with the progression of breast cancer. As the disease progresses, NK activity declines and, conversely, declining NK activity and counts predicts breast cancer progression.

Stress Hormones and Cancer Progression

Mechanisms have been proposed whereby the neuroendocrine correlates of stress may promote neoplastic growth. Stress hormones may suppress immune resistance to tumors or act via differential effects on gluconeogenesis in healthy versus tumor cells. Tumor cells may become resistant to the catabolic action of cortisol, which inhibits the uptake of glucose in numerous cell types. In such cases, energy would be preferentially shunted to the tumor and away from normal cells by cortisol. Several studies have found an association between the stress-related elevation

of glucocorticoids and more rapid tumor growth in animals. Another hypothesis suggests that HPA axis hormones may promote the expression of breast cancer oncogenes due to activation of proopiomelanocortin (POMC) genes by corticotropin releasing hormone (CRH). The POMC promoter has homology with POMC transcription factors that bind to human MAT-1 breast cancer oncogenes. Also, fasting insulin levels have been found to predict subsequent mortality from breast cancer, whereas levels of insulin-like growth factor binding proteins predict recurrence but not mortality. Thus, there are a variety of means by which depression-related abnormalities in HPA axis function could affect the rate of cancer progression. One leading hypothesis is an association between HPA axis and immune dysfunction because glucocorticoids are potentially immunosuppressive.

Stress and Tumor Immunosurveillance

Although some aspects of antitumor immune responses are still controversial, there is now ample evidence to show that under certain conditions the immune system plays a crucial and protective role by detecting and eliminating tumors. Recently, Dunn et al. introduced the concept of tumor immunoediting, which hypothesizes that the immune system plays a critical role not only in detecting and eliminating tumors but also in shaping the extent of immunogenicity or nonimmunogenicity of tumors that ultimately escape immunosurveillance and result in clinically observable malignant disease. Evidence for the immunosurveillance hypothesis comes from (1) studies that show an increased incidence of virally induced cancers, such as Kaposi's sarcoma and non-Hodgkin's lymphoma, in patients undergoing immunosuppressive treatments; (2) studies of patients with immunodeficiency diseases; and (3) studies using selective cytokine- or leukocyte-deficient mice. However, immunosurveillance is also thought to play an important role in the elimination of cancers that do not have a viral etiology (e.g., malignant melanoma, non-Kaposi's sarcoma, and lung tumors), and higher incidence ratios for these cancers are observed in immunocompromised hosts.

All major leukocyte subtypes, including macrophages, granulocytes, NK cells, helper and cytolytic T cells, and B cells, are thought to play a role in immune surveillance. However, under certain conditions, regulatory/suppressor leukocytes such as subsets of CD25+ and CD4+ T cells and natural killer T (NKT) cells, may inhibit antitumor immune responses. Specific cytokines and chemokines are thought to exert direct and indirect protective (suppression of tumor development and progression)

as well as harmful (promotion of tumor growth and metastasis) effects. Moreover, in contrast to protective, regulated, antitumor immune responses, dysregulated or chronic immune responses directed against pathogens or self-antigens can induce or exacerbate tumorigenesis through chronic tissue damage and activation.

Prolonged or chronic stress has been shown to affect most immune parameters thought to be important for tumor surveillance. Therefore, stress may affect immune function and susceptibility to cancer in several different ways. First, in general, chronic stress and stress hormone exposure have been associated with the suppression of NK activity, T-cell responses, and cell-mediated immune reactions; decreased telomere length and telomerase activity in circulating leukocytes; increased tumor metastases; enhanced growth of transplanted tumors; and increased tumor number, size, and density of skin cancer (squamous cell carcinoma). Second, stress may also suppress antiviral immune responses. The suppression of immunity to transforming viruses could significantly enhance the risk of tumor formation and progression. Third, chronic stress may contribute to immune dysregulation and promote or sustain chronic pro-inflammatory states that, in turn, induce or exacerbate cancer. Fourth, glucocorticoids are potentially immunosuppressive, so the effects stress and hypercortisolemia may include functional immunosuppression as well, as has been shown extensively in animals and for which there is a growing body of evidence in humans. This, in turn, could influence the rate of cancer progression.

Psychosocial-Biobehavioral Intervention and Cancer

Understanding how psychosocial factors such as stress can suppress protective immune function may help in the design of psychosocial-biobehavioral interventions that may buffer against stress- and depression-induced effects, thereby improving patient quality of life and potential survival. Five of eleven published randomized trials demonstrate such an effect. The principles of such stress management techniques include helping cancer patients construct a new network of social support; encouraging the expression and management of the range of emotions typical of such stress-inducing illness, including anger, fear, and sadness; helping people face their fears of dying and death; and reordering life priorities. Other helpful components include improving social relationships; enhancing doctor-patient communication; educating the patient about coping skills, the disease, and its treatment; and teaching anxiety

and pain management techniques. Such interventions clearly improve patient quality of life and may have a positive effect on survival time via effects on endocrine, circadian, and immune physiology.

Further Reading

- Andersen, B. L., Farrar, W. B., Golden-Kreutz, D., et al. (1998). Stress and immune responses after surgical treatment for regional breast cancer. *Journal of the National Cancer Institute* 90(1), 30–36.
- Ben-Eliyahu, S. (2003). The promotion of tumor metastasis by surgery and stress: immunological basis and implications for psychoneuroimmunology. *Brain, Behavior & Immunity* 7(supplement 1), S27–S36.
- Ben-Eliyahu, S., Yirmiya, R., Liebeskind, J. C., et al. (1991). Stress increases metastatic spread of a mammary tumor in rats: evidence for mediation by the immune system. *Brain, Behavior & Immunity* 5(2), 193–205.
- Derogatis, L. R., Morrow, G. R., Fetting, J., et al. (1983). The prevalence of psychiatric disorders among cancer patients. *Journal of the American Medical Association* 249(6), 751–757.
- Dhabhar, F. S. and McEwen, B. S. (1997). Acute stress enhances while chronic stress suppresses cell-mediated immunity in vivo: a potential role for leukocyte trafficking. *Brain, Behavior & Immunity* 11, 286–306.
- Dunn, G. P., Bruce, A. T., Ikeda, H., et al. (2002). Cancer immunoediting: from immunosurveillance to tumor escape. *Nature Reviews Immunology* 3(11), 991–998.
- Epel, E. S., Blackburn, E. H., Lin, J., et al. (2004). Accelerated telomere shortening in response to life stress. *Proceedings of the National Academy of Sciences USA* 101(49), 17312–17315.
- Epel, E. S., McEwen, B., Seeman, T., et al. (1999). Psychological stress and lack of cortisol habituation among women with abdominal fat distribution. *Psychosomatic Medicine* 61, 107.
- Glaser, R. and Kiecolt-Glaser, J. K. (2005). Stress-induced immune dysfunction: implications for health. *Nature Reviews Immunology* 5(3), 243–251.
- Glaser, R., Kiecolt-Glaser, J. K., Malarkey, W. B., et al. (1998). The influence of psychological stress on the immune response to vaccines. *Annals of the New York Academy of Sciences* 840, 649–655.
- Goodwin, P. J., Ennis, M., Pritchard, K. I., et al. (2002). Fasting insulin and outcome in early-stage breast cancer: results of a prospective cohort study. *Journal of Clinical Oncology* 20(1), 42–51.
- Irvine, D., Brown, B., Crooks, D., et al. (1991). Psychosocial adjustment in women with breast cancer. *Cancer* 67(4), 1097–1117.
- Kiecolt-Glaser, J. K. and Glaser, R. (1999). Psychoneuroimmunology and cancer: fact or fiction? *European Journal of Cancer* 35(11), 1603–1607.
- Kissane, D. W., Grabsch, B., Love, A., et al. (2004). Psychiatric disorder in women with early stage and advanced breast cancer: a comparative analysis. *Australian and New Zealand Journal of Psychiatry* 38(5), 320–326.
- Klein, G. and Klein, E. (1998). Immunology. Sinking surveillance's flagship. *Nature* 395(6701), 34–41.
- Licinio, J., Gold, P. W. and Wong, M. L. (1995). A molecular mechanism for stress-induced alterations in susceptibility to disease. *Lancet* 346(8967), 104–106.
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* 338(3), 171–179.
- Omne-Ponten, M., Holmberg, L. and Sjoden, P. O. (1994). Psychosocial adjustment among women with breast cancer stages I and II: six-year follow-up of consecutive patients. *Journal of Clinical Oncology* 12(9), 1778–1782.
- Posener, J. A., Schildkraut, J. J., Samson, J. A., et al. (1981). Diurnal variation of plasma cortisol and homovanillic acid in healthy subjects. *Psychoneuroendocrinology* 21(1), 33–38.
- Riley, V. (1981). Psychoneuroendocrine influences on immunocompetence and neoplasia. *Science* 212(4499), 1100–1109.
- Romero, L., Raley-Susman, K., Redish, D., et al. (1992). A possible mechanism by which stress accelerates growth of virally-derived tumors. *Proceedings of the National Academy of Sciences USA* 89, 11084.
- Sapolsky, R. M. and Donnelly, T. M. (1985). Vulnerability to stress-induced tumor growth increases with age in rats: role of glucocorticoids. *Endocrinology* 117(2), 662–666.
- Sapolsky, R., Krey, L. and McEwen, B. S. (1985). Prolonged glucocorticoid exposure reduces hippocampal neuron number: implication for aging. *Journal of Neuroscience* 5, 1222–1227.
- Sephton, S. E., Sapolsky, R. M., Kraemer, H. C., et al. (2000). Early mortality in metastatic breast cancer patients with absent or abnormal diurnal cortisol rhythms. *Journal of the National Cancer Institute* 92(12), 994–1000.
- Spiegel, D. (1999). A 43-year-old woman coping with cancer [clinical conference]. *Journal of the American Medical Association* 282(4), 371–378.
- Spiegel, D. (2002). Effects of psychotherapy on cancer survival. *Nature Reviews Cancer* 2(5), 383–389.

Immune System, Aging

M A Horan

University of Manchester Medical School, Salford, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 536–540, © 2000, Elsevier Inc.

Immune Systems
Pattern of Immune Aging
Causes of Immune Aging
Aging and Infections

Glossary

<i>Autoantibody</i>	Immunoglobulin whose specificity is directed at self.
<i>C-reactive protein</i>	A small protein of the pentraxin family, secreted and catabolized only by the hepatocytes (liver parenchymal cells). It was named for its ability to bind with pneumococcal C-polysaccharide. Its synthesis is under transcriptional control, regulated by pro-inflammatory cytokines (especially interleukin 6), and it is one of the proteins that characterize the acute phase response.
<i>Cytokine</i>	A small protein or glycoprotein that generally acts locally to alter cell behavior or function. Cytokines may be considered locally acting hormones. The major classes of cytokines are interleukins, interferons, and tumor necrosis factors.
<i>Endotoxin</i>	Lipopolysaccharides from the walls of gram-negative bacteria that can activate lymphocytes and mononuclear phagocytes.
<i>Epstein-Barr virus</i>	A herpesvirus that, after a primary infection, remains latent in B lymphocytes and may be reactivated at some later time.
<i>Lymphocyte</i>	A small cell derived from the bone marrow that, after appropriate processing and maturation, can participate in specific immune responses. B lymphocytes (B cells) further differentiate into plasma cells, which secrete immunoglobulin (antibody).
<i>Monoclonal antibody</i>	An immunoglobulin of a single specificity produced by an expanded clone of plasma cells.

Mononuclear phagocyte

A cell of bone marrow origin, larger than a lymphocyte, that can phagocytose foreign material and present it to lymphocytes to bring about an immune response. Examples are monocytes, macrophages, and Kupffer cells.

Immune Systems

Two interactive immune systems are recognized: innate (natural) immunity and acquired (adoptive or specific) immunity. Innate immunity comprises polymorphonuclear leukocytes (neutrophils, eosinophils, and basophils), natural killer (NK) cells, and mononuclear phagocytes, and it uses the complement cascade as its main soluble protein effector mechanism. It also uses numerous recognition molecules, including C-reactive protein (CRP), serum amyloid protein (SAP), and mannose-binding protein (MBP). These molecules have been selected during evolution to bind carbohydrate structures that do not occur on eukaryotic cells and so differentiate potentially harmful invaders from innocuous self.

Acquired immunity employs several subtypes of lymphocytes and uses antibody as its effector protein. Antibody and the T-cell receptor (TCR) are the recognition structures. B lymphocytes can recognize protein, carbohydrate, and simple chemical structures, whereas T lymphocytes seem to recognize only peptides. Clones of lymphocytes with receptors of sufficient affinity are triggered by antigen to proliferate and differentiate into the various effector cells: T-helper (Th) cells, cytotoxic T cells, suppressor T cells, and plasma cells (which secrete antibody).

After an immune response subsides, specific antibody often persists in the blood for many years and even decades. This implies that humoral effector cells (plasma cells) persist and continue to secrete antibody. In contrast, the antibody response at mucosal surfaces is generally short-lived (a few months to a year). Effector T cells do not persist for very long, but antigen-specific clones remain expanded as memory lymphocytes, which can differentiate quickly into effector cells if the antigen is encountered again. In general, memory responses are most effective in protecting against systemic infections (e.g., measles, mumps, poliomyelitis, and smallpox). Localized infections at mucosal surfaces (e.g., rotavirus, respiratory

syncytial virus, and rhinovirus) can recur before memory lymphocytes can differentiate into effector cells, although the subsequent episodes of disease are usually less severe.

The flexible nature of specific immunity poses the problem of differentiating innocuous (self) antigens from harmful ones, a problem that innate immunity does not have. One way of overcoming this is to delete potentially self-reactive clones during maturation in the thymus (for T cells) and at a yet unknown site for B cells. Further specificity is assured by interactions with innate immunity. For T cells, antigen presentation is the critical step. The uptake of antigen into antigen-presenting cells (e.g., mononuclear phagocytes and dendritic cells) is determined by the presence of carbohydrate moieties that are recognized by innate immunity. For B cells, the interaction with innate immunity seems to be their membrane receptors for the C3d component of complement, which is found in association with CD19. CD19 is needed for antibody production for T-cell-dependent antigens and amplifies signaling after the antigen binds membrane immunoglobulin. Covalent binding of C3d to microbial carbohydrate structures facilitates this process. Different pathogens require different types of response for their elimination: type 1 responses (characterized by dependence on macrophages for phagocytosis and intracellular killing) and type 2 responses (macrophage independent, relying on non-cytotoxic antibodies, mast cells, and eosinophils). Th cells are required for both types of response, but these exist as two subtypes, called Th1 and Th2. Th1 cells produce interleukin (IL)-2, IL-3, tumor necrosis factor (TNF)- α , and interferon (IFN)- γ , and this pattern of cytokines promotes type 1 responses. Macrophages are activated and B cells are switched to IgG1 production (in humans), which binds macrophage Fc receptors and activates complement, both of which promote phagocytosis. Activated macrophages produce IL-12, TNF- α , and IFN- γ . This pattern of cytokines (especially IL-12) induces the activation and proliferation of NK cells, the differentiation of naïve T cells into Th cells, and the expression of the Th1 phenotype.

Th2 cells promote type 2 responses. Th2 cells require IL-4 to prime naïve cells. Once activated, they produce IL-3, IL-4, IL-5, IL-6, IL-10, and IL-13. This pattern of cytokines leads to the growth and activation of mast cells and eosinophils. It also switches B lymphocytes to produce IgG4 (and IgE) and inhibits macrophage activation. This pattern of response is typical of helminth infections, and it may also be important for the downregulation of type 1 responses.

Pattern of Immune Aging

Of all body systems, the immune system is probably the best understood, and studies of it can be traced to the early part of this century (and even earlier). Perhaps the first publication was by Peter Ludwig Panum, a Danish physician best remembered for the discovery of endotoxins. He reported an outbreak of measles in the Faroe Islands in 1846. The islands had been free of measles since the previous epidemic in 1781. The reported outbreak affected 75–95% of the population, although “of the many aged people still living on the Faroes who had had measles in 1781, not one was attacked a second time.”

The classic view of immune aging (immunosenescence) is of an immunodeficiency state that predisposes the host to infectious diseases and possibly also neoplasms. The state is often attributed to the involution of the thymus gland and is thought to be characterized by the decreased proliferation of T lymphocytes and impaired Th activity, which lead to impaired cell-mediated and humoral responses to T-cell-dependent antigens. B-lymphocyte numbers and function (antibody production) seem to change little, although in the intact organism antibody titers to a variety of vaccine antigens are generally reduced. Paradoxically, there is an increased incidence of autoantibodies (inevitably of low titer) and benign monoclonal B-lymphocyte proliferations with monoclonal antibody production. Interestingly, older people in Japan have a very low incidence of these benign monoclonal proliferations, suggesting the importance of genetic factors in outbred populations such as humans.

With the discovery of cytokines and their regulatory importance for many cell types (including the cells of the immune system), new insights have emerged concerning the mechanisms underlying immunosenescence. The patterns of cytokine secretion by T lymphocytes from young and old mice clearly differ. In the young, IL-2 and IL-4 predominate, whereas in the old it is IL-10 and INF- γ . The young pattern also characterizes Th1 responses and the old pattern the Th2 responses. It is also known that naïve T cells (which have never encountered antigen) exhibit Th1 responses and that memory T cells (which persist after antigen exposure) exhibit Th2 responses. Memory T cells are known to accumulate during aging. Furthermore, the change in the balance of Th1-to-Th2 responses can be partially reversed by thymic hormones and the steroid dehydroepiandrosterone. However, the altered Th1–Th2 balance is not explained simply by the accumulation of memory T cells. Th1 function clearly declines during aging. For example, old mice are unable to develop

tolerance to Staphylococcal enterotoxin B and die from shock, whereas young mice develop a tolerance and do not die. Thus, thinking has changed from the classic view of immunosenescence toward a view based on immune dysregulation, predominantly the disturbed balance of Th1–Th2 responses, a change that also appears to occur in humans. Increased Th1 and reduced Th2 responses would be expected to predispose to infections, but what causes this change?

Causes of Immune Aging

Clearly, the accumulation of memory T cells could explain some of the change, but it cannot account for it entirely. In outbred populations such as humans, there are genetic polymorphisms in the promoter regions of many (if not all) cytokine genes and the selective survival of particular genotypes could play a part but cannot explain the change in an inbred strain. One intriguing suggestion is that inflammation could be a significant cause.

Inflammation

Of all inflammatory cells, macrophages have been given the closest scrutiny. When stimulated, they secrete a variety of inflammatory mediators, including the pro-inflammatory cytokines IL-1, IL-6, and TNF- α . *In vivo* administration of endotoxin to mice induces the production of TNF- α , IL-1, IL-6, and INF- γ , and the response is much exaggerated in old animals. IL-6 is produced constitutively, and circulating levels are higher in the older (mice and humans) subjects than in the young. Treating old mice with antibodies to IL-6 reduces the magnitude of the Th2 pattern. IL-6 is known to orchestrate the so-called acute phase response, including the production of CRP in the liver. CRP is familiar to most readers as a clinical marker for inflammation and may itself exacerbate tissue damage, but why is the constitutive production of IL-6 higher in the old? What signals might be present in them that may not be present in the young?

Aging as a Pro-inflammatory State

The evidence presented in this article argues strongly for an exaggeration of the macrophage component of inflammation. Is there evidence that other inflammatory cells are similarly affected? The cell type that arrives earliest at a site of infection or tissue damage is the neutrophil. The effects of aging on neutrophil function are controversial – some *in vitro* studies show impaired function, whereas others show no impairment. In contrast, *in vivo* studies are unambiguous in showing that the neutrophil influx is greater

in the old. One interesting study in mice with experimental pneumonia revealed a greater neutrophil influx in the old, associated with a higher death rate. This finding is entirely consistent with the author's studies of endotoxin administration to aged rats. Older humans also seem to experience such exaggerated inflammatory responses. For example, the adult respiratory distress syndrome and the systemic inflammatory response syndrome are both largely conditions of the old, and both are characterized by unrestrained inflammation. More controlled studies of experimental cutaneous wound healing in humans also show more rapid and more marked neutrophil influx into the site of injury in older subjects. This leads to high collagen turnover and probably slower wound healing but a restoration of dermal architecture (unlike the disorganized collagen deposition in the young) and a cosmetically more appealing end result. Interestingly, this effect can be reversed by prior treatment with estrogen, at least in older women.

Replicative Senescence

Studies on aging murine and human T-cell clones have shown that a population remains with normal responses in the midst of a much larger population with reduced responses, including reduced proliferation. It has been suggested that this might be an *in vivo* manifestation of replicative senescence (Hayflick limit), thought to occur in other types of aging cells *in vitro*. *In vitro*, such senescent cells fail to express CD28, have shortened telomeres, and are clearly memory cells (and express CD45RO). *In vivo*, there is known to be a decline in CD28+, with telomere shortening as well as a minimal proliferative capacity. The majority of these CD28+ are CD8+. Almost identical changes have been documented in acquired immunodeficiency syndrome (AIDS), in which they may be due to prolonged or repeated episodes of human immunodeficiency virus (HIV)-induced T-cell proliferation.

Reactivation of latent viruses It has been proposed that the accumulation of T cells showing the characteristics of replicative senescence of aging occurs by a similar mechanism to that in AIDS, that is, periodic or chronic restimulation of memory T cells. The idea that immune aging might result from the presence of immune activators is not new. Because such episodes of antigenic stimulation might also induce nonspecific T-cell replication (bystander replication), their effect on the integrity of the aging immune system may have been underestimated. Candidates for such chronic restimulation may include latent viruses

such as the Epstein–Barr virus (EBV) or pathogens with epitopes that are cross-reactive with self-antigens present on T cells (e.g., influenza A). The consequence of bystander lymphocyte senescence may be that fewer functional CD8⁺ T cells may be available to combat infections and cancer. EBV is acquired asymptotically in childhood, as a symptomatic disease in adolescence (glandular fever and infectious mononucleosis), or as a new infection in old age. Of the seven EBV seroepidemiological studies that include older people, all have shown the highest titers of antibodies in the old. Although this suggests that reactivation does indeed occur in old age, the symptoms of glandular fever are absent. EBV can also be detected in saliva. Twenty percent of people over the age of 50 excrete the virus, but the prevalence of viral shedding in a truly aged population is unknown. Immunosuppression is known to increase the prevalence of virus shedding to over 90%. Symptomatic reactivation has been associated with Hodgkin's disease, nasopharyngeal carcinoma, oral hairy leukoplakia, lymphocytic interstitial pneumonia, and B-cell lymphoproliferative disease. EBV reactivation syndromes have never been investigated in older people, and morbidity could arise from the inappropriate production of cytokines (particularly in people with certain promoter region polymorphisms) and the premature onset of replicative senescence.

Aging and Infections

When an organism arrives at a potential portal of entry into the host (usually an epithelial surface), three outcomes are possible: the organism is eliminated, the organism proliferates successfully but fails to invade tissues, or tissue invasion occurs. For most organisms, tissue invasion is a *sine qua non* for infection, although some induce changes in the host predominantly through the elaboration of toxins (e.g., cholera). When tissue invasion occurs, an acute inflammatory response is triggered, and it is usually by the recognition of the systemic consequences of this that we diagnose infections. The host feels unwell, anorexic, and somnolent (thought to be induced mainly by IL-6). The hypothalamic thermostat is reset to a higher level and the host feels cold and attempts to raise the core temperature to this new level (fever). This is achieved by reducing heat loss by vasoconstriction and behavioral means (e.g., extra clothes) and by shivering. The metabolic rate is increased, although this is a much more important mechanism for small mammals than for large ones. Uncontrolled studies suggest that the old often fail to produce a fever response, but more careful studies suggest that this may not be the case. First, care must be taken

with the measurement of body temperature. Using inadequate instruments at sites of heat loss (e.g., mouth or axilla) almost guarantees the failure to detect fever. The rectum and the tympanic membranes are more reliable sites, and electronic and infrared thermometers equilibrate much faster than mercury-in-glass ones. Second, the ambient temperature may be too low for the host to overcome, and the core temperature may rise into the febrile range only in a warm environment. Finally, experimental studies with several pyrogenic cytokines suggest that the hypothalamic thermostat of the old may be desensitized (dose–response curve shifts to the right) but that the maximal response is unaffected by aging. Another important sign of infection is the detection of an elevation in the circulating white blood cell count (mainly neutrophils), and there is no evidence that aging impairs this response. Likewise, older people also increase the levels of CRP in a pattern indistinguishable from similarly infected young people.

Comorbid factors such as malnutrition, glucocorticoid administration, and so forth can certainly impair host responses and do so in both young and old. Further, diseases at other sites are common in the old and may obscure the characteristic findings of infections. Age-related changes in several body systems predispose to more general types of presentation, known as the giants of geriatrics: immobility, falls, delirium, and incontinence. Thus, it is important to consider infection as a possible diagnosis in any older person whose condition changes acutely, whether the characteristic features are present or not.

Are Older People Predisposed to Acquire Infections?

This question is very difficult to answer. Theoretically, age-related immune changes could well predispose to acquiring infections, but they are not the only factors involved. To acquire an infection requires exposure to a potential pathogen, and this exposure is greatly dependent on environmental and lifestyle factors. For example, people who do not practice unprotected sexual intercourse and who have never had a blood transfusion are highly unlikely to be exposed to HIV. To acquire a common cold or influenza requires contact with an infected person; those living a solitary lifestyle, not traveling on crowded public transport, and not visiting places where people congregate (school, work, football matches, and so forth) may have considerably reduced risks of exposure to potential pathogens. However, those living in relatively closed environments (e.g., nursing homes) may be at an extremely high risk if a potential pathogen is introduced (e.g., tuberculosis, methicillin-resistant *Staphylococcus aureus*, *Clostridium difficile*,

or rotaviruses). Very few hard data exist that make any direct comparisons across age groups. Perhaps the most reliable data come from the American National Health Interview Surveys. These have shown the community-dwelling aged to be at a considerably lower risk of developing colds or influenza than any other age group and at only a modestly increased risk of developing pneumonia. What is not in doubt is that, once infected, the old will be sicker and more likely to die.

See Also the Following Articles

Aging and Stress, Biology of; Cytokines; Infection; Lymphocytes; Macrophages.

Further Reading

- Ashcroft, G. S., Horan, M. A. and Ferguson, M. W. J. (1998). Aging and cutaneous wound healing. In: Horan, M. A. & Little, R. A. (eds.) *Injury in the aging*, pp. 147–153. Cambridge, UK: Cambridge University Press.
- Effros, R. B. (1998). Human genetics '98: aging and senescence; replicative senescence in the immune system: impact of the Hayflick limit on T-cell function in the elderly. *American Journal of Human Genetics* 26, 165–181.
- Horan, M. A. and Ashcroft, G. S. (1997). Ageing, defence mechanisms and the immune system. *Age and Ageing* 26 (supplement 4), 15–119.
- Horan, M. A. and Parker, S. G. (1998). Infections, aging and the host response. In: Horan, M. A. & Little, R. A. (eds.) *Injury in the aging*, pp. 126–146. Cambridge, UK: Cambridge University Press.

Immune System See: Immune Response.

Immunity

F Chiappelli

UCLA School of Dentistry, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by F Chiappelli and Q N Liu, volume 2, pp 541–546, © 2000, Elsevier Inc.

T-Cell Proliferation

Migration

Regulation

Conclusion

*Suppressors
of cytokine
signaling
(SOCS)*

level by the expression of FoxP3, the P3 member of the first forkhead-box (FOX) transcription factor, which functions as a transcriptional repressor at the nuclear factor for activation of T cell (NFAT)/activating protein (AP)-1 sites in cytokine gene promoters, and functionally Treg act to impair proliferative and IL-2 production responses, thus blunting T cell-mediated immunity.

A family of seven known cytoplasmic proteins, the suppressor of cytokine signaling, which modulate cytokine-mediated regulation of immunity.

Glossary

Immunological synapse Complex assemblage of rafts that integrates the signaling cascades for the amplification and regulation of T cell activation.

Rafts Microdomains determined by and anchored to the cytoskeleton of T cells, and important in coordinating the signaling response in T cell-mediated immunity.

Regulatory T cells (Treg) Characterized phenotypically by the expression of the α chain of the interleukin (IL)-2 receptor, CD25, at the molecular

level by the expression of FoxP3, the P3 member of the first forkhead-box (FOX) transcription factor, which functions as a transcriptional repressor at the nuclear factor for activation of T cell (NFAT)/activating protein (AP)-1 sites in cytokine gene promoters, and functionally Treg act to impair proliferative and IL-2 production responses, thus blunting T cell-mediated immunity.

Immunity confers protection of the organism against pathogens. Innate immunity concerns humoral (e.g., complement) and cellular events (e.g., natural killer cells) that do not require priming by antigens. Antigen-specific immunity addresses that branch of cellular (e.g., cytotoxic T lymphocytes and plasma B cells) and humoral immune surveillance (e.g., cytokine patterns, immunoglobulins) that depends upon the phenotypic (e.g., memory CD4+ T cells, CD4+CD45RO+, and B cells) and functional stage maturation of

lymphocytes following presentation of the antigen by the major histocompatibility complex (MHC).

Immunity comprises important innate and adaptive immune responses. Viruses, bacteria, or classical parasites all probe the limit of immunity, and of the three basic parameters of immunology – specificity, tolerance, and memory. The practical specificity repertoire of T and B cells is probably in the order of 10^4 – 10^7 specificities expressed by T cells or by neutralizing antibodies. Tolerance is best defined by rules of reactivity to eliminate infections while avoiding destruction of normal cells by complete elimination of T cells that are specific for antigens persisting in the blood and lymphohemopoietic system. Induction of a T-cell response is the result of antigens newly entering the lymph nodes or the spleen, initially in a local fashion and exhibiting an optimal distribution kinetics within the lymphohemopoietic system. Antigens that stay outside lymphatic tissues are generally immunologically ignored. Immunity is regulated by antigen dose, time, and relative distribution kinetics. Memory relates to the host's resistance following a second exposure to the same agent, and it correlates with antigen-dependent maintenance of circulating antibody titers, and with T-cell responses.

In a previous article, we discussed innate and antigen-dependent immunity in general. Recent research developments have established the central role of T cells in regulating immune surveillance. This article focuses on novel knowledge pertaining to T cell-mediated immunity. The vast domain of autoimmunity is not addressed in this article.

T-Cell Proliferation

T lymphocytes, 70–80% of the peripheral blood mononuclear cells, express the cluster of differentiation (CD)-3, but lack CD14 (i.e., monocyte-specific) and CD20/21 markers (i.e., B cell-specific). CD3 is proximal to the T-cell receptor (TcR), constituted by the α and β chains in circulating T cells, and by the γ and δ chains in the mucosal lymphoid tissue. T cells are stimulated *in vivo* via TcR, and *in vitro* by immobilized anti-CD3 antibody and the CD28 co-stimulus, lest the cells fail to proliferate and enter anergy, a state of stunted cell activation and proliferative capability.

Rafts

Lipids and proteolipids in the T-cell membrane are organized into rafts, microdomains determined by and anchored to the cytoskeleton. Rafts compose the immune synapse, a complex assemblage that integrates the signaling cascades. The synapse amplifies and regulates the activation of T cells. The assembly

of the immune synapse at the antigen-presenting site depends on the interactions between rafts and the actin cytoskeleton, which controls the coalescence of smaller raft components into the larger immune synapse complex to integrate cellular responses to diverse signals. Phosphoinositide 1,5 bisphosphate (PIP2) regulates raft cytoskeletal structure by regulating proteins that control actin dynamics. Upon stimulation, PIP2 is converted to PIP3 and diacylglycerol. Ceramide mediates this clustering and insertion into the lipid bilayer by acetylation motifs.

Raft constituents change, define and determine the process of T cell activation, and maturation, as for instance the increase in raft-associated glycosphingolipid GM1. Compared to resting (CD69–CD25–) or naïve T cells (CD45RA+), activated (CD69+CD25+) or memory T cells (CD45RO+) have significantly higher GM1 levels. CD3/CD28 stimulation promotes recruitment into rafts of the Src family kinase Lck, which co-localizes in rafts with glycosylphosphatidyl inositol (GPI)-anchored proteins, the adaptor protein LAT and Ras, but not with nonraft membrane proteins (e.g., protein tyrosine phosphatase, CD45). β -endorphin, which is released during stress and which significantly impairs PIP2 metabolism and associated kinase activity during T-cell activation, may alter rafts structure.

Membrane Markers of Cell Activation

T-cell activation results in change of morphology to blast cells with large cytoplasm-to-nucleus ratio, loss of CD62L, the peripheral lymph node homing receptor, expression of the integrin lymphocyte function-associated antigen (LFA)-1, and replacement of CD45RA with the maturation marker CD45RO. Activated T cells express CD69, the activation inducer molecule, a very early activation antigen (MLR-3, Leu-23), several hours before CD25 (Figure 1). CD69, a member of the natural killer cell family of signal-transducing receptors, is a type II transmembrane glycoprotein with a C-type lectin-binding domain in its extracellular domain. CD69 controls and regulates the expression of the interleukin (IL)-2 gene by its action downstream to the nuclear factor for activation of T cell (NFAT) and activating protein (AP)-1 transcription complexes.

Following expression of CD69, the IL-2 and CD25 genes are activated. CD25, the chain of the IL-2 receptor is expressed in high density on the cell membrane. Rafts regulate immune signaling in part by segregating cytokine receptors. The β chain of the IL-2 and the IL-15 receptors, for example, is enriched in rafts. This enrichment functionally ensures cytokine selectivity and specificity. Rafts thus assist in the

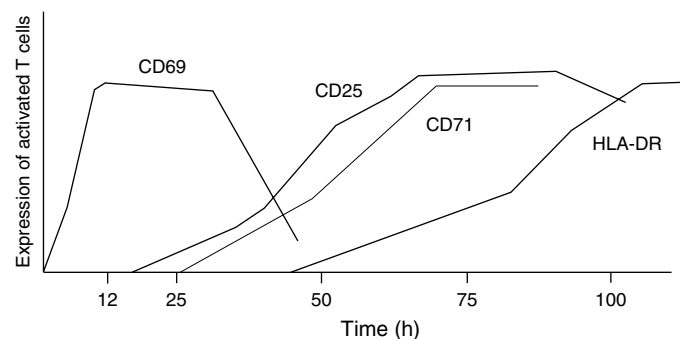


Figure 1 Comparative kinetics of expression of the membrane markers of activation, CD69, CD25, CD71, and human leukocyte antigen (HLA)-DR in activated healthy human adult T cells.

recruitment of CD25 to the vicinity of the signaling β chain, and promote formation of the (α, β, γ) high affinity IL-2 receptor complex. Activated T cells express the transferrin receptor, CD71, which is critical for the transport of iron from the extracellular environment into the cell during traversal of the DNA synthesis (S) phase. The expression of these and other membrane activation markers, including the exopeptidase CD26 (dipeptidyl peptidase IV) and CD137, the T cell co-stimulatory molecule (4-1BB) antigen, is transient, and exhibits finely controlled kinetics.

Traversal of the Cell Cycle

Successful traversal of the cell cycle depends on finely regulated biochemical pathways. At each critical juncture that signifies passage from one stage to the next, gatekeeper proteins allow or disallow the activated cell to progress. These molecular regulators are proteins, the cyclins, which associate with their respective kinase enzymes. Cyclin/cyclin-dependent kinase cdk phosphorylate key proteins that directly or indirectly modulate gene transcription and allow the activated T cell to move forward in the cycle. Cyclin D consists of a family whose members contribute to the regulation of G1/S transition. In T cells, cyclin D3 associates with either cdk4 or cdk6 late in Gap1 (G1). Cyclin D3-cdk4/6 kinase phosphorylates the retinoblastoma protein (Rb), which then releases the transcription factor E2F required for entry into the S phase and for DNA replication. For instance, at 10^{-9} M, *Panax ginseng* (Ginsenoside [20R]-Rg[3]Rc) blunts protein tyrosine kinase activity by 50%, impairs the phosphorylation of the retinoblastoma protein (Rb), and the expression of cyclin D3 (Figure 2), and leads to a threefold decrease in IL-2 production in response to CD3/CD28 stimulation. Expression of cyclin D3, and of the proliferating cell nuclear antigen (PCNA), and the phosphorylation of Rb are blunted in T cells by 0.1% ethanol, but enhanced by 12 ng ml $^{-1}$ prolactin.

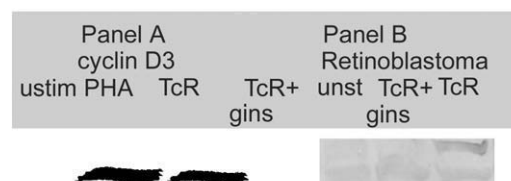


Figure 2 Modulation of cell cycle traversal by ginseng. Normal human T cells were stimulated *in vitro* with immobilized antiCD3/antiCD28 antibody, or the T-cell mitogen phytohemagglutinin, in the presence or the absence of 10^{-9} M ginsenoside Rc (Sigma) for 24 h. Cell extracts were obtained and used for assessment of cyclin D3 (panel A) or retinoblastoma (Rb) protein phosphorylation (panel B) by Western blotting with commercial monoclonal antibodies (Sta Cruz Biochemicals) to measure traversal of G $_1$. Note: Rb phosphorylation is measured as the ratio between the ratio of the upper, ppRb, and the lower pRb bands. TcR, T-cell receptor; PHA, phytohemagglutinin.

At 10^{-8} – 10^{-7} M, the plasma levels of individuals taking ginseng herbal tea, ginseng significantly ($p < 0.05$) stimulates the expression of CD69, and induces a 60% increase in calcium/calmodulin-dependent kinase (CaM kinase) activity. Ion channels are indeed critical in the control of the membrane potential and calcium signaling, and modulate signal transduction pathways following stimulation.

The process and dynamics of cell cycle traversal is monitored by propidium iodide/bromodeoxy uridine (PI/BrdU) immunofluorescent flow cytometry analysis to obtain a timeline sequence monitoring of cells as they traverse G1, S, G2, and mitosis (M). At 10^{-6} M, ginseng has no effect on CD69 expression (4.04 ± 1.34 -fold increase versus 4.30 ± 1.70 -fold increase), on traversal of the cell cycle by PI/BrdU analysis (Figure 3), or on IL-2 production. By contrast, 10^{-6} M mercuric chloride (HgCl $_2$) blunts cell cycle traversal by accumulating activating T cell subpopulations in late S (% cells in S at 48 h, 76.0 ± 13.2 with control versus 91.0 ± 5.0 with HgCl $_2$, $p < 0.07$),

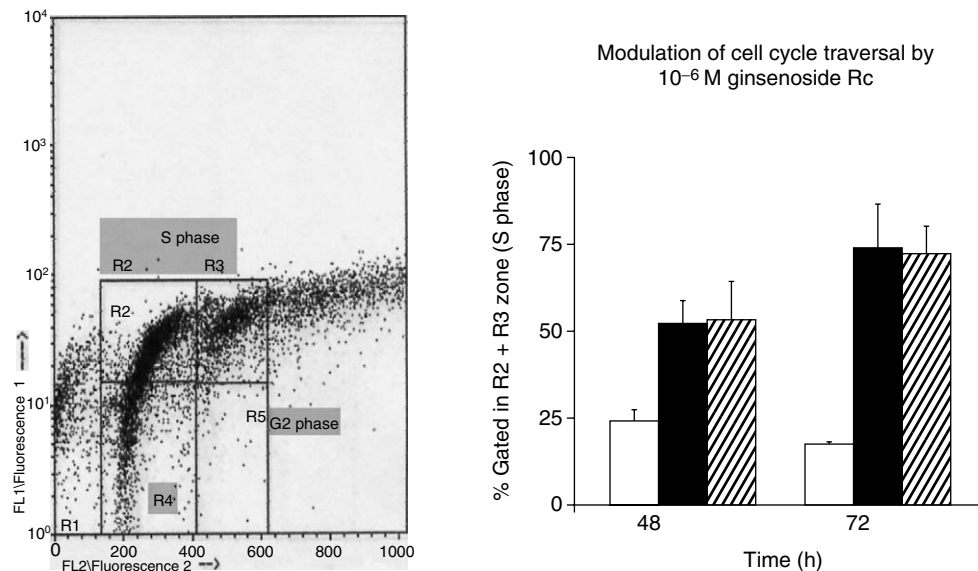


Figure 3 Modulation of cell cycle traversal by ginseng. Normal human T cells were stimulated in the presence or absence of 10^{-6} M ginseng, and the cultures were stained sequentially with propidium iodide (nuclear stain) and bromodeoxyuridine (DNA stain), followed by dual fluorescence flow cytometric analysis. Panel A shows the raw representation of the data from a representative culture, with highlighted the five regions (R1, R2, R3, R4, and R5) that constitute the population of cells traversing the phases of the cell cycle (G_1 , S, G_2 , and M). Panel B shows summary data of the effect of ginseng on cell cycle traversal as monitored by dual fluorescence propidium iodide/bromodeoxy uridine (PI/BrdU). Open bars, control unstimulated; closed bars, control stimulated; hatched bars, stimulated in the presence of 10^{-6} M ginsenoside.

which corresponds to a 50% drop in cell number traversing G_2 and M. However, 0.2% ethanol produces an accumulation of cells in late G_1 (2.5-fold increase, compared to control).

T-Cell Memory – Cytokine Patterns

Monoclonal antibodies against CD45RA can serve to purify naïve T cells from memory T cells ($93.94 \pm 5.51\%$ pure). Dual fluorescence flow cytometry is then used to monitor the maturation of T cells, i.e., the change in proportion from the naïve (CD45RA⁺) to the memory phenotype (CD45RO⁺) following CD3/CD28 stimulation. Kolmogorov-Smirnov analysis quantifies phenotypic T cell maturation. This approach has demonstrated that β -amyloid significantly ($p < 0.05$) blunts the ability of T cells to mature phenotypically.

Cytokines regulate immunity. One aspect of functional maturity of T cells, the variety of cytokines, depends on the initial antigen, and upon the physiological milieu (e.g., modulatory hormones and neuropeptides consequential to the stress response status of the organism). T helper (TH)1 and TH2 patterns of cytokine production are intertwined in terms of both their regulation and their relevance for the overall modulation of immune surveillance in health and disease.

Normal aging and aging-related dementia (e.g., dementia of the Alzheimer's type, DAT) are associated

with significant changes in TH1 and TH2 cytokine production by circulating T cells. Intracellular cytokine staining, following stimulation of T cells in the presence of an inhibitor of protein transport thus retaining the cytokines inside the cell, and double-label immunofluorescence flow cytometry, reveals that normal subjects of 55–70 years of age have trace levels of cytoplasmic IL-10, interferon (IFN) γ , or IL-12 in T cells. Cohorts with DAT show significantly ($p < 0.05$) elevated levels of cytoplasmic IL-10 (2.67 ± 1.15 ng ml⁻¹), IFN γ (3.25 ± 2.25 ng ml⁻¹), and IL-12 (1.78 ± 19.00 ng ml⁻¹) in T cells. In older (>70 years of age) patients with DAT, these levels drop to 1.99 ± 0.61 ng ml⁻¹, $p < 0.06$; 1.25 ± 0.05 ng ml⁻¹, $p < 0.09$; 1.04 ± 0.29 ng ml⁻¹, $p < 0.01$ for cytoplasmic IL10, IFN γ , and IL-12 in T cells, respectively. In addition, the process of immunological aging is evidently distinct in patients with DAT compared to normal subjects, and cytoplasmic levels of IL-10, IFN γ , and IL-12 in T cells rise to levels similar to their age-cohorts diagnosed with DAT in older normal subjects to 1.99 ± 0.70 ng ml⁻¹, 2.75 ± 1.26 ng ml⁻¹, and 1.33 ± 2.07 ng ml⁻¹, respectively.

These changes can be attributed to a plethora of immunophysiological mechanisms within the context of stress and neuroendocrine immunity. The hypothesis has been put forward that age-associated alterations in immunity are associated with the reduction in

hypothalamic dopaminergic tone with aging, related alterations in pituitary hormone, such as the anterior pituitary hormone, prolactin, and interference with the regulation of immunity. For example, 10–15 ng ml⁻¹ prolactin increased IL-2 production by stimulated T cells *in vitro* close to twofold (8.15 ± 2.15 ng ml⁻¹ versus 4.17 ± 1.22 ng ml⁻¹, respectively, $p < 0.09$). The prolactin effect is specific to Th1 cytokines, and is not obtained with the production of the Th2 cytokine, IL-4.

A third pattern of cytokines refers to the effect of certain growth factors (e.g., granulocyte macrophage colony stimulating factor; GM-CSF) in modulating the increase the production of bone marrow-derived antigen-presenting dendritic cells, and the production of IL-23 by these cells. IL-23 belongs to the IL-12 cytokine family, which shares p40, and binds to the IL12 β 1/IL-23R receptor complex to play a critical role in end-stage inflammation via signaling of the Janus tyrosine kinases (JAK) and signal transducers and activators of transcription (STAT) (JAK/STAT) pathway. Administration of IL-11 downregulates IL-12. IL-23 augments production of IL-17 by T cells, which becomes the significant cytokine in the involvement of T cells in inflammation. IL-17 leads to decreased neutrophilia in animal models, and *in vitro* administration of IL-17 synergizes the effects of tumor necrosis factor (TNF)- α for GM-CSF induction, for increased intercellular adhesion molecule (ICAM)-1 (CD54) expression by CD34+ progenitors, for increased neutrophil maturation, and for increased proinflammatory IL-6 and prostaglandin (PG)E2 and granulocyte-colony stimulating factor (G-CSF) growth factor production by keratinocytes and endothelial cells. IL-17 also favors increased secretion of the migration-inducing cytokine, IL-8. In human U937 cells, and probably in all lymphoid and myeloid cells, the JAK/STAT signaling pathway is responsible for transducing signals from the IL-17 receptor. Human vascular embryonic cells (HUVEC) express a homolog of IL-17R, as do ductal epithelial cells of human salivary glands, suggesting a significant role for IL-17 in immunity in the oral mucosa. In fact, in inflamed gingival tissue, IL-17 is the predominant cytokine in tissue proximal to 4–5 mm diseased periodontal pockets. IL-17 plays a critical role in immunity and mucosal inflammation, by modulating the effects of GM-CSF and other growth factors.

Migration

Cell migration is critical to immune surveillance. The migration of T cells is brought about by the concerted

action of specialized factors and receptors, and is regulated by chemokines. Chemokines belong to three groups: C-X-C, C-C, and C depending on the number and position of cysteine. The first group (C-X-C) contains IL-8, growth-regulated oncogene (GRO) α , and GRO β and other chemokines with chemoattractant properties mainly, but not exclusively, towards neutrophils. The second (C-C) and third (C) groups comprise substances like monocyte chemoattractant protein (MCP)1, MCP2, MCP3, and macrophage inflammatory protein (MIP)1 α and MIP1 β , which mainly work on macrophages, lymphocytes, and eosinophils.

Recruitment of immune cells generally occurs by upregulating cell-surface adhesion molecule expression, such as endothelial leukocyte adhesion molecule (ELAM)-1 and increased ICAM-1, thereby enhancing cell adherence to the endothelial surface, and facilitating cell diapedesis through endothelial vessel walls. IL-8 mediates the recruitment and activation of immune cells in inflamed tissue, and increases endothelial permeability *in vitro*.

Plasma levels of IL-8 are raised in subjects with chronic alcoholism without liver disease and even more elevated in subjects with alcoholic hepatitis, who also show a rise in C-X-C chemokines. Alcohol abuse is a typical physiological stress that alters the distribution of monocytes and macrophages, and certain T-cell subsets, including a sharp decline in circulating CD4+CD62L+.

Regulation

Healthy human adult peripheral blood T cells express either CD4 or CD8. CD4 allows T cells to adhere by means of TcR to antigen-presenting cells (APC) that expose antigen in association with MHC-class 2. CD8 favors binding of the T cells to APC that present antigen via MHC-class 1. Most CD4+CD3+ cells act as TH cells, but a few CD4+ T cells are endowed with cytotoxic activity. CD8+ T cells are generally cytotoxic T lymphocytes, but they produce cytokines and growth factors that contribute to T cell-mediated immune regulation.

Membrane and Molecular Level

CD4 and CD8 are complexed lck, are members of the immunoglobulin gene superfamily, but are significantly divergent in their structure: CD4 is a monomer, while CD8 consists of two chains and occurs as a homodimer (α/α) or heterodimer (α/β). lck interacts with the CD4's 20 cytoplasmic relatively acidic residues (408–421), and with CD8 α 's 10 cytoplasmic residues (194–203) via ionic forces and by cysteine

residues. The CD8 β chain does not bind to lck. T cell stimulation leads to internalization of CD4:lck and its dissociation. The homodimeric form of CD8:lck may be internalized, but does not dissociate. The heterodimeric form of CD8:lck is not internalized, and does not dissociate. That is to say, active lck kinase is released in the cytoplasmic environment only following CD4+ cell activation.

A related src family member, fyn, associates primarily, but not exclusively with CD3, and other peptides generally implicated in the regulation of cell proliferation. The carboxyl domain of src family members exhibits a high degree of homology and contains the catalytic site. The N-terminal is unique to each peptide. The catalytic domain exhibits an adenosine triphosphate (ATP)-binding site at lysine 273, a site for autophosphorylation at tyrosine 394, and a site for tyrosine dephosphorylation mediated by CD45, or other phosphatases.

Related molecular events involved in the regulation of T cell stimulation include Akt (phosphoinositide 3-kinase, protein kinase B, PKB), the tumor suppressor, PTEN (phosphatase and tensin homolog deleted on chromosome 10, which removes phosphates primarily from lipids), Ras, mitogen-activated protein kinase (MAPK)/extracellular signal-regulated kinase (MEK), the dual threonine and tyrosine recognition kinase that phosphorylates and activates MAPK, and the novel tyrosine kinase inhibitor peptide, Tkip, a mimetic of SOCS-1 that binds to the autophosphorylation site of JAK2 and inhibits phosphorylation of STAT1 α . Tkip could serve as a potentially important therapeutic target for both constitutive and IL-6-induced STAT3 activation.

Suppressors of cytokine signaling (SOCS) constitute a family of seven known cytoplasmic proteins that regulate the feedback attenuation of cytokine-mediated signal transduction. IL-2 stimulation of T cells induces signaling via JAK/STAT, and SOCS-3 inhibits the JAK/STAT cytokine signaling cascade by directly binding to it, and by preventing access to the signaling complex. IL-2 enhances SOCS-3 messenger ribonucleic acid (mRNA) and protein synthesis, and its phosphorylation on tyrosine residues 204 and 221. SOCS-3 blocks the Ras/MEK activation pathway for c-fos and c-jun. Phosphorylation of SOCS-3 destabilizes its bond to proteasomes via elgulin B and C, and leads to proteasomal destruction. In the 3T3 fibroblast model, activation of SOCS-1 and SOCS-3 inhibits focal adhesion kinase (FAK) activity *in vitro* and tyrosine phosphorylation of FAK. Activation of SOCS promotes polyubiquitination and degradation of FAK in a SOCS box-dependent manner. Together, these events lead to inhibited FAK-dependent signaling events, including cell motility on fibronectin.

Phosphorylation of FAK is an essential regulatory step in T-cell migration.

Another important pathway related to SOCS is the P3 member of the first forkhead-box (FOX) transcription factor (FoxP3). FoxP3 functions as a transcriptional repressor, targeting composite NFAT/AP-1 sites in cytokine gene promoters. The region responsible for NFAT inhibition has now been mapped to the N-terminus of FoxP3. Introduction of FoxP3 into conventional mouse T cells converts these cells to the regulatory T cell (Treg) phenotype (CD4+CD25+).

Cellular Level

Treg are defined phenotypically as CD4+CD25+ T cells. They constitute 7–10% of CD4+ T cells in peripheral blood from healthy donors, and can rise to up to 20% in patients with lymphoma. Depletion of CD4+CD25+ cells significantly ($p < 0.05$) enhances the proliferative response of the remaining CD4+CD25– population, but addition of sorted CD4+CD25+ cells to the CD4+CD25– T cell population blunts proliferative activity. Sorted CD4+CD25+ T cells are susceptible to apoptotic cell death, associated with low Bcl-2 expression, and prevented by IL-2 salvage.

Neuropilin-1 (Nrp1) is a plasma transmembrane receptor (120 kD) involved in axon guidance, angiogenesis, and the activation of T cells. It is a ligand to the class 3 semaphorin subfamily and the heparin-binding forms of vascular endothelial growth factor and placenta growth factor. It acts as another specific surface marker for CD4+CD25+ Treg cells. Nrp1 mediates contacts between the dendritic cells and T cells via homotypic interactions and is essential for the initiation of the primary immune response. It is constitutively expressed on the surface of CD4+CD25+ Treg cells independently of their activation status. Nrp1 is downregulated in naïve CD4+CD25– T cells following TcR stimulation.

Functionally, Treg are endowed with suppressing T-cell activation. Expression of FoxP3, the transcription factor for the scurf protein product, is a master regulatory gene for the development and function of CD4+CD25+ subset. When transduced into T cells, FoxP3 recapitulates the hyporesponsiveness of naïve T cells, including impaired proliferation and production of cytokines including IL-2 and IL-10 upon TcR stimulation. FoxP3 is less efficient in reprogramming memory T cell subset into Treg cells, and is a reliable marker for functional Treg. Expression on the plasma membrane the glucocorticoid-induced TNF receptor (GITR) by Treg cells strongly suggests an important role for SOCS in neuroendocrine-immune regulation.

Different subsets of Treg cells have been identified. Naïve peripheral CD4⁺CD25[−] T cells can be induced *in vitro* into CD25⁺ anergic/suppressor cells through co-stimulation TcR and TGF- β , which induces Foxp3 gene expression. These TcR-challenged CD4⁺CD25[−] naïve can transit toward a regulatory T cell with Treg phenotype and potent immunosuppressive function. Treg are also divided into subsets according to cell-surface expression of CD62L. The suppressive function of the CD62L⁺ subset is far more potent on a per cell basis, can be expanded far more easily in culture, and is more responsive to chemokine-driven migration, compared to the CD62L[−] Treg population. Treg cells may be differentiated based on their expression of chemokine receptors, which may contribute to determine their migration properties: e.g., the traffic to and retention in bone marrow of certain Treg cells through CXCR4/CXCL12 signals.

Clinical Implications

Chemokine receptors serve as co-receptors for HIV-1 infection by binding to certain viral envelope proteins. For example, CXCR4 favors X4 HIV-1 strain tropism, and CCR5 favors tropism of R5 strain. Treg from healthy donors that express the HIV-co-receptor CCR5 are highly susceptible to HIV infection and replication. It is not surprising, therefore, that FoxP3-transduced T cells are markedly more susceptible to HIV infection, and that HIV-positive individuals with a low percentage of CD4⁺ and higher levels of activated T cells have greatly reduced levels of FoxP3⁺CD4⁺CD25⁺ T cells.

CD4⁺CD25⁺ T cells expand significantly in peripheral blood of HIV-seropositive patients receiving highly active antiretroviral therapy (HAART), and exhibit phenotypic (e.g., CD25), molecular (e.g., FoxP3), and functional characteristics (e.g., impaired proliferative and IL-2 production responses) of Treg cells. HIV-seropositive patients under HAART treatment supplemented with intermittent IL-2 show an increase of up to 79% (median, 61%) in CD4⁺CD25⁺CD69[−] cells, with a sharp increase in the proportion of cells expressing the CD45R0 memory phenotype, and Ki67 (proliferating cell antigen) expression.

Treg cells suppress Th1 cytokine responses, and ensure polarization of the Th2 pattern of cytokines during infection, such as schistosomiasis. A rat population of constitutive CD4⁺CD25⁺ T cells has been described that corresponds to Treg cells described in mice and humans, indicating that Treg phenomena are widespread among mammals.

Treg and T cells in general are critical to the regulation of immunity. We recall that T cells generally

express either one of two major subtypes of the TcR. Most T cells in the peripheral blood and in lymphoid organs express the heterodimeric α/β chains TcR. The majority of T cells that express the TcR composed of γ/δ proteins are enriched in epithelial tissues. This distinction is evident to a lesser degree in humans, compared to other mammals (e.g., murine model). Human peripheral blood contains less than 5% of T cells that bear the γ/δ TcR subtype, and in the oro-pharyngo-gastrointestinal mucosa, γ/δ T cells predominate. T cells that express the α/β TcR contribute to the immune surveillance of mucosal tissues. For example, the demonstration that T-cell clones derived from peripheral lymph nodes, and expressing α/β TcR, preferentially home to gingival tissue following infection with the dental plaque micro-organism, *Actinobacillus actinomycetemcomitans*. This migration is blocked by anti-very late antigen (VLA)-4 antibodies. Constitutive Treg cells emerge in the peripheral blood from the thymus, and play a critical role in the control of intestinal inflammation. Lymph node-derived Treg cells can effectively counter established, severe, and progressive colitis. Syngeneic Treg cells obtained from spleens and administered to mice with colitis contribute to the resolution of the *lamina propria* inflammation, and to the reappearance of normal intestinal architecture. Treg cells proliferate in the mesenteric lymph nodes and the inflamed colon of mice under treatment. Another plaque-related organism, *Fusobacterium nucleatum*, induces a threefold increase in CD69 expression by CD4⁺ T cells, and hence in the generation of inducible Treg, compared to a 4.15-fold activation in mitogen-stimulated control cultures. HgCl₂ at micromolar concentrations suppresses the *F. nucleatum*-mediated stimulation of CD69-expressing Treg by over 90%, and significantly ($p < 0.05$) blunts the production of IL-2. At supramicromolar concentrations (e.g., 10^{-5} M), HgCl₂ significantly ($p < 0.05$) increases the expression of CD25 by T cells by almost threefold. Taken together, these observations indicate that Treg from the peripheral and lymphoid system play an important role in the immune surveillance of inflammatory processes of mucosal immunity.

Conclusion

This article has discussed novel mechanisms of the regulation of T cell-mediated immunity. Future research in genomic, proteomic, interactomic, and molecular cartography in general will elucidate the fundamental pathways in immunity and autoimmunity, and reveal new and improved modes of therapeutic intervention for immune-related diseases.

Acknowledgments

Support for our research efforts was obtained from the National Institutes of Health, the Alzheimer's Foundation, the MacArthur Foundation, the UCLA Psychoneuroimmunology program, and the UCLA Schools of Medicine and Dentistry.

Further Reading

- Campos, M. and Godson, D. L. (2003). The effectiveness and limitations of immune memory: understanding protective immune responses. *International Journal of Parasitology* 33, 655–661.
- Chen, W., Jin, W., Hardegen, N., et al. (2003). Conversion of peripheral CD4⁺CD25⁻ naive T cells to CD4⁺CD25⁺ regulatory T cells by TGF- β induction of transcription factor Foxp3. *Journal of Experimental Medicine* 198, 1875–1886.
- Chiappelli, F., Kavelaars, A. and Heijnen, C. J. (1992). β -endorphin effects on membrane transduction in human lymphocytes. *Annals of the New York Academy of Sciences* 650, 211–217.
- Chiappelli, F. and Liu, N. Q. (2000). Immunity. In: Fink, G. (ed.) *Encyclopedia of stress* (vol. 2), pp. 541–547. New York: Academic Press.
- Chiappelli, F., Kung, M. A., Stefanini, G. F., et al. (2000). Alcohol and immune function. In: Gershwin, M. E., German, J. B. & Keen, C. L. (eds.) *Nutrition and immunology: principles and practice*, Chap. 21, pp. 261–274. New Jersey: Humana Press.
- Chiappelli, F., Prolo, P., Iribarren, J., et al. (2006). Alzheimer's disease: new frontiers for the XXI century. In: Welsh, E. M. (ed.) *Advances in psychology research*, pp. 233–258. New York: Nova Science Publisher.
- Chiappelli, F., Prolo, P., Fiala, M., et al. (2006). Allostasis in HIV-1 infection and AIDS. In: Minagar, P. A. & Shapshak, P. (eds.) *Neuro-AIDS*. New York: Nova Science Publisher (in press).
- Cogoli-Greuter, M., Lovis, P. and Vadrucchi, S. (2004). Signal transduction in T cells: an overview. *Journal of Gravitational Physiology* 11, 53–56.
- Crotty, S. and Ahmed, R. (2004). Immunological memory in humans. *Seminars in Immunology* 16, 197–203.
- Eisenlohr, L. C. and Rothstein, J. L. (2005). Antigen processing and presentation. *Cancer Treatment Research* 123, 3–36.
- Esche, C., Stellato, C. and Beck, L. A. (2005). Chemokines: key players in innate and adaptive immunity. *Journal of Investigative Dermatology* 125, 615–628.
- Larsen, L. and Ropke, C. (2002). Suppressors of cytokine signalling: SOCS. *Acta Pathologica, Microbiologica, et Immunologica Scandinavica* 110, 833–844.
- Liu, E., Cote, J. F. and Vuori, K. (2003). Negative regulation of FAK signaling by SOCS proteins. *EMBO Journal* 22, 5036–5046.
- Maffei, A. and Harris, P. E. (1998). Peptides bound to major histocompatibility complex molecules. *Peptides* 19, 179–198.
- Magee, T., Pirinen, N., Adler, J., et al. (2002). Lipid rafts: cell surface platforms for T cell signaling. *Biological Research* 35, 127–131.
- Matsuzaki, J., Tsuji, T., Imazeki, I., et al. (2005). Immunosteroid as a regulator for Th1/Th2 balance: its possible role in autoimmune diseases. *Autoimmunity* 38, 369–375.
- Meiri, K. F. (2004). Membrane/cytoskeleton communication. *Sub-Cellular Biochemistry* 37, 247–282.
- Nishimura, M. I., Roszkowski, J. J., Moore, T. V., et al. (2005). Antigen recognition and T-cell biology. *Cancer Treatment Research* 123, 37–59.
- Piccirillo, C. A. and Shevach, E. M. (2004). Naturally-occurring CD4⁺CD25⁺ immunoregulatory T cells: central players in the arena of peripheral tolerance. *Seminars in Immunology* 16, 81–88.
- Reith, W., LeibundGut-Landmann, S. and Waldburger, J. M. (2005). Regulation of MHC class II gene expression by the class II transactivator. *Nature Reviews Immunology* 5, 793–806.
- Romagnani, S. (1992). Human TH1 and TH2 subsets: Regulation of differentiation and role in protection and immunopathology. *International Archives of Allergy and Immunology* 98, 279–285.
- Taams, L. S., Smith, J., Rustin, M. H., et al. (2001). Human anergic/suppressive CD4⁺CD25⁺ T cells: a highly differentiated and apoptosis-prone population. *European Journal of Immunology* 31, 1122–1131.
- Trowsdale, J. (2005). HLA genomics in the third millennium. *Current Opinions in Immunology* 17, 498–504.
- van der Merwe, P. A. (2002). Formation and function of the immunological synapse. *Current Opinions in Immunology* 14, 293–298.
- van Seventer, G. A., Salmen, H. J., Law, S. F., et al. (2001). Focal adhesion kinase regulates β 1 integrin-dependent T cell migration through an HEF1 effector pathway. *European Journal of Immunology* 31, 1417–1427.
- Watford, W. T., Hissong, B. D., Bream, J. H., et al. (2004). Signaling by IL-12 and IL-23 and the immunoregulatory roles of STAT4. *Immunological Reviews* 202, 139–156.
- Weiss, L., Donkova-Petrini, V., Caccavelli, L., et al. (2004). Human immunodeficiency virus-driven expansion of CD4⁺CD25⁺ regulatory T cells, which suppress HIV-specific CD4 T-cell responses in HIV-infected patients. *Blood* 104, 3249–3256.
- Yagi, H., Nomura, T., Nakamura, K., et al. (2004). Crucial role of FOXP3 in the development and function of human CD25⁺CD4⁺ regulatory T cells. *International Immunology* 16, 1643–1656.
- Zou, L., Barnett, B., Safah, H., et al. (2004). Bone marrow is a reservoir for CD4⁺CD25⁺ regulatory T cells that traffic through CXCL12/CXCR4 signals. *Cancer Research* 64, 8451–8455.

Immunomodulation See: Neuroimmunomodulation.

Impact of Terrorism on the Development of Mental Health Symptoms

R Yehuda

Bronx VA Medical Center, New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction

Relationship between Terrorism and Other Traumatic Experiences

Differentiating between Acute Mental Health Symptoms That Will Probably Abate and Those That Will Develop into Long-Term or Pathological Mental Health Consequences

Risk Factors for Chronic Posttraumatic Symptoms

The Role of Biological Studies in Helping to Identify Pathological Responses

Introduction

Because the broad goals of terrorism are ultimately psychological, it is imperative to understand the magnitude and scope of the psychological casualties in the wake of terrorism and their biological correlates so as to identify the appropriate means of immunizing people to these effects. Although the majority of people exposed to terrorism experience acute distress, the most important observation to emerge from longitudinal studies of people exposed to terrorism is that these symptoms are transient for the great majority. It therefore becomes important from a public health perspective to be able to predict those at greatest risk for long-term problems. Those at risk for long-term symptoms appear to have a more intense reaction at the time of the trauma (e.g., peritraumatic dissociation, panic, or related emotional distress), associated with a more negative appraisal of danger. It is not clear whether and to what extent pretraumatic risk factors (e.g., prior adversity, family history of psychopathology, cognitive risk factors, or preexisting personality traits) influence the intensity of the peritraumatic response to trauma, but presumably these risk factors are more relevant in situations in which exposure is less severe. Posttraumatic risk factors, such as lack of social support, also seem to

be important predictors of psychopathology, but the extent to which these reflect pre- or even peritraumatic risk factors is currently unknown.

In considering the strength of the information available regarding the psychological consequences of terrorism and their biological correlates, it is important to point to gaps in our knowledge with respect to these issues that currently preclude a thorough understanding of the psychological and biological consequences of terrorism. The first concerns whether psychological outcomes related to terrorism are comparable with those observed in the case of other traumatic events such as combat, interpersonal violence, and natural disasters, whose immediate and long-term consequences have been well established. The second concerns whether the immediate effects of terrorism are pathological, that is, requiring mental health intervention. In the absence of empirical data regarding these gaps in knowledge, governmental and privately sponsored interventions following terrorist attacks are modeled after those developed for responding to natural disasters. Thus, it is important to supply the missing information so that responses to terrorism can be tailored to the specific needs of survivors.

Relationship between Terrorism and Other Traumatic Experiences

Terrorism is a prototypic traumatic event (i.e., it meets the objective and subjective criteria as defined according to the *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn., DSM-IV) for those who directly experience a threat to their life or physical integrity or who experience loss, most notably, the sudden loss of a loved one. However, terrorism is not only about life threat to single individuals, or even a small group of people; it is designed to instill fear in society at large. Thus, in addition to those who experience the effects of terrorism firsthand, there are collateral effects. People who were not in proximity to the target where the terrorist event occurred and people who were not directly affected by the loss of

someone important to them can also be affected in terms of mental health outcomes, particularly because such events often receive repeated coverage on television and other media outlets. It is not known whether mental health consequences in people indirectly exposed to terrorism are qualitatively or quantitatively different from those who have been directly exposed. The additive effect of anticipatory anxiety is also not known, yet the real threat of imminent attack may be related to how quickly those exposed to terrorism can recover from its effects.

In applying knowledge from what has already been learned about the effects of other traumatic stressors, it is clear that a key variable in considering pathological outcomes is the passage of time. Although there are certainly immediate symptoms following any traumatic event, there is debate about whether and how to intervene in the acute aftermath of any trauma because, in most cases, such symptoms abate within weeks and months, and the consequences of interfering with the natural healing processes are unknown. One difference between terrorism and other traumatic events is that a single act of terrorism may represent the beginning or continuation of a situation or threat. In the context of an ongoing terrorist threat, identifying a persistent disorder may be more accurately defined only after the immediate threat of terrorism is substantially reduced in reality. This caveat does not suggest that people do not have mental health needs that should be met within this time frame; rather, it acknowledges that pathological responses that are defined by persistence and chronicity need to be considered in the context of actual ongoing threat.

Differentiating between Acute Mental Health Symptoms That Will Probably Abate and Those That Will Develop into Long-Term or Pathological Mental Health Consequences

The high rate of spontaneous recovery in the aftermath of terrorism raises an important question with respect to how initial symptoms should be understood and treated. If the initial symptoms are understood as transient reflections of distress, it is not clear that such symptoms require mental health intervention. Alternatively, it can be argued that treating early symptoms, even in those who will experience spontaneous recovery, might help prevent long-term responses in the 7–12% who generally remain symptomatic or deteriorate as the broad population begins to improve. Clarifying the causes of high levels of immediate and long-term symptoms

will no doubt lead to ideas about potential preventative treatments because it is now clear that the intensity of the exposure to trauma does not sufficiently predict who will develop or sustain symptoms.

In the absence of a linear relationship between acute stress reactions and longer-term adjustment, there are two theoretical models for pathological responses in the aftermath of a trauma. The first defines a pathological response to trauma by the persistence of the same symptoms that within a specified period of time immediately following the event would otherwise be considered normal. The second defines posttraumatic pathology by the presence of a fundamentally different initial response to stress, resulting in persistent symptoms of hyperarousal, recollection of intrusive events, and avoidance of reminders. Both models are consistent with the idea that posttraumatic mental health consequences are based on a failure to recover from the universal distress in the immediate aftermath of trauma. In the first case, the persistence of symptoms probably results from post-exposure factors that prevent recovery, such as lack of social support or other coping resources. In contrast, the second model implies that the initial symptoms are fundamentally different at the outset and probably result from pre- or peritraumatic factors.

Supporting the idea that persistent responses are simply continuations of initial reactions are the findings that people who show high-magnitude post-traumatic stress disorder (PTSD) symptoms (i.e., intrusive, avoidance, and hyperarousal) in the first few days and from 1 to 2 weeks after the event are those at greatest risk for subsequent symptom severity. However, many researchers performing longitudinal studies of acute trauma survivors have not been able to identify specific symptoms that can convincingly predict pathological mental health outcomes. More promising clinical predictors appear to be the presence of a panic attack and/or intense dissociation during or immediately after a trauma, which suggests additional or fundamentally different immediate responses. Both panic attacks and dissociative responses reflect catastrophic or negative interpretations of the event. Such interpretations may further perpetuate arousal associated with stress responses and reduce the body's ability to achieve homeostasis following the trauma. The presence of catastrophic attributions (e.g., that posttraumatic symptoms signify a major problem and will not ultimately resolve) provides a way of drawing a distinction between being terrorized and being terrified. People exposed to a traumatic event are not traumatized unless they are deeply distraught at the time of the event and then make

catastrophic interpretations that may result in fundamentally different biological stress responses.

Indeed, although trauma exposure certainly results in the release of stress hormones such as catecholamines and cortisol, it has been noted that dissociation and panic attacks are specifically associated with increased catecholamine levels. It may be that a peritraumatic panic attack or other intense distress during and immediately after the traumatic event leads to higher levels of stress hormones that might, in turn, strengthen traumatic memories and increase the probability of intrusive recollections. On the other hand, extreme reactions of panic and dissociation may be facilitated by relatively lower levels of stress hormones, such as cortisol, that play an important role in containing the catecholamine response to stress.

Risk Factors for Chronic Posttraumatic Symptoms

The finding that only a proportion of those exposed to trauma develop short- and long-term symptoms justifies an exploration of both the factors that increase the risk for and the factors that might serve to protect individuals from developing such symptoms following trauma exposure. Although there is not much known about the long-term consequences of terrorism specifically, there is information about factors that contribute to the development of chronic PTSD following exposure to other types of traumatic events. Putative risk factors for PTSD can describe the characteristics of the event or the response to the event or the characteristics of the people who experience those events. Indeed, a wide variety of risk factors, ranging from situational or environmental factors to familial or even genetic risk factors, for PTSD have now been identified. The identification of specific patterns of risk factors is relevant to an understanding of different types of acute reactions and for providing prophylaxis.

It appears on the basis of retrospective studies that those at greatest risk for developing chronic symptoms following a traumatic event are people with prior exposure to trauma, specific cognitive factors, and certain preexisting personality traits, such as proneness to experiencing negative emotions and having poor social supports. However, increasingly, family and genetic studies are pointing toward the role of genetic contributions to posttraumatic stress reactions. The adult children of Holocaust survivors with PTSD show a greater prevalence of PTSD to their own traumatic events than the adult children of Holocaust survivors without PTSD. It is difficult

to know the extent to which increased vulnerability to PTSD in family members of trauma survivors with PTSD can be attributed to being raised by a parent with this disorder, to exposure of the fetus to the mother's stress hormones during pregnancy, or to genetic factors. Evidence that family resemblance with respect to PTSD may be at least partly explained by genes comes from twin studies in which concordance for PTSD was examined in monozygotic and dizygotic twins. However, identifying those at risk will depend on identifying how these genetic factors interact with patterns or risk factors such as childhood trauma, low educational attainment, personal or family history of anxiety or mood disorders, history of heavy alcohol use in the months prior to the incident, poor social supports in the posttraumatic period, and greater levels of exposure during the incident together with symptoms in the peritraumatic period.

The Role of Biological Studies in Helping to Identify Pathological Responses

There appears to be a distinct set of biological alterations in trauma survivors who develop chronic posttraumatic symptoms. Some of these reflect changes in stress-responsive systems that are qualitatively different than those predicted on the basis of the stress literature. Furthermore, the alterations in individuals with PTSD have been found to be distinct from those of similarly exposed individuals without PTSD and also different from those found in other psychiatric disorders, such as mood and other anxiety disorders. Together, these findings suggest that the biology of PTSD is not simply a reflection of a normative stress response but, rather, a pathological one.

For example, in contrast to the normal fear response, which is characterized by a series of biological reactions that help the body cope with, and gradually recover from, stress (e.g., high cortisol levels), some recent prospective biological studies have demonstrated that individuals who develop PTSD symptoms appear to have attenuated cortisol increases in the acute aftermath of a trauma compared to those who do not develop this disorder. Moreover, people who develop PTSD show higher heart rates in the emergency room and at 1 week posttrauma compared to those who ultimately recover, suggesting a greater degree of sympathetic nervous system activation. Thus, it may be that long-term symptoms following trauma are facilitated by a failure to contain the normal stress response at the time of the trauma, resulting in a cascade of biological alterations that lead to intrusive, avoidance, and hyperarousal symptoms.

This difficulty may occur because of pretraumatic biological risk factors.

Basic science research on both memory consolidation and fear conditioning have demonstrated that heightened adrenergic activation can promote the consolidation and retrieval of fear-provoking memories. Thus, extreme sympathetic arousal at the time of a traumatic event may result in the release of stress-related neurochemicals (including norepinephrine and epinephrine), mediating an overconsolidation of trauma memories. The enhanced elevation of catecholamines in the immediate aftermath of a traumatic event may increase the probability of intrusive recollections in the first few days and weeks posttrauma.

One of the major gaps in our knowledge concerns the interplay of biological responses at the time of a traumatic event and cognitive factors. Indeed, it is easy to see how an increased catecholaminergic response to trauma could be the proximal cause of intense panic. Furthermore, it seems plausible that people who find themselves in a more intense biological state of fear might be likely to appraise a situation as more immediately dangerous. On the other hand, it may be that preexisting cognitive factors mediate the catecholaminergic response to the trauma that leads to the panic. These preexisting cognitive factors, in turn, may or may not be the cause, result, or correlate of a preexisting biological alteration that sets the stage for an extreme response.

Clarifying the causes of high levels of immediate and long-term symptoms will no doubt lead to ideas about potential preventative treatments. For example, to the extent that panic reactions are associated with increased catecholamine responses at the time of trauma, aggressive intervention with adrenergic-blocking agents (such as propranolol) or cortisol or cognitive-behavioral stress management techniques emphasizing relaxation rather than any form of retelling (i.e., reexposure, as in debriefing) may be

the most appropriate immediate interventions for people who panic in the hours immediately following a traumatic experience. For people who are not in a position to receive intervention within several hours or days posttrauma, it will be necessary to determine the impact of posttraumatic risk factors such as lack of social support on the longitudinal course of pathological responses to trauma.

Further Reading

- Breslau, N., Chilcoat, H. D., Kessler, R. C., et al. (1999). Previous exposure to trauma and PTSD effects of subsequent trauma: results from the Detroit Area Survey of Trauma. *American Journal of Psychiatry* 156, 902–907.
- Bryant, R. A. and Panasetis, P. (2001). Panic symptoms during trauma and acute stress disorder. *Behaviour Research Therapy* 39, 961–966.
- Galea, S., Ahern, J., Resnick, H., et al. (2002). Psychological sequelae of the September 11 terrorist attacks. *New England Journal of Medicine* 346, 982–987.
- Keane, T. M., Kaufman, M. L. and Kimble, M. O. (2001). Peritraumatic dissociative symptoms, acute stress disorder, and the development of posttraumatic stress disorder: causation, correlation or epiphenomena. In: Sanchez-Planell, L. & Diez-Quevedo, C. (eds.) *Dissociative states*, pp. 21–43. Barcelona: Springer-Verlag.
- Schelling, G., Kilger, E., Roozendaal, B., et al. (2004). Stress doses of hydrocortisone, traumatic memories, and symptoms of posttraumatic stress disorder in patients after cardiac surgery: a randomized study. *Biological Psychiatry* 55, 627–633.
- Yehuda, R. (1999). *Risk factors for posttraumatic stress disorder*. Washington, DC: American Psychiatric Press.
- Yehuda, R. (2002). Post-traumatic stress disorder. *New England Journal of Medicine* 346, 108–114.
- Yehuda, R., Engel, S. M., Brand, S. R., et al. (2005). Transgenerational effects of posttraumatic stress disorder in babies of mothers exposed to the World Trade Center attacks during pregnancy. *Journal of Clinical Endocrinology and Metabolism* 90, 4115–4118.

Impotence, Stress and

M R Gignac, G M Rooker and J K Cohen
Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2,
pp 547–550, © 2000, Elsevier Inc.

Overview
Physiology of Erection
Erectile Dysfunction
Evaluation and Treatment

Glossary

<i>Erectile dysfunction (ED)</i>	Failure to obtain sufficient erectile rigidity of the penis to allow vaginal penetration, also known as impotence.
<i>Erection</i>	Tumescence or engorgement of the corporal bodies leading to elongation, widening, and rigidity of the penis.
<i>Intrinsic/extrinsic stress</i>	Psychological forces a person experiences. Intrinsic stress, for example, could result from guilt over marital infidelity; extrinsic stress, for example, could result from the death of one's parents.
<i>Libido</i>	The desire to have sex.
<i>Nocturnal erection</i>	Normal physiological erection of the penis that occurs subconsciously during the rapid eye movement sleep cycle and, when present, is thought to exclude organic disease as a cause of erectile dysfunction.

Overview

The inability to obtain sufficient penile rigidity to allow vaginal penetration, regardless of the cause, is termed impotence or is perhaps better described as erectile dysfunction (ED). ED is an exceedingly common problem. In fact, nearly half of all men over the age of 40 suffer from some degree of ED at various times. ED is also strongly associated with advancing age. ED is commonly related to medical conditions as well as psychological disorders, and most physicians agree that in many cases there is significant overlap. Stress in the medical literature has been defined as physical or psychological forces experienced by a person. An example of a physical force leading to ED is severe vascular disease; this leads to failure of the physical erectile mechanism. A typical psychological stress leading to ED is depression, which may not

only be the cause but also the result of ED. The role stress plays in the development of ED is significant and often difficult to define.

Physiology of Erection

Neurological and Psychological Aspects

The events that lead to normal penile erection are complex and involve psychological and physiological mechanisms. Volitional erection usually begins with psychological arousal. This involves all of the senses (sight, touch, smell, taste, and hearing) as well as imagination. The sensory input arrives to special areas of the conscious brain via afferent neural pathways. In the case of positive input, the signals are used to create the sensations and fantasy associated with sexual arousal. This process can be modulated by stress and hormonal imbalance. Next are the descending pathways from the brain, leading to erection centers in the spinal cord. From the spinal cord, the pathway leads to the paired cavernosal nerves running alongside the prostate (neurovascular bundles), which are responsible for the induction of erection. This series of neural connections can collectively be termed the efferent pathway. Stimulation of the genital organs sends signals to the spinal erection center, some of which are transmitted up to the brain and others of which lead to the reflex stimulation of the cavernosal nerves. Thus, three types of penile erection are thought to exist. Psychogenic erection is the result of conscious sensory input and fantasy. Reflexogenic erection occurs by a reflex pathway in the spinal cord initiated by genital stimulation. Finally, nocturnal erections occur by an as yet unknown mechanism.

Mechanical Aspects: Penile Hemodynamics

The physiological mechanics of penile erection are fairly well understood. Essentially, the efferent neural activation leads to a change in the blood flow pattern to the penis whereby more blood is trapped in the paired erectile bodies (corpora cavernosum). Thus, there is greater inflow compared to outflow, and the penis elongates, increases in diameter, and develops sufficient rigidity for vaginal penetration. This requires normally functioning arterial, cavernosal, and venous systems. It is obvious on the basis of our current understanding of sexual function that a complex interaction must occur between the psychological and physical aspects of erection and there are many potential areas in which stress can affect the outcome.

Erectile Dysfunction

For treatment purposes, clinicians commonly classify patients as having either organic or nonorganic causes of ED. This division roughly corresponds to physical and psychological causes, respectively.

Organic and Nonorganic Erectile Dysfunction

Organic causes of ED relate to specific malfunctions in the physiological mechanisms of erection, in other words, any pathological process that affects the afferent input, efferent output, or mechanics of erection. For example, multiple sclerosis can disrupt the afferent and efferent pathways, vascular disease or antihypertensive medications can interfere with the erectile hemodynamics, and diabetes may affect multiple physiological mechanisms.

Nonorganic ED refers to problems that cannot be explained easily by the failure of physiological mechanisms; it is also commonly referred to as psychogenic impotence. This type of ED has defied universally accepted classification; however, it has been subdivided into five major categories: type 1, anxiety, fear of failure (e.g., performance anxiety); type 2, depression; type 3, marital conflict, strained relationship; type 4, ignorance and misinformation, religious scruples (e.g., about normal anatomy or sexual function); and type 5, obsessive-compulsive personality (e.g., anhedonia).

Stress and Erectile Dysfunction

The role stress plays in the causation of ED has been difficult to study and therefore information is lacking. Several commonsense inferences can be made. First, the presence of ED, regardless of the cause, leads to psychological stress such as fear of underlying disease and loss of self-image. Second, intrinsic and extrinsic causes of psychological stress can lead to physiological changes altering the afferent and efferent pathways as well as the mechanical function of erection. For example, anxiety states can be associated with higher circulating levels of catecholamines and these neurotransmitters can affect the hemodynamics of the penis, causing failure of the normal erectile mechanism. This may explain some cases of ED due to performance anxiety. Finally, emotional tone can have a strong influence on how the senses are interpreted, potentially turning positive signals into inhibition of erection; thus, stress can diminish libido profoundly.

Evaluation and Treatment

History and Physical Examination

The evaluation of patients with ED has generally been performed by either urologists or psychiatrists. With changing attitudes toward sexuality and the advancement in the science of ED, many different types of

doctors now possess the appropriate evaluation and management skills. Still, however, sexual dysfunction often remains overlooked in the treatment of many patients. The initial evaluation begins with a thorough history and physical examination. In cases in which psychological factors seem to predominate, the careful elucidation of the history is especially important and often is the only means by which the psychological stress can be discovered. The encounter is best suited to a quiet, comfortable examining room and the history taken with the patient fully clothed, followed by the physical examination. This allows the interaction to occur in a minimally intimidating format. Elements of the history that are relevant include a specific description of the degree of dysfunction, duration, circumstances associated with the onset, coexisting medical problems, loss of libido, past surgeries, injuries, medications, psychiatric problems, and temporal relationship to major life events. On occasion, it is informative to interview the partner separately. The physical examination should focus on signs of systemic disease; thorough neurological evaluation; careful inspection of the genitalia, including rectal examination; and assessment for hormonal and vascular diseases.

Special Testing

On the basis of the initial interview, additional or second-line studies may be needed. These additional tests could include nocturnal penile tumescence (NPT) monitoring, injection cavernosometry and cavernosography, hormonal evaluation, and personality questionnaires.

Nocturnal penile tumescence monitoring The application of NPT monitoring is simple and can usually be performed at home by the patient. Either a one-time-use snap gauge device or an electronic monitor is applied to the penis at bedtime. The presence of normal nocturnal erections is considered proof that the mechanical aspects of penile erection are functioning appropriately. Both devices can measure the presence of an erection and, to some degree, the quality. The more sophisticated electronic monitor provides additional information, such as the number and duration of erections.

Injection cavernosometry and cavernosography Injection cavernosometry and cavernosography are invasive tests in which a vasoactive drug (e.g., prostaglandin) is injected directly into the body of the penis. The drugs work by stimulating the same hemodynamic changes as in normal erection. Before and after injection, pressures within the corporal bodies can be measured to determine an objective degree of tumescence, high-resolution ultrasound can be used to assess vascular changes, and intravenous contrast can be administered for the evaluation of venous leak.

These tests are used to detect arterial (inflow) and venous (outflow) diseases.

Both NPT and injection testing are intended to evaluate the physiological mechanism of erection. The use of these tests is directed not only by the suspected diagnosis but also, and perhaps more important, by potential therapeutic alternatives. Stated differently, extensive physiological testing affects the choice of treatment options infrequently and therefore is indicated only in select cases.

Hormonal evaluation Suspicion of a hormonal imbalance needs to be confirmed by blood tests assessing the function of the pituitary (prolactin), thyroid (thyroid hormones T_3 and T_4 and thyrotropin), and testicle (testosterone). Common abnormalities causing decreased libido include hypogonadism and hypothyroidism. It is important to note that psychological stress can decrease libido without lowering testosterone or thyroid hormone levels. If testosterone is found to be low, then prolactin should be measured; a significant elevation of prolactin can indicate a pituitary problem.

Personality questionnaires Another second line type of testing that is seldom used by urologists and occasionally used by psychiatrists is personality questionnaires. Many have been developed (e.g., the Minnesota Multiphasic Personality Inventory), and they are intended to elicit information similar to a sexual history. The use of this technique may offer advantages in assessing the degree of psychological stress. Many clinicians feel that a detailed personal interview is equally effective, if not more so, and requires less interpretation skills than these instruments. This explains their infrequent use.

Treatment

For the successful management of ED, it is necessary to understand the degree to which both psychological and physiological factors are contributing. For instance, the successful correction of anxiety disorder is helpful, but ignoring concomitant vascular disease will result in treatment failure. Likewise, the placement of a penile prosthesis can restore an adequate erection, but if the patient has a depressed mood he may not be interested in having sex. Treatment is therefore directed toward both the correction of physiological malfunctions and behavioral modification. Furthermore, managing the physiological and psychological factors must be considered in the context of available treatment options. A precise cause can often be determined, particularly when physiological disease predominates, but cannot be precisely corrected given the current limitations in treatment. Thus, in most cases of ED, treatment is directed at the symptoms and not the specific cause.

Physiological treatments There are several approaches to the physiological treatment of ED. These options, from least to most invasive, include an external vacuum erection device (VED); medications in oral pill, intraurethral pellet, and intracavernosal injection; and surgical prosthesis placement. The VED works by creating negative pressure around the penis, thereby causing blood engorgement and erection. It is useful in all types of ED and leaves no permanent changes. Various medications have been used and all affect penile hemodynamics. Oral medications leave no permanent change and effectiveness depends on the cause of ED. Typically, injection therapy is considered most reliable, but it can ultimately lead to corporal fibrosis. Penile prostheses are mechanical devices that are surgically implanted to replace the cavernosal tissue. Implants come in a variety of types, including semirigid and inflatable. In general, treatment is offered from the simplest to the most invasive, depending on the patient's wishes and therapeutic effectiveness.

A major breakthrough in ED medicament treatment was provided by the discovery that phosphodiesterase 5 (PDE5) inhibitors were effective in helping to induce penile erection. PDE5 inhibitors increase the concentration of cyclic guanosine monophosphate in the corpus cavernosum: this results in smooth muscle relaxation (vasodilation) with consequent increased inflow of blood into the corpus cavernosum and penile erection. In 1998, the FDA approved a potent PDE5 inhibitor, sildenafil ("Viagra"), as an oral treatment for ED. Since then several clinical trials, including three independent double blind controlled trials in patients with erectile dysfunction, have demonstrated that sildenafil was significantly more effective than placebo in producing sustained improvements in confidence, self-esteem, and almost normal erectile function. Other effective PDE5 inhibitors, tadalafil ("Cialis") and vardenafil ("Levitra"), are now also available and widely used for the treatment of ED.

Psychological treatment Incorporated into any ED treatment plan must be a provision for the psychological component. This is particularly relevant for ED associated with stress and is probably most useful in patients without significant organic disease. First, any discovered psychological diagnosis, such as depression, should be treated appropriately. Second, sex therapy counseling should be used, not only to uncover functional problems but also to direct intervention and alter negative patterns of behavior. Sex counseling is usually performed by therapists specially trained in sex therapy. The session or sessions will probably require both partners to participate and may involve assignments and exercises to be performed at home. The goals of this intervention are to identify and correct behavioral patterns and to develop coping skills

for sources of internal and external stress. In many cases of stress-associated ED, a combination of both physiological and psychological therapies are used.

See Also the Following Article

Sexual Dysfunction.

Further Reading

- Allen, J. R. (1987). Psychiatric treatment of erectile dysfunction. *Journal of the Oklahoma State Medical Association*, 80, 19.
- Althof, S. E., O'Leary, M. P., Cappelleri, J. C., Crowley, A. R., Tseng, L.-J. and Collins, S. (2006). Impact of erectile dysfunction on confidence, self-esteem and relationship

- satisfaction after 9 months of sildenafil citrate treatment. *The Journal of Urology* 175, 2132–2137.
- Blake, D. J. (1984). Issues in the psychiatric diagnosis and management of impotence. *Psychiatric Medicine* 2, 109.
- Hawton, K., Catalan, J. and Fagg, J. (1992). Sex therapy for erectile dysfunction: characteristics of couples, treatment outcome, and prognostic factors. *Archives of Sexual Behavior* 21, 161.
- Lue, T. F. (1998). Physiology of penile erection and pathophysiology of erectile dysfunction and priapism. In: Walsh, P. C., et al. (eds.) *Campbell's urology* (7th edn., pp. 1157–1179). Philadelphia: Saunders.
- Tiefer, L. and Schuetz-Mueller, D. (1995). Psychological issues in diagnosis and treatment of erectile disorders. *Urologic Clinics of North America* 22, 767.

Impulse Control

I-M Blackburn

Newcastle Cognitive and Behavioral Therapies Centre,
Newcastle upon Tyne, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 551–553, © 2000, Elsevier Inc.

Problems Associated with Uncontrolled Anger

Assessment of Anger

Methods of Anger Control

Conclusion

Glossary

Cognitive behavior therapy

A short-term psychotherapy for emotional disorders that aims at symptomatic, affective, and behavioral improvement through the evaluation and modification of an individual's dysfunctional way of processing information.

Idiographic measures

Measures that are not standardized but are devised to assess problems specific to an individual.

Nomothetic scales

Scales validated and standardized in normals and specific comparison groups, yielding norms for different client populations.

Impulse control can relate to several aspects of behavior, for example, anger, substance abuse, repeated self-harm, or suicidal attempts. Anger control is the primary concern of this article.

Problems Associated with Uncontrolled Anger

Anger can be seen as a normal emotional reaction experienced universally by adults and children, as well as by animals. As such, it serves a useful purpose in letting others know that their behavior is aversive so that they can alter their behavior if amicable or respectful relationships are to be maintained or restored. Anger directed at the self allows one to register unpleasant arousal, perhaps at transgressing one's own rules, again allowing for changes to occur. Anger or hostility can be seen as being composed of affective (irritability, full blown anger), physiological (increased arousal), cognitive ("you should not do that," "you are overstepping the mark"), and behavioral (sarcasm, verbal abuse, assault) components. It is when these interacting components reach the point at which anger and hostility are no longer controllable and are expressed in ways that entail negative social and interpersonal consequences that they become a problem. Thus, in the clinic, patients complain of their increased irritability, of anger being provoked by minor triggers, and of their fear of catastrophic consequences. In forensic settings, anger and impulse control are often the main treatment targets.

Anger and hostility have not been widely studied in the clinical literature, relative, say, to depression and anxiety. However, the rising rates of violent crime among adolescents, acts of uncontrolled anger as seen in road rage, violent familial discord, child abuse, and assaults have made uncontrolled anger and hostility a topical subject in the media. Moreover, poorly

managed anger has been linked convincingly with a number of health problems, such as coronary artery disease and essential hypertension. Friedman and Rosenman postulated that type A behavior, characterized by striving for achievement, competitiveness, vigilance, impatience, and easily aroused hostility, predicted coronary heart disease (CHD). The wealth of research that followed has implicated anger and hostility as the component of the type A complex that best predicts CHD and mortality in general.

Several explanations have been put forward to explain the link between hostility/anger and health. The psychophysiological reactivity model proposes that hostile individuals react to stressors with extreme cardiovascular and neuroendocrine responses, which, over time, lead to the development of CHD and CHD symptoms such as myocardial infarction and angina. The psychosocial vulnerability model implicates increased interpersonal stressors and decreased social support, both at work and in daily life. The transactional model integrates the first two models – the hostile individual elicits negative reactions from others, which helps to confirm his or her views of others as antagonistic, increasing hostile tendencies further. Thus, there is a reciprocal relation between the cognitive and behavioral aspects of hostility, leading to increasingly exaggerated cardiovascular and endocrine reactions to intensifying interpersonal stressors. In a study of male patients with postmyocardial infarction and unstable angina, Hall and Davidson found that patients who had higher scores on a measure of hostile style, as rated during a standard interview, perceived the interviewer as more aggressive, although this perception was not corroborated by an independent rater.

Assessment of Anger

The assessment of anger in research studies is usually by standard interview or by nomothetic scales. Some studies have used physiological measures, such as systolic blood pressure in response to potentially anger-arousing stimuli; others have used observers' ratings of facial expressions or behavioral measures, such as hand dynamometer readings.

Interview methods are not validated easily, are time-consuming to administer, and are costly in terms of interviewer training and scoring. Self-rated questionnaires, tested psychometrically for their validity and reliability, have been most widely used. The Cook–Medley hostility scale relates more to cognitive than behavioral aspects of anger; in particular, high scores reflect a negative view of disparagement by others. The Buss–Durkee hostility inventory attempts to measure angry cognitions as well as the verbal and physical expression of anger. Anger, as a trait, can

be assessed by the trait anger scale, which measures cognitive and affective aspects of anger. Novaco developed the anger inventory to measure anger in response to a wide range of potential provocations. Deffenbacher and colleagues have devised a number of idiographic measures to assess anger in day-to-day living; the most intense, ongoing source of anger for the individual; and anger-related physiological arousal. These three idiographic measures are correlated positively with one another and with other indices.

More recently, Snyder and colleagues have developed a scale to measure hostile automatic thoughts (HAT). Automatic thoughts, as described by Beck, are the characteristic self-statements or images that flash into individuals' minds in reaction to external and internal triggers. They are not controlled, occur as reflexes, and are the products of assumptions, attitudes, and beliefs regarding ourselves and the world that derive from our past experiences. The HAT consists of 30 items, rated on five-point likert scales (1, not at all; 5, all the time), reflecting hostile thoughts that involve physical aggression, derogation, and revenge toward other people. It has been shown to be psychometrically valid and reliable.

Self-report measures are easy to administer and score, but they are not immune to faking if respondents desire to present themselves favorably. Hostility is generally considered to be socially undesirable, and low scores on hostility scales may be the result of faking good. This methodological issue may be resolved using a variety of scales that have been shown to be psychometrically sound, as well as using behavioral and observer ratings.

Methods of Anger Control

The majority of anger treatment outcome studies have used a cognitive-behavioral approach, involving methods aimed at reducing the occurrence of anger, regulating levels of arousal when angry, developing new skills in the face of provocation, and preventing conflictual situations from escalating. These treatment approaches have variously been labeled as stress inoculation training, based on Meichenbaum's methods for modifying self-statements in anxious patients; inductive social skills training, based on the principles of cognitive therapy; or cognitive relaxation coping skills. All these approaches have a lot in common, putting more or less emphasis on different aspects of cognitive-behavior therapy (CBT). Treatment can be applied individually or in groups.

Engaging the patient in treatment is of primary importance because patients with more anger are typically suspicious, avoidant, and provoking. Patients need to develop trust in the therapist, learn

to differentiate between different types and levels of affect by self-observation and self-monitoring, recognize the costs of uncontrolled anger, and thus become motivated to change. Following this preparatory phase of treatment, various cognitive and behavioral methods are used. Cognitive restructuring aims to help patients modify their appraisals of threat and provocation, develop appropriate problem-solving thinking, adopt coping self-statements, and modify their basic attitudes regarding themselves and others.

To reduce arousal, methods such as controlled breathing, muscle relaxation, using relaxation imagery, and engaging in alternative tension-reducing activities are taught. Clients need to develop coping skills other than violent expressions of anger. These may include strategic withdrawal, negotiation, appropriate verbal and nonverbal communication, and respectful assertiveness.

In a recent meta-analytic study, Beck and Fernandez analyzed 50 studies incorporating 1640 subjects, published from 1970 to 1995, that used CBT. They reported a grand mean weighted effect size of 0.70, indicating that the average CBT client did better than 76% of untreated subjects in terms of anger reduction. This effect was statistically significant, robust, and relatively homogeneous across studies.

Conclusion

Anger, as other negative impulses, can be viewed as an affective stress reaction with cognitive and behavioral determinants. It can lead, when uncontrolled, to important catastrophic social, interpersonal, and physical health outcomes. Although ignored for a long time compared to other negative emotions such as depression and anxiety, it has been the subject of increasing interest in psychological and psychotherapeutic research.

See Also the Following Articles

Anger; Cognitive Behavioral Therapy; Hostility; Stress Management and Cardiovascular Disease; Type A Personality, Type B Personality.

Further Reading

- Beck, R. and Fernandez, E. (1998). Cognitive behavioral therapy in the treatment of anger: a meta-analysis. *Cognitive Therapy and Research* 22, 63–74.
- Buss, A. H. (1961). *The psychology of aggression*. New York: John Wiley.
- Cook, W. W. and Medley, D. M. (1954). Proposed hostility and pharisaic virtue scales for the MMPI. *Journal of Applied Psychology* 38, 414–418.
- Deffenbacher, J. L. and Stark, R. S. (1992). Relaxation and cognitive-relaxation treatments of general anger. *Journal of Counseling Psychology* 39, 158–167.
- Deffenbacher, J. L., Story, D. A., Brandon, A. D., et al. (1988). Cognitive and cognitive-relaxation treatments of anger. *Cognitive Therapy and Research* 12, 167–184.
- Deffenbacher, J. L., Thwaites, G. A., Wallace, T. L., et al. (1994). Social skills and cognitive-relaxation approaches to general anger reduction. *Journal of Counseling Psychology* 41, 386–396.
- Friedman, M. and Rosenman, R. H. (1974). *Type A behavior and your heart*. New York: Knopf.
- Hall, P. and Davidson, K. (1996). The misperception of aggression in behaviorally hostile men. *Cognitive Therapy and Research* 20, 377–389.
- Novaco, R. W. (1975). *Anger control*. Lexington, MA: Heath.
- Novaco, R. W. (1979). The cognitive regulation of anger and stress. In: Kendall, P. C. & Hollon, S. D. (eds.) *Cognitive behavioral interventions, theory, research, and procedures*, pp. 241–285. New York: Academic Press.
- Snyder, C. R., Crowson, J. J., Houston, B. K., et al. (1997). Assessing hostile automatic thoughts: development and validation of the HAT scale. *Cognitive Therapy and Research* 21, 477–492.
- Spielberger, C. D. (1988). *State-trait anger expression inventory*. Orlando, FL: Psychological Assessment Resources.

Incest

A P Mannarino

Allegheny General Hospital, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A P Mannarino, volume 2, pp 554–557, © 2000, Elsevier Inc.

Incest: Definition and Descriptive Characteristics

Diagnosis

The Psychological Impact of Incest

Treatment of Incest Victims

Glossary

Incest

A type of sexual abuse in which the abuser is a family member.

Nonoffending parent

The nonabusive parent, who typically has custody of the child.

Perpetrator

The family member who abuses the child.

Posttraumatic stress disorder (PTSD)

A stress disorder that includes three categories of psychiatric symptoms in response to a traumatic life event.

Sexual abuse

Sexual exploitation involving physical contact between a child and another person.

Incest: Definition and Descriptive Characteristics

Sexual abuse can be defined as sexual exploitation involving physical contact between a child and another person. Exploitation implies an inequality of power between the child and the abuser on the basis of age, physical size, and/or the nature of the emotional relationship. There can also be noncontact sexual abuse, which includes exposing a child to pornography, taking photographs or videotaping a child for sexual purposes, or sexual exploitation through the Internet.

Incest is a type of sexual abuse in which the abuser is a family member. The abuser may be a parent, step-parent, grandparent, older sibling, or other close relative. Sociological surveys by Finkelhor and others have suggested that 25–30% of all girls and 10–15% of all boys are sexually abused by the age of 18. The proportion of sexual abuse that is incestuous in nature is not absolutely known, but in clinical studies it seems to account for approximately one-half of all reported cases.

Incest is not limited to sexual intercourse. It also includes oral or anal intercourse, fondling of the breasts or genitals, and sexualized, inappropriate kissing. The empirical literature suggests that approximately 50% of incest cases involve oral, anal, or vaginal intercourse while the other half involve genital fondling. Violence is not typically perpetrated against a child victim as part of an act of incest, although threats of violence are common. A child may be coerced to participate in a sexual experience through the promise of rewards or the threat of punishment. In some families, children who are sexually abused may also be physically abused or witness domestic violence.

Children may be victimized by incest at any age. Despite the public misconception that only girls are victimized, boys may also be subjected to an incestuous relationship. The great majority of perpetrators of incest are male, with statistics suggesting that males account for 90% or more of reported cases. The severity of incest can vary from one episode of genital touching to hundreds of episodes of sexual abuse, including intercourse, which may extend over many years. Although incest is commonly reported across all socioeconomic classes, studies by Finkelhor have demonstrated that poor children are more vulnerable to this type of victimization.

Diagnosis

In the nomenclature of the American Psychiatric Association (*Diagnostic and Statistical Manual*, 4th edn.-TR), sexual abuse or incest is not usually considered to be a primary diagnosis (axis I) but is classified

as a psychosocial stressor (axis IV) that may contribute to the primary diagnosis. (However, in situations in which the sexual abuse of the child is the primary focus of clinical attention, it can be coded on axis I.) Children who are victimized by incest do not manifest a unitary syndrome of traits, characteristics, or symptoms. Although, as a group, incest victims display more emotional and behavioral symptoms than normal children, these problems are typically diverse in nature. Moreover, approximately 30% of all incest victims are asymptomatic. Sexual preoccupation or sexualized behaviors are more common in incest victims than nonvictims, but the latter group may exhibit these problems as well. Overall, the general conclusion is that incest cannot be diagnosed based solely on the symptom presentation.

On medical examination, less than 20% of all incest victims display significant clinical findings. Even in cases in which the victim has been vaginally or anally penetrated, significant physical findings may not be present. Accordingly, the absence of significant physical findings does not mean that incest has not occurred.

Determining that incest has occurred is a complex legal and clinical problem. With very young children, they may engage in sexualized behaviors and only disclose alleged abuse after being questioned by the nonoffending parent. Older children may disclose the alleged victimization to a friend, relative, teacher, or other trusted adult. In the great majority of incest cases, the alleged perpetrator initially denies the victimization. In the absence of significant medical findings or a perpetrator confession, it is ultimately very difficult to prove that incest has occurred.

Interviewing the alleged victim is the most frequent method used to determine whether incest has been perpetrated. The nonoffending parent typically provides background data and information about what the child has previously disclosed. If the alleged perpetrator is a parent, he or she may also be interviewed; however, it is recommended that this be done at a separate time from the alleged victim. Most experts recommend that a comprehensive evaluation, including interviews with both parents, be conducted when incest is disclosed in the context of divorce or custody issues.

Over the past 15 years, many concerns have been generated regarding the procedures used to interview alleged sexual abuse and incest victims. Questions have been raised about children's memory and their ability to recall accurately what allegedly occurred. Laboratory research related to this issue has provided mixed results, with some studies suggesting that children do accurately recall the core features of significant events and other studies indicating that children

can produce inaccurate information, particularly in response to repetitive or coercive-like questions. In this regard, criticism sometimes has been directed at child protective service workers, mental health consultants, and other investigative interviewers for allegedly pressuring children to make abuse disclosures or for using inappropriate interviewing techniques (i.e., leading questions). Unfortunately, with the exception of some highly publicized cases, it is impossible to know how often investigations of sexual abuse and incest have been contaminated by improper interview strategies.

In light of the significant issues related to interview methods, some professional guidelines have been promulgated to address this problem. For example, the American Professional Society on the Abuse of Children has developed guidelines for the evaluation of suspected sexual abuse in children, and the American Academy of Child and Adolescent Psychiatry has established practice parameters for the forensic evaluation of children and adolescents who may have been physically or sexually abused. Although it is impossible to know to what degree investigative interviewers are aware of and follow these guidelines, in the past few years there seem to be fewer concerns about inappropriate interviewing procedures.

The Psychological Impact of Incest

As mentioned earlier, children who have been victimized through an incestuous relationship may display a variety of emotional and/or behavioral difficulties subsequent to disclosure. Despite the diversity of the symptom presentation of this population, a number of studies have reported that 20–50% of sexual abuse and incest victims meet full or partial criteria for PTSD. The symptoms of PTSD fall into three categories: re-experiencing symptoms (i.e., intrusive thoughts, nightmares, and flashbacks), avoidance symptoms (i.e., avoid thoughts or feelings about the trauma, emotional detachment, and inability to recall important aspects of the trauma), and symptoms of increased arousal (i.e., sleep disturbance, hypervigilance, and difficulty concentrating). Other psychiatric sequelae of incest may include clinical depression, other anxiety disorders, low self-esteem, sexually reactive behaviors, and, for adolescents, substance abuse problems.

Recent studies have suggested that the frequency and severity of psychiatric symptoms in sexual abuse and incest victims may be mediated by their attributions and perceptions related to the victimization. Thus, victims who feel different from other children, who blame themselves for the abuse, who feel that others do not believe what they say, or have a decreased amount of interpersonal trust are at higher risk for

more serious symptomatology. In particular, children who feel that they are to blame for the victimization are more vulnerable to experiencing depressive disorders.

Legal proceedings may also have an impact on outcome. There is some evidence that when legal proceedings are delayed, emotional/behavioral symptoms in sexual abuse victims tend to increase. Also, when children have to testify in court and are subjected to a very hostile cross-examination or are required to testify multiple times, psychiatric symptomatology appears to worsen.

The psychological impact of being victimized by an incestuous relationship may be moderated by several factors. A number of major studies have demonstrated that emotional support of the victim from the nonoffending parent, which includes not blaming the child for the victimization, has a highly beneficial effect and is correlated with fewer psychiatric symptoms in the victim. Also, as we would anticipate, many nonoffending parents have highly intense emotional reactions to their child's disclosure of incest. These emotional reactions may include anxiety, guilt, anger, and shame. Recent empirical investigations have shown that, when nonoffending parents are able, over time, to resolve their emotional upset related to their child's victimization, this correlates with a reduction in the frequency and severity of the child's emotional and behavioral symptoms.

Treatment of Incest Victims

The treatment of incest may involve a variety of interventions, including individual and/or group therapy for the child victim, supportive therapy with the nonoffending parent, counseling for the perpetrator, and family therapy. Space does not permit a discussion of perpetrator treatment, except to say that most incest perpetrators initially deny victimizing the child and/or minimize its impact and typically require long-term therapy to address their difficulty in taking responsibility for the abuse, cognitive distortions, and issues related to potential relapse.

With the increased attention to and reporting of sexual abuse over the past two decades in the United States, a proliferation of clinical programs has occurred that focus on victim treatment. Unfortunately, the efficacy of the interventions offered in most of these programs has not been investigated. Nonetheless, there is now strong empirical evidence that Trauma-Focused Cognitive Behavioral Therapy (TF-CBT) is an efficacious treatment for sexual abuse and incest victims. Specifically, TF-CBT has been found to be superior to nondirective supportive therapy in reducing PTSD, depressive symptoms, shame, and general behavioral difficulties in the victimized child and

parental distress and depressive symptoms in the non-offending caretaker. This pattern of results has been replicated in a large multisite study of sexual abuse victims, most of whom had also been subjected to other traumatic events.

There are a number of core components of TF-CBT. These include psychoeducation, stress reduction strategies (e.g., relaxation training), creating a trauma narrative, cognitive processing to address inappropriate attributions (i.e., self-blame), and the development of safety and other coping skills.

Work with the nonoffending parent is an essential component of TF-CBT for sexual abuse and incest victims. Critical interventions with the nonoffending parents include addressing their attributions about the sexual abuse, enhancing their support of the victimized child, and behavioral management strategies.

In some treatment programs for incest families, family therapy is the intervention of choice. This is true even though family treatment has not been widely investigated empirically as a potentially effective therapeutic strategy for incestuous families. Family therapy for incest is based on the belief that family dynamics either contribute to the occurrence of the incestuous situation, help maintain the secrecy of the incestuous relationship, or can be changed to prevent future episodes of victimization. The perpetrator of the abuse may be included in family treatment, especially if there is a possibility of family reunification. However, the inclusion of the perpetrator in family sessions would typically occur only after the perpetrator has experienced extensive individual and/or group treatment and has been willing to accept total responsibility for the victimization.

See Also the Following Articles

Child Sexual Abuse; Familial Patterns of Stress; Sexual Assault.

Further Reading

American Academy of Child and Adolescent Psychiatry (1997). Practice parameters for the forensic evaluation of children and adolescents who may have been physically

- or sexually abused. *Journal of the American Academy of Child and Adolescent Psychiatry* 36, 432–442.
- American Professional Society on the Abuse of Children (1990). *Guidelines for psychosocial evaluation of suspected sexual abuse in young children*. Chicago: Author.
- American Psychiatric Association (2000). *Diagnostic and statistical manual* (4th edn.). Washington, DC: American Psychiatric Press.
- Bruck, M. and Ceci, S. J. (1998). Reliability and credibility of young children's reports: From research to policy and practice. *American Psychologist* 53, 136–151.
- Cohen, J. A., Deblinger, E. and Mannarino, A. P. (2005). Trauma-focused cognitive behavioral therapy for sexually abused children. In: Hibbs, E. & Jensen, P. (eds.) *Psychosocial treatments for child and adolescent disorders: empirically-based strategies for clinical practice*. (2nd edn., pp. 743–765). Washington, DC: American Psychological Association.
- Cohen, J. A., Deblinger, E., Mannarino, et al. (2004). A multisite, randomized controlled trial for sexually abused children with PTSD symptoms. *Journal of the American Academy of Child and Adolescent Psychiatry* 43, 393–402.
- Cohen, J. A. and Mannarino, A. P. (1996). A treatment outcome study for sexually abused preschool children: initial findings. *Journal of the American Academy of Child and Adolescent Psychiatry* 35, 42–50.
- Deblinger, E., Lippman, J. and Steer, R. (1996). Sexually abused children suffering PTSD symptoms: initial treatment outcome findings. *Child Maltreatment* 1, 310–321.
- Finkelhor, D. and Baron, L. (1986). Risk factors for child sexual abuse. *Journal of Interpersonal Violence* 1, 43–71.
- Mannarino, A. P. and Cohen, J. A. (1996). Abuse-related attributions and perceptions, general attributions, and locus of control in sexually abused girls. *Journal of Interpersonal Violence* 11, 162–180.
- Runyan, D. K., Hunter, W. M. and Everson, M. D. (1992). Impact of legal intervention on sexually abused children. *Journal of Pediatrics* 113, 647–653.
- Saywitz, K. J., Goodman, G. S. and Lyon, T. D. (2002). Interviewing children in and out of court: current research and practice implications. In: Myers, J. E. B., Berliner, L., Briere, J., Hendrix, C. T., Tenny, C. & Reid, T. A. (eds.) *The APSAC handbook on child maltreatment* (2nd edn., pp. 349–377). Thousand Oaks, CA: Sage.
- Saywitz, K. J., Mannarino, A. P., Berliner, L., et al. (2000). Treatment for sexually abused children and adolescents. *American Psychologist* 55, 1040–1049.

Income Levels and Stress

S V Subramanian and I Kawachi

Harvard School of Public Health, Boston, MS, USA

© 2007 Elsevier Inc. All rights reserved.

Absolute Income and Health
Relative Income and Health
Income Inequality and Health
Income and Health: The Role of Social Comparisons

Glossary

<i>Absolute income</i>	The measure of the actual income of the individual; linked to the absolute income hypothesis.
<i>Absolute income hypothesis</i>	The idea that health of individuals depends on their own (and only their own) level of income.
<i>Comparative group</i>	The social group to which an individual compares him- or herself.
<i>Income inequality</i>	The measure of the extent of income dispersion in a society or a community; measures such as Gini coefficient are summary measures that quantify the extent of income dispersion in a society.
<i>Membership group</i>	The social group of which an individual feels a part.
<i>Neomaterialist theory</i>	A theory that emphasizes the importance of access to tangible material conditions, both elementary (e.g., food, shelter, and access to services and amenities) and nonelementary (e.g., car and home ownership and access to telephones and the Internet).
<i>Normative group</i>	The social group from which a person takes his or her standards or norms; this could be either the membership or the comparative group.
<i>Psychosocial effects</i>	The direct or indirect effects of stress stemming from either being lower on the socioeconomic hierarchy or living in conditions of relative socioeconomic disadvantage, which facilitates constant social comparisons.
<i>Relative deprivation</i>	The condition in which individuals or groups subjectively perceive themselves to be unfairly disadvantaged compared to others who they perceive as having similar attributes and deserving similar rewards (typically their reference groups or groups to whom they typically compare themselves).
<i>Relative income</i>	The measure of the income of the individual in relation to the incomes of

others in the society or group in which the individual resides. It is linked to the relative income hypothesis.

Relative income hypothesis

The idea that health depends not just on an individual's own level of income but also on the incomes of others in society.

Social comparison

The metric that results from an individual's self-comparison of some other individual or group of individuals on certain dimension, such as income, education, and material possessions.

Differential exposure to stressors is one of the links through which income is associated with health. Income can be conceptualized in absolute or relative terms, with each having distinct mechanisms through which income is hypothesized to influence stressful experiences, thereby affecting health.

Absolute Income and Health

The absolute income hypothesis posits that:

$$h_i = f(y_i), \quad f'_i > 0, \quad f''_i < 0 \quad (1)$$

where h_i is an individual's level of health, and y_i is that individual's own level of income. The relationship between individual income and individual health is concave; that is, every additional dollar is associated with diminishing returns to health. The relationship of individual income to health outcomes is often described as a gradient – at each level of income, individuals experience better health than those immediately below them. On the other hand, to describe the association between income and health as a gradient is possibly misleading because the relationship has been shown to be nonlinear; that is, there are diminishing returns to health improvement with additional rises in income. The concave relationship between income and health has been corroborated in numerous empirical demonstrations, in affluent countries as well as in poor countries. The absolute income hypothesis has been closely associated with the neomaterialist interpretation of the relationship between income and health, which posits that income matters for health because it confers on individuals the ability to purchase goods and services (e.g., health insurance, housing, and nutritious foods) that are necessary for maintaining health, even though the absolute income hypothesis could equally well have a psychosocial effect.

We can also posit a societal or community-level analog of the absolute income hypothesis, such that:

$$h_{ij} = f(Y_j, y_{ij}) \quad (2)$$

where the health of individual i in community j is now hypothesized to be a function of not just his or her own income (y_{ij}) but also of the average income level of the community (Y_j) in which the individual resides. Indeed, a neomaterialist interpretation of the absolute income hypothesis need not be restricted to the availability of income at the individual level; it can also reflect the presence or absence of collective infrastructure at a population or community level. Deprived communities (i.e., communities with high levels of absolute poverty) are less likely to have services and amenities (e.g., health clinics and supermarkets) that are essential for protecting and promoting the health of their residents. Even if an individual resident is not poor, his or her health may be detrimentally affected by living in such a community. A growing number of multilevel studies have now demonstrated that community poverty is linked to higher mortality, higher morbidity, and a more adverse pattern of health behaviors, even after taking into account individual incomes.

Relative Income and Health

In contrast to the absolute income hypothesis, the relative income hypothesis posits that:

$$h_i = f(y_i - y_r) \quad (3)$$

where health of an individual i is a function of the term $(y_i - y_r)$, which denotes the relative gap between an individual's income y_i and the income of some reference population y_r . The reference population could be the income of coworkers, neighbors, or the national population. The concept of relative deprivation was first rigorously formulated by Runciman. According to Runciman, "we can roughly say that A is relatively deprived of X when (i) he does not have X, (ii) he sees some other person or persons, which may include himself at some previous or expected time, as having X (whether or not this is or will be in fact the case), (iii) he wants X, and (iv) he sees it as feasible that he should have X" (Runciman, 1966: 10).

Runciman further conceptualizes the concept of relative deprivation in terms of magnitude, frequency and degree. "The magnitude of the relative deprivation is the extent of the difference between the desired situation and that of the person desiring it (as he sees it). The frequency of a relative deprivation is the proportion of a group who feel it. The degree of a

relative deprivation is the intensity with which it is felt" p.10.

Runciman went on to elaborate three types of reference groups. The first is the membership group, of which an individual feels a part; the second is the normative group, "the group from which a person takes his standards" (Runciman, 1966: 12); and the third is the comparative group, which is the group to which an individual compares him- or herself. Significantly, these different sources of relative comparisons may well coincide, as in Runciman's example of members of the middle class who aspire to be aristocracy and who try to emulate the behavior of the upper class (the normative reference) as well as to attain their status (the comparative reference). Building on these sources of relative comparisons, he further proposed two forms of relative deprivation, depending on the source: egoistic deprivation (relative deprivation of an individual with respect to his or her own group) and fraternalistic deprivation (relative deprivation of an individual's own group compared to other groups in society).

Note that relative deprivation theory, as the phrase suggests, presumes that individuals typically make relative comparisons of themselves to those who are better off than they are and not to those who are worse off. The latter comparison is referred to as relative gratification or relative satisfaction. Indeed, relative deprivation measures happen to be correlated with relative satisfaction.

Although the theory of relative deprivation is quite sophisticated, the empirical identification of reference groups has proved tricky. Empirically, the relative income hypothesis has been operationalized using variations of [equation \(3\)](#), with an important operationalization provided by Yitzhaki, following Runciman's theory. Thus, for a person i with income y_i who is part of a reference group with N people, Yitzhaki's index is given as:

$$RD_i = (1/N) \sum_j (y_j - y_i), \quad \forall y_j > y_i \quad (4)$$

where the amount of relative deprivation (RD) for individual i is the sum of the incomes of j individuals who have incomes higher than that individual. The summation in [equation \(4\)](#), $\sum_j (y_j - y_i)$, is divided by the number of people in the reference group, N , making the measure invariant to size. If income is the object of relative comparison, then we can assess relative income deprivations by groups, based on, for instance, the state of residence, age group, race, educational attainment, gender, or occupation.

The British Whitehall Study of civil servants is frequently cited as evidence showing the health effects of relative deprivation (even though the study did

not include any specific or explicit measures of relative deprivation). In that study, relatively lower-ranked civil servants were shown to have three times the mortality rate as the highest ranked civil servants, despite having access to the same National Health Service. Moreover, none of the subjects in the civil service cohort was deprived in an absolute sense; that is, they all had adequate nutrition, housing, and access to other material amenities. The mortality differential between civil service grades was thus attributed – at least in part – to the adverse psychosocial consequences of relative deprivation. Indeed, systematic tests suggest that people do take into account relative income position when making real-life decisions such as choosing between two earning schemes. For instance, Frank and Sunstein (2001) hypothesized two possible scenarios. In scenario A, the individual earns \$110,000 per year and others earn \$200,000; in scenario B, the individual earns \$100,000 per year (less than in world A), but others earn only \$85,000. Contrary to traditional economic perspective (wherein world A should be preferable because it offers higher absolute consumption), a substantial number of respondents opted for world B.

As we mentioned earlier, the absolute income hypothesis tends to be interpreted in materialist terms, whereas the relative income hypothesis tends to be interpreted in psychosocial terms. This dichotomy is unwarranted. For instance, the absolute income hypothesis is equally compatible with the psychosocial interpretation. That is, if lower income is associated with more stress (e.g., inability to pay the bills and to plan for the future), then an individual could end up with an income–health gradient even if no social comparison is taking place.

Analogously, the relative deprivation hypothesis could also be interpreted both ways (neomaterial as well as psychosocial). That is, an individual's social comparison to a reference group could be stressful (the psychosocial interpretation) or he or she may be unable to function in society because of limited access to material commodities (e.g., car ownership, Internet access, or air conditioning). For example, rising community living standards are often associated with an enlargement of the range of consumer goods that are necessary for an individual to function as a member of that community. Many consumer goods, such as central heating/air conditioning, telephone, and access to the Internet, have followed this trajectory, starting out as luxury goods and eventually ending up as necessities. Even if an individual is not deprived in any absolute sense (i.e., has adequate nutrition and housing), relative deprivation (compared to the standard basket of commodities that everyone else has access to) could adversely affect his or her health. In this sense, relative deprivation may affect health

because it reflects the ability of affected individuals to access material goods and services that in turn promote well-being. This underscores the futility of distinguishing material from psychosocial effects of income – any effect of income on health could be simultaneously operating through material as well as psychosocial pathways.

Income Inequality and Health

The community-level analog of relative deprivation leads us to the idea of income inequality, which can be stated as:

$$h_{ij} = f(I_j, y_{ij}) \quad (5)$$

where I_j refers to summary measure of income distribution (e.g., the Gini coefficient) for community j in which the individual resides. Income inequality and relative deprivation are conceptually related because, as income inequality rises, those at the bottom end of the distribution become more relatively deprived. Various measures are available to quantify the extent of income inequality in a given community or society; of these, the Gini coefficient is frequently used. Algebraically, the Gini coefficient is defined as one-half of the arithmetic average of the absolute differences between all pairs of incomes in a population, the total then being normalized on the mean income. If incomes in a population are distributed completely equally, the Gini value is 0; and if one person has all the income (the condition of maximum inequality), the Gini is 1.0. The Gini coefficient can also be illustrated through the use of Lorenz curve (Figure 1). On the horizontal axis, the

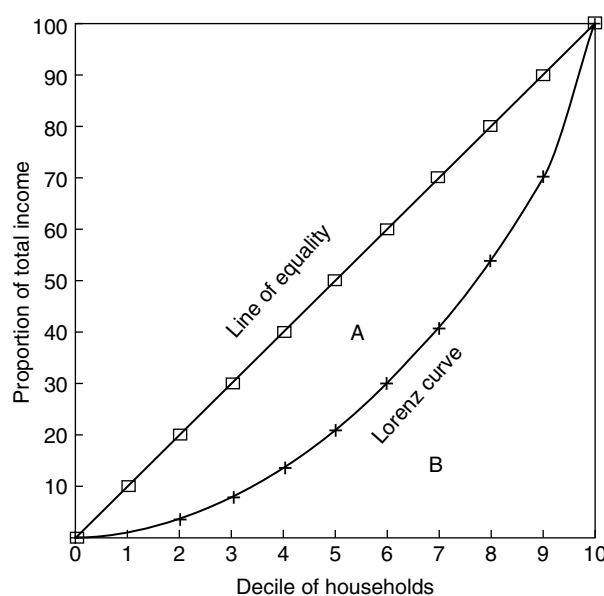


Figure 1 Lorenz curve.

population (in this case, households) is sorted and ranked according to deciles of income, from the low-est-decile group to the top-decile group. The vertical axis then plots the proportion of the aggregate income within that community accruing to each group. Under conditions of perfect equality in the distribution of income ($Gini = 0$), each decile group would account for exactly 10 percent of the aggregate income and the Lorenz curve would follow the 45° line of equality. In reality, the Lorenz curve falls below the 45° line of equality because the bottom groups in the income distribution earn considerably less than their equal shares. In [Figure 1](#), for instance, the bottom one-half of households accounts for just 20% of the aggregate income. The degree to which the Lorenz curve departs from the 45° line of equality is a measure of income inequality. As it turns out, the Gini coefficient is the ratio of the area between the Lorenz curve and the 45° line of equality (A) and the area of the triangle below the 45° line of equality (A + B).

There is some debate as to whether income inequality is a valid community analog of relative deprivation at the individual level. Some argue that Gini coefficient is the measure of relative deprivation. For instance, the relative deprivation curve that represents the gaps between an individual's income and the incomes of all individuals richer than him or her (as a proportion of mean income) turns out to be the Gini coefficient. Others have presented a generalization of Gini, the s-Ginis, that can be interpreted as indices of relative deprivation. Indeed, it has been stated that "the Gini coefficient may be interpreted as a transformation of an index of 'envy' evaluated through the comparison of the income of an individual and that of all other individuals richer than him/her" (Atkinson and Bourguignon, 2000: 43).

As mentioned earlier, the relationship between individual income and health status is concave, such that each additional dollar of income raises individual health by a decreasing amount. The concave relationship between income and health has important implications for the aggregate-level relationship between income distribution and average health achievement, as noted by Rodgers. As illustrated in [Figure 2](#), in a hypothetical society consisting of just two individuals – a rich one (with income x_4) and a poor one (with income x_1) – transferring a given amount of money (amount $x_4 - x_3$) from the rich person to the poor person will result in an improvement in average health (from y_1 to y_2) because the improvement in the health of the poor person more than offsets the loss in health of the rich person. Indeed, it is possible that, by transferring incomes from the relatively flat part of the

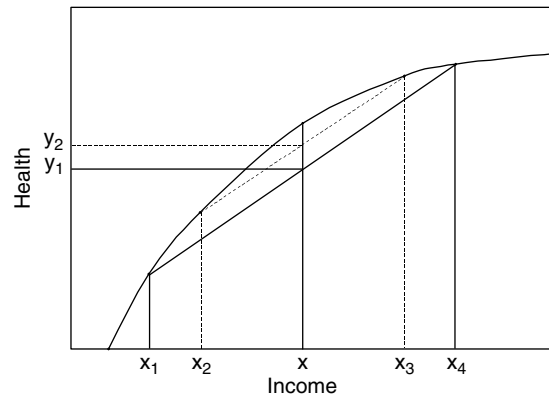


Figure 2 The individual-level relationship between income and health.

income–health curve, there may be no loss in health for the wealthy person.

Consequently, researchers have posited that an aggregate relationship between the average health status of a society and the level of income inequality in a society could be observed if the individual-level relationship between income and health (within the society) is concave. That is, the aggregate relationship between income inequality and health may be observed simply due to the underlying concave functional form of the individual income–health relationship and assuming an x amount of transfer of money from rich to the poor. Indeed, such a transfer also implies a reduction in the income inequality level in that particularly society, and as such, the society with the narrower distribution of income will have a better average health status, all other things being equal. It is worth emphasizing that, if the relationship between income and health at the individual level is linear (not concave), then a transfer of income from the rich to the poor will reduce the level of income inequality but will not lead to improvements in the average health status of that society.

Occasionally, this expected relationship between income distribution and the average health status of a population (which is a direct function of the concave relationship between individual income and health) has been described as a statistical artifact of the concave relationship between individual income and health. The use of the term artifact is misleading here because it suggests that the potential for improving the health of the poor through income redistribution is a statistical illusion. Indeed, there is nothing artifactual about improving the health of the poor (and hence, average population health) through income and wealth redistribution – indeed the success of much philanthropy (e.g., donating money to provide vaccines to the world's poor) rests on the validity of this assumption.

In addition to the concavity effect just described, researchers have posited an additional contextual effect of income inequality on health. This is the hypothesis that the distribution of income in society, over and above individual incomes as well as societal average income, matters for population health such that individuals (regardless of their individual incomes) tend to have worse health in societies that are more unequal. Thus, income inequality *per se* may be damaging to public's health by causing a *downward shift* in the income–health curve. Thus, the income inequality hypothesis is viewed as an independent contextual effect akin to a social pollution effect.

Income and Health: The Role of Social Comparisons

Wilkinson argued that income inequality is harmful to health because it leads to invidious social comparisons among individuals in society and that these comparisons, in turn, may act as a stressor in the form of feelings of shame, anxiety, and social exclusion and, as such, impact physiological systems. The economist Juliet Schor described how widening inequalities have given rise to a culture of upward social comparisons (and the attendant frustration of aspirations). Thus, a mechanism linking income inequality to health involves psychosocial pathways that have adverse physiological consequences. Models of the direct effects of stress on physiological systems include allostatic load, which describes the wear and tear on the organism caused by exposure to daily adverse life circumstances. Stress may also affect health indirectly by leading to a more adverse profile of behaviors, such as smoking and excess drinking. This line of explanation links the community income inequality and individual health through individual relative deprivation (including those related to income). Researchers interested in adult health and mortality have argued that inequality worsens adult health through social comparison. They have argued that income inequality affects health because ranking low in the social hierarchy produces negative emotions such as shame and distrust that lead to worse health via neuroendocrine mechanisms and stress-induced behaviors such as smoking, excessive drinking, and drug use. Although the human body is remarkable in adapting to occasional bouts of stress, they have argued that cumulative exposures to stressors (such as chronic poverty, low socioeconomic status, and constant social comparisons) eventually result in strain on regulatory systems.

Although we have discussed here the role of social comparisons as a source of stress, there are other

mechanisms that link income and health that can also be explained through stress-based explanations. For instance, a chronic exposure to unpredictable social and economic circumstances (associated with low income) can also be a source of stress in addition to inhibiting access to important material goods necessary to sustain desirable levels of well-being. The challenge for researchers on income and health is to elucidate the specific pathways through which the different conceptualizations of income may impact on health through differential types of stressors.

Acknowledgments

S. V. Subramanian is also supported by a National Institutes of Health/National Heart, Lung, Blood Institute Career Development Award (1 K25 HL081275).

See Also the Following Articles

Economic Factors and Stress; Health and Socioeconomic Status; Social Status and Stress.

Further Reading

- Atkinson, A. B. and Bourguignon, F. (2000). Introduction: income distribution and economics. In: Atkinson, A. B. & Bourguignon, F. (ed.) *Handbook of income distribution*, (vol. 1), pp. 1–58. Amsterdam: Elsevier Science.
- Deaton, A. (2003). Health, inequality, and economic development. *Journal of Economic Literature* 41, 113–158.
- Duclos, J.-Y. (2000). Gini indices and the redistribution of income. *International Tax and Public Finance* 7, 141–162.
- Frank, R.H. and Sunstein, C.R. (2001). Cost-benefit analysis and relative position. *University of Chicago Law Review*, 68: 323–374.
- Kakwani, N. (1984). The relative deprivation curve and its applications. *Journal of Business and Economic Statistics* 2, 384–394.
- Kawachi, I. (2000). Income inequality and health. In: Berkman, L. F. & Kawachi, I. (ed.) *Social epidemiology*, pp. 76–94. New York: Oxford University Press.
- Kawachi, I. and Kennedy, B. P. (2003). *The health of nations*. New York: The New Press.
- Marmot, M. (2002). The influence of income on health: views of an epidemiologist. *Health Affairs* 21(2), 31–46.
- McEwen, B. S. and Seeman, T. S. (1999). Protective and damaging effects of mediators of stress: elaborating and testing the concepts of allostasis and allostatic load. *Annals of the New York Academy of Sciences* 896, 30–47.
- Rodgers, G. B. (1979). Income and inequality as determinants of mortality: an international cross-section analysis. *Population Studies* 33, 343–351.

- Runciman, W. G. (1966). *Relative deprivation and social justice*. London: Routledge.
- Schor, J.B. (1998). *The overspent American*. New York: Basic Books.
- Subramanian, S.V., Belli, P., and Kawachi, I. (2002). The macroeconomic determinants of health. *Annual Review of Public Health* 23: 287–302.
- Subramanian, S.V., and Kawachi, I. (2004). Income inequality and health: what have we learned so far. *Epidemiologic Reviews* 26: 78–91.
- Subramanian, S.V., and Kawachi, I. (2006). Whose health is affected by income inequality? A multilevel interaction analysis of contemporaneous and lagged affects of state income inequality on individual self-rated health in the United States. *Health and Place* 12: 141–156.
- Subramanian, S.V., Kawachi I. (2006). Being well doing well: on the importance of income for health. *International Journal of Social Welfare* 15, S1: S13–S22.
- Wagstaff, A. and van Doorslaer, E. (2000). Income inequality and health: what does the literature tell us? *Annual Review of Public Health* 21, 543–567.
- Wilkinson, R. G. (1996). *Unhealthy societies – the afflictions of inequality*. London: Routledge.
- Wilkinson, R.G. (2005). *The impact of inequality*. New York: New Press.
- Yitzhaki, S. (1979). Relative deprivation and the Gini coefficient. *Quarterly Journal of Economics* 93(2), 321–324.

Indigenous Societies

W W Dressler

University of Alabama, Tuscaloosa, AL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by W A Dessler, volume 2, pp 558–564, © 2000, Elsevier Inc.

Introduction

Stress in Unacculturated Indigenous Societies

Stress and Acculturation in Indigenous Societies

Stress and Migration

Conclusion

Glossary

<i>Acculturation</i>	The process by which a society is culturally influenced by another society.
<i>Culture-bound syndromes</i>	Local idioms for the expression of distress, consisting of socially recognized illness categories that have no acknowledged counterpart in Western biomedicine.
<i>Cultural consonance</i>	The degree to which an individual approximates in his or her own behavior the prototypes for behavior encoded in shared cultural models.
<i>Extended kin support systems</i>	Systems in which help or assistance is most appropriately sought from individuals who are related by common ancestry or marriage.
<i>Indigenous societies</i>	Societies that are native to a particular locale, in which people speak a distinctive language and have a distinctive social heritage and which are often technologically less complex. Usually, the way of life of the people is on the periphery of the

major world industrial societies. Indigenous societies can also be referred to as traditional societies.

Modernization

A special case of acculturation in which an indigenous society is undergoing economic change or development.

Status

incongruence

A discrepancy between an individual's aspirations for social status or prestige and his or her resources for actually achieving those aspirations.

Introduction

An indigenous or traditional society is characterized as native to a specific region, with a distinctive language and way of life. Also, the way of life of the people is peripheral to global capitalist market systems. This does not mean that such a society is unaffected by global market systems; in fact, one major source of stress in indigenous societies is the impact of economic change emanating from those larger systems. Nor does this mean that an indigenous society cannot be embedded within a modern, industrial society (e.g., Native Americans). Rather, in an indigenous society, culture and related systems of social organization are structured more in terms of local context and local systems of meaning, and less in terms of the middle-class values of industrial society. Understanding stress and its effects in indigenous societies requires an examination of both social arrangements that generate stresses within the indigenous social structure and the way in which traditional culture and social structure interact with outside influences to generate stresses.

Stress in Unacculturated Indigenous Societies

Anthropologists conventionally describe the relative degree of external influence on a society or community along a continuum of acculturation. A society that is relatively unacculturated or traditional has been minimally influenced by processes of modernization. Usually this means that households tend to practice a mix of economic pursuits for subsistence (i.e., raising food directly for consumption within the household) and for exchange in local markets. With respect to material lifestyles, although there is some access to imported consumer goods, these often are primarily related to subsistence activities (e.g., agricultural implements, outboard motors). Most goods consumed by a household are produced locally. What wage labor exists is usually within the community, and there is little formal education. Social relationships tend to be dominated by kinship. Systems of kinship can range from large groups formed around descent from a common ancestor (or unilineal descent groups) to somewhat more loosely structured kindreds that are like large ego-centered social networks. Finally, belief systems reflect local meanings and understandings, even when there have been modernizing influences (e.g., missionaries).

Patterns of morbidity and mortality within traditional societies provide one clue to patterns of stress in those societies. Generally speaking, rates of high blood pressure, coronary artery disease, stroke, and cancer tend to be very low. In traditional societies, patterns of morbidity and mortality tend to be dominated by infectious and parasitic disease, especially in childhood, and by trauma in adulthood. Although any generalizations must be tempered by reference to local ecological conditions, in many traditional societies life expectancy beyond 5 years of age is comparable to life expectancy in industrial societies, and in the aged in these societies there is little evidence of the kind of pathologies (e.g., atherosclerosis) associated with aging in industrial societies.

What has been most illuminating in the study of stress in unacculturated societies is the study of culture-bound syndromes. Culture-bound syndromes are local idioms of distress. They can be thought of as culturally appropriate ways of experiencing and expressing distress arising from stressful social relationships. Some attempts have been made to equate the culture-bound syndromes with Western psychiatric diagnoses, but recent evidence indicates that there is not a direct correspondence between culture-bound syndromes and biomedical psychiatric diagnoses; it is probably more useful to think of

culture-bound syndromes and Western psychiatric diagnoses as comorbid.

A classic example of a culture-bound syndrome is *susto* in Latino societies of Central and South America. The individual suffering from *susto* experiences a loss of energy, difficulty in maintaining customary activities, frequent spells of crying, and diffuse somatic symptoms such as loss of appetite and sleep disturbance. As the name in Spanish implies, *susto* is attributed to a sudden fright (e.g., seeing a snake), at which time the soul of the individual leaves the body and wanders freely.

Research has shown that the distribution of *susto* is socially patterned. The prevalence is higher in females, tends to increase with age, and tends to be higher in relatively poorer communities. Furthermore, the greatest risk of *susto* has been found among persons experiencing difficulty in enacting common social role expectations. This usually arises from a lack of specific kinds of social resources (e.g., not having a large kinship network) that can be called on in carrying out expected role obligations (such as contributions to community work groups). Similar findings have been obtained in research on other culture-bound syndromes such as *ataques de nervios* in Puerto Rico, *debilidad* in the Andean highlands, heart illness in Iran, and *nervos* in Brazil. Culture-bound syndromes occur when individuals are low in cultural consonance; that is, they are unable to approximate in their own behaviors the prototypes for behaviors that are encoded in widely shared cultural models.

The experience of a culture-bound syndrome is important in two respects. First, it makes meaningful and intelligible the experience of social stress both to the person suffering the stress and to his or her social network. Second, in many societies, there are cultural practices, including healing rituals and participation in religious organizations, that deal directly with the syndrome and the underlying difficulties in social relationships. The aim is to mend the tear in the fabric of social relationships and hence end the individual's suffering.

The existence of these beliefs and practices that are helpful in ameliorating cultural stresses may in part account for patterns of morbidity and mortality in indigenous societies. As noted previously, the diseases conventionally associated with stress in industrial societies are relatively less important in indigenous societies. Also, in most cases an increase of blood pressure with age is not observed in indigenous societies. Other patterns of disease distribution that are taken for granted in industrial societies are not

observed in traditional societies. One of the more striking of these is the association of blood pressure with African descent ethnicity. While it is assumed that persons of African descent have higher blood pressures, in fact this is true primarily for those of African descent in the Western hemisphere, and more specifically in societies in which Africans had formerly been enslaved. So, for example, when communities of African descent are compared, communities in Africa have the lowest average blood pressures, communities in the West Indies have intermediate average blood pressures, and African American and African Brazilian communities have the highest average blood pressures. Recent findings show also that *quilombo* or refugee communities descended from escaped slaves in isolated areas of Brazil have low average blood pressures and exhibit little increase of blood pressure with age.

These patterns suggest two things. First, there may be a relatively higher level of social integration in indigenous societies, along with practices that help to moderate the impact of social stressors, that account for the lower prevalence of conditions and diseases associated with stress in industrial societies. Second, the process of social change leading to the modern industrial state may itself generate profound social stresses that contribute to the distinctive pattern of morbidity and mortality in those societies.

At the same time, it is important not to romanticize life in indigenous societies, in the sense of overemphasizing social integration and cohesion, because stresses will be generated within any system of social relationships. What is important to specify across different cultural contexts is the process by which social stresses are generated and the ability of indigenous support systems to deal with those stresses.

Promising results in this regard are emerging from research on hormones and neurotransmitters associated with the stress process. Newer techniques of data collection under difficult field conditions, along with techniques for the analysis of those data, are beginning to show how variation in social behavior within traditional societies is associated with inter- and intra-individual variation in stress hormones, which in turn is associated with acute illness. For example, research in a peasant village in the West Indies has shown that men who are perceived by their peers to emulate the ideals of manhood in this community have lower circulating levels of cortisol. Similarly, children growing up in families that are closer to the cultural ideal of the family have lower circulating cortisol levels and experience fewer acute illnesses. This research, coupled with research on local idioms of distress, suggests that stresses in

unacculturated societies are deeply embedded in the system of social relationships that organizes everyday life.

Stress and Acculturation in Indigenous Societies

Societies that are undergoing acculturation are those that are being influenced by other social and cultural systems. In the study of stress and disease, the effect of modern industrial societies on local sociocultural systems has been of particular interest. There is considerable imbalance in this type of acculturation, because of the unequal power and influence that modern industrial states exert on traditional societies. The terms modernization and development have been used to describe this kind of influence.

Modernization in traditional societies was initiated by colonial expansion and has been particularly prominent since World War II and related processes of globalization. This influence has not been inadvertent. The aim has been to take advantage of both physical and social resources in developing societies.

These changes have had large effects on local social systems, including a transition from subsistence occupations to wage labor occupations, the replacement of indigenous languages by European languages, increased urbanization, increased emphasis on formal education, decreased emphasis on traditional social relationships, especially kinship, and substantial changes in indigenous belief systems. Everyday life can change at a rapid pace in modernizing contexts, the result being a stressful lack of consonance between traditional culture and the demands of modern life.

Specific and general aspects of this modernization process have been found to be associated with increasing rates of chronic diseases. Urbanization has been found to be associated with increasing rates of hypertension, independent of changes in diet and physical activity. Some investigators have used summary measures of acculturation both for individuals and for communities. When communities are ranked along a continuum of traditional, intermediate, and modern (depending on aggregate characteristics of the population by the variables noted previously), rates of hypertension, obesity, diabetes, and coronary artery disease consistently increase in communities with higher levels of acculturation or modernization. Also, daily circulating levels of hormones such as cortisol appear to increase in association with acculturative stress.

The results have been somewhat less consistent using measures of acculturation operationalized at the level of the individual. The pattern can still be

observed in many studies, but the strength and the consistency of the associations are not as great. This has led some researchers to speculate that the linear model of stress and acculturation is not specified well enough to describe the process at the individual level.

Because the general model of stress and acculturation has not worked very well at the level of the individual, researchers have adapted the stress model, as it has been developed for European and American populations, and applied it to communities experiencing change and development. A major challenge in this research has been to identify factors that generalize across different cultural contexts and to distinguish those from factors that are culturally specific. There is emerging evidence to suggest that at least one set of social stressors generalizes across modernizing societies, primarily because these stressors are generated by the modernization process itself. In most societies undergoing modernization, there is an increased availability of Western consumer goods. Frequently the ownership of these goods becomes highly valued as symbols of status or prestige, often supplanting traditional indicators of higher status. This by itself contributes to the climate of change in a modernizing community. In addition, however, aspirations for the lifestyles of the Western middle class can quickly outstrip the ability of a developing economy to provide the kinds of jobs and salaries necessary to maintain such a lifestyle. Therefore, a kind of status incongruence can occur, in which the desire to attain and maintain a Western middle-class lifestyle exceeds an individual's economic resources for such a lifestyle. This kind of status incongruence has been found to be related to psychiatric symptoms, high blood pressure, elevated serum lipids, the risk of diabetes, and immunological status in developing societies in Latin America, the Caribbean, and Polynesia.

Resources for coping with stressors have been found to be more culturally variable. Social support systems are a good case in point. Social support can be defined as the emotional and practical assistance an individual believes is available to him or her during times of felt need; the social network in which this assistance is available is the social support system. In research in Europe and North America, generally speaking, emphasis has been placed on the nature of the assistance or social support transactions, rather than on who might provide that assistance. There is a growing body of cross-cultural research indicating that, in many societies, who provides the assistance is critical in determining the relationship between social support and health. This is probably a reflection of a

continuing importance of kinship in defining who is and who is not an appropriate individual with whom to enter into a social relationship.

For example, in Latin American societies people have traditionally lived in large extended families organized around a father and his married sons (known as a patrilineal extended family). In addition to these extended family relationships, there is a social practice known as *compadrazgo*, through which individuals, especially men, establish formal, kinship-like relationships with unrelated persons (known as fictive kinship). The term *compadrazgo* literally means co-parenthood, and this carries the expectation of mutual support. These ties of fictive kinship are used to establish economically and politically important alliances. Research has shown that men who perceive greater amounts of support from both their extended and their fictive kin have lower blood pressures. Women are expected to restrict themselves to the household and domestic duties. Not surprisingly, with respect to health status, women benefit primarily from support available within the household. These studies show that the definition and effects of social supports are closely related to cultural and social structural factors and can only be understood within that context.

Other forms of stress resistance show similar differences cross-culturally, although the evidence is not quite as consistent as with social support. For example, in European and North American studies, a direct action coping style has been found to be helpful in moderating the effects of stressful events or circumstances. In this style of coping, attempting to directly confront and alter stressful circumstances contributes to better health status. In some research in traditional societies, this same relationship has been observed. In others, the opposite effect has been observed, i.e., a direct action coping style actually exacerbates distress and poor outcomes. In the specific case of this coping style, this probably has to do with the actual resources available to individuals and families in coping with stressors. When social and economic resources are meager at best, the belief that an individual can truly change the circumstances that are often thrust upon him- or herself may, in the long run, be deleterious.

A major source of stress in societies undergoing modernization is the increase in socioeconomic inequity that accompanies modernization. Generally speaking, the distribution of wealth becomes more unequal, resulting in marked social stratification, whereas previously such stratification was, if not absent, at least muted. Recent studies point to the effects of such stratification on cultural consonance as a

potent source of stress. The capacity of an individual to achieve higher cultural consonance is severely compromised by lower socioeconomic status. Lower cultural consonance has been found to be associated with higher blood pressure and psychological distress and to mediate the effects of socioeconomic status in contexts of modernization. The inability to act on these widely shared cultural models is a potent source of stress in developing societies.

Stress and Migration

The other way in which social and cultural change can influence indigenous communities is through migration. This century has seen remarkable movements of people from traditional societies to North America and Western Europe, along with internal migration within developing societies that takes migrants into cosmopolitan urban centers in their own societies. Rarely can migration be considered an individual matter. Rather, it is much more common for entire communities of migrants to become established in their host country. This usually occurs because migration follows patterns established through social networks, especially kinship networks. Individuals and households take advantage of kin-based social support systems established in host countries in order to establish themselves there.

Research suggests that a similar pattern of stress and response develops in cases of migration like that observed in developing societies. That is, migrants come to the new setting with aspirations for a new life, aspirations that are reinforced in host countries through the depiction of middle-class lifestyles in advertising and other media forms. At the same time, migrants typically occupy the lowest levels of socioeconomic status in their host countries. Therefore, the ability of the migrant to amass the economic resources necessary to achieve a middle-class lifestyle is severely compromised. This status incongruence again has been found to be associated with chronic disease risk factors such as blood pressure and glucose levels.

Patterns of social support that can help an individual to cope with this kind of social stressor will often again be found in the kin support system. There are, however, several additional complications. First, kin support systems will often be fragmented. Only some people will migrate, not entire support systems. This can mean that some households and individuals are socially isolated, lacking even the most basic social supports. This can also mean that a large burden can be placed on a support system that is fragmented, and that the resulting demands for support are simply too

much for the system to bear. Second, the migrant, especially to North America, is entering a highly competitive and individualistic society. The kinds of mutual rights and obligations entailed by kinship have ceased to be recognized as strongly, for example, in the United States as in traditional societies in Latin America or Southeast Asia. Therefore, in some specific situations, the kin support system can come to be a source of stress and tension, as opposed to a resource for resisting stress.

The complications entailed here are illustrated by migrants from Samoa to northern California. Traditionally organized into large extended kin groups represented by chiefs, Samoans have transplanted their social organization to some U.S. urban centers. Traditionally, the chief (or *matai*) controls economic decision making for the entire extended family. In the American urban setting, however, the economic demands of daily life for an individual household can conflict with the decisions made for the extended family, causing this traditional form of social organization to become a source of stress. At the same time, within the larger extended family, core systems of kin-based social support have emerged, especially involving networks of adult siblings. Individuals and households with a strong support system of this kind have better health status in spite of social stressors such as status incongruence. This specific case illustrates the importance of understanding the adaptation of migrants in a host society in relation to their traditional cultural context, as well as the demands of the new social setting.

Some research on migrants from the developing world has shown the long-term effects of major life events experienced by families in the society of origin. This has been observed in immigrants from Latin America and Southeast Asia who have been exposed to protracted civil war and related conflicts. For example, it was found that persons who had had relatives kidnapped or murdered by nonmilitary death squads in Guatemala had continuing high levels of anxiety and depression years after the event. Similarly, anxiety and depression levels in migrants from Southeast Asia to the United States were associated with time spent in refugee camps, independently from other stressors and demographic control variables. These major crises associated with large-scale political events and circumstances can have effects that continue well after the initial, acute stages.

Conclusion

The study of stress and indigenous societies has helped to illuminate various aspects of the stress

process. Perhaps the most important has been the clear demonstration of the link between the stress process and the social and cultural context in which individuals and families live. When stressors and resistance resources are examined only within a single cultural context, it can appear as if individual differences in exposure to stressors or in access to resistance resources are the only key to understanding the process. What is lost is the recognition that what counts as a stressor or a resistance resource is itself a function of the cultural context and related social influences. Furthermore, the relationship between stressors, resources, and outcomes can also be modified by social and cultural context. Comparing the stress process in different cultural contexts has been integral to revealing this aspect of the process.

Future research must be explicitly comparative in scope in order to expand on findings produced thus far. For example, research in developing societies suggests that the most important stressors in those societies are actually a function of the development process itself, such as status incongruence. Put differently, this aspect of the stress process is comparable across different settings. On the other hand, the most important resources for resisting stress appear to be specific to the local setting, as in the way in which systems of social support are structured by the existing systems of social organization. Continuing to refine these studies, including better measurements of the physiological dimensions of stress, will increase our understanding of human adaptation.

See Also the Following Articles

Cultural Factors in Stress; Cultural Transition; Economic Factors and Stress; Evolutionary Origins and Functions of the Stress Response; Familial Patterns of Stress; Health

and Socioeconomic Status; Industrialized Societies; Minorities and Stress; Psychosocial Factors and Stress; Social Capital; Social Status and Stress; Social Support.

Further Reading

- Decker, S., Flinn, M., England, B. G., et al. (2003). Cultural congruity and the cortisol stress response among Dominican men. In: Wilce, J. M., Jr. (ed.) *Social and cultural lives of immune systems*, pp. 147–169. London and New York: Routledge.
- Dressler, W. W. (1994). Cross-cultural differences and social influences in social support and cardiovascular disease. In: Shumaker, S. A. & Czajkowski, S. M. (eds.) *Social support and cardiovascular disease*, pp. 167–192. New York: Plenum Publishing.
- Dressler, W. W. (1999). Modernization, stress and blood pressure: new directions in research. *Human Biology* 71, 583–605.
- Dressler, W. W. (2004). Culture, stress and cardiovascular disease. In: Ember, C. R. & Ember, M. (eds.) *The encyclopedia of medical anthropology*, pp. 328–334. New York: Kluwer Academic/Plenum Publishers.
- Dressler, W. W. and Bindon, J. R. (2000). The health consequences of cultural consonance. *American Anthropologist* 102, 244–260.
- Flinn, M. V. and England, B. G. (1997). Social economics of childhood glucocorticoid stress response and health. *American Journal of Physical Anthropology* 102, 33–54.
- Janes, C. R. (1990). *Migration, social change and health*. Stanford, CA: Stanford University Press.
- McDade, T. W. (2001). Lifestyle incongruity, social integration, and immune function among Samoan adolescents. *Social Science and Medicine* 53, 1351–1362.
- Panter-Brick, C. and Worthman, C. W. (eds.) (1999). *Hormones, health, and behavior*. Cambridge: Cambridge University Press.
- Rubel, A. J., O'Neill, C. W. and Ardon, R. C. (1984). *Susto: a folk illness*. Berkeley and Los Angeles, CA: University of California Press.

Indoleamines See: Serotonin; Serotonin in Stress; Serotonin Transporter Genetic Modifications.

Industrialized Societies

J Siegrist

University of Duesseldorf, Duesseldorf, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 565–569, © 2000, Elsevier Inc.

Process of Industrialization and Epidemiological Transition
Stressful Social Environments and Health
Concluding Remarks

Glossary

<i>Epidemiological transition</i>	The changing pattern of most prevalent diseases from infectious to chronic degenerative diseases.
<i>Formal rationality</i>	The purposeful calculation of the most efficient means to achieve a goal where rational choices are made to maximize individual benefit.
<i>Gender roles</i>	The cluster of social expectations and norms guiding typically male vs. typically female behavior.
<i>Health-related lifestyle</i>	A collective pattern of health-promoting or health-damaging behavior based on routinized choices people make about food, exercise, hygiene, safety, relaxation, and so on. This pattern is structured by specific needs, social norms, and constraints.
<i>Social disintegration</i>	A weakening of social ties and an absence of cooperative exchange resulting in a lack of social rules and trust and, ultimately, in a breakdown of a social order.

Process of Industrialization and Epidemiological Transition

The industrial revolution started in Great Britain during the second half of the eighteenth century and has subsequently spread across a number of economically advanced countries. This revolution, together with the political revolution initiated in France, must be considered one of the most radical changes in human history. Within a very short time period, fundamental ways of living and working were transformed from a basically agricultural society to a society that is driven increasingly by technology and market economy. With the invention of engines and machines, productivity was increased in an unprecedented way.

A large part of the workforce, employed formerly as peasants and craftsmen, moved to urban areas to engage in unskilled or semiskilled work. The rapid growth of cities and the formation of an industrial workforce contributed to two subsequent, most significant demographic trends: (1) a take-off period of population growth with declining infant mortality rates and (2) a decline in fertility with slow population growth in combination with increased life expectancy. At the same time, the social institutions of marriage and family underwent a marked transformation, and a new division of labor between the sexes started to develop.

After severe poverty and economic exploitation in the early stages of industrial capitalism, economic progress and the development of a welfare state were experienced by a growing proportion of industrial populations. This progress included the availability of healthier food, increased opportunities of energy consumption, better housing, transport, education, and general hygiene. In terms of population health, the process of industrialization was associated with a marked increase in life expectancy and a change in the pattern of prevailing diseases. Overall, 2 to 3 years are added to life expectancy at birth with each decade passing in advanced societies. This increase cannot be attributed to the fact that people, once becoming older, live substantially longer than in previous generations. Rather, increases in life expectancy were, and still are, mainly due to a reduction of death rates at earlier ages and, thus, to increased probabilities of survival. The most substantial decline has been in infant mortality, followed by reductions in childhood and midlife mortality. A dramatic reduction in mortality from infectious diseases played a major role in this process. The typical pattern of leading causes of disability and death in advanced societies is characterized by chronic conditions such as coronary heart disease, cancer, stroke, accidents, and metabolic and mental disorders. This change is referred to as the epidemiological transition.

While the development of modern medicine and related improvements of health care are a direct outflow of the broader processes of industrialization and modernization, it has been demonstrated that the impact of medicine on this major change in the pattern of morbidity and mortality has been limited. In fact, the bulk of the decline in mortality from infectious diseases came before medicine had developed effective forms of treatment and prevention. This means that public health measures and factors related to a general

rise in living standards may have accounted for this change more significantly than medicine.

With this epidemiological transition from infectious to degenerative diseases, medicine and public health face new challenges. A substantial part of these challenges directly or indirectly concerns, the impact of stressful social environments on health. The following section discusses three different aspects of this impact in more detail.

Stressful Social Environments and Health

Stressful Experience and Health-Related Lifestyles

Many of the aforementioned degenerative diseases are considered diseases of affluence because they were prevalent to the extent that life became more comfortable. For instance, with technological advances and with the expansion of commercially available food and drugs, a number of behavioral health risks were programmed, e.g., access to a car reduced physical activity significantly, and, more importantly, it also increased the number of injuries and deaths from accidents. Frequent consumption of fat and meat, at first available to the wealthier groups, contributed to elevated blood lipids, overweight, and associated health risks. Spending money on drugs such as cigarettes or alcohol increased the risk of cancer and heart disease, among others. Consequently, during the first stage of the epidemiological transition, these behaviorally induced risks associated with a wealthier lifestyle were more prevalent among socially and economically privileged groups. At a later stage of the process of industrialization and modernization, however, this social pattern changed: diseases of affluence increasingly became the diseases of the poor. In contemporary societies, this is clearly the case in almost all northern, western, and central European countries, in the United States, and in Canada. The same pattern may be experienced in the near future in countries that are currently exposed to a process of rapid industrialization.

Hence, substantial social inequalities in health are observed, mainly due to higher rates of the former affluent diseases among lower socioeconomic groups.

To explain this shift in the social distribution of a broad range of degenerative diseases and their behavioral risk factors (e.g., smoking, alcohol consumption, overweight, hypertension, metabolic disorders, cardiovascular diseases, some manifestations of cancer, AIDS), the concept of health-related lifestyle needs to be described more precisely.

Health-related lifestyles are collective patterns of health-promoting or health-damaging behavior based on routinized choices people make about food, exercise, hygiene, safety, relaxation, and so on. In many cases, these choices form a coherent pattern that is structured by specific needs, attitudes, and social norms and by the constraints of a group's socioeconomic condition. Once established among a sociocultural group, such behavioral patterns may be transferred from one generation to the next through socialization processes or they may be adopted by model learning from peer groups during adolescence. Unfortunately, the latter is often the case with respect to health-damaging behaviors (smoking, alcohol and drug consumption, unsafe sex practices, etc.). As these behaviors are experienced as relief and reward in stressful psychosocial circumstances they may be reinforced easily and, thus, trigger careers of addiction later in life.

Alternatively, with higher levels of education and knowledge and with the availability of alternative means of coping with stressful psychosocial circumstances, health-damaging lifestyles may be abandoned more easily. This is most probably what happened to the wealthier populations in advanced societies in the recent past and what contributed, and still contributes, to the widening gap in life expectancy between socioeconomic groups.

To date, there are still powerful constraints toward maintaining health-damaging lifestyles among less privileged social groups. A major constraint concerns the exposure to adverse living and working environments. A large number of epidemiologic studies document associations between the amount of exposure to these unfavourable circumstances and the prevalence of unhealthy behavior. For instance, with respect to cigarette smoking, conditions such as heavy workload, shiftwork, social isolation at work, and job instability were identified. Moreover, people who suffer from unfavorable social comparison, e.g., regarding education, income, or general social standing, are at risk of exhibiting a health-damaging life-style. In all these cases, relief and reward experience in terms of some form of addictive behavior may be needed to mitigate a stressful experience.

Some of these processes operate in more subtle ways. It is well known that the processes of primary socialization vary considerably according to the parent's socioeconomic status. In particular, children born into families with a low educational level face more difficulties in deferring their gratifications and in pursuing long-lasting goals in future life. Their sense of self-efficacy and self-esteem is low, and they are more likely to experience their environment as

uncontrollable and fateful. Thus, when facing frustrations, they may be less capable of coping in an active, effective way. Rather, they may seek relief in passive behavior, including drug consumption and related health risks.

This knowledge, derived from sociological and psychological research, has direct implications for health promotion activities and the design of preventive programs. It is evident that comprehensive approaches tackling structural and individual conditions are needed to change health lifestyles in a successful way.

Unfortunately, influential economic interests operate in almost all advanced societies to counteract these measures, most evidently the tobacco and alcohol industry. Advertising activities in mass medias and diffusion of new fashions through opinion leaders are effective means to maintain drug consumption. Efforts to equalize gender roles and to assimilate female attitudes and activities to behaviors that were typical attributes of the traditional male gender role resulted in an unprecedented increase in cigarette consumption among young women. Moreover, women in recent years, as compared to the 1950s, are more likely to work in occupations that were once almost exclusively male. As a consequence of these large-scale societal changes, some of the most advanced industrial societies, such as the United States, are moving toward greater equality in mortality between the sexes.

Impact of Social Disintegration on Health

Industrialized societies are characterized by a high level of mobility. A substantial part of this mobility is due to migration. Economic constraints and wars are the major causes of large migration waves that have shaped the life of many generations throughout the processes of industrialization and modernization. Social mobility is a second component of these secular trends. As mentioned briefly, the onset of early industrialization in Europe and the United States has been facilitated by a broader sociocultural development of modernization that promoted formal rationality and individualism as the dominant modes of thinking and behaving in western societies. Formal rationality is the purposeful calculation of the most efficient means to achieve a goal. Rational choices are made to maximize individual benefit. With the expansion of individual rights, duties, and life chances, traditional ways of maintaining group solidarity, social norms, and obligations lost their significance and attraction. Individual striving for an improved standard of living has become the dominant motivation

overriding traditional social constraints. Thus, a large proportion of people was, and currently is, subject to intragenerational social upward (and downward) mobility in occupational life.

For a long time, social mobility and increased individualism were considered the ultimate benefits of modernization, as legitimized by substantial gains in the quality of life. However, more recently, awareness of the limits, risks, and possible health adverse effects of this development has been growing. One line of thinking along these lines concerns the negative effects on health produced by social disintegration.

Social disintegration is an often reported consequence of enhanced social mobility and related gain in individual freedom. It becomes manifest through a weakening of close social ties (e.g., among family members) and a decrease in reciprocal social exchange or solidarity. Under these conditions, the stability of social relationships is threatened, and commitments held among affiliates may lose their significance. Social disintegration is often associated with a state of social anomie, i.e., a lack of rules and orientations guiding interpersonal exchange. In a society characterized by anomie, individual behavior cannot be predicted by social norms any longer as these norms have lost their validity.

A growing number of investigations conducted by social epidemiologists and medical sociologists document adverse health effects among people who experience social disintegration. Lack of social cohesion or social support, shrinking social networks, and elevated proportions of societal members suffering from social isolation, social exclusion, or individual struggle are different facets of this general trend. For instance, in a cross-cultural perspective analyzing data from 84 different societies, one study concluded that with greater involvement in a money economy a systematic increase in mean blood pressure was evident, even after adjusting for important confounders such as salt consumption and body mass index. With greater involvement in modern money economy, individual competition and striving for benefit became dominant at the expense of cooperative social relationships. In a review of more recent investigations it was concluded that at least eight community-based prospective studies in industrialized societies reveal an association between social disintegration and mortality rates, usually death from all causes. In particular, mortality risks are elevated among people who lack social ties to others. Even at a more distant, macrosocial level, current evidence suggests that a decline in overall mortality, now familiar in most advanced societies, is prevented in an economy that is maximizing

shareholder profit at the expense of investments into community life, culture, and social security.

Although the precise mechanisms that transform stressful experience resulting from social disintegration into increased susceptibility to fatal or nonfatal bodily disease are far from clear, impressive evidence is available from social experiments in mammals. It illustrates the critical role of rewarding and securing experiences that social exchange, affiliation, and affection play in regulating the body's basic physiological processes. It may well be that the cultural canon of modern industrialized societies heavily weighted toward formal rationality, control, and individual striving substantially diminishes the health gain associated with progress in human evolution. This latter conclusion is supported further by a third line of investigation focusing on the impact of modern working life on health.

Stress at Work and Health

The most profound change evoked by the process of industrialization concerns the nature of human work. For a large part of the workforce, work and home become two different places and two different types of social organization. In all economically advanced societies, work and occupation in adult life are accorded primacy for the following reasons. First, having a job is a principal prerequisite for continuous income and, thus, for independence from traditional support systems (family, community welfare, etc.). Increasingly, the level of income determines a wide range of life chances. Second, training for a job and achievement of occupational status are important goals of socialization. It is through education, job training, and status acquisition that personal growth and development are realized, that a core social identity outside the family is acquired, and that goal-directed activity in human life is shaped. At the same time, occupational settings produce the most pervasive continuous demands during one's lifetime, and exposure to harmful job conditions is an important determinant of disability and premature death in midlife.

The nature of work has changed dramatically over the past hundred years. Industrial mass production no longer dominates the labor market. This is due, in part, to technological progress, in part to a growing number of jobs available in the service sector, and, most importantly, in person-based service occupations and professions. In this latter case, unlike in the industrial sector, the potential of rationalization and related cut down in personnel is limited. However, even within industrial work, major changes have taken place. With

less focus on physically strenuous work, with improved safety measures against chemical and physical hazards at work, and with more demands for psychomentally stressful jobs (e.g., information processing, controlling, coordinating activities, often performed under time pressure) new health risks emerged. Meanwhile, there has been a recognition that the importance of work for health goes beyond traditional occupational diseases as there is growing evidence of adverse health effects produced by psychomentally and socio-emotionally stressful jobs.

Before describing these health risks briefly, additional important changes in occupational life need to be mentioned. They are related to the structure of the labor market. There has been an increasing participation of women in the labor market, an increase in short term or part-time working and other forms of flexible job arrangements, and, most importantly, an increase in job instability and structural unemployment. This latter trend currently hits the economically most advanced societies as well as economies that lag behind. Overall, a substantial part of the economically active population is confined to insecure jobs, to premature retirement, or to job loss. Loss of job, as well as continued experience of job insecurity, was shown to be associated with an elevated risk of mortality in a number of prospective studies. In view of the rather high prevalence of these labor market conditions, the adverse effects on health produced by the economy are obvious.

Even within a continuously employed workforce, exposure to stressful psychosocial work environments carries the risk of ill health. Examples of this latter trend are shift work, irregular working hours and overwork, social isolation at work, conflicting demands, lack of control, and monotony. People working in jobs that are defined by a combination of at least some of these characteristics are at risk of developing chronic disease conditions such as hypertension, coronary heart disease, low back pain, or depression at least twice as often compared to people working in more favorable jobs. Additional evidence documents substantially elevated health risks among workers and employees who suffer from an imbalance between high efforts spent at work and low rewards received in terms of money, esteem, or promotion prospects. In many cases, such unfavorable job conditions are maintained, e.g., due to lack of alternatives in the labor market.

Concluding Remarks

In conclusion, despite substantial progress in life expectancy and the standard of living, the process

of industrialization in technology, science, economy, and sociopolitical development has produced a substantial health burden that is attributable to an increasingly stressful social environment. Three aspects of this development were analyzed: (1) the impact of psychosocially adverse circumstances on health-damaging lifestyles; (2) health risks associated with social disintegration; and (3) afflictions of stressful working life and labor market conditions on well-being and health. It seems unlikely that these trends will be diminished in the near future as the process of economic and sociocultural globalization is progressing at a high pace. Moreover, pressures originating from population growth in rapidly developing parts of the world, new socioenvironmental risks, and man-made disasters may aggravate rather than moderate this burden. At the same time, with scientific progress in understanding the determinants and dynamics of stress-related health risks and with new worldwide public health efforts evolving, there is still hope that modernization and industrialization of human societies result in a sustainable humane and more healthy future.

See Also the Following Articles

Economic Factors and Stress; Gender and Stress; Health Behavior and Stress; Indigenous Societies; Social Support; Workplace Stress.

Further Reading

- Blaxter, M. (1990). *Health and lifestyles*. London: Routledge.
- Cockerham, W. C. (1998). *Medical sociology* (7th edn.). Upper Saddle River, NJ: Prentice Hall.
- Henry, J. P. (1997). *Cultural change and high blood pressure*. LIT Publishing, Germany: Münster.
- Henry, J. P. and Stephens, P. M. (1977). *Stress, health, and the social environment*. Berlin: Springer.
- Hobfoll, S. E. (1998). *Stress, culture, and community*. New York: Plenum.
- Karasek, R. A. and Theorell, T. (1990). *Healthy work: stress, productivity, and the reconstruction of working life*. New York: Basic Books.
- Marmot, M. G. and Wilkinson, R. (1999). *Social determinants of health*. Oxford: Oxford University Press.
- Weber, M. (1958). *The protestant ethic and the spirit of capitalism*. New York: Scribner.

Infection

H Friedman and S H Pross

University of South Florida College of Medicine,
Tampa, FL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by H Friedman and S Specter, volume 2, pp 570–580, © 2000, Elsevier Inc.

Introduction

Central Nervous–Immune–Endocrine System Interactions
Stress Hormones and Infections
Bacterial Infections and Steroids
Virus Infections and Stress
Discussion and Conclusion

Glossary

<i>Adrenocorticotropic hormone (ACTH)</i>	A secreted serum factor from the adrenal glands that is considered a stress hormone.
<i>Cytokines</i>	Soluble proteins secreted by one type of cell that interacts and activates immune cells.

Legionella pneumophila

An opportunistic bacterium that causes Legionnaire's disease and primarily infects macrophages, a class of immune cells responsible for host resistance to infectious diseases.

Lentiviruses

A retrovirus, such as human immunodeficiency virus, that primarily infects immune cells and can induce acquired immunodeficiency syndrome.

Opportunistic infections

Infections caused by otherwise lowly pathogenic microbes in individuals who have some immunodeficiency.

Respiratory viruses

Small microorganisms that preferentially infect the cells of the respiratory system.

Introduction

Stress is a major factor in the progression of infectious diseases in humans as well as in experimental and domestic animals. In this regard, it is now known that the central nervous system (CNS), the endocrine system, and the immune system are complex systems that interact with one another. In particular, the brain and CNS exert a profound effect on the immune

system, including mechanisms necessary for resistance to microbial infection. Stress dysregulates the immune response by affecting the interplay of these systems. New knowledge developed during the last few decades through experimental and clinical studies provide evidence that many immune alterations are due to stressful events that provoke health changes. Stressors, including specific stressful events, can stimulate the activities of the hypothalamic-pituitary-adrenal (HPA) axis as well as the sympathetic nervous system to enhance an individual's ability to adapt physiologically in response to a threat (Figure 1). Furthermore, psychological stress ensues when events or environmental demands exceed an individual's perceived ability to cope. Stress can increase susceptibility to infectious agents, influence the severity of infectious diseases, diminish an immune response to a vaccine, reactivate a latent herpesvirus, and in general depress resistance to opportunistic infectious organisms that otherwise should not cause overt disease. Moreover, stressful events and the stress they evoke also may substantially increase the production of pro-inflammatory cytokines associated with a wide range of diseases, especially chronic immune diseases.

Not only the CNS is complex; there is now much new knowledge about the complexity of the immune

system, which consists of both cellular and humoral elements. Many immune responses are mediated by soluble factors such as cytokines produced by lymphoid cells and by factors important for control of the CNS, including steroids and various neuropeptides and neurohormones produced by neuronal cells (Figure 1). Important mechanisms whereby stress influences susceptibility or the more rapid progression of infectious diseases are becoming clearer because of new experimental and clinical studies by investigators worldwide. Specifically, the current explosion of new knowledge concerning the nature and mechanism of the immune response system has focused attention on the important role of various lymphoid cells for resistance versus susceptibility to microbes, especially intracellular pathogens such as opportunistic bacteria, viruses, fungi, and even parasites. It is now known that the brain and CNS in general have profound effects on immunity. The interaction between the brain and immune system entails a complex series of activities involving responses of many cells to multiple stimuli.

Central Nervous-Immune-Endocrine System Interactions

The modulation of the immune response by the CNS is mediated by a complex network of bidirectional interactions between the nervous, endocrine, and immune systems. The HPA axis and autonomic nervous systems provide two key pathways for immune system dysregulation. For example, stress may activate the sympathetic adrenomedullary (SAM), as well as the HPA axis, and thereby result in the release of pituitary and adrenal hormones, including catecholamines (epinephrine and norepinephrine), ACTH, cortisol, growth hormone, and prolactin, which are influenced by stressful events, including negative emotions. Each of these soluble factors results in quantitative and qualitative changes in immune activity. For example, external factors, including stress, increase cortisol levels, which may induce adverse immunological alterations. Immune cells express receptors for one or more factors associated with the HPA and SAM axes, generally called hormones. Immunomodulation by hormones occurs through two pathways: directly, through binding to a specific cell receptor, or indirectly, by the induction of cytokines such as interferon (IFN- γ), interleukins (ILs), and pro-inflammatory cytokines such as tumor necrosis factor (TNF- α). Such cytokines interact with various target cells. Furthermore, the interaction between the CNS and the immune system is bidirectional. For example, immune cell-derived ILs such as IL-1 affect the production of corticotropin releasing hormone

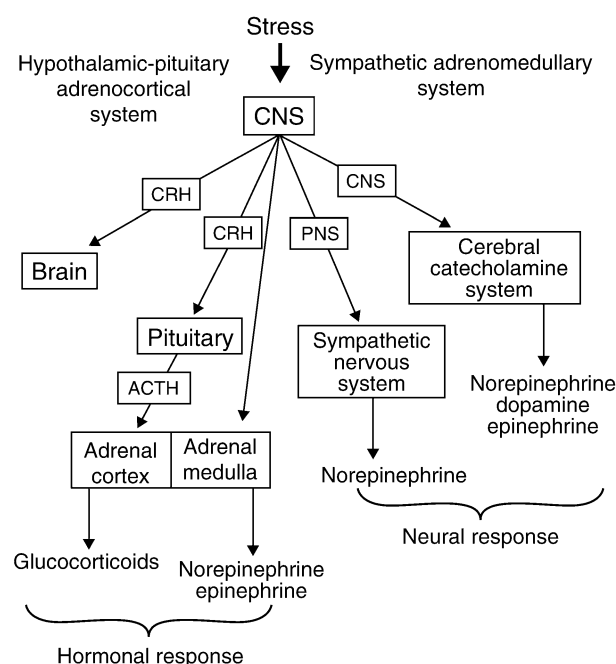


Figure 1 Arrangement of the hypothalamic-pituitary-adrenal and sympathetic adrenomedullary systems. The central noradrenergic system is activated in parallel with the peripheral sympathetic nervous system, and it is the central nervous system counterpart. The cerebral catecholamine system (i.e., dopamine and epinephrine) is activated under stress. ACTH, adrenocorticotrophic hormone; CNS, central nervous system; CRH, corticotropin releasing hormone; PNS, peripheral nervous system.

(CRH) by the hypothalamus in the brain. An increase in stress hormone levels results in the further dysregulation of immune function. Similarly, lymphocytes produce many hormones such as ACTH, prolactin, and growth hormone.

Resistance to microbes, including bacteria, viruses, and parasites, depends on both innate and adaptive immune responses, consisting of a variety of cell types such as T and B lymphocytes; phagocytic monocytic cells, such as dendritic cells and macrophages; neutrophils; and polymorphonuclear cells. Soluble factors produced mainly by lymphoid cells, including ILs, IFNs, and other cytokines, are essential for host resistance. Immune resistance is based on primary and secondary antibody responses and cell-mediated immunity due to T lymphocytes such as Th1 effector and Th2 helper cells, as well as Th3 regulatory cells. Such cells produce the many humoral factors important for immunoregulatory and protective pro-inflammatory activity, especially to intracellular opportunistic microbial pathogens.

Stress Hormones and Infections

As already indicated, many types of signals to the brain (e.g., visual, tactile, and emotional) cause the hypothalamus to release CRH, and, in turn, this stimulates the pituitary gland to release ACTH into the circulation, which then stimulates the adrenal cortex to produce glucocorticoids. ACTH inhibits the response of lymphocytes to antigenic or mitogenic stimulation, depresses cytokine responses by lymphocytes, and suppresses natural killer (NK) cell activity and antibody responses to a variety of antigens, including those associated with microorganisms. It is widely recognized that increased glucocorticoid production is by no means the only response to stress. The levels of several neuropeptides, including vasoactive intestinal peptides, substance P, prolactin, and growth hormones are increased, as are catecholamines such as epinephrine and norepinephrine. Localized increases in catecholamine levels probably account for the suppression of lymphocyte responses to mitogens or antigens. The immune system also returns signals to the CNS to close this circle. As previously stated, immune cells synthesize a number of neuropeptide hormones, including ACTH and β -endorphin. Moreover, cytokines produced by lymphocytes and macrophages have effects that are signals to the brain. Various ILs, including TNF- α , stimulate the hypothalamus, resulting in the induction of fever and release of corticosteroid hormones. Receptors for neuropeptides are present on lymphoid cells (Table 1), and neuropeptides derived from neural or immune cells stimulate leukocytes (Tables 2–3).

Table 1 Receptors for neuropeptide hormones on cells of the immune system^a

<i>Hormone</i>	<i>Receptor site</i>
ACTH	Rat spleen T and B cells
β -Endorphin	Rat spleen T and B cells
Calcitonin gene-related peptide	Rat spleen T and B cells
CRH	Human PBL
Growth hormone	Rodent spleen, thymus, human PBL
Growth hormone-releasing hormone	Rodent spleen, human PBL
Luteinizing hormone-releasing hormone	Rat thymocytes
Prolactin	Human and mouse T and B cells
Somatostatin	Human PBL
Substance P	Mouse T and B cells
Thyrotropin	Human neutrophils, monocytes, B cells

^aACTH, adrenocorticotrophic hormone; CRH, corticotropin releasing hormone; PBL, peripheral blood leukocytes.

Table 2 Stimuli for leukocyte-derived peptide hormones^a

<i>Hormone</i>	<i>Stimulator</i>
Arginine vasopressin	Unknown
Corticotropin and endorphin	Newcastle disease virus Lymphotropic viruses Bacterial lipopolysaccharide CRH Arginine vasopressin IL-1
Follicle-stimulating hormone	Con A
Growth hormone	Growth hormone-releasing hormone Growth hormone
Luteinizing hormone	Luteinizing hormone-releasing hormone
Metenkephalin	Con A
Neuropeptide Y	Unknown
Parathyroid hormone-related protein	Human T-cell lymphotropic virus
Prolactin	Con A
Somatostatin	Hypersensitivity
Substance P	Unknown
Thyrotropin	Staphylococcal enterotoxin A Thyrotropin-releasing hormone
Vasoactive intestinal peptide	Histamine liberators Hypersensitivity

^aCon A, concanavalin A; CRH, corticotropin releasing hormone; IL, interleukin.

Many studies have shown that increases in stress, especially those resulting in increased levels of stress hormones such as corticosteroids, alter the susceptibility of experimental animals to infectious diseases significantly (Table 4). The study of effects of

Table 3 Modulation of immune responses by neuropeptides^a

Neuropeptide	Immune response
Arginine vasopressin	T-cell helper function
Corticotropin	Functions for IFN- γ production
	Antibody synthesis
	IFN- γ production
Endorphins	B-lymphocyte stimulation
	Antibody synthesis
	Mitogenesis
	NK-cell activity
Growth hormone	Cytotoxic T-cell activity
	Mitogenesis
Growth hormone-releasing hormone	Inhibit NK-cell activity
	Inhibit chemotactic response
Luteinizing hormone	Cytokine formation
Prolactin	IL-2 receptor production
Somatostatin	Reduces IFN- γ production
Substance P	Stimulate chemotaxis
	Stimulate proliferation
	Modulate cytokine levels
Thyrotropin	Antibody synthesis
	Co-mitogen
	Co-mitogenic with Con A

^aCon A, concanavalin A; IFN, interferon; NK, natural killer.

Table 4 Effect of glucocorticoids on microbial immunity^a

Organism	Glucocorticoid effect
<i>Bordetella pertussis</i>	Promotes initial protective effect, then promotes microbial survival
<i>Candida albicans</i>	Exacerbates infection
<i>Escherichia coli</i>	Decreases inflammation but not bacterial clearance
<i>Legionella pneumophila</i>	Exacerbates infection
<i>Listeria monocytogenes</i>	Increases susceptibility, impairs monocyte function
<i>Mycobacterium avium</i>	Impairs macrophage function
<i>Mycobacterium leprae</i>	Alters IFN- γ effects on monocytes
<i>Mycobacterium tuberculosis</i>	Increases cellular infiltration, increases bacterial growth
<i>Neisseria gonorrhoeae</i>	Increases infection rate
<i>Plasmodium berghei</i>	Decreases parasitemia
<i>Pneumocystis carinii</i>	Increases susceptibility
<i>Propionibacterium acnes</i>	Exacerbates infection
<i>Pseudomonas aeruginosa</i>	Increases susceptibility
<i>Staphylococcus aureus</i>	Ameliorates IP route infection
<i>Streptococcus faecalis</i>	Increases resistance
<i>Trichinella spiralis</i>	Decreases nematode clearance, cellular infiltration
<i>Yersinia pestis</i> (<i>Pasteurella pestis</i>)	Increases invasive infection, bacteremia

^aIFN, interferon; IP, intraperitoneal.

neuroendocrine-immune interactions on microbial pathogenesis and immunity during the course of infectious disease is an important area of interdisciplinary research, and various studies showed that

physiological stress and psychiatric illness can compromise immune function and alter susceptibility to infectious agents (Table 5). It is widely known that patients with psychiatric disorders, including schizophrenia and even depression, are more likely to have infectious diseases. This may be associated with altered HPA activity or physiological dysfunction related to immunomodulation associated with drugs that affect the brain. Unconventional life styles, eating disorders, poor hygiene, and other factors associated with severe mental health problems may contribute to altered immune function. In this regard, it is now recognized that the recreational use and abuse of illegal drugs such as marijuana, cocaine, heroin, and other opiates, as well as more widely used legal substances such as alcohol and tobacco, have marked effects on the CNS and the immune system. Individuals who use these drugs frequently are more susceptible to opportunistic microorganisms. The detrimental effects of such abused drugs on immunity in both humans and animals have now been extensively studied. These drugs are known to increase corticosteroid levels in both humans and experimental animals, either at the time of addiction or during withdrawal, and this effect has been related to altered immune resistance to microbial infection. Addiction to a psychoactive substance (as in the stimulation of the HPA axis) induces enhanced adrenal corticosteroid secretion, followed by consequential increases in serum glucocorticoids and the activation of the sympathetic nervous system. These responses result in immune dysregulation and increased susceptibility to microbial infections, especially to opportunistic pathogens.

Numerous early clinical observations reported that individuals often become more susceptible to infections following stressful situations. However, the concept of stress was not defined sharply. A stressor was often defined as any type of external or internal challenge, visual, tactile, or emotional, that disrupts the physiological equilibrium or homeostasis of an individual. For example, a number of studies showed that infection by mycobacteria, which cause tuberculosis, is increased by stress.

Alterations in the susceptibility to other opportunistic bacteria, as well as to viruses and parasites in general, have been studied extensively. Increases in infections due to stress occur not only in human populations but also in farm animals and experimental animals. There is now a large body of biomedical literature showing that stress alters susceptibility to infection. The mechanism of the stress includes effects on humoral function such as the production of glucocorticoids. The effects of stresses on two types

Table 5 Effects of stress on microbial immunity

Pathogen	Animal species	Stressor	Effect
<i>Escherichia coli</i>	Chicken	Cold	Increases resistance
	Chicken	Social	Increases resistance
<i>Hymenolepis nana</i>	Mouse	Predator	Increases reinfection rate
<i>Mycobacterium avium</i>	Mouse	Restraint	Increases susceptibility, impairs macrophage function
<i>Mycobacterium tuberculosis</i>	Rabbit	Tumbling	Increases susceptibility
<i>Plasmodium berghei</i>	Mouse	Electric	Increases survival
<i>Salmonella enteritidis</i>	Chicken	Fasting	Decreases antibody response
<i>Salmonella typhimurium</i>	Mouse	Crowding	Increases susceptibility

of infective diseases, bacterial and viral, are discussed next as examples.

Bacterial Infections and Steroids

Legionella pneumophila, an opportunistic ubiquitous bacterium, is the causative agent of Legionnaire's disease. Although almost all individuals are exposed to *Legionella*, only a small percentage of individuals exposed to these organisms are infected clinically. Indeed, when *Legionella* was first discovered in the outbreak of Legionnaire's disease in Philadelphia in the late 1970s, most of the 2500 conventioners exposed to this organism from the air conditioning system in the Bellevue Stratford Hotel developed significant titers to these bacteria. However, only approximately 200 individuals became clinically ill, and a significant number of these people died. These individuals had certain factors associated with their susceptibility, including increased alcohol consumption, cigarette smoking, (in general) being older, and being more immunocompromised. It is now well recognized that *L. pneumophila* causes clinical infection mainly in immunocompromised individuals, including patients who are treated with immunosuppressive drugs in hospitals for cancer chemotherapy or for organ transplantation.

Increases in susceptibility to infection by *Legionella* associated with steroid use have been studied in detail (Table 4). Although steroid therapy for various diseases has had revolutionary therapeutic effects and is widely recognized as constituting a major success for modern medicine, it is often associated with serious side effects; steroids have both beneficial and nonbeneficial features, and these may be inseparable. The anti-inflammatory and immunosuppressive activities of steroids are considered essential for the therapy of certain diseases, including autoimmune diseases, and to prevent organ rejection following transplantation. However, steroids may also result in an increased susceptibility to infections, especially by an opportunistic bacterium such as *L. pneumophila*. Although the spectrum of microbes causing infections in patients

during steroid therapy is not completely clear, experimental animal studies showed that steroid treatment is associated with increased infections caused by certain groups of opportunistic bacteria, including *Legionella*. The general consensus concerning the mechanism whereby steroids induce altered susceptibility to bacterial infection is that steroids suppress the immune response system.

The most important host defense mechanism against invading intracellular opportunistic bacteria is phagocytosis by macrophages, followed by intracellular killing and digestion of the microbes. It is now clear that steroids suppress the number and function of macrophages and neutrophils (Figures 2–3). For example, experimental studies have shown that peritoneal macrophages from hydrocortisone-treated mice are more susceptible to infection and killing by *L. pneumophila in vitro*. The steroid-involved impairment of antimicrobial activity by macrophages from mice infected with *Legionella* is generally similar to that reported for *Listeria monocytogenes*. For example, *in vivo* treatment of mice with hydrocortisone acetate affects the ability of peritoneal macrophages from the animals to kill *L. pneumophila in vitro*, as well as other microorganisms such as *Listeria monocytogenes*, *Streptococcus pyogenes*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Escherichia coli*, and *Pseudomonas aeruginosa*.

The treatment of macrophages with steroids *in vitro* may affect antimicrobial immunity directly. For example, the ability of macrophages treated with steroids to kill *L. pneumophila* and organisms such as *Staphylococci in vitro* is markedly reduced. The specific mechanisms involved in the steroid impairment of macrophages, however, are unclear. In general, the intracellular killing of bacteria and other microbes by macrophages requires lysosomal migration to and fusion with the microbe-containing phagosome. It is probable that the mechanism of the steroid-induced impairment of antimicrobial activity of monocytes or macrophages is associated with the stabilization of lysosomes by steroids, which prevents their fusion with phagosomes. However, the possibility exists that

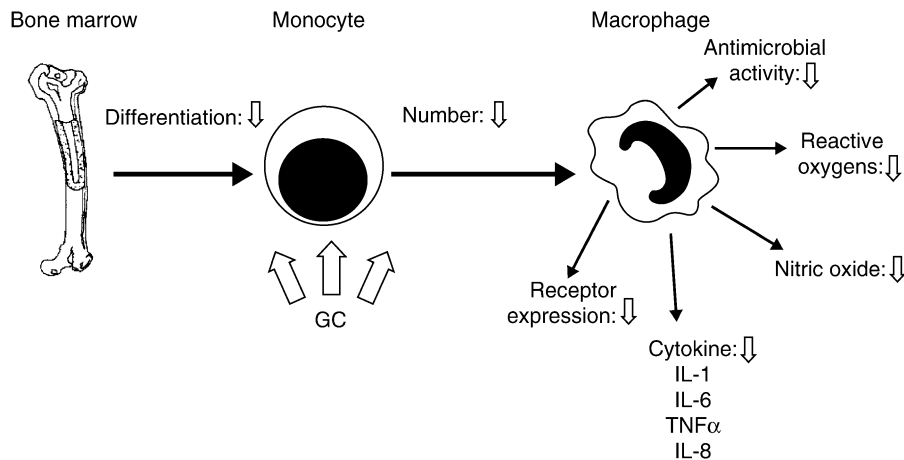


Figure 2 The effect of steroids on macrophages and monocytes. Glucocorticoids have a direct action on the number and function of phagocytic cells, including cytokine production and receptor expression. ↓, decrease; GC, glucocorticoid; IL, interleukin; TNF, tumor necrosis factor.

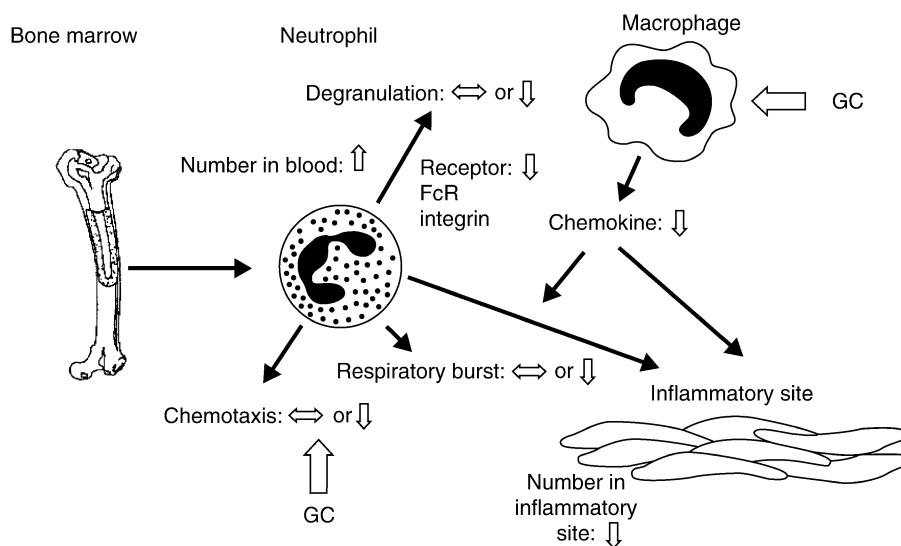


Figure 3 Prevention of neutrophil accumulation at local inflammatory site by steroid treatment. The chemotactic activity of neutrophils from steroid-treated animals or humans was enhanced, inhibited very weakly, or unaffected, except for neutrophils from inflammatory exudates, which showed reduced chemotactic activity. ↓, decrease; ↑, increase; ↔, minimal or no change; FcR, Fc receptor; GC, glucocorticoid.

steroids induce the impairment of the antimicrobial activity of macrophages by suppressing the synthesis of one or several specific proteins of the killing system of the phagocytes. Nevertheless, several observations have indicated that macrophages treated with steroids ingest microorganisms at a normal rate and, after phagocytosis, secrete the usual reactive oxygen intermediates and exhibit normal fusion of lysosomes and phagosomes (i.e., the oxidative killing systems). Steroid-treated macrophages react to a cytokine such as IFN and are normally primed to secrete reactive oxygen intermediates on challenge. Thus, there have been reports negating the likelihood there is a direct

deleterious effect of steroids on the antimicrobial activity of macrophages.

Virus Infections and Stress

Virus infection can be exacerbated by stress hormones such as cortisone. For example, hydrocortisone-treated mice show an increased susceptibility to challenge with viruses such as influenza virus, Coxsackie virus, or even mouse hepatitis virus. Peritoneal macrophages from cortisone-treated mice are almost 1000 times more susceptible to challenge with the mouse hepatitis virus *in vitro* than are macrophages from

untreated mice. Resistant macrophages from normal mice succumb to cytolysis by mouse hepatitis virus when cultured with spleen cells from cortisone-treated mice but not with spleen cells from normal mice, suggesting that a soluble product from lymphoid cells may be involved in increased susceptibility of the macrophages due to steroids.

Various studies have reported the association between naturally occurring stresses and individuals' susceptibility to upper respiratory tract viral infections. Although most of these studies provide some support for the hypothesis that stress is associated with an increased risk of upper respiratory viral infections, it is not yet clear how this occurs. Longitudinal studies performed with large populations of individuals suggest that stress is a significant factor in increasing the susceptibility to viral infection.

A number of studies have been performed testing the possibility that stress is involved in experimentally induced upper respiratory tract viral infections. Psychologically stressed, experimentally infected individuals had significantly higher levels of upper respiratory tract infections, but there were no differences in the incidence of infections or the amount of shed virus in nasal secretions between these and unstressed individuals. Several studies reported that stress precedes the onset of upper respiratory tract viral infection episodes, especially in children; for example, several studies reported a significant increase in minor stresses a few days before the onset of upper respiratory tract infection symptoms. The likelihood exists that the severity of the symptoms of upper respiratory tract infections rather than the incidence of infection may be affected by different types of psychological stress. So far, it is apparent that more clinical studies must be performed before we can fully understand the relationship between stress and upper respiratory tract viral infections in patient populations. In addition, little is known about how the timing of the stressor and viral exposure influences the development of these infections. There is some evidence suggesting that stress occurring after experimental viral exposure may increase susceptibility; however, there are differing published findings, some indicating an increased and some a decreased susceptibility even when stress precedes viral exposure.

Studies have been reported concerning stress and progression of acquired immunodeficiency syndrome (AIDS) in human immunodeficiency virus (HIV)-positive subjects. For example, a study published by investigators from the University of Miami School of Medicine, supported by the National Institute of Mental Health, showed that HIV-positive high-risk subjects had less evidence of progression to clinical AIDS when psychological intervention techniques,

such as stress reduction by counseling, music therapy, or exercise, were used. This study supported the view that psychological stress in HIV-positive individuals can result in a more rapid progression to clinical AIDS and that psychological stresses were often related to lack of an effective support system and with the notification of HIV-positive status. Such findings suggested that psychological stress and its intervention may be associated with the progression of chronic life-threatening viral infections, such as HIV.

There are also well-known studies concerning the effects of stress on herpesvirus exacerbation. A landmark study performed at the Ohio State University Medical School reported that medical students taking a routinely scheduled examination had increased psychological stress immediately preceding the examination and that this was associated with decreased immune function and the increased recurrence of herpesvirus titers, indicative of the reactivation of the herpesvirus. The study was particularly well controlled, consisting of medical students who were generally healthy and immunologically mature (i.e., approximately the second decade of life). All the students had normal immune functions, but at the time of the examination they were shown by psychological tests to have a significant stress level and increased serum glucocorticoids. Simultaneously, they showed a decrease in an immune function important in resistance to viral infection, NK cell activity and T-cell levels. Many of the students had an exacerbation of latent herpes shortly after the examination (with its associated stress).

Note that temporal relations often cannot be controlled readily in population studies. Furthermore, little is known about the differential effects of acute (short-term) versus chronic (long-term) or repetitive stressors. The clinical literature suggests that chronic stress may have particularly strong pernicious effects on health. But it is difficult to undertake true experimental studies of stress exposure in human populations because it is difficult to randomly assign subjects to defined stressor conditions. Although challenge viral exposure in experimental studies appears to be relatively controlled, it is often not because of independent variables and because stress is difficult to assess and manipulate. Furthermore, little is known about the biological pathways that may be responsible for the association between stress and increased viral infection in patient populations.

The effect of psychosocial interventions in ameliorating a number of infectious problems in patients with diseases as diverse as cardiovascular disease and cancer is currently being investigated. Many investigators believe that appropriate psychosocial intervention may result in reduced stress and thus

reduced viral infections. Nevertheless, most studies to date with patient populations and viral infections suggest only that there is an association between stress and increased susceptibility to infection.

Discussion and Conclusion

It is now widely accepted that stress plays an important role in altering the susceptibility to infectious agents such as bacteria, viruses, and even parasites. Furthermore, stress is also acknowledged to be involved in increasing the susceptibility to the progression of autoimmune diseases and malignancies. Many studies have suggested that the stress hormones, especially cortisone and ACTH, are important in altering the immune mechanisms, especially in the lymphoid cells, which are important in the resistance to opportunistic microorganisms such as intracellular bacteria and viruses.

Various studies showed that a variety of stressors increase the susceptibility of experimental animals to bacterial infections. Some of the most defined studies to date were performed with mycobacteria, organisms that cause tuberculosis in humans and in domestic and experimental animals. Infection by mycobacteria is known to be increased by stress; for example, a number of widely known studies showed that mice that were normally resistant to mycobacterial infection showed increased susceptibility when given certain stressors. This has been related to increases in glucocorticoid levels, which are known to affect lymphoid cell function.

Glucocorticoids produced by the adrenal cortex under the influence of ACTH from the pituitary gland in the brain affect both humoral and cellular immune responses. Mice treated with glucocorticoids often show increased cellular infiltration following infection with mycobacteria and increased growth of the bacteria. Glucocorticoids have been shown to increase the susceptibility of experimental mice to infectious agents such *Listeria monocytogenes* and *Pneumocystis carinii*. However, glucocorticoids promote the survival of *Bordetella pertussis*, the causative agent of whooping cough in humans, although initially this hormone provides some protection against challenge infection. Furthermore, cortisone ameliorates infection by *Staphylococcus aureus* in experimental animals, but has been reported to increase the susceptibility to *Pseudomonas aeruginosa* in humans and animals and also to exacerbate infection by the ubiquitous fungus *Candida albicans*. Thus, there is a marked variability in the effects of corticoids on infection, depending on the organisms involved; the reasons for this remain unknown. For example, various studies showed that infection with

Propionibacterium acnes is exacerbated by corticosteroids but that infection with *Streptococcus faecalis* is decreased by them. Furthermore, animals infected with *Plasmodium*, the cause of malaria, showed decreased parasitemia due to corticosteroids.

It is clear that corticosteroids decrease the accumulation of neutrophils at a local inflammatory site. Such neutrophil accumulation is important in clearance of or resistance to bacterial infection. It has also been reported that the chemotactic activity of neutrophils from steroid-treated animals or humans can be enhanced, inhibited very weakly, or unaffected; however, most studies showed that neutrophils obtained from inflammatory exudates in steroid-treated mice showed reduced chemotactic activity. Such data indicate that the poor accumulation of neutrophils in inflammatory sites may be mediated by the inhibition of the release of local chemotactic factors rather than a direct effect on neutrophils. This possibility seems likely because steroid treatment does not inhibit neutrophil accumulation induced by the injection of a chemotactic complement component into human skin. It has been demonstrated that the treatment of purified human neutrophils with the synthetic steroid deoxymethasone does not inhibit neutrophil chemotaxis in response to a range of concentrations of known chemoattractants.

There are several mechanisms for gene regulation by the steroid-receptor complex. Glucocorticoid response elements regulate the rate of transcription initiated by promoters of the corresponding glucocorticoid response genes. Another possible regulation mechanism is the interaction with DNA-binding proteins associated with regulatory elements of DNA, such as the glucocorticoid modulatory element protein. Thus, it appears that steroids regulate the transcription of certain genes coding for proteins that augment steroid effects.

The infection of animals with the ubiquitous bacterium *L. pneumophila* is modulated markedly by steroids. Thus, steroid levels are an important factor in the increased susceptibility of organ-transplant patients, who are often immunosuppressed with steroids prior to the transplant and thus may become more susceptible to an opportunistic infection such as *Legionella* in hospitals. Viral infections have also been shown to be exacerbated by stress hormones, including cortisone. It is well known that hydrocortisone-treated mice show an increased susceptibility to challenge with a variety of viruses, including influenza virus, Coxsackie virus, and mouse hepatitis virus. It seems likely that, because stress is often associated with increased corticosteroid levels only, changes in the level of this hormone may be related to the apparent increased susceptibility of individuals to an upper respiratory tract viral infection following a

stressful event. Results obtained using animal models show an increased susceptibility to challenge infection with bacteria or viruses, and studies of humans, infected either naturally or experimentally, strongly suggest a relationship between stress-induced immunomodulatory hormones and increased susceptibility to infection.

The complex interrelationship between the nervous and the immune system parallels the association of the susceptibility or resistance to infections by various microorganisms under conditions of stress. This is even more likely because it is known that the immune response is modulated markedly by a wide variety of neuropeptides associated with the CNS, not only corticotropin but also other hormones such as endorphins, prolactin, growth hormone, vasopressin, and somatotrophin. Neuropeptides affect a wide variety of immune responses, including antibody formation, mitogenic responses of lymphocytes, cytotoxicity by T cells, NK cell production and activity, and cytokine production. Thus, neuropeptides, hormones, and their receptors, present not only in the CNS but also on immune cells, are important factors in immunomodulation related to altering the susceptibility to infection by microorganisms, especially intracellular opportunistic bacteria and viruses. However, because of the complexity of both the CNS and the immune system, it is difficult to isolate the various pathways that are important in this altered susceptibility to infection associated with stress. Further studies in numerous laboratories, including those mainly concerned with the neurological system and those concerned with infectious diseases and immunity, will undoubtedly provide a greater understanding of the role of stress in infection.

See Also the Following Articles

Antibody Response; Desensitization; Disease, Stress Induced; Herpesviruses; HIV Infection/AIDS; Immunity; Leukocyte Trafficking and Stress; Natural Killer (NK) Cells.

Further Reading

- Ader, R. and Cohen, N. (1993). Psychoneuroimmunology: conditioning and stress. *Annual Review of Psychology* **44**, 53–71.
- Ader, R., Felten, D. and Cohen, N. (1991). *Psychoneuroimmunology* (2nd edn.). San Diego, CA: Academic Press.
- Akira, S. and Kishimoto, T. (1992). IL-6 and NF κ B in acute phase response and viral infection. *Immunological Reviews* **127**, 26–41.

- Baue, A. F. (1989). Neuroendocrine response to severe trauma and sepsis. In: Faist, E., Ninnemann, J. & Green, D. (eds.) *Immune consequences of trauma, shock and sepsis*, pp. 17. New York: Springer-Verlag.
- Bayle, J. E., Guellati, M., Ibos, F., et al. (1991). *Brucella abortus* antigen stimulates the pituitary-adrenal axis through the extra-pituitary B-lymphoid system. *Progress in Neuroendocrinimmunology* **4**, 99–112.
- Beisel, W. R. (1975). Metabolic response to infection. *Annual Review of Medicine* **26**, 9–29.
- Blalock, I. E. (1992). Production of peptide hormones and neurotransmitters by the immune system. In: Blalock, J. E. (ed.) *Neuroimmunoendocrinology* (2nd edn.). Basel: Karger.
- Friedman, H., Klein, T. W. and Friedman, A. L. (1996). *Psychoneuroimmunology, stress and infection*. Boca Raton, FL: CRC Press.
- Friedman, H., Klein, T. W. and Specter, S. (1996). *Drugs of abuse, immunity and infections*. Boca Raton, FL: CRC Press.
- Friedman, H., Newton, C. and Klein, T. W. (2003). Microbial infection, immunomodulation and drugs of abuse. *Clinical Microbial Reviews* **16**, 209–220.
- Glaser, R. and Kiecolt-Glaser, J. K. (2005). Stress-induced immune dysfunction: implications for health. *Nature Reviews Immunology* **5**, 243–252.
- Hart, B. L. (1988). Biological basis of the behavior of sick animals. *Neuroscience and Biobehavioral Reviews* **12**, 123–129.
- Jefferies, W. M. (1991). Cortisol and immunity. *Medical Hypotheses* **34**, 198–215.
- Kass, E. H. and Finland, M. (1958). Corticosteroids and infections. *Advances in Internal Medicine* **9**, 45–61.
- Padgett, D. A. and Glaser, R. (2003). How stress influences the immune response. *Trends in Immunology* **24**, 444–453.
- Rabin, B. S. (1999). *Stress, immune function, and health: the connection*. New York: Wiley-Liss & Sons.
- Schattner, A. and Schaffner, T. (1987). Glucocorticoid-induced impairment of macrophage antimicrobial activity: mechanisms and dependence on the state of activation. *Review of Infectious Diseases* **9**, 5620.
- Sheridan, J. F., Dobbs, C., Brown, D., et al. (1994). Psychoneuroimmunology: stress effects on pathogenesis and immunity during infection. *Clinical Microbial Reviews* **7**, 200.
- Stone, A. and Bovbjerg, D. (1994). Stress and humoral immunity: a review of human studies. *Advances on Neuroimmunology* **4**, 49.
- Stone, E. A. (1975). Stress and catecholamines. In: Friedhoff, A. J. (ed.) *Catecholamines and neuropsychopharmacology*, (vol. 2), p. 31. New York: Plenum Press.
- Webster, J. I., Tonelli, L. and Sternberg, E. M. (2002). Neuroendocrine regulation of immunity. *Annual Review of Immunology* **20**, 126–163.

Inflammation

G Z Feuerstein, R R Ruffolo, C Coughlin, J Wang and D Miller

Wyeth Research, Collegeville, PA, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction

Immune and Inflammatory Cells in Acute Inflammation:

Role of Endothelium and Adhesion Molecules

Inflammation, Artherosclerosis, and Cardiovascular Diseases

Inflammation and Cancer

Inflammation and Metabolic Syndrome

Inflammation and Respiratory Diseases

Inflammation and Joint Diseases

Inflammation and Sepsis

Summary

Glossary

Apoptosis

A process of cell death that results from the activation of specific signaling pathways that lead to the intracellular disruption of mitochondrial integrity, activation of cascade of proteases (e.g., caspases), and cell disintegration while maintaining the cell membrane intact.

Chemokines

A family of polypeptides that comprises about 50 members and 20 different receptors. This family of polypeptides is all involved in leukocyte traffic. Their diverse biological actions include interactions with viruses and cell receptors in numerous tissues.

Chemotaxis

The orientation and movement of a cell along a specific chemical concentration. Positive chemotaxis denotes movement of the cell toward the high-concentration site of the chemical; negative chemotaxis denotes movement of the cell away from the source of the chemical. Chemotactic properties are prevalent in all forms of leukocytes and in a wide variety of substances released in sites of tissue injury or activated cells.

Complement system

The entire functionality of approximately 20 distinct plasma proteins, their receptors, and ancillary/regulatory proteins that are responsible not only for immune cytolysis but also for reactions such as anaphylaxis, phagocytosis, opsonization, and hemolysis. Two major pathways for activation of the complement system are known: the classic pathway

Cytokines

in which all components C1–C9 participate in the reaction and the alternative pathway in which only part of the plasma factors participate along with others that are not essential in the classic reaction.

Nonantibody peptides and polypeptides that are released by activated cells (e.g., T cells) and transfer signals to other cells. Lymphokines (cytokines produced by lymphocytes) and monokines (cytokines produced by monocytes) are cytokines.

Inflammation

A localized protective response elicited by injury or destruction of tissues, which serves to destroy, dilute, or wall off (sequester) both the injurious agent and the injured tissue. It is characterized in the acute form by the classic signs of pain, heat, redness, swelling, and loss of function. Histologically, it involves a complex series of events, including the dilation of the arterioles and capillaries, extravasation of fluids, and migration of plasma proteins and leukocytes into the injured area.

Interleukins (ILs)

A generic term for a group of multifunctional cytokines that are produced by a variety of lymphoid and nonlymphoid cells. Originally believed to play a role only in leukocyte biology, ILs are now recognized as the products of diverse nonlymphoid cells that also act on diverse receptors on multiple cells outside the lymphoid system. Over 20 ILs have been identified to date that act via different receptors with large divergent and cross-activation among the IL family members. ILs are involved in many disease processes and are targets for promising drug discovery.

Leukocytes

Colorless blood cells capable of ameboid movement; they are capable of moving directionally in response to chemical stimuli (chemotaxis). Leukocytes are divided morphologically into granular cells (basophils, eosinophils, and neutrophils) and nongranular cells (lymphocytes and monocytes). Leukocytes are the primary cells that launch immune and inflammatory reactions. Leukocytes are present in the systemic circulation and in special lymphoid tissue such as the lymph node, bone marrow, and thymus.

Lymphocytes

Mostly mononuclear, nonphagocytic leukocytes found in the blood and

lymphoid tissue that make up the primary body of immune competent cells. Lymphocytes belong to two primary classes – T cells (thymus dependent) and B cells (Bursa dependent) – but several other subclasses with various functional and biochemical markers are known. Lymphocytes function via direct cell–cell interaction (specialized receptors) and via humoral mediators such as cytokines, interleukins, and chemokines as well as antibodies (a primary B-cell function). T cells are primarily responsible for cell-mediated immunity.

Mast cells

A connective tissue cell that elaborates highly reactive mediators such as cytokines, histamine, and enzymes, which can modulate matrix proteins such as chymotrypsin and tryptase. Mast cells (or mastocytes) are believed to play an important role in allergic rhinitis and asthma and, via enzymes released, to contribute to inflammation and atherosclerosis.

Matrix metalloproteinases (MMPs)

Zinc-dependent endopeptidases known for their ability to cleave one or several constituents of the extracellular matrix. MMPs consist of a large family of proteases that share common structural and functional elements.

Metabolic syndrome

A clinical condition that manifests four or five risk factors for cardiovascular and metabolic disorders. These include obesity, high glucose levels, high blood pressure, and aberrant plasma lipid levels (high triglycerides and low high-density lipoproteins).

Monocytes

Mononuclear phagocytic leukocytes that originate in the bone marrow. Monocytes are the source of tissue macrophages, which are abundant in lung and liver.

Nonsteroidal anti-inflammatory drug (NSAID)

Small organic (synthetic) chemicals that have been developed to suppress inflammation of various mechanisms, locations, and durations. They are commonly compounds that do not act like the steroid hormones (e.g., glucocorticoids). This class of medicines contains primarily inhibitors of the cyclooxygenase enzymes (which produce pro-inflammatory metabolites of the fatty acid arachidonate) and displays various proportional inhibitory capacities toward the cyclooxygenase 1 or 2 enzymes. Many such agents are marketed for pain relief from conditions such as arthritis and other muscular-skeletal conditions.

Sepsis

A condition that results from a harmful or host response to infection. Sepsis is marked by endotoxemia (the presence of bacterial or other toxins in the blood) and the activation of multiple systems such as the immune, cardiovascular, respiratory, coagulation, and complement systems. Sepsis can lead to multiple organ failure and death.

Introduction

Inflammation is a fundamental reaction of the body to injury. It is a complex biological phenomenon that comprises diverse cellular, neural, hormonal, and humoral systems. Inflammation involves the vascular system as well as blood-borne components linked via a complicated network of mediators. Inflammation can be the result of endogenous or exogenous injurious events. Inflammation is fundamentally viewed as a purposeful reaction aimed to repair, salvage, and restore the structure, function, and integrity of the injured tissue, organ, and organism well-being at large.

However, inflammation may also become a destructive or pathological condition that underlies numerous diseases in humans. Such diseases include bone and cartilage diseases (rheumatoid arthritis and osteoarthritis), skin diseases (psoriasis), central nervous system (CNS) diseases (multiple sclerosis and Alzheimer's disease), vascular diseases (atherosclerosis), cardiac diseases (heart failure and myocardial infarction), respiratory diseases (asthma and chronic obstructive pulmonary disease, COPD), gastrointestinal diseases (inflammatory bowel disease and Crohn's disease), and chronic unresolved infectious diseases (e.g., tuberculosis). In recent years, inflammation is increasingly recognized as a fundamental contributing element in many human diseases that have not been traditionally associated with inflammation. Conditions such as obesity, diabetes, metabolic syndrome, stroke, and cancer are all recognized as harboring various immune and inflammatory cells and mediators.

Inflammation is invariably linked to general homeostasis systems (e.g., coagulation and fibrinolysis) as also to the immune, complement, and the nervous system. As such, intervention in or modulation of specific pro-inflammatory mediators might result in broader consequences than the intended anti-inflammatory objective. Moreover, many mediators of inflammation are actually part of the normal physiological repertoire of molecules and cells that secure essential function and health. Therefore, the modulation (blockade or enhancement) of many recognized pro-inflammatory mediators, although

potentially to providing medical benefits, could also harm. Thus a fine balance needs to be preserved in many cases so as not to disrupt functionalities that require the same mediators to be at a level context and duration that, if not secured, could lead to pathological conditions on their own.

The advances made in understanding the cellular elements and mediators that initiate, propagate, and amplify various forms of inflammation have already yielded important therapeutics. Nonsteroidal anti-inflammatory drugs (NSAIDs), steroids, and, more recently, anti-cytokines and interleukin mimetics have an important role in the treatment of inflammation. It is, however, important to keep in mind that no anti-inflammatory therapeutic introduced into medical practice is free of adverse effect.

Immune and Inflammatory Cells in Acute Inflammation: Role of Endothelium and Adhesion Molecules

Acute injury to tissues, such as bruises, blunt-force trauma, or infectious agents, elicit a typical inflammatory reaction. The early response consists of local neurovascular reaction and release of preformed mediators that induced vasodilation, microvascular permeability, and extravasation of plasma components into the injured zone. This early stage is followed by the recruitment and migration of leukocytes from the systemic circulation with primary role of the polymorphonuclear (PMN) cells. The PMN cells are followed by monocytes, which together with the PMS cells and tissue macrophages provide the local tissue with the ability to conduct phagocytosis of invading pathogens, engulf cell debris (e.g., necrotic cells), and set the stage for tissue repair. The recruitment of leukocytes into the injury zone is governed by two key processes: chemotaxis and cell adhesion. The directional influx of leukocytes into the injured tissue is dependent on mediators such as chemokines or complement factors, which, along with cytokines and interleukins, also activate the leukocytes. In addition, the endothelium along the microvessels associated with the injury zone express adhesion molecules (in addition to constitutive ones) that enable the interaction with ligands that are expressed on the leukocytes. These cell adhesion molecules act as a cascade of events – from loose rolling, loose adhesion, committed adhesion, and final migration mostly through the endothelium junction. The key adhesion molecules are the vascular cell adhesion molecule (VCAM-1); intercellular adhesion molecule (ICAM-1), selectins (E, P, and L), and platelet-endothelial cells adhesion molecules (PECAMs).

Inflammation, Artherosclerosis, and Cardiovascular Diseases

The major pathological process that underwrites heart attacks, strokes, and peripheral vascular diseases is arteriosclerosis. Although the role of lipids (cholesterol in particular) in initiation and propagation of the arteriosclerotic pathology is well documented, inflammation is considered to be a major pathological feature in the evolution of this vascular disease. Lipid-loaded cells are the hallmark in any vascular atheroma, or plaque, which are macrophages (derived from blood monocytes) that phagocytosed modified cholesterol within lipoproteins particles. This process leads to the activation of the macrophages and release of cytokines, chemokines, interleukins, and other pro-inflammatory mediators such as arachidonic acid metabolites (lipoxygenase and cyclooxygenase). Apoptosis of excessively loaded macrophages (also called foam cells) is believed to release numerous inflammatory mediators that exacerbate the inflammatory process and increase the likelihood of its rupture. Most significantly in this regards are matrix metalloproteinases (MMPs; especially MMP-2 and -9), which are believed to reduce the fibrous cap of the atheroma and enhance plaque rupture, leading to the initiation of thrombosis and vascular occlusion. The atheromatous plaque is also rich in other inflammatory cells such a mast cells and T cells. The vascular inflammation part of the atherosclerosis process is viewed as an important therapeutic target and a component of the efficacy of major anti-atherosclerosis drugs such as the statins.

Inflammation and Cancer

A causal link between cancer and chronic inflammation has been demonstrated in the more than 1 million cancer patients linked to viruses and bacteria. Infection with human papillomaviruses, HCV, HBV; *Helicobacter pylori*; and Epstein-Barr virus are major risk factors for the development of cervical, gastrointestinal (GI), and hepatocellular cancer and lymphoma, respectively. The link between inflammation and cancer is not limited to infectious agents; chronic inflammation from cigarette smoke, asbestos, inflammatory bowel disease, silica, and so on contributes to the development of malignancy in the lungs, bladder, and GI tract. The activation of innate inflammation has been shown to contribute to cancer progression, including tumor necrosis factor (TNF)- α , and nuclear factor (NF)- κ B signaling. The disruption of the adaptive (or antigen-specific) immune response to tumors has been shown to be crucial to carcinogenesis. Many studies have documented the anergic antitumor

immune responses contrasting with the favorable prognosis conferred by the infiltration of tumors with cytotoxic killer T cells in several malignancies. In contrast, infiltration with supportive immune elements, such as growth-promoting macrophages, confers a worse prognosis in some malignancies. Finally, the association between multiple inflammatory biomarkers, such as interleukin (IL)-8, macrophage inflammatory protein (MIP)-1a, and various chemokines, whose expression levels correlate directly with worse prognosis, underlies the intimate relationship between inflammation and cancer development and progression.

Inflammation and Metabolic Syndrome

Metabolic syndrome emerges as the twenty-first-century epidemic that threatens both developed and less developed nations. Although the original concept defining metabolic syndrome focused on insulin resistance, namely, the poor responsiveness of target cells to insulin's ability to advance glucose metabolism, it has been recently recognized that inflammatory mediators are likely to be causally associated with the state of insulin resistance and, hence, metabolic syndrome. Most notable in this regard are the cytokines TNF- α , and IL-6. These cytokines, and primarily TNF- α , reduce the efficacy of insulin to enhance glucose uptake into the cells, a situation that manifests itself as hyperglycemia, a hallmark of the insulin resistant condition. Obesity, a key feature of the metabolic syndrome, is considered a pro-inflammatory state. Obesity is associated with increased oxidative stress, a biochemical condition prone to activating NF- κ B, NADPH, and other factors. Oxidative stress-induced inflammation increases endothelial leukocyte adhesion molecules that enhance leukocyte engagement with the endothelium, a primary step for their entry into the tissue, their activation, and the further release of inflammatory mediators. Excessive caloric intake, from carbohydrates or fats, is considered to significantly contribute to the inflammation state that drives key elements of metabolic syndrome.

Inflammation and Respiratory Diseases

Chronic obstructive pulmonary disease (COPD) is a chronic lung disorder that is mostly the result of smoking and other environmental toxic agents. The disorder is characterized by the accumulation of inflammatory cells such as macrophages and neutrophils, which are activated to release numerous mediators. The inflammation resides within the terminal bronchioles (small airways) and alveoli (small vesicular spaces where oxygen and CO₂ exchange takes

place). The persistent inflammation ultimately leads to the destruction of these pulmonary structures and respiratory failure. Prominent in the COPD pathomechanism are the inflammatory products of arachidonic acid and MMPs released by the activated inflammatory cells. MMP-9 and, in particular, MMP-12 play a predominant role in COPD by degrading the matrix proteins such as elastin, a matrix protein that is vital to the recoil of the bronchioles (small airways) and to alveolar health. MMP-12-mediated elastin degradation is believed to lead to emphysema, a state of destruction of the normal architecture of the lung respiratory elements. MMP-12 also plays a role in macrophage recruitment. Promising data emanating from recent research suggest that selective inhibitors of MMP-12 might be useful drugs in the treatment and prevention of COPD.

Bronchial asthma, a disease of lung airway spasm that leads to severe respiratory distress and severe forms to death, has strong inflammatory components. Arachidonic acid metabolites of potent inflammatory properties figure prominently in this disease. Lipoxigenase products called leukotrienes, prostanoids (prostaglandins and thromboxane A₂), and isoprostanes have been implicated in this disorder. Indeed, anti-inflammatory agents such as glucocorticosteroids have been a cornerstone of asthma-attack prevention, largely due to their inhibition of inflammatory eicosanoids. In addition, specific inhibitors of cysteinyl leukotriene have proven to serve additional therapeutic efficacy.

Inflammation and Joint Diseases

Several joint and bone diseases, such as rheumatoid arthritis and osteoarthritis, are probably the result of degenerative (the former) or immune and inflammatory (the latter) processes. Rheumatoid arthritis is a chronic systemic disease primarily affecting the joints in which severe chronic inflammation leads to the destruction of the joint structure and, ultimately, functional disability. Misdirected inflammation is the cardinal feature of rheumatological diseases. It involves both the initiation and maintenance phases of the disease, and the processes and players (cells and mediators) involved are similar to that in other diseases. Major cells in the immune system (i.e., neutrophils, macrophages, dendritic cells, B cells, and T cells) are the major culprits in the development and exacerbation of the disease. Although autoimmune mechanisms (including the generation of antibodies against joint tissue components) are believed to play a major causative role, the chronic condition is probably sustained by numerous inflammatory mediators, including cytokines, ILs, chemokines,

MMPs, prostanoids, and reactive oxygen radicals. Traditional anti-inflammatory agents such as the NSAIDs have been widely used to alleviate pain in and swelling of the joint, but their efficacy in mitigating the fundamental pathological process remained questionable. Likewise, the more recent class of anti-rheumatoid arthritis drugs, the COX-2 inhibitors, have a proven efficacy, yet at the cost of some increased risk for a cardiovascular adverse event. More recently, disease-modifying anti-arthritis drugs (DMARDs) have been successfully introduced into the medical management of rheumatoid arthritis. Such agents neutralize the cytokine TNF- α . Other cytokines and ILs (e.g., IL-6, IL-1, and IL-18) as well as interferon (IFN)- γ and endothelium leukocyte adhesion molecules (ICAM-1, VCAM-1, and selectins) are also targeted by therapeutic agents, as well as angiogenesis promoters such as vascular endothelium growth factor (VEGF).

The role of inflammation in osteoarthritis is less clear, although many of the cytokines already discussed are believed to play a key role in the destruction of bone and cartilage. Cartilage destruction probably involves cytokine activation of aggrecanase, an enzyme that destroys a major constituent of the normal cartilage.

Inflammation and Sepsis

Sepsis is a complex clinical syndrome that results from the harmful host response to infection; it results from the failure of the innate and acquired immune mechanisms to resolve the infectious condition and the damage induced by the toxins elaborated by the infectious agent. The bacterial toxins activate many of the inflammatory cells (in particular, neutrophils and monocytes/macrophages) but also other cells such as the endothelium. As a result of the toxins' action on the inflammatory cells, numerous preformed and newly elaborated mediators are produced, including oxygen radicals, bioactive lipid mediators, cytokines, chemokines, enzymes, and pro- and anticoagulation factors. Although many of the inflammatory cell products are helpful in confining and eliminating the infectious agent and its toxins, multiple tissues, organs, and system are also affected. Oxygen radical have bactericidal action, yet also affect the vasculature. Chemokines and cytokines modify the endothelium to support the fast recruitment of inflammation cells into the tissue to remove the invading agent while ILs activate and enhance the neutrophils, lymphocytes, and endothelium function. The inflammation cells also play a role in tissue repair and rehabilitation via the generation and release of growth factors that enhance

the regeneration of the damaged circulation, the matrix elements, and tissue integrity. The immune and inflammation cells and mediators are the most critical systems that secure defenses in sepsis. The failure of these defenses could lead to organ failure and death.

Summary

Inflammation plays a key role in acute injuries as well as chronic diseases. The immune and inflammatory reactions serve as the major defense mechanism against numerous external injurious events and pathological events that emerge as a result of endogenous aberrant genetic and biochemical imbalances. However, immune and inflammatory reactions can also evolve as a cause of diseases such as autoimmune diseases; granulomatous conditions; cancer; respiratory, joint, and bone diseases; and atherosclerosis. Thus, anti-inflammatory drugs emerge as an important source of novel therapeutics for diverse diseases such as cancer, metabolic syndrome, atherosclerosis, joint and bone diseases, and respiratory tract disorders.

Further Reading

- Barish, G. D., Narkar, V. A. and Evans, R. M. (2006). PPAR δ : a dagger in the heart of the metabolic syndrome. *Journal of Clinical Investigations* **116**, 590–597.
- Bloomgarten, Z. T. (2005). Inflammation, atherosclerosis, and aspects of insulin action. *Diabetes Care* **28**, 2312–2319.
- Cohen, J. (2002). The immunopathogenesis of sepsis. *Nature* **420**, 885–891.
- Dandona, P., Aljada, A., Chaudhuri, A., et al. (2005). Metabolic syndrome: a comprehensive perspective based on interaction between obesity, diabetes and inflammation. *Circulation* **111**, 1448–1454.
- Fontech, A. N. and Wykle, R. L. (2004). Arachidonate remodeling and inflammation. In: Parnham, M. J. (ed.) *Progress in Inflammation*, pp. 1–243. Basel: Birkhauser Verlag.
- Lagente, V., Manoury, B., Nenan, S., et al. (2005). Role of matrix metalloproteinases in the development of airway inflammation and remodeling. *Brazilian Journal Medical Biological Research* **38**, 1521–1530.
- Libby, P. and Theroux, P. (2005). Pathophysiology of coronary artery disease. *Circulation* **111**, 3481–3488.
- Morgan, D. W., Forssmann, U. J. and Nakada, M. T. (2004). Cancer and inflammation. In: Parnham, M. J. (ed.) *Progress in inflammation*, pp. 1–201. Basel: Birkhauser Verlag.
- Moser, B., Letts, G. and Neote, K. (2006). Chemokine biology – basic and clinical application. Vol. 1: Immunology of chemokines. In: Parnham, M. J. (ed.) *Progress in inflammation research*, pp. 1–205. Basel: Birkhauser Verlag.

- Rolin, S., Masereel, B. and Dogne, J.-M. (2006). Prostanoids as pharmacological targets in COPD and asthma. *European Journal of Pharmacology* 533, 89–100.
- Schottenfeld, D. and Beebe-Dimmer, J. (2006). Chronic inflammation: a common and important factor in the pathogenesis of neoplasia. *Journal of Clinical Investigations* 56, 69–83.
- Simmons, D. L. (2006). What makes a good anti-inflammatory drug target? *Drug Discovery Today* 11 (5–6), 210–219.
- Van den Berg, W. B. and Miossec, P. (2004). Cytokines and joint injury. In: Parnham, M. J. (ed.) *Progress in inflammation*, pp. 1–291. Basel: Birkhauser Verlag.
- Weizmann, M. N. and Pacifici, R. (2006). Estrogen deficiency and bone loss: an inflammatory tale. *Journal of Clinical Investigations* 116, 1186–1193.
- Zelcer, N. and Tontonoz, P. (2006). Liver X receptors as integrators of metabolic and inflammatory signaling. *Journal of Clinical Investigations* 116, 607–614.

Inhibition See: Emotional Inhibition.

Injuries See: Acute Trauma Response.

Insomnia See: Sleep, Sleep Disorders, and Stress.

Instinct Theory

R Gardner Jr

University of Wisconsin, Madison, WI, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R Gardner Jr., volume 2, pp 581–582, © 2000, Elsevier Inc.

The word instinct stems from a Latin root meaning impulse or instigation. Historically, the term differentiated innate from learned actions. Instincts were untaught, genetically preprogrammed, stereotypic, goal-directed behaviors and behavior chains shaped by natural selection. They exhibited more complexity than reflexes and occurred mostly in nonhuman animals. Nobel Prize-winning ethologist Konrad Lorenz described fixed motor patterns; releasers, the triggering stimuli of instinctive responses; and imprinting, wherein a newly hatched bird followed and attached

itself to any moving object located before it, replacing its mother.

The concept of inherited behavior remains important. However, Nikos Tinbergen, who shared the 1972 Nobel Prize with Lorenz, pointed out that the popular use of the term instinct has three meanings: (1) behavior driven from within (e.g., the mating instinct emerges in the spring), (2) nonlearned behavior (e.g., the newborn knows instinctively how to suckle its mother's nipple), and (3) unpremeditated behavior (e.g., the driver instinctively puts his foot on the brake). The term instinct, therefore, is often omitted in works of present-day ethologists, sociophysicologists, behavioral ecologists, and other workers in animal and human behavior.

Prior to the ethologists, the theoretician–psychoanalyst Sigmund Freud used the word instinct (also libido and drive) to carry forward his machine metaphors of

the mind (which had been popular in the nineteenth century), employing, for example, hydraulic and expended-fuel metaphors (as Lorenz also did for a time). Later John Bowlby, in contrast, defined instinct as behavioral-system actions that follow a recognizably similar pattern typical of the species, that have survival value, and that emerge even without the ordinary opportunities for learning.

Machine models hold little interest for modern theorists and researchers. Indeed, explosions of biological research on all fronts have rendered distinctions between learned and innate knowledge impossible. The capacity for learning is clearly inherited, a property that is expanded more in humans than in other animals. Moreover, research in molecular biology, notably in deciphering the genomes of many living forms, has demonstrated numerous conserved genes held in common by species, from bacteria to flies to fish to humans. Many influence behavior. Modern mammals exhibit reactions that ancestral vertebrates surely also had, for example, the flight-or-flight response.

Many anciently determined behaviors hinge on conspecific recognition and interaction (conspecifics are members of a same species), and these entail learning and behavioral modifications to varying extents. Certainly mating, social rank hierarchy, nurturance, eliciting nurturance, making in/out-group distinctions, and play behaviors must be included. These behaviors go awry in psychiatric disorders; examples include

1. Persecutory delusions, in which the patient has the conviction that he or she is targeted by hostile out-group members.
2. Mania, which involves the patient behaving in a commanding (dominant) fashion inappropriate for circumstances.
3. Depression, which involves the patient behaving like someone of low rank.
4. Paraphilias, which can be redefined as displaced mating urges.
5. Autism, in which children elicit nurturance poorly from parents and do not play spontaneously with other children.

Humans possess a brain three times heavier than their nearest relatives among the great apes, probably from greater sociality, which seems to enhance adaptation more than it costs in childbirth complications. However, 98.5% of primate genomes remain identical. Inherited neurons, altered to varying extents by experience, clearly determine all behaviors. Neuroscientist-psychiatrist Eric Kandel stressed that all learning reflects altered gene transcription and neuronal action. Learned behavior, once the antonym of instinct, now instead implies additional choices for the individual. Psychiatric leader Daniel X. Freedman noted that biological psychiatry is a redundancy; there is no unbiological psychology or sociology. All animal behavior, including human action, represents genetically programmed performances influenced by individual experience.

See Also the Following Articles

Evolutionary Origins and Functions of the Stress Response; Fight-or-Flight Response.

Further Reading

- Allman, J. M. (1999). *Evolving brains*. New York: Scientific American Library.
- Bowlby, J. (1969). *Attachment and loss* (vol. 1). New York: Basic Books.
- Gardner, R. (1997). Sociophysiology as the basic science of psychiatry. *Theoretical Medicine* 18, 335–356.
- Gardner, R. (2001). Affective neuroscience, psychiatry and sociophysiology. *Evolution & Cognition* 7, 25–30.
- Kandel, E. R. (1998). A new intellectual framework for psychiatry. *American Journal of Psychiatry* 155, 457–469.
- Lorenz, K. (1965). *Evolution and modification of behavior*. Chicago: University of Chicago Press.
- Price, J. S., Gardner, R. and Erickson, M. (2004). Depression, anxiety and somatization as appeasement displays. *Journal of Affective Disorders* 79, 1–11.
- Tinbergen, N. (1969). Instinct. In: *Encyclopedia Britannica*, (vol. 12), pp. 323–324. Chicago: William Benton.
- Verhulst, J., Gardner, R., Sutton, B., et al. (2005). The social brain in clinical practice. *Psychiatric Annals* 35 (10), 803–815.

Integrative Medicine (Complementary and Alternative Medicine)

D Spiegel

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D Spiegel, volume 1, pp 512–514, © 2000, Elsevier Inc.

Introduction
Acupuncture
Biofeedback
Nutritional and Herbal Treatments
Homeopathy
Hypnosis
Manipulative Therapies
Mindfulness-Based Stress Reduction
Relaxation Training
Conclusion

Glossary

<i>Acupuncture</i>	The insertion of thin needles at specific meridians or points in the body thought to affect the balance of the body's ability to manage pain and other medical illnesses, a practice originally developed in China.
<i>Homeopathy</i>	A treatment based on the idea that tiny amounts of substances can change the properties of water in a way that triggers the body's ability to respond to a disease.
<i>Hypnosis</i>	A state of highly focused attention, coupled with reduced awareness of experiences at the periphery of attention, and a heightened responsiveness to social cues.
<i>Integrative medicine</i>	The integration of many alternative or culturally sanctioned techniques into mainstream Western medical practice.

Introduction

While modern medical science has provided cures for many bacterial infections, life-extending treatments for viral infections such as HIV, and symptom-reducing interventions for heart disease, and has converted many kinds of cancer from a terminal to a chronic disease, it is especially striking that large numbers of Americans now turn to allied medical traditions and recently developed treatments for help with disease. More people than ever are living with serious illness, which is a chronic and severe stressor. Thus,

traditional medicine's very success has created a whole new set of problems. The acute disease model emphasizes diagnosis, definitive treatment, and cure. In some situations this works, but in many it does not. The leading killers of Americans, heart disease, stroke, and cancer, are by and large chronic and progressive rather than curable illnesses. The application of a curative model, in a setting in which disease management is often the most productive approach, leaves both doctors and patients frustrated. Rather than feeling satisfied when a disease is well managed, doctors and patients often feel disappointed that it has not been cured, and there is little emphasis on managing pain, anxiety, fatigue, and other debilitating aspects of illness and its treatment. The oldest adage of medicine is that our job is to "cure rarely, to relieve suffering often, and to comfort always." At the end of the twentieth century, it is as though that adage has been rewritten, that our job is now to "cure always, relieve suffering if one has the time, and leave the comforting to someone else."

Patients do indeed turn elsewhere for comfort. Complementary and alternative medicine professionals are one such place. They spend on average 30 min with their patients, compared with the typical 7 min allocated for physician–patient visits. Americans spend more out of pocket on complementary and alternative treatments than on visits to outpatient physicians or hospital care. An illness can be a lonely journey, and patients crave people who understand what the journey is like and who can stay with them during its course. Thus, the apparent appetite for complementary and alternative medicine is stimulated by the lack of help in coping with serious illness in modern medical care, which has become increasingly technological. This problem is being intensified by the business constraints in modern American medicine, which include frustration at obtaining appropriate reimbursement, increased paperwork burden, limited continuity of care, and second-guessing of treatment decisions. This can only further desiccate the doctor–patient relationship and limit the opportunities for medical compassion to enhance medical care.

Close to one-half of Americans utilize some form of alternative treatment, a substantial increase from the already sizeable figure of one-third in 1990. However, more than two-thirds of them utilize standard medical treatment as well, suggesting that as far as the patients are concerned, these interventions are complementary rather than alternative. A study in the

Kaiser health-care system indicated that the majority of complaints that lead to seeking alternative services are pain (52%) and anxiety (25%). Thus, patients seek complementary care for sensible reasons, bringing problems that are not life-threatening and that are often ignored or minimized in mainstream medical treatment. However, the relationship between complementary and standard medical care is complicated by the fact that three-quarters of those individuals who use alternative services do not tell their primary care physicians that they were obtaining these alternative or complementary services, so their use of them had the potential to undermine or at least complicate their medical care.

These data provide support for the use of the term integrative rather than complementary or alternative. The idea is to integrate the best of other treatment traditions into modern medical care. The NIH now has a National Center for Complementary and Alternative Medicine (NCCAM) that conducts studies and disseminates information on such treatments. Centers for integrative medicine have been opened at a number of prominent academic and proprietary medical centers. The modalities described briefly below have been shown to be helpful in stress management.

Acupuncture

This ancient Chinese technique is primarily employed in the United States for chronic pain, and to a lesser extent for habit problems or substance abuse. Given the interaction between pain and stress in chronic illness, the pain relief obtained through this technique can also reduce secondary stress and anxiety. Needles are inserted in locations thought to mediate pain perception and healing, usually alone, but sometimes with either mild electrical stimulation (electroacupuncture) or the combustion of a herbal preparation (moxibustion). Acupuncture analgesia has been found to be reversed by administration of naloxone, an opiate antagonist, so mediation of its effects via endogenous opiates is possible. Many patients find that they achieve a sense of well-being after acupuncture treatment that provides them with a respite from the stress of chronic illness.

Biofeedback

Biofeedback involves the use of telemetry recording peripheral muscle tension, skin temperature, or skin conductance (reflecting sweating) to inform people about their physiological responses and their ability to modulate them. This approach has been effective in treating tension and migraine headaches, managing stress, and treating psychophysiological disorders

such as Reynaud's disease, which involves vasoconstriction. Training with the equipment is utilized to teach patients how to enhance their ability to manage physiological responses to stress, tension, or other symptoms and thereby feel less secondary anxiety when a stressful situation arises.

Nutritional and Herbal Treatments

There are longstanding traditions of herbal treatments that contain a mixture of generally low-potency and complex combinations of substances that may have specific therapeutic benefit in some cases, or that may be ineffective or even harmful in others. Digitalis, a mainstay in the treatment of congestive heart failure and other cardiac disease, was originally the active component of foxglove, used as a herbal treatment for dropsy. Research supported by NCCAM has shown that omega-3 fatty acids, found in some fish oils, may be helpful in the management of autoimmune inflammatory diseases. Arthritis and other such diseases may be helped by a combination of interventions, including gentle movement such as *tai chi*, a Chinese technique of exercise designed to redirect energy flow in the body, and pain management including hypnosis and biofeedback. However a recent large study by NCCAM found that the herb hypericum, also known as St. John's Wort, did not appear effective in relieving the symptoms of depression, as had been found in previous studies. L-tryptophan, an amino acid precursor of serotonin, has been used in attempt to increase serotonin levels in the brain. Low serotonin has been linked to depression and impulsive behavior, and the most widely used antidepressants, selective serotonin reuptake inhibitors, increase the activity of serotonin in the synapses between neurons in the brain. However, one manufacturer of these dietary supplements produced batches with a small amount of a very harmful contaminant that produced inflammatory muscle pain in many and a number of deaths as well. While their product has been withdrawn from the market, it is an example of how apparently benign dietary supplements can be harmful. The use of some herbs may enhance peoples' sense of control over their disease, independent of specific effects, and thereby at least temporarily alleviate disease-related stress.

Homeopathy

This treatment regimen, developed in Germany and the United States, utilizes serial dilutions of substances thought to impart therapeutic properties in the water or alcohol that may initially mildly provoke and then help the body to heal symptoms of disease. Homeopathy and also other techniques such as

traditional Chinese medicine emphasize a complex assessment of all bodily systems, treating imbalances in these systems and not just focusing on specific diseased parts of the body. Studies conducted through NCCAM and elsewhere have not proven specific effectiveness of homeopathic remedies beyond what might be expected from placebo response.

Hypnosis

There are a variety of well-established and thoroughly studied methods for teaching patients how to increase their control over mental distress and concomitant somatic discomfort. Hypnosis, for example, is a state of aroused, attentive focal concentration accompanied by a reduction in peripheral awareness. It has proven effectiveness in facilitating anxiety control and inducing analgesia. The techniques most often employed involve a combination of physical relaxation and imagery that provides a substitute focus of attention for the painful sensation. Hypnosis can be quickly induced in a matter of seconds, and patients can be taught self-hypnosis to provide ongoing anxiety and pain control. The ability of hypnotizable individuals to dissociate, or separate psychological from somatic response, can be used to maintain physical relaxation even in the face of emotional distress. This can also help patients to restructure their understanding of the situation that is making them anxious, in part because they are approaching it while in an unusually relaxed state. This is an important concept in many complementary treatments for anxiety such as meditation, massage, biofeedback, and hypnosis: the patient may not be able to alter the stressor itself, but can enhance his or her control over his or her psychological and somatic response to the stressor. This can, in turn, enhance the array of options available to the patient for dealing with that stressor, since he or she is more in control of somatic symptoms and less preoccupied with the effect of the stressor.

Self-hypnosis, for example, has been shown to be quite useful in helping medical patients suffering from procedure anxiety. A recent randomized trial employing hypnosis during interventional radiology procedures demonstrated that patients offered training in hypnosis had less pain, fewer procedural interruptions, less cardiovascular instability, and less anxiety while utilizing only 12% as much analgesic medication from a patient-controlled analgesia (PCA) device.

Manipulative Therapies

Approaches such as osteopathy, chiropractic, and therapeutic massage approach symptoms such as pain and limitation of movement, especially involving

musculoskeletal disorders, through careful assessment of anatomical problems and physical manipulation of them. In chronic and severe but difficult-to-treat problems such as lower back pain, such approaches have shown effectiveness equivalent to more invasive surgical treatment in many cases. They may also help to reduce stress related to pain and disability. These treatment approaches are based upon the idea that one way that stress is experienced and can be treated is in the body. Psychological tension can manifest itself in physical tension and misalignment or dysfunction of parts of the body. The goal is to restore more normal structure–function connections in the body. Realignment physical movement is believed to reduce adverse effects of psychological as well physical stress.

Mindfulness-Based Stress Reduction

Similarly, other techniques involving breathing exercises, self-monitoring and regulation, and meditation, several based on adaptations of Buddhist meditation practices, have proven quite effective in stress management in the medical setting. Jon Kabat-Zinn has developed a model that involves teaching patients with pain, anxiety, and other illness-related problems a mindfulness meditation practice designed to help them

1. control their attentional processes
2. induce physical comfort and relaxation
3. maintain a present orientation
4. enhance empathic connection with others

This method helps patients learn to experience more choice in the management of their perceived stress and to reduce anxiety about the future by more richly immersing themselves in the range of their present experiences, both positive and negative. This helps them accept but put into perspective worries about the future by maintaining enjoyment of the present.

Research with this technique has produced promising results. Kabat-Zinn and colleagues demonstrated in a randomized trial that psoriasis patients who listen to an audiotape teaching mindfulness-based stress reduction had significantly more rapid healing of their skin lesions than control patients who listened to music. Thus, mindfulness training may influence the course of disease as well as reducing stress-related anxiety and pain.

Relaxation Training

Other techniques based upon meditative traditions have also been introduced for stress reduction. Benson

and colleagues have popularized an approach that involves brief exercises designed to help patients manage daily stress and have termed the approach the relaxation response. Simple, brief relaxation exercises have been helpful in managing such problems as mild to moderate hypertension and are recommended in such situations prior to or in conjunction with antihypertensive medication treatment.

Conclusion

There is little doubt that Americans are voting with their pocketbooks and their feet for integrative approaches. Recent research has shown that some of these approaches produce specific therapeutic benefits, while others clearly do not. Even if relief of symptoms is a small but real effect, when the big picture is overwhelming, small amounts of mental and/or physical control can go a long way toward reassuring patients that they can more effectively manage the course of their life with disease. Stress is a normal part of life, and a natural if unwelcome part of the course of disease. The extent to which stress produces distress can be modified by a variety of treatments that engage patients as participants in their care. While results vary, these approaches should be viewed as an opportunity rather than a threat, a potentially useful addition to the management of stress in health care.

See Also the Following Articles

Disease, Stress Induced; Hypnosis; Meditation and Stress; Relaxation Techniques.

Further Reading

- Doan, B. (1998). *Alternative and complementary therapies*. New York: Oxford University Press.
- Eisenberg, D. M., Davis, R. B., et al. (1998). Trends in alternative medicine use in the United States, 1990–1997: results of a follow-up national survey. *Journal of the American Medical Association* 280, 1569–1575.
- Gore-Felton, C., Vosvick, M., et al. (2003). Alternative therapies: a common practice among men and women living with HIV. *Journal of the Association of Nurses in AIDS Care* 14(3), 17–27.
- Lang, E. V., Benotsch, E. G., et al. (2000). Adjunctive non-pharmacological analgesia for invasive medical procedures: a randomised trial. *The Lancet* 355, 1486–1490.
- Luskin, F. M., Newell, K. A., et al. (1998). A review of mind-body therapies in the treatment of cardiovascular disease. Part 1: implications for the elderly. *Alternative Therapies in Health and Medicine* 4(3), 46–61.
- Newell, S. A., Sanson-Fisher, R. W., et al. (2002). Systematic review of psychological therapies for cancer patients: overview and recommendations for future research. *Journal of the National Cancer Institute* 94(8), 558–584.
- Spiegel, H. and Spiegel, D. (2004). *Trance and treatment: clinical uses of hypnosis*. Washington, D.C.: American Psychiatric Publishing.
- Vincent, C. and Furnham, A. (1996). Why do patients turn to complementary medicine? An empirical study. *British Journal of Clinical Psychology* 35, 37–48.
- Wetzel, M. S., Eisenberg, D. M., et al. (1998). Courses involving complementary and alternative medicine at US medical schools. *Journal of the American Medical Association* 280(9), 784–787.
- Zeltzer, L. K., Tsao, J. C., et al. (2002). A phase I study on the feasibility and acceptability of an acupuncture/hypnosis intervention for chronic pediatric pain. *Journal of Pain and Symptom Management* 24(4), 437–446.

Interactions Between Stress and Drugs of Abuse

P V Piazza and M Le Moal

Université de Bordeaux, Bordeaux, France

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 586–594, © 2000, Elsevier Inc.

Introduction

Influence of Stressors on Drug Self-Administration

Factors Influencing Effects of Stressors

Mechanisms of Stress Action

Does Stress Induce a Drug-Prone Phenotype?

Conclusions

Glossary

Dopamine

A neurotransmitter that, with norepinephrine and serotonin, belongs to the class of monoamines. In the central nervous system there are two principal groups of dopaminergic neurons, one in the mesencephalon projecting to a large number of brain structures and the other in the hypothalamus, controlling hormonal function in the hypophysis. Mesencephalic dopaminergic neurons are divided into two major subgroups: (1) The nigrostriatal neurons (A9) that have cell bodies in the substantia nigra-

pars compacta (SNc) and that principally project to the dorsal striatum and the caudate nucleus and (2) the meso-limbic-cortical neurons (A10) with cell bodies in the ventral tegmental area (VTA) and that project to the ventral striatum, nucleus accumbens, and cortical and limbic structures, such as the amygdala and the hippocampus.

Drug self-administration

Intravenous drug self-administration is one of the most used models for studying drug abuse in animals. This test measures the learning of an operant response, such as pressing a lever, which stimulates drug infusion, i.e., this test measures the capacity of pharmacological substances to act as positive reinforcers.

Glucocorticoids

Consisting of cortisol in humans and corticosterone in rodents, this is the final step of the activation of the hypothalamo-pituitary-adrenal axis. These hormones have large effects in the periphery, where they modify energy metabolism and the activity of the immune system. Glucocorticoids also act at the level of the central nervous system. These hormones cross the blood-brain barrier easily and bind to two types of specific intracellular receptors, which are hormone-activated transcription factors. Glucocorticoids also seem to have direct membrane effects, although the mechanism of this action remains unclear.

Negative reinforcer

A stimulus that induces a subject to respond in a manner that interrupts the response. Foot shock is the most common negative reinforcer used in animals. Stressors are usually stimuli that can also act as negative reinforcers.

Positive reinforcer

A stimulus that increases the probability of a response. For example, positive reinforcers are food, a receptive sexual partner, and drugs of abuse.

Introduction

This article analyzes the interactions between stress and the responses to drugs of abuse. In particular, it focuses on intravenous drug self-administration, one of the principal experimental models of drug abuse. Much of our information on the interaction between stress and drugs of abuse is based on experimental studies and in particular from research in rats. Although several authors have proposed that stress is related positively to abuse of both opioid and psychostimulant drugs, a firm causal link between stress and drug abuse cannot be established on the

basis of studies in humans. There are two main reasons for this. First, the variables under investigation, i.e., the history and the experiences of the subject, in humans cannot be manipulated directly under controlled experimental conditions, but only assessed indirectly on the basis of retrospective self-reports of stress. Second, a reliable measure of the outcome of stress on drug abuse implies that all individuals have equal access to drugs under identical environmental conditions. Again, this experimental setting is impossible to achieve in humans. As will be reviewed, this type of experimental investigation has revealed some of the biological factors that mediate stress-induced propensity to self-administer drugs of abuse.

Influence of Stressors on Drug Self-Administration

Definition of Stress

Stress is a very vague concept often used with different meanings that principally refer to a subjective state with a strong negative connotation. Because subjective states cannot be investigated directly in animals, experimental research on stress has focused principally on the behavioral and biological responses to the forced exposure to stimuli or situations that are normally avoided by the individual. This is why stimuli used as experimental stressors are very often identical to aversive stimuli used as negative reinforcers in learning tasks. Consequently, stress, as studied in animals, probably corresponds to "the internal status induced by the exposure to threatening and aversive stimuli." However, what is principally studied in stress research is a panoply of responses induced by aversive stimuli.

Model of Intravenous Drug Self-Administration

The experimental study of drug abuse has been made possible by the discovery that animals self-administer drugs intravenously. This behavior shares many similarities with drug use in humans. First, animals self-administer practically all the drugs that induce addiction in humans. Second, the patterns of self-administration of the different drugs in humans and animals are similar. Finally, the large individual differences in the response to drugs of abuse that characterize humans are also found in animals.

Self-administration measures positive reinforcing effect of drugs, i.e., the learning of an operant response, such as pressing a lever, that has as a consequence the delivery of a drug infusion. In general, changes in the rate of the response reinforced by the

drug are considered to reflect changes in the sensitivity to its addictive properties.

Self-administration is usually studied using three different and complementary approaches: acquisition, retention, and reinstatement. The influence of different forms of stress on the three aspects of self-administration will be analyzed separately.

Influence of Stress on the Acquisition of Self-Administration

Acquisition studies evaluate the propensity of an individual to develop drug self-administration. For this purpose, the effects of first contact with low doses of the drug are studied, and parameters such as threshold or rate of acquisition are taken into account.

In adult animals, artificial and physical stressors, such as repeated tailpinch (Figure 1, top) and electric foot shock, facilitate the acquisition of the self-administration of psychostimulants, such as cocaine and amphetamine. Food restriction is another physical stressor that facilitates the acquisition of psychostimulant, opiate, and alcohol self-administration.

The acquisition of psychostimulant self-administration is also enhanced in male and female rats exposed to an aggressive biological 'congener' and in male rats raised in colonies in which there is a high level of social competition, i.e., colonies in which male rats fight to establish and maintain the social hierarchy that will determine access to females. Facilitation of the acquisition of morphine and alcohol oral self-administration, as well as of heroin intravenous self-administration is also induced by social isolation.

Early life events, such as prenatal stress, also increase the propensity to develop amphetamine self-administration. Such an effect has been observed in adult rats whose mothers had been submitted to a restraint procedure during the third and fourth weeks of gestation.

Influence of Stress on Retention

Retention studies are performed after prolonged training with high doses of drug and allow estimation of changes in (i) the sensitivity to the reinforcing effects of drugs and (ii) the motivation to self-administer drugs. The sensitivity to drugs is assessed by performing dose-response curves. In this case, the dose of the drug per infusion is varied, usually between sessions, and the rate of responding is recorded. Motivation for drugs is evaluated principally by means of progressive ratio schedules. In this case, the dose is maintained constant and the ratio requirement (number of responses necessary to obtain

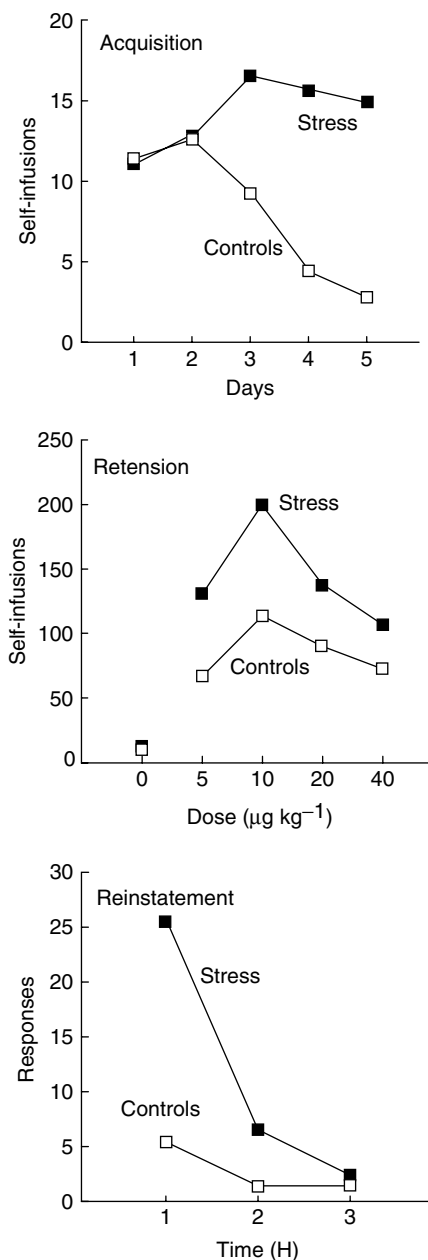


Figure 1 Examples of the effects of stress on intravenous self-administration. Top: effects of tail pinch on the acquisition of the self-administration of amphetamine ($10\mu\text{g}/\text{infusion}$). Center: effects of food restriction on the dose-response curves of etonitazene self-administration. Bottom: influence of electric foot shock on the reinstatement of heroin self-administration. Stress facilitates the acquisition of self-administration, induces an upward shift of the dose-response function, and triggers the reinstatement of drug-seeking behavior. Redrawn with permission from, respectively, Piazza, P. V., Deminiere, J. M., Le Moal, M., et al. (1990). Stress- and pharmacologically-induced behavioral sensitization increases vulnerability to acquisition of amphetamine self-administration. *Brain Research* **514**, 22–26; Carrol, M. E. and Meich, R. A. (1984). Increased drug-reinforced behavior due to food deprivation. *Advances in Behavioral Pharmacology* **4**, 47–88; and Shaham, Y. and Stewart J. (1995). Stress reinstates heroin-seeking in drug-free animals: an effect mimicking heroin, not withdrawal. *Psychopharmacology* **111**, 334–341.

one infusion) is increased progressively. The highest ratio reached by the subjects (breaking point) is considered an index of its motivation to self-administer drugs.

Changes in dose–response functions for psychostimulant and/or opiate self-administration have been observed after social stress, social isolation, and food restriction (Figure 1, center). In all these cases, the stressors induce an increase in the rate of response over a large range of doses, resulting in a vertical upward shift of the dose–response function. This type of shift in dose–response curves suggests an increase in the reinforcing efficacy of drugs in stressed animals. This idea is confirmed by progressive ratio studies. Thus, over a large range of doses, the breaking point for heroin self-administration has been consistently higher for animals subjected to repeated electric foot shock than for control rats.

A large variability has been observed in the effect of social isolation. Indeed, higher, slightly higher, equal, or lower sensitivities to cocaine have all been described in socially isolated rats as compared to grouped rats. The conditions that were considered as controls could also contribute to such variability. In fact, the rate of self-administration of control animals, and not the one of isolated rats, is what really differentiates some of these reports. Not surprisingly, the effects of social isolation appeared weaker when other conditions that facilitate self-administration, such as priming injections of the drug and pretest periods of food restriction, were used concomitantly.

Influence of Stress on Reinstatement

Reinstatement studies are considered a measure of relapse in drug intake. In this case, the behavior is studied after the extinction of previously acquired self-administration. Extinction is usually obtained by substituting the drug with a vehicle solution. Following extinction, the administration of stimuli known to induce craving for the drug in abstinent human addicts, such as low doses of the drug, will reinstate responses to the stimulus that elicited the infusion of the drug during acquisition. The rate of response to this stimulus is the principal measure of reinstatement.

Stressors increase reinstatement of drug self-administration. In rats in which the response to heroin (Figure 1, bottom) or cocaine has been extinguished by substituting the drug with a saline solution, repetitive electric foot shock induces the reinstatement of response to the drug. Stress-induced reinstatement seems to be a very robust and well-documented phenomenon, and it is of comparable intensity to the one induced by other experimental manipulations known

to induce reinstatement, such as the infusion of low doses of the self-administered drugs.

Factors Influencing Effects of Stressors

Predictability and Intensity

Predictability and intensity seem important variables in mediating the effects of stress on drug self-administration. Unpredictable stressors have higher effects than predictable ones. Furthermore, stressors, especially when acute, need to be of a certain intensity in order to increase self-administration. The relationship between the intensity of the stressor and its outcome on drug seeking is also shown by reinstatement studies. Thus, the rate of reinstatement has been found to be a function of the duration of the shocks. However, in the case of prolonged exposure to stress, the duration of the stress seems less important. Similar facilitatory effects on self-administration have been found for stressors that are continuous and for those that are of short duration and administered repeatedly.

Physical Versus Psychological Threats

Although ‘physical’ stressors involving a noxious stimulus such as electric foot shock and tail pinch increase self-administration, physical aggression per se does not seem to be a necessary condition for mediating stress effects. In fact, pure psychological stress is also able to increase drug self-administration. For example, it has been shown that witnessing another rat receiving foot shocks facilitates the acquisition of cocaine self-administration. Similarly, in social aggression experiments, the facilitation of self-administration is also found when the intruder is always protected by a screen grid and consequently never submitted to physical attacks.

Time Contingencies

Both acute and repeated exposure to stress can increase the propensity to develop drug self-administration. However, acute and chronic stress are influenced differently by temporal contingencies. In the case of acute stressors, a short interval between the stressful event and the exposure to the drug appears to be crucial. In contrast, after a prolonged exposure to stress, the interval between the termination of the stress and the exposure to the drug does not seem to be a relevant variable. In this case a facilitation of drug self-administration has also been found when the stress is terminated weeks before the start of the self-administration session. These observations are important from a pathophysiological point of view.

They indicate that repeated stress can induce long-lasting modifications that result in a drug-prone state that becomes independent from the actual presence of the stressor.

Mechanisms of Stress Action

Different mechanisms could mediate the effects of stress on drug self-administration and much remains to be done on this field of research. We will focus principally on the possibility that stressors sensitize the endogenous system of reward by an action on dopaminergic neurons and glucocorticoid hormones. Mesencephalic dopaminergic neurons, particularly their projection to the nucleus accumbens, are considered one of the principal substrates of drug-reinforcing effects. Glucocorticoid hormones are the final step of the activation of the hypothalamic-pituitary-adrenal axis. The secretion of glucocorticoids by the adrenal gland is considered one of the principal biological adaptations to external aggression.

Clearly, an interaction between glucocorticoids and dopamine should not be considered as the only possible mechanism mediating stress effects, but as the best studied of only one of the likely possibilities. **Figure 2** shows schematically the interactions among stress, glucocorticoids, and dopaminergic neurons.

Sensitization of the Reward System

One of the mechanisms by which stress could modify the sensitivity of an individual to drugs of abuse is by modifying the activity of those neurobiological systems that mediate the addictive properties of drugs.

Role of mesencephalic dopaminergic neurons Mesencephalic dopaminergic neurons seem to be a likely candidate for such a neurobiological system. The role of these neurons in drug abuse has been emphasized in several good reviews. In particular, the release of dopamine in the nucleus accumbens is considered one of the major mechanisms of the addictive properties of drugs. Practically all drugs of abuse increase

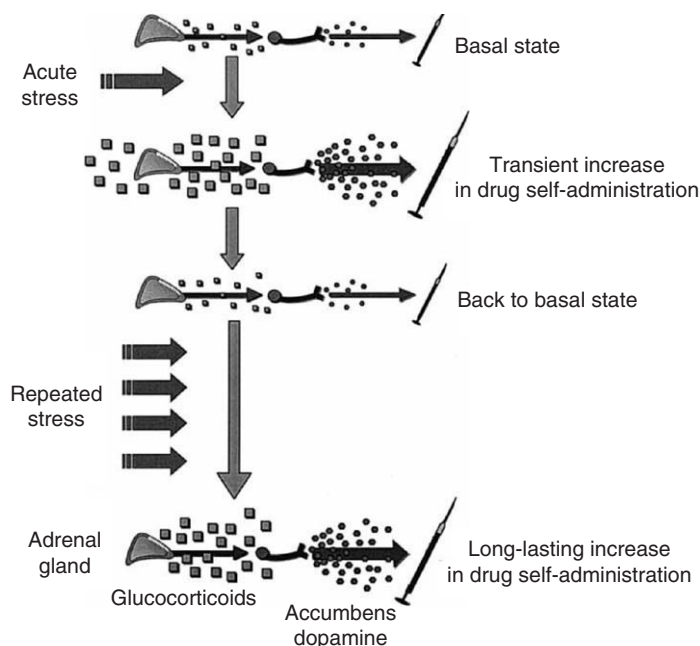


Figure 2 Possible pathophysiological mechanisms of the increase in drug self-administration induced by acute and repeated stress. The concentrations of glucocorticoids (large squares) determine the level of dopamine release in the nucleus accumbens (small squares). In basal conditions (basal state), glucocorticoid secretion and dopamine release are low, as is sensitivity to drugs of abuse. Acute stress increases glucocorticoid secretion, which, by enhancing the release of dopamine, results in an increase in the sensitivity to the reinforcing effects of drugs of abuse. However, activation of the negative feedback that controls the secretion of these hormones brings the system back to basal levels within 2 h. The increase in the concentrations of glucocorticoids induced by repeated exposure to stress will progressively result in a long-lasting increase in the secretion of these hormones and in the release of dopamine in the nucleus accumbens. These changes will, in turn, determine a long-lasting increase in the sensitivity to the reinforcing effects of drugs of abuse. The transient increase in glucocorticoids and dopamine observed after acute stress may explain why, in this case, an increase in drug self-administration is found only if the exposure to drugs closely follows the stressor. The long-lasting increase in the activity of these two biological factors could explain why, after repeated stress, an increase in the sensitivity to drugs is also found long (weeks) after the end of the stressor.

dopamine release in this brain area, and an increase or a decrease in dopaminergic activity in the nucleus accumbens corresponds, respectively, to an increase or a decrease in the reinforcing properties of drugs. Several types of acute stress transiently increase dopamine release in the nucleus accumbens. In parallel, repeated stress induces a long-lasting increase in the release of this neurotransmitter, particularly of drug-induced dopamine release.

Role of glucocorticoid hormones Glucocorticoid hormones seem to be the link between environmental experiences and mesencephalic dopaminergic neurons. In basal conditions, the administration of glucocorticoids at levels that are in the stress range can increase dopamine release in the nucleus accumbens, whereas the suppression of glucocorticoid secretion has the opposite effect. During stress, a selective blockade of stress-induced secretion of glucocorticoids reduces dopamine release by around 50%. Finally, the development of the long-lasting sensitization of the dopaminergic response to psychostimulants and opioids induced by stress is suppressed by the blockade of corticosterone secretion.

That glucocorticoids hormones, through their action on dopamine release, could increase the behavioral responses to drugs is supported by a supplementary set of data. First, administration of glucocorticoids, at levels that are in the stress range, increases the propensity to self-administer amphetamine. Second, the blockade of stress-induced corticosterone secretion induces a downward shift in the dose–response curve for cocaine self-administration, i.e., changes opposite to those observed after stress. Third, the blockade of corticosterone secretion can reverse the stress-induced sensitization of the dopaminergic and behavioral responses to cocaine.

Mechanisms of reinstatement The hypothalamic-pituitary-adrenal axis and dopaminergic neurons also seem to play a role in reinstatement, but in a way that differs from the one described for drug self-administration.

For example, a concomitant blockade of D1 and D2 dopaminergic receptors is needed to block stress-induced reinstatement, whereas selective antagonists of either receptor suffice to decrease the positive reinforcing effects of drugs. In parallel, although the injection of glucocorticoids alone can induce reinstatement, blockade of the secretion of these hormones does not prevent the stress-induced reinstatement of heroin self-administration. In contrast, the blockade of glucocorticoid secretion seems to block the stress-induced reinstatement of cocaine self-administration.

Stress-Induced Impulsivity

Changes in impulsivity could be an alternative mechanism by which stress modifies drug self-administration. A deficit in the inhibitory processes that normally operate to control the reward effect of a stimulus is considered an important dimension of impulsivity. This type of impairment could contribute to the increase in drug self-administration induced by stress. This idea is prompted by the effects of social aggression on the rate of responding during time-out periods, i.e., fixed time periods that follow each drug infusion and during which responses are not expected. During time-out, control animals learn quickly to reduce response rates and to wait for the next period of drug availability, whereas stressed rats maintain high rates of response during these time periods. The possibility that an inhibitory deficit is involved in excessive drug consumption is also backed by the finding that the inability to delay gratification predicts alcohol self-administration.

Although stress-induced impulsivity seems an interesting mechanism, further experiments should specifically test the actual occurrence of such a change in behavior. At the neurochemical level it will be important to correlate such potential changes in behaviors with neurobiological modifications involved in impulsivity, such as a decrease in serotonergic activity.

Does Stress Induce a Drug-Prone Phenotype?

The Individual-Centered Vision of Addiction

It is a common observation that although a large number of people take drugs for variable periods of time, only few of them develop a true addiction, i.e., a compulsive drug use that becomes the main goal-directed activity of the subject. One of the theories that has been used to explain such individual differences in the development of addiction postulates the existence of a drug-prone phenotype, i.e., drug abuse is a pre-existing pathological condition revealed by the drug. Addiction would appear only in certain individuals because their biological phenotype will generate a ‘pathological response’ to drugs. This pathological response would make the addictive properties of drugs much greater in some subjects, increasing their likeliness to develop drug abuse.

Knowledge about Drug-Prone Phenotypes

Research in rodents has identified some of the biological factors that could determine a drug-prone

phenotype. As indicated earlier, large individual differences in the propensity to develop drug self-administration have been described in laboratory rats, and the propensity of an individual animal to develop self-administration correlates with its behavior in other experimental situations. For example, animals that show the highest locomotor activity when exposed to a novel environment, a procedure that can be seen as a model of mild stress, also show the highest propensity to self-administer amphetamine.

Animals predisposed to drug addiction show a faster acquisition of drug self-administration, an upward shift of the dose-response function, and reach higher ratio requirements. They also show a higher release of dopamine into the nucleus accumbens in response to a drug challenge and a longer glucocorticoid secretion in response to stress. Finally, the higher dopaminergic activity of animals that are spontaneously drug prone seems to be glucocorticoid dependent. These rats are more sensitive to the dopaminergic effect of glucocorticoids, and blockade of glucocorticoid reduces their dopaminergic hyperactivity.

Stress-Induced Phenotype

Data presented in this article suggest that stressful experiences are one of the determinants of a drug-prone phenotype. Indeed, stressors can induce a long-lasting increase in drug self-administration, and such an effect occur place at an early stage of life, as exemplified by prenatal stress. Further support for this idea arises from the comparison of the characteristics of individuals who are spontaneously drug prone with those of subjects in which such a predisposition has been induced by stress. Indeed, dopaminergic hyperactivity and a dysregulation of glucocorticoid hormone activity, similar to that observed in subjects susceptible to drug addiction, have also been found in individuals made vulnerable by stress.

Conclusions

In conclusion, stress experiences increase the propensity of an individual to develop drug self-administration, inducing a drug-prone phenotype. Stress-induced increases in the secretion of glucocorticoids, which in turn enhance the drug-induced

release of dopamine in the nucleus accumbens, seem to be an important substrate of such an effect of stress.

These observations highlight the importance of environmental experiences in the etiology of drug abuse. This confirms the hypothetical role of stress suggested by correlative studies in humans. Second, the observations support the idea that drug abuse is not simply induced by chronic drug intake, but that the phenotype of the individual plays an important role in determining the development of such a pathological behavior.

See Also the Following Article

Drug Use and Abuse.

Further Reading

- Carrol, M. E. and Meich, R. A. (1984). Increased drug-reinforced behavior due to food deprivation. *Advances in Behavioral Pharmacology* 4, 47–88.
- de Kloet, E. R. (1992). Brain corticosteroid receptors balance and homeostatic control. *Frontiers of Neuroendocrinology* 12, 95–164.
- Kalivas, P. W. and Stewart, J. (1991). Dopamine transmission in the initiation and expression of drug- and stress-induced sensitization of motor activity. *Brain Research Reviews* 16, 223–244.
- Piazza, P. V. and Le Moal, M. (1996). Pathophysiological basis of vulnerability to drug abuse: Role of an interaction between stress, glucocorticoids and dopaminergic neurons. *Annual Review of Pharmacology and Toxicology* 36, 359–378.
- Piazza, P. V. and Le Moal, M. (1997). Glucocorticoids as an endogenous substrate of reward. *Brain Research Reviews* 25, 359–372.
- Piazza, P. V. and Le Moal, M. (1998). The role of stress in drug self-administration. *Trends in Pharmacological Sciences* 19, 7–74.
- Schuster, C. R. and Thompson, T. (1969). Self-administration and behavioral dependence on drugs. *Annual Review of Pharmacology* 9, 483–502.
- Shaham, Y. (1996). Effect of stress an opioid-seeking behavior: Evidences from studies with rats. *Annals of Behavioral Medicine* 18, 255–263.
- Wise, R. A. (1996). Neurobiology of addiction. *Current Opinion in Neurobiology* 6, 243–251.

Interferons and Interleukins

See: Cytokines; Immune Response; Multiple Sclerosis; Prostaglandins; Chronic Social Stress: GR Sensitivity in Leukocytes.

J

Job Insecurity: The Health Effects of a Psychosocial Work Stressor

J E Ferrie

University College London Medical School, London, UK

P Martikainen

University of Helsinki, Helsinki, Finland

© 2007 Elsevier Inc. All rights reserved.

Job Insecurity and the Labor Market

The Job Insecurity Concept

Mechanisms

Psychological Health

General Measures of Physical Health

Health Outcomes of Interest to the Organization

Effects beyond the Individual and the Workplace

Who Experiences Job Insecurity: The Public Health Impact

Summary

Glossary

Downsizing

The most common name given to the process whereby an organization reduces its workforce through natural wastage, early retirement, or redundancy.

Employment security

While job security represents the ability to remain in a particular job, employment security represents the likelihood of being able to remain in paid employment, even if this is a succession of jobs.

Job insecurity

In general terms, the discrepancy between the level of security a person experiences and the level he or she prefers. It can cover both the loss of any valued condition of employment and the threat of total job loss. The depth of the job insecurity experience is dependent on the perceived probability and perceived severity of the job loss. Job insecurity has a large subjective appraisal element that is highly context dependent,

and the job insecurity experience may effect employees who do not lose their job as much as those who do. Job insecurity arising from the threat to a particular job may lead to loss of employment security if subsequent jobs prove hard to find. Job insecurity can be self-reported or externally attributed. The population in studies of self-reported job insecurity is composed of individuals who report their job as insecure. In studies of attributed job insecurity, the study population is deemed to be under threat of job loss by the researchers. Associations observed in studies of attributed job insecurity are likely to be underestimates, as the study population will contain respondents who do not perceive themselves to be under threat.

Primary labor market

The highly regulated section of the labor market in which employees have traditionally enjoyed long-term, secure employment with many benefits, such as occupational pensions, maternity cover, etc.

Psychological contract

Worker's beliefs regarding the terms and conditions of her or his employment. Fairness and good faith are implied in the psychological contract, and breaches have implications for the worker's trust in the organization, performance, and behavior.

Secondary labor market

That section of the labor market in which wages are low and employment rights are few. Such employment is often seasonal, part-time, or temporary and is frequently used to buffer short-term changes in labor requirements.

Sickness presenteeism

Describes the phenomenon that occurs when an employee, despite a level of ill health that should prompt rest and absence, still turns up at work.

Job Insecurity and the Labor Market

Over the past 25 years, large changes have taken place in the structure of the labor market in most industrialized countries. The major determinants of this restructuring have been deindustrialization, technological innovation, globalization, and a commitment to a free market economy, including the privatization of public services. There has been a marked shift from economies driven by manufacturing toward those dominated by employment in services. Technological innovation has allowed many routine functions largely to be replaced by electronic devices (such as cash dispensers), and global information systems have facilitated the rapid transfer of work to newly industrializing countries where labor is cheaper. This deindustrialization and technological change have generally been accompanied by the expansion and strengthening of the role of market forces. Many workers, who previously thought that they had a job for life, now feel much less secure, and job insecurity has become more widespread in all the countries belonging to the Organisation for Economic Co-operation and Development (OECD) since the mid-1980s. Like all social transformations, these changes have the potential to affect the health of individuals and populations.

The Job Insecurity Concept

Research during the major recessions of the 1930s and 1980s provided unequivocal evidence of the adverse effects of unemployment on health, particularly mental health. However, although it is a major determinant of unemployment, much less work has been done on job insecurity. This is partly because job insecurity did not excite attention until it affected white-collar workers in the 1990s and partly because it is less obvious that job insecurity could affect health, as, unlike unemployment, it involves no actual loss of income or status. However, some research has suggested that job insecurity may have detrimental consequences equal to those of job loss. This is consistent with the central proposition in stress research that anticipation of a stressful event provokes as much, if not more, anxiety than the event itself. For workers accustomed to long-term, secure employment, the threat of job loss involves a breach of the psychological contract and generates uncertainty, which, in addition to emotional reactions, can trigger changes in other psychosocial stressors, such as job involvement. The level of stress generated by job insecurity depends on the perceived probability and perceived severity of the threatened job loss. As job insecurity in the primary labor market has increased

and research on psychosocial stressors has gained credence and prominence, interest in the effects of job insecurity has also developed.

Job security is a vital concern for both employees and their organizations. The general concept has been defined by Jean Hartley and colleagues as the discrepancy between the level of job security a person experiences and the level he or she prefers. Job insecurity can also be generated by deteriorating employment conditions and career opportunities, so some investigators have extended the concept to include loss of any valued condition of employment. These definitions encompass large numbers of workers who have insecure jobs, often seasonal, part-time, or temporary. For workers in this secondary labor market, job security is not part of the psychological contract and insecurity is an integral part of their work experience. However, widespread downsizing, privatization, outsourcing, mergers, and closures have led to the suggestion that job insecurity, even for employees in the primary labor market, is no longer a temporary break in an otherwise predictable work life pattern, but is becoming a structural feature of the new labor market. This new employment insecurity brings employees in the primary labor market closer to those in the secondary labor market.

The difficulty in studying the health effects of job insecurity among workers in the secondary labor market is working out whether poor health outcomes can be attributed to job insecurity and unemployment, or whether it is those already in poorer health who are selected into the secondary labor market. For this reason much of the published research has concentrated on workers in the primary labor market.

Job insecurity can be self-reported or externally attributed. Studies of self-reported job insecurity examine workers who say their job is insecure, while studies of attributed job insecurity examine workers deemed to be under threat of job loss because downsizings or workplace closures are expected or are taking place. There is a strong association between self-reported and attributed job insecurity. However, as stress levels are determined by the perceived probability and perceived severity of the job loss, self-reported job insecurity is generally considered to be the more potent stressor.

Mechanisms

Work examining potential explanations of any association between job insecurity and health remains patchy. Studies to date have implicated at least the following: pre-existing ill health; personal characteristics, such as pessimism; material factors, such as financial hardship; non-work psychosocial stressors,

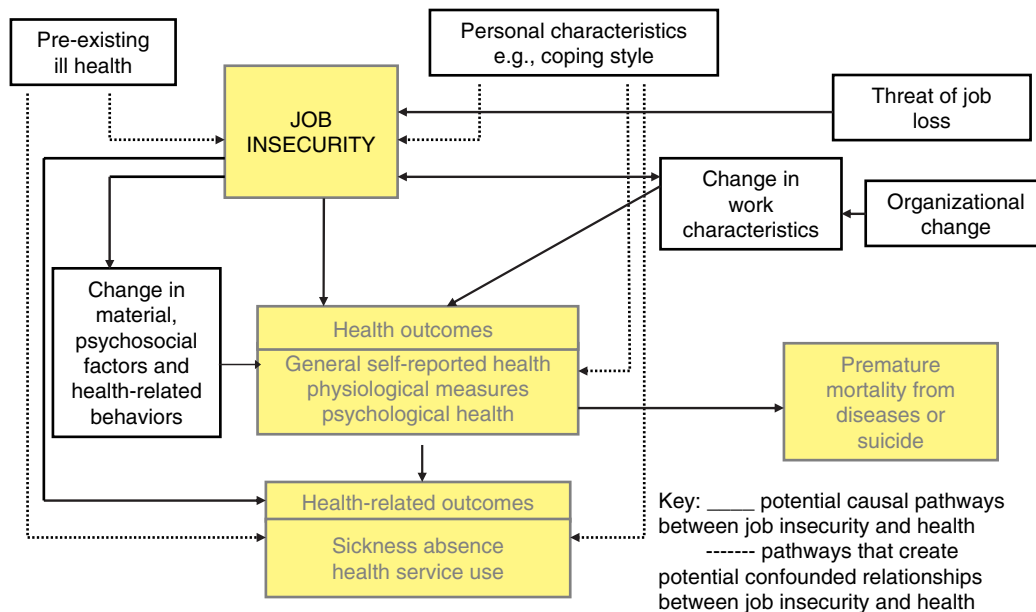


Figure 1 Possible pathways in the association between job insecurity and morbidity. Adapted from Ferrie, J. E., Shipley, M. S., Marmot, M. G., Martikainen, P., Stansfeld, S. and Davey Smith, G. (2001). Job insecurity in white-collar workers: towards an explanation of associations with health. *Journal of Occupational Health and Psychology* 6, 26–42.

such as life events; other psychosocial and physical characteristics of the work environment, such as decreased job involvement; and health-related behavior (Figure 1).

The personal characteristics of pessimism, vigilance, neuroticism, mastery, and self-esteem have been found to explain part of the association between self-reported job insecurity and ill health and health-related outcomes. One of the main explanations of associations between unemployment and health is loss of income. Although job insecurity involves no loss of income, pre-existing deprivation and fear of financial insecurity have been shown to explain part of the association between self-reported job insecurity and self-reported psychological and physical health outcomes. Much has also been made of the potential for job insecurity to interact with other material and psychosocial stressors outside work and to precipitate increases in smoking and drinking.

It seems obvious that organizational change, which involves a reduction in the size of the workforce, in addition to generating job insecurity, is likely to result in changes to other characteristics of the work environment. This has been examined in a number of studies on changes in other psychosocial work characteristics, in particular, loss of job satisfaction, job control, and social support at work. Other work has shown a synergistic adverse effect on health in workers with both high job insecurity and high job strain. Despite this work, much of the association between job insecurity and ill health remains unexplained and

may be due either to direct effects of stress on the autoimmune system or to unobserved health-related selection.

Psychological Health

Most studies that have examined the effects of self-reported job insecurity on health have looked at psychological ill health as an outcome, often as the only outcome. Every study has documented consistent adverse effects on all measures of psychological health, although a combined analysis of the association between job insecurity and mental health outcomes in 37 study samples has estimated the effect size for self-reported psychological ill health to be medium rather than large, with a standardized effect of approximately 25%.

Good evidence for an association between job insecurity and poor psychological health has come from the observation of a dose-response relationship; that is, the higher the level of job insecurity, the greater the increase in psychological ill health. However, a dose-response relationship would be observed whether job insecurity increased psychological ill health or whether ill health increased the likelihood of reporting of job insecurity. In addition, self-reporting of both exposure (job insecurity) and outcome (psychological health) could contaminate the results if some participants have a tendency to report their situation, both work and health, in a negative light. Stronger evidence for the causal precedence of job insecurity

comes from the study of change in job security. For example, it has been shown that, compared to employees whose jobs remain secure, psychological ill health increases among those who report a loss of job security. Studies that follow workers over time also show that self-reported job insecurity acts as a chronic stressor. This means that workers exposed to ongoing job insecurity experience the highest levels of psychological ill health.

Almost all studies of attributed job insecurity have documented an association between insecurity and psychological ill health. Findings from the most notable exception, an early factory closure study in the United States, surprised the investigators. "In the psychological sphere the personal anguish experienced by the men and their families does not seem adequately documented by the statistics" (Cobb, S and Kasl, S. V. (1977). *Termination: the consequences of job loss*. Cincinnati: DHEW-NIOSH Publication no. 77-224, National Institutes for Occupational Safety and Health, p. 180). The observation was attributed to imperfect measurement techniques, as numerous adverse psychological outcomes in individuals were documented in an eloquent narrative account of the closing. Another notable exception comes from a longitudinal study of white-collar civil servants, which took advantage of a natural experiment. In such experiments, a naturally exposed group is compared to an unexposed reference group. This means that there is no plausible selection process involved, and factors such as poor health or age are not influential in increasing the risk of job insecurity. Inferences about causality are strong under such conditions. In this particular natural experiment, well after the initial survey, one of the 20 departments in the study was unexpectedly sold to the private sector, a transfer of business in which most of the workforce lost their jobs. However, no effects on psychological health were seen in either sex in the run-up to the sale. A possible explanation of these seemingly contradictory findings recently emerged from a study that took advantage of a similar natural experiment in the Netherlands. One of the government agencies in the study was threatened with closure after the initial survey. Psychological ill health among employees in this agency was compared with that for employees in agencies unaffected by threat of closure, adjusting for ill health at baseline. Over the 13 months following the closure announcement, the number of new cases of psychological ill health in the closure group was significantly higher than in the nonclosure group. However, the increase in psychological distress among those working in the agency targeted, but who did not think they were at risk of job loss, was not significant; thus, self-reported rather

than self-attributed job insecurity appeared to have been driving the association.

General Measures of Physical Health

Evidence is growing that self-reported job insecurity has adverse effects on self-reported general physical health measures, independent of effects on psychological health, although a combined analysis of the association in 19 study samples estimated the effect size for self-reported general physical health measures to be small, with a standardized effect of approximately 16%.

Reasonably consistent findings have been obtained for a number of health outcomes in both cross-sectional and longitudinal studies. Self-reported job insecurity has been shown to be associated with poor self-rated health, sleep disturbance, chronic insomnia, fatigue, migraine, colds and flu-like symptoms, longstanding illness, and musculoskeletal disorders. Evidence of a dose-response relationship and evidence from studies of change have provided more convincing evidence of a causal association, and there is evidence that self-reported job insecurity acts as a chronic stressor in relation to physical health. The small number of studies that have been able to examine associations between attributed job insecurity and self-reported measures of physical health have observed associations similar to those for self-reported job insecurity. However, there appears to be no evidence that attributed job insecurity acts as a chronic stressor, so it is possible that the finding for self-reported job insecurity may be due to a general tendency for negative reporting.

Very few studies have looked at the effect of either category of job insecurity on biomedical risk factors. There appears to be only two studies that have examined the association between self-reported job insecurity and blood pressure. Both studies found a significant increase in blood pressure compared with control workers who did not report job insecurity. Two studies of attributed job insecurity similarly found significant increases in blood pressure, but another two studies that examined a change in attributed security observed either no significant differences or effects in some groups but not others. Studies that have examined effects on body weight similarly report mixed findings, with no change as well as gain and loss in mean weight being observed. Resolution of these apparently divergent findings may lie in the observation that job insecurity has bidirectional effects on weight. A study of occupational factors and 5-year weight change among Danish men showed that job insecurity increased the likelihood of weight gain among obese men and weight loss among men who were lean.

Apart from two Scandinavian studies, there has been even less interest in biomedical risk factors other than blood pressure and body weight. Attributed job insecurity has been associated with nonsignificant increases in the stress hormone, cortisol, and in cholesterol in a mainly female blue-collar workforce, and levels of adrenal hormones have been shown to be raised among women at the bottom of the white-collar employment hierarchy. However, evidence on cholesterol from a longitudinal study of white-collar civil servants is mixed, and it remains uncertain what clinical relevance the observed changes in biomedical risk factors may have. Hardly any work has examined the effect of job insecurity on diagnosed diseases. Three studies have documented an association between self-reported job insecurity and various measures of coronary heart disease, and an association has been observed between attributed job insecurity and risk of ischemia (abnormal ECG or diagnosed angina) relative to unexposed employees. However, overall the evidence is limited. While the findings were indicative of adverse changes, these were modest and not significant after adjustment for major confounding somatic and behavioral coronary risk factors.

In part because job insecurity was rarely studied until the recession of the 1990s, there has been little research on the effect of job insecurity on mortality from diseases or suicide. Studies of self-reported job insecurity have not been documented, and there is no compelling evidence yet of an association between attributed job insecurity and premature death. Most work prior to the 1990s was a spin-off of studies of unemployment and health and found no significant differences between groups of workers exposed and unexposed to attributed job insecurity. More recently, a factory closure study in New Zealand examined mortality at the end of an 8-year follow-up as an outcome under the conditions of a natural experiment. The study found no statistically significant difference in the death rate for any cause between workers in the closing plant and in the control plant. This was by far the largest factory closure study to date, with nearly 2000 in both the study group and the control group. However, the number of deaths after 8 years was still relatively small, and instability in the sector, which led to the closing of the control plant by the end of the follow-up, may mean that these findings are minimum likely estimates of effects. Although temporary workers are generally considered to be part of the secondary labor market, recent work using Finnish data has looked at the effect on mortality of transfer from temporary to permanent employment. Moving from temporary to permanent employment is associated with a significantly lower risk of death, while remaining in temporary

employment is associated with a significantly higher risk of death than being continuously in permanent employment. While moving the fittest temporary workers into permanent employment is likely to explain part of this effect, change in exposure to job insecurity may also contribute.

Consistent with the evidence of effects on psychological health, a review of eight studies of unemployment and suicide that used longitudinal data for individuals showed that in all but one there was greater job instability and occupational problems among suicides compared to nonsuicides. In the New Zealand factory closure study discussed previously, there was a significant, 2.5-fold increased risk of serious self-harm that led to hospitalization or death from suicide, and out of 46 men who lost their job during a U.S. factory closure, two committed suicide. Although this was 30 times the number expected, the number of people in the study population was too small to draw firm conclusions.

Health Outcomes of Interest to the Organization

A small number of studies have looked at the effect of self-reported job insecurity on safety at work. The main finding from this research is that job insecurity increases workplace injuries and accidents through detrimental effects on employee safety motivation and safety compliance. Both self-reported and attributed job insecurity have been associated with increases in short spells and medically certified spells of sickness absence from all causes, musculoskeletal disorders, and trauma. And a longitudinal study in Finland has shown that downsizing increases sickness absence more among low-income than high-income employees. In contrast to these findings, studies of British civil servants and Finnish workers on fixed-term contracts have shown that in some cases attributed job insecurity is associated with decreases in sickness absence, even when the workforce is known to have a higher level of ill health. This phenomenon is called sickness presenteeism and tends to occur when workers feel that being absent from work may increase their chance of redundancy. Unsurprisingly, sickness presenteeism appears to be more common among workers in the secondary labor market.

Job insecurity has also been widely studied in the organizational and managerial literature as a determinant of a number of outcomes that have adverse consequences for organizations, such as commitment, work effort, and productivity. Findings from the Netherlands have shown that not only is job insecurity a strong determinant of job change, but it is often

the most valuable individuals that are most inclined to seek other job alternatives.

Effects beyond the Individual and the Workplace

A downstream effect of the stress caused by self-reported and attributed job insecurity is an increase in health service use: general practitioner (GP) consultations, hospital referrals, hospital attendance, and use of prescribed medicines, in particular, tranquilizers and antidepressants. However, as for sickness absence, effects can go either way, so there is also evidence that those with high levels of job insecurity may forgo a medical consultation or neglect care of themselves for fear of missing work. An in-depth study of effects on health service use also showed that these effects were not restricted to the workers themselves. GP consultations, hospital referrals, and attendance also increased significantly in the workers' families. Other effects on the family include increases in work-family conflict and tension in the home. Evidence of this spill-over confirms results from studies of both categories of job insecurity that have reported adverse effects on family members, family relationships and the behavior of children.

Who Experiences Job Insecurity: The Public Health Impact

Job insecurity is not evenly distributed throughout the working population. Along with unemployment, it falls more heavily on those with fewer skills, those with less education, and those at a disadvantage in the labor market such as immigrants and the disabled. Several studies have now demonstrated a strong inverse association between job insecurity and socioeconomic position, measured as occupational grade, social class, or education. This means that those at the top of the social hierarchy, however measured, are less likely to experience job insecurity than those at the bottom. Such inverse gradients have been demonstrated in specific occupational groups and in general population samples. There is also evidence that the effect of job insecurity on health is highest among those in the lowest socioeconomic group.

Studies of large companies or government employees – contexts in which most studies on the health effects of job insecurity have been carried out – may underestimate the extent of job insecurity in the general working population. A survey of a large sample of the working population in Taiwan found that job insecurity affected 50% of workers, and a study in Switzerland in 1997 showed that 10% of the working population reported a high level of job insecurity.

These may be two extreme examples, but an OECD study showed that in 1996 experience of job insecurity was very common, with on average 67% of workers unable to agree strongly with the statement “my job is secure.” With the experience of job insecurity being this common, any detrimental effects of job insecurity on health can have major repercussion on the health of the working population as a whole.

Summary

Unequivocal evidence of a causal association between job insecurity and health would require longitudinal studies, with information from a period of secure employment for those subsequently exposed to job insecurity, in addition to a well-matched control group who remained in secure employment. Such prerequisites mean that opportunities for ideal studies are rare. Consequently, although research interest in the health consequences of job insecurity has increased, strong evidence of a causal link between job insecurity and many health outcomes remains limited. Despite this, there is now firm evidence of a causal association between both self-reported and attributed job insecurity and all measures of psychological ill health, and evidence of an association with a number of self-reported physical health outcomes.

Self-reported and attributed job insecurity are associated with adverse changes in most biomedical risk factors that have been examined, but effects are mostly nonsignificant. There is an indication that attributed job insecurity is associated with an increase in heart disease, but no evidence yet of an association between job insecurity and death from diseases or suicide. Job insecurity has been shown to increase workplace injuries and accidents through detrimental effects on employee safety motivation and safety compliance. Most studies show that attributed job insecurity is accompanied by an increase in self-certified and medically certified sickness absence. However, lower rates of sickness absence as well as sickness presenteeism among temporary workers and employees facing imminent threat have also been demonstrated. Both self-reported and attributed job insecurity result in increased use of health services and medications among both those directly affected and their families.

It has been proposed that the effect of job insecurity on health may be mediated via changes in health-related behaviors. However, there is little evidence to support or refute this hypothesis. Although some studies have found the effect of job insecurity on health to be partially explained by other work characteristics such as job satisfaction, social support at work, and job control, much of the association

between job insecurity and ill health remains unexplained and may be due either to direct effects of stress on the autoimmune system or to unobserved health-related selection.

Recent work has demonstrated a synergistic adverse effect on health in workers exposed to both high job insecurity and high job strain. In a context of downsizing and reliance on outsourced labor, perceptions of insecurity are likely to be widespread and, as fewer people do more, occur alongside increasing demands and perceptions of low control. If current trends continue, the number of employees exposed to both job insecurity and job strain will probably increase. Despite its limitations, existing research on the stress of job insecurity has much to contribute to policies that will improve the health of workers and their families and prevent adverse effects on the organizations that employ them.

See Also the Following Articles

Demand–Control Model; Effort–Reward Imbalance Model; Unemployment, Stress and Health; Workplace Stress.

Further Reading

- Bartley, M. and Ferrie, J. E. (2001). Unemployment, job insecurity and health – glossary. *Journal of Epidemiology and Community Health* 55, 776–781.
- Burchell, B., Day, D., Hudson, M., Ladipo, D., Mankelow, R., Nolan, J., et al. (1999). *Job insecurity and work intensification*. York: Joseph Rowntree Foundation.
- Ferrie, J. E., Marmot, M. G., Griffiths, J. and Ziglio, E. (eds.) (1999). *Labour market changes and job insecurity: a challenge for social welfare and health promotion*. European Series No. 81. Copenhagen: World Health Organisation.
- Hartley, J., Jacobson, D., Klandermans, B. and Van Vuuren, T. (1991). *Job insecurity: coping with jobs at risk*. London: Sage Publications Ltd.
- Quinlan, M., Mayhew, C. and Bohle, E. (2001). The global expansion of precarious employment, work disorganisation, and consequences for occupational health: a review of recent research. *International Journal of Health Services* 31, 335–414.
- Platt, S., Pavis, S. and Akram, G. (1998). *Changing labour market conditions and health: a systematic literature review (1993–1998)*. Dublin: European Foundation for the Improvement of Living and Working Conditions.

K

Kidney Function

W J Welch

Georgetown University, Washington, D.C., USA

© 2007 Elsevier Inc. All rights reserved.

Introduction

Renal Vascular Resistance (RVR)

Sodium Reabsorption

Does Chronic Stress Cause Hypertension by a Renal Mechanism?

Summary

Glossary

Renal sympathetic nerve activity Net activity (measured in Hz) of renal nerve bundles supplying the kidney.

Renal vascular resistance Vascular resistance of blood flow through the kidney, calculated as systemic mean arterial blood pressure factored by renal blood flow.

Sodium retention Excess reabsorption of Na^+ from various segments of the nephron, resulting in increased extracellular volume.

Introduction

The kidney's primary function is to maintain homeostasis by the efficient elimination of excess fluid and solute. This key function is regulated by both extrarenal and intrarenal mechanisms that control the kidney's ability to maintain proper fluid and electrolyte balance. The sympathetic nervous system (SNS) is an integral part of the homeostatic control system that maintains normal kidney function, with an extensive network of nerves identified in most renal tissues. The SNS modulates short-term regulation of renal function, and under normal conditions renal sympathetic nerve activity is low. During acute stress, however, elevated renal sympathetic activity depresses renal function. Chronic stress studies have been less conclusive and are difficult to design due to additional

pathologies induced by stress, such as hypertension and cardiovascular complications. Major renal effects of mental stress include increased renal vasoconstriction, causing greater renal vascular resistance (RVR), and increased Na^+ transport, causing greater Na^+ retention. Both of these effects may contribute to the development of hypertension, and some investigators propose that this may be the primary mechanism of human essential hypertension. Therefore, most studies on the renal effects of stress have focused on its relationship to hypertension. Though the effects of stress are more pronounced in hypertensive humans and animals, stress increases renal vasoconstriction and Na^+ retention in normal subjects and animals, as well.

Renal Vascular Resistance (RVR)

The systemic blood pressure determines the renal perfusion pressure (RPP) and is a major determinant of RVR. The kidney has a unique ability to autoregulate its own blood flow that is efficient over a wide range of RPPs. Renal blood flow (RBF) changes very little over a large range of RPPs, therefore maintaining a stable RVR. This is due to the exquisite control of vascular tone in the primary renal resistance vessels, the preglomerular afferent arterioles and the postglomerular efferent arterioles. However, many hormonal and neural disturbances can alter autoregulation, often targeting these arterioles, leading to inappropriate increases in RVR. Increased sympathetic activity increases arteriolar resistance and RVR. Increased RVR exacerbates hypertension via its contribution to overall increased total peripheral resistance (since RBF is 20–25% of cardiac output) and its ability to promote Na^+ reabsorption. Stress can induce both neural and hormonal changes that affect RVR.

Renal Sympathetic Nerve Activity (RSNA) and RVR

Multiple studies using several types of acute stress in rats have shown that RBF is reduced and blood

pressure (BP) and RVR over the period of stress are elevated. Renal sympathetic nerve activity (RSNA) when measured is increased and is now associated with the stress-induced changes in RVR. Renal sympathetic nerves act primarily by release of norepinephrine (NE), and nerve endings extend to most renal structures, including arterioles and tubules. In studies designed to test the effects of RSNA alone, higher RSNA frequencies within the range observed during stress increased RVR. The mechanism of RSNA stimulation on vascular tone is linked to increased release of NE and higher circulating epinephrine, acting on α - and β -adrenergic receptors, modulated by Ang II.

Experiments that stimulated RSNA show that there are at least three renal neuroeffector targets: renin-containing juxtaglomerular cells, the tubule, and the arterial vascular resistance. The sequence of RSNA input shows that renin is released at lower levels of RSNA, tubular reabsorption is increased at slightly higher intensities, and renal vasoconstriction occurs at even higher intensities. Each system appears to be regulated differently over a range of RSNA. Nonetheless, it appears that only moderate stress is sufficient to increase RSNA to levels that can alter major renal effectors. The role of RSNA on the renal contribution to hypertension is less clear. In studies in which periodic or prolonged periods of stress increased RSNA, there are only short-term effects on BP.

Many of the RSNA pacing studies have been limited to anesthetized rats, which may have altered basal RSNA. However, recent radiotelemetry techniques developed to measure RSNA in conscious rabbits exposed to stress have confirmed the role of renal nerves on increased RVR and have been used to explore the mechanisms. In response to air jet or white noise stress, RSNA was increased by 55 and 40%, respectively, within 1–2 min, and returned to normal by 9 min in conscious rabbits. Following treatment with rilmenidine, an imidazoline I (1) receptor agonist that reduces sympathetic-induced vascular tone, basal RSNA was reduced. However, the early response to air jet stress was reduced by 35%, whereas rilmenidine had no effect on white noise stress. Therefore, stress stimulates RSNA, but different stressor pathways may involve different central processing.

Increased RSNA impairs renal autoregulation, but the effects appear to be greater in hypertensive animals, suggesting a critical interaction with BP and RPP. However, stress also increases RSNA and BP in normotensive animals. WKY rats exposed to two different strategies of territorial stress (TS) developed higher BP over 5 months compared to control,

unstressed rats. There was no evidence that changes in renal function contributed to the hypertension. Indeed, renal function in spontaneously hypertensive rats (SHR) in which higher BP correlates to RSNA is not different from normotensive WKY. However, in this study TS rats had higher levels of serum corticosterone and catecholamines early in the study, but the values were similar to unstressed rats after 10–12 weeks. This suggested that early exposure to stress might have long-term effects. To test this idea, young normotensive rats were exposed to acute injection of the α -1 adrenoceptor agonist phenylephrine, which transiently increased BP and decreased glomerular filtration rate (GFR). After recovery, the animals were normotensive, but developed hypertension with reduced GFR when exposed to a high-salt (HS) diet, 8–10 weeks later. Together, these studies suggest that stress-induced surges in catecholamines, rather than constant stress, may enhance the susceptibility to other environmental stimuli, such as HS intake, and lead to hypertension and renal injury.

Based on the results from rat denervation studies, chronic increases in RSNA impact renal function. Under normal conditions, RSNA is low and renal denervation does not lead to changes in RBF or RVR. However, when RSNA is high, such as in the SHR or in heart failure rats, renal denervation increases RBF and decreases RVR. This occurs despite the similar RBF measured in intact SHR and WKY rats. Denervation does not alter RBF in WKY. This shows that chronic stress/increased RSNA tends to suppress RBF and increase RVR in SHR. However, it is not clear whether this effect contributes to the development of hypertension in SHR.

The role of stress on RVR has also been studied in human models, where stress-induced renal vasoconstriction also increases RSNA. RSNA is determined by measuring plasma or renal outflow of NE. In healthy volunteers, a short period of mental stress was accompanied by increased renal release of NE and epinephrine (Epi) and renin release, which led to a 48% increase in RVR. Renal arteries also appear to be more sensitive than other visceral vessels to stress in humans. Renal artery (RA) and small mesenteric artery (SMA) resistance and blood velocity were measured by simultaneous pulse and echo Doppler ultrasound flowmetry in 16 young female volunteers for 3 min before and 3 min during a mental stress test. RA vascular resistance was increased and RA blood velocity was decreased during the first minute of stress. SMA resistance was modestly increased, but velocity was unaffected by stress. This suggests that renal blood flow is exquisitely sensitive to mental stress and the accompanying increase in RSNA.

Hormonal Modulators

In addition to direct actions of renal nerves and activation of α -adrenergic receptors, multiple vasoactive hormones may contribute to stress-induced RVR. Endothelin is a peptide hormone that acts on receptors in endothelial cells and on vascular smooth muscle cells, generating vasoconstriction. In a series of human trials in elderly and young women, acute mental stress (30 min) increased BP, heart rate (HR), plasma NE, effective renal plasma flow, and RVR, but its effect on RVR was more prolonged in the elderly group. Urinary excretion of endothelin-1 was elevated in both groups, whereas excretion of vasodilating prostaglandins and nitric oxide (NO) was high only during the mental stress period in elderly subjects but was sustained at a higher excretion levels in young subjects. The balance between these opposing vasoactive hormones may therefore determine the level of stress-induced RVR. These studies provide good evidence that the balance between renal vasoconstrictors and vasodilators is invoked in the kidney's ability to autoregulate blood flow during mental stress. However, studies that evaluate the direct effects of endothelin infusion do not support a major role for endothelin as a potent vasoconstrictor during stress.

Several studies have shown that renin is released into the renal circulation during various forms of stress. However, it is not clear whether the levels of renin and locally produced Ang II are sufficient to contribute to the increased vasoconstriction that accompanies increased sympathetic activity. One report suggests that Ang II modulates the effects of catecholamine activation of adrenergic receptors, but the interaction has not been carefully investigated.

The role of NO was examined in adult borderline hypertensive rats (offspring of SHR and Wistar rats) during 5 weeks of crowding stress. Compared to unstressed controls, BP, HR, and nitric oxide synthase (NOS) activity in the aorta were higher. However, NOS activity in the heart, adrenal gland, and kidney was not affected. Superoxide (O_2^-), which scavenges NO, may contribute to RSNA. The higher RSNA activity in SHR compared to WKY was normalized by tempol, a powerful antioxidant, which reduced BP and RSNA to a greater extent in SHR than in WKY. However, the effect of reactive oxygen species on other influences on renal function could not be ruled out in this study.

The renal levels of neuropeptide Y (NPY) and calcitonin gene-related peptide (CGRP) were measured following a period of mental stress (word conflict test) or after low-dose adrenalin infusion in 12 healthy volunteers. Mental stress caused a threefold

increase in NE renal overflow but had no effect on NPY or CGRP, whereas Epi infusion increased levels of all neurotransmitters.

The pro-inflammatory transcription factor nuclear factor- κ B (NF- κ B) has been linked to psychosocial stress as a possible mediator of organ damage, specifically in the kidney. NF- κ B activation in renal tissue was greater in tissue from patients with renal complications. NF- κ B in peripheral blood mononuclear cells (PBMC), which can be easily measured, correlated with the degree of albuminuria in diabetic nephropathy patients and was reduced by antioxidant therapy. Similar levels of NF- κ B activation in PBMC were observed in healthy volunteers during a 15-min psychosocial stress test. The activation of NF- κ B in PBMC correlated with the stress-induced increases in catecholamines and cortisol. Though the long-term effects of stress on NF- κ B have not been tested, these acute studies provide a possible mechanism of kidney damage caused by stress.

Sodium Reabsorption

A major hypothesis for the development of hypertension invokes the role of the kidney in the inappropriate retention of solute and fluid resulting in increased fluid volume. In order to maintain homeostatic control of body fluids during periods of hypertension, the kidney excretes greater levels of urine and solute, resetting a lower extracellular fluid volume. However, the increases in BP seen during periodic stress does not always generate this pressure-natriuresis response, partially due to the increased RSNA, which simultaneously increases Na^+ reabsorption. As some studies suggest, the inability to excrete more solute and fluid during periods of increased BP may predispose those individuals to hypertension. The observation that acute stress can cause Na^+ retention has led to the popular theory that human essential hypertension is primarily due to chronic stress mediating excess extracellular fluid volume via a renal mechanism. Acute stress clearly increases RSNA, and several animal models of hypertension show increased renal sympathetic activity. However, RSNA is determined by postganglionic fibers and is controlled by the arterial baroreceptor reflex, which is not associated with long-term BP control. Though recent evidence shows a stronger correlation of RSNA and long-term control, RSNA generated by stress provides primarily acute regulation of renal function-induced hypertension.

Animal studies show that acute stress consistently increases RSNA and decreases Na^+ excretion in normotensive rats and dogs and does so to an even

greater extent in hypertensive rats. The neurohormonal regulation of these effects has been the focus of several studies. In conscious rats, a 10-min air jet stress increased BP and RSNA and reduced urine volume and Na^+ excretion by three- to fourfold (Figure 2). After denervation, urine volume and Na^+ excretion were not affected by stress, demonstrating the dependency of neural activity. Treatment with an AT1-R antagonist also prevented a decrease in urine flow or Na^+ excretion after stress, though RSNA was not affected. This suggests that Ang II mediates the antidiuresis and antinatriuresis induced by stress without affecting nerve activity. Previous studies showed that increases in RSNA led to greater Na^+ reabsorption, which was mediated by α -2 receptors. Inhibition of Ang II also decreased the RSNA effect, suggesting that Ang II interacts with α -2 receptors to regulate Na^+ reabsorption. Therefore, changes in RSNA induced by mental stress may have similar regulation by Ang II and α -2 receptors.

Results from human studies are less consistent but do show the effects of stress, and the associated increases in sympathetic activity on Na^+ retention are more closely associated with hypertension than in normotensive humans. In a well-controlled human subject trial, two 20-min periods of stress increased BP and decreased Na^+ excretion in 10 normotensive adult males, without altering GFR, measured by the clearance of inulin, or renal plasma flow (RPF), measured by the clearance of para-aminohippurate (PAH). Fractional Na^+ reabsorption was increased by stress and was primarily due to increases in reabsorption in the proximal tubule (PT), based on fractional lithium clearance measurements. Rilminedine, an imidazoline-I1 receptor antagonist, lowered basal BP but did not alter the effects of stress.

In a study that tested the interaction of hypertension and stress on renal function, three groups were evaluated: normotensive offspring of hypertensives (HP), hypertensive subjects (HT), and normotensive offspring of normotensives (NP). Basal BP, which differed between groups, was elevated by stress in all groups. However, the BP increased more in HP than in the other groups (Figure 1). Basal GFR and RBF were higher in HP than in the other groups. However, GFR was reduced by stress in HP and HT but not in NP, and RBF was reduced in HT and NP but not in HP. More interesting, however, were the responses to stress in normal subjects. Stress increased the filtration fraction, which suggests that efferent arteriolar resistance was elevated. There was also a paradoxical decrease in Na^+ excretion given that BP was higher. Again, the decrease in Na^+ reabsorption was attributed to increased PT reabsorption, based on the changes in lithium clearance.

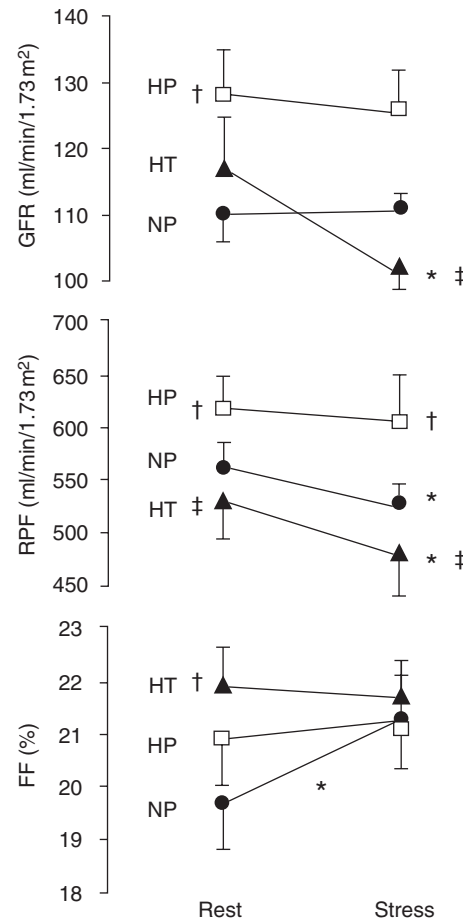


Figure 1 The effect of stress on GFR, RPF, and filtration fraction (FF) in normotensive (NP), normotensive with hypertensive parents (HP), and hypertensive (HT) subjects. From Ducher, M., et al. (2002). *American Journal of Hypertension* 15, 346–350. © 2002 American Journal of Hypertension, Ltd.

However, Na^+ retention induced by stress is not always observed in normal subjects. In three groups designed to test the potential for the development of hypertension, the effects of mental stress on Na^+ excretion were evaluated in normotensive young males with and without a family of hypertension and compared to hypertensive males. Each subject was studied with or without 1 week of therapy with an angiotensin converting enzyme inhibitor (cilazapril). A 30-min period of mental stress led to increased BP, HR, and GFR and decreased RBF in all groups. However, Na^+ excretion was increased in both normotensive groups, eliciting the appropriate pressure-natriuresis response, whereas the hypertensive group failed to change. Angiotensin converting enzyme therapy restored the increased Na^+ excretion in the hypertensive group, suggesting that the failed natriuretic response is due to the observed elevation of Ang II in hypertensive subjects. Together, these studies suggest that individuals who respond to stress by decreasing

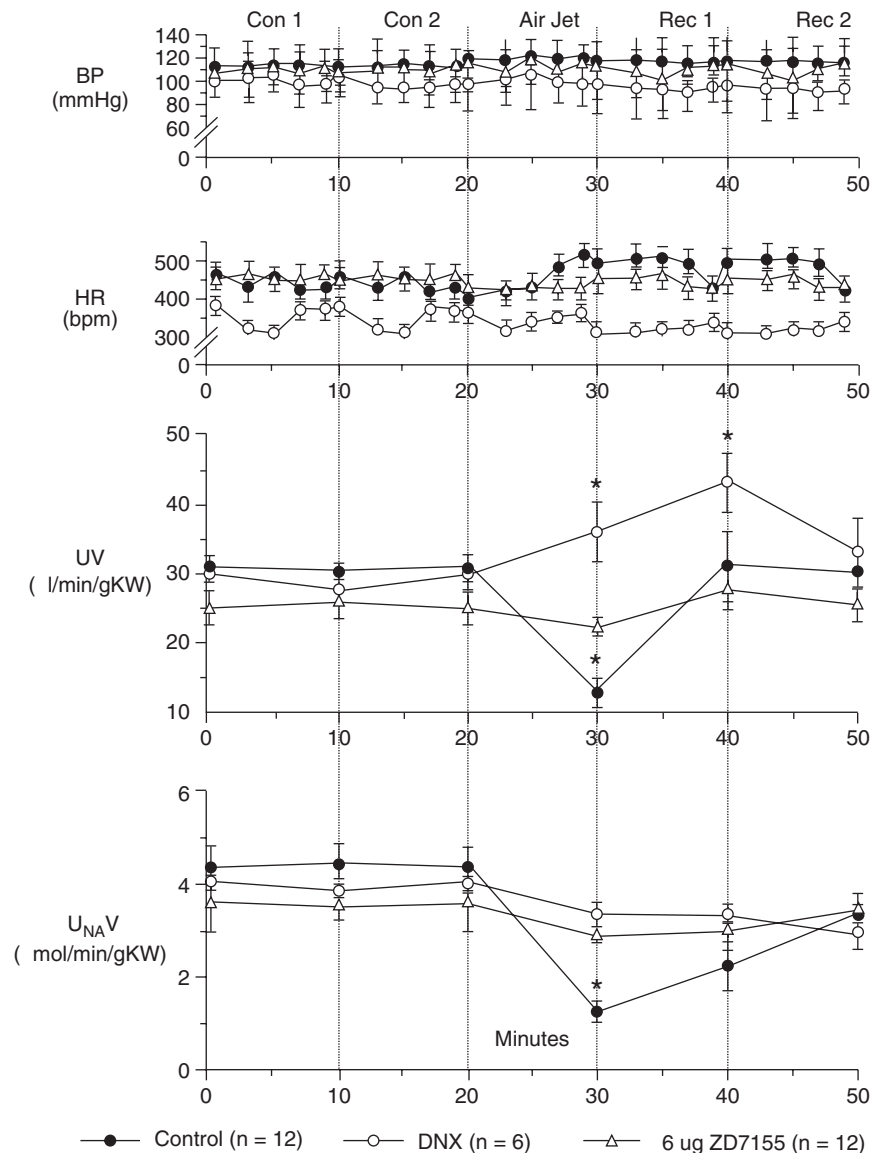


Figure 2 Effects of 10-min air-jet stress on BP, HR, urine flow rate (UV), and Na^+ excretion rate (U_{NaV}). UV and U_{NaV} are reduced during stress in conscious control rats. This effect is blocked by either denervation (DNX) or by Ang II, type 1 receptor blocker (ZD7155). From Veelken, R., et al. (1996). *Hypertension* 28, 825–832.

Na^+ excretion are predisposed to Na^+ retention and hypertension, whereas normotensive humans with the appropriate natriuresis response to stress are less likely to progress to hypertension via stress.

Does Chronic Stress Cause Hypertension by a Renal Mechanism?

As discussed previously, most of the studies that have shown increased RSNA, RVR, and Na^+ reabsorption to stress were observed during acute stress. Under these acute conditions the increased RVR and renin release may contribute to the acute rise in BP. However, it is difficult to assign a role for Na^+ retention in

acute hypertension, which occurs almost immediately. Therefore, caution should be observed when using acute studies to interpret chronic stress-induced hypertension. Studies in normotensive rats and mice have shown that prolonged stress for as much as 4–5 months increases BP during this period. Mice have shown progressive arteriosclerosis, myocardial hypertrophy, increased catecholamine levels, and greater angiotensin sensitivity, since Ang II blockade prevented hypertension. However, there were no measurements of Na^+ balance during the development of hypertension to evaluate the kidney's role. In hypertensive models, such as SHR, renal function and volume status also rarely differ. The use of prehypertensive models

has suggested a stronger role for the kidney in the development of hypertension but has not controlled adequately for the genetic backgrounds of these models. A more valuable set of data has been made in renal denervation studies, which have indicated that antinatriuresis and renal vasoconstriction may be due to the elevated RSNA in SHR.

Another issue of concern is that stress in rodents and humans may differ substantially. Human studies have not always found renal dysfunction in acute stress in normal subjects. This would be expected, since genetic predisposition appears to be necessary for the failure of the kidneys to adjust to high BP with appropriate natriuresis. It is possible that some normotensive subjects used in these studies had unknown genetic predisposition. Rodent studies have consistently shown the lack of renal compensation to acute stress, which is surprising but may not relate to the kidney's role, if any, in chronic stress. Indeed, chronic stress in rodents has been associated with cardiac and vascular remodeling.

The proper resolution of this question, at least in animals, will require a reproducible chronic stress model in rodents with noninvasive measurements of BP and renal function that can measure the critical components of renal influences on the development of hypertension over a chronic period. Lacking this type of information, the current evidence does not strongly support a role for changes in renal function contributing significantly to chronic stress-induced hypertension.

Summary

In summary, acute mental stress increases RSNA, BP, RVR, and Na^+ reabsorption in both animal and human studies. These effects are exaggerated in borderline hypertensive or hypertensive subjects, and many effects are blocked by Ang II inhibition, though multiple hormonal systems may be involved. The effects of chronic mental stress are more controversial, with few consistent observations. Though RSNA and general sympathetic activity is increased

in SHR, there are rarely any reports of renal dysfunction that would be expected if stress had permanent effects on renal function. The concept of mental stress altering renal function that leads to hypertension needs further experimental evidence.

See Also the Following Articles

Angiotensin; Blood Pressure; Hypertension; Regional Blood Flow, Stress Effects; Salt Appetite; Vasoactive Peptides.

Further Reading

- Barrett, C. J., Ramchandra, R., Guild, S. J., et al. (2003). What sets the long-term level of renal sympathetic nerve activity: a role for angiotensin II and baroreflexes? *Circulation Research* **92**, 1330–1336.
- Bierhaus, A., Wolf, J., Andrassy, M., et al. (2003). A mechanism converting psychosocial stress into mononuclear cell activation. *Proceedings of the National Academy of Science USA* **100**, 1920–1925.
- Castellani, S., Ungar, A., Cantini, C., et al. (1999). Impaired renal adaptation to stress in the elderly with isolated systolic hypertension. *Hypertension* **34**, 1106–1111.
- Di Bona, G. F. (2004). The sympathetic nervous system and hypertension: recent developments. *Hypertension* **43**, 147–150.
- Di Bona, G. F. and Kopp, U. C. (1997). Neural control of renal function. *Physiological Reviews* **77**, 75–197.
- Ducher, M., Bertram, D., Pozet, N., et al. (2002). Stress-induced renal alterations in normotensive offspring of hypertensives and in hypertensives. *American Journal of Hypertension* **15**, 346–350.
- Schneider, R. E., Veelken, R., Schobel, H., et al. (1997). Glomerular hypertension during sympathetic nervous system activation in early essential hypertension. *Journal of the American Society of Nephrologists* **8**, 893–900.
- Schneider, M. P., Klingbeil, A. U., Schlaich, M. P., et al. (2001). Impaired sodium excretion during mental stress in mild essential hypertension. *Hypertension* **37**, 923–927.
- Veelken, R., Hilgers, K. F., Stetter, A., et al. (1996). Nerve-mediated antidiuresis and antinatriuresis after air-jet stress is modulate by angiotensin II. *Hypertension* **28**, 825–832.

Korean Conflict, Stress Effects of

C A Goguen

National Center for PTSD, USA

M J Friedman

Dartmouth College, Hanover, NH, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 595–598, © 2000, Elsevier Inc.

Critical Issues in Korean Conflict Trauma Research

Combat Exposure

Lasting Effects of Combat Exposure

Summary

Prior to the Korean conflict (1950–1953), combat-related posttraumatic stress reactions, variously termed battle fatigue, war neurosis, and combat exhaustion, were considered a temporary condition best treated close to the combat zone. Influential theoretical perspectives included Freudian concepts of defense mechanisms as adaptations to stress-induced intrapsychic conflicts and, later, Selye's general adaptation syndrome, stress reactions viewed as a three-stage physiological response. Psychiatry during World War II and the Korean conflict viewed traumatic war neurosis as a complex phenomenon resulting from "multiple physical and psychic forces struggling for emotional control . . . the stimuli of battle itself evoked a defensive process that sustained men in combat . . . the lonely, fearful battle environment forces individuals to join together for protection and emotional support" (Glass, 1954: 728). The military unit was seen as a key source of support and motivation for soldiers to return to active combat. Therefore, all but the most severe cases of psychiatric casualties were treated at the front lines.

Further reinforcing the belief that combat-related posttraumatic stress reactions were short-lived, the publication of the first edition of *Diagnostic and Statistical Manual of Mental Disorders* (DSM-I) in 1952 established responses to either combat or civilian catastrophe invoking overwhelming fear, "gross stress reaction," as a temporary condition. Any protracted course of symptoms was attributed to an unrelated neurosis. Subsequently, the 1968 edition, DSM-II, dropped "gross stress reaction," relegating combat-related posttraumatic stress reactions to an adult adjustment problem.

Critical Issues in Korean Conflict Trauma Research

Nevertheless, the most prevalent service-connected disability in a sample of former Korean prisoners of war (POWs) was anxiety neurosis (19.2%). Few researchers at the time examined the psychological impact of combat-related traumatic stress on Korean veterans. In fact, most clinical and empirical research on the possible long-term psychological sequelae of the Korean conflict did not begin in earnest until the 1980s, three decades later, and only after a substantial amount of research had been conducted on posttraumatic stress disorder (PTSD) among Vietnam veterans. These retrospective studies all suffer from sampling problems that make it difficult to generalize from their findings to all Korean conflict veterans. Some studies did not differentiate Korean veterans from World War II veterans, whereas all the remaining studies focused exclusively on treatment-seeking veterans or POWs, each a highly select group.

Combat Exposure

Despite limited data, certain trends can be seen. It is clear that some Korean combat veterans were exposed to high levels of war-zone stress. Indeed, Blake and colleagues found no significant differences for combat exposure or posttraumatic stress symptomatology between Korean combat veterans and either World War II or Vietnam combat veterans. In an early study, Korean veterans reported that their combat-related posttraumatic stress symptoms not only persisted but actually worsened over time.

One of the best-studied groups of Korean combat veterans are former POWs. Like non-POW Korean combat veterans, Korean POWs reported that they had continued to experience posttraumatic stress symptoms years, even decades, after the trauma. Consistent with studies demonstrating a dose-response relationship between traumatic exposure and posttraumatic stress symptoms, Korean POWs reported higher levels of posttraumatic symptoms than non-POW Korean combat veterans. Whereas Korean POWs experienced less combat exposure than non-POW Korean veterans, the combat they had experienced before their capture was fierce with heavy casualties. After capture, they were sent on death marches and were subjected to brainwashing techniques and long periods of physical deprivation, interrogation, and solitary confinement. Thirty-eight percent of Korean POWs died during captivity.

Lasting Effects of Combat Exposure

Despite the circumscribed number of empirical investigations, the existing database suggests that, in addition to PTSD, Korean POWs and combat veterans have experienced long-term consequences of traumatic exposure, including alcohol abuse, depression, and general psychiatric distress. Further, higher mortality rates, especially due to accidents, appear to occur among Korean POWs than are expected for U.S. males. With respect to their World War II counterparts, Korean veterans had a higher lifetime prevalence of suicide attempts, more frequent occurrences of drug abuse-related hospitalizations and outpatient treatment, and greater use of mental health outpatient treatment. However, fewer Korean veterans met diagnostic criteria for PTSD. Nevertheless, Korean POWs reported higher rates of posttraumatic stress symptomatology and depression than World War II POWs.

Compared to Vietnam counterparts, Korean veterans reported lower levels of alcohol or drug abuse and fewer occurrences of drug abuse-related hospitalizations and outpatient treatment. Korean veterans were more likely than Vietnam veterans to recognize, early in their disorder, the link between their symptoms and their combat exposure. However, less use of PTSD treatment programs was found among Korean veterans.

Summary

Relative to Vietnam and World War II veterans, few researchers have included Korean veterans in their studies. A caveat should be exercised when drawing conclusions from studies of highly selective samples of war-era groups that differ in important areas such as age, premilitary socioeconomic status, time since combat exposure, POW status, repatriation, and social attitudes. Nevertheless, it seems apparent that traumatized Korean veterans suffer chronic debilitating posttraumatic psychological consequences of their war-zone experiences.

See Also the Following Articles

Combat Stress Reaction; Persian Gulf War, Stress Effects of; Prisoners of War; Vietnam Veterans, Postwar Experiences and Health Outcomes; War-Related Posttraumatic Stress Disorder, Treatment of.

Further Reading

American Psychiatric Association (1952). *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Association.

- American Psychiatric Association (1968). *Diagnostic and statistical manual of mental disorders* (2nd edn.). Washington, DC: American Psychiatric Association.
- Archibald, H. C. and Tuddenham, R. D. (1965). Persistent stress reaction after combat. *Archives of General Psychiatry* 12, 475–481.
- Beebe, G. W. (1975). Follow-up studies of World War II and Korean conflict prisoners. II: Morbidity, disability, and maladjustments. *American Journal of Epidemiology* 101, 400–422.
- Blake, D. D., Keane, T. M., Wine, P. R., et al. (1990). Prevalence of PTSD symptoms in combat veterans seeking medical treatment. *Journal of Traumatic Stress* 3, 15–27.
- Druley, K. A. and Pashko, S. (1988). Posttraumatic stress disorder in World War II and Korean combat veterans with alcohol dependence. *Recent Developments in Alcoholism* 6, 89–101.
- Engdahl, B. E., Harkness, A. R., Eberly, R. E., et al. (1993). Structural models of captivity trauma, resilience, and trauma response among former prisoners of war 20 to 40 years after release. *Social Psychiatry and Psychiatric Epidemiology* 28, 109–115.
- Foy, D. W., Sippelle, R. C., Rueger, D. B., et al. (1984). Etiology of posttraumatic stress disorder in Vietnam veterans: analysis of premilitary, military, and combat exposure influences. *Journal of Consulting and Clinical Psychology* 52, 79–87.
- Glass, A. J. (1954). Psychotherapy in the combat zone. *American Journal of Psychiatry* 110, 725–731.
- Goodwin, J. (1987). The etiology of combat-related posttraumatic stress disorders. In: Williams, T. (ed.) *Posttraumatic stress disorders: a handbook for clinicians*, pp. 1–18. Cincinnati, OH: Cincinnati Disabled American Veterans.
- Keehn, R. J. (1980). Follow-up studies of World War II and Korean conflict prisoners. III: Mortality to January 1, 1976. *American Journal of Epidemiology* 111, 194–211.
- Lee, K. A., Vaillant, G. E., Torrey, W. C., et al. (1995). A 50-year prospective study of the psychological sequelae of World War II combat. *American Journal of Psychiatry* 152, 516–522.
- Mellman, T. A., Randolph, C. A., Brawman-Mintzer, O., et al. (1992). Phenomenology and course of psychiatric disorders associated with combat-related posttraumatic stress disorder. *American Journal of Psychiatry* 149, 1568–1574.
- Page, W. F., Engdahl, B. E. and Eberly, R. E. (1991). Prevalence and correlates of depressive symptoms among former prisoners of war. *Journal of Nervous and Mental Disease* 179, 670–677.
- Rosenheck, R. and Fontana, A. (1994). Long-term sequelae of combat in World War II, Korea and Vietnam: a comparative study. In: Ursano, R. J., McCaughey, B. G. & Fullerton, C. S. (eds.) *Individual and community responses to trauma and disaster*, pp. 359. New York: Cambridge University Press.
- Selye, H. (1950). *The physiology and pathology of exposure to stress*. Montreal, Canada: Acta.

- Sutker, P. B. and Allain, A. N. (1996). Assessment of PTSD and other mental disorders in World War II and Korean conflict POW survivors and combat veterans. *Psychological Assessment* 8, 18–25.
- Sutker, P. G., Winstead, D. K., Galina, Z. H., et al. (1991). Cognitive deficits and psychopathology among former prisoners of war and combat veterans of the Korean conflict. *American Journal of Psychiatry* 148, 67–72.
- U.S. Veterans Administration/Office of Planning and Program Evaluation (1980). *POW: study of former prisoners of war*. Washington, DC: U.S. Government Printing Office.



Lactation See: Oxytocin; Prolactin and Stress.

Learned Helplessness

D M Isaacowitz

Brandeis University, Waltham, MA, USA

M E P Seligman

University of Pennsylvania, Philadelphia, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
M E P Seligman and D M Isaacowitz, volume 2, pp 599–602,
© 2000, Elsevier Inc.

Original Observations and Theory
Learned Helplessness and Psychopathology in Humans

Glossary

<i>Explanatory style</i>	A person's characteristic ways of explaining the causes of the events that happen in his or her life.
<i>Shuttle box</i>	Experimental apparatus for animals with two compartments separated by a barrier that can be adjusted by the experimenter.
<i>Uncontrollable reinforcer</i>	A reinforcer whose probability is the same whether an emitted response is present or absent. In this situation, the responding organism cannot control the reinforcer.

Original Observations and Theory

Learned helplessness refers to the motivational, cognitive, and emotional deficits that may follow from an organism's exposure to uncontrollable stressors. The theory arose from the observation that, after experiencing inescapable shock over which they had no control, dogs in the laboratory displayed a variety of behavioral deficits. In the original experiments on the learned helplessness phenomenon, the following

design was typically used. The dogs were first given unpredictable and inescapable electric shocks in a Pavlovian hammock. Then, 24 h later, the dogs were placed in an experimental shuttle box. In this shuttle box, there were two compartments, and the dog was given 10 trials of escape/avoidance training. Shock was given in both compartments, but the termination of the shock took place when the dog jumped from one compartment to the other.

Approximately two-thirds of the dogs placed in the shuttle box after experiencing inescapable shock did not learn the escape/avoidance behavior; that is, they never learned that they could end the shock simply by jumping to the other compartment. Instead, they struggled at first and then stopped attempting to take action that might terminate the shock. The vast majority of dogs who had not experienced inescapable shock prior to entering the shuttle box had no trouble learning the escape/avoidance procedure to terminate the shock. The deficit in many of the previously shocked dogs was psychological rather than physical because these same dogs were able to run out of the shuttle box when the exit was opened.

This learned helplessness effect observed in the previously shocked dogs (i.e., they passively accepted later shock and did not learn to avoid it) was caused by the uncontrollability of the shocks experienced in the Pavlovian hammock. This was demonstrated by Seligman and Maier in 1967 using a yoked experimental design with three groups of dogs. The first group could press a panel in the hammock to terminate the shock. The second group received the same duration of shock as the first group, but could not control the onset or termination of its shock in the hammock. The third group received no shock prior to being placed in the shuttle box. The yoked group,

which had received uncontrollable shock in the hammock, showed the strongest deficits in learning in the shuttle box. The dogs that had received the same amount of shock but could press to control the termination of the shock did not show these deficits.

The deficits observed in helpless dogs fall into three categories. First, motivational deficits exist because the dogs stop initiating voluntary behaviors, such as jumping from one compartment to the other. Cognitive deficits also appear in that the dogs do not learn that their responses have been effective even when they have indeed caused the desired effect. Finally, the dogs show transient helplessness effects that dissipate over time, suggesting that helplessness may be a passing emotional response. Interestingly, exposure to controllable shock in the shuttle box prior to experiencing uncontrollable shocks appears to immunize dogs against later helplessness deficits and forcibly exposing dogs to the appropriate response contingency (i.e., that jumping to the other compartment terminates the shock) can eliminate these deficits.

According to the researchers, the dogs who received uncontrollable, inescapable shock learned that outcomes were independent of their responses. When organisms experience uncontrollable outcomes, they may notice this contingency and learn that the outcomes are independent. They will then expect outcomes to be independent of their responses in the future. This expectation involves a cognitive representation of the contingency. The purpose of voluntary action is to cause certain outcomes, so a belief in response–outcome independence will reduce the organism’s motivation to engage in voluntary responses. Because forming this expectation of independence is an act of learning, this cognitive representation will interfere proactively with future attempts to learn about response–outcome dependence. Finally, the fear that follows a traumatic event may be replaced with negative emotions when the organism realizes its lack of control in the situation. Thus, learned helplessness theory seeks to explain the three primary deficits observed through a cognitive pathway. In their 1976 review paper, Maier and Seligman reviewed alternative hypotheses for the observed deficits and argued that the helplessness account fits most parsimoniously with the available data.

Learned helplessness is not specific only to dogs in the shuttle box. Deficits have also been documented in cats, rats, and humans. These extensions of learned helplessness to animals other than dogs have been empirically and theoretically useful. One study found that helpless rats were less likely to reject cancerous tumors than nonhelpless rats, suggesting an important connection between helplessness and immune function. In humans, several studies used a yoked

paradigm in which participants (college students) in the first group could turn off a loud noise by pressing a button. Students in the second group heard the same noise for the same duration of time as those in the first group, but the sound was not contingent on their pressing a button. A final control group did not hear any noise. In one study conducted by Hiroto, students then put their hand in a shuttle box; they were shocked but could escape by moving their hand to the other side. Just as in the dogs, the students in the uncontrollable group showed helplessness deficits; they did not learn the avoidance strategy and just took the shock. In another study by Miller and Seligman, students who heard the uncontrollable inescapable noise had more trouble solving difficult anagrams than did those who heard escapable noise and those who heard no noise at all.

Learned Helplessness and Psychopathology in Humans

Although these experiments demonstrate that learned helplessness deficits can be produced in humans, a large amount of research has investigated whether the learned helplessness model can be thought of as an analog for various psychological disorders in humans. The most work has been done on depression and posttraumatic stress disorder (PTSD), so we consider here how learned helplessness might be useful for the understanding and treatment of these two types of psychopathology. Although our focus in the remainder of this article is on links between learned helplessness and mental health, there is an additional, smaller body of work linking learned helplessness and physical health in humans.

Depression

In his 1975 book *Helplessness*, Seligman asserted that the learned helplessness phenomenon appeared to be very much like human depression. Specifically, the motivational, cognitive, and emotional deficits that appeared in helpless dogs seemed to mimic the symptoms of reactive depression, in which a human gets depressed after a major life stressor. Certainly the apparent emotional withdrawal and passivity shown by helpless dogs and college students paralleled depression; in addition, decreased motivation and increased cognitive difficulty, also parts of the helplessness experience, are major symptoms of depression. Based on these and other similarities, Seligman hypothesized that both learned helplessness and reactive depressions result from the expectancy that responses and outcomes are independent. Thus, gaining a sense of control would serve as an important part of treatment for depression.

Although the expectancy of future response–outcome independence served as a helpful stimulant for research on the cognitive predictors of depression in humans, there were several flaws with this original learned helplessness conceptualization of depression. First, just as not every dog became helpless after uncontrollable shock, not every human becomes depressed after experiencing stressful events with uncontrollable outcomes. Second, the theory could not distinguish between a depressed person feeling that he or she personally could not effect outcomes and the feeling that the response–outcome independence was true universally.

Abramson, Seligman, and Teasdale therefore offered a reformulated learned helplessness model of depression in their 1978 paper. This cognitively oriented reformulation centered on the notion that, when a life stressor takes place, people do not simply notice the contingencies involved. Rather, they ask themselves why the event happened. People tend to have usual ways of answering this question, and this is called their explanatory style. Causes for events generated by people tend to fall along three separate dimensions: (1) people may feel the cause is due to themselves or to other people or luck, (2) they will either attribute the cause to factors that are stable in time or temporary, and (3) the cause may affect many life domains or just the ones involved most specifically with the stressor. The theory follows a diathesis-stress model of depression onset – it claims that people who have a pessimistic explanatory style, in which negative events are explained with internal, stable, and global causes, will be especially vulnerable to depression when faced with uncontrollable life stressors.

There has been a large amount of empirical support for this model, using both questionnaires (primarily the Attributional Style Questionnaire) and content analysis of verbatim data and a wide range of experimental methodologies. A usual study followed a population prospectively over time to determine whether pessimists were more likely to get depressed after experiencing negative life events. In one series of studies, pessimistic college students experienced the most enduring depressive symptoms after doing poorly on a midterm. A meta-analysis supported the conclusion that a pessimistic explanatory style relates to depression, especially with regard to explanations for negative events. However, the vast majority of studies used either college students or adult psychiatric patients; there is some indication that the relationship between pessimism and depression may be slightly different in older adults. Nonetheless, prevention programs based on teaching pessimistic children and college students to dispute their automatic pessimistic thoughts and to become more realistically

optimistic have produced lower rates of significant depressive symptoms in participants at high risk for depression. In addition, there is some evidence that successful cognitive-behavioral psychotherapy can cause a shift in explanatory style such that clients become more optimistic.

Post Traumatic Stress Disorder

Despite the reformulated learned helplessness model's prominence in research on cognitive approaches to human depression, it is the original learned helplessness model that has primarily (although not exclusively) influenced research on PTSD in humans. The behaviors that follow inescapable shock in animals appear to mimic many of the symptoms of PTSD in humans. As already noted, animals exposed to inescapable shock show cognitive, motivational, and emotional deficits. Similarly, humans exposed to traumatic stressors may show a variety of sequelae, including anxiety and hyperreactivity, isolation, and anhedonia, that are similar to the sequelae observed in helpless animals. After inescapable shock, animals experience a depletion of catecholamine in their central nervous system. Van der Kolk and his colleagues posit that catecholamine depletion, especially norepinephrine (NE) depletion, is also involved in the negative symptomatology of PTSD, such as lack of motivation. They also posit that an augmentation of locus coeruleus pathways causes intrusive PTSD symptoms, such as nightmares and flashbacks.

More important to van der Kolk and his colleagues' work is the analgesia response observed in helpless animals, such that they do not appear to experience pain in response to subsequent stressors presented within a short period of time after the administration of inescapable shock. This response is due to endogenous opioids in the body. Interestingly, there are many parallels between the symptoms of opiate withdrawal and PTSD, leading van der Kolk and colleagues (1985: 320) to hypothesize that "both are, at least in part, due to central noradrenergic hyperactivity associated with a relative decrease in brain opioid receptor binding." This hypothesis is supported by evidence that trauma survivors often put themselves in situations similar to the traumatic event, perhaps in an attempt to evoke endogenous opioid release and thus a sense of calm and control. However, this sense of calm is followed by symptoms of opiate withdrawal, such as hyperreactivity, thus immersing PTSD sufferers in a cycle of reexposure and withdrawal.

Other research on helpless animals finds that the NE depletion may be the outcome of a more immediate increase in NE when a helpless animal is exposed to a mild stressor. The hippocampus of the helpless

animal appears to be conditioned by the inescapable, helplessness-inducing trauma to respond with increased NE to later lesser stressors. This increase in NE secretion immediately after later stressors could be part of the same general catecholamine dysregulation following traumatic stress as the NE depletion reported by van der Kolk.

In addition to the neurochemical and behavioral evidence for a learned helplessness approach to PTSD based on uncontrollable stress, evidence also exists supporting a relationship between helpless explanatory style and PTSD symptoms. McCormick and colleagues studied explanatory style among patients addicted to alcohol and/or gambling and found that patients with PTSD had a more pessimistic explanatory style than those who did not have PTSD. Although correlational, these results suggest that the cognitive reformulation of learned helplessness may also contribute to an understanding of the phenomenology of and potential treatments for PTSD. Moreover, recent work suggests that learning experiences can immunize animals against becoming helpless in the face of stress and that posttrauma therapeutic interventions centered on learning to predict and control stressors can be helpful in the treatment of PTSD in humans. This work suggests that the learned helplessness framework may be useful for both prevention and treatment of stress-related psychological disorders.

See Also the Following Articles

Cytokines, Stress, and Depression; Depression Models.

Further Reading

- Abramson, L. Y., Seligman, M. E. P. and Teasdale, J. D. (1978). Learned helplessness: critique and reformulation. *Journal of Abnormal Psychology* 87, 49–74.
- Hiroto, D. S. (1974). Locus of control and learned helplessness. *Journal of Experimental Psychology* 102, 187–193.
- Isaacowitz, D. M. and Seligman, M. E. P. (2001). Is pessimistic explanatory style a risk factor for depressive mood among community-dwelling older adults? *Behaviour Research and Therapy* 39, 255–272.
- Maier, S. F. and Seligman, M. E. P. (1976). Learned helplessness: theory and evidence. *Journal of Experimental Psychology: General* 105, 3–46.
- Miller, W. R. and Seligman, M. E. P. (1975). Depression and learned helplessness in man. *Journal of Abnormal Psychology* 84, 228–238.
- Overmier, J. B. and Murison, R. (2005). Trauma and resulting sensitization effects are modulated by psychological factors. *Psychoneuroendocrinology* 30, 965–973.
- Peterson, C. and Bossio, L. M. (1991). *Health and optimism*. New York: Free Press.
- Seligman, M. E. P. (1975). *Helplessness*. San Francisco: W. H. Freeman & Company.
- Seligman, M. E. P. (1990). *Learned optimism*. New York: Pocket Books.
- Seligman, M. E. P. and Maier, S. F. (1967). Failure to escape traumatic shock. *Journal of Experimental Psychology* 74, 1–9.
- Sweeney, P. D., Anderson, K. and Bailey, S. (1986). Attributional style in depression: a meta-analytic review. *Journal of Personality & Social Psychology* 50, 974–991.
- Van der Kolk, B., Greenberg, M., Boyd, H., et al. (1985). Inescapable shock, neurotransmitters, and addiction to trauma: toward a psychobiology of post-traumatic stress disorder. *Biological Psychiatry* 20, 314–325.

Learning and Memory, Effects of Stress on

M Lindau

Uppsala University, Uppsala, Sweden

O Almkvist

Karolinska Institutet and University of Stockholm,
Stockholm, Sweden

A H Mohammed

Karolinska Institutet, Stockholm, and Växjö University,
Växjö, Sweden

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
M Lindau, O Almkvist, and A H Mohammed, volume 2,
pp 603–609, © 2000, Elsevier Inc.

Introduction

Effects of Stress on Learning and Memory in Animals
Memory, Learning, and Their Anatomical Bases
in Humans

Encoding in Stressful Situations

Effects of Anticipated Stress on Memory

Possible Mechanisms of the Relation between Stress
and Memory

Summary

Glossary

<i>Stressor</i>	Stimuli that puts an extra demand on the organism.
<i>Stress hormones</i>	Hormones of the adrenal glands that serve to maintain bodily homeostasis in response to challenging external and emotional environment. These hormones are the glucocorticoids and epinephrine.
<i>Corticosterone</i>	A steroid hormone in many species including rodents produced by the adrenal cortex and involved in response to stress and immune reaction.
<i>Cortisol</i>	A steroid hormone in humans produced by adrenal cortex, involved in response to stress and has direct influence on the nervous system.
<i>Hippocampus</i>	A region of the cerebral cortex lying in the basal medial part of the temporal lobe thought to be important for learning and memory.
<i>Amygdala</i>	A group of nuclei involved in the medial anterior part of the temporal lobe regulating emotion and certain types of learning.
<i>Emotion</i>	Mental state of arousal.

Memory

Encoding of information, storage and recall.

Learning

A relatively permanent change in cognition as a consequence of the acquisition of new information and experience, directly influencing behaviour.

Introduction

Conceptually, the term stress describes a load on the organism. A stressful situation puts extra demands upon the system for cognitive or physical productivity. A stressor, which might be external or internal, can be anything that puts extra demands upon the system, such as a specific problem, issue, challenge, or personal conflict. A stress reaction is the individual's response to a given stressor on the physiological, cognitive, and/or emotional plane. Strain occurs when the organism is overloaded, exposed to prolonged stress, or placed in situations that the system can not cope with. The psychological consequence of strain is fatigue or exhaustion. Traumatic, salient memories that the individual would like to forget but can not may result in a neurotic fixation on the memories of the trauma, i.e., posttraumatic stress disorder (PTSD).

Physiologically, stress might be defined as a state of physiological and emotional arousal of the organism. It leads to the activation of the autonomic nervous system and the hypothalamic-pituitary-adrenal axis. The main signaling hormones of these systems are (nor)adrenaline, corticotropin-releasing hormone, and cortisol (corticosterone in most rodents).

It is generally accepted that stress affects learning and memory processes and that stress might improve as well as impair the acquirement of information. How is this possible? One answer to this question is that there are at least three types of stressors, physical, cognitive, and emotional, which affect different regions of the brain. Another answer lies in the progress that has been made in the knowledge of the effects of stress on the different stages of learning and memory processing.

Although the mnemonic system in animals is considerably more restricted than that in humans, much complementary knowledge about the effects of stress on learning and memory in humans has been gained through animal studies. The following section gives an account of some of the basic research about this topic in animals.

Effects of Stress on Learning and Memory in Animals

Animals learn about the relationship or association between two events. In classical conditioning, the animal learns to associate a stimulus such as a tone (conditioned stimulus) with food (unconditioned stimulus). In operant or instrumental conditioning, the animal learns to associate a response, for instance, pressing a lever (instrumental response), with food reward (reinforcement). It has been suggested by some researchers (e.g., Anthony Dickinson) that associative learning mechanisms have been shaped by evolution to enable animals to detect and store information about real causal relationships in their environment.

Stress Hormones and Learning

The impact of stress hormones in modulating associative learning and cognitive function in animals has been a subject of continued investigation. The studies that have documented the relationship of levels of the stress hormones with cognitive function, such as attention, perception, and memory, have been done mostly in rodents, in which investigators have examined the influence of stress hormones on acquisition, consolidation, and retrieval of information. Glucocorticoids are the stress hormones secreted by the adrenal glands in animals. In primates and dogs, the naturally occurring glucocorticoid is cortisol, whereas in rodents and birds it is corticosterone.

Fluctuations in corticosterone levels can be said to reflect emotional states related to stress. In experimental animals, changes in hormonal levels can be achieved by pharmacological means or environmental manipulations and their effects assessed. Several studies have shown that stress and glucocorticoids influence cognitive function.

The administration of low levels of corticosterone improves performance in learning tasks in animals. However, whereas short-term exposure to low levels of corticosterone can enhance cognitive function, it is known that short-term exposure to higher levels as well as long-term effects of high levels of corticosterone (e.g., through sustained exposure to restraint stress) have deleterious effects on learning, memory, and cognitive function. The biphasic effects of corticosterone on memory formation have been demonstrated in behavioral studies performed in chicks and rats.

Young chicks have a tendency to peck spontaneously at small salient objects in the environment. This natural behavior of chicks has been exploited to study learning. Chicks will peck spontaneously at bright beads, and if the bead is dipped in a distasteful

substance, the chicks will peck once at the bead and show a strong disgust response. Subsequently, they will avoid pecking a similar but dry bead and learn this avoidance behavior. The strong emotional fear component involved in this passive avoidance response in chicks prompted Sandi and Rose to explore the role of stress hormones in learning in the chick. In tests of passive avoidance learning in the chick, they found that an inverted U-shaped curve emerged, in which low levels (0.1 μg) and higher levels (2 μg) of corticosterone were without effect, whereas an intermediate dose (1 μg) improved learning. Similarly, in rodents, an inverted U shape is seen between the dose of administration of corticosterone and performance in learning tasks. It has thus been seen that low doses of corticosterone can improve learning, whereas continuous exposure to glucocorticoids results in neuronal death in the hippocampus and in impaired learning ability. Aged rats that are impaired in cognitive function have elevated levels of corticosterone and increased hippocampal neuronal loss when they are compared with aged rats that are not impaired in cognitive function.

The effects of corticosterone on cognitive function are mediated through binding of the stress hormones to specific receptors in the brain. These receptors, known as glucocorticoid receptors, have been found to be abundant in the hippocampus, a brain region that is critically involved in modulating learning and memory. Another region in the brain that participates in cognitive function is the amygdala, which also has moderate amounts of receptors for corticosterone. The actions of corticosterone in the hippocampus and amygdala can be induced by the administration of selective drugs that interact with the glucocorticoid receptors. Drugs include those that enhance the effects of corticosterone (glucocorticoid agonists) or those that block or attenuate the effect of corticosterone (glucocorticoid antagonists). These drugs have provided powerful tools in dissecting the role of stress hormones in the hippocampus and amygdala in modulating cognitive functions in rodents. Pharmacological or environmental manipulation of the glucocorticoid receptors influences cognitive function in rats.

Hippocampus and Learning in Animals

Considerable evidence shows that the hippocampus mediates spatial learning in rodents. Spatial learning in rodents can be examined by the use of a water maze, known as the Morris maze. The procedure involves placing an experimental animal in a large pool of water of about 1.5 m in diameter. The animal has to use spatial cues in the experimental room to

locate an invisible escape platform that is hidden 2 cm below the surface of the opaque water. After a few days of testing, the animals learn to locate the platform within seconds. It has been demonstrated that rats perform poorly in this test if hippocampal function is impaired. Thus, aged rats with pathological changes in the hippocampus show impaired spatial learning in the water maze compared to aged rats that do not show pathological hippocampal changes. Aged rats that had been exposed to repeated restraint stress have a reduction in glucocorticoid receptors in the hippocampus. This results in, among other things, an impaired spatial learning ability. In contrast, animals that have higher levels of glucocorticoid receptors in the hippocampus show an increased ability for spatial learning. Administration of a glucocorticoid receptor antagonist in the brain appears to impair information processing during spatial learning.

To sum up, spatial memory in animals is dependent on hippocampal function. Through their action in the hippocampus, stress hormones facilitate spatial information processing so that low levels of corticosterone improve learning, whereas higher levels impair performance.

Amygdala and Learning in Animals

Many studies have concentrated on the effects that stress hormones have on the hippocampus and its relationship to learning. Work by James McGaugh and colleagues has begun to explore the role of stress hormone receptors in the amygdala on learning. The amygdala is a region of the brain that plays a crucial role in acquisition and expression of fear responses. This brain region is thought to influence memory storage processes in other brain regions, such as the hippocampus and cortex. An interaction of stress hormones and amygdala nuclei appears to be important in modulating memory consolidation for emotional events. Binding of glucocorticoid receptors in the basolateral nucleus is known to affect memory storage. It was found that administration of the glucocorticoid receptor agonist in this area immediately after training enhanced retention of a passive avoidance response. Many studies have examined the effects of glucocorticoids on the acquisition of newly acquired information. In a test of memory retrieval of long-term spatial learning memory, it was found that stress caused impaired performance, and this impairment could be related to circulating levels of corticosterone. These findings point to a critical involvement of the amygdala in regulating stress hormone effects on learning and memory.

To conclude, the amygdala is a region of the brain known to be importantly involved in emotions. Stress

hormones can affect the memory storage of emotional events by their actions on the amygdala. The impact of stress on the hippocampus and amygdala is highlighted by emerging evidence from animal studies that revealed that stress causes changes in hippocampal plasticity and can thus modify the information storage mechanism in the hippocampus and that the amygdala is involved in the emergence of stress-associated changes in the hippocampus.

Summary

Stress hormones have biphasic effects on learning and memory in animals. Low levels of stress hormones improve performance, whereas higher levels have a deleterious effect on learning. Two brain regions important in regulating learning and memory are the hippocampus and amygdala, which contain receptors for stress hormones. The effects of stress hormones on learning are mediated by these receptors in the hippocampus and amygdala. Research in rodents has revealed that emotional arousal regulates learning through the action of stress hormones, which activate receptors in the hippocampus and amygdala to affect memory formation.

Memory, Learning, and Their Anatomical Bases in Humans

Conceptualization of Learning and Memory

The contemporary conceptual frameworks of learning and memory emanate from experimental psychology. According to this model, human memory is not a unitary system, but is composed of a series of interdependent brain systems characterized by different features in terms of time for the registration of new information (immediate/extended in time), type of information (implicit/declarative or explicit), accessibility (immediate/search), amount of information stored (limited/unlimited), spontaneous duration of information in memory (limited/unlimited), and awareness of information (yes/no). These distinct memory systems appear to be subserved by separate neural networks, as shown by brain behavior studies and prior studies demonstrating specific effects on memory due to lesions in the brain.

One major division of memory is that between implicit or nondeclarative memory and declarative (explicit) memory. Nondeclarative memory includes procedural memory, classical conditioning, the perceptual representation system, and priming. In line with phylogenesis and ontogenesis, one early type is procedural memory, which is expressed as skilled behavior or motor procedures, which appear to be

suberved by regions in the basal ganglia. One division of procedural memory is classical conditioning, which is the same as a stimulus–response connection. The hippocampus and certain parts of the cerebral cortex have been identified as important areas for this memory system. Another type of implicit memory is the perceptual representation system, which is concerned primarily with the acquisition and use of perceptual objects of the world. The predominantly responsible brain areas are the primary association cortices. Another branch of implicit memory is priming, which refers to an unconscious ability to process a stimulus due to prior exposure. Declarative memory can be described as the conscious recollection of facts and events (Krystal et al., 1995). To this memory system belongs short-term memory or working memory, semantic memory, and verbal as well as visuospatial episodic memory. Short-term or working memory processes and stores temporarily a very limited amount of information. Semantic memory covers acquisition and use of factual knowledge about the world and is mapped by association areas in the brain. Information unique in time, location, and person is involved in storage and recall of information from episodic memory, for which the medial temporal lobes are one crucial point of the subserving neural networks. Another important aspect of memory is learning. Some authors equate long-term storage with learning, for example, calling it the process by which information that may be needed again is stored for recall on demand (Peterson, 1967). A broader definition is that learning is the same as the modification of behavior according to earlier experiences.

The Anatomy of Human Memory

The hippocampus thus plays an important role in several of the subdivisions of the two main memory systems. In the hippocampus, stimuli from the external world are classified and stored, momentarily in short-term memory and for longer periods of time in long-term memory, and the spatial and temporal dimensions of the experiences are recorded. It is doubtful, however, that the hippocampus is responsible for the more permanent consolidation of memories. Adjacent to the hippocampal system is the amygdala, which plays an important role in the processing of emotional information in the brain. The amygdala consists of several distinct nuclei, whose functional properties vary. The basolateral amygdaloid complex appears to be the nucleus that is involved most crucially in memory performance. However, it is not critical for the consolidation of a memory trace per se. Instead, it has been suggested that the

amygdala assigns emotional meaning to incoming free-floating feelings, integrates this emotionally meaningful information with associated emotional memories, and, through projections to the hypothalamus, hippocampus, and basal forebrain, directs emotional behavior. The amygdala also contributes to the enhancement of a memory trace that is related to emotional responses such as stress.

The frontotemporal area is the center for the executive function, short-term memory, and personality. This area of the brain has an elaborate network of connections to most other regions of the brain and a high concentration of stress hormone receptors, which has been found to assure the role of the frontal lobes in stress, learning and memory.

Vulnerability of Human Memory

The current conceptualization of human memory posits that the order of memory acquisition during ontogenesis, as described earlier, is reversed when the brain is exposed to various adverse influences such as aging, toxins, inflammatory agents, and stress. The meaning of this is that more sophisticated memory systems, such as declarative memory, are hit at an early stage, whereas more basic and ontogenetically earlier systems, such as procedural memory, are afflicted later on. Of course, the reversed order of degradation does not hold when there is a highly localized attack on the brain, as is the case for brain trauma. However, it is well known that episodic memory is highly sensitive for aging as well as for disease processes such as Alzheimer's disease. The next section describes changes that take place in relation to memory performance when humans are exposed to stress.

Encoding in Stressful Situations

Encoding is closely connected with attention and vigilance, which both influence the encoding of information into declarative as well as nondeclarative memory. Attention, like vigilance, is highly sensitive to stress. To be able to encode an all-embracing picture of an event, it is necessary that attention be unrestricted. During a stressful situation the attentional span tends to be narrowed to the most salient stimuli in the situation, whereas more peripheral stimuli are neglected. It has also been suggested that the hippocampus seems to shut down under severe stress, which may contribute to an incomplete encoding of the stressful situation. It has been demonstrated experimentally that people who have been exposed to pictures of a bloody car accident tend to remember central trauma-related items better than more

peripheral details. This laboratory finding is related to a real-world phenomenon called weapon focusing. Witnesses to crimes committed with visible weapons have sharper memories of the weapon than of other aspects of the situation, including the face of the perpetrator, for which the memories are more diffuse. The explanation for this is that the gun is the carrier of the most emotionally salient information and therefore captures the attention at the expense of other aspects. This phenomenon has been found to be particularly pronounced in people reacting with anxiety at the sight of weapons.

Short-term memory, which demands that information be held in consciousness momentarily, has been found to be extra sensitive to stress because the situation demands a focus on stress-relevant stimuli and therefore attention tends to be biased. On the contrary, overlearned abilities, such as procedural memory, for which the memory traces are well consolidated, are found to be more stress resistant than recently acquired memories, which are destroyed more easily. In a laboratory test of declarative memory, spatial thinking, and procedural memory tested with a word stem priming task, high levels of cortisol impaired declarative memory and spatial thinking, whereas procedural memory was spared.

One reason why overlearned skills are more difficult to erase is that repetition contributes to a better memory. Thus, it appears that more recently registered material in long-term storage and short-term memory is more sensitive to stress.

Effects of Anticipated Stress on Memory

It is not always being in the stressful situation that causes emotional arousal; the anticipation of the incoming stressor may have the same effect. In an experiment by Lupien and colleagues, a group of 14 healthy elderly persons was given a nonstressful task and a stressful task. The nonstressful condition consisted of a computer task. The stressful condition consisted of a videotaped public-speaking task, for which the discourse was said to be evaluated by experts. Before and after each condition subjects were given a declarative memory task and a nondeclarative word completion task. Results showed that in the responders (individuals who react to stress with an increase in cortisol levels), high levels of cortisol were secreted long before the stressful condition, which was not the case in the nonresponders (no change in cortisol level due to stress). This signifies that the anticipation of the incoming stressor played a more important role in the secretion of cortisol than the actual stressor. The responders also

performed worse on the declarative memory task than the nonresponders, indicating that even the anticipation of stress has a detrimental effect on declarative memory. Nondeclarative memory was spared. The stress did not affect the nondeclarative memory significantly, either in the whole group or in the subgroups.

One concluding remark about the effects of stress on learning and memory is that stress seems to influence the modulation of memories through its effects on attention and vigilance during the encoding of information. This may be true both for being in a stressful situation and for anticipating of a fear-provoking event. Knowledge about the biological and psychological mechanisms behind the effects of stress on attention and memory storage may help individuals gain control over these processes in order to perhaps make use of the positive effects of stress on memory and learning and avoid the negative.

Possible Mechanisms of the Relation between Stress and Memory

It is possible to differentiate among physiological, cognitive, and emotional stress, which sometimes may partly overlap. These types of stress might be analyzed from two angles: context and time, and whether the stress is good or bad for learning and memory. Two types of learning that are enhanced by stress are classical fear and eyeblink conditioning, as well as processes related to learning about threatening stimuli. In the standard paradigm of eyeblink conditioning, the Pavlovian eyeblink conditioning, a human is exposed to an air puff to the eye, which evokes an involuntary eyeblink reflex as an unconditioned response (UR). When the air puff is directly preceded by a sound as the conditioned stimulus (CS), the human learns the association between the tone and the air puff and blinks in response to the tone, thus anticipating the air puff. The blink is the conditioned response (CR) and is an example of associative learning. Here we are dealing with a physiological stressful event that generates, not disturbs, new associative learning. Physical stressors are generally said to activate lower brain regions involved in, e.g., pain responses, whereas psychological stressors engage the limbic structures.

Concerning the effects of stress on cognition, particularly learning and memory, one of the greatest challenges for neuropsychological research today is to determine the effects of stress on the different stages of the memory process. It is suggested that stress has differential effects on distinct phases of the learning and memory processes. Consolidation

of information is found to be enhanced by stress if the stress occurs concomitantly in time and in the same context as the thing to be remembered. The recall of information is impaired, however, if the system is exposed to stress shortly before the retrieval of information. One illustrative example of these conjunctions is given by Joëls et al. People passing a stressful examination may have difficulties in recalling the earlier learned, required information (impaired retrieval), but at the same time the traumatic experience of not remembering is burned deeply into their memories (enhanced consolidation). Retrieval difficulties have been interpreted as a negative consequence of the release of corticosteroid hormones. Inversely, the release of corticosteroid hormones facilitate new learning, which tends to compete with or overwrite the earlier acquired information. The competition between the access to earlier information and current challenges has been conceived as a kind of fruitful adaptation, since knowledge about new threats (i.e., risk of failure on the examination) may enhance the chances of future survival, in this context a better preparation for the next examination.

These findings are congruent with results showing that exposition for a psychological stressor preserves or enhances the memory for the emotional aspect of an event, while it simultaneously disrupts the memory for non-emotional or neutral aspects of the same event. This pattern may be explained by the fact that emotional and non-emotional memories are modulated by different structures of the brain. Encoding and retrieval of emotional memories are associated with the amygdala, whereas the encoding and retrieval of non-emotional memories are associated with the hippocampal formation. This pattern might also explain the gaps in memory and general memory impairment seen in patients suffering from traumatic memories.

Summary

Stress refers to an extra load on the organism. Neuroendocrinologically, stress reactions are characterized by the release of the stress hormones corticosterone in animals and cortisol in humans. Normally, increasing demands lead to increasing performance. However, this relationship holds up to a certain optimal level; thereafter, increasing demands lead to decreasing performance. This appears to be true for animal as well as human behavior in the domain of learning and memory. The mnemonic process in animals is characterized by associations between a few or several elements. Memory and learning in humans are more complex and may be

analyzed in two main systems: declarative and non-declarative memory. In animals as well as in humans, the same cerebral structures have been identified as crucial in the mnemonic process: the hippocampus and the amygdala. The hippocampus is critical for learning and memory, whereas the amygdala is responsible for the transmission of emotional information during the mnemonic process. Stress hormone receptors in the hippocampus and amygdala make these cerebral structures responsive to stress. The frontal lobes are also involved in the memory process, through the region's accumulation of stress hormone receptors and rich connections to other parts of the brain. It is possible to differentiate among physiological, cognitive, and emotional stress, which might be analyzed from a time and context perspective and with respect to whether they are good or bad for memory. Generally, declarative memory is more sensitive to stress than nondeclarative memory. Stress has different effects on the declarative memory depending on which phase in the memory process that it occurs.

See Also the Following Articles

Amygdala; Cognition and Stress; Hippocampus, Overview; Memory and Stress; Memory Impairment.

Further Reading

- de Quervain, D. J., Roozendaal, B. and McGaugh, J. L. (1998). Stress and glucocorticoids impair retrieval of long-term spatial memory. *Nature* **394**, 787–790.
- Joëls, M., et al. (2006). Learning under stress: how does it work? *Trends in Cognitive Science* **10**, 153–158.
- Kim, J. J., Song, E. U. and Kosten, T. A. (2006). Stress effects in the hippocampus: synaptic plasticity and memory. *Stress* **9**, 1–11.
- Krystal, J. H., et al. (1995). Toward a cognitive neuroscience of dissociation and altered memory functions in posttraumatic stress disorder. In: Friedman, M. J., Charney, D. S. & Deutch, A. Y. (eds.) *Neurobiological and clinical consequences of stress*, p. 249. Philadelphia, PA: Lippincott-Raven.
- Lazarus, R. S. and Folkman, S. (1986). Cognitive theories of stress and the issue of circularity. In: Appley, M. H. & Trumbull, R. (eds.) *Dynamics of stress*, p. 63. New York: Plenum.
- McEwen, B. S. and Sapolsky, R. M. (1995). Stress and cognitive function. *Current Opinion in Neurobiology* **5**, 205–216.
- Payne, J. D., et al. (2006). The impact of stress on neutral and emotional aspects of episodic memory. *Memory* **14**, 1–16.
- Peterson, L. R. (1967). Short-term memory. In: McGaugh, J. L., Weinberger, N. M. & Whalen, R. E. (eds.) *Psychobiology: readings from Scientific American*, pp. 141. San Francisco, CA: Freeman.

- Pryce, C., Mohammed, A. H. and Feldon, J. (eds.) (2002). Environmental manipulations in rodents and primates: insights into pharmacology, biochemistry and behaviour. *Pharmacological and Biochemical Behavior* 73, Special Issue.
- Sandi, C. and Rose, S. P. (1994). Corticosterone enhances long-term retention in one-day-old chicks trained in a weak passive avoidance learning paradigm. *Brain Research* 647, 106–112.
- Shalev, A. Y. (1996). Stress versus traumatic stress. In: van der Kolh, B. A., McFarlane, A. C. & Weisaeth, L. (eds.) *Traumatic stress*, pp. 93. New York: Guildford Press.
- Shors, T. J. (2006). Stressful experience and learning across the lifespan. *Annual Review of Psychology* 57, 55–85.
- Yerkes, R. M. and Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Computational and Neurological Psychology* 18, 459–482.

Lebanon, Studies of Stress in See: Combat, Acute Reactions to; Combat Reaction, Chronic; Korean Conflict, Stress Effects of; Posttraumatic Stress Disorder – Clinical; Posttraumatic Stress Disorder Overview; Traumatic Stress and Posttraumatic Stress Disorder, the Israeli Experience; War-Related Posttraumatic Stress Disorder, Treatment of.

Left Ventricular Mass

T G Pickering

New York Presbyterian Hospital, New York, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 610–611, © 2000, Elsevier Inc.

ness of its walls. The size of the ventricle normally varies with the size of the person, so for comparative purposes it is customary to normalize left ventricular mass to body surface area, as the left ventricular mass index.

Glossary

Echocardiography

A noninvasive technique for visualizing the structure and function of the heart using ultrasound. In its simplest form, a single beam (m mode) of ultrasound is aimed at the heart from a transducer placed on the chest wall, and reflections are detected from the cardiac structures hit by the beam. Two-dimensional echocardiography uses an array of beams that give a two-dimensional picture of the heart. Because traces can be recorded over several cardiac cycles, it is possible to gain information about the function of the heart as well as its structure.

Left ventricular hypertrophy

A left ventricular mass index that is above the upper limit of normal for the individual's gender and body size.

Left ventricular mass

Mass of the muscle of the left ventricle of the heart, calculated from echocardiographic measurements of the size of the chamber of the ventricle and the thick-

ness of its walls. The size of the ventricle normally varies with the size of the person, so for comparative purposes it is customary to normalize left ventricular mass to body surface area, as the left ventricular mass index.

Left ventricular (LV) mass is the weight of the left ventricle, typically estimated using echocardiography, and is thought to represent the cumulative effect of blood pressure on the heart. It is normally closely related to body size and is greater in men than in women. It also increases with age. It correlates more closely with blood pressure measured over prolonged periods of time (e.g., over several years in the Framingham heart study or by 24-h blood pressure monitoring) than with routine office measurements, but this correlation does not exceed 0.5, so that only 25% of the variance of LV mass can be explained by the blood pressure level. LV hypertrophy (defined as an LV mass that exceeds the normal range) is seen in about 30% of people with hypertension and is more common in older and more obese patients. The importance of LV mass also lies in the fact that it is an independent risk factor for cardiovascular morbidity; in fact, no other biological variable apart from age predicts cardiac risk better than LV hypertrophy. This may be partly because it represents a superior measure of the effects of increased blood pressure on the circulation, but it may also have direct pathological

significance. Thus a high LV mass is associated with a decreased coronary flow reserve (i.e., the heart outgrows its blood supply) and it also makes the heart more susceptible to arrhythmias. Treatment of hypertension with either drugs or lifestyle measures (e.g., weight reduction) can reduce LV mass, and there is some evidence that the regression of LV hypertrophy is associated with an improved survival.

LV mass is usually measured by echocardiography, which is considered to be the gold standard. This is a noninvasive ultrasound test performed by holding an ultrasonic probe on the chest wall, which emits a single beam (m mode) or a two-dimensional array of beams, and detects the reflected beams from the tissues of the heart. Magnetic resonance imaging also enables visualization of the heart and may give an even more accurate assessment, but is also much more expensive.

While blood pressure is one of the major determinants of LV mass, there are others that may be of equal significance. Twin studies have shown that there is a genetic component, which may be linked closely to the genetic regulation of body weight. LV mass is related to salt intake, and the activity of the sympathetic nervous system and renin-angiotensin systems may also increase it; in healthy young adults, plasma angiotensin II is one of the determinants of LV mass. Like 24-h ambulatory blood pressure, measures of LV geometry, including LV mass and wall thickness, have provided important information about the significance of behavioral factors relating to heart disease due to hypertension. Studies have shown that occupational stress (measured as job strain) in men is associated not only with hypertension, but also with increased LV mass. The propensity to show exaggerated blood pressure responses to stress may also be associated with greater LV mass in African-American and white adolescents and adults. What remains uncertain is the causal relationship between these observations: (1) it could be that the increased LV mass is associated with structural changes in the arteries, which are responsible for the increased reactivity; (2) the left ventricle could produce an exaggerated response leading to the increased reactivity; or (3) the increased reactivity could be the cause of the increased LV mass. For the same level of clinic blood pressure, African-Americans have a greater LV mass

than white people. The reason for this is unknown, but one contributory factor could be the comparatively higher nocturnal blood pressure in black people.

See Also the Following Articles

Blood Pressure; Cardiovascular System and Stress; Heart Disease/Attack.

Further Reading

- Allen, M. T., Matthews, K. A. and Sherman, F. S. (1997). Cardiovascular reactivity to stress and left ventricular mass in youth. *Hypertension* 30, 782–787.
- de Simone, G., Ganau, A., Verdecchia, P., et al. (1994). Echocardiography in arterial hypertension: when, why and how? *Journal of Hypertension* 12, 1129–1136.
- Esler, M. (1996). The relation of human cardiac sympathetic nervous activity to left ventricular mass: commentary. *Journal of Hypertension* 14, 1365–1367.
- Harrap, S. B., Dominiczak, A. F., Fraser, R., et al. (1996). Plasma angiotensin II, predisposition to hypertension, and left ventricular size in healthy young adults. *Circulation* 93, 1148–1154.
- Koren, M. J., Devereux, R. B., Casale, P. N., et al. (1991). Relation of left ventricular mass and geometry to morbidity and mortality in uncomplicated essential hypertension. *Annals of Internal Medicine* 114, 345–352.
- Koren, M. J., Mensah, G. A., Blake, J., et al. (1993). Comparison of left ventricular mass and geometry in black and white patients with essential hypertension. *American Journal of Hypertension* 6, 815–823.
- Langenfeld, M. R. and Schmieder, R. E. (1995). Salt and left ventricular hypertrophy: what are the links? *Journal of Human Hypertension* 9, 909–916.
- Levy, D., Garrison, R. J., Savage, D. D., et al. (1990). Prognostic implications of echocardiographically determined left ventricular mass in the Framingham Heart Study. *New England Journal of Medicine* 322, 1561–1566.
- Schnall, P. L., Devereux, R. B., Pickering, T. G., et al. (1992). The relationship between “job strain,” workplace diastolic blood pressure, and left ventricular mass index: a correction. *Journal of the American Medical Association* 267, 1209.
- Verhaaren, H. A., Schieken, R. M., Mosteller, M., et al. (1991). Bivariate genetic analysis of left ventricular mass and weight in pubertal twins (the Medical College of Virginia twin study). *American Journal of Cardiology* 68, 661–668.

Leishmania, Stress Response in

M Shapira

The Ben Gurion University of the Negev,
Beer Sheva, Israel

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
M Shapira, volume 2, pp 612–617, © 2000, Elsevier Inc.

Leishmania: The Organism and Its Life Cycle
Axenic Amastigotes: The Stress Response as a
Differentiation Trigger
Gene Regulation in *Leishmania*: Studying Heat Shock
Genes as Models
Heat Shock and Cell Death in *Leishmania*
Redox Control, Oxidative Stress, and *Leishmania*
Summary

Glossary

<i>Amastigote</i>	The life stage that resides within mammalian hosts, mostly in macrophages. An obligatory intracellular form.
<i>Cutaneous leishmaniasis</i>	A form of leishmaniasis that produces skin lesions, usually self-healing (<i>L. major</i> , <i>L. amazonensis</i>).
<i>Leishmania</i>	A protozoan parasite that belongs to the Kinetoplastida order. The parasites cycle between sand fly vectors and mammalian hosts, causing a wide range of clinical manifestations.
<i>Metacyclic form</i>	The infective life stage that resides in the sand fly vector, mainly in the proboscis.
<i>Mucocutaneous leishmaniasis</i>	A form of leishmaniasis in which the mucosal organs, primarily the nose and mouth, are attacked (<i>L. braziliensis</i> , <i>L. amazonensis</i>).
<i>Promastigote</i>	The parasite life stage that resides within the intestinal tract of the sandfly vector.
<i>Visceral leishmaniasis</i>	A systemic form of leishmaniasis in which visceral organs, mainly the liver and spleen, serve as targets for infection by parasites. Visceralization occurs in infections by certain species, such as <i>L. donovani</i> , <i>L. infantum</i> , and <i>L. ethiopica</i> . The disease is lethal if left untreated.

Leishmania: The Organism and Its Life Cycle

Leishmania parasites are flagellated diploid protists that belong to the order Kinetoplastida. This order is named after the kinetoplast, a unique organelle that harbors the mitochondrial DNA. Parasites of the genus *Leishmania* are the causative agents of a broad spectrum of human diseases. Different species of *Leishmania* vary in their clinical manifestations, ranging from lesions of the skin and mucosal organs to the lethal visceral form. Drug treatment is available, though toxic, and there is yet no safe vaccine available. *Leishmania* are distributed in large areas of the old and new world, causing both morbidity and mortality.

The parasites lead a digenetic life cycle, migrating between insect vectors and mammalian hosts. They exist as extracellular flagellated promastigotes in the alimentary tract of the sand fly vector. Promastigotes from the midgut of the sand fly are noninfective; however, upon proliferation they differentiate from procyclics into a virulent metacyclic form and migrate to the front part of the intestinal track. During the blood meal of female flies, the parasites are transmitted to the mammalian host, where they become obligatory intracellular amastigotes that reside within the phagolysosome vacuoles in macrophages. At all life stages *Leishmania* encounter hostile environments, which are characterized by high hydrolytic and proteolytic activities. Moreover, the parasites are exposed to growth conditions that differ significantly between vector and host in temperature, pH, nutrients, and serum components. Survival of *Leishmania* is facilitated by mechanisms that allow tolerance to pH and temperature shifts and resistance to complement lysis and to oxidative stress in the macrophage. Transformation includes morphogenesis and changes in metabolism and in enzymatic activities, as well as in the abundance of surface molecules. The adaptation necessary for parasite survival is acquired by alterations in gene expression, which can be modulated by the changing environment. The ability of *Leishmania* parasites to withstand the wide range of unfavorable conditions leads to their high virulence and therefore has been extensively studied.

Axenic Amastigotes: The Stress Response as a Differentiation Trigger

It has been well established that stage transformation of *Leishmania* parasites from promastigotes to amastigotes can be induced axenically upon exposure to elevated temperatures and reduced pH, conditions that mimic the environment in the mammalian host. Thus, in several *Leishmania* species, stage transformation may occur without a requirement for internalization within macrophages or other cells of the immune system. This has been shown for species of the new world, such as *L. amazonensis*, as well as for species that are distributed in the old world, such as *L. donovani*. Titration of the response to a gradual temperature elevation shows that axenic transformation is temperature restricted. It involves the increased synthesis of heat shock proteins; however, above a certain temperature threshold, translation of all other cellular proteins stops completely, preventing proper differentiation. Thus, axenic differentiation is restricted to a narrow range of temperatures, and above a certain threshold the parasites will die rather than transform. The temperature threshold varies for different species: the heat shock response in the visceralizing *L. donovani* parasites occurs at 39°C, whereas in *L. amazonensis*, a cutaneous disease agent, the stress response is induced at 33°C.

A host-free differentiation system for *L. donovani* was used to investigate the initial process of differentiation, by monitoring the induction of additional stage-specific genes. Within 1 h of exposing promastigotes to the differentiation signals that combine temperature elevation (37°C) and exposure to acidic pH (5.5), expression of the amastigote-specific A2 protein family was observed. At 5 h, morphological changes such as elimination of the flagellum and rounding up of the cells were observed, and the process was completed 7 h later. This morphological transformation occurred synchronously, while the cells were arrested at G1. Sequential exposure to elevated temperature (for 24 h) and acidic pH conditions indicated that heat was responsible for the growth arrest and the acidic pH was required for the actual completion of differentiation into amastigotes. Other inducers of the stress response such as ethanol and azetidine-2-carboxylic acid (a synthetic proline analog) were capable of replacing the heat treatment as a differentiation signal.

Expression of heat shock proteins is tightly associated with stage differentiation in *Leishmania*. Although members of the Hsp83 and Hsp70 proteins are abundantly expressed throughout the life cycle, the steady-state level of transcripts encoding these proteins increases in amastigotes, along with their

preferential translation. Exposure of *Leishmania* parasites to geldanamycin, a drug that binds the ATPase site of Hsp90 and inactivates it, induces a stress response and triggers the differentiation process from the promastigote to amastigote life stage. In the presence of geldanamycin, expression of amastigote-specific genes such as A2 increases, and the cells transform into aflagellated and rounded amastigote-like cells. This effect of the drug is specific to *Leishmania* and is not observed with *Trypanosoma cruzi*, for example, suggesting that the differentiation signals may vary between these two kinetoplastids.

As indicated, Hsp90 and Hsp70 are expressed in all life stages, although their regulatory pattern changes in response to the changing environment. However, it has been shown that thermotolerance is associated with Hsp100, which is expressed mainly in the amastigote stage of *Leishmania*. Knock-out mutants depleted from Hsp100 delay the development of skin lesions in infected mice, but did not have an apparent effect on growth of promastigotes. Virulence is partly recovered upon introduction of the missing gene into the knock-out parasite strain. The association of Hsp100 with thermotolerance has been shown for other organisms such as yeast, and it appears that this protein plays a similar role in *Leishmania*.

Gene Regulation in *Leishmania*: Studying Heat Shock Genes as Models

Genome sequencing projects of *L. major*, *T. cruzi* and *T. brucei* have been accomplished, and the sequencing of additional trypanosomatid genomes is currently under way (*L. infantum* and *L. braziliensis*). Protein-coding genes are arranged in huge unidirectional clusters that cover large portions of the chromosomes and are transcribed polycistronically. The mature, monocistronic transcripts are excised by *trans*-splicing and polyadenylation. *Trans*-splicing involves the addition of a short, capped leader to the 5' end of the RNA, which is donated by an independent RNA molecule, denoted spliced leader RNA (SL RNA). Thus, capping of mRNAs is independent of the transcribing RNA polymerase, as in other eukaryotes.

Adaptation of trypanosomatids to different environments requires a tight control of gene expression. However, as a consequence of polycistronic transcription of protein coding genes, developmental regulation occurs posttranscriptionally by RNA processing, RNA stability, and translation efficiency. This has been established using heat shock genes as a model system, and unlike in most eukaryotes, transcriptional activation of these genes has been completely ruled out in *Leishmania*. The absence of transcriptional

activation of protein coding genes fits well with the genomic organization of protein coding genes in *Leishmania* and their polycistronic transcription.

Splicing during Heat Shock

In higher eukaryotes, RNA splicing is disturbed by heat shock. Thus, heat shock genes often lack introns, except for Hsp83 in *Drosophila*. As a result, splicing of the hsp83 RNA is improved if a mild heat shock precedes a more drastic temperature stress. In *Leishmania*, temperature elevation also affects *trans*-splicing in a stage-specific manner. For example, tubulin expression is altered during the life cycle. In the amastigote stage, the flagellum is adsorbed and the subpellicular microtubules do not reach the tip of the cells, causing them to round up. In accordance, processing intermediates of α -tubulin mRNAs accumulate at elevated temperatures. Counterwise, processing intermediates of Hsp70 can be seen only at 26°C, and not at mammalian like temperatures. This is in agreement with the increased steady-state levels of Hsp70 mRNAs that are observed in amastigotes.

RNA Stability during Heat Shock

Expression of stage-specific genes can be achieved by their differential stability at altered temperatures. Thus, transcripts of promastigote-specific genes are rapidly degraded at elevated temperatures, whereas transcripts encoding for heat shock genes, such as Hsp70 and Hsp90, are differentially stabilized at elevated temperatures. The stability determinants of these transcripts are found in their 3' untranslated regions (UTRs), although they could not be associated with a specific sequence element. Almost any deletion in the 3' UTR eliminated the preferential stability of the Hsp83 mRNA at elevated temperatures, suggesting that stabilization was acquired due to a specific RNA structure, which was sensitive to any change in the sequence of the 3' UTR. Temperature control over RNA stability was observed for other genes that are differentially expressed throughout the life cycle, such as the gene encoding for a promastigote-specific protein kinase A (PKA). The PKA transcript has a much shorter $T_{1/2}$ at mammalian-like temperatures as compared to 26°C, the temperature at which promastigotes are grown. Thus, it is abundant in promastigotes but can hardly be detected in amastigotes. Control of transcript abundance during the life cycle of *Leishmania* by temperature and pH shifts was reported for the amastigote-specific A2 gene and the promastigote-specific surface protease gp63.

Preferential Translation at Elevated Temperatures

Similar to higher eukaryotes, *de novo* synthesis of heat shock genes in *Leishmania* increases dramatically at elevated temperatures, as observed in metabolic labeling experiments that measure directly the nascent incorporation of amino acids into polypeptide chains. This pattern of regulation can be transferred onto a transfected reporter gene, which is flanked by genomic sequences that include the UTRs on both sides of the coding gene, along with adjacent processing signals. Using a series of reporter constructs, it was shown that unlike in other eukaryotes, preferential translation in *Leishmania* is controlled mainly by the 3' UTR of the Hsp83 gene, whereas the 5' UTR alone can improve the effect of the 3' UTR but cannot function alone in directing preferential translation. A similar observation was made for preferential translation of the Hsp70 gene, which is also controlled by the 3' UTR. Furthermore, two types of 3' UTR were found for Hsp70 genes, which differ in their ability to promote differential accumulation and translation at elevated temperatures, enabling expression at a broad range of conditions.

Translation regulation that is based on the 3' UTR varies from what we know for Hsp70 in higher eukaryotes, where preferential translation is directed by the 5' UTR. Great efforts were invested in understanding the mechanism that promotes and increases translation of heat shock genes at elevated temperatures in higher eukaryotes. Preferential translation cannot be assigned to binding of a specific protein factor to the 5' UTR, and different reports have suggested potential mechanisms for directing preferential translation. For example, the 5' UTR of heat shock genes is rich in adenosines and thus has minimal secondary RNA structures. Since heat shock impairs the ability to unwind RNA structures, which are common in 5' UTRs of non-heat shock genes, only mRNAs with minimal structures can be efficiently translated. Another report showed that the 5' UTR of Hsp70 contains boxes of sequence complementarity with the 18S rRNA, which may facilitate ribosome shunting, a mechanism by which the 40S ribosomal subunit binds the cap and bypasses large segments of the mRNA to reach the initiation codon. Preferential translation of *Drosophila* Hsp90 is also controlled by the 5' UTR, but nucleotides close to the first AUG were implicated as essential for preferential translation at elevated temperatures.

In *Leishmania*, the 3' UTR of the Hsp83 transcript appears to confer preferential translation at elevated temperatures. Deletion analysis mapped the regulatory region that controls translation of Hsp83 in *Leishmania* to sequences 201–472 of the 886-nt-

long 3' UTR. This element is required for preferential translation of a fused reporter gene at elevated temperatures. Interestingly, it functions irrespective of transcript stability. However, the minimal fragment that can induce preferential translation of a reporter gene is larger and requires the presence of the proximal half of the 3' UTR (nts 1–472). Furthermore, mutations in the 3' UTR that eliminate the increased stability of the mRNA do not affect their preferential translation at elevated temperatures, as long as the regulatory region is present in the 3' UTR. The requirement for a relatively large regulatory region supports the possibility that a discrete RNA structure is required for promoting preferential translation during heat shock. However, to date, no specific proteins that bind to this region have been identified, and it is not clear how this 3' UTR element functions.

Another 3' UTR control element sized 450 nts directs preferential translation of amastigote-specific genes and appears to be conserved between the 3' UTR of amastin and Hsp100 of *Leishmania*. The large size of this control region may also support a potential involvement of a specific RNA structure in regulating translation of amastigote-specific genes at elevated temperatures and acidic pH. RNA structure is known to be temperature sensitive and can even serve as a thermosensor molecule in prokaryotes, as well as in eukaryote organisms.

Heat Shock and Cell Death in *Leishmania*

Environmental stress conditions may cause multiple cell damages that include protein misfolding. Heat shock proteins can function as dedicated molecular chaperones in assisting the unfolding of damaged proteins and their proper refolding during recovery from stress. However, when repair is beyond reach, cells must be eliminated. Thus, events that are triggered by cell stress are closely associated with the decision process that targets a cell for its death. Programmed cell death (PCD) is a process that prevails in multicellular organisms, but an increasing number of reports indicate that single-celled organisms can die following a process that partly resembles PCD. Several apoptotic-like processes have been described in *Leishmania*; however, the issue of apoptotic death in this parasite is still controversial. Apparently, exposure of *Leishmania* to extreme conditions triggers stage differentiation rather than cell death. Despite their resistance to a broad range of environmental conditions, it is possible that upon passing a certain threshold, a PCD-like process will be induced. Preliminary reports indicate that stationary phase parasites of *L. donovani* present some of the basic features

typical of PCD in response to treatment with antileishmanial drugs such as amphotericin B. These include nuclear condensation, nicked DNA in the nucleus, DNA ladder formation, increase in plasma membrane permeability, decrease in the mitochondrial membrane potential, and induction of a specific cysteine proteinase activity. Exposure to extreme temperatures also resulted in cell death, which could be rescued in part by overexpression of Bcl-X(L). Despite these results, major differences are observed between the apoptotic death in higher eukaryotes and the death process in *Leishmania*, trypanosomes, and other unicellular protists, such as *Tetrahymena* and *Dictyostelium*. It appears that lower eukaryotes possess the basic machinery that can drive cell death; however, the pathway of this process in these organisms is different from that found in multicellular organisms. Further studies are required to understand whether these PCD features are associated with the life cycle of *Leishmania* parasites.

Redox Control, Oxidative Stress, and *Leishmania*

Another type of stress faced by *Leishmania* parasites is oxidative stress. *Leishmania* parasites reside within macrophages, and cells of the immune system can inhibit the oxidative burst of their host cells. However, despite this feature, a strong system for detoxification of oxyradicals is also required. Thiol metabolism varies between *Leishmania* and their invertebrate vector or mammalian host cells. Redox control in kinetoplastid parasites is mediated by trypanothione, a compound that substitutes glutathione in defending the cell against chemical and oxidative stress. Trypanothione, the major reducing thiol molecule in *Leishmania* and other trypanosomatids, consists of a spermidine moiety that is linked with two glutathione molecules. It is of major importance for parasite survival, since the parasites reside within mammalian macrophages, which specialize in generating oxyradicals. It is also interesting to note that trypanothione and its metabolic pathway serve as a target for the popular antimony-based drugs. Moreover, trypanosomes lacking trypanothione reductase are avirulent and show increased sensitivity to trivalent arsenicals and oxidative stress.

Kinetoplastids, including *Leishmania*, do not encode for the glutathione S-transferase gene, and they also lack the enzyme catalase. Instead, they express the trypanothione S-transferase, which in *L. major* is associated with the eukaryotic translation elongation factor 1B (eEF1B). This association could have impli-

cations for translational control by sensing the redox state of thiol groups on proteins in the cell.

Summary

Leishmania is a parasitic protozoan that cycles between an invertebrate vector and a mammalian host and thus experiences a broad range of environmental conditions as part of its natural life cycle. Environmental stresses serve as a trigger for the parasite to differentiate from one life form to another, and as an obvious outcome, heat shock genes are abundantly expressed. Due to their conserved mode of regulation among eukaryotes, heat shock genes serve as an ideal system to study the unique molecular features of *Leishmania*, a very early and interesting eukaryote.

See Also the Following Articles

Apoptosis; Heat Shock Response, Overview; Oxidative Stress.

Further Reading

- Ahmed, R. and Duncan, R. F. (2004). Translational regulation of Hsp90 mRNA. AUG-proximal 5'-untranslated region elements essential for preferential heat shock translation. *Journal of Biological Chemistry* **279**, 49919–49930.
- Alzate, J. F., Barrientos, A. A., Gonzalez, V. M. and Jimenez-Ruiz, A. (2006). Heat-induced programmed cell death in *Leishmania infantum* is reverted by Bcl-X(L) expression. *Apoptosis* **11**, 161–171.
- Argaman, M., Aly, R. and Shapira, M. (1994). Expression of the heat shock protein 83 in *Leishmania* is regulated post transcriptionally. *Molecular and Biochemical Parasitology* **64**, 95–110.
- Barak, E., Amin-Spector, S., Gerliak, E., Goyard, S., Holland, N. and Zilberstein, D. (2005). Differentiation of *Leishmania donovani* in host-free system: analysis of signal perception and response. *Molecular and Biochemical Parasitology* **141**, 99–108.
- Bates, P. (1993). Axenic culture of *Leishmania* amastigotes. *Parasitology Today* **9**, 143–146.
- Charest, H., Zhang, W. W. and Matlashewski, G. (1996). The developmental expression of *Leishmania donovani* A2 amastigote-specific genes is post-transcriptionally mediated and involves elements located in the 3'-untranslated region. *Journal of Biological Chemistry* **271**, 17081–17090.
- Davis, M. C., Ward, J. G., Herrick, G. and Allis, C. D. (1992). Programmed nuclear death: apoptotic-like degradation of specific nuclei in conjugating *Tetrahymena*. *Developmental Biology* **154**, 419–432.
- Debrabant, A. and Nakhasi, H. (2003). Programmed cell death in trypanosomatids: is it an altruistic mechanism for survival of the fittest? *Kinetoplastid Biology and Disease* **2**, 7.
- Fairlamb, A. H., Blackburn, P., Ulrich, P., Chait, B. T. and Cerami, A. (1985). Trypanothione: a novel bis(glutathionyl)spermidine cofactor for glutathione reductase in trypanosomatids. *Science* **227**, 1485–1487.
- Folgueira, C., Quijada, L., Soto, M., Abanades, D. R., Alonso, C. and Requena, J. M. (2005). The translational efficiencies of the two *Leishmania infantum* HSP70 mRNAs, differing in their 3'-untranslated regions, are affected by shifts in the temperature of growth through different mechanisms. *Journal of Biological Chemistry* **280**, 35172–35183.
- Garlapati, S., Dahan, E. and Shapira, M. (1999). Effect of acidic pH on heat shock gene expression in *Leishmania*. *Molecular and Biochemical Parasitology* **100**, 95–101.
- Hess, M. A. and Duncan, R. F. (1996). Sequence and structure determinants of *Drosophila* Hsp70 mRNA translation: 5'-UTR secondary structure specifically inhibits heat shock protein mRNA translation. *Nucleic Acids Research* **24**, 2441–2449.
- Ivens, A. C., et al. (2005). The genome of the kinetoplastid parasite *Leishmania major*. *Science* **309**, 436–442.
- Johansson, J., Mandin, P., Renzoni, A., Chiaruttini, C., Springer, M. and Cossart, P. (2002). An RNA thermometer controls expression of virulence genes in *Listeria monocytogenes*. *Cell* **110**, 551–561.
- Jolly, C. and Morimoto, R. I. (2000). Role of the heat shock response and molecular chaperones in oncogenesis and cell death. *Journal of the National Cancer Institute* **92**, 1564–1572.
- Krobitsch, S., Brandau, S., Hoyer, C., Schmetz, C., Hubel, A. and Clos, J. (1998). *Leishmania donovani* heat shock protein (100). Characterization and function in amastigote stage differentiation. *Journal of Biological Chemistry* **273**, 6488–6494.
- Lee, N., Bertholet, S., Debrabant, A., Muller, J., Duncan, R. and Nakhasi, H. L. (2002). Programmed cell death in the unicellular protozoan parasite *Leishmania*. *Cell Death and Differentiation* **9**, 53–64.
- Liang, X. H., Haritan, A., Uliel, S. and Michaeli, S. (2003). Trans and cis splicing in trypanosomatids: mechanism, factors, and regulation. *Eukaryotic Cell* **2**, 830–840.
- Martinez-Calvillo, S., Yan, S., Nguyen, D., Fox, M., Stuart, K. D. and Myler, P. J. (2003). Transcription of *Leishmania major* Friedlin chromosome 1 initiates in both directions within a single region. *Molecular Cell* **11**, 1291–1299.
- McGarry, T. J. and Lindquist, S. (1985). The preferential translation of *Drosophila* hsp70 mRNA requires sequences in the untranslated leader. *Cell* **42**, 903–911.
- McNicoll, F., Muller, M., Cloutier, S., et al. (2005). Distinct 3'-untranslated region elements regulate stage-specific mRNA accumulation and translation in *Leishmania*. *Journal of Biological Chemistry* **280**, 35238–35246.
- Morita, M. T., Tanaka, Y., Kodama, T. S., Kyogoku, Y., Yanagi, H. and Yura, T. (1999). Translational induction of heat shock transcription factor sigma32, evidence for a built-in RNA thermosensor. *Genes and Development* **13**, 655–665.

- Purdy, J. E., Donelson, J. E. and Wilson, M. E. (2005). Regulation of genes encoding the major surface protease of *Leishmania chagasi* via mRNA stability. *Molecular and Biochemical Parasitology* **142**, 88–97.
- Quijada, L., Soto, M., Alonso, C. and Requena, J. M. (2000). Identification of a putative regulatory element in the 3'-untranslated region that controls expression of HSP70 in *Leishmania infantum*. *Molecular and Biochemical Parasitology* **110**, 79–91.
- Shapira, M., McEwen, J. G. and Jaffe, C. L. (1988). Temperature effects on molecular processes which lead to stage differentiation in *Leishmania*. *EMBO Journal* **7**, 2895–2901.
- Siman-Tov, M. M., Aly, R., Shapira, M. and Jaffe, C. J. (1996). Cloning from *Leishmania major* of a developmentally regulated gene, *c-lpk2*, for the catalytic subunit of the cAMP-dependent protein kinase. *Molecular and Biochemical Parasitology* **77**, 201–207.
- Vickers, T. J. and Fairlamb, A. H. (2004). Trypanothione S-transferase activity in a trypanosomatid ribosomal elongation factor 1B. *Journal of Biological Chemistry* **279**, 27246–27256.
- Wiesig, M. and Clos, J. (2001). Heat shock protein 90 homeostasis controls stage differentiation in *Leishmania donovani*. *Molecular and Biological Cell* **12**, 3307–3316.
- Wyllie, S., Cunningham, M. L. and Fairlamb, A. H. (2004). Dual action of antimonial drugs on thiol redox metabolism in the human pathogen *Leishmania donovani*. *Journal of Biological Chemistry* **279**, 39925–39932.
- Yueh, A. and Schneider, R. J. (1996). Selective translation initiation by ribosome jumping in adenovirus-infected and heat-shocked cells. *Genes and Development* **10**, 1557–1567.
- Zilka, A., Garlapati, S., Dahan, E., Yaolsky, V. and Shapira, M. (2001). Developmental regulation of heat shock protein 83 in *Leishmania*. 3' processing and mRNA stability control transcript abundance, and translation is directed by a determinant in the 3'-untranslated region. *Journal of Biological Chemistry* **276**, 47922–47929.

Leptin, Adiponectin, Resistin, Ghrelin

A Gavrilu, D Barb and C S Mantzoros

Beth Israel Deaconess Medical Center, Boston, MA, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction

Adipose Tissue

Gastrointestinal Tract: Ghrelin

Conclusion

Glossary

<i>Adipocytokine</i>	Cytokines secreted by adipose tissue (i.e., leptin, adiponectin, and resistin).
<i>Adiponectin</i>	Hormone secreted by fat cells that modulates several physiological processes, such as glucose and fatty acid metabolism. Decreased plasma adiponectin levels are associated with obesity, insulin resistance, type 2 diabetes mellitus, and atherosclerosis.
<i>Adipose tissue</i>	Specialized connective tissue composed of fat cells (adipocytes), which store droplets of fat in the form of triglycerides and have many other important metabolic roles.
<i>Adipostatic peptide</i>	An agent that prevents adipocyte/fat cell accumulation.

Anorexia nervosa

Anorexigenic peptide
Cachexia

Cytokine

Ghrelin

Insulin

Leptin

An eating disorder that occurs most frequently in adolescent females and is characterized by the lack or loss of appetite, known as anorexia. Other features include excessive fear of becoming overweight, body image disturbance, significant weight loss, refusal to maintain minimal normal weight, and amenorrhea. An agent that transiently depresses food intake.

Extreme weight loss, usually associated with chronic disease.

A protein secreted by inflammatory leukocytes but also by nonleukocytic cells; cytokines act as intercellular mediators and differ from classic hormones in that they are produced by a number of tissues or cell types rather than by specialized glands.

A peptide produced and secreted from stomach that stimulates growth hormone release and, in addition, increases appetite and food intake.

A protein hormone secreted by the pancreas that plays a major role in the regulation of glucose metabolism, promoting the cellular use of glucose; also an important regulator of protein and lipid metabolism.

*Orexigenic
peptide
Resistin*

A hormone secreted by fat cells and implicated in the regulation of food intake and energy balance, depressing appetite and increasing energy expenditure. An agent that stimulates food intake.

A hormone secreted by fat cells that may play a role in insulin resistance and inflammation associated with obesity.

Introduction

Obesity has become a public health problem due to its increasing prevalence and associated morbidity and mortality, mainly from cardiovascular disease, type 2 diabetes, and certain malignancies. In addition to a genetic predisposition, environmental factors, physical inactivity, a high-energy high-fat diet, smoking, and stress (which results in dysregulation of endogenous hormones) strongly contribute to the development of this multifactorial disorder.

Obesity is known to be associated with insulin resistance and development of the metabolic syndrome. However, not all obese patients are insulin-resistant and total fat mass *per se* is not the most important factor contributing to insulin resistance but, rather, the presence of central or visceral obesity. Central obesity has been associated with the dysregulation of the hypothalamic-pituitary-adrenal (HPA) axis and the sympathetic nervous system. The stress hormones of the HPA axis (i.e., cortisol) and the sympathetic system (catecholamines) oppose the effects of insulin, and this may promote insulin resistance. A chronically stressed HPA axis leads to decreased circadian variations in circulating cortisol levels, and chronic glucocorticoid exposure may lead to increased food intake and obesity.

In parallel with the increasing prevalence of obesity, there has been a tremendous increase in research activity in a global effort to better understand energy balance and the regulation of metabolism and body weight. Although a biological basis for body fat regulation was first proposed more than 50 years ago, the hypothesis of an endocrine feedback loop regulating body weight remained elusive until the discovery of the adipocyte-secreted hormone leptin in 1994. This discovery has dramatically changed our understanding of adipose tissue from an inert depot of energy storage to an active organ producing and secreting various molecules that regulate metabolism and energy homeostasis, such as leptin, adiponectin, and resistin (see Table 1). In addition, it has recently been recognized that other peripheral organs, such as the gastrointestinal tract, secrete molecules that also convey information on energy intake and balance. For exam-

ple, ghrelin, a novel gastrointestinal tract-secreted peptide, not only stimulates the secretion of growth hormone (GH) from the pituitary but also plays an important role in energy homeostasis, causing a positive energy balance by stimulating food intake.

The currently accepted model of energy homeostasis proposes that peripheral signals, such as molecules secreted into the circulation from the gut and the adipose tissue, are received and integrated by the central nervous system together with other signals/modulators of food intake and energy homeostasis, such as visual stimuli, smell, the presence of food, social habits or social behavior, and stress, with the final purpose to match food intake with energy expenditure. Signals from both adipose tissue and other peripheral organs not only are processed centrally in the brain to control body metabolism and energy but may also influence other physiological functions, such as the neuroendocrine and reproductive systems.

Various central and peripheral factors involved in energy homeostasis are being continuously discovered, and it is expected that a better understanding of these mechanisms will lead to effective treatments for the control of metabolism and obesity. This review discusses the potential role of four recently discovered molecules involved in metabolism and energy homeostasis: leptin, adiponectin, resistin, and ghrelin.

Adipose Tissue

The adipose tissue is considered today to be a very active and complex endocrine organ. Adipocytes produce and secrete a variety of bioactive molecules, including adipokines, which act at both the local (paracrine/autocrine) and systemic (endocrine) levels to exert diverse metabolic functions. These adipokines include leptin, tumor necrosis factor (TNF)- α , and interleukin (IL)-6, which have already been intensively studied, as well as newer molecules, such as adiponectin and resistin, which may prove to have important roles in controlling metabolism and energy homeostasis.

Leptin

Leptin (from Greek *leptos*, thin) is a 167-amino-acid protein discovered in 1994 as the product of the obese (*ob*) gene. Although known to have numerous other physiological functions in mice and humans, leptin is primarily involved in the control of the neuroendocrine and immune response to energy deprivation. Leptin is produced predominantly by adipose tissue and is secreted into the circulation at levels that are directly proportional to fat mass. Leptin is also expressed at low levels in the hypothalamus,

Table 1 Adipocytokines and ghrelin: effects on energy homeostasis^a

<i>Molecule</i>	<i>Site of production</i>	<i>Circulating levels</i>	<i>Target organ/tissue</i>	<i>Effect on insulin sensitivity</i>	<i>Effect on energy expenditure</i>	<i>Stress interaction</i>	<i>Other effects</i>
Adiponectin	Adipose tissue (primary)	Decreased in obesity, type 2 DM, CVD, malignancies	Muscle	Increases insulin sensitivity Stimulates glucose uptake Stimulates fatty acid oxidation Decreases gluconeogenesis	Increases energy expenditure Decreases body weight without affecting food intake (ICV administration)	Exo- or endogenous glucocorticoids inhibit adiponectin secretion	Anti-inflammatory Anti-atherogenic Anti-carcinogenic
	Cardiomyocytes ^b Placenta ^b		Liver Hypothalamus (PVN nucleus)				
Ghrelin	GI tract: stomach (enterochromaffin cells) (primary)	Increased acutely before each meal	Pituitary (GH-releasing cells)	May reduce insulin sensitivity	Decreases energy expenditure	Ghrelin stimulates ACTH release and activates HPA	Increases gastric motility and acid secretion
	Hypothalamus (ARC nucleus)	Decreased in obesity	Hypothalamus (NPY and AgRP neurons)	Decreased glucose disposal rate during ghrelin infusion	Increases appetite and stimulates food intake	Acute psychological stress raises plasma ghrelin in rodents	Reduces locomotor activity
		Elevated in anorexia nervosa or cachexia		Increases GH release Enhances the use of carbohydrates and reduces fat use	Reduces thermogenesis	Activation of gut sympathetic nerves stimulate ghrelin secretion	
Leptin	Adipose tissue (primary)	Increased in proportion to body fat	Hypothalamus (ARC) Brain stem	Increases insulin sensitivity Stimulates glucose uptake Stimulates free fatty acid oxidation	Increases energy expenditure Decreases appetite, food intake, and body weight Increases sympathetic activity Stimulates thermogenesis	Suppresses the activity of the HPA axis (CRH, ACTH, cortisol) Glucocorticoids stimulate leptin secretion, but excess may contribute to leptin resistance	Stimulates reproductive and thyroid functions Stimulates immunity Inhibits bone formation
	Hypothalamus ^b Pituitary ^b Placenta ^b Skeletal muscle ^b Gastric and mammary epithelia ^b						
Resistin	Adipose tissue Mononuclear leukocytes Macrophages	Variable in obesity	Liver ? Muscle ? Hypothalamus ?	Decreases insulin sensitivity in rodents (impairs insulin action on hepatic glucose production and may decrease fatty acid oxidation) Controversial in humans	May decrease energy expenditure in rodents Controversial in humans	Unknown	Inhibits adipocyte differentiation in rodents May have pro-inflammatory effects

^aACTH, adrenocorticotrophic hormone; AgRP, agouti-related protein; ARC, arcuate nucleus; CRH, corticotropin releasing hormone; CVD, cardiovascular disease; DM, diabetes mellitus; GH, growth hormone; GI, gastrointestinal; HPA, hypothalamic-pituitary-adrenal; ICV, intracerebroventricular; NPY, neuropeptide Y; PVN, paraventricular nucleus.

^bAt low levels.

pituitary, placenta, skeletal muscle, and gastric and mammary epithelia. Leptin is found in the peripheral circulation in both a free and a bound form; the free form crosses the blood–brain barrier, binds to the long isoform of a specific leptin class 1 cytokine receptor (also known as ObRb) and activates mainly the janus kinase (JAK)–signal transducer and activator of transcription (STAT) and phosphatidylinositol 3 (PI3)–AKT pathways to convey information to the central nervous system regarding the amount of energy stored in adipose tissue and acute changes in caloric intake. Leptin receptors are predominantly found in the hypothalamus, mainly in the arcuate nucleus but also in the brain stem and other regions of the central nervous system. The activation of hypothalamic receptors stimulates the production of anorectic peptides, such as α -melanin-stimulating hormone (α -MSH) and cocaine and amphetamine-regulated transcript (CART), and inhibits the production of orexigenic peptides, such as neuropeptide Y (NPY) and agouti-related peptide (AgRP), resulting in the inhibition of appetite. In addition, leptin binding to hypothalamic receptors results in increased sympathetic activity, which stimulates free fatty acid oxidation and thermogenesis in brown adipose tissue, thus favoring energy expenditure in mice.

Leptin expression is mainly controlled by the amount of body fat mass; caloric restriction and weight loss results in decreased circulating leptin levels, whereas overfeeding and increasing adipose tissue mass results in increased leptin levels. Leptin production and secretion is also increased by insulin, glucocorticoids, TNF- α , and estrogens; it is decreased by β_3 -adrenergic activity, androgens, free fatty acids, growth hormone, and peroxisome proliferator activated receptor (PPAR)- γ agonists.

The direct involvement of leptin in the regulation of energy intake and expenditure (by signaling the adequacy of energy stores to the brain) was initially demonstrated in studies in mice. Mutations in the obese (*ob*) gene (leptin gene) and the diabetes (*db*) gene (leptin receptor gene) result in morbid obesity and diabetes, whereas the administration of leptin to leptin deficient (*ob/ob*) and normal rodents results in significant weight loss. However, diet-induced obese mice, the mouse model of obesity that most closely corresponds to garden-variety obesity in humans, are not able to lose weight significantly after leptin administration, indicating that these hyperleptinemic mice are leptin resistant, probably due to receptor or postreceptor defects.

Similarly, only a distinct minority of morbidly obese human subjects have congenital leptin deficiency due to mutations of the *ob* gene, which result in hyperphagia and early-onset, rapid weight gain. The

administration of exogenous leptin to these patients results in the normalization of their hyperphagia, decreased caloric intake, and a significant loss of fat mass. Rare inactivating mutations of the leptin receptor gene have also been reported, which, similar to leptin gene defects, manifest clinically with hyperphagia and extreme obesity, as well as hypogonadotropic hypogonadism, but they are also associated with growth delay and secondary hypothyroidism. Our studies have shown that increasing leptin levels precede the onset of puberty, suggesting that leptin may be a permissive factor for the onset of puberty in humans.

Although exogenous leptin administration is an effective treatment, resulting in significant weight loss in obese leptin-deficient subjects, the majority of the cases of human obesity do not respond to leptin therapy. It is now known that most obese individuals have increased leptin production and secretion from adipocytes as well as high circulating leptin levels, features that are consistent with a state of leptin resistance. Several mechanisms are involved in the leptin resistance of obese individuals, including defective transport of leptin across the blood–brain barrier into the brain and/or reduced hypothalamic leptin signaling, in part due to the upregulation of specific inhibitors of leptin signaling. A complete understanding of the pathogenesis of leptin resistance in humans may open new pathways for the treatment of obesity.

Interestingly, a small study reported that exogenous leptin administration to replace leptin levels to preweight-loss levels, prevented the regaining of weight, and promoted the loss of fat mass in a small group of subjects participating in a weight loss program. These findings need to be confirmed by larger studies.

There is accumulating evidence that leptin has complex neuroendocrine functions, regulating the activity of the HPA and also the hypothalamic-pituitary-gonadal and -thyroid axes. It has been reported that leptin suppresses the activity of the HPA; that is, decreased leptin levels during fasting/starvation are associated with increased activation of the HPA and cortisol secretion. In accordance with this, in critically ill patients the nocturnal fall in leptin levels is associated with a paradoxical increase of nocturnal adrenocorticotrophic hormone (ACTH) and cortisol levels. Leptin-deficient mice and humans fail to go through puberty, and leptin administration reverses this abnormality. We have shown that the decreased leptin levels found in food deprivation are responsible for the starvation-induced suppression of the hypothalamic-pituitary-gonadal axis as well as changes in the function of other neuroendocrine axes in humans. In addition, extremely thin women with hypothalam-

ic amenorrhea have low leptin levels and exogenous leptin administration normalizes their neuroendocrine and reproductive functions. Thus, it seems that leptin may act as the critical link between adipose tissue and not only hypothalamic centers regulating energy homeostasis but also the reproductive system, indicating whether adequate energy reserves are present for normal reproductive function. Further larger trials are needed to conclusively prove whether leptin replacement treatment can ameliorate or treat the neuroendocrine dysfunction and also the immune abnormalities, osteoporosis, and infertility in women with hypothalamic amenorrhea.

In addition to signaling in the central nervous system, leptin also acts in the periphery and has important roles in regulating insulin action and lipid metabolism. Other peripheral endocrine effects include the direct regulation of immune function, hematopoiesis, angiogenesis, and direct and indirect effects on bone metabolism.

Taken together, these findings suggest that leptin plays a significant role in the central regulation of energy balance and that leptin is a critical signal for neuroendocrine function in humans; this raises the possibility that leptin is indeed not only an adipostatic hormone but also a stress-related hormone crucial for survival.

Adiponectin

Adiponectin (Acrp30, apM1 protein, adipoQ, or GBP28) is a 247-amino-acid protein discovered simultaneously by four different research groups in the mid-1990s. Adiponectin is produced and secreted exclusively by differentiated adipocytes and represents one of the most abundant serum proteins that circulates in tightly associated trimers and higher-order oligomers. Eight adiponectin isoforms with variable potency have been isolated, which bind and activate at least two adiponectin receptors. Adiponectin receptor 1 (AdipoR1) is expressed ubiquitously but most abundantly in skeletal muscle and has a high affinity for globular adiponectin and a very low affinity for full-length adiponectin; adiponectin receptor 2 (AdipoR2) is found predominantly in the liver and has an intermediate affinity for both forms.

Adiponectin decreases with increasing overall and central adiposity and increases with prolonged weight reduction. In humans, although short-term fasting does not influence circulating levels of adiponectin, prolonged caloric restriction results in weight loss and increased adiponectin levels. In addition, daily caloric intake, macronutrient intake, and high fat intake do not correlate with and do not affect circulating adiponectin levels in humans, except possibly in obese individuals, but a Mediterranean-type diet,

which has beneficial effects on morbidity and mortality, increases adiponectin levels. Interestingly, studies in rodents have revealed that central and peripheral adiponectin administration increases energy expenditure by stimulating free fatty acid oxidation and glucose uptake, thus reducing body weight and visceral adiposity without affecting energy intake.

Although its role has not been definitively established, there is accumulating evidence that adiponectin regulates insulin sensitivity and is also involved in inflammation and atherosclerosis. Initial animal studies showed that adiponectin-knockout mice develop insulin resistance either independently of diet or only after high-fat and high-sucrose diet, whereas adiponectin treatment improves their insulin resistance. In addition, lipoatrophic and obese mice have low adiponectin levels and insulin resistance, which improve partially after exogenous adiponectin administration. Moreover, the decrease in circulating adiponectin levels parallels the development of insulin resistance and type 2 diabetes in a longitudinal study of rhesus monkeys.

Studies in humans showed that there is a negative correlation of adiponectin levels with fasting glucose, insulin, and insulin resistance and a positive correlation with insulin sensitivity, independent of body-mass index (BMI). We also know that the adiponectin gene has the same location as one of the diabetes susceptibility genes on chromosome 3q27, and several longitudinal studies showed that human subjects with lower baseline adiponectin levels are more likely to develop type 2 diabetes, independent of adiposity. In contrast, type 1 diabetic patients may have elevated adiponectin levels compared to nondiabetic individuals, whereas chronically administered insulin does not have an effect on adiponectin levels.

Taken together, all these findings show that adiponectin levels are low in states of insulin resistance, such as obesity and type 2 diabetes, and adiponectin levels increase with improved insulin sensitivity after weight reduction or treatment with insulin-sensitizing drugs (thiazolidindiones). It is currently thought that adiponectin represents a link among obesity, insulin resistance, and various other metabolic abnormalities associated with obesity.

Adiponectin improves insulin sensitivity through multiple cellular and molecular mechanisms that are not entirely known at present. An important enzyme stimulated by adiponectin is AMP-activated protein kinase (AMPK), which, in turn, phosphorylates several intracellular proteins to alter various cellular functions, including the enhancement of insulin-stimulated glucose uptake into fat and skeletal muscle as well as the suppression of hepatic glucose production and stimulation of fatty acid oxidation. By

increasing fatty acid oxidation through AMPK activation, adiponectin decreases circulating free fatty acids, resulting thus in improved insulin action. Adiponectin may also improve insulin sensitivity by inducing the activation of the insulin signaling system. Moreover, by stimulating nitric oxide (NO) formation, adiponectin may augment vascular blood flow to promote glucose uptake.

In addition to its insulin-sensitizing effects, adiponectin is also related to the lipid profile, another independent risk factor for cardiovascular disease. Previous studies reported a strong positive association between adiponectin and high-density lipoprotein (HDL) cholesterol and a negative association with serum triglyceride levels. However, there is no independent relationship between adiponectin and low-density lipoprotein (LDL) or total cholesterol. Adiponectin may lead to a favorable lipid profile by stimulating fatty acid oxidation, an effect probably mediated by AMPK. However, one study reported increased fatty acid oxidation in muscle cells in adiponectin-knockout mice, and a cross-sectional study of humans found that plasma adiponectin levels did not correlate with lipid oxidation, as measured by energy expenditure and respiratory quotient. Thus, further studies are needed to understand adiponectin's effects on fatty acid oxidation.

Adiponectin may also exert important anti-inflammatory and atheroprotective effects. It inhibits the production and action of several inflammatory factors involved in atherosclerosis, including TNF- α , and also suppresses monocyte adhesion and macrophage transformation to foam cells within the vascular wall. In addition, adiponectin inhibits smooth muscle cell proliferation in the vascular wall by inhibiting growth factors that promote hyperplasia and promotes vasodilation and angiogenesis by increasing NO production in endothelial cells through AMPK activation. In conclusion, adiponectin's involvement in cardiovascular disease is probably multifactorial, including its effects on risk factors such as insulin resistance and dyslipidemia but also a direct effect on the vascular system.

Furthermore, there is emerging mechanistic and epidemiological evidence supporting a protective role of adiponectin in malignancies. *In vitro* treatment of acute myelomonocytic leukemia cell lines with adiponectin suppressed the growth of myelomonocyte cells, induced apoptosis, and downregulated *Bcl-2* gene expression. More recently, adiponectin has been shown to be a direct inhibitor of angiogenesis *in vivo* and *in vitro* through the activation of caspases in endothelial cells. Consistent with the *in vitro* studies, which suggest that adiponectin is an important negative regulator of hematopoiesis and the immune

system, *in vivo* intratumoral administration of murine adiponectin resulted in significant suppression of T241 fibrosarcoma tumor growth. Adiponectin receptors are expressed by prostate, hepatocellular carcinoma, and various other cancer cell types (breast, endometrial, colon, and neuroblastoma cell lines). Recent experiments suggest that adiponectin may act on tumor cells directly through the activation of c-jun NH₂-terminal kinase, a mammalian MAPK family involved in the regulation of cell proliferation and apoptosis, and also through the suppression of the transcription factor signal transducer and activator of transcription (STAT)3, which also regulates cell proliferation, differentiation, survival, and apoptosis. Therefore, given its antiangiogenesis properties and the inactivation of STAT3 reported in obesity and associated malignancies, adiponectin may prove to be an effective novel anticancer agent.

Epidemiological evidence also suggests that adiponectin plays an important role in the insulin-related pathways of carcinogenesis, independently of other known risk factors for cancers. We and others have recently reported that circulating adiponectin levels are inversely associated with risk for endometrial, breast, prostate, gastric, and colorectal cancer. In two recent case-control studies, we reported that women with high BMI and low plasma adiponectin had a 6.5-fold increased risk of endometrial cancer, independent of the possible effects of insulin-like growth factor (IGF)-1, IGF-2, IGF binding protein (IGFBP)-3, leptin, BMI, and other known risk factors of the disease. We also found in a prospective study in men that individuals in the highest quintile of adiponectin levels had an approximately 60% reduced risk for colorectal cancer compared to the lowest quintile. Significantly, these associations were also independent of body size and physical activity.

In addition, adiponectin levels may also correlate inversely with tumor size, histological grade, and disease stage in gastric and prostate cancer. Large association and prospective studies to prove adiponectin's role in the development of malignancies as well as mechanistic studies to address potential roles of adiponectin in cancer pathogenesis are needed.

Resistin

Resistin (adipose-tissue-specific secretory factor or FIZZ3) is a 94-amino-acid polypeptide with 11 cysteine residues. Resistin is almost entirely expressed and excreted from white adipose tissue in mice and circulates as a dimer. A single dicysteine residue is required for dimerization, whereas the remaining 10 cysteine residues are involved in determining the structure of the resistin monomeric unit.

Similar to adiponectin, resistin has been postulated to link obesity with insulin resistance, but these two molecules probably have opposite functions. In 2001, Steppan and colleagues reported that resistin secretion is decreased by the antidiabetic drug rosiglitazone and is increased in diet-induced and genetic mouse models of obesity. Moreover, the administration of antiresistin antibody improves blood sugar and insulin action in obese mice, and the administration of recombinant resistin impairs glucose tolerance and insulin action in normal mice. More recent studies have shown that resistin may increase hepatic insulin resistance by inhibiting AMPK and may also increase hepatic glucose production, thus controlling glucose and lipid metabolism. In addition, resistin has been postulated to regulate adipose tissue mass by inhibiting adipocyte differentiation and thus limit adipose tissue formation in response to increased energy intake. These observations have not been confirmed by other studies, however; thus, the role of resistin in mice remains controversial.

Human data are even more conflicting. Although resistin is expressed at very low levels in human adipose cells, high levels are expressed in circulating mononuclear leukocytes and macrophages. Similar to its distribution in rodents, resistin is found in human sera, and several studies (but not all) reported increased levels in obese and diabetic patients and a positive association with insulin resistance. One study showed that human resistin expression, but not serum levels, is proportional to insulin resistance, and several studies failed to show an association with insulin resistance. Resistin may have pro-inflammatory effects because recent *in vitro* studies found that resistin induces the expression of adhesion molecules such as vascular cell adhesion molecule (VCAM)-1 and intercellular adhesion molecule (ICAM)-1 in human endothelial cells.

Further studies are required to determine whether stress influences resistin levels and to elucidate the role of resistin in regulating insulin sensitivity. In addition, whether resistin influences obesity either directly or indirectly by altering glucose and insulin levels and whether resistin plays a role in inflammation associated with obesity warrant further investigation.

Gastrointestinal Tract: Ghrelin

Ghrelin, a 28-amino-acid lipophilic peptide with a labile octanoyl fatty acid side chain at the serine residue 3, was originally discovered by Kojima and coworkers during their search for an endogenous growth hormone secretagogue. This hormone is expressed primarily in specialized enterochromaffin cells (called X/A-like cells), which are located mainly

in the mucosa of the fundus of the stomach. In addition, ghrelin-immunoreactive cells are also found in the intestine, with their concentration gradually decreasing from duodenum to colon; they are also found in the pancreatic islets, brain, and kidney. Ghrelin-producing neurons are located in the hypothalamus, more specifically in the hypothalamic arcuate nucleus, an important region controlling appetite, but also in previously uncharacterized neurons adjacent to the third ventricle, which send efferent fibers to NPY and AgRP neurons, stimulating the release of these orexigenic peptides.

Ghrelin binds to the growth hormone secretagogue receptor (GHS-R), a typical G-protein-coupled receptor, the activation of which by ghrelin results in increased intracellular calcium via the activation of the phospholipase C pathway. GHS-Rs are expressed in various regions of the brain, predominantly in the growth hormone-releasing somatotroph cells of the pituitary gland. Ghrelin stimulates the release of growth hormone from the pituitary in a dose-dependent manner and acts synergistically with growth hormone-releasing hormone (GHRH). In addition, ghrelin has a weak prolactin (PRL) and ACTH-releasing activity, but more important ghrelin can modulate gonadotropin-releasing hormone (GnRH) pulses, possibly being one of the signals that turns off the gonadal axis during starvation. The expression of ghrelin in the pituitary is high after birth and declines with puberty.

In addition to being a potent growth hormone secretagogue, ghrelin is a potent appetite stimulant. Ghrelin administration by intracerebroventricular, intravenous, or subcutaneous injection increases food intake and body weight, decreases fat use, and increases gastric acid output in a dose-dependent manner. Administered to rats by intracerebroventricular injections, ghrelin induces Fos protein expression, a marker of neuronal activation, in regions of primary importance in the regulation of feeding, including NPY neurons and AgRP neurons, but also in the neurons of the solitary tract nucleus and the dorsomotor nucleus of vagus. The latter may sustain a role for ghrelin in the central regulation of gastric acid secretion and gastric motility via vagal cholinergic pathways.

Ghrelin circulates in the bloodstream in two forms, acylated and desacyl ghrelin, with the acylated form being the biologically active form of ghrelin. Serum levels of ghrelin depend on the state of nutrition and are negatively correlated with the BMI. Accordingly, circulating ghrelin levels are decreased in obesity but elevated during fasting and in anorexia nervosa and cachexia, a finding that may reflect a compensatory mechanism. Acute psychological stress raises plasma ghrelin in the rat, and the activation of gut sympathetic

nerves directly stimulate ghrelin secretion. Also, an increase in plasma ghrelin leads to the release of ACTH and the activation of the HPA axis. Recent human data demonstrated an inverse correlation between ghrelin and leptin that is independent of adiposity, but there is no regulation of ghrelin by leptin administration over the short term (a few hours to a few days).

In normal humans, plasma ghrelin levels rise immediately before each meal and in response to diet-induced weight loss and fall to minimum levels 1 h after feeding. This fall in ghrelin levels after a meal is correlated with insulin peak, suggesting that insulin may be a negative regulator of ghrelin secretion in the postprandial state. This fall in ghrelin levels after a meal is absent or blunted in obese humans, a fact that, researchers propose, may contribute to the development of their obesity. Significantly, obese patients who undergo gastric bypass surgery have significantly lowered ghrelin levels after surgery, which, on the basis of observational studies, has been proposed to contribute to maintaining decreased weight after bypass surgery.

Finally, what are potential clinical applications of ghrelin? It has been suggested that blocking or neutralizing ghrelin action may prove to be a reasonable pharmacological approach to reverse a chronically positive energy balance, such as in obesity. However, in contrast to predictions made from the pharmacology of ghrelin, studies in ghrelin-null mice showed that the deletion of ghrelin impairs neither growth nor appetite. In fact, ghrelin-null mice are neither anorexic nor dwarfs and display normal responses to starvation and diet-induced obesity, strongly suggesting that ghrelin is not critically required for viability, fertility, growth, appetite, or fat deposition. Thus, antagonists of ghrelin may not prove to be very successful anti-obesity agents. In contrast, exploring its orexigenic activity for the treatment of eating disorders such as anorexia nervosa or the replacement of ghrelin in ghrelin-deficient states such as cachexia or other catabolic states seems to be a more reasonable approach.

Conclusion

This brief review has covered recent advances in the area of molecules regulating energy homeostasis

and metabolism – leptin, adiponectin, resistin, and ghrelin. Most probably, we can expect other related hormones to be added to this list in the near future. In addition, synthetic analogs, agonists, or antagonists are expected to be developed by the pharmaceutical industry for the management of states of energy excess, such as obesity and type 2 diabetes, and states of energy deficiency, such as anorexia nervosa and hypothalamic amenorrhea.

See Also the Following Articles

Feeding Circuitry (and Neurochemistry); Food Intake and Stress, Non-Human; Ghrelin and Stress Protection; Glucocorticoid Negative Feedback; Metabolic Syndrome and Stress; Nutrition; Obesity, Stress and; Orexin.

Further Reading

- Ahima, R. S. (2005). Central actions of adipocyte hormones. *Trends in Endocrinology and Metabolism* **16**, 307–313.
- Barb, D., Pazaitou-Panayiotou, K. and Mantzoros, C. S. (2006). Adiponectin: a link between obesity and cancer. *Expert Opinions on Investigational Drugs* **15**, 917–931.
- Bjorbaek, C. and Kahn, B. B. (2004). Leptin signaling in the central nervous system and the periphery. *Recent Progress in Hormone Research* **59**, 305–331.
- Bluher, S. and Mantzoros, C. S. (2004). The role of leptin in regulating neuroendocrine function in humans. *Journal of Nutrition* **134**, 2469S–2474S.
- Brennan, A. M. and Mantzoros, C. S. (2006). Drug insight: the role of leptin in human physiology and pathophysiology-emerging clinical applications. *Nature Clinical Practice: Endocrinology & Metabolism* **2**(6), 318–324.
- Buren, J. and Eriksson, J. W. (2005). Is insulin resistance caused by defects in insulin's target cells or by a stressed mind? *Diabetes/Metabolism Research and Reviews* **21**(6), 487–494.
- Chan, J. L. and Mantzoros, C. S. (2005). Role of leptin in energy-deprivation states: normal human physiology and clinical implications for hypothalamic amenorrhoea and anorexia nervosa. *Lancet* **366**, 74–85.
- Gale, S. M., Castracane, V. D. and Mantzoros, C. S. (2004). Energy homeostasis, obesity and eating disorders: recent advances in endocrinology. *Journal of Nutrition* **134**, 295–298.
- Haslam, D. W. and James, W. P. (2005). Obesity. *Lancet* **366**, 1197–1209.

- Havel, P. J. (2004). Update on adipocyte hormones: regulation of energy balance and carbohydrate/lipid metabolism. *Diabetes* 53(supplement 1), S143–S151.
- Kadowaki, T. and Yamauchi, T. (2005). Adiponectin and adiponectin receptors. *Endocrine Review* 26, 439–451.
- Kojima, M. and Kangawa, K. (2005). Ghrelin: structure and function. *Physiological Reviews* 85, 495–522.
- Munzberg, H., Bjornholm, M., Bates, S. H., et al. (2005). Leptin receptor action and mechanisms of leptin resistance. *Cellular and Molecular Life Sciences* 62, 642–652.

- Shetty, G. K., Barb, D. and Mantzoros, C. S. (2006). Nutrients and peripherally secreted molecules in regulation of energy homeostasis. In: Mantzoros, C. S. (ed.) *Obesity and diabetes*, pp. 69–86. Totowa, NJ: Humana Press.
- Steppan, C. M. and Lazar, M. A. (2004). The current biology of resistin. *Journal of Internal Medicine* 255, 439–447.

Leukocyte Trafficking and Stress

S Hong

University of California, San Diego, La Jolla, CA, USA

M U Goebel

University of Duisburg-Essen, Essen, Germany

P J Mills

University of California, San Diego, La Jolla, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction

What Is Leukocyte Trafficking?

Acute Stress Paradigms

Chronic Stress Paradigms

Underlying Mechanisms and Mediators

Potential Clinical Implications

Summary

Glossary

Cellular adhesion molecules (CAMs)

Cell surface molecules that mediate leukocyte migration, homing, and cell-to-cell interaction. Types of adhesion molecules are the integrins, selectins, cadherins, and proteins of the immunoglobulin superfamily.

Cluster of differentiation (CD)

A uniform nomenclature system developed for distinguishing lineage or differentiation stage of leukocytes by identifying surface proteins that are recognized by a group of monoclonal antibodies.

Cytokines

Soluble proteins produced mostly by activated immune cells during effector phases. They regulate and mediate immune and inflammatory responses.

Enzyme-linked immunosorbent assay (ELISA)
Flow cytometry

A serological assay that detects the concentration of bound antigen or antibody through a photometer.

A cellular research tool to characterize individual cells or other microparticles by size, granularity, and surface or intracellular markers using fluorescent-labeled antibodies.

Granulocytes

A subset of leukocytes that originate in the bone marrow and include neutrophils, eosinophils, basophils, and mast cells that possess abundant cytoplasmic granules. About 90% of the peripheral blood cells are granulocytes.

Inflammation

Process that begins with cytokines produced by activated immune cells upon recognition of the antigen presence.

Leukocytosis

An increase of leukocyte (white blood cells) numbers in the peripheral circulation.

Lymphocytes

A subset of leukocytes, including natural killer (NK), T, and B cells, that are major mediators of adaptive (specific) immunity in mammals. They originate in the bone marrow.

Monocytes

A subset of leukocytes that mature to become macrophages and that are phagocytic and produce cytokines.

Introduction

The immune system is the body's defense against disease-causing agents. It is complex yet well orchestrated and nearly self-regulatory. It provides surveillance throughout the body via circulation of white

blood cells. The circulation of immune cells is an ongoing process under normal conditions (e.g., no infection), but the trafficking pattern changes under pathogenic conditions to accommodate the mounting of an immune response.

It is of great interest to researchers that there are other stimuli (that are not immunogenic) that lead to changes in trafficking of immune cells; one of these stimuli is stress. Stressors are acute immune system activators that lead not only to modulation in immune functions but also to leukocyte influx into the circulation. The effects of acute stress on leukocyte trafficking are mostly transient and fairly well understood, but the possible lasting effects of chronic stress on leukocyte mobilization are still largely unknown. While decades of research has helped delineated the effects of stress on leukocyte populations, it is only recently that the mechanisms underlying the phenomenon are being revealed, including various neurohormones and their respective receptors and cellular adhesion molecules.

This article reviews what is currently known about the effects of acute and chronic stressors on leukocyte trafficking, including discussion of mechanisms and potential clinical implications. It focuses on the evidence of changes in circulating leukocytes rather than in other tissues or lymphoid organs. The terms traf-

ficking, demargination, leukocytosis, and mobilization are used interchangeably.

What Is Leukocyte Trafficking?

Surveillance and Inflammation

Even though only 2% of total leukocytes are in the peripheral circulation at any given time, they provide a diagnostic account of the functional status of the immune system. The majority are homed in lymphoid organs and the rest are diffusely distributed in other organs. During infection or injury migration of immune cells to different sites of antigen localization within the body is critical to host defense. Recognition of antigen leads to cytokine production, causing inflammation in the area that facilitates recruitment of circulating leukocytes to the infected site. Many of the mechanisms responsible for regulating these types of fundamental responses of the immune system, including cellular adhesion molecules (CAMs), are also central to the regulation of the immune response to stressors.

Major Players: Cellular Adhesion Molecules

Initial trafficking, rolling, and firm adhesion of leukocytes to sites of inflammation are mediated by cytokines and CAMs (e.g., L-selectin, LFA-1, VLA-4, and CD44), which are homing receptors on the immune cells, and their ligands on endothelial cells (molecules that binds to a macromolecule, such as

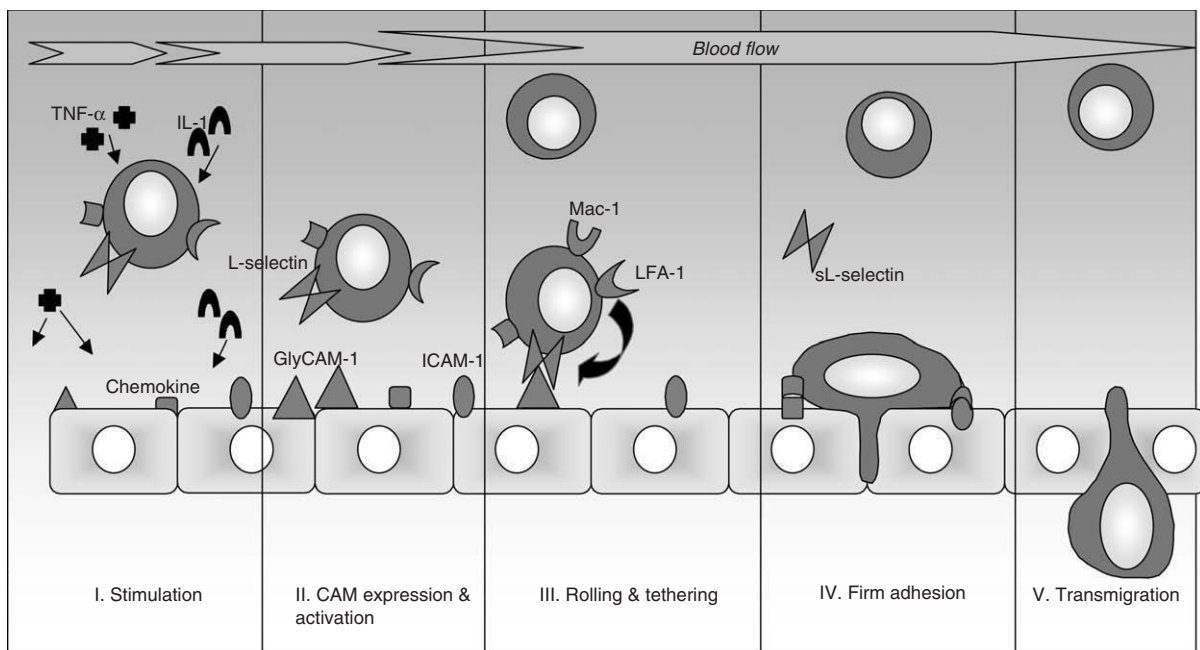


Figure 1 Schematic diagram of stages in migration and adhesion of leukocytes and cell adhesion molecules. (I, II) Stimulating factors including pro-inflammatory cytokines result in activation of leukocytes and endothelial cells and lead to increased expression and activation of adhesion molecules on both endothelial and immune cells. (III) Selectins (e.g., L-selectin) bind to ligands on the endothelium (e.g., GlyCAM-1), causing leukocytes to be captured, roll, and tether. (IV) Activation and binding of integrins (e.g., LFA-1) to ligands (e.g., ICAM-1) leads to attachment and further activation of leukocytes. (V) Leukocytes transmigrate through the endothelial wall.

a receptor, e.g., E-selectin, GlyCAM-1, ICAM-1, and VCAM-1). Inflammatory stimuli (e.g., cytokines tumor necrosis factor- α [TNF- α] and interleukin-1 [IL-1] and chemokines [chemotactic cytokines]) and antigen recognition by immune cells lead to the expression of these molecules on both immune cells and endothelial cells. Leukocytes then attach to endothelial cells lining the infected areas (firm adhesion) and penetrate into the tissue (transmigration).

Adhesion of leukocytes to the endothelium involves multiple processes. First, the initial contact of leukocytes to the endothelium is mediated by L-selectin (CD62L), tethering them loosely to carbohydrate ligands like GlyCAM-1 as they roll along the surface, as shown in [Figure 1](#). Until the leukocyte gets activated, this rolling process is reversible. Second, during activation, chemokines are released by the endothelium and trigger integrin LFA-1 (CD11a) expression on the leukocyte surface to bind tightly to the intercellular adhesion molecule ICAM-1 (CD54) on the endothelium. This event leads to an eventual transendothelial migration (extravasation). CAMs, such as L-selectin and ICAMs, are shed in a soluble form during the process and can further participate in biological processes. As mentioned previously, cytokines play a crucial role in trafficking and homing of leukocytes: inflammatory conditions, which activate macrophages and monocytes to release cytokines such as TNF- α , further activate the endothelium. This leads to upregulation of endothelial adhesion molecules, such as ICAM-1 and the vascular VCAM-1, which further attract leukocytes. Other cytokines, such as IL-2, IL-4, and IL-10, are suggested to have inhibitory effects through reduction of CAMs. The expression of CAMs is assessed by quantifying adhesion molecules on the surface of circulating leukocytes (using flow cytometry) and by determining the levels of their circulating soluble form (typically by enzyme-linked immunosorbent assay [ELISA]).

Acute Stress Paradigms

Numerous studies demonstrate that the immune system is responsive to acute psychological events. One example of immune reactivity to stress is the increased numbers of leukocytes in circulation. Evidence for this dates back to the 1930s, when Farris described the relationship among emotional stress, academic exams, and increases in the number of circulating leukocytes. A relatively parallel phenomenon in response to acute physical stress has also been documented. During the past two decades, with the more sophisticated research tools and techniques available, investigators have been actively examining the

nuances of these phenomena, as well as researching the mechanisms that regulate them.

Psychological Stress

Researchers have used a wide variety of naturalistic and standardized stressors to investigate the phenomenon of acute psychological stress-induced leukocytosis. These stress paradigms include mental and emotional stressors such as uncontrollable stimuli (e.g., shock and noise), laboratory stressors (e.g., the color-word Stroop test and mental arithmetic test), and real-life stressors (e.g., academic examinations, piloting airplanes, giving a public speech, and parachute jumping). Regardless of the stress model used, psychological stress leads to leukocytosis at varying degrees depending on the type and duration of the stressor. The literature, for example, shows marked increases in the circulation of natural killer (NK) and CD8⁺ T cell subsets, a decrease in the CD4⁺/CD8⁺ ratio, and small increases in helper T and B cells. Bosch and colleagues recently reported that a subset of NK cells is more responsive to a speech stressor such that the number of NK cells expressing low levels of CD56 (cytotoxic) tripled but the number of NK cells expressing high levels of CD56 (regulatory) did not change following a stressor. Together, such findings have raised several questions, including why selected subsets of leukocytes are more responsive to stress, and from which sites the leukocytes are being released? Regarding the first question, studies suggest that it is in part attributable to a different degree or profile of CAM expression and receptors for neurotransmitters and stress hormones on different cell types. Regarding the second question, immune cells are released from the so-called marginal pools as well as from the spleen, which is a primary reservoir of circulating leukocytes, especially following sympathetic nervous system activation.

Psychological stress also leads to changes in CAM expression on peripheral (demarginated) leukocytes. It is consistently shown that the phenomenon of leukocyte redistribution in response to acute stress can be differentiated according to L-selectin (CD62L) expression. As seen in [Figure 1](#), L-selectin serves a crucial role in a leukocyte's initial rolling on the endothelium. Findings suggest that it is primarily lymphocytes not expressing L-selectin (CD62L⁻) that markedly increase in the circulation following acute stress (either psychological or physical). CD8⁺CD62L⁻ T cells, for example, show a two- to threefold increase in the circulation following acute stress as compared to CD8⁺CD62L⁺ T cells, which show either no change or a more moderate increase. A similar CD62L^{-/+} response differentiation is shown

for both CD4⁺ and NK lymphocytes. This phenomenon most likely results from a combination of events, including a preferential release of CD62L⁻ lymphocytes from the spleen and marginal pools, an increased rate of adhesion of CD62L⁺ lymphocytes, and a downregulation or shedding of CD62L from the leukocyte's surface upon further activation.

Physical Stress

In addition to the phenomenon of psychological stress-induced leukocytosis, researchers report dramatic increases of immune cells in the circulation after a bout of exercise. It should be made clear here that a bout of exercise should be distinguished from longitudinal exercise training (see Physical Stress Models and Adaptation below). And, in this article, acute exercise is discussed as a physiological stressor, given the cardiopulmonary, metabolic, and hemodynamic demands of exercise on the individual. In addition, a large body of literature on the effects of exercise on the neuroendocrine system supports the notion that exercise is a biologically plausible paradigm to investigate stress-induced modulation of the immune system.

Changes in circulating cell numbers in response to exercise vary among subpopulations of leukocytes and are dependent upon the intensity and duration of the exercise challenge. There is a dose response in exercise-induced leukocytosis such that mild exercise results in smaller changes in cell numbers. Marked increases in numbers of most lymphocyte subsets, monocytes, and granulocytes are consistently observed. Similar to psychological stress, small or no changes are typically seen in numbers of B lymphocytes, even after exercise of a moderate to high intensity. Changes in NK cell and neutrophil numbers after exercise are most dramatic; the number of NK cells can increase two- to fourfold above resting levels. Interestingly, a small subset of monocytes that also expresses CD16 (CD14⁺CD16⁺) (CD16 is largely expressed on NK cells) is shown to be mobilized to the circulation during exercise at a larger degree than regular monocytes (CD14⁺CD16⁻). Not only cardiovascular exercise but also resistance exercise using relatively small muscle groups can significantly affect circulating immune cells.

In spite of the consistent findings on associations between exercise intensity and leukocytosis, exhaustive and prolonged exercise (e.g., a marathon or triathlon) results in decreased immune cell numbers (leukocytopenia; especially T and NK cells) in the circulation. Although the reason for this trafficking pattern after prolonged exercise is unclear, there are several possible explanations. It may well be that the initial leukocytosis for the first 20–30 min of a marathon may be followed by a gradual decline,

resulting in leukocytopenia as the high-intensity exercise continues. As is discussed later, sustained activation of the hypothalamic-pituitary-adrenocortical (HPA) axis, resulting in high cortisol levels, may be a mediating factor for decreased immune cell numbers after a marathon. In addition, muscular tissue injury during a marathon or triathlon may also be a contributing factor, although whether the leukocytopenia seen after a marathon reflects the redistribution of the cells to the fatigued or damaged muscle tissue is not known.

During exercise, leukocytes may demarginate to the peripheral blood compartment in part by increased hemodynamic forces that detach cells that normally loosely adhere to the vessel walls at rest. Cells that are mobilized into the circulation during exercise appear to be from the secondary lymphoid organs (the spleen and lymph nodes) rather than the primary lymphoid organs (the thymus and bone marrow), given findings that a greater percentage of activated CD62L⁻ memory T cells demarginate as compared to naive cells. Thus, exercise-induced leukocytosis appears to be a redistribution of cells from the spleen and lymph nodes rather than proliferation or differentiation of them from primary sources.

As with the psychological stress responses, CD62L⁻ lymphocytes appear highly responsive to an acute exercise challenge. Granulocytes and lymphocytes show a significant decrease in the surface density of CD62L following exercise. A similar shedding process may also occur with the adhesion molecule ICAM-1. Such a shedding of ICAM-1 might lead to a further demargination of CD62L⁻ leukocytes back into the circulation. Studies demonstrate an increase in levels of soluble ICAM-1 following exercise, supporting such a conjecture.

Pharmacological Stress

In order to test the hypothesis that leukocytosis following psychological or physical stressors may be mediated by sympathetic nervous system activation, investigators have examined leukocyte trafficking by administering sympathetic (adrenergic) agonists. Pharmacological infusion studies of such agonists provide direct mechanistic evidence for the stress-leukocytosis phenomenon. Infusion of the nonspecific β -adrenergic agonist isoproterenol (a pharmacological agent that leads to similar physiological effects as norepinephrine), for example, leads to increases in CD8⁺ T and NK cell numbers in blood. Isoproterenol also leads to increased numbers of CD62L⁻ but not CD62L⁺ CD8⁺ cells as well as a decrease of CD62L density and increased CD11a density. Activated/memory T cells appear more responsive to the infu-

sion as compared to naive T lymphocytes. Studies using β -adrenergic antagonists such as propranolol to block sympathetic agonist effects in combination with stressors provide further evidence of adrenergic regulation of leukocytosis. Ingestion of propranolol for 5 days leads to a blunted stress-induced lymphocytosis to psychological and exercise stressors. For example, the $CD8^+CD62L^-$ T cell response to exercise is dramatically attenuated by treatment with propranolol. Together, these findings indicate that acute stress-induced leukocyte trafficking is mediated in part by β -adrenergic receptor activation.

Chronic Stress Paradigms

Psychological Stress

It is widely believed that more enduring stressors lead to downregulation of immune function and increased disease susceptibility. Studies report immune suppression under such chronic stressors as marital strife, bereavement, long-term caregiving, living in an unfavorable environment such as near a nuclear power plant, and environmental disasters such as floods and hurricanes. Immunosuppressive effects of chronic stress include a decrease in circulating leukocyte NK cell number and activity at rest, a lower cytokine release rate, and enhanced antibody titer for latent viruses such as herpes simplex virus type 1 and Epstein-Barr virus. Kiecolt-Glaser et al. reported that medical students who experience high stress levels and poorer social support during academic examination exhibit very low NK cell activity.

As with acute stress studies, CAMs appear to have an important role in mediating the effects of chronic stress on leukocytes. For example, one study was conducted on CD62L expression on T lymphocytes among elderly (average age of 75 years) spousal caregivers of Alzheimer patients experiencing high stress levels due to increased caregiving burden, who were compared to a group of caregivers with a relatively low stress burden from caregiving. The highly stressed caregivers showed significantly lower numbers of circulating $CD8^+CD62L^-$ and $CD4^+CD62L^-$ T cells at rest and during acute psychological stress but no differences in $CD62L^+$ T cells as compared to the control group. Lower numbers of $CD62L^-$ T cells in the circulation indicate that there is decreased availability of activated/memory T cells. In addition, this deficit in activated/memory T lymphocytes, cells that are critical in mounting immune responses, may imply a stress-associated accelerated decline of immunity in these older individuals, which, in turn, might have adverse consequences for health.

Physical Stress Models and Adaptation

It is of great interest to many researchers to determine if chronic exercise training or regular physical activity leads to a favorable modulation of the immune system. Contrary to findings of immunosuppression and potential increased disease susceptibility following chronic psychological stress, long-term regular exercise typically leads to improvements in health. It is unclear, though, whether regular exercise leads to enhanced immunity or changes in immune cell distribution that would support observations of improved health. In general, findings on the effects of chronic exercise (training) or physical fitness on resting peripheral blood leukocyte numbers and adhesion molecule expression are inconclusive, with no consistent effects shown for $CD4^+$, $CD8^+$, or B lymphocyte numbers among studies.

However, physical fitness might affect leukocyte trafficking in response to acute stressors. It has been shown that physically fit individuals exhibit attenuated lymphocytosis in response to a speech stressor and to a moderate exercise challenge (Figure 2). This blunted response to a stressor is not derived from the differences in cell numbers at rest and is only observed among T lymphocytes, not in monocytes, granulocytes, or NK cells. In addition, a significant attenuation of demargination is found for $CD4^+$ memory T cells that do not express the activation marker HLA-DR after exercise in fit individuals. Such findings imply that the effects of fitness on leukocytosis is cell type and activation stage specific.

Underlying Mechanisms and Mediators

How do stress-induced changes in leukocyte trafficking occur? How is selective and cell type-dependent mobilization explained? What factors influence CAM expression during leukocyte trafficking? There are two main underlying physiological pathways that address these questions.

Sympathetic Nervous System Activation

There is widespread direct communication between the sympathetic nervous system (SNS) and the immune system. Adrenergic projections and sympathetic nerve terminals are found in many organs of the immune system, including the spleen. Sympathetic agonists (catecholamines) are rapidly released from nerve endings following acute SNS activation and act directly on α - and β -adrenergic receptors. β_2 -adrenergic receptors are found on all white blood cells, including lymphocytes, neutrophils, and monocytes. Activation of these receptors leads to lymphocytosis. Adherence

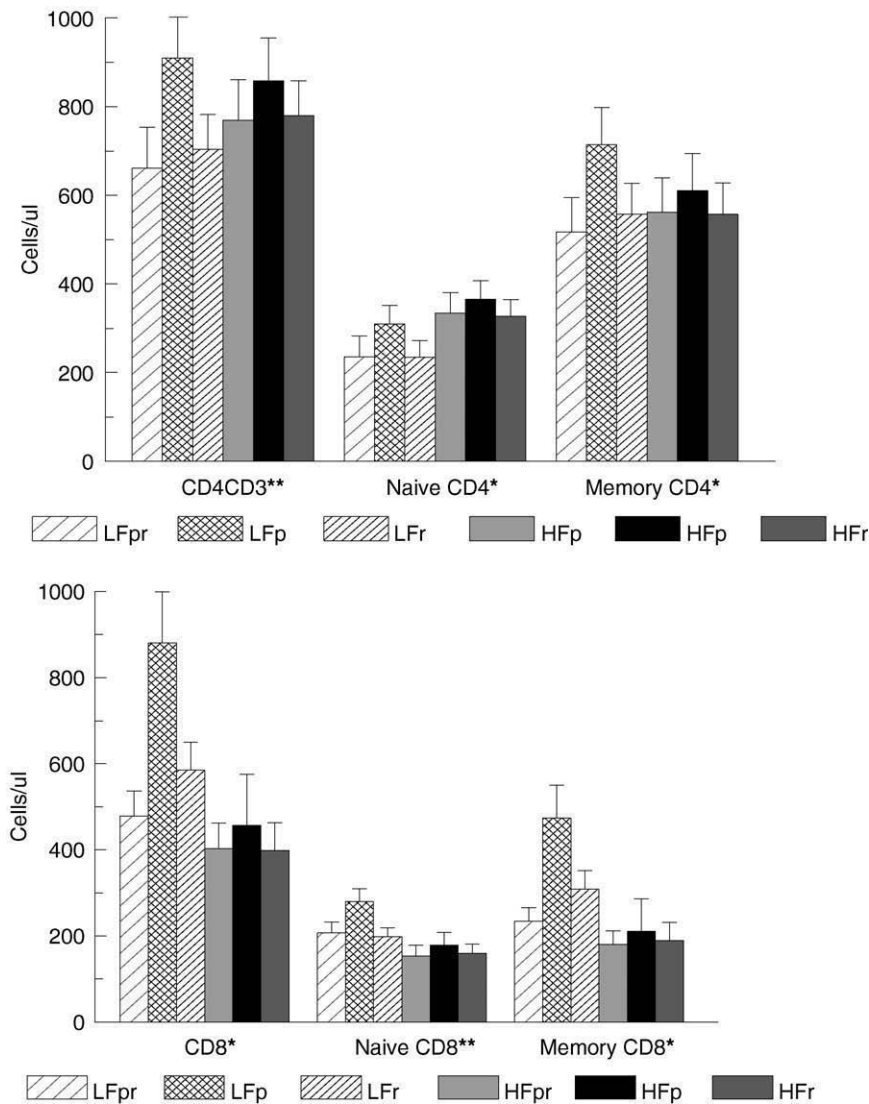


Figure 2 Naive and memory $CD4^+$ and $CD8^+$ T lymphocyte numbers pre-, post-, and 10-min postexercise in high- vs. low-fitness groups. Physically fit subjects showed reduced $CD4^+$ and memory $CD4^+$ T cell responses (indicated by *) but not in naive $CD4^+$ T cell responses. Fit subjects showed smaller $CD8^+$ naive and memory cell responses. LF, low fitness; HF, high fitness; pr, pre-exercise; p, postexercise; r, recovery; * and ** denote significant ($p < 0.05$ and $p < 0.01$, respectively) time by group interactions (mean $\pm$ standard error).

of NK cells to endothelial cells is decreased in a dose-dependent manner after adding β_2 -adrenergic agonists *in vitro*. As described previously, infusing the β -adrenergic agonist isoproterenol leads to demargination and trafficking of leukocyte subsets similar to those seen following a stressor. Catecholamines lead to increased recirculation of lymphocytes into the blood by decreasing the affinity of the lymphocyte to the vessel wall, and possibly through a shedding of adhesion molecules such as CD62L and ICAM-1 from the cell surface.

The HPA Axis

Neuroendocrine responses to stressors include increased levels of glucocorticoids in circulation as a result of HPA activation. Glucocorticoids are known for their immunosuppressive effects. Their pharmacological administration has wide-ranging effects, including the modulation of metabolic processes and immunosuppression. Glucocorticoids are widely used in the treatment of many inflammatory conditions, influencing the migration of immune cells to infected tissues. Studies demonstrate that glucocorticoids reduce the number of circulating leukocytes and down-regulate cytokine production and chemotaxis, as well as downregulate adhesion molecules such as E-selectin

and ICAM-1. *In vivo* studies demonstrate time-related effects of dexamethasone on ICAM-I expression, in which prolonged treatment (but not single acute doses) significantly reduces ICAM-1 expression on monocytes.

Mediators

As with many other physiological processes, inter-individual factors such as age, diet, gender, ethnicity, and psychosocial characteristics can affect leukocyte trafficking in response to stressors. Depression, lack of social support, or hostile personality, for example, can all lead to altered immune cell responses to acute stress. Although studies examining the effects of gender as a possible mediator of leukocyte trafficking during stressors have been inconsistent, studies examining the effects of reproductive hormones do indicate an effect. A low- versus a high-carbohydrate meal leads to attenuated leukocyte and neutrophil trafficking in response to exercise. Although aging appears to be associated with a general decline in immune function, including a loss of naive T lymphocytes, few studies have systematically investigated whether age influences stress-induced leukocyte trafficking. As discussed earlier, physical fitness is a very important mediator of leukocyte responses to both exercise and psychological stressors.

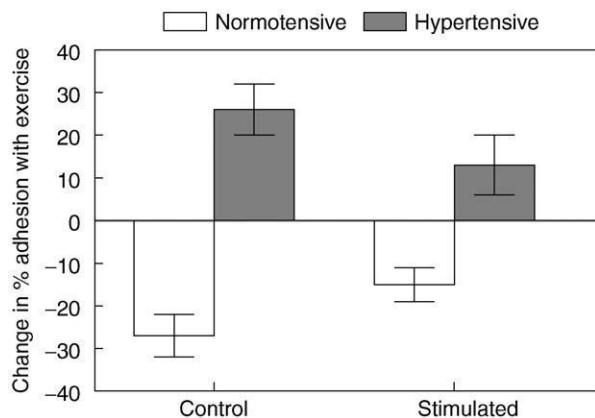


Figure 3 The effects of hypertension on the leukocyte-endothelial adhesion response to exercise. Percentage change of peripheral blood mononuclear cell (PBMC) adhesion to human umbilical venous endothelial cells (HUVEC) under control and cytokine-stimulated conditions at rest and in response to acute exercise in hypertensive patients and normotensive subjects. In response to exercise, PBMC adhesion to HUVECs was increased in hypertensives but decreased in normotensives under control and stimulated conditions ($p < 0.03$). Reprinted from Mills PJ, Maisel AS, Ziegler MG, et al. 2000. Peripheral blood mononuclear cell-endothelial adhesion in human hypertension following exercise, *Journal of Hypertension*. 18(12), 1804 with permission from the publisher.

Potential Clinical Implications

Are there potential clinical implications associated with exaggerated or perhaps diminished leukocyte trafficking responses to stressors? A few studies have endeavored to address this issue by studying clinical populations. For example, hypertension, or high blood pressure, is a well-established risk factor for atherosclerosis, which is an inflammatory disease. Patients with hypertension have elevated endothelial expression of the adhesion molecules ICAM-1, VCAM-1, E-selectin, and P-selectin. This enhanced CAM activation attracts leukocytes and monocytes to adhere to the endothelial surface, supporting inflammation. As previously discussed, the SNS plays a prominent role in leukocyte trafficking and CAM expression. Elevated sympathetic nerve activity and circulating catecholamines levels are common in hypertension. It is reasonable, therefore, to hypothesize that the sympathetic mediation of leukocyte trafficking and cell adhesion may be accentuated in hypertension. Incidentally, in addition to direct SNS activation effects, greater hemodynamic forces found in hypertension also contribute to the phenomenon of endothelial activation and increased expression of endothelial CAMs.

Findings show a greater expression of LFA-1 (CD11a) on circulating lymphocytes at rest and in response to stressors in hypertensive patients. In addition, the adhesion of lymphocytes to endothelial cells in hypertensives is enhanced in response to the stressor of exercise but reduced in people with normal blood pressure (Figure 3). Such differences in the adhesion characteristics of lymphocytes in hypertensives following stress might have clinical relevance, as adherent mononuclear cells mediate endothelial injury associated with atherosclerosis as well as with ischemic disorders. However, whether such stressor-related effects on leukocytes are related to the etiology and/or progression of vessel disease in hypertension has yet to be demonstrated. More research is needed on the possible clinical relevance of stress-induced changes of immune cell adhesion before the significance of such observations can be fully understood. Stress-induced leukocyte trafficking also merits further investigation of its clinical implications in other clinical populations, including patients with cancer, multiple sclerosis, and rheumatoid arthritis.

Summary

Circulation of immune cells throughout the body is a critical immunosurveillance process for host defense. Many studies demonstrate mobilization of immune

cells in response to acute and chronic psychological and physical stressors. Stress-induced leukocytosis is cell subset specific and there is a dose response based on the intensity of the stressor. CAMs, including CD62L, CD11a, and ICAM-1, play a crucial role in the phenomenon. Both the SNS and the HPA axis help regulate immune cell responses to a variety of stressors. Numerous important questions still need to be answered, including the identity of interindividual factors that significantly mediate immune cell responses to stressors, as well as the potential clinical implications of this interesting effect of stressors on the trafficking of immune cells in the blood.

See Also the Following Articles

Desensitization; Infection; Lymphocytes; Natural Killer (NK) Cells.

Further Reading

Bosch, J. A., Berntson, G. G., Cacioppo, J. T. and Marucha, P. T. (2005). Differential mobilization of

functionally distinct natural killer subsets during acute psychological stress. *Psychosomatic Medicine* 67(3), 366–375.

Goebel, M. U. and Mills, P. J. (2000). Acute psychological stress and exercise and changes in peripheral leukocyte adhesion molecule expression and density. *Psychosomatic Medicine* 62(5), 664–670.

Hong, S., Johnson, T. A., Farag, N., Guy, H., Matthews, S. and Mills, P. J. (2005). Attenuation of T lymphocyte demargination and adhesion molecule expression in response to moderate exercise in physically fit individuals. *Journal of Applied Physiology* 98(3), 1057–1063.

Kiecolt-Glaser, J. K., McGuire, L., Robles, T. F. and Glaser, R. (2002). Psychoneuroimmunology and psychosomatic medicine: back to the future. *Psychosomatic Medicine* 64, 15–28.

Mills, P. J., Farag, N. H., Hong, S., Kennedy, B. P., Berry, C. C. and Ziegler, M. G. (2003). Immune cell CD62L and CD11a expression in response to a psychological stressor in human hypertension. *Brain Behavior & Immunity* 17(4), 260–267.

Sackstein, R. (2005). The lymphocyte homing receptors: gatekeepers of the multistep paradigm. *Current Opinion in Hematology* 12(6), 444–450.

Leukocytes See: Leukocyte Trafficking and Stress.

Life Events and Health

P Surtees and N Wainwright

Strangeways Research Laboratory and University of Cambridge, Cambridge, UK

© 2007 Elsevier Inc. All rights reserved.

Measurement
Health Associations
Individual Differences

Glossary

Annual prevalence

Life event

Major depressive disorder

The proportion of people identified with a disease or other characteristic at any time during the period of a year.

A discrete incident or circumstance that may disrupt life, may require adaptation, and may be associated with, or result in, physical and/or emotional health change. A health state characterized by sustained feelings of sadness, hopelessness, helplessness, and worthlessness that is often also accompanied by suicidal thoughts, an inability to concentrate or to take pleasure in activities that were once enjoyable, and sleep, appetite, and memory problems.

Kiecolt-Glaser, J. K., McGuire, L., Robles, T. F. and Glaser, R. (2002). Psychoneuroimmunology and psychosomatic medicine: back to the future. *Psychosomatic Medicine* 64, 15–28.

Mills, P. J., Farag, N. H., Hong, S., Kennedy, B. P., Berry, C. C. and Ziegler, M. G. (2003). Immune cell CD62L and

CD11a expression in response to a psychological stressor in human hypertension. *Brain Behavior & Immunity* 17(4), 260–267.

Sackstein, R. (2005). The lymphocyte homing receptors: gatekeepers of the multistep paradigm. *Current Opinion in Hematology* 12(6), 444–450.

Leukocytes See: Leukocyte Trafficking and Stress.

Life Events and Health

P Surtees and N Wainwright

Strangeways Research Laboratory and University of Cambridge, Cambridge, UK

© 2007 Elsevier Inc. All rights reserved.

Measurement

Health Associations

Individual Differences

Glossary

<i>Annual prevalence</i>	The proportion of people identified with a disease or other characteristic at any time during the period of a year.
<i>Life event</i>	A discrete incident or circumstance that may disrupt life, may require adaptation, and may be associated with, or result in, physical and/or emotional health change.
<i>Major depressive disorder</i>	A health state characterized by sustained feelings of sadness, hopelessness, helplessness, and worthlessness that is often also accompanied by suicidal thoughts, an inability to concentrate or to take pleasure in activities that were once enjoyable, and sleep, appetite, and memory problems.
<i>Prospective cohort study</i>	A study population defined according to risk factor exposure and prospectively followed up, for example, to ascertain incident disease endpoints.
<i>Record linkage study</i>	A study that brings together information included in two or more records based upon a unique identifying system.

On January 17, 1995, at 5:46:52 a.m., the southern part of Hyogo Prefecture in Japan was struck, without warning, by the Hanshin-Awaji earthquake, which measured 7.2 on the Richter scale and lasted for around 20 s. During the 3 months following this unexpected catastrophic natural disaster, in the districts of Awaji Island near the earthquake epicenter, heart attack rates increased by 50% and stroke rates by 90% in comparison with the same observation period during the previous year. Survivors in areas with the greatest loss and destruction of housing experienced the highest subsequent heart attack rates. While the occurrence of a profoundly stressful life event, such as the Hanshin-Awaji earthquake, provides a clear and dramatic incident from which to assess subsequent health, the attribution of health change to life events experienced more commonly across a lifetime is less straightforward.

Measurement

Life event assessment techniques have evolved to include the use of checklists and interview methods and have been used in a wide variety of research designs to aid understanding of the risk of life events to health.

Checklists

During the late 1940s, Thomas Holmes and colleagues at the University of Washington started to investigate the relationship between adjustment to social stress and illness onset using a self-administered

instrument, the Schedule of Recent Experience (SRE). The SRE included a list of 43 life change events intended to represent most circumstances experienced by the patient groups studied, and to which differing degrees of readjustment were required. A subsequent scaling study was designed to take account of both the intensity and the relative time judged to be needed to adjust to each life event on the list. This Social Readjustment Rating Scale (SRRS) included the life events ranked according to these degrees of readjustment – subsequently termed life change units (LCUs). The event of highest rank was death of spouse (with a mean value score of 100), relative to marriage, with a mean value score of 50, and, for example, death of close friend, which scored 37. Typically the LCUs for all events reported on the checklist to have been experienced during a given time period (e.g., the past year) were summed to give a measure of the cumulative social stress experienced. With varying degrees of modification, the SRRS has served as a model of a questionnaire format for assessing life event experience in relation to health. For example, in psychiatry, scaling studies focused on the capacity of life events to induce distress and were scaled in terms of their associated upset or undesirability. While the checklist approach provided a powerful impetus to study the association between life events and illness, these techniques attracted criticism. This stemmed from suggestions that the circumstances to be rated were defined vaguely, resulting in variability in item interpretation, that symptoms were included in the event lists, and that the assumption of additivity of event effects was untested. Despite these criticisms, the list techniques can provide a reliable assessment of the occurrence of specified adverse experiences, although they may be limited in their capacity to provide details of the context surrounding their occurrence.

Interviews

The Life Events and Difficulties Schedule (LEDS), an interview method for the assessment of stressful life events, was developed in the late 1960s and over the subsequent 30 years by George Brown and colleagues at London University. In contrast to the early respondent-based checklist assessment techniques, this approach was based on using the interviewer as the measuring instrument. Through the LEDS, ratings of adverse experience followed from consideration of both the context and meaning of the immediate circumstances associated with a stressful situation. Use of the LEDS was designed to provide measures of the long-term threat of life events experienced, to attempt to date their occurrence accurately, and

to rate ongoing long-term difficulties. In addition, LEDS-based research drew attention to the importance of distinguishing the extent to which life events and long-term difficulties could be illness-related in order to help clarify their etiological significance for illness onset.

However, Bruce Dohrenwend and colleagues, at Columbia University, New York, raised fundamental criticisms of the LEDS core ratings of contextual threat. These criticisms were based upon their view that the contextual ratings represented a nonexplicit distillation of the event, the social circumstances surrounding its occurrence, and the personal history of the person who experienced the event, into one measure. Dohrenwend and colleagues argued that these situational and personal variables may themselves be important risk factors for illness, therefore creating ambiguity in understanding precisely which aspects of the stress process were associated with the illness outcomes of interest. They attempted to address these perceived problems with the LEDS approach through developing their own theoretical framework and measurement approach, using the Structured Event Probe and Narrative Rating method (SEPRATE) of life events measurement. The SEPRATE semistructured interview method was designed to provide measures that distinguished between the life event, the ongoing social situation present before event occurrence (for example, employment circumstances), and the personal dispositions and characteristics of the person who had experienced the life event. While the LEDS and SEPRATE share similarities in methodological approach, they differ markedly in their representation and rating of life event experience.

Health Associations

Efforts to synthesize research findings relating life events to health outcomes have been limited by the rarity of studies that meet stringent research design requirements. A commonly used set of criteria to help establish causality was proposed by the epidemiologist and statistician Sir Austin Bradford Hill. These criteria include consideration of the strength of the association, whether it is consistent (i.e., repeatedly demonstrated in different populations), whether there is evidence of temporality (i.e., the presence of the risk factor precedes onset of disease), and whether a dose-response relationship exists (i.e., disease risk increases monotonically with increasing exposure). These criteria have not yet been fully met for any health outcome following from stressful life event experience. However, increasingly persuasive evidence

is now accumulating in relation to the onset, and subsequent prognosis, of some health outcomes.

Life Events and Psychological Health

Evidence from both prospective cohort and record linkage studies has consistently shown an association between the experience of stressful life events and the subsequent onset of episodes of major depressive disorder. The strength of this association has been shown to vary according to the way in which life events were assessed: generally stronger associations were reported when events were assessed using the LEDS approach, rather than through use of the list techniques such as the SRE. These studies have also suggested evidence of a dose–response relationship, with the risk of depressive disorder onset increasing broadly in step with the severity of the events experienced. A dose–response association between the number of recently reported stressful life events and the annual prevalence of depressive disorder is illustrated in [Figure 1](#).

Record linkage studies, through their use of very large populations, provide a potentially powerful research design with which to investigate the relationship between severely stressful events and psychiatric health. For example, one study linked records from the Danish Civil Registration System with the Danish Psychiatric Central Register to investigate the association between death of a child and subsequent

parental admission to psychiatric hospital. Based on over a million persons identified in the registers and almost 12 000 deaths of children aged under 18 years, the study revealed a 67% increased risk of hospital admission for any psychiatric condition among the parents who had lost a child, with the risk being greatest for mothers and most pronounced for admission with a depressive condition.

The relationship between chronic enduring or progressively deteriorating stressful circumstances and psychiatric health is complex and difficult to study. For example, changes in the Japanese economy during the late 1980s, subsequently associated with escalating bankruptcy rates, has been interpreted as a significant contributory cause of increases in the male suicide rate, stemming from the high proportion of suicide notes in which economic circumstances were the stated reason for suicide. The Japanese male suicide rate increased from around 25 per 100 000 during the mid 1990s to over 35 per 100 000 at the beginning of the twenty-first century, when over 30 000 Japanese per year committed suicide, by then around twice the rate observed in the United States.

Life Events and Physical Health

Clinicians have long suspected stressful life events to be associated with the onset of physical disease (for example, this was noted by Sir James Paget, the surgeon and pathologist, in his *Clinical Lectures and Essays* in 1875). The assessment of stressful life event experience within chronic disease epidemiological study settings presents substantial methodological challenges related to cost, the need to limit participant burden, and the requirement of asking perhaps tens of thousands of study participants, ideally when they are all healthy, to recall their experience of stressful events over a particular time period. In consequence, within large-scale prospective cohort studies, the assessment of stressful life events has been largely restricted either to the use of short event lists or to specific events (e.g., bereavement). Record linkage research designs have investigated the risk of chronic disease and mortality associated with specific stressful events; for example, following the death or serious illness of a child or loss of a partner or twin. Both prospective cohort and record linkage studies have now provided evidence that under some circumstances such events may increase risk for congenital malformations, heart disease, and perhaps some cancers, just a few among many physical health outcomes studied.

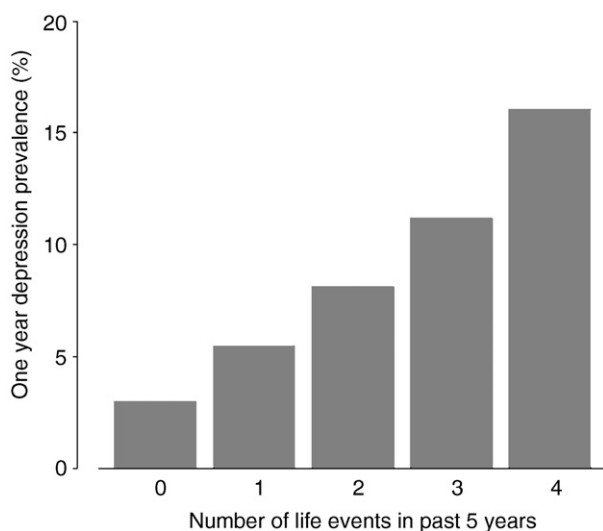


Figure 1 Annual prevalence of major depressive disorder according to the number of moderately or extremely upsetting (non-health-related) life events reported during the previous 5 years. Based upon 20 921 participants in the European Prospective Investigation into Cancer (EPIC)-Norfolk study in the United Kingdom.

Heart disease Sudden and particularly catastrophic life events, such as the 1995 Hanshin-Awaji

earthquake in Japan, the 1999 Ji-Ji earthquake in Taiwan, and the terrorist attacks on the World Trade Center in New York on September 11, 2001, have been associated with transient increases in blood pressure, increases in ventricular arrhythmias among patients with cardioverter defibrillators, and increases in the incidence of fatal and nonfatal cardiovascular events. Further, use of national registers in Denmark has shown that experiencing the unexpected death of a child (under 18 years) is associated with an increased subsequent risk of a heart attack (myocardial infarction) in the bereaved parents. The association between chronic stressful circumstances and cardiovascular events is less certain, but there is an increasing international consensus that a combination of high work demands and low job control is associated with cardiovascular events in certain occupational groups and that caregiving responsibilities for a disabled or ill partner for 9 or more h per week is associated with an increased risk of incident heart disease events in women.

Cancer Studies meeting stringent research design criteria that have been devoted to evaluating life events as possible determinants of cancer incidence and progression are rare. Evidence for the association between life events and cancer incidence is often contradictory, based on studies that often rely on imprecise case ascertainment methods, may be prone to recall bias in their assessment of life event histories, are able to include relatively few participants that develop cancer, and may be unable to include adjustment for factors known to be associated with cancer outcomes. Some of these limitations were overcome in a prospective study of a twin cohort of 10 808 women in Finland. During 15 years of follow-up, 180 incident breast cancers were diagnosed, and the results suggested that life events experienced during the 5 years before entry into the study were associated with an increased risk of breast cancer during follow-up. Three specific life events were found to be associated with an increase in breast cancer risk: divorce or separation, death of a husband, and death of a close relative or friend. The association between widowhood and divorce and subsequent cancer risk compared to married people has been investigated in a large Swedish study. A relatively consistent pattern of cancer risk was found in the divorced compared to the widowed (both with increases and decreases in risk, according to cancer site) that collectively may reflect changes in lifestyle following partner loss. Similar conclusions have been suggested to account for the evidence from record linkage studies that have found a slightly elevated overall cancer risk in mothers following death of a child.

Individual Differences

There are profound individual differences in the health consequences of stressful life events and difficult circumstances. These differences are most evident through studies of the health outcomes of people subsequent to their exposure to very stressful circumstances, including shipwreck, concentration camp experience, and terrorist attacks. Some people adapt to stressful circumstances more successfully than others. Not all people exposed to adverse or even traumatic circumstance experience health change. Identification of what distinguishes people in the way they are able to manage stressful life events can aid understanding of disease susceptibility and inform the design and delivery of effective care to reduce the long-term health impact of events. For example, a brief questionnaire designed to distinguish the extent to which people are optimistic about their lives and future expectations has shown that optimists have a reduced risk from dying due to cardiovascular causes over a 15-year period and after consideration of traditional risk factors. More direct evidence of individual differences in the capacity to adapt to stressful life events has followed from work originally inspired by Aaron Antonovsky's study of survivors of the Holocaust. This work has revealed how, despite their horrific experiences, some Holocaust survivors had subsequently been able to rebuild their lives successfully. Antonovsky defined sense of coherence (SOC), a theoretical construct based upon these observations, as a flexible and adaptive dispositional orientation enabling successful coping with adverse

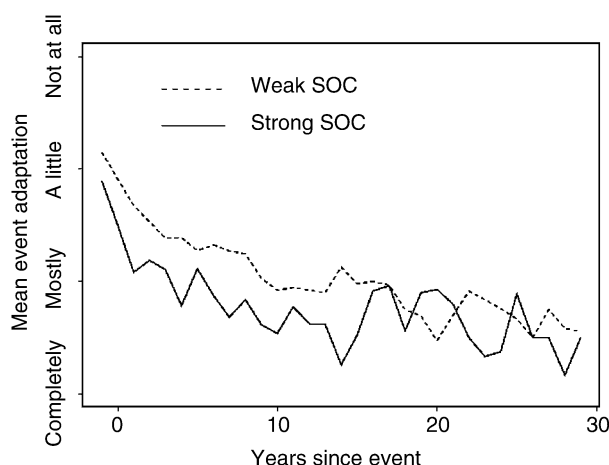


Figure 2 Mean reported adaptation to death of spouse or partner according to the elapsed time since the event occurred and according to sense of coherence (SOC). Based upon 20 921 participants in the European Prospective Investigation into Cancer (EPIC)-Norfolk study in the United Kingdom. Mean adaptation score was calculated from responses to the question "Do you feel that you have got over this now?" (1 = completely, 2 = mostly, 3 = a little, 4 = not at all).

experience. **Figure 2** shows that individuals with a weak SOC report that they take longer to get over the adverse effects of the death of their spouse or partner than those with a strong SOC and illustrates work that has provided support for the idea that SOC distinguishes adaptive capacity to adverse life event experience.

See Also the Following Articles

Cancer; Concentration Camp Survivors; Earthquakes, Stress Effects of; Heart Disease/Attack; Life Events Scale; Major Depressive Disorder; Terrorism.

Further Reading

- Antonovsky, A. (1987). *Unraveling the mystery of health*. San Francisco, CA: Jossey-Bass.
- Cooper, C. L. (ed.) (2005). *Handbook of stress medicine and health* (2nd edn.) Boca Raton, FL: CRC Press.
- Dohrenwend, B. P. (1998). *Adversity, stress and psychopathology*. New York: Oxford University Press.
- Giltay, E. J., Kamphuis, M. J., Kalmijn, S., Zitman, F. G. and Kromhout, D. (2006). Dispositional optimism and the risk of cardiovascular death. The Zutphen Elderly Study. *Archives of Internal Medicine* **166**, 431–436.

- Hansen, D., Lou, H. C. and Olsen, J. (2000). Serious life events and congenital malformations: a national study with complete follow-up. *Lancet* **356**, 875–880.
- Harris, T. (2000). *Where inner and outer worlds meet. Psychosocial research in the tradition of George W. Brown*. London: Routledge.
- Holmes, T. H. and Rahe, R. H. (1967). The social readjustment rating scale. *Journal of Psychosomatic Research* **11**, 213–218.
- Kessler, R. C. (1997). The effects of stressful life events on depression. *Annual Review of Psychology* **48**, 191–214.
- Li, J., Laursen, T. M., Precht, D. H., Olsen, J. and Mortensen, P. B. (2005). Hospitalization for mental illness among parents after the death of a child. *New England Journal of Medicine* **352**, 1190–1196.
- Lillberg, K., Verkasalo, P. K., Kaprio, J., Teppo, L., Helenius, H. and Koskenvuo, M. (2003). Stressful life events and risk of breast cancer in 10,808 women: a cohort study. *American Journal of Epidemiology* **157**, 415–423.
- Stansfeld, S. and Marmot, M. (2002). *Stress and the heart: psychosocial pathways to coronary heart disease*. London: BMJ Books.
- Surtees, P. G., Wainwright, N. W. J. and Khaw, K.-T. (2006). Resilience, misfortune and mortality: evidence that sense of coherence is a marker of social stress adaptive capacity. *Journal of Psychosomatic Research* **61**, 221–227.

Life Events Scale

E Wethington

Cornell University, Ithaca, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by E Wethington, volume 2, pp 618–622, Elsevier Inc.

General Assessment Issues

Major Life Events

Stress Appraisal

Chronic Stressors

Hassles or Daily Events

Glossary

<i>Chronic stressor</i>	An ongoing set of consistent, persistent physically or emotionally arousing demands on an organism; also referred to as a long-term difficulty.
<i>Daily event</i>	See hassle.

<i>Distal stressors</i>	Major events and stressors that occur earlier in life, usually in childhood or adolescence.
<i>Event rating, contextual</i>	An estimate of the severity of a stressor based on the objective demands posed for a particular individual who undergoes a more or less demanding stressor of that type.
<i>Event rating, normative</i>	A population-based estimate of the relative severity of a particular class of stressors.
<i>Hassle</i>	Minor life event whose effects disperse within a day or two; also referred to as a daily event.
<i>Life event</i>	A discrete occurrence believed to arouse a need for adjustment or to create negative emotions and physical reactions.
<i>Proximal stressors</i>	Recent events or stressors (e.g., within the past year).
<i>Risk factor</i>	A condition, such as stressor exposure, predisposing an individual to develop illness.

<i>Stress appraisal</i>	The individual evaluation of a stressor for its impact on well-being.
<i>Stressor exposure</i>	Exposure to life events, chronic stressors, or hassles in the environment.
<i>Threat, contextual</i>	The degree or type of demand posed by a life event or stressor.

A life events scale is a comprehensive list of external events and situations (stressors) that are hypothesized to place demands that tend to exceed the capacity of the average person to adapt. The difficulty in adaptation leads to physical and psychological changes or dysfunction, creating risk for psychological disorder or physical disease.

General Assessment Issues

It is now generally recognized that exposure to stressors in daily life, or over the course of life, is one of the most critical components of health and well-being. However, there remains ambiguity about the dimensions of stressor exposure that are involved in this process, specifically the types of stressors that have the more deleterious effects on health and that are risk factors for the development of physical and psychological illness. Variations in stressor assessment reflect different theories of how and why stressor exposure poses a risk to health.

Life events scales are the most commonly used means of naturalistic assessment of exposure to stressors, in distinction to the manipulation of stressor exposure in experimental research. Assessment of stressor exposure has been dependent on well-established research traditions in stress assessment. The first is the environmental perspective, which focuses on characteristics of events as risks for developing disease. The second is the psychological perspective, which focuses on perceptions and appraisals of threats. A third, and increasingly influential, is the life course perspective, which considers (1) the accumulation of stressors across the life span of individuals, (2) whether exposure to stressors early in life has long-term impact on health and well-being into adulthood, and (3) whether group differences in exposure to stressors (poor versus middle class; minority versus majority ethnic/racial status) are implicated in group-level differences in morbidity and mortality.

Four types of naturalistic stressor assessment predominate in the research literature: exposure to out-of-the-ordinary, demanding events, such as divorce, that have the capacity to change the patterns of life or arouse very unpleasant feelings (life events); self-reports of perceived stressfulness and appraisals of threat posed by events (stress appraisals); enduring

or recurrent difficulties and strains in an area of life (chronic stressors); and exposure to smaller, relatively minor, and normally less emotionally arousing events whose effects disperse within hours or within a day or two (daily events or hassles). Although conceived of primarily as comprehensive lists of exposure to discrete events outside the organism, life events scales often combine two or more of these approaches, sometimes including appraisals of stressors, measuring dimensions such as severity, frequency, controllability, and expectedness of events.

A frequent goal of research on stressor exposure is to establish the relationship between exposure to stressors and some health outcome in order to evaluate stress as a risk factor for disease. If the research question is to estimate the size of the association between stressor exposure and distressed mood, then the usual procedure is to use a retrospective, simple count of likely stressful life events occurring in the months prior to the measurement of mood. However, if the researcher is interested primarily in the relationship between stressor exposure and the onset, timing of onset, or recurrence of onset of a major clinical disorder, then the stressor exposure measure shifts to a determination of the type, severity, and timing of events and preexisting chronic stressors that preceded the onset of disorder. Studies of daily variation in mood or of the onset of more transient physical disorders such as colds or body aches use measures of short-term exposure to relatively minor stressors, daily events or hassles, sometimes in combination with measures of preexisting chronic stressors. Studies focusing on diseases that develop over long periods of time, such as heart disease, tend to use measures of exposure to long-term, recurrent, and persistent chronic stressors in major life roles. In practice, these methods are increasingly combined as more investigators apply the life course perspective to research on health and disease. There is considerable interest in combining the measurement of more recent (or proximal) stressors with retrospective measurement of more distal stressors, particularly severely threatening stressor exposures occurring in childhood and adolescence.

Major Life Events

Two contrasting methods of life event measurement have developed: checklist measures and personal interview measures. The latter uses qualitative probes in order to specify more precisely the characteristics of life events believed to produce the actual risk of illness and the timing of life events in relation to the outcome studies. Because of their ease of administration and the amount of exploratory health research

conducted, checklist measures predominate in studies of stressor exposure.

Checklist measures were derived from an environmental perspective on stress proposing that the basis of experienced stress is an event that brings about a demand for social, physical, or psychological readjustment. The earliest of these perspectives was the life change-readjustment paradigm, developed by Holmes and Rahe in 1967. Other theoretical perspectives on stress, such as those developed by Lazarus, Dohrenwend, and Brown, have augmented their approach in significant ways.

A typical checklist measure consists of a series of yes/no questions, asking participants to report whether any event of situation like the one described has occurred over a past period of time (from a few months to a year or more). The typical checklist measure is self-administered, although there are exceptions. Many checklist measures have been elaborated to assess more detailed information about the time of occurrence and severity. For example, some ask respondents to report the date of occurrence. Others ask for a brief written description of the situation or who else was involved in the situation. Still others ask respondents to rate the relative stressfulness of the event (from "not at all" to "very"). Self-report descriptive questions are used to estimate the severity of the event and its likely relationship to an outcome. However, the typical checklist measure does not rely on descriptive information. A typical measure sums the absolute number of event reports or sums investigator-assigned severity ratings for the events reported. The investigator-assigned ratings, termed normative, are survey-established estimates of the relative severity of stressors such as deaths, job losses, marital problems, and moves. Checklist methods yield a summary score of the estimated, normatively derived stressfulness of changes experienced over a period of time. Related to the latter point, checklist measures tend to be used in studies that use continuous symptom or mood measures as outcomes, although they are also sometimes used in exploratory studies of the onset of illness.

The ancestor of most current measures is the Holmes and Rahe social readjustment rating scale. This measure included both positive and negative events because Holmes and Rahe theorized that change per se was associated with changes in health status. Over time, checklist (and other) life events measures have moved away from the change-readjustment paradigm to an emphasis on negative or undesirable events. This trend was based on much replicated findings that undesirable events are more predictive of health problems than positive events are. Life event scales intended for use in the general

population have become more comprehensive and inclusive of events occurring to women, minorities, and other populations. Concomitantly, researchers have developed comprehensive events lists for age groups and special populations, including children, adolescents, and the elderly.

Despite their popularity, checklist measures have also been much criticized, most of these criticisms casting doubt on their reliability and validity as measures of stressor exposure. Criticisms include the vagueness and generality of the questions, which may lead respondents to misreport or overreport the occurrence of events or to report minor events as major. Life event scales have also been criticized for including events that are confounded with feeling states or psychiatric illness. There is also considerable evidence that life events scales suffer from recency bias: respondents are more likely to recall events that occur in the last month or two than those that occurred a year ago. Recall of events, moreover, is affected by mood at the time of recall.

The early development of personal interview methods for assessing life events used a theoretical perspective distinct from the change-readjustment paradigm, which informed the life events checklists. The major developer of interview methods, George W. Brown, proposed that social and environmental changes (and anticipations of those changes) that threaten the most strongly held emotional commitments are the basis for experienced severe stress. This perspective also holds that severe, rather than minor, stressors threaten health, distinguishing it from both change-readjustment and hassles paradigms.

Interview measures are not used commonly, primarily because of their greater expense and complexity. Investigators tend to use them under the following circumstances: (1) when more precise severity ratings are required; (2) where the relative timing of stress exposure and illness onset is critical to the study; (3) when the occurrence of an event, or series of events, may be related to respondent's preexisting illness or behavior (formally termed, independence from disorder); and (4) when there is interest in the specific effects of different types and severities of stressors on specific disease outcomes. Promoters of interview measures also claim that they are more comprehensive, reliable, and valid than checklist measures, although the debate is unresolved on this point. The complexity of interview methods precludes their use in lower-budget, very-large-sample (2000+), or more exploratory studies.

The Life Events and Difficulties Schedule (LEDS), developed by Brown and Harris, is the most widely used personal interview method. There are also several other personal interview methods, some based on

different rating systems (e.g., Paykel's Interview for Recent Life Events and Dohrenwend's Structured Event Probe and Narrative Rating Method). All are semistructured survey instruments assessing a wide variety of stressors, many including more distal stressors that may have occurred in childhood or adolescence. The interviews consist of a series of questions asking whether certain types of stressors have occurred over the past 12 months (or longer) and a set of guidelines or optional probes for determining more objective features of those events. The probes are used by investigators to rate the long-term contextual threat of events; contextual-threat rating is the key component of these methods because the occurrence of a severely threatening event is believed to pose a risk for the onset of illness. Contextual ratings for different types of events are documented in dictionaries and available to those who complete training in the interview methods.

Stress Appraisal

Appraisal plays an important part in life event scales, as well as theories of the stress process, most notably the transactional model of stress developed and refined by Lazarus and Folkman. The appraisal of stressors is believed to underlie the emotional experience of stress. Measures of appraisal focus on the degree to which an event threatens well-being or threatens to overwhelm a person's resources to cope.

Life event scales differ in whether they explicitly include appraisal as a component or explicitly exclude appraisal. Several measures of stressor exposure consist of assessments of how well an individual coped with stressors (although these measures are not predominant in health research). Checklist and interview measures of life events, for the most part, aim to exclude appraisal from stressor exposure assessment, although some explicitly do not. Those that exclude appraisal do so because of concerns that stress appraisal is confounded with the health and psychological outcomes that stressor exposure is supposed to predict. Indeed, researchers have speculated that some stress appraisals are caused by underlying, persistent mood disturbance or ill health, rather than vice versa. The transactional model of stress posits that appraisals can be measured separately.

Chronic Stressors

Recent research on stress and illness has turned toward emphasizing the role of persistent, continuous, or regular exposure to stressors as an important risk factor for the development of disease. There are four

main methods by which researchers measure chronic stress: (1) multi-item self-report interviews and questionnaires covering ranges of situations believed to produce chronic stressor exposure in a given setting, such as work-role demands, conflicts, and ambiguity; (2) intensive longitudinal surveys, such as daily-diary or experience-sampling methods, primarily using structured questions at regular (diary) or random (experience-sampling) intervals; (3) personal interview (contextual) measurement in life event interview instruments such as the LEDS; and (4) third-party observations of those believed to be at risk for stress exposure. The self-report interview or questionnaire is the most widely used method. Multi-item measures of chronic-stressors are an important class of life events scales.

A typical measure of chronic stressors is a set of questions designed to assess either the frequency or the severity of commonly occurring stressors in important life roles, such as work, marriage, and parenting. The strength of the multi-item approach is its grounding in the detailed, multidimensional assessment of environmental factors known to produce chronic stress, such as work and family. A weakness of the approach for some investigations is that the questions may be confounded with stress appraisal, particularly those scales assessing stressor severity.

Hassles or Daily Events

The research perspective in which measures of hassles were developed, the transactional model of stress (Lazarus), differs significantly from the research perspective in which major life event scales were developed. The hassles paradigm focuses attention on the potentially deleterious ways in which minor stressors, even those whose effects are relatively fleeting, can have negative impacts on health. Measures of hassles are life events scales, but assess exposure to minor hassles such as traffic jams, missed appointments, and quarrels with children. Although such events might in fact be major events for some people, the assumption of the hassles paradigm is that they are made more or less severe by the ways in which a person appraises and copes with the event.

Methods assessing daily events or hassles rely on diary methods of collection. These methods require participants to keep records of small events occurring over a given period of time, usually 1 day or a 24-h period, although weekly diaries have also been used. There are multiple styles of measurement: (1) open-ended approaches, which ask respondents to describe bothersome events of the day (not unlike personal interview measures, but usually self-administered

rather than conducted in person); (2) structured question approaches, which are simple yes/no response questions modeled on life events checklists; and (3) combinations of checklists with open-ended questions. Hassles scales are almost always administered in tandem with measures of stress appraisal.

Current hassles scales share the strengths and weaknesses of related approaches to the measurement of more major life events. Because diary methods of data collection rely on written and self-report, there is concern that they may confound objective event occurrence with unmeasured appraisal process. Almeida's Daily Inventory of Stressful Events uses telephone interviews and investigator ratings of narrative event descriptions to reduce some types of appraisal bias. Experience sampling methods, which involve beeping a respondent at random times during the day and asking questions about the current situation, have emerged as an alternative method for measuring daily stressor exposure.

See Also the Following Articles

Environmental Factors; Life Events and Health.

Further Reading

- Almeida, D. M., Wethington, E. and Kessler, R. C. (2002). The Daily Inventory of Stressful Events (DISE): an investigator-based approach for measuring daily stressors. *Assessment* 9, 41–55.
- Brown, G. W. and Harris, T. O. (1978). *Social origins of depression: a study of depressive disorder in women*. New York: Free Press.
- Brown, G. W. and Harris, T. O. (eds.) (1989). *Life events and illness*. New York: Guildford.
- Cohen, S., Kessler, R. C. and Gordon, L. U. (eds.) (1995). *Measuring stress: a guide for health and social scientists*. New York: Oxford University Press.
- Dohrenwend, B. S., Krashnoff, L., Ashkenasy, A. R., et al. (1978). Exemplification of a method for scaling life events: The PERI life events scale. *Journal of Health and Social Behavior* 19, 205–229.
- Dohrenwend, B. P., Raphael, K. G., Schwartz, S., Stueve, A. and Skodol, A. (1993). The structured event probe and narrative rating method for measuring stressful life events. In: Goldberger, L. & Breznitz, S. (eds.) *Handbook of stress: theoretical and clinical aspects*, pp. 174–199. New York: Free Press.
- Elder, G. H., George, L. K. and Shanahan, M. (1996). Psychosocial stress over the life course. In: Kaplan, H. B. (ed.) *Psychosocial stress: perspectives on structure, theory, life course, and methods*, pp. 247–292. San Diego, CA: Academic Press.
- Holmes, T. H. and Rahe, R. H. (1967). The social readjustment rating scale. *Journal of Psychosomatic Research* 11, 213–218.
- Kanner, A. D., Coyne, J. C., Schaefer, C., et al. (1981). Comparison of two methods of stress measurement: daily hassles and uplifts versus major life events. *Journal of Behavioral Medicine* 4, 1–39.
- Lazarus, R. S. (1999). *Stress and emotion*. New York: Springer.
- Lazarus, R. S. and Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.
- Paykel, E. S. (1997). The interview for recent life events. *Psychological Medicine* 27, 301–310.
- Stone, A. A. and Shiffman, S. (2000). Capturing momentary self-report data: A proposal for reporting guidelines. *Annals of Behavioral Medicine* 24, 236–243.

Lifestyle Changes See: Life Events and Health; Life Events Scale.

Lockerbie Air Crash, Stress Effects of

H Livingston

Dykebar Hospital, Paisley, UK

M Livingston

Southern General Hospital, Glasgow, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by H Livingston and M Livingston, volume 2, pp 629–633, © 2000, Elsevier Inc.

Researching the Impact of Disasters on Elderly People –
Some Methodological Issues

What Happened to the Elderly at Lockerbie

What Do We Know about the Response of Elderly People
to Experiencing Disasters?

Future Research

Glossary

Depression

A disorder of lowered mood typically associated with the loss of some or all of the following: interest, energy, pleasure, appetite, weight, concentration, and self-esteem. Disturbance of sleep, feelings of guilt, and wishing for death are also typical of depression.

Posttraumatic stress disorder (PTSD)

An anxiety disorder triggered by a traumatic and threatening event or events. Symptoms of PTSD are characteristically related to reexperiencing the event, the avoidance of triggers associated with the event, and increased arousal.

Researching the Impact of Disasters on Elderly People – Some Methodological Issues

Pan-Am flight 103 exploded in mid-air on December 21, 1988, over the small Scottish town of Lockerbie. This disaster continues to provoke media interest, not least because of the controversy over who were the perpetrators. The fireball of burning aviation fuel not only caused the death of all on board the aircraft but also death and injury to the inhabitants of the town. Survivors often had the harrowing experience of witnessing the mutilated bodies of the victims scattered over a wide area of the locality. The town hall was turned into a temporary mortuary as an army of police, firefighters, and other rescue services

descended on Lockerbie. Disasters such as Lockerbie, the Hillsborough football stadium tragedy in England, and Piper Alpha, in which an oil rig in the North Sea caught fire, provide unique opportunities to research mental disorder. The main factor in the etiology of the subsequent psychiatric problems – the traumatic stressor – is known in a disaster. This is an unusual occurrence, not only in mental health research but also in psychiatric clinical practice.

However, there are a number of special methodological problems in postdisaster research. Obviously disaster research must be retrospective, an assessment of an event that has occurred generally as a random act of nature or as the result of unforeseen human error. The presence of a large number of health-care professionals and other services brought to the site of the disaster may affect the research outcome in several ways. (In the case of Lockerbie the town's population increased from 3500 to 10,000 virtually overnight.) The health-care and support services may be effective and succeed in reducing the level of distress experienced by the disaster victims, for example through the effects of early psychological intervention programs, although these remain controversial. In addition, the psychological impact of the disaster on the rescue and recovery workers themselves is often considerable, contributes to the overall rate and level of psychological distress and psychopathology that results, and is often little acknowledged by the workers themselves until they become dysfunctional at work. This is often because of a culture in the rescue services such as the police and the armed forces that encourages a macho image and in which the acknowledgment of emotional problems is perceived by staff workers to have an adverse effect on careers. Many of the staff workers in these services are young and ill-prepared for tasks such as collecting and identifying human remains, having never previously witnessed death and dying. They are required to work at a disaster site that may be some distance from their own homes and usual family and friend support networks. The consequent media attention and the possibility of compensation for psychological damage as well as physical injury and material loss are additional factors that have to be considered in a disaster research design.

The method of assessment is important also in postdisaster studies. A desire for compensation may (in theory at least) lead to a desire on the part of

respondents to fake bad, or generate higher than warranted scores of distress. This is a phenomenon that is of even greater concern in the Internet era because of the ready access to questionnaires listing, for example, the phenomena of PTSD. In our view, therefore, such research should include not only self-report rating-scale measures but also some more objective evaluations such as clinical diagnostic assessments.

PTSD may persist for many years. A large study of Vietnam veterans found that the prevalence of PTSD was 15% after 19 years. It seems reasonable to expect that a significant number of disaster victims will continue to experience the psychiatric consequences of trauma into old age. This was certainly our impression when we were invited to examine survivors of the so-called Clydebank Blitz, the heavy German bombing of a small Scottish industrial town, important for the manufacture of ships and armaments, in March 1941 during World War II. The survivors of this event were recently encouraged to appeal to the benefits authorities in the United Kingdom for a war pension to compensate for the psychiatric consequences of exposure to two nights of intensive bombing that flattened large areas of the town and led to many casualties. Although not a systematic research study, our clinical impression was that many of these elderly people continued to experience psychiatric symptoms associated with an experience of traumatically induced stress more than 50 years prior to the clinical assessment, including full-blown PTSD.

A review has also examined the differing impact of a range of traumatic occurrences, such as the Holocaust, World War II, and the Vietnam and Korean wars. Curiously, elderly survivors of the Holocaust appeared to function well in life, unlike the combat veterans. In both groups, vivid imagery of what took place persisted, as well as disruption of sleep and a lack of trust in relationships. Furthermore avoidance of potential stressors that carried actual or symbolic relevance to the original traumatizing experience took place because these were likely to provoke a relapse or exacerbation of symptoms.

When the Lockerbie disaster occurred in 1988, there were no published studies on how traumatic stress impacted on elderly people at the time of the stress or soon after. It is possible to speculate that older people, faced with more loss-making events in a longer life, would either be better able to cope as a result or less so because of being sensitized by prior trauma. Solicitors acting for those involved in the Lockerbie explosion sought psychiatric input to assist the insurers in handling claims arising from the psychiatric damage sustained by the victims of the events

of December 1988. We decided therefore to study the clinical impact of traumatic stress on elderly people in the context of this Scottish disaster.

What Happened to the Elderly at Lockerbie

The initial assessment took place approximately 1 year after the incident, but avoided the anniversary of that event; the review took place 2 years later. Subjects were selected from a larger sample of survivors who had claimed that they had been traumatized by the explosion on the basis of a General Health Questionnaire (GHQ) score above the case cut-off level. The 31 elderly people in this category were seen, and their assessments were compared with 24 GHQ-case-designated younger survivors. Nineteen (61%) of the elderly were available for review, 2 had died, 1 was in the hospital for unrelated health problems, 1 had moved, 3 refused to participate, and 5 did not reply. Diagnosis was assigned using the *Diagnostic and statistical manual of mental disorders* (3rd edn. rev.; DSM-III-R). This classification system has since been revised, and DSM-IV is now in use. The psychological impact was measured by the Leeds scales, which assess anxiety and depression; the GHQ28, a measure of global psychological functioning; and the Revised Impact of Events scale (RIE), which quantifies posttraumatic symptoms.

The older survivors did not differ significantly from the younger in GHQ, Leeds, and RIE responses. Average GHQ28 scores were well above case level (score 4/5) 1 year after the disaster in both groups (younger 9.83 [SD, 8.20]; older 11.76 [SD 8.74]; not significant). Anxiety rather than depression tended to predominate on the GHQ and the Leeds scales. The younger and older survivors recorded high scores on the RIE intrusion (younger 15.09 [SD 9.44]; older 18.71 [SD 12.84]; not significant) and avoidance (younger 15.08 [SD 9.99]; older 17.71 [SD 11.76]; not significant) subscales. Clinically, all but five elderly survivors fulfilled DSM-III criteria for PTSD, as did all the younger subjects (Fisher test, not significant).

At the follow-up 2 years later, the proportion of GHQ cases among the elderly survivors had declined from 71 to 47%, but this was not statistically significant. The mean Leeds depression (initially 7.40 [SD 4.40]; at follow-up 3.76 [SD 2.59]; $p < 0.004$) and anxiety (initially 15.47 [SD 10.60]; at follow-up 4.87 [SD 4.87]; $p < 0.001$) ratings had declined significantly (see [Figure 1](#)). Both avoidance behavior and the level of intrusive thoughts (RIE) had also declined significantly (paired $t = 2.05$, $p = 0.07$; paired $t = 3.69$, $p = 0.003$, respectively) (see [Figure 2](#)). The proportion

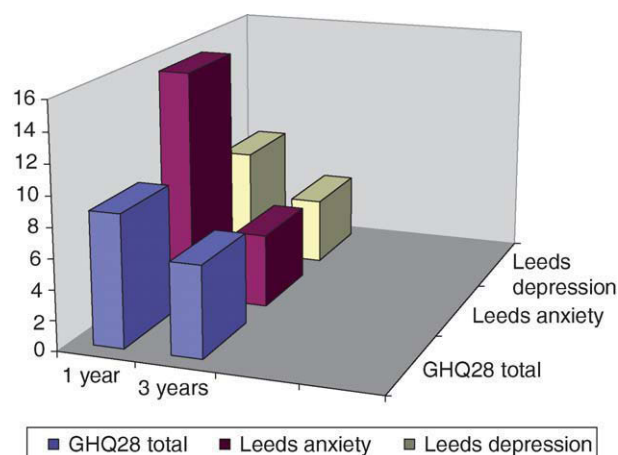


Figure 1 Emotional distress in elderly Lockerbie survivors, Leeds scales and General Health Questionnaire.

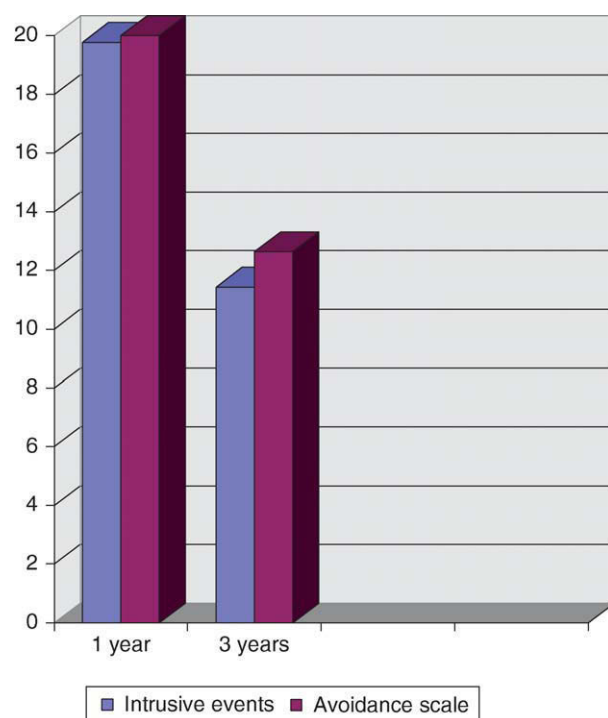


Figure 2 Impact of the disaster on elderly survivors, Revised Impact of Events scale.

of PTSD cases had declined significantly from 14/19 (74%) to 3/19 (16%) ($\chi^2 10.64$, $df = 1$; not significant). Significantly, no new cases of PTSD had developed during the follow-up period. Persistent PTSD was associated with high Leeds anxiety and GHQ anxiety scores (anxiety and insomnia subscale and social dysfunction subscale). Furthermore, there was a positive correlation between persistent PTSD and a diagnosis of depression on DSM-III criteria, although none between a PTSD diagnosis and individual RIE items.

What Do We Know about the Response of Elderly People to Experiencing Disasters?

At the time we undertook our studies of elderly survivors following the Lockerbie disaster we were unaware of any other work on elderly people in such situations apart from occasional single case studies. Subsequently, reports on elderly victims of two other disasters appeared: the 1988 earthquake in Armenia and the 1995 Hanshin-Awaji earthquake in Japan.

The Armenian earthquake study was undertaken approximately 18 months after that event and involved the assessment of a total of 179 elderly and younger subjects. In young and old, as might be expected, the most traumatized were those who were closest to the earthquake epicenter. Age did not affect the level of distress suffered, regardless of the location of the victim relative to the epicenter, but elderly people tended to experience more symptoms indicative of emotional hyperarousal than did the young and to have less evidence of intrusive symptoms.

Following the Japanese Hanshin-Awaji earthquake, a study of the early psychological impact of the disaster on its victims was carried out that divided the subjects into two groups using an age cut-off of 60 years. Sixty-seven younger subjects and 75 older subjects were seen 3 weeks after the disaster. Fifty younger subjects and 73 older subjects were subsequently reviewed at 8 weeks after the disaster. The older people showed a significant decrease in symptoms in contrast with those under 60. The authors attribute this to the presence of more extensive social networks and some previous disaster experience among the older group of subjects, although it could also be argued on strong clinical grounds that previous experience of severe stress sensitizes rather than protects the sufferer from future stress related reactions.

A subsequent review looked at the effects of PTSD in groups of elderly people who were World War II veterans, were Holocaust survivors, and had experienced a natural disaster in later life. This review suggested that PTSD symptoms in the elderly can be persistent or intermittent. The severity of the trauma and premorbid psychiatric illness predisposed this group to the development of PTSD, and good psychosocial support seemed to protect against it. Older adults did not seem any more predisposed to developing PTSD than younger adults, and the symptoms appeared to be similar across the age groups.

Sadly, several other large-scale natural disasters have occurred recently – in 2004–2005 alone, the Asian tsunami, Hurricane Katrina, and the South Asian earthquake. To date, no studies of the responses of elderly people to these events have been published.

Obviously, the nature of disasters is such that they do not permit prospective study. Our sample of elderly victims was small, but remains the only long-term prospective follow-up of elderly people after a natural disaster. The multiple standardized assessment techniques used, including the application of diagnostic criteria in all cases, showed a remarkable degree of internal consistency. The major methodological flaw is that all our subjects were engaged in the process of claiming compensation from the airline on whose aircraft the explosion took place. We found that the elderly subjects had an initial emotional response to the disaster that was similar to the younger subjects, meeting the psychological case criteria in high numbers and recording high levels of anxiety and diagnosed PTSD. The level of distress diminished significantly over the 3-year time period of the studies, although one in six of the elderly subjects still had PTSD at follow-up, comparable to the findings in a study of the survivors (of all ages) of an Australian brush fire. No links between a PTSD diagnosis and specific aspects of the disaster were found at Lockerbie, although the disaster had a number of unique aspects, including a delay in recognizing that it was a terrorist act and in bringing the suspects to justice, which might be expected to adversely impact survivors' coming to terms with the event (analogous to mourning).

Future Research

Future work should attempt to assess a complete population of elderly survivors following a disaster early on and to review them over the long-term to obtain a clear idea of how traumatic stress in such situations affects older people. This work, just like the disaster-recovery effort, may require international organization, involving the creation of a disaster assessment team that can be mobilized at short notice. Such a team would be preequipped with a research design and appropriate measures. To maintain objectivity, the team needs to avoid being part of the therapeutic input or, indeed, the assessment process if a medical-legal issue is involved, such as a compensation claim. At the same time, the researchers have to strive to avoid artificially creating more stress by prolonging unnecessarily the process of assessment for an already bewildered and traumatized group of people.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Aging and Stress, Biology of; Community Studies; Disaster Syndrome; Disasters and Mass Violence, Public, Effects of; Earthquakes, Stress Effects of; Posttraumatic Stress Disorder, Delayed.

Further Reading

- American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd edn.). Washington, DC: American Psychiatric Association.
- Goenjian, A. K., Najarian, L. M., Pynoos, R. S., et al. (1994). Post-traumatic stress disorder in elderly and younger adults after the 1988 earthquake in Armenia. *American Journal of Psychiatry* 151(6), 895–901.
- Goldberg, D. (1978). *The manual of the General Health Questionnaire*. Windsor, UK: National Foundation for Educational Research.
- Hilton, C. (1997). Media triggers of post-traumatic stress disorder after the second world war. *International Journal of Geriatric Psychiatry* 12(8), 862–867.
- Horowitz, M., Wilner, N. and Alvarez, W. (1979). Impact of events scale: a measure of subjective stress. *Psychosomatic Medicine* 41, 209–218.
- Kato, H., Asukai, N., Miyake, Y., et al. (1996). Post-traumatic symptoms among younger and elderly evacuees in the early stages following the 1995 Hanshin-Awaji earthquake in Japan. *Acta Psychiatrica Scandinavica* 93(6), 477–481.
- Kulka, R. A., Sclenger, W. E. and Fairbank, J. A. (1990). *Trauma and the Vietnam War generation*. New York: Brunner/Mazel.
- Livingston, H. M., Livingston, M. G., Brooks, D. N., et al. (1992). Elderly survivors of the Lockerbie air disaster. *International Journal of Geriatric Psychiatry* 7, 725–729.
- Livingston, H. M., Livingston, M. G. and Fell, S. (1994). The Lockerbie disaster: a 3 year follow-up of elderly victims. *International Journal of Geriatric Psychiatry* 9, 989–994.
- McFarlane, A. C. (1989). The aetiology of post traumatic morbidity, predisposing, precipitating and perpetuating factors. *British Journal of Psychiatry* 154, 221–228.
- Sadavoy, J. (1997). Survivors: a review of the late-life effects of prior psychological trauma. *American Journal of Geriatric Psychiatry* 5(4), 287–301.
- Snaith, R. P., Bridge, G. W. K. and Hamilton, M. (1976). The Leeds scales for the self assessment of anxiety and depression. *British Journal of Psychiatry* 128, 156–165.
- Weintraub, D. and Ruskin, P. E. (1999). Posttraumatic stress disorder in the elderly: a review. *Harvard Review of Psychiatry* 17(3), 144–152.

for an already bewildered and traumatized group of people.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Aging and Stress, Biology of; Community Studies; Disaster Syndrome; Disasters and Mass Violence, Public, Effects of; Earthquakes, Stress Effects of; Posttraumatic Stress Disorder, Delayed.

Further Reading

- American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd edn.). Washington, DC: American Psychiatric Association.
- Goenjian, A. K., Najarian, L. M., Pynoos, R. S., et al. (1994). Post-traumatic stress disorder in elderly and younger adults after the 1988 earthquake in Armenia. *American Journal of Psychiatry* 151(6), 895–901.
- Goldberg, D. (1978). *The manual of the General Health Questionnaire*. Windsor, UK: National Foundation for Educational Research.
- Hilton, C. (1997). Media triggers of post-traumatic stress disorder after the second world war. *International Journal of Geriatric Psychiatry* 12(8), 862–867.
- Horowitz, M., Wilner, N. and Alvarez, W. (1979). Impact of events scale: a measure of subjective stress. *Psychosomatic Medicine* 41, 209–218.

- Kato, H., Asukai, N., Miyake, Y., et al. (1996). Post-traumatic symptoms among younger and elderly evacuees in the early stages following the 1995 Hanshin-Awaji earthquake in Japan. *Acta Psychiatrica Scandinavica* 93(6), 477–481.
- Kulka, R. A., Schlenger, W. E. and Fairbank, J. A. (1990). *Trauma and the Vietnam War generation*. New York: Brunner/Mazel.
- Livingston, H. M., Livingston, M. G., Brooks, D. N., et al. (1992). Elderly survivors of the Lockerbie air disaster. *International Journal of Geriatric Psychiatry* 7, 725–729.
- Livingston, H. M., Livingston, M. G. and Fell, S. (1994). The Lockerbie disaster: a 3 year follow-up of elderly victims. *International Journal of Geriatric Psychiatry* 9, 989–994.
- McFarlane, A. C. (1989). The aetiology of post traumatic morbidity, predisposing, precipitating and perpetuating factors. *British Journal of Psychiatry* 154, 221–228.
- Sadavoy, J. (1997). Survivors: a review of the late-life effects of prior psychological trauma. *American Journal of Geriatric Psychiatry* 5(4), 287–301.
- Snaith, R. P., Bridge, G. W. K. and Hamilton, M. (1976). The Leeds scales for the self assessment of anxiety and depression. *British Journal of Psychiatry* 128, 156–165.
- Weintraub, D. and Ruskin, P. E. (1999). Posttraumatic stress disorder in the elderly: a review. *Harvard Review of Psychiatry* 17(3), 144–152.

Loss Trauma

A Bifulco

Royal Holloway, University of London, London, UK

© 2007 Elsevier Inc. All rights reserved.

Introduction

Loss Trauma Events

Responses to Loss Trauma

Childhood Loss Trauma

Adult Loss Trauma

Treatment for Loss Trauma

Resilience to Loss Trauma

Glossary

Complicated bereavement A pathological response to loss in which the usual bereavement process is impeded leading to psychopathology.

Eye movement desensitization and reprocessing (EMDR) Loss

A treatment for trauma responses involving desensitization and rapid eye movements.

Loss trauma

Death or long-term separation from a close other, but also experiences with loss of self-identity, security, or trust in others.

Posttraumatic stress disorder (PTSD) Traumatic event

A type of event combined with a pathological response, which involves characteristics of both trauma and loss.

A classified anxiety disorder, with distinctive symptoms, specifically in response to a traumatic event.

A type of severe life event, often involving horror, which has fundamental implications for one's safety and place in the world. Can be single or multiple.

Traumatic grief

A pathological response to death of a close other with symptoms similar to PTSD.

Introduction

The investigation of loss trauma is a relatively new psychological subdiscipline, based on the specification of a particular type of traumatic event (that of the death of a close person), with the identification of specific pathological responses, which, although similar to those identified in relation to other traumatic events, have identifiable symptoms associated with a grief response. Its identification as a new focus of study is marked by a new journal (*Journal of Loss and Trauma*) and by organizations such as the National Institute for Trauma and Loss in Children. It is also marked by variants of treatment protocols, which attend specifically to elements of the loss in trauma. However, there is as yet no widely recognized classification of traumatic grief as a disorder. The word trauma has a dual meaning, referring both to an event or circumstance and to the psychological impact of that event on an individual's well-being. In this article these meanings will be differentiated by traumatic event and trauma response.

Loss Trauma Events

A traumatic event is a type of life event, that is, an external circumstance that happens at a defined time. However, as trauma, it constitutes an event with a particularly high level of threat and unpleasantness, which has implications for a person's sense of safety and place in the world. Thus, it is usually sudden and unpredicted and is often accompanied by circumstances that involve horror or intense shock. Trauma events are usually public (e.g., witnessing a suicide, being involved in a terrorist bomb attack, being in the line of a major flood or tsunami) and are common in combat and war situations with both soldiers and civilians. However, trauma can also be private, for example, physical or sexual assault, which involves threat to life or personal integrity. Trauma events can be single or multiple, with different pathological responses associated with them. Loss trauma events are often single when deaths of others are involved. However, a context of severe illness, caring responsibility, or enforced separation may constitute a multiple or chronic course to the traumatic experience.

Loss trauma identifies traumatic circumstances specifically with loss events, usually the death of a close person. This usually involves a parent, partner, child, or close relative or friend, and usually under unexpected or horrific circumstances. Thus, suicides, accidents involving physical mutilation, or circum-

stances with a random aspect such as freak accidents would all be encompassed in loss trauma events. Loss trauma can also include losses of children or babies, for example, birth of a stillborn child or a child with multiple handicaps or deformities. As a category of event, loss events not only can involve deaths of close others and separations from a close other (e.g., parent, partner, child), but also can involve psychological losses such as loss of security (e.g., loss of job or home), sense of self (e.g., loss of reputation or physical losses such as limbs), or a meaningful relationship (e.g., revelation or betrayal). However, there has been less investigation of these wider losses in terms of loss trauma.

Responses to Loss Trauma

Any traumatic and major loss event will have some immediate psychological consequences such as shock, distress, denial, or numbing and possibly flashbacks to the event. However, in a small proportion of cases, posttraumatic stress disorder (PTSD) will ensue, a classified psychological anxiety disorder. Complicated bereavement responses may set in whereby the bereavement process is incomplete, the loss is unresolved, and traumatic grief may occur. Traumatic grief is seen as a variant of other disorders such as depression and PTSD and does not have a distinct classification in standard psychiatric manuals. However, there is ongoing debate about whether traumatic grief should constitute a disorder in its own right.

Posttraumatic Stress Disorder (PTSD)

PTSD involves response to a traumatic event described as actual or threatened death or serious injury, or a threat to the physical integrity of self or others. PTSD is one of several anxiety disorders. It was originally used for explaining combat reactions in terms of shell shock but is now linked with a range of traumatic events. PTSD symptoms exist for at least 1 month and include preoccupation with the horror of incident, hypervigilant scanning of the environment for potential threat or danger, avoidance of situations, futility about the future, shattered world view, excessive anger, cognitive dysfunction, and prolonged impaired social or occupational functioning. PTSD is often typed by whether the trauma provoking it is acute or chronic, that is, single exposure or multiple incidents and exposure. The latter is considered the most damaging in terms of functioning, particularly when occurring in childhood. In terms of prevalence, while it is estimated that most people (90%) experience at least one traumatic event in a lifetime, only 10% develop a history of PTSD. This is more likely

after certain events. For example, prevalence ranges from as high as 50% in response to rape to less than 1% after hearing of serious injury to close relative or friend.

Complicated or Pathological Bereavement

When the response to loss (usually the death of a close person) does not follow the usual progression through the stages of grief work leading to acceptance, it is considered a pathological bereavement response. It impedes the development of new ties and inhibits adaptation to the loss both personally and socially. When this occurs in childhood, it is argued, it impedes adult functioning and leads to long-term problems. However, not all researchers accept the inevitability of grief work and claim that adaptive responses can occur without evidence of following particular stages in the process. Grief work thus may be required only following more severe losses or for individuals with high levels of distress or vulnerability to loss.

Traumatic Grief

Traumatic grief is a relatively new term that combines trauma with bereavement or grief responses. It is provoked by the death of a significant other and includes symptoms similar to PTSD but specifically focused on the lost person, including intrusive, distressing preoccupation with the deceased, hypervigilant scanning of the environment for cues of the deceased, the wish to be reunited with the deceased, separation anxiety features, futility about the future, difficulty acknowledging the death, shattered world view, and anger together with impaired social functioning.

Studies have examined traumatic loss among groups such as elderly caregivers of terminally ill spouses, young adults who lost a friend to suicide, and parents who lost a child in a traffic accident. These have shown ensuing symptoms in two categories: (1) separation distress, involving preoccupation with thoughts of the deceased, longing, searching, loneliness, and impaired functioning, and (2) traumatic distress, feeling disbelief, anger, mistrust, and detachment from others.

Peritraumatic Dissociation

Peritraumatic dissociation is a dissociated response to loss with expression of psychological shock, as indicated by a sense of reality having been abruptly altered. Thus, being in a daze, feeling disoriented, numb, or on automatic pilot, and experiencing the surroundings as unreal are typical, together with star-

ing into space or feeling like a spectator. This response is usually associated with traumatic grief. The greater the experience of shock and overwhelming emotion, the more likely the dissociation, with PTSD following afterward. This is similar to the symptoms described in complicated bereavement or traumatic grief.

Childhood Loss Trauma

Loss trauma including loss of parents, loss of partners, and loss of close friends may be expected to occur at any life stage. However, when such losses occur off-time (i.e., nonnormatively), this may be a contribution to their traumatic nature and impact. Thus, losses of parents in childhood, losses of children, and widowhood in early- or mid-adulthood are all candidates for traumatic loss.

Psychoanalytic approaches to depression emphasize its relationship to loss experience, particularly in childhood. As a result there has been much investigation of childhood loss of a parent and vulnerability to adult disorders such as depression. Detailed investigation of childhood deaths of parents has shown that these have marked long-term impact only when neglectful or abusive parenting follows, or when the loss is considered aberrant in terms of abandonment. When death of a parent in childhood occurs in otherwise caring and supported circumstances, no long-term dysfunction occurs.

Childhood trauma has been extensively studied by Leonie Terr, who identified different symptoms following from single-blow or multiple-blow trauma. Type I trauma, that following a sudden single shock, was found to evince symptoms such as full repeated memories, strangely visualized or repeated, repetitive behaviors, trauma-specific fears, omens, and misperceptions. In contrast, type II trauma, following multiple or chronic shock, evinced denial and numbing, dissociation, and rage. In both instances childhood trauma led to changed attitudes about people, aspects of life, and the future.

Adult Loss Trauma

The most typical adult losses concern loss of a child or pregnancy and loss of a partner. These events can occur at any age, but death of a partner is clearly more common (and more normative) in later life. Again, the traumatic nature of the losses include suddenness and horror. PTSD is observed after pregnancy loss, and given that most losses of pregnancy are sudden and unexpected, these can constitute trauma, although those pregnancies that are lost in later weeks (or taken to term and ending in stillbirth) or those in which the fetal loss involves pain and bloodshed are

more likely to constitute loss trauma. However, even with miscarriage in earlier weeks, a small proportion (7%) develops PTSD symptoms. However, more common responses are depression (30–50%) or anxiety (32%) occurring within 6 months of miscarriage. Research into widowhood (e.g., by Colin Murray-Parkes in the 1970s) showed individual variation in responses, but showed complicated bereavement with aspects such as intrusive thoughts, avoidant behaviors, yearning, loneliness, and loss of interest. Further work with elderly widows following an extensive caring role showed distinct traumatic grief responses.

Treatment for Loss Trauma

The core of treatment for trauma involves being able to talk about the traumatic experience within a therapeutic trusting relationship. The dilemma for the therapist is that reminders of trauma can trigger symptoms of PTSD, so symptoms can in the short term be worsened. Thus, trauma treatment must balance the degree of processing (of negative information) with containment (by the therapeutic relationship) in a secure setting. There are various approaches to treating trauma, which are not necessarily specific to loss trauma.

Critical Incident Stress Debriefing

Debriefing was originally developed for limited use as a brief group intervention to help mitigate psychological distress among emergency response personnel. This was then developed to be used on an individual rather than a group basis. However, global application of debriefing has been shown to be relatively ineffective and may at times impede recovery. This may be because it has been removed from its original purpose as a group bonding exercise for emergency workers, but also because it may impede the natural support that individuals can receive from those close at times of emergency. It has been argued that formalized debriefing can pathologize normal reactions and thus neutralize normal coping responses.

Cognitive-Behavioral Treatment (CBT)

CBT treatments used for affective disorder can help individuals to understand and manage anxiety associated with the trauma-related stimuli. Exposure

treatment involving repeated confrontations with memories of the traumatic stressor in therapeutic settings has shown successful outcomes, although not specifically adapted for trauma responses. CBT can be helpful in conjunction with other techniques such as eye movement desensitization and reprocessing.

Eye Movement Desensitization and Reprocessing (EMDR)

EMDR was discovered by Francine Shapiro; it is a desensitization process that involves bringing the traumatic memory to mind while the patient simultaneously moves the eyes from side to side, following the movement of the therapist's finger back and forth in front of the face. After a series of eye movements the patient is instructed to let go of the traumatic images and say what comes to mind. Over the course of successful treatment, the emotional distress abates, traumatic images become less intrusive, and beliefs about the self become more positive. There is general agreement that the technique is successful in conjunction with other therapeutic techniques, but the precise mechanism for its action is not well understood, although it may directly work on brain mechanisms.

Dialectic Behavior Therapy (DBT)

DBT was invented by Marcia Linehan and although originally used mainly for borderline personality disorder in which self-harm is involved, it was found to also be effective for trauma. The therapy requires learning skills for coping with painful emotional states that lead to self-destructive behavior. Behavioral skills and emotional control are enhanced through emphasis on actively teaching and reinforcing adaptive behavior. It can be conducted both on an individual psychotherapy basis and in group meetings. Again, it is not specifically designed for loss trauma but has elements required to deal with some traumatic grief responses.

Resilience to Loss Trauma

Experience of traumatic events is common, but experience of trauma psychopathology is fairly rare. Therefore, it is argued, resilience must exist in the face of encountering trauma or, conversely, some

individuals are vulnerable or sensitized to loss trauma. As with other psychological disorders, factors influencing vulnerability to PTSD include lack of social support, lack of education, family background, prior psychiatric history, and aspects of trauma response such as dissociative states.

It has been argued that resilience to loss trauma has been underestimated because of research focus on individuals developing disorders, which biases the field toward vulnerable individuals. For example, absence of grief, usually pathologized among patients, may in fact indicate resilience in hardier individuals. It is argued that there is little empirical support for the notion of delayed grief and that some individuals may adapt quickly to loss without going through grieving stages. Similarly, PTSD-type symptoms are common in the first few weeks after a traumatic event, but in most people these disappear and the full syndrome does not occur, thus indicating some process of resilience. Resilient and recovering individuals are often put into same category, since chronicity of symptoms is a key marker for pathological reactions and more resilient individuals adapt more quickly and may be perceived to have been treated rather than have drawn on personal resilience factors.

See Also the Following Articles

Childhood Stress; Childbirth and Stress; Dissociation; Familial Patterns of Stress; Grieving; Posttraumatic Stress Disorder in Children; Posttraumatic Stress Disorder, Delayed; Posttraumatic Therapy; Trauma and Memory; Trauma Group Therapy.

Further Reading

- Allen, J. (2005). *Coping with trauma: hope through understanding* (2nd edn.). Washington, D.C: American Psychiatric Publishing.
- Bifulco, A., Harris, T. and Brown, G. W. (1992). Mourning or early inadequate care? Reexamining the relationship of maternal loss in childhood with adult depression and anxiety. *Development and Psychopathology* **4**, 433–449.
- Bonnano, G. A. (2004). Loss, trauma and human resilience. *American Psychologist* **59**, 20–28.
- Bowlby, J. (1980). *Loss sadness and depression: vol. 3: Attachment and loss*. London: Hogarth Press and Institute of Psychoanalysis.
- Figley, C. R., Bride, B. E. and Mazza, N. (eds.) (1997). *Death and trauma: the traumatology of grieving*. Washington, D.C: Taylor Francis.
- Herman, J. L. (1992). *Trauma and recovery*. London: Pandora.
- Horowitz, M. J. (1986). *Stress response and syndromes: PTSD, grief and adjustment disorders* (3rd edn.). Northvale, NJ: Jason Aronson.
- Parkes, C. M. (1986). *Bereavement studies of grief in adult life*. Harmondsworth, UK: Penguin.
- Prigerson, H. G., Shear, M. K., Frank, E., Silberman, R. and Reynolds, C. F. (1997). Traumatic grief: a case of loss-induced trauma. *American Journal of Psychiatry* **154**, 1003–1009.
- Terr, L. C. (1990). *Too scared to cry*. New York: Harper & Row.
- Van der Kolk, B. A., McFarlane, A. C. and Weisaieth, L. (eds.) (1996). *Traumatic stress: the effects of overwhelming experience over mind, body and society*. New York: Guilford.

Lymph Nodes

T L Whiteside

University of Pittsburgh Cancer Institute,
Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
T L Whiteside, volume 2, pp 634–640, © 2000, Elsevier Inc.

Architecture of Lymph Nodes

Lymphocyte Migration within the Lymph Node

Antigen-Driven Interactions in the Lymph Node

Significance of Lymph Nodes

Interactions between the Central Nervous System
and Lymph Nodes

Glossary

*B cell-
dependent
area*

Primary and secondary follicles located in the superficial cortex, underneath the subcapsular sinus. B cells migrate to and are organized into follicles.

Cortex

A major area of the lymph node consisting of a superficial cortex and a deeper region, the paracortex.

*Fibroblast
reticular cell*

Network composed of collagen-containing reticular fibers that are produced by

<i>(FRC)</i> <i>network</i>	FRCs. It forms the infrastructure of the lymph node.
<i>Lymph node</i> <i>(LN)</i>	A key organ of the peripheral immune system.
<i>Medulla</i>	The second major area of the lymph node located beneath the paracortex, closest to the hilum, and subdivided into medullary cords and medullary sinus, which drains into the efferent lymphatic vessel.
<i>Paracortical cords</i>	The paracortex is made up of multiple cords. Each is a functional repeating unit of the cortex, which stretches from a medullary cord to the base of a B cell follicle. It consists of the central high endothelial venule (HEV) surrounded by a perivenular channel, a corridor, and a cortical sinus. HEVs are specialized segments of vessels localized deep in the paracortex, whose luminal surface is lined by high cuboidal endothelial cells serving as transmigration points for lymphocytes emigrating from the bloodstream to the lymph node. Perivenular channels are narrow spaces along the HEVs into which lymphocytes arrive, having crossed the endothelium, and where they seem to accumulate. To reach the corridor, they must next cross the barrier of overlapping fibroblast reticular cells, which form the perivenular sleeve. In the corridor, lymphocytes are able to migrate more freely and interact with antigen-presenting cells lining the corridor walls. The cortical sinus is a fluid-filled compartment that bounds the outer surface of the paracortical cords and connects the subcapsular sinus with the medullary sinuses.
<i>T cell-dependent area</i>	The interfollicular regions and the paracortex of the lymph node. T cells migrate preferentially to these anatomical locations.

Lymph nodes (LNs) are fundamental functional units of the immune system. They are the organs in which immune cells transact their business. Distributed strategically throughout the body, they are localized close to major junctions of the lymphatic channels. Naive lymphocytes recirculate continuously through the peripheral lymphoid organs, including LNs, passing from blood to lymph and then back to blood via the thoracic duct. LNs play a key role in collecting information necessary for entering lymphocytes to make a decision about whether to respond to an antigen or to become tolerant and/or die, thus avoiding autoimmunity. Lymph draining the extracellular spaces of the body carries antigens from tissue sites to the nearest LN, where afferent lymphatic vessels converge. Antigens thus delivered to the LN are trapped, screened, and presented to those rare immune cells in the pool of T and B lymphocytes migrating through the node that are capable of antigen recognition. The process of recognition involves Ig receptor (e.g., 1 in 100 000 cells) on B cells and T cell receptor (TCR) on T lymphocytes and initiates a series of events orchestrated within the milieu of the LN, culminating in generation of an antigen-specific immune response. Virtually all primary immune responses take place in LNs and other secondary lymphoid organs, such as spleen or gut-associated lymphoid tissue. Hence, LNs are extraordinarily important sites of interactions between immune cells and antigens and are especially adapted to facilitate these interactions.

Architecture of Lymph Nodes

LNs are small, oval structures (about 0.5 cm or less in diameter) with complex architecture, which differs

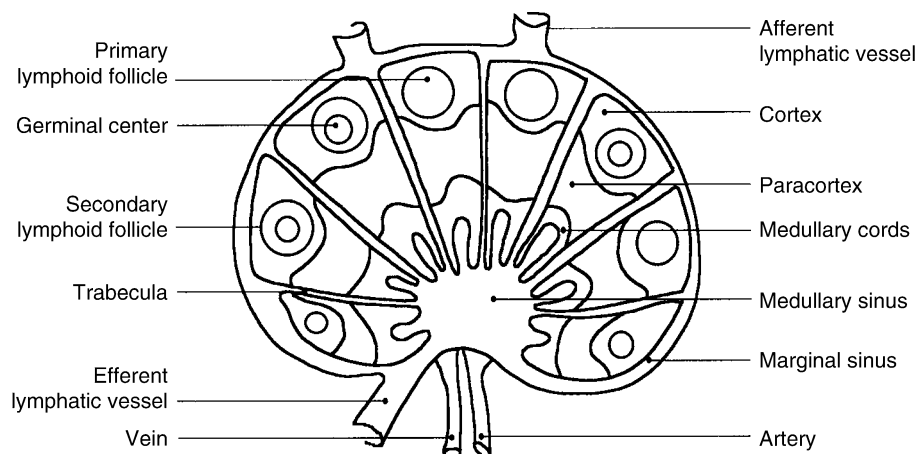


Figure 1 A schematic view of a lymph node in cross-section, showing its various components.

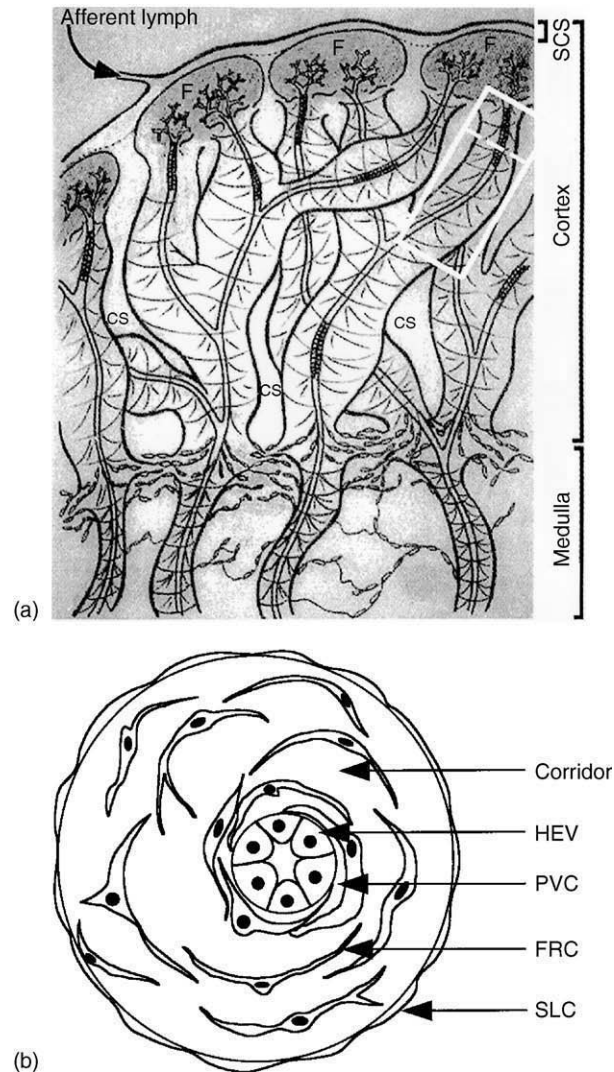


Figure 2 Functional units of a lymph node. (a) Paracortical cords in a longitudinal view. A white box indicates a representative cord. Note an extensive network of reticular fibers surrounding each postcapillary venule and high endothelial venules (HEV) shown as darker regions of the venules. Paracortical cords begin at the base of follicles (F), run the length of the cortex, and coalesce into medullary cords at the corticomedullary junction. Spaces around paracortical cords are cortical sinuses (CS). SCS, subcapsular sinus at the margin of the LN. Reproduced with permission from Kelly (1975). (b) A cross-section of a paracortical cord, indicating a control HEV surrounded by perivascular channels (PVC) and corridors supported by a framework of fibroblastic reticular cells (FRC) and bounded by sinus-lining cells (SLC).

between species. The basic structural elements of lymph nodes are the cortex and the medulla (Figure 1). A section through a LN shows the outermost capsule; the subcapsular sinus draining afferent lymphatic vessels; the cortex, consisting of an outer region made up of B lymphocytes organized into lymphoid follicles and of deeper paracortical cords made up mainly of T lymphocytes; the medulla, organized into medullary cords, consisting of strings of macrophages and

antibody-secreting plasma cells; and medullary sinus connected with efferent lymphatic vessels. This basic architecture of LNs undergoes dramatic changes during an antigenic challenge or in response to neural or inflammatory stimuli, and it is this remarkable plasticity of the critical architectural elements of a lymph node that facilitates immune cell interactions.

The cortex of a lymph node in larger animals and humans is subdivided into 10 or more lobules separated by fibrous bands called trabeculae (Figure 1). Each lobule has the follicle with a germinal center, the B lymphocyte-dependent area, and the paracortex corresponding to the T cell-dependent area. This area accommodates paracortical cords, which form the fundamental structural unit of the lymph node paracortex. Figure 2a provides a lateral view of paracortical cords, each with a central postcapillary high endothelial venule (HEV) bordered by the cortical sinus. Paracortical cords begin at the base of follicles and run parallel to each other to the corticomedullary junction, where they coalesce into medullary cords. In cross-section, a paracortical cord is a cylinder composed of sinus-lining cells at its periphery, concentric rings of barrier-type cells called fibroblastic reticular cells (FRCs), and a central HEV (Figure 2b). The concentric arrangement of FRCs creates a perivenular channel. The spaces between the rings are filled with lymphocytes and other cells and FRCs from an outermost layer of the channel, enclosing it like a sleeve. Outside of the sleeve is a fluid-filled corridor bounded and supported by a network of FRCs. The corridors are 10–25 μm in diameter, and they join with and split from each other, creating no barriers and allowing for efficient cell migration. The boundary of flat sinus-lining cells separates the corridors from the surrounding paracortical sinus. A representative radius of a cord is 50–100 μm or roughly 10 lymphocyte diameters. However, paracortical cords are pliable structures, capable of stretching to easily accommodate masses of migrating lymphocytes.

The outermost aspect of a lobule, situated underneath the subcapsular sinus, is the B-dependent area of a LN. It is composed of tightly packed B lymphocyte formations referred to as primary follicles. Germinal centers are formed within follicles on antigenic stimulation of B cells and represent areas of extensive B cell proliferation. Occasional T cells are also present. An important cell resides in the follicular area, the so-called follicular dendritic cell (FDC), which is responsible for trapping and presenting antigens to B cells. The FDC is a highly specialized type of DC that binds immune complexes via Fc and complement receptors, internalizes them, and holds them for long periods of time. This function of FDC is crucial for memory B cell development. The germinal centers are

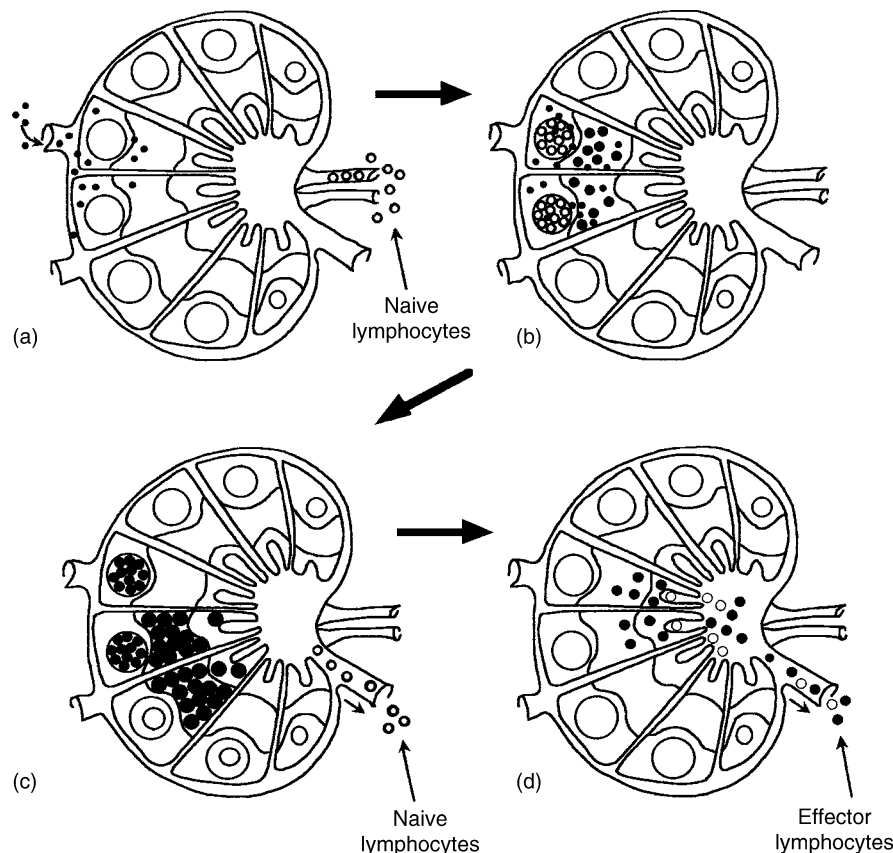


Figure 3 A diagram illustrating continuous trafficking of lymphocytes through a LN. (a) Antigen enters the nodes through afferent lymphatics, and small naive lymphocytes enter via the blood vessels from the bloodstream. In the cortex, lymphocytes enter through HEV (not shown). (b) Once in the node, T cells migrate to T cell-dependent areas, whereas B cells migrate to follicles. Antigen is taken up and processed by dendritic cells (not shown). (c) Lymphocytes that recognize the antigen stop recirculating, enlarge or blast and begin to proliferate. Masses of blasting T cell are seen filling the paracortex. Lymphocytes that do not recognize antigen continue migrating and leave the LN via the efferent lymphatic vessel. (d) After about a week, plasma cells appear in the medulla, and effector lymphocytes and antibodies leave the LN via the efferent lymphatic vessel.

surrounded by dense formations of small memory B cells. Interfollicular areas are populated by T cells and represent T-dependent regions of the LN. Follicles provide the microenvironment for the exposure of B cells to trapped antigens and for their proliferation. They are the anatomic site for somatic mutations. Although B cell differentiation to plasma cells occurs mostly in the deep paracortex and medullary cords, follicles are the site of B cell affinity maturation and memory development.

The medulla of a LN is a loose network of medullary cords and sinuses situated in the hilar area and separated loosely from the paracortex by a network of connective tissue fibers. Medullary cords contain both T and B cells and most of the plasma cells in the LN. B cells complete their maturation into antibody-producing plasma cells in the medulla, and lymph exiting via the efferent lymphatics is enriched in antibodies. In addition, scavenger macrophages line

the medullary sinuses and are responsible for the removal of particulate antigens found in the lymph.

Lymph vessels carrying interstitial fluid deliver antigens to the LN. The lymph enters from afferent lymph vessels into the subcapsular sinus and is channeled around and through trabecular sinuses, across the LN parenchyma toward a medullary network of sinuses. From there, it can leave the LN via efferent lymph vessels and travel to the next LN in the chain. The intranodal lymph channels are enclosed by FDCs, sinus-lining endothelial cells, and macrophages, which sample the lymph and transport antigenic material for presentation to B and T cells. Antigens in the lymph do not have free access to the lymphocyte compartment, and FDCs serve to select and transport lymph-borne molecules to be presented to T and B cells that home to LNs from the blood.

Overall, the architecture of a LN is characterized by a high degree of compartmentalization, which appears to be necessary in order to maintain the

constant and directed flow of lymphocytes from one compartment to another and to allow for cellular interactions to occur at the time and place best suited for optimal immune responses.

Lymphocyte Migration within the Lymph Node

Figure 3 is a diagrammatic representative of trafficking of lymphocytes in, within, and out of the LN. Lymphocyte migration is a continuous process. Naive lymphocytes enter the node from the bloodstream via specialized structures called HEVs in the paracortical cords. Adhesion molecules on the surface of lymphocytes are instrumental in their binding to and the subsequent transmigration through the high cuboidal endothelium. Antigens enter the LN through the afferent lymphatics and are trapped by specialized cells (dendritic cells [DCs]) for presentation to lymphocytes. Only the rare lymphocytes recognizing the antigen via the TCR or BCR are activated and stop recirculating. Lymphocytes that do not recognize antigen continue migrating through the LN and leave it via the efferent lymphatic vessel (**Figure 3c**).

Lymphocytes entering the paracortex cords from blood via HEV move outward through perivenular channels made up of concentric rings of FRCs. Lymphocytes spend 10–100 min moving through these channels, which are rich in extracellular matrix (ECM) components and provide an excellent substrate for lymphocyte migration. To reach the cord, lymphocytes must once again transmigrate, this time through tightly overlapping FRCs enclosing the perivascular channels like a sleeve. Once within the next compartment, called the corridor, the route of lymphocytes depends on whether they are T or B cells. T cells migrate to the T-dependent areas, whereas B cells migrate to the follicles. It is not known how the cells select their route or how their traffic is directed. Lymphocytes end their journey through the cord by migrating between the sinus-lining cells into the cortical sinus.

The migration of lymphocytes through the cortex proceeds radially outward, from the HEV lumen to the outer sinuses. It is remarkable that during this journey lymphocytes manage to traverse across three cellular boundaries: the HEV endothelium, the FRC boundary in perivenular channels, and the sinus-lining cell. In addition, lymphocytes migrate more or less up and down within the cortex. They do so by crawling along the corridor wall or sinus boundaries, both richly expressing the ECM components. This oriented migration within the cortex is crucial for lymphocyte interactions with antigen-presenting cells (APCs) and for the connection between the follicular

areas and the medulla. Also, the ability of differentiated effector T cells or plasma cells to leave the LNs via efferent lymphatics depends on the open routes of migration within the cord.

Antigen-Driven Interactions in the Lymph Node

The presence of an antigen dramatically alters the appearance and size of the LN. Upon entering the node environment through afferent lymphatics, antigens are taken up by sessile APCs referred to as interdigitating dendritic cells (IDCs) and found in the cortex. These cells migrate via lymph from tissues to the corridor, where they attach to FRCs, providing an optimal milieu for lymphocyte–APC interactions. Aside from enlarging the surface available for these interactions, IDCs are rich in costimulatory molecules and cytokines. Lymphocytes typically spend 3–6 h in the corridor space. Thus, migrating T lymphocytes have the opportunity to screen IDCs immobilized on the corridor walls for antigens (peptides) presented by the major histocompatibility complex (MHC). When the TCR on the T cell encounters the peptide it recognizes, activation of the T cell occurs. This process can occur anywhere in the corridor space, allowing migrating T cells an opportunity for screening multiple IDCs along the way. Thus, corridors of the paracortex are the ideal sites for interactions of IDCs, T cells, and B cells, where the generation of the immune response is facilitated by efficient cell-to-cell contact. They also provide a suitable environment for the proliferation of activated T cells.

When a T cell traversing the paracortex meets its cognate antigen, the TCR transmits activation signals to the nucleus, and T cell proliferation begins within the corridor. Simultaneously, B cells migrating through the paracortex or moving along the axis of the corridor are also likely to encounter antigen recognized by the Ig receptor and to initiate proliferation. In an environment rich in cytokines secreted by activated T cells, B cells undergo differentiation. The clonal expansion of antigen-specific T and B cells leads to their accumulation in the corridors and to marked distention of the cords. The size of a LN increases dramatically during an immune response. In addition to ballooning, paracortical cords undergo elongation in reactive LNs, and endothelial cells in HEV also proliferate. Thus, the entire size of the LN increases several-fold in a relatively short period of time. Conversely, when the depletion of lymphocytes occurs, e.g., during thoracic drainage, the diameter of the cord shrinks. Large differences in the size of the cortex due to changes in lymphocyte numbers are the basis for a common perception that reactive LNs are large and

nonreactive LNs are small. The enormous plasticity of LNs depends on their infrastructure of collagen fibers and FRCs, forming a loose reticular network amenable to shrinking and extension. The LN also has a network of nerve endings that connects it to the central nervous system (CNS). In an enlarged reactive LN, changes in the migration rate of lymphocytes occur: the traffic along the long axis of the cords becomes prominent, with effector T cells and differentiating B cells moving toward the medulla and other B cells, those destined to become memory cells, moving toward the follicles. This is possible because extended cords allow for easy access and enhanced lymph flow to the interfollicular area at the base of follicles. Similarly, medullary cords and sinus are also extended in reactive LNs and are able to accommodate more lymphocytes now migrating toward the efferent lymphatic vessel in the hilum. Before lymphocytes exit the LN, they spend a period of time in the medulla. Here, in the environment rich in cytokines and in the presence of helper T cells, B cell maturation into antibody-secreting plasma cells takes place. Antigen-specific effector T cells and plasma cells leave the LN through the efferent lymphatic vessel to enter the bloodstream and eventually to reach the site of infection or repopulate sites of antibody production, respectively.

Significance of Lymph Nodes

Because lymph-borne antigens enter the LN continuously and immune responses are generated to a variety of antigens at all times, immune cells leaving the LN contain effector cells primed to deal with a variety of different antigens. The LN environment is structured and organized to facilitate entry and migration of naive lymphocytes to areas of antigen presentation, to promote antigen recognition, to allow for activation and proliferation of antigen-specific T and B cells, and, finally, to assure the availability of immune effector cells to sites of infection. This series of events is accomplished in 4 to 6 days in primary immune responses and in a much shorter period of time in a secondary response. The presence in the follicular compartment of memory B cells surrounded by interfollicular T cells assures that anamnestic responses are initiated promptly at the site of antigen entry. Migration of activated T and B cells from the follicular area along the length of extended paracortical cords to medulla proceeds rapidly during secondary immune responses and may be facilitated by chemokines, by cytokines produced by activated lymphocytes or accessory cells, and perhaps by the directional flow of fluids in the corridors. The regulation of directed migration of lymphocytes during primary or secondary immune responses is a subject

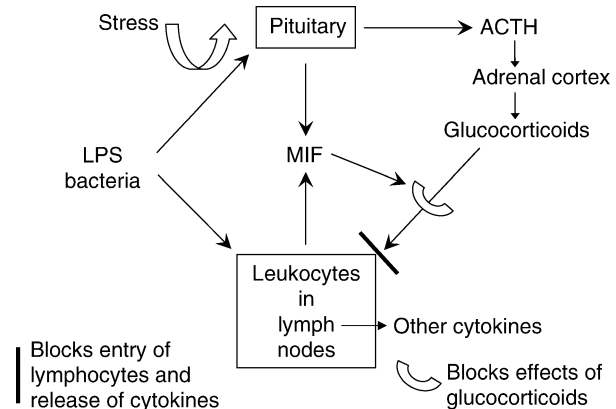


Figure 4 A schematic representation of stress-induced, glucocorticoid-mediated effects on lymphocyte entry into LN and on cytokine production by lymphocytes present in LN. ACTH, adrenocorticotropic hormone; MIF, migration inhibitory factor, LPS, lipopolysaccharide.

of intense investigation. To home to the LN, lymphocytes must execute a series of adhesion and signaling events that involve a variety of molecules. First, tethering and rolling of lymphocytes entering HEV of LNs is mediated by L-selectin (CD62L) and its ligands, such as MECA-79. Next, arrest of rolling lymphocytes is mediated by the integrin leukocyte function-associated antigen 1 (LFA1), interacting with ICAM1 and ICAM2 expressed on HEV. Activation of integrins on lymphocytes is mediated by the CC chemokine ligand 21 (CCL21, SLC, 6CKine) constitutively expressed by HEV cells. It binds to its receptor (CCR7) on lymphocytes. A second CCR7 ligand, CCL19 (ELC, MIP3 β) is expressed by the lymphatic endothelial and interstitial cell but not by HEV. Other integrin-activating signals involving chemokines and their receptors exist, and it is thought that T cells and B cells respond to different signals in HEV. Activation-dependent adhesion steps allow lymphocytes to stick firmly to HEV and to emigrate through the vessel wall and enter the LN environment. The LN represents a series of distinct micro environments, each adapted to optimize the function of immune cells. The key element is migration of these cells from one compartment to another. It is important to note that the migration of lymphocytes within the node can proceed in either of two directions, from HEV outward to the sinus in the paracortical cord or from the follicular area to the medullary sinus at the hilum. Naive lymphocytes entering the node via HEV take the first route; once activated and/or proliferating, lymphocytes tend to take the second path. The regulated movement of fluids and cells within the LN compartments is a wonderful example of the efficiency of the immune system in

executing primary and secondary antigen-specific responses.

Interactions between the Central Nervous System and Lymph Nodes

Recent advances in the field of neuroimmunology have shown that the CNS and the immune system are intimately linked. The CNS influences responses of the immune system through the release of humoral factors such as cortisol or epinephrine from the hypothalamic-pituitary-adrenocortical (HPA) axis. Glucocorticoids released as part of a systemic stress response influence (downregulate) immune responses profoundly. Communication between nervous and immune systems is also maintained by the efferent neuronal sympathetic nervous system in lymphoid organs. LNs are innervated so that a close association exists between autonomic nerve terminals and immune cells. Local synthesis and colocalization of norepinephrine, neuropeptides, such as neuropeptide Y, and opioids have been documented in the nerve terminals of nerve fibers located in LNs. Immune cells present in this microenvironment have specific surface receptors for neurotransmitters and neuropeptides and are responsive to signals transmitted via the receptors. In addition to efferent nerve terminals, LNs contain afferent nerve fibers characterized by the expression of receptors for cytokines or other factors produced by immune cells. Thus, bidirectional traffic of biochemical mediators between the CNS and immune cells regulates their interactions.

An example of chemically mediated signal transmission in response to stress is shown in [Figure 4](#). In this schema, stress or infectious agents induce the release of glucocorticoids from the HPA axis. Glucocorticoids have potent anti-inflammatory and immunosuppressive effects, including the downregulation of lymphocyte functions and of their migration into the LN. In the presence of cortisol, immune cells begin to produce a cytokine, migration inhibitory factor (MIF), which counteracts inhibitory effects of glucocorticoids at the local level. The pituitary also produces and releases MIF into the systemic circulation. MIF suppresses glucocorticoid production and restores the immune balance.

Trafficking of lymphocytes is fundamental to the generation of immune responses. Glucocorticoids inhibit lymphocyte migration into LNs and disrupt lymphocyte recirculation.

Lymphopenia observed in the presence of steroids is not due to death of lymphocytes but to redistribution of the circulating lymphocyte pool. Lymphocytes that normally enter LNs now migrate away from

them, with the net result being the downregulation of lymphocyte entry into the LNs. The molecular mechanism responsible for this effect is a glucocorticoid-induced decrease in expression of L-selectin (CD62L), also known as a lymph node homing receptor, and of CD44 Hermes protein on the surface of lymphocytes. Because these adhesion molecules are involved in lymphocyte-HEV interactions and are necessary for lymphocyte entry into LNs, dysregulation of their expression naturally interferes with homing. It is probable that these interactions are not direct but are mediated via cytokines, whose local production is inhibited by glucocorticoids (see [Figure 4](#)). The profound effects of stress on the immune system are partly mediated via neurotransmission in LNs. Thus, the understanding of LN architecture and of lymphocyte migration patterns in LNs is necessary for the appreciation of stress-induced alterations in immune responses. The various aspects of the CNS-immune system communication and, especially, the role of cytokines in orchestrating these interactions are being intensely investigated today.

See Also the Following Articles

Cytokines; Immune Cell Distribution, Effects of Stress on; Immune Response; Lymphocytes.

Further Reading

- Anderson, A. O. and Shaw, S. (1995). Lymphocyte trafficking. In: Rich, R. R., Fleischer, T. A., Schwanz, B. D., Shearer, W. T. & Strober, W. (eds.) *Clinical immunology*, pp. 39–49. St. Louis, MO: Mosby.
- Calandra, T. and Bucala, R. (1997). Macrophage migration inhibition factors (MIF): a glucocorticoid counter-regulator within the immune system. *Critical Reviews in Immunology* 17, 77–88.
- Cyster, J. G. (1999). Chemokines and cell migration in secondary lymphoid organs. *Science* 286, 2098–2102.
- Eskandari, F. and Sternberg, E. M. (2002). Neural-immune interactions in health and disease. *Annals of the New York Academy of Science* 966, 20–27.
- Friedman, E. M. and Lawrence, D. A. (2002). Environmental stress mediates changes in neuroimmunological interactions. *Toxicological Science* 67, 4–10.
- Gretz, J. E., Anderson, A. O. and Shaw, S. (1997). Cords, channels, corridors and conduits: Critical architectural elements facilitating cell interactions in the lymph node cortex. *Immunology Reviews* 156, 11–24.
- Kelly, R. H. (1975). Functional anatomy of lymph nodes. I. The paracortical cords. *International Archives of Allergy and Applied Immunology* 48, 836–849.
- Straub, R. H., Westermann, J., Scholmerich, J. and Falk, W. (1998). Dialogue between the CNS and the immune system in lymphoid organs. *Immunology Today* 19, 409–413.

See Also the Following Articles

Cytokines; Immune Cell Distribution, Effects of Stress on; Immune Response; Lymphocytes.

Further Reading

- Anderson, A. O. and Shaw, S. (1995). Lymphocyte trafficking. In: Rich, R. R., Fleischer, T. A., Schwanz, B. D., Shearer, W. T. & Strober, W. (eds.) *Clinical immunology*, pp. 39–49. St. Louis, MO: Mosby.
- Calandra, T. and Bucala, R. (1997). Macrophage migration inhibition factors (MIF): a glucocorticoid counter-regulator within the immune system. *Critical Reviews in Immunology* 17, 77–88.
- Cyster, J. G. (1999). Chemokines and cell migration in secondary lymphoid organs. *Science* 286, 2098–2102.
- Eskandari, F. and Sternberg, E. M. (2002). Neural-immune interactions in health and disease. *Annals of the New York Academy of Science* 966, 20–27.
- Friedman, E. M. and Lawrence, D. A. (2002). Environmental stress mediates changes in neuroimmunological interactions. *Toxicological Science* 67, 4–10.
- Gretz, J. E., Anderson, A. O. and Shaw, S. (1997). Cords, channels, corridors and conduits: Critical architectural elements facilitating cell interactions in the lymph node cortex. *Immunology Reviews* 156, 11–24.
- Kelly, R. H. (1975). Functional anatomy of lymph nodes. I. The paracortical cords. *International Archives of Allergy and Applied Immunology* 48, 836–849.
- Straub, R. H., Westermann, J., Scholmerich, J. and Falk, W. (1998). Dialogue between the CNS and the immune system in lymphoid organs. *Immunology Today* 19, 409–413.
- Von Andrian, U. H. and Mempel, T. R. (2003). Homing and cellular traffic in lymph nodes. *Nature Reviews Immunology* 3, 867–878.
- Warnock, R. A., Askari, S., Butcher, E. C. and von Andrian, U. H. (1998). Molecular mechanisms of lymphocyte homing to peripheral lymph nodes. *Journal of Experimental Medicine* 187, 205–216.
- Wrona, D. (2005). Neural-immune interactions: an integrative view of the bidirectional relationship between the brain and immune systems. *Journal of Neuroimmunology* 172, 38–58.

Lymphocyte Trafficking and Cell Adhesion Molecules (CAM) See: Leukocyte Trafficking and Stress; Lymph Nodes.

Lymphocytes

N R Rose

Johns Hopkins University School of Medicine,
Baltimore, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition
article by N R Rose, volume 2, pp 641–645,
© 2000, Elsevier Inc.

Introduction

Origin and Development

B Lymphocytes

T Lymphocytes

Natural Killer (NK) Cells

Glossary

Anamnestic response An accelerated secondary immune response.

B cell Large, metabolically active B cells.

lymphoblasts

Class I major histocompatibility complex (MHC)

The MHC expressed on all nucleated cells.

Class II major histocompatibility complex (MHC)

The MHC expressed mainly on antigen-presenting cells.

Cytokines

A diverse group of multifunctional proteins that play a key role in a variety of biological processes, including cell growth and activation, inflammation, immunity, hematopoiesis, tumorigenesis, tissue repair, fibrosis, and morphogenesis. Chemokines are a subgroup involved in chemotaxis. Cytokines are synthesized in a variety of cell types, including lymphocytes and cells of the mononuclear phagocyte system.

Endocytosis

Ingestion of molecules by the cell.

Epitope

Antigenic determinant.

Erythropoietin

The growth factor required for the generation of erythrocytes from the pluripotential stem cell.

Granulocyte/monocyte-colony stimulating factor (GM-CSF)

The growth factor required for the generation of granulocytes and monocytes from the pluripotential cell.

Immuno-globulin
Stem cell

The antibody containing fraction of serum.

A self-renewing undifferentiated cell capable of developing into specialized cells.

Introduction

In vertebrates, the specialized function of acquired or adaptive immunity is assigned to the immune system. Like other specialized organs, the immune system comprises a number of differentiated cells that contribute to its special function; unlike most other organs, however, the immune system is distributed throughout the body (see **Immune Response**). This decentralized anatomical structure best serves its purpose of providing immunological surveillance and protection. The distinguishing features of the acquired immune response are its specificity (the ability to distinguish one antigenic molecule from another) and its memory. The immune system must recognize and respond to the vast universe of antigenic determinants and must learn to respond more vigorously at a second encounter. The cells specialized to serve these critical functions are the lymphocytes. This article focuses on the major population of lymphocytes: B lymphocytes, T lymphocytes, and natural killer (NK) cells.

Origin and Development

Figure 1 diagrams the early development of the hematopoietic system, of which the immune apparatus is a component. All of the blood cells develop from a single multipotential or pluripotential stem cell. Phylogenetically this cell arises in the yolk sac. In mammals, the pluripotential stem cell transfers to the liver during embryonic development, but later, during gestation and birth, it migrates to the bone marrow. Like all stem cells, it is capable of self-renewal but can differentiate into a number of specialized progenitor or precursor cells. These include the erythrocyte progenitor, which eventually develops into the erythrocytes; the megakaryocyte progenitor, which is the eventual source of platelets; and the myeloid progenitor of the granulocytes and monocytes, which is the source of all of the granulocytic leukocytes, such as polymorphonuclear neutrophils, eosinophils, and basophils and of the monocytes of the blood and the macrophages and dendritic cells of the tissues. Finally, the lymphocyte progenitor is the source of B lymphocytes, T lymphocytes, and NK cells. The production of these progenitor cells, as well as their continued differentiation into the mature functional cells of the lymphopoietic system, is driven by a family of peptide messengers, referred to collectively as colony-stimulating factors. At least two of them, erythropoietin, which promotes the production of erythrocytes, and granulocyte/monocyte-colony stimulating factor (GM-CSF), which fosters production

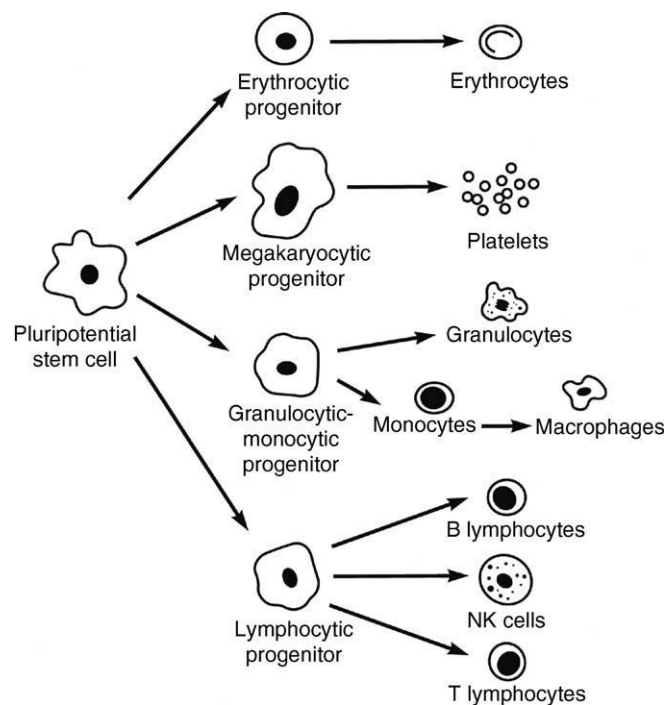


Figure 1 Development of the lympho-hematopoietic system.

of the granulocyte/monocyte lineage, have been isolated and used clinically to treat deficiencies of the respective cell populations.

B Lymphocytes

As indicated previously, B lymphocytes arise from the pluripotential stem cell in the fetal liver and later in the bone marrow. Directed by several stimulating factors, it matures into a functional B lymphocyte. These cells comprise approximately half of the approximately 10^{12} lymphocytes in the human body. They represent about 20% of the circulating lymphocytes but a majority of those found in the organized lymphoid tissues, such as the spleen and lymph nodes. The critical event in B cell maturation is the production of an antigen-specific receptor and its expression on the cell surface.

The B cell receptor (BCR) is a membrane-bound form of the immunoglobulin (Ig) molecule. The clonal selection theory of the immune response requires that there be an enormous diversity of these cell surface receptors, enough to bind any antigenic determinant or epitope that the B cell may encounter. The BCR is a heterodimeric molecule consisting of two heavy and two light chains, both of which have variable regions that contribute to the epitope-specific binding site. The required diversity of the receptors is produced by a specialized process of recombination, wherein one of a large set of variable (V) genes combines with one of a smaller number of diversity (D) genes (in the immunoglobulin heavy chain) and joining (J) genes to produce the variable regions of both the heavy and light chains. The pairing of a heavy and light immunoglobulin chain in a random manner further expands the diversity of BCR on different B cells. Another limited set of genes encodes the constant (C) region of the immunoglobulin molecule. Although initially apart on the chromosome, the V(D)J complex is joined with a C region gene to produce the immunoglobulin heavy chain or light chain. The assembly of immunoglobulin genes depends upon two recombination activator genes, RAG-1 and RAG-2. After completion of the immunoglobulin gene arrangement, the BCR is expressed on the cell surface with a μ heavy chain, representative of the IgM class of antibody. As the cell continues its maturation, a second class of immunoglobulins, IgD, composed of the same V(D)J segment but a δ constant region gene, is also expressed at the B cell surface. Finally, the mature B cell completes recombination by secreting immunoglobulin products with the δ constant region (i.e., IgG antibodies), the α constant region (IgA antibodies), or the ϵ constant region (IgE) antibodies. Circulatory antibodies of the IgG class are

mainly responsible for acquired protective immunity; IgA antibodies produced at mucosal surfaces serve immune functions in secretions; IgE antibodies are associated with allergic and hypersensitivity reactions.

A small population of B cells, termed B-1 cells, develops early during ontogeny and expresses the CD5 cell surface marker. The cells secrete low- to moderate-affinity IgM antibodies that are often broadly cross-reactive, including reactions to common pathogens and self-antigens. The antibodies are referred to as natural antibodies because their production does not appear to require the presence of a specific antigen. They may provide some initial protection against infectious agents.

The mature B cell is activated by binding an antigen that expresses the complementary (cognate) epitope recognized by its BCR. The activation process requires cross-linking of membrane BCR. The receptor is associated with a number of auxiliary molecules such as CD19, CD20, and CD21, which act in concert with the receptor and serve as distinguishing markers of B cells. These molecules transmit activation signals to the interior of the cell after receptor aggregation takes place on the cell surface.

This event sets in motion a series of metabolic steps, manifest mainly by the activation of tyrosine kinases, leading to tyrosine phosphorylation of a set of substrates that control intracellular signals. Early events of activation lead to development from a resting mature B lymphocyte into a metabolically active B lymphoblast with enlarged cytoplasm containing multiple mitochondria. The B cell blasts then divide repeatedly to produce a clone, each member of which bears the parental BCR and, therefore, expresses the same epitope specificity.

Upon binding to the BCR, antigen is endocytosed and processed for presentation to specialized T cells (see below). In addition to presenting antigens, B cells secrete a number of cytokines such as IL-10, IL-12, tumor necrosis factor- α (TNF- α), and lymphotoxin, which subsequently act to direct or regulate immune responses (see **Cytokines**).

Some of the activated B cells differentiate into plasma cells, which are the source of circulating antibody, whereas others differentiate into memory cells, which give rise to memory response upon rechallenge of the individual with the same antigen. The hallmarks of the immune response are both its specificity for the antigen that induced its formation and the enhanced immune response or anamnestic response on rechallenge.

Since a single B cell bears multiple BCRs with identical specificities, cross-linkage of the receptors requires that an antigenic molecule express more than one identical epitope. A number of microbial

carbohydrates fulfill this requirement, and so B cells can be stimulated directly by the capsular polysaccharides of such microorganisms as pneumococci and meningococci. It is much less likely, however, that a globular protein will express more than one copy of an epitope. Under these circumstances, B cell activation requires help from a T cell capable of recognizing the same antigen. The endocytosed antigen is fragmented within the B cell, and the peptide fragments are loaded into the groove of a specialized cell surface glycoprotein, the class II major histocompatibility complex (MHC) molecule. The resultant class II/peptide fragment complexes are then inserted into the B cell surface. These complexes can be recognized by the antigen-specific receptors of the subset of T cells, referred to as CD4⁺ or helper T cells.

The two-way interaction between T cells and B cells is critical for the development of a vigorous, adaptive immune response. On the one hand, B cells concentrate antigens and effectively present them to T cells that recognize the particular antigen. On the other hand, activated CD4⁺ T helper cells produce cytokines that foster the activation and differentiation of B cells into plasma cells. When the T cell interacts with the MHC II/peptide fragment, it induces division of the B cell through the agency of costimulating surface molecules expressed by the T cell. CD154 of the T cell, for instance, engages CD40 of the B cell. In addition, T cell-derived cytokines, such as IL-4, IL-5, and IL-6, direct the B cell into an antibody-forming plasma cell. As the B cell progresses into the immunoglobulin-producing process, it undergoes a class switch; that is, the heavy chain C region, which originally expressed the specificities of the IgM and IgD classes, differentiates into cells that produce IgG, IgA, or IgE. This process allows the production of antibodies with differing biological functions while retaining the same antigen-combining site. At the same time, there is an increased affinity of the antibodies resulting from the extraordinarily high level of somatic mutation of the V genes, providing antibodies of high efficiency in combining with their antigenic targets.

T Lymphocytes

T lymphocytes constitute the second major class of lymphocytes, comprising approximately 70% of lymphocytes in the peripheral circulation and essentially all of the lymphocytes in the lymphatic circulation. They also derive from the hematopoietic stem cell through the lymphocyte progenitor, but undergo a process of differentiation in the thymus before they are seeded to peripheral lymphoid tissues and blood (see **Autotolerance**), (see **Thymus**). Like B cells,

mature T cells have a great diversity of antigen-specific cell surface receptors. T cell receptors (TCRs) are heterodimers made up of α and β chains or, rarely, γ and δ chains. Knowledge of the function of the latter subset of T cells is still limited.

Unlike the BCR, the TCR remains at the cell surface and is not shed into the medium. It is associated with a set of transmembrane proteins, referred to as the CD3 complex, which play a critical role in signal transduction. Through their agency, the T cell undergoes a process of activation similar to that described for the B cell. Like the BCR, the TCR is organized into a variable and constant portion. Furthermore, the same recombinational mechanisms are involved in the assembly of the V region genes in T cells and B cells. The process is controlled by RAG-1 and RAG-2, which together are required for T cell maturation. Antigen primed T cells may differentiate directly into effector or regulatory T cells. Following clonal expansion, a proportion of the antigen-stimulated T cells differentiates into long-lived memory cells. Memory T cells are responsible for the ability of the immune system to mount an anamnestic response when they encounter the same antigen. Thus, memory T cells play a central role in maintaining long-term protection against infection. In humans, memory T cells are usually CD45 RA⁻ CD45 RO⁺. Unlike effector T cells, they are not susceptible to programmed cell death following activation. Consequently, they survive for long periods of time in spleen and lymph nodes without undergoing division. Following a second encounter with their cognate antigen, memory T cells multiply rapidly and produce large amounts of cytokines so that even low levels of antigen can produce a vigorous secondary response. Memory B cells can be divided into central memory cells that inhabit secondary lymphoid organs and effector memory cells that migrate to inflamed areas.

T cells exert both regulatory and effector functions. The major regulatory T cells are the CD4⁺ helper T cells, referred to previously. Their regulatory function depends on costimulatory molecules, such as CD154, as well as the production of particular cytokines. Some CD4⁺ T cells develop into cells that primarily secrete the cytokines IL-4, IL-5, IL-6, and IL-13 (referred to as T_h2 cells) or into T cells that produce mainly IL-2, interferon- γ (IFN- γ), and lymphotoxin (referred to as T_h1 cells). The T_h2 cells help B cells to develop into antibody-producing cells, whereas T_h1 cells are primarily inducers of cell-mediated responses. Cell-mediated immunity is responsible for protection against intracellular microorganisms and involves the activation of macrophages as well as the generation of cytotoxic T cells. Cytotoxic T cells mediate the direct killer functions

of cell-mediated immunity; they carry the CD8 marker in place of CD4 and are capable of lysing target cells that express the requisite antigen together with matched class I MHC. Virus-infected tissue cells, for example, may be eliminated from the host's cells by virus-specific CD8 T cells.

Under normal circumstances, the immune system is kept under strict control by a variety of regulatory mechanisms. Regulatory or suppressor T cells were recognized a number of years ago, but their identity remained elusive. Recently, several populations of T cells with regulatory capability have been identified. They include naturally occurring T regulatory cells (Tregs), which were first identified by their ability to prevent spontaneous autoimmune disease. They are characterized by the cell surface marker CD4⁺CD25⁺ and the transcription factor FoxP3. The critical role of the FoxP3 factor is attested by the demonstration that CD4⁺CD25⁺ T cells do not develop a regulatory function in FoxP3-deficient mice. In addition, retroviral transduction of FoxP3 into CD4⁺CD25⁻ cells converts them into functional regulatory T cells. CD4⁺CD25⁺ Tregs are anergic *in vitro*, failing to proliferate following nonspecific or antigen-specific stimulation. They can, however, mount an antigen-dependent proliferative response *in vivo*. The T cell growth factor IL-2 appears to play a critical role in their development, since mice genetically deficient in IL-2 have reduced numbers of Tregs. The distinguishing feature of Tregs is their ability to suppress a broad range of innate and adaptive immune responses; the suppression requires direct contact between the regulatory T cell and its target.

In addition to naturally occurring Tregs, naive T cells can be induced to differentiate into regulatory T cells following an encounter with their cognate antigen. These cells differ from the natural CD4⁺CD25⁺ Tregs in that they do not express FoxP3 and act primarily by the production of immunosuppressive cytokines such as IL-10 and transforming growth factor- β (TGF- β). They are referred to as Tr1 cells. These induced regulatory T cells are favored by antigen presentation through immature dendritic cells or by environments in which high levels of suppressive cytokines are present, such as at mucosal surfaces. Because induced Tr1 cells secrete these immunosuppressive cytokines, they can downregulate responses to other nearby antigens, a process called bystander suppression.

Recent studies have provided evidence of a number of other regulatory T cell populations. For example, CD8⁺ T cells can mediate immune regulatory functions including allograft rejection. Although their

mechanism of action is still unclear, they may induce suppression by cytolysis of helper CD4⁺T cells.

NKT cells represent a distinct lineage of T cells that are characterized by an invariant TCR light chain as well as by the presence of characteristic NK cell markers. This invariant TCR does not recognize peptide presented in the usual fashion by class I or class II MHC molecules, but instead recognizes lipids and glycolipids presented by the nonclassical CD1 molecule. It appears that NKT cells are capable of both upregulating and downregulating immune responses.

Finally, a population of regulatory T cells bearing the CD3⁺ marker but lacking CD4⁺ or CD8⁺ has been described. The cells are referred to as double-negative T cells, but their mode of action is still unknown.

T cell recognition differs from that of B cells. The BCRs recognize free antigens based on their complementary conformations. In contrast, T cells respond to antigens only on the surface of other cells, because they recognize the complex of a peptide fragment bound in the groove of a class II or class I MHC glycoprotein. CD4⁺ T cells recognize peptides in class II MHC complexes, whereas CD8⁺ T cells recognize peptides in class I complexes. In the case of class II recognition, the MHC/peptide fragment complex is generated within an antigen-presenting cell (APC). They may be B cells, as described previously, macrophages, interdigitating dendritic cells of the lymph nodes, circulating dendritic cells, or Langerhans cells of the skin (see **Lymph Nodes**). In all of these cells, endocytosed proteins are fragmented by proteolytic enzymes within the endosomal/lysosomal compartment, and the resulting short peptides are loaded into the class II MHC groove. The complexes are then expressed at the surface of the cell, where they can be recognized by those CD4⁺ T cells that bear the requisite TCR. The recognition is based not on the conformation of the antigen, but rather on key amino acids.

CD8⁺ T cells recognize antigens and peptides in conjunction with class I MHC molecules. Generally, these antigens are derived from internally synthesized proteins, including viral proteins (see **Herpesviruses**). The fragments are produced in cytoplasmic vesicles by proteolysis, translocated into the endoplasmic reticulum, and bound to the class I MHC molecules, which are then brought to the cell surface. There they can be recognized by CD8⁺ T cells expressing the appropriate receptors. Generally, the antigenic epitopes loaded into class I MHC are approximately nine amino acids in length.

Two signals, specific and nonspecific, are required for T cell activation. Specificity is provided by the

interaction of the TCR with an MHC/peptide fragment complex. A second, non-antigen-specific costimulating signal is given by membrane glycoprotein of the APC B7-1 (CD80) or B7-2 (CD86). Through their interaction with CD28 on resting T cells, these molecules trigger activation and clonal expansion of the T cell. Engagement of the TCR in the absence of B7 costimulation results in clonal anergy. Another receptor expressed on activated T cells, CTLA-4, binds B7 and may serve as a regulatory signal.

A number of other costimulatory receptors have recently been described, including inducible costimulator intracellular (ICOS) and programmed death-1 (PD1), which regulate the immune response.

Natural Killer (NK) Cells

NK cells are an important component of the innate as well as the adaptive immune system. In fact, in animals that lack the RAG-1 or RAG-2 genes, NK cells provide a substantial degree of protection against infection. They comprise about 10% of the circulating lymphocytes. Although closely related to T cells, they lack the conventional TCRs, but express two specialized receptors. One receptor allows them to recognize virally infected cells or tumor cells, permitting them to destroy such cells. They also express receptors for the host's class I MHC molecules, which inhibit lytic activity. Thus, the NK cell is disarmed in the presence of its corresponding MHC molecule. Often, tumor cells or virally infected cells diminish expression of class I MHC molecules, thus avoiding the attack of cytotoxic T cells. They may then become targets for killing by NK cells, which are no longer shut off by recognition of the host MHC class I molecules. In addition, some NK cells express receptors for the constant portion of the immunoglobulin molecule. Armed with an IgG, the NK cell then has acquired a specific receptor to direct its cytotoxic effects, lysing target cells that bear the appropriate antigen. This process is referred to as antibody-dependent, cell-mediated cytotoxicity (ADCC). Thus, the NK can mediate its cytotoxic actions through two different recognition mechanisms. NK cells also produce cytokines that regulate the development of acquired immunity, for example, IFN- γ , and stimulate macrophages to produce cytokines, such as TNF and

IL-12, which enhance the development of cell-mediated immune responses.

See Also the Following Articles

Autotolerance; Cytokines; Herpesviruses; Immune Response; Lymph Nodes; Thymus.

Further Reading

- Bendelac, A., Rivera, M. N., Park, S. H. and Roark, J. H. (1997). Mouse CD1-specific NK1 T cells: development, specificity and function. *Annual Review of Immunology* **15**, 535–535.
- Bentley, G. A. and Mariuzza, R. A. (1996). The structure of the T cell antigen receptor. *Annual Review of Immunology* **14**, 563.
- Jonuleit, H. and Schmitt, E. (2003). The regulatory T cell family: distinct subsets and their interrelations. *Journal of Immunology* **171**, 6323–6327.
- Kisielow, P. and von Boehmer, H. (1995). Development and selection of T cells: facts and puzzles. *Advances in Immunology* **58**, 87–209.
- Kronenberg, M. and Rudensky, A. (2005). Regulation of immunity by self-reactive T cells. *Nature Insight* **435**, 598–603.
- Lanier, L. L., Corliss, B. and Phillips, J. H. (1997). Arousal and inhibition of human NK cells. *Immunological Reviews* **155**, 145–154.
- Ohashi, P. S. (2002). T-cell signaling and autoimmunity: molecular mechanisms of disease. *Nature Reviews Immunology* **2**, 427–438.
- O'Shea, J. J., Ma, A. and Lipsky, P. (2001). Cytokines and autoimmunity. *Nature Reviews Immunology* **21**, 37–45.
- Pitkänen, J. and Peterson, P. (2003). Autoimmune regulator: from loss of function to autoimmunity. *Genes and Immunity* **4**, 12–21.
- Reth, M. (1992). Antigen receptors on B lymphocytes. *Annual Review of Immunology* **10**, 97–121.
- Sakaguchi, S. (2004). Naturally arising CD4⁺ regulatory T cells for immunologic self-tolerance and negative control of immune responses. *Annual Review of Immunology* **22**, 531–562.
- Salomon, B. and Bluestone, J. A. (2001). Complexities of CD28/B7: CTLA-4 costimulatory pathways in autoimmunity and transplantation. *Annual Review of Immunology* **19**, 225–252.
- Shevach, E. M. (2002). CD4⁺ CD25⁺ suppressor T cells: more questions than answers. *Nature Reviews Immunology* **21**, 389–400.

M

Macrophage Antimycobacterial Activity, Effects of Stress on

C S Boomershine and B S Zwillig

The Ohio State University, Columbus, OH, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 657–662, © 2000, Elsevier Inc.

Introduction

Stress Influences Mycobacterial Infection

Summary

Glossary

Nramp (*natural resistance associated macrophage protein*) A phagolysosomal membrane product encoded by the *Nramp1* gene that confers resistance to mycobacterial growth during the early nonimmune phase of infection. Susceptible mice differ from resistant mice in a nonconservative glycine-to-aspartic acid substitution within a predicted transmembrane domain of Nramp1.

Introduction

Tuberculosis, a worldwide public health problem, has also increased significantly in the United States. The increase in tuberculosis is partly due to depressed cellular immunity associated with HIV infection. However, tuberculosis has also increased independently of HIV infection. Only since 1997, following mobilization of the public health community, has the incidence of tuberculosis in the United States begun to resume its downward course. Tuberculosis is a highly complex disease with multiple determinants for disease progression and severity. The susceptibility of humans to mycobacterial disease is determined partly by genetic

differences. One of the most important genetic factors determining resistance to mycobacterial infection is the ability of host macrophages to control the growth of the tubercle bacilli.

Innate Resistance against Mycobacteria

The innate resistance of macrophages to mycobacterial infection has been studied most thoroughly in mice. The ability of mice to control mycobacterial growth in the spleen, following intravenous injection of *Mycobacterium bovis* BCG, maps to a region on mouse chromosome 1. Mice resistant to mycobacterial growth are termed Bcg^r, whereas those strains in which mycobacteria grow exponentially are termed Bcg^s. The resistance of Bcg^r mice is mediated by the *Nramp1* gene, which has been cloned in both mice and humans. *Nramp1* codes for a phagolysosomal membrane product termed natural resistance associated macrophage protein. Susceptible mice differ from resistant mice in a nonconservative glycine-to-aspartic acid substitution within a predicted transmembrane domain of Nramp1. While the mechanism of Nramp-mediated resistance remains ill defined, work from the authors' laboratory suggests that Nramp1 acts as a transporter of iron into the macrophage phagolysosome.

Although the susceptibility of an individual to mycobacterial disease is influenced strongly by innate immunity, genetic differences alone cannot account for the recent rise in the incidence of tuberculosis. The vast majority (90–95%) of people exposed to pathogenic strains of mycobacteria are able to control the growth of the organism without any disease manifestations. While some macrophages are able to limit mycobacterial growth, they are not sufficient to resolve a primary mycobacterial infection. Fortunately, macrophages are the major antigen-presenting cells of the body. Enzymes in the macrophage phagolysosome digest microorganisms killed by the respiratory burst and some microbial peptides become associated

with glycoproteins encoded by genes of the major histocompatibility complex (MHC). Peptides coupled to MHC class II glycoproteins are cycled to the cell surface where they serve as the sole basis for antigen recognition by CD4⁺ T lymphocytes. These macrophages then travel to regional lymph nodes where they present this processed antigen to T cells.

Specific Immune Response to Mycobacterial Infection

T-cell activation results in the development of cell-mediated immunity and enhanced reactivity to the purified protein derivatives (PPD) of mycobacteria. Activated Th cells then travel to the area of the primary focus and release lymphokines that cause monocytes to migrate into the site and differentiate into specialized histiocytic cells termed epithelioid giant cells. The primary lesion becomes organized into a granuloma, with macrophages containing ingested mycobacteria located centrally surrounded by epithelioid giant cells and activated T cells. The production of interferon (IFN)- γ by activated T cells results in macrophage activation and enhanced antimicrobial activity resulting from the activation of IFN- γ responsive genes. Mycobacteria may persist within macrophages of the granuloma for decades, but in immunocompetent individuals, further multiplication and spread are usually confined. The relative strength of the cell-mediated immunity of the host ultimately determines if resumption in multiplication of the mycobacteria will occur. The consequence of a resumption in the multiplication of bacilli leads to tuberculosis with subsequent tissue destruction, systemic dissemination of disease, and, when left untreated, death. Reactivation of disease has been attributed to several environmental factors, including protein malnutrition associated with chronic alcoholism, immunosenescence associated with aging, and stress.

Stress Influences Mycobacterial Infection

One of the first observations that stressful life events can affect disease pathogenesis was made in studies of tuberculosis by Ishigami. Ishigami found decreased phagocytic cell activity during periods of emotional stress in chronically ill, tuberculous school children. He postulated that the stressful school environment led to the children's immunodepressed state and their increased susceptibility to tuberculosis. Contemporaneous studies have proven that physical and psychological stress can modulate susceptibility to tuberculosis as well as general immune function. Physiologic responses to stressors are mediated by

both the hypothalamic-pituitary-adrenal (HPA) axis and the sympathetic nervous system (SNS).

HPA Axis Effects on Mycobacterial Infection

The HPA axis mediates the stress response by producing a cascade of hormones in response to a variety of stimuli, ultimately resulting in the enhanced production of glucocorticoids by the adrenal glands. Glucocorticoids then enter the circulation and exert their effects through binding to cytoplasmic receptors. Upon binding, the glucocorticoid-receptor complex is internalized and translocated to the nucleus where it binds to glucocorticoid response elements (GREs). The receptor-charged GREs enhance the transcription initiation of glucocorticoid-regulated genes influencing the amount of critical proteins within the cell. Glucocorticoids have been shown to suppress the immune response. Glucocorticoid hormones cause a species- and cell type-specific lysis of lymphocytes. They suppress the inflammatory response by decreasing the number of circulating leukocytes and the migration of tissue leukocytes, inhibiting fibroblast proliferation, and inducing lipocortins that blunt the production of prostaglandins and leukotrienes, which are potent anti-inflammatory molecules.

Mononuclear cells possess high-affinity (type II) glucocorticoid receptors and represent one of the best studied primary target cells for glucocorticoids. Studies have shown that the production of proinflammatory cytokines interleukin (IL)-1, IL-6, and tumor necrosis factor (TNF)- α by activated macrophages is blocked at the transcriptional and posttranscriptional level by glucocorticoids. The downregulation of these potent mediators of inflammation underlies some of the immunosuppressive and anti-inflammatory actions of glucocorticoids, but does not address the effect glucocorticoids have on macrophage-T-cell interactions. Due to the requirement for macrophage MHC class II expression in antigen presentation for the induction of an immune response, the authors began their work by evaluating the effect of restraint stress on the expression of MHC class II glycoprotein expression in murine macrophages.

MHC class II glycoprotein expression is induced in mice following the injection of antigen or microorganisms *in vivo* or by treatment of macrophages with recombinant IFN- γ *in vitro*. It was found that macrophages from resistant strains of mice are able to persistently express MHC class II following infection of the mice with *M. bovis* (strain BCG). It was shown subsequently that the treatment of macrophages from Bcg^r mice with rIFN- γ induced persistent MHC class II expression, whereas macrophages from

Bcg^s mice expressed this glycoprotein only transiently. The regulation of MHC class II expression by macrophages from resistant and susceptible mice also differed. Glucocorticoid treatment of macrophages from susceptible mice resulted in decreased MHC class II expression, whereas MHC class II expression by macrophages from resistant mice was unaffected.

The regulation of MHC class II expression by glucocorticoids raised the possibility that stress could modulate class II expression following activation of the HPA axis. This hypothesis was confirmed by demonstrating that restraint stress suppressed MHC class II expression by murine peritoneal macrophages and that this suppression coincided with peak corticosterone levels in plasma. The authors' previous work that showed strain-specific differences in MHC class II glycoprotein expression was extended by showing that the stress-induced suppression of MHC class II expression was limited to certain strains of mice. Bcg^r mice were resistant to the restraint stress-induced decrease in MHC class II glycoprotein expression observed in Bcg^s mice. The resistance of Bcg^r mice was not due to a decrease in HPA axis activation, as corticosterone suppressed the induction of MHC class II by macrophages from both resistant and susceptible mice.

Adrenalectomy ameliorated the suppressive effects of restraint stress on MHC class II expression in Bcg^s mice by abrogating the increase in corticosterone that results from HPA axis activation. Later experiments showed that the difference in MHC class II expression by macrophages from resistant and susceptible mice was due to a difference in MHC class II mRNA stability (Figure 1). Corticosterone decreased the stability of MHC class II mRNA in macrophages from susceptible mice but not in macrophages from resistant mice. The stability of mRNA of other IFN- γ -induced genes was affected similarly. Thus, the increased mRNA stability of macrophages from Bcg^r mice may account for the pleiotrophic effects that have been attributed to *Nramp1*.

The differential regulation of MHC class II glycoprotein expression by glucocorticoids could account for differences in tuberculosis susceptibility. The role of HPA axis activation in the control of the growth of mycobacteria in humans has been the subject of some discussion. Glucocorticoid treatment of human macrophages suppressed antimicrobial activity markedly. However, Rook and co-workers showed that dexamethasone increased the susceptibility of monocytes from some individuals to mycobacterial growth, but other donor monocytes were unaffected by glucocorticoid treatment.

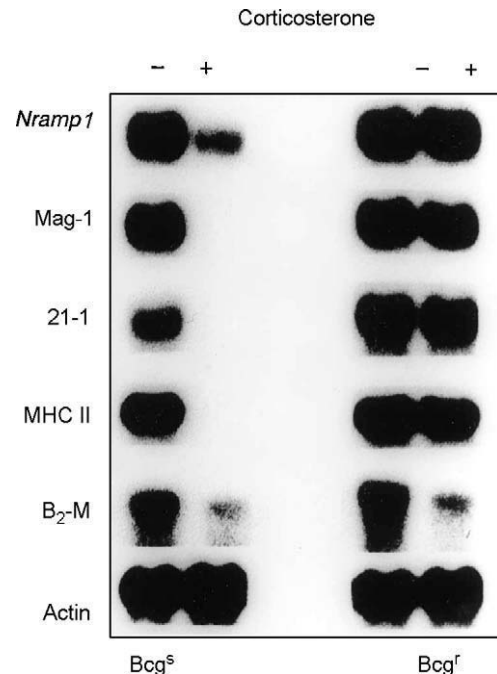


Figure 1 Stabilized expression of mRNA is associated with mycobacterial resistance controlled by *Nramp1*. Splenic macrophages from Bcg^r and Bcg^s mice were stimulated with rIFN- γ for 24 h followed by treatment with corticosterone. After 20 h, total RNA was extracted, and the effect of corticosterone on the expression of rIFN- γ -induced genes was determined by Northern analysis. Copyright 1997 by the American Society for Microbiology. Reprinted with permission from Brown *et al.* (1997).

These studies, together with previous work on MHC class II expression, led the authors to conduct a series of experiments to determine if mycobacterial growth, which is under *Nramp1* control, may also be affected differentially by HPA axis activation. It was found that activation of the HPA axis increased the growth of mycobacteria in susceptible mice but did not affect the ability of resistant mice to limit growth of the bacilli (Figure 2). The failure of restraint stress to increase the susceptibility of Bcg^r mice was not due to unresponsiveness of this strain of mouse to HPA axis activation. HPA axis activation resulted in an increase in plasma corticosterone and ACTH levels in both strains of mice. The role of corticosterone in mediating the effects of HPA axis activation was demonstrated in three ways. First, adrenalectomy abolished the effect of HPA activation on mycobacterial growth. Second, treatment of Bcg^s mice with the glucocorticoid receptor antagonist RU 486 abrogated the effects of restraint stress. Finally, the implantation of timed-release pellets that elevated plasma corticosterone levels to those attained by restraint increased the growth of mycobacteria in adrenalectomized susceptible mice. A subsequent *in vitro* study of macrophages

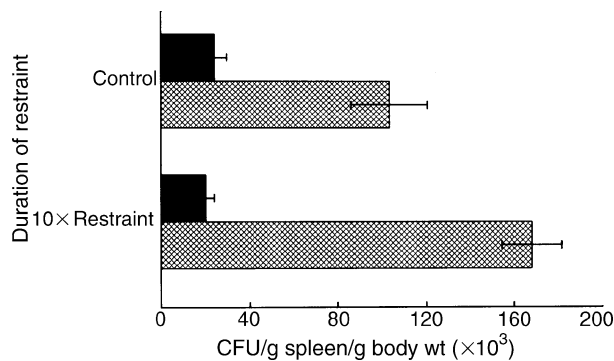


Figure 2 Regulation of mycobacterial growth by the hypothalamic-pituitary-adrenal axis: differential responses of *Mycobacterium bovis* (strain BCG)-resistant and -susceptible mice. Mice were restrained for ten 18-h cycles after infection with 5×10^4 CFU of *M. avium*. The numbers of CFU in the spleens were determined 12 days after infection. Data are the means $\pm$ SD for seven animals per group. The difference between growth of mycobacteria in the spleens of Bcg^r (■) mice versus growth in those of Bcg^s (▨) mice was significant ($P < 0.001$). Copyright 1993 by the American Society for Microbiology. Reprinted with permission from Brown *et al.* (1993).

from Bcg^r and Bcg^s mice showed that corticosterone addition decreased the antimycobacterial activity of susceptible macrophages, whereas corticosterone had no effect on resistant macrophages (Figure 3).

Sympathetic Nervous System Effects on Mycobacterial Infection

Stress results in activation of the HPA axis as well as the SNS. The authors have also reported that epinephrine treatment of peritoneal macrophages suppressed MHC class II expression induced by rIFN- γ . Other observations found that restraint stress also increased epinephrine and norepinephrine levels in the spleen. Numerous other studies have reported that macrophage function can be suppressed or stimulated by catecholamines. It was found that catecholamine treatment of macrophages infected with mycobacteria inhibited intracellular growth of the bacilli. By using specific adrenoceptor agonists and antagonists, it was shown that this effect of catecholamines was mediated via the α_2 -adrenergic receptor. The mechanism that accounts for this stimulation of macrophage functional capacity by catecholamines is under investigation. The use of macrophages harvested from Bcg^s mice in this study demonstrates that susceptible macrophages can be induced to control mycobacterial growth.

It has been reported that epinephrine treatment of *Mycobacterium avium*-infected macrophages increased IL-10 production. This effect was mediated by the β_r -adrenoceptor and could be mimicked by the addition of cAMP. IL-10 is a cytokine produced by a

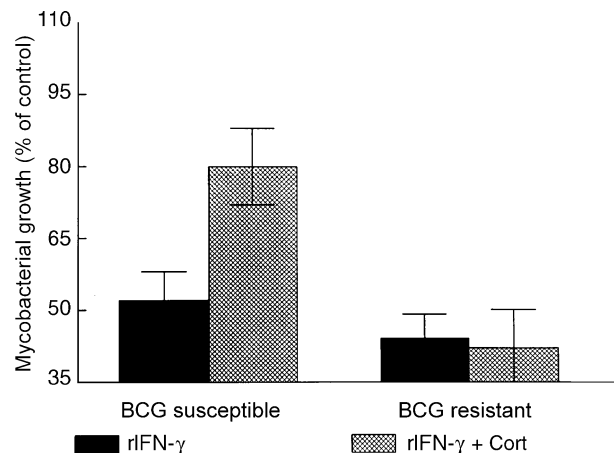


Figure 3 Cytokine-mediated activation of macrophages from *Mycobacterium bovis* (strain BCG)-resistant and -susceptible mice: differential effect of corticosterone on antimycobacterial activity and expression of the *Nramp1* gene. Splenic macrophages were treated with rIFN- γ for 24 h prior to infection with 4×10^5 CFU of *M. avium*. After 24 h to allow for phagocytosis and following removal of unphagocytized bacteria, the cultures were treated with 10^{-6} M corticosterone + rIFN- γ or just rIFN- γ for 5 days prior to lysis of the macrophages and pulsing of the bacteria with [³H] uracil. Data are expressed as the percentage of the control value, which represents the amount of radioactivity incorporated by the bacteria after infection of cultures of macrophages treated with rIFN- γ divided by that of cultures not treated with rIFN- γ . The effect of corticosterone on the antimycobacterial activity of rIFN- γ -activated macrophages from Bcg^s mice was significant ($P \leq 0.03$). Copyright 1995 by the American Society for Microbiology. Reprinted with permission from Brown *et al.* (1995).

variety of immune cells that has a diverse array of actions. IL-10 production could play a role in tuberculosis by affecting the production of IFN- γ by T cells. IL-10 has been shown to inhibit MHC class II expression on monocytes and to inhibit T-cell proliferation directly. Preliminary studies in the authors' laboratory have also found that catecholamine treatment can inhibit macrophage activity. It has been shown that the addition of catecholamines during the infection of rIFN- γ -primed macrophages reduced reactive nitrogen intermediate production. Studies on the regulation of catecholamine-induced modulation of macrophage function are currently underway.

Summary

Differences in the resistance of certain mouse strains to mycobacterial growth have allowed us to study the effect activation of the HPA axis has on the innate mechanisms of mycobacterial resistance. It has been shown that macrophages from *M. bovis* BCG-resistant and -susceptible mice differ in their susceptibility to HPA axis-mediated effects. MHC class II glycoprotein expression by macrophages from Bcg^s mice decreased in response to increased plasma corticosterone levels

induced by restraint. Restraint of Bcg^r mice led to a similar rise in corticosterone levels, but MHC class II expression remained unaffected. Restraint stress also increased mycobacterial growth in susceptible mice but did not affect the ability of resistant mice to limit growth of the bacilli. These results suggest that activation of the HPA axis may account for the increase in susceptibility of Bcg^s mice to mycobacterial growth. This suppression may be particularly apparent in individuals who express the mutation in *Nramp1* that results in susceptibility to mycobacterial disease and may account for the increase in the incidence of active tuberculosis among susceptible populations. Other work in the authors' laboratory has shown that catecholamine treatment of susceptible macrophages increased their ability to inhibit the intracellular growth of mycobacteria. The significance of these findings to mycobacterial disease is under investigation.

See Also the Following Articles

Cytotoxic Lymphocytes; Immune Response; Infection; Macrophages.

Further Reading

- Ader, R., Felten, D. L. and Cohen, N. (eds.) (1991). *Psychoneuroimmunology* (2nd edn.). San Diego: Academic Press.
- Beato, M. (1989). Gene regulation by steroid hormones. *Cell* **56**, 335.
- Blackwell, J. M. (1996). Structure and function of the natural-resistance-associated macrophage protein (Nramp1), a candidate protein for infectious and autoimmune disease susceptibility. *Molecular Medicine Today* **2**, 205–211.
- Bloom, B. R. (1992). Tuberculosis, back to the frightening future. *Science* **358**, 538–539.
- Brown, D. H., Lafuse, W. P. and Zwillling, B. S. (1995). Cytokine-mediated activation of macrophages from *Mycobacterium bovis* BCG-resistant and -susceptible mice: Differential effects of corticosterone on antimycobacterial activity and expression of the *Bcg* gene (candidate *Nramp*). *Infection and Immunity* **63**, 2983–2988.
- Brown, D. H., Lafuse, W. P. and Zwillling, B. S. (1997). Stabilized expression of mRNA is associated with mycobacterial resistance controlled by Nramp1. *Infection and Immunity* **65**, 597–603.
- Brown, D. H., Sheridan, J., Pearl, D. and Zwillling, B. S. (1993). Regulation of mycobacterial growth by the hypothalamic-pituitary-adrenal axis: Differential responses of *Mycobacterium bovis* BCG-resistant and -susceptible mice. *Infection and Immunity* **61**, 4793–4800.
- Centers for Disease Control and Prevention (1995). Tuberculosis morbidity—United States, 1994. *MMWR Morbidity and Mortality Weekly Report* **44**, 387–389, 395.
- Chen, L., Boomershine, C. S., Wang, T., Lafuse, W. P. and Zwillling, B. S. (1999). Synergistic interaction of catecholamine hormones and *Mycobacterium avium* results in the induction of interleukin-10 mRNA expression by murine peritoneal macrophages. *Journal of Neuroimmunology* **93**, 149–155.
- Collins, F. M. (1989). Mycobacterial disease: Immunosuppression and acquired immunodeficiency syndrome. *Clinical Microbiology Reviews* **2**, 360–377.
- Dannenberg, A. M. (1993). Immunopathogenesis of pulmonary tuberculosis. *Hospital Practice* **1**, 51–58.
- Fiorentino, D. F., Bond, M. W. and Mosmann, T. R. (1989). Two types of mouse T helper cells. IV. Th2 clones secrete a factor that inhibits cytokine production by Th1 clones. *Journal of Experimental Medicine* **170**, 2081–2095.
- Ishigami, T. (1919). The influence of psychic acts on the progress of pulmonary tuberculosis. *American Review of Tuberculosis* **2**, 470–484.
- Kuhn, D. E., Baker, B. D., Lafuse, W. P. and Zwillling, B. S. (1999). Differential iron transport into phagosomes isolated from the RAW264.7 macrophage cell lines transfected with Nramp1^{Gly159} or Nramp1^{Asp169}. *Journal of Leukocyte Biology* **66**, 113–119.
- Lalani, I., Bhol, K. and Ahmed, R. (1997). Interleukin-10: Biology, role in inflammation, and autoimmunity. *Annals of Allergy Asthma and Immunology* **79**, 469–481.
- Malefyt, R., Haanen, J., Spits, H., et al. (1991). IL-10 and viral IL-10 strongly reduced antigen specific T cell proliferation by diminishing the antigen presenting capacity of monocytes via down regulation of class II major histocompatibility complex expression. *Journal of Experimental Medicine* **174**, 915–924.
- Malefyt, T., Yssel, H. and de Vries, J. E. (1993). Direct effects of IL-10 on subsets of human CD4⁺ T cell clones and resting T cells. *Journal of Immunology* **150**, 4754–4765.
- Malo, D., Vogan, K., Vidal, S., et al. (1994). Haplotype mapping and sequence analysis of the mouse Nramp gene predict susceptibility to infection with intracellular parasites. *Genomics* **23**, 51–61.
- Miles, B. A., Lafuse, W. P. and Zwillling, B. S. (1996). Binding of alpha-adrenergic receptors stimulates the anti-mycobacterial activity of murine peritoneal macrophages. *Journal of Neuroimmunology* **71**, 19–24.
- Rook, G. A. W., Steele, J., Ainsworth, M. and Leveton, C. (1987). A direct effect of glucocorticoid hormones on the ability of human and murine macrophages to control the growth of *M. tuberculosis*. *European Journal of Respiratory Disease* **71**, 286–291.
- Safley, S. A. and Ziegler, H. K. (1994). Antigen processing and presentation. In: Zwillling, B. S. & Eisenstein, T. K. (eds.) *Macrophage-Pathogen Interactions*, pp. 115–130. New York: Dekker.
- Sheridan, J. F., Dobbs, C., Jung, J., et al. (1998). Stress-induced neuroendocrine modulation of viral pathogenesis and immunity. *Annals of the New York Academy of Sciences* **840**, 803–808.
- Stead, W. W. (1992). Genetics and resistance to tuberculosis. *Annals of Internal Medicine* **116**, 937–941.
- Vespa, L., Johnson, S. C., Aldrich, W. A. and Zwillling, B. S. (1987). Modulation of macrophage I-A expression: Lack of effect of prostaglandins and glucocorticoids on

- macrophages that continuously express I-A. *Journal of Leukocyte Biology* 41, 47–54.
- Wiegshauss, E., Balasubramanian, V. and Smith, D. W. (1989). Immunity to tuberculosis from the perspective of pathogenesis. *Infection and Immunity* 57, 3671–3676.
- Zwilling, B. S. (1994). Neuroimmunomodulation of macrophage function. In: Glaser, R. & Kiecolt-Glaser, J. (eds.) *Handbook of Human Stress and Immunity*, pp. 53–76. New York: Academic Press.
- Zwilling, B. S., Brown, D., Christner, R., et al. (1990). Differential effect of restraint stress on MHC class II expression by murine peritoneal macrophages. *Brain Behavior and Immunity* 4, 330–338.
- Zwilling, B. S., Brown, D., Feng, N., Sheridan, J. and Pearl, D. (1993). The effect of adrenalectomy on the restraint stressed induced suppression of MHC class II expression by murine peritoneal macrophages. *Brain Behavior and Immunity* 7, 29–35.
- Zwilling, B. S., Brown, D. and Pearl, D. (1992). Induction of major histocompatibility complex class II glycoproteins by interferon- γ : attenuation of the effects of restraint stress. *Journal of Neuroimmunology* 37, 115–122.
- Zwilling, B. S., Kuhn, D. E., Wikoff, L., Brown, D. and Lafuse, W. (1999). Role of iron in *Nramp1*-mediated inhibition of mycobacterial growth. *Infection and Immunity* 67, 1386–1392.
- Zwilling, B. S., Vespa, L. and Massie, M. (1987). Regulation of I-A expression by murine peritoneal macrophages: differences linked to the Bcg gene. *Journal of Immunology* 138, 1372–1376.

Macrophages

W P Fehder

New York University, New York, NY, USA

F Tuluc

Children's Hospital of Philadelphia, Philadelphia, PA, USA

W-Z Ho and S D Douglas

Children's Hospital of Philadelphia and University of Pennsylvania Medical School, Philadelphia, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by W P Fehder, W-z Ho and S D Douglas, volume 2, pp 663–668, © 2000, Elsevier Inc.

Origin and Function of Monocytes/Macrophages
Stress and Monocytes/Macrophages

Glossary

Activated macrophage

A macrophage that has been stimulated either biochemically or immunologically and exhibits an increase in one or more functional activities or the appearance of a new functional activity.

Cytokines

A class of low-molecular-weight soluble protein molecules released by immune cells during the course of immune response that serve to mediate and regulate the amplitude and duration of immune/inflammatory responses in the organism. Cytokines are produced during the effector phases of natural and specific immunity in a brief self-limited event. Many of the same cytokine substances are produced by different cells and may act on

many different cell types. Cytokines exert their action by binding to target cell-surface receptors, which may be on the surface of the secreting cell (autocrine action), on the surface of a nearby cell (paracrine action), or on the surface of a cell in a remote location (endocrine action) reached through the bloodstream. Cytokines that are principally synthesized by leukocytes and act primarily on leukocytes.

Large, irregularly shaped white blood cells measuring 25–50 μm in diameter that are the major differentiated cells of the mononuclear phagocyte system.

A class of immune cells that have a common lineage that includes monoblasts, promonocytes, monocytes, and macrophages.

A G-protein-coupled receptor that mediates the biological responses to substance P that has been demonstrated on T cells, B lymphocytes, neutrophils, mast cells, and monocytes/macrophages.

An 11-amino-acid peptide present in the central nervous system that is involved in the pathogenesis of nausea/emesis, pain, mood disorders, stress, anxiety, reinforcement, neurotoxicity, and neurogenesis. It is a member of the tachykinin family of neuropeptides. Outside the central nervous system, SP regulates intestinal motility and secretion, causes vasodilation, and serves as a modulator of a number of important immunological functions. Circulating levels of SP become elevated in stressful situations; monocytes stimulated

Interleukins (ILs)

Macrophages

Mononuclear phagocytes

Neurokinin (NK-1) receptor

Substance P (SP)

Tissue macrophages

by SP produce and release cytokines such as tumor necrosis factor α , interleukin-1, and interleukin-6.

Members of the monocyte/macrophage family that have migrated from the circulation into a body tissue, remaining adherent for several months and ultimately migrating to nearby draining lymph nodes for disposal at the end of their useful life. Some tissue macrophages may survive for the lifetime of the individual.

Macrophages are a diverse group of leukocytes (white blood cells) that share a common origin in the bone marrow and enter the circulation as incompletely differentiated monocytes. Macrophages are the major differentiated cells of the monocyte/macrophage system.

Origin and Function of Monocytes/Macrophages

The monocyte/macrophage system comprises bone marrow monoblasts and promonocytes, peripheral blood monocytes, and tissue macrophages. The peripheral blood monocytes differentiate into macrophages that are capable of phagocytosis of particles, cells, bacteria, and fungi; they also have major roles in antigen presentation and cytokine secretion. Monocytes remain in the circulatory system or migrate to different tissues where they increase in size, becoming tissue macrophages. Tissue macrophages become part of the architecture of specific tissues where they may remain, performing specific protective functions and functions unique to the organ. These tissue macrophages are given names according to their location (Table 1). Tissue macrophages may be stimulated by chemical messengers to break away from their attachment and migrate along with mobile macrophages and monocytes to areas of inflammation by

a process known as chemotaxis. The term monocyte/macrophage system is virtually synonymous with, and has replaced in the nomenclature, the reticulo-endothelial system.

Macrophages perform an essential function in the immune system by attacking and destroying invading bacteria, viruses, and other injurious substances, participating in a wide range of physiological and pathological processes. Although many micro-organisms are phagocytized and destroyed with ease by macrophages, certain pathogens such as listeria, salmonella, brucella, mycobacteria, chlamydia, rickettsia, leishmania, toxoplasma, trypanosoma, and legionella are able to parasitize nonactivated macrophages and replicate within them. Only when macrophages have been activated against these specific intracellular pathogens are they able to inhibit or destroy them. Monocytes and macrophages play an active role in the clearance and inactivation of most viral pathogens. Paradoxically, these same cells also represent a major target cell and infectious reservoir for many viruses. The human immunodeficiency virus (HIV-1) infects and replicates within monocytes and macrophages by binding to the CD4 and CCR5 receptors that are present on the macrophage cell membrane.

The macrophages in tissues and body cavities are a cell population that is regularly renewed by the influx of monocytes manufactured in the bone marrow and circulated via the blood. The control of monocyte production, release of monocytes into the blood, monocyte migration and extravasation into tissues, and finally differentiation into tissue macrophages are accomplished by means of soluble chemical factors. These soluble chemical factors are termed cytokines and are themselves influenced by stress-associated factors such as adrenocorticotrophic hormone (ACTH), cortisol, adrenergic hormones, and neuropeptides such as substance P (SP), vasoactive intestinal peptide (VIP), and proopiomelanocortin (POMC). Macrophages secrete cytokines that produce effects on other immune cells, as well as receptors for cytokines produced by other immune cells, thereby promoting bidirectional communication. The major cytokines that produce effects on the monocyte/macrophage cells are: interferon (IFN- γ), IFN- α , IFN- β , and interleukin (IL-1). Monocyte/macrophages produce and secrete IL-1, IL-6, and tumor necrosis factor (TNF- α).

Phagocytosis and Microbicidal Functions

The term macrophage was first used more than 100 years ago by Elie Metchnikoff to describe the large mononuclear phagocytic cells he observed in tissues. Through the process of phagocytosis, macrophages are

Table 1 Tissue-localized macrophage/monocyte cells

Location	Cell type
Bone marrow	Monoblasts, promonocytes, monocytes, macrophages
Peripheral blood	Monocytes
Skin	Langerhans cell and histiocyte
Lung	Alveolar macrophage
Bone	Osteoclast
Liver	Kupffer cell
Central nervous system	Microglia
Lymphoid organs	Tissue macrophages, dendritic cells
Spleen	Red-pulp macrophages
Gastrointestinal tract	Tissue macrophages
Serous cavities	Pleural and peritoneal macrophages

one of the most effective components of nonspecific immunity. Macrophages also play two essential roles in specific immunity, first as professional antigen presenting cells and then, with proper stimulation, as immune effector cells. Once activated by the immune system, macrophages become more powerful than neutrophils in their ability to phagocytize. Macrophages are able to engulf much larger particles such as whole red blood cells and malarial parasites or as many as 100 bacteria. Phagocytized particles are encapsulated and digested by intracellular proteolytic enzymes contained in lysosomes and other cytoplasmic granules. Macrophage lysosomes are unusual in that they contain lipases that aid in the digestion of the thick lipid bacterial membranes. Once the engulfed particles are digested, macrophages can extrude residual products. Macrophages perform many essential roles in the immune response such as the direct killing of damaged cells, bacteria, and viruses and antigen presentation. Macrophages produce and respond to cytokines and other chemicals central to both the inflammatory and the specific immune response. There are populations of macrophages that live throughout the life span of the organism.

Cytokine Mediation of Macrophage Response

Macrophages express major histocompatibility complex (MHC) class II on their surface and are important antigen-presenting cells for both B and T lymphocytes. Macrophages, when activated by cytokines produced by T helper (TH) cells, become effector cells of cell-mediated immunity. Macrophages are the first cells involved in an immune reaction and help orchestrate immune events both by the production of chemical signals (cytokines and SP) and by the processing and presentation of antigens to subsets of T lymphocytes that may either enhance or suppress specific immune responses. Activated macrophages also produce cytokines, such as IL-1, TNF- α , and IL-6, that augment T-cell function, as well as IL-8, macrophage colony-stimulating factor (M-CSF), granulocyte macrophage colony-stimulating factor (GM-CSF), and granulocyte colony-stimulating factor (G-CSF) that stimulate polymorphonuclear neutrophil (PMN) and monocyte activity. The ability of macrophages and lymphocytes to stimulate one another's functions provides an important amplification mechanism for specific immunity. Thus, macrophages play an important and central immunoregulatory role.

Toll-like Receptors

Toll-like receptors, identified in macrophages in mammals, are a relatively new family of membrane receptors able to recognize specific patterns of pathogen

components, such as bacterial endotoxins and viral nucleic acids. Toll-like receptor (TLR)-4 is part of a molecular complex that recognizes and mediates intracellular responses to lipopolysaccharide (LPS). Pathogen recognition by TLRs activates the innate immune system through different signaling pathways, causing inflammatory responses such as cytokine production. Impaired adrenal stress response occurs in TLR-2-deficient mice, suggesting that TLR-2 and, possibly, TLR-4 may constitute an important link between the immune and endocrine stress systems at both the central and peripheral levels, particularly during inflammation and sepsis.

Stress and Monocytes/Macrophages

An examination of the role of macrophages in the modulation of immune function by stress is particularly compelling. A representation of the interactions among various stressors and the immune system is provided in [Figure 1](#).

Macrophages play a pivotal role in immune responses, first as professional antigen-presenting cells stimulating the expansion of one of two sets of CD4⁺ T-cell populations (TH1 or TH2) and finally, when activated themselves, functioning as major effector cells in cell-mediated immunity. The TH1 and TH2 responses are highly integrated and contraregulated, relying on two distinctly different profiles of cytokine production. Many of the cytokines produced by one or the other subpopulation have positive effector functions and, in addition, negatively regulate the activity of the alternate subgroup. The negative regulation may be at the level of either T-cell expansion or T-cell effector functions, including macrophage activation. Thus, an immune response in which TH1 cells are stimulated downregulates the co-stimulation of TH2 cells. Conversely, an immune response in which TH2 cells are stimulated downregulates both the co-stimulation of TH1 cells and their effector activity.

TH1 cells produce cytokines associated with aspects of classic cell-mediated immunity, producing IFN- γ , IL-2, IL-3, GM-CSF, TNF- α , and TNF- β , which activate macrophages. IFN- γ induces oxygen burst and nitric oxide production, which are essential for microbicidal reactions. IL-3 activates macrophages, causing increased antigen presentation. IFN- γ and IL-2 both stimulate macrophages to produce TNF- α , which in turn further stimulates macrophage activation in a cascade of events.

TH2 cells produce cytokines (IL-4, IL-5, and IL-6) that are associated with the B-cell responses of humoral immunity. IL-4 acts as a growth factor specific for TH2 cells, acts as an inhibitory factor for the

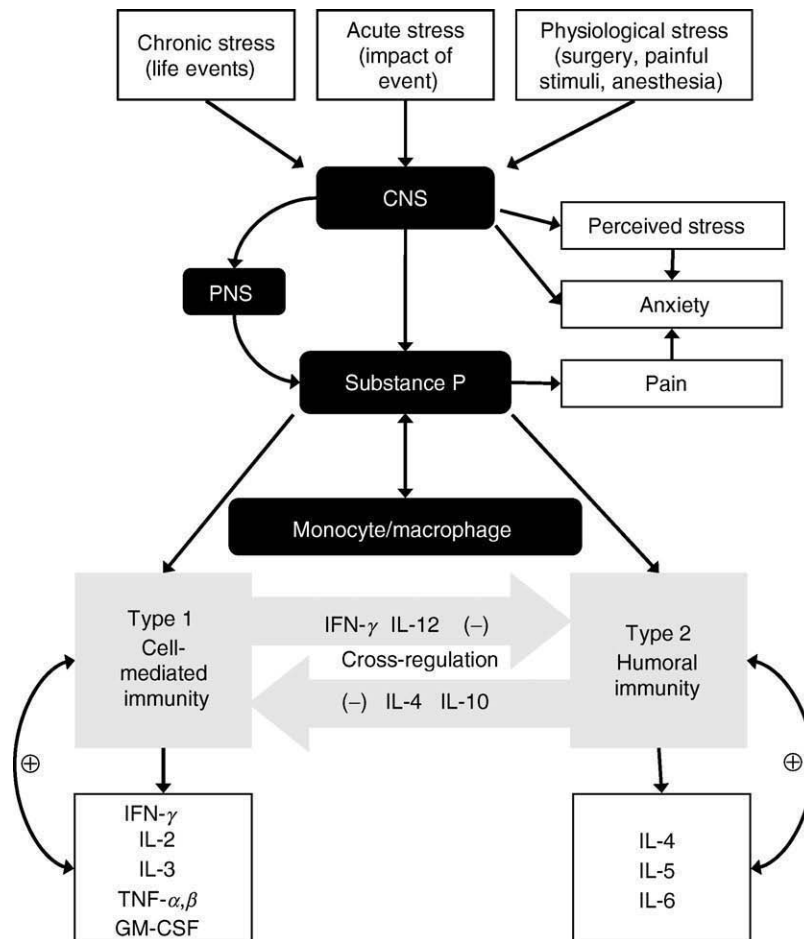


Figure 1 Stress effects on monocyte/macrophages. The effects of stress on the immune system are mediated by the neuropeptide substance P. The cells of the monocyte/macrophage system have receptors for and are able to synthesize and release substance P. Thus, they play a central role in mediating these effects. CNS, central nervous system; GM-CSF, granulocyte macrophage colony-stimulating factor; IFN, interferon; IL, interleukin; PNS, peripheral nervous system; TNF, tumor necrosis factor.

stimulation of TH1 cells by IL-2, and also effectively inhibits IFN- γ activation of macrophages. When appropriately stimulated by TH2 cytokines and antigen, B cells give rise to antibody-secreting plasma cells. The antigen-antibody complex serves as a powerful stimulator of the monocyte/macrophage phagocytosis of these circulating complexes.

In a mouse surgical-sponge model of inflammation, acute restraint-induced stress caused increase in leukocyte trafficking to the site of immune activation. Mice acutely stressed before sponge implantation presented 200–300% higher macrophage, neutrophil, natural killer cell, and T-cell infiltration than did nonstressed animals. Stress-induced increases in leukocyte trafficking may be responsible for immunoprotection during trauma, surgery, and infection, but it may also exacerbate immunopathological responses during inflammatory diseases, such as arthritis, cardiovascular disease, psoriasis, multiple sclerosis, and gingivitis.

Substance P as a Mediator of Stress Response

A major mediator of the effects of stress on monocyte/macrophage function is the neurotransmitter peptide SP, which is widely distributed in the central, peripheral, and enteric nervous systems. SP may have a major role in several inflammatory disease processes such as rheumatoid arthritis, bullous pemphigus, asthma, and inflammatory bowel diseases such as Crohn's disease and ulcerative colitis, as well as in the transmission of pain. SP may be the chemical messenger of a neurogenic alarm system that serves to communicate stress in a reciprocal manner among the central nervous system (CNS), endocrine system, and immune system. SP is associated with the sympathetic nervous system and is essential to the maintenance of adrenal catecholamine secretion from the adrenal medulla in times of stress.

Studies of surgical patients demonstrate that the stress of surgery shifts the TH1/TH2 cytokine balance

toward TH2, suggesting that cell-mediated immunity was downregulated. Furthermore, the highest immune alteration was associated with the most stressful and invasive surgery. Several experimental animal studies indicate an association between stress and levels of SP. Two studies of SP in the brain tissue of rats that had been subjected to stressful stimuli found decreases in SP levels, which the authors attributed to tissue SP depletion following the increased release of SP. Experimental models of induced stress such as restraint stress and sequential removal of group-housed rats were found to be associated with increased levels of SP in the periaqueductal gray area of the brain. Insulin-induced hypoglycemia or mild electric foot shock were shown to increase the release of both catecholamines and SP from rat adrenomedullary cells. The intracerebroventricular administration of SP in conscious rats elicited an integrated cardiovascular, behavioral, and endocrine response in a pattern consistent with an integrated stress response in rodents. Experimental evidence in first-time parachutists and patients undergoing diagnostic procedures demonstrated increased levels of circulating SP associated with acute psychological stress. The high levels of circulating SP remained elevated long after the anxiety associated with the acute stressor diminished.

Macrophages, endothelial cells, eosinophils, and Leydig cells have also been identified as nonneuronal sources of SP. The presence of specific SP receptors on the surface of monocytes and the ability of macrophages to produce SP suggest its role in the autocrine regulation of SP gene expression.

The biological responses to SP are mediated by the NK-1 receptor, a 407-amino-acid G-protein-coupled glycosylated receptor that has been demonstrated on T cells, B lymphocytes, monocytes, macrophages, neutrophils, and mast cells. There is a connection between SP in the peripheral nervous system and the immune system. There is direct innervation of the spleen by SP-containing nerve fibers in splenic areas adjacent to lymphocytes, macrophages, and granulocytes, with evidence of its role in the traffic of lymphocytes from lymph nodes. SP modulates the effector phase of adaptive immunity and serves as a co-stimulator in the activation of monocytes, macrophages, T cells, and B cells, with the subsequent release of regulatory cytokines by these cells.

Effect of Stress on Monocytopoiesis

Macrophages synthesize and release two hematopoietic growth factors, M-CSF and GM-CSF, in response to external stimuli such as particles that are phagocytosed and endotoxin. Furthermore, macrophages

release IL-1 and TNF, which in turn can induce fibroblasts and endothelial cells to produce M-CSF and GM-CSF, which stimulate the production of mononuclear phagocytes. The spiral of increased inflammatory response is further magnified by the synthesis of TNF- α by mononuclear phagocytes in response to GM-CSF stimulation. Thus, macrophages play a pivotal role in regulating immune response. SP induces the production of hematopoietic growth factors IL-1, IL-6, and TNF- α from human monocytes. Bone marrow is innervated by nerve fibers that release SP in concentrations that induce tissue macrophages resident in bone marrow to release IL-1, IL-3, and GM-CSF.

Furthermore, macrophages express NK-1 receptors and serve as a nonneuronal source of SP. The release of SP by macrophages serves as a paracrine mechanism (by stimulating nearby macrophages) for amplifying the effect of stress on monocytopoiesis. Thus, stress acting through SP as a mediator may play a physiological role in regulating monocytopoiesis through two mechanisms: stimulation of the direct release of hematopoietic growth factors by monocytes and paracrine amplification.

Effect of Stress on Monocyte/Macrophage Function

The manner and extent to which stress influences immune response is, then, of vital concern. Stress may either enhance or inhibit macrophage function. Experimental evidence shows that SP plays a role in modulating immune response at each of the succeeding levels of response to foreign antigen. SP induces the release of stem cell factor and IL-1 in bone marrow stroma, thereby mediating hematopoiesis. A summary of the effects of stress on the monocyte/macrophage system is provided in [Table 2](#).

During acute inflammatory reactions, the number of circulating monocytes is increased with enhanced bone marrow production. This is mitigated by the decreased time that the monocytes are in the circulation due to efflux into inflammatory exudates. The SP-containing neurons have direct contact with immune and hematopoietic cells in the primary and secondary lymphoid organs. Thus, traffic of lymphocytes may be regulated by the secretion of neuroactive peptides such as SP from nerves.

Macrophages and monocytes increase phagocytosis and increase chemotaxis in response to SP. Increased levels of IL-6 are released by peritoneal macrophages in response to increased levels of SP produced by mice exposed to cold-water stress. Macrophages from humans and guinea pigs have receptors for and respond to SP-releasing cytokines

Table 2 Effects of stress on monocytes/macrophages

Source	Model	Effect
Animal (<i>in vitro</i>)	Macrophages exposed to salmonella	Increased SP receptors
Animal (<i>in vivo</i>)	Mice exposed to salmonella given SP antagonist	Increased mortality
	Mice exposed to cold-water stress	Increased SP; increased IL-6 release by macrophages
	Stressed rats (restraint and sequential removal)	Increased levels of SP in periaqueductal gray area of brain
Human (<i>in vitro</i>)	Bone marrow stroma cells exposed to SP	Increased stem cell factor, IL-1
	Monocytes/macrophages exposed to SP	Increased IL-1, IL-6, IL-7, IL-10, IL-12, TNF- α ; chemotaxis, and phagocytosis
Human (<i>in vivo</i>)	Acute psychological stress in parachutists	Increased SP
	Acute psychological stress in medical diagnostic procedures	Increased SP, CD8 ⁺
	Sickle cell patients in vaso-occlusive crisis	Increased SP correlated with pain; increased IL-8

IL, interleukin; SP, substance P; TNF, tumor necrosis factor.

IL-1, IL-6, and TNF- α . Macrophages in mice exposed to oral salmonella doses expressed higher levels of SP receptors and the administration of the SP antagonist spantide II resulted in higher mortality than in non-SP-antagonized mice. Treatment with spantide II was also associated with significantly reduced levels of IL-12 and IFN- γ . Thus, SP serves to amplify ongoing immune reactions and serves as a mediator in stress-induced immune reactions.

See Also the Following Articles

Cytokines; Immune Response; Macrophage Antimicrobial Activity, Effects of Stress on; Neuroimmunomodulation.

Further Reading

Bornstein, S. R., Zacharowski, P., Schumann, R. R., et al. (2004). Impaired adrenal stress response in toll-like receptor 2-deficient mice. *Proceedings of the National Academy of Sciences USA* 101, 16695–16700.

Chancellor-Freeland, C., Zhu, G., Kage, R., et al. (1995). Substance P and stress-induced changes in macrophages. *Annals of the New York Academy of Sciences* 771, 472–484.

Douglas, S. and Ho, W. (2006). Morphology of monocytes and macrophages. In: Lichtman, M. A. (ed.) *Williams hematology* (7th edn., pp. 960–970). New York: McGraw Hill.

Hassan, F. and Douglas, S. (1990). Stress-related neuroimmunomodulation of monocyte-macrophage functions in HIV-1 infection. *Clinical Immunology and Immunopathology* 54, 220–227.

Lewis, C. and McGee, J. (1992). *The macrophage*. Oxford: Oxford University Press.

Ho, W. Z. and Douglas, S. D. (2004). Substance P and neurokinin-1 receptor modulation of HIV. *Journal of Neuroimmunology* 157, 48–55.

Lai, J. P., Ho, W. Z., Kilpatrick, L. E., et al. (2006). Full-length and truncated neurokinin-1 receptor expression and function during monocyte/macrophage differentiation. *Proceedings of the National Academy of Sciences USA* 103, 7771–7776.

McGillis, J., Mitsuhashi, M. and Payan, D. (1990). Immunomodulation by tachykinin neuropeptides. *Annals of the New York Academy of Sciences* 594, 85–94.

Perry, V. (1994). *Macrophages and the nervous system*. Austin, TX: R. G. Landes.

Viswanathan, K. and Dhabhar, F. S. (2005). Stress-induced enhancement of leukocyte trafficking into sites of surgery or immune activation. *Proceedings of the National Academy of Sciences USA* 102, 5808–5813.

Major Depressive Disorder

A B Negrão and P W Gold

National Institute of Mental Health, Bethesda, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 669–675, © 2000, Elsevier Inc.

Stress System

Syndromes of Major Depression

Glossary

<i>Amygdala</i>	Brain region involved in the expression and acquisition of conditioned fear.
<i>Atypical depression</i>	Major depressive disorder with a predominance of somatic features (increased appetite, sleep, and weight gain) that are the inverse of what is seen in melancholia.
<i>Cushing's disease</i>	A disorder characterized by the clinical effects of increased circulating glucocorticoids caused by a pituitary ACTH-producing microadenoma.
<i>Major depression</i>	A psychiatric disorder characterized by depressed mood, loss of pleasure in most activities, and symptoms that affect cognitive function, appetite, weight control, sleep patterns, and sexual function.
<i>Melancholia</i>	A historical term that designates the presence of decreased appetite, weight loss, complete lack of pleasure, insomnia, and early morning worsening of symptoms in a patient with major depression.

Stress System

The stress response consists of a simultaneous repertoire of behavioral and physiological adaptations that promote survival during threatening situations. The most apparent are fear-related behaviors that include anxiety, decreased exploration (or freezing), increased vigilance, and intense focus on the threatening stimulus. Concurrently, there is evidence of physiological activation of both the hypothalamic-pituitary-adrenal (HPA) axis and the sympathetic nervous system. Metabolically, a catabolic state ensues that mobilizes fuel for the stressed body site or the brain, and endocrine programs for growth and reproduction are inhibited. Finally, there is systematic inhibition of vegetative functions whose execution would impair the likelihood of survival during a life-threatening crisis (e.g., feeding, sexual behavior, and sleep).

This section focuses on the core stress systems, the corticotropin-releasing hormone (CRH) (including the HPA axis) and the locus ceruleus–norepinephrine (LC–NE) systems. We also discuss the importance of other relevant neural substrates to the stress response and major depression.

Corticotropin-Releasing Hormone

Several discrete CRH systems have been identified in the brain. CRH was first discovered as the hypothalamic hormone that plays the principal role in the regulation of the HPA axis. Intrahypothalamic CRH pathways contribute to the inhibition of feeding and sexual behavior and to the inhibition of endocrine programs for growth. Parvocellular CRH neurons in the paraventricular nucleus (PVN) of the hypothalamus send terminals to brain stem arousal and autonomic centers. An extrahypothalamic CRH system has been identified in the amygdala and is important in modulating aversively charged emotional memory. The administration of a CRH antagonist directly into the central nucleus of the amygdala (CEA) diminishes fear conditioning significantly. These CRH pathways represent the anatomical context for the observation that the intracerebroventricular administration of CRH recapitulates many of the metabolic, endocrine, autonomic, and behavioral features of the stress response. While it is well known that cortisol participates in a classic negative feedback loop to restrain the HPA axis, amygdala and descending hypothalamic CRH pathways are actually activated by cortisol. Cortisol also contributes significantly to the inhibition of programs for growth and reproduction seen during stress, as well as to feedback restraint on an activated immune system ([Figure 1](#)).

Locus Ceruleus–Norepinephrine System

The LC–NE system originates in the mid-pons and sends a dense network of terminals to many cortical and subcortical structures. The LC–NE system serves globally as a generalized warning system, inhibits vegetative functions such as grooming, eating, and sleeping, and contributes to accompanying increases in autonomic outflows. Ascending norepinephrine pathways project to the PVN CRH neurons and exert excitatory effects on the HPA axis. Brain stem norepinephrine, in turn, also influences other key structures that work together with the core stress system components, including the amygdala.

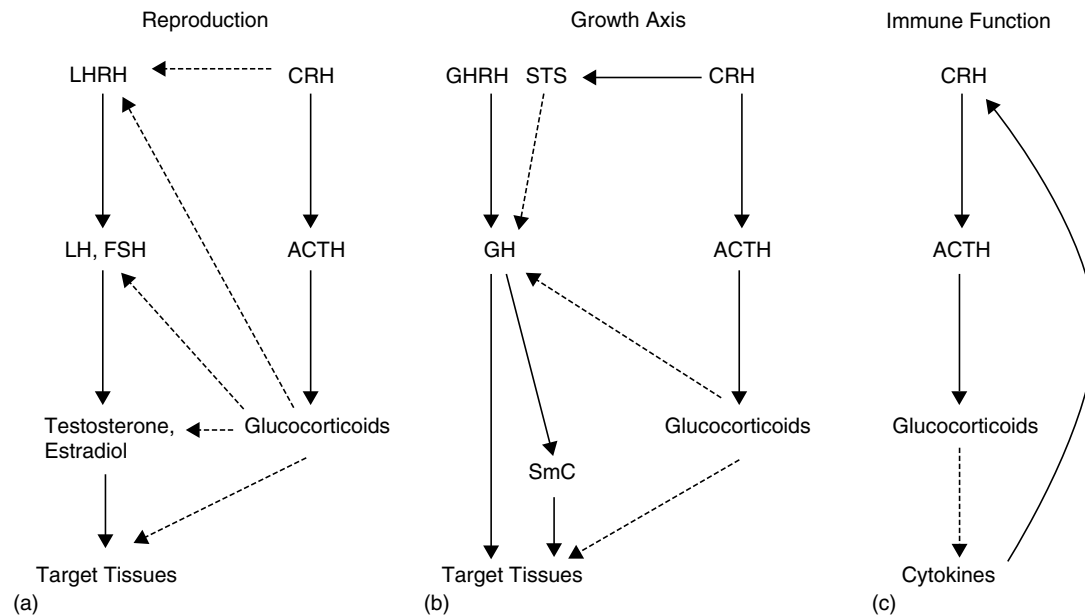


Figure 1 A representation of the interactions between the hypothalamic-pituitary-adrenal axis and other neuroendocrine systems: (a) reproductive axis, (b) the growth axis, (c) immune function. LHRH, luteinizing hormone-releasing hormone; CRH, corticotropin-releasing hormone; LH, luteinizing hormone; FSH, follicle-stimulating hormone; ACTH, corticotropin; GHRH, growth hormone-releasing hormone; STS, somatostatin; GH, growth hormone; SmC, somatomedin C. Dotted lines indicate inhibition and solid lines indicate stimulation.

CRH/LC-NE Systems and Interactions with Neural Substrates of Relevance to the Stress Response and Depression

Activation of the stress system sets into motion a range of fear-related behaviors and promotes the encoding of conditioned fear. The amygdala is the key structure involved in conditioned fear, and, as noted, CRH in the central nucleus of the amygdala in the rat plays an important role in this phenomenon. Activation of the amygdala is associated with activation of the HPA axis, most probably via direct projections from the CEA to the PVN. There is also evidence that the amygdala can activate the HPA axis indirectly via projections to brain stem catecholaminergic centers.

The classic effect of glucocorticoids on the CRH system is the negative feedback effect exerted on the HPA axis. Glucocorticoid receptors in the hippocampus play an important role in mediating this negative feedback restraint. However, not all glucocorticoid effects on the CRH system are inhibitory. Thus, glucocorticoids actually enhance the expression of CRH in the CEA and increase fear-related behaviors. This glucocorticoid effect provides the context for a feed-forward loop between the amygdala and the HPA axis.

The core stress systems and the amygdala are also linked closely to the medial prefrontal cortex (MPC). This complex structure is important in promoting the extinction of conditioned fear encoded in the amygdala and, by doing so, exerts restraint on stress system activity. Pathways from the MPC to the brain stem

also inhibit brain stem autonomic centers. This structure also sends the only direct cortical connections to the hypothalamus, and the activation of glucocorticoid receptors located in the MPC inhibits the HPA axis. It has been shown that the chronic hypercortisolism of Cushing's disease is associated with a significant decrease in the size of the prefrontal cortex. This finding suggests the possibility of another feed-forward loop in the context of sustained hypercortisolism, a state in which the inhibitory actions of the MPC on the hypothalamus would be impaired due to cortisol-mediated atrophy.

The medial prefrontal cortex is also involved in a critical array of complex behaviors. This structure helps decide whether a particular situation is likely to result in reward or punishment and is involved in the shifting of affect from one state to another based on internal and external cues and in orchestrating sequences of sophisticated repertoires of motor activity and behavior. Activation of the amygdala takes the medial prefrontal cortex off line, thus inhibiting complex, innovative behaviors and promoting the kinds of autonomic, stereotyped repertoires that would be most adaptive in an immediately life-threatening situation.

Syndromes of Major Depression

Definition

Depression affects up to 15% of women and 8% of men worldwide. According to the World Health

Organization, depression is a greater source of morbidity for women than any other illness. The *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (DSM-IV), edited by the American Psychiatric Association, provides clinicians and researchers with criteria to diagnose someone who suffers from depression. Major depressive disorder, as the condition is named, is characterized by depressed mood or inability to experience pleasure and at least four of the following associated symptoms: feelings of worthlessness, suicidal thoughts, diminished concentration, and changes in weight, appetite, psychomotor activity, and sleep. The symptoms have to be severe enough to cause impairment in the patient's daily activities, and they have to last at least 2 weeks in order for the diagnosis to be firmed. The causes of the disorder are not known. Stressful life events play a role in the course of the disorder, and there are studies demonstrating the importance of heritable factors.

Phenotypic Subtypes of Major Depression

Major depression is a disorder that is clinically heterogeneous. Although the core symptoms (depressed mood and inability to experience pleasure) need to be present, there are many distinct clinical presentations of the disorder. Differences in the subtypes of major depression are determined by the combination of the associated features. Two subtypes of major depression, namely, major depression with melancholic features (melancholia) and major depression with atypical features (atypical depression), have received the greatest attention from investigators (Table 1).

Although the term depression implies a suppression of thought and feeling, melancholia is instead a state of hyperarousal and fear often stemming from a profound sense of personal unworthiness and pessimism about future prospects. Any loss in cognitive and emotional range is counterbalanced by the intensity of this anxiety, which coincides with, and is perhaps inseparable from, a loss of pleasure in everyday activities. Patients with severe melancholia seem to be in especially close contact with painfully charged memories of past losses and failures. The emotional intensity of melancholia is matched by indices of physiological hyperarousal (insomnia, early morning awakening, and hypercortisolemia). Patients also experience inhibition of vegetative functions: loss of appetite and libido, inhibition of growth, and sex steroid secretion.

In contrast to melancholia, patients with atypical depression often feel much less alive than usual and experience a profound inertia and fatigue that often isolates them and interferes with normal functioning. Characteristically, they sleep excessively, eat a lot,

Table 1 Characteristics of major depression with melancholic features (melancholia) and atypical features (atypical depression)

Characteristic	Melancholia	Atypical depression
Vegetative functions	Insomnia, anorexia, weight loss, diminished libido	Hypersomnia, increased appetite, weight gain
Behavior	Anxious, preoccupied with concerns about deficient self, preferential access to dysphorically charged memories	Partial reaction to positive stimuli, intense fatigue, diminished sense of connection to self
Stress system Medical consequences	Hyperactive Osteoporosis, increased cardiovascular morbidity and mortality, prefrontal cortex abnormalities	Hypoactive Autoimmune disorders?
Response to treatment	Similar response to different classes of antidepressants	Preferentially to MAOIs ^a

^aMonoamine oxidase inhibitors, an antidepressant medication that acts on the catabolic path of monoamines.

and usually gain weight. Like patients with melancholia, they have a diminished libido and a blunted capacity to experience everyday pleasure, although they may react temporarily to positive external and internal clues (mood reactivity).

Major Depression as a Disorder of Adaptation

As reviewed previously, the stress response system is a stereotyped, adaptive repertoire that ensures survival during threatened homeostasis. Major depression, by definition, is associated with relatively unshifting affect and simultaneous changes in fundamental biological processes that regulate sleep, food intake, metabolism, and reproduction. Similarities between the stress response and the phenomenology of melancholia prompted us to first postulate that the clinical and biochemical manifestations of this disorder reflect a dysregulation in stress system mediators. Changes seen during the stress response are under strict control by counterregulatory mechanisms that serve to limit the response in time and magnitude. Melancholia can be seen as a state of chronically and excessively activated stress response, possibly due to an escape of the usual regulatory influences. In that sense, melancholia would be a disorder of adaptation, more specifically, one in which the stress response required for survival becomes dysregulated and is the very source of the

illness. In contrast, many symptoms found in atypical depression can reflect a pathological inactivation of stress system mediators. The remainder of this article reviews the lines of evidence indicating the convergence between a dysregulation of the stress response and the biochemical and neurobiological findings in both subtypes of major depression.

Major Depression with Melancholic Features

Hypercortisolemia and hypothalamic CRH secretion

Sustained hypercortisolism has been demonstrated repeatedly in patients with major depression and is perhaps the best documented finding in the biology of psychiatric disorders. Many lines of evidence indicate that hypercortisolism of major depression is due to the hypersecretion of hypothalamic CRH neurons. Data obtained in brains taken from patients who had suffered with affective illness reveal overexpression of CRH in the hypothalamus.

Norepinephrine secretion Patients with melancholic depression show sustained and profoundly elevated levels of cerebrospinal fluid (CSF) NE when sampled through an indwelling lumbar drain for 30 h. Moreover, CSF NE levels were highly elevated during sleep, indicating that the hypersecretion of centrally directed norepinephrine was not an artifact of the conscious distress of the illness. Increased basal norepinephrine spillover into arterialized venous blood was noted in patients with major depression with melancholic features. We have subsequently replicated this finding in arterial plasma and have shown that patients with melancholic depression show significantly higher norepinephrine spillover than controls in response to psychological and pharmacological stimuli. Our patients also showed significantly higher basal heart rates and heart rate increases to psychological and pharmacological stressors.

Pathophysiology of melancholic depression The reciprocal connectivity of the structures that participate in the modulation of the core stress system suggests that similar phenotypes can be generated by primary defects in any of these structures, including the prefrontal cortex, the amygdala, or the CRH/LC-NE systems. For example, data showed a decrease in the volume and metabolic activity of the left subgenual prefrontal cortex in patients with major depression. Theoretically, an initial decrease in the function of the MPC could potentially enhance amygdala activation as well as the activation of the HPA axis and brain stem noradrenergic neurons. An activated amygdala and hypothalamic CRH centers would further activate brain stem arousal and autonomic centers.

An increased noradrenergic drive from the brain stem could potentiate the activity of the HPA axis. Hypercortisolism in major depression could be enhanced further by the capacity of glucocorticoids to damage or destroy hippocampal glucocorticoid receptor-containing neurons that mediate negative feedback on the HPA axis. In this regard, Sheline and colleagues found that patients with depressive episodes in the past had diminished bilateral hippocampal volumes compared to controls matched by age and medical conditions.

Taken together, a number of feed-forward loops among key components of the stress system could emerge that would intensify the stress system dysfunction in major depression, leading to a diminished activity of the medial prefrontal cortex and activation of the amygdala, HPA axis, and brain stem noradrenergic centers. The final result could be a syndrome associated with a relatively stereotyped affect, increased fear, and activation of the HPA axis and sympathetic nervous system. Such a syndrome would closely resemble the syndrome of melancholic depression.

Long-term medical consequences Patients with major depression exhibit a number of neuroendocrine and autonomic changes that could potentially predispose them to long-term medical consequences such as osteoporosis and premature ischemic heart disease. Hypercortisolism and suppression of the growth hormone and reproductive axes during depressive episodes would be expected to result in loss of bone mineral density. Unfortunately, bone mineral density, once lost, is not regained easily. In a study of premenopausal women (average age 41) with past or current major depression, almost 40% showed a loss of bone mineral density at the hip or spine that was at least 2 standard deviations below peak bone age, thus meeting American Osteoporosis Society criteria for osteoporosis. Because peak bone age is reached at around age 30, this means that these patients had lost more than an average of 1.5% of bone mineral density per year. Thus, by the time they enter menopause, these patients could have severe osteoporosis with losses as great as 35% below peak bone mass. Because each 10% decrease in bone mineral density is associated with a doubling of the fracture rate, the bone loss seen associated with depression represents an additional risk to fractures in aging women.

Patients with major depression show an overall increase in mortality regardless of age (independent of suicide). This relates, in part, to the fact that the relative risk for ischemic heart disease in patients with major depression is two or more and this increased risk is independent of traditional risk factors such

as hypertension, dyslipidemia, and smoking. Many of the changes noted in depression could contribute to this enhanced susceptibility, including hypercortisolism and inhibition of the growth hormone and reproductive axes (leading to, among other things, increased visceral fat mass), with the resultant insulin resistance and its sequelae. Norepinephrine and epinephrine spillover studies showed that patients with depression had higher basal and stimulated heart rates and elevated norepinephrine spillover rates in the context of normal epinephrine secretion. These data strongly suggest increased cardiac sympathoneural activity in patients with major depression, another likely risk factor for ischemic heart disease. Thus, major depression with melancholic features is not only a disorder that affects fundamental parameters that define our humanity, such as self-regard, but also the product of sustained dysregulation in global neuroendocrine and autonomic functions acting on the myriad of peripheral substrates subjected to their influences.

Major Depression with Atypical Features

Biochemical manifestations and pathophysiology Relatively few studies have been conducted on the neurobiology of atypical depression; most have focused on HPA axis regulation. We studied patients with Cushing's disease, a rare illness characterized by severe hypercortisolism and a high incidence of depression. We demonstrated that depression in those patients was almost all of the atypical subtype. After having shown that the profound hypercortisolism of Cushing's disease was attributable entirely to a pituitary rather than central defect, we then focused on the function of hypothalamic CRH-containing neurons in Cushing's disease. In a series of studies, we showed that hypothalamic CRH-containing neurons in patients with Cushing's disease were suppressed by long-standing hypercortisolism. This confluence of atypical depression with a suppressed CRH system led us to hypothesize that some forms of atypical depression were associated with a pathological suppression rather than activation of arousal producing stress mediators.

In subsequent studies, we determined the pattern of plasma ACTH and cortisol responses to ovine CRH in patients with known hypothalamic CRH deficiency, secondary either to traumatic pituitary stalk section or to hypothalamic neoplasia. These patients often showed blunted and delayed plasma ACTH and cortisol responses, indicating that the adrenal cortex had been hypostimulated in these subjects. To test the hypothesis that patients with fatigue states and classic forms of atypical depression might show evidence of

hypothalamic CRH deficiency, patients with seasonal affective disorder and patients with chronic fatigue syndrome were studied. Seasonal affective disorder is characterized by the classical phenotype of atypical major depression. Both of these groups showed plasma ACTH and cortisol responses to CRH similar to those seen in patients with hypothalamic CRH deficiency. We have also shown this pattern in patients with atypical depression studied in the postpartum period.

In summary, these observations raise the question that the presentation of atypical symptoms (lethargy, fatigue, hypersomnia, and hyperphagia), as seen in Cushing's disease and in atypical depression, might be associated with hyposecretion of hypothalamic CRH. More studies will be needed to definitively document this important point.

Potential medical implications of HPA axis dysfunction in atypical depression If a centrally mediated hypoactivity of the HPA axis were documented in patients with atypical depression, this finding could have implications both for the clinical features of atypical depression, such as lethargy, fatigue, and hyperphagia, and for susceptibility to long-term medical consequences. Data show that peripheral inflammatory mediators such as interleukin-6 and tumor necrosis factor activate hypothalamic CRH neurons and the hypothalamic-pituitary axis. This circuit, during inflammatory stress, has been shown to promote HPA activation and glucocorticoid-mediated restraint of the immune response. The putative role of hypothalamic CRH and subsequent pituitary-adrenal activation is as a counterregulatory restraint on the immune response. Thus, it has been postulated that during the stress of a fight-or-flight situation, inflammatory mediators recruit the HPA axis to prevent the immune response from overshooting in the context of an anticipated injury and powerful inflammatory stimulus.

In *in vivo* studies, we and others have shown that adrenalectomized rats given an inflammatory stimulus show a profoundly exaggerated inflammatory response that, in some instances, resulted in death. In addition, we have shown that susceptibility to inflammatory disease in Lewis rats is associated with the hypoactivity of hypothalamic CRH in response to inflammatory mediators and other stimuli. Lewis rats also show significant attenuations in behavioral and other physiological responses to inflammatory mediators and other stressors. In the light of these data, we propose that patients with atypical depression could represent the tip of the iceberg of subjects who complain of lethargy, fatigue, and apathy, as well as an array of low-level but persistent complaints

such as muscle pain, joint pain, and allergic phenomena. We are currently continuing to test this hypothesis utilizing new paradigms that will hopefully provide sufficient resolution to detect subtle hypoactivity of the HPA axis secondary to a change in CRH function.

See Also the Following Articles

Anxiety; Corticotropin Releasing Factor (CRF); Cushing's Syndrome, Medical Aspects; Cushing's Syndrome, Neuropsychiatric Aspects; Depression and Manic-Depressive Illness.

Further Reading

- Chrousos, G. P. and Gold, P. W. (1992). The concepts of stress and stress system disorders: overview of physical and behavioral homeostasis. *Journal of the American Medical Association* 267, 1244–1252.
- Chrousos, G. P. and Gold, P. W. (1998). A healthy body in a healthy mind – and vice versa – the damaging power of “uncontrollable” stress. *Journal of Clinical Endocrinology and Metabolism* 83, 1842–1845.
- Gold, P. W., Goodwin, F. K. and Chrousos, G. P. (1988). Clinical and biochemical manifestations of depression: relation to the neurobiology of stress (1). *New England Journal of Medicine* 319, 348–353.
- Gold, P. W., Goodwin, F. K. and Chrousos, G. P. (1988). Clinical and biochemical manifestations of depression: relation to the neurobiology of stress (2). *New England Journal of Medicine* 319, 413–420.
- Herman, J. P. and Cullinan, W. E. (1997). Neurocircuitry of stress: central control of the hypothalamo–pituitary–adrenocortical axis. *Trends in Neuroscience* 20, 78–84.
- Michelson, D., Stratakis, C., Hill, L., Reynolds, J., Galliven, E., Chrousos, G. P. and Gold, P. W. (1996). Bone mineral density in women with depression. *New England Journal of Medicine* 335, 1176–1181.
- Sapolsky, R. M., Krey, L. C. and McEwen, B. S. (1986). The neuroendocrinology of stress and aging: the glucocorticoid cascade hypothesis. *Endocrine Reviews* 7, 284–301.
- Sternberg, E. M. and Gold, P. W. (1997). The mind-body interaction in disease. Special issue: Mysteries of the mind. *Scientific American* 8–15.
- Sternberg, E. M., Hill, J. M., Chrousos, G. P., Kamilaris, T., Listwak, S. J., Gold, P. W. and Wilder, R. L. (1989). Inflammatory mediator-induced hypothalamic–pituitary–adrenal axis activation is defective in streptococcal cell wall arthritis-susceptible Lewis rats. *Proceedings of the National Academy of Sciences USA* 86, 2374–2378.

Male Partner Violence

M Ingram, N P Yuan and M P Koss
University of Arizona, Tucson, AZ, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M P Koss and M Ingram, volume 2, pp 676–681,

© 2000, Elsevier Inc.

Prevalence
Psychological Impact
Physical Impact
Prevention
Intervention
Conclusion

Glossary

Physical abuse Physical violence directed toward an intimate partner including pushing, slapping, punching, kicking, choking, assault with a weapon, tying down, and restraining;

Psychological abuse

Sexual assault

threats of bodily harm and threats of violence directed at the woman's children; and behaviors liable to cause fear of physical safety such as leaving a woman in a dangerous place, driving fast and recklessly to frighten her, and refusing to help when she is sick or injured.

Abusive behaviors whose damage is primarily emotional including physical and social isolation, extreme jealousy, possessiveness, and removing access to communication or transportation; deprivation, degradation, and humiliation; emotional distress inflicted through threats to harm companion animals or livestock; and intimidation, intense criticizing, insulting, belittling, and ridiculing.

Forced sexual acts or sexual degradation, such as coercing a woman to perform sexual acts against her will; hurting her during sex or assaulting her genitals, including using objects intravaginally, orally, or anally; pursuing sex when she is not fully conscious or otherwise un-

Stalking

able to consent; insisting on sex without protection against pregnancy or sexually transmitted diseases; and forcing her to have sex with others either privately or while being observed by the partner; often part of violent relationships.

A pattern of repeated harassing or threatening behavior including following a woman, appearing at her home or place of business, making harassing phone calls, leaving written messages or objects, and vandalizing her property, even in the absence of threat of serious harm.

Male partner violence is broadly defined to be inclusive of the various types of behaviors that are used to coerce, control, and demean women. According to a report on intimate partner violence by the U.S. Department of Justice, 20% of all nonfatal violence against women in 2001 was committed by current or former spouses or boyfriends, whereas intimate partners committed 3% of nonfatal violence against men. Women are at a greater risk of assault, including rape and homicide, by a husband, ex-husband, boyfriend, or ex-boyfriend than they are by an acquaintance or stranger. Some methodologies generate data showing equal rates of intimate partner violence by men and by women; however, women are more likely to sustain minor or serious injuries. Violence within an abusive relationship is characterized as a recurrent experience when a woman is repeatedly placed in physical danger or controlled by the threat or use of force. Specifically included in these definitions, along with physical assault, are sexual assault, battering, threats, intimidation designed to obtain power and control by inducing fear in the victim, and stalking. The United Nations, the American Medical Association, and the American Psychological Association Presidential Task Force on Violence and the Family have established definitions of violence that recognize a pattern of physical, sexual, and/or psychological acts aimed at harming, intimidating, or coercing someone who is or was involved in an intimate relationship with the perpetrator.

Prevalence

Over the past 25 years, researchers have documented that violence by male intimates is one of the most prevalent and serious threats to the well-being of women in the United States and internationally. The landmark 1998 National Violence Against Women survey (NVAW), consisting of telephone interviews with 16,000 men and women, found that 25% of the women surveyed had been raped and/or physically assaulted in their lifetime by a current or former

spouse, cohabiting partner, or date. The National Crime Victimization Survey (NCVS), an annual general survey of crime, indicated that the rate of intimate partner violence against women dropped by 49% between 1993 and 2001 but that the rate in 2001 was still high. There was an estimated 691,710 nonfatal incidents of violence committed by current or former spouses, boyfriends, or girlfriends, and the victims were primarily women (85%). These estimates, however, must be understood in the context of the limitations of the methods used to obtain them. National estimates based on telephone surveys are likely to be conservative due to the exclusion of the homeless, those too poor to own a telephone, people with limited English, recent immigrants who may be fearful of the government, and women whose partners control their ability to freely receive calls. Household-based samples mean that several living situations are missed in which women may be at high risk for violence including prisons, psychiatric hospitals, military settings, and colleges. These and other differences in methods contribute to substantial variation in estimate size across surveys.

Physical Assault

The NVAW documented that 22% of women were physically assaulted by current or former intimate partners during their lifetime. The majority of assaults were relatively minor, involving pushing, grabbing, and shoving. Few incidents involved more extreme forms of aggression, such as kicking or beating, threatening with a knife or gun, or using a weapon.

Sexual Assault

NVAW have shown that more than 7% of women report being raped by a current or former intimate partner in their lifetime. Those subjected to physical abuse are especially vulnerable to sexual abuse, with 26% to 52% of physically abused women reporting also being raped. Men who are both physically and sexually abusive are consistently found to perpetrate the most violent and severe assaults on their partners, with the majority of victims being repeatedly assaulted during the course of the relationship. Women who are victims of both physical and sexual violence are also at greater risk for abuse during pregnancy as well as at greater risk of homicide.

Psychological Abuse

Emotional abuse is generally concomitant with physical or sexual assault. In a recent study of primary-care patients, all women who were physically abused also reported emotional abuse in that relationship. The Canadian Centre for Justice Statistics National

Survey on Violence Against Women included questions on jealousy, limiting contact with family or friends, belittling and insulting female partners, preventing access to family income, and insisting on knowing a partner's whereabouts at all times. Approximately one-third of ever-married women in Canada reported that their current or previous partner was emotionally abusive in at least one of these five ways. In the majority of cases, emotional abuse was concomitant with physical or sexual abuse; only 18% of women who experienced emotional abuse reported no physical violence by a partner.

Stalking

Stalking by male partners also threatens the health of women. Based on a definition that includes experiencing fear, approximately 5% of women (compared to about 1% of men) have been stalked by a current or former intimate partner in their lifetime, according to the NVAW. Among women who were stalked by a former or current partner, 81% were also physically assaulted and 31% were sexually assaulted by the same man. Stalking frequently escalates among women who have left their abusive partners. The likelihood of being stalked by a male partner increases with the duration of time that the woman has been separated from the abusive relationship.

Risk Factors

Various factors increase the likelihood that a woman will be assaulted by a male partner. Research has identified several demographic risk factors. They include being younger in age; being single, divorced, or separated; having lower income; and having less education. In addition, ethnic minorities report higher rates of male partner violence compared to Whites. The NVAW indicated that American Indians/Alaska Native women had the highest rates, followed by African Americans. White and Hispanic women reported similar rates and Asian/Pacific Islanders reported the lowest rates of victimization by an intimate partner. These differences may reflect, in part, socioeconomic differences. However, researchers have found that differences in risk exist even after controlling for such variables. It is of note that studies consisting of only Native American samples have shown that some tribes have higher rates of partner violence, whereas others do not. Immigrant women experiencing partner violence are also a vulnerable population because they are often isolated in a foreign country, in constant fear of deportation, and feel at the mercy of their spouse to gain legal status. A history of childhood physical abuse, alcohol-related problems, drug use, and certain personality characteristics also increase the likelihood that women will be assaulted

by a male partner. Certain characteristics of the male partner also increase the risk of victimization, including childhood exposure to intimate partner violence, partner alcohol-related problems, and emotionally abusive and controlling behaviors.

Psychological Impact

Victims of male partner violence are subject to physical, emotional, and perhaps sexual violence over an extended period of time. The impact on a woman's mental health is thus not derived solely from specific acts of violence but rather from the stress of living with recurrent, escalating, and unpredictable attacks. Not surprisingly, the severity and frequency of violence is associated with an increase in mental health symptoms, and symptoms have been shown to decrease when a woman is free from the abuse.

Posttraumatic Stress Disorder

Posttraumatic stress disorder (PTSD), as defined in the Diagnostic and Statistical Manual of Mental Disorders, fourth edition, has become the predominant diagnosis used to describe the various symptoms of psychological trauma exhibited by partner violence victims, with intensity of exposure being the primary etiological variable for PTSD. Rates of PTSD among abused women range from 31 to 84%. The occurrence of sexual assault in marriage is associated with greater trauma, more severe depression, and lower self-esteem than is found among victims of physical abuse alone. The most common symptoms of the disorder among abused women are nightmares, intrusive memories of abuse, avoidance of reminders of the abuse, and anxiety symptoms. Victims may also exhibit cognitive symptoms such as difficulty concentrating, problems with memory, and anger.

There is concern within the mental health profession and among victim advocates regarding the constructive use of PTSD as a diagnostic tool. PTSD pathologizes the victim's response to violence rather than focus the source of the trauma. In treating the symptoms of PTSD, the causal role of violence is neglected, placing victims at risk of misdiagnosis, inappropriate medication, and even possible further abuse. In addition, the victim may perceive the diagnosis as evidence of her own inadequacy rather than as condemnation of the violent behavior being perpetrated on her. Research has shown that access to both social support and victim advocacy services leads to a reduction in PTSD symptoms among abused women.

Depression

Prior to the introduction of PTSD, depression was the preferred diagnosis for victims of partner violence.

Today, depression continues to be prevalent among victims of intimate partner violence with rates varying between 39 and 83%. Women are more likely to be diagnosed with either depression or PTSD, and an investigation of the co-occurrence in physically abused women found that only 17% met the criteria of both. Depression is more likely among victims when psychological and sexual abuse occur within a physically violent relationship. Having low levels of social support has also been associated with depressive symptoms among victims.

Although it is imperative to investigate, document, and respond to the impact of trauma on the mental health of victims, the attempt to diagnose a woman trapped in a violent relationship with a mental disorder may be counterproductive.

Other Psychiatric Disorders

Victims of male partner violence are at greater risk of alcohol abuse and illicit drug use than nonabused women. Although partner violence may involve alcohol consumption by the female victim, or by the perpetrator and victim, increased alcohol use and illicit drug use by women frequently follows victimization. Compared to nonvictims, victims are also at higher risk for exhibiting eating disorders and more likely to report attempted suicide.

Physical Impact

Researchers and health-care providers worldwide single out partner violence as a major public health threat to women. The most severe physical outcome is death. Approximately 30% of murders of women that occurred between 1976 and 1996 were committed by male partners. The majority of these women were abused by a partner prior to being killed. Female victims also experience high rates of short-term and long-term health consequences. Children are also often affected because of prenatal or postnatal exposure to violence. Therefore, it is not surprising that women with a history of partner abuse demonstrate increased health-care use. Female victims of partner violence seek care for violence-related injuries, general health problems, and mental health problems.

Acute Injuries

Male partner violence is the largest single cause of injury to women requiring emergency medical treatment. Estimates of emergency department visits by women due to partner assault are between 13 and 25%. In a Bureau of Justice Statistics report on violence-related injuries treated in hospital emergency departments, only 14% of the injuries treated in

women were inflicted by strangers compared to 29% in men. Physically abused women experience injuries to multiple parts of the body, including injuries to the face, head, neck, throat, breast, chest, and abdomen. Battered women are more likely to have abrasions and contusions or pains for which no physiological cause can be discovered. Fractures, dislocations, and lacerations occur with about equal frequency in abused and nonabused women. One of the most troublesome aspects of male partner violence is the recurrent and often escalating nature of the abuse. Women who frequently present for medical treatment of trauma are likely to have been treated previously for injuries from abuse. Victims of partner violence have on average seven treatment episodes in the emergency service compared to two among nonbattered women. The pattern of escalation is also evident in homicide rates. In one study, 14% of the women who were killed had been in the health-care system for violence-related injuries. Approximately 47% had received health-care services for some type of problem during the year before they were killed.

Chronic Health Problems

Women of all ages with a history of past or current partner abuse are disproportionately represented among women receiving a variety of services in primary-care settings. The injuries, emotional stress, and fear associated with male partner violence can contribute to the development of chronic health problems. They include chronic pain (e.g., headaches and back pain) and central nervous system symptoms, such as fainting and seizures. Women who have been abused report higher rates of gastrointestinal symptoms and functional gastrointestinal disorders, such as irritable bowel syndrome and nonulcer dyspepsia, compared to other women. These disorders have no known pathogenesis or established treatment and are the most common chronic gastrointestinal conditions seen in primary care, as well as the most common reason for referral to gastroenterologists. In various clinical samples, between 44 and 67% of all female gastroenterology patients report a history of sexual violation, and between 32 and 48% of female patients had experienced severe physical trauma in childhood or adulthood. Gynecological problems are also long-lasting outcomes of male partner violence. Symptoms and conditions include HIV infection, other sexually-transmitted diseases, vaginal bleeding or infections, decreased sexual desire, genital irritation, painful intercourse, chronic pelvic pain, and urinary tract infections. The likelihood of having a gynecological problem is three times greater among women who are abused by spouses.

Birth Outcomes

Because young women in their most active childbearing years are at highest risk for abuse, abuse during pregnancy deserves special mention. Women who are abused and pregnant are more likely to engage in poor maternal health behavior. They frequently delay prenatal care. Compared to nonabused women, women who are abused are twice as likely to delay prenatal care until the third trimester. Pregnant women who are abused report higher levels of alcohol, tobacco, and drug use than other pregnant women. Pregnancy complications may include preterm labor, low birth weight, spontaneous abortion, anemia, and infections. It is also important to mention the effects of intimate partner violence on children. Children who observe violence against their mothers are at higher risk of emotional, behavioral, physiological, cognitive, and social problems.

Prevention

Primary Prevention

The aim of primary prevention, as applied in public health, is the reduction of the risk of obtaining a disease or health condition. Specific to male partner violence, primary prevention focuses on learning not to use violence. Primary prevention strategies include changing social norms promoting violence against women, changing male roles, promoting the status of women, and increasing the public's awareness of the problem of male partner violence. Other approaches focus on training to improve relationship, conflict resolution, and anger management skills.

Secondary Prevention

Secondary prevention occurs when an individual has a disease or health condition, but health consequences have not yet developed. One approach for partner violence is to conduct violence screening in health-care settings and non-health-care settings, such as workplaces, schools, and military bases. Intimate partner violence screening facilitates early identification and treatment referrals for female victims. Specific to perpetrators, secondary prevention efforts include training to deescalate aggression and counseling and hotlines for violent partners.

Intervention

The response to male partner violence must include effective interventions, not only in specific cases of victimization but in communities and from a societal perspective. Beyond providing victim services, communities need to develop a coordinated response

across different types of agencies, and public policy needs to counter the systemic issues related to male partner violence.

Mental Health Services

Mental health clinicians have the potential to play a vital role in assisting the recovery process for survivors of partner violence. However, to avoid retraumatizing the survivors, clinical services must focus on providing specific interventions to counter the impact of trauma rather than on the mental disorders of the victim. For a woman living in a violent situation, mental health practitioners should prioritize her physical safety and assisting her to access victim services as a first step in addressing mental health concerns. Key to this process is assisting women to identify and develop formal and informal support networks that will help facilitate situational realities associated with both living in or leaving a violent relationship. For this to be feasible, strong linkages need to be fostered between clinical services and women's services, such as advocacy centers, domestic violence shelters, and rape crisis organizations. Additional resources are also needed to ensure that women who are using victim services, for example, those living in shelters, are able to access appropriate clinical services.

Coordinated Community Response

Mental health services should also be addressed within the framework of implementing a coordinated community response to partner violence. The development of coordinated community response is an outgrowth of the efforts of women's advocates to make reforms within the often-fragmented criminal justice system. The objective of the coordinated community response is to ensure that components of the system work together to form a consistent and systematic response that shifts responsibility away from victims and holds perpetrators of violence responsible for their behavior. Women's safety is considered of primary importance. Within this framework, emphasis is placed on batterer rehabilitation programs and support and advocacy for victims, but not specifically on the provision of clinical services for women suffering from trauma. Accessibility to immediate, appropriate clinical services could serve to avert the escalation of trauma symptoms.

Public Policy and Advocacy

Examples of proactive public policy are demonstrated in the Violence Against Women Act (VAWA) of 1994, which sponsored legislation to support increased police personnel, higher levels of prosecution, more

severe penalties for perpetrators, and extended victim services. VAWA also recognized that immigration law placed abused women at risk and provided relief by giving abused immigrant women the means to establish lawful permanent residence without the help of a spouse. VAWA 2000 strengthened protections for battered immigrant women, sexual assault survivors, and victims of dating violence. Appropriate interpretation and enforcement of mandatory arrest laws is another area of public-policy intervention that should be pursued. Coupled with an advocacy agenda is the need for a public health response. This perspective, emphasizing the establishment of screening protocol, early interventions, and multiple approaches to prevention, offers hope of reducing the incidence of partner violence and minimizing its effects.

Conclusion

Women's Advocates Incorporated calculated in 2002 that intimate partner violence costs the U.S. economy \$12.6 billion on an annual basis in legal services, direct medical care, policing, lost productivity, and victim services. Efforts have been made in the past decade to address this issue through initiatives such as mandatory arrest and sentencing laws within the criminal justice system and the development and implementation of screening protocols by health and social service organizations. However, the fact that thousands of women are injured or killed each year by an intimate partner indicates the need for social norms to be addressed. There is much to be done in terms of creating a comprehensive systematic response to male partner violence both in the United States and internationally, and the lack of such response contributes to the perception of societal support or even encouragement of violence. Although victims and perpetrators must both have access to appropriate intervention, the prevention of male partner violence is of paramount importance and requires action on a multiple levels.

See Also the Following Articles

Aggressive Behavior; Crime Victims; Familial Patterns of Stress; Marital Conflict; Posttraumatic Stress Disorder – Clinical; Pregnancy, Maternal and Perinatal Stress, Effects of; Sexual Assault; Violence.

Further Reading

- Campbell, J. C. (2002). Health consequences of intimate partner violence. *Lancet* 359, 1331–1336.
- Coker, A. L. (2004). Primary prevention of intimate partner violence for women's health: a response to Plichta. *Journal of Interpersonal Violence* 19, 1324–1334.
- Crowell, N. A. and Burgess, A. W. (eds.) (1996). *Understanding violence against women, Panel on Research on Violence against Women, Committee on Law and Justice, Commission on Behavioral and Social Sciences and Education, National Research Council*, Washington, DC: National Academy Press.
- Field, C. A. and Caetano, R. (2004). Ethnic differences in intimate partner violence in the U.S. general population: the role of alcohol use and socioeconomic status. *Trauma, Violence, & Abuse* 5, 303–317.
- Humphreys, C. and Stephens, J. (2004). Domestic violence and the politics of trauma. *Women's Studies International Forum* 27(5–6), 449–570.
- Koss, M. P., Goodman, L. A., Browne, A., Fitzgerald, L. F., Keita, G. P. and Russo, N. F. (1994). *No safe haven: violence against women at home, work, and in the community*. Washington, DC: American Psychological Association Press.
- Koss, M. P., Bailey, J. A., Yuan, N. P., et al. (2003). Depression and PTSD in survivors of male violence: research and training initiatives to facilitate recovery. *Psychology of Women Quarterly* 27, 130–142.
- Max, W., Rice, D. P., Finkelstein, E., et al. (2004). The economic toll of intimate partner violence against women in the United States. *Violence & Victims* 19, 259–272.
- Plichta, S. B. (2004). Intimate partner violence and physical health consequences: policy and practice implications. *Journal of Interpersonal Violence* 19, 1296–1323.
- Rennison, C. M. (2003). *Intimate partner violence, 1993–2001*. Washington, DC: Bureau of Justice Statistics, Crime Data Brief, National Institute of Justice.
- Tjaden, P. and Thoennes, N. (2000). *Extent, nature, and consequences of intimate partner violence*. Washington, DC: Office of Justice Programs, National Institute of Justice, U.S. Department of Justice.
- Violence Against Women Act of 1994. (1994). Pub. L. No. 103–322, Title IV, 108 Stat. 1902. Washington, DC: Government Printing Office.
- Waters, H., Hyder, A., Rajkotia, Y., Basu, S., Rehwinkel, J. and Butchart, A. (2004). *The economic dimensions of interpersonal violence*. Geneva: World Health Organization.

Marital Conflict

P T McFarland and A Christensen

University of California, Los Angeles, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P T McFarland and A Christensen, volume 2, pp 682–684, © 2000, Elsevier Inc.

What Is Conflict?

What Happens During Conflict?

What Are the Consequences of Conflict?

Glossary

Demand–withdraw interaction

A pattern of interaction in which one member of a couple attempts to engage in the discussion of relationship issues and pressures, nags, demands, or criticizes the other while the other member attempts to avoid such discussions and becomes silent, withdraws, or acts defensively.

Process of conflict

The interaction that takes place around a conflict of interest.

Psychoneuro-immunology

The study of the interactions among the psychological, neurological, and immunological systems.

Structure of conflict

The conflict of interest between spouses; the incompatibility of needs, desires, or preferences that characterize a couple's struggle.

What Is Conflict?

Conflict in marital relationships is normal and inevitable. Indeed, conflict occurs to a much greater extent in marriage than in any other long-term relationship. Although some theorists have suggested that conflict may serve constructive or functional purposes, such as allowing for the disclosure of feelings and the creative resolution of problems, the focus in the marital literature has generally been on the destructive or dysfunctional consequences of conflict.

An important distinction is made in the marital literature between the structure of conflict and the process of conflict. The structure of conflict refers to the conflict of interest between the spouses, that is, the incompatibility of needs, desires, or preferences that characterize the couple's struggle. The process of conflict refers to the actual interaction that occurs between the spouses regarding the conflict of interest.

It is important to note that a particular conflict of interest may give rise to a variety of conflictual processes. Couples may engage in overt conflict about the issue, or they may choose to avoid discussion about the issue altogether. Some theorists suggest that the decision to engage or to avoid is determined by spouses' beliefs about the likelihood that their efforts at resolving the conflict will be successful. Others have suggested that factors such as commitment to the relationship and negative feelings about the partner may also influence engagement or avoidance in conflict.

What Happens During Conflict?

Assuming that the management of conflict affects the quality and stability of the relationship, marital researchers have focused more of their attention on the process of conflict than on the structure of conflict. Researchers often study marital conflict by having couples discuss a conflict of interest in a laboratory setting for a predetermined amount of time (typically 10–15 min) while the resulting interaction is audiotaped or videotaped. Trained observers then code the problem-solving interaction for categories or dimensions of interest. Although the same type of discussion leads to more intense negative affect at home than it does in the laboratory, the observational research paradigm described does produce interactions that are similar to those experienced by couples in naturalistic settings.

One of the objectives of marital research has been to identify conflict behaviors that are associated with relationship satisfaction and thus differentiate between distressed and nondistressed couples. Although a number of differences have been identified, there are a few consistent findings across studies. First, distressed couples display higher rates of negative behaviors during conflictual interactions than nondistressed couples. These negative behaviors include criticism, complaints, hostility, defensiveness, denial of responsibility, and withdrawal. Second, compared to nondistressed spouses, distressed spouses exhibit greater reciprocity of negative behavior. That is, when a distressed spouse behaves in a negative manner, the other spouse tends to respond in kind, which escalates the conflict and leads to a cycle of negativity that is difficult to break. Third, nondistressed couples display higher rates of positive behaviors, such as approval and humor, during conflictual interactions than distressed couples. Similar to this cross-sectional

research, longitudinal research has also demonstrated the deleterious effect of negative interactions on future satisfaction and stability.

A pattern of conflictual interaction that has received much empirical attention and has been found to be highly correlated with marital dissatisfaction is the demand–withdraw interaction pattern. In this pattern, one spouse (the demander) attempts to engage in the discussion of relationship issues and pressures, nags, demands, or criticizes the other while the other spouse (the withdrawer) attempts to avoid such discussions and becomes silent, withdraws, or acts defensively. Although both self-report and observational studies have demonstrated a gender linkage in roles in this pattern, with wives generally in the role of demander and husbands generally in the role of withdrawer, this gender effect is moderated by the specific issue that is being discussed. Specifically, the gender linkage in roles is apparent during discussions in which wives' issues are being addressed, but this linkage is not apparent during discussions in which husbands' issues are being addressed. Also, the pattern often occurs when a structural conflict of interest exists between partners whereby the partners have a differential investment in change on an issue. The spouse who desires a change that can only be achieved through the cooperation of the other (e.g., spending more time together) will generally take on a demanding role during discussion, whereas the spouse who seeks no change or change that can be achieved unilaterally (e.g., spending less time together) will generally take on the withdrawing role because discussion will only create an argument or undesired change.

In addition to studying overt behavior, marital researchers have also examined covert factors that may impact marital conflict and satisfaction, such as affective processes and perceptual differences between spouses. One area that has received much attention in the marital literature concerns the attributions that spouses make in accounting for events that occur in marriage. There is accumulating evidence indicating that attributions for negative events can promote conflict, such as when negative partner behavior is attributed to the selfish intent of the partner. Such attributions have been found to be related to less effective problem-solving behavior and to higher rates of negative behavior during problem-solving tasks. Moreover, research has demonstrated that attributions for negative partner behavior affect a spouse's own behavior toward the partner.

Recent years have seen a growing emphasis on the application of social psychological theories to the study of marital conflict. Attachment theory is one such theory that has been the focus of increased attention in marital conflict research. Attachment

theory asserts that mental models of the self and other develop in the context of early parent–child interactions and that these models influence the communications and reactions of spouses to partner behaviors. Research findings generally suggest that spouses with secure attachment styles are more likely to compromise when dealing with conflict, whereas those with insecure attachment styles tend to employ less productive approaches to marital conflict.

What Are the Consequences of Conflict?

Conflict not only has an effect on the marital relationship itself, but it can also affect the health of individual partners. A growing body of research documents the association between marital conflict and the physical and mental well-being of spouses. Although married people tend to be healthier than those who are not married, marital conflict has been linked to an increase in physical illnesses such as rheumatoid arthritis and cardiac problems. Marital conflict also increases the risk for mental disorders. The association between marital conflict and depression is well established, as is the association between marital conflict and the physical and psychological abuse of spouses. Some of the most compelling data relating marital conflict to mental disorders are seen in the strong linkage between marital conflict and substance abuse.

The link between marital discord and poorer health outcomes has been systematically explored over the past decade. Negative and hostile behaviors during marital conflict are associated with various physiological effects in spouses, such as elevations in blood pressure and heart rate and changes in endocrine and immune functioning. Across studies, these negative physiological effects are more severe and persistent for wives than for husbands. Research in the field of psychoneuroimmunology suggests that there are negative long-term health consequences for spouses who are unable to physiologically recover from marital conflicts or cannot adapt physiologically to repeated conflicts.

In addition to its negative effects on spouses, marital conflict is related to problems in family functioning, including parenting and sibling relationships. Moreover, research has demonstrated the effects of marital conflict on the parent–child relationship, such as attachment difficulties and parent–child conflict, and to a wide range of adjustment problems in children, including internalizing problems such as depression and anxiety and externalizing problems such as aggression, delinquency, and conduct disorders. Exposure to frequent and severe marital conflict, such as physical aggression, seems to be particularly disturbing for children.

Not only is marital conflict itself a stressor, but other life stressors can affect marital conflict. For example, greater work demands among air traffic controllers and unemployment in blue-collar workers have been shown to be associated with more negative marital interactions. Stressful events may increase conflict by diminishing the capacity of spouses to provide support to one another, even as it increases their need for support. Such events may also affect marital interactions by giving rise to new conflicts of interest or by exacerbating old ones.

See Also the Following Articles

Marital Status and Health Problems; Marriage.

Further Reading

- Aubry, T., Tefft, B. and Kingsbury, N. (1990). Behavioral and psychological consequences of unemployment in blue-collar couples. *Journal of Community Psychology* 18, 99–109.
- Beach, S. R. H., Fincham, F. D. and Katz, J. (1998). Marital therapy in the treatment of depression: toward a third generation of therapy and research. *Clinical Psychology Review* 18, 635–661.
- Bradbury, T. N. and Fincham, F. D. (1992). Attributions and behavior in marital interaction. *Journal of Personality and Social Psychology* 63, 613–628.
- Burman, B. and Margolin, G. (1992). Analysis of the association between marital relationships and health problems: an interactional perspective. *Psychological Bulletin* 112, 39–63.
- Christensen, A. and Heavey, C. L. (1990). Gender and social structure in the demand/withdraw pattern of marital conflict. *Journal of Personality and Social Psychology* 59, 73–81.
- Cummings, E. M. and Davies, P. T. (2002). Effects of marital conflict on children: recent advances and emerging themes in process-oriented research. *Journal of Child Psychology and Psychiatry* 43, 31–63.
- Ewart, C. K., Taylor, C. B., Kraemer, H. C., et al. (1991). High blood pressure and marital discord: not being nasty matters more than being nice. *Health Psychology* 10, 155–163.
- Gottman, J. M. (1979). *Marital interaction: experimental investigations*. New York: Academic Press.
- Grych, J. H. and Fincham, F. D. (1990). Marital conflict and children's adjustment: a cognitive-contextual framework. *Psychological Bulletin* 108, 267–290.
- Halford, W. K. and Markman, H. J. (eds.) (1997). *Clinical handbook of marriage and couples intervention*. London: John Wiley.
- Hazan, C. and Shaver, P. R. (1994). Attachment as an organizational framework for research on close relationships. *Psychological Inquiry* 5, 1–22.
- Kiecolt-Glaser, J. K., Malarkey, W. B., Chee, M. A., et al. (1993). Negative behavior during marital conflict is associated with immunological down-regulation. *Psychosomatic Medicine* 55, 395–409.
- Kiecolt-Glaser, J. K., Glaser, R., Cacioppo, J. T., et al. (1997). Marital conflict in older adults: endocrinological and immunological correlates. *Psychosomatic Medicine* 59, 339–349.
- Repetti, R. L. (1989). Effects of daily workload on subsequent behavior during marital interaction: the roles of social withdrawal and spouse support. *Journal of Personality and Social Psychology* 57, 651–659.
- Robles, T. F. and Kiecolt-Glaser, J. K. (2003). The physiology of marriage: pathways to health. *Physiology and Behavior* 79, 409–416.

Marital Status and Health Problems

I M A Joung

Erasmus University, Rotterdam, Netherlands

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 685–691, © 2000, Elsevier Inc.

Health Differences among Marital Status Groups
Theories about the Explanation
Selection on Health and Determinants of Health
Social Causation, Intermediary Factors, and Biological Pathways

Glossary

<i>Determinants of health</i>	Factors associated with health and illness, such as socioeconomic status and alcohol consumption.
<i>Intermediary factors</i>	Determinants of health through which marital status affects health outcomes.
<i>Unmarried</i>	People who have never been married, are divorced, or are widowed.
<i>Selection</i>	The theory about health differences between marital status groups in which it is assumed that health affects marital status.
<i>Social causation</i>	The theory about health differences between marital status groups in which it is assumed that marital status affects health.

Marital status is associated with all kinds of health outcomes: both subjective health states (illness, e.g., self-perceived health) and objective health states (disease, e.g., clinically diagnosed conditions), both mental and physical health, and both morbidity and mortality. Health differences between marital status groups are generally assumed to result from both an effect of health on marital status (selection) and an effect of marital status on health (social causation). Marital status affects health through several intermediary factors; of these, psychosocial factors, especially psychosocial stress, occupy a central position.

Health Differences among Marital Status Groups

The general pattern for health differences among marital status groups is that married people have the fewest health problems, followed by never-married and widowed people, whereas divorced people have the most health problems. This ranking is found both among men and women; however, the differences among the marital status groups are generally larger among men than among women.

In most studies in which separated people are distinguished as a separate group, it is found that they have as many as or even more health problems than divorced people. However, in many studies on health differences among marital status groups, separated people have not been treated as a separate marital status category but have been grouped with either divorced or married people.

In many Western countries, nonmarital cohabitation is of increasing importance. The few studies that have addressed health differences between married people and people living in consensual unions have found either no health differences or that the health of people in consensual unions compared somewhat unfavorably with that of married people but favorably with that of the unmarried, noncohabitating people. The remainder of this article focuses on health differences among married, never-married, divorced, and widowed people.

Mortality

Mortality differences among marital status groups were described in the nineteenth century. For instance, in 1853 William Farr examined age-specific mortality rates for never-married, married, and widowed people living in France. He found that the mortality rates of married men and women compared favorably to those of never-married and widowed men and women. For example, at the ages 40–50 mortality rates were 17.7, 10.3, and 20.1 per 1000 for never-married, married, and widowed men, respectively.

Since then numerous researchers in many countries have looked at mortality differences among marital status groups. A considerable number of studies have been performed on this subject in the last decades. The results from these studies indicate very consistently that married people have lower mortality rates than unmarried people.

The sizes of the mortality differences between married and unmarried groups and the rankings of the mortality rates of the three unmarried groups have changed over time. For instance, in the Netherlands in the period of 1869–1872, the differences in total mortality between married and never-married women were rather small, and under the age of 40 never-married women actually had lower mortality rates than married women. This was mainly due to the high childbirth mortality rate among married women. In the following decades, the differences in total mortality between married and never-married women increased as, childbirth mortality rate, among others, decreased.

The relationship between marital status and mortality also shows some differences between countries. Hu and Goldman explored mortality differences between marital status groups in 16 industrialized countries for the period of 1950–1980. They found that, for the majority of countries, divorced men experienced the highest mortality rates among men and that, in about half of the countries, the same pattern existed for women. Generally, the mortality rates of divorced men and women were 2–2.5 and 1.5–2 times, respectively, those of their married counterparts. In Japan, however, the mortality rate of never-married men was as high as the rate of divorced men and never-married women had much higher mortality rates than divorced women. Also, mortality rate differences between the never-married and married in Japan far exceeded the largest mortality rate differences by marital status in the other countries (being more than three times that of the married group).

The differences in total mortality rates among marital status groups are more or less mirrored in the differences in rates for specific causes of death. For the majority of causes of death, the largest differences are found between the rates for divorced and married people. In general, never-married and widowed people have higher rates of cause-specific mortality than married people, alternately occupying the position of the group with the second highest mortality rate. In comparison with the differences found for the total mortality rate, the relative size of the differences in the rates of mortality due to cardiovascular diseases and cancer is somewhat smaller, but it is considerably larger for the mortality rates due to external causes of death (e.g., accident and suicide). [Table 1](#) shows the

age-adjusted relative risks of divorced men and women compared to married men and women for total mortality and mortality from a number of specific causes of death.

Morbidity

More recently, differences in morbidity, short-term disability, and health-care use among marital status groups have been studied. Morbidity differences largely have patterns similar to mortality rate differences: married people have the lowest rates, divorced people have the highest rates, and never-married and widowed people have rates in between. The size of the differences among marital status groups is generally larger for indicators of subjective health (e.g., self-perceived general health and subjective health complaints) than for indicators of objective health (e.g., chronic diseases) and is larger for mental than for physical health.

Morbidity differentials are by and large reflected in the patterns for health-care use and short-term

disability. Widowed and divorced people have higher rates for physician contacts, hospital admissions, and use of medicines and report more bed days and days of restricted activity than married people. Never-married people, however, seem to have a lower health-care use and less bed disability days than married people, despite their higher morbidity rates.

Theories about the Explanation

Several explanations have been suggested for the association between marital status and health. First, it has been put forward that the association might be due to data errors and thus be an artifact. Although data errors might explain part of the association between marital status and health found in earlier studies, more recent studies have shown convincingly that there are genuine health differences among marital status groups. Other explanations that have been proposed can be grouped into two main theories: selection theory and social causation theory. According to selection theory, the relatively good health of married people is the result of the selection of healthy people into and unhealthy people out of the married state. According to social causation theory, marriage has a health-promoting or a health-protective effect, whereas the unmarried state has adverse health effects. Thus, in selection theory, health status precedes marital status; in social causation theory, marital status precedes health status. A graphic representation of these explanations is shown in [Figure 1](#). In order to establish whether or to what extent health differences precede differences in marital status or vice versa, longitudinal data are required. Selection theory and social causation theory are not mutually exclusive, and it is generally accepted that a combination of

Table 1 Relative risks of divorced men and women compared to their married counterparts in the Netherlands, 1986–1990^a

	<i>Divorced men</i>	<i>Divorced women</i>
Total mortality	1.6	1.5 ^b
Cardiovascular diseases	1.5	1.4
Cancer	1.2	1.3
External causes	3.8	3.0
Infective diseases	2.2	2.1
Cirrhosis of the liver (with mention of alcohol)	9.1	6.0
Diabetes mellitus	2.2	1.4

^aSource: Statistics Netherlands.

^bFor example, a relative risks of 1.5 means that the mortality rate of divorced women is 1.5 times that of married women.

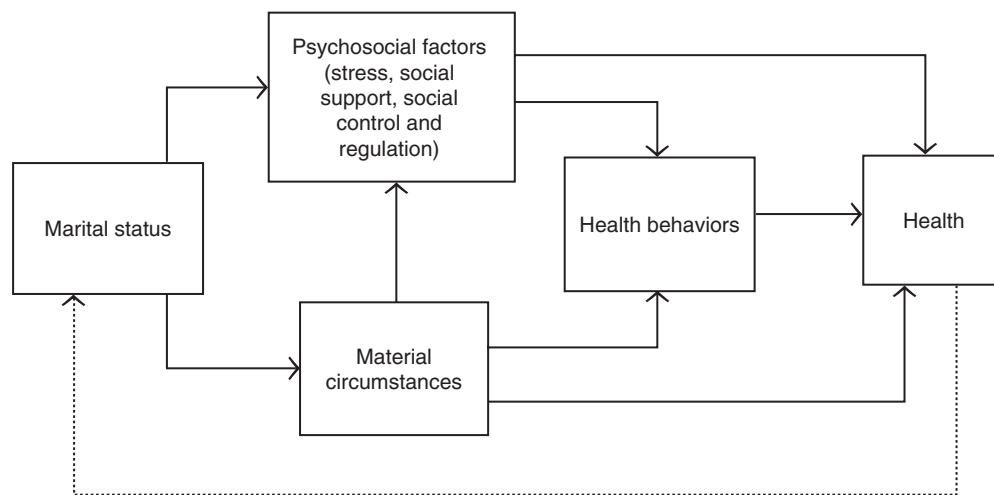


Figure 1 A representation of the explanations for health differences among marital status groups. →, social causation; -.->, selection.

selective and causal factors are responsible for the observed health patterns in marital status groups.

Selection on Health and Determinants of Health

According to selection theory, the relatively good health of married people is the result of the selection of healthy people into and unhealthy people out of the married state, thus increasing the relative amount of unhealthy people in the unmarried states. A distinction can be made between direct selection and indirect selection. In direct selection, health itself is the selection criterion; in indirect selection, the determinants of health are the selection criteria.

The process of selection could occur with regard to first marriages as well as other marital transitions (divorce, bereavement, and remarriage) and could cause the health differences by marital status in several ways. Selection in partner choice is the most straightforward mechanism: unhealthy people may be less attractive marriage partners and thus may either not be chosen or, if illness develops during marriage, might be discarded as marriage partners. Selection may also operate through assortive mating, the fact that people generally tend to marry partners with traits that they themselves possess, such as physical attractiveness. Assortive mating could also include health status and does not so much influence whether a person marries as whom the person marries. If, indeed, unhealthy people are more likely to marry unhealthy others, we could find that unhealthy married people are more likely to become widowed because their unhealthy partner is at greater risk of mortality. In addition, it is conceivable that relationships in which both partners are unhealthy are more stressful and therefore more prone to dissolution. Finally, with regard to the transition from the married to the widowed state, selection may also operate through processes independent of partner choice due to health considerations, for instance, through both spouses developing health problems after marriage for reasons such as a joint unfavorable environment (e.g., shared material deprivation or unhealthy behaviors). In this case, the health differences between married and widowed people are not caused by the conditions of widowhood itself (social causation) but are, instead, based in already existing health differences between those who will become widowed and those who will remain married (selection).

Direct Selection

It has been demonstrated in several longitudinal studies that direct selection is operative in marital

transitions. However, the evidence is still partial and sometimes inconsistent. There are indications that the presence of disease decreases the marriage probabilities of never-married people and increases the divorce probabilities of married people. It has also been found that direct selection might be able to explain a considerable part of the health differences among marital status groups. However, other studies have been unable to demonstrate selection effects, and sometimes even evidence for adverse selection was found (i.e., that healthier people were less likely to marry).

Indirect Selection

As stated earlier, selection may operate not only through the exclusion of unhealthy people from marriage (direct selection) but also through selection on a wide range of determinants of health (indirect selection). The determinants of health that may operate in this way are, for instance, socioeconomic status, physical appearance (e.g., body length and obesity), and health-related habits (e.g., alcohol consumption). Research has demonstrated that indirect selection is indeed operative in marital transitions and that its direction is mostly, but not always, in accordance with the health differences observed among marital status groups. Adverse selection occurs, for instance, with regard to educational level among women – people with a higher educational level have more favorable health outcomes than those with a lower educational level, but women with a high educational level are overrepresented among never-married and divorced women.

Social Causation, Intermediary Factors, and Biological Pathways

According to social causation theory, health depends on marital status. Marital status may, on the one hand, affect the etiology of health problems: the married state may prevent people from becoming ill, whereas the unmarried state may be the cause of a decline in resistance to diseases. Marital status may, on the other hand, once health problems have developed, affect the course and outcome of the disease. Ample evidence from longitudinal studies shows that differences in marital status are associated with subsequent health differences; thus social causation mechanisms explain (part of the) health differences among marital status groups. The effect of marital status on health is not assumed to be a direct effect but to be intermediated by psychosocial factors (e.g., stress and social relationships), material circumstances (e.g., financial situation and housing conditions), and health behaviors (e.g., smoking and alcohol consumption).

Psychosocial Factors

Psychosocial stress is related causally to illness and mortality. Lazarus and Folkman defined psychosocial stress as “a particular relationship between the person and the environment that is appraised by the individual as taxing or exceeding his or her resources and endangering his or her well-being.” Psychosocial stress varies among marital status groups. First, bereavement and divorce are stressful life events in themselves. On the social readjustment rating scale of Holmes and Rahe, a scale of 43 life events ordered on the basis of the assumed intensity and length of time necessary to accommodate to the life event, bereavement and divorce rank first and second, respectively. Understandingly, the loss of a beloved person is an important source of stress itself. Feelings of failure, lowered self-esteem, and sense of incompetence, which are often experienced by divorced people, also evoke stress. Furthermore, the many concurrent changes in the lives of bereaved or divorced people, such as lowered income, change in parental responsibilities, forced move to other housing, or the loss of familiar activities and habit systems, also contribute highly to the total amount of stress experienced.

Second, the differences in psychosocial stress among marital status groups are also caused by mechanisms other than the stressful character of the event of bereavement or divorce. Negative societal attitudes toward the marital status a person occupies can be a source of stress. Although societal attitudes toward alternatives to marriage have become more liberal in recent years, prejudices and stereotyped images of never-married people and divorced people still exist. Uncertainty about social roles is also a source of stress, and certainty about social roles differs according to marital status. Marriage provides people with clear and socially acceptable roles, whereas divorce lacks clearly defined norms.

Other psychosocial factors may also contribute to the effect of marital status on health. Several aspects of social relationships are associated with health status and differ among the marital status groups. First, social integration (the existence and quantity of social relationships) is related directly to health. Psychological and sociobiological theories suggest that the mere presence of, or sense of relatedness with, another organism may promote health through relatively direct motivational, emotional, or neuroendocrinal effects. Married people are, on average, more socially integrated for several reasons: they have at least one social tie in their spouse; their children constitute additional social ties, which most never-married people lack; and their social network is expanded with the social ties of their spouse.

Consequently, the loss of a spouse also means a break in the social network.

Second, social support (the functional nature or quality of social relationships) has been found to be related directly and indirectly to health. With regard to the latter, social support is assumed to buffer the negative health effects of stress, thus modifying the relationship between stress and health. In this case, social support has no beneficial effect on the health of people who experience little stress, but the beneficial effects of social support increase with increasing stress. Social support is the aid that is transmitted between social network members. A distinction is often made between emotional and instrumental support. Emotional support includes expressions of affection, admiration, respect, or affirmation; instrumental support is the provision of advice, information, or more practical assistance. The availability and quality of social support differ by marital status. Partner relationships are, in general, more supportive than are other types of relationships. This also means that in cases of bereavement or divorce an important provider of social support is lost.

Finally, differences in social relationships among marital status groups result in differences in social regulation and control. Married people experience more social control; wives, in particular, try to influence their spouses' health behaviors. Furthermore, married people have a more regulated life, which facilitates health-promoting behaviors such as proper sleep, diet, and exercise; the moderate use of alcohol; the adherence to medical regiment; and seeking appropriate medical care.

Material Circumstances

Differences in material resources among marital status groups constitute another intermediary of the relationship between marital status and health. People who share a household profit from economies of scale in the purchasing and use of housing and other goods and services. In addition, changes in marital status are often associated with large changes in material resources. Bereavement and divorce are, especially among women, often accompanied by a decline of their financial situation and, consequently, a deterioration in other structural living circumstances (e.g., housing). This is most obvious in situations in which the husband is the sole wage earner. However, even in cases of double incomes the wife is likely to be worse off financially after divorce or bereavement. Women are generally married to men of a higher or equivalent educational level, and in many countries men earn on average more than women of an equivalent educational level earn. The fact that the material

situation of women generally deteriorates after a divorce does not necessarily imply that men gain materially from divorce. In a divorce, possessions are divided between the former spouses, the ex-husband might be obliged to pay alimony, and economies of scale are lost. This might put divorced men in a materially disadvantaged position compared to married men. Widowed women are generally better off financially than divorced women, even with comparable levels of household income, because divorced women lose their house and must divide assets with their ex-husband, whereas widowed women more often keep the family home and other financial assets intact.

Health Behaviors

Differences in health behaviors among marital status groups are also seen as an intermediary of the association between marital status and health. Marriage often has a deterrent effect on negative health behaviors, such as smoking, excessive alcohol consumption, substance use, and other risk-taking behavior, and marriage promotes an orderly lifestyle. Differences in sexual and reproductive behavior among the marital status groups have been mentioned as explanations for differences in mortality rates from several malignant neoplasms – the excessive mortality rate among never-married women from breast cancer, uterine cancer, and ovarian cancer; the higher mortality rate among divorced women from cervical cancer; and the lower mortality rate among never-married men from prostate cancer.

Finally, differences in the use of health services among marital groups have been mentioned (e.g., married people are more inclined to use preventive health services) and in compliance with required prolonged treatment (married people are more willing to undertake the required treatments for diabetes mellitus and tuberculosis).

Interrelationships between Intermediary Factors

Psychosocial factors, material circumstances, and health behaviors have been mentioned as possible intermediary factors in the effects of marriage on health. However, the effects of marriage on health through intermediary factors cannot be viewed as three independent pathways. Several interrelationships could exist among the three intermediary factors. In the conceptual model of these interrelationships is shown in [Figure 1](#), marital status is not assumed to have a direct effect on health behavior but to influence health behavior only indirectly through psychosocial factors and material conditions.

Psychosocial factors may have direct health effects or may operate through an effect on health behaviors.

Unmarried people experience higher levels of psychosocial stress. Cigarette smoking and alcohol consumption are palliative coping responses to psychosocial stress. In addition, social support is important in health behavior changes; partner support, for instance, is beneficial to smoking-cessation maintenance. Finally, the more regulated life of married people facilitates healthy behaviors, and married people attempt to influence the health behaviors of their spouses.

Also, the material circumstances associated with marital status may have direct health effects, increase psychosocial stress, or operate through changes in health behaviors. With regard to the latter, unhealthy eating habits and a lack of recreation possibilities could, in part, be determined by an individual's financial position.

Evidence from several studies shows that psychosocial factors, material circumstances, and health behaviors act separately as intermediary factors in the relationship between marital status and health. The relative importance of these three groups of intermediary factors has been addressed only recently. Tentative results suggest that psychosocial factors are the most important intermediary factor in health differences between marital status groups among men, whereas material circumstances constitute the major intermediary factor of health differences among women.

Pathways of Intermediary Factors to Health

The effects of health behaviors such as smoking, alcohol consumption, obesity, and lack of physical activity on health are relatively well known. For instance, smoking is, in the long term, associated with cardiovascular diseases, chronic obstructive lung diseases, and many malignant neoplasms such as cancer of the lung, bronchus, trachea, larynx, pancreas, and bladder. Excessive alcohol consumption is associated with diseases of the liver, stomach, and central nervous system and is related to external causes of death, such as traffic accidents and suicide. Obesity and lack of physical exercise are both associated with cardiovascular diseases and conditions of the locomotor system; obesity is also associated with the development of non-insulin-dependent diabetes mellitus.

There is also a considerable amount of research on the relationship between psychosocial stress and health. Stress may cause direct physiological changes in the endocrine, immune, and autonomic nervous systems. Evidence shows that these physiological changes increase susceptibility to infectious diseases, cancer, and cardiovascular diseases. It has also been suggested that marital status, mediated by psychosocial stress, enhances susceptibility to diseases in

general rather than having specific etiological effects. This theory of generalized susceptibility might explain why marital status is associated with many apparently different causes of disease and death.

The direct links between social relationships and health are more speculative. Evidence suggests that biological and psychological mechanisms are involved. A variety of studies of animals and humans suggest that the mere presence of, and especially affectionate physical contact with, another similar or nonthreatening organism reduces cardiovascular and other forms of physiological reactivity. The psychological mechanisms are related to, but partially independent of, the biological mechanisms and might, for instance, be affective in nature (if there is a basic human need for relationships or attachments, people will feel better psychologically when that need is fulfilled, with physiological consequences).

The direct health effects of material circumstances (i.e., health effects not mediated by psychosocial stress or health behaviors) are also of a more speculative nature. In previous centuries, poverty might have been the cause of starvation, death from hypothermia, or increased susceptibility to infectious diseases because of undernourishment. However, the contribution of these causes of death to overall mortality are minimal in current Western societies. However, it is conceivable that unfavorable material circumstances still increase risks for respiratory infections or the transmission of infectious agents through effects on living conditions (damp houses, presence of fungal mold, and crowding). The results of studies on this subject are inconclusive, however. The health effects of material circumstances are presumably mediated mainly by changes in psychosocial stress and, to a lesser extent, by changes in health behaviors.

A final comment with regard to the social causation theory is that we should keep in mind that marital relationships are not always health enhancing. Marital relationships may also have negative health effects,

whereas divorce may have positive health effects. Marital problems are a source of stress, and social control in an environment in which hazardous health behaviors are valued could have unfavorable consequences. With regard to the positive health effects of divorce, it has, for instance, been found that people improve their health after divorce, depending on the stress of the marriage. In addition, several studies have reported that unhappily married people are less healthy than divorced people and happy married people. On the average, however, marriage is associated favorably with intermediary factors and with health.

See Also the Following Articles

Bereavement; Economic Factors and Stress; Familial Patterns of Stress; Gender and Stress; Marital Conflict; Psychosocial Factors and Stress; Social Support.

Further Reading

- Burman, B. and Margolin, G. (1992). Analysis of the association between marital relationships and health problems: an interactional perspective. *Psychological Bulletin* 112, 39–63.
- Joung, I. M. A. (1996). *Marital status and health: descriptive and explanatory studies*. Alblaserdam, Netherlands: Haveka.
- Joung, J. M. A., van de Mheen, H., Stronks, K., et al. (1998). A longitudinal study of health selection in marital transitions. *Social Science & Medicine* 46, 425–435.
- Joung, I. M. A., Stronks, K., van de Mheen, H., et al. (1997). Contribution of psychosocial factors, material circumstances and health behaviours to marital status differences in self-reported health. *Journal of Marriage and the Family* 59, 476–490.
- Stroebe, W. and Stroebe, M. S. (1987). *Bereavement and health*. New York: Cambridge University Press.
- Waldron, I., Weis, C. C. and Hughes, M. E. (1997). Marital status effects on health: are there differences between never married women and divorced and separated women? *Social Science & Medicine* 45, 1387–1397.

Marriage

A C Yoneda and J Davila

State University of New York at Stony Brook,
Stony Brook, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by J Davila, volume 2, pp 692–698, © 2000, Elsevier Inc.

The Relation between Stress and Marriage
Major Models of Stress and Marital Functioning
Specific Types of Stressors Associated with Marital
Functioning
Methodological and Measurement Issues

Glossary

<i>ABCX and double ABCX models</i>	Models of family adaptation that suggest that successful adaptation is a function of the nature of the stressor, the family's resources for coping with the stressor, how they appraise the stressor, and the extent to which the family has experienced an accumulation of stress over time.
<i>Adaptive strategies</i>	Behaviors that spouses and couples use to negotiate and cope with problems or stressors; include skills in communication, problem solving, and provision of social support.
<i>Infidelity</i>	Sexual or emotional unfaithfulness to a spouse.
<i>Infertility</i>	Persistent inability to conceive a child for at least 1 year.
<i>Marital functioning</i>	A broad term that includes how spouses and couples think about, feel about, and behave in their marriage.
<i>Marital stress generation model</i>	A model, adapted from research on unipolar depression, that was designed to explain how marital stressors are created by troubled spouses and then negatively impact the subsequent individual well-being of spouses.
<i>Marriage</i>	The legal union of a man and woman as husband and wife.
<i>Vulnerability stress adaptation (VSA) model</i>	A model of marital outcomes that suggests that marital dissatisfaction results from the interplay among spouses' individual vulnerabilities, stressors the couple must face, and the skills the spouses and couples have for negotiating and coping with stressors.

This article is based on the definition of marriage in the United States as of February 2005. Because this definition may change in time to encompass more

diverse sets of couples, the types of stressors faced by married couples may also change. Marriage is currently defined as the legal union of a man and woman as husband and wife. Recent U.S. census data indicate that approximately 80% of all Americans will marry at some point in their lives. However, more than 50% of all first marriages will end in divorce, and subsequent remarriages are even more likely to end in divorce. Relationship problems are one of the most common reasons why people in the United States seek counseling, and the development of prevention and intervention programs to combat marital distress and dissolution is proceeding rapidly.

The Relation between Stress and Marriage

Most people agree that stress plays a role in marital functioning, and there is clear evidence that stressful events and circumstances can have a negative impact on marriages. One of the difficulties in this field of study arises from problems inherent in defining the relations between stress and marriage. One distinction that must be drawn is whether the stressors are internal or external to the marriage. An external stressor is something that originates outside the marriage, such as unemployment or health problems. An internal stressor is something that originates inside the marriage, such as marital conflict or divergent needs and desires. A second distinction that must be drawn is between effects of stressors on marriage and effects of marriage on stressors. The effects of stressors on marriage are how stressors, either external or internal, influence marital functioning and the quality and stability of the marriage. The effects of marriage on stressors are how marital functioning affects the experience of stressors, again either external or internal. The experience of stressors, in this case, is how well an individual copes with stressors and the impact the stressor has on the individual in terms of his or her own well-being. The field is still expanding its understanding of the bidirectional relationship between stress and marriage, and advances in this domain are likely to follow in the coming years.

Major Models of Stress and Marital Functioning

There are presently three major models that were designed explicitly to explain the relations between stress and marriage.

The Vulnerability-Stress-Adaptation Model of Marital Outcomes

The VSA model was developed to understand precursors to marital dissatisfaction and marital dissolution. Prior to the development of this model, the study of marital outcomes was dominated by a behavioral and social-learning perspective on marriage, which maintains that marital dysfunction arises because couples do not have adequate skills to negotiate their problems. However, many felt that this perspective was limited, particularly because it failed to focus on the individual qualities of spouses and on stressful life experiences that might affect spouses' abilities to interact successfully with their partners. Hence, the VSA model was developed as an integrative perspective that incorporates three factors: enduring vulnerabilities (e.g., individual characteristics such as personality, psychological symptoms, and family of origin experiences) that spouses bring to the marriage, stressful events (e.g., external stressors) that individuals and couples encounter and experience, and adaptive processes (e.g., skills that couples have for communicating, problem solving, and providing support to one another) that may help the couple negotiate and cope with individual and marital difficulties and stressors.

The VSA model suggests that marital dissatisfaction results from the interplay of these three factors, and most directly from repeated experiences with poor adaptation (e.g., poor communication and social support). Once a couple is dissatisfied, it is then at risk for marital dissolution. Stressful life events in the VSA model are seen as factors that impact or interact with a couple's capacity for successful problem negotiation or social support in predicting marital dissatisfaction. For example, stressful life events will lead to dissatisfaction when a couple does not have the adaptive skills to successfully negotiate or cope with the stressor. Stressful life events might also erode a couple's capacity for successful coping (i.e., successful adaptation) if the events are severe enough. The VSA model also suggests that stressful life events can be created or exacerbated by spouses' enduring vulnerabilities and the inability to successfully negotiate problems. For example, people with certain personalities might be particularly prone to generating stress, and couples with poor communication skills may inadvertently contribute to the stressful events with which they then must contend. Hence, in this model, stressful life events play a very important role in marital dissatisfaction and dissolution.

Research largely supports the hypothesis that good adaptive processes moderate the impact of life stressors on marital functioning. This is particularly true

for the moderating role of social support. Some research also indicates that problem-solving behaviors moderate the association between stressors and marital functioning. For example, Cohan and Bradbury found that levels of stressful life events were associated with emotions and behaviors displayed while couples were discussing marital problems. Furthermore, the association between life events and later marital satisfaction was moderated by problem-solving behavior. Higher levels of life events were associated with lower levels of marital satisfaction in couples when, for example, wives did not effectively express their anger about marital problems.

The ABCX and Double ABCX Models of Family Stress

Hill's 1949 ABCX model of family adaptation has largely guided family stress theory. His theory was based on his study of servicemen separated from their families during World War II, but the model is sufficiently general to be considered applicable to all families and marriages. Hill suggested that stressors have the potential to produce change in the family system and that family adaptation differs in response to stressors. He proposed that family adaption (X) was a function of the nature of the stressor (A), the family's resources (B), and the family's appraisal of the stressor (C). Regarding the nature of the stressor, the stressfulness of an event depends on exactly what the nature of the event is – some events are inherently more stressful than others. Family resources are similar to adaptive processes and enduring vulnerabilities in the VSA model. For example, family resources may be material, such as income; intrapersonal, such as education or family of origin experiences; or interpersonal, such as communication skills or family cohesion. More resources should minimize, and fewer resources should maximize, the negative consequences of stressful life events. The family's appraisal of the stressor determines how the family views the stressful life event. Families who view the event as a threat should fare worse than families who view the event as a challenge.

The Double ABCX model expands the ABCX model by including a mechanism to explain the development of family adaptation or distress over time. That is, whereas the ABCX model focuses on adaptation at one point in time in response to a concurrent stressor, the Double ABCX model includes the accumulation of stress over time as a predictor of longer-term adaptation. The notion is that family crises do not occur in isolation. They occur in conjunction with other preexisting stressful circumstances, and they initiate new stressful circumstances. Thus, the Double

ABCX model encompasses the accumulation of stressful circumstances (aA), existing resources and additional resources developed by responses to earlier stressors (bB), appraisals of the original stressor and the aftermath of the family's attempt to deal with it (cC), and longer-term family adaptation (xX). These two models have gained much empirical support in studies of families with diverse stressors, including families coping with wartime separations and families coping with children who have mental disorders. These models have also been used to inform the assessment and treatment of children and families facing severe life stressors. Although not explicitly designed to understand stress in marriage, these models have clear applications to marital outcomes in that when successful adaptation is not achieved, marital distress and dissolution may follow.

The Marital Stress Generation Model

The marital stress generation model was designed to explain how marital stressors (internal stressors) are created by spouses and how marital stressors impact the individual well-being of spouses. Hence, it differs from the other two models that focus on external stressors and how couples and families cope with them. The marital stress generation model proposes that symptomatic spouses (e.g., spouses with depressive symptoms) perceive their partners negatively and are deficient in their capacity to communicate with, negotiate with, or support their partners. These negative perceptions and deficiencies lead spouses to experience more stressors within their marriage, and these stressors subsequently maintain or exacerbate the symptomatic spouse's symptoms.

The marital stress generation model developed out of the stress generation model of unipolar depression, which stated that depressed people contribute to the occurrence of interpersonal stressors, which then contribute to further depression. Research has consistently supported this model and has suggested that a stress-generation process may not be specific to depression. Research on the marital stress generation model supports the model, at least for wives. For example, Davila et al. found that wives with depressive symptoms expect their husbands to be critical of them and then solicit and provide support in a negative manner when interacting with their husbands. This behavior leads to increases in marital stress, which in turn leads to increases in wives' depressive symptoms. Hence, the model describes how marital interactions (e.g., involving the provision and receipt of social support), marital stress and individual well-being can affect one another and keep spouses mired in a cycle of marital dysfunction and distress. In a

similar vein, some researchers hypothesize that in maritally discordant wives uncertainty about the future of the relationship and ongoing concerns about spousal disagreements may elicit and further perpetuate anxiety symptoms, and that reduced positive reinforcement from spousal interactions and helplessness about the ability to change the relationship may elicit and perpetuate depressive symptoms.

Summary

The three models described here address different types of stressors and different segments of the stress-marital functioning association. However, they have at least one significant commonality in that they all focus on the adaptive strategies used by couples to negotiate problems or to support one another. This highlights the fact that the manner in which couples cope with stressors is critical to their ultimate marital outcome.

Specific Types of Stressors Associated with Marital Functioning

A number of specific types of stressors and their association with marital functioning have been studied. Samplings of the major stressors are presented. Not all of these programs of study were conceived of as research on stress and marriage, but, by their nature, the variables they examine are stressful events and fit within the stress paradigms just described. Initial investigations of many of these stressors focused on direct associations between the stressor and marital functioning. Research has moved toward more integrated models, as described, particularly with respect to how the manner in which couples negotiate or cope with stressors affects associations between stress and marital functioning. Although such knowledge is expanding, further research on the mechanisms of the association between stress and marital functioning is still necessary.

Work-Related Stressors and Unemployment

Work-related stressors and unemployment (especially those ensuing financial hardship) are consistently shown to be associated with poorer marital and family functioning. Job loss and unemployment decrease the individual well-being of spouses, decrease marital satisfaction, and increase marital dissolution. Consistent with the models described earlier, research suggests that these negative effects can result from poor adaptation within the family (e.g., increased conflict and poor social support) and from increased depression among spouses. For example, financial strain can lead to increases in depressive symptoms, which lead

to reductions in spousal support and increases in marital distress. Work-related stressors show similar associations. For example, couples have more negative marital interactions on workdays that are reported as more stressful and when daily workload is high. Different types of jobs may also be more or less stressful. For example, shift work or jobs involving separations may be more stressful for some couples than are other types of work. The negative impact of long working hours on marriage is receiving greater attention as the length of the typical work week increases. Related to this, a number of studies have examined differences between blue-collar and white-collar workers facing job stress, and they suggest that blue-collar marriages may be at greater risk than white-collar marriages, possibly because of the different strategies they use to cope with job stress. However, much variance exists within blue- and white-collar jobs. In summary, there is clear evidence that work-related stress, and the manner in which couples negotiate it, can have a serious negative impact on marriage.

Military Service

Military couples and families often deal with a unique set of stressors, such as frequent relocation, relationship maintenance during separation, reunion, and numerous other factors above and beyond those dealt with by nonmilitary couples. A number of studies, many deriving from early tests of the ABCX model of family stress, indicate an association between aspects of military service and marital functioning. For example, wives' perceptions of marital adjustment declined during husbands' military-related absences, and wives and children reported increased emotional distress during absences. In addition, higher levels of marital quality before husbands' military-related absences led to more successful family reintegration at the end of the absence. The higher levels of military life stress are also associated with lower levels of general well-being and global life satisfaction among military wives. Hence, being married to a man in the military and the stressors that come along with that may have negative outcomes for wives, although this is likely to be moderated by family adaptation. Newer lines of research are investigating the role of military women as well as couples in which both partners serve in the military. The implications of such roles for marital and familial relationships are still in an early stage of examination.

Health Problems

A relatively large literature has indicated that health problems are significant marital stressors and that

marital quality has a significant impact on health outcomes. Longitudinal studies indicate that unhappy marriages are associated with morbidity and mortality. Although the relationship between poor marital quality and health problems is bidirectional and thus difficult to disentangle, examining the consequences is nonetheless of paramount importance. Numerous health problems have been studied, including chronic pain; cancer; cardiac disease; pulmonary disorders; renal, endocrine, and gastrointestinal disorders; and neurological disorders. In general, this research indicates that health problems both affect and are affected by marital functioning. Furthermore, social support may play an important stress-buffering role for individuals and couples. Coronary heart disease is one of the most well-researched stressors, and it provides a good example of how health problems and marital functioning can be related. For example, research has indicated that spouses of cardiac patients have higher levels of psychological distress and more marital tension postrecovery. Research has also indicated that spousal support (e.g., intimacy, and emotional and instrumental support) is associated with better outcomes for coronary heart disease patients, including lower mortality rates, whereas marital discord and conflict leads to worse outcomes. The association between coronary heart disease and marital functioning is considered so important that relationship-based interventions for cardiac patients and their partners are now emerging. Of course, different health problems may pose different challenges for individuals and couples. Hence, the associations between health stressors and marital functioning may be different depending on the health problem studied. Abrasive marital interactions appear to have negative consequences on endocrinological and immunological functioning. Studies examining negative behaviors during marital conflict have found that such activities are associated with increased levels of epinephrine, norepinephrine, growth hormone, and adrenocorticotrophic hormone (ACTH), as well as greater immunological change over the following 24 h. Wives appear to be more reactive physiologically to marital conflict than husbands. Longitudinal studies have further supported such findings, with higher levels of stress hormones in newlyweds foreshadowing marital distress and dissolution 10 years later.

Infertility

Infertility is a significant stressor that affects approximately 10% of couples, according to recent estimates. Research suggests that infertility can have numerous negative effects for women and men, including lowered self-esteem, increased depression and

anxiety, and feelings of guilt and shame. Some of these negative effects may be even more prominent among women. For example, studies have found that associations exist between infertile women who desire children (and have no social or biological children) and substantial and significant long-term psychological distress. Clinical and anecdotal evidence suggests that infertility also negatively affects marital satisfaction and marital functioning, particularly sexual satisfaction. The diagnosis, investigation, and/or management of infertility may interact with a couple's sexuality and sexual expression by exacerbating sexual problems, or the sexual problems themselves may be a contributory factor in childlessness. It is certain that couples dealing with infertility are faced with many issues that can strain a marriage beyond sexual difficulties alone, such as financial difficulties (from the high costs of fertility treatments) and reconceptualization of future plans and social roles. However, the empirical data indicate that couples with fertility problems are no more dissatisfied with their relationships than are other couples. In addition, there is evidence that infertility has positive effects on some couples. Successful coping with infertility can result in increased commitment and closeness. Hence, the effects of infertility on marital outcomes may depend on how couples adapt to or cope with the stressor. In line with this, data indicate that couples will fare worse when spouses have incompatible ways of coping with infertility.

Transition to Parenthood

Numerous studies have suggested that marriage suffers during the transition to parenthood. For example, marital satisfaction tends to decline and marital problems and conflict can increase, particularly in wives. However, research that contradicts these findings also exists. Methodological issues are problematic in prior studies showing decline; there are distortions such as the assessment of spouses during the third trimester of pregnancy and the failure to include comparison groups of couples at similar points in marriage who were not having children. Studies that include comparison groups and that assess spouses before, during, and after pregnancy have failed to find differences between the couples who were transitioning to parenthood and those who were not. There is some evidence, however, that marital factors affect which couples will fare better or worse during the transition to parenthood. For example, existing marital problems may lead spouses and couples to experience a more stressful transition, and, to the extent that spouses' child care and housekeeping

expectations for their partners are not met, couples may view their marriages more negatively. Hence, the transition to parenthood may not be equally stressful for all couples, but may instead depend on how couples negotiate the transition.

Psychopathology

Much research indicates that psychopathology is intrinsically linked with marital functioning. Although psychopathology is conceived of as an enduring vulnerability in the VSA model of marital outcomes and as a cause and an outcome of marital stress in the marital stress generation model, it may be reasonable to also consider psychopathology as a stressor, particularly when a spouse is in the active phase of a disorder. Various types of psychopathology, including depression, substance abuse, anxiety disorders, and schizophrenia, are associated with significant marital dysfunction and divorce. As such, they are stressors that have major implications for marital outcomes, largely because they interfere with important aspects of marital functioning. For example, skillful communication, adaptive emotion regulation, and the ability to be physically and emotionally available and supportive are important to good marital outcomes, but they are often impaired by psychopathology. Research consistently suggests that greater levels of psychopathology are associated with lower levels of marital satisfaction. It has also been found that depression in one spouse is significantly associated with lower marital satisfaction in both spouses. The association between psychopathology and poor marital functioning is so prominent that forms of marital and couple therapy are used to treat many disorders, highlighting the importance of examining the effects of psychopathology on both partners.

Infidelity

Infidelity is an important internal marital stressor that has recently garnered a great deal of attention. Although a paucity of empirical research on this stressor currently exists, both clinicians and scholars have recognized the significance of infidelity to couples. For example, Cano and O'Leary found the discovery of infidelity to be associated with increased levels of clinical depression in the noninvolved spouse. Given the significant negative role of psychopathology in marriages, as noted in a prior section, these findings highlight the importance of investigating infidelity as a marital stressor. Currently, researchers are investigating the differing effects of emotional versus sexual infidelity on wives and husbands. Findings are mixed as to whether gender differences exist

in response to different types of infidelity. Present research has investigated differences in husbands and wives in terms of the amount of distress caused or the likelihood of forgiveness depending on the type of infidelity. It appears that men's distress is primarily predicted from the type of infidelity (greater distress from sexual infidelity than emotional), whereas women's distress is predicted from a number of factors (e.g., relationship commitment and whether the person the husband is having the affair with is a new person vs. an ex-partner). Relationship commitment appears to be highly associated with the likelihood to forgive for both husbands and wives. However, more research is needed on the interactions between gender differences and the type of infidelity.

Ethnic Differences in the Relations between Stress and Marriage

Research on ethnic differences in the association between stress and marriage is still in its early stages, but indicates that couples of varying ethnicities are likely to experience different types of stressors consistent with their ethnic/cultural backgrounds. These stressors may be altered or amplified in couples of differing ethnicities and cultural backgrounds. Ethnic minority couples also face different stressors at a societal level than do Caucasian couples (e.g., stereotyping and stigmatization), suggesting that couples of different ethnicities may face stressors that present unique challenges to the success of their marriages. Much of the existing research focuses on differences between Caucasian and African American couples and families; however, there is an increasing trend toward including Latino and Asian couples and families. In general, research indicates some similarities between couples of various ethnicities but also some potentially important differences. Stressors related to financial hardship and social networks may play a particularly important role in African American and Latino marriages. Because of these sorts of findings, some authors are now calling for a more comprehensive model of marital outcomes for various ethnic minority couples, similar to the models described earlier, that include the role of stressors and how couples adapt to them. A great deal of additional research is necessary before the ethnicity-specific associations between stressors and marital outcomes for all groups are more fully understood. Research on interracial marriages and their potentially unique stressors is also garnering an increasing amount of attention. Such research suggests higher levels of marital challenges due to lack of cultural understanding, racial pressure, and lack of familial support. However, much more research is needed in this area.

Methodological and Measurement Issues

The study of stress in general has been fraught with measurement and methodological problems, and many of these problems emerge in the study of stress and marriage. For example, both external and internal stressors typically have been assessed using self-report measures. People are asked to report whether a stressor has occurred and how much stress or negative impact was associated with the occurrence of the event. At issue here is the extent to which people are reliable reporters, particularly when assessing amount of stress or negative impact. These reports can be biased by mood. Hence, distressed or unhappy spouses may report higher levels of stress and negative impact than nondistressed spouses, and this may artificially inflate associations between stress and poor marital functioning. One solution to this problem draws on the contextual threat procedures developed by George Brown and colleagues. Participants are interviewed and interviewers or coders make ratings of objective stress, as opposed to spouses' subjective stress ratings. This procedure is used frequently in the study of stress and psychopathology, and it could easily be applied to studies of marriage and stress. However, it is a time-consuming procedure. Another option is the use of daily diaries. Although such methods rely on self-report data, if we are interested in data obtained through self-report, such as examining each partner's account of various life events, daily-diary information may prove very useful. Notably, in daily-diary studies there is a dramatic reduction in the likelihood of retrospection because the amount of time elapsed between an experience and the account of the experience is minimized. The recall of the frequency of various stressors and marital interactions, and the immediate response to such events, therefore, may be greatly improved. Another consideration for marital researchers is whether it is the number of stressors that spouses experience or the stressfulness of stressors (e.g., the nature of the stressor) that affects marriage. Similarly, there are numerous types of stressors, including discrete events, chronic strains, and daily hassles, which might affect marital functioning in different ways. Another issue in stress research that applies to marital research is the reliance on retrospective reports of stress. Current models of stress and marriage imply longitudinal associations between the two, which suggests that stressors should be examined prospectively over time. Finally, it may be important to study specific types of stressors individually rather than aggregating unrelated events. All stressors may not show similar associations with marital functioning.

Further Reading

- Booth, A., Crouter, A. C. and Clements, M. (2001). *Couples in conflict*. Mahwah, NJ: Lawrence Erlbaum.
- Bradbury, T. N. (1998). *The developmental course of marital dysfunction*. New York: Cambridge University Press.
- Cann, A. and Baucom, T. R. (2004). Former partners and new rivals as threats to a relationship: infidelity type, gender, and commitment as factors related to distress and forgiveness. *Personal Relationships* 11, 305–318.
- Cano, A. and O'Leary, K. D. (2000). Infidelity and separations precipitate major depressive episodes and symptoms of nonspecific depression and anxiety. *Journal of Consulting and Clinical Psychology* 68, 774–781.
- Cano, A., O'Leary, K. D. and Heinz, W. (2004). Short-term consequences of severe marital stressors. *Journal of Social and Personal Relationship* 21(4), 419–430.
- Cohan, C. L. and Bradbury, T. N. (1997). Negative life events, marital interaction, and the longitudinal course of newlywed marriage. *Journal of Personality and Social Psychology* 73(1), 114–128.
- Cutrona, C. E. (1996). *Social support in couples: marriage as a resource in times of stress*. Thousand Oaks, CA: Sage.
- Davila, J., Bradbury, T. N., Cohan, C. L., et al. (1997). Marital functioning and depressive symptoms: evidence for a stress generation model. *Journal of Personality and Social Psychology* 73, 849–861.
- Demo, D. H. and Allen, K. R. (2000). *Handbook of family diversity*. London: Oxford University Press.
- Dillaway, H. and Broman, C. (2001). Race, class and gender differences in marital satisfaction and divisions of household labor among dual-earner couples. *Journal of Family Issues* 22(3), 309–327.
- Drummet, A. R., Coleman, M. and Cable, S. (2003). Military families under stress: implications for family life education. *Family Relations* 52(3), 279–287.
- Fu, X., Tora, J. and Kendall, H. (2001). Marital happiness and inter-racial marriage: a study in a multi-ethnic community in Hawaii. *Journal of Comparative Family Studies* 32(1), 47–60.
- Hammen, C. L. (1991). The generation of stress in the course of unipolar depression. *Journal of Abnormal Psychology* 100, 555–561.
- Hetherington, E. M. (1999). *Coping with divorce, single parenting, and remarriage: a risk and resiliency perspective*. Mahwah, NJ: Lawrence Erlbaum.
- Hill, R. (1949). *Families under stress*. Westport, CT: Greenwood Press.
- Karney, B. R. and Bradbury, T. N. (1995). The longitudinal course of marital quality and stability: a review of theory, method, and research. *Psychological Bulletin* 118, 3–34.
- Kreider, Rose M. (2005). Number, timing, and duration of marriages and divorces: 2001. In: *Current population reports*, pp. 70–97. Washington, DC: U.S. Census Bureau.
- L'Abate, L. (1998). *Family psychopathology: the relational roots of dysfunctional behavior*. New York: Guilford.
- Leiblum, S. R. (1997). *Infertility: psychological issues and counseling strategies*. New York: John Wiley & Sons.
- Lundquist, J. H. and Smith, H. L. (2005). Family formation among women in the U.S. military: evidence from NLSY. *Journal of Marriage & Family* 67(1), 1–13.
- McCubbin, H. I. and Patterson, J. M. (1983). The family stress process: the double ABCX model of adjustment and adaptation. *Marriage and Family Review* 6, 7–37.
- McQuillan, J., Greil, A. L., White, L., et al. (2003). Frustrated fertility: infertility and psychological distress among women. *Journal of Marriage & the Family* 65(4), 1007–1018.
- Revenson, T. A., Kayser, K. and Bodenmann, G. (2005). *Couples coping with stress: emerging perspectives on dyadic coping*. Washington, DC: American Psychological Association.
- Robles, T. F. and Kiecolt-Glaser, J. K. (2003). The physiology of marriage: pathways to health. *Physiology & Behavior* 79, 409–416.
- Root, M. P. (2001). *Love's revolution: interracial marriage*. Philadelphia: Temple University Press.
- Schmalings, K. B. and Sher, T. G. (2000). *The psychology of couples and illness*. Washington, DC: American Psychological Association.
- Story, L. B. and Bradbury, T. N. (2004). Understanding marriage and stress: essential questions and challenges. *Clinical Psychology Review* 23, 1139–1162.
- Whisman, M. A., Uebelacker, L. A. and Weinstock, L. M. (2004). Psychopathology and marital satisfaction: the importance of evaluating both partners. *Journal of Consulting and Clinical Psychology* 72(5), 830–838.

Maternal Deprivation

R L Huot, C O Ladd and P M Plotsky

Emory University School of Medicine, Atlanta, GA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 699–707, © 2000, Elsevier Inc.

Introduction

Maternal Deprivation Models

Consequences of Maternal Deprivation

Biological Significance

Glossary

Corticotropin releasing hormone (CRH)

A 41-amino-acid peptide and the primary regulator of the pituitary-adrenal system that is activated in response to stressors. CRH also acts as a neuromodulator in corticolimbic and brain-stem areas and, thus, is believed to participate in the coordination of the mammalian neuroendocrine, behavioral, autonomic nervous system, and immune responses to a stressor.

Hypothalamic-pituitary-adrenal (HPA) axis

A system that transduces interoceptive and exteroceptive signals into an endocrine response, culminating in the synthesis and secretion of adrenal glucocorticoids that act throughout the body to mobilize energy substrates such as glucose for use by the individual in escape or fighting.

Maternal deprivation

A neonatal manipulation in which rat pups are briefly handled and separated from the dam for prolonged periods of time ranging from 3 to 24 h. This procedure is typically implemented between postnatal days 1 and 21, with either single or repeated exposures.

Stress hypo-responsive period (SHRP)

A period from postnatal days 3 to 14 when neonatal rat pups show attenuated pituitary-adrenal responses to stressful stimuli that would cause a response in the adult. This period is characterized by low levels of circulating glucocorticoids, although a majority of this steroid is in free form due to the low production of corticosteroid-binding globulin during this age period. It has been postulated that the maintenance of low glucocorticoid levels is necessary during this period for the normal development of the HPA axis and the central nervous system, especially the hippocampus.

Early experience is associated with long-term behavioral, neuroendocrinological, cognitive, and central nervous system (CNS) changes in rodents, guinea pigs, and primates. The specific effects of positive or negative stimuli during the neonatal period are believed to be dependent on the genetic background of the individual, developmental status of the CNS at the time of exposure, duration of the stimulus, type and severity of the experience, and other factors such as the availability of social support. A number of research laboratories have employed the paradigms of neonatal handling and neonatal maternal deprivation as probes of CNS plasticity during the postpartum period. In general, these paradigms exert opposite influences, with maternal deprivation being associated with long-term increases in anxietylike behavior, HPA axis responsiveness to stressors, and changes in the functional characteristics of the central neurocircuits involved in the perception of and response to environmental events. Due to the relative stability of the changes resulting from neonatal maternal deprivation, as well as the ability of antidepressant treatment to reverse these changes transiently, this paradigm has been explored as an animal model of affective disorders.

Introduction

Initial studies in the late 1950s and 1960s by Levine, Denenberg, Ader, and others showed that neonatal handling or infantile stimulation led to long-term adaptations in behavior and the stress response. Subsequently, numerous paradigms of prenatal stress, neonatal maternal deprivation, and isolation rearing have been developed by numerous laboratories to systematically study the effects of adverse early experience on the development of behavior, endocrine function, and the CNS. This review is primarily limited to the effects of neonatal maternal deprivation in the rat. Before introducing the neonatal deprivation models, we provide a brief review of the developmental status of the neonatal CNS, the HPA axis and stress response, and the role of maternal behavior in development.

Central Nervous System Plasticity in the Neonate

Genetic and environmental factors interact to guide CNS development throughout the life span. Although the basic structure and connectivity of the brain are specified by genetic factors, environmental factors

play a key role in programming the patterns and density of connectivity that underlie the individual organism's perception of and responsiveness to its environment. These adaptations occur via a variety of processes, including activity-dependent plasticity and synaptic pruning. Frequently, this plasticity involves neurocircuits mediating the various aspects of the stress response, including the regulation of HPA axis activity.

Because many species spend their lifetimes within a narrow geographical range with respect to their birthplace, this interactive process may provide an adaptive advantage, serving as a mechanism by which an individual organism may optimize its neurocircuitry and develop behaviors appropriate for survival. As illustrated in [Figure 1](#) for the rat, many neurodevelopmental processes continue during and after the neonatal period. Furthermore, many of these processes are sensitive to circulating glucocorticoid hormone levels. Therefore, exposure to adverse conditions during the neonatal period can modulate the CNS developmental process through multiple mechanisms. This postpartum period of increased vulnerability to environmental factors varies among species, depending on the state of CNS developmental maturity at birth.

The Hypothalamic–Pituitary–Adrenal Axis and Regulation of the Stress Response

Internal and external stimuli (i.e., stressors) conveying actual or perceived threat, pain, metabolic imbalance, or failure of expectations are processed via widely distributed pathways and converge on the paraventricular nucleus (PVN) of the hypothalamus. Neurons in this structure produce the peptides corticotropin releasing hormone (CRH) and arginine vasopressin (AVP), which are transported to the median eminence nerve terminals and are released into the hypophyseal-portal circulation, the functional link between the CNS and the endocrine system. At the anterior pituitary gland, these peptides stimulate the synthesis and release of adrenocorticotrophic hormone (ACTH) into the systemic circulation. Circulating ACTH evokes the synthesis and secretion of glucocorticoids (cortisol in primates and corticosterone in rodents) from the adrenal cortex, which serve to mobilize energy substrates during stress. Plasma glucocorticoid levels are controlled tightly by glucocorticoid-mediated negative feedback on the secretion of CRH and ACTH by the activation of cytoplasmic mineralocorticoid receptors (MR) localized in the

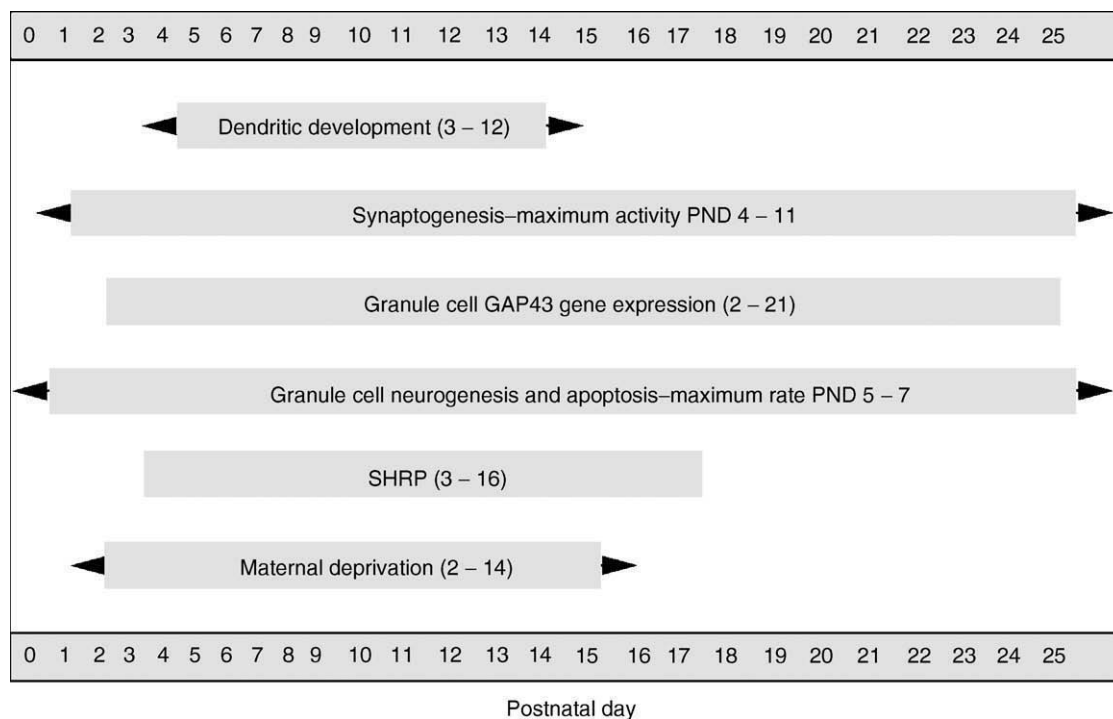


Figure 1 Developmental processes occurring during the neonatal period in rodents. Maternal deprivation paradigms overlap a significant portion of the stress hyporesponsive period (SHRP). Throughout the neonatal period, myelination, neurogenesis, and apoptosis, as well as synaptogenesis and synaptic pruning, are occurring. All these mechanisms represent targets for epigenetic and environmental modification of the basic genetic plan for CNS development.

hippocampus and septum and/or the activation of cytoplasmic glucocorticoid receptors (GR) located in the adenohypophysis and throughout the CNS.

HPA axis responses are adaptive and essential for survival in adulthood, but may be maladaptive in infancy while the brain is immature. In the normally developing rodent the brain is protected from catabolic and other effects of glucocorticoids due to a reduced sensitivity of the HPA axis to activation during the neonatal period, postnatal days (PND) 3–16, a developmental epoch known as the SHRP. This reduced responsiveness is not absolute and may be overcome by strong stressors, including maternal deprivation.

Role of Maternal Behavior in Postnatal Development

The early environment in many mammalian species is determined largely by the mother. In many primate species, including humans, early interactions between the mother and the infant serve as a template for all subsequent relationships. The neonatal rat is completely dependent on its mother for nutrition, temperature regulation, grooming, and regulation of micturition for the first 2 weeks of life. Alterations in the quality of maternal care have been shown to exert profound and long-lasting influences on the development of stress responsiveness and health in rodent offspring. This dependence and the subtle complexities of maternal behavior led Hofer to postulate the concept of the mother as a hidden regulator of physiological and behavioral development in the rat.

In many mammals, maternal care is quite stereotypic and involves a number of distinct behaviors. Maternal behavior in the rat consists of nest building, retrieval of the pups into the nest, nursing periods associated with an arched-back posture to encourage suckling, and anogenital licking to stimulate urination and defecation. Levine and colleagues showed that particular maternal behaviors modulate the development of the HPA axis at different levels and serve, in part, to reduce the responsivity of the HPA axis beginning a few days after birth and lasting until the pups are more independent.

Maternal Deprivation Models

Rodent Models of Maternal Deprivation

Numerous maternal deprivation models have been developed to study the effects of prolonged separations of infants from their mothers on behavioral and endocrine responsiveness. These deprivation manipulations vary greatly among research groups, ranging from a single 24-h deprivation to repeated maternal deprivation episodes lasting 12, 8, 6, or 3 h from PND

1–2 through PND 14–21. In addition, some paradigms require the isolation of pups from one another during the separation period, whereas in others the pups remain together as a litter during the period of separation.

In an open colony setting, the dam typically leaves her pups for periods of 5–15 min in order to forage for food and water; however, in some cases the dam may be away for periods of 60–180 min. Mimicking these observations, Plotsky and Meaney developed a maternal deprivation model in which Long Evans rat litters were standardized to eight male pups (or four male and four female pups) on PND 2 and then assigned randomly to one of four neonatal rearing conditions during PND 2–14, inclusive: (1) brief handling and 180 min per day of maternal separation (HMS180); (2) brief handling and 15 min per day of maternal separation (HMS15); (3) nonhandled (HMS0); and (4) animal facility rearing (AFR), in which litters were handled twice weekly beginning on PND 4 for cage changing. Bedding material in the home cages was replaced only partially (50%) with clean bedding once per week when the dams and pups were out of the cage (HMS15 and HMS180). In all deprivation studies, the dams were moved to an adjacent cage when they were off the nest in the morning and then the pups were placed in a transfer cage and moved to an incubator in another room. At the end of the daily separation, both HMS15 and HMS180 litters were returned to their home cage, rolled in the native bedding, and then reunited with their dam. All litters were weaned on PND 21–23 and group housed (two to three per cage) until adulthood (> PND 60).

Comparisons have been made between the maternally deprived animals (HMS180) and a control group consisting of nonhandled (HMS0), animal facility-reared (AFR), or handled (HMS15) rats. Despite the wide variations in the maternal deprivation protocols, adult animals that experienced neonatal maternal deprivation typically display high anxietylike behavior and HPA axis hyperresponsiveness, whereas adults that experienced neonatal handling exhibit reduced anxietylike behavior and conserved HPA axis responses to stressors. In this review, the control and experimental groups are presented as in the studies, but it should be kept in mind that the selection of the control group may have a profound effect on the conclusions drawn from the data.

Primate Models of Maternal Deprivation

The majority of nonhuman primate maternal deprivation studies have been performed in rhesus monkeys or squirrel monkeys. The specific nature of the maternal deprivation protocols varies widely in these studies from severe and prolonged maternal

deprivation to less profound periods of deprivation. Maternal separation in primates is associated with increased oral behaviors and other stereotypic movements, anxietylike behavior, cognitive impairment, aggression, deficits in bonding, and exaggerated HPA axis function.

Consequences of Maternal Deprivation

Changes in Maternal Behavior

The dams of pups exposed to the HMS15 protocol exhibited increased maternal behavior, such as licking and grooming of the pups, as compared to the AFR group, whereas prolonged maternal deprivation (HMS180) was associated with reduced and deranged maternal behavior. Dams of the HMS180 group retrieved their pups with a longer latency on reunion and took longer to begin to nurse the pups as well as to lick and groom them. This decreased maternal behavior was largely responsible for the long-term resistance to stress as demonstrated in cross-fostering studies. In support of the role of maternal behavior, Meaney and colleagues demonstrated an inverse correlation between the amount of time the dam spent in licking, grooming, and arched-back nursing of her pups and subsequent adult stress responsiveness. Furthermore, the importance of anogenital licking for the development of the normal HPA axis stress response has been shown.

Hypothalamic-Pituitary-Adrenal Axis Adaptations

Handled, nonhandled, and maternally deprived rats show distinct and stable differences in their HPA response to stressors as adults, as summarized in Table 1. In adult rats, basal levels of ACTH and corticosterone typically do not differ among the rearing groups described earlier. However, in response to a psychological stressor (e.g., novel environment or air puff startle), exaggerated ACTH and corticosterone responses are apparent in maternally deprived and nonhandled rats as compared to handled or animal facility reared rats. Using a single 24-h period of neonatal maternal deprivation, van Oers and colleagues have shown that the age at which maternal deprivation occurred altered the outcome with respect to HPA axis function. Deprivation at PND 3 resulted in hyperresponsiveness of the pups at PND 20, whereas deprivation at PND 11 resulted in hypo-responsiveness. Similar experiments will be necessary to elucidate the neurocircuits and mechanisms mediating aspects of the overall neuroendocrine, behavioral, and cognitive effects of neonatal maternal deprivation.

Table 1 HPA axis effects of neonatal handling and maternal deprivation as compared to nonhandled rats

Variable	Neonatal handling ^a	Neonatal maternal deprivation ^b
HPA axis function		
Basal ACTH	↔	↔
Basal corticosterone	↔	↔
ACTH stress response	↓	↑
Corticosterone stress response	↓	↑
Dexamethasone resistance	↔	↑
CRH receptor binding		
Paraventricular nucleus	↔	↑
Pituitary	↔	↓
CRH content		
Paraventricular nucleus	↔	↑
Stalk median eminence	↓	↑
CRH mRNA		
Paraventricular nucleus	↓	↑
GR mRNA		
Hippocampus	↑	↓
Paraventricular nucleus	↔	↔
Pituitary	↔	↔
Prefrontal cortex	↑	↓

^aHMS15 (brief handling and 15 min per day of maternal separation).

^bHMS180 (brief handling and 180 min per day of maternal deprivation).

At the level of the hypothalamus, maternally deprived and nonhandled rats show significantly greater expression of CRH mRNA, specifically in the PVN, compared to handled rats. In addition, these rats have an elevated peptide content of CRH and AVP in the stalk median eminence as well as increased CRH and AVP concentrations in the hypophyseal-portal circulation compared to handled rats. Also, maternally deprived and nonhandled rats show increased CRH receptor binding in the PVN compared to handled rats. This latter adaptation may reflect the increased expression of CRHR1-receptor mRNA in the PVN as a consequence of the stress of maternal deprivation. Thus, the brains of these animals exhibit increased CRH as well as potentially increased responsiveness to this CRH.

Impaired glucocorticoid-mediated negative feedback may contribute to the elevated expression of CRH as well as to the exaggerated secretion of ACTH and corticosterone seen in deprived rats, whereas an increased feedback tone may underlie the reduction in stress responsiveness observed in handled rats. The expression of the two types of glucocorticoid receptors is extremely vulnerable to neonatal experience. The type 1 or mineralocorticoid receptor (MR) is a high-affinity receptor that shows a high degree of saturation even under basal conditions. It is believed to be

involved in the maintenance of basal HPA activity. In contrast, the type 2 or glucocorticoid receptor (GR) is a low-affinity receptor that becomes activated in the presence of elevated levels of glucocorticoids typical of stress. As adults, the density of hippocampal GR binding sites in the experimental groups is handled = AFR > nonhandled > maternally deprived; the levels of GR mRNA follow a similar pattern. An increased expression of GR in the hippocampus and prefrontal cortex may increase the negative feedback tone on the HPA axis, thus providing a mechanism by which handled rats dampen their stress response. Conversely, the reduction of hippocampal GR in maternally deprived rats may, in part, explain their HPA axis hyperresponsiveness to stressors. Reports also indicate that hippocampal MR mRNA levels are increased in HMS180 maternally deprived rats, perhaps accounting for the normal circadian trough and peak ACTH and corticosterone levels in these animals. These changes in GR and MR binding and expression are apparent from at least PND 21 in the HMS180 model.

However, in response to 24 h of maternal deprivation on PND 3, Levine's group found decreased binding for both MR and GR in the CA1 region of the hippocampus in males at PND 48. No changes were observed in binding of the MR in females at PND 48 or of mRNA for either MR or GR at PND 60. When deprivation was applied on PND 11 and the brains collected the next day, one study showed a decrease in both MR and GR mRNA in CA1, whereas another study showed a decrease only in the MR mRNA in CA1 and no changes in the GR mRNA. Thus, although these studies confirm the high degree of plasticity of the glucocorticoid receptors in the brain in response to early experience, the actual effects of maternal deprivation seem variable over time. Studies of the time sequence in glucocorticoid receptor expression changes in response to maternal deprivation are needed to sort these out among maternal deprivation paradigms.

Behavioral Consequences

Stress has long been associated with increased anxiety- or fearlike behavior. Using multiple tests of these behaviors (e.g., elevated plus maze, novel open-field locomotion, defensive withdrawal, and acoustic startle), many groups have shown that maternal deprivation is associated with increased anxiety- and fearlike behavior beginning in the neonatal period and persisting throughout adulthood.

When placed in a brightly lit, novel open field, rats initially spend more time near the walls and show reduced locomotor activity. HMS180 maternally deprived rats continue to display neophobia long after the AFR and HMS15 animals have habituated to the

novel environment. Similar results were found in an open-field exploration test as well as in the latency to eating in a conflict test following overnight food deprivation, suggesting a heightened fearfulness in maternally deprived rats exposed to a novel arena.

In the defensive withdrawal test, HMS180 rats show an increased latency to exit from a safe compartment into the open-field arena and demonstrate an increased number of reentries into the cylinder as well as a greater total amount of time spent in the cylinder compared to AFR and HMS15 rats. Similarly, in the elevated plus maze, maternally deprived rats spent less time in the open arms compared to either handled or AFR rats. Nonhandled rats show a level of anxietylike intermediate between maternally deprived and handled rats in these tests. Thus, it appears that maternally deprived rats show greater anxiety and neophobia than do handled (HMS15), nonhandled (HMS0), or normally reared (AFR) rats.

Early experience can also affect an adult animal's cognitive performance. Infantile handling has been shown to prevent cognitive impairments as the animal ages. At 24 months, but not at 12 months, handled rats learn to find a hidden platform in the Morris water maze faster than nonhandled rats. This may be due to less exposure to glucocorticoids throughout the animals' life because elevated levels of these stress hormones are related to cognitive impairments. It has been hypothesized that maternally deprived rats should be impaired on tasks of learning and memory. Data from the authors' laboratory support this hypothesis; middle-age (120 days) maternally deprived rats show a slight, but significant, impairment on the learning of the Morris water maze.

Overall, these behavioral observations provide compelling support for the thesis that maternal deprivation during the neonatal period increases anxiety- or fearlike behavior that persists throughout adulthood. Thus, sensory experiences encountered during a critical window of development, particularly those associated with mother-infant interactions, significantly affect individual perception of the environment and responses to it throughout life. It is postulated that this process arises through a combination of stimulus-induced plasticity during postnatal development, activation of extrahypothalamic CRH neurocircuits, and glucocorticoid effects mediated through a cascade of transcriptional changes in critical sensory and limbic systems.

Extrahypothalamic Effects

Considering the effects of maternal deprivation on behavior and endocrine responses to psychological stressors, it would not be surprising to find a number of adaptations in the CNS, especially in systems

involved in the regulation of stress and fear. Studies in the HMS180 maternal deprivation paradigm have shown that maternal deprivation and handling affect extrahypothalamic CRH systems, GABA and benzodiazepine pathways, and norepinephrine circuits as outlined in Table 2.

The involvement of extrahypothalamic CRH systems is particularly salient because these circuits appear to coordinate the neuroendocrine, behavioral, sympathetic nervous system, and cognitive responses to stressors. Furthermore, there are substantial interactions between these CRH-containing neurocircuits and brain-stem noradrenergic systems. CRH mRNA levels in the central nucleus of the amygdala and the bed nucleus of the stria terminalis were increased in HMS180 maternally deprived rats and decreased in HMS15 handled rats compared to nonhandled rats. The CRH mRNA levels in these brain regions of the HMS15 rats were equivalent to those of animal facility-reared animals.

These CRH-containing pathways project to brain-stem areas such as the locus ceruleus, where CRH activates locus ceruleus noradrenergic neurons. This stress-induced activation may serve to increase autonomic nervous system information flow, enhance vigilance behavior, and provide additional drive to the HPA axis. In addition, the apparent reduction in α_2 -adrenergic receptor binding in the locus ceruleus and nucleus of the solitary tract implies that, once activated, the firing rate of the locus ceruleus neurons may be difficult to shut off. Interestingly, HMS180 maternal deprivation was also associated with decreased central benzodiazepine (CBZ) receptor binding in the locus ceruleus and the nucleus of the solitary tract compared to handled rats. Thus, increased CRH and reduced CBZ, GABA_A, and α_2 -adrenergic receptor binding all provide mechanisms for general hyperresponsiveness to psychological stressors as a result of maternal deprivation. It is not known whether these same changes result from other maternal deprivation paradigms.

Table 2 Extrahypothalamic effects of neonatal rearing paradigms

Variable	Neonatal paradigm ^a
<i>CRH receptor binding</i>	
Locus ceruleus	HMS180 > HMS0 > HMS15 = AFR
Raphe	HMS180 > HMS0 > HMS15 = AFR
<i>CRF content</i>	
Bed nucleus of the stria terminalis	HMS180 > HMS0 > HMS15 = AFR
Central nucleus of the amygdala	HMS180 > HMS0 > HMS15 = AFR
Hippocampus	HMS180 = HMS0 = HMS15 = AFR
Locus ceruleus	HMS180 = HMS0 = HMS15 = AFR
<i>CRF mRNA</i>	
Bed nucleus of the stria terminalis	HMS180 > HMS0 > HMS15 = AFR
Central nucleus of the amygdala	HMS180 > HMS0 > HMS15 = AFR
Hippocampus	HMS180 = HMS0 = HMS15 = AFR
GR mRNA	
Hippocampus	AFR = HMS15 > HMS0 > HMS180
Prefrontal cortex	AFR = HMS15 > HMS0 > HMS180
<i>α_2-Adrenergic receptor binding</i>	
Locus ceruleus	HMS15 > AFR > HMS0 > HMS180
Nucleus of the solitary tract	HMS15 > AFR > HMS0 > HMS180
<i>GABA_A receptor binding</i>	
Amygdala	HMS180 = HMS0 = HMS15 = AFR
Locus coeruleus	AFR = HMS15 > HMS0 = HMS180
Nucleus of the solitary tract	AFR = HMS15 > HMS0 = HMS180

^aAFR, animal facility-reared control; HMS0, nonhandled; HMS15, handled (brief handling and 15 min per day of maternal separation); HMS180, maternal deprivation (brief handling and 180 min per day of maternal deprivation).

Biological Significance

Accumulating evidence indicates that there are critical developmental windows during which the developing brain is very susceptible to environmental influences. It is postulated that cues related to the environment are transmitted primarily via the primary caregiver, which in most species is the mother. Transmission may be through factors transferred in the milk, aspects of maternal behavior, or both. Maternal-infant interactions and other sensory experiences during these sensitive periods sculpt the final form of the nervous system via multiple mechanisms. The processes of activity-dependent plasticity and hormonal programming act to fine-tune the genetically determined neurocircuitry by altering the strength of connections, changing the morphology of neurons, and altering the expression critical signaling molecules. These changes then lead to robust differences in how that individual decodes and responds to its environment.

Observations of rats exposed to the HMS180 maternal deprivation paradigm suggest a likely cascade initiated by adverse early experience. However, the sequence of this cascade is currently unknown. The initial events lead to impaired function of two primary inhibitory systems: the glucocorticoid-mediated negative feedback system and the GABAergic-CBZ system. Subsequent steps in the cascade involve the dysregulation of CRH and monoaminergic systems.

The relative loss of glucocorticoid tone, as well as of inhibitory GABAergic-CBZ tone in the amygdaloid complex and the locus coeruleus, appears to

disinhibit ascending noradrenergic drive, thus further increasing CRH secretion and gene expression in hypothalamic and extrahypothalamic circuits. Overall, these processes underlie the complex neuroendocrine, behavioral, and cognitive changes observed in response to maternal deprivation in rodents.

Like young rats, both monkey and human babies exhibit elevated glucocorticoid responses during periods of brief maternal separation. Furthermore, early adverse experience in humans also leads to long-term changes in behavior and stress responsiveness. Childhood stressors, including the loss of a parent, neglect, or child abuse, have been associated with an increased vulnerability to mood and anxiety disorders. Models of neonatal maternal deprivation in rats may serve to elucidate the CNS changes occurring in response to adverse early experience in the human and may provide insights into effective prevention and treatment strategies. The mechanisms uncovered by these animal studies may provide a bridge between psychological and psychiatric constructs and their central systems and molecular neuroscience mechanisms.

See Also the Following Article

Childhood Stress.

Further Reading

- Arling, G. L., Harlow, H. F., Sackett, G. P., et al. (1976). A 10-year perspective of motherless-mother monkey behavior. *Journal of Abnormal Psychology* 85, 341–349.
- Bhatnagar, S., Shanks, N., Plotsky, P. M. and Meaney, M. J. (1996). Hypothalamic-pituitary-adrenal responses to stress in neonatally handled and nonhandled rats: differences in facilitatory and inhibitory neural pathways. In: McCarty, R., Aguilera, G., Sabban, E. & Kvetnansky, R. (eds.) *Stress: molecular genetic and neurobiological advances*. New York: Gordon and Breach Science Publishers.
- Caldji, C., Francis, D., Sharma, S., et al. (in press). The effects of early rearing environment on the development of GABA_A and central benzodiazepine receptor levels and novelty-induced fearfulness in the rat. *Journal of Neuroscience*.
- Caldji, C., Tannenbaum, B., Sharma, S., et al. (1998). Maternal care during infancy regulates the development of neural systems mediating the expression of fearfulness in the rat. *Proceedings of the National Academy of Sciences USA* 95, 5335–5340.
- Calhoun, J. B. (1962). *The ecology and sociology of the Norway rat*. Bethesda, MD: H. E. W. Public Health Service.
- Cullinan, W. E., Herman, J. P., Helmreich, D. L. and Watson, S. J. (1995). A neuroanatomy of stress. In: Friedman, M. J., Charney, D. S. & Deutch, A. Y. (eds.) *Neurobiological and clinical consequences of stress: from normal adaptation to PTSD*, pp. 3–25. Philadelphia: Lippincott-Raven.
- Dunn, A. J. and Berridge, C. W. (1990). Physiological and behavioral responses to corticotropin-releasing factor administration: is CRF a mediator of anxiety and of stress responses? *Brain Research Reviews* 15, 71–100.
- Francis, D., Diorio, J., LaPlante, P., et al. (1996). The role of early environmental events in regulating neuroendocrine development. *Annals of the New York Academy of Sciences* 794, 136–152.
- Francis, D. D. and Meaney, M. J. (1999). Maternal care and the development of stress responses. *Current Opinion in Neurobiology* 9, 128–134.
- Hall, F. S. (1998). Social deprivation of neonatal, adolescent, and adult rats has distinct neurochemical and behavioral consequences. *Critical Review of Neurobiology* 12, 129–162.
- Hennessy, M. B. (1997). Hypothalamic-pituitary-adrenal responses to brief social separation. *Neuroscience and Biobehavioral Reviews* 21, 11–29.
- Hofer, M. A. (1994). Early relationships as regulators of infant physiology and behavior. *Acta Paediatrica* (supplement) 397, 9–18.
- Jacobson, L. and Sapolsky, R. M. (1991). The role of the hippocampus in feedback regulation of the hypothalamic-pituitary-adrenal axis. *Endocrine Review* 12, 118–134.
- Kolb, B. and Whishaw, I. Q. (1998). Brain plasticity and behavior. *Annual Review of Psychology* 49, 43–64.
- Ladd, C. O., Huot, R. L., Thirivikraman, K. V., Nemeroff, C. B., Meaney, M. J. and Plotsky, P. M. (2000). Long term behavioral and neuroendocrine adaptations to adverse early experience. In: Mayer, E. & Saper, C. (eds.) *Progress in brain research: the biological basis for mind body interactions* (vol. 122), pp. 79–101. Amsterdam: Elsevier.
- Levine, S. (1994). The ontogeny of the hypothalamic-pituitary-adrenal axis: the influence of maternal factors. *Annals of the New York Academy of Sciences* 746, 275–288.
- Levine, S., Wiener, S. G. and Coe, C. L. (1993). Temporal and social factors influencing behavioral and hormonal responses to separation in mother and infant squirrel monkeys. *Psychoneuroendocrinology* 18, 297–306.
- Liu, D., Tannenbaum, B., Caldji, C., et al. (1997). Maternal care, hippocampal glucocorticoid receptor gene expression and hypothalamic-pituitary-adrenal responses to stress. *Science* 277, 1659–1662.
- McEwen, B. S., de Leon, M. J., Lupien, S. J. and Meaney, M. J. (1999). Corticosteroids, the aging brain and cognition. *Trends in Endocrinology and Metabolism* 10, 92–96.
- Pellow, S., Chopin, P., File, S. E., et al. (1985). Validation of open:closed arm entries in an elevated plus maze as a measure of anxiety in the rat. *Journal of Neuroscience Methods* 14, 149–167.
- Plotsky, P. M. (1991). Pathways to the secretion of adrenocorticotropin: a view from the portal. *Journal of Neuroendocrinology* 3, 1–9.

- Plotsky, P. M. and Nemeroff, C. B. (1998). Molecular mechanisms regulating behavior. In: Jameson, J. L. (ed.) *Textbook of molecular medicine*, pp. 979–988. New Jersey: Humana Press.
- Plotsky, P. M., Owens, M. J. and Nemeroff, C. B. (1998). Psychoneuroendocrinology of depression. *Psychoneuroendocrinology* **21**, 293–307.
- Sapolsky, R. M. and Meaney, M. J. (1986). Maturation of the adreno-cortical stress response: neuroendocrine control mechanisms and the stress hyporesponsive period. *Brain Research Reviews* **11**, 65–76.
- Spencer-Booth, Y. and Hinde, R. A. (1971). Effects of brief separations from mothers during infancy on behavior of rhesus monkeys 6–24 months later. *Journal of Child Psychology and Psychiatry* **12**, 157–172.
- Takahashi, L. K., Kalin, N. H., Vandenburgt, J. A., et al. (1989). Corticotropin-releasing factor modulates defensive-withdrawal and exploratory behavior in rats. *Behavioral Neuroscience* **103**, 648–654.
- Valentino, R. J., Curtis, A. L., Page, M. E., et al. (1998). Activation of the locus ceruleus brain noradrenergic system during stress: circuitry, consequences, and regulation. *Advances in Pharmacology* **42**, 781–784.
- Van Oers, H. J., de Kloet, E. R. and Levine, S. (1998). Early vs. late maternal deprivation differentially alters the endocrine and hypothalamic responses to stress. *Developmental Brain Research* **111**, 245–252.
- Weinstock, M. (1997). Does prenatal stress impair coping and regulation of hypothalamic-pituitary-adrenal axis? *Neuroscience and Biobehavioral Reviews* **21**, 1–10.
- Wheal, H. V., Chen, Y., Mitchell, J., et al. (1998). Molecular mechanisms that underlie structural and functional changes at the postsynaptic membrane during synaptic plasticity. *Progress in Neurobiology* **55**, 611–640.

Medical Profession and Stress

K G Power and V Swanson

University of Stirling, Stirling, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 708–712, © 2000, Elsevier Inc.

Sources of Occupational Stress for Medical Professionals
Mediators and Moderators of Occupational Stress
Stress Outcomes for Medical Professionals
Summary

Glossary

<i>Consultant</i>	A specialist medical practitioner providing secondary care, normally in a hospital setting.
<i>General practitioner</i>	A doctor working in primary care in the community; the first point of contact for medical cases of all kinds.
<i>Mediators/moderators</i>	Intervening variables that alter (increase or decrease) the relationship between stressors and strains.
<i>Strains</i>	Physical, psychological, or behavioral outcomes of stressors.
<i>Stressors</i>	External demands or stimuli acting on an individual, perceived as taxing his or her resources.
<i>Transactional models</i>	Models that conceptualize stress as a dynamic interaction between the person and his or her environment.

Sources of Occupational Stress for Medical Professionals

Intrinsic Factors

Although medical work can be routine and repetitive, doctors have high levels of responsibility, necessitating intensive skill use. A wrong decision or diagnosis may put a life at risk or lead to litigation. Doctors must be familiar with a vast area of knowledge and increasingly complex technology, but often act and make decisions in isolation. The corollary of autonomy of decision making is therefore a heavy burden of responsibility and a high degree of accountability for errors. Many aspects of medical work are emotionally stressful, e.g., dealing with death, trauma, suffering, and serious illness and communicating difficult and emotional decisions to patients and relatives. There may be personal moral and ethical dilemmas involved in patient treatment, e.g., in dealing with abortion. Environmental factors may increase personal risk, including exposure to viral infections or contagious diseases such as hepatitis and HIV. There is risk of physical injury from abusive or violent patients, such as in emergency departments, from psychiatrically disturbed patients, or during home visits. Although clinical aspects of medical work can be a source of stress, intrinsic factors have been shown to contribute less to occupational stress than extrinsic or organizational factors. Rigorous training generally equips doctors to deal with clinical issues, facilitating the development of effective coping mechanisms to deal with

clinical responsibilities. However, they may be less well equipped to cope with the hassles of day-to-day management and increased bureaucratization of health care. The repetitive or trivial nature of some patient consultations may be a source of irritation or hassle, and the expectations or demands of patients and their relatives may be perceived as unrealistic.

Workload

Doctors' workload and constant time pressure have been identified as major sources of occupational stress. Several factors may be considered, including overall volume of work, patient list sizes, number of consultations, home visits, unsocial hours and night work, time pressure, and time spent on call. Being on call is described as being in a state of heightened alert, during which it is difficult to relax. Dealing with night calls may result in disturbed sleep, leading to impaired physical and mental functioning. Howie and colleagues assessed the impact of workload in general practitioners (GPs) and found that larger than average list sizes, time pressures (e.g., appointments running late), and more patient-centered attitudes in doctors correlated positively with stress. Faster doctoring and poor time management were associated with less recognition of patients' psychosocial problems and higher rates of prescribing.

Organizational factors such as overbooked and late-running surgeries can lead to shortened consultations and reduced quality of care for patients, whereas improved time management leads to reduced stress for GPs. Agius and co-workers identified demands on time as a major stressor for consultants, including interruptions, meeting deadlines, finding time for research and teaching demands, having so much to do that everything cannot be done well, and conflicts between work and family time as main components. Compared with doctors in hospital settings, general practitioners working in the community may feel constantly visible and on call, leaving less time for self or family. In some countries, the development of GP cooperatives to manage on-call requirements addresses the issue of 24-h commitment to patient care, potentially reducing a major source of occupational stress.

Relationships at Work

Relationships with colleagues are a potential source of stress for doctors. In particular, relationships between medical students and consultants may be a source of stress. Furthermore, the previously competitive style of much medical training is at odds with recent emphasis on the health-care team in general practice. Team working may also introduce interprofessional role conflicts, uncertainties or ambiguities,

and a sense of a diminished role for doctors more accustomed to autocratic or autonomous working styles.

Gender Issues

Male and female doctors have different sources and levels of occupational stress. Despite intakes of women to some medical schools now standing at 50% or higher, women are underrepresented in some specialties (e.g., surgery) and hold fewer consultant grades overall. Patronage in some specialties and the old boys network are still considered to be important parts of the career system discriminating against women. The premise that an increased entry of females to the profession would lead to changes in the structural hierarchy of medical care has not been fulfilled. In India, where sex role segregation is commonplace, medical specialty choice is strongly associated with gender-based stereotypes, with female physicians dealing almost exclusively with women's and children's medical problems.

Role Conflicts: Home-Work Interface

Stress in the home-work interface is common for medical professionals. Stress between work and home is bidirectional and may affect male and female doctors differently depending on the complexity of domestic and occupational roles at different life stages. People-intensive and emotionally demanding medical work may leave the individual feeling drained and uncommunicative with spouse or family, and domestic relationship problems may adversely affect work performance. Female doctors fare less well than male colleagues in maintaining successful marriages, with approximately one-third of female doctors in the United Kingdom remaining single and divorce rates being higher for female than for male doctors.

Mediators and Moderators of Occupational Stress

Transactional psychological models emphasize the importance of individual differences in determining the link between stressors and stress outcomes or strains. Structural characteristics of both the job and the individual medical practitioner influence the stressor-strain relationship.

Structural Factors

The characteristics and type of medical practice (e.g., single-handed, group practice; government funded or profit based) or medical specialty may make them more or less demanding and threatening to doctors' general well-being. Some comparative studies have

identified general practice as one of the most stressful areas of medicine. Anesthetics and psychiatry have also been singled out for being respectively more and less stressful than other specialties. However, evidence regarding the stressfulness of different types of medical specialty is not conclusive. Other studies comparing psychiatric morbidity in doctors in different specialties found no differences among gastroenterologists, surgeons, radiologists, and oncologists or among consultant physicians, surgeons, and radiologists.

Dispositional Factors

The psychological characteristics of individuals who chose a medical career may predispose them to experience work-related psychological distress. Cooper and colleagues suggested that a type A personality predicted reduced mental well-being in GPs, a finding confirmed in other studies, although only a small percentage of variance in strains tends to be explained by type A personality. An internal locus of control has been associated with reduced levels of stress in medical professionals, and neuroticism has been found to correlate highly with occupational stress, emotional exhaustion, depersonalization, and lowered job-related achievement in hospital consultants.

Coping

Doctors possess more resources for coping than most occupational groups due to their length of education and training, status, and financial security. However, few studies have considered coping in doctors. The association between neuroticism and emotion-focused coping strategies used by consultants was linked to a negative appraisal of job stressors by Deary and colleagues. Although maladaptive coping behaviors such as drug or alcohol abuse or work-related strategies such as increased prescribing have been examined, adaptive coping has been studied infrequently. Social support has been identified as an important intervening variable. Increased quality and quantity of social support are associated with a reduction in perceived stress, increased job satisfaction, and improved mental well-being in both male and female doctors, although female doctors utilize a wider range of social support and home- or work-based coping strategies than males.

Stress Outcomes for Medical Professionals

Job Satisfaction

Job dissatisfaction and lowered morale are commonly described outcomes of stress in medical professionals,

with moderate negative correlations between job satisfaction and occupational stress being reported frequently. Job dissatisfaction is associated with poor work performance, mental ill health, and inappropriate prescribing. Female doctors have consistently been shown to report greater job satisfaction than males.

Physical Ill Health

Although medical knowledge is no guarantee of good health, doctors have improved rates of morbidity and mortality in comparison with the general population and other professionals. However, mortality for some causes (e.g., cirrhosis of the liver, road accidents) is substantially higher for doctors than for other comparable professional groups, and concerns have been expressed regarding the tendency of doctors to self-treat or neglect their own health. Because doctors have multiple sources of self-referral, the collection of accurate morbidity data may be more difficult for doctors than for other groups.

Mental Ill Health

Several studies have suggested that doctors have higher rates of mental distress, anxiety, and depression than norms. Because of long working hours, heavy workloads, and the emotional demands of medical practice, younger doctors and those in training may be more vulnerable to psychological distress than older colleagues. However, direct causal links between work stress and mental illness in doctors have not been established firmly. Many studies fail to take into account baseline gender differences in the prevalence of anxiety and depression. Some studies suggest that female doctors suffer from more mental health problems than male doctors or other female professionals, whereas others find no significant differences between female and male doctors or between female doctors and other professional women in stress-related mental distress.

Measures of burnout, encapsulating depersonalization and emotional exhaustion, may be particularly sensitive indicators of strain in medical professionals. One Australian study found that young male GPs were more likely to suffer burnout than young female GPs, suggesting that this may be due to a more empathetic and psychosocial approach to patient care in female doctors. However, other studies have found no gender differences in doctors' burnout. Furthermore, the concept of burnout may overlap a dispositional tendency toward negative affectivity and/or neuroticism, leading to confounding of dependent and independent variables in studies using burnout measures.

Behavioral Outcomes: Alcohol and Drug Use

Doctors generally exhibit a greater prevalence of alcohol and drug abuse than comparable social groups. A main source of drug misuse is via self-prescribing. Psychiatric admissions for alcoholism for male doctors in Scotland were 2.7 times greater than for a comparison group of social class 1 males, with 58% of psychiatric hospitalizations for male doctors in the 45 to 54 age group attributed to alcoholism. GPs in single-handed practices are also overrepresented in hospital admissions for doctors with drug and alcohol problems. General practitioners and some medical specialties, notably anesthetics, have higher rates of alcoholism and drug abuse than others. It has been argued that males in general are more likely to use alcohol as a means of coping with stress than females. This gender difference is reflected in studies of GPs. However, one longitudinal study of junior house officers, albeit in a mixed-gender sample, found no relationship between reported drinking habit and self-reported stress or depression.

Suicide

Suicide rates for doctors in the United States are twice that of the general population and three times that of other professions. Certain medical specialties may have an increased suicide risk, including anesthetics, pharmacy, and psychiatry, although a U.K. study confirmed the excess mortality of doctors from suicide in doctors under 40, but found no difference by medical specialty. Female doctors are also three to four times more likely to commit suicide than females in the general population. An explanation of the elevated suicide rate for doctors is necessarily speculative. The availability of means for suicide (e.g., access to drugs), technical skills, and knowledge ensuring the act is successful may be relevant. Personality characteristics of doctors, including obsessive-compulsive behavior and proneness to depression, may also be important, along with high rates of alcohol and drug abuse and the stress of the doctor's occupational role.

Impact on Patient Care

Although the media image of the overworked, sleep-deprived doctor persists, it has proved methodologically difficult to investigate the impact of stress on doctors' job performance and patient care. Long working hours and sleep deprivation affect work performance in junior hospital doctors, being linked to lowered arousal, attentional deficits, and poor information recall, with fatigued subjects giving less consideration to risks they take. Mechanic's classic

studies of doctors related work frustration to short cuts in work performance, including inadequate examinations and prescribing. Prescribing behavior is one objective indicator of performance, and dysfunctional prescribing is related to job dissatisfaction in GPs. However, reaching agreement as to what constitutes efficient prescribing is difficult.

Summary

A review of studies investigating the extent and impact of occupational stress on doctors indicates that it is pervasive across specialties and cultures. It is therefore crucial for the future well-being of both doctor and patient to identify and manage stress in medical professionals. However, the measurement of stress and strain in this group, as in other occupational groups, has methodological problems. Measurement tools mainly comprise cross-sectional self-report questionnaire surveys and semistructured interview techniques. Response rates vary considerably, questioning representativeness, with some studies being small scale and parochial. The use of a wide range of measures and methods makes comparison among studies, specialties, and cultures difficult. Many studies employing standardized measures of mental ill health, such as the General Health Questionnaire, Maslach Burnout Inventory, and Hospital Anxiety and Depression Scale, report positive correlations between stress and psychological distress, particularly anxiety and depression.

However, few studies have addressed physical or behavioral aspects of the stress response or have considered individual differences as mediators/moderators of stressors/strains for medical professionals. This lack of methodological sophistication is an important omission. For example, where gender is taken into account, females, compared with male doctors, report greater job satisfaction, different behavioral outcomes (e.g., less alcohol and drug use), and a greater use of coping strategies. Similarly, dispositional factors such as neuroticism, which have been clearly associated with strain outcomes such as job dissatisfaction and mental ill health, are often not considered. A shift away from empirical/descriptive stimulus-response models toward more sophisticated transactional conceptualizations in the study of stress in medical professionals is required. This more sophisticated approach has implications for selection, training, and continued professional development of doctors, enabling doctors, managers, and educators to improve specific skills for coping with stress in those medical professionals who need them most. This has relevance for both medical professionals and the patients in their care.

See Also the Following Articles

Caregivers, Stress and; Workplace Stress.

Further Reading

- Agius, R. M., Blenkin, H., Deary, I. J., Zealley, H. and Wood, R. A. (1996). Survey of perceived stress and work demands of consultant doctors. *Occupational and Environmental Medicine* 53, 217–224.
- Cooper, C. L., Rout, U. and Faragher, B. (1989). Mental health, job satisfaction and job stress among general practitioners. *British Medical Journal* 298, 366–370.
- Deary, I. J., Blenkin, H., Agius, R. M., Endler, N. S., Zealley, H. and Wood, R. (1996). Models of job-related stress and personal achievements among consultant doctors. *British Journal of Psychology* 87, 3–29.
- Firth-Cozens, J. (1987). Emotional distress in junior house officers. *British Medical Journal* 295, 533–535.
- Grol, R., Mokkink, H., Smits, A., et al. (1985). Work satisfaction of general practitioners and the quality of patient care. *Family Practice* 2, 128–135.
- Howie, J. G. R., Hopton, J. L., Heaney, D. J. and Porter, A. M. D. (1992). Attitudes to medical care. The organization of work and stress among general practitioners. *British Journal of General Practice* 42, 181–185.
- Mechanic, D. (1972). General medical practice: some comparisons between the work of primary care physicians in the United States and England and Wales. In: *Public expectations and health care*. New York: Wiley.
- Richardsen, A. M. and Burke, R. J. (1991). Occupational stress and job satisfaction among physicians: sex differences. *Social Science and Medicine* 33, 1179–1187.
- Riska, E. and Wegar, K. (1993). Women physicians: a new force in medicine? In: Riska, E. & Wegar, K. (eds.) *Gender, work and medicine*. London: Sage.
- Sutherland, V. J. and Cooper, C. L. (1993). Identifying distress among general practitioners: predictors of psychological ill-health and job dissatisfaction. *Social Science and Medicine* 37, 575–581.
- Swanson, V., Power, K. G. and Simpson, R. J. (1998). Occupational stress and family life: a comparison of male and female doctors. *Journal of Occupational and Organizational Psychology* 71, 237–260.
- Winefield, H. R. and Anstey, T. J. (1991). Job stress in general practice: practitioner age, sex and attitudes as predictors. *Family Practice* 8, 140–144.

Meditation and Stress

J L Kristeller

Indiana State University, Terre Haute, IN, USA

© 2007 Elsevier Inc. All rights reserved.

Meditation: Overview and History

Types of Meditation

Theory and Mechanisms

Research Evidence and Clinical Application

Conclusions

Glossary

*Loving
kindness
meditation*
*Mantra
meditation*

A group of guided meditative practices directed toward cultivating compassion toward oneself and others.

Also mantram. A widely practiced form of concentrative meditation in which a word or short phrase is repeated, usually silently, to oneself. In traditional usage, the mantra carries with it spiritual significance. Om, referring to the presence of the absolute within, is one of the most familiar mantras.

Shamatha

Sustained quieting or calming of the mind; meditative practices that are directed toward doing so.

Vedanta

Refers to the Vedas, part of the foundation of Hinduism, as interpreted in the mid-1800s by the Indian mystic Sri Ramakrishna, who promulgated an ecumenical world view of religion and meditative practice. The Vedanta Society was founded by Swami Vivekenanda, his chief disciple, in the United States in 1893, and maintains a number of centers today.

*Vipassana
Yoga*

Insight or mindfulness meditation.

To yoke or link the mind/spirit and body. Meditation is one traditional means by which this occurs. Hatha yoga, referring to practice of physical postures, is one branch of the traditional Yogic system.

Meditation: Overview and History

The therapeutic use of meditation may be unique within contemporary psychology, as it draws explicitly on non-Western traditions to inform both the

practice and the conceptual understanding of meditation effects. Awareness of Hindu and Buddhist meditative and yogic traditions entered Europe and the United States from India, Japan, and China beginning in the 1800s, influencing the Theosophist movement, which adopted the term meditation to refer to a wide range of related practices. Interest was further sparked by the 1883 Parliament of the World's Religions; some Zen and Hindu-based practices in the United States can still be traced directly to this conference, through the influential writings of Alan Watts on Zen Buddhism and the Vedanta Society. The first documented secular adaptation of meditative practices to therapeutic use was Johannes Schultz's development of autogenic training in 1932, which used slow-paced breathing and mindful awareness of physical sensations to facilitate deep relaxation experiences. The writings of D. T. Suzuki on Zen practice stimulated interest in meditation in the 1950s by psychodynamic therapists, particularly Fromm and Horney, and fueled openness to more secular forms of meditation, such as transcendental meditation, a decade later. There is increasing recognition of similarities in meditative and contemplative practices across world religious traditions, including Christianity, Judaism, and Islam, in both process and purpose.

Types of Meditation

Meditative techniques are generally classified into two types: concentrative meditation and mindfulness meditation. Concentrative techniques involve sustained attentional focus on a specific object, such as a mantra (a repeated sound or word), with a goal of complete absorption, suspension of everyday preoccupations, and calming the mind, whereas mindfulness meditation cultivates sustained attentive awareness to whatever may emerge into conscious experience, without reaction or analysis. In addition, guided meditation, in which attention is directed to a particular type of experience (i.e., an emotion, such as anxiety, a physical feeling, such as pain, or a type of thought), can be considered a third category that may have particular value in regard to therapeutic effects. Tibetan Buddhist practices are particularly rich in complex guided meditations. However, these distinctions are a matter of degree: awareness cannot occur without sustained attention, and sustained attention or concentration cultivates awareness. Most traditional meditative approaches incorporate aspects of both concentrative and mindfulness practice. Focused concentration may be more powerful in inducing trancelike states. Yet, as active thinking and feeling processes are quieted, similar states of transcendent experience may be achieved through a

wide range of practices, as evident from phenomenological descriptions of meditative experience across traditions.

Concentrative Meditation

A wide variety of traditional concentrative techniques exist, but the most common involve focus on the breath (e.g., in Buddhist Shamatha meditation), focus on a word or phrase (e.g., the mantra), or visualization of an object (e.g., as in Tibetan mandala meditation). Transcendental meditation (TM), the most widely researched version of Hindu mantra or concentrative meditation that has been used therapeutically, was created in the 1960s for teaching in the West by the Maharishi Mahesh Yogi; other influential Hindu-based meditation groups in the United States include the Vedanta Society, founded in 1893, and the Himalayan Society, founded by Swami Rama in 1971. Recently, Catholic Centering Prayer, as being revived from early Christian practices by Father Thomas Keating, has garnered increasing attention, and is strikingly similar in many respects to other mantra meditations. In TM, the mantra is generally a Sanskrit word or phrase assigned by a TM teacher and is intended to have a spiritual, although not religious, meaning. TM practice involves sitting quietly in a relaxed posture for 20 min twice a day, with closed eyes. The mantra is repeated silently to oneself, with attention returned to it as thoughts wander to other experiences. Unlike some concentrative techniques, TM practitioners are instructed that attention is to be directed in an effortless manner.

Benson further secularized this practice in the 1970s for research and clinical applications, coining the term the relaxation response, and suggesting that practitioners use the word "one" as a mantra, but other words or phrases, whether of a secular or spiritual meaning, may be used. Eknath Easwaran, an Indian scholar and meditation teacher based in California, developed a transcultural approach that uses both mantra-based meditation and longer memorized textual passages; although he is deceased, his work is increasingly being utilized in treatment programs. These concentrative approaches have all been integrated into more comprehensive treatment programs.

Mindfulness Meditation

Mindfulness techniques have been introduced primarily through Vipassana or insight meditation drawn primarily from southeast Asian Theravadan practices (Thai, Vietnamese, and Burmese), Japanese Zen *shikantaza* ("just sitting"), and visualization and other techniques from the very rich Tibetan traditions.

Body-focused awareness is also cultivated in a number of traditions; the word yoga means to connect (or yoke) the mind and body. Influential venues by which these traditions entered the United States were a number of Zen monasteries founded beginning in the 1960s; the Naropa Institute in Boulder, CO, founded in 1974 by the Tibetan monk Chogyam Trungpa Rinpoche; and the Insight Meditation Society, founded in 1975 by Jack Kornfield, a psychologist who studied Buddhism in Thailand, and his associate Sharon Salzberg, among others. The prolific writings of Thich Nhat Hanh, a Vietnamese Zen monk, have been particularly influential in regard to loving kindness meditation and walking meditation. A contemporary Asian mindfulness meditation retreat program, developed by S. N. Goenka in Burma and India, uses a traditional 10-day intensive practice format and has been explored for use in prison populations and with drug addicts in both India and the United States.

The Mindfulness-Based Stress Reduction (MBSR) program, developed by Jon Kabat-Zinn in 1981 at the University of Massachusetts Medical Center, is the most widely used and researched therapeutic approach drawing on mindfulness meditation techniques; it was influenced by both Vipassana and Korean Zen traditions, in combination with yoga and body-focused meditation. The MBSR program uses the breath as an attentional focus but also emphasizes cultivating detached, nonjudgmental awareness of whatever arises into consciousness. Practice is generally a single sitting meditation session per day of 40–45 min, though this may be shortened as appropriate for use with children or other groups. Maintaining an erect posture is emphasized, consistent with a goal of staying alert. Engaging a trancelike state is avoided. Variations on the MBSR program include Mindfulness-Based Cognitive Therapy (MBCT), developed by Teasdale, Segal, and Williams, and Mindfulness-Based Eating Awareness Therapy (MB-EAT), developed by Kristeller. MBCT incorporates cognitive therapy techniques designed to address depressive thinking; MB-EAT incorporates cognitive and self-regulation approaches to management of compulsive overeating and obesity. Linehan's Dialectical Behavior Therapy incorporates brief mindfulness meditation practices into an intensive group treatment program for borderline personality disorder.

Theory and Mechanisms

Meditation involves the intentional cultivation of moment-to-moment, nonjudgmental awareness of one's present experience, whether narrowly or more

broadly focused. The goal of these practices is to develop an ability to bring stable and nonreactive awareness to one's internal (cognitive-affective-sensory) and external (social-environmental) experiences and thereby heighten the capacity to bring conscious control to responses and reactions. Within Buddhist psychology, meditation is identified as the most important means to bringing forth skillful action, speech, livelihood, effort, understanding, and thought. It is through meditation that higher wisdom in all areas of life is sought and achieved. These dimensions map fairly well onto the range of meditation effects that have been explored scientifically, including physical relaxation, emotional well-being, behavioral regulation, improvement in relation to self, growth in compassion toward others, and spiritual enlightenment or a sense of transcendence.

The psychology of meditation practice must therefore consider how it is that a group of relatively simple attentional and cognitive practices may bring about this range of effects. The means through which this is thought to occur within Buddhist psychology is again compatible with contemporary learning theory. Within contemporary Buddhist psychology, distress is viewed as a function of conditioning leading to cravings and aversions, which create *dukkha* (suffering or stress) and are incompatible with higher or wiser levels of functioning. The traditional goal of meditation practice is release from craving or conditioning. This is strikingly parallel to learning models of stress, as expressed in anxiety disorders, addictive behaviors, and depressed mood or cognitions.

Several fundamental mechanisms are likely to be involved in early and intermediate stages of practice. Advanced stages may engage more complex processes related to altered states of consciousness that are somewhat distinct, or that, as Austin posits, may be on a continuum with more fundamental shifts in neuropsychological functioning. First, focusing attention on a neutral repetitive object, whether it be the breath, a mantra, or a repeated chant, disengages attention from other targets that may be stress related. This process may produce stress-reducing effects similar to those of any distracter but is different in that the mind is not then caught up in an alternative source of attention. It is not uncommon with initial practice for individuals to note how profoundly relaxing they find the experience. The next step is cultivation of the ability to rest the attention on a neutral stimulus, such as the breath or the mantra, in a more sustained manner. This is intended to generalize to increased ability to sustain attention on other objects (sometimes referred to as bare attention). Linked to this process, and heightened in mindfulness meditation, is observing the occurrence of

patterns of habitual or conditioned responding, a type of reflective self-monitoring. Initial practice often brings awareness of the chattering mind, the undisciplined pattern of racing thoughts. Sustained attention then leads to an active disengagement of usual reactivity, analysis, or letting the attention move unbidden to the next association, emotion, or response. Simply observing experiences such as pain sensations, anxiety, or cravings, without reacting, may share similarities with systematic desensitization, in that repeated exposure, without either withdrawal or reinforcement, interrupts the feedback loop of the conditioning process. Therapeutic value may be amplified by guided meditations that utilize specific objects as targets for mindfulness practice, such as the experience of pain, fear, anxiety, eating, or compassion. Another value in maintaining sustained attention is similar to what Marlatt describes in the addiction literature as urge surfing: coming to the realization that by observing an experience or impulse in a sustained way, one discovers that these experiences are not constant, but rise and then fall away.

One hallmark of highly conditioned reactive responses is that they occur so quickly or are so powerful that alternatives cannot emerge or are not experienced as sustainable. As automatic reactive patterns are disengaged, the potential emerges for alternative choices that may have been lying dormant. Meditators often report that they observe the

alternative choices that arise as fresh and in some way unexpected, yet emerging from their own capabilities, rather than being directed or prescribed from the outside. Even at early stages of practice, individuals will note that it becomes easier to introduce a moment of silence or space into daily choices and that making a healthier decision no longer feels like as much of a struggle as it once did.

Figure 1 illustrates several of these steps in regard to mindfulness practice. Awareness is heightened in regard to stress responses (1a), such as anger or overeating (Y), the trigger (X) for that response, and the connections between the trigger and the response (1b). Resting attention with this awareness then acts to disengage or disconnect the trigger from the response (2a) and the feedback/reinforcement value (2b) from the response to the trigger. When these links are disconnected, access to alternative choices (A, B, or C) becomes available (3); with further practice, a novel or wiser response (W) may emerge (4).

Cultivating bare attention also appears to disengage a sense of self-identification with the experience, i.e., being able to say “I feel the pain, but I am not my pain.” This disidentification of self from experience is considered to play a primary mediational role within some meditative traditions; one counterpart in contemporary psychology is the recognition that the constructed self, along with critical self-judgment, is often central to distress and psychopathology. This

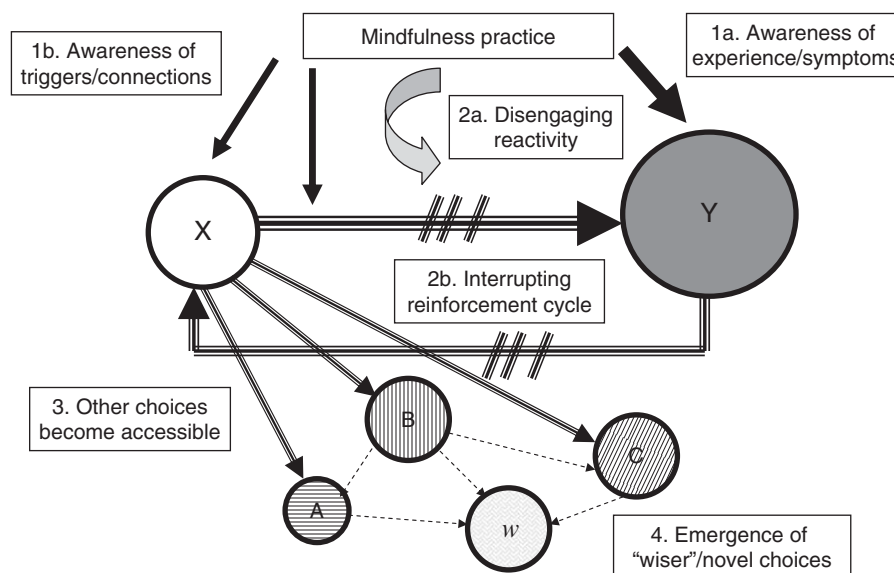


Figure 1 Illustration of several steps in regard to mindfulness practice. Awareness is heightened in regard to stress responses (1a), such as anger or overeating (Y), the trigger (X) for that response, and the connections between the trigger and the response (1b). Resting attention with this awareness then acts to disengage or disconnect the trigger from the response (2a) and the feedback/reinforcement value (2b) from the response to the trigger. When these links are disconnected, access to alternative choices (A, B, or C) becomes available (3); with further practice, a novel or wiser response (W) may emerge (4).

disidentification may be accompanied by a sense of liberation, a release from operating on automatic, whether from the usual chatter of the conscious mind or from more challenging patterns, such as anger, drinking, or compulsive eating.

Spiritual growth as a functional result of meditation practice may also occur due to disengagement of more immediate survival needs. Spiritual engagement is being increasingly recognized as a potent protective factor when individuals are under high stress conditions. While understanding the neuropsychological processes underlying spiritual or mystical experience is at an exploratory stage, meditative traditions universally aim to cultivate such experiences; disengagement of ordinary cognitive processing appears to be a necessary, if not sufficient, foundation for engaging spiritual or mystical experience.

Research Evidence and Clinical Application

Background

Contemporary research on meditation effects is generally dated from early studies in the 1950s on physiological effects in Indian yogis and Japanese Zen masters. Research on the clinical application of meditation is usually considered to date from the early work with TM or from the publication of Benson's *The Relaxation Response* in 1973. Murphy and Donovan's annotated 1996 bibliography summarizes more than 1000 studies on meditation.

It is useful to organize research evidence, both experimental and clinical, around the six domains noted earlier: cognitive/attentional, physiological/health, emotional, behavioral, relation to self/others, and spiritual. Clearly, these domains intersect and overlap, but such an approach helps in considering underlying mechanisms. For example, traditional perspectives on meditation practice have tended to focus on self-transformation and spiritual growth, yet it is apparent that substantial health and emotional effects in novice meditators can be obtained independently of such changes. The level of experience of research participants may vary from thousands of hours of meditation experience to individuals just being introduced to a given practice. Methodological quality varies greatly, and until recently, few studies met criteria for well-controlled randomized clinical trial research. Yet Grossman's meta-analytic review of 20 well-designed MBSR intervention trials found substantial effect size estimations of approximately 0.50, across various samples and problem areas. Neuroimaging techniques may prove particularly

valuable in identifying possible underlying processes, given the inherent introspective nature of the processes involved. Somewhat surprisingly, there are very few studies comparing concentrative and mindfulness meditation (or other meditation) techniques.

Individual differences have also been of interest; an early model suggested that individuals higher on cognitive anxiety might respond better to meditation techniques than to more body-centered relaxation interventions. However, several attempts to confirm this model either failed to do so or found that the variance accounted for was very small. Clinically, some individuals who drop out of meditation practice report that the experience is boring; conversely, others have difficulty with racing thoughts that they are unable to quiet. Beginning with shorter time periods, use of a mantra, or other techniques, such as breath counting, can help with these problems. Although the basics of meditation practice appear relatively simple, it is generally strongly emphasized that one should have a meaningful personal practice before using meditation therapeutically with others.

Efforts to use self-reporting to measure mindfulness have greatly accelerated in the past few years. Baer's factor analytic research has identified five facets of mindfulness that can be measured in both meditators and nonmeditators, including attending to experience, nonjudging, nonreactivity to inner experience, describing experience in words, and acting with awareness; this research is still in the formative stage. The revival of interest in spirituality as a psychological construct is also contributing to measurement of this aspect of meditation practice.

Cognitive and Attentional Effects

Most of the research on attentional and cognitive effects of meditation has investigated concentrative and TM approaches. Although evidence is not entirely consistent, there appears to be an improvement in field independence, reaction time, sustained attention, and ability to detect detail in rapidly presented material. Additional research is needed to investigate the role of meditation in affecting the cognitive processes hypothesized to mediate other effects.

To date, clinical application within this domain has been limited, but preliminary evidence suggests that there is value in using mindfulness meditation in treating attention-deficit disorder. Clinically, one of the notable cognitive effects is an experience of inner silence, of quieting the mind to the point that inner self-talk is not experienced. Beginning meditators often express frustration that they cannot silence their thoughts, not understanding that such silence is the result of practice and of the ability to disengage,

and is difficult, if not impossible, to force. The occurrence of such silence may be an important marker of a shift in neuropsychological functioning and deserves further investigation.

Physiological and Health Effects

Within the domain of physiological effects, the most robust research findings are related to a shift in balance of less sympathetic to greater parasympathetic autonomic nervous system function, accompanied by feelings of relaxation, decreases in blood pressure, and lowered levels of stress hormones. This shift is essentially what Benson has referred to as the relaxation response. There are also primary physiological effects due to slowed breathing; respiration rates of two to four cycles per minute have been shown in separate studies by Benson and by Lehrer to lead to marked physiological changes, including increases in body temperature, in Tibetan and Rinzai Zen monks, two traditions in which such slow-paced breathing is cultivated. Central nervous system effects have primarily focused on changes in alpha and theta rhythm dominance during practice. The increase in theta activity is sometimes interpreted as indicating that meditators are simply falling asleep; while this may occur, a pattern of sustained (nondescending) theta suggests a distinct state of consciousness. More recent work is documenting increased synchronization of activity across regions of the brain and heightened gamma activity in advanced meditators.

Impact on health through lowering of cortisol levels, lowering of blood pressure in both adults and adolescents, and improved management of pain and immune functioning deserve further exploration. Chronic pain patients have been shown to respond remarkably well to the rigorous combination of sitting meditation, gentle hatha yoga, and body awareness practices, with decreases in pain experience and increases in pain tolerance. A randomized clinical trial documented improvement in fibromyalgia patients, but these effects were not replicated using an attentional-support control comparison group. Evidence is limited regarding improvement in immune functioning in cancer patients enrolled in an MBSR program, but research by Carlson and colleagues found improvement in several indicators of immune response in breast and prostate cancer patients, including interleukin (IL)-4. They also found improved diurnal profiles in salivary cortisol, which is associated with survival time. A notable use of guided mindfulness meditation practice has been in conjunction with standard treatment of psoriasis, leading to markedly faster clearing rates of this chronic skin disease.

Emotional Effects

Meditation can be considered one of the few tools for systematic cultivation of emotional equanimity, an advanced level of stress and affect tolerance. Improvement in emotional well-being has been documented in multiple studies with both concentrative and mindfulness meditation practice, and in both novice and more advanced meditators. Reduced anxiety, anger, and depression and heightened experience of well-being have been documented in a number of studies involving both clinical and non-clinical populations. The MBSR program has been used successfully to treat individuals with diagnosed anxiety disorders, producing long-term effects for those with panic attacks and agoraphobia. There is a growing literature showing the value of MBCT in preventing relapse in individuals with three or more previous episodes of major depressive illness. The MBCT program integrates cognitive-behavioral components into the structure of the MBSR program, in addition to using guided meditations in regard to depressive thoughts.

Meditation practice, however, may not uniformly result in a positive emotional experience. Because one of the mechanisms involved is interrupting dominant conditioned responses, emotionally painful experiences or even flooding with uncovered memories of traumatic events may emerge when constructed mechanisms for suppression are disengaged. In the general population, the incidence of these experiences is very low, but this may increase substantially in psychiatric populations, requiring tailoring of the meditation practice and guided support from the therapist. The value of mindfulness techniques, including brief meditation, with borderline personality disorder is now well supported in research on Linehan's Dialectic Behavior Therapy. There is a growing clinical literature by such individuals as Epstein, Rubin, and Germer and colleagues on the integration of meditation practice into psychodynamic psychotherapy.

Meditation and Behavior

Improved behavioral regulation may be the result of several factors: the interruption of chains of behavioral reactions as awareness is cultivated, improved emotional regulation, increasing receptivity to behavioral/lifestyle recommendations, or learning to tolerate and ride out waves of craving, rather than responding impulsively. Much of the research literature focused on behavior change has investigated the effects of TM on reductions of drug, smoking, and alcohol intake, with claims made of substantial

improvement in long-term abstinence, depending on the amount and duration of practice. While suggestive, randomized clinical trials are needed to investigate the effects of both concentrative and mindfulness-based meditative interventions. The MB-EAT program has preliminary support of effectiveness in treating binge eating disorder, as does Dialectical Behavior Therapy.

Relation to Self and Others

Cultivating compassion toward oneself and toward others is central to Buddhism; cultivating the capacity to disengage from self-preoccupation is a central goal of meditative practice in all contemplative traditions. Overidentification of self with either positive or negative experience may contribute to stress and interfere with self-growth. Research evidence has supported the development of self-acceptance when under stress in more advanced meditators in both concentrative and mindfulness traditions. Such self-acceptance may come through disengaging from usual preoccupations and cultivation of insight, and may be heightened through use of loving kindness meditations. Increases in both self-growth and empathy have been observed in couples, therapists, and other health professionals enrolled in relatively brief (4- to 8-week) meditation programs.

Profound transformative experiences have been documented in incarcerated individuals enrolled in Goenka's week-long intensive meditation retreats in prisons in India and the United States. Evidence is accruing that advanced meditators practicing within Tibetan traditions that involve extensive guided meditation on cultivation of compassion exhibit strikingly higher left prefrontal cortex activity that underlies positive emotion. It is unclear whether such effects emerge inherently from intensive practice, are mediated through attentional and cognitive disengagement of self-preoccupation, or require domain-specific guided meditations, such as loving kindness practices.

Spiritual Well-Being

Recent research on spiritual well-being in medical populations suggests that it is a powerful factor in overall adjustment, despite high levels of stressors. Cultivation of spiritual growth and well-being is fundamental to virtually all meditative traditions, but different emphasis is placed on the centrality of achieving spiritual or transcendent experience across different traditions; furthermore, the degree to which this is viewed as a secular goal (as within TM or some schools of Buddhism and Hinduism) or as an inherently religious goal (as within Judeo-Christian

contemplative traditions) varies substantially. Research with novice meditators using both concentrative and mindfulness practices is beginning to document modest shifts in spiritual well-being, using psychometrically sound self-report measures.

Concentrative meditative practices are recognized to more readily cultivate trance states even in novice meditators, and induction of mystical experience is being examined in neuroimaging studies in highly experienced meditators. D'Aquili and Newberg's research on advanced meditators has garnered substantial attention in positing a neurophysiological basis for suspending a sense of the self and engaging the experience of union with a higher being.

Conclusions

Meditation practice, as it is currently understood, engages mechanisms that promote effects far broader than basic physical or emotional relaxation. To summarize, meditation can affect the stress response in a number of ways: first, it serves as a neutral distractor; second, mindfulness meditation teaches the capacity for self-observation of patterns of experience; third, the capacity to sustain awareness without responding or reacting is heightened. With yet more practice, conditioned reactions and responses to inner experience gradually weaken; finally, in the course of this uncoupling, meditation appears to facilitate the emergence of more integrative responses. Specific therapeutic goals may be facilitated by directing meditative awareness toward the target of concern, such as anxiety symptoms or pain management. By understanding meditation practice as engaging these processes, it becomes clearer how a range of therapeutic effects may occur as a function of a relatively simple set of practices, much as other principles, such as cognitive restructuring, behavioral techniques, or insight-oriented psychotherapies, are applied to a wide range of therapeutic issues.

Both concentrative and mindfulness approaches have developed and gained acceptance over the past 30 years as valuable components of treatment programs for a wide range of disorders and presenting issues. The literature on meditation, as it can be incorporated into therapeutic modalities, continues to expand. Research also continues to explore underlying mechanisms, in ways that may inform a greater understanding of self-regulatory processes in general. As the theoretical framework for understanding meditation psychotherapeutically evolves, it becomes more evident that meditation cannot be dismissed either as an esoteric spiritual/religious practice or as only one of many relatively simple relaxation techniques.

See Also the Following Articles

Cognitive Behavioral Therapy; Relaxation Techniques; Stress Management, CAM Approach.

Further Reading

- Austin, J. H. (1998). *Zen and the brain*. Cambridge, MA: MIT Press.
- Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J. and Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment* 13(1), 27–45.
- Benson, H. (1975/2000). *The relaxation response*. New York: Harper Press.
- Bodian, S. (1999). *Meditation for dummies*. Foster City, CA: IDG Books Worldwide, Inc.
- Davidson, R. J., Kabat-Zinn, J., Schumacher, J., Rosenkranz, M., Muller, D., Santorelli, S. F., et al. (2003). Alterations in brain and immune function produced by mindfulness meditation. *Psychosomatic Medicine* 65(4), 564–570.
- Epstein, M. (1995). *Thoughts without a thinker: psychotherapy from a Buddhist perspective*. New York: Basic Books, Inc.
- Germer, C. K., Siegel, R. D. and Fulton, P. R. (eds.) (2005). *Mindfulness and psychotherapy*. New York: Guilford Press.
- Goleman, D. (1977/1996). *The meditative mind. Varieties of meditative experience*. Los Angeles, CA: Tarcher Books.
- Grossman, P., Niemann, L., Schmidt, S. and Walach, H. (2004). Mindfulness-based stress reduction and health benefits. A meta-analysis. *Journal of Psychosomatic Research* 57(1), 35–43.
- Kabat-Zinn, J. (1990). *Full catastrophe living*. New York: Delacorte Press.
- Murphy, M., Donovan, S. and Taylor, E. (1999). *The physical and psychological effects of meditation: a review of contemporary research with a comprehensive bibliography, 1981–1996*. Sausalito, CA: Institute of Noetic Science.
- O’Connell, D. F. and Alexander, C. N. (eds.) (1994). Self-recovery: treating addictions using transcendental meditation and Maharishi Ayer-Veda. *Alcoholism Treatment Quarterly* 11(1–4), Special Issue.
- Segal, Z. V., Williams, J. M. G. and Teasdale, J. D. (2003). *Mindfulness-based cognitive therapy for depression. A new approach for preventing relapse*. New York: The Guilford Press.
- Travis, F. T. (2001). Transcendental meditation technique. In: Craighead, W. E. & Nemeroff, C. B. (eds.) *The Corsini encyclopedia of psychology and behavioral science* (3rd edn., pp. 1705–1706). New York: John Wiley & Sons.
- Walsh, R. and Shapiro, S. (2006). The meeting of meditative disciplines and Western psychology. *American Psychologist* 61, 227–239.
- Walton, K. G., Schneider, R. H., Nidich, S. I., et al. (2002). Psychosocial stress and cardiovascular disease. Part 2: Effectiveness of the Transcendental Meditation program in treatment and prevention. *Behavioral Medicine* 28(4), 106–123.

Melanocortin Receptors See: Feeding Circuitry (and Neurochemistry).
--

Membrane Glucocorticoid Receptors

P J Gasser and M Orchinik

Arizona State University, Tempe, AZ, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 713–720, © 2000, Elsevier Inc.

Glucocorticoid Receptors

Glucocorticoid Receptors in Liver

Glucocorticoid Receptors in Lymphocytes

Glucocorticoid Receptors in Adrenal Gland

Glucocorticoid Receptors in Pituitary

Glucocorticoid Receptors in Brain

Glucocorticoid Transport Proteins

Corticosteroid-Binding Globulins

Signal Transduction by Membrane Glucocorticoid Receptors

Glossary

<i>Corticosteroid-binding globulin (CBG or transcortin)</i>	A plasma protein that specifically binds corticosteroids.
<i>Corticotropin releasing factor (CRF)</i>	A 41 amino acid residue hypothalamic peptide that mediates the neural control of pituitary adrenocorticotropin secretion.
<i>Dexamethasone</i>	A synthetic glucocorticoid.
<i>EC₅₀</i>	The dose of a hormone or drug necessary to produce 50% of the maximal response.
<i>GABA</i>	γ -Aminobutyric acid, an inhibitory neurotransmitter.
<i>GABA_A receptor</i>	A membrane protein complex that forms an agonist-gated chloride ion channel and mediates fast inhibitory synaptic transmission.
<i>K_D</i>	Equilibrium dissociation constant. The concentration of ligand at which 50% of the receptors are occupied.
<i>Nongenomic</i>	A mechanism of steroid hormone action that does not initially involve modulation of gene transcription.
<i>RU38486 (Mifepristone)</i>	A synthetic progesterone receptor antagonist that binds with high affinity to the cytosolic glucocorticoid receptor.
<i>TBPS (t-Butylbicyclophosphorothionate)</i>	A competitor for binding of GABA to the GABA _A receptor.
<i>Triamcinolone acetate</i>	A synthetic glucocorticoid.

Glucocorticoid Receptors

Cellular and organismal responses to glucocorticoid hormones are complex and integrative, involving immediate, short-term, and long-term components. Target cell responses to glucocorticoids are initiated by two main mechanisms. In the classical, genomic mechanism, glucocorticoids interact with intracellular receptors, which act as ligand-gated transcription factors. The ligand–receptor complex then regulates the transcription of specific genes by interaction with hormone response elements and/or AP-1 transcriptional regulatory sites. This mechanism produces responses that are usually apparent in hours or days.

In the other, nongenomic, mechanism, glucocorticoids initiate immediate or very short latency responses in target cells. These short latencies preclude the genomic mechanism of transcriptional activation/inhibition and altered translation, although genomic responses may follow rapid responses. It is believed that these responses are mediated by glucocorticoid receptors that are integral to or associated with target cell plasma membranes. Membrane-associated glucocorticoid receptors have not been cloned, but several membrane-associated glucocorticoid-binding proteins have been purified, and there are numerous reports of nongenomic physiological and behavioral responses to the hormones (Table 1). Receptor binding studies have characterized glucocorticoid binding to plasma membranes in various tissues from several species (Table 2), and some of the cellular signaling mechanisms activated by the hormones have been elucidated. Multiple classes of glucocorticoid-binding sites have been revealed in plasma membranes: (1) membrane-associated glucocorticoid receptors distinct from intracellular glucocorticoid receptors, (2) membrane-associated forms of intracellular glucocorticoid receptors, (3) membrane receptors for corticosteroid-binding globulins (CBG), (4) membrane glucocorticoid-transporting proteins, and (5) other membrane proteins that also bind glucocorticoids. These proteins have been distinguished pharmacologically, based on binding specificity profiles for various natural and synthetic ligands, and immunologically, based on cross-reactivity with antibodies for intracellular glucocorticoid receptors. This article reviews the membrane glucocorticoid receptors and describes what is known of the mechanisms by which they initiate rapid cellular/organismal responses.

Glucocorticoid Receptors in Liver

The liver, a major target organ for glucocorticoid hormones during stress, shows both immediate and long latency physiological responses to glucocorticoid hormones and has both cytosolic and membrane glucocorticoid receptors. Radioligand binding and photo-affinity labeling reveal high-affinity ($K_D = 7\text{--}12\text{ nM}$) and low-affinity ($K_D = 90\text{--}344\text{ nM}$) membrane-binding sites for both dexamethasone and cortisol (see Table 2) in rat and chicken hepatocytes. The sites appear to be integral membrane proteins, as they can be extracted by detergent, but not by high salt. The sites for the two glucocorticoids also appear to be distinct from each other. Low-affinity dexamethasone binding is attributed to a 45-kDa protein, whereas two membrane proteins, 52 and 57 kDa, are responsible for cortisol binding. Glucocorticoid-binding sites have also been demonstrated visually in rat hepatocyte membranes. Gold-conjugated corticosterone-bovine serum albumin (BSA) specifically labels the apical, but not the lateral, membranes of hepatocytes (see Table 2).

During stress, glucose stores in the liver are mobilized rapidly by the activation of glycogen-degrading enzymes in the hepatocytes. This effect is largely due to the actions of catecholamines, but glucocorticoids also play a role. Corticosterone, dexamethasone, and cortisol increase the activity of glycogen phosphorylase, a glucose-liberating enzyme, in rat and chicken hepatocytes within 10 min of administration (see Table 1). Glucocorticoids act in this case via a Ca^{2+} -mediated signaling process, as the effect is diminished in Ca^{2+} -depleted medium. Other rapid effects of glucocorticoids in hepatocytes include increases in the phosphorylation of histone and nonhistone proteins and general activation of translational machinery, including increased synthesis of intracellular glucocorticoid receptors (see Table 1). These rapid effects may facilitate long-term, genomic responses to glucocorticoids.

Glucocorticoid Receptors in Lymphocytes

The immune system is another major target of glucocorticoid hormones. Membrane-mediated effects of glucocorticoids and membrane-associated forms of the intracellular glucocorticoid receptor have been demonstrated in various lymphocytes. In rat mast cells and basophils, cortisol inhibits the antigen-induced release of histamine and attenuates the antigen-induced release of Ca^{2+} from intracellular stores (see Table 1). This effect may be related to the potent antiallergic actions of glucocorticoids.

Studies with antibodies have demonstrated that a population of human lymphoma cells possesses a

membrane-associated form of the intracellular glucocorticoid receptor (see Table 2). Cells that possess this form of the receptor are particularly sensitive to glucocorticoid-induced cytolysis.

Glucocorticoid Receptors in Adrenal Gland

Both the adrenal cortex and the adrenal medulla are targets of rapid glucocorticoid hormone action. Rapid effects in both areas appear to be related to the feedback regulation of adrenal hormone release. Corticosterone decreases the secretion of adrenal steroids from frog interrenal gland by 45–80%. This effect is not mimicked by dexamethasone and is not blocked by the specific glucocorticoid receptor antagonist RU43044 (see Table 1), indicating that this ultra-short feedback regulation of adrenal steroid secretion is mediated by another type of glucocorticoid receptor.

A membrane glucocorticoid-binding site in bovine adrenal cortex has high affinity ($K_D = 77\text{ nM}$) for corticosterone and very low affinity ($\text{IC}_{50} = 25\text{--}50\text{ }\mu\text{M}$) for dexamethasone, triamcinolone, and cortisol (see Table 2).

Chromaffin cells of the adrenal medulla secrete catecholamines in response to a variety of neurotransmitters, including acetylcholine. These cells are exposed to very high concentrations of glucocorticoids in venous blood from the adrenal cortex. At high concentrations (up to $10\text{ }\mu\text{M}$), glucocorticoids modulate the responses of adrenal chromaffin cells to acetylcholine and to K^{+} -induced depolarization rapidly (see Table 1). Dexamethasone immediately diminishes the acetylcholine-induced current when applied extracellularly to dispersed guinea pig chromaffin cells, but has no effect when applied intracellularly. In PC12 cells from rat medullary tumors, corticosterone and corticosterone-BSA inhibit depolarization-induced intracellular Ca^{2+} release by a mechanism that involves G-protein-mediated calcium signaling.

Glucocorticoid Receptors in Pituitary

Several hypothalamic peptide hormones can cause cells of the anterior pituitary to release adrenocorticotrophic hormone (ACTH), which in turn stimulates the release of glucocorticoid hormones from the adrenal cortex. Glucocorticoids regulate pituitary production/secretion of ACTH in a complex system of feedback inhibition. This inhibition occurs on rapid (seconds to minutes), delayed intermediate (10–15 min to hours), and delayed slow (hours to days) time courses. Rapid feedback inhibition of pituitary ACTH secretion by

Table 1 Selected membrane-mediated effects of glucocorticoids

Hormone	Effect	Tissue	Dose (EC_{50}) ^a	Organism	Reference
Liver					
Corticosterone/ dexamethasone	↑ glycogen phosphorylase activity	Hepatocytes	1–10 μ M B; 10–1000 nM Dex	Rat	Gomez-Munoz et al. (1989). <i>Biochem. J.</i> 262 , 417–423
Cortisol	↑ glycogen phosphorylase activity	Hepatocytes		Chicken	Sancho et al. (1986). <i>Biochim. Biophys. Acta</i> 88 , 249–255
Adrenal					
Corticosterone	↓ depolarization-induced calcium release	PC12 cells	100 nM–10 μ M	Human	Lou and Chen (1998). <i>Biochem. Biophys. Res. Commun.</i> 244 , 403–407
Dexamethasone	↓ acetylcholine-induced current	Chromaffin cells	10 μ M	Guinea pig	Inoue and Kuriyama (1989). <i>Am. J. Physiol.</i> 257 , C906–C912
Corticosterone	↓ steroid release	Interrenal tissue	10 μ g/ml	Frog	Netchitailo et al. (1991). <i>J. Mol. Endocrinol.</i> 6 , 249–255
Intestine					
Cortisol	↑/↓ GABA-induced contractility	Ileum	1–10 pM (↑); 10–1000 nM (↓)	Guinea pig	Ong et al. (1987). <i>Neurosci. Lett.</i> 82 , 101–106.
Endometrium					
Dexamethasone	↑ actin assembly	Endometrial cells	0.1 μ M	Human	Koukouritaki et al. (1997). <i>J. Cell. Biochem.</i> 65 , 492–500
Immune system					
Cortisol	↓ histamine release	Mast cells	3 μ M	Rat	Dukhanin et al. (1996). <i>Inflamm. Res.</i> 45 , S15–S16
Pituitary					
Corticosterone, cortisol, dexamethasone	↓ CRF-stimulated ACTH release	Pituitary	100 nM	Rat	Widmaier and Dallman (1984). <i>Endocrinology</i> 115 , 2368–2374
Dexamethasone	↓ AVP-stimulated ACTH release	Anterior pituitary	1 μ M	Rat	Nicholson and Gillham (1989). <i>J. Endocrinol.</i> 122 , 545–551
Corticosterone	↓ CRF-stimulated ACTH release	Pituitary	45 μ g/kg	Dog	Keller-Wood and Bell (1988). <i>Am. J. Physiol.</i> 255 , R344–R349
Cortisol/dexamethasone	↓ prolactin release	Rostral pars distalis	10 nM–1 μ M	Tilapia	Borski et al. (1991). <i>Proc. Natl. Acad. Sci. USA</i> 88 , 2758–2762
Brain/neuronal tissue					
Corticosterone	↑ Ca^{2+} uptake	Synaptosomes	130 nM	Rat	Sze and Iqbal (1994). <i>Neuroendocrinology</i> 59 , 457–465
Cortisol/corticosterone	↑ tryptophan uptake ↓ calcium current	Hippocampal CA1 neurons	1 pM–100 nM (F) 10 pM–100 nM (B)	Guinea pig	Neckers and Sze (1975). <i>Brain Res.</i> 93 , 123–132 ffrench-Mullen (1995). <i>J. Neurosci.</i> 15 , 903–911
Corticosterone	↓ vasopressin release	Hypothalamic slices	100 nM–100 μ M	Rat	Liu and Chen (1995). <i>Brain Res.</i> 704 , 19–22
Corticosterone	↑/↓ somatostatin release	Hypothalamic neurons	100 nM–1 μ M	Rat	Estupina et al. (1997). <i>Exp. Brain Res.</i> 113 , 337–342

Cortisol/corticosterone	↓ CRF release	Hypothalamic synaptosomes	750 nM (B) 6–60 nM (F)	Sheep	Edwardson and Bennett (1974). <i>Nature</i> 251 , 425–427
Cortisol	↑/↓ single-unit activity	Medial preoptic septal neurons	Iontophoresis	Rat	Kelly et al. (1977). <i>Exp. Brain Res.</i> 30 , 53–64
Corticosterone	↑/↓ firing rate	Pontine reticular neurons	2.5–3 mg/kg	Rat	Dubrovsky et al. (1985). <i>J. Neurosci. Res.</i> 14 , 117–128
Corticosterone	↑/↓ spontaneous and evoked activity	Hypothalamus, amygdala, midbrain	25–33 mg/kg	Rat	Feldman et al. (1983). <i>Exp. Neurol.</i> 80 , 427–438
Corticosterone	↓ population spike amplitude	Hippocampal pyramidal cells	100 nM–10 μM	Rat	Vidal et al. (1986). <i>Brain Res.</i> 383 , 54–59
Corticosterone	↓ population spike amplitude	Hippocampal pyramidal cells	0.05–5 nM	Rat	Talmi et al. (1992). <i>Neuroendocrinology</i> 55 , 257–263
Corticosterone	↑/↓ amplitude of sciatic-evoked responses	Pontine reticular neurons	2.5–3 mg/kg	Rat	Dubrovsky et al. (1985). <i>J. Neurosci. Res.</i> 14 , 117–128
Corticosterone/cortisol	↑/↓ activity	Hypothalamic neurons	Iontophoresis	Rat	Saphier and Feldman (1988). <i>Brain Res.</i> 453 , 183–190
Cortisol	↑/↓ sciatic-evoked responses	Hypothalamic neurons (MBH)	Iontophoresis	Rat	Mandelbrod et al. (1981). <i>Brain Res.</i> 218 , 115–130
Corticosterone/cortisol	↑/↓ TBPS binding	Synaptosomes	10–100 nM (↑); >100 nM (↓)	Rat	Majewska et al. (1985). <i>Brain Res.</i> 339 , 178–182
Cortisol	↓ GABA-induced depolarization	Spinal ganglion neurons	5 μM–1 mM	Bullfrog	Majewska (1987). <i>Brain Res.</i> 418 , 377–382 Ariyoshi and Akasu (1986). <i>Brain Res.</i> 367 , 332–336 Ariyoshi and Akasu (1987). <i>Brain Res.</i> 435 , 241–248
Corticosterone	↑/↓ activity	Medullary neurons	11–25 μg/animal	Newt	Rose et al. (1993). <i>Neuroendocrinology</i> 57 , 815–824 Rose et al. (1995). <i>Neuroendocrinology</i> 62 , 406–417
Behavioral effects					
Corticosterone	↓ courtship behavior		0.6 nM	Newt	Orchinik et al. (1991). <i>Science</i> 252 , 1848–1851 Moore and Miller (1984). <i>Horm. Behav.</i> 18 , 400–410
Corticosterone	↑ perch hopping		26 ng/ml	Sparrow	Breuner et al. (1998). <i>Gen. Comp. Endocrinol.</i> 111 , 386–394
Corticosterone	↑ locomotor response to novelty		2.5–5 mg/kg	Rat	Sandi et al. (1996). <i>Eur. J. Neurosci.</i> 8 , 794–800
Corticosterone	Stimulation of behavior during social challenge		2 mg/kg	Rat	Haller et al. (1997). <i>J. Neuroendocrinol.</i> 9 , 515–518
Corticosterone	↑ lordosis		200 μg/animal	Rat	Kubli-Garfias (1990). <i>Horm. Behav.</i> 24 , 443–449
Cortisol	↑/↓ aggression		1–10 μM	Hamster	Hayden-Hixson et al. (1991). <i>J. Neuroendocrinol.</i> 3 , 613–622

^aB, corticosterone; F, cortisol; Dex, dexamethasone.

Table 2 Selected glucocorticoid-binding sites associated with plasma membranes

<i>Hormone</i>	<i>K_D</i>	<i>Organism</i>	<i>Reference</i>
Liver			
Cortisol	12 nM, 344 nM	Rat	Ibarrola et al. (1996). <i>Biochim. Biophys. Acta</i> 1284 , 41–46 ^a
Dexamethasone	7 nM, 90 nM	Rat	Howell et al. (1989). <i>Biochem. J.</i> 260 , 435–441 ^a
Cortisol	1.4 nM	Rat	Suyemitsu and Terayama (1975). <i>Endocrinology</i> 96 , 1499–1508
Corticosterone	None available	Rat	Spindler et al. (1991). <i>J. Steroid Biochem. Mol. Biol.</i> 39 , 315–322
Cortisol	4.5 nM	Chicken	Trueba et al. (1987). <i>Int. J. Biochem.</i> 19 , 957–962
Pituitary			
Corticosterone	3.2 nM	Rat	Koch et al. (1978). <i>J. Endocrinol.</i> 79 , 215–222
Corticosterone	None available	Mouse (AtT-20 cells)	Harrison et al. (1979). In: Leavitt and Clark (eds.) <i>Steroid hormone receptor systems</i> , pp. 423–443. New York: Plenum
Adrenal cortex			
Corticosterone	77 nM	Calf	Andres et al. (1997). <i>Cell. Mol. Life Sci.</i> 53 , 673–680
Brain			
Corticosterone	120 nM	Rat (synaptosomes)	Towle and Sze (1983). <i>J. Steroid Biochem.</i> 18 , 135–143
Corticosterone	6 nM (via CBG)	Vole (synaptosomes)	Orchinik et al. (1997). <i>J. Steroid Biochem. Mol. Biol.</i> 60 , 229–236
Corticosterone	0.5 nM	Newt (synaptosomes)	Orchinik et al. (1991). <i>Science</i> 252 , 1848–1851
Corticosterone	24.2 nM	Sparrow (synaptosomes)	Breuner et al. (1997). <i>Soc. Neurosci. Abst.</i> 23 , 1076
Lymphocytes			
Corticosterone	None available	Human (leukemic cells)	Gametchu et al. (1993). <i>FASEB J.</i> 7 , 1283–1292

^aIndicates that the glucocorticoid-binding protein has been purified.

glucocorticoids is thought to be mediated by membrane receptors. The stimulation of ACTH release by CRF, oxytocin, vasopressin, and insulin is diminished rapidly by glucocorticoids, including corticosterone, cortisol, and dexamethasone (see Table 1). These effects are not completely understood, but appear to involve the attenuation of secretagogue-induced increases in second messenger accumulation. Some fast feedback effects, with latencies of 15–20 min, have been attributed to rapid protein synthesis.

Glucocorticoid-binding sites have been characterized in plasma membranes from rat and mouse pituitary cells (see Table 2). The rat membrane receptor has high affinity ($K_D = 3.2$ nM) for corticosterone and does not bind dexamethasone. Mouse pituitary tumor (AtT-20) cell membranes specifically bind corticosterone, but not dexamethasone or triamcinolone. This binding site appears to be a glycoprotein, as treatment of cells with proteases abolishes glucocorticoid binding and treatment with neuraminidase enhances binding.

The release of other pituitary hormones can also be modulated by glucocorticoids. In pituitaries of the teleost fish tilapia, cortisol inhibits the release of prolactin rapidly, apparently via the membrane-mediated modulation of Ca^{2+} and cAMP levels (see Table 1).

Glucocorticoid Receptors in Brain

Stress effects on neuronal activity and animal behavior are widely varied, stressor specific, and dependent

on the endocrine/sensory history of the organism. Rapid glucocorticoid actions on neurophysiology and behavior show the same variability. In the brain, glucocorticoids can modulate both spontaneous and evoked neuronal activity rapidly. The direction of this modulation (increase or decrease) depends on the identity of the target neuron and is often context dependent. In the rat hypothalamus, for instance, corticosterone increases the activity of oxytocin-secreting neurons, but is mainly inhibitory to vasopressin neurons and neurons projecting to the median eminence (see Table 1). In the newt, corticosterone inhibits the responses of naive medullary reticular neurons to sensory stimuli, but it potentiates sensory responses of the same neurons when they have been previously exposed to vasotocin (see Table 1). Glucocorticoid-induced increases in sparrow perch-hopping behavior are dependent on photoperiod.

The best correlations between glucocorticoid-binding characteristics of membrane receptors and doses of glucocorticoids necessary for rapid physiological/behavioral effects have been demonstrated in the brains of rats and newts. In rat brain synaptosomes, corticosterone causes an increased uptake of Ca^{2+} and tryptophan (see Table 1). The EC_{50} for this effect (130 nM) is almost identical to the K_D (120 nM) determined for glucocorticoid binding to synaptic plasma membranes from the rat brain. In newts, corticosterone causes a rapid inhibition of male courtship behavior (see Table 1). This effect has an EC_{50} of 0.6 nM and is not mimicked by dexamethasone.

The K_D for corticosterone binding to newt synaptic plasma membranes is 0.5 nM, and dexamethasone does not compete with corticosterone for this binding site.

Other effects of glucocorticoids in the brain include rapid modulation of neuropeptide release by hypothalamic neurons, modulation of spontaneous neuronal activity in various brain regions, and modulation of neuronal responses to sensory stimuli (see [Table 1](#)). It is likely that these rapid neuronal responses to glucocorticoids mediate rapid changes in behavior during stress.

Glucocorticoid Transport Proteins

There is still debate over the mechanisms by which glucocorticoid hormones cross plasma membranes. Certain membrane glucocorticoid-binding proteins also transport the steroids across the plasma membrane. One major membrane steroid transporter, P-glycoprotein, binds corticosterone, cortisol, and dexamethasone and transports them out of cells. This protein is expressed at high levels in epithelia of adrenal cortex, kidney proximal tubule, biliary hepatocytes, brain capillaries, small intestine, and colon.

Rat liver plasma membranes possess a bidirectional glucocorticoid transporter that is specific for corticosterone and cortisol. Immunochemical methods demonstrate that this binding site is distinct from P-glycoprotein. Specific transmembrane transport of glucocorticoids has also been demonstrated in membrane vesicles from pituitary and placenta.

Corticosteroid-Binding Globulins

Of the pool of glucocorticoid hormones circulating in the blood, a substantial fraction occurs bound to plasma proteins. Adrenal steroids (cortisol, corticosterone, and deoxycorticosterone) bind specifically and with high affinity to corticosteroid-binding globulin (or transcortin). While its functions are still unclear, CBG does interact specifically with a plasma membrane receptor. Specific, high-affinity CBG receptors have been purified from plasma membranes of rat and human liver and human endometrium and trophoblast cells. The interaction of the CBG–cortisol complex with trophoblast cell membranes causes activation of adenylate cyclase and accumulation of cAMP.

High affinity ($K_D = 6$ nM) binding of corticosterone to vole synaptic plasma membranes appears to involve one form of CBG. Specific binding to vole membranes was abolished by the saline perfusion of

brains before binding studies, and vole plasma displayed a corticosterone-binding profile similar to that of synaptic plasma membranes (see [Table 2](#)).

Signal Transduction by Membrane Glucocorticoid Receptors

Pharmacological manipulations of nongenomic glucocorticoid effects (see [Table 1](#)) have revealed some aspects of the signal transduction mechanisms by which the hormones modulate cell physiology rapidly. These signaling mechanisms are nearly as diverse as the hormone effects. Existing data indicate that glucocorticoids can initiate rapid responses via (1) modulation of intracellular Ca^{2+} , often via G-protein-mediated pathways, and (2) activation/inhibition of other signal transduction systems.

Modulation of Calcium Signaling

A large body of evidence indicates that glucocorticoids initiate rapid responses in many target cells by the modulation of intracellular calcium signaling. Cortisol inhibits N- and L-type Ca^{2+} currents in hippocampal neurons via a pertussis toxin-sensitive G-protein and protein kinase C (see [Table 1](#)). In rat brain synaptosomes, corticosterone enhances depolarization-induced Ca^{2+} influx rapidly and potentiates calmodulin-induced activation of voltage-sensitive calcium channels (see [Table 1](#)). Fish pituitary cells, rat lymphocytes, and PC12 cells respond to glucocorticoid treatment with rapid decreases in intracellular Ca^{2+} accumulation. In PC12 cells, this effect is diminished by pertussis toxin and by inhibitors of protein kinase C.

Many rapid effects of glucocorticoids are modulated by changes in extracellular calcium concentration. For example, the corticosterone-induced inhibition of vasopressin release from hypothalamic slices is potentiated in Ca^{2+} -rich medium and decreased in Ca^{2+} -depleted medium (see [Table 1](#)). The same manipulation of extracellular Ca^{2+} concentration attenuates/enhances the corticosterone-induced decrease in population spike amplitude in hippocampal neurons (see [Table 1](#)). The corticosterone-induced increase in hepatocyte glycogen phosphorylase activity is diminished in Ca^{2+} -depleted medium.

Other Signaling Pathways

A variety of other signal transduction pathways have been implicated in the mediation of rapid glucocorticoid actions. The diversity of these pathways reflects the diversity of rapid glucocorticoid actions. Roles have been proposed for rapid glucocorticoid signaling via the GABA_A receptor system, the cytoskeleton, the

cAMP and cGMP pathways, and the nitric oxide system. Certain rapid effects of glucocorticoids result from the interaction of the hormones with other membrane proteins. One such protein is the GABA_A receptor, a membrane protein complex that forms an agonist-gated chloride channel. Most studies indicate that metabolites of adrenal steroids, rather than the hormones themselves, modulate GABA_A receptor function. However, there are reports of interactions between glucocorticoids and GABA_A receptors. Glucocorticoids alter physiological responses to GABA and modulate agonist binding to the GABA_A receptor (see [Table 1](#)). Other intracellular signaling pathways have been implicated in mediating the rapid effects of glucocorticoids. In human endometrial cells, dexamethasone induces rapid actin polymerization without concomitant changes in actin synthesis, implicating the cytoskeleton as another rapid glucocorticoid effector system (see [Table 1](#)). Cortisol inhibits cAMP phosphodiesterase in bovine testis and cardiac muscle homogenates and enhances guanylate cyclase activity in various tissues, indicating other possible mechanisms of rapid glucocorticoid signaling. Inhibitors of nitric oxide synthase prevent corticosterone-induced increases in the locomotor activity of rats in a novel environment (see [Table 1](#)).

See Also the Following Articles

Corticosteroid-Binding Globulin (Transcortin); GABA (Gamma Aminobutyric Acid); Glucocorticoids, Overview.

Further Reading

- Antoni, F. A., Hoyland, J., Woods, M. D. and Mason, W. T. (1992). Glucocorticoid inhibition of stimulus-evoked adrenocorticotrophic release caused by suppression of intracellular calcium signals. *Journal of Endocrinology* **133**, R13–R16.
- Fant, M. E., Harbison, R. D. and Harrison, R. W. (1979). Glucocorticoid uptake into human placental membrane vesicles. *Journal of Biological Chemistry* **254**, 6218–6221.
- Gee, K. W., McCauley, L. D. and Lan, N. C. (1996). A putative receptor for neurosteroids on the GABA_A receptor complex: the pharmacological properties and therapeutic potential of epalons. *Critical Review of Neurobiology* **9**, 207–227.
- Grote, H., Ioannou, I., Voigt, J. and Sekeris, C. E. (1993). Localization of the glucocorticoid receptor in rat liver cells: evidence for plasma membrane bound receptor. *International Journal of Biochemistry* **25**, 1593–1599.
- Harrison, R. W. and Yeakley, J. (1979). Evidence for glucocorticoid transport into AtT-20 cells. *Molecular and Cellular Endocrinology* **15**, 13–18.
- Howie, J. G. R., Hopton, J. L., Heaney, D. J. and Porter, A. M. D. (1992). Attitudes to medical care. The organization of work and stress among general practitioners. *British Journal of General Practice* **42**, 181–185.
- Lackner, C., Daufeldt, S., Wildt, L. and Allera, A. (1998). Glucocorticoid-recognizing and-effector sites in rat liver plasma membrane: kinetics of corticosterone uptake by isolated membrane vesicles. III. Specificity and stereospecificity. *Journal of Steroid Biochemistry and Molecular Biology* **64**, 69–82.
- Majewska, M. D. (1987). Antagonist-type interaction of glucocorticoids with the GABA receptor-coupled chloride channel. *Brain Research* **418**, 377–382.
- Mechanic, D. (1972). General medical practice: some comparisons between the work of primary care physicians in the United States and England and Wales. In: *Public expectations and health care*. New York: Wiley.
- Orchinik, M., Moore, F. L. and Rose, J. D. (1994). Mechanistic and functional studies of rapid corticosteroid actions. *Annals of the New York Academy of Sciences* **746**, 101–114.
- Richardsen, A. M. and Burke, R. J. (1991). Occupational stress and job satisfaction among physicians: sex differences. *Social Science and Medicine* **33**, 1179–1187.
- Riska, E. and Wegar, K. (1993). Women physicians: a new force in medicine? In: Riska, E. & Wegar, K. (eds.) *Gender, work and medicine*. London: Sage.
- Rose, J. D., Kinnaid, J. R. and Moore, F. L. (1995). Neurophysiological effects of vasotocin and corticosterone on medullary neurons: implications for hormonal control of amphibian courtship behavior. *Neuroendocrinology* **62**, 406–417.
- Schmidtke, J., Wienker, T., Flugel, M. and Engel, W. (1976). *In vitro* inhibition of cyclic AMP phosphodiesterase by cortisol. *Nature* **262**, 593–594.
- Strel'chyonok, O. A. and Avvakumov, G. V. (1991). Interaction of human CBG with cell membranes. *Journal of Steroid Biochemistry and Molecular Biology* **40**, 795–803.
- Sutherland, V. J. and Cooper, C. L. (1993). Identifying distress among general practitioners: predictors of psychological ill-health and job dissatisfaction. *Social Science and Medicine* **37**, 575–581.
- Swanson, V., Power, K. G. and Simpson, R. J. (1998). Occupational stress and family life: a comparison of male and female doctors. *Journal of Occupational and Organizational Psychology* **71**, 237–260.
- Winefield, H. R. and Anstey, T. J. (1991). Job stress in general practice: practitioner age, sex and attitudes as predictors. *Family Practice* **8**, 140–144.

Memory and Stress

S J Lupien

McGill University and Geriatric Institute of Montreal,
Montreal, Canada

F S Maheu

National Institute of Mental Health, Bethesda, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by S J Lupien and S Brière, volume 2, pp 721–727,
© 2000, Elsevier Inc.

Emotion, Stress, and Memory

Emotion and Stress: Two Different Entities?

Memory for Emotionally Arousing Events

Memory after Stress

Conclusion

Glossary

<i>Acute stress</i>	Mental or bodily tension resulting from internal or external factors that tend to alter the existing equilibrium.
<i>Arousal</i>	A condition that varies in a continuum from a low point in sleep to a high point represented by intense excitement.
<i>Emotion</i>	A state of excited feeling of any kind. Basic emotions presumably are universal, are biologically based, are present from infancy, and have associated facial expressions and primitive action scripts. Basic emotions include love, joy, anger, sadness, and fear.
<i>Stress response</i>	The total set of reactions to the stressor.
<i>Stressor</i>	Any physical or psychological agent that causes the disequilibrium. Physical stressors are situations that can operate in the absence of consciousness and include disturbances of the internal environment (anoxia, hypoglycemia, etc.), external extremes (heat and cold), and physical strain (exercise or injury). Psychological stressors are situations for which there is a consciously perceived potential for harm and include stimuli that affect emotion and result in fear, anxiety, or frustration.

Emotion, Stress, and Memory

Emotionally arousing and stressful experiences are often cited as the cause of many psychological and physical problems. Many of us have experienced

emotionally arousing and stressful experiences at one point or another of our life and have noted that these experiences can have important effects on our memory. Declarative memory, which is the type of memory we refer to during the course of this article, is defined as the conscious or voluntary recollection of previously learned information. We can have forgotten an important meeting or anniversary due to work overload, or else we can have a vivid recollection of a car accident or any other emotionally arousing experience. Because of their impact on our lives, we have a tendency to pay more attention to the negative effects of stress on our memory and to forget that, under certain conditions, emotionally arousing and stressful experiences can also have a positive impact on memory function. The goal of this article is to describe the positive and negative effects of emotionally arousing and stressful experiences on memory function and to delineate the psychological and biological determinants of these effects.

Emotion and Stress: Two Different Entities?

Emotion and stress share many characteristics. A stressful experience will often cause a particular emotion (e.g., surprise, fear, or joy), and particular emotions can create stressful situations (e.g., blushing due to extreme timidity can cause a stressful situation). Moreover, an emotion possesses many of the properties of a stressor. First, it often has an identifiable source. Second, it is usually brief and leads to an intense and conscious experience of short duration. Finally, an emotion creates bodily reactions (e.g., increase in heart rate or perspiration) that are similar to those induced by a stressor, and both states act by increasing arousal. Because of these similarities between emotion and stress, most of the literature on emotion, stress, and memory intermixes the effects of emotion and those of stress on memory function. However, emotion and stress are two different entities. Although a stressful experience will almost always trigger a specific emotion, a particular emotion does not always elicit a stress reaction. As far as laboratory setting is concerned, emotion and stress differ in the way they are induced and, thus, in the way they influence memory in humans. Hence, emotions are usually induced by the presentation of emotional words, films, or pictures, whereas stress is usually induced by putting the individual in a social situation known to create a stress (e.g., public

speaking). Because of this important difference between the experimental paradigms used to measure the effects of emotion and stress on memory, different questions have been asked. The induction of emotion has been used to measure memory for emotionally arousing events, whereas the induction of stress has been used to measure the specific effects of stress on subsequent memory function.

Memory for Emotionally Arousing Events

It is well known that what we encode and remember from an event depends primarily on the attention that we devoted to this event and its components. If we do not pay attention to what we are reading right now, there is less chance that we will remember it at a later time than if we gave all our attention to a lecture. This is because the more attention given to an event, the higher the probability that this event will be elaborated (relating the information from this event to other situations and related concepts in memory) at the time of encoding. Research on memory has shown that events that are poorly elaborated (shallow processing) at the time of encoding are remembered less well than events that are deeply elaborated (deep processing) at the time of encoding.

The level of attention devoted to an event at the time of encoding will greatly depend on the emotional salience of this event. Most of us remember what we were doing and with whom we were at the time we learned about the World Trade Center attacks, but the majority of us may have difficulty remembering what we were doing and with whom we were 13 days before or after the attacks. This flashbulb phenomenon may be explained by the fact that the emotion (e.g., surprise or anger) that was triggered by the announcement of the World Trade Center attacks directed the totality of our attention to the event, leading to a deeper elaboration and, thus, to an optimization of our memory for this event.

Memory for Central versus Peripheral Details

Studies of trauma victims have reported how vividly the traumatic event is recalled in the absence of any memory for information surrounding the traumatic event. A closely analogous situation appears in the field of law enforcement and describes the weapon focus phenomenon. Witnesses to violent crimes demonstrate a weapon focus effect in which the weapon captures most of the victim's attention, resulting in a reduced ability to recall other details of the scene and to recognize the assailant at a later time. This phenomenon has been explained by Easterbrook's cue utilization theory, which suggests that emotionally

arousing events narrow subjects' attention and lead them to attend only to the center of an event and to exclude more peripheral information. In general, laboratory experiments have confirmed this hypothesis. For example, as reported by Christianson, subjects viewed a series of slides in which the emotional valence of one critical slide in the series was varied. In the neutral version, the critical slide showed a woman riding a bicycle. In the emotional version, the same woman was seen lying injured on the street near the bicycle. In both versions, a peripheral car was seen in the distant background. The results showed that the central detail information (woman and bicycle) was retained better in the emotional condition but that the peripheral information (car in background) was retained better in the neutral event. In general, the pattern of results obtained from various studies shows that the central aspects of an emotional event (those events in the scope of attention) are well retained in memory, whereas memory for peripheral details (out of the scope of attention) are poorly remembered. Recently, Cahill and collaborators pushed this analysis further and reported gender-related influences in the recall of central versus peripheral information from an emotional story. Men and women with high male-related traits on the Bem Sex-Role Inventory show better memory for central aspects of an emotional story, as opposed to peripheral details. On the other hand, women and men with high female-related traits on the Bem Sex-Role Inventory have a better memory for peripheral details of an emotional story, as opposed to central information.

Immediate versus Delayed Memory for Emotional Events

Studies have examined the relation among emotion, retention interval, and memory. What these studies have shown so far is that emotionally arousing events delay forgetting. In a highly cited 1963 study, Kleinsmith and Kaplan presented subjects with pairs of words to learn. Some of the words were neutral, whereas others had a strong emotionally negative component (e.g., rape, mutilation). The memory of the word pairs was tested 2 min or 1 week after presentation. Memory performance tested shortly after learning (i.e., short-term declarative memory) generally refers to consolidation processes occurring within the first 3 h after learning (i.e., early phase), whereas memory performance tested after an extensive delay (i.e., long-term declarative memory) generally refers to the consolidation processes taking place 3 h post-learning (i.e., late phase) and that persist for at least a day and involve gene transcription and protein

synthesis. Kleinsmith and Kaplan's study showed that at short intervals memory was poorer for emotional words, but at long intervals memory was better for emotional words. These results have been replicated many times, and the majority of studies have shown enhanced memory for emotionally arousing material, provided that memory is tested at longer retention intervals.

Possible Mechanisms Underlying the Increased Memory for Emotionally Arousing Events

Psychological as well as biological mechanisms have been suggested to explain how emotionally arousing events might enhance memory. At the psychological level, one of the most frequently cited references on the relation between emotionally arousing events and memory is the Yerkes–Dodson law. In a series of studies published in 1908, Yerkes and Dodson reported that arousal improved performance in a linear manner for easy tasks, whereas for difficult tasks they found that an inverted-U-shaped curve related better performance with medium arousal and worse performance with low or high arousal. However, many field and laboratory studies suggest that individuals retain information from high arousing events quite well, a finding that goes against the Yerkes–Dodson law predictions. Thus, Easterbrook nuanced Yerkes–Dodson arousal hypothesis, suggesting the cue utilization theory. According to this theory, increased arousal due to emotionally arousing events acts by narrowing attention to the central details of a scene, leading to better recall of central details compared to peripheral details. Based on this model, Cahill recently suggested that men's ability to better remember central emotional information was attributable to their higher use of the right amygdala when processing emotional material (the right brain hemisphere being biased for a global/holistic analysis of the information). On the other hand, women's remembering more precisely peripheral emotional information was attributable to their higher use of the left amygdala during the processing of emotional material (the left brain hemisphere being biased for local, finer analysis of the information).

Finally, Walker used the arousal hypothesis in conjunction with Hebb's concept of reverberating circuits in order to explain the retention interval effect of emotional memory. He suggested that emotional information, because it creates a high level of arousal at the time of learning, produces a more active consolidation process and leads to a stronger memory trace and better recall at longer intervals. In contrast, because it creates a high level of arousal, emotional information also acts by inhibiting recall at

shorter interval due to the fact that the process of consolidation is active (and thus inaccessible) at this time.

Contrary to psychological explanations, biological explanations of memory for emotionally arousing events have presented extensive data indicating that emotionally arousing experiences produce a retrograde enhancement of memory for that experience via a variety of hormones released during these experiences. Hence, during emotionally arousing events, there is a secretion of the peripheral catecholamines epinephrine and norepinephrine by the adrenal medulla, a secretion of central norepinephrine by the locus ceruleus and a secretion of corticosteroids by the adrenal glands. The peripheral catecholamines enhance memory by reaching the vagus nerve, nucleus of the solitary tract, and locus ceruleus. Central norepinephrine is then secreted by the locus ceruleus, activating noradrenergic neurons throughout the brain. Significantly, central norepinephrine is also triggered as soon as an emotionally arousing event occurs, independently of peripheral catecholamines. Corticosteroids, being liposoluble, enhance memory by easily crossing the blood–brain barrier and reaching the brain. Adrenergic hormones and corticosteroids may facilitate memory consolidation for emotional information through their interactions with noradrenergic and corticosteroid receptors located in the amygdala, which, in turn, modulates hippocampal activity and enhances consolidation for emotionally arousing material.

Animal and human studies have confirmed the role of these hormones in the memory-modulating effects of emotionally arousing events. Thus, in rodents postlearning stimulation of the noradrenergic system enhances (but postlearning blockade inhibits) long-term declarative memory of an inhibitory avoidance task. Likewise, posttraining injections of moderate doses of synthetic corticosteroids enhance and pretraining corticosteroid synthesis inhibition impairs long-term expression of inhibitory avoidance in animals. In humans, prelearning blockade of central β -adrenergic receptors or prelearning corticosteroid synthesis inhibition impairs long-term declarative memory for emotionally arousing material, whereas prelearning or postlearning stimulation of the noradrenergic or corticosteroid systems enhances it. Although psychological and biological explanations of the arousal hypothesis were used to explain the positive effects of emotion on memory function, research performed to this day shows that arousal is significant as an intervening variable only when the source of the arousal (in this case, the emotionally arousing event) is directly related to the information to be remembered. As we will see next, when the source

of emotion is not directly related to the information to be remembered, other psychological and biological mechanisms come into play and have a stronger impact on memory function than arousal itself.

Memory after Stress

Influences of Acute Stress on Subsequent Memory Function and Possible Psychological Effects Underlying These Influences

Remembering with great accuracy a particular emotionally arousing event is very different from performing various tasks involving memory in a day-to-day situation when we are faced with stress. We may remember with great accuracy the events in our lives surrounding the World Trade Center attacks, but we may also have forgotten the anniversary of a loved one due to work stress. Originally, the effects of a stressor on subsequent declarative memory for material unrelated to the source of stress were studied using noise as a stressor. Indeed, human beings possess a limited amount of attention that can be allocated to a particular task at any one time. Information overload is said to occur when this capacity is exceeded, that is, when there is too much information to process at the same time. When this happens, the individual may have difficulty discriminating between and attending to the relevant and irrelevant information to encode new events in memory. Studies measuring the effects of noise on memory reported that noise impairs the recall of the information encoded at the time of noise stress. This result has been interpreted as showing that noise acts by overloading the limited processing capacity of the individual, leading to impairments in processing task-relevant stimuli because of an increased distraction induced by task-irrelevant stimuli (in this case, noise). Because the individual is unable to discriminate between relevant and irrelevant information when under noise stress, memory impairment appears and is represented by a poorer recall of task-relevant stimuli.

Effects of Acute Stress on Subsequent Memory Function and Possible Physiological Mechanisms Underlying These Effects

In order to better understand the mechanisms by which stress might modulate subsequent memory processing of material unrelated to the source of the stressor, researchers aimed at defining and understanding, more specifically, memory processing, stress-related physiological activation, and the links between these two. Research over the past 40 years revealed that memory is not a unitary and passive process. In the field of cognitive psychology, the individual is viewed

as being active, constructive, and organized, not as being a passive recipient of environmental stimulation or overstimulation. Thus, the individual is seen as actively constructing a view of reality, as selectively choosing some aspects of experience for further attention, and as attempting to commit some of that information to memory. Hence, to better understand the influence of stress on subsequent memory processing of information unrelated to the stressor, it is thus necessary to determine the series of stages that make up the given act of memorization and to measure the exact effects of stress on specific components of this information processing system. Declarative memory is temporally defined in three phases; information must first be attended to (arousal and attentional processes) and encoded (short-term/working memory; phase 1) before it is consolidated (phase 2) into long-term memory and recalled (phase 3).

Physiological mechanisms have been suggested to explain the effects of stress on memory phases. The brain is the organ that gives an interpretation to a stressful experience, and it is also responsible for mounting the physiological response to the stressor. Important physiological components of the stress response include the hypothalamic-pituitary-adrenal (HPA) axis (with corticosteroids being the end product), the sympathoadrenal system, and the central noradrenergic system. The HPA axis, along with the sympathoadrenal system, governs metabolic responses to the slings and arrows of everyday life. Both these systems support a range of metabolic responses that serve to ensure the availability of sufficient energy substrates in circulation during periods when there is inevitably increased cellular activity in vital organs and an accompanying increase in the requirement for fuel. The corticosteroid system and the central noradrenergic system have been studied in relation with the memory-modulating effects of stress because these two systems, when activated, can modulate activity in noradrenergic and corticosteroid receptors in the frontal lobes, hippocampus, and amygdala, three brain structures involved in declarative memory processes.

Rodent studies have shown that when a laboratory stressor (e.g., tail shocks, water immersion, or restraint stress) is administered at various times before or after learning, as well as before recall, stress-induced elevations in corticosteroid levels modulate declarative memory according to an inverted-U-shaped function. Optimal declarative memory for material unrelated to the stressor (e.g., inhibitory avoidance protocols or spatial water-maze tasks) occur at moderate levels of stress and stress-induced increases in circulating corticosteroid levels, whereas lower (i.e., boredom or drowsiness) or higher stress

levels and stress-induced increases in circulating corticosteroid levels are less effective or may even impair declarative memory performance on these tasks.

In humans, when a laboratory stressor (e.g., a public speaking task combined with a public arithmetic task) is administered before learning or retrieval, high corticosteroid levels following stress are associated with memory impairments for material unrelated to the stressor, such as neutral words lists. Recently, studies measuring the influence of stress on memory for emotional material unrelated to the stressor reported more heterogeneous findings. Thus, when a laboratory stressor was presented before learning or retrieval of emotional and neutral information unrelated to the stressor, high corticosteroid levels following stress were associated with memory impairments for emotional information (whether positive or negative), whereas they had no influence on memory for neutral material. However, two other studies showed that stress administered before or after learning enhanced memory for emotional material, whereas it had no impact or impaired subsequent memory for neutral information.

Altogether, these results show that stress-related elevations in corticosteroids can have different effects on subsequent memory for material unrelated to the stressor. The effects of emotionally arousing and/or stressful events on declarative memory vary according to the nature of the material to be remembered, with elevated levels of corticosteroids enhancing memory for the emotionally arousing event itself but leading, more often than not, to memory impairments for material unrelated to the source of the stressor.

The time of day (morning vs. afternoon) and levels of circulating corticosteroids at the time of testing could also be important factors influencing the effects of stress-related elevations in corticosteroids on subsequent memory for material unrelated to the stressor. Corticosteroid receptors differ in terms of their affinity for circulating levels of corticosteroids. Mineralocorticoid receptors (MRs) have a 6- to 10-time higher affinity for corticosteroids than glucocorticoid receptors (GRs). A wealth of evidence now demonstrates that the activation of MRs is mandatory for the successful acquisition of environmental cues necessary to encode information, whereas the activation of GRs is necessary for the long-term memory consolidation of this information.

Endogenous levels of corticosteroids and, thus, the activation of the MRs and GRs, significantly vary during the day, with higher endogenous levels of corticosteroids in the morning (AM) phase than in the evening (PM) phase. Consequently, the addition of a stressful and emotionally arousing event, which

by itself will trigger a significant increase in the endogenous levels of corticosteroids, should have a differential impact on the activation of MRs and GRs as a function of time of day. In the AM phase, most of the MRs and approximately half of the GRs are activated, whereas, in the PM phase, most of the MRs and about one-tenth of the GRs are activated. If we apply a stressor in the AM phase, the endogenous increase in corticosteroid levels will act by saturating GRs, but the same stressor applied in the PM phase will act by activating approximately half of the GRs. Because stress-induced elevations in corticosteroid levels have been shown to modulate declarative memory for material unrelated to the stressor according to an inverted-U-shaped function, the differential activation of MRs and GRs at different times of the day imply that a stressor applied in the morning should impair memory function (right-hand-side of the inverted U-shaped curve), whereas the same stressor applied in the PM phase should increase or have no impact on memory (left-hand-side or top of the inverted-U-shaped curve). Moreover, given that stress has differential effects as a function of the emotional valence of the material to be learned, it can be suggested that the application of a stressor in the AM versus PM phase should have a different effect on memory for emotionally arousing or neutral information unrelated to the source of the stressor.

Individual differences could also be an important factor explaining the discrepancy in the effects of stress-related elevations in corticosteroids on subsequent memory for material unrelated to the stressor. Indeed, high and low corticosteroid responders can show different memory performance following stress. Individual differences could be attributable to genetic factors, corticosteroid sensitivity differences and lifespan cortisol exposure differences, gender, or personality.

Finally, very few studies measured the influence of the central noradrenergic system, activated following stress, on subsequent memory for material unrelated to the stressor. Animal studies reported that stress-induced elevations in norepinephrine levels (following foot shocks) modulate declarative memory according to an inverted-U-shaped function, with optimal levels enhancing (and lower or higher levels of norepinephrine impairing) memory for information unrelated to the source of stress (e.g., passive aversive conditioning). In humans, one recent study showed that the blockade of peripheral and central noradrenergic β receptors before the administration of a stressor did not impair memory for material unrelated to the source of stress. These results suggest that the β -adrenergic system is not implicated in the effects of stress on subsequent declarative memory

function, contrasting with the well-established role of this system during the memorization of events that are emotionally arousing in nature. Further studies in humans are needed to determine the exact role played by the central noradrenergic system in the effects of stress on subsequent memory for information unrelated to the stressor.

Conclusion

In sum, important differences in mnemonic performance exist between memory for emotionally arousing events and memory for material unrelated to a stressor that is processed after stress. Interestingly, whether the subject is memorizing the emotionally arousing event itself or material unrelated to the stressor, both the central noradrenergic system and the corticosteroid system are at play, although they seem to act differently in the recall of neutral and emotional information. Experimental factors (e.g., nature of material to remember and time of day) and individual differences (e.g., personality and genetics) are also important components to consider because they can all influence the direction of the effects of emotion and stress on memory. Future studies considering and controlling for these factors and measuring the interactive and/or differential effects of the noradrenergic and corticosteroid systems should prove highly valuable to our understanding of the effects of emotionally arousing and stressful experiences on human declarative memory processes.

See Also the Following Articles

Cognition and Stress; Corticosteroids and Stress; Learning and Memory, Effects of Stress on; Self-Esteem, Stress and Emotion; Glucocorticoid Effects on Memory: the Positive and Negative.

Further Reading

- Abercrombie, H. C., Kalin, N. H., Thurow, M. E., et al. (2003). Cortisol variation in humans affects memory for emotionally-laden and neutral information. *Behavioral Neuroscience* 117, 505–516.
- Arnsten, A. F. T. (2000). Through the looking glass: differential noradrenergic modulation of prefrontal cortical function. *Neural Plasticity* 7, 133–144.
- Buchanan, T. W. and Lovullo, W. R. (2001). Enhanced memory for emotional material following stress-level cortisol treatment in humans. *Psychoneuroendocrinology* 26, 307–317.
- Cahill, L. (2003). Sex-related influences on the neurobiology of emotionally influenced memory. *Annals of the New York Academy of Science* 985, 163–173.

- Cahill, L. and Alkire, M. T. (2003). Epinephrine enhancement of human memory consolidation: interaction with arousal at encoding. *Neurobiology of Learning and Memory* 79, 194–198.
- Cahill, L., Gorski, L. and Le, K. (2003). Enhanced human memory consolidation with post-learning stress: interaction with the degree of arousal at encoding. *Learning and Memory* 10, 270–274.
- Cahill, L., Gorski, L., Belcher, A., et al. (2004). The influence of sex versus sex-related traits on long-term memory for gist and detail from an emotional story. *Consciousness and Cognition* 13, 391–400.
- Christianson, S. A. (1992). Emotional stress and eyewitness memory: a critical review. *Psychological Bulletin* 112, 284–309.
- Domes, G., Heinrichs, M., Reichwald, U., et al. (2002). Hypothalamic-pituitary-adrenal axis reactivity to psychological stress and memory in middle-aged women: high responders exhibit enhanced declarative memory performance. *Psychoneuroendocrinology* 27, 843–853.
- Domes, G., Heinrichs, M., Rimmele, U., et al. (2004). Acute stress impairs recognition for positive words – association with stress-induced cortisol secretion. *Stress* 7, 173–181.
- Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review* 66, 183–201.
- Elzinga, B. M., Bakker, A. and Bremner, J. D. (2005). Stress-induced cortisol elevations are associated with impaired delayed, but not immediate recall. *Psychiatry Research* 134, 211–223.
- Golberger, L. and Breznitz, S. (1993). *Handbook of Stress: Theoretical and Clinical Aspects* (2nd Edition). New York: Free Press.
- Jelicic, M., Geraerts, E., Merckelbach, H., et al. (2004). Acute stress enhances memory for emotional words, but impairs memory for neutral words. *International Journal of Neuroscience* 114, 1343–1351.
- Kandel, E. R. (2001). The molecular biology of memory storage: a dialogue between genes and synapses. *Science* 294, 1030–1038.
- Kleinsmith, L. J. and Kaplan, S. (1963). Paired associate learning as a function of arousal and interpolated interval. *Journal of Experimental Psychology* 65, 190–193.
- Kuhlmann, S., Piel, M. and Wolf, O. T. (2005). Impaired memory retrieval after psychological stress in healthy young men. *Journal of Neuroscience* 25, 2977–2982.
- Lupien, S. J., Fiocco, A., Wan, N., et al. (2005). Stress hormones and human memory function across the lifespan. *Psychoneuroendocrinology* 30, 225–242.
- Maheu, F. S., Joobar, R., Beaulieu, S., et al. (2004). Differential effects of adrenergic and corticosteroid hormonal systems on human short- and long-term declarative memory for emotionally arousing material. *Behavioral Neuroscience* 118, 420–428.
- Maheu, F. S., Joobar, R. and Lupien, S. J. (2005). Declarative memory after stress in humans: differential involvement of the β -adrenergic and corticosteroid systems. *Journal of Clinical Endocrinology and Metabolism* 90, 1697–1704.

- Maheu, F. S., Collicutt, P., Kornik, R., et al. (2005). The perfect time to be stressed: a differential modulation of human memory by stress applied in the morning or in the afternoon. *Progress Neuropsychopharmacology and Biological Psychiatry* 29(8), 1281–1288.
- McGaugh, J. L. (2000). Memory – a century of consolidation. *Science* 287, 248–251.
- Milner, B., Squire, L. R. and Kandel, E. R. (1998). Cognitive neuroscience and the study of memory. *Neuron* 20, 445–468.
- Reul, J. M. H. M. and de Kloet, E. R. (1985). Two receptor systems for corticosterone in rat brain: microdistribution and differential occupation. *Endocrinology* 117, 2505–2512.
- Roozendaal, B. (2002). Stress and memory: opposing effects of glucocorticoids on memory consolidation and memory retrieval. *Neurobiology of Learning and Memory* 78, 578–595.
- Sauro, M. D., Jorgensen, R. S. and Teal Pedlow, C. (2003). Stress, glucocorticoids, and memory: a meta-analytic review. *Stress* 6, 235–245.
- Takahashi, T., Ikeda, K., Ishikawa, M., et al. (2004). Social stress-induced cortisol elevation acutely impairs social memory in humans. *Neuroscience Letters* 363, 125–130.
- Walker, E. L. (1958). Action decrement and its relation to learning. *Psychological Review* 65, 129–142.
- Wolf, O. T., Schommer, N. C., Hellhammer, D. H., et al. (2002). Moderate psychosocial stress appears not to impair recall of words learned 4 weeks prior to stress exposure. *Stress* 5, 59–64.
- Wust, S., Van Rossum, E. F., Federenko, I. S., et al. (2004). Common polymorphisms in the glucocorticoid receptor gene are associated with adrenocortical responses to psychological stress. *Journal of Clinical Endocrinology and Metabolism* 89, 565–573.
- Yerkes, R. M. and Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-information. *Journal of Comparative Neurology and Psychology* 18, 459–482.

Memory Impairment

A J Parkin[†]

University of Sussex, Sussex, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 728–731, © 2000, Elsevier Inc.

Memory in Posttraumatic Stress Disorder
Other Stress-Related Memory Disorders
Neurophysiological and Neuroanatomical Aspects of the
Effect of Stress on Memory
Summary

Glossary

Declarative memory Any memory task that requires a subject to retrieve a specific past event; used interchangeably with explicit memory and episodic memory.

Episodic memory Any memory task that requires a subject to retrieve a specific past event; used interchangeably with explicit memory and declarative memory.

Explicit memory Any memory task that requires a person to retrieve a specific past event; used

Flashback

interchangeably with declarative memory and episodic memory.

Sudden reemergence of a traumatic memory that has previously been inaccessible; characteristic of posttraumatic stress disorder (PTSD).

Fugue

Stress-induced loss of memory in which the person loses personal identity and may even assume a new one.

Hippocampal formation

Name given to the hippocampus proper plus associated structures, including the dentate gyrus and entorhinal cortex.

Hippocampus

Region of the medial temporal lobes involved critically in the formation of new memories. Particularly vulnerable to the effects of cortisol but also to other neuromodulators that increase with the onset of stress.

Hysterical amnesia

Loss of memory for a specific period associated with trauma. Difficult to distinguish from PTSD.

Implicit memory

Any memory task that tests a person's memory for an event without direct reference to that event; used interchangeably with nondeclarative memory and sometimes with procedural memory.

Multiple personality

Memory disorder associated with history of abuse. If genuine, it is thought to be linked to stress, but many believe it has a strong iatrogenic component.

[†]Deceased.

<i>Nondeclarative memory</i>	Any memory task that tests a person's memory for an event without direct reference to that event; used interchangeably with implicit memory and sometimes with procedural memory.
<i>Priming</i>	A widely used measure of nondeclarative memory in which prior exposure of a stimulus can facilitate subsequent processing of that stimulus without the subject's awareness.
<i>Procedural memory</i>	Strictly speaking, this defines any form of memory that is not subject to conscious access but is sometimes used interchangeably with implicit memory and nondeclarative memory.
<i>Trier social stress test</i>	A widely used experimental manipulation for increasing stress in which subjects undertake a public-speaking exercise.
<i>Wechsler memory scale revised</i>	The most widely used instrument for assessing deficits in declarative memory. It tests both verbal and nonverbal memory.

Memory in Posttraumatic Stress Disorder

Posttraumatic stress disorder (PTSD) is now a recognized condition that presents in an individual following a traumatic event. It is observed most commonly in combat veterans but can also be found in people who have been involved in accidents, been the victim or witness of a violent event, or been subjected to sexual abuse as a child. The effects of PTSD on memory are somewhat complex. Typically the sufferer is amnesic for the traumatic experiences, but, from time to time, flashbacks occur in which the traumatic experience is recreated vividly in the individual's mind. It is not clear what determines the occurrence of a flashback, but one determining factor appears to be mood. Memory is known to be sensitive to state dependency in that a memory is more likely to be retrieved if the individual's mood state matches that experienced at the time of the traumatic experience. There is also some suggestion that the recall of traumatic memories has a dissociative quality in that the person attributes the traumatic recall to an aspect of their self that is in some way detached from their normal ongoing self.

There is now abundant evidence that PTSD victims also have an impaired ability to remember new information. For example, Vietnam veterans with PTSD performed more poorly than controls on components of the Wechsler Memory Scale, and this type of effect has been shown on a variety of other tests of memory. Adolescent victims of violence in Beirut with PTSD

have been shown to achieve lower academic performance than their counterparts without PTSD. Victims of PTSD also exhibit memory gaps ranging from minutes to hours. Examples include one person who remembered nothing between walking down a street in Boston and waking up in Texas and another who disappeared from a psychiatric unit and then awoke to find himself in battle dress in the woods. Also, consistent with the state dependency view of PTSD, victims of PTSD show a greater recall of trauma-related words but a poorer recall of neutral and positive affect words.

Other Stress-Related Memory Disorders

While PTSD is the most prominent psychiatric disorder of memory linked to stress, there are other conditions that have rather different qualities. The first of these is what has been termed hysterical amnesia. There are great similarities here with PTSD in that the patient has a temporary loss of memories for a particular adverse event. The only difference is that the disorder appears associated with less overall psychiatric morbidity than PTSD.

Far less common, although a great favorite with fiction writers, is the fugue state. In this dissociative disorder the patient loses his or her identity and, in many cases, assumes a new one. There are very few detailed studies of fugue because the condition can ameliorate quite quickly and a number of cases turn out to be malingering. However, in a well-documented case, a young man was deeply disturbed by the death of his grandfather and remained in a fugue state for some time, exhibiting only fragments of memory about his own past. However, while watching the funeral scene in the film *Shogun*, his memory of himself came back, again stressing the importance that state dependency may play in the mediation of stress-related memory dysfunction.

Stress can also be implicated in perhaps the strangest of memory disorders – multiple personality disorder (MPD). Here the person assumes a number of different personalities that have differing relationships with one another. Thus, personality A may know about personality B but not vice versa, and so on. There is massive controversy about MPD, and many clinicians believe that the condition is essentially iatrogenic – generated by clinicians in suggestible clients and also faked by knowledgeable individuals, a good example being the famous case of the hillside strangler. Taking a more moderate view, it is possible to argue that in genuine cases of MPD, various personalities absorb different aspects of a stressful past. In one case, for example, a young woman had a

history of sexual abuse and drug taking. Discussion of these different aspects of her past was only possible via different personalities that she presented.

Neurophysiological and Neuroanatomical Aspects of the Effect of Stress on Memory

There is now a great deal of knowledge about why stress can cause a loss of memory function. Stress is known to increase the levels of glucocorticoids, which in turn has adverse effects on memory. There is a disorder known as the Morbus-Cushing syndrome that results in abnormally high levels of cortisol production; patients with this disorder exhibit memory dysfunction as one of their primary symptoms. Cortisol is widely held to exert its adverse effects on a brain region known as the hippocampus. This structure, located bilaterally in the medial temporal lobes, is critically involved in the initial consolidation of memory and is also thought to mediate the storage of memories for some time after their initial registration. Animal studies have shown that increased cortisol levels result in a loss of hippocampal neurons and in a decrease in dendritic branching. This loss is thought to be attributable to an increased vulnerability of the hippocampus to endogenously generated amino acids. Further evidence for the effects of cortisol on the hippocampus comes from studies showing that a major physiological index of hippocampal function, long-term potentiation, is affected by the administration of cortisol.

Further evidence for cortisol involvement in stress-related memory dysfunction comes from work involving the endogenous indole isatin, which exists in particularly high levels in the hippocampus and cerebellum. At relatively low levels it has an anxiogenic effect and results in an increase in circulating levels of cortisol. It also negates the anxiolytic effects of atrial natriuretic peptide and its associated memory enhancement effects.

Other research also points toward a critical relationship between stress and hippocampal function. Monkeys exposed to stress as a result of overcrowding were found to have hippocampal damage, and there is similar evidence in humans. One study used magnetic resonance imaging to compare hippocampal volume in Vietnam veterans with PTSD and healthy subjects. The PTSD veterans were found to have an 8% reduction in the volume of the right hippocampus, with no changes in adjacent brain structures. Similarly, with positron emission tomography, Vietnam veterans with PTSD were found to have a reduced blood flow in the hippocampus, again indicating a degree of malfunction. The same research

group has found a selective 17% reduction in the left hippocampus of adult survivors of sexual abuse.

The just-mentioned studies point to permanent changes in brain structure caused by stress and are presumably mediated, at least in part, by prolonged exposure to elevated cortisol levels. However, there is also evidence that increased cortisol levels can also cause temporary memory impairment. In one study, human volunteers were administered the synthetic steroid dexamethasone. The volunteers were then given two types of memory test involving either declarative memory or nondeclarative memory. Declarative memory is what we typically think of as memory – the ability to remember lists of words, recognize faces, and so on. The effect of dexamethasone was to reduce the performance on this type of memory dramatically. Unlike declarative memory, where the individual has to reflect directly on a previous event, nondeclarative memory tests an individual's retention indirectly. A favorite method is known as priming. Here a subject is exposed to a list of words and is asked, for example, to rate each word for pleasantness. No mention is made of having to remember any of the words. Subsequently, the subject is given a test, such as "Tell me the first word that comes to mind beginning with the letters PL." If the word placid was in the previous word list, there is a strong probability that this word will be generated in preference to other possibilities. This indicates that the subject has retained information about the word list. Interestingly, these effects appear independent of subjects' declarative memory and are shown reliably by people who have a dense amnesia for declarative information. The administration of dexamethasone did not affect nondeclarative memory. This type of effect has been reported in other studies and indicates that the effects of glucocorticoids are limited to declarative memory and, by implication, that the hippocampus, with its vulnerability to enhanced cortisol levels, is not involved in nondeclarative memory performance.

Aging has been associated with an increased vulnerability to stress, which has led to the possibility that the effects of age on memory may, in part, be mediated by cortisol levels. One study employed the Trier social stress test. Essentially, this test involves subjects being asked to speak in public as a means of inducing variable amounts of stress depending on the individual. Fear of public speaking was found to enhance cortisol levels in varying degrees, which in turn predicted recall performance, with higher cortisol levels predicting poorer recall. A comparable non-stressful task had no effect on performance. Interestingly, the best predictor of the relation between cortisol levels and memory performance was a

measure obtained 60 min before the stressful event, thus indicating that the anticipation of stress is a crucial factor in determining memory performance.

While this account has emphasized the role of stress on memory via changes in cortisol levels, a variety of other neurotransmitters and neuropeptides are released during stress. These include adrenocorticotrophic hormone (ACTH), dopamine, acetylcholine, endogenous opiates, vasopressin, oxytocin, and γ -aminobutyric acid. Of these, vasopressin and oxytocin are of particular interest in relation to the occurrence of flashbacks in PTSD. One study used a structured imagery technique designed to solicit the retrieval of traumatic memories in Vietnam veterans. The administration of vasopressin enhanced the recall of trauma, whereas oxytocin inhibited trauma recall, relative to the placebo.

A final issue is whether the available research on stress and memory has any bearing on the false memory debate. This debate has caused enormous controversy – the central point being whether memories of abuse could remain buried for years and then suddenly reemerge. Available evidence certainly makes an *a priori* case that the stress caused by continual abuse would be expected, via the mediation of cortisol, to have a deleterious effect on an individual's recall for the period of abuse. What is less clear is whether the mechanism used to account for flashbacks in PTSD could be adapted to explain recovered memory. It is a very problematic issue, not in the least because there is abundant evidence that so much recovered memory is generated by inappropriate hypnosis-based memory work induced by very biased therapists.

Summary

Stress has a range of effects on memory, ranging from minor lapses of memory in normal individuals to the sometimes devastating effects of PTSD. The major change in brain anatomy associated with prolonged exposure to stress is shrinkage of the hippocampus.

This is widely attributed to increased cortisol levels experienced during stressful periods.

PTSD, the most notable stress-related memory disorder, is characterized by impaired declarative memory and normal nondeclarative memory. Other stress-related memory disorders are hysterical amnesia, fugue, and multiple personality disorder. Transient increases in cortisol levels induced either synthetically or via stress induction procedures result in impaired declarative memory but normal nondeclarative memory. The sudden and vivid availability of traumatic memories in PTSD may be due to the special influence of vasopressin, which is also released during periods of stress.

See Also the Following Articles

Amnesia; Cognition and Stress; Hippocampus, Overview; Multiple Personality Disorder; Posttraumatic Stress Disorder, Neurobiology of.

Further Reading

- Bremner, J. D. and Narayan, M. (1998). The effects of stress on memory and the hippocampus throughout the life cycle: implications for childhood development and aging. *Developmental Psychopathology* 10, 871–885.
- Bremner, J. D., et al. (1995). Functional neuroanatomical correlates of the effects of stress on memory. *Journal of Traumatic Stress* 8, 527–553.
- Lupien, S. J., et al. (1997). Stress-induced declarative memory impairment in healthy elderly subjects: relationship to cortisol reactivity. *Journal of Clinical Neuroendocrinology and Metabolism* 82, 2070–2075.
- McGaugh, J. L. (1989). Involvement of hormonal and neuromodulatory systems in the regulation of memory storage. *Annual Review of Neuroscience* 12, 255–287.
- Parkin, A. J. (1997). *Memory and amnesia* (2nd edn.). Hove, UK: Psychology Press.
- van der Kolk, B. A., et al. (eds.) (1996). *Traumatic stress: the effects of overwhelming experience on mind, body and society*. New York: Guildford Press.

Menopause and Stress

N E Avis

Wake Forest University School of Medicine,
Winston-Salem, NC, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
N E Avis, volume 2, pp 732–735, © 2000, Elsevier Inc.

Menopause as a Stressful Life Event

Impact of Stress on Menopause Symptomatology

Stress and Ovarian Function

Conclusions

Glossary

<i>Menopause</i>	The permanent cessation of menses. The standard epidemiological definition of natural menopause is 12 consecutive months of amenorrhea, in the absence of surgery or other pathological or physiological cause (e.g., pregnancy, lactation, excessive exercise).
<i>Perimenopause</i>	That period of time immediately prior to menopause when the endocrinological, biological, and clinical features of approaching menopause begin through the first year after the final menstrual period. It is characterized by increased variability in menstrual cycles, skipped menstrual cycles, and hormonal changes. Early perimenopause is defined as menses in the previous 3 months, but changes in regularity. Late perimenopause is defined as no menses in the previous 3 months, but menses in the preceding 11 months.
<i>Premenstrual syndrome (PMS)</i>	A disorder characterized by a set of hormonal changes that trigger disruptive physical and emotional symptoms in a significant number of women for up to 2 weeks prior to menstruation.
<i>Surgical menopause</i>	Menopause induced by a surgical procedure that stops menstruation. Women who have both ovaries or the uterus with or without removal of the ovaries are generally included in this category.

Menopause as a Stressful Life Event

Whether or not menopause is viewed as a stressful life event can be looked at in two ways: (1) women's attitudes toward menopause, and (2) women's affective responses to menopause.

Attitudes toward Menopause

Cross-cultural and anthropological studies provide evidence that the meaning of menopause varies greatly across cultures. How a society views menopause is influenced by how it views aging and women in general. Menopause is often viewed as a positive event in women's lives in non-Western cultures, where menopause removes constraints and prohibitions imposed upon menstruating women. In countries where women have low status or are not allowed to show sexuality (such as in India), menopause is seen positively, as it provides freedom to go out in public and do things usually forbidden to women. Among South Asian women, the end of childbearing and the menstrual cycle is welcomed. While social status is tied to motherhood, it is motherhood that is valued and not biological fertility itself.

Among both Mayan and Greek women, menopause is seen as a positive event, although for different reasons. Mayan women marry young, do not practice birth control, and spend most of their reproductive years either pregnant or lactating. Pregnancy is viewed as dangerous and stressful, and menopause frees women from restrictions and pregnancy. While Greek women attempt to curtail family size and often use abortion as a means of birth control, menopause also frees Greek women from taboos and restrictions. A postmenopausal Greek woman is allowed to participate fully in church activities, as she is no longer viewed as a sexual threat to the community. Both Mayan and Greek women report better sexual relationships with their husbands following menopause, as the fear of pregnancy is eliminated. In other cultures, women give menopause little thought. In Papago culture, menopause may be completely ignored, to the extent that the language contains no word for menopause. In Japan there is no word to describe hot flashes. As Lock points out, the lack of a Japanese word to describe hot flashes is remarkable in a language that is infinitely more sensitive than English in its ability to describe bodily states (e.g., there are more than 20 words to describe the state of the stomach).

In Western societies, women are valued for sexual attractiveness and do not face restrictions found in other cultures. Aging, especially among women, is not revered, but rather is viewed quite negatively. In these societies menopause takes on a very different meaning. Despite these negative societal views, women themselves do not hold such negative attitudes.

A review of research on women's attitudes toward menopause conducted across a wide range

of populations and cultures shows that women consistently feel relief about the cessation of menses and do not agree that they become less sexually attractive following menopause. Women consistently report that they are glad to no longer deal with menstruation, accompanying PMS or menstrual cramps, fear of pregnancy, and purchase of feminine products. Thus, the end of menstruation, rather than bringing on a sense of psychological loss, is often met with relief. While women may feel a decrease in sexual desire or frequency, which they attribute to aging as much as to menopause, women overwhelmingly disagree with the statement that postmenopausal women are less feminine or attractive. Studies consistently show that postmenopausal women generally have a more positive view of menopause than premenopausal women. Thus, women's own attitudes toward menopause are not as negative as those of the medical profession or Western societies as a whole.

Women's Response to Menopause

Another way to examine whether menopause can be viewed as a stressful life event is to look at the impact of menopause on mental health. Earlier reports based on patient or clinic populations often perpetuated the perception that women become depressed and irritable and suffer a host of other symptoms during menopause. These studies, however, suffered from numerous methodological problems. First, such patient-based samples are highly biased. Fewer than half of menopausal women seek treatment, and those who do tend to report more stress in general and suffer more from clinical depression, anxiety, and psychological symptoms than nonpatient samples. Second, these studies often did not differentiate between women who experienced surgical and those who experienced natural menopause. Women who undergo surgical menopause have a very different experience of menopause, in that they experience more sudden hormonal changes, as well as a surgical procedure. A number of studies have found that women who have had a surgical menopause report more distress than women who experience a natural menopause. Finally, these studies often used menopausal checklists that asked women to check off symptoms that they experienced due to menopause. When asked in this way, women will check off any symptom they have that fits their beliefs about what women experience during menopause. This method thus perpetuates existing stereotypes of the menopause experience.

Cross-sectional epidemiological studies of community-based samples of women do not show consistent evidence of a relation between menopause and depression or other negative moods in the general

population. More recently, several longitudinal studies have shown an increased risk of depressive symptoms during perimenopause. However, it is important to note that only a minority of women appear to experience an increase in negative mood or depression during perimenopause, and other factors, such as stress and socioeconomic factors, are more related to depressed mood than hormones or menopausal status. Longitudinal data suggest that prior depression is the primary factor related to depression during menopause. Studies have also found that women who report greater mood disturbances at menopause also report menstrual cycle or reproductive-related problems. These include previous or current premenstrual symptoms or complaints, dysmenorrhea, and postpartum depression. Researchers have also found that social circumstances and stress account for much of the mood effects during the menopause transition. Studies that have included measures of stress or life changes have typically found that social factors are highly related to mood, and often more so than menopause status.

Thus, for the majority of women, menopause does not appear to be a stressful life event. However, like other life events or changes, this does not mean that menopause is not stressful for some women. It has been hypothesized that women who find menopause stressful are those who see it as the end of reproduction or a sign of old age. Women with less education or knowledge about menopause generally have more negative attitudes, as do women who have more symptoms in general or worse mental health.

Impact of Stress on Menopause Symptomatology

Vasomotor symptoms (hot flashes/flushes and night sweats) are the primary symptoms associated with menopause. Estimates of the incidence of hot flashes from population studies in the United States and worldwide have ranged from 24 to 93%. Studies of menopause in different cultures reveal wide cultural variation in symptom reporting. For example, hot flashes are uncommon in Mayan women, and Japanese and Indonesian women report far fewer hot flashes than women in Western societies. Because women across cultures differ in terms of their diet, physical activity, number of pregnancies, use of contraception, as well as attitudes toward menopause, it has been difficult to assess the reason(s) for this variation. However, even within a culture, a high degree of variability of symptom reporting is found among women, suggesting considerable individual variation in symptom experience. What differentiates symptomatic from asymptomatic women is not well understood.

Women experiencing other sorts of stress are likely to notice or amplify menopausal symptoms, as they would any symptom. Epidemiological studies have shown greater vasomotor symptoms associated with less education, interpersonal stress, more general symptom reporting, and negative attitudes toward menopause. A relationship between stress and hot flashes has been shown in several correlational and treatment studies. One study has shown that laboratory-induced stressors correlate with women's objectively measured hot flashes. However, hot flashes did not concentrate around the actual stressor, suggesting that stress may potentiate (rather than cause) hot flashes by decreasing the threshold for the triggering of hot flashes at the hypothalamic level. Findings that women who report more symptoms premenopause also report more menopausal symptoms are consistent with the notion that some people have greater sensitivity to symptoms.

Thus, consistent with other research on stress and symptoms, there is some evidence that women experiencing stress for other reasons may have a tendency to notice or report greater menopausal symptoms. However, stress is only one of many factors that appear to impact menopausal symptoms.

Stress and Ovarian Function

Impact of Stress on Ovarian Function

Appropriate secretion of hypothalamic gonadotropin-releasing hormone (GnRH) is necessary for ovarian cyclicity. Decreased GnRH is a common cause of anovulation. Research in both humans and monkeys suggests that stress desynchronizes the GnRH neuronal network. High levels of stress have been shown to lead to altered menstrual function in premenopausal women. The impact of stress on perimenopausal women has not been well studied, but there is some evidence that marked increases in stress may alter menstrual cycles in the short term. It has also been hypothesized that high levels of stress may accelerate menopause. Recent research, however, has shown conflicting findings, with one study finding some evidence that stress accelerated menopausal age among women who were already experiencing irregular cycles or who were African-American, and another study finding that women who reported more psychological symptoms had a longer perimenopause.

Ovarian Function and Stress Responsivity

It has also been hypothesized that reproductive hormones influence the response to stress. Several researchers have examined the impact of ovarian

function on reactivity to stress. These studies have largely examined cardiovascular reactivity in response to laboratory stresses among women in differing stages of ovarian functioning. While studies have shown greater responses among postmenopausal or surgically menopausal women as compared to premenopausal women, more recent research among women with experimentally suppressed ovarian function has not shown greater responses to stress. This remains an area of investigation.

Conclusions

Although menopause itself is a physiological event, a woman's response to menopause depends on cultural, behavioral, psychosocial, as well as physiological factors. How an individual woman responds to menopause is a complex interaction of physiology, her current life circumstances, attitudes/concerns about fertility, the culture in which she lives, and her history of responding to physiological changes/symptoms. On the whole, menopause is not a stressful event for most women. Concurrent stresses from other sources, however, may exacerbate a woman's response to menopausal symptoms.

See Also the Following Articles

Menstrual Cycles and Stress; Premenstrual Dysphoric Disorder; Stress Induced Anovulation.

Further Reading

- Avis, N. E. (1996). Women's perceptions of the menopause. *European Menopause Journal* 3(2), 80–84.
- Avis, N. E. (2000). Is menopause associated with mood disturbances? In: Lobo, R. A., Kelsey, J. & Marcus, R. (eds.) *Menopause: biology and pathobiology*, pp. 339–352. New York: Academic Press.
- Avis, N. E., Crawford, S. L. and McKinlay, S. M. (1997). Psychosocial, behavioral, and health factors related to menopause symptomatology. *Women's Health: Research on Gender, Behavior, and Policy* 3(2), 103–120.
- Barsom, S. H., Mansfield, P. K., Kock, P. B., Gierach, G. and West, S. G. (2004). Association between psychological stress and menstrual cycle characteristics in perimenopausal women. *Women's Health Issues* 14, 235–241.
- Berga, S. L. (1996). Stress and ovarian function. *The American Journal of Sports Medicine* 24(6), S36–S37.
- Berga, S. L. and Louks, T. L. (2005). The diagnosis and treatment of stress-induced anovulation. *Minerva Ginecologica* 57, 45–54.
- Beyene, Y. (1986). Cultural significance and physiological manifestations of menopause, a biocultural analysis. *Culture, Medicine & Psychiatry* 10, 47–71.

- Bromberger, J. T., Matthews, K. A., Kuller, L. H., et al. (1997). Prospective study of the determinants of age at menopause. *American Journal of Epidemiology* 145, 124–133.
- Freeman, E. W., Sammel, M. D., Liu, L., et al. (2004). Hormones and menopausal status as predictors of depression in women in transition to menopause. *Archives of General Psychiatry* 61(1), 62–70.
- Freeman, E. W., Sammel, M. D., Lin, H., et al. (2005). The role of anxiety and hormonal changes in menopausal hot flashes. *Menopause* 12(3), 258–266.
- Freeman, E. W., Sammel, M. D., Lin, H. and Nelson, D. B. (2006). Associations of hormones and menopausal status with depressed mood in women with no history of depression. *Archives of General Psychiatry* 63(4), 375–382.
- Gold, E. B., Colvin, A., Avis, N. E., et al. (2006). Longitudinal analysis of vasomotor symptoms and race/ethnicity across the menopausal transition: Study of Women's Health Across the Nation (SWAN). *American Journal of Public Health* 96(7), 1226–1235.
- Kronenberg, F. (1990). Hot flashes: epidemiology and physiology. *Annals of the New York Academy of Sciences* 592, 52–86.
- Lock, M. (1986). Ambiguities of aging: Japanese experience and perceptions of menopause. *Culture, Medicine & Psychiatry* 10, 23–46.
- Nicol-Smith, L. (1996). Causality, menopause, and depression: a critical review of the literature. *British Journal of Medicine* 313, 1229–1232.
- Pearlstein, M. D., Rosen, K. and Stone, A. B. (1997). Mood disorders and menopause. *Endocrinology and Metabolism Clinics of North America* 26(2), 279–294.
- Shively, C. A., Watson, S. L., Williams, J. K., et al. (1998). Stress-menstrual cycle and cardiovascular disease. In: Orth-Gomer, K., Chesney, M. A. & Wenger, N. K. (eds.) *Women, stress, and heart disease*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Sommer, B., Avis, N., Meyer, P., et al. (1999). Attitudes toward menopause and aging across ethnic/racial groups. *Psychosomatic Medicine* 61, 868–875.
- Swartzman, L. C., Edelberg, R. and Kemmann, E. (1990). The menopausal hot flush: symptom reports and concomitant physiological changes. *Journal of Behavioral Medicine* 13(1), 15–31.
- World Health Organization Scientific Group. (1996). Research on the menopause in the 1990's. *WHO Technical Services Report Series* 886.

Menstrual Cycles and Stress

R Suri

UCLA Neuropsychiatric Institute and Hospital, Los Angeles, CA, USA

L Altshuler

UCLA Neuropsychiatric Institute and Hospital and West LA Veterans Administration Medical Center, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R Suri and L Altshuler, volume 2, pp 736–741, © 2000, Elsevier Inc.

The Menstrual Cycle

Biological Effects of Stress on Menstruation

Anorexia Nervosa and Amenorrhea

Exercise, Stress, and the Menstrual Cycle

Effects of Menstrual Cycle on Stress, Mood, and Psychiatric Symptoms

Consequences of Amenorrhea

Treatment Issues

Glossary

<i>Amenorrhea</i>	The absence of menses for 3 months or longer.
<i>Anorexia nervosa</i>	An eating disorder characterized by the refusal to maintain a minimally normal body weight; symptoms also include the absence of at least three consecutive menstrual cycles.
<i>Corticotropin releasing hormone (CRH)</i>	A 41-amino-acid peptide secreted by the paraventricular nucleus of the hypothalamus, which stimulates pituitary secretion of adrenocorticotrophic hormone and, consequently, cortisol secretion by the adrenal cortex.
<i>Functional hypothalamic amenorrhea</i>	Amenorrhea associated with factors such as psychogenic stressors, excessive exercise, or weight loss.
<i>Gonadotropin-releasing hormone (GnRH)</i>	A decapeptide hormone that plays a significant role in reproductive physiology through stimulation of follicular stimulating hormone and luteinizing hormone by the pituitary. GnRH is secreted by neurons of the preoptic and arcuate

nuclei of the hypothalamus in a pulsatile manner, with the amplitude and frequency of its secretions under the regulation of estrogen, progesterone, catecholamines, and neuropeptides.

Secondary amenorrhea

Amenorrhea that occurs when the patient has menstruated in the past but has stopped for at least 3–6 months. Causes can include pregnancy, lesions of the hypothalamus or pituitary (from surgery, infection, irradiation, chemotherapy, infarction, or tumor), emotional stress, endocrine disease (e.g., thyroid, adrenal, or diabetes), or medications.

Stress can be conceptualized as a state of disharmony or threatened homeostasis, with the patient's adaptive response being specific to the stressor or being more generalized. Although the demands and complexity of society continue to increase, the physiological mechanisms for coping with adversity have not changed appreciably over the past thousands of years, and the physiological responses to social pressures, rapid changes, and increased information are often similar to responses to physical danger and threats to survival. The effects of stress can be profound, resulting in the mobilization of adaptive behaviors and peripheral functions and the inhibition of biologically costly behaviors such as reproduction. For many animals, environmental stress, or conditions of high population density, are associated with chronic anovulation and infertility. During unfavorable circumstances, the lack of reproduction allows females to focus scarce resources on survival, improvement in overall condition, and investment in existing offspring. Although animal studies have provided information on the hormonal pathways through which stressors impact reproductive function, these mechanisms are poorly understood in humans. Nevertheless, for women, psychosocial stressors such as bereavement or separation from family, in the absence of a psychiatric diagnosis, have been linked to abnormalities of the menstrual cycle. Rates of menstrual irregularities increase in proportion to chronic stress, which, if severe enough, can completely inhibit the female reproductive system. Although the nature of the actual stressor can be significant, the response of the individual to the stressor, as well as certain cognitive, behavioral, or personality traits, may cause some women to be more vulnerable to the stress of daily living than other women and may play a more significant role in affecting menstruation than the actual stressor itself. The impact of stress can range from a lack of menstrual cycles to persisting cycles of marginal quality.

The Menstrual Cycle

The menstrual cycle is defined by the regular, repetitive monthly occurrence of ovulation throughout a woman's reproductive life. Menarche, defined as the onset of the first menses, occurs in the United States between the ages of 9 and 12, with a mean of 12.6 years. The mean duration of the cycle is 28 days, and the cycle is divided into two functional phases: the follicular (or proliferative) phase and the luteal (or secretory) phase. Polymenorrhea refers to menstrual cycles that occur at intervals of less than 21 days, whereas oligomenorrhea describes menstrual cycles occurring at intervals greater than 35 days. Menstrual cycles tend to be most irregular in the first 2 years after menarche and in the 3–5 years before menopause (perimenopause).

The follicular phase of the menstrual cycles starts at the first day of menses, continues until ovulation, and is characterized by a variable length, low basal body temperature, development of ovarian follicles, vascular growth of the endometrium, and the secretion of estrogen from the ovary. The luteal phase begins with ovulation, continues until the beginning of menses, and is characterized by a relatively constant duration of 12–16 days, an elevated basal body temperature, the formation of the corpus luteum in the ovary with the secretion of progesterone and estrogen, and preparation of the endometrium for implantation of an embryo.

Regular menstrual cycles occur as a result of complex, coordinated interactions among the central nervous system, the hypothalamus, the pituitary, and the ovary. Specifically, menstrual cyclicity occurs as a result of the pulsatile secretion of gonadotropin-releasing hormone (GnRH) by neurons in the preoptic and arcuate nuclei of the hypothalamus. GnRH then acts on the pituitary to stimulate release of the gonadotropins luteinizing hormone (LH) and follicular stimulating hormone (FSH), which influence the selection and maturation of the follicle for ovulation as well as secretion of ovarian steroid hormones that act on the genital tract. GnRH is secreted in a pulsatile manner at approximately 1 pulse per h in the follicular phase and 1 pulse per 2–3 h in the luteal phase, and the amplitude and frequency of secretions are influenced by feedback from estrogen and progesterone as well as catecholamines, dopamine, and norepinephrine within the brain. Estrogen levels rise in the follicular phase, reaching a peak approximately 24–36 h before ovulation. LH increases steadily until the midcycle when there is a surge, and ovulation occurs approximately 10–12 h after the LH peak. In the absence of pregnancy, decreasing steroid levels result in the constriction of the arteries

supplying the upper two-thirds of the endometrium, with subsequent shedding of the degraded endometrial tissue.

Biological Effects of Stress on Menstruation

Amenorrhea is defined by the absence of menses for 3 or more months. The failure of menarche to occur is referred to as primary amenorrhea; secondary amenorrhea refers to the discontinuation of menses after the onset of menarche. Prevalence rates for secondary amenorrhea range from 8.5% among women ages 13–18 years, 7.6% among women ages 15–24 years, and 3% among women ages 25–34 years.

Identifiable or organic causes of secondary amenorrhea and anovulation include a pituitary adenoma or hypothalamic tumor; however, the most common cause of amenorrhea is not the result of an identifiable organic source but is, rather, associated with psychological and lifestyle variables, including excessive exercise, low weight, affective and eating disorders, substance use, and stress. This condition is known as functional hypothalamic amenorrhea (FHA). The function of the hypothalamus is to maintain homeostasis and promote adaptation by regulating the neuroendocrine axis. It receives afferent neural information from other brain areas, as well as signals from the periphery that cross the blood-brain barrier, and then releases hormones that influence the synthesis and release of specific pituitary hormones in an integrated fashion, coordinating endocrine responses with internal and external demands. Changes in activity at one hypothalamic area can result in profound changes in the function of other areas. Psychophysiological factors, for instance, affect central neuroregulatory networks that subsequently disrupt the GnRH pulse generator and impact reproductivity. Specifically, stress can result in a decrease in GnRH drive, with the frequency and amplitude of the pulsatile release of GnRH falling out of the normal range. In general, catecholamines, epinephrine, and norepinephrine increase GnRH release, whereas β -endorphins decrease GnRH release. Although plasma concentrations of GnRH are too low to measure, LH levels correlate well with those of GnRH, and its secretory patterns can be used for information on hypothalamic-pituitary-gonadal (HPG) activity. LH pulses normally have a constant amplitude with a frequency of 2–6 h in the follicular phase and a more variable amplitude with a frequency of 2–6 h in the luteal phase. Studies have shown that, in women with the presence of both a stressful life event and a psychiatric diagnosis, spontaneous LH pulse

frequency is inhibited. A study of 67 patients with hypothalamic secondary amenorrhea found that the onset of the menstrual disorder was correlated with a psychosocial stressor in close to half of the patients. When both a psychiatric diagnosis and a stressful life event are present, LH pulse frequency can be reduced by as much as 50% of that for normally menstruating women. Studies have also demonstrated that the presence of an anxiety or depressive disorder reduces pulse amplitude significantly. However, for women with amenorrhea without a psychiatric disorder or a stressful life event, the pulsatile release of LH is similar to that in normally menstruating women in the follicular phase. Thus psychogenic factors appear to influence LH release, reflecting effects on HPG activity and GnRH secretion. With significant declines in pulsatile GnRH secretion, the pituitary secretion of LH and FSH is reduced, compromising folliculogenesis and ovulation.

A possible mechanism that may link the GnRH pulse generator and stress involves the hypothalamic-pituitary-adrenal (HPA) axis, which is activated by stress. Women with secondary amenorrhea often display hypercortisolism, and in monkeys and rats, the administration of corticotropin releasing hormone (CRH) results in an acute decrease in pulsatile GnRH and gonadotropin release. CRH also inhibits LH by central effects, partially mediated by central endogenous opiates. Stress stimulates the expression of CRH and vasopressin. In monkeys, vasopressin has been shown to decrease LH secretion acutely when administered intracerebroventricularly.

Further evidence suggesting involvement of the HPA axis has been provided by Berga, who demonstrated that cortisol concentrations, which reflect HPA activity, are elevated in women with FHA. In women with other causes of anovulation, cortisol concentrations were found to be comparable to those observed in women who were eumenorrheic with biochemical evidence of ovulation. Women with FHA who resumed ovulation had cortisol concentrations similar to eumenorrheic women and lower than women with persistent FHA.

Anorexia Nervosa and Amenorrhea

Much of the literature on stress and menstruation focuses on women with anorexia nervosa. Anorexia nervosa is characterized by a refusal to maintain body weight at or above the minimally normal weight for age and height, an intense fear of gaining weight even though underweight, disturbance in the way in which one's body weight or shape is experienced, and the absence of at least three consecutive menstrual cycles.

The disorder has a prevalence rate of 0.5–1.0% among females of late adolescence and early adulthood, although individuals with eating disorders that do not meet the full criteria for anorexia are more common. Anorexia accounts for 15–34.5% of patients with amenorrhea.

Clinical traits of anorexia often include an obsessive preoccupation with dieting, a desire to regress to prepubertal body habits to disguise femininity, and intense, obsessive-compulsive personality traits that result in hyperachievement. Mean age of onset for the disorder is 17 years, with possible bimodal peaks at ages 14 and 18. The onset of illness is often associated with a stressful life event, such as leaving home for college or the breakup of a relationship. The onset of dieting may be related to the stressor or may begin more insidiously, and amenorrhea can precede or be associated with weight loss. Unless the psychological stress is also addressed, menstrual cycles can remain abnormal, despite adequate weight gain. Follow-up studies at 2 years have found that psychological impairment persists in one-third to one-half of patients, with one-third of patients experiencing recurrent affective disorders and one-fourth of patients not regaining menses or attaining 75% of ideal body weight.

Females with anorexia nervosa demonstrate a pattern of 24-h gonadotropin secretion that is similar to prepubertal girls, with reversion to adult patterns following adequate weight gain and a change in behavior. Compared with normally menstruating women, the LH response to GnRH is diminished in the follicular phase. For women with secondary amenorrhea or infertility and weight loss, but without anorexia, similar but less severe changes in LH are seen.

Exercise, Stress, and the Menstrual Cycle

Strenuous, regular physical activity has also been associated with menstrual irregularities, oligomenorrhea, amenorrhea, or delayed menarche. Up to 50% of competitive runners, 25% of recreational runners, and 12% of cyclists and swimmers demonstrate amenorrhea. Both weight loss and stress have been implicated in the cause of exercise-associated amenorrhea through the disruption of GnRH secretion. Stress can be metabolic, psychological, or a combination of both, and its effects may be partly mediated by the increased β -endorphin levels that impact the amplitude and frequency of LH pulses. In addition, the release of CRH in endurance exercise can inhibit gonadotropin secretion and activate the release of corticosteroids and androgenic steroids.

Although psychological stress may result in erratic menstrual cycles, its role in exercise-related amenorrhea appears to involve initiation more than continuation.

Effects of Menstrual Cycle on Stress, Mood, and Psychiatric Symptoms

In addition to the influence of stress on the menstrual cycle, monthly fluctuations of reproductive hormones can impact a woman's vulnerability to the experience of stress. Some studies, for example, reported a four-fold increase in the prevalence of suicide attempts and allergic bronchial asthma attacks in the late luteal phase of the menstrual cycle.

The levels of estrogen, progesterone, and their metabolites fall during the late luteal phase of the menstrual cycle and remain low during the early follicular phase. These gonadal steroids modulate the central neurotransmitter function of serotonin, dopamine, norepinephrine, and γ -aminobutyric acid and thus, as their levels vary across the menstrual cycle, may influence symptoms of irritability, anxiety, depression, and fatigue. For example, the progesterone metabolites allopregnanolone and pregnanolone may have anxiolytic properties through their barbiturate-like effect on γ -aminobutyric acid receptors, with decreased levels of these metabolites correlating with increased tension and anxiety.

Premenstrual asthma, or the worsening of asthma symptoms and pulmonary function in relation to menstruation, has been described in approximately one-third of women. A number of studies reported increases in asthma prior to menstruation and, to a lesser extent, during menstruation. Mirdal suggested a relationship among lowered resistance to stress, lowered resistance to infections, and increased bronchial hyperreactivity as background etiological factors for the occurrence of premenstrual asthma.

The premenstrual and menstrual phases of the menstrual cycle have been associated with an increased vulnerability to an exacerbation of psychiatric symptoms, as well as a peak in psychiatric admissions. Literature in this area is limited to case reports and small studies, which have found evidence of premenstrual exacerbation in schizophrenia, depression, bulimia nervosa, anxiety disorders, and substance abuse. For example, in a prospective study of 27 women with research diagnostic criteria for major, minor, or intermittent depression responding to treatment with imipramine or phenelzine, one-fourth of the women experienced premenstrual recurrence of symptoms of low mood, anhedonia, anxiety, increased appetite, and hypersomnia, which remitted

following menses. For some women, an increased consumption of alcohol and marijuana occurs premenstrually. A study of 14 women with alcohol dependence found an association between alcohol intake and severity of premenstrual symptoms. Similar findings were reported in a prospective study of 21 women with a history of regular marijuana use. Smoking frequency has also been shown to vary across the menstrual cycle, with approximately 70% of women reporting an increase in the 7–10 days before their menstrual period. Researchers postulate that reproductive hormones may influence psychiatric symptoms by direct effects on central neurotransmitter function, as well as by causing the serum concentrations of medication to vary across the menstrual cycle.

Consequences of Amenorrhea

The hypoestrogenism associated with secondary amenorrhea can have long-term effects, the most significant of which is osteoporosis. Osteoporosis results from an imbalance between bone formation and bone resorption. Rapid increases in bone density are seen during puberty, with peak bone density of the spine and hip occurring by age 20. Females ages 14–18 with amenorrhea, especially if associated with weight changes, anorexia, or excessive exercise, risk decreased bone mineral density compared with normally menstruating young women. Amenorrheic athletes, for instance, have a significantly reduced spinal bone mineral content compared with controls with normal menstrual cycles. The peak bone density of premenopausal women influences the onset of osteoporosis later in life; thus, as life expectancy increases, the potential morbidity from osteoporosis also increases.

Treatment Issues

Stress can have a profound impact on menstruation and reproduction, ranging from cycles of marginal quality to chronic anovulation. Although mediated through biological actions on the hypothalamus, GnRH, and the HPA axis, these effects are potentially reversible. In managing women with functional hypothalamic amenorrhea, the psychological, behavioral, and psychiatric variables that influence hypothalamic function must be considered. Although pharmacological interventions can supply appropriate sex steroid exposure, restore hormonal balance, or induce ovulation, hormonal treatments can prevent the recognition of important emotional, behavioral, or psychiatric symptoms. Many women with FHA may have a potentially reversible condition

that can be addressed with education, changes in lifestyle, psychotherapy, assistance in identifying and coping with stressors, or attention to a psychiatric condition. Cognitive-behavioral therapy, for instance, has been shown to result in the recovery of ovarian function for women with FHA who do not have a psychiatric disorder or engage in excessive exercise. Such interventions can help ameliorate the long-term consequences of stress and anovulation, as well as the health hazards associated with relative hypoestrogenism.

Finally, in evaluating a woman of reproductive age, the relationship between her symptoms and her menstrual cycle should be considered. Emotional symptoms can vary across the menstrual cycle because reproductive hormones can influence vulnerability to stress as well as mood changes and anxiety. With an awareness of these fluctuations, women may benefit from a greater sense of predictability and control over their symptoms. In addition, overmedication or undermedication with pharmacological agents can be avoided.

See Also the Following Articles

Amenorrhea; Eating Disorders and Stress; Estrogen; Menopause and Stress; Premenstrual Dysphoric Disorder.

Further Reading

- Barnea, E. R. and Tal, J. (1991). Stress-related reproductive failure. *Journal of In Vitro Fertilization and Embryo Transfer* 8, 15–23.
- Berga, S. L. (1996). Functional hypothalamic chronic ovulation. In: Adashi, E. Y., Rock, J. A. & Rosenwaks, Z. (eds.) *Reproductive endocrinology, surgery and technology* (vol. 1), pp. 1061–1075. Philadelphia: Lippincott-Raven.
- Berga, S. L. (1996). Stress and ovarian function. *American Journal of Sports Medicine* 24, S36–S37.
- Berga, S. L., Daniels, T. L. and Giles, D. E. (1997). Women with functional hypothalamic amenorrhea but not other forms of anovulation display amplified cortisol concentrations. *Fertilization and Sterilization* 67, 1024–1030.
- Berga, S. L. and Girton, L. G. (1989). The psychoneuroendocrinology of functional hypothalamic amenorrhea. *Psychiatric Clinics of North America: Women's Disorders* 12, 105–116.
- Berga, S. L., Marcus, M. D., Loucks, T. L., et al. (2003). Recovery of ovarian activity in women with functional hypothalamic amenorrhea who were treated with cognitive behavioral therapy. *Fertilization and Sterilization* 80, 976–981.
- Buskirk, E. R., Mendez, J. and Durfee, S. (1985). Effects of exercise on the body composition of women. *Seminars in Reproductive Endocrinology* 3, 9–16.

- Carpenter, S. E. (1994). Psychosocial menstrual disorders: stress, exercise and diet's effect on the menstrual cycle. *Current Opinion in Obstetrics and Gynecology* 6, 536-539.
- Christia, J. S., Loyd, J. A. and Davis, D. E. (1965). The role of endocrines in the self regulation of mammalian population. *Recent Progress in Hormone Research* 21, 501-578.
- Chrousos, G. P. and Gold, P. W. (1992). The concepts of stress and stress system disorders. *Journal of the American Medical Association* 267, 1244-1252.
- Collins, A., Eneroth, P. and Landgren, B. M. (1985). Psychoneuroendocrine stress responses and mood as related to the menstrual cycle. *Psychosomatic Medicine* 47, 512-527.
- Dalton, K. (1964). The influence of menstruation on health and disease. *Proceedings of the Royal Society of Medicine* 57, 18-20.
- Drinkwater, B. L. (1984). Bone mineral content of amenorrheic athletes. *New England Journal of Medicine* 311, 277-281.
- Facchinetti, F., Fava, M., Fioroni, L., et al. (1993). Stressful life events and affective disorders inhibit pulsatile LH secretion in hypothalamic amenorrhea. *Psychoneuroendocrinology* 18, 397-404.
- Feingold, K. R., Gavin, L. A., Schambelan, M., et al. (1990). Female endocrinology. In: Andreoli, T. E., Carpenter, C. C., Plum, F. & Smith, L. H. (eds.) *Cecil essentials of medicine* (2nd edn., pp. 478-486). Philadelphia: Saunders.
- Freeman, E. W., Purdy, R. H., Coutifaris, C., et al. (1993). Anxiolytic metabolites of progesterone: correlation with mood and performance measures following oral progesterone administration to healthy female volunteers. *Neuroendocrinology* 58, 478-484.
- Fries, H., Nillius, S. and Pettersson, F. (1974). Epidemiology of secondary amenorrhea. II: A retrospective evaluation of etiology with special regard to psychogenic factors and weight loss. *American Journal of Obstetrics and Gynecology* 118, 473-479.
- Glick, R., Harrison, W., Endicott, J., et al. (1991). Treatment of premenstrual dysphoric symptoms in depressed women. *Journal of the American Medical Women's Association* 46, 182-185.
- Hendrick, V., Altshuler, L. L. and Burt, V. K. (1996). Course of psychiatric disorders across the menstrual cycle. *Harvard Review of Psychiatry* 4, 200-207.
- Herzog, D. B. and Copeland, P. M. (1985). Eating disorders. *New England Journal of Medicine* 313, 295-303.
- Marshall, J. (1989). Regulation of gonadotropin secretion. In: Degroot, L., Besser, G., Marshall, J., et al. (eds.) *Endocrinology*. Philadelphia: Saunders.
- Melo, N. K., Mendelson, J. H. and Lex, B. W. (1990). Alcohol use and premenstrual symptoms in social drinkers. *Psychopharmacology* 101, 448-455.
- Melo, N. K., Mendelson, J. H. and Palmieri, S. L. (1987). Cigarette smoking by women: interactions with alcohol use. *Psychopharmacology* 93, 8-15.
- Mirdal, G. M., Petersson, B., Weeke, B., et al. (1998). Asthma and menstruation: the relationship between psychological and bronchial hyperreactivity. *British Journal of Medicine and Psychology* 71, 47-55.
- Nepomnaschy, P. A., Welch, K., McConnell, D., et al. (2004). Stress and female reproductive function: a study of daily variations in cortisol, gonadotropins, and gonadal steroids in a rural Mayan population. *American Journal of Human Biology* 16, 523-532.
- Podolsky, E. (1963). The woman alcoholic and premenstrual tension. *Journal of the American Medical Women's Association* 18, 816-818.
- Reame, N., Sauder, S., Kelch, R., et al. (1984). Pulsatile gonadotropin secretion during the human menstrual cycle: evidence for altered frequency of gonadotropin-releasing hormone secretion. *Journal of Clinical Endocrinology and Metabolism* 59, 328-337.
- Rivier, C. and Vale, W. (1984). Influence of corticotropin releasing factor on reproductive function in the rat. *Endocrinology* 114, 914-921.
- Sapolsky, R. M., Romero, L. M. and Munck, A. U. (2000). How do glucocorticoids influence stress responses?: integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrine Reviews* 21, 55-89.
- Schachter, M. and Shoham, Z. (1994). Amenorrhea during the reproductive years: is it safe? *Fertilization and Sterilization* 62, 1-16.
- Shalts, E., Xia, L., Xiao, E., et al. (1994). Inhibitory effects of arginine-vasopressin on LH secretion in the ovariectomized rhesus monkey. *Neuroendocrinology* 59, 336-342.
- Shangold, M. M. (1985). Exercise and amenorrhea. *Seminars in Reproductive Endocrinology* 3, 35-43.
- Taylor, J. W. (1974). The timing of menstrual-related symptoms assessed by a daily symptom rating scale. *Acta Psychiatrica Scandinavica* 60, 87-105.
- Tilbrook, A. J., Turner, A. I. and Clarke, I. J. (2000). Effects of stress on reproduction in non-rodent mammals: the role of glucocorticoids and sex differences. *Reviews of Reproduction* 5, 105-113.
- Tilbrook, A. J., Turner, A. I. and Clarke, I. J. (2002). Stress and reproduction: central mechanisms and sex differences in non-rodent species. *Stress* 5, 83-100.
- Warren, M. P. (1996). Clinical review 77: evaluation of secondary amenorrhea. *Journal of Clinical Endocrinology and Metabolism* 81, 437-442.
- Xiao, E. and Ferin, M. (1997). Stress-related disturbances of the menstrual cycle. *Annals of Medicine* 29, 215-219.

Mental Stress Testing

P G Saab

University of Miami, Coral Gables, FL, USA

K A Kline

Virginia Military Institute, Lexington, VA, USA

J R McCalla

University of Miami, Coral Gables, FL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P G Saab and K A Kline, volume 2, pp 742–746, © 2000, Elsevier Inc.

Introduction

Conducting Mental Stress Testing

Task Differences in Responses to Mental Stressors

Individual Differences in Responses to Mental Stressors

Other Considerations

Conclusion

Glossary

<i>α-Adrenergic activity</i>	Reflecting activity of sympathetic nervous system receptors located in the vasculature; stimulation leads to constriction of the blood vessels.
<i>β1-Adrenergic activity</i>	Reflecting activity of sympathetic nervous system receptors located in the heart; stimulation leads to increases in heart rate and contraction of the heart muscle.
<i>Cardiac output</i>	Quantity of blood ejected from the heart in liters per minute.
<i>Endothelin-1</i>	Potent vasoconstrictor peptide released by vascular endothelial cells.
<i>Hemodynamic</i>	Pertaining to changes in blood flow reflecting changes in the underlying components of blood pressure, cardiac output, and total peripheral resistance.
<i>Left ventricular mass</i>	Size of the heart muscle of the left ventricle; enlargement of the left ventricle is a risk factor for cardiovascular morbidity and mortality.
<i>Nitric oxide</i>	Potent vasodilator substance released by vascular endothelial cells.
<i>Psychophysiological</i>	Pertaining to the study of physiological responses to psychological stimuli (e.g., stressors).
<i>Serotonin transporter gene Promoter polymorphism (5HTTLPR)</i>	A coding variant in the promoter of the human serotonin transporter gene which results in low transcriptional activity of this gene and is thought to be a genetic susceptibility factor for depression.
<i>Total peripheral resistance</i>	The resistance to blood flow in the systemic blood vessels.

Vagal tone

Index of parasympathetic nervous system influences on the cardiovascular system.

Introduction

Mental stress testing is typically used to examine the physiological effects of exposure to acute stressors, i.e., tasks, presented in the laboratory. Widely used as a research technique, mental stress testing is usually intended to distinguish groups with varying risk for cardiovascular diseases. To this end, mental stress testing functions as a provocative tool for unmasking information that is not available from assessments made under static conditions alone. As such, the psychophysiological adjustments that occur in the context of mental stress testing can enhance understanding of regulatory adjustments due to exposure to stress. Furthermore, these adjustments have implications for the reactivity hypothesis, which states that individuals who consistently display exaggerated cardiovascular responses (i.e., reactivity) to stressors may be at increased risk for the development of hypertension and/or coronary heart disease. The present discussion is limited to cardiovascular responses, although mental stressors also affect other systems such as those pertaining to neuroendocrine and immune function.

Conducting Mental Stress Testing

Baseline, Stress, and Recovery Periods

Mental stress testing requires assessment during resting baseline as well as during stress conditions. Regulatory mechanisms operative under resting conditions become modified as individuals are exposed to stressors. Examination under baseline conditions in the laboratory is requisite since it provides a comparison against which stressor-induced responses can be evaluated. Baseline measures may be affected by a variety of factors, including level of consciousness, habituation to the laboratory and procedures, time of day, posture, and prior substance ingestion (e.g., cardioactive agents such as caffeine, nicotine, and certain medications). As such, the baseline serves as a control level; a true basal level is not likely to be achieved under these circumstances. In relation to control levels, stressor-induced responses provide a sample of dynamic adjustments under varying conditions. In addition, recovery to baseline levels following the

termination of the stressor is often examined, as it provides further information about homeostatic mechanisms. To minimize extraneous influences and to ensure precise stimulus control, procedures for acquiring the baseline, stressor, and recovery data must be highly standardized. Efforts to enhance standardization have included scripted and/or taped instructions and computer presentation of stressor stimuli.

Types of Laboratory Stressors

Laboratory studies have utilized a wide range of stressors that differ to the extent that they elicit psychological versus physical stress. Social stressors include speech stressors (preparing and presenting a speech), interviews designed to elicit general or specific emotional responses, and tasks intended to mimic real-world social interactions, such as discussions of current events and role-play involving interpersonal conflict. Nonsocial psychological stressors include, but are not limited to, mental arithmetic (e.g., serial subtractions), mirror tracing (tracing a star looking only at its reflection in a mirror), reaction time (e.g., rapidly pressing a button in response to specific stimuli), the Stroop color-word interference test, video games, puzzle-solving tasks, and films (often consisting of graphic footage). Another commonly used stressor, the cold pressor test, is a physical stressor with a measurable psychological component. This task typically involves immersing a limb in ice water or placing an ice pack on the forehead. Finally, tasks such as static (e.g., squeezing a handgrip device) or dynamic (e.g., walking on a treadmill, riding a stationary bicycle) exercise are generally regarded as pure physical stressors and, thus, are not the focus of this article.

Several methods of classifying tasks have been proposed. A classic categorization scheme is Lacey's notion of sensory intake versus sensory rejection. Sensory intake tasks (e.g., reaction time) refer to stressors that require attention to visual or auditory cues in the external environment, while sensory rejection tasks (e.g., mental arithmetic) require mental work facilitated by blocking out external distractions. The most widely used classification system was proffered by Obrist, who distinguished laboratory stressors on the basis of the behavioral dimensions required for task completion, i.e., active versus passive coping. Active coping tasks provide an opportunity to influence the outcome of the situation by performing a specific behavior, while passive coping tasks require passive endurance of the stressor. Tasks such as shock avoidance reaction time, mental arithmetic, speech, video games, and the Stroop color-word interference test are considered active coping stressors. Examples of

passive coping tasks include the cold pressor test and film viewing.

Task Selection

Mental stress testing may involve presenting individuals with one or more laboratory stressors. It is recommended that task selection be guided by the specific questions that the research intends to address. For example, the research question would determine whether a stressor that elicits responses that are variable or uniform in magnitude is chosen. An explicit rationale for stressor choice also requires articulation. Selection decisions are often made on the basis of the behavioral demands associated with the stressors. To illustrate, an investigator may wish to compare responses to stressors involving active coping with those requiring passive coping. This would be particularly critical for studies comparing Black and White Americans. The literature indicates that Blacks, who as a group have a higher prevalence of hypertension than Whites, are more responsive to passive coping stressors, whereas Whites are more reactive to active coping stressors.

The relevance of stressors to the groups under study also requires consideration. Stressors, by virtue of the domains they tap, may hold differential relevance for one group relative to another. For example, a large body of research comparing males and females (children and adults) and employing cognitive and/or achievement-oriented stressors demonstrated that males show exaggerated physiological responses compared to females. Given the differential gender-related expectancies for performance that are likely in such situations, it is probable that those types of tasks may have been less relevant for the females. Therefore, it is essential to select stressors that are expected to be equally pertinent to participants.

When there is interest in the relationship between mental stress and a particular psychosocial variable, a fair test requires the inclusion of a task that is relevant to the specific construct. To illustrate, if one is interested in the relationship between trait anger and responses to mental stress, it is incumbent upon the investigator to include an anger-relevant task. This situation will provide an opportunity for individual differences to emerge.

Consideration also needs to be given to the response patterns evoked by the stressor when making stressor choices. Investigators often wish to compare stressors that elicit different hemodynamic responses. Such decisions are aided by the availability of noninvasive methods (e.g., impedance cardiography) to assess hemodynamic parameters such as cardiac output and total peripheral resistance during mental stress testing.

Task Differences in Responses to Mental Stressors

Research indicates that mental stressors evoke integrated response patterns across several parameters rather than an isolated response in a sole parameter. Consequently, a vast array of stressors elicits relatively few response patterns. Situational stereotypy, the propensity for a specific type of mental stressor to elicit a particular response pattern, is well known. Pattern 1 and Pattern 2 response patterns have been identified. Pattern 1 is characterized by skeletal muscle vasodilation, elevated heart rate, augmentation of blood pressure primarily as a function of increases in cardiac output, enhanced β 1-adrenergic activity, and decreased vagal tone. This myocardial response pattern is thought to characterize mental stressors involving active coping, sensory rejection, or mental work. Pattern 2 is associated with skeletal muscle vasoconstriction, smaller changes in heart rate, increases in blood pressure largely due to augmented total peripheral resistance, heightened α -adrenergic activity, and increased vagal tone. Stressors thought to elicit this vascular response pattern include those that involve passive coping, passive avoidance, inhibition, sensory intake, and vigilance. The use of adrenergic pharmacologic blockade has been informative with respect to clarifying the degree to which mental stressor response patterns are mediated by β -adrenergic and α -adrenergic receptor activity.

While Pattern 1 and Pattern 2 responses are often associated with active and passive coping stressors, respectively, some active coping tasks such as certain versions of mental arithmetic and speech (i.e., presentation period, not preparation) stressors are often associated with more of a mixed response pattern, composed of myocardial and vascular responses. It is well known that several variables influence and alter hemodynamic response patterns. A Pattern 1 response can be modified by a variety of factors, including instructional sets, expectations, appraisals, controllability, incentives, effort, predictability, individual differences, and the like, resulting in mixed response patterns.

Individual Differences in Responses to Mental Stressors

Responder Types

In addition to situational stereotypy, response stereotypy, which reflects the propensity for an individual to display a consistent response pattern across stressors, is also operative. A promising approach to the investigation of individual differences in reactivity to

stressors has been the identification of participants on the basis of their response patterns. The literature shows that some individuals respond in a consistent hemodynamic manner across active coping and passive coping mental stressors. In this regard, among men, Black Americans are more prevalent among vascular responders, whereas White Americans are more prevalent among myocardial responders across various stressors. Alternatively, evidence also supports the view that individuals may respond in a variable manner across myocardial and vascular dimensions to specific stressors. It should be noted, however, that for certain stressors, due to their stimulus characteristics, situational stereotypy might overwhelm the response stereotypy of the individual. For example, the cold pressor test typically elicits a Pattern 2 response pattern despite an individual's response stereotypy. Nonetheless, there is considerable individual variability with respect to how that stressor is tolerated and experienced.

Psychosocial Variables

The literature linking psychosocial variables to individual differences in cardiovascular reactivity assumes that when confronted with a stressor, individuals possessing certain psychosocial characteristics display exaggerated cardiovascular responses. The most commonly studied constructs include Type A behavior pattern and hostility/anger. A meta-analysis of the Type A reactivity literature concluded that Type A males, but not females, exhibit greater heart rate and systolic blood pressure reactivity to a variety of tasks than their Type B counterparts. These findings were dependent upon the type of task used, and the strength of findings varied with the method of Type A behavior pattern assessment (i.e., interview or questionnaire). Numerous studies have examined relationships among hostility and/or anger and cardiovascular reactivity. Qualitative and quantitative reviews of this literature conclude that hostility/anger is consistently related to reactivity to social/interpersonal stressors (e.g., involving provocation) but not to reactivity to standard nonsocial stressors.

Other Considerations

Prolonged Recovery

There is growing support for the notion that prolonged recovery has pathophysiological consequences. Evidence of delayed blood pressure recovery is available for hypertensives, as well as for groups at risk for future hypertension, such as borderline hypertensives, Blacks, and individuals with a family history

of hypertension. Furthermore, prolonged recovery from mental stress in borderline hypertensives has been shown to predict future sustained hypertension. Several recent studies have provided evidence of concurrent and prospective associations between recovery and outcome measures in adults and adolescents. Overall, research indicates that blood pressure and/or heart rate recovery may contribute uniquely to the prediction of ambulatory or resting cardiovascular measures.

Given increased recognition of the importance of recovery, investigation of the psychosocial correlates of post stressor responses is of interest. While a number of psychosocial factors have been examined in this regard, a promising variable is post stressor rumination, particularly when it involves the experience of anger. Studies have demonstrated delayed recovery from tasks involving harassment relative to nonharassment conditions, as well as more rapid recovery when participants are distracted and, thus, prevented from ruminating following emotionally provocative tasks.

Individual-Task Interactions

Understanding the potency of the mental stressor requires considering the interaction of the individual with the mental stressor. Although it is assumed that the stressor is perceived as stressful by the participants, in practice, this may not be the case. Appraisal of the stressor may be influenced by several situational variables, including the social context of the experiment, task ease, controllability, and novelty. In addition, subject variables such as experience, ability, involvement, and fatigue are also likely to impact perceptions. The individual's preexisting attitudes toward, experience with, and ability to perform similar tasks in real-life situations may also influence perception of the stressor. Stressor appraisal consequently will affect the extent to which the individual is engaged by the stressor. Engagement and involvement with the stressor are key in determining its efficacy and may reciprocally influence the effort that the individual expends when confronted with the stressor.

Stability

The notion that cardiovascular responses to mental stressors are relatively stable across tasks and over time is a critical assumption of the reactivity hypothesis. To make the case that exaggerated stress-induced responses have a detrimental impact on the cardiovascular system, it is first necessary to demonstrate that reactivity is a reproducible phenomenon. With regard to the stability of reactivity, there is increasing

evidence of adequate test-retest reliability for cardiovascular responses to laboratory stressors, particularly with task standardization, presentation of alternate forms of the stressor, aggregation of responses across tasks and testing sessions, and relatively short test-retest intervals. Furthermore, meta-analytic reviews report stability of responses over periods of years in adult and pediatric-aged samples.

The adequate evaluation of the health implications of recovery also requires reliable recovery measures. Research shows that the test-retest reliability of recovery using definitions, such as the time required to return to baseline levels following stressor termination, does not yield adequate test-retest reliabilities. Other findings suggest that the reliability of recovery scores might be improved by aggregating measures across tasks. Future research needs to be directed at identifying factors that affect long-term stability of both reactivity and recovery.

Laboratory-to-Field Generalization

A second implicit assumption of the reactivity hypothesis is that an individual's cardiovascular responses to laboratory stressors are representative of responses to stressors encountered in everyday life. Evidence for the laboratory-to-field generalizability of reactivity remains marginal. Critics maintain that generalizability is adversely affected by the failure to utilize ecologically valid tasks in mental stress testing protocols. Given the lack of accepted standards to demonstrate that a stressor is ecologically valid, the ecological validity of a stressor is typically based on face validity rather than on psychometric characteristics. For example, it has been suggested that the use of interpersonal or social stressors might improve generalization, as real-life stressors are likely to involve social factors; the data on this issue remain mixed. Future research needs to be directed at specifying the properties that determine a stressor's ecological validity. When making choices about whether a laboratory stressor should be used as a predictor, it is likely that the investigator would be best served by selecting a task that elicits stable cardiovascular responses over the long term. Given the chronicity of naturally occurring stressors, predictability might be enhanced when prolonged laboratory stressors, rather than the typically brief stressors, are employed.

Left Ventricular Mass

While definitive support for the reactivity hypothesis would require demonstrating an association for cardiovascular reactivity and the development of hypertension and/or coronary heart disease, establishing a

relationship with proximal medical endpoints would support the validity of the reactivity construct. The use of proximal endpoints, however, would not obviate the need for definitive longitudinal studies, which are not without their own limitations. Nonetheless, increasing attention needs to be directed at examining the association of cardiovascular reactivity evoked by mental stressors and proximal endpoints, such as increased left ventricular mass, a potent predictor of coronary heart disease. To date, the literature suggests that the relationship is modest at best. Further research examining this association is warranted, particularly as it pertains to left ventricular modeling and the contributions of responder type, ethnicity, and gender.

Genetic Contributions

The potential contributions of heritable factors to individual differences in cardiovascular reactivity have long been recognized. It has recently been proposed that future studies should involve systematic consideration of genetic, as well as environmental, factors as moderators or mediators of the association between reactivity and health outcomes. Technological advances are enabling the exploration of the role of specific genes in the determination of individual differences in reactivity. Some of the more promising candidates include, but are not limited to, the adrenergic receptor genes (i.e., β_1 , β_2 , α_1 , and α_2), the endothelin-1 and endothelin-1 receptor A genes, nitric oxide synthase genes, and the serotonin transporter polymorphism, i.e., 5HTTLPR. For instance, a few studies have demonstrated associations of polymorphisms of the β_1 - and β_2 -adrenergic receptor genes with blood pressure reactivity and/or resting or stress levels. Similarly, the 5HTTLPR polymorphism has been associated with greater heart rate and/or blood pressure responses. There is, however, conflicting evidence regarding whether greater reactivity is observed in individuals with one or two long 5HTTLPR alleles or in individuals with two short alleles, as well as whether such relationships are moderated by gender.

Conclusion

Mental stress testing typically involves the assessment of physiological responses in a laboratory setting during exposure to mental stressors and baseline and recovery conditions. Stressors range from psychological to physical and social to nonsocial. Stressors differ in the behavioral demands required of participants. Tasks also elicit certain well-defined physiological response patterns. Researchers should employ

stressors that most effectively address their research questions, taking into consideration the appropriateness of behavioral task demands, the significance of tasks for the groups being studied, and the relevance of tasks to psychosocial variables or response patterns of interest.

With few exceptions, there may be considerable individual variability in responses to mental stressors. Sources of variability include individual differences in the propensity to exhibit a particular response pattern, psychosocial characteristics, and a variety of preexisting and situational factors influencing individual-task interactions. In conducting mental stress testing, a number of additional areas of inquiry are worthy of consideration. These include the reproducibility of responses across stressors and testing sessions, the generalizability of laboratory responses to field settings, the predictability of proximal medical endpoints, and genetic contributions.

See Also the Following Articles

Autonomic Nervous System; Blood Pressure; Cardiovascular System and Stress; Hypertension; Type A Personality, Type B Personality.

Further Reading

- Blascovich, J. and Katkin, E. S. (eds.) (1993). *Cardiovascular reactivity to psychological stress and disease*. Washington, D.C: American Psychological Association.
- Kamarck, T. and Lovallo, W. R. (2003). Cardiovascular reactivity to psychological challenge: Conceptual and measurement considerations. *Psychosomatic Medicine* 65, 9–21.
- Krantz, D. S. and Manuck, S. B. (1984). Acute psychophysiological reactivity and risk of cardiovascular disease: A review and methodologic critique. *Psychological Bulletin* 96, 435–464.
- Lacey, J. I., Kagan, J., Lacey, B. C., et al. (1963). The visceral level: Situational determinants and behavioral correlates of autonomic response patterns. In: Knapp, P. H. (ed.) *Expression of the emotions in man*, pp. 161–196. New York: International Universities Press.
- Manuck, S. B. (1994). Cardiovascular reactivity in cardiovascular disease: “Once more unto the breach” *International Journal of Behavioral Medicine* 1, 4–31.
- Obrist, P. A. (1981). *Cardiovascular psychophysiology: a perspective*. New York: Plenum Press.
- Saab, P. G., Llabre, M. M., Hurwitz, B. E., et al. (1992). Myocardial and peripheral vascular responses to behavioral challenges and their stability in black and white Americans. *Psychophysiology* 29, 384–397.
- Schneiderman, N., Weiss, S. M. and Kaufmann, P. G. (eds.) (1989). *Handbook of research methods in cardiovascular behavioral medicine*. New York: Plenum Press.

- Schwartz, A. R., Gerin, W., Davidson, K. W., et al. (2003). Toward a causal model of cardiovascular responses to stress and the development of cardiovascular disease. *Psychosomatic Medicine* 65, 22–35.
- Snieder, H., Harshfield, G. A., Barbeau, P., et al. (2002). Dissecting the genetic architecture of the cardiovascular and renal stress response. *Biological Psychology* 61, 73–95.
- Treiber, F. A., Kamarck, T., Schneiderman, N., et al. (2003). Cardiovascular reactivity and development of preclinical and clinical disease states. *Psychosomatic Medicine* 65, 46–62.
- Turner, J. R., Sherwood, A. and Light, K. C. (eds.) (1992). *Individual differences: Cardiovascular response to stress*. New York: Plenum Press.

Metabolic Syndrome

L Keltikangas-Järvinen

University of Helsinki, Helsinki, Finland

© 2007 Elsevier Inc. All rights reserved.

What Is Metabolic Syndrome?

Origins of MetS

MetS and Environmental Stress

Development of MetS in Childhood

MetS and Inherited Temperament

Glossary

<i>Innate temperament</i>	A biologically rooted behavioral disposition that is stable across time and situations.
<i>Metabolic syndrome (MetS)</i>	A constellation of physiological and metabolic abnormalities.

What Is Metabolic Syndrome?

Metabolic syndrome (MetS), also called insulin resistance syndrome, is a cluster of various physiological and metabolic abnormalities including hyperinsulinemia, hyperglycemia, hypertension, a decreased plasma concentration of high-density lipoprotein (HDL) cholesterol, an increased plasma concentration of very low density lipoprotein (VLDL) triglyceride, glucose intolerance, and abdominal obesity. Insulin resistance is the primary metabolic defect of this syndrome, with compensatory hyperinsulinemia being the common denominator ultimately responsible for other changes of this constellation. In addition to insulin resistance, abdominal obesity is a key contributor to the development of MetS (see **Figure 1**).

MetS is a huge public health problem worldwide that is responsible for a growing number of premature deaths throughout the world. Its prevalence is

approximately 16% among Caucasians. However, comparison across different countries is difficult as there are different age structures of the population, and even though a World Health Organization (WHO) expert committee published a definition of MetS, there are other, differing definitions available.

MetS is significant because it plays an important role in the etiology of coronary heart disease (CHD) and non-insulin-dependent diabetes mellitus (type II diabetes). Type II diabetes is one of the most prevalent and serious metabolic diseases in the world, and among middle-aged men, CHD is still the leading cause of mortality in Western countries.

Origins of MetS

The exact cause of MetS is not known. It is generally accepted that a genetic predisposition, possibly an interaction of several genes, is an essential factor in the genesis of this disorder. In addition to a genetic predisposition, lifestyle factors such as physical inactivity, habitual alcohol drinking, and smoking play a role in the development of MetS. There is also compelling evidence that stress of diverse origins has a significant role in the development of MetS. It is well established that the development of MetS is indisputably related to activity and/or dysregulation in the two major physiological and neuroendocrine stress systems, the sympathetic-adrenomedullary (SA) system and the hypothalamic-pituitary-adrenomedullary (HPA) system, a key component being the activation of the HPA axis.

The primary role of SA and HPA systems is increasing the energy available for action; activity of these systems is needed in everyday performance. The label of stress systems is based on the fact that they are very sensitive to physical and mental stress. Stress usually refers to overactivity or dysregulation of those systems, not to their ordinary functioning.

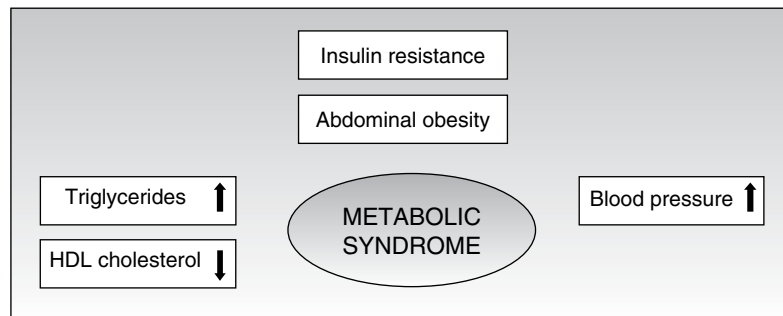


Figure 1 Contributing factors of metabolic syndrome.

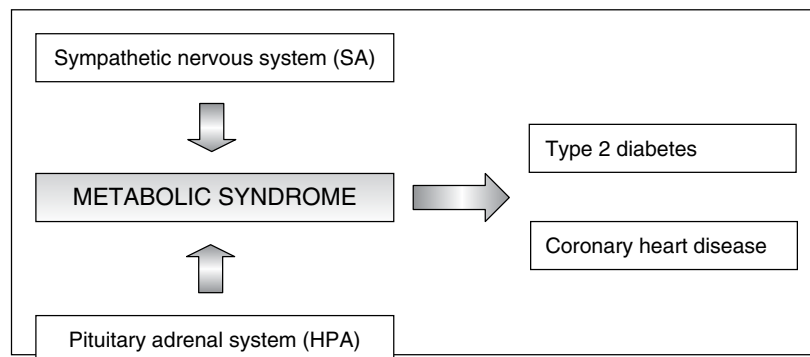


Figure 2 Stress contributes to the development of diabetes and coronary heart disease through metabolic syndrome.

SA and HPA systems are both known to exert marked insulin resistance. This is especially true with responses of serum cortisol, as it regulates glucose levels in the blood, with all of the accompanying consequences. To summarize, mental stress might contribute to the development of MetS and subsequently increase the risk of CHD and diabetes via the SA and HPA axes, with increased cortisol secretion being of basic importance (see [Figure 2](#)).

MetS and Environmental Stress

Evidence of the contribution of environmental stress to MetS via increased reactivity of the stress axis was first derived from animal experiments. It was shown that in standardized experimental stress situations, macaque monkeys reacted with helplessness and loss of control (i.e., with a behavioral stress reaction) and responded with an increased level of cortisol and detached MetS symptoms, especially body fat. Comparable findings have been reported with rats exposed to chronic, uncontrollable environmental stress.

In line with animal studies, it was shown in laboratory experiments with human beings that persons with a high level of MetS symptoms were likely to express elevated responses of serum cortisol during task-induced mental stress.

Research has predominantly emphasized the role of the HPA axis when searching for the mechanism mediating the relationship between stress and MetS. Enhanced sympathetic activity (SA activity) is, however, important as well because chronic sympathetic overactivity with hemodynamic consequences (increased heart rate and pulse pressure) may have pathogenic importance in the development of MetS.

It is also known that long-lasting stress predisposes a person to stronger acute stress reactions, both mental and physical; that is, chronically stressed individuals are likely to show greater reactivity to and prolonged recovery from short-term challenging tasks.

Evidence of the influence of stress on MetS has mostly been based on laboratory experiments, whereas real-life situations, that is, the influence of everyday mental stress, especially long-lasting stress, have not been well studied. The risk characteristic that has been systematically shown to be related to both reactivity of the HPA axis and MetS is vital exhaustion. Vital exhaustion is seen as an indicator of long-lasting mental stress and is characterized by feelings of excessive fatigue, loss of energy, increased irritability, and demoralization. It is also well established that vital exhaustion is a risk factor for CHD, especially for myocardial infarction. The highest risk here, however, is not related to a high level of cortisol but to

hypocortisolemia, a state in which both mental and physiological reservoirs are unavailable. This questions the common argument that high risk would be related to high cortisol secretion.

We also know that several coronary-prone personality characteristics, such as type A behavior (i.e., a behavior pattern consisting of aggression, impatience, high striving for achievement, ambitiousness, competitiveness), hostility, and anger, are related to perceived stress and reactivity of SA and the HPA axis. They are also related to components of MetS, even though the empirical findings have not been consistent.

If environmental stress is likely to increase the level of MetS, there is also an inverse effect: environmental support is able to decrease the level of MetS. It is known that changes in perceived social support are associated inversely with changes in the MetS level; that is, increasing support decreases the level of MetS and vice versa. The effect of social support in stress and diseases is widely documented. It increases the overall positive effect, feelings of self-worth and acceptance. It may exert stress-buffering effects by affecting the appraisal of potential stressors or by bolstering coping efforts.

Development of MetS in Childhood

It is generally accepted that atherosclerosis, the underlying process of CHD, originates in childhood and adolescence and that CHD risk factors are already present then. In addition, fasting insulin levels in early adolescence have been shown to predict type II diabetes in young adulthood.

The MetS precursors also appear to be present as early as childhood. It has been shown that the essential parameters of MetS express strong clustering and high tracking, even in healthy children and

adolescents. Consequently, the potential effect of stress on the parameters of MetS even in childhood and adolescence has evoked interest. Until recently, the studies were few in number. There are, however, findings that suggest that a stress-prone and inappropriate childhood environment predisposes a child to a high level and early tracking of the parameters of MetS. Sources of environmental childhood stress are, of course, likely to stem from the family environment, especially from parental child-rearing practices. It is well established that a mother's hostile child rearing, consisting of a strict disciplinary style, a tendency to emotionally reject the child, and a feeling that the child is a burden on her, preventing her self-fulfillment, indicates an environmental stress that, in some studies, has been shown to be correlated with an elevated level of a child's MetS parameters.

With regard to the child's own characteristics, some components of type A behavior in adolescence predict changes in the level of MetS, with aggressive competitiveness being especially associated with an increase of that syndrome. In addition, some findings suggest that childhood temperament, especially hyperactivity and negative emotionality, is associated with and predict a later elevated level of MetS (see [Figure 3](#)).

MetS and Inherited Temperament

The findings previously described are related to the question of whether inherited temperament could play a role as a stress buffer or make a person vulnerable to stress and consequently explain a variance of MetS. Temperament refers to biologically rooted individual differences in behavioral reactions and dispositions that are present early in life and are relatively stable across time and situations. Temperament

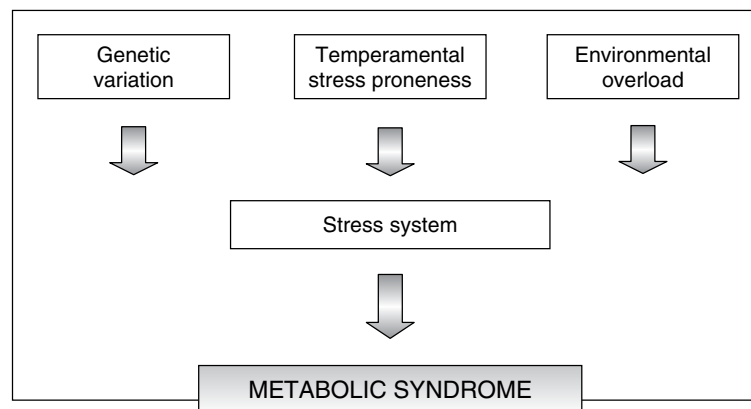


Figure 3 Origins of metabolic syndrome.

is an inherited core of personality; personality is developed by the interaction of inherited temperament and environment, mostly parental child-rearing practices.

Temperament has its origins in brain structures, and individual differences in temperament reflect innate differences in neuroendocrine and autonomic nervous system functions. Therefore, a biologically determined temperament has been viewed as one of the sources responsible for individual differences in stress proneness. It has been suggested to explain the variance in events that are experienced as stressful, and especially why daily troubles have greater health consequences for some individuals than others. There is, indeed, increasing evidence that differences in SA and HPA systems, i.e., differences in major stress systems, are attributable to different temperament profiles.

Findings suggest that childhood temperament plays a direct and indirect role in a development of MetS; that is, it interacts with the environment, meaning that some childhood temperamental factors (e.g., hyperactivity and negative emotionality) are related to an adulthood risk for MetS. In addition, childhood temperament also has an indirect effect. It has been suggested that the influence of hostile maternal child-rearing practices on the child's MetS level may be mediated by the child's temperament. Children with a difficult temperament, conceptualized in terms of negative emotionality (aggression and anger proneness), high activity, and low sociability, are especially vulnerable. Children with this type of temperament are especially prone to high levels of MetS when faced with hostile child rearing.

Interest in the role of adulthood temperament in individual stress reactions has been increasing of late. Findings have conflicted slightly with previous opinions on risk personality. A coronary-prone personality has traditionally been described in terms of hostility, depression, social aloneness, and type A behavior. In contrast, risk temperament is not primarily related to these kind of features; the most vulnerable individuals (i.e., persons who expressed systematically high levels of MetS) are those dealing with oversocialization and overresponsibility. These persons are highly sensitive to social cues of approval or the praise of others, have a strong tendency to conform to peer pressures, and are sensitive to the rejection of others, leading to excessive reward-seeking behaviors such as overworking in response to social rejection or frustration. As workers, they are persistent overachievers who frequently push themselves to exhaustion.

Issues concerning the role of temperament in physiological and neuroendocrine stress reactivity are still

new and thus in process. The findings, however, have been promising for explaining the complicated relationship between mental stress and MetS. The findings also suggest that psychological and physiological stress symptoms might have the same source – it may be possible in the future to find a solid neurobiological basis for both psychological and physiological stress reactions.

See Also the Following Articles

Atherosclerosis; Heart Disease/Attack; Stress, Insulin Resistance and Type II Diabetes; Environmental Stress, Effects on Human Performance; Type A Personality, Type B Personality; Hypothalamic-Pituitary-Adrenal.

Further Reading

- Björntorp, P. (1991). Metabolic implications of body fat distribution. *Diabetes Care* **14**, 1132.
- Björntorp, P. (1992). Regional fat distribution: implications for type II diabetes. *International Journal of Obesity* **16** (supplement 4), S19–S27.
- Björntorp, P. and Rosmond, R. (2000). The metabolic syndrome – a neuroendocrine disorder? *British Journal of Nutrition* **83**(supplement 1), S49–S57.
- Black, P. H. (2003). The inflammatory response is an integral part of the stress response: implications for atherosclerosis, insulin resistance, type II diabetes and metabolic syndrome X. *Brain, Behavior, and Immunity* **17**, 350–364.
- DeFronzo, R. A. and Ferrannini, E. (1991). Insulin resistance: a multifaceted syndrome responsible for NIDDM, obesity, hypertension, dyslipidemia, and atherosclerotic vascular disease. *Diabetes Care* **14**, 173–194.
- Hansel, B., Giral, P., Nobecourt, E., et al. (2004). Metabolic syndrome is associated with elevated oxidative stress and dysfunctional dense high-density lipoprotein particles displaying impaired antioxidative activity. *The Journal of Clinical Endocrinology and Metabolism* **89**, 4963–4971.
- Hjemdahl, P. (2002). Stress and the metabolic syndrome: an interesting but enigmatic association. *Circulation* **106**, 2634–2636.
- Julius, S. and Gudbrandsson, T. (1992). Early association of sympathetic overactivity, hypertension, insulin resistance, and coronary risk. *Journal of Cardiovascular Pharmacology* **20**, S40–S48.
- Keltikangas-Järvinen, L., Räikkönen, K., Hautanen, A., et al. (1996). Vital exhaustion, anger expression, and pituitary and adrenocortical hormones. *Arteriosclerosis, Thrombosis, and Vascular Biology* **16**, 275–280.
- Lteif, A. and Mather, K. (2004). Insulin resistance, metabolic syndrome and vascular diseases: update on mechanistic linkages. *Canadian Journal of Cardiology* **20**(supplement B), 66B–76B.
- Ravaja, N. and Keltikangas-Järvinen, L. (1995). Temperament and metabolic syndrome precursors in

- children: a three-year follow up. *Preventive Medicine* 24, 518–527.
- Rosmond, R. (2005). Role of stress in the pathogenesis of the metabolic syndrome. *Psychoneuroendocrinology* 30, 1–10.
- Räikkönen, K., Matthews, K. A. and Kuller, L. H. (2002). The relationship between psychological risk attributes and the metabolic syndrome in healthy women: antecedent or consequence? *Metabolism* 51, 1573–1577.
- Seematter, G., Binnert, C., Martin, J. L., et al. (2004). Relationship between stress, inflammation and metabolism. *Current Opinion in Clinical Nutrition and Metabolic Care* 7, 169–173.

Metabolic Syndrome and Stress

R Rosmond

Björndammsterrassen 41, Partille, Sweden

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Björntorp and R Rosmond, volume 2, pp 747–749, © 2000, Elsevier Inc.

The Impact of Stress in Animal Studies

Effects of Stress in Humans

The Physiology of Stress

The Long-Term Effect of Stress on the Human Body

Concluding Remarks

Glossary

<i>Metabolic syndrome</i>	A well defined group of metabolic risk factors such as insulin resistance, atherogenic dyslipidemia, central (abdominal) obesity and hypertension.
<i>Hypothalamic-pituitary-adrenal axis</i>	The hormonostatic endocrine axis that regulates cortisol production and secretion.
<i>Sympathetic nervous system</i>	A division of the autonomic nervous system which mediates the response of the body to stress.

In recent years, scientists have recognized that some risk factors for cardiovascular disease and type 2 diabetes cluster together in certain people. The causal risk factors involved in the significant increase in cardiovascular disease and type 2 diabetes stem mostly from a now well-defined group of multiple metabolic risk factors called the metabolic syndrome (MetS). These risk factors include elevated insulin levels (insulin resistance), central (abdominal) obesity, atherogenic dyslipidemia, and hypertension. While these combined risk factors do not usually cause overt disease symptoms, it is obvious that persons with at least several of these correlated factors

greatly increase the likelihood of atherosclerosis, heart disease, stroke, diabetes, kidney disease, and ultimately even premature death. Complicated genetic factors such as race, age, family weight, and diabetes history that interact equally with complicated cultural, dietary, psychosocial stress, exercise, and smoking factors, result in a heterogeneous at-risk patient group yet surprisingly consistent poor prognosis.

As MetS is gaining wider recognition in the clinic as a serious life-threatening condition, researchers and doctors are struggling to formalize the connections among all the diverse symptoms, risk factors, and correlated conditions. Of all the factors influencing MetS, chronic psychosocial stress is likely the most pervasive and controllable. Stress has long been recognized as a major contributor to hypertension, endocrine dysfunction, diabetes, and elevated cortisol-related conditions such as gluconeogenesis, increased serum glucose levels, decreased insulin production, perpetuation of catecholamine vasoconstriction, increased arterial blood pressure, fat and protein mobilization, and centralization of body fat. Considering the overlap in symptomatology between stress-correlated conditions and MetS, connecting the two, at least heuristically, is not a stretch. Indeed, solid evidence now exists that suggests the two conditions have reciprocal effects, with chronic stress including MetS and increased intra-abdominal fat and other hormonal dysfunction augmenting cortisol-induced stress and somatopsychic effects.

The Impact of Stress in Animal Studies

The activity of the hypothalamic-pituitary-adrenal (HPA) axis has been studied in several animal species in a variety of models of psychosocial stress. Not surprisingly, the majority of studies indicate that the HPA axis is activated in low-ranking animals in hierarchical social groups and in animals that have been defeated by a conspecific. However, activity and reactivity of the HPA axis has been shown

to be modulated by a variety of different factors, including the species, gender, and behavioral style of the individuals.

Socially subordinated animals are generally more reactive to a novel stressor compared to their dominant counterparts, as shown in social groups of mice, rats, hamsters, guinea pigs, squirrel monkeys, and olive baboons. Taken together, these data indicate that social subordination and defeat appear to be stressful and lead to HPA axis activation. Chronic social stress can lead to long-term changes in HPA axis activity, including persistent elevations in basal glucocorticoids, abnormal responses to subsequent stressors, and impaired feedback regulation. For the most part, these effects are seen most clearly in subordinated animals housed in stable social groups; however, similar responses have been observed in dominant animals in such groups, and also in animals of all ranks in unstable social grouping.

Effects of Stress in Humans

Social stress is generally viewed as a major factor in the etiology of a variety of psychopathologies such as depression and anxiety, in addition to its effects on male and female reproduction, immune functioning, heart disease, etc. Social stress in people is often evaluated in terms of the number and magnitude of life events that an individual experience, and a general conclusion from this approach is that a plethora of moderately stressful events can have as great an impact as a few major events. Another important index that is strongly associated with the number of stressful events that are likely to be experienced is social status. Low social status is regarded as having an impact on almost every area of the individual's life. Furthermore, the ranking difference itself, and the meaning assigned by the individual to his or her status with reference to others, may provide stress that is additional to (or interactive with) the material consequences of low status.

The Physiology of Stress

The stress response occurs through two major pathways in the human body: the central nervous system (CNS) and the endocrine system. The sympathetic nervous system (SNS) of the CNS that is involved when a stressor is encountered is the SNS stimulation of the adrenal glands, which, releases the catecholamines epinephrine (adrenaline) and norepinephrine (noradrenaline). The resulting fight-or-flight response via catecholamine release from end target neurons has an immediate effect by increasing heart rate, force of

myocardial contraction and cardiac output, arterial blood pressure, blood coagulation, and serum glucose.

Simultaneously with nervous system stimulation, the hypothalamus and the pituitary glands secrete hormones into the general circulatory system. Once in the bloodstream, they arrive at the adrenal glands, which then secrete glucocorticoids, mainly cortisol, into the bloodstream to assist meeting the demands of the stressors. This endocrine pathway of hormones increases body metabolism, increases fatty acids in blood, increases blood pressure, and retains extra sodium, which results in increased water retention.

The Long-Term Effect of Stress on the Human Body

Since the early 1980s, a wealth of information has accumulated which illustrates that activation of the HPA axis leads to suppression of both the growth axis and the reproductive axis including activation of the sympathetic system. Chronic increase in cortisol, catecholamines, and chronic suppression of the growth and reproductive axes leads to a number of adverse health effects, and their sequelae result in increased morbidity and mortality.

Cortisol interferes at several levels of insulin action. In addition, cortisol inhibits insulin secretion from pancreatic β -cells. In primary cultured adipocytes, synthetic glucocorticoids (dexamethasone) induce progressive insulin resistance by sequentially regulating multiple aspects of the insulin-responsive glucose transport system. In the early stages, dexamethasone impairs the ability of insulin to translocate intracellular glucose transporters (GLUT 4) to the cell surface. Some, but not all, studies indicate that glucocorticoids may increase hepatic glucose metabolism. This diabetogenic effect of glucocorticoids is well characterized at least in animal models. In skeletal muscle, cortisol inhibits the glycogen synthase. Recent data indicate that subjects with insulin resistance do have an increased number of glucocorticoid receptors (GRs) in muscle.

In summary, initially hyperglycemia is accompanied by an increased insulin response to a glucose load, but eventually β -cell response becomes insufficient, and the unopposed actions of stress-related counterregulatory catabolic hormones (cortisol, catecholamines) serve to exacerbate the metabolic disturbances and contribute to the biochemical changes seen. This simplified account provides a framework within which the association between stress and insulin resistance can be understood.

Cortisol regulates adipose-tissue differentiation, function, and distribution, and in excess, causes

visceral obesity. Glucocorticoids exert their cellular action by complexing with a specific cytoplasmic GR, which in turn translocates to the nucleus and binds to specific sites on chromatin. GRs belong to the superfamily of nuclear transcription factors. Cortisol exerts chronic effects on lipid metabolism. One of the most striking effects in humans is the centralization of body fat seen in Cushing syndrome, exhibited as severe truncal obesity. Excess cortisol promotes the activity of lipoprotein lipase (LPL), the primary enzyme responsible for conversion of lipoprotein triglyceride into free fatty acids and accumulation of triglycerides in adipocytes. The density of GRs, assessed by ligand-binding techniques and hybridization techniques with GR cRNA probes, is particularly high in intra-abdominal adipose tissue as compared to other regions. Furthermore, the antilipolytic action of insulin during the prevailing hyperinsulinemic state causes a blunted lipid mobilization, resulting in a condition with increased lipid accumulation. Moreover, in the presence of insulin, cortisol has a marked stimulatory effect on LPL activity in human adipose tissue *in vitro*. This effect involves both an increased level of LPL mRNA, leading to increased relative LPL synthesis, and additional posttranslational regulation.

The primary reason for considering stress as an etiologic factor relates to the influence of the SNS and neuroendocrine factors in hypertension, although high salt intake in some subjects may contribute through volume expansion. The principal hypothesis is that the stress-induced activation of hypothalamic, sympathohormonal regions occurs repeatedly over time, leading to cardiovascular adjustments that increase hypertension risk. Evidence relevant to this hypothesis derives from both animal and human experimental studies. Complications such as myocardial infarction and stroke are not directly due to elevated pressures, but to the resulting structural changes in the heart and blood vessels. One of the structural consequences of hypertension, hypertrophy of the left ventricle, is the strongest predictor, other than advancing age, of cardiovascular morbidity and mortality.

Concluding Remarks

The human body has many effector systems including the SNS and the HPA axis for maintaining homeostasis and connecting the brain with the periphery. Overall these systems interact both indirectly and

directly. The HPA axis determines most of the acute and prolonged effects of stressors. The secretory end product of the HPA axis, cortisol, is kept within an optimal range through the feedback action of cortisol interacting with neural control mechanisms. Distressing events or situations evokes these two pathways of the stress response. However, each person's unique combination of heredity, life experience, personality, and ability to cope are all involved in the perception of an event and the meaning attached to it. Rather than confronting life-threatening stressors, life today centers around events that hold symbolic meaning; an upcoming test, a disagreement with a friend, or the boredom of a job.

Individuals with persistent psychosocial and socioeconomic handicaps such as anxiety and depression, poor economy, low education, and unemployment are more likely to suffer from frequent challenges of the SNS and the HPA axis. Thus, psychosocial stress preceded by genetic susceptibility that result in a disruption of central regulatory systems all adds up to less coping ability, followed by insulin resistance, atherogenic dyslipidemia, central (abdominal) obesity, and hypertension, that is, the metabolic syndrome.

Further Reading

- Björntorp, P. and Rosmond, R. (2000). Obesity and cortisol. *Nutrition* **16**, 924–936.
- Björntorp, P., Holm, G., Rosmond, R., et al. (2000). Hypertension and the metabolic syndrome: closely related central origin? *Blood Pressure* **9**, 71–82.
- Charmandari, E., Tsigos, C. and Chrousos, G. (2005). Endocrinology of the stress response. *Annual Reviews in Physiology* **67**, 259–284.
- Dallman, M. F., Pecoraro, N. C. and la Fleur, S. E. (2005). Chronic stress and comfort foods: self-medication and abdominal obesity. *Brain Behavior and Immunity* **19**, 275–280.
- Henry, J. P. (1992). Biological basis of the stress response. *Integrative Physiological and Behavioral Science* **27**, 66–83.
- Hjemdahl, P. (2002). Stress and the metabolic syndrome: an interesting but enigmatic association. *Circulation* **106**, 2634–2636.
- Miller, D. B. and O'Callaghan, J. P. (2002). Neuroendocrine aspects of the response to stress. *Metabolism* **51**, 5–10.
- Rosmond, R., Dallman, M. F. and Björntorp, P. (1998). Stress-related cortisol secretion in men: relationships with abdominal obesity and endocrine, metabolic and hemodynamic abnormalities. *Journal of Clinical Endocrinology and Metabolism* **83**, 1853–1859.

Metals, Oxidative Stress, and Brain Biology

D I Finkelstein, T Lynch, S Wilkins, R A Cherny and A I Bush

The Mental Health Research Institute of Victoria,
Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

Introduction

Biometals in the Developing Organism

Biometals in the Adult

Parkinson's Disease

Alzheimer's Disease

Glossary

Alzheimer's disease (AD)

A progressive form of dementia with symptoms that include impaired memory, usually followed by impaired thought and speech. AD is histologically characterized by the deposition of extracellular amyloid plaques, the major constituent being the amyloid- β peptide (A- β).

Biometals

Metals found in living organisms that have defined physiological functions that serve to maintain normal cellular processes. These include copper (Cu), zinc (Zn), iron (Fe), and manganese (Mn).

Dopamine (DA)

A neurotransmitter that normally acts to reinforce behavior. Drugs of addiction act by causing dopamine release; a lack of dopamine results in the symptoms of Parkinson's disease. L-Dopa treatment provides the precursor for DA that partially compensates for the reduction of DA observed in Parkinson's disease.

Oxidative stress

The damage to cells caused by reactive oxygen species (ROS; e.g., hydrogen peroxide). Metals such as iron and copper are capable of redox chemistry in which a single electron may be accepted or donated by the metal, reacting with oxygen and thereby producing radicals and ROS. Oxidative stress is imposed on cells as a result of an increase in oxidant generation, a decrease in antioxidant protection, or a breakdown in oxidative damage-repair mechanisms.

Parkinson's disease (PD)

A debilitating disease occurring only in humans and resulting from a predominant loss of dopamine neurons. The symptoms include muscle rigidity, tremor, slow and shuffling gait, masklike facial expressions, and eventually cognitive decline and death. Some neurons

Reactive oxygen species (ROS)

contain Lewy bodies that consist of intracellular inclusions of α -synuclein.

Natural by-products of the normal metabolism of oxygen that have important roles in cell signaling. During times of biological stress, ROS levels can increase, resulting in significant damage to, or death of neurons.

Substantia nigra (SN)

A specific part of the brain that is made up of neurons that contain the chemical neurotransmitter dopamine; loss of these cells occurs in Parkinson's disease.

Toxicological metals

Metals with no known normal biological function; these include lead and aluminum. These metals can interact with many proteins, resulting in the dysfunction of many biological processes.

Introduction

The brain is an exceptional tissue in that it concentrates biometals (copper, zinc, iron, and manganese) for its normal functional metabolism. Some of these processes include maintaining enzymatic and mitochondrial function, immunity, myelination, learning and memory, co-release with neurotransmitters, and binding with prion proteins. The importance of metals in antioxidant defense mechanisms should also be highlighted; for example, Cu/Zn superoxide dismutase protects against oxidative stress. The concentration of biometals in the brain, together with the brain's high levels of oxygen, appears to make the central nervous system vulnerable to a host of metal-centered pathologies.

Biometals in the Developing Organism

Deficiencies in biometals result in increased mortality, abnormal development, and poor growth. The accumulation of iron in the brain appears to be handled differently than the other biometals. Whereas copper, zinc, and manganese can be absorbed by the brain after the organ has developed, the neonatal period is critical for the establishment of normal iron content in the adult brain. Iron accumulates in the brain only while it is still growing; for rats and mice this occurs in the first few postnatal weeks, whereas in humans the period begins during the third trimester of pregnancy and is maintained throughout the first year of life. Iron, copper, zinc, and manganese are found in the milk of humans and other mammals. Mammalian milk meets the crucial

supply requirements needed for the normal development of the neonate, with these biometals varying in concentration during the course of lactation. Interestingly, the iron content of milk is not reduced even when the iron status of the mother is poor, indicating tight regulation of the iron content of milk. Toxic milk syndromes have been observed in mice that have genetic defects in the copper or zinc pathways resulting in the abnormal delivery of these biometals, with significant biological consequences. Deliberately increasing the exposure of neonates to iron in this critical period has been shown to elevate iron levels in the adult brain.

Biometals in the Adult

Early studies used the term “trace element” to designate elements that are essential for normal function, including zinc, manganese, copper, iodine, iron, cobalt, molybdenum, tin, selenium, and chromium. The minimum nutritional requirements for individuals are hotly debated in the popular and scientific literature and are not discussed here. What is clear is that various genetic mutations can cause a disruption in the homeostasis of these elements, resulting in dramatic biological consequences (e.g., Segawa’s syndrome and iron; Wilson’s disease and copper). Biometals normally bind to a range of specific proteins (e.g., Cu/Zn and Mn bind to superoxide dismutase; Cu binds to the prion protein). Poisoning occurs when the levels of biometals exceed the body’s capacity to bind them to specific metal-binding proteins. Two outcomes likely to result from this are that (1) the metal binds to an incorrect protein, causing dysfunction in a biological system or (2) the presence of such metals in an unbound form causes oxidative stress, which damages cells.

Parkinson’s Disease

Parkinson’s disease (PD) is characterized by the loss of dopaminergic neurons of the substantia nigra (SN) and the deposition of intracellular inclusion bodies. The principal component of these deposits is the α -synuclein protein that is expressed throughout the brain, and mutations in this protein (A30P and A53T) are responsible for familial forms of the disease. The association between iron and PD is multifaceted and complex; too little iron leads to neurological deficits, yet too much appears to result in neuronal loss. There is long-standing evidence of a selective increase in iron in the SN of patients with PD as well as an accumulation in animals following artificially induced lesions to the SN. Iron is required for the synthesis of dopamine (DA): tyrosine hydroxylase, a nonheme

iron-containing enzyme uses molecular oxygen to hydroxylate tyrosine, resulting in the formation of L-Dopa, the natural biochemical precursor of DA and the most common medication for the treatment of PD. Mutations in the iron-binding site of tyrosine hydroxylase have been identified in cases of L-dopa-responsive PD and Segawa’s syndrome (a rare hereditary progressive dystonia with diurnal fluctuation). Other redox-active metals such as manganese have been associated with neurological disorders. Among the neurological effects of manganese poisoning (linked to occupational exposure) is an irreversible Parkinsonian-like syndrome. The specific mechanisms of the toxicity are not known.

Redox-active iron, α -synuclein, and DA are abundant in the SN. The three elements colocalized at one anatomical location can interact to generate toxic reactive oxygen species (ROS). It is apparent, however, that none of these elements alone is sufficient to perpetuate the neurodegeneration of PD: rather, the combination of iron, α -synuclein, and DA together generates toxic products. For instance, DA is a potent chelator of transition metals and can generate hydrogen peroxide in the presence of iron (or copper) and α -synuclein. Under these conditions, α -synuclein molecules link to form soluble aggregates known as oligomers. These oligomers are thought to be deposited as the microscopically visible intracellular inclusion bodies called Lewy bodies. This pathological mechanism has parallels with Alzheimer’s disease, in which current theory suggests that soluble oligomeric intermediates, rather than the visible aggregates (plaques), are the cause of neurodegeneration.

Alzheimer’s Disease

Alzheimer’s disease (AD) is characterized by the deposition of extracellular amyloid plaques, the major constituent being the amyloid- β protein (A- β) that is cleaved from the membrane-bound amyloid precursor protein. There is considerable evidence for an imbalance in the homeostasis of biometals in AD, which is reflected by specific alterations in metal levels in the brain (in areas such as the hippocampus, amygdala, neocortex, and olfactory bulb). These observations take into account the varying background of biometals that exists throughout life. Not only do the levels of biometals change with age, but so do the concentrations of different metal transport and storage proteins. Whereas metal levels in the brains of laboratory mice rise dramatically with age, studies of brain-metal content in aging humans are less conclusive. Interestingly, of the major biometals only zinc appears to be elevated in tissue from the

human AD brain. However, copper, zinc, and iron are present in very high concentrations, bound to the aggregated A- β that makes up the amyloid plaques. Although the normal function of A- β is unknown, it has been proposed that the biochemical environment of the aging brain provides conditions in which the A- β protein pathologically binds to copper and zinc, causing it to form oligomers that foster unregulated redox activity and toxic free-radical formation. A paradoxical consequence of the buildup of metals in the amyloid plaques is that the brain cells become deficient in these crucial biometal components, rendering them more susceptible to oxidative stress.

Further Reading

Barnham, K. J., Masters, C. L. and Bush, A. I. (2004). Neurodegenerative diseases and oxidative stress. *Nature Reviews Drug Discovery* 3(3), 205–214.

Berg, D., Gerlach, M. B., Youdim, K. L., et al. (2001). Brain iron pathways and their relevance to Parkinson's disease. *Journal of Neurochemistry* 79(2), 225–236.

Blennow, K., de Leon, M. J. and Zetterberg, H. (2006). Alzheimer's disease. *Lancet* 368(9533), 387–403.

Huang, X., Atwood, M., Hartshorn, G., et al. (1999). The A peptide of Alzheimer's disease directly produces hydrogen peroxide through metal ion reduction. *Biochemistry* 38(24), 7609–7616.

Lynch, T., Cherny, R. A. and Bush, A. I. (2000). Oxidative processes in Alzheimer's disease: the role of abeta-metal interactions. *Experimental Gerontology* 35(4), 445–451.

Lonnerdal, B., Keen, C. L. and Hurley, L. S. (1981). Iron, copper, zinc, and manganese in milk. *Annual Review of Nutrition* 1, 149–174.

Religa, D., Strozzyk, R. A., Cherny, I., et al. (2006). Elevated cortical zinc in Alzheimer disease. *Neurology* 67(1), 69–75.

Metastasis

M K Demetrikopoulos

Institute for Biomedical Philosophy, Atlanta, GA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M K Demetrikopoulos, volume 2, pp 750–751, © 2000, Elsevier Inc.

Physical Stress and Tumor Progression in Animals
 Psychological Stress and Tumor Progression in Animals
 Psychological Stress and Tumor Progression in Humans
 Modifiers of Tumor Progression
 Conclusion

Glossary

<i>Interferon</i>	Glycoproteins produced by cells infected with a virus which inhibit viral growth and may be useful to treat tumors.
<i>Metastasis</i>	The spread of malignant tumor to a distant site.
<i>Natural killer (NK) cell</i>	A white blood cell that can kill certain types of cancer cells.
<i>Prospective</i>	An experimental design which is forward-looking.
<i>Retrospective</i>	A design which looks back and reviews aspects of a clinical situation.

Rotational stress

Physical stressor whereby subject's cage is gently, but continuously, moved in a circle on a horizontal plain.

Tumor progression

A measure of the cancer's metastasis, size, and/or weight.

While there is ample evidence which suggests that stress can have immunomodulatory effects, prediction of disease outcome based on generalizations from studies which examine *in vitro* immune function is difficult since the direction and degree of immunomodulation has been shown to be dependent on both the stimulus parameters employed and the immune variables under study. Therefore, research has been directed at examining disease-specific outcomes such as the effects of stress on cancer. This article will provide a brief overview of the effect of stress on cancer by examining how a variety of physical and psychological factors can effect tumor progression in both animal subjects and human patients. Additionally, the article will suggest possible mechanisms for this phenomena by presenting several known modifiers of tumor progression (i.e., hormones and neurochemicals) that are relevant in a discussion of the physiological consequences of stress. While it is well beyond the scope of this article to fully review this

topic, several illustrative examples will be presented which were selected in order to demonstrate the complexity of this subject or because of their clinical relevance.

Physical Stress and Tumor Progression in Animals

Physical stressors are generally more easily defined than psychological stressors since they involve the imposition of a specific external condition or stimulus. It is important to be aware, however, that the imposition of these conditions will likely initiate a cascade of both physiological and psychological reactions to the stimulus.

There are a variety of physical stressors that have been shown to effect tumor growth in animal subjects including exposure to cold, changes in light/dark cycles, dietary restrictions, rotation, restraint, forced swim, and surgery. The complexity of the physiological and psychological cascade that follows the imposition of a rather simplistic physical stressor, such as rotation, is illustrated by a study by Zorz et al. (2002: 368). Similar to their findings with restraint stress, they found a circannual rhythm in the expression of metastasis with rotational stress increasing metastasis volume in June and decreasing it in February compared to controls. The antitumor agent cyclophosphamide was able to block metastasis in the control mice but not in the mice in the rotational stress condition. Furthermore, the survival time of control mice in February was twice as long as in June. In all groups, survival time was inversely correlated with number of metastases, and number of CD3+ and CD4+ splenic T lymphocytes. The authors suggested that T-lymphocyte-mediated immune responses that vary seasonally were modulated by the rotational stress.

Psychological Stress and Tumor Progression in Animals

Psychological stress models provide a mechanism for studying the effects of psychological phenomena on tumor progression and serve as potential models for the effects of stress and psychiatric illness on cancer since the physical components of the stressors are either minimal or appropriately controlled for. A variety of psychological stressors have been shown to be important modifiers of tumor progression including handling, inescapable shock, housing, and social dominance. An important test of the psychological component of a physiological reaction is whether or not the response can be conditioned.

For example, Wu et al. (2001: 1) examined the effects of social isolation on liver metastasis in mice. Metastasis formation was quicker in the isolation stress condition such that isolated mice had five times the incidence of metastasis on day 7. The effect of isolation was evident across several measures such that metastasis was more likely in the isolated stress condition even with a smaller tumor challenge; survival time was shorter in the stress mice; and stressed mice did not respond as well to cisplatin chemotherapy. The fact that there was an effect on survival is particularly important since it would suggest that the effects are biologically meaningful to cancer development.

Psychological Stress and Tumor Progression in Humans

The relationship of psychological phenomena with cancer in human subjects has been reviewed in a series of papers presented in *Social Science Medicine* special issue *Cancer and the Mind*, 1985. Geyer (1993: 1545) found more severe life events and chronic difficulties in women with malignant breast cancer compared to women with benign lumps. Similarly, Hatch et al. (1990: 397) presented data which suggested that an increased rate of cancer living near the Three Mile Island nuclear accident was due to the stressfulness of the event and not due to the minimally increased radioactivity levels. While these studies suggest that a possible relation between psychological phenomena and tumor initiation may occur, the results are difficult to interpret given the fact that the data are mostly retrospective in nature. A prospective intervention by Spiegel et al. (1989: 888) examined the effect of a psychosocial intervention on survival time of women with metastatic breast cancer and found that survival time was essentially doubled in the intervention group compared to the control group. While this study did not address the effects of psychological phenomena on the early stages of cancer development, it demonstrated that psychosocial factors can modify the progression of cancer in patients and effect the survival time of patients with advanced disease. A more recent study by Butow et al. (2000: 469) demonstrated similar findings in that they showed that patients, who reported that they had minimized the impact of cancer in their lives, survived metastatic breast cancer longer.

Modifiers of Tumor Progression

A variety of factors are important in stress-induced tumor progression including hormonal modulation, neurochemical modulation, and the use of psychiatric

pharmaceuticals. However, determining the influence of any particular variable is difficult due to the interrelationships of these systems. Additionally, many of these compounds are known to have diurnal, monthly, and yearly cycles which may interact in various ways with their stress-induced patterns of expression thus contributing to the difficulty of elucidating their individual contribution to mediating stress-induced changes in tumor progression.

In spite of the complexity just suggested, it has been demonstrated that a variety of hormonal factors contribute to stress-induced tumor progression including corticosterone, melatonin, estrogen, and prolactin among others. For example, the effects of estrogen have been suggested to be important for some forms of breast cancer. Experimental evidence by several investigators have begun to explore the underlying mechanism of this phenomenon including examination of the effects of tumor exposure during different phases of the menstrual cycle.

In addition to the contribution of hormonal factors on tumor progression, a variety of neurochemical factors act as modifiers of tumor progression including opiates and catecholamines. This line of research is especially clinically relevant since pain management is an important consideration in cancer patients. Gaining an understanding of the central nervous system pathways which can effect stress-induced changes in tumor progression will likely lead to improvements in pain management that do not contribute to tumor development. For example, a study by Gaspani et al. (2002: 18) demonstrated that the analgesic drug tramadol blocked the increased metastatic effect of surgery stress in rats. They suggested this was mediated by natural killer (NK) cells since tramadol increased NK cell activity in nonoperated subjects. This is in marked contrast to the typical suppression in NK activity that has been repeatedly demonstrated with morphine and other opiates used for patients undergoing surgery for cancer.

Furthermore, since the effects of stressors on tumor progression are likely due to the activation of psychological modifiers, it may be possible to intervene by the use of anxiolytic or antidepressant agents. In fact, stress models often demonstrate an amelioration of the detrimental effects of short-term stressors by anxiolytic agents and an amelioration of the effects of long-term or uncontrollable stressors by the use of antidepressant agents.

Hormonal and neurochemical agents may effect tumor progression directly by causing changes in tumor growth and also indirectly by modification of immune-related functioning. In fact, the majority of hormonal and neurochemical changes that occur following stress have known consequences on immune

cells. For example, the effects of opiates on tumor progression have been shown to be related to changes in NK cell functioning. In addition to this, stress effects on tumor progression may be due to changes in NK number as well as due to changes in T and B cell number and/or functioning. Additionally, Wu et al. (2000: 1) have demonstrated that factors such as vascular endothelial growth factor, hepatocyte growth factor, and tumor necrosis factor α may be involved in metastasis by mediating the overexpression of normal endothelial cell functions such as migration, invasion, and tube formation.

Conclusion

The relation of stress to cancer has been suggested since at least the time of Galen and there have been many experimental studies which have attempted to elucidate this phenomenon since then. In testimony to the complexity of the subject, two earlier reviews came to vastly different conclusions as to the role of stress in cancer. In the first review, Sklar and Anisman (1981: 369) present a hypothesis that seems to be well supported by recent findings such that the chronicity of the stressor, the prior experience of the subject, controllability, and social context seem to be the major determining factors in predicting the direction of tumor progression. Additionally, the exact timing of an acute stressor in relation to the presentation of the tumor can dramatically effect the experimental outcome. They suggest that these factors may be producing enhanced tumor development by depleting brain catecholamines and increasing acetylcholine. In the second review, Justice (1985: 108) focused on the types of tumors studied and suggested that tumors be differentiated into viral and nonviral tumors; and the stress response could be thought of as having stages including a rebound or recovery stage following the presentation of an acute stressor. Justice hypothesized that viral tumors were enhanced during stress, were diminished during rebound or recovery to stress, and were the only type of tumors mediated through the immune system. Conversely, nonviral tumors were regulated outside of the immune system and were diminished during stress and enhanced during the rebound to stress. While there continues to be some studies which support this hypothesis, there are many studies that are contradictory. Yet, his review is an important reminder that it is necessary to understand the nature of the tumor under study when designing and interpreting experimental studies. Another review by Demetrikopoulos et al. (1999: 417) presents a more thorough discussion of the relevance of these older reviews to the more recent findings and presents an overview of the current

literature. A review by Ben-Eliyahu (2003: S27) examines a subset of this literature in detail and examines the effects of surgical stress on metastasis. This review explores the effects of cell-mediated immunity and explores possible therapeutic approaches that could be employed to reduce the effects of the surgical stress.

Overall, both physical and psychological stressors have been shown to effect tumor progression with most enhancing tumor growth and/or development. The mediators of this phenomenon are many and include a variety of both hormonal and neurochemical factors that are most likely utilizing central nervous system pathways that are important in anxiety and depression. Furthermore, the mechanisms underlying the effects of hormonal and neurochemical mediators on cancer include direct effects on tumors as well as modification of immunological functioning.

Further Reading

- Ader, R., Felten, D. L. and Cohen, N. (eds.) (2001). *Psychoneuroimmunology* (3rd edn.). New York: Academic Press.
- Ben-Eliyahu, S. (2003). The promotion of tumor metastasis by surgery and stress: immunological basis and implications for psychoneuroimmunology. *Brain Behavior and Immunity* 17, S27–S36.
- Butow, P. N., Coates, A. S. and Dunn, S. M. (2000). Psychosocial predictors of survival: metastatic breast cancer. *Annals of Oncology* 11, 469–474.
- Demetrikopoulos, M. K., Weiss, J. M. and Goldfarb, R. H. (1999). Environmental factors: stress and disease. In: McEwen, B. S. (ed.) *Handbook of Physiology: Coping with the Environment*, pp. 417–512. Oxford: Oxford University Press.
- Gaspani, L., Bianchi, M., Limiroli, E., et al. (2002). The analgesic drug tramadol prevents the effect of surgery on natural killer cell activity and metastatic colonization in rats. *Journal of Neuroimmunology* 129, 18–24.
- Geyer, S. (1993). Life events, chronic difficulties and vulnerability factors preceding breast cancer. *Social Science and Medicine* 37, 1545–1555.
- Hatch, M. C., Beyea, J., Nieves, J. W., et al. (1990). Cancer near the Three Mile Island nuclear plant: radiation emissions. *American Journal of Epidemiology* 132, 397–412.
- Justice, A. (1985). Review of the effects of stress on cancer in laboratory animals: importance of time of stress application and type of tumor. *Psychological Bulletin* 98, 108–138.
- Sklar, L. S. and Anisman, H. (1981). Stress and cancer. *Psychological Bulletin* 89, 369–406.
- Speigel, D., Bloom, J. R., Kraemer, H. C., et al. (1989). Effect of psychosocial treatment on survival of patients with metastatic breast cancer. *Lancet* 2, 888–891.
- Vallejo, R., Hord, E. D., Barna, S. A., et al. (2003). Perioperative immunosuppression in cancer patients. *Journal of Environmental Pathology, Toxicology and Oncology* 22, 139–146.
- Wu, W., Murata, J., Hayashi, K., et al. (2001). Social isolation stress impairs the resistance of mice to experimental liver metastasis of murine colon 26-L5 carcinoma cells. *Biological and Pharmaceutical Bulletin* 24, 772–776.
- Wu, W., Murata, J., Murakami, K., et al. (2000). Social isolation stress augments angiogenesis induced by colon 26-L5 carcinoma cells in mice. *Clinical and Experimental Metastasis* 18, 1–10.
- Zorzet, S., Perissin, L., Rapozzi, V., et al. (2002). Seasonal dependency of the effects of rotational stress and cyclophosphamide in mice bearing Lewis lung carcinoma. *Brain Behavior and Immunity* 16, 368–382.
- Social Science and Medicine: Cancer and the Mind*. (1985). 20. [A series of reviews].

Metirapone: Basic and Clinical Studies

E A Young

University of Michigan, Ann Arbor, MI, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by E A Young, volume 2, pp 752–756, © 2000, Elsevier Inc.

Clinical Uses

Clinical Research Studies

Basic Science Studies

Glossary

<i>11-Deoxycortisol</i>	The immediate steroid precursor for cortisol.
<i>11-β-Hydroxylase</i>	The enzyme responsible for the conversion of 11-deoxycortisol to cortisol.
<i>Cushing's disease/syndrome</i>	An endocrine disease due to excessive cortisol secretion. The syndrome may be due to a pituitary adenoma secreting intact adrenocorticotrophic hormone as well as the larger precursor forms (Cushing's disease; pituitary-dependent Cushing's) or to an adrenal adenoma secreting cortisol (pituitary-independent Cushing's syndrome).
<i>Depression</i>	Major depression as defined by DSM-IV, consisting of a collection of symptoms of at least two weeks duration leading to impairment in day to day functioning.
<i>Hypercortisolemia</i>	Increased plasma cortisol, which occurs in Cushing's patients and in a subset of patients with depression.
<i>Ovine corticotropin releasing hormone (oCRH)</i>	The protein sequence of the hypothalamic CRH found in sheep, in which the hormone was initially isolated.
<i>Panhypopituitarism</i>	A disease characterized by the failure of multiple endocrine systems at the pituitary level, usually secondary to the injury of the pituitary.

Metirapone is an 11-β-hydroxylase enzyme inhibitor that consequently blocks the last step in the synthesis of cortisol in the adrenal, which results in the secretion of 11-deoxycortisol, an inactive adrenal steroid (Figure 1). By blocking cortisol synthesis, metirapone leads to the inhibition of glucocorticoid negative feedback, and this opens up the closed feedback loop of the hypothalamic-pituitary-adrenal (HPA) axis, leading to rebound adrenocorticotrophic hormone

(ACTH) secretion. The presence of an open feedback loop allows the researcher to investigate the hypothalamic and pituitary components of the axis under controlled conditions. Metirapone is used primarily in human studies, whereas adrenalectomy is generally used in studies of experimental animals to accomplish the same goals. However, a few studies have investigated the effects of metirapone in rodents.

Clinical Uses

In clinical endocrinology, metirapone is used primarily in the diagnosis of Cushing's syndrome to distinguish excessive cortisol secretion secondary to excessive ACTH from an adrenal cause for cortisol hypersecretion in the absence of ACTH hypersecretion. In cases of a pituitary adenoma, the ACTH response is augmented tremendously with metirapone blockade of cortisol synthesis, whereas hypercortisolemia from an adrenal adenoma is accompanied by a lesser ACTH response. With the advent of newer, more sensitive ACTH assays, as well as magnetic resonance imaging technology, this test may be less critical to the diagnosis of Cushing's syndrome than previously thought. In addition, metirapone is used in the diagnosis of panhypopituitarism to determine if the patient is capable of responding to a stimulus to the HPA axis. In the past, metirapone has been used as a treatment for hypercortisolemia in Cushing's disease, but the blockade of cortisol is too short term and the side effects limit its usefulness. In addition, one open-label study suggested that it was helpful in patients with depression. A recent placebo-controlled study suggested its potential usefulness in the treatment of refractory cases of major depression; these patients did not necessarily manifest baseline hypercortisolemia. Currently, ketoconazole is the preferred glucocorticoid synthesis inhibitor for the treatment of hypercortisolemia.

Clinical Research Studies

In addition to its standard clinical use, metirapone is a useful research tool. In normal subjects given metirapone between 8 a.m. and 2 p.m., brisk rebound ACTH secretion is observed 2 h following metirapone administration, at which point cortisol levels have fallen to 3–4 μg dl⁻¹ (Figure 2). In normal subjects, the administration of metirapone in the late afternoon and evening (between 4 and 10 p.m.) does not lead to rebound ACTH secretion. This is an indication that hypothalamic drive (endogenous

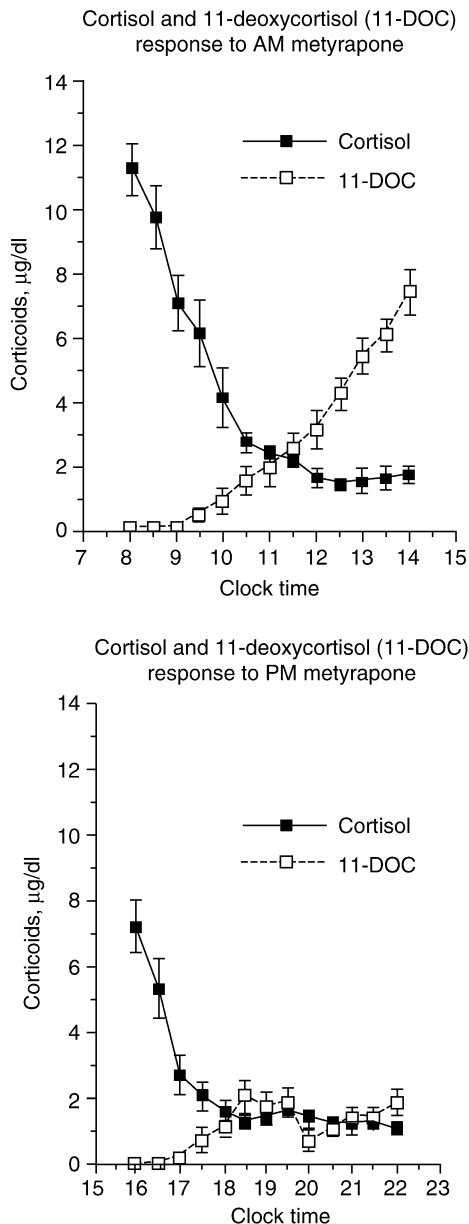


Figure 1 Administration of metyrapone leads to the blockade of cortisol production and secretion of 11-deoxycortisol, shown for both morning and evening administration in normal subjects.

corticotropin releasing hormone, CRH) is inactive during this time. Alternatively, it may be that the low levels of cortisol still present under metyrapone blockade (between 1 and 2 $\mu\text{g dl}^{-1}$) are able to exert the necessary inhibitory effect through high-affinity mineralocorticoid receptors. Using this same approach in depressed patients, we were able to demonstrate increased hypothalamic drive in the evening in depressed patients, whereas their ACTH response to metyrapone in the morning is relatively the same as controls. These studies suggest that depressed patients demonstrate increased brain CRH in the

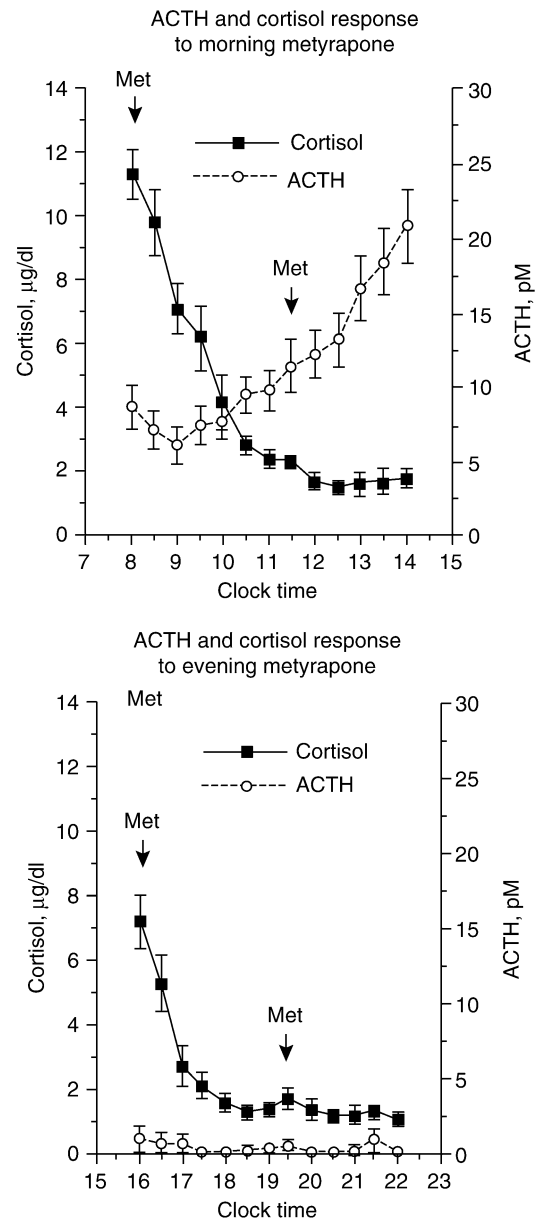


Figure 2 ACTH and cortisol response to the administration of metyrapone in the morning and evening in normal subjects. Following the blockade of cortisol production, rebound ACTH secretion is observed in the morning but not in the evening.

evening. A recent report of ours also found a link between excessive response to metyrapone and failure to respond to a 6-week course of fluoxetine. In addition to studies of major depression, metyrapone has been used as a probe of HPA axis function in anxiety disorders, in which the data are mixed. One study of posttraumatic stress disorder showed increased ACTH secretion, but two other studies showed a normal ACTH response. One study of panic disorder also demonstrated a normal response.

Metyrapone has been used by a number of investigators to block glucocorticoid negative feedback in

conjunction with CRH challenge studies. Metyrapone is particularly useful in reducing the effects of high baseline cortisol on the response to CRH and thus evaluating the true corticotroph response. The original studies by de Bold and colleagues demonstrated an increased ACTH response to oCRH in normal subjects given metyrapone compared to the same subjects given oCRH without metyrapone pretreatment. De Bold and colleagues were further able to characterize the effects of cortisol infusion following metyrapone plus oCRH challenge, thus demonstrating that the rise in cortisol induced by oCRH plays a significant role in terminating the ACTH response to oCRH. From these studies, it appears that both baseline cortisol and oCRH-induced increases in cortisol affect the ACTH response to oCRH.

Because patients with major depression demonstrate increased baseline cortisol, their ACTH response to oCRH is decreased. This has been observed in all studies. A number of investigators have examined the response to oCRH challenge under metyrapone in depressed patients. Studies by von Bardeleben and colleagues compared the response to oCRH under metyrapone in depressed patients to that of normal and depressed subjects given oCRH alone and concluded that elevated baseline cortisol was responsible for the diminished response to oCRH in depressed patients. However, these studies did not include normal subjects given oCRH plus metyrapone. Subsequent studies by Young and coworkers and Lisansky and coworkers using controls treated with metyrapone plus CRH confirmed a normal-to-augmented response to oCRH, thus arguing against CRH receptor downregulation in depression.

Metyrapone has also been used to evaluate the kinetics of ACTH inhibition by cortisol following cortisol infusion. Using metyrapone to augment the ACTH response, investigators have infused cortisol to study the fast feedback regulation of ACTH, although metyrapone pretreatment is not necessary to demonstrate fast feedback in humans.

Basic Science Studies

As stated previously, the majority of basic studies used adrenalectomy as a stimulus to ACTH secretion. However, in attempts to reduce the surgical stress involved in adrenalectomy, a few investigators have used metyrapone as a pharmacological adrenalectomy. These studies confirmed that metyrapone administration leads to an increase in both CRH and arginine vasopressin (AVP) in the hypophysial portal blood. Following treatment with metyrapone, a short-term rise was observed in plasma ACTH and β -endorphins

in rats, with a return to normal levels within 4 h. Plasma levels remained low at 24 h. However, by 72 h, ACTH secretion was again elevated. This same pattern was observed to be independent of whether metyrapone was administered during the peak of circadian activation (the evening) or the trough of activation (the morning). The measurement of pro-opiomelanocortin (the ACTH precursor) mRNA in the anterior pituitary and CRH mRNA in the paraventricular nucleus of the hypothalamus also demonstrated at least a 24-h lag time in the effects of metyrapone on steady-state mRNA levels. This lag in the plasma secretory response to the removal of negative feedback in rodents was also observed by Jacobson and colleagues using a different model. This finding contrasts to studies in humans (described earlier) in which rebound ACTH secretion was observed relatively promptly following the removal of glucocorticoid negative feedback during the peak of circadian activation.

See Also the Following Articles

Adrenocorticotrophic Hormone (ACTH); Corticotropin Releasing Factor (CRF); Cushing's Syndrome, Medical Aspects; Cushing's Syndrome, Neuropsychiatric Aspects.

Further Reading

- de Bold, C. R., Jackson, R. V., Kamilaris, T. C., et al. (1989). Effects of ovine corticotropin-releasing hormone on adrenocorticotropin secretion in the absence of glucocorticoid feedback inhibition in man. *Journal of Clinical Endocrinology and Metabolism* 68, 431–437.
- Jacobson, L., Akana, S. F., Cascio, C. S., et al. (1989). The adrenocortical system responds slowly to removal of corticosterone in the absence of concurrent stress. *Endocrinology* 124, 2144–2152.
- Jahn, H., Schick, M., Kiefer, F., et al. (2004). Metyrapone as additive treatment in major depression: a double-blind and placebo-controlled trial. *Archives of General Psychiatry* 61, 1235–1244.
- Kellner, M., Otte, C., Yassouridis, A., et al. (2004). Overnight metyrapone and combined dexamethasone/metyrapone tests in post-traumatic stress disorder: preliminary findings. *European Neuropsychopharmacology* 14, 337–339.
- Kellner, M., Schick, M., Yassouridis, A., et al. (2004). Metyrapone tests in patients with panic disorder. *Biological Psychiatry* 56, 898–900.
- Lisansky, J., Peake, G. T., Strassman, R. J., et al. (1989). Augmented pituitary corticotropin response to a threshold dose of human corticotropin releasing hormone in depressive pretreated with metyrapone. *Archives of General Psychiatry* 46, 641–649.

- Plotsky, P. L. M. and Sawchenko, P. E. (1987). Hypophyseal-portal plasma levels, median eminence content, and immunohistochemical staining of corticotropin-releasing factor, arginine vasopressin, and oxytocin after pharmacological adrenalectomy. *Endocrinology* **120**, 1361–1369.
- von Bardeleben, U., Stalla, G. L. K., Mueller, O. A., et al. (1988). Blunting of ACTH response to CRH in depressed patients is avoided by metyrapone pretreatment. *Biological Psychiatry* **24**, 782–786.
- Young, E. A., Altemus, M., Lopez, J. F., et al. (2004). HPA axis activation in major depression and response to fluoxetine: a pilot study. *Psychoneuroendocrinology* **29**, 1198–1204.
- Young, E. A., Haskett, R. F., Grunhaus, L., et al. (1994). Increased circadian activation of the hypothalamic pituitary adrenal axis in depressed patients in the evening. *Archives of General Psychiatry* **51**, 701–707.
- Young, E. A., Lopez, J. F., Murphy-Weinberg, V., et al. (1997). Normal pituitary response to metyrapone in the morning in depressed patients: implications for circadian regulation of CRH secretion. *Biological Psychiatry* **41**, 1149–1155.

MHC (Major Histocompatibility Complex) See: Lymphocytes.

Migraine

N M Ramadan

Rosalind Franklin University of Medicine and Science,
North Chicago, IL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by N M Ramadan, volume 2, pp 757–770, © 2000, Elsevier Inc.

Introduction
Epidemiology
Neural Basis of Head Pain
Genetics
Migraine Mechanisms
Classification of Headache Disorders
Clinical Presentations and Differential Diagnosis of
Migraine
Management
Summary

Glossary

Cephalad Toward the head.
Channelopathy A condition resulting from many scenarios (membrane ion channels that open when they should not, channels that do not open very well or at all, channels that stay open too long, misplaced channels, and a lack of channels or too many channels) that potentially can be deleterious.

Comorbidity

The association between migraine and another condition that is unlikely to be random or causal.

Cortical spreading depression

A local neuronal phenomenon that is elicited by chemical, physical, or electrical stimuli; characterized by a short burst of neuronal electrical activity and followed by a short-lived (minutes) marching wave of slow neuronal suppression.

Incidence

The proportion of new cases of an illness in a defined population over a set time period.

Migraine

A primary headache disorder manifesting as episodic, usually unilateral severe head pain with nausea, vomiting, and light and noise sensitivity.

Neurogenic inflammation

A sterile inflammatory response that is caused by an injurious stimulus of peripheral neurons and that leads to the release of substances (e.g., neuropeptides) that alter vascular permeability.

Phonophobia

An unpleasant feeling when exposed to noise.

Photophobia

An unpleasant feeling when exposed to light.

Prevalence

The proportion of a given population with a disease over a defined period of time. Lifetime prevalence of migraine is the proportion of individuals in the general population who have ever had migraine during their life. Prevalence is the

	average disease incidence multiplied by its duration.
<i>Trepanation</i>	An old surgical procedure in which a hole is drilled or scraped into the skull, leaving the membrane around the brain intact.

Introduction

Migraine is a painful and disabling headache disorder that has plagued humanity for thousands of years. Julius Caesar, England's Queen Mary, Thomas Jefferson, Claude Monet, Vincent van Gogh, Georges Seurat (visual disturbances of migraine aura, the Seurat effect), Alexander Graham Bell, Friedrich Nietzsche, Peter Tchaikovsky, and Sigmund Freud all suffered from migraine. Over the centuries, migraine has been transformed from a disease that is caused by evil spirits to a vascular and then a neurovascular disorder with well-defined epidemiology, pathology, pathophysiology, pharmacology, therapeutics, and genetics. Several thousand years BC, trepanation was a common practice used to relieve migraine by releasing demons from the head. Egyptians bound a clay crocodile, holding grain in its mouth, to the head of headache patients, using a strip of linen that bore the Gods' names. Galen considered migraine a visceral pain disorder when he said, "How constantly do we see the head attacked with pain when yellow bile is contained in the stomach: as also the pain forthwith ceasing when the bile has been vomited." (Koehn, 1826) Tenth century Arabian physicians used garlic and hot iron on temple incisions (two-site venesection) to relieve head pain.

More recently, Sir Thomas Willis believed that migraine was caused by vasodilatation, arguing that headache symptoms were related to slowly ascending spasms beginning at the peripheral ends of the nerves. Similarly, Erasmus Darwin (Charles Darwin's grandfather) believed in vasodilation-induced migraine and proposed to treat it with centrifugation in order to force the blood from the head to the feet. Edward Living commented in "On Megrim, Sick-headache, and Some Allied Disorders" that the migraine, similar to epilepsy, is a malady of the nervous system caused by nerve-storm, which may be the first account of cortical spreading depression (CSD) as an initiator of migraine. Sir William Gowers proposed dividing the treatment of headache into prophylactic and episodic. He advocated continuous treatment with drugs to render attacks less frequent and treatment of the attacks themselves. In the late 1800s, the potent neurotoxin ergot, which comes from a fungus that grows on rye, was used to treat unilateral headache. Later in the early 1900s, the active chemical ergotamine was

discovered and became the preferred treatment for migraine.

The modern understanding of migraine started with the work of Harold Wolff, a migraine sufferer himself, who performed elegant clinical experiments that supported the vascular theory of headache. Wolff and Ray published in 1940 a detailed map of pain-sensitive structures in the brain, which provided the basis for the link between arterial pathology and referred head pain. Contemporaneously, Leao discovered CSD in animals, which later was equated with human visual aura. Today, this phenomenon can be imaged and measured, and it is forming the backbone for the development of several classes of potential antimigraine drugs. To this end, it is important to mention seminal research from Boston (Cutrer and Sanchez del Rio), Detroit (Welch and Cao), Los Angeles (Woods), and elsewhere, which provided us with a window for analyzing and quantifying CSD in humans.

In the 1960s and beyond, Sicuteri, Lance, and Anthony stressed the role of serotonin in migraine. Their work was later amplified by Saxena, Humphrey, Martin, and others, who discovered the triptans. Also, the Copenhagen group led by Olesen drew attention to a mismatch between cerebral blood flow (CBF) changes and symptoms of migraine, which challenged the vasoconstriction-vasodilation theory. The same group evaluated the role of nitric oxide (NO) in inducing headache, and now NO synthase (NOS) inhibition is becoming a prime target in averting migraine.

The understanding of migraine mechanisms and therapeutics evolved over the last 2 decades with experiments that Moskowitz and fellow researchers conducted. The activation of the trigeminal system in animals was shown to induce plasma protein extravasation and neurogenic inflammation, and many antimigraine therapies blocked this response. Similarly, the initial work of Edvinsson and Goadsby was amplified in the last 10–15 years, culminating in advancing calcitonin gene-related peptide (CGRP) antagonists for migraine. Also, important work from Detroit, Michigan, and Bologna, Italy, supported the theory that the brain-energy metabolism in migraine sufferers is defective and that magnesium deficiency may contribute to the disease. These observations opened the field to testing magnesium, co-enzyme Q10 (CoQ10), and riboflavin (B2) for migraine prevention. Apropos, John Graham modernized the preventive treatment of migraine with the introduction of methysergide, Oscar Reinmuth's group drew attention to the role of propranolol in migraine, and recent studies from Welch and colleagues and Ferrari and coworkers suggested that migraine may be a progressive disorder that may warrant early and aggressive preventive therapy.

Finally, the potential role of peripheral and central sensitization in migraine, which Burstein and colleagues are unfolding, is opening new windows into migraine therapeutic strategies.

Migraine is not an orphan in the era of genomic discoveries. French and Dutch researchers led the effort demonstrating that some forms of migraine result from defective calcium trafficking across its channels and provided a new avenue for exploring migraine mechanisms and therapies. A knockin model of abnormal mutations now provides a model of familial hemiplegic migraine and susceptibility to CSD. Also, the discovery of mutations in the Na-K adenosine triphosphatase (ATPase) gene is steering us to new pathophysiological and pharmacological directions. Indeed, excessive synaptic glutamate (Glu) is a by-product of these mutations, which predisposes the migraine brain to hyperexcitability, sensitization, and/or neuronal cell death.

Epidemiology

Prevalence estimates of headaches vary significantly among studies. For example, the prevalence of headache disorders in Europe and the United States ranges from 13% (1-year prevalence in people with serious headaches) to 93% (lifetime prevalence in the United Kingdom, based on a telephone interview). Several factors contribute to the observed variability, including the time periods used for estimating prevalence (e.g., 1-year vs. lifetime), targeted age range and other sociodemographic factors, methods of ascertaining headache diagnoses, and cooccurrence of different headache types. Regardless of the study methods, the majority of population-based studies of headache disorders in adults indicate a female preponderance, which is not unexpected because most chronic pain disorders with a remitting and relapsing course (e.g., fibromyalgia and irritable bowel syndrome) are more prevalent in women.

Rasmussen et al. conducted a large 1991 epidemiological study (1000 people, 25–64 years old) on the prevalence of headache disorders in Copenhagen County, Denmark. Tension-type headache was the most common, followed by fasting headache, migraine, and headache attributed to nose or sinus disease (Table 1).

Epidemiology of Migraine

Few longitudinal studies have investigated the incidence of migraine. A 12-year follow-up of a Dutch general population sample ($n = 549$) determined that the incidence of migraine was 8.1 per 1000 person-years (male : female = 1 : 6). Data from a 2005 prevalence

Table 1 Lifetime prevalence of headache disorders in Denmark^a

Type	Prevalence (%)
Primary headache disorders	
Tension-type headache	78.0
Migraine	16.0
Secondary headache disorders	
Fasting	19.0
Nose and sinus disease	15.0
Head trauma	4.0
Nonvascular intracranial disease (e.g., brain tumor)	0.5

^aFrom Rasmussen, B. K., Jensen, R., Schroll, M., et al. (1991), Epidemiology of headache in a general population – a prevalence study, *Journal of Clinical Epidemiology* **44**, 1147–1157.

study in people 12–29 years old revealed slightly higher numbers – the incidence of migraine with aura was 14.1 per 1000 person-years and that of migraine without aura was 18.9 per 1000 person-years.

The incidence of migraine is two to three times higher in females than in males, and the peak age of onset is 5–9 years in males and 12–13 years in females for migraine with aura. For migraine without aura, the figures are 10–11 years for males and 14–17 years for females. The incidence of migraine may be increasing, particularly in women 10–49 years old.

In contrast to the paucity of studies of migraine incidence, several reports from various countries across the globe addressed migraine prevalence and established that it is a common disorder with particular predilection to Whites and women. Prior to 1988, population estimates of migraine prevalence varied from 3 to 35%. The application of the International Headache Society (IHS) 2004 diagnostic criteria led to more robust and consistent data on migraine epidemiology and allowed a better assessment of demographic influences such as race, socioeconomic status, gender, and geographical location. Indeed, several studies have now provided gender-dependent estimates of the 1-year prevalence of migraine: approximately 3–22% in women and 1–16% in men. The prevalence of migraine is highest in North America and lowest in Africa (Figure 1).

Population studies have confirmed that migraine is more prevalent in females than in males older than 12 years (female : male = 2–3 : 1), regardless of race or geographical location. Migraine in children is almost as prevalent in boys as it is in girls.

The prevalence of migraine follows an inverted-U curve with advancing age (Figure 2). This phenomenon is significantly more pronounced in females. The peak prevalence of migraine is in the fourth and fifth decades; it declines substantially thereafter.

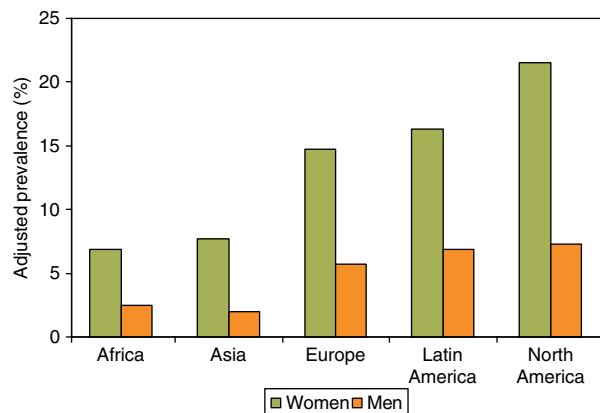


Figure 1 Prevalence of migraine across continents. From Lipton, R. B. and Bigal, M. E. (2005), *Migraine: epidemiology, impact and risk factor for progression*, *Headache* 45(supplement 1), S3–S13.

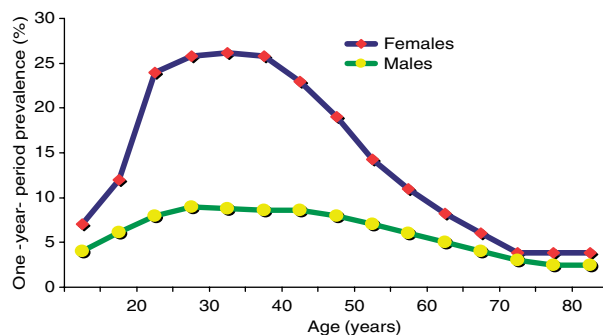


Figure 2 Prevalence of migraine by age. From Stewart, W. F., Lipton, R. B., Celentano, D. D., et al. (1992), *Prevalence of migraine headache in the United States: relation to age, income, race, and other sociodemographic factors*, *Journal of the American Medical Association* 267, 64–69.

Migraine varies with race. Whites are more commonly affected (prevalence = 20.4% in women and 8.6% in males) than African Americans (16.2% in women and 7.2% in men) or Asian Americans (9.2% in women and 4.2% in men). Contrary to a widely held belief, migraine is more common in low-socio-economic groups and in less-educated individuals. The use of care data indicate, however, that people in higher social strata and better-educated migraine sufferers are more likely to seek medical advice.

Neural Basis of Head Pain

Triggers of head pain in different headache disorders may vary, but the substrates for pain generation, stimulus conduction through the peripheral and central nociceptive pathways, recognition and perception of the experience, and memory and reaction to pain are largely similar whether the type of headache is primary (e.g., migraine) or secondary (e.g., trauma).

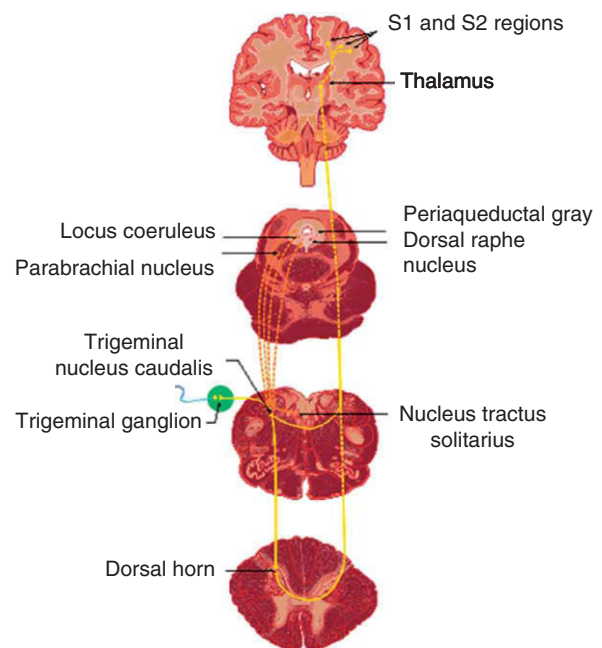


Figure 3 Nociceptive and modulatory pathways. Pathways in yellow represent afferent projections subserving pain; pathways in orange represent descending, mainly inhibitory projections.

A discussion of trigger mechanisms is beyond the scope of this article, but a brief overview of the final common pathways in pain are described next.

Neuroanatomy

The pathways of head pain can be anatomically dissected into four distinct locales (**Figure 3**):

1. Sensing organs such as the skin and blood vessels.
2. Sensory-discriminative regions, which include the trigeminal nucleus caudalis (TNC), thalamus, and the primary (S1) and secondary (S2) sensory cortices.
3. Emotion, memory, and behavioral response networks. These include the insula and hippocampus, which encode the memory and recognition of pain; cingulate, which modulates the affective and motivational aspects of pain; and precentral cortex, cerebellum, and cingulate, which are involved in the reaction to the painful experience.
4. Modulatory, predominantly inhibitory regions such as the hypothalamus, basal forebrain, and various brain-stem structures.

Head pain is generated when pain-sensitive intracranial structures are stimulated. These include vascular structures such as the circle of Willis and proximal middle cerebral arteries, neuronal structures including the trigeminal and glossopharyngeal nerves, and the cranial meninges, particularly the perivascular components of the dura. Pain signals are transmitted

from these nociceptive structures cephalad, mostly via A δ and C fibers that contain neurotransmitters such as Glu, CGRP, substance P (SP), and neurokinin A (NKA). Middle meningeal artery nociceptive fibers travel through the first division of the trigeminal nerve (V1) to the ipsilateral trigeminal ganglion (TG), whereas fibers from the middle cranial fossa dura run in V3 to the ipsilateral TG. First-order neuron fibers in TG then project centrally onto secondary sensory neurons in the TNC, which in turn send their projections to the contralateral nucleus ventralis posteromedialis (VPM) of the thalamus. Projections of the nociceptive pathway from the trigeminal brain-stem nuclear complex include the trigeminal lemniscus, hypothalamic projections, pontine parabrachial nucleus (PBN), reticular formation, and the nucleus of the tractus solitarius (NTS). PBN feeds back into the trigeminal brain-stem nuclear complex and is involved in nociceptive control. Central processing of incoming pain signals beyond the thalamus include cortical areas such as the cingulate, insular, S1, and S2, as well as the cingulate gyrus and insula. Cortical neurons that are activated during head and peripheral pain include localizing and discriminating pain neuronal pools, which receive afferents from VPM, and those involved in the effective components of pain, which receive afferents from the medial thalamus.

Descending, predominantly inhibitory neuronal fibers project from the frontal lobe cortex to the hypothalamus and the periaqueductal gray (PAG) and from there descend to the rostral ventromedial medullary nuclear complex (RVM, the raphe magnus and adjacent reticular formation), which ultimately projects to the medullary and spinal dorsal horns.

Neurochemistry

The PAG–RVM system contains serotonin (5-HT) excitatory neurons that activate inhibitory interneurons of the substantia gelatinosa and lamina II of the spinal trigeminal nucleus. Electrical stimulation of either PAG or RVM or the exogenous injection of opioids, which activate the 5-HT system, results in inhibitory activity in the nociceptive neurons in the dorsal horn. Independently of PAG–RVM, a diffuse noxious inhibitory control (DNIC) system (not shown in [Figure 3](#)) directly reduces the firing of wide-dynamic range (WDR) neurons of the medullary dorsal horn cells following noxious stimulation from parts of the body that are remote from their receptor field.

The neurochemicals that are involved in pain derive from the vascular endothelium of nociceptive structures, which contains both vasoconstrictive (e.g., thromboxane, superoxide ions, and endothelin) and

vasodilator substances (e.g., NO and prostacyclin), and from the adventitia–media junction of intracranial vessels, which are rich in neuropeptide-positive terminals (e.g., CGRP and SP).

Neurophysiology

Nociceptive activation induces peripheral sensitization, which is predominantly mediated by enhanced Na channels opening. This process perpetuates discharges along the nociceptive terminals and into TG with an ongoing cycle of enhanced Na-channel opening, excitatory neurotransmitter release, and further neuronal activation. With time, peripheral sensitization, perhaps through excessive N-methyl-D-aspartate (NMDA) Glu receptor activation, leads to a central sensitizing process initially at the TNC level and subsequently at the thalamic level and beyond. The phenomenon of central sensitization manifests physiologically as spontaneous activity in otherwise quiescent neurons, reduced threshold to activation, and expansion of the neuronal receptor field following distal stimulation.

Peripheral and central sensitization could continue unless various protective neuromodulatory mechanisms become operant. Modulation is mediated peripherally through the presynaptic inhibition of neurotransmitter release or centrally through several possible mechanisms, including the inhibition of afferent input. At the peripheral level, serotonergic presynaptic receptors of the 5-HT_{1D} and 5-HT_{1F} types and some metabotropic Glu receptors participate in suppressing neurotransmitter release. Centrally, TNC neurons are inhibited by local inhibitory interneurons (e.g., GABAergic), by activation of presynaptic inhibitory receptors (5-HT_{1D}, 5-HT_{1B}, 5-HT_{1B}, and metabotropic), or by descending inhibitory fibers (e.g., noradrenergic, adenosinergic, glutamatergic, serotonergic, and GABAergic). The descending modulatory system itself is under the control of various hypothalamic–pituitary outputs, such as estrogen levels, and melatonin cycles, which have direct relevance to the roles of sleep and menstrual cycles in migraine and other painful disorders.

Genetics

The genetics of most headache disorders remain elusive. On the other hand, some major advances in the genetics of migraine have been published in the last decade or so.

There is mounting evidence that a subform of migraine (familial hemiplegic migraine, FHM) is due to a membrane ion channelopathy or ionopathy ([Table 2](#)). A net gain of function of the calcium

Table 2 Some genetic defects in familial hemiplegic migraine that might predispose to hyperexcitable neurons^a

<i>Chromosome</i>	<i>Gene mutations locale</i>	<i>Potential biological or physiological consequence</i>
19p13	CaCNA1A (Ca _v 2.1)	↑ Ca ²⁺ into presynaptic neuron, causing neurotransmitter release (e.g., glutamate)
1q23	α ₂ subunit of Na ⁺ /K ⁺ ATPase	Defective clearing of glutamate from synaptic cleft
2q24	Neuronal Na channel	Enhanced EPSP
5p13	EAAT1 glutamate transporter	Defective clearing of glutamate from synaptic cleft

^aCaCNA1A, α₁ subunit of calcium channel gene; ATPase, adenosine triphosphatase; Na, sodium; EPSP, excitatory postsynaptic synapse; EAAT1, excitatory amino acid transporter type 1.

channel leads to an excessive intracellular calcium influx into the presynaptic glutamatergic terminal, Glu release, and persistent postsynaptic activation that predisposes to cortical spreading depression (CSD). Indeed, a knockin mouse model of a CaCNA1A mutation that predisposes animals to CSD and transient hemiplegia has been developed. Alternatively, a defective Na pump reduces the efficiency of Glu transporters in clearing Glu, resulting in excessive Glu at the synaptic cleft and subsequently in continued postsynaptic excitation. Interestingly, a heterozygous mutation on the gene encoding for the Glu transporter excitatory amino acid transporter type 1 (EAAT1) was discovered in 2004 in a patient with seizures, migraine, and alternating hemiplegia. Last, hyperexcitability can be caused by a neuronal Na-channel mutation, reported in 2005.

Genetic defects in FHM might not be extrapolated to the more common forms of migraine with and without aura, especially with the absence of consistent evidence of genetic mutations or functional polymorphism at the known FHM sites, but nearby loci are implicated. Furthermore, a 2005 genomewide scan of people with common forms of migraine confirmed genetic susceptibility at sites previously reported with FHM. In addition, new susceptibility loci were reported (3q and 18p11), which deserve further exploration.

Migraine Mechanisms

The mechanisms of migraine pain and its associated neurological, gastrointestinal, and autonomic manifestations are multiple and complex. Attacks could manifest as an isolated symptom (e.g., migraine aura without headache) or a sign and symptom complex (e.g., migraine with aura), which could take on a range of phenotypic expressions (e.g., emotional lability as part of a premonitory symptom complex followed by pain or a full-blown multiphasic attack manifesting as depressive symptoms, followed by a visual aura followed by pain with nausea and photophobia and phonophobia, and terminating with social-withdrawal feelings and tiredness). Therefore, deciphering migraine mechanisms requires a

comprehensive understanding of the isolated manifestations and explanations for the event sequences.

Migraine is a neurobiological disorder of altered neuronal excitability and defective adaptive mechanisms to stressors of the system. The susceptibility to migraine is an interplay between genetic factors and environmental elements (e.g., stress and weather change). The migraine cascade can be conceptualized as a sequence of events occurring in a series, starting with the activation of hypothalamic pathways that trigger the premonitory symptoms and followed by cortical neuronal activation and CSD. Neuronal activation releases various substances such as hydrogen ions and NO, which in turn activate nociceptive terminals that, on the one hand, cause the release of substances such as CGRP and, on the other hand, transmit neuronal impulses into the medulla to the thalamus and up to the cortex for perception and reaction to the headache. CGRP, SP, and NKA participate in neurogenic vasodilation and inflammation, leading to peripheral sensitization and an abnormal response to previously innocuous stimuli (e.g., blood-vessel pulsations becoming painful). Often, peripheral sensitization can culminate in central sensitization with the activation of ancillary trigeminal nociceptive (e.g., allodynia in the distribution of the first division of the trigeminal nerve) and paratrigeminal pathways (e.g., photophobia and phonophobia).

The female preponderance of migraine sufferers provides some insight into the mechanisms of headache and pain. A discussion of the role of sex hormones and migraine mechanisms is quite complex and beyond the scope of this article, but a few highlights are noteworthy. Women detect trigeminal pain at a lower threshold than men, and their tolerance for the painful stimulus may be lower. Some researchers observed that these gender differences may be related to higher sensory-receptor density on the peripheral nerve terminals in women. Furthermore, estrogen receptors (ERs) are distributed on various neuronal pools that participate in migraine pain trafficking (e.g., TNC, hypothalamus, amygdala, PBN, and PAG). Lastly, ERs modulate neuropeptides in the trigeminal system that are involved in migraine (e.g., CGRP, SP, and NKA).

Table 3 Comorbidity of migraine and systemic and psychiatric disorders

System	Condition
Gastrointestinal	Irritable bowel syndrome
Neurological	Epilepsy
	Multiple sclerosis
Psychiatric	Affective disorders
	Anxiety disorders
	Neuroticism
	Sleep disorders
	Substance abuse
	Suicide
Pulmonary	Asthma
Vascular and cardiovascular	Hypertension
	Hypotension
	Ischemic stroke
	Myocardial infarction
Others	Allergies
	Raynaud's phenomenon

Migraine is comorbid with psychiatric conditions such as generalized anxiety disorders and depression (Table 3), and migraine attacks often are triggered by stress and lack of sleep. It has been suggested that bidirectional signal transmission through the trigemino-vascular system accounts for both the migraine symptoms and the induction of these symptoms by triggers such as stress. Also, migraine shares similar mechanisms with chronic pain when comorbid psychiatric conditions such as depression could result from functional alterations in the reward-aversion pathways.

Migraine patients have some personality traits such as neuroticism, which could lead to altered coping mechanisms. Indeed, one study suggested that a migraine sufferer's coping mechanisms include amplified physical symptoms, social litigation, and preoccupation with stress. Also, migraine patients are less calm, more irritable, do not relax as well as a nonmigraine healthy controls, and often respond to physical symptoms, including pain, with internal tension.

Classification of Headache Disorders

The development of a headache disorder classification that is intuitive, psychometrically robust, and easy-to-use enables a better understanding of conditions associated with headache, improves patients' management, and provides a standardized approach to headache research. Headache disorders can be classified broadly as primary (e.g., migraine), and secondary (e.g., posttraumatic). Indeed, the *International Classification for Headache Disorders* (2nd edn.;

ICHD-II) recognizes this dichotomy (Table 4). Primary headaches are classified largely on the basis of symptoms, whereas secondary headaches are grouped etiologically. Both acute and chronic forms of primary and secondary headache disorders are recognized (e.g., chronic migraine and chronic posttraumatic headache).

Clinical Presentations and Differential Diagnosis of Migraine

Migraine is characterized by episodes of head pain that is aggravated by movement, with associated nausea, photophobia, and phonophobia. Head pain in migraine is unilateral in approximately 60% of patients and throbbing in many but not all. Up to one-third of migraine patients experience transient neurological symptoms (i.e., aura) that herald or accompany the pain. Most migraine attacks last approximately 1 day (4–72 h in adults; they are of shorter duration in children), and their frequency is quite variable. The frequency of migraine is approximately once per month in population-based studies and three times per month in clinic-based cohorts. In the general population, 2–3% suffer chronic migraine, which is defined as attacks that recur more than 15 times per month. Last, the majority of migraine sufferers are disabled with their attacks.

A careful history and appropriate physical and neurological examination establish most causes of headache disorders. For example, the presence of associated symptoms (e.g., nausea and photophobia) helps to differentiate migraine from tension headache. Important elements of the history that help in establishing the diagnosis include (1) age of onset; (2) duration; (3) time to peak severity; (4) severity and headache-related disability; (5) frequency; (6) location; (7) preceding and accompanying symptoms, if any; (8) precipitating and aggravating factors; (9) circadian or diurnal patterns; and (10) relieving factors. It is also crucial to obtain a full headache medication history, including over-the-counter remedies.

The clinical presentations of migraine with and without aura are often characteristic (Table 5). On the other hand, the clinical presentations of secondary headache types mimicking migraine with aura are not unique (Table 6), with a few exceptions (e.g., benign or idiopathic intracranial hypotension). Also, the symptom complex of migraine and that of other primary headaches and secondary headaches overlap (Table 7). An algorithm that helps in the differentiation between primary and secondary headache disorders uses red flags (called SSNOOPs, an acronym based on the first letter of items in the

Table 4 Second International Classification for Headache Disorders (ICHD-II)^a

Category	Type
Primary headache disorders (categories 1–4)	Migraine, tension-type headache, cluster headache and other trigeminal cephalgias, other primary headache (e.g., primary stabbing headache, primary cough headache)
Secondary headache disorders (categories 5–12)	Posttraumatic headache; headache attributed to cranial or cervical vascular disorders; headache attributed to nonvascular intracranial disorder; headache associated with substances or their withdrawal; headache attributed to infection; headache attributed to disorder of homeostasis; headache or facial pain associated with disorder of cranium, neck, eyes, ears, nose sinuses, teeth, mouth, or other facial or cranial structures; headache attributed to psychiatric disorder
Cranial neuralgias and central causes of facial pain (category 13)	Trigeminal neuralgia; glossopharyngeal neuralgia; nervus intermedius neuralgia; superior laryngeal neuralgia; nasociliary neuralgia; supraorbital neuralgia; occipital neuralgia; neck-tongue syndrome; external compression headache; cold stimulus headache; constant pain caused by compression, irritation or distortion of cranial nerves or upper cervical roots by structural lesions; optic neuritis; ocular diabetic neuropathy; head and facial pain attributed to herpes zoster; Tolosa-Hunt syndrome; ophthalmoplegic migraine; anesthesia dolorosa; central poststroke pain; facial pain attributed to multiple sclerosis; persistent idiopathic facial pain; burning mouth syndrome
Other headache, cranial neuralgia, central or primary facial pain	Headache unspecified

^aFrom the Headache Classification Subcommittee, International Headache Society (2004), *Special issue on the international classification of headache disorders (2nd edn.)*, *Cephalalgia* **24**(supplement 1).

Table 5 Diagnostic criteria for migraine^a

Migraine without aura	A. At least five attacks fulfilling B–D B. Headache attacks lasting 4–72 h (untreated or unsuccessfully treated) C. Headache has at least two of the following characteristics: unilateral location, pulsating quality, moderate or severe pain intensity, aggravation by or causing avoidance of routine physical activity (e.g., walking or climbing stairs) D. During headache at least one of the following: nausea and/or vomiting, photophobia and phonophobia E. Not attributed to another disorder
Migraine with typical aura	A. At least two attacks fulfilling B–C B. Aura consisting of at least one of the following, but no motor weakness: fully reversible visual symptoms including positive features (e.g., flickering lights, spots, or lines) and/or negative features (i.e., loss of vision), fully reversible sensory symptoms including positive features (e.g., pins and needles) and/or negative features (i.e., numbness), fully reversible dysphasic speech disturbance C. At least two of the following: headache [1.21] fulfilling criteria B–D for migraine without aura; begins during the aura or follows aura within 60 min; headache [1.22] that does not fulfill criteria B–D for migraine without aura; begins during the aura or follows aura within 60 min; headache [1.23] does not occur during aura nor follows aura within 60 min D. Not attributed to another disorder

^aFrom the Headache Classification Subcommittee, International Headache Society (2004), *Special issue on the international classification of headache disorders (2nd edn.)*, *Cephalalgia* **24**(supplement 1).

following list) as the major branching point (**Figure 4**). SSNOOPs are:

- Systemic symptoms (fever or weight loss)
- Secondary risk factors: underlying disease (HIV or systemic cancer)
- Neurological symptoms or abnormal signs (confusion, impaired alertness or consciousness, neck stiffness, clumsiness, or weakness)
- Onset: sudden, abrupt, or split-second (first or worst)

- Older: new onset and progressive headache, especially in middle age (>50 years old)
- Previous headache history or headache progression: pattern change, first headache or different (change in attack frequency, severity, or new clinical features).

A variation of SSNOOP is SSNOOPP, which considers pain with local tenderness (e.g., jaw or temporal artery) as an additional red flag.

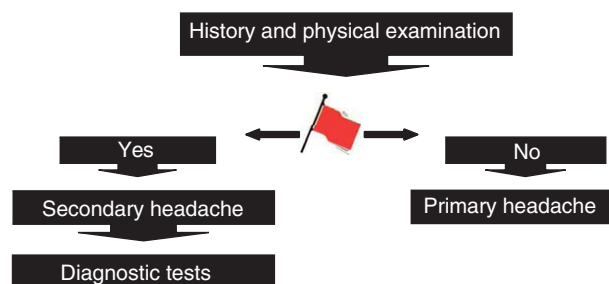
Table 6 Differential diagnosis of headache with transient neurological symptoms mimicking migraine with aura^a

Cerebrovascular disease	Dissection, ischemic stroke and transient ischemic attacks, hemorrhagic stroke, AVM, primary CNS vasculitis, cerebrovenous sinus/dural thrombosis, postpartum angiopathy
Drug-induced	Nitric oxide-releasing compounds (e.g., nitroglycerin, sildenafil), illicit drugs
Endocrine/metabolic disorders	Hypoglycemia
Inherited/genetic disorders	CADASIL, MELAS, OTC deficiency
Nonvascular intracranial disorders	Focal encephalitis and aseptic meningitis, primary or metastatic brain tumors, idiopathic intracranial hypertension, HANDL (PLP), systemic vasculitis (e.g., SLE and aPL syndrome), HIV
Traumatic disease	Posttraumatic headache, postwhiplash injury
Others	Ictal headaches, whiplash injury, cervicogenic headache, chiari malformation type I, glaucoma

^aaPL, antiphospholipid; AVM, arteriovenous malformation; CADASIL, cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy; CNS, central nervous system; HANDL, headache, neurologic deficits, and cerebrospinal fluid lymphocytosis; HIV, human immunodeficiency virus; MELAS, mitochondrial encephalopathy with lactic acidosis and stroke-like episodes; OTC, ornithine transcarbamylase; PLP, pseudomigraine with lymphocytic pleocytosis; SLE, systemic lupus erythematosus.

Table 7 Overlap in symptoms of migraine pain and secondary headaches

<i>Symptom</i>	<i>Some examples of headache conditions</i>
Unilateral headache	Trauma, cluster headache, trigeminal autonomic cephalgia, migraine, cerebrovascular disease, idiopathic jabbing headache, sinusitis, eye/jaw disease, cervical spine disease, giant cell arteritis
Throbbing pain	Migraine, cerebrovascular disease, hypertensive headaches
Exacerbation with movement	Migraine, headache attributed to conditions that increase intracranial pressure, posttraumatic headache, Chiari malformation type I, benign cough headaches
Photophobia	Migraine, cluster headache, meningitis, posttraumatic headache, subarachnoid hemorrhage
Nausea, vomiting	Headache attributed to conditions that increase intracranial pressure, meningitis, intracranial hemorrhage, trauma

**Figure 4** Approach to the diagnosis of headache. Red flag represents any of the SSNOOPs. Adapted from the Neurology Ambassador Program with permission.

General medical and surgical histories, social and family histories, medications intake, and review of systems may provide clues to the diagnosis. For example, transient visual obscuration and intracranial noises may point to idiopathic intracranial hypertension. An antecedent motor vehicle accident within 6–12 months of the headache onset may indicate posttraumatic headache or whiplash headache. General and neurological examinations are an integral component of the assessment of patients with headaches. Papilledema, carotid bruits, neck limitation of motion, second trigeminal nerve branch or occipital triggers, tender temporal artery, and focal

or lateralizing signs can all provide important clues to the diagnosis.

Once the characteristic symptoms of a migraine headache are identified and red flags are not elicited, imaging studies are likely to reveal a significant intracranial abnormality in only approximately 0.2% of patients. A magnetic resonance imaging (MRI) study is the preferred test when imaging is indicated.

Management

General Principles

General management strategies for migraine follow the TEEM[®] principles, which are:

- Trust and establish a strong rapport with the patient.
- Educate the patient about the disease/condition, treatment plan, and compliance. A major reason for failed therapy in headache disorders is noncompliance and/or unmatched or unrealistic expectations (e.g., expecting a cure for migraine).
- Empower the patient.
- Measure.

It is crucial to establish an accurate diagnosis based on a detailed history and physical examination and

aided in special circumstances by neuroinvestigative techniques such as an MRI. It is also very important to systematically assess and follow up on the disease severity and disease-related disability by encouraging patients to maintain headache diaries in order to identify headache patterns, triggers, associated symptoms, response to intervention, and so on. The elicitation during the history of migraine comorbid conditions (Table 3) provides either opportunities to simplify treatment regimens (e.g., valproate for migraine and bipolar disease) or limitations on the use of certain medications (e.g., propranolol for migraine sufferers with asthma).

Pharmacotherapy of Migraine

Migraine pharmacotherapy can be divided into (1) the treatment of acute pain, (2) the treatment of associated symptoms, (3) traditional preventive agents, and (4) preemptive approaches when attacks have predictable triggers (e.g., menstruation). Preventive therapy for migraine is indicated for patients with (1) frequent or disabling attacks, (2) a prolonged aura, or (3) a poor response or intolerance to acute therapy.

Current acute therapy (Table 8) for migraine includes simple analgesics (e.g., acetaminophen) and compound analgesics (e.g., combined acetaminophen, aspirin, and caffeine; acetaminophen with codeine),

nonsteroidal anti-inflammatory drugs (NSAIDs, e.g., ibuprofen), migraine-specific therapies such as ergots (ergotamine and dihydroergotamine, DHE) and triptans (serotonin type-1B and 1D, 5-HT_{1B/D}, agonist), and adjuvant therapies such as metoclopramide for gastrointestinal disturbances. Over-the-counter non-specific therapies are well suited for low-severity migraine, but patients with more disabling attacks are better served by using migraine-specific agents. When at-home therapy fails, parenterally administered drugs such as DHE, which is often combined with other drugs such as prochlorperazine or metoclopramide, may be needed. The choice of acute or preventive therapy should be guided by evidence, when available.

Preventive therapy is ideally aimed at reducing attack frequency, severity, duration, and overall impact and at achieving synergy with abortive therapy to improve its effectiveness. Pharmacological prevention includes β -adrenergic blockers, anticonvulsant or antiepileptic drugs, antidepressants, calcium channel blockers, and NSAIDs. Evidence supports the use of group 1 drugs (Table 9) as the first-choice therapies in migraine prevention.

Nonpharmacological complementary and alternative (integrative) migraine therapy includes trigger-avoidance and physical and behavioral interventions.

Table 8 Acute migraine pharmacotherapy^a

<i>Drug class/mechanism</i>	<i>Drug/compound</i>	<i>Daily dose range (mg)</i>
Analgesics, simple	Acetaminophen	1000
Analgesics, compound	Acetaminophen + ASA + caffeine	600 + 400 + 200
	Acetaminophen + codeine	(250–500) + 30
	Isometheptane + dichloralphenazone + acetaminophen	(65–130) + (100–200) + (325–650)
Ergots	Nasal dihydroergotamine	2–4
	Ergotamine	1–6
Nonsteroidal anti-inflammatory drugs	Aspirin (ASA)	650–1000
	Diclofenac	50–100
	Flurbiprofen	100
	Ibuprofen	400–1200
	Ketoprofen	75–150
	Mefenamic acid	500
	Naproxen	750–825
	Rofecoxib	25–50
	Tolfenamic acid ^b	200
	Almotriptan	6.25–12.5
Triptans	Eletriptan	40–80
	Frovatriptan	2.5–5.0
	Naratriptan	1.0–2.5
	Rizatriptan	5–10
	Sumatriptan	25–100
	Zolmitriptan	2.5–5.0

^aOral and nasal therapies with at least one randomized controlled trial demonstrating efficacy of the active compound versus either placebo or another comparator. Reprinted from Ramadan, N. M. and Buchanan, T. M. (2006), New and future migraine therapy, *Pharmacology & Therapeutics* 112(1), 199–212, with permission from Elsevier.

^bNot available in the United States.

Table 9 Migraine preventive therapy^a

	Group				
	1	2	3	4	5
Efficacy	+++	+ to ++	+++	+++	—
Evidence	Good	Limited	Consensus	Good	Good
Adverse events	Mild to severe	Mild to moderate	Mild to concerning	Concerning	
Examples	Amitriptyline, valproate, propranolol, topiramate, timolol	Verapamil, vitamin B ₂ , candesartan	Nortriptyline, phenelzine	Methysergide	Nifedipine

^a+ indicates efficacy established (+++ is highest); — indicates that there is no evidence for efficacy as migraine prophylaxis. Adapted from Buchanan, T. M. and Ramadan, N. M. (2006), Prophylactic pharmacotherapy for migraine headaches. *Seminars in Neurology* **26**(2), 188–198. Reprinted by permission.

A systematic review in 2000 led to the conclusion that relaxation training, biofeedback, and cognitive-behavioral therapy are proven treatments for migraine, but the evidence for physical treatments such as cervical manipulation and transcutaneous electrical nerve stimulation (TENS) is not as strong. Largely uncontrolled clinical trials support a role for acupuncture in migraine prevention, but two recent randomized trials failed to demonstrate that acupuncture is more effective than sham intervention.

Summary

Significant advances in the acute and prophylactic management of migraine headaches have emerged in recent years, providing relief for millions of migraine sufferers. Given that approximately 6% of men and 17% of women experience migraines, many untreated patients may benefit from proper diagnosis and treatment. Various treatment regimens for the prevention of migraines have proved effective, but long-term prognosis remains less well delineated.

Acknowledgment

The authors wish to thank Ms. Joyce Lenz for her editorial assistance in preparing this manuscript.

See Also the Following Article

Pain.

Further Reading

Aizenman, E. and Sanguinetti, M. C. (2002). Channels gone bad: reflections from a Tapas bar. *Neuron* **34**, 679–683.
 Borsook, D., Becerra, L., Carlezon, W. A., Jr., et al. (2006). Reward-aversion circuitry in analgesia and pain: implications for psychiatric disorders. *European Journal of Pain* **11**, 7–20.

Buchanan, T. M. and Ramadan, N. M. (2006). Prophylactic pharmacotherapy for migraine headaches. *Seminars in Neurology* **26**(2), 188–198.

Burstein, R. (2001). Deconstructing migraine headache into peripheral and central sensitization. *Pain* **89**, 107–110.

Burstein, R. and Jakubowski, M. (2005). Unitary hypothesis for multiple triggers of the pain and strain of migraine. *Journal of Comparative Neurology* **493**, 9–14.

Buzzi, M. G. and Moskowitz, M. A. (2005). The pathophysiology of migraine: year 2005. *Journal of Headache Pain* **6**, 105–111.

Dichgans, M., Freilinger, T., Eckstein, G., et al. (2005). Mutation in the neuronal voltage-gated sodium channel SCN1A in familial hemiplegic migraine. *Lancet* **366**, 371–377.

Goadsby, P. (1997). Pathophysiology of migraine: a disease of the brain. In: Goadsby, P. J. & Silberstein, S. D. (eds.) *Headache*, pp. 5–24. Boston: Butterworth-Heinemann.

Headache Classification Subcommittee, International Headache Society (2004). *Special issue on the International Classification of Headache Disorders* (2nd edn.) *Cephalalgia* (supplement 1).

Hu, X. H., Markson, L. E., Lipton, R. B., et al. (1999). Burden of migraine in the United States: disability and economic costs. *Archives of Internal Medicine* **159**, 813–818.

Jen, J. C., Wan, J., Palos, T. P., et al. (2005). Mutation in the glutamate transporter EAAT1 causes episodic ataxia, hemiplegia, and seizures. *Neurology* **65**, 529–534.

Koehn, C. G. (ed.) (1826). *Claudii Galeni opera omnia. vol. xii. De loci affectis*. Leipzig: In officina Car. Cnoblochii.

Lea, R. A., Nyholt, D. R., Curtain, R. P., et al. (2005). A genome-wide scan provides evidence for loci influencing a severe heritable form of common migraine. *Neurogenetics* **6**, 67–72.

Lipton, R. B. and Bigal, M. E. (2005). Migraine: epidemiology, impact and risk factor for progression. *Headache* **45**(supplement 1), S3–S13.

Lyngberg, A. C., Rasmussen, B. K., Jorgensen, T., et al. (2005). Incidence of primary headache: a Danish epidemiologic follow-up study. *American Journal of Epidemiology* **161**, 1066–1073.

- Martin, V. T. and Behbehani, M. M. (2001). Toward a rational understanding of migraine trigger factors. *Medical Clinics of North America* 85, 911–941.
- Quality Standards Subcommittee, American Academy of Neurology (1994). The utility of neuroimaging in the evaluation of headache patients with normal neurologic examinations. *Neurology* 44, 1353–1354.
- Ramadan, N. M. and Buchanan, T. M. (2006). New and future migraine therapy. *Pharmacology & Therapeutics* 112(1), 199–212.
- Rasmussen, B. K., Jensen, R., Schroll, M., et al. (1991). Epidemiology of headache in a general population – a prevalence study. *Journal of Clinical Epidemiology* 44, 1147–1157.
- Stewart, W. F., Lipton, R. B., Celentano, D. D., et al. (1992). Prevalence of migraine headache in the United States: relation to age, income, race, and other sociodemographic factors. *Journal of the American Medical Association* 267, 64–69.
- Stovner, L. J. and Scher, A. I. (2006). Epidemiology of headache. In: Olesen, J., Goadsby, P. J., Ramadan, N. M., Tfelt-Hansen, P. & Welch, K. M. A. (eds.) *The headaches* (3rd edn., pp. 17–25). Philadelphia: Lippincott Williams & Wilkins.
- van den Maagdenberg, A. M., Pietrobon, D., Pizzorusso, T., et al. (2004). A Cacna1a knockin migraine mouse model with increased susceptibility to cortical spreading depression. *Neuron* 41, 701–710.

Relevant Websites

- Campbell, J. K., Penzien, D. B. and Wall, E. W. (2000). Evidence based guidelines for migraine headache: behavioral and physical treatments. <http://www.aan.com>.
- Frishberg, B. M., Rosenberg, J. H., Matchar, D. B., et al. (2000). Evidence-based guidelines in the primary care setting: neuroimaging in patients with non-acute headache. <http://www.aan.com>.

Mineralocorticoid Receptor Polymorphisms

R H DeRijk and E R de Kloet

Leiden University Medical Center and Leiden University, Leiden, Netherlands

© 2007 Elsevier Inc. All rights reserved.

Mineralocorticoid Receptor Gene
Genetic Variants in the Mineralocorticoid Receptor

Glossary

<i>Aldosterone</i>	The classic salt-retaining steroid hormone secreted from the zona glomerulosa of the adrenal cortex; its release is stimulated by the renin–angiotensin system and rising plasma potassium concentrations.
<i>Glucocorticoid hormones</i>	Hormones synthesized and secreted from the zona fasciculata by the adrenal gland in response to stress and also during the diurnal rhythm. The main endogenous glucocorticoids are corticosterone in rodents and cortisol in humans.
<i>Haplotypes</i>	Combinations of several single-nucleotide polymorphism (SNP) alleles on the same chromosome. Here an allele is one variant of a sequence on the DNA, for

Mineralocorticoid receptors (MRs)

Single-nucleotide polymorphism (SNP)

Splice variant

example, a certain SNP. Often several haplotypes can be found in a gene due to unique combinations of SNPs.

Intracellular receptors by which aldosterone acts in epithelial cells to promote sodium transport while mediating the glucocorticoid-like effects in nonepithelial cells. In the brain, MRs bind the glucocorticoid hormones (corticosterone and cortisol) and mediate the effect of these steroids on appraisal processes and the onset of the stress response.

A variation in DNA sequence that occurs with a frequency of at least 1%. SNPs can result in amino acid changes, if they are located in the coding region. If they are not in the coding region, they can influence gene regulation by effects on, for example, gene splicing or the promoter region or through the introduction of a stop codon.

mRNA splicing occurs following transcription leading to alternative different mature mRNA molecules, so-called splice variants. Splicing can be tissue specific and context dependent with effects on mRNA expression levels and potentially leading to alternative different proteins.

<i>Transactivational properties</i>	Effects different factors have on the control of gene expression.
<i>Zinc fingers</i>	A highly conserved protein domain consisting of about 30 amino acid residues enabling corticosteroid receptors to bind DNA. Steroid receptors have typically two of these domains, encoded by exon 3 and 4. Each domain contains four cysteine residues interacting with one zinc ion, which is crucial for the stability of this domain.

Mineralocorticoid receptors (MRs) have similar affinity for the naturally occurring glucocorticoids (corticosterone and cortisol) and the mineralocorticoid aldosterone. In the epithelia, such as in the kidney, colon, sweat glands and circumventricular organ of the brain, these MRs selectively bind aldosterone, in spite of the approximately 100- to 1000-fold higher concentrations of glucocorticoid hormone in humans. MR selectivity for aldosterone is maintained through the activity of 11 β -hydroxysteroid dehydrogenase (11 β HSD) type 2, which converts cortisol and corticosterone to inactive metabolites. Through this mechanism, aldosterone can regulate the electrolyte balance, blood volume, and blood pressure. In nonepithelial cells residing in the heart, blood vessels and brain, the MRs are nonselective due to virtual absence of 11 β HSD type 2. Accordingly, MRs are exposed to both aldosterone and excess glucocorticoid hormone.

In addition to the regulation of the electrolyte balance, aldosterone and MRs have been implicated in vascular inflammation and damage and in fibrosis in the heart. These actions seem to be relatively independent from the regulation of blood pressure and electrolyte balance. In the brain, the activation of MRs can induce drinking behavior, salt craving, and other effects aimed at reestablishing the electrolyte balance. However, MRs also play a crucial role in the maintenance of cellular stability, regulation of the hypothalamic-pituitary-adrenal (HPA) axis, behavioral adaptation, and cell homeostasis. The MRs involved in the latter functions are predominantly located in the hippocampus, where the receptors are activated by cortisol or corticosterone, due to the relative absence of 11 β HSD type 2. Finally, recent data indicate the expression of the MRs in adipose tissue, regulating adipocytes differentiation and thermogenesis.

On the binding of the ligand, aldosterone or cortisol/corticosterone, MRs translocate from the cytosol to the nucleus and bind as a homodimer to hormone responsive elements (HREs). Transcriptional activation occurs through direct interactions with

transcription factors and recruitment of co-regulator proteins. This recruitment is governed in part by the ligand and by the HRE. Co-regulator complexes are involved in chromatin remodeling and the recruitment of numerous other factors that are part of the transcriptional machinery.

Genetic variants of the MR gene (i.e. splice variants, deletions, insertions, or SNPs) probably have effects on these functions. These effects include not only an action on the electrolyte balance and cardiovascular control but also on vascular inflammation, behavior, HPA axis control, and homeostatic control mechanisms.

Mineralocorticoid Receptor Gene

The MR (NR3C2) is a member of the nuclear receptor family and resembles the glucocorticoid receptor (GR; NR3C1). In the DNA-binding domain, amino acid identity between MR and GR is approximately 94%, in the C-terminus 57%, and in the N-terminus less than 15%. Transcriptional regulation is mediated by the N-terminal part (exon 2, region AF-1a and AF-1b), plus a small part of the ligand-binding domain (C-terminus, AF-2). DNA binding involves residues in exons 3 and 4. Furthermore, exons 5–9 are involved in ligand binding, dimerization (with exons 3–4), heat shock protein (Hsp)90 binding, nuclear localization, and transcriptional regulation.

The MR gene is located on chromosome 4 (4q31.1) and has a size of approximately 350 kbp (this can be viewed on the Ensembl website). On a regular basis, new SNPs are being identified. The full protein contains 984 amino acids and has a size of approximately 107 kDa. Two (possibly three) exon 1s exist, which can join exon 2 (see [Figure 1](#)).

Genetic Variants in the Mineralocorticoid Receptor

Genetic Variability

Genetic variants can result from mutations causing deletions, insertions, or just changes of the nucleotide at a given position. Depending on the position in the genome, different effects on the phenotype can be observed. Genetic variants can affect protein structure or regulation through different mechanisms. Even single base pair substitutions can cause dramatic changes to the phenotype; however, the vast majority of such changes which reach appreciable frequencies in the population (referred to as single nucleotide polymorphisms, SNPs) are benign.

The number of different SNPs in the human genome is probably in the range of 3–4 million out of

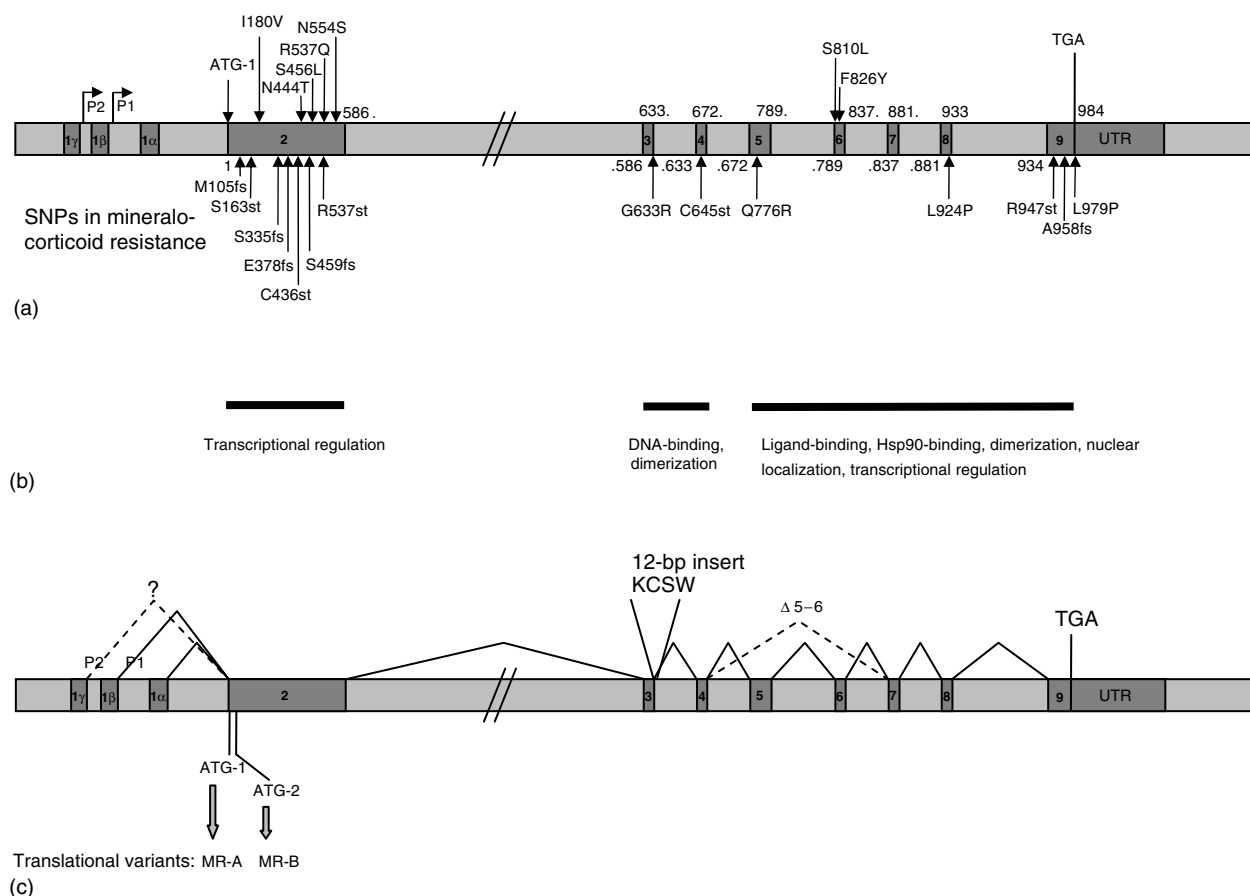


Figure 1 Human MR variants. a, Reported single-nucleotide polymorphisms (SNPs) leading to amino acid changes; b, the three functional domains; c, splice variants and translational variants. The number of the exon is in bold: 1 α , 1 β , 1 γ , 2–9. The exons are flanked by upper and lower numbers, which indicate the codon (or amino acid), ending with 984 in exon 9. If the number is followed by a dot, the codon spans both exons. Translation starts in the beginning of exon 2 and ends in exon 9 (TGA), followed by an untranslated region (UTR). In a (upper part), the indicated SNPs leading to amino acid changes are shown (from the Ensembl website), with the addition of S810L (associated with hypertension in pregnancy). In a (lower part), specific SNPs are shown that are found to be associated with mineralocorticoid resistance. These have a very low frequency and are only found in specific kindreds. The effects of these SNPs can involve a frameshift (fs) or the introduction of a premature stop codon (st). In b, the three functional domains are shown: the transcriptional active domain, the DNA-binding domain, and the ligand-binding domain (which has several other functions). In c, although both exon 1 α and 1 β can join exon 2, this is not clear for exon 1 γ . In addition to most common annealing pattern (exons 2–9), exon skipping of exons 5 and 6 occurs, resulting in a frameshift with a new termination codon ending in exon 7 (position 2836). At the end of exon 3, 12 extra base pairs can be inserted due to alternative splicing (coding for Lys-Cys-Ser-Tyr), and joined with exon 4; this form is designated MR4+. KCSW, lysine-cysteine-serine-tryptophan.

a total of approximately 3 billion base pairs, counting only SNPs with a minor allele frequency of more than 1%. SNPs make up 90% of all human genetic variations, and two of every three SNPs substitute cytosine (C) with thymine (T), due to a high level of deamination of cytosine. These changes in nucleotide sequence give rise to alternative forms of a gene, named alleles. Combinations of several alleles on the same chromosomes are called haplotypes. However, not all possible combinations of the variants in a stretch of DNA are observed because of specific allele association or so-called linkage disequilibrium (LD). Often blocks of several SNPs are found, giving rise to several haplotypes in a gene.

SNPs are increasingly used to study the involvement of a certain gene and its protein product, in a complex phenotype, as is presented in psychiatric diseases. The association of an SNP in a given gene with a certain phenotype indicates the involvement of the gene. Alternatively, the carriers of a certain SNP/haplotype could represent a vulnerable (or protected) phenotype. Moreover, the determination of genetic variants could be used for more individual treatment of patients based on their specific needs, with respect to treatment efficacy and side effects.

However, several pitfalls are encountered when using variants in genetic research. Phenotyping is crucial because the effect of genetic variants are

often moderate, which also calls for the use of large numbers of subjects in association studies. The elucidation of the molecular mechanism of a functional variant will strengthen the *a priori* hypothesis and will also help identify which genetic variant in a given haplotype causes the observed phenotype. Finally, the replication of an association in an independent cohort of subjects will overcome the often observed false-positive associations.

Variants of the Human Mineralocorticoid Receptor Gene

Splicing variants Several alternative splicing variants exist at the 5' end of the gene: exon 1 α , exon 1 β , or exon 1 γ can join exon 2 (Figure 1c). The existence of exon 1 γ is still not completely clear. Upstream of exon 1 α and 1 β , many regulatory sites have been identified, and this region (P2 and P1) is considered the proximal promotor. The effect of the alternative use of exon 1s is still unresolved. However, during development a differential regulation and response to adrenalectomy of 1 α , 1 β , or 1 γ is observed in the rat hippocampal CA1, CA2, and CA3 pyramidal cell fields and the dentate gyrus (DG) neurons.

In the human MR, a 12-bp insert directly adjacent to the 3' end of exon 3, as a result of alternative splicing, has been found. The result is four extra amino acids (lysine-cysteine-serine-tryptophan; KCSW) in between the two zinc (Zn)-fingers involved in DNA binding. No clear effect of this splice variant, especially on DNA binding, was found. It is, however, expressed in most human tissues, with relative levels varying from 5 to 50%, compared to the most common MR.

Another human MR isoform is derived from skipping exons 5 and 6 with a resultant premature stop, giving rise to a protein of 75 kDa, retaining DNA-binding capacity and ligand-independent transcriptional activity. The high level of expression of this splice variant makes it a potential important modulator of MR effects.

Translational variants also exist as a result of the alternative use of different translational start sites. Two different forms of the MR, designated MR-A and MR-B have recently been described, with different transcriptional activities.

Cardiovascular control Research into the molecular mechanism of mineralocorticoid resistance or pseudohypoaldosteronism (PHA1), characterized by urinary salt wasting and hypotension, has revealed several inactivating mutations in the MR gene. Two forms of PHA1 exist: an autosomal dominant form

that is mild and a recessive form that is more severe due to defects in the epithelial sodium channel. In the autosomal dominant form, the target organ defect is confined to the kidney and clinical expression varies from asymptomatic to moderate. It may be severe at birth, with a failure to thrive, but symptoms often remit with age. The frequency of these mutations, including frameshifts, splice-site mutations, and premature stop codons, is quite low, which accords with the fact that PHA1 is a relatively rare disease. The usefulness of these mutations lies primarily in the understanding of MR protein activity with respect to DNA binding, translocation, and transactivational properties.

An activating mutation in the MR has been detected in patients with severe hypertension, which is exacerbated in pregnancy. The mutation, serine to leucine (S810L; Figure 1a), was found to alter the MR ligand specificity; progesterone and other steroids lacking the C-21-OH group became agonists.

Stress responsiveness Mutations of the MR may compromise the role of this receptor in the control of neuronal stability and stress responsiveness. Indeed Brown Norway (BN) rats are, compared to Fisher rats, insensitive to adrenalectomy because there is no weight loss and resistance to saline intake. In the BN-rat MR gene, a tyrosine to cysteine (Y73C) substitution was detected. This mutation revealed a greater transactivational activity of aldosterone and, interestingly, also of progesterone, suggesting that the MR is constitutively more active in the BN rat. It is of interest to note that BN rats have a long life span and seem to have a relatively stable HPA axis, compared to other rat strains.

In humans, some SNPs have a high enough frequency to be potentially important for stress-related disorders, such as anxiety and depression. We found the MR 180V (MRI180V, isoleucine to valine; allele frequency 11%) to be associated with measures of depression in an elderly cohort. In human 180V carriers (i.e., the heterozygote I180V and the homozygote V180V genotypes), cortisol levels and heart rate were enhanced in response to psychological challenge evoked in the Trier Social Stress Test (Trier Social Stress Test), whereas the SNP was not associated with changes in blood pressure in response to salt loading as part of the Weinberger's test. Moreover, *in vitro* the response of the MR 180V allele to cortisol was attenuated, which was not observed with aldosterone. The finding suggests that such a subtle SNP can have profound effects on stress responsiveness with potentially important consequences for stress-related pathology.

See Also the Following Articles

Aldosterone and Mineralocorticoid Receptors; Corticosteroid Receptors; Steroid Hormone Receptors; Genetic Polymorphisms in Stress Response; Trier Social Stress Test.

Further Reading

- Arriza, J., Weinberger, C., Cerelli, G., et al. (1987). Cloning of human mineralocorticoid receptor complementary DNA: structural and functional kinship with the glucocorticoid receptor. *Science* **237**, 268–275.
- de Kloet, E. R., Joëls, M. and Holsboer, F. (2005). Stress and the brain: from adaptation to disease. *Nature Reviews Neuroscience* **6**(6), 463–475.
- DeRijk, R. H., Wust, S., Meijer, O. C., et al. (2006). A common polymorphism in the mineralocorticoid

receptor modulates stress responsiveness. *Journal of Clinical Endocrinology and Metabolism* Oct. 3 [Epub ahead of print].

- Geller, D. S. (2005). Mineralocorticoid resistance. *Clinical Endocrinology* **62**, 513–520.
- Vazquez, D. M., Lopez, J. F., Morano, M. I., et al. (1998). Alpha, beta, and gamma mineralocorticoid receptor messenger ribonucleic acid splice variants: differential expression and rapid regulation in the developing hippocampus. *Endocrinology* **139**, 3165–3175.
- Zennaro, M.-C. and Lombes, M. (2004). Mineralocorticoid resistance. *Trends in Endocrinology and Metabolism* **15**(6), 264–270.

Relevant Website

Ensembl website. <http://www.ensembl.org>.

Mineralocorticoids, Overview See: Adrenal Cortex; Aldosterone and Mineralocorticoid Receptors.

Minorities and Stress

I Mino

Harvard University, Cambridge, MA, USA

W E Profit

Los Angeles, CA, USA

C M Pierce

Harvard University, Cambridge, MA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by I Mino, W E Profit, and C M Pierce, volume 2, pp 771–776, © 2000, Elsevier Inc.

Basic Issues

Disadvantaged Minorities

Treatment Considerations

Glossary

Culture

Refers to the dominant set of symbolic codes (linguistic, moral, aesthetic) and material practices (dietary, behavioral) that characterize a group. Culture may

refer to an entire society's codes and practices, as when reference is made to the American or Japanese culture.

Deculturated

When people lose their culture or cannot use it because of changed circumstances.

Discrimination

A negating selection and differential comparison not based on merit.

Ethnicity

Refers to the particular reference group for individuals with a shared heritage, e.g., Puerto Rican, American Jewish. Some reference groups may be voluntary, such as those based on religion. Others may be involuntarily ascribed to the persons, such as those based on race. A part of a population, numerically less than 50%, differing from others in some characteristics and often subjected to different treatment.

Minority

Prejudice

An injurious and intolerant preconceived judgment.

Racism

Intentional or unintentional bias directed at individuals or groups based upon notions of the superiority or inferiority of skin color.

Stress

In minority communities, stress occurs when an individual or group's resources are overextended, even to the point of causing paralysis, apprehension, disintegration, and ineffectiveness from perceived and/or actual duress and assaults from the majority community.

Basic Issues

The basic issues in discussing minorities and stress include defining minority, differentiating the individual's stress from the stress of his or her minority group, and deciding which minority variables are best studied.

Who Is a Minority?

Everyone can claim multiple minority memberships. In fact, an individual may only or best be described in terms of such memberships. For instance, there are over 6000 languages in the world. Everyone's native language is a minority language. No one can claim majority status in terms of nationality, age, occupation, religion, avocation, or talent. Some minority memberships bring honor, others dishonor. Minority memberships may be forced or sought. They can be voluntary or involuntary. Many memberships are visible, often relating to appearance, including bodily features, uniforms, and insignia. Language usage can suggest a minority status. However, many minority memberships are invisible and can even be kept secret, such as migrant status, transnational allegiance, racial identity, and political and social affiliations.

Cultural and social forces determine the saliency of minority memberships. Humans create hierarchies, which in turn breed class distinctions, whereby minorities are made through a process of exclusion or inclusion. Rules of membership can be flexible or rigid, fluid or static. In a given community one might be in the most favored economic group but the least favored social group. Many minority memberships find favor or disfavor from legal regulations.

No matter which minority memberships one has, there are two conditions that must be negotiated for each minority membership. On the one hand, an accommodation for being acculturated and/or deculturated is made. On the other hand, some people or agency has to accept you as a member of the group. Such certification, written or unwritten, can be meticulous or arbitrary. The evaluation may not be accurate, complete, or satisfying to either the assessor or the assessed. Furthermore, certification of legitimacy may vary from the vantage point of various individuals, groups, and organizations.

Most individuals probably give little thought to most of their minority memberships. Therefore, they have not rank-ordered them. Yet, they are aware that the saliency and importance in their lives for any such membership is in a constant state of flux. The importance of a minority membership depends on the combination of specific, general, immediate, and far-flung circumstances at any moment and place.

These ever-shifting interpersonal interactions and intrapersonal considerations mean that in some sense an individual's minority status is situational. Furthermore, the situation depends both on what the individual perceives and on what other persons and groups perceive, even when there is agreement about exactly what minority membership is being scrutinized. Misperceptions between parties can range from being innocuous to being life-threatening. A common cause for misperceptions comes from different opinions about the speed, type, and quantity of acculturation. Also, views may differ significantly about how much a person or group is acculturated, should be acculturated, or can be acculturated.

Seldom is it possible to address these misperceptions among parties without the withering burden of stereotypical thinking by all participants. Thus, an obstacle in contemplating or negotiating a minority or majority issue is that all parties have confusion about what is unique to the individual versus what is a stereotype about a collective group.

The importance of and investment in any minority membership varies from person to person and even from time to time, whether or not one is a member of the in-group. Therefore, even with unequivocal certainty that one is a minority or one is dealing with a minority, the same stressors do not exert uniform force among the actors. Everyone attaches different valences to any minority status, even those that occupy his or her most intense concern. Furthermore, it is usually not possible to ascertain how much any participant identifies with, accepts, or knows about any minority membership that is under discussion.

Is Stress the Same for the Individual and the Group?

Stress results when resources are overextended and overwhelmed. There is the threat of dissolution with accompanying fears of abandonment and inability to escape. Stress is the reciprocal of support. Theoretically, perfect biological, sociological, and psychological support would equal no stress. Stress may be acute, chronic, or intermittent, with dimensions of intensity and duration.

Individuals in a group under stress will have varying perceptions and adjustments. These differences are mediated by both existing background factors

and factors peculiar to the stress. For instance, age, state of health, prior relevant experience, and faith are examples of attributes that help govern response to stress.

The number and general condition of a group under stress, as well as the quality of its leadership and its ability to be cohesive and cooperative, are important determinants of the group's outcome. Resourcefulness and a fierce will to survive increase the odds for a favorable outcome. The ability to take decisive, independent action is important.

In groups under stress, behavior of individuals speaks to the important aspect of group morale. Among such behaviors are the willingness to share, carefulness never to endanger the group, and the ability to perform multiple tasks.

Some individuals in any group thrive better than others. This seems true even when all the participants have similar backgrounds and quality of health. These individuals, for whatever reasons, may have more tolerance for uncertainty and ambiguity. Likewise, they may be more able to consider the consequences of the uncontrollable, the unpredictable, and the unexpected, with a more sanguine philosophy. In a mundane extreme environment such as a racial group living under oppression, such individuals may be more accepting of high risk and low reward, as the price for their own and their group's survival. Such persons are the group's everyday heroes, who demonstrate the greatest hope in the group that relief and amelioration are in the future.

Stressful events, due to their limited duration and more discrete circumstances, may be easier to study than stressful backgrounds. Groups and individuals in a railroad crash may be more easily understood and supported than individuals undergoing daily stress. In stressful events the individuals are more conscious of and willing to work for the common survival of the group. This shared specific stressful event will become another minority membership for each participant.

However, individual minority background experiences are filled with so much randomness, variability, and strategy that they are not easily delineated for study. There is a wide range of responses by individual members exposed to the same general trauma. Proximate factors such as personal loss and pain, differences in exactly what was witnessed, and the quantity of available and accessible resources required are never identical. This alone determines a plethora of heterogeneous responses by minority individuals under stress.

Which Minority Variables Are Best Studied?

The great bulk of minority studies are performed on populations considered disadvantaged. They verify in

detail that, in general, life is more terrible for these subjects. The data are abundant in this verification, whether addressing health disparities, inequalities in the law, unfair housing and employment, discrepancies in educational opportunity and achievement, harshness in the criminal justice system, insensitivity to the needs of children, etc. The studies focus on vulnerable populations and indicate negative patterns of service, response, and rates of problems. They document that under all manner of geographic, demographic, and political climates these minorities fare poorly.

There is an awareness of, but underpublished interest in, the strengths of individuals in these groups. The strengths the groups had to have in order to survive are less emphasized. Probably methodologies have not evolved that can capture the extensiveness of life events that permit and sustain the host of strengths that differentiate disadvantaged minority people.

A related issue is the huge thrust on aspects of resiliency in these populations, without much consideration of resistance. Studies of resilience imply post-traumatic phenomena. Resistance inquiries seek to understand the defenses against the offenses delivered to the disadvantaged populations, in which trauma continues to accumulate but is deflected, minimized, or diluted.

Disadvantaged Minorities

Social, cultural, economic, and political pressures determine what minority memberships are designated as disadvantaged. They may share features as the result of such designation, but each majority community would elect, in its own manner, which groups are targeted and what form of disenfranchisement and handicap is appropriate.

General Characteristics

A society may have written and unwritten rules that designate a group as disadvantaged. Such groups are disenfranchised in numerous ways. It is permitted and expected that the group will be abused and exploited. They occupy the margins of the society, often fractionated and segregated. The groups are likely to be the source of both amusement and wrath for the general population. While given substantially less in the community, it is demanded that they do more for less reward.

Petitions for succor by the disadvantaged prompt exasperation, perplexity, and astonishment. Social machinery is elaborated and sustained that is designed to control, dominate, and disrespect the unworthies. As such, the designated disenfranchised group struggles for its security, safety, and livelihood.

The disadvantaged are tyrannized and easily disregarded, discounted, and trivialized. Even so, they are feared as social or economic competitors. Regardless of their powerlessness, they are dismissed as unproductive noncontributors. They may be resented on grounds of their standards and values. They are given blame but little sympathy for an improper identity. Often they are accused of inviting their own catastrophe. In many cases, such as being female or having colored skin, there is no way the person can banish that identifier.

An individual member of a disadvantaged minority may live a lifetime with little or no direct, overt discrimination. Of course, many are victims of hatred, including physical assault, destruction of possessions, deliberate neglect, and even death.

More common are everyday minor, but cumulative, aggressions and insults. For instance, a White couple approaches a well-dressed Black male in a luxury hotel lobby and requests a taxi. They assume, incorrectly, that the man is an employee. The lifelong toll from these microaggressions, both physiologically and psychologically, remains unstudied.

An abundance of comparative data demonstrates the hardship and adversity suffered by disadvantaged minorities. For instance, between 1996 and 2002, the gap between American White and Black wealth grew from 10:1 to 14:1. In the United States, White family wealth now is 14 times the median net worth of Black families. Blacks make twice as many emergency room and outpatient visits as Whites. They make a third fewer visits to a doctor's office, however. In 2003, the percentage of Whites with advanced degrees was twice that of Blacks with advanced degrees. The same numbers hold in the comparison of White and Black college graduates.

There are clear burdens and risks imposed on groups designated as disadvantaged. In general, they are treated unfairly, and their efforts to improve their conditions generate a significant portion of the world's stress and violence.

The historical circumstances of each country determine which groups are disadvantaged and what efforts are made to promote intergroup peace and prosperity. In the United States, as elsewhere, there are numerous disadvantaged minorities. Due to demographic considerations, the most important issue for the United States in the twenty-first century may relate to disadvantaged minorities. It is expected that during this century Whites will become the minority population. The reduced proportion of Whites probably will be accompanied by significant alterations in the influence and hegemony of Whites. How Whites, especially White males, respond to this shift, may be the defining activity of the United States in this century.

Typological Classification

Any classification scheme is both arbitrary and limited. It follows that there can be no complete agreement, since every citizen can use his or her own system. Overall, in the United States, those generally considered by the general public to be disadvantaged minority groups meet a few broad criteria. First, there is evidence, usually with academic credibility, that the designated status does bring uncommon burdens to both the designated member and the general society. Second, the designated minority usually has a long and continuous history regarding getting legal relief from specified ills. Third, in the view of the public, sometimes including members of the minority itself, the designated group inspires fear, requires restraint, and often is deserving of mild to severe punitive actions.

There are those who are constitutionally handicapped. People are born into this category. These include women as well as selected racial/ethnic groups. The groups selected by the United States government include Blacks, Hispanics, Asians, and Native Americans.

The other large category, which most of the public considers disadvantaged, makes up a disparate miscellany. In this general group are migrants, the disabled, criminals, religious and cult community groups, especially those distant from core Christian organizations, and people with a variety of conditions in which there is some disagreement as to how much genetics is at fault. These conditions include obesity, homosexuality, alcohol and drug addiction, and sexual deviations.

All the groups in this classification scheme have been persecuted. For instance, Blacks were enslaved, Native Americans were objects for official genocide, and Hispanics lost large tracts of land to the United States. In general, in the United States, cross-racial interactions, particularly between White and colored races, has brought conflict and stress.

Special Obstacles

Designated minorities meet ongoing stress in an attempt to overcome obstacles in pursuing an occupation, finding adequate housing, and attracting political benefits. Always they remain sensitive as to whether they are welcome or merely tolerated. There is the ceaseless focus on how much, how fast, and in which ways the person and the group accommodate to or resist oppression.

All the separate disadvantaged minorities in the categorization have their own issues. For instance, in the racial/ethnic group there is enormous ambivalence about the cost of integration in such terms as increased intermarriage and loss of unity and leadership as segregated housing diminishes.

As another example, many migrants remain closely tied to their country of origin. The economy of entire countries would be seriously impaired without the input of dollars remitted by migrant workers in the United States. Some migrants in the United States exercise voting rights in other countries. Strong transnational loyalty has to be balanced against charges of diluted loyalty from host nationals.

In this age of wide and rapid communication, designated disenfranchised groups know, as studies show, that bad behavior by one of them has a disproportionate and negative impact on the memory and attitude of majority citizens. The problem of having little control over portrayal by the media becomes a vexing issue of some magnitude.

All the disadvantaged groups protest that the rules are constantly and covertly altered. The fundamental belief is that their space, time, energy, and mobility should be limited and serve majority purposes. This makes the quest for equality all the more elusive, since rules once learned and a system once understood become mystifying and obfuscating.

An advantaged population tends not to see any special, unmerited, or unjustified privilege in what it does. Similarly, it may not calculate any disadvantage from having privilege. Not counting the cost to themselves of disenfranchising whole groups brings serious consequences to the majority. However, if the police have permission to exercise tactics in the disadvantaged community that would not be allowed in the advantaged community, it means that the entire community has elements of a police state. Well over 25% of Americans are in a disadvantaged group. The United States operates under the stress of a huge disadvantage. Any body politic that functions at less than 100% potential is handicapped. The oppressing portion of a society has an obstacle to its own achievement when it does not count the cost of oppressive behavior in everything from total gross national product to intellectual, cultural, and technical achievement, as well as personal freedom.

Treatment Considerations

History teaches that minority groups do best when they themselves are unrelenting in seeking redress from wrongs inflicted upon them. The broad societal pressures are exerted on and felt by all members of a disadvantaged minority. Yet it is hard to imagine how any one-on-one process, even if achieved, could eliminate a minority group's stress. Accordingly, treatment possibilities must be conceived on a grand scale. Only a broad approach can be sketched.

In the foreseeable future, direct medical interventions seem less promising than possible sociological

and psychological approaches. The aim in these approaches is to provide increased hope and increased esteem to both the whole disadvantaged group and its individual members. In order to accomplish this objective, the groups themselves must somehow articulate and agree upon who is in their group and what the group cherishes. It must decide what to preserve and what to relinquish in accomplishing intergroup amity. The tactics required to reach these goals involve decreasing stress and increasing adaptation. Political and social action is obligatory.

Decreasing Stress

The disadvantaged minorities and their allies will need to continue and enlarge concerted, dedicated, social, and political actions to decrease stress. Diverse, multiple means must be instituted to teach the group, especially its youngsters, to anticipate the future. To do this, the group will have to mobilize structures at home, in community organizations, at schools, and at churches. All media channels, especially the Internet, TV, and radio, must be enlisted to help. The substantive debate about the details of the subject matter needs participation by persons who will inform themselves, seek input from the community, and be mindful of the history of the group.

Stress research would predict that supplying ways to help oneself and securing more support would be curative. Among topics, for instance, for colored groups would be propaganda analysis to learn to modulate messages that promote defeatism, futility, and loss of self-respect; lifestyle education to indicate the available steps that need to be and can be taken to increase the quality and quantity of life; and cross-racial interactions to emphasize aspects of group dynamics that aid or hinder interpersonal relations.

Perhaps special focus for many groups needs to be on networking and applying demographic remedies. For instance, in the United States, the racially/ethnically disadvantaged compared to the general population tend to be younger, more segregated, and more urban. A firm command of information of this sort will allow framing the future in terms of possible strengths and weaknesses.

Increasing Adaptation

Hopelessness and helplessness are the chief enemies to those under stress. The overall need is to supply support. Corollary to this need is to maximize adaptation. That is, one must cope.

For members of disadvantaged groups and the group as a whole, adaptation, historically, has required education as the critical step toward opportunity.

Acquisition of exploitable skills and unquestioned competency in these skills are more difficult to obtain by disadvantaged minorities. Therefore, much of their social action and networking has to be aimed at securing education.

Community members can facilitate adaptation for each other by countermeasures against noxious environmental stimuli: simple acts of encouragement, providing positive feedback, and giving credit for coping. Similarly, coping is facilitated by encouraging the members under stress, particularly youngsters, to explore options and to be cautious about foreclosing on any opportunity. Community members in a cohesive, cooperative community can augment coping by helping their community provide proper sites and environment for relaxation, another requirement for those in need of surcease. A community may underestimate its ability to contribute to coping. Yet by finding ways to make the community as stable as possible and by trying to reduce feelings of isolation the coping process is facilitated.

People cope better when they do not romanticize the past; when they have a present, clear, and certain mission; and when they appreciate the importance and urgency of that mission. Community members, by their positive attitudes and expectations, can contribute to the individual's enthusiasm for the direction of his or her future.

In coping, helplessness and hopelessness are to be avoided. Their presence may be manifested clinically, particularly by depression and anxiety. Furthermore, depression and anxiety may be comorbid with addiction, violence, obesity, diabetes, and cardiovascular disease. Medical surveillance to prevent illness and medical intervention to treat illness are important considerations in helping disadvantaged minorities to cope.

It is here that ongoing research holds promise to help disadvantaged populations. Advances in molecular biochemistry, endocrinology, genetics, and neurobiology can bring great relief to individuals and families of disadvantaged minorities. Being sure that

medical care is accessible will depend somewhat on the political and social activities of the group itself.

See Also the Following Articles

Cultural Factors in Stress; Cultural Transition; Ethnicity, Mental Health; Racial Harassment/Discrimination.

Further Reading

- Bell, C. C. (2004). *The sanity of survival: reflections on community mental health and wellness*. Chicago, IL: Third World Press.
- Cohen, S., Kessler, R. C. and Gordon, L. U. (eds.) (1995). *Measuring stress: a guide for health and social scientists*. New York: Oxford University Press.
- Cornish, E. (2004). *Futuring: the exploration of the future*. Bethesda, MD: World Future Society.
- Green, B. L., Lewis, R. K. and Bediako, S. M. (2005). Reducing and eliminating health disparities: a targeted approach. *Journal of the National Medical Association* 97, 25–30.
- Gruenewald, T. L., Kemney, M. E., Aziz, H., et al. (2004). Acute threat to the social self: shame, social self-esteem and cortisol activity. *Psychosomatic Medicine* 66, 915–924.
- Hobfoll, S. E. (1998). *Stress, culture, and community: the psychology and philosophy of Stress*. New York: Plenum Press.
- Jackson, J. S. and Volckens, J. (1998). Community stressors and racism: structural and individual perspectives on racial bias. In: Arriaga, X. B. & Oskamp, S. (eds.) *Addressing community problems: psychological research and interventions*, pp. 19–51. Thousand Oaks, CA: Sage Publications.
- National Geographic Society. (2004). *Atlas of the world* (8th edn.). Washington, D.C.: National Geographic Society Press.
- Smedley, B. D., Stith, A. Y. and Nelson, A. R. (2003). *Unequal treatment: confronting racial and ethnic disparities in healthcare*. Washington, D.C.: National Academy Press.
- Special Report. (2004). The state of African American health. *The Crisis* 111, 17–35.
- U.S. Census, Bureau. (2004). *Statistical abstract of the United States: 2004–2005* (124th edn.). Washington, D.C.: U.S. Government Printing Office.

Mitochondria

I Manoli

National Institutes of Health, Bethesda, MD, USA, and
University of Athens, Athens, Greece

S Alesci

National Institutes of Health, Bethesda, MD, USA

G P Chrousos

University of Athens, Athens, Greece

© 2007 Elsevier Inc. All rights reserved.

Background on Mitochondria

Mitochondria as Targets for Stress Hormones and
Mediators

Mitochondrial Adaptation and Dysfunction during Acute
and Chronic Stress

Seeking Mitochondrial Anti-stress Treatments

Glossary

Apoptosis

The active process of controlled cellular self-destruction, which is intrinsically programmed. It plays an important role in the development of multicellular organisms and in the regulation and maintenance of cell populations in tissues depending on physiological and pathological conditions. It can be initiated by the mitochondrion in response to energy deficiency, oxidative stress, and increased Ca^{2+} , among other stimuli.

Mitochondria

Intracellular organelles containing their own genetic material and the enzymatic machinery for major metabolic pathways, including the citric acid cycle, fatty acid oxidation, and oxidative phosphorylation, and for apoptosis.

Oxidative phosphorylation

The coupling of electron transport in the respiratory chain to ATP synthesis via the proton gradient and ATP synthase in the presence of oxygen, which occurs at the inner mitochondrial membrane.

Oxidative stress

The imbalance between reactive oxygen species production and antioxidant defenses, resulting in cellular damage.

Reactive oxygen species (ROS)

Oxygen-containing chemical species with unpaired electrons that are highly reactive and function in cell-signaling processes but also, at higher level, can damage cellular macromolecules and participate in apoptosis.

Background on Mitochondria

Mitochondria play a pivotal role in cell homeostasis, housing multiple metabolic pathways (including the tricarboxylic acid-, β -oxidation- and urea cycle) and the oxidative phosphorylation machinery that generates energy in the form of ATP. Research during the last decade has extended the prevailing view of mitochondrial function well beyond their critical role in supplying energy. Mitochondria synthesize heme and steroid hormones; are important for intracellular Ca^{2+} metabolism and signaling; generate the majority of cellular reactive oxygen species (ROS); and serve as the cell's gatekeeper for apoptosis. They contain multiple copies of circular mitochondrial DNA (mtDNA) and function under dual genetic control. Both nuclear and mtDNA encode genes that modulate mitochondrial protein synthesis and import, enzymatic activities, biogenesis, and apoptosis.

Most of the ATP necessary to supply energy to tissues and organs is generated by the mitochondrial oxidative phosphorylation system (OXPHOS). Under aerobic conditions, electrons derived from the oxidation of glucose or fatty acids are transferred through the respiratory chain complexes I–IV. At each of the enzymatic complexes I, III, and IV, protons are pumped out of the mitochondrial matrix into the intermembrane space, generating an electrochemical gradient used by the fifth enzymatic complex (ATP synthase) to drive ATP synthesis.

Mitochondria consume more than 90% of the cell's oxygen and are therefore the main site of ROS production. ROS, including superoxide, hydrogen peroxide, and hydroxyl radicals, are generated as a by-product of mitochondrial activity, primarily at complexes I, II, and III of the respiratory chain. Normally, 2–5% of O_2 used in the mitochondria is metabolized into superoxide. The superoxide radicals are converted to hydrogen peroxide by superoxide dismutase (SOD), which can further form the highly reactive hydroxyl radical. Mitochondrial (Mn-SOD and glutathione peroxidase) and cytosolic (Cu-SOD and catalase) antioxidant enzymes help to scavenge the ROS and limit their toxicity.

ROS play a significant role in the regulation of cell-signaling processes and in cytoprotection; thus their controlled generation is necessary for cell survival. On the other hand, when produced in excess of cellular antioxidant reserves, lipid peroxidation,

mtDNA damage, OXPHOS dysfunction, and damage to Fe–S-containing enzymes ensue. Oxidative stress to nuclear or mtDNA results in strand breaks and base modifications, followed by cellular dysfunction, mutagenesis, and carcinogenesis.

Mitochondria are involved in cell death by at least three general mechanisms: the release of proteins that trigger the activation of the caspase family of proteases, the disruption of oxidative phosphorylation and ATP production, and the modification of the cellular reduction–oxidation (redox) potential. Each of those mechanisms can result in programmed cell death (apoptosis) or cell necrosis. The initiation of apoptosis occurs via the opening of the permeability transition pore (mtPTP) located in the inner mitochondrial membrane. The mtPTP is composed of the adenine nucleotide translocator (ANT), voltage-dependent anion channel (VDAC), Bax, Bcl2, cyclophilin D, and benzodiazepine receptor. The opening of the mtPTP leads to swelling of the mitochondrial inner membrane and the rupture of the outer membrane, followed by the release of

cytochrome c, procaspases 2,3,4, and caspase-activated deoxyribonuclease (CAD) into the cytoplasm. Cytochrome c activates apoptotic protease activating factor (Apaf-1), which first activates procaspase 2 and 9, and subsequently caspases 3, 6, and 7, inducing apoptosis. Another caspase-activating protein released by mitochondria is the apoptosis-inducing factor (AIF). AIF and endonuclease G are transported to the nucleus and initiate chromatin condensation and degradation. Depending on the amount of cytochrome c and caspase inhibitors available, the cell undergoes either necrosis or apoptosis. The Bcl-2 family proteins exhibit pro- and/or anti-apoptotic properties (Figure 1).

Given their critical role in cell physiology, it is obvious that mitochondria are among the first responders to various stressful stimuli challenging cell homeostasis. They are responsible for meeting the enormous energy demands of the fight-or-flight response in vital tissues. Among other equally important effects, they control the fever response by the modulation of thermogenesis, balance the host immune

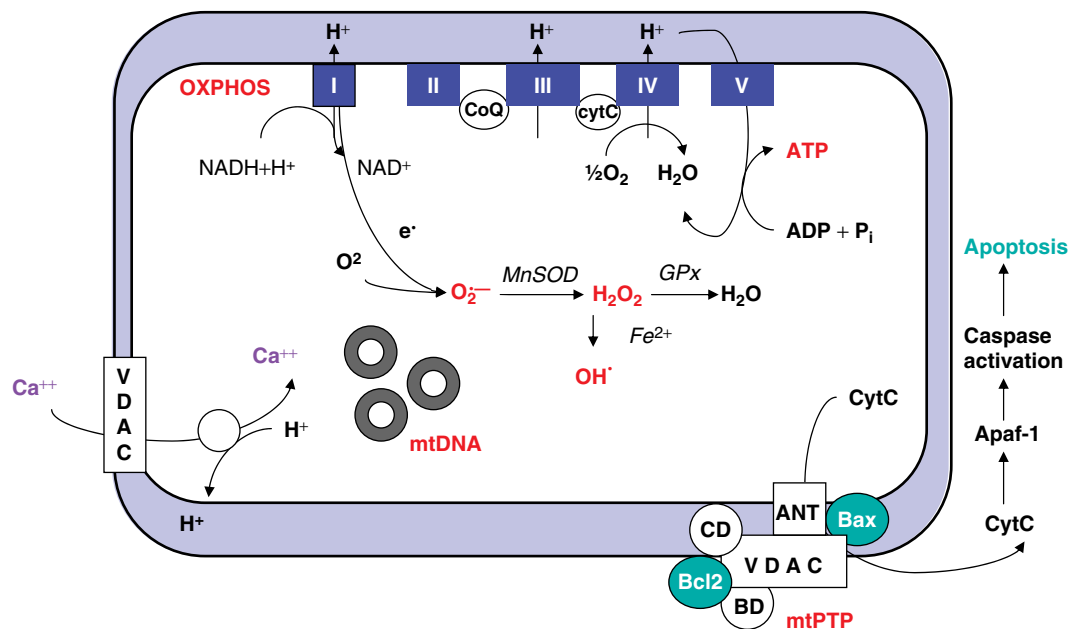


Figure 1 Mitochondrial functions. The figure illustrates (1) energy (ATP) production at the oxidative phosphorylation chain (OXPHOS) in the inner mitochondrial membrane, (2) reactive oxygen species (ROS) generation in the mitochondrial matrix and the enzymes Mn superoxide dismutase (MnSOD) and glutathione peroxidase (GPx), and (3) induction of apoptosis through the opening of the mitochondrial permeability transition pore (mtPTP). The OXPHOS complexes are complex I (NADH–ubiquinone oxidoreductase), complex II (succinate–ubiquinone oxidoreductase), complex III (ubiquinol–cytochrome c oxidoreductase), complex IV (cytochrome c oxidase), and complex V (ATP synthase). Ca^{2+} defuses through the outer membrane via the voltage-dependent anion channel (VDAC). The VDAC together with adenine nucleotide translocator (ANT), Bax, and cyclophilin D (CD) are components of the mtPTP. The release of cytochrome c (CytC) in the cytoplasm activates apoptotic protease activating factor (Apaf-1) and subsequently caspases that eventually lead to proteolysis and apoptosis.

response during infection by deciding on the fate of the affected cells, and adjust the oxidative stress response for cytoprotective and signaling purposes.

Mitochondria as Targets for Stress Hormones and Mediators

Several important mediators of the stress response, namely hormones (glucocorticoids, GCs, and catecholamines), immune mediators (cytokines), and heat shock proteins (Hsps), among other factors, exert multiple effects on mitochondrial biogenesis, metabolism, ROS generation, and apoptosis.

Glucocorticoids

GCs play a major role in survival during stress. Increased cortisol levels raise the resting energy expenditure and metabolic rate by as much as 200% during acute stress. Despite the significant hyperglycemia, the observed decrease in the respiratory quotient (the ratio of the carbon dioxide produced to oxygen consumed) suggests that a significant portion of that energy comes from the oxidation of lipids. GCs also stimulate protein degradation in muscle and lymphoid tissue to provide amino acids for gluconeogenesis in the liver.

In addition to the well-described molecular mechanism of GC action, via binding to glucocorticoid receptors (GRs) in the cytoplasm, and subsequent translocation of the GC-GR complexes to the nucleus for the trans-activation or -repression of nuclear gene targets, nongenomic mechanisms, such as the activation of phosphatidylinositol 3 (PI3) kinase and other cytosolic signaling pathways are increasingly recognized. Mitochondrial function can be affected directly or indirectly through both these mechanisms. Furthermore, GRs were recently described inside the mitochondria of several cell types, and mtDNA contains nucleotide sequences that are highly similar to the nuclear DNA GR binding consensus sequences. Thus, it is possible that GCs stimulate mtDNA transcription by direct interaction via the mitochondrial GR.

Indeed dexamethasone, a synthetic GC, stimulates the transcription of mtDNA-encoded genes (COX II and cyt-b) in hepatoma cell lines; a rise of mitochondrial transcripts (COX I, II, and III) was also found in colon epithelium of rats after dexamethasone injection. A recent study of rat skeletal muscle suggests an increase in mitochondrial biogenesis, mtDNA transcription (COX II and III), and cytochrome c activity after treatment with dexamethasone for 3 days. The mechanism underlying the observed transcriptional induction remains unclear because no increases in

mitochondrial transcription factor A (Tfam) or nuclear respiratory factor (NRF-1 and 2) have been observed in these studies.

Several studies suggest the modulation of mitochondrial metabolic activities by GCs in a biphasic mode. Low doses of or early exposure to GCs induce mitochondrial biogenesis and the enzymatic activity of selected subunits of the respiratory chain complexes, whereas higher doses or prolonged exposure cause increased ROS generation and respiratory chain dysfunction. Those effects depend on the target organ and developmental stage of the organism. Our recent transcriptional screening for mitochondrial targets of GC action in the human skeletal muscle revealed the induction of mitochondrial translation and energy metabolism-related genes after a shorter exposure to dexamethasone and the down-regulation of genes involved in lipid metabolism, induction of ROS, and apoptosis after prolonged exposure.

Clinical studies of the effects of chronic exposure to GCs have shown increased lactate production following aerobic exercise (suggesting anaerobic glycolysis due to mitochondrial dysfunction), decreased complex-I activity and oxidative damage-induced alterations in skeletal muscle mtDNA and nuclear DNA after long-term GC treatment. Moreover clinical symptoms resembling mitochondrial disorders, such as external ophthalmoplegia, have also been described under such conditions.

ROS generation follows exposure to GCs in different tissues. Hydrogen peroxide is described as a potent mediator of glucocorticoid effects in skeletal muscle cells, stimulating myogenesis and myotube formation at low doses while inducing apoptosis and cell death at higher concentrations or after prolonged treatment. GCs increase oxidative stress damage to neurons, in part by increasing glutamate and calcium and in part by decreasing antioxidant enzymes, specifically in the hippocampal neurons. Hydrogen peroxide is also synthesized in large amounts by neutrophils and macrophages during the oxidative burst, when it aids in destroying engulfed microorganisms. Moreover, various apoptotic stimuli, such as tumor necrosis factor (TNF)- α , have been shown to increase the levels of mitochondria-derived hydrogen peroxide. The increased production of hydrogen peroxide or superoxide after prolonged exposure to GCs in various tissues leads to mtDNA oxidative damage and apoptosis.

Interestingly, antioxidant compounds (i.e., N-acetyl-cysteine, vitamin C, and vitamin E) and enzymes (i.e., SOD and catalase) can prevent those adverse effects in skeletal muscle and vascular endothelium, both *in vitro* and *in vivo*. Those findings suggest a

primary pathogenetic role of ROS in GC-induced tissue damage.

GCs are important regulators of T-cell growth, differentiation, and apoptosis. Their profound anti-inflammatory and immunosuppressive activities place synthetic GCs among the most commonly prescribed drugs worldwide, especially for the treatment of hematological malignancies, such as leukemia and lymphoma.

The current evidence suggests that GC-induced cell death is linked to the *de novo* synthesis of one or several gene products. Genes described as being up- or downregulated during GC-induced apoptosis include c-myc139 and T-cell-death-associated gene (TDAG)8, Bcl-x_L, IκB, GILZ, and GITR. Interaction with other transcription factors, such as nuclear factor (NF)-κB, activator protein (AP)-1, nuclear factor of activated T cells (NF-AT), cAMP response element binding protein (CREB), and signal transducer and activator of transcription (STAT)5 is also critical. On the other hand, GC binding to its receptor rapidly induces ceramide generation via the activation of the phosphatidyl-inositol-specific phospholipase C (PIPLC) and acidic sphingomyelinase (aSMase) enzymes, which is necessary for caspase 9 activation.

A vast amount of data indicates that the disruption of mitochondrial integrity is a prerequisite for GC-induced apoptosis. This is tightly regulated by the balance between pro-(the BH-3 only: Bim, Bid, and Bad) and anti-apoptotic (Bcl-2 and Bcl-x_L) Bcl-2 family members, resulting in the activation of Bax and Bak, the opening of the mtPTP in the outer mitochondrial membrane, and the release of cytochrome c. The multimeric complex containing cytochrome c, caspase 9, and Apaf-1 (called the apoptosome) activates caspase 3, which finally leads to apoptosis.

Catecholamines

Catecholamines are the primary mediators of the fight-or-flight response. Norepinephrine is the major neurotransmitter in the peripheral sympathetic nervous system, whereas epinephrine is the primary hormone secreted by the adrenal medulla. The release of both is increased during stress. Their effect on target tissues is mediated by 6α- and 3β-adrenoceptor subtypes.

Epinephrine affects mitochondrial metabolism mainly by mobilizing body energy stores to augment the availability of substrates for oxidation. Epinephrine stimulates both lipolysis and glycogenolysis (most notably in the skeletal muscle) by interacting with the β₁- and β₂-adrenoceptors. It also acts through the G_s-protein-coupled β₂-adrenergic receptor to

stimulate adenylyl cyclase activity and cAMP production, leading to the activation of the protein kinase A (PKA). This signaling pathway is responsible for the modulation of numerous processes and notably gene transcription through the phosphorylation of CREB.

Another major mitochondria-targeted effect of catecholamines is the enhancement of thermogenesis, which occurs primarily in the brown adipose tissue, which is highly innervated by sympathetic nerve terminals. During this process, energy derived from the oxidation of fuel substrates is dissipated as heat rather than being stored as ATP (uncoupling). Uncoupling occurs via β₃-adrenergic stimulation and the downstream regulation of specific uncoupling proteins (UCPs). The inner mitochondrial membrane in brown adipose tissue contains UCP-1. Other known UCPs include UCP-2, which is widely expressed, and UCP-3, which seems to be expressed primarily in the skeletal muscles, heart, and brown adipose tissue, which are controlled by thyroid hormone and leptin. Several pharmaceutical companies have worked on the development of β₃-adrenoceptor agonists to stimulate UCP-1 in brown adipocytes and increase energy expenditure in obese patients. Unfortunately, these drugs often have adverse side effects on the cardiovascular system and poor bioavailability.

The peroxisome proliferator-activated receptor γ co-activator (PGC-1)α controls adaptive thermogenesis in adipose tissue and skeletal muscle by stimulating mitochondrial biogenesis and oxidative phosphorylation. Stress, fasting, and exercise activate the PGC-1α through different signaling pathways, including β-adrenergic (β₃/cAMP), Ca²⁺-dependent, NO and mitogen-activated protein kinase (MAPK) pathways. PGC-1α, in turn, activates the expression of nuclear respiratory factor (NRF)-1 and -2, estrogen-related receptor (ERR)-α, and PPAR-α, which regulate the mitochondrial biogenesis and fatty acid oxidation pathways. NRF-1 and NRF-2 regulate the expression of Tfam, a nuclear-encoded transcription factor that binds regulatory sites on mtDNA and is essential for the replication and transcription of the mitochondrial genome. Furthermore, NRF-1 and NRF-2 regulate the expression of nuclear genes encoding respiratory chain subunits and other proteins required for mitochondrial function.

Cytokines

The overproduction of inflammatory cytokines, with the subsequent induction of NO, platelet activating factor, and prostaglandins have been implicated in the changes leading to the multiorgan failure associated with, for example, sepsis. This pro-inflammatory

state is described as the systemic inflammatory response syndrome (SIRS).

Cytokines like TNF- α can induce apoptosis in several cell populations, whereas other cytokines such as interleukin (IL)-1 and IL-6 often inhibit apoptosis. The release of these pro-inflammatory cytokines is a major component of the stress response to infection, ischemia/reperfusion injury, and trauma. The intracellular signaling pathways involved in TNF- α - or Fas ligand (FasL)-induced apoptosis appear to be caspase 3-dependent.

Exciting new findings describing mitochondria localization of functional proteins required for the activation of interferon 1 expression in response to viral infections highlight the central role played by mitochondria in balancing the host immune response to viral infection and possibly to other pathogens.

Heat Shock Proteins

Hsps constitute a highly conserved and functionally interactive network of chaperones, some of which are constitutively expressed and others that are induced by various environmental, physical, and chemical stressors that challenge the intracellular milieu and cause damage to DNA and proteins. Chaperones are proteins that disaggregate, refold, and renature misfolded proteins, thus protecting the cell from the damage imposed by stressful stimuli. If the stressor is heat shock, the induced chaperones are called heat shock proteins.

The induction of Hsps following cellular damage can prevent apoptosis. Hsp27, 70, and 90 have been implicated in protection against apoptosis induced by multiple signals, such as heat shock, nutrient withdrawal, ROS, endoplasmic reticulum stress, proteasome inhibition, UV radiation, and chemotherapy-induced DNA damage. Hsps promote cell survival by preventing mitochondrial outer-membrane permeabilization and subsequent cytochrome c release, caspase activation, and apoptosome assembly. In addition, Hsp60 and Hsp70 interact with the outer mitochondrial membrane-specific translocase (TOM), contributing to the importing of nuclear-encoded proteins into the mitochondrial matrix. It may well be that the beneficial effects of fever initially described by Hippocrates actually relate to increased Hsp expression and their protective effects.

Mitochondrial Adaptation and Dysfunction during Acute and Chronic Stress

The acute response to stressful stimuli is characterized by the massive release of stress mediators. Their

coordinated interaction aims at (1) maintaining the effective circulation (and hence oxygenation) of the brain, cardiac muscle, and skeletal muscle; (2) increasing energy generation by the release of substrates (i.e., glucose, fatty acids, and amino acids) from body fuel storages (i.e., the liver, adipose tissue, and skeletal muscle); and (3) optimizing ATP availability to vital tissues at the expense of others (i.e., the gonads and gastrointestinal tract).

Tissue oxygen consumption and total energy expenditure are profoundly increased during the initial phase of the acute stress response. Rapid elevations in GC and catecholamine levels stimulate mitochondrial oxygen consumption and ATP formation to sustain the substantial hypermetabolic state. Mitochondria respond by modulating the expression and activity of certain OXPHOS subunits, or by increasing significantly their size and numbers (mitogenesis). Yet the exact mechanisms by which mitochondria fulfill this increased energy demand remain poorly understood.

Although the initial burst in mitochondrial function is necessary for survival, excessive mitochondrial activation during critical illness, such as trauma, surgery, or sepsis, can be detrimental to the cell ([Figure 2](#)). Stress-induced hyperglycemia, resulting from increased glucose mobilization and insulin resistance, is a common finding in these conditions, and its severity correlates strongly with increased mortality risk. Several lines of evidence support the concept that mitochondrial impairment, caused by glucose toxicity, is the underlying mechanism for the disturbance in oxygen use, called cytopathic hypoxia. Cellular glucose overload can exceed the capacity of the respiratory chain leading to (1) increased ROS generation and oxidative damage to the mitochondrial complexes I and IV, and the antioxidant enzyme MnSOD; (2) the impaired detoxification of superoxide and other free radicals; (3) the shunting of glucose into toxic pathways (i.e., polyol and hexosamine); and (4) increased apoptosis. Recent studies show that strict glycemic control with insulin therapy can prevent and correct mitochondrial ultrastructural and functional abnormalities in surgical intensive-care patients.

Prolonged physical and, interestingly, psychological stress can induce oxidative damage. In a recent study, psychological stress, both perceived stress and chronicity of stress, in mothers of children suffering from chronic illness was shown to be associated with increased markers of oxidative stress in their circulation (F₂-isoprostanes levels and isoprostanes/vitamin E ratio), in addition to telomere shortening. Telomeres are DNA sequences necessary for DNA replication, which shorten at cell division at a rate related to the levels of oxidative stress. Once telomeres are

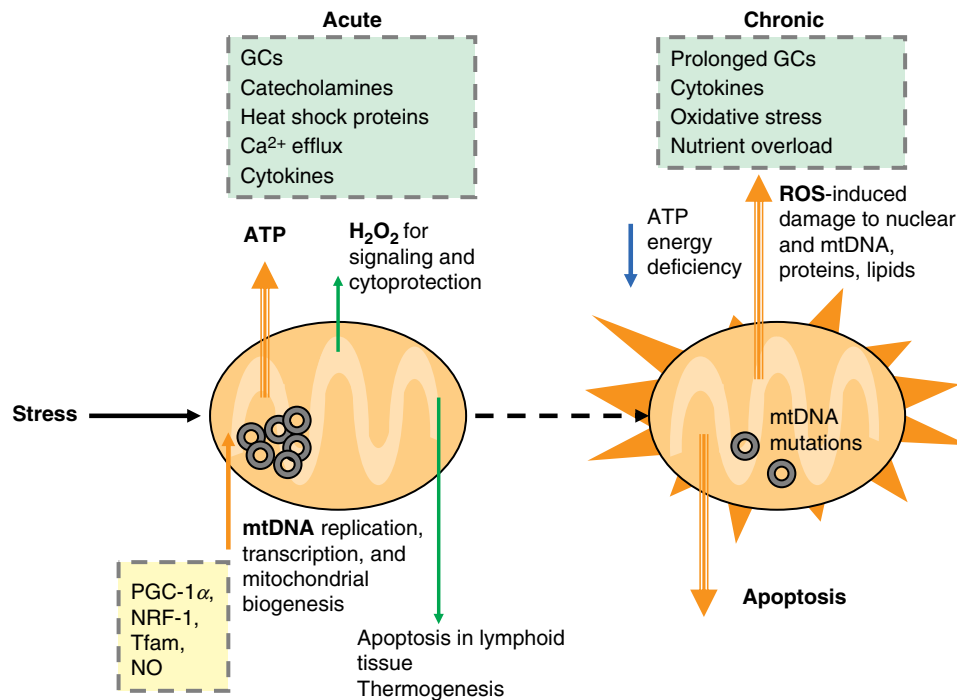


Figure 2 Mitochondrial stress response under acute and chronic stress. Early exposure to glucocorticoids (GCs) or other stressful stimuli and their respective mediators induce mitochondrial biogenesis and the enzymatic activity of selected subunits of the respiratory chain complexes to meet the increased energy demands of the cell. Controlled production of reactive oxygen species (ROS), apoptosis, and thermogenesis serve as protective mechanisms against infection or other cell damage. Prolonged stress can exceed mitochondrial reserves, leading to respiratory chain dysfunction and decreased ATP production, increased ROS generation, and significant apoptosis and/or necrosis. mtDNA, mitochondrial DNA; NRF, nuclear respiratory factor; PGC, peroxisome proliferator-activated receptor γ co-activator; Tfam, mitochondrial transcription factor A.

shortened to a critical length, cells enter a stage of replicative senescence; hence, telomere length is a determinant of cell senescence and longevity. Although not examined in the study, it is expected that changes in the mtDNA, which is far more prone to oxidative damage, could also arise under those conditions. These findings have implications for the molecular mechanisms at the cellular level by which stress may promote the earlier onset of age-related diseases.

Psychological stress can lead to oxidative stress, but the opposite seems to hold true also. An intriguing recent study in an animal model showed a strong positive correlation between the expression levels of two antioxidant enzymes, glyoxalase and glutathione reductase 1, and anxiety levels. To determine whether those genes have a causal role in the genesis of anxiety, genetic manipulation was performed to induce or inhibit their activity. The local overexpression of these genes in the mouse brain resulted in increased anxiety-like behavior that was decreased with their inhibition. The authors propose oxidative stress as a new and unsuspected player in the complex control of anxious behavior, which, if exaggerated, in humans can be expressed as panic disorder, obsessive-compulsive disorder, posttraumatic stress disorder,

social and other specific phobias, and generalized anxiety disorder.

The metabolic syndrome (central/visceral obesity, insulin resistance, dyslipidemia, hypertension, pro-inflammatory, and prothrombotic syndrome) has been associated with hyperactivity of the hypothalamus-pituitary-adrenal axis. The pathogenesis remains unknown, but both genetic and environmental factors have an important share. It is proposed that underlying genetic polymorphisms or alterations in nuclear and, potentially more important, in mitochondrial DNA, in combination with exposure to environmental stresses such as excessive caloric intake, exhaust mitochondrial reserves and lead to the imbalance and failure of energy metabolism.

Support for this theory comes from numerous animal studies, across several species, which demonstrate the strong beneficial effects of calorie restriction on life expectancy. Reduced workload on the respiratory chain leads to improved energy output and decreased oxidative stress and subsequent damage. In this regard, a study that selected rats for their low versus high intrinsic capacity to run on a treadmill over 11 generations – hence, rats with low-versus high-power mitochondria – demonstrated that

cardiovascular risk factors, such as obesity, hyperglycemia, hyperlipidemia, hypertension, and insulin resistance, segregate with low aerobic capacity and reduced expression of genes required for mitochondrial biogenesis and oxidative phosphorylation.

Defects in mitochondrial fatty acid oxidation can lead to the intracellular accumulation of fatty acid metabolites that disrupt insulin signaling and possibly also secretion from the pancreatic β -cells, leading to insulin resistance and ultimately failure. Defects of mitochondrial oxidative phosphorylation and mitochondrial copy number are indeed observed in the skeletal muscle of elderly subjects and of healthy relatives of individuals with type 2 diabetes. Moreover, alterations in mtDNA, specific mutations in mitochondrial tRNA, or decreased expression of genes regulating mitochondrial function, such as PGC-1 α , are all associated with type 2 diabetes.

Transgenic animal models, in which antioxidant enzymes are overexpressed, such as manganese superoxide dismutase (MnSOD) or Cu,ZnSOD in *Drosophila*, exhibit a significantly extended life span. To the contrary, in knockouts for MnSOD, a 100% increased incidence of cancer was observed, whereas knockouts for mtDNA polymerase γ (POLG), the enzyme responsible for mtDNA proofreading, experience a significant acceleration in their aging process, with cardiomegaly, lipodystrophy, osteoporosis, anemia, and alopecia – but, interestingly, an absence of oxidative stress. The later study and others challenge the free-radical theory of aging and pose further questions in the fascinating field of mitochondrial pathology.

Seeking Mitochondrial Anti-stress Treatments

Therapeutic options for mitochondrial dysfunction are currently very limited. Strategies are being developed for mitochondria-targeted therapeutics, using their distinct biophysical and biochemical properties. The high negative potential of the mitochondrial matrix allows for selective targeting of positively charged compounds. Furthermore, the presence of specific transporter-dependent delivery for pro-drugs that can be subsequently activated by mitochondrial enzymes enable the design of highly specific agents.

The most effective of such attempts to date has been the development of mitochondria-targeted antioxidants. Mito-vitamin E and mito-coenzyme Q are coupled to a triphenylphosphonium cation, which accumulates several hundredfold in the mitochondrial matrix due to the large membrane potential, which is negative inside. First-line evidence from *in vitro* applications and initial animal studies on cardiac

ischemia-reperfusion injury have been very promising. Further support for mitochondria-targeted antioxidants comes from a transgenic mouse model, in which the overexpression of the antioxidant enzyme catalase targeted specifically to mitochondria (vs. to peroxisomes or the nucleus) significantly reduced oxidative damage in both nuclear and mitochondrial DNA, resulting in increased murine life span and delayed cardiac pathology and cataract development.

We can envision a rapid increase in the efforts to exploit the unique properties of mitochondria for the development of selective therapeutic agents aimed at increasing mitochondrial biogenesis and energy output while concomitantly minimizing mitochondrial Ca²⁺ overload, ROS production, and oxidative stress.

Acknowledgments

This work was supported by the Intramural Research Programs of the National Human Genome Research Institute and by the National Institute of Mental Health. The authors have no conflicts of interest related to this work.

Further Reading

- Balaban, R. S., Nemoto, S. and Finkel, T. (2005). Mitochondria, oxidants, and aging. *Cell* 120(4), 483–495.
- Chinopoulos, C. and Adam-Vizi, V. (2006). Calcium, mitochondria and oxidative stress in neuronal pathology: novel aspects of an enduring theme. *FEBS Journal* 273(3), 433–450.
- Chrousos, G. P. (1998). Stressors, stress, and neuroendocrine integration of the adaptive response: the 1997 Hans Selye memorial lecture. *Annals of the New York Academy of Sciences* 851, 311–335.
- Chrousos, G. P. (2004). The HPA axis and the stress response. *Endocrine Research* 26(4), 513–514.
- Epel, E. S., Blackburn, E. H., Lin, J., et al. (2004). Accelerated telomere shortening in response to life stress. *Proceedings of the National Academy of Sciences USA* 101(49), 17312–17315.
- Goldenthal, M. J. and Marin-Garcia, J. (2004). Mitochondrial signaling pathways: a receiver/integrator organelle. *Molecular and Cellular Biochemistry* 262(1–2), 1–16.
- Jezek, P. and Hlavata, L. (2005). Mitochondria in homeostasis of reactive oxygen species in cell, tissues, and organism. *International Journal of Biochemistry and Cell Biology* 37(12), 2478–2503.
- Kino, T., Charmandari, E. and Chrousos, G. P. (eds.) (2004). *Special issue on glucocorticoid action: basic and clinical implications. Annals of the New York Academy of Sciences*. 1024.
- Lee, S. H. K. D., Tanaka, M. and Wei, Y. (eds.) (2004). *Special issue on mitochondrial pathogenesis: from genes*

- and apoptosis to aging and disease. *Annals of the New York Academy of Sciences*. 1011.
- Lemasters, J. J. and Nieminem, A. L. (2001). *Mitochondria in pathogenesis*. New York: Kluwer Academic/Plenum.
- Lowell, B. B. and Shulman, G. I. (2005). Mitochondrial dysfunction and type 2 diabetes. *Science* 307(5708), 384–387.
- Manoli, I., Le, H., Alesci, S., et al. (2005). Monoamine oxidase-A is a major target gene for glucocorticoids in human skeletal muscle cells. *FASEB Journal* 19(10), 1359–1361.
- Oberholzer, C., Oberholzer, A., Clare-Salzler, M., et al. (2001). Apoptosis in sepsis: a new target for therapeutic exploration. *FASEB Journal* 15(6), 879–892.
- Pacak, K., Aguilera, G., Sabban, E., et al. (eds.) (2004). *Special issue on stress: current neuroendocrine and genetic approaches*. *Annals of the New York Academy of Sciences*. 1018.
- Psarra, A. M., Solakidi, S. and Sekeris, C. E. (2006). The mitochondrion as a primary site of action of steroid and thyroid hormones: presence and action of steroid and thyroid hormone receptors in mitochondria of animal cells. *Molecular and Cellular Endocrinology* 246(1–2), 21–33.
- Scheffler, I. E. (1999). *Mitochondria*. New York: Wiley-Liss.
- Sheu, S. S., Nauduri, D. and Anders, M. W. (2006). Targeting antioxidants to mitochondria: a new therapeutic direction. *Biochimica et Biophysica Acta* 1762(2), 256–265.
- Wallace, D. C. (2005). A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a dawn for evolutionary medicine. *Annual Review of Genetics* 39, 359–407.

Monoamine Oxidase

P Huez-Diaz and I W Craig

King's College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

Stress and the Cycle of Violence

Genetic Factors in Aggression and Violence

MAOA as a Candidate Gene

The Role of Stress and Abusive Upbringing in Aggression and Violence

Conclusion

Glossary

<i>Allele</i>	One of two or more possible forms of a gene. An individual may possess two identical, or two different, alleles, except for those of the genes on the X chromosome in males because they possess only a single copy of this chromosome.
<i>Antisocial behavior (ASB)</i>	A pervasive pattern of disregard for and violation of the rights of others occurring since age 15 years old.
<i>Cycle of violence</i>	The observed relationship between violent parenting and violence in the parents' offspring.
<i>Major effect gene</i>	A gene whose influence can be detected without the need to control for environmental factors.
<i>Null mutation</i>	A mutation that completely eliminates the function of the gene.
<i>Promoter</i>	The section of DNA that precedes a gene and controls its activity.

Stress and the Cycle of Violence

A common observation in studies of ASB and aggression in humans is the cycle of violence, in which there is a tendency for individuals who were exposed to abusive parents and/or difficult childhoods to manifest antisocial personalities in adulthood. Not all individuals exposed to abusive childhoods, however, develop antisocial and violent tendencies. The challenge is, therefore, to establish the extent to which genetic predispositions and the extent to which stressful environments are responsible for eliciting the violent behavior. Recent research suggests that there is an interaction between the two that may underpin this repeating pattern.

Genetic Factors in Aggression and Violence

Results from more than 50 behavioral genetic studies indicate that, on average, approximately 41% of the variance of ASB is due to genetic factors, approximately 16% is due to shared environmental factors and approximately 43% is due to nonshared environmental factors. The same overall pattern appears to hold for violence. Epidemiological studies have also consistently shown that the risk of aggression is greater in males than in females, and this raises the possibility that genetic and environmental influences contribute differentially to the risk of aggressive behavior depending on gender. Twin studies indicate that common environmental factors, sex-specific genes,

and the interaction between the two play an important role in increasing liability toward aggressive behavior in males, even though females have a higher heritability and males a higher common environmental liability.

Evidence for a direct role of testosterone levels, which could be genetically determined, in promoting aggression in humans is not clear-cut. This leaves open the possible influence of other specific genetic factors that may contribute to increased aggression in males; recent interest has focused on the monoamine oxidase A gene, *MAOA*, in this regard.

***MAOA* as a Candidate Gene**

The main role for the monoamine oxidase (*MAOA*) enzyme is thought to be in degrading serotonin following its reuptake from the synaptic cleft, although it is also capable of degrading both norepinephrine and dopamine. It therefore plays a key role in the regulation of nerve transmission, and alterations in its activity produced by pharmacological intervention or through genetic variants are likely to have profound effects on behavior. Indeed, drugs inhibiting its activity have long been employed in the treatment of behavioral disorders, particularly depression.

A wide range of genetic variants of the *MAOA* gene exists in the general population. One of the most significant is a DNA motif localized in the promoter 1.2 kb upstream from the *MAOA* coding region, comprising a 30-bp repeat existing in 3, 3.5, 4, or 5 copies (Figure 1). It has been shown that the copy number has a significant effect on the level of the gene's expression. The two predominant forms observed in the population are those with three or four repeats of the motif, with the three-copy version having reduced transcriptional activity compared to the four-copy allele. The gene is located on the X chromosome, which means that females have two copies but males have only one. Any defect in the gene in males

is therefore exposed because they are unaffected by the status of a second copy.

A possible direct role for *MAOA* in predisposing violence, ASB, and conduct disorders was suggested by studies of males from a Dutch pedigree who exhibited a complex behavioral syndrome (including impulsive aggression). All affected individuals suffered from a mutation in the *MAOA* gene that resulted in zero activity of the enzyme. A role for *MAOA* deficiency in promoting aggression is further supported by studies on mutant mice that had deletions of portions of their monoamine oxidase gene, resulting in the lack of enzyme activity in the brain and liver. Behavioral studies on the adult male mice indicated heightened aggression in response to intruders and also increased inappropriate courtship behaviors. These have been the only reports to date of the phenotypic consequences of null mutations. The existence of high- and low-activity *MAOA* alleles in the human population, however, raises the intriguing possibility that lower levels of the enzyme may predispose male individuals to ASB and/or violent behavior.

There have been a variety of reports suggesting a role for variants at the *MAOA* locus in the context of violence and/or ASB, particularly in relation to subgroups of males with alcoholism and attention deficit and hyperactivity disorders. Generally, it appears that individuals with the low-activity allele are at risk, but there is no overall consensus concerning its possible status as a major effect gene. Nevertheless, a significant and intriguing interaction between stressful environments and low-activity variants of *MAOA* in promoting violence and ASB has recently emerged.

The Role of Stress and Abusive Upbringing in Aggression and Violence

Given the observed relationship between abusive environments and antisocial outcomes, the *MAOA*

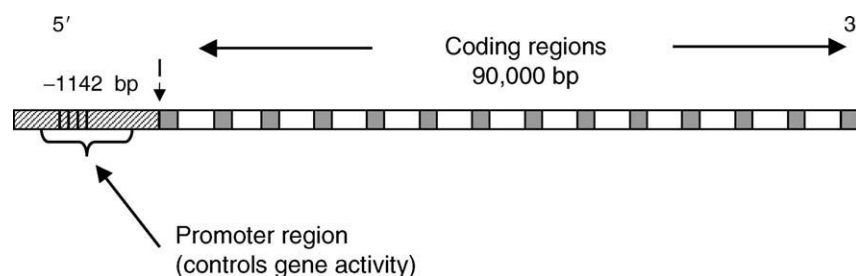


Figure 1 Representation of the monoamine oxidase gene (*MAOA*). The promoter region (in diagonal lines) controls the levels of gene activity and contains a 30-bp repeat (often referred to as a variable-number tandem repeat, VNTR). Different copy numbers of the basic motif have different effects on transcription. Three- and four-copy motifs are the most common in the population, with the former conferring lower activity than the latter. Rare alleles with 2-, 3.5-, and 5-copy motifs are also observed. I indicates position of transcription initiation (RNA synthesis). Exons (coding regions) are in gray (not to scale).

gene became a candidate for a genetic predisposing factor that may interact with stress induced by adverse early rearing experiences to promote ASB. Caspi and associates, in 2002, were the first to study the interactions between environment and MAOA activity variants in the etiology of ASB in males by investigating the patterns of antisocial outcomes in a range of male children who had been maltreated. A broad range of measures for ASB was followed in the males from a longitudinal cohort of approximately 800 individuals who were age 26 at the time of data collection. The information available relevant to ASB included the commission of violent crimes, a personal disposition toward violence, and adolescent conduct disorder (assessed according to *Diagnostic and Statistical Manual*, 4th edn., criteria). Whichever measure of antisocial outcomes was examined, a significant association between maltreatment and ASB conditional on the individual's MAOA genotype was observed. Whereas maltreated males with low-activity alleles were significantly at risk of conviction for violent crime or of exhibiting other ASB, those with high-activity alleles were not. An increasing tendency to antisocial acts by those with more extreme childhood maltreatment was observed for males with the low-activity allele, but not for those protected by a high-activity allele.

Similar observations have since been reported in studies employing data from white male twins from the community-based longitudinal Virginia Twin Study for Adolescent Behavioural Development. Several other studies have been completed, and an analysis of all the information currently available is broadly supportive of an interaction between variants of this gene and stressful environments in predisposing ASB.

It is of interest that an analogous promoter variant system of this X chromosome gene has been discovered in rhesus macaque monkeys. A variable-repeat motif of 18 bp has been identified and shown to regulate the expression of MAOA. Of particular relevance to the putative role of functional MAOA variants in humans is that male macaques with the low-activity form were found to be more aggressive in competing for food.

The physiological relationship between the stress caused by abusive upbringing and the apparent protection provided by the high-activity allele for MAOA is not yet clear. Stress is known to elicit an increase in levels of transcription factors, such as c-Fos. These factors, in turn, may upregulate genes encoding components of neurotransmitter pathways, and it is possible that the apparent risk conferred by the low-activity allele of MAOA reflects a relative inability to respond to the effects of stress arising from maltreatment.

Conclusion

The lack of a functional copy of the MAOA gene in humans and mice results in ASB and aggression, and the less-functional fundamental variation of the gene appears to have related behavioral consequences. In particular, there is a complex interaction between the response to chronic stress and functional variants of the gene and the likelihood of individuals developing ASB or violent behavior. Longitudinal studies of the effects of childhood maltreatment and its interaction in males with high- or low-activity genetic variants of the MAOA gene suggest that the high-activity allele may confer protection in males because carriers are much less likely to develop ASB and aggressive behavior than are those with low-activity alleles.

See Also the Following Articles

Aggression; Domestic Violence; Gene Environment Interactions in Early Development.

Further Reading

- Brunner, H. G., Nelen, M., Breakefield, X. O., et al. (1993). Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A. *Science* 262, 578–580.
- Cases, O., Seif, I., Grimsby, J., et al. (1995). Aggressive behaviour and altered amounts of brain serotonin and norepinephrine in mice lacking MAOA. *Science* 268, 1763–1766.
- Caspi, A., McClay, J., Moffitt, T. E., et al. (2002). Role of genotype in the cycle of violence in maltreated children. *Science* 297, 851–854.
- Craig, I. W. (2005). The role of monoamine oxidase A, MAOA, in the aetiology of antisocial behaviour: the importance of gene environment interactions. In: Bock, G. & Goode, J. (Eds.) *Molecular mechanisms influencing aggressive behaviours* (Novartis Foundation Symposium 268), pp. 227–237. New York: John Wiley & Sons.
- Deckert, J., Catalano, M., Sygailo, Y. V., et al. (1999). Excess of high activity monoamine oxidase A gene promoter alleles in female patients with panic disorder. *Human Molecular Genetics* 8, 621–624.
- Denney, R. M., Koch, H. and Craig, I. W. (1999). Association between monoamine oxidase A activity in human male skin fibroblasts and the genotype of the MAO promoter-associated variable number tandem repeat. *Human Genetics* 105, 541–551.
- Foley, D. L., Eaves, L. J., Wormley, B., et al. (2004). Childhood adversity, monoamine oxidase A genotype, and risk for conduct disorder. *Archives of General Psychiatry* 61, 738–744.
- Kim-Cohen, J., Caspi, A., Taylor, A., et al. (in press). MAOA, early adversity, and gene-environment interaction

- predicting children's mental health: new evidence and a meta-analysis. *Molecular Psychiatry*.
- Moffitt, T. E., Caspi, A., Rutter, M. and Silva, P. A. (2001). *Sex differences in antisocial behaviour: conduct disorder, delinquency and violence in the Dunedin Longitudinal Study*. Cambridge, UK: Cambridge University Press.
- Newman, T. K., Syagailo, Y. V., Barr, C. S., et al. (2005). Monoamine oxidase A gene promoter variation and rearing experiences influences aggressive behaviour in rhesus monkeys. *Biological Psychiatry* 57, 167–172.
- Sabol, S. Z., Hu, S. and Hamer, D. (1998). A functional polymorphism in the monoamine oxidase A gene promoter. *Human Genetics* 103, 273–279.
- Vierikko, E., Pulkkinen, L., Kaprio, J., et al. (2003). Sex differences in genetic and environmental effects on aggression. *Aggressive Behaviour* 29, 55–68.

Relevant Website

<http://www.ojp.usdoj.gov/bis>.

Motor Vehicle Accidents, Stress Effects of

T C Buckley

Boston VA Medical Center and Boston University
School of Medicine, Boston, MA, USA

E B Blanchard

University at Albany–State University of New York,
Albany, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2,
pp 777–780, © 2000, Elsevier Inc.

Introduction

Psychological Disorder Following MVAs

What Predicts Psychopathology Following MVAs?

What Predicts Remission of Psychological Disorder?

Conclusion

Glossary

<i>Comorbidity</i>	The simultaneous presence of two disorders in the same individual.
<i>Longitudinal studies</i>	Studies that assess individuals at multiple points in time in a prospective manner.
<i>Prevalence</i>	Total number of cases of a particular disorder over a given period of time.
<i>Psycho-pathology</i>	Difficulty with psychological functioning (e.g., depression, anxiety, alcohol abuse).
<i>Spontaneous remission</i>	The alleviation of signs and symptoms of a particular disorder in the absence of formal treatment.

Introduction

The negative impact that motor vehicle accidents (MVAs) have on society is staggering. Total property damage, health-care expenditure, and hours lost from

employment have a cost to the U.S. economy of over 150 billion dollars each year. Moreover, the individual psychosocial consequences of MVAs can be very debilitating. Motor vehicle accident survivors often deal with complicated issues for the rest of their lifetime subsequent to an accident that may have lasted a total of 3 s. It is not uncommon for MVA survivors to face such complicated issues as bereavement, chronic pain, permanent disability, financial problems, loss of employment, depression, a variety of anxiety problems, and legal difficulties. Space limitations prevent us from covering all of these topics in detail. This article does not cover such interesting topics as the role that litigation plays in posttrauma recovery, the experimental investigation of posttrauma psychiatric symptoms, and the role of psychosocial treatments for MVA survivors. This article summarizes the literature on the chief behavioral problems that follow MVAs: specific phobia, posttraumatic stress disorder (PTSD), and major depression.

Psychological Disorder Following MVAs

Specific Phobia

A specific phobia is a behavioral disorder in which individuals manifest an excessive and unrealistic fear of a situation or object. Exposure to the feared situation will invariably provoke an anxiety response (e.g., situationally bound panic attack) from the phobic individual. Phobic individuals will deliberately avoid their feared situation or else endure the situation with great distress. Phobic fears can be seen in individuals with seemingly no traumatic history with the feared object or situation. Phobic fears can also develop via classical conditioning when aversive conditions are paired with the phobic object (which would not have been the target of phobic behavior at the time

of conditioning). MVAs can serve as a strong, single-trial conditioning event that leads to a phobic fear of driving-related situations in some individuals.

The experience of a MVA can be very frightening. Quite often, individuals will report that they experienced an extreme sense of life threat during the course of their accident. Thus, MVAs can serve as a single-trial conditioning event that will come to elicit anxiety responses (e.g., panic attacks) in future driving-related situations. For some individuals, the conditioned response is so severe that driving becomes exceedingly difficult. A small set of MVA survivors will discontinue driving following their MVAs, whereas others restrict their driving behavior relative to pre-MVA driving behavior.

The prevalence of a specific phobia of driving-related situations following MVAs is difficult to estimate. Some investigators have adopted a strict definition of driving phobia, wherein a complete elimination of driving or severe restriction of driving (i.e., only drives to work) defines phobic driving behavior. Others have adopted less restrictive criteria, defining driving phobias (sometimes termed driving reluctance) as a change in postaccident driving behavior characterized by increased anxiety related to some or all aspects of travel and a reduction in total number of miles traveled. Estimates of severe driving phobia (i.e., elimination or near-elimination of driving) are about 10–15% of MVA survivors. Estimates of less severe driving phobia or driving reluctance are between 38 and 77%.

Despite the general lack of diagnostic and prevalence data, one thing is clear: MVAs change the driving behavior of MVA survivors greatly. In countries where dependence on a private automobile is great, this can be a costly issue. Fortunately, exposure-based behavioral treatments can be very efficacious ways to treat driving phobia and reluctance to drive.

Posttraumatic Stress Disorder

PTSD is currently classified as an anxiety disorder by the American Psychiatric Association. It is a condition that follows exposure to an event that involves severe threat to one's life or physical integrity (e.g., wartime trauma, rape, MVAs). Following such life-threatening events, some individuals meeting the diagnostic criteria for PTSD will continue to reexperience their traumatic event in one of the following ways: psychological distress upon exposure to trauma-reminiscent cues, physiological arousal upon exposure to trauma-reminiscent cues, intrusive recollections of the trauma, trauma-related dreams that disrupt sleep, and flashbacks of the event. They also will evidence behavioral avoidance of trauma-related stimuli in much the same way that phobic individuals

will, as well as cognitive avoidance (trying not to think about the MVA). Depressive-type symptoms such as a restricted range of emotional experience and a loss of interest in pleasurable activities are also part of the symptomatic criteria. Finally, individuals meeting the diagnostic criteria for PTSD may have chronic hyperarousal, evincing such symptoms as sleep disturbance, exaggerated startle response, and difficulty concentrating.

Given the fear, sense of life threat, and extent of physical injury associated with severe MVAs, it is not surprising that a certain number of MVA survivors will develop PTSD. Prevalence estimates of PTSD among MVA survivors range from 1 to 46%. These discrepancies are due in large part to sample recruiting methods and diagnostic procedures. Self-referred clinical populations yield prevalence estimates of 40% or higher. However, more representative samples of consecutive accident admissions in hospital settings (i.e., emergency rooms) consistently yield prevalence rates near 20%. After reviewing the literature during the course of our work with MVA survivors, it is our impression that the most accurate prevalence estimate of PTSD following accidents is about 20%.

Studies that have examined this prevalence issue have varied in terms of time post-MVA that the patients were assessed. It is clear from longitudinal research that prevalence rates can vary greatly as a function of time postaccident. While many individuals will meet the criteria for PTSD in the first 1–4 months post-MVA, approximately 50% of those cases will spontaneously remit without formal intervention over the course of 6–12 months posttrauma. Thus, for MVA survivors who meet symptomatic criteria for PTSD shortly following trauma, many will experience a remission of symptoms without formal intervention. This spontaneous remission phenomenon is similar to that seen in PTSD populations subsequent to other traumas such as rape and wartime atrocity. What distinguishes those who remit vs. those who do not is covered in a later section of this article.

Major Depressive Disorder

Unipolar major depression is the most common mood disorder. Major depressive episodes are discrete periods of time (usually months to years) when individuals experience a variety of cognitive, affective, and physiological symptoms such as depressed mood, anhedonia, difficulty with concentration, feelings of worthlessness, sleep disturbance, appetite disturbance, psychomotor agitation/retardation, fatigue, and suicidal ideation or attempts.

Individuals often experience major depressive episodes in the months following their MVA. One

empirical question is whether the depression is a consequence of the MVA or whether the presenting individuals had problems with major depression prior to their MVA. Longitudinal studies suggest that a large number of individuals who experience major depressive episodes following MVAs are experiencing depression for the first time in their life. The incidence of depression during the time periods of these studies is greater than one would expect in non-MVA representative population samples during the same period of time. In addition, the onset of episodes is quite often close in time to the occurrence of the MVA. Considering these two findings, it is tempting to attribute causal status to MVAs for inducing major depressive episodes in some individuals. In addition, individuals who have had difficulties with major depression prior to their accident quite often experience another episode of depression following their MVA.

Thus, a large portion of the cases of major depression noted are a direct consequence of the MVA itself. Not surprisingly, major depression is a common comorbid condition with PTSD. The limited amount of research on the impact of comorbid depression with PTSD indicates poorer outcomes in major role functioning in individuals with comorbidity relative to individuals with one condition or the other.

What Predicts Psychopathology Following MVAs?

Phobic avoidance, PTSD, and major depression are more likely to occur post-MVA in individuals who are in more severe accidents. The severity of MVAs has generally been judged in two ways. First, subjective ratings of perceived life threat have been used as an index of severity of accident. Individuals who report a greater fear of death at the time of the MVA generally have more post-MVA psychopathology than those with a lower sense of life threat.

Second, the severity of MVAs has been assessed via more objective methods by quantifying the extent of physical injury incurred during the accident. Some studies have found this variable to have predictive power, in that the more severe the physical injuries, the more likely the individuals were to have behavioral problems. Other studies have failed to find such a relationship. Measurement differences may account for a large part of this variation. Some studies simply dichotomize the presence or absence of injury to predict psychological outcome in their data analyses, whereas other studies have quantified physical injury in a dimensional manner, assessing the range and severity of injuries through the use of instruments with good psychometric properties. Studies that use

these more reliable and valid physical assessment instruments are those that demonstrate the predictive utility of assessing physical injury. It is our impression that the degree of physical injury is a valid predictor of post-MVA psychopathology.

Consistent with research in other areas of trauma (i.e., rape trauma, wartime trauma), a history of previous trauma seems to be predictive of who will develop post-MVA psychopathology and who will not. Those with histories of previous trauma are more likely to have post-MVA psychopathology than those without previous trauma histories. The mechanisms that drive this cumulative trauma phenomenon are not known. Suffice it to say that previous trauma indicates a poor prognosis for survivors of serious MVAs.

Also consistent with the non-MVA trauma literature is the fact that previous psychopathology (preexisting depression, obsessive-compulsive disorder, etc.) indicates a poor prognosis for survivors of MVAs. Preexisting axis I disorders increase the likelihood of trauma-related diagnoses such as PTSD, specific phobia with traumatic onset, and major depression. This could be due to similar predispositions such as biological vulnerability, poor coping skills, and/or a lack of social support systems. It does appear that high levels of perceived social support are protective against the development of post-MVA psychopathology. These findings are entirely consistent with the literature that has assessed the relationship between social support and a variety of stress-related difficulties ranging from behavioral problems to immune functioning.

Gender has consistently been found to be predictive of psychopathology following MVAs. Women develop PTSD, major depression, and specific phobia at greater rates than men following MVAs. It may be that men are likely to cope with their problems by self-medicating with alcohol or drugs and thus present with different psychiatric diagnoses. The epidemiological literature as a whole would suggest that men are more likely to use alcohol and drugs as coping mechanisms.

What Predicts Remission of Psychological Disorder?

At a time when health-care resources are at a premium, a question of great interest to clinical researchers is, among MVA survivors with psychopathology shortly following an MVA, what variables will identify those individuals whose psychiatric symptoms will remit without intervention? Because treatment resources may not be available to all MVA victims (especially with accidents involving large numbers of individuals, i.e., bus accidents), the identification of

predictors of recovery is a very important applied question.

Not surprisingly, the degree of psychiatric symptomatology present shortly after the occurrence of MVA is predictive of remission of symptoms. Individuals with greater numbers of symptoms and more severe symptoms are less likely to show remission of diagnostic status than those who meet diagnostic criteria for disorder in the months post-MVA but have lower levels of symptomatology. The level of initial physical injury is also predictive of the remission of psychiatric symptoms. Those who are injured more severely at the time of MVA are less likely to show remission of psychiatric symptoms over time.

The rate of physical recovery, as assessed by self-report measures, is also predictive of psychiatric symptom remission. Those whose physical injuries recover at a slower rate also show a slower remission of psychiatric symptoms. The causal relationship between these variables is unknown. Some investigators have speculated that physical injury status, especially pain, exacerbates and maintains psychiatric symptoms, whereas others have argued that the presence of psychiatric symptoms makes people more sensitive to pain and thus more likely to inflate self-report scores of pain. This is a particularly vexing dilemma when trying to assess the rate of physical and emotional recovery in MVA survivors because many of their injuries are soft tissue injuries (e.g., cervical and lumbar spine strains), which do not lend themselves to quantifiable measurement other than by self-report methods.

Finally, the presence of behavioral avoidance shortly after trauma signifies a poor prognosis. Individuals who show a high degree of driving reluctance or avoidance shortly after their MVA are less likely to show a remission of psychiatric problems such as PTSD than those who force themselves to drive. It may be the case that self-administered exposure helps break the cycle of anxious responding in the presence of trauma-reminiscent cues that characterize PTSD.

Conclusion

Empirical investigation of the psychosocial consequences of MVAs is a growing field of research. Unfortunately, MVAs are a ubiquitous phenomenon in industrialized countries. The psychosocial consequences of MVAs are severe and many. Blanchard and Hickling is a valuable reference for clinicians who seek information about the psychosocial treatment of MVA survivors.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Acute Trauma Response; Avoidance; Major Depressive Disorder; Posttraumatic Stress Disorder, Delayed.

Further Reading

- Blanchard, E. B. and Hickling, E. J. (1997). *After the crash: assessment and treatment of motor vehicle accident survivors*. Washington, D.C.: American Psychological Association.
- Buckley, T. C., Blanchard, E. B. and Hickling, E. J. (1996). A prospective examination of delayed onset PTSD secondary to motor vehicle accidents. *Journal of Abnormal Psychology* 105, 617–625.
- Ehlers, A., Mayou, R. A. and Bryant, B. (1998). Psychological predictors of chronic posttraumatic stress disorder after motor vehicle accidents. *Journal of Abnormal Psychology* 107, 508–519.
- Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M. and Nelson, C. B. (1995). Posttraumatic stress disorder in the National Comorbidity Survey. *Archives of General Psychiatry* 151, 888–894.
- Kuch, K., Cox, B. J. and Evans, R. J. (1996). Posttraumatic stress disorder and motor vehicle accidents: a multidisciplinary overview. *Canadian Journal of Psychiatry* 41, 429–434.
- Taylor, S. and Koch, W. J. (1995). Anxiety disorders due to motor vehicle accidents: Nature and treatment. *Clinical Psychology Review* 15, 721–738.

Mourning See: Bereavement.

Mucosal Secretory Immunity, Stress and

J A Bosch

University of Birmingham, Birmingham, UK and
University of Illinois at Chicago, Chicago, IL, USA

D Carroll

University of Birmingham, Birmingham, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
D Carroll, C Ring, and A Winzer, volume 2, pp 781–785,
© 2000, Elsevier Inc.

Introduction

An Outline of Mucosal Secretory Immunity

Neuroendocrine Regulation of Salivary S-IgA Secretion

Stress and S-IgA

Stress and Innate Secretory Immunity

Stress and Microbial Colonization

Conclusion and Future Perspectives

Glossary

<i>Antigen</i>	Molecule that initiates a response of the immune system; abbreviation of antibody generator.
<i>Immunoglobulin or antibody</i>	Protein that is produced by a specialized set of white blood cells (B lymphocytes) that has the capacity to specifically bind to antigens. This binding labels antigens for destruction by molecules and cells of the immune system.
<i>Infection</i>	Invasion and multiplication of microorganisms, which may lead to tissue damage and disease.
<i>Mucosa</i>	The mucus-coated lining of all body cavities and passages that form the interface with the outside world.
<i>Secretory immunity</i>	Protection against infection and noxious substances that is provided by the proteins present in the mucous secretions that coat the mucosa.

Introduction

Both human and animal studies have provided convincing evidence that psychosocial stress is associated with increased susceptibility to infectious disease. It has been estimated that 90 to 95% of all infections initiate at the mucosa; the mucus-secreting lining of all body cavities and passages that form the interface with the outside world. Examples are the soft oral tissues, the respiratory, gastrointestinal, and genitourinary tracts, and the eyes. Hence, investigating the

effects of stress on the immunological protection of the mucosa is an obvious starting point for research attempting to explain the link between stress and susceptibility to infectious diseases.

The mucosal surfaces are protected by the secretory immune system, which entails the secretions of a large number of local glands, including the salivary glands, the lacrimal glands, and the glands of the respiratory and gastrointestinal tracts. These glands secrete a wide variety of protective proteins (e.g., immunoglobulins, mucins, cystatins, lysozyme, lactoferrin) that constitute a first line of defense that prevents infection and disease by interfering with the entry and multiplication of microorganisms and noxious substances. This mucosal secretory immune system is under autonomic nervous system control, and it is well established that both acute and protracted psychological stress perturb autonomic nervous system activity. Thus, autonomic nervous system regulation of mucosal secretory immunity affords a plausible pathway by which psychological stress may increase disease susceptibility.

This article provides an overview of the research on stress and mucosal secretory immunity. Most research in this area has used saliva as a model system for studying mucosal secretory immunity. Moreover, of the many protective proteins that constitute secretory immunity, one protein, secretory immunoglobulin A (S-IgA), has attracted particular attention.

An Outline of Mucosal Secretory Immunity

The Secretory Immunoglobulins

Immunoglobulins (also called antibodies) are proteins that are secreted by specialized cells of the immune system, the B lymphocytes. Antibody production by the B lymphocytes is induced by foreign elements (called antigens), such as toxins, allergens, bacteria, and viruses. Upon interaction with an antigen, B lymphocytes rapidly divide and transform into immunoglobulin-producing plasma cells. A unique feature of immunoglobulins is that they have the capacity to bind (i.e., recognize) specific antigens, which helps with neutralizing or destroying this antigen. The most abundant immunoglobulin found in mucosal secretions is secretory immunoglobulin A (S-IgA).

S-IgA is produced locally by plasma cells that have migrated into the glandular tissues. Plasma cells in the salivary glands produce a polymeric form of IgA (as opposed to IgA found in plasma, which is

monomeric). This polymeric IgA (pIgA) is transported and secreted into saliva as S-IgA. The formation and secretion of S-IgA is a two-step process. First, the pIgA locks on to a receptor molecule, referred to as the polymeric-immunoglobulin receptor (pIgR), present on the basolateral surface of a glandular cell. The pIgR is translocated from the basolateral to the apical surface of the glandular cell, where it is cleaved off and secreted into saliva as S-IgA. Accordingly, both IgA secretion by the plasma cells and the availability of pIgR for translocation by the glandular cells are rate-determining steps in the secretion of S-IgA. The pIgA receptor itself can also be translocated and is subsequently secreted as secretory component (SC).

Human IgA occurs in two isotypic forms, IgA1 and IgA2. The most notable difference between the two is that IgA1 is more vulnerable to degradation by bacteria, which makes this IgA isotype less effective. This may explain why the less vulnerable IgA2 subclass is more abundant in mucosal secretions; whereas the IgA1 subclass makes up some 90% of the total IgA in serum, it comprises 60% of S-IgA in saliva and only 35% of S-IgA in the lower gastrointestinal tract.

Innate Secretory Immunity

S-IgA is only one of many protective proteins secreted in the fluids covering the mucosa. The secretion of

these other, innate immune factors (e.g., mucins, cystatins, lysozyme, lactoferrin) is also under strong neurohormonal control. Accordingly, these antimicrobial secretory proteins are also good candidates for psychoneuroimmunological investigation, although they currently received insufficient attention. [Table 1](#) lists several innate secretory factors that have been studied in this arena.

Neuroendocrine Regulation of Salivary S-IgA Secretion

As indicated, the synthesis and release of the secretory proteins by salivary glands and other mucosal secretory exocrine glands is under autonomic nervous system control, providing a sound neurobiological basis for anticipating an effect of stress on salivary secretory immunity. The preganglionic autonomic centers in the brain stem that regulate salivary gland activity receive direct inhibitory and excitatory inputs from brain areas that are part of recognized stress circuits and centers for homeostatic regulation. Studies with rats have shown that stimulation of both the sympathetic and parasympathetic autonomic nerves that innervate the salivary glands gives a rapid increase of S-IgA secretion into saliva, although the effect of sympathetic stimulation appears more robust than that of parasympathetic stimulation. Similar rapid

Table 1 Measures of innate secretory immunity that have been used in stress studies

<i>Secretory protein</i>	<i>Where found</i>	<i>Functions</i>
MUC5B (also denoted as MG1)	Saliva, nasal fluid, bronchial mucus, middle ear secretions, gall bladder epithelia	A mucin, i.e., a high-molecular-weight glycoprotein that consists largely (>75%) of carbohydrate. MUC5B is a major constituent of mucous gels covering the mucosa, where it protects against noxious substances and desiccation. Binds to a limited number of microorganisms. Host defense functions of MUC5B are enhanced by the formation of complexes with other secretory proteins.
MUC7 (also denoted as MG2)	Saliva, bronchial mucus	A mucin; binds and aggregates a large number of bacteria and fungi. Possibly also inactivate viruses.
Lactoferrin	Breast milk, nasal fluid, bronchial mucus, saliva, ocular fluid, and gastrointestinal secretions; also secreted by neutrophils	Inhibits microbial growth by binding iron, thereby depleting microorganisms from this essential substrate. Has anti-inflammatory effects and kills bacteria and fungi.
Lysozyme	Present in nearly all mucosal secretions including nasal fluid, bronchial mucus, saliva, ocular fluid, and gastrointestinal secretions; also secreted by neutrophils	Bacteriocidal by breaking down a cell wall component of gram-positive bacteria. Has also antimicrobial functions (growth-inhibiting, antiviral) through other mechanisms.
Cystatin S	Saliva, bronchial mucus	Cystatin S is a cystatin, i.e., protein that inhibits the activity of cysteine proteinases, thereby inhibiting enzymes that, among other things, are involved in virus replication and tissue invasion by bacteria.
α -Amylase	Saliva and pancreatic secretion. Low concentrations have also been demonstrated in other bodily fluids, including blood plasma, bronchial secretions, and tears	Starch-degrading enzyme. Influences the growth and tissue adhesion of streptococcal bacteria. Salivary α -amylase has been proposed as a measure of adrenergic activity, although this claim is still controversial.

increases in S-IgA secretion are observed in the gastrointestinal tract after stimulation of the local autonomic nerves. Autonomic stimulation of the salivary glands primarily affects the IgA translocation process and has little, if any, effect on IgA release by B lymphocytes.

It is still unclear whether the classic sympathetic transmitter norepinephrine is a major factor in the sympathetic stimulation on salivary S-IgA secretion. For example, the enhanced S-IgA output during moderate physical exercise, a strong sympathetic stimulus, is not attenuated by either alpha- or beta-adrenergic blockade in humans. An effect of sympathetic neuropeptides, which are coreleased with norepinephrine, has been demonstrated in animals. Other pharmacological studies in animals suggest that parasympathetic stimulation of S-IgA secretion may involve the classic parasympathetic transmitter acetylcholine and various neuropeptides.

Whereas autonomic stimulation may rapidly affect S-IgA levels, slower effects (>24 h) have been reported for glucocorticoids. The release of glucocorticoid hormones is considered central to the stress response. In rats, the synthetic glucocorticoid dexamethasone reduces both total salivary S-IgA and antigen-specific S-IgA levels. In contrast to salivary S-IgA levels, serum IgA levels are increased. Moreover, the expression of SC, the molecule that transports S-IgA into the secretions, is also increased. The effects of glucocorticoids on S-IgA secretion in humans remain to be studied, and since human and rodent S-IgA systems differ in several ways, findings for rodents should be generalized with caution.

In addition to direct stimulation of glandular S-IgA secretion, autonomic transmitters and glucocorticoids may indirectly affect the secretion of S-IgA by affecting immunological processes in the lymphoid tissues. Lymphoid organs are innervated by the autonomic nerves, and the residing immune cells carry functional receptors for the various autonomic transmitter substances. Indeed, autonomic nervous system activation has been shown to influence various aspects of local immune processing, including antigen presentation, antibody production, and cell migration. Thus, autonomic transmitters may also affect the S-IgA response by influencing the immunological priming of B lymphocytes in the lymphoid organs.

Stress and S-IgA

The Critical Distinction between Acute and Protracted Stress

The results of research on stress and S-IgA appear inconsistent at first glance. For example, some studies

Table 2 Results of academic examination studies that measured S-IgA^a

	<i>S-IgA concentration (μg/ml)</i>	<i>S-IgA secretion (μg/min)</i>
Samples taken during or close to a single exam (i.e., acute stress)		
McClelland et al., 1985	↑	
Evans et al., 1994		↑
Bosch et al., 1996, 1998	↑	↑
Bristow et al., 1997	↑	↑
Spangler et al., 1997	↑	
Huwe et al., 1998	↑	↑
Samples taken during the extended exam period (i.e., protracted stress)		
Jemmott et al., 1983	↓	↓
Kiecolt-Glaser et al., 1984	↔	
Jemmott et al., 1988	↓	↓
Mouton et al., 1989	↓	↓
Li et al., 1997	↓	
Deinzer et al., 1998	↓	↓
Deinzer et al., 2000	↔ ^b	↔

↑, increase; ↓, decrease; ↔, no change; empty cells indicate not determined.

^aStudies are separated on the basis of whether samples were taken close to (i.e., during, or minutes before or after) a single exam or sometime during the extended exam period. For complete references, see Bosch et al., 2004.

^bIn this study a decrease was observed 1 and 2 weeks after the exam period.

report that S-IgA decreases during academic examination stress, whereas other studies report an increase. Research within psychoneuroimmunology has repeatedly demonstrated that acute and protracted stressors can generate opposing results on the same immunological parameters, and this also seems to apply to S-IgA. To illustrate this, Table 2 categorizes academic examination studies on the basis of whether saliva samples were collected close to (i.e., during, or minutes before or after) a single examination or sometime during the extended examination period. Sorting studies according to this single criterion reveals a remarkably consistent picture: all studies in which the samples were collected close to an actual examination were associated with increases in S-IgA, whereas nearly all of the remaining studies found decreases. Clearly, the acute stress of an imminent exam increases salivary S-IgA levels, whereas protracted academic stress is associated with a decrease. Note that these distinct S-IgA responses are independent of when the nonstress reference, or baseline, measurement is taken.

S-IgA and Chronic Stressors

There is a robust association between protracted (i.e., weeks to years) forms of stress, impaired immune

function, and health. Typical examples of such stressors are marital disruption, bereavement, caregiving for a severely ill spouse, and unemployment. Given the association between these types of stressors, immunity, and health, it is surprising that very few studies have examined S-IgA in this context. Nonetheless, a fairly consistent picture emerges from those that have. For example, lower salivary S-IgA levels have been observed in individuals distressed by living in parts of the United States stricken by environmental disasters (e.g., the Three Mile Island nuclear power plant, toxic waste sites in Delaware) than in individuals living in safer environments. Likewise, the academic stress studies cited previously (see [Table 2](#)) indicate that protracted academic pressures can lower salivary S-IgA levels.

A number of S-IgA studies have investigated chronic stress exposure by using self-report inventories. Such questionnaires typically request subjects to check a list of recent major life events (e.g., death of a spouse, divorce, loss of job) or minor daily hassles (e.g., minor conflicts and annoyances at home or work), or require subjects to give a general assessment of the stressfulness of their lives. The validity of this approach is established by prior research demonstrating that individuals scoring highly on such questionnaires exhibit increased susceptibility to a variety of infectious and inflammatory conditions, including respiratory infections and oral infectious diseases. The evidence provided by these questionnaire-based S-IgA studies also tends to support the hypothesis that higher levels of stress exposure and perceived stress are associated with lower levels of salivary S-IgA.

Despite consistent indications that chronic stressors lower S-IgA levels, methodological considerations temper a firm conclusion at this stage. For example, none of the studies controlled for potential sources of confounding, such as health behaviors (e.g., smoking, alcohol consumption), oral health, or medication usage. Further, many of the studies, including those with null results, tested small samples and/or used questionnaires with inferior psychometric properties. Thus, although largely supportive of the hypothesis that chronic stress lowers salivary S-IgA, there is a pressing need for a large-scale, well-controlled study to accurately assess the impact of perceived chronic stress on salivary S-IgA.

S-IgA and Acute Stress

Examination stress and other naturalistic stressors
The results of the academic examination studies summarized in [Table 2](#) indicate that the acute stress (of an imminent or ongoing examination) is associated with increased S-IgA levels. The fact that acute forms of

stress increase S-IgA levels is also confirmed by research using other stress paradigms. For example, S-IgA secretion is elevated in air traffic controllers following a stressful work shift. Also, S-IgA secretion is elevated in soccer coaches watching their team play. It should be noted that such acute increases are transient; S-IgA returns to baseline values typically within 1 h after acute stress exposure.

Laboratory stressors Further support for the hypothesis that acute stress increases salivary S-IgA is provided by studies that used conventional laboratory stressors such as computer games, mental arithmetic, and time-paced memory tests. These experiments, which are performed under highly controlled conditions, use stressors of short duration (5–30 min) that typically require mental effort and elicit robust sympathetic nervous system activation. The results consistently show increases in S-IgA concentrations ($\mu\text{g/ml}$) but tend to be somewhat less consistent regarding increases in S-IgA secretion rate ($\mu\text{g/min}$). This indicates that concomitant reductions in saliva flow rate contribute to the concentration increases during acute stress.

More recent studies have demonstrated that under some conditions, which include specific emotional states (e.g., disgust) and stimuli (e.g., pain, cold), S-IgA levels may transiently decrease. The mechanisms causing these rapid decreases have yet to be elucidated. Hedonic tone per se does not seem to determine the direction of the acute changes in S-IgA; both positive and negative emotional states (induced by imagery, listening to music tapes, or watching humorous sketches) have been associated with increases in S-IgA. Thus, apart from a few studies using painful stimuli or stimuli that induce disgust, virtually all types of acute states and virtually all experimental manipulations elicit an acute and transient rise in salivary S-IgA.

Stress and Oral Immunization

The research described thus far has concerned the effects of stress on total S-IgA levels. A few studies also investigated the effects of psychosocial stress on the daily variation in specific S-IgA, i.e., S-IgA that binds to a specific antigen. In these studies, participants were required to swallow a capsule containing rabbit albumin each morning for 8–12 weeks, and the daily variations in specific S-IgA (i.e., S-IgA directed against the rabbit protein) in saliva were measured in association with variations in daily mood and events. Antigen-specific S-IgA levels were lower on days with higher levels of negative mood and undesirable events and were elevated on days with more positive moods

and desirable events. Considering the pivotal role of the secretory immune defenses in protecting against infection, oral immunization provides an important and underutilized model for studying the effects of psychosocial factors on immune function.

The validity of the conventional method of measuring total S-IgA levels (as opposed to antigen-specific S-IgA levels described previously) has been a topic of heated debate. This debate, however, rests on the assumptions that S-IgA is a readout parameter for specific immune responses directed toward a particular antigen. If this were the case, only the levels of S-IgA molecules that are specifically directed against particular antigens would matter for host protection, rendering the measurement of total S-IgA levels less meaningful. However, total S-IgA for a large part consists of polyreactive antibodies (i.e., recognizing many different antigens). This broad specificity of S-IgA antibodies implies that total immunoglobulin levels, and not just antigen-specific levels, are important to mucosal protection. This idea is consistent with the evidence showing that total S-IgA levels predict the risk for respiratory and oral infections. Thus, the measurements of total and antigen-specific S-IgA levels seem to be immunologically meaningful ways of determining mucosal protection.

Effects on Secretory Component and S-IgA Subclasses

During acute stress in humans, the secretion of S-IgA appears to be regulated independently from the secretion of its transporter molecule, SC. For example, the passive stress induced by viewing a surgical video produced a rapid decrease in S-IgA secretion but an increase in the secretion of SC. During conventional active stress, a time-paced memory task, both S-IgA and SC secretion increased. It is as yet unclear whether these acute increases are due to increased endothelial transport or to a combination of increased IgA availability and transport into saliva. Both stress and physical exercise studies show that the rapid increases in S-IgA mainly reflect increases in the IgA1 subclass. The mechanism responsible for this subclass-specific effect is unknown.

Stress and Innate Secretory Immunity

Stress and Mucins

Two human studies have shown that acute stress enhances the secretion in saliva of the mucins MUC5B and MUC7. These findings are in line with the results of animal studies showing that increased gastrointestinal and respiratory mucin secretion

occurs in response to water immersion and restraint stress. These animal experiments also found effects on related biological processes, such as mucin synthesis and posttranslational processing (e.g., sulfatation, glycosylation). Pharmacological studies show that synthesis, posttranslational modification, and secretion of salivary mucins results, at least in part, from beta-adrenergic stimulation.

Stress-induced increases in salivary mucin may have consequences that benefit the host by enhancing mucosal barrier function, but also some that benefit microorganisms. For instance, the acute stress-induced increase in salivary MUC5B secretion enhances the adherence (*in vitro*) of *Helicobacter pylori*, a causative agent of duodenal ulcers and stomach cancer. Again, the effects of acute and protracted stressors may be quite different. For example, protracted stress decreases colonic mucin secretion in rats. This decrease in mucin secretion enhances mucosal permeability, thereby potentiating reactivation of experimental colitis induced by a chemical irritant.

Stress and α -Amylase

Studies in the early 1980s reported that salivary α -amylase concentration increases during relaxation and decreases with acute stress. However, later studies were unable to replicate these findings and instead found that acute stress increases salivary α -amylase concentration and secretion rate. These effects are again thought to result from stress-induced sympathetic activation, which would be consistent with the findings of studies showing that α -amylase secretion is increased by administration of adrenergic agonists, electrical stimulation of the local sympathetic nerves, and physical exercise. Moreover, increases in salivary α -amylase correlate with serum norepinephrine and cardiac left ventricular ejection time, a measure of cardiac sympathetic drive. Nonetheless, recent claims that salivary α -amylase is a valid noninvasive measure of adrenergic activity should be regarded with caution. First, α -amylase is also secreted in response to nonadrenergic sympathetic transmitters, i.e., various neuropeptides. Second, α -amylase secretion is also stimulated by parasympathetic stimulation, and parasympathetic stimulation augments the effects of sympathetic stimulation.

Stress and Other Mucosal Secretory Proteins

In humans, acute stressors that elicit sympathetic-parasympathetic coactivation increase the secretion of salivary cystatin S and lactoferrin, but no or only minimal effects are seen in response to a stress exposure that evokes sympathetic activation in

conjunction with a vagal withdrawal (a time-paced memory test). This finding is consistent with the notion that the two branches of the autonomic nervous system have a synergistic effect on secretory gland activity; that is, although sympathetic activation is the main stimulus for glandular protein secretion, the effects are strongly augmented by concurrent parasympathetic activity. As for more protracted stress, a decrease in salivary lysozyme secretion is observed during an examination period. The association between self-reported stress and salivary lysozyme levels appears inconsistent: one study found a negative correlation, whereas another did not find any association.

Stress and Microbial Colonization

A remaining, and critical, issue is whether the effects of stress on secretory immunity translate into functional changes, such as affecting processes that are involved in microbial colonization. Such functional effects might explain the alterations in mucosal microflora, seen in both humans and animals, under various forms of psychological strain, such as depression, space flight simulations, maternal separation, familial strains, animal fighting and relocation, and life events. Such effects may also explain the close association between stress and oral diseases with a bacterial etiology, such as periodontal disease and dental caries.

Aggregation is a process by which bacteria are clumped together through interactions with the host's secretory proteins so that these microorganisms cannot effectively adhere to the mucosal surfaces. Hence, a decreased bacterial aggregation indicates a decreased capacity for mucosal microbial clearance. Examination stress reduces the aggregation by saliva of the oral bacterium *Streptococcus gordonii*. A reduced streptococcal aggregation is also observed in rats administered physiological doses of the beta-adrenergic agonist isoproterenol for 1 week. This indicates that both brief and prolonged sympathetic activation can negatively affect the capacity of saliva to aggregate oral bacteria.

Microbial adhesion is a process by which microbes gain a stable foothold in the host; it constitutes a first and essential step in infection. Acute laboratory stress in humans enhances saliva-mediated adhesion of oral bacteria (viridans streptococci) and non-oral bacteria (*Helicobacter pylori*). Laboratory stress also affects the saliva-mediated adhesion of microorganisms to each other. Animal studies demonstrate that glucocorticoid administration is associated with an increased bacterial adherence to the intestinal mucosa and a

reduced luminal S-IgA. Similarly, in rats, 1 week of physiological doses of the beta-agonist isoproterenol increased the salivary adherence of *Streptococcus mutans*. Thus, stress, sympathetic activation, and elevated glucocorticoid levels promote the adhesion of microorganisms to mucosal surfaces.

Conclusion and Future Perspectives

Psychological stress has a variety of effects on mucosal salivary secretory immunity in humans, including altered secretion of adaptive and innate immune factors, reduced bacterial aggregation, increased bacterial adherence, and altered microbial coadherence. There is still little known about the effects of stress on secretory immunity in secretions other than saliva. The scant research in this area, mostly in rodents, confirms and extends the salivary findings in humans; there are effects on protein secretion, synthesis, and posttranslational modification, as well as on functional changes such as enhanced bacterial adherence and increased intestinal permeability.

In addition to an emphasis on salivary immunity, there has been a disproportional focus on S-IgA. The fact that individuals deficient in S-IgA, the most common form of immunodeficiency, are, for the most part, in good health suggests that other secretory factors may be important.

More attention also needs to be paid also to the effects of protracted stress exposure on secretory immunity. Knowledge about the effects of chronic stress on innate secretory immunity is practically absent, and the studies on chronic stress and S-IgA, although painting a fairly consistent picture, suffer from several methodological shortcomings. Further, the underlying mechanisms have yet to be fully elucidated. A promising model for studying the effects of chronic stress on adaptive secretory immunity is mucosal immunization.

Finally, researchers should extend their horizons to include the study of the functional implications of stress-induced changes in mucosal protein secretion. A number of microbiological methods are available to investigate the effects of immunosecretory changes on microbial colonization processes, which include viral invasion, bacterial adherence, bacterial growth, and bacterial aggregation. This approach, rarely considered in the realm of psychoneuroimmunology, may reveal the mechanisms by which immunosecretory changes affect the mucosal microflora and susceptibility to infection. Moreover, this methodology will help to strengthen the proposition that stress, via its effects on secretory immunity, has important health consequences.

Acknowledgment

Support for this work was provided by grant 1-RO3-DE-16726-01 from the National Institute of Dental and Craniofacial Research to J. A. B.

See Also the Following Articles

Acute Stress Response; Experimental; Autonomic Nervous System; Immune Function, Stress-Induced Enhancement; Immune Response; Immune System, Aging; Lymphocytes; Sympathetic Nervous System.

Further Reading

- Bosch, J. A., Ring, C., de Geus, E. J. C., Veerman, E. C. I. and Amerongen, A. V. (2002). Stress and secretory immunity. *International Review of Neurobiology* 52, 213–253.
- Bosch, J. A., Turkenburg, M., Nazmi, K., et al. (2003). Stress as a determinant of saliva-mediated adherence and coadherence of oral and non-oral microorganisms. *Psychosomatic Medicine* 65(4), 604–612.
- Bosch, J. A., de Geus, E. J. C., Veerman, E. C. I., Hoogstraten, J. and NieuwAmerongen, A. V. (2003). Innate secretory immunity in response to laboratory stressors that evoke distinct patterns of cardiac autonomic activity. *Psychosomatic Medicine* 65(2), 245–258.

- Bosch, J. A., de Geus, E. J. C., Ring, C., NieuwAmerongen, A. V. and Stowell, J. R. (2004). Academic examinations and immunity: academic stress or examination stress? *Psychosomatic Medicine* 66(4), 625–626.
- Brandtzaeg, P. and Johansen, F. E. (2005). Mucosal B cells: phenotypic characteristics, transcriptional regulation, and homing properties. *Immunological Reviews* 206, 32–63.
- Norderhaug, I. N., Johansen, F. E., Schjerven, H. and Brandtzaeg, P. (1999). Regulation of the formation and external transport of secretory immunoglobulins. *Critical Reviews in Immunology* 19(5–6), 481–508.
- Proctor, G. B. and Carpenter, G. H. (2002). Neural control of salivary S-IgA secretion. *International Review of Neurobiology* 52, 187–212.
- Schenkels, L. C. P. M., Veerman, E. C. I. and Nieuw Amerongen, A. V. (1995). Biochemical composition of human saliva in relation to other mucosal fluids. *Critical Review in Oral Biology and Medicine* 6(2), 161–175.
- Soderholm, J. D. and Perdue, M. H. (2001). II. Stress and intestinal barrier function. *American Journal of Physiology Gastrointestinal and Liver Physiology* 280(1), G7–G13.
- Teeuw, W., Bosch, J. A., Veerman, E. C. and Amerongen, A. V. (2004). Neuroendocrine regulation of salivary IgA synthesis and secretion: implications for oral health. *Biological Chemistry* 385(12), 1137–1146.

Multi Drug Resistance P Glycoprotein and other Transporters

E C M de Lange

Leiden University, Leiden, Netherlands

© 2007 Elsevier Inc. All rights reserved.

ATP-Binding Cassette Transporter Superfamily
 Multidrug Resistance Transporter Proteins
 Multidrug Resistance Transporter Proteins and Pharmacokinetics
 Other Physiological Functions of Multidrug Resistance and Other ATP-Binding Cassette Transporters
 Multidrug Resistance: ATP-Binding Cassette Transporter Modulation and Regulation of Expression

Glossary

ATP-binding cassette (ABC) transporters A large and multifunctional family of structurally related membrane proteins. ABC transporters use the energy of ATP

hydrolysis to pump substrates across cell membranes and are involved in the transmembrane transport of various substances.

The ability of cells to develop resistance to a broad range of structurally and functionally unrelated drugs after being exposed to (one of) these drugs. The mechanism is that the efflux transport by these proteins leads to lower intracellular concentrations, thereby protecting the cells and tissues from potential toxic compounds from the external milieu.

A group of MDR ABC transporters that are structurally related to but distinct from P-glycoproteins; mainly transports anionic derivatives of drugs when conjugated to glutathione/glucuronide. This family of MRP proteins comprises nine members in humans, designated MRP1–MRP6.

Multidrug resistance (MDR)

Multidrug resistance-related proteins (MRPs)

P glycoprotein (Pgp) The first discovered MDR transporter and most well-known ABC transporter; has been shown to transport a variety of structurally unrelated compounds across cellular membranes.

ATP-Binding Cassette Transporter Superfamily

The ABC transporters comprise a large and multi-functional family of structurally related membrane proteins that are located in the plasma membrane of the cells or in the membrane of various cellular organelles. ABC transporter proteins use the energy of ATP hydrolysis to translocate various molecules across these barriers, and some other ABC transporters use ATP to form specific membrane channels. The ABC transporters serve a variety of physiological roles, and also some inherited diseases appear to be linked to mutations in the genes encoding for the ABC transporters.

At present, approximately 50 ABC transporters are known. Many synonyms exist for these transporters due to their history of discovery. Recently, a consistent nomenclature has been introduced on the basis of the organization of domains and amino acid homology of these ABC transporters. In the new nomenclature, the ABC transporters are divided into seven distinct sub-families: ABCA, ABCB, ABCC, ABCD, ABCE, ABCF, and ABCG. [Table 1](#) presents the classification of these families and their members (including the synonyms), together with their expression, endogenous and exogenous substrates, functions, and relation to human disorders. It is clear that the ABC transporters govern the kinetics of chemical communication in the body, which makes the knowledge of their specific actions in homeostatic and feedback mechanisms extremely important for further insight into the body's functioning at the molecular level under normal and challenged (stress) conditions.

Multidrug Resistance Transporter Proteins

ABC transporters have been the subject of intense scrutiny as potential mediators of clinical drug resistance. The identification of the MDR1-encoded P-glycoprotein (Pgp) approximately 20 years ago, allowed us to recognize that reduced intracellular accumulation of anticancer agents can result in significant degrees of drug resistance. Over the years, Pgp has become the prototype of the MDR transporters. Among the ABC transporters, two other major groups can be identified as being involved in intrinsic and/or acquired MDR of cells: the multidrug

resistance-related protein (MRP) family and the breast cancer resistance protein (BCRP). Many drugs that have been developed for the treatment of human diseases are substrates for these multidrug transporters. Because of that, the degree of expression and the functionality of these transporters can directly affect the therapeutic effectiveness of such agents. Treatment failure due to MDR was first found in connection with cancer, but later in time it was also found in connection with other conditions such as some autoimmune disorders and infectious diseases ([Table 1](#)).

P-Glycoprotein

Pgp is the most well-known of the ABC transporters and was the first to be identified in humans, in which it plays a critical role in drug resistance in the treatment of cancers. Numerous investigations with many drugs have demonstrated that Pgp has an important role in determining the concentration–time profiles of Pgp substrates in the different parts of the body. In general, Pgp preferentially extrudes large hydrophobic, positively charged molecules. This transporter also is involved in the transport of certain cytokines. Pgp may even play a role in allograft rejection and in the inhibition of apoptosis. Another possible role of Pgp lies in the pathogenesis of Alzheimer's disease by the direct interaction of Pgp on amyloid- β 40 and amyloid- β 42 or by Pgp's influencing the accumulation of these proteins. Also, an interesting role for Pgp is indicated in the neuroendocrine functioning and regulation of the hypothalamic-pituitary-adrenocortical (HPA) axis because Pgp is involved in the efflux of certain natural and synthetic glucocorticoids from the brain.

Multidrug Resistance-Related Protein Family

The MRPs (or the ABCCs) have a broad tissue distribution and are able to extrude negatively charged anionic drugs and neutral drugs conjugated to glutathione, glucuronate, or sulfate. Some MRPs are able to transport neutral drugs if co-transported with glutathione. The total MRP family now includes MRP1–MRP9 (and ABCC7 and ABCC8). MRP1 was the first to be discovered and is actually a prototype glutathion conjugate pump that transports a variety of drugs conjugated to glutathione, sulfate, or glucuronate; anionic drugs and dyes; neutral-basic amphiphatic drugs; and even oxyanions. The most important substrate of MRP1 is the endogenous leukotriene LTC₄. MRP1 also contributes to the blood–cerebrospinal fluid (CSF) barrier at the level of the choroid epithelial cells. MRP2 and MRP3 have overlapping substrate specificities with MRP1 but different tissue distributions, and they appear also to exhibit drug resistance.

Table 1 ATP binding cassette transporter families and members^a

Family	Member (synonyms); most common name	Tissue expression	Physiological substrates; other substrates; inhibitors/modulators	Function	Genetic mutations related to human disease
ABCA	ABCA1 (ABC1, TGD, HDLDT1, CERP)	Ubiquitous (plasma membrane)	Phospholipids, cholesterol	Removal of phospholipids and cholesterol on to high-density lipoprotein particles	Tangier disease, atherosclerosis
	ABCA2 (ABC2)	Brain, kidney, lung, heart, lysosomal membrane	<i>Estramustine?</i>	Steroid transport?	
	ABCA3 (ABC3, ABCC)	Lung			
	ABCA4 (ABCR, ABC10, STGD)	Retina (rod photoreceptors)	<i>N-retinylidene-phosphatidyl-ethanolamine, retinaldehyde</i>		Stargardt disease, retinitis pigmentosa, macular dystrophy, cone-rod dystrophy
	ABCA5	Muscle, heart, testes			
	ABCA6	Liver			
	ABCA7	Spleen, thymus			
	ABCA8 (ABCR)	Ovary			
	ABCA9	Heart			
	ABCA10	Muscle, heart			
	ABCA11	Stomach			
	ABCA12	Low in all tissues			
ABCB	ABCB1 (MDR1, PGP, PGY1, GP170) P-Glycoprotein	Epithelial cells of kidney, liver, choroid plexus, intestine; endothelial cells of brain capillaries	IL-2, IL-4, IFN- γ , amyloid β -40, amyloid β -42, glucosylceramide, platelet-activating factor <i>Actinomycine D, aldosterone, amprenavir, bisantrene, calcein-M, citalopram, colchicine, corticosterone, cortisol, cyclosporin A, dexamethasone, digoxin, domperidone, doxorubicin, daunorubicin, enaminonen, erythromycine, etoposide, FK506, indinavir, loperamide, lovastatin, morphine, nelfinavir, ondansetron, paclitaxel (taxol), phenytoin, prednisolone, quinidine, rifampin, ritonavir, saquinavir, sparfloxacin, teniposide, terfenadine, (99m)-Tc-tetrofosfine, valspodar (PSC833), verapamil, vinblastine, vincristine</i> <i>Inhibitors:</i> Agosterol A, amiodarone, amitriptylin, amprenavir, biricodar (VX-710), chlorpromazin, cyclosporin A, diltiazem, dipyridamol, fluphenazin, GF120918, gramicidine D, LY335979, mefipriston, midazolam, nelfinavir, pimozin, pluronic L61, progesterone, promethazin, propafenone, propranolol, quinidine, reserpine, ritonavir, saquinavir, spironolactone, staurosporin, tamoxifen, trifluoperazin, triflupromazin, V-104, valinomycine, valspodar (PSC 833), verapamil	Multidrug resistance	

	ABCB2 (TAP1, PSF1, RING4, ABC17)	Ubiquitous (ER)	Peptides	Peptide transport (into the ER)	Juvenile-onset psoriasis; immune deficiency
	ABCB3 (TAP2, PSF2, RING11, ABC18)	Ubiquitous (ER)	Peptides	Peptide transport (into the ER)	Immune deficiency
	ABCB4 (PGP3, MDR3(2), PFIC-3)	Liver (hepatocytes)	Long-chain phosphatidyl-choline	Phosphatidylcholine transport	Progressive familial intrahepatic cholestasis-3, intrahepatic cholestasis of pregnancy
	ABCB5 (MTABC3)	Ubiquitous			
	ABCB6 (ABC14, UMAT, MTCA3)	Ubiquitous in mitochondria	Iron?	Iron transport	
	ABCB7 (ABC7, ATMIP, ASAT)	Ubiquitous in mitochondria	Iron?	Fe/S cluster transport	Sideroblastic anemia with ataxia
	ABCB8 (MABC1)	Mitochondria	Intermediates in heme/biosynthesis?		
	ABCB9	Lysosomes			
	ABCB10 (MTABC2)	Mitochondria			
	ABCB11 (sPGP, BSEP, PGY4)	Liver (hepatocytes)	Bile salts (e.g., taurocholate) <i>Paclitaxel (taxol)</i>	Bile salt transport	Progressive familial intrahepatic cholestasis-2
ABCC	ABCC1 (MRP1, MRP) Multidrug resistance related protein	Ubiquitous in normal tissues (mainly lung and choroid plexus epithelial cells) and brain capillary endothelial cells	17- β -glucuronyl estradiol, S-glutathionyl prostaglandin A2, glutathione disulfide, 4-hydroxynonenol-GS, leukotriene C4 (LTC4)-GS, leukotriene D4 (LTD4)-GS, leukotriene E4 (LTE4)-GS, NNAL-O-glucuronide lipid peroxidation products, nitroquinoline 1-oxide-GS <i>Aflatoxin B1-epoxide-GS, arsenite, arsenate (sodium), BCECF, calcein, chlorambucil-GS, colchicine, CPT-14, daunorubicin, DHEAS, 2,4-dinitrophenyl-GS, doxorubicin, epirubicin, ethacronic acid-GS, etoposide, fluorescein, flutamide, hydroxyflutamide, glucuronosyl etoposide, herbicides, heavy metals, indomethacin (-GS), melphalan-GS, methotrexate, rhodamine, ritonavir, sequinavir, ⁹⁹Tc-Sestambi, SN-38, SN-38-GS, SNARF, tobacco nitrosamines, ⁹⁹Tc-tetrofosmin, vinblastine, vincristine</i> <i>Inhibitors:</i> Agosterol A, benzbromarone, budesonide, biricodar (VX-710), cyclosporin A, indomethacin, probenecid, genistein, quercetin, sulfinylpyrazone, verapamil	Drug resistance, ubiquitous GS-X pump, immune response involving cysteneyl leukotrienes	
	ABCC2 (MRP2, cMOAT)	Liver, intestine, kidney, brain endothelial cells (plasma membrane)	Bilirubin-glucuronides, GSSH, GSH, including cotransport, estradiol-17- β -o-glucuronide, acidic bile salts <i>Anionic drug conjugates, cisplatin, doxorubicin, epirubicin, etoposide, indinavir, phenytoin, ritonavir, saquinavir, sulfinylpyrazome, vinblastine</i> <i>Inhibitors/modulators:</i> Leukotriene C4, probenecid	Organic anion efflux, hepatobiliary extrusion of amphipathic ions	Dubin–Johnson syndrome

Continued

Table 1 Continued

<i>Family</i>	<i>Member (synonyms); most common name</i>	<i>Tissue expression</i>	<i>Physiological substrates; other substrates;^b inhibitors/modulators</i>	<i>Function</i>	<i>Genetic mutations related to human disease</i>
	ABCC3 (MRP3, cMOAT-2)	Liver, bile ducts, gut, adrenal cortex of kidney, intestine, placenta	Bile salts, leukotriene C4 <i>Anionic drug conjugates, acetaminophen glucuronide methotrexate, etoposide, teniposide</i> <i>Inhibitors/modulators:</i> Probenecid	Drug resistance	
	ABCC4 (MRP4, MOAT-B)	Many tissues (plasma membrane)	Cyclic nucleotides, conjugated steroids and bile acids, DHEAS nucleotide analogs, 9-(2-phosphophenyl methoxyethyl)-adenine, prostaglandins, taurothiocholate 3-sulfate, thiopurine monophosphates <i>Nucleotide analogs, organic anions, nucleosides, monophosphates, zidovudine, other antiretroviral agents</i> <i>Inhibitors/modulators:</i> Probenecid	Nucleoside transport	
	ABCC5 (MRP5, MOAT-C)	Ubiquitous	Cyclic nucleotides, nucleotide analogs, 9-(2-phosphono- methoxyethyl) adenine, thiopurine monophosphates <i>Nucleotide analogs, organic anions nucleoside monophosphates, mitoxantrone, topotecan, doxorubicin</i> <i>Inhibitors/modulators:</i> Probenecid fumitremorgin C, GF120918	Nucleoside transport	
	ABCC6 (MRP6)	Kidney, liver (plasma membrane)	<i>BQ-123? (acidic peptide)</i>	Elastic tissue homeostasis	Pseudo-xanthoma elasticum
	ABCC7 (CFTR)	Exocrine tissues		Cystic fibrosis transmembrane conductance regulator (cAMP-dependent chloride channel)	Cystic fibrosis, CBAVD, pancreatitis, bronchiectasis
	ABCC8 (SUR1)	Pancreas		K(ATP) channel regulation	Familial persistent hyperinsulinemia hypoglycemia of infancy
	ABCC9 (SUR2)	Heart, muscle		Sulfonurea receptor	Dilated cardiomyopathy
	ABCC10 (MRP7)	Ubiquitous (low)			
	ABCC11 (MRP8)	Testis and breast, lower in all other tissues			
	ABCC12 (MRP9)	Brain and testis, lower in all other tissues			

ABCD	ABCD1 (ALD, ALDP)	Many tissues (peroxisomes)	VLCFA	VLCFA transport regulation	
	ABCD2 (ALD1, ALDR)	Many tissues (peroxisomes)	VLCFA	Adrenoleukodystrophy	
	ABCD3 (PMP70, PXMP1)	Many tissues (peroxisomes)	VLCFA		
	ABCD4 (PMP69, P70R)	Many tissues (peroxisomes)	VLCFA		
ABCE	ABCE1 (OABP)	Ovary, testes, spleen		Oligoadenylate-binding protein	
ABCF	ABCF1	Ubiquitous			
	ABCF2 (ABC50)	Ubiquitous			
	ABCF3	Ubiquitous			
ABCG	ABCG1 (ABC8, human white)	Ubiquitous		Cholesterol transport	
	ABCG2 (ABCP, MXR, BCRP)	Mainly in liver, placenta, small intestine; also in heart, lung, skeletal muscle, kidney, spleen, thymus, brain capillaries	Heme, porphyrins, flavanoids, pheophorbide, sulfated estrogens	Drug resistance, stem-cell differentiation, survival of stem cells (under hypoxic conditions).	HIV treatment?
	Breast cancer resistance protein		<i>Anthracyclines, bisantrene, diflomotecan, etoposide, flavopiridol, imatinib mesylate (STI-571, Gleevec), imatib, irinotecan, SN-38, methotrexate, mitoxantrone, lamivudine, rhodamine 123, teniposide, topotecan, zidovudine, Inhibitors/modulators</i> CI1033, futremorgirin C, GF120918, hammerhead ribozyme, HIV protease inhibitors, imatinib (STI-571, Gleevec) pantoprazole, prazosin, reserpine, ZD1839 (Iressa)		
	ABCG4 (white2)	Liver			
	ABCG5 (steroline 1)	Liver, intestine		Sterol transport, biliary cholesterol secretion/absorption	Sitosterolemia, sterol accumulation, atherosclerosis
	ABCG8 (steroline 2)	Liver, intestine		Sterol transport, biliary cholesterol secretion/absorption	Sitosterolemia, sterol accumulation, atherosclerosis

^aBCECF, 2',7'-bis(carboxyethyl)-5(6)-carboxyfluorescein; CBAVD, congenital bilateral absence of the vas deferens; CFTR, cystic fibrosis transmembrane conductance regulator; DHEAS, dehydroepiandrosterone; ER, endoplasmic reticulum; GSH, glutathione; GSSH, glutathione persulfide; GS-X, GSH conjugate; HIV, human immunodeficiency virus; IFN, interferon; IL, interleukin; PMEA, 9-(2-phosphophenyl methoxyethyl)-adenine; VLCFA, very long-chain saturated fatty-acyl-CoA.

^bOther substrates appear in italics.

MRP4 and MRP5 broaden the spectrum of drug resistance to nucleotide analog drugs; MRP6 plays a role in elastic tissue homeostasis. The functions of MRP7–MRP9 are largely unknown.

Breast Cancer Resistance Protein

BCRP (or ABCG2) is an ABC half transporter that displays drug resistance. In general, the BCRP transporter preferentially extrudes large hydrophobic, positively charged molecules. BCRP is involved in MDR in cancer, especially with regard to acute myeloid leukemia. Also, it plays a role in the survival of stem cells under hypoxic conditions and might play a role in regulating stem-cell differentiation.

Multidrug Resistance Transporter Proteins and Pharmacokinetics

Both the absorption and body distribution of drugs depend on their passage through diverse barriers. A number of MDR transporters limit the absorption of drugs via the gastrointestinal tract and the distribution of drugs into the brain via the blood–brain barrier; excretion is also governed by these transporters.

Oral Absorption

Efflux transporters such as Pgp and the MRPs in the intestinal may form the rate-limiting barrier to oral drug absorption. The poor bioavailability characteristics of many chemotherapeutic agents for cancer and human immunodeficiency virus (HIV) protease inhibitors may be explained by the fact that those drugs are among the best substrates for these transporters. The co-administration of the relevant substrates with the inhibitors/modulators of these transporters is a possible route to enhancing bioavailability. Of interest is the systematic site-specific expression of Pgp and MRPs along the various parts of the gastrointestinal tract.

Brain Distribution

The brain is the tissue most frequently targeted by drugs. To be effective, drugs must be transported to the site of action. The interface between the blood and brain is composed of the blood–brain barrier (brain capillary endothelial cells) and the blood–CSF barrier (choroid plexus epithelial cells).

The presence of these efflux transporters at the blood–brain/CSF barriers explains the fact that many drugs exhibit poor brain distribution despite their favorable lipophilic character; the barriers are substrates for the efflux transporters. Thus, in many studies Pgp, especially (expressed at the luminal face

of the brain capillary endothelial cells), has been shown to be responsible for lowering the brain distribution of many drugs. MRP1 is more specifically localized in the basolateral membrane of the choroid epithelial cells, preventing CSF penetration of its substrates. Pgp is also expressed at the choroid epithelial cells, but in a manner that governs influx transport. Other ABC transporters, such as MRP2 and BCRP, have also been implicated in protecting the brain tissue against exogenous compounds.

Drug Excretion

The liver and kidneys play an important role in the excretion of drugs, the main routes being through the biliary tract and renal proximal tubule, respectively. In both the liver and kidneys, Pgp and other multidrug transporter proteins are variably expressed in the different suborgan structures and cell linings.

Functional Genetic Polymorphism

In addition, functional genetic polymorphisms, as have been identified for the Pgp gene, may influence the absorption, distribution, and excretion of the transporters substrates.

Other Physiological Functions of Multidrug Resistance and Other ATP-Binding Cassette Transporters

The ABC transporters govern a broad spectrum of physiological functions because they are involved in the transmembrane transport of various substances that exhibit a wide variety of chemical structures; examples are endogenous compounds such as ions, amino acids, peptides, sugars, vitamins, steroid hormones, bile acids, and phospholipids and also many exogenous compounds. Next we briefly describe a number of these physiological roles.

Defense against Oxidative Stress, Cancer, and Inflammation

Adequate glutathione (GSH) levels are essential to life and needed to fight infections and cancer cells. The MRPs that have been functionally characterized so far act as ATP-dependent export pumps for conjugates with glutathione, glucuronate, or sulfate. The most important substrate of MRP1 is the endogenous glucuronide conjugate leukotriene C4 (LTC4), a *de novo* synthesized mediator of inflammation. MRP1 and MRP2 also mediate the co-transport of unconjugated amphiphilic compounds, together with free glutathione. MRP3 preferentially transports glucuronides but not glutathione S-conjugates or

free GSH. In addition, MRP1 and MRP2 contribute to the control of the intracellular glutathione disulfide levels and may thereby play an essential role in the response to oxidative stress when the activity of glutathione disulfide reductase becomes rate limiting. Thus, the members of the MRP family serve as export pumps that prevent the accumulation of anionic conjugates and glutathione disulfide in the cytoplasm, and they therefore play an essential role in detoxification and defense against oxidative stress.

Lipid Metabolism

The transport of specific molecules across lipid membranes is also an essential factor in the physiology of all living organisms, and a large number of specific transporters have evolved to carry out this function. Many ABC genes play a role in the maintenance of the lipid bilayer and in the transport of fatty acids and sterols within the body. Actually, the ABCA1 transporter is responsible for delivering cholesterol to nascent high-density lipoprotein (HDL) particles, a process known as reverse cholesterol transport. Enhancing ABCA1 activity, either through changes in gene expression or activity, is a plausible therapeutic strategy for increasing HDL levels. Dietary cholesterol is the primary contributor to the risk of developing cardiovascular disease for a subset of patients. The half-transporters ABCG5 and ABCG8 form a heterodimer responsible for regulating intestinal cholesterol absorption and the excretion of cholesterol in the liver. Increasing ABCG5/G8 expression or activity is a viable strategy for reducing dietary cholesterol. The expression of a large number of ABC transporters in monocytes/macrophages and their regulation by cholesterol flux indicate that these transporter molecules are potentially critical players in chronic inflammatory diseases such as atherosclerosis.

Antigen Presentation

The transporter associated with antigen processing 1 (TAP1; ABCB2) and TAP2 (ABCB3) are half-transporters that serve to transport peptides into the endoplasmic reticulum. At that site, they can be complexed with class I human leukocyte antigen (HLA) molecules for presentation on the cell surface. TAP expression is required for the stable expression of the HLA class I proteins. The TAP complex transports amino acid peptides with a preference for Phe, Leu, Arg, and Tyr at the C-terminus, similar to the specificity of the HLA class I proteins. Several DNA viruses such as herpes simplex virus express molecules that interfere with antigen expression by disrupting the function of the TAP complex. TAP2 (ABCB3) functions as a heterodimer with TAP1.

Regulation of Insulin Secretion

Mutations in ABCC8 may result in familial persistent hyperinsulinemic hypoglycemia of infancy, demonstrating its role in the regulation of insulin secretion. Actually ABCC8 forms ATP-sensitive potassium channel together with an inwardly rectifying potassium channel. The closure of ATP-sensitive potassium channels in pancreatic islet β -cells initiates a cascade of events that leads to insulin secretion. ABCC8 therefore seems to be responsible for regulating glucose metabolism by insulin secretion.

Chloride Secretion

Chloride is the most abundant anion and the predominant permeating species. The chloride channels are crucial for transepithelial transport and the control of water flow, and they often provide unexpected permeation pathways for a large variety of anions. ABCC7 (CFTR, cystic fibrosis transmembrane) is a cAMP-dependent chloride channel and is responsible for chloride secretion in the gastrointestinal mucosa. The major toxins that cause secretory diarrheas act by increasing the levels of cAMP or cGMP. This results in a secondary increase in active chloride secretion through ABCC7, whereas sodium and water follow passively. Compounds that inhibit ABCC7 appear to be well suited as treatments for both secretory diarrhea and cholera.

Potassium Channel Regulation

ATP-sensitive potassium (KATP) channels, as inhibited by intracellular ATP, play key physiological roles in many tissues. KATP channels set membrane excitability in response to stress challenge and preserve cellular energy-dependent functions. ABCC9, together with an inwardly rectifying potassium channel, forms a KATP channel. This complex may have a vital role in securing cellular homeostasis under stress. Mutations in ABCC9 may lead to KATP channel dysfunctions and have been related to susceptibility to dilated cardiomyopathy.

Bile Salt Transport

Canalicular bile secretion is the rate-limiting step in bile formation. The canalicular membrane contains ABC proteins that transport biliary compounds against steep concentration gradients across the canalicular membrane. Pgp excretes hydrophobic cationic compounds, whereas ABCB4 (Pgp3 or MDR3) acts as a phospholipid flippase for phosphatidyl choline. In addition to biliary excretion of phospholipids, bile is also the major pathway for the elimination of cholesterol, mostly derived from HDLs in plasma.

ABCC2 (MRP2) is an important (bilirubin) conjugate export pump. A major liver bile-salt transporter in the liver is ABCB11 (BSEP, bile salt export pump). The production of bile is also critically dependent on the coordinated regulation and function of this BSEP. The endothelin antagonist bosentan inhibits BSEP and leads to the intracellular accumulation of cytotoxic bile salts, and bile-salt-induced liver damage. Similarly, troglitazone, an insulin sensitizer, induces cholestasis and hepatotoxicity via competitive inhibition of the BSEP-mediated bile salt transport.

Multidrug Resistance: ATP-Binding Cassette Transporter Modulation and Regulation of Expression

Modulation

MDR modifying agents, which inhibit the function of the MDR transporters by either competitive or noncompetitive mechanisms, are good candidates for the pharmacological modulation of transporter functionality, currently still mostly focused on cancer treatment. The modulators are intended to increase the cytotoxic action of MDR-related drugs by preventing the extrusion of anticancer drugs from the target cells to significantly improve the treatment effectiveness.

To date, three generations of modulators/inhibitors exist, mostly acting on Pgp. The first generation comprises calcium channel blockers (verapamil, diltiazem, and azidopine), quinine derivatives, calmodulin inhibitors (trifluoroperazine and chlorpromazine), and the immunosuppressive agent cyclosporin A. The modulators of the second generation consists of derivatives of the first-generation compounds, for example R-verapamil and PSC-833 (from cyclosporin). Because of differences in the substrate-specificities and inhibitor-sensitivities of the different MDR proteins expressed in different tumor cells, proper therapeutic intervention requires an advanced diagnosis and targeted modulator agents. The third generation of MDR modifiers was devised to specifically interact with a particular MDR transporter. One example is the reversins, which are small hydrophobic peptide derivatives that are supposed to interact with Pgp with high affinity and selectivity; their clinical efficiency has yet to be proven.

Regulation of Expression

A number of environmental stimuli are known to affect the expression of the Pgp (*mdr1*) genes. In humans, it has been suggested that *mdr1* may function as a heat shock gene and its expression may be

modulated by the numerous agents that trigger these signaling pathways. It seems likely that cellular stress and Pgp expression are closely linked because this transporter plays a crucial role in the protection of cells from toxic products released during environmental stress. More specifically, it has been found that certain steroid hormones enhance Pgp expression. Enhanced expression does not necessarily lead to enhanced efflux functionality because examples also exist of enhanced expression but decreased functionality.

See Also the Following Articles

Aging and Stress, Biology of; Cancer Treatment; Cholesterol and Lipoproteins; Corticosteroids and Stress; Cytokines, Stress, and Depression; Dexamethasone Suppression Test (DST); Environmental Factors; Genetic Factors and Stress; Hippocampus, Corticosteroid Effects on; Immune Response; Metabolic Syndrome; Oxidative Stress; Stress, Definitions and Concepts of; Synthetic Glucocorticoids; Perinatal Dexamethasone.

Further Reading

- Borst, P. and Oude Elferink, R. (2002). Mammalian ABC transporters in health and disease. *Annual Review of Biochemistry* 71, 537–592.
- Dean, M., Rzhetsky, A. and Allikmets, R. (2001). The human ATP-binding cassette (ABC) transporter superfamily. *Genome Research* 11(7), 1156–1166.
- Glavinas, H., Krajcsi, P., Cserepes, J., et al. (2004). The role of ABC transporters in drug resistance, metabolism, and toxicity. *Current Drug Delivery* 1, 27–42.
- Haimeur, A., Conseil, G., Deeley, R. G., et al. (2004). The MRP-related and BCRP/ABCG2 multidrug resistance proteins: biology, substrate specificity and regulation. *Current Drug Metabolism* 5(1), 21–53.
- Kerb, R., Hoffmeyer, S. and Brinkmann, U. (2001). ABC drug transporters: hereditary polymorphisms and pharmacological impact in MDR1, MRP1 and MRP2. *Pharmacogenomics* 2(1), 51–64.
- Liscovitch, M. and Lavie, Y. (2002). Cancer multidrug resistance: a review of recent drug discovery research. *IDrugs* 5, 349–355.
- Meijer, O. C., Karssen, A. M. and de Kloet, E. R. (2003). Cell- and tissue-specific effects of corticosteroids in relation to glucocorticoid resistance: examples from the brain. *Journal of Endocrinology* 178(1), 13–18.
- Müller, M. B., Keck, M. E., Binder, E. B., et al. (2003). ABCB1 (MDR1)-type P-glycoproteins at the blood-brain barrier modulate the activity of the hypothalamic-pituitary-adrenocortical system: implications for affective disorder. *Neuropsychopharmacology* 28, 1991–1999.
- Renes, J., de Vries, E. G., Jansen, P. L., et al. (2000). The (patho)physiological functions of the MRP family. *Drug Resistance Update* 5, 289–302.

- Schmitz, G. and Kaminski, W. E. (2001). ABC transporters and cholesterol metabolism. *Frontiers in Bioscience* 6, D505–D514.
- Štefkova, J., Poledne, R. and Hubacek, J. A. (2004). ATP-binding cassette (ABC) transporters in human metabolism and diseases. *Physiology Research* 53, 235–243.
- Stein, W. D. (2002). Reversers of the multidrug resistance transporter P-glycoprotein. *Current Opinion in Investigative Drugs* 3(5), 812–817.
- Sukhai, M. and Piquette-Miller, M. (2000). Regulation of the multidrug resistance genes by stress signals. *Journal of Pharmacy and Pharmaceutical Sciences* 268–280.
- Tan, B., Piwnica-Worms, D. and Ratner, L. (2000). Multi-drug resistance transporters and modulation. *Current Opinion in Oncology* 12(5), 450–458.
- Wijnholds, J., Evers, R., van Leusden, M. R., et al. (1997). Increased sensitivity to anticancer drugs and decreased inflammatory response in mice lacking the multidrug resistance-associated protein. *Nature Medicine* 1275–1279.

Multiple Personality Disorder

R P Kluft

Temple University, Philadelphia, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R P Kluft, volume 2, pp 786–790, © 2000, Elsevier Inc.

Introduction
History
Definition and Characteristic Findings
Etiology
Diagnosis
Treatment
Miscellaneous Concerns
Concluding Remarks

Glossary

Alter

A distinct identity or personality state with its own characteristic and relatively enduring pattern of perceiving, relating to, and thinking about the environment and self. An alter has its own identity, self-representation, autobiographic memory, and sense of ownership of its own thoughts, feelings, and actions. Some synonyms are personality, personality state, subpersonality, alter personality, and dis-aggregate self-state.

Dissociation

“Disruption in the usually integrated functions of consciousness, memory, identity, or perception” (American Psychiatric Association, 2000: 519); the segregation of some subsets of information from other subsets of information in a relatively rule-bound manner (Spiegel, 1986: 123).

Ego state

“An organized system of behavior and experience whose elements are bound together by some common principle but that is separated from other such states by boundaries that are more or less permeable” (Watkins and Watkins, 1993: 278). The creation (or worsening) of a problematic condition as a result of the interventions of a helping professional.

Iatrogenesis

Integration

The unification of all personalities into a single coherent identity; the cessation of separateness of a particular personality or group of personalities; the interventions made by a therapist to bring together two or more personalities at a particular point in time.

Personality

An enduring pattern of behavior, adaptation, and inner experience that is pervasive over time and across situations (in general psychology); synonym for alter (in dissociative disorders).

Resolution

An outcome in which a multiple personality patient has achieved smooth function, continuity of contemporary memory, and the relief of distressing symptoms due to the more facile cooperation and collaboration of the remaining personalities. Some personalities may have integrated.

Shift

A change in the personality system such that while the personality or alter in apparent executive control has not changed or been replaced, it is being influenced by different alters behind the scenes than it had been, leading to minor changes in the appearance, actions, and speech of the alter in apparent executive control.

Switch

A transition from one personality being in apparent executive control to another personality being in apparent executive control.

Introduction

Multiple personality (dissociative identity disorder in the United States) has been described as difficult to understand, difficult to diagnose, difficult to treat, and difficult to discuss objectively not only because of its complexity, but also because of the controversies that often surround it. Its study draws upon the literatures of dissociation, hypnosis, memory, cognitive psychology, social psychology, development, trauma, attachment and psychoanalysis, among others, requiring researchers and clinicians to grapple with concerns and issues that often appear remote and obscure.

History

Conditions in which an individual's customary personality or way of being is replaced by an alternative (and often very different) personality or way of being, whose activities are unknown to the customary personality, were virtually worldwide and commonplace prior to the modern scientific era. Possession states were attributed to the intrusion of various forces, spirits, angels, demons, gods, or ancestors. Judeo-Christian possession states included forms in which the intruding entity took over completely, leaving the possessed person without memory for the intruding entity's actions, and forms in which both customary personality and intruding entity were aware of one another and contended for control (somnambulistic and lucid possession, respectively).

As possession was discredited (circa 1775–1800), cases with these phenomena were reported and described in secular terms (e.g., exchanged identity, doubling of the personality). By the 1830s, French physician Antoine Despine and his physician son and nephew had observed and treated series of such patients. Utilizing contemporary methods and magnetism, a precursor of hypnosis, successful treatments and integrations were achieved. Throughout the nineteenth century there was considerable interest in multiple personality.

However, rising forces in psychiatry and psychology eclipsed its study. Bleuler subsumed multiple personality under schizophrenia. Freud abandoned dissociation for repression. Janet was marginalized by psychoanalysis and the discrediting of his mentor, Charcot. Behaviorism disregarded multiple personality. Within a generation, multiple personality was largely forgotten except as a historical curiosity.

Two books, *The Three Faces of Eve* and *Sybil*, brought multiple personality to the attention of North American readers. A small but increasing number of North American clinicians began to report

numbers of such patients (1970s–80s). Sensitized by feminism and the plight of Viet Nam veterans to the sequelae of abuse for women, children and combat veterans, they understood multiple personality as a chronic posttraumatic dissociative condition. Cases were recognized with increasing frequency. This experience was duplicated in The Netherlands and elsewhere.

In the 1990s there were highly polarized debates about whether multiple personality was iatrogenic, instigated and sustained by clinicians' interest in motivating patients to demonstrate the condition's phenomena, and whether the abuses alleged by patients, often recalled after years of apparent amnesia, were false, suggested by leading questions or subtle expressions of interest.

It remains unclear whether multiple personality can be created by iatrogenic factors alone. Some reports indicate that efforts to create conditions similar to multiple personality have been undertaken by various intelligence agencies. Some social psychology manipulations can cause subjects to demonstrate some characteristics of multiple personality under experimental conditions. However, no evidence demonstrates that clinicians exert influences similar to the interventions of either covert agencies or laboratory experiments. While some individuals are subject to develop false memories, many memories that became available after long absence from awareness have proven accurate.

Definition and Characteristic Findings

Criteria

The American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition, text revision (DSM-IV-TR) offers four diagnostic criteria for dissociative identity disorder (multiple personality) (Table 1).

Phenomena

Multiple personality involves problems with identity, memory, thinking, containment, cohesive conation, and the switch process. Problems of identity may include the presence of other identities (alters), depersonalization, an absence of or confusion about identity, alters' impingements on one another, and the combined presence of two or more alters. Problems of memory may include amnesia for past events, losing blocks of contemporary time, uncertainty about whether events did or did not occur, fragmentary recall, and experiencing events as dream-like, unreal, or of uncertain reality. There are often amnesia barriers across the alters, which may be aware or

Table 1 DSM-IV diagnostic criteria for dissociative identity disorder^a

- A. The presence of one or more distinct identities or personality states (each with its own relatively enduring pattern of perceiving, relating to, and thinking about the environment and self).
- B. At least two of these identities or personality states recurrently take control of the person's behavior.
- C. Inability to recall important personal information that is too extensive to be explained by ordinary forgetfulness.
- D. The disturbance is not due to the direct physiological effects of a substance (e.g., blackouts or chaotic behavior during alcohol intoxication) or a general medical condition (e.g., complex partial seizures). Note: in children, the symptoms are not attributable to imaginary playmates or other fantasy play.

^aFrom the American Psychiatric Association (2000), p. 529. Reprinted with permission from the Diagnostic and Statistical Manual of Mental Disorders, Copyright 2000. American Psychiatric Association.

unaware of one another or display directional amnesia. In directional amnesia, alter A may know about alter B and be aware of B's thoughts and activities, while B knows little or nothing about A, is unaware of A's thoughts, and has no recall of what happens when A is in control.

Problems of thinking include cognitive errors, magical thinking, and the toleration of mutually incompatible ideas and perceptions without appreciating the impossibility of their coexisting (trance logic). They stem from distortions of reality made to accommodate to intolerable childhood circumstances.

Problems of containment include the intrusions of alters into one another and leakage of the memories, feelings, and sensations of one alter into the awareness and experience of another.

Problems of cohesive conation (will and intentionality) refer to difficulties in the distribution of executive control across the alters. Patients' sense of control of themselves and their actions may be compromised, along with their sense of ownership, responsibility, and voluntary control of themselves and their actions.

Problems of the switch process refer to changes of executive control that are jarring, incomplete, or so frequent that the person cannot concentrate or pay attention consistently enough to accomplish necessary functions and tasks.

Personalities

Personalities, personality states, subpersonalities, personifications, self-states, and alters are synonyms. They are relatively stable and enduring entities with fairly consistent ways of perceiving, relating to, and thinking about the environment and self. They are experienced as having their own identities,

self-images, personal histories, and senses of ownership of their own activities and mental contents. Personalities usually but not inevitably have names. Some names may refer to attributes or emotions or humiliating insults. Usually a rather passive, guilt-ridden, dependent, masochistic, and depressed personality in apparent control most of the time bears the legal name. Alternate personalities may differ in name, age, race, gender, sexual orientation, areas of knowledge, predominant affect, vocabulary, and apparent intelligence. Alters may or may not be aware of one another. Alters may relate to one another in a potentially complex inner world. They constitute ongoing sets of parallel processes; that is, as one alter is dealing with the outside world, another may be making comments on or to the first, while others are interacting with one another, oblivious to external events. Neuropsychophysiological studies often demonstrate significant differences in different types of personality.

Personalities arise as desperate coping strategies, initially serving adaptive and defensive purposes. Achieving a degree of autonomy and persisting beyond the situations to which they were responses, they may become increasingly problematic and disruptive. Alters are formed by repudiating a sense of self and a representation of self in interactions with others that have become intolerable, severing empathic connection and erecting boundaries between the self that has become intolerable and a more adaptive identity created to manage the adaptation believed to be required. This new alter, envisioned and experienced as real, is accepted and interpreted as real. Its alternate autobiographic memories are endorsed as real.

For example: Lois, a 5-year-old girl, is molested by her previously warm and loving Uncle Ben. She is deeply attached to her uncle and does not want to lose him, afraid of the consequences of telling her parents, confused about the feelings and sensations arising in her, wishing this molestation had never occurred, and wanting someone to rescue her. [Table 2](#) illustrates numerous coping strategies and alters that might be created to embody them. Strategies strongly influence the transformations of identity and autobiographic memory that alters will embody. For example, Bad Lois must repudiate knowledge that would demonstrate that Lois is an innocent victim. Louis not only must disregard attributes and experiences that would compel him to acknowledge he is a little girl, but also may need to hallucinate having a male body with appropriate musculature, facial hair, and genitalia.

Alters often experience themselves as having relationships with one another in an inner world that

Table 2 Coping strategies and alter formation of 'Lois'^a

<i>Cognitive coping strategy</i>	<i>Alter created</i>
This did not happen	A Lois who knows, and a Lois who does not
I must have deserved it	Bad Lois, whose behavior would explain trauma as punishment
I must have wanted it	A sexual alter, Sherrie
I can control it better if I take charge	An aggressively sexual alter, Vickie
I would be safe if I were a boy	Louis, Lois' male twin
I wish I were a big man	Big Jack, based on some person of power who could prevent this
I wish I were the one who could hurt someone and not be hurt	Uncle Ben, or a more disguised identification with the aggressor
I wish I could feel nothing	Jessie, who endures all yet feels nothing
I wish someone could replace me	The Girls, who encapsulate specific experiences of trauma unknown to Lois
I wish someone would comfort me	Angel, with whom Lois imagines herself to be while the body is being exploited and the Girls are experiencing the trauma

^aFrom Kluft (1999), p. 5.

may be experienced as possessing reality and importance that is equally or more compelling than the external world. Inner world events may be reported as if they had occurred in external reality, and vice versa.

Comorbidity

Multiple personalities commonly suffer additional mental disorders. It may be difficult to be sure whether the diagnostic criteria for some of these other disorders are satisfied by symptoms emerging from multiple personality or whether they indicate co-occurring diagnoses that require treatments of their own. Posttraumatic stress disorder, major depression, various substance abuses, borderline personality disorder, other anxiety and affective disorders, somatoform disorders, sexual dysfunctions, eating disorders, and other personality disorders commonly co-occur. Symptoms of another disorder may be found in all or most of the personalities, but sometimes only in particular personalities. Distinguishing between comorbidities and look-alike epiphenomena can prove challenging.

Prognosis may be determined more by the treatability of comorbid conditions than by the multiple personality. For example, a multiple personality with posttraumatic stress disorder and a depression that responds well to medication has a much better prognosis than one with posttraumatic stress disorder, anorexia nervosa, rapid-cycling bipolar disorder, and borderline personality disorder.

Etiology

The predominant theory of the etiology of multiple personality holds that this condition comes into being when a child with high hypnotic potential and problematic support from caretakers and/or attachment issues experiences overwhelming traumatic stress, usually on a repeated basis, and is not protected,

comforted, and supported in a way that alleviates the child's distress. The iatrogenesis hypothesis was noted previously. Many scholars consider both alternatives plausible and do not assume that either precludes the possibility of the other.

Data consistent with the predominant theory include the high hypnotizability of this patient group, documentation of abuse in 95% of two cohorts of young dissociative patients, and epidemiological studies in many nations using reliable and valid screening and diagnostic instruments finding a similar incidence of multiple personality (4–6% of psychiatric inpatients) in many nations with very different degrees of awareness of and sympathy toward the diagnosis of multiple personality. If clinicians with special interest in multiple personality act in such a way as to create it, an assumption of the iatrogenesis hypothesis, significantly higher percentages should be encountered in countries in which clinicians are accused of applying such pressures.

Observations consistent with the iatrogenesis hypothesis are the development of transient multiple personality-like phenomena in laboratory settings, reports of the creation of artificial multiple personality by intelligence agencies, and the fact that multiple personality patients have been known to create additional personalities in response to therapists' suggestions. Those most experienced with multiple personality hold that multiple personality cannot be created by therapists under normal clinical circumstances, but that additional personalities can be created in response to perceived pressures from a therapist.

Diagnosis

It was long thought that multiple personality was rare and sufficiently flamboyant to make its presence self-evident. Multiple personality has proven to be far

from rare, usually covert, and commonly misdiagnosed for an average of 6.8 years. With the demonstration of widespread false negative diagnoses, diagnostic efforts were accorded greater importance.

The Natural History of Multiple Personality

While a small percentage (6%) of adult multiple personalities are florid and overt most of the time, the vast majority either take pains to avoid drawing attention to their situations and/or have only windows of overtness, usually in the context of psychosocial stress or encountering situations that in some way are analogous to overwhelming childhood events or that force them to contend with an abusive figure from their childhood (or someone who is perceived as similar to such a figure). Usually personalities either are relatively quiescent or exert their influence from behind the scenes by shifting or intruding into or otherwise influencing (by persuasion or threat) the personality in apparent executive control. They may contrive to pass as the usually apparent personality.

Multiple personality usually is clandestine in children and difficult to distinguish from various turmoils and early manifestations of psychotic disorders in adolescents. Young adults often work to avoid detection and deny manifestations of the condition. As individuals mature, they usually enter a complex matrix of work, family, and relationships and are less closely connected with important figures of their childhood. Less immediately pressured to keep childhood issues out of awareness and increasingly experiencing situations that are related to or analogous to childhood experiences (including intercurrent adult trauma), they often become symptomatic and seek treatment for what appears to be anxiety or depression. Treating clinicians may recognize or suspect dissociative phenomena.

Diagnostic Approaches

Considering the usually covert nature of multiple personality, it cannot be assumed that it will declare itself in typical clinical interviews or mental status examinations, or even in therapy sessions. Clinicians have derived lists of suggestive signs of multiple personality. Experience-derived suggestive signs include common findings in multiple personalities' histories and indications of dissociated behavior (i.e., having been given many different diagnoses, being told by others of disremembered out-of-character behavior, finding objects, productions, or writing in one's possession that one cannot recall acquiring or creating). Loewenstein developed a mental status that studies six relevant symptom clusters: (1) indications of

multiple personality processes at work (e.g., differences in behavior, linguistic indications, switching/shifting); (2) signs of the patient's high hypnotic potential (e.g., enthrallment, trance logic, out-of-body experiences); (3) amnesia; (4) somatoform symptoms; (5) posttraumatic stress disorder symptoms; and (6) affective symptoms.

Instruments have been developed to screen patients for dissociative phenomena; high scores or particular patterns of response suggest that further evaluation is needed. Structured diagnostic interviews have enabled actual diagnostic assessment. The most well-known of these are the Dissociative Experiences Scale (Bernstein and Putnam) for screening and the Structured Clinical Interview for the Diagnosis of DSM-IV Dissociative Disorders – Revised (Steinberg) for diagnostic purposes.

Differential Diagnosis

The differential diagnosis includes other dissociative disorders, psychotic disorders, affective disorders, borderline personality disorder, partial complex seizures, factitious disorders, and malingering. Evaluators knowledgeable about all of these disorders usually find differential diagnosis to be straightforward. Evaluators unfamiliar with dissociative disorders commonly encounter difficulties because phenomena of many of these conditions overlap.

Treatment

Approaches and Modalities

Treatment resembles a stage-oriented trauma therapy, modified to include work with alters. However, work with some alters and issues may be quite advanced and work with others may be just beginning while still others have not yet been discovered. The treatment may be understood as a series of short-term therapies imbricated within the process of a long-term psychotherapy. Virtually every therapeutic model and modality has been applied to multiple personality. In practice, advocates of different theoretical models and treatment approaches find themselves making many similar interventions. In successful treatments, the pragmatic realities of dealing with multiple personality tend to determine what is done.

Treatment Stages

In the three-stage model of modern trauma therapy outlined by Herman, a phase of safety, in which the patient receives sanctuary and support and is strengthened, is followed by a phase of remembrance and mourning, in which the mind's representation of

its traumatic experiences is explored, processed, and mastered and in which the losses and consequences associated with traumatization are grieved. The mind is reintegrated, and roles and functions are resumed in a phase of reconnection.

In the nine-stage treatment of multiple personality (Kluft, 1999a,b) with multiple personality (1) the psychotherapy is established and (2) preliminary interventions are made to establish safety, develop a therapeutic alliance that includes the alters, and enhance the patient's coping capacities. Then follows (3) history gathering and mapping to learn more about the alters, their concerns, and how the system of alters functions. Then is it possible to begin (4) the metabolism of trauma within and across the alters. As the alters share more, work through more, communicate more effectively with one another, and achieve more mutual awareness, identification, and empathy, their conflicts are reduced, as is contemporary amnesia. They increasingly cooperate and experience some reduction of their differences and senses of separateness. This is called (5) moving toward integration/resolution. More solidified stances toward one's self and the world are reached in (6) integration/resolution. Smooth and functional collaboration among the alters, usually including the blending of several personalities, is called a resolution. Blending all alters into a subjective sense of smooth unity is an integration. Then the patient focuses on (7) learning new coping skills, working out alternatives to dissociative functioning, and resolving other previously unaddressed concerns. Issues continue to be processed, and mastery without resort to dysfunctional dissociation is pursued in (8) solidification of gains and working through. Finally, treatment tapers, and the patient is seen at increasingly infrequent intervals in a stage of (9) follow-up.

Typical Issues

Treatment may prove arduous and challenging to patient and therapist alike. Work with traumatic material can be upsetting and destabilizing. Worse than the material itself is the pain of integrating what patients learn into their perceptions of their relationships with significant others who may appear to have been perpetrators of previously unremembered mistreatment. This is complicated by the fact that it is usually not possible to either validate or invalidate most of the traumatic memories that emerge. Patients should be informed about the possibility that material that emerges and may be useful for treatment may not prove to be accurate.

Processing traumatic memories has been controversial because the accuracy of initially unavailable

memories has been challenged and the affects experienced in association with this processing may cause upset and trigger self-destructive impulses. Occasionally decompensation occurs. Some multiple personalities cannot tolerate such work. However, thus far, reported successful recoveries to the point of integration have involved processing traumatic memories. Studies have demonstrated that many recovered memories of multiple personality patients have been confirmed, and some have been proven inaccurate. Current opinion suggests that deliberate processing of traumatic memories should not proceed unless patients have demonstrated adequate strength and stability for such work. All others should be treated supportively, addressing traumatic memories only when they are intrusive, are disruptive, and cannot be put aside.

Patients typically have periods of wanting to disavow everything said in therapy, trying to banish painful memories of trauma, betrayal, and loss associated with important people in their lives in order to retain relationships and a sense of safety within those valued relationships. Tact, containment, and circumspection are required from therapist and patient alike. The patient should be protected from becoming overwhelmed by and lost in the traumatic material, and treatment should be paced to safeguard the patient's safety and stability. All pressures to get everything out and over with must be resisted.

The alter system is designed to facilitate escape from pain and difficulty or, failing that, to reframe or disguise it. Alters often reenact scenarios that (in their perceptions) are tried and true methods of keeping pain at bay, even if they disrupt the patient's treatment, life, and relationships. Containing and/or minimizing such events is an ongoing challenge to the treatment, rarely completely successful until therapy is well advanced.

Some have advised against working with alters lest separateness be reified and reinforced by their being addressed directly and individually, and recommend speaking in ways that always convey that the patient is a single person. In practice, working directly with alters often may make them more prominent transiently, but as they are worked with, empathized with, and helped to communicate with other alters, their separateness is eroded. All published series of successful integrations have been contributed by authors who work directly with alters.

The therapist should treat all of the personalities with respect, simultaneously appreciating the immediacy, forcefulness, and defensive aspects of their entrenched and subjectively compelling senses of separateness, and that all express aspects of a single

individual, whose personality structure is to have multiple personalities. Interventions to contain alters' dysfunctional behaviors, aggressiveness toward other personalities, self-destructiveness, and irresponsible autonomy (e.g., failing to care for children, who may be seen as belonging to another personality) may prove necessary. The therapist may call upon personalities to work on their particular issues in the treatment and to facilitate their cooperation with the treatment and one another.

Treatment must respect the entirety of the patient's concerns. Specific multiple personality treatment may be deferred repeatedly to address pressing concerns and other mental health issues. For example, a woman with multiple personality whose child develops cancer is not in a position to pursue trauma work.

Patient Subgroups

Some multiple personality patients cannot achieve the goals of the early stages of treatment and move on to process traumatic events and bring personalities together. Those with severe comorbidity and maladaptive character issues may never achieve enough stability to follow that path. Their treatment will involve ongoing efforts to enhance safety and strength and will defer trauma work unless material is intrusive and unavoidable. Therapy prioritizes better coping strategies and enhanced cooperation among alters. An intermediate group will go through phases in which trauma work can be done and phases in which it cannot be tolerated. Therapists should be flexible and err on the side of caution whenever uncertain about the safety of trauma work.

Hypnosis

Multiple personality is associated with high hypnotizability, a stable trait. Hypnosis may occur because a clinician performs an induction procedure, because an individual induces self-hypnosis, or because spontaneous trance is triggered by some stimulus or activity, such as strong emotion or meditation. Therefore, therapists cannot control whether hypnosis occurs in the course of a therapy. Bearing this in mind, therapists should avoid making remarks that might suggest that particular events have taken place or that particular persons were involved in doing harmful things, because they may be heard as powerful hypnotic suggestions.

Notwithstanding such concerns, hypnotic interventions can play a major role in stabilizing multiple personality patients, managing abreactions, containing powerful emotions, accessing alters for therapeutic work, and promoting integration.

Miscellaneous Concerns

Forensic Aspects of Multiple Personality

The relevance of multiple personality as a defense varies tremendously depending upon (1) diverse requirements for the insanity defense in different jurisdictions and (2) the phenomenology in particular multiple personalities. Criminal acts may have taken place because of the impact of multiple personality upon a person's capacity to know his or her own mind and control its various aspects, without thereby satisfying legal definitions of insanity.

The assessment of defendants claiming to have or thought to have multiple personality should proceed with a keen eye toward discerning both factitious and malingered presentations, informed by a state-of-the-art knowledge of multiple personality drawn from the modern clinical literature.

The Impact of Controversy

Controversy magnifies the burden of having, treating, or researching multiple personality. Multiple personality patients must contend with the impact of efforts to deny the reality of their condition and the credibility of what they believe to be their autobiographic memories, attacks on the competence and credibility of those who treat them, and criticism and rejection from those who believe, either accurately or inaccurately, that the multiple personality patient is revealing their secrets or making scandalous accusations against them or those they love in treatment. Therapists treating such patients must contend with ongoing questioning of their efforts, methods, and intentions and accusations that they are encouraging iatrogenesis and confabulated false memories. Researchers have difficulty obtaining funds to expand our knowledge of this condition and may be endangering the trajectory of their careers by working in this area.

Concluding Remarks

Notwithstanding the complexity and difficulty associated with multiple personality, this disorder often proves treatable to a satisfactory outcome and offers unique insights into the impact of overwhelming stressors upon human memory and identity and into the neuropsychophysiology of mental structures and functions. It is moving into the mainstream of the mental health professions and should be of concern to all mental health professionals.

See Also the Following Articles

Coping Skills; Dissociation; Hypnosis.

Further Reading

- Bernstein, E. and Putnam, F. W. (1986). Development, reliability, and validity of a dissociation scale. *Journal of Nervous and Mental Disease* **174**, 727–734.
- Boon, S. and Draijer, N. (1993). *Multiple personality in the Netherlands: a study on reliability and validity of the diagnosis*. Amsterdam: Swets & Zeitlinger.
- Coons, P. M. (1980). Multiple personality: diagnostic considerations. *Journal of Clinical Psychiatry* **141**, 330–336.
- Kluft, R. P. (1999a). Current issues in dissociative identity disorder. *Journal of Practical Psychiatry and Behavioral Health* **5**, 3–19.
- Kluft, R. P. (1999b). An overview of the psychotherapy of dissociative identity disorder. *American Journal of Psychotherapy* **55**, 289–319.
- Kluft, R. P. (2005). Diagnosing dissociative identity disorder. *Psychiatric Annals* **35**, 633–643.
- Kluft, R. P. and Fine, C. G. (eds.) (1993). *Clinical perspectives on multiple personality disorder*. Washington, D.C.: American Psychiatric Press.
- Loewenstein, R. J. (1991). An office mental status examination for complex chronic dissociative symptoms and multiple personality disorder. *Psychiatric Clinics of North America* **14**, 567–604.
- Putnam, F. W. (1989). *Diagnosis and treatment of multiple personality disorder*. New York: Guilford.
- Putnam, F. W. (1997). *Dissociation in children and adolescents: a developmental perspective*. New York: Guilford.
- Ross, C. A. (1997). *Dissociative identity disorder: diagnosis, clinical features, and treatment of multiple personality*. New York: Wiley.
- Spiegel, D. (1986). Dissociating damage. *American Journal of Clinical Hypnosis* **29**, 123–131.
- Steinberg, M. (1994). *Structured clinical interview for DSM-IV dissociative disorders, revised*. Washington, D.C.: American Psychiatric Press.
- Tutkun, H., Sar, V., Yargiuc, I., Ozpulat, T., Yanik, M. and Kiziltan, E. (1998). Frequency of dissociative disorders among psychiatric inpatients in a Turkish university clinic. *American Journal of Psychiatry* **155**, 800–805.
- Watkins, H. and Watkins, J. (1993). Ego-state therapy in the treatment of dissociative disorders. In: Kluft, R. P. & Fine, C. G. (eds.) *Clinical perspectives on multiple personality disorder*, pp. 277–299. Washington, D.C.: American Psychiatric Press.

Multiple Sclerosis

A T Reder

University of Chicago, Chicago, IL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A T Reder, volume 2, pp 791–795, © 2000, Elsevier Inc.

Multiple Sclerosis – Overview

Immune Abnormalities in Multiple Sclerosis

Neuroendocrine and Sympathetic Nervous Systems in Multiple Sclerosis

Stress Inhibits Experimental Allergic Encephalomyelitis

Could Stress Affect Immune Function?

Do Neurosurgical Lesions or Physical Trauma Incite Multiple Sclerosis Plaques?

Does Stress Cause Multiple Sclerosis? Actual Data from Multiple Sclerosis Patients

Medical-Legal Implications

Glossary

Cytokine A protein secreted by the immune cells that affects immune responses.

Experimental allergic encephalomyelitis (EAE)

An autoimmune disease caused by the injection of brain proteins (myelin oligodendrocyte protein, MOG; proteolipid protein, PLP; myelin basic protein, MBP); in some ways, a model of multiple sclerosis.

Human leukocyte antigen (HLA) Interferon (IFN)- γ

A protein on the membrane of most cells that is recognized by immune cells.

A cytokine secreted by lymphocytes that is involved in Th1 immune responses. Treatment with this IFN causes exacerbations of multiple sclerosis.

Oligoclonal bands

Immunoglobulin G of restricted heterogeneity, which forms several discrete bands on electrophoresis of cerebrospinal fluid. Myelin-forming cells of the central nervous system.

Oligodendroglia Plaque

Inflammatory and scarred (gliosis) central nervous system lesions in multiple sclerosis.

T helper 1 (Th1) cells

T helper lymphocytes that secrete IL-2 and IFN- γ and mediate delayed-type hypersensitivity immunity.

T helper 2 (Th2) cells

T helper lymphocytes that secrete IL-4 and IL-10 and amplify humoral immunity.

Ask a roomful of multiple sclerosis (MS) patients whether MS attacks are caused by stress, and 9 out of 10 will say “of course.” However, objective and quantitative measurement is difficult for both MS exacerbations and stress. Could stress affect immunity and could immune changes in turn affect MS? More important, does stress directly affect MS?

Multiple Sclerosis – Overview

Multiple sclerosis is a disease caused by intermittent perivenular inflammation of areas of the central nervous system (CNS). These plaques affect every function controlled by the brain and spinal cord.

MS usually appears between the age of 20 and 50, in women twice as often as men, and is most common in people of northern European ancestry. A complex interaction of multiple genes and the environment may lead to MS. HLA-DR2 predisposes to MS in northern Europeans, but other HLA types are linked to MS in Sardinia, Turkey, and the Middle East. MS is more common in areas far from the equator. Migration studies suggest that an environmental influence before puberty also affects the development of MS. Viruses are the main environmental culprits, but no single virus has been convincingly linked to MS.

Three factors consistently increase MS exacerbations: (1) virus infections (there may also be some effect of prostate, bladder, and sinus infections, which provoke MS symptoms and possibly cause exacerbations), (2) delivery of a baby (there are more exacerbations during the several months postpartum than during pregnancy), and (3) IFN- γ (injections of IFN- γ causes a fourfold increase in MS flares).

Immune Abnormalities in Multiple Sclerosis

The immune system is abnormal in MS. Inflammatory cells destroy oligodendroglia and some neurons in MS plaques. Macrophages far outnumber lymphocytes in these lesions. The inflammation in brain tissue is reflected by small numbers of activated immune cells in the cerebrospinal fluid (CSF; T cells outnumber B cells and macrophages in CSF) and excessive levels of immunoglobulins (IgG; seen in oligoclonal bands). Some T cells and some antibodies recognize virus or brain antigens (myelin oligodendrocyte protein, MOG; proteolipid protein, PLP; myelin basic protein, MBP; heat shock proteins; S-100 protein; and α B-crystallin), but most recognize other targets (e.g., nonsense antibodies). Cytokines in the CSF such as tumor necrosis factor (TNF)- α and interleukin (IL)-6 correlate with disease activity and blood–brain barrier (BBB) damage. In culture, very high doses of IFN- γ and

TNF- α kill oligodendroglia; at lower doses, TNF- α still interferes with oligodendroglial function.

The peripheral blood is also abnormal in MS. Changes in the blood sometimes precede clinical attacks. Several weeks before exacerbations, mitogen-activated white blood cells secrete excessive amounts of IFN- γ and TNF- α , and lymphocytes show more IFN- γ -induced $[Ca^{2+}]_i$ influx. All of these peripheral immune changes could amplify the antibrain immune response.

During attacks, suppressor cell function drops, IL-12 increases and is likely to induce IFN- γ , and expression of the costimulatory CD80 (B7.1) protein increases on lymphocytes. These changes may lead to antibrain immune reactions. Treatments for MS partially reverse immune abnormalities, and reduce exacerbations by 33%. Interferon-beta reduces inflammatory cytokines, increases suppressor cell function, and prevents immune cells from entering the brain. Glatiramer acetate shifts cells from pro-inflammatory (Th1) to anti-inflammatory (Th2) immunity. Both agents may also induce neurotrophic factors in lymphocytes.

Neuroendocrine and Sympathetic Nervous Systems in Multiple Sclerosis

There are other abnormalities in MS possibly linked to stress responses. MS and Crohn's disease occur together more often than expected. Both are relapsing/remitting diseases with possible viral/bacterial triggers, show similar immune abnormalities, and have anecdotal links to stress. However, typical autoimmune diseases such as systemic lupus erythematosus are not linked to MS.

There is an increased incidence of depression in MS patients and possibly in their families. The diagnostic laboratory test for depression – the dexamethasone suppression test – is abnormal in active MS, at least as often as in depression.

Some patients have elevated cortisol and enlarged adrenal glands. Steroids do not inhibit proliferation and do not induce glucocorticoid receptor translocation in MS lymphocytes compared to normal lymphocytes. Responses to corticotropin releasing hormone (CRH) are normal, but arginine vasopressin (AVP) stimulation of the pituitary is subnormal. These abnormalities may all be a consequence of chronic high cortisol levels caused by stress or inflammation, which could change the set point for steroid responses. (Rats susceptible to EAE have similar neuroendocrine abnormalities; EAE-sensitive strains are less responsive to steroids.)

Glucocorticoid therapy curtails acute inflammation in MS plaques. Steroid therapy may correct the

abnormal endogenous glucocorticoid feedback that otherwise allows unchecked immune responses. As a corollary, some MS patients may need higher levels of steroid therapy than normal subjects to inhibit immune functions. Note that a sudden drop in steroid levels may enhance immunity.

The sympathetic nervous system (SNS) is abnormal in MS, especially in patients with progressive disease. MS patients with progressive disease often have cold purple legs, reflecting decreased control of blood vessel tone. These patients have diminished R-R wave variation on an electrocardiograph (EKG) during inspiration and reduced sympathetic skin responses. This response is a sympathetic discharge after a peripheral stimulus. All these abnormalities are likely to be from a loss of sympathetic control caused by MS plaques in the spinal cord.

The SNS innervates immune organs. The loss of sympathetic tone may have consequences for the immune system in MS. When nerves are cut, the target tissue develops denervation supersensitivity to the relevant neurotransmitter. In MS, β -adrenergic receptors (β -AR) and responses to adrenergic agonists (which induce cAMP) are increased in CD8⁺, CD28⁺ T lymphocytes (suppressor T cells). These cells are responsible for the reduced concanavalin A-induced suppressor function seen during the exacerbations and progression of MS. Stress elevates catecholamines and could theoretically destabilize immune regulation if MS lymphocytes were overly sensitive to adrenalin.

Muscarinic receptors are increased on CD4⁺ lymphocytes (T helper cells). Vagal output could potentially affect the development of thymic tolerance or activate T helper cells. Stress effects on the parasympathetic nervous system are not well studied.

Stress Inhibits Experimental Allergic Encephalomyelitis

Immunization with brain proteins (MBP, PLP, and MOG) causes autoimmune EAE in one animal model of MS. Some stress effects on EAE could therefore be instructive. However, the murine immune system has significant differences from the human immune system, the initial antigenic trigger in MS is unknown, and some drugs have completely opposite effects in EAE and MS. For these reasons, the effects of stress in EAE may not always correspond to the effects in MS.

Glucocorticoids suppress EAE. Merely picking up a rat ameliorates EAE. Handling stress triggers a surge of corticosterone, the rat homolog of human cortisol. Restraint stress also suppresses acute EAE, but it does not affect established EAE. In parallel with these

endogenous steroid effects, treatment with dexamethasone markedly inhibits all stages of EAE. Extending these data to humans suggests that stress should inhibit MS. Anecdotal reports that MS improves after skydiving fit into this line of thought, but remain anecdotes.

When stress is stopped, there is a rebound of severe inflammation in EAE. Similarly, if the dexamethasone used to suppress EAE is suddenly withdrawn, rats develop very severe disease within 5 days. This rebound worsening is seen following a 3-day pulse of dexamethasone even in rats that have recovered from EAE. A pulse of high-dose steroids is a standard treatment for MS exacerbation, yet clinical rebound is seldom seen in MS. Again, EAE differs from MS.

Catecholamines also inhibit EAE through β_2 -adrenergic receptors on lymphocytes. Terbutaline, a β_2 -adrenergic agonist; misoprostal, a prostaglandin E analog; and pentoxifylline, a phosphodiesterase III inhibitor, all elevate cAMP. These agents increase levels of cAMP in lymphocytes, inhibit Th1 responses, and reduce EAE severity.

Could Stress Affect Immune Function?

Stress appears to inhibit the immune system. After intense studying for final exams, medical students are more likely to develop upper respiratory tract infections. Similarly, bereavement (loss of a loved one), divorce, caring for a spouse with Alzheimer's disease, and surgery for breast cancer are associated with excess infections and decreased *in vitro* immune function.

Acute stress and chronic stress both affect immunity. Acute stress causes a rise in serum catecholamines and cortisol; these suppress immune function and also affect immune cell distribution and trafficking. Chronic stress enforces the hypothalamic AVP response to stress, but reduces steroid responses, and so may increase cellular immunity. Products of this stress response affect immunity. Agents that elevate cAMP in lymphocytes inhibit EAE. When used to treat MS exacerbations, glucocorticoids speed up recovery from exacerbations but have no long-term benefit. Steroids cause apoptosis of T cells that are in contact with astrocytes in inflammatory sites. This argues that stress should prevent MS exacerbations. However, although terbutaline does decrease fatigue, treatment of MS patients with cAMP elevators has no significant effect on disease activity.

Two factors could boost immune function after stress. A rebound in immune function is possible after stress is withdrawn. Second, virus infections appear to trigger MS exacerbations. One of three upper respiratory tract infections is followed by an

MS flare, possibly because viruses induce IFN- γ and other inflammatory cytokines. It therefore seems logical to invoke a chain of events that ultimately cause attacks of MS – stress, immunosuppression, virus infection, and then an exacerbation of MS.

Do Neurosurgical Lesions or Physical Trauma Incite Multiple Sclerosis Plaques?

Theoretically, direct trauma to brain tissue damages the BBB and could be a nidus for an MS plaque. Stress, frequently associated with surgery or injury could also potentially affect inflammation at the site of trauma. What is the evidence?

One to 2 days after direct trauma to a rat brain or spinal cord, mononuclear cells invade the damaged area. Seven days after injury, lymphocytes isolated from lymph nodes induce adoptive-transfer EAE. At later times, lymph node cells are not encephalitogenic, indicating that endogenous suppressor mechanisms prevent further immune reactivity to the brain. Thus, there is a critical period following trauma when immune cells are predisposed to attack the brain. A critical period is more difficult to define in MS, but some studies have looked at exacerbations during the period following trauma.

EAE is prone to develop at an electrolytic lesion – damage to the BBB, heat (releasing heat shock proteins), hemorrhage (including white blood cells), and inflammation are potential mechanisms. Moreover, trauma can induce diffuse changes in the brain. At 4–7 days after trauma, the BBB becomes leaky in the gray matter contralateral to the insult.

Early reports suggested that trauma caused the onset of MS. Contrary to a few case reports, anesthesia or spinal taps do not trigger attacks of MS. Most experts now believe that there is no convincing connection between trauma and exacerbations of MS. There remains some debate, although it mainly takes place in courtrooms.

MS plaques are twice as frequent in the cervical spinal cord as other cord sites. It has been argued that flexion and extension of the neck stretch the cord and predispose to MS lesions. Some also suggest that the effect would be more pronounced at the level of a herniated disk. All of these arguments are irrelevant in the thoracic area, which is only slightly mobile compared to the cervical cord and is infrequently the site of herniated disks. The thoracic cord is a relatively frequent site of MS lesions, but is thinner than the cervical cord so the number of plaques per gram of tissue is not so different. A related condition, transverse myelitis, is actually more common in the thoracic area. Because two diseases can have additive effects, the removal of a symptomatic disk

in an MS patient can decrease some neurological symptoms.

Direct trauma to the CNS under controlled conditions was studied in MS patients with severe tremors. In the 1970s, several patients treated with thalamic ablation developed MS plaques along the needle track. Other patients, however, had similar surgery but had no lesions along the track. More recent treatment of tremor with thalamic lesions or stimulation has not been reported to cause MS plaques. The discrepancy between old and new studies may lie in the neurosurgical technique. Hot electrodes could generate heat shock proteins (B. G. W. Amason, personal communication), which cross-react with MBP. However, the older studies were anecdotal and at the time imaging was nonexistent; plaques may have been present before the needle was inserted.

The biopsy of a large, active MS plaque should be a provocateur of nearby inflammation. As with other recent neurosurgical procedures in MS patients, there are no reports of biopsies being followed by exacerbations.

Does Stress Cause Multiple Sclerosis? Actual Data from Multiple Sclerosis Patients

All the arguments described so far suggesting that stress and immunity cause MS exacerbations, are theoretical. The only way to resolve these speculations is through controlled studies of the effect of stress on the course of MS.

There is a large literature on the effect of physical and psychological stress on MS. In one of the first recorded cases of MS, Augustus d'Este wrote in 1822 (Murray 2005) that decreased vision followed the funeral of “a relative for whom I had the affection of a son.” Much of the early literature is anecdotal or is based on retrospective, uncontrolled studies.

MS symptoms are intermittent, have highly variable severity, and involve disparate areas of the nervous system. Moreover, even without new active inflammation, MS symptoms can be mimicked if old subclinical lesions are compromised. Elevation of body temperature, external heat or humidity, and infections (e.g., upper respiratory, urinary tract, lung, prostate) lower the safety factor for electrical impulses along a nerve and amplify the prior loss of function in plaques. There are also many patient reports of acute worsening with menses, anxiety, stress, fatigue, malaise, and depression. Thus, various forms of stress could make MS symptoms worse without changing immunity or immune function.

These pseudoexacerbations are likely to confound the clinical exam and history of true MS

exacerbations. MS symptoms are also quite subject to selective memory in the patient's search for the cause of the exacerbation. The best way to examine any connection between stress and MS is with controlled prospective studies by investigators with extensive MS experience. Only a few have been performed.

For example, in controlled prospective studies Sibley and Siva showed that emotional stress is not linked to MS exacerbations. They also found no convincing link between physical trauma and MS exacerbations. Buljevac and Ackerman reported a link, but recall bias may have contributed to this result. Severe stress, such as gunshot wounds and SCUD missile attacks, actually seemed to reduce exacerbations.

Medical-Legal Implications

Stress or trauma often coincides with MS exacerbations. Trauma to the brain can cause CNS dysfunction, and it is often difficult to sort out whether trauma or MS is responsible for a given symptom. Moreover, stress or trauma could amplify the effects of prior MS lesions. Confounding factors such as elevated temperature, infections, and fatigue could also intensify MS symptoms.

Some argue that in some patients, at some times, stress could modify immunity and possibly trigger attacks of MS. However, this position is highly speculative and probably wrong in a disease with very complex clinical symptoms. There is no clear evidence that stress or trauma causes the onset of MS or that stress or trauma triggers attacks of MS.

See Also the Following Articles

Brain Trauma; Central Stress Neurocircuits; Immune Suppression.

Further Reading

- Ackerman, K. D., Heyman, R., Rabin, B. S., et al. (2002). Stressful life events precede exacerbations of multiple sclerosis. *Psychosomatic Medicine* 64, 916–920.
- Andersen, B. L., Farrar, W. B. and Golden-Kreutz, D. (1998). Stress and immune responses after surgical treatment for regional breast cancer. *Journal of the National Cancer Institute* 90, 30–36.
- Anlar, B., Karaszewski, J. W., Reder, A. T., et al. (1992). Increased muscarinic cholinergic receptor density on CD4⁺ lymphocytes in progressive multiple sclerosis. *Journal of Neuroimmunology* 36, 171–177.
- Arnason, B. G. W. and Reder, A. T. (1994). Interferons and multiple sclerosis. *Neuropharmacology* 17, 495–547.
- Buljevac, D., Hop, W. C. J., Reedeker, W., et al. (2003). Self reported stressful life events and exacerbations in multiple sclerosis: prospective study. *British Medical Journal* 327, 646–651.
- Burgerman, R., Rigamonti, D., Ranelle, J. M., et al. (1992). The association of cervical spondylosis and multiple sclerosis. *Surgical Neurology* 38, 265–270.
- Karaszewski, J. W., Reder, A. T., Anlar, B., et al. (1993). Increased high affinity beta-adrenergic receptor densities and cyclic AMP responses of CD8 cells in multiple sclerosis. *Journal of Neuroimmunology* 43, 1–7.
- Levine, S., Strebel, R., Wenk, E. J., et al. (1962). Suppression of experimental allergic encephalomyelitis by stress. *Proceedings of the Society for Experimental Biology and Medicine* 109, 294–298.
- Michelson, D. and Gold, P. W. (1998). Pathophysiologic and somatic investigations of hypothalamic-pituitary-adrenal axis activation in patients with depression. *Annals of the New York Academy of Sciences* 840, 717–722.
- Murray, T. J. (2005). *Multiple sclerosis: The history of a disease*. New York: Demo.
- Popovich, P. G., Stokes, B. T. and Whitacre, C. C. (1996). Concept of autoimmunity following spinal cord injury: possible roles for T lymphocytes in the traumatized central nervous system. *Journal of Neuroscience Research* 45, 349–363.
- Poser, C. M. (1994). The role of trauma in the pathogenesis of multiple sclerosis: a review. *Clinical Neurology and Neurosurgery* 96, 103–110.
- Reder, A. T. and Arnason, B. G. W. (1985). Immunology of multiple sclerosis. In: Vinken, P. J., Bruyn, G. W., Kiawans, H. L. & Koetsler, J. C. (eds.) *Handbook of clinical neurology: demyelinating diseases*, pp. 337–395. Amsterdam: Elsevier Science Publishers.
- Reder, A. T., Thapar, M. and Jensen, M. (1994). A fall in serum glucocorticoids provokes experimental allergic encephalomyelitis: implications for treatment of inflammatory brain disease. *Neurology* 44, 2289–2294.
- Sibley, W. A., Bramford, C. R., Clark, K., et al. (1991). A prospective study of physical trauma and multiple sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry* 54, 584–589.
- Siva, A., Radhakrishnan, K., Kurland, L. T., et al. (1993). Trauma and multiple sclerosis: a population-based cohort study from Olmsted County, Minnesota. *Neurology* 43, 1878–1882.

Multiple Trauma

P N Soucacos

University of Athens School of Medicine, Athens,
Greece

E O Johnson

University of Ioannina School of Medicine, Ioannina,
Greece

© 2007 Elsevier Inc. All rights reserved.

Multiple Trauma: Background

Physiological Responses to Trauma

The Neuroendocrine Stress Response to Trauma

Time-Dependent Changes in the Stress Response

Immune System Response to Trauma

Posttraumatic Psychological Complications

Glossary

<i>Hemorrhagic shock</i>	Significant acute blood loss leading to hypovolemia and hypotension.
<i>Open fracture</i>	A compound fracture with an open wound down to the bone; usually associated with severe vascular impairment.
<i>Posttraumatic stress disorder (PTSD)</i>	The development of characteristic symptoms after a traumatic event, including numbed responsiveness to stimuli, autonomic and cognitive dysfunction, and dysphoria.

Multiple Trauma: Background

In most Western societies, traumatic injuries represent the third leading cause of death after cardiovascular diseases and tumors. Of the accidents that occur, approximately 60% take place at home or during sport activities, 27% take place at work, 9% are related to traffic accidents, and approximately 4% are the result of violent actions. However, of the polytraumatic injuries, over 70% are attributed to traffic accidents, followed by accidents at home (20%) and work (10%). More than one-half of the patients with multiple trauma have fractures, dislocations, or both.

In the polytrauma patient, the management of orthopedic injuries can have profound effect on the patient's ultimate functional recovery and may be life- or limb-saving. Life- and limb-threatening musculoskeletal problems include hemorrhage from wounds and fractures, infections from open fractures, limb loss from vascular damage and compartment syndrome, and loss of function from spinal or peripheral neurological injuries. The frequency of

pulmonary complications, adult respiratory distress syndrome (ARDS), fat emboli, and pneumonia has been correlated to the timing and type of treatment of long bone fractures, with a statistically significant increase in morbidity, pulmonary complications, and length of hospitalization when stabilization of a major fracture was delayed. Consequently, since the early 1990s, emphasis has been on early total care of multiply injured patients, including fracture stabilization.

Multiple trauma injuries are frequently complicated with open fractures. Open fractures types IIIB and IIIC are extremely severe injuries that often may lead to limb amputation. These fractures are usually caused by a high-energy impact that results in extensive bony comminution or segmental bone loss, pronounced soft tissue injury including extensive skin loss, tendon and nerve damage, muscular and periosteal stripping from the bone (type IIIB), and severe circulatory compromise of the extremity related to the complete ischemia produced secondary to trauma of the major vessels (type IIIC).

The incidence of amputation rate of type IIIC fractures ranges from 60 to 100%. The aim today is not just to salvage the limb that has sustained this compound injury but to produce a functional painless extremity with at least protective sensation. Several variables play a decisive role in determining the outcome and success of preserving a limb. These include the extent and severity of vascular injury, the bone and soft tissue damage, the type and duration of limb ischemia, the patient's age, the time elapsed from the initial injury to surgery, and concomitant organ injuries. To determine which limbs can be salvaged and which should be amputated primarily, numerous scales using a variety of criteria have been proposed assessing the severity of injury. They include the mangled extremity syndrome, the mangled extremity severity score, and NISSA (N = nerve, I = ischemia, S = soft tissue injury, S = skeletal injury, S = shock, and A = age of patient).

Patients with multiple injuries pose a significant challenge. Optimal management is essential early on after trauma to avoid long-term complications, sepsis, organ failure, and subsequent increased late mortality. The principles of management now include simultaneous assessment and resuscitation, with complete physical examination and diagnostic studies as appropriate to establish priorities for life-saving surgery. Trauma management can be divided into four periods ([Table 1](#)). The primary goal during the acute

Table 1 Periods of trauma management

	<i>Time after trauma</i>
Acute or resuscitation period	1–3 h
Primary or stabilization period	1–72 h
Secondary or regeneration period	3–8 days
Tertiary or rehabilitation period	after 8 days

or resuscitation period is to establish adequate ventilation, maintain circulation (intravascular volume and cardiac function), and assess neurological state. The primary or stabilization period of treatment starts when all vital functions have been stabilized by achieving adequate ventilation, hemodynamic stability, and control of intracranial or internal hemorrhage. During this stage, priorities for surgical treatment after resuscitation are given to brain injuries, eye and facial injuries, progressive compression of the spinal cord, visceral injuries, open fractures, fractures with concomitant injuries to major vessels, pelvic ring fractures, and unstable spine fractures. The secondary period is a phase of regeneration. Reconstructive treatments during this phase include secondary wound closure and soft-tissue reconstruction, osteosynthesis of upper extremity fractures, and complex joint reconstruction. After eight days, in the tertiary period, the prognosis of the polytrauma patient is usually apparent. If recovery continues, final reconstructive operations can be performed, including bone grafting at sites of massive bony defects, special soft-tissue reconstructions, definitive closure of amputation sites, and any remaining procedures that were postponed. In cases with organ dysfunction or respiratory distress (e.g., ARDS), further surgical procedures cannot be considered.

Physiological Responses to Trauma

Early death following multiple traumatic injuries is usually attributed to primary brain injury or significant blood loss (hemorrhagic shock), whereas late traumatic death is usually the result of secondary brain injury and host defense failure. The host defense response system is triggered by a series of insults that occur primarily or secondarily (Table 2). The host defense response is characterized by the local and systemic release of various factors (hormonal mediators, pro-inflammatory cytokines, acute phase proteins, etc.) that make up the systemic inflammatory response.

Adjustments in homeostatic mechanisms occur following trauma in an effort to maintain homeostasis and ensure wound healing. Multiple factors, including

Table 2 Primary and secondary physiological insults

<i>Primary insults</i>	<i>Secondary insults</i>
Hypoxia	Ischemia
Hypotension	Reperfusion injury
Organ and soft tissue injury	Compartment syndrome
Fractures	Operative interventions infections

Table 3 Acute and prolonged responses to traumatic stress

<i>Acute responses</i>	<i>Prolonged responses</i>
Activation of blood-clotting mechanisms	Immunological alterations: mobilization of leukocytes, increased T cells and macrophages, increased acute-phase plasma proteins
Shift of body fluids from extravascular compartment to restore blood volume	Inflammatory cells invade injured area
Blood flow redistributed to ensure perfusion of vital organs	Increased fibroblasts for collagen scaffolds for wound repair
Respiratory and renal functions compensate acid–base balance	

diminished blood volume, tissue underperfusion, extensive tissue damage, and invasive infections initiate the stress-response cascade via the neuroendocrine system. As a result of these initial physiological adjustments, tissue perfusion is maintained, and vital organ function and wound repair are supported.

Major traumatic injury is associated with rapid blood loss, tissue underperfusion, massive cellular damage, and disturbances to vital organ function. Response to the traumatic stress include local changes at the injury site and system changes that occur either acutely or gradually with have a prolonged responses (Table 3).

The events that take place following trauma are generally graded – the more severe the injury, the greater the response. In addition, the responses to trauma change over time and can be divided into acute-phase and chronic-phase changes. The acute-phase changes occur immediately after injury and are characterized by a fall in metabolic functions and a decrease in core temperature and by increased levels of stress response hormones. Many of the changes observed acutely are related to hypovolemia. With the restoration of blood flow and with time, the patient's responses change. In the chronic phase, the metabolic rate rises, body temperature increases, and blood insulin levels reach normal or can be increased (Table 4). The chronic phase following injury is characterized by a hypermetabolic state. The degree of hypermetabolism (increased oxygen production) is generally related to the severity of the injury. Patients

Table 4 Time course in responses to multiple trauma

	<i>Acute phase</i>	<i>Chronic phase</i>
Blood glucose	Elevated	Normal or slightly elevated
Glucose production	Normal	Increased
Free fatty acids	Elevated	Normal
Insulin levels	Low	Normal or elevated
Catecholamines	Elevated	High to normal
Blood lactate	Elevated	Normal
Oxygen consumption	Low	Elevated
Cardiac output	Low	Increased
Core temperature	Low	Elevated

with long-bone fractures exhibit a 15–25% increase in metabolic rate, whereas those with multiple injuries increase metabolic needs by 50%. Patients with severe burn injuries covering over 50% of their body may demonstrate a resting metabolic rate that is twice basal levels.

Multiple trauma injury is a clinical condition that results in a marked activation of neuroendocrine stress response aimed at restoring hemodynamic and metabolic homeostasis. The stress response is activated by diverse physical conditions (e.g., ischemia and glucopenia) and is primarily mediated by the hypothalamic-pituitary-adrenal (HPA) axis and sympathetic systems. The activation of the stress response system appears to be critical for survival from injury.

The Neuroendocrine Stress Response to Trauma

Traumatic injury poses a significant psychological and physiological threat, and it provokes an immediate neuroendocrine response. Neural input from the cerebral cortex, injured tissues, and receptors detecting fluid loss results in increased adrenocorticotrophic hormone (ACTH), growth hormone (GH), prolactin, and arginine vasopressin (AVP) release from the pituitary and in a general activation of the sympathetic nervous system with increases in epinephrine and norepinephrine concentrations. Secondary changes include the stimulation of cortisol and aldosterone. The duration of these responses generally depend on the severity of the injury and varies among hormones. The autonomic nervous system provides a rapidly responding mechanism to control a wide range of functions, including cardiovascular, respiratory, gastrointestinal, renal, endocrine, and other systems.

The stress system consists of three central and two peripheral components. The central components of the stress system comprise the locus ceruleus in the

reticular formation, which regulates arousal; the brain-stem centers of the autonomic system, which regulates sympathetic/adrenomedullary function; and the hypothalamic paraventricular nucleus (PVN), which regulates adrenocortical function. The peripheral components consist of the sympathetic/adrenomedullary system and the pituitary-adrenal axis. The interactions among the various components of the stress system are numerous and complex. Corticotropin releasing hormone (CRH)-secreting neurons of the lateral paraventricular nucleus project toward the arousal and sympathetic system in the hindbrain, and conversely, catecholaminergic fibers from the locus ceruleus and central sympathetic system project via the ascending noradrenergic bundle to the PVN in the hypothalamus. The activation of the CRH neurons in the PVN results in the release of CRH into the hypophyseal portal system and the stimulation of the arousal and sympathetic centers in the brain stem in a positive reverberating feedback loop.

Both endorphins and ACTH secreted by the pro-opiomelanocortin (POMC) neurons of the arcuate nucleus exert inhibitory effects on CRH secretion. As POMC neurons are stimulated by CRH, they provide another negative feedback control loop on the HPA axis. ACTH stimulates the release of cortisol from the adrenal cortex, the latter which feeds back in a negative fashion, both at the level of the pituitary and of the hypothalamus, forming the long feedback loops. The activation of the stress system has direct consequences on the function of other major systems, particularly those responsible for reproduction, growth, and immunity.

Today, the role of AVP in the control of ACTH secretion and its involvement in various stress response paradigms is generally accepted. AVP is released from two sites in the hypothalamus; the parvocellular division of the PVN, where CRH is also formed, and from the magnocellular neurons of the supraoptic nucleus (SON) and the PVN. Recent evidence supports an intricate interaction of AVP with CRH and with glucocorticoids, the inhibitory feedback component of the HPA axis.

It is clear that AVP plays an important role in the stress response. Chronic stress paradigms have shown with relative consistency a shift in the CRH: AVP ratio that may relate to a differing feedback sensitivity of AVP to glucocorticoid feedback restraint or a greater responsivity of AVP, over CRH, to chronic stimulatory stress input. AVP serum concentrations significantly increase in the critically ill patient population. However, the lack of a correlation between AVP concentrations and hemodynamic parameters suggests a complex dysfunction of the vasopressinergic system in critical illness.

Extremely low corticosteroid-binding globulin (CBG) levels in early-stage multitrauma is reflected in a concomitant higher free cortisol index, indicative of higher free cortisol levels. This suggests that CBG plays an active role in the glucocorticoid response to severe stress and the regulation of cortisol availability to target tissues.

Time-Dependent Changes in the Stress Response

It appears that the hypothalamus and anterior pituitary function may be altered differently in the first few hours or days (acute phase) compared to 7–10 days (chronic phase) after disabling multiple traumatic injuries. During the first few hours or days after multiple traumatic insult, circulating GH levels become elevated and the normal GH profile is altered (peak GH levels and interpulse concentrations are high, and the GH pulse frequency is elevated). After a prolonged period, the changes observed with the somatotropic axis change; GH secretion pattern becomes random or chaotic, with a considerable reduction in pulses.

Within a couple of hours after trauma, serum 3,5,3' triiodothyronine (T_3) levels decrease, whereas thyroxine (T_4) and thyroid-stimulating hormone (TSH) levels rise briefly. Later, in the chronic phase, TSH levels are low to normal, and T_3 and T_4 concentrations are low. Prolactin levels show a pronounced increase following trauma. In the chronic phase, on the other hand, prolactin secretion is blunted. It has been suggested that this may play a role in the anergic immune dysfunction and in the increased susceptibility to infection observed at this time.

The stress-induced hypercortisolism triggered by trauma is associated with elevated ACTH release, which, in turn, is augmented by CRH, cytokines, and the noradrenergic system. Circulating aldosterone also rises markedly, and this may be attributed, at least in part, to the activated renin-angiotensin system. Hypercortisolism acutely shifts carbohydrate, fat, and protein metabolism, so that energy becomes available to the vital organs, such as the brain, and so that anabolism can be delayed. Intravascular fluid retention and enhanced AVP response provide hemodynamic advantages for the fight-or-flight stress response. Finally, the hypercortisolism induced by acute trauma may reflect an attempt by the organism to dampen its own inflammatory response. There is some evidence of a biphasic time course with the eventual dissociation of ACTH and cortisol concentrations following major trauma or surgery. Although there is some evidence of a possible role of endothelin

and/or atrial natriuretic peptide (ANP) in this dissociation of concentrations of cortisol and ACTH, the mechanism or significance of this in the regulation of HPA axis function in the critically ill or stressed is not clear.

In the chronic phase, serum ACTH levels drop and are low, whereas cortisol levels remain high. This suggests that cortisol release in this later stage may be driven by a different pathway. On the other hand, circulating levels of adrenal androgens (dehydroepiandrosterone sulfate, DHEAS), which has an immunostimulatory effect on T helper (Th)-1 cells are low during the chronic state. Some suggest that sustained hypercortisolism in the presence of low levels of DHEAS and prolactin in the chronic phase could provoke an imbalance between immunosuppressive and immunostimulatory pathways, which could increase susceptibility to infections.

Taken together, the neuroendocrine stress response differs in the acute and chronic phases following traumatic injury. Initially following trauma, the anterior pituitary actively releases hormones into the circulation, whereas anabolic target organ hormones are inactivated in the periphery. In the chronic state, the pulsatile release of anterior pituitary hormones becomes reduced, presumably because of reduced hypothalamic stimulation. The end result is diminished activity of target tissues and impaired anabolism.

Immune System Response to Trauma

The increased stress response after traumatic injury is associated with altered immune function and decreased immunity. The immune function of most patients is suppressed as a result of their traumatic injuries. Immunological changes after multiple trauma correlate with the degree of tissue damage, as well as with severity of hemorrhage and ischemia. Trauma has been found to significantly depress polymorphonuclear and monocyte functions, with both the phagocytic and bactericidal functions of these cell types being depressed. For the most part, immune changes are attributed to excessive cortisol levels.

Corticosteroids represent one of the most potent endogenous anti-inflammatory agents known. They have the capacity to inhibit and suppress virtually all critical inflammatory and immune cell functions, even at physiological concentrations and particularly during the early development of the immune/inflammatory response. At the molecular level, they inhibit the production of most inflammatory mediators, including interleukin (IL)-1, tumor necrosis factor (TNF), phospholipase A₂, and prostaglandins. Hence, the stress-induced enhanced production and secretion of glucocorticoids appears to counterregulate and

suppress excessive immune/inflammatory cell activation and mediator production that could otherwise result in self-induced tissue injury. This suggests a critical role for the corticosteroids in maintaining physiological homeostasis during the adaptive response to noxious stressors.

Several neurohormonal systems and immunological factors appear to play a fundamental role in these communicatory links, which enable the peripheral immunological apparatus to signal and interact with the brain and participate in maintaining immunological homeostasis. This is facilitated by an extensive coexpression of receptors and overlap of the synthesis of a multitude of hormonelike molecules that regulate cell function. In this regard, IL-1, a major product of activated macrophages, has been extensively studied. IL-1 receptors are expressed not only on peripheral T lymphocytes but also in the brain, where, in addition, it is synthesized. In addition to regulating a wide range of neural functions, including arousal, body temperature, sleep, food intake, reproductive behavior, and mood, that appear to be under direct central nervous system (CNS) control, some of the actions of IL-1 on peripheral immune function also appear to be mediated by IL-1 in the CNS. However, the relation of CNS IL-1 to peripheral immune system-derived IL-1 remains to be elucidated. IL-1 is also a potent stimulator of the HPA axis, resulting in the release of corticosteroids. Corticosteroids, in turn, feed back to inhibit macrophage release of IL-1. Other examples of this intricate communication between the immune and nervous system are rapidly being characterized.

Recent evidence suggests that there are two major pathways by which the CNS can regulate the immune system. The first pathway is through neurohormonal systems, particularly the HPA, or stress, axis. Other neurohormonal systems that appear to also play important roles include the hypothalamic-pituitary-gonadal (HPG) axis, the hypothalamic-pituitary-thyroid (HPT) axis, and the hypothalamic-growth hormone axis. The second major pathway is the autonomic nervous system and the release of norepinephrine and acetylcholine. Glucocorticoids inhibit the functions of virtually all inflammatory cells via the alteration of transcription of cytokine genes (TNF, IL-1, and IL-6) and inhibition of the production of arachidonic acid-derived pro-inflammatory substances, such as leukotrienes and prostaglandins.

Conversely, the immune system modulates CNS function through various factors, particularly cytokines, which can act independently or synergistically. The CNS contains IL-1 receptors in areas that control the acute-phase response. In addition, IL-1 stimulates

the production of endothelial cell prostaglandins, which in turn, induce CRH release from the median eminence. In contrast, IL-6 acts as a potent stimulator of ACTH and cortisol release. In healthy individuals, this bidirectional regulatory system forms a negative feedback loop, which keeps the CNS and immune system in balance.

The efferent sympathetic/adrenomedullary system apparently participates in an important way in the interactions of the adrenal axis and the immune/inflammatory reaction by being reciprocally connected with the CRH system, by receiving and transmitting humoral and nervous immune signals from the periphery, by innervating lymphoid organs, and by reaching all sites of inflammation via the postganglionic sympathetic nerve fibers.

Posttraumatic Psychological Complications

Multiple trauma injury poses a significant challenge to patients' perception of control over their environment. The various and multiple stressors presented not only by the traumatic episode itself but from the sequelae of events following the accident (multiple surgeries, diagnostic tests, etc.) diminish patients' perception of control, resulting in additional, although subjective, activation of the stress response.

Psychological complications are an important and persistent factor after injury and are associated with adverse effects on the daily activities of the patient. Multiple-trauma patients have been reported to suffer from loss of initiative, irritability, mood swings, feelings of guilt, among others, assessed using the State/Trait Anxiety Scale. The prevalence of posttraumatic stress disorder (PTSD) and symptoms of depression and anxiety in severely injured polytrauma patients is not clear. It appears, however, that the incidence of PTSD is low following multiple traumatic injuries in victims who were healthy before experiencing the trauma. Nonetheless, a substantial proportion of severely injured accident victims develop some form of psychiatric morbidity.

During the first few hours and days following a serious multitraumatic injury, most patients have at least short periods of anxiety with dissociative symptoms, such as derealization, occurring for a short duration in approximately 15% of the patients. Over the following weeks and months, the rates of PTSD reported range from 8 to 39%. A few long-term follow-up studies (28 months) report psychiatric morbidity, mostly depressive disorders, in approximately 22% of trauma victims and PTSD in 8% after 5 years.

Further Reading

- Barton, R. N. (1987). The neuroendocrinology of physical injury. *Baillières Clinical Endocrinology and Metabolism* 1(2), 355–374.
- Johnson, E. O., Kamilaris, T. C., Chrousos, G. P., et al. (1992). Mechanisms of stress: a dynamic overview of hormonal and behavioral homeostasis. *Neuroscience and Biobehavioral Reviews* 16, 115–130.
- Keel, M. and Trentz, O. (2005). Pathophysiology of polytrauma. *Injury* 236(6), 691–705.
- Smith, E. M. and Blalock, J. E. (1989). A molecular basis for interaction between the immune and neuroendocrine systems. *International Journal of Neuroscience* 38, 455–465.
- Soucacos, P. N., Beris, A. E., Xenakis, T. A., et al. (1995). Open type IIIB and IIIC fractures treated by an orthopaedic microsurgical team. *Clinical Orthopaedics* 314, 59–66.
- Van den Berghe, G. (2000). Novel insights into the neuroendocrinology of critical illness. *European Journal of Endocrinology* 143, 1–13.
- Willmore, D. W. (1991). Homeostasis: bodily changes in trauma and surgery. In: Sabiston, D. C. (ed.) *Textbook of surgery*. Philadelphia: W. B. Saunders.

Muscular Dystrophy See: Myopathy.

Musculoskeletal Problems and Stress

S Svebak

Norwegian University of Science and Technology,
Trondheim, Norway

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
S Svebak, volume 2, pp 796–802, © 2000, Elsevier Inc.

Epidemiology of Musculoskeletal Pain

Muscle Fiber Contraction, Hardness, and Pain

Sperry's Principle: Brain Control of Muscle Tension

Motivation, Emotion, Muscle Tension, and Performance

Personality and Sensitivity to Emotional versus Ergonomic Loads

Central Nervous System Mechanisms of Pain Sensitivity

Coping with Musculoskeletal Problems: Prevention and Rehabilitation

Glossary

Brain

neuro-transmitters

Emotional load

Biochemical substances that modulate the efficacy of nerve-impulse transmission in the brain.

The psychosocial causes of unpleasant and enduring emotions induced by the perceived or extrinsic nature of work-load, inducing psychobiological changes

Ergonomic load

Fear avoidance

Myalgia

Myofascial pain

Striate muscles

Tendonitis

that can cause bodily complaints and illness.

The physical causes of biomechanical strain induced by frequent or enduring postural work demands.

Passivity due to fear of the pain expected from movements.

Pain from the skeletal muscles.

Pain from the fasciae that cover the skeletal muscles.

Skeletal muscles named after their microscopically observable striated pattern of muscle fiber arrangements. These fibers are of two major kinds: one is designed for aerobic metabolism and the other for anaerobic metabolism (in which oxygen is not involved). Individuals differ in their proportion of fiber types, mainly due to genetic disposition.

Pain from the tendons of the skeletal muscles.

Skeletal muscles represent the only case in which the brain can exert some degree of volitional control over the activity level of an organ. Despite this fact, much pain, stiffness, and discomfort are presented to the health-care system as originating in the skeletal muscles. It is perceived by the patient as beyond control.

Epidemiology of Musculoskeletal Pain

Most musculoskeletal problems present as nonspecific pain, that is, pain that has no obvious and verifiable cause. Often it is due to a complex constellation of ergonomic, physiological, motivational, and emotional load factors that act on biological mechanisms involved in pain perception. Not all pain is attributable to a particular body area. A major diagnostic criterion in fibromyalgia is the presence of widespread pain. However, the most prevalent types of muscle pain involve the neck, shoulders, and lower back; the extremities may also be involved (mostly the gastrocnemius muscle of the legs and the flexor muscles of the forearms).

A major proportion of muscle pain is due to myalgia, myofascial pain, and pain from the tendons of muscles, referred to as tendonitis. In the lower back, the muscles involved are the *musculus (m.) erector spinae*, *m. iliocostalis lumborum*, and *m. gluteus medius*. In the neck and shoulders, the *m. trapezius*, *m. deltoideus*, and muscles covered by these superficial muscles are also prevalent sources of pain.

Muscle pain is not always taken to the health-care system. In most cases, the pain and related discomfort is mild and involves transient disability or the pain is treated by the consumption of over-the-counter analgesics. It has been estimated that, on average, 30–40% of musculoskeletal pain relates to work and that severe episodes of low back pain may occur in less than 5% of the population. However, the 1-year prevalence of musculoskeletal pain may be as high as 50% among adults in Western societies. In a recent study involving 70% of a county population ages 20 and above ($N = 64,690$), the 1-year prevalence, including less severe episodes of muscle pain, was as high as 50.1% in females and 42.6% in males. Continuous pain and discomfort for at least 3 months were most prevalent in the neck, shoulders, and lower back (females: neck = 19.4%, shoulders = 21.1%, lower back = 17.4%; males: neck = 14.5%, shoulders = 16.8%, lower back = 14.2%). In preadolescents, persistent pain may be as prevalent as 15%, with pain in the neck and lower extremities accounting for a major proportion of this pain, particularly among the girls.

There was a shift toward multidisciplinary paradigms in musculoskeletal pain research after 1985 with the increasing acknowledgment of psychological and psychosocial factors. On the behavioral side, more children and adults are physically inactive in everyday life than ever before, although mental activity may be high in front of a visual display unit or television screen. On the other end of this skewed distribution of physical activity level is a small

percentage of the adolescent and young adult population that is extremely active in competitive sports. Both lifestyles increase the risk of musculoskeletal pain but for different reasons. Inactivity causes the atrophy of skeletal muscle and poor circulation, resulting in poor workload tolerance; sport injuries cause pain due to physical overload, mishaps, and accidents.

Muscle Fiber Contraction, Hardness, and Pain

Human muscles are of three kinds. They are distinguished mainly by the anatomical microstructure that can be microscopically observed by histological inspection. (1) Striate (skeletal) muscles have a striated pattern, reflecting the special arrangement of muscle fibers side by side; they contract on stimulation by nerves in the pyramidal and extra-pyramidal systems. (2) Smooth muscles induce movements in internal organs, including the digestive tract and blood vessels; they lack a striated pattern and contract on stimulation by the autonomic nervous system. (3) The heart muscle has a semi-structured arrangement of fibers and is stimulated by nerves in a special section of the autonomic nervous system. A number of other factors modulate the neural stimulation of muscles, including neurotransmitters, hormones, and the acid–base balance of the body.

Muscles can also be classified according to the dimensionality of their movements. Skeletal muscles induce one-dimensional movements due to the attachment of their tendons to the skeleton across joints. In contrast, smooth muscles induce two-dimensional movements by the compression of the diameter of a tube (e.g., blood vessel); the heart muscle fiber contractions induce three-dimensional movements to reduce the globular volume and act as a pump.

Skeletal muscles harden with contraction. Activation from complete relaxation to maximum voluntary contraction (MVC) presents a substantial change in muscle hardness that is easily palpable, whereas a shift from relaxation to 5% of MVC can be accurately assessed only by the use of modern pressure technology and electromyography (EMG). In EMG, surface and implanted electrodes pick up neuroelectrical impulses generated by motor nerve activation. Research has supported a relationship between hardness and tenderness in the trapezius muscle of patients with chronic tension-type headache. This relationship is present also on days without pain and, therefore, indicates a permanent alteration of muscle hardness and tenderness that is different from that of healthy controls. Also, the mechanism is more enduring and

extensive than those responsible for the perception of pain from these muscles.

Skeletal muscles are composed of a mixture of aerobic and anaerobic fibers (see [Figure 1](#)). There are remarkable individual differences in the ratio of aerobic (type 1) to anaerobic (type 2) fibers among individuals. These differences are genetically given. A subset of the type 2 fibers, the type 2A fibers, can perform aerobic metabolism when forced to do so, whereas the type 2 B fibers are nearly always anaerobic. Oxygen is needed in aerobic metabolism, and carbon dioxide is a waste product. Oxygen is not involved in anaerobic metabolism, and lactic acid is a waste product that can be used by type 1 fibers for aerobic metabolism. This is to say that the blood circulation brings oxygen to the aerobic fibers and washes out harmful carbon dioxide from the local tissue, whereas the waste product from type 2 fibers, lactic acid, can be exploited by neighboring aerobic fibers and, therefore, puts less demand on the systemic circulation. Aerobic fitness training makes aerobic fibers more efficient at aerobic metabolism, but it does not increase the number of genetically given aerobic fibers. Elite marathon runners may have 80% of aerobic fibers in their leg muscles, whereas elite sprint runners have around 80% of anaerobic fibers in their leg muscles.

The genetic differences are often forgotten when people explain why they intuitively feel they are not made for aerobic activities such as long-distance running, cross-country skiing, and other endurance activities at a moderate intensity. Conversely, they also shed light on why others display an intuitive distaste for anaerobic activities such as sprinting and downhill skiing. Most individuals have a good balance between type 1 and 2 muscle fibers and, therefore, are capable of performing aerobic as well as

anaerobic work to a reasonable degree. Some individuals are strongly genetically biased toward a high tolerance for anaerobic work only, whereas others have a remarkable tolerance for aerobic work production without experiencing muscle pain.

Sperry's Principle: Brain Control of Muscle Tension

The sole product of brain function, including our thinking mechanism and the related mental functions of motivation, emotion, attention, and memory, is muscular coordination. Skeletal muscles make up most of the body mass in humans. Consequently, perceptual-motor control of the skeletal muscles presents the most extensive challenge to the brain's capacity for information processing. The perceptual component in motor control is a less extensive challenge than is the challenge of motor coordination. This perspective on the brain's control of body functions is named Sperry's principle.

The musculoskeletal system is the biological substrate for volitional behavior in the common sense. The dynamics of motor activity is also the basis for displays of expressive behavior, despite the often unintended nature of such displays, and other elements of body language including body posture. The way we sit on a chair and stand by the desk involves a much more automatic coordination of complex muscle tension patterns than is involved in the volitional movement of a finger to release a button. Walking toward a good friend involves the activation of a whole range of muscle coordination, including automated control of walking, the expression of a friendly display as well as raising a hand to say hello. The ballistic smoothness of such movements in space is taken care of by integrated input from the cerebellum.

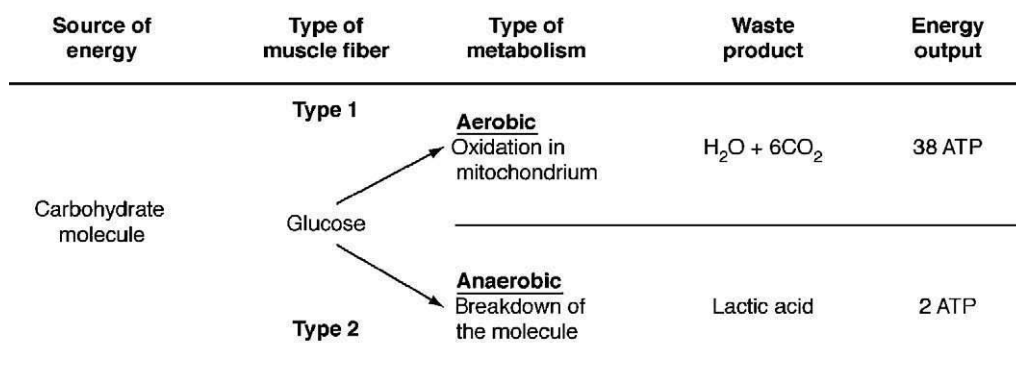


Figure 1 Major differences in metabolic energy production between aerobic (type 1) and anaerobic (type 2) muscle fibers. Note that lactic acid is a waste product from type 2 fibers and can be available for subsequent aerobic metabolism in type 1 fibers. ATP, adenosine triphosphate (the work horse in muscle contractions).

The Pyramidal Pathway of Muscle Control

The most remarkable connection between the cortex of the brain and the skeletal muscles is called the pyramidal tract. Nerve fibers in this tract originate in the motor cortex and synapse with a second nerve fiber in the spinal tract to make the nerve impulse from the cortex initiate contraction in a group of skeletal muscle fibers. There is no other example of a monosynaptic connection from the cortex of the brain to a target organ in the body. This arrangement is responsible for the precise control of intended movements with minimal interference from other parts of the brain. The brilliance of this volitional control can become remarkable with practice, as professional musicians, and in elite athletes.

A motor nerve fiber from the brain to the muscle terminates by splitting up into several endings to stimulate simultaneously a group of muscle fibers in a single motor unit (SMU). All fibers in one SMU are of the same kind (type 1 or 2). The number of fibers may vary between approximately 10 and 200, with small SMUs in muscles involved in precise motor control, such as in the muscles of the fingers, and with large SMUs involved in the postural muscles of the trunk.

The Extra-Pyramidal Pathways to the Skeletal Muscles

Body posture is regulated via the extra-pyramidal pathways from the brain to the skeletal muscles. These influences originate subcortically and, therefore, are subject to automated induction of tension patterns. Several such pathways are integrated as parts of the basal ganglia between the thalamus and the cortex. One extra-pyramidal pathway originates in the reticular formation of the medulla oblongata and is a powerful modulator of skeletal muscle tension levels throughout the body. This pathway provokes muscle tension with increasing anxiety, alertness, and levels of effort to assure a stabilizing platform for intended actions as part of the attentional demands and flight-or-fight adaptive mechanisms of the hypothalamus and limbic system.

All extra-pyramidal pathways to the skeletal muscles are multisynaptic and can receive input from several other subcortical connections. This fact explains the chronic nature of many tension-related muscle problems and why they are influenced by a number of psychological processes that are beyond volitional control. It also explains the frequent need for an adjustment of lifestyle, including physical exercise habits and interpersonal competence, in the rehabilitation of musculoskeletal-pain patients, as well as why it is a time-consuming challenge to recover from musculoskeletal-pain problems.

Motivation, Emotion, Muscle Tension, and Performance

Brain activation of muscle tension is often due to task involvement, which is influenced strongly by motivational processes. Long ago, Malmö demonstrated that the intensity of motivational involvement in perceptual-motor tasks provokes increasingly higher tension levels in muscles that are not strictly involved in the demands that are extrinsically given by the nature of the task. One example is the operation of a joystick by the dominant forearm; muscle tension gradients build up in the nondominant forearm, which is never called on by the extrinsic nature of that task. He demonstrated that such generalized gradients of muscle tension tend to become more marked with increasing interest, anxiety, and effort. The quality of the performance may also improve with the increasing gradients of tension.

Quality versus Intensity of Motivational States

Apter and Svebak reviewed a series of EMG experiments on muscle tension patterns provoked by the performance of a perceptual-motor task in a playful versus serious-minded motivational state. They concluded from these experiments that playful task involvement with high effort provokes a pattern biased toward high pyramidal activation of muscles called on by the extrinsic nature of the task and low extra-pyramidal tension gradients in background muscles. Serious-minded task involvement consistently provoked gradients of high tension buildup in the seemingly passive muscles of the nondominant forearm, the legs, and the neck when the dominant forearm was actively performing a perceptual-motor task. These states of motivational involvement can be manipulated with some success by the use of extrinsic incentives such as the promise of a monetary bonus for a good performance or the threat of aversive electric shock for a poor performance. Incentives improve performance, whereas the qualitative distinction between serious-mindedness and playfulness has been reflected less clearly in the quality of perceptual-motor performance.

Playfulness and serious-mindedness are facets of personality. Some individuals are dominated more of the time by one of these motivational states, whereas others reside in the opposite state. When extrinsic incentives are not manipulated, motivational dominance explains why some people present with enduring and seemingly unprovoked muscle tension patterns that are not present in others performing the same type of perceptual-motor task. When extrinsic incentives are manipulated, they may interact with aspects of personality, which explains some of the, often paradoxical, differences of workload tolerance across

individuals. The incentive that is perceived as a threat by one person may be welcomed as an exciting challenge by another.

Personality and Sensitivity to Emotional versus Ergonomic Loads

Patients with chronic low back pain (i.e., daily pain for more than 3 months) have elevated scores on somatization disorder without meeting the *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (DSM-IV) criteria for this diagnosis. Some of these pain patients also tend to report the incidence of lifetime major depression as well as alcohol dependence. Other personality factors can increase the risk of back pain when challenged by the nature of the workload. One example is a high emotional load. Hospital staff members in some wards are exposed to high emotional loads when working with chronically ill children and their parents or when employed as a midwife, situations in which a life-threatening crisis can present without much warning. Employees who score high on neuroticism, with a tendency to overrespond with anxiety, worry, and other dysphoric moods when faced with emotional stressors, are at increased risk of suffering from pain in the neck and shoulders. In contrast, staff members in wards with a predominantly high ergonomic load are at increased risk of low back pain. This risk is mediated, provided they score high on personality traits reflected in habitual impatience and the willingness to respond on impulse with a high motor effort when coping with such a load.

Personality can increase the risk of musculoskeletal pain in ways that distinguish ergonomic from emotional load factors. They can also add to local tissue damage and cause muscle pain as a result of nested vicious circles of algogenic processes (see [Figure 2](#)). With ergonomic load, the brain activates muscles predominantly via the pyramidal pathways to induce intentional movements. The speed and force components can induce muscle and tendon ruptures as well as pain from a high spinal load on disks, which respond with hernias that may irritate local nerves and give rise to irradiating pain. An increased risk of such ruptures occurs when high force is provoked by an impulsive brain discharge into the striate muscles with a composition of fibers designed for low enduring loads (type 1 fibers). Ruptures of tendons can be provoked, and pain from inflammatory processes may be provoked as secondary consequences of these ruptures also involving the fasciae of muscles.

With emotional load, the extra-pyramidal pathways from the brain to the muscles become more

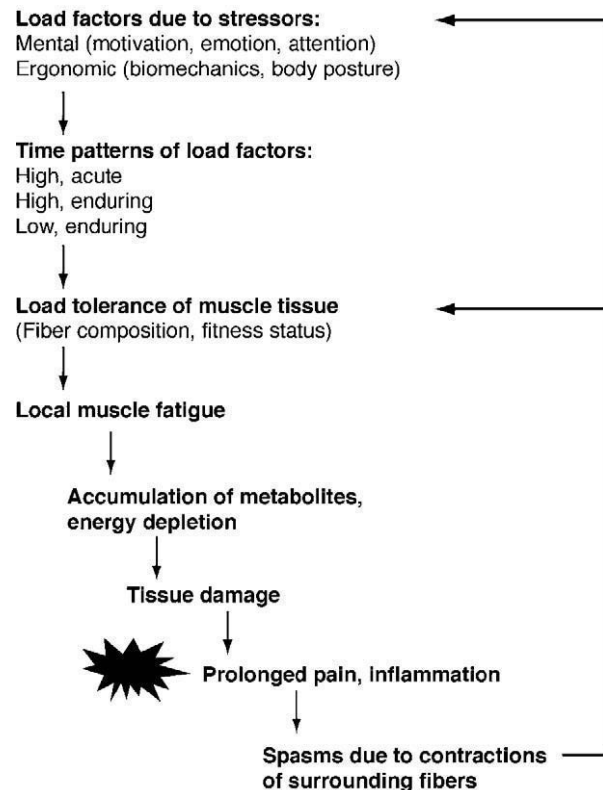


Figure 2 Relationship among factors in musculoskeletal load, their time and intensity parameters and their tissue tolerance. Arrows indicate hypothesized patterns of nested vicious circles in musculoskeletal pain.

activated and induce enduring tension in muscles not strictly called on by the extrinsic nature of motor tasks. These activation patterns are likely to have side-effects on blood circulation through the muscles, resulting in local ischemia. The brain activation pattern may also be in conflict with the local tolerance for enduring a low workload due to a bias toward anaerobic type 2 muscle fibers, which are quickly exhausted. The resulting tissue damage provokes inflammation that gives rise to pain.

Central Nervous System Mechanisms of Pain Sensitivity

Several brain processes can significantly influence pain sensitivity. The perception of pain is always subjective. Pain can be perceived with no biologically definable cause. In contrast, large tissue damage may not always give rise to the perception of severe pain. Also, pain thresholds tend to fluctuate throughout the day and over weeks. One reason for these poor relationships between the degree of tissue damage and pain perception is purely due to a shift in the focus of attention. Pain is mediated in unemployment

where body awareness can be facilitated in the lack of extrinsic attentional demands. Hypnotic and meditative techniques guide a shift in the focus of attention away from pain and may also induce relaxation.

Brain Neurotransmitters

Several neural pathways and chemical neurotransmitters of the brain modulate sensitivity to pain. Until recently, most brain research on pain sensitivity concentrated on the thalamus; now it has become evident that many areas of the brain modulate pain perception. The spinothalamic tract is the most predominant nociceptive pathway for this cortical input, but the spinoreticular pathway is also acknowledged to be important, and they both terminate in the somatosensory cortex where there are small receptive areas that explain the perception of pain in discrete areas of the body but not the diffuse nature of most clinical pain. A third pathway is the spinomesencephalic pathway that terminates in the periaqueductal gray matter of the midbrain, an area rich in opioid receptors. All these pathways modulate pain perception.

The main neurotransmitters in the cortex are glutamate (pyramidal cells that are excitatory) and γ -aminobutyric acid (GABA; nonpyramidal cells that are inhibitory). Anxiolytic medication acts on GABA to facilitate neuroelectric inhibition. The cortical degeneration of glutamate receptors is a monotonous process from birth to the onset of puberty. Early onset means there will be a higher density of glutamate receptors than late onset. Subcortical serotonergic neurons inhibit pain perception. Serotonergic hyposecretion tends to provoke dysphoric moods such as irritability and depression. Endorphins act like morphine and moderate pain sensitivity (via the opioid receptors). Their secretion is believed to increase with good moods and to decrease with bad moods. However, personality factors may play a more important role than moods in the psychological regulation of the endorphins, and personality as well as mood states may influence other pain-modulating transmitters more than they influence endorphin secretion. Enkephalins are also endogenous opioids and decrease pain sensitivity.

In 1992, the endogenous cannabinoid (cannabis-like molecules) brain system was discovered and, since then, has attracted much research due to its ubiquitous receptor in the brain. These molecules bind to the CB₁ receptor that is densely distributed in areas related to motor control, emotional control, cognition and motivated behavior. These new insights may provide cannabinoid receptor antagonists and improve clinical approaches to a diversity of problems including pain, obesity, anxiety and drug addiction.

Spinal and Peripheral Mechanisms

The gate control theory of pain sensitivity was proposed by Melzack and Wall in 1965 to include free nerve endings in the deeper layers of the skin. These nerve endings respond to touch by inhibiting the spinal transmission of pain signals from the body to the brain. The pleasant component of being touched can also induce a shift in mental state that, subsequently, induces increased pain thresholds via changes in pain neurotransmitters of the brain. A therapeutic circle of events is indicated by the fact that the spinal pain-transmission peptide, substance P, is regulated via the secretion of serotonin in the brain. All mental influences that raise the secretion of brain serotonin also increase the concentration of substance P in the brain, which, again, increases spinal serotonin, which, in turn, reduces spinal substance P and therefore reduces pain sensitivity.

Coping with Musculoskeletal Problems: Prevention and Rehabilitation

There is probably no better way to prevent nonmalignant musculoskeletal pain and discomfort than to develop sensory-motor competence from early childhood and to maintain this competence throughout life. And yet there is no guarantee. A healthy ego development is based on the acquisition of sensory-motor skills. Through play, work, and sports, experience gradually builds to provide a solid basis for the individual to define those motor activities that are intrinsically felt to be right. There is no better way to find out what one's own skeletal muscles are good at and where their tolerance limits are. It is more important to health for an individual to maintain a level of physical fitness that is felt to be intrinsically rewarding than to meet some socially mandated standard prescription such as jogging three times a week for at least 20 min. Not all individuals are genetically equipped with skeletal muscles made for aerobic exercise, but all are made to be active.

Work-related musculoskeletal pain and discomfort often reflect a mixture of biomechanical (postural), physical (climatic factors and lightening noise), and psychosocial (mental demands and psychosocial relationships) exposures. An extensive review of ergonomic intervention studies by Westgaard and Winkel concluded that successful outcomes are due to (1) organizational culture interventions with high commitment of the stakeholders, (2) the use of multiple interventions to reduce identified risk factors, and (3) the modification of risk factors for the individual workers at risk, using measures that ensure the active involvement and support of the worker.

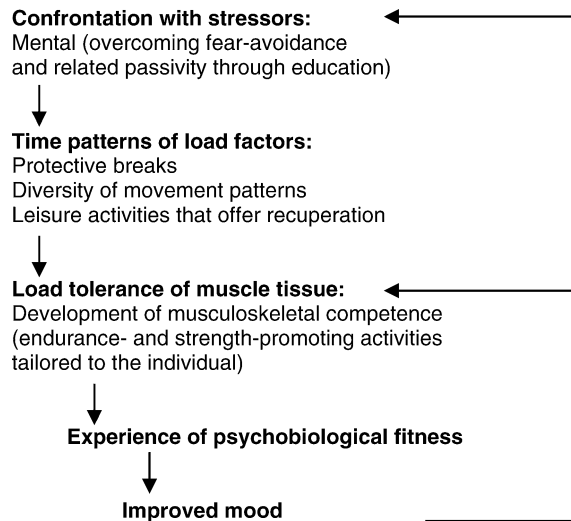


Figure 3 The dynamic relationship among important elements in rehabilitation of musculoskeletal pain. Arrows indicate hypothesized patterns of nested self-rewarding benign circles in the psychobiological development of musculoskeletal competence.

Multidisciplinary intervention programs have been recommended for nonmalignant musculoskeletal pain since the 1970s. A review of 65 treatment outcome studies for back pain was published in 1992. These studies included a diversity of treatments across the psychological, medical, physical, and occupational fields. They lasted, on average, 7 weeks with a range of from 1 to 31 weeks. Hours spent in treatment ranged from 4 to 264, with an average of 96 h. The site of pain in most studies was reported to be the back or in heterogeneous locations. The results from a meta-analysis clearly support the conclusion that multidisciplinary treatments for nonmalignant chronic musculoskeletal pain are superior to single-discipline treatment, waiting list, and no treatment. And the effects have proven to be stable over time. Interestingly, the effects also extend beyond the reduction of pain. Patients report improved mood, a high frequency of return to work, and a reduced need for support from the health-care system after multidisciplinary treatments (Figure 3). These early findings are strongly supported in recent research.

Patients must be encouraged to be physically active; the recommendation of a passive lifestyle is not a professional treatment. In this encouragement, fear avoidance may be the major obstacle. This concept was introduced in 1983 by Lethem et al. to account for a passive lifestyle in patients with chronic musculoskeletal pain, in which the avoidance of movements that are expected to give rise to pain leads to the maintenance or exacerbation of fear. The effectiveness of rehabilitation is improved when the patient learns to orient him- or herself away from patterns of passive avoidance and toward active confrontation with pain and the related fear. Multidisciplinary

treatments of groups of outpatients with various diagnoses, of various ages, and of both genders are best at overcoming pain-related fear. And professionals, beyond being academically skilled, also should be socially competent to provide emotional security as well as facilitate motivation and fun.

Further Reading

- Apter, M. J. and Svebak, S. (1992). Reversal theory as a biological approach to individual differences. In: Gale, A. & Eysenck, M. W. (eds.) *Handbook of individual differences: biological perspectives*, pp. 323–353. Chichester: John Wiley.
- Ashima, M., Bendtsen, L., Jensen, R., et al. (1999). Muscle hardness in patients with chronic tension-type headache: Relation to actual headache state. *Pain* 79, 201–205.
- Bru, E., Mykletun, R. and Svebak, S. (1993). Neuroticism, extraversion, anxiety and Type A behaviour as mediators of neck, shoulder and lower back pain in female hospital staff. *Personality and Individual Differences* 15, 485–492.
- Bruehl, S., Carlson, C. R., Wilson, J. F., et al. (1996). Psychological coping with acute pain: an examination of the role of endogenous opioid mechanisms. *Journal of Behavioral Medicine* 19, 129–142.
- Flor, H., Fydrich, T. and Turk, D. C. (1992). Efficacy of multidisciplinary pain treatment centers: a meta-analytic review. *Pain* 49, 221–230.
- Lethem, J., Slade, P. D., Troup, J. D. G., et al. (1983). Outline of fear avoidance model of exaggerated pain perceptions. *Behavior Research and Therapy* 21, 401–408.
- Malmo, R. B. (1975). *On emotions, needs, and our archaic brain*. New York: Holt, Rinehart and Winston.
- Mikkelsen, M., Salminen, J. J. and Kantiainen, H. (1999). Non-specific musculoskeletal pain in preadolescents: prevalence and 1-year persistence. *Pain* 73, 29–35.
- Norton, P. J., Asmundson, G. J. G., Norton, G. R., et al. (1999). Growing pain: 10-year research trends in the study of chronic pain and headache. *Pain* 79, 59–65.
- Sperry, R. W. (1952). Neurology and the mind-brain problem. *American Scientist* 40, 291–312.
- Svebak, S., Hagen, K. and Zwart, J.-A. (in press). One-year prevalence of chronic musculoskeletal pain in an adult Norwegian county population: relations with age and gender. The Nord-Trøndelag Health Study. *Journal of Musculoskeletal Pain* 14 (in press).
- Storror, S., Moen, J. and Svebak, S. (2004). Effects on sick-leave of a multidisciplinary rehabilitation programme for chronic low back, neck, or shoulder pain: comparison with usual treatment. *Journal of Rehabilitation Medicine* 36, 12–16.
- Vlaeyen, J. W. and Linton, S. J. (2000). Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. *Pain* 85, 317–332.
- Westgaard, R. H. and Winckel, J. (1997). Ergonomic intervention research for improved musculoskeletal health: a critical review. *International Journal of Industrial Ergonomics* 20, 463–500.

Music Therapy *See: Integrative Medicine (Complementary and Alternative Medicine).*

Myopathy

G A Small

Drexel University of Medicine, Philadelphia, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G A Small, volume 2, pp 803–809, © 2000, Elsevier Inc.

Muscle Structure
Diseases of Muscle
Differential Diagnosis of Myopathy
Electrodiagnosis
Muscle Biopsy
Therapy
Conclusion

Glossary

<i>Electro-diagnosis</i>	A method of evoking electrical responses or recording spontaneous electric responses from nerves and muscles, as well as from the neuromuscular junction, in order to localize pathology to one or more of these specific regions.
<i>Fatigue</i>	A diminishing in muscle strength that is measurable but often without clear anatomic or organic explanation.
<i>Muscular dystrophy</i>	A progressive, untreatable muscle disease defined by scarring and inflammation on histology. The condition is generally recognized as inherited.
<i>Myalgia</i>	Pain attributed to a muscular origin.
<i>Myositis</i>	The inflammation of muscle creating pain, fatigue, and weakness. It is generally proximally predominant.

Myopathy defines a category of neurological diseases affecting muscle proteins and muscle cell substructure, muscle membrane, and other subcellular chemical constituents involved in the transduction of carbon-bond energy into muscle contraction.

Muscle Structure

Skeletal muscle contracts voluntarily, requiring the transduction of a nerve action potential to a chemical message at the neuromuscular junction. This elicits an end-plate potential in the muscle membrane, bringing the membrane to its electrical threshold, which results in calcium release throughout the muscle cell, initiating conformational changes in subcellular proteins (actin, myosin, and troponin) and shortening of the muscle cell, generating force.

Muscle tissue histochemically is a mosaic of muscle proteins surrounded by a muscle membrane, called endomysium, and grouped into fascicles by perimysial connective tissue. The entire muscle is enclosed by a connective tissue cocoon continuous with the perimysium, known as the epimysium, which in turn is continuous with tendon tissue. This bonding of the entire muscle substructure across joints directs the vector of contractile force of each muscle cell in a similar direction, resulting in a coordinated contraction of muscle and flexion or extension of a limb.

Muscle histologically appears as a mosaic due to the dark or light staining characteristics of subcellular proteins involved in fast or slow muscle contraction. Whereas some muscles are dominated by one type and twitch quickly, others are dominated by slower contracting cells, conserving energy. The adaptive advantage of fast- or slow-twitch muscle fibers has a teleological explanation in the response of the organism to threats in the environment. Muscles that have a predominance of fast-twitch fibers are the limb and eye muscles, allowing the organism to respond quickly either toward prey or away from predators, determining whether the organism survives by getting nourishment or avoiding danger. These muscle cells twitch with a great force, but they cannot sustain that force for long. Slow-twitch fibers can contract with reasonable force indefinitely, as long as the pathways for the energy breakdown of fatty acids exist. Such muscles are involved in sustained activity and not necessarily in fight, fright, or flight activity.

Electrical impulses generate neurotransmitter release into the neuromuscular junction, resulting in electrical responses at the muscle membrane, which are carried to all parts of the muscle cell through a complicated system of invaginations of the muscle membrane known as the T tubule system and sarcoplasmic reticulum. These electrical responses are then transduced into the release of calcium from the sarcoplasmic reticulum into the cytoplasm of the muscle cell, known as the sarcoplasm. The interaction of calcium ions with filaments of the muscle proteins actin and myosin allow these strands, lying parallel, to be drawn against one another longitudinally along the muscle cell axis to shorten it. This is an ATP-dependent process. Damage to signal transduction along this pathway results in weakness that may be abrupt and permanent, fluctuating, or relentlessly progressive, depending on the pathophysiology and location of the damage.

Fatigue can result from the functional breakdown at any one of these levels, even at the mental or central nervous system level. Fatigue implies that normal forces are being initially generated and cannot be maintained. It can be measured more easily than its cause can be specifically elicited. Stress to any part of the nerve, neuromuscular junction, and muscle, as well as muscle subcellular structure, can exceed the capacity of the system elements to work synergistically, resulting in diminished muscle strength, which is perceived as fatigue by the individual.

Diseases of Muscle

The variety of muscle diseases reflects the specific localization of pathology in the muscle substructure. Scores of separate muscular dystrophies, inflammatory, and metabolic conditions have been described.

Muscular Dystrophies

Dystrophinopathies Dystrophinopathies are defined by abnormalities of the large structural protein dystrophin, which was first described in the late 1980s, revolutionizing the study of Duchenne muscular dystrophy. Abnormalities of the structural protein depend on genetic problems in the gene for dystrophin at Xp21. Duchenne and Becker muscular dystrophies fall into this category.

Duchenne muscular dystrophy, the most common muscular dystrophy in children, is inherited in an X-linked recessive manner. Females carry the gene, and their male offspring have a 50% chance of manifesting the disease, depending on whether they inherit the abnormal X chromosome. Although the symptoms generally do not appear until early childhood, serum muscle enzymes are frequently elevated

in infancy. The upper extremities are usually affected later than the lower ones in a proximal to distal fashion. Walking is generally not possible by puberty. During late adolescence, the respiratory muscles are significantly compromised and the cardiac muscle is affected. These two manifestations, as well as the general danger of immobility, resulting in deep vein thrombosis, pneumonia, and skin breakdown, are the general causes of mortality. Dystrophin is absent from muscle tissue.

Becker muscular dystrophy differs from Duchenne muscular dystrophy in that the age of onset is later, the rate of progression slower, and the disease is not terminal. Abnormal dystrophin occurs in muscle. Becker muscular dystrophy refers to numerous abnormal dystrophy subtypes. The lethal disturbance in the dystrophin gene, resulting in the low or abnormal translation of effective dystrophin, is unique for each subtype. The muscle substructure is then prone to inflammation, scarring, and the general pathological characteristics necessary for histological diagnosis.

Myotonic muscular dystrophy This is the most common muscular dystrophy in adults. The pathophysiology is unclear; however, the genetics of this very commonly underdiagnosed condition is different from the dystrophinopathies. Myotonic dystrophy is inherited in an autosomal dominant pattern. It is a multisystem disease that includes cardiomyopathy, cataracts, endocrine problems, and central neurogenic hypoventilation. The prevalence is 1 in 20,000. Variations in the phenotype of the disease occur despite the classical description of individuals with this syndrome as having neck flexor atrophy along with ptosis; some patients have a normal physiognomy. Some patients endure muscle cramps, whereas others have a disease incompatible with a normal life span due to cardiopulmonary problems. All suffer cataracts by age 40. It is very slowly progressive in most patients.

By definition, myotonia is a problem of impaired muscle relaxation secondary to the depolarization of the muscle membrane and impaired repolarization. Characteristic electrodiagnostic findings are waxing and waning repetitive motor-unit action potentials with the characteristic audible sound of a dive bomber. This electrodiagnostic finding is not specific for myotonic muscular dystrophy but is present in all the myotonic diseases, suggesting that the disease itself is a muscle membrane disorder. In the early 1990s, the gene for myotonic dystrophy was mapped to chromosome 19, but the gene product has not been identified clearly. The likely candidate is a protein kinase involved in the function of numerous tissues within the body, thus explaining the pleiotropism of

the disorder. Myotonic muscular dystrophy is one of the trinucleotide repeat sequence diseases. The repetitive appearance of three nucleotides in sequence may occur up to thousands of times within the gene, resulting in genomic instability.

Congenital forms This designation includes obscure, slightly progressive, or nonprogressive diseases clearly present at birth. Patients may manifest mental retardation, seizures, and anatomical brain abnormalities. Muscle pathology is variable, including rodlike structures and central cores. Heritability is varied.

Inflammatory Myopathies

Polymyositis Polymyositis is an inflammatory T-cell-mediated dysfunction of muscle cells, with high serum creatine phosphokinase (CPK), white cell invasion of muscle cells, and progressive proximal weakness. It responds well to steroids.

Dermatomyositis Dermatomyositis is a B-cell-mediated disease of muscle and muscle vasculature, also affecting the skin. It is associated controversially with underlying malignancy.

Inclusion body myositis Patients with clinical manifestations of polymyositis who do not respond to prednisone probably have inclusion body myositis (IBM). With more advanced pathological examinations, many victims of steroid unresponsive polymyositis have been found to have vacuolated muscle fibers, mononuclear cell invasion, and amyloid deposits and tubular filaments by electron microscopy. The discovery of amyloid in muscle fibers, along with inflammation, prompted a trial of immunosuppression to determine whether the suppression of the inflammatory response would result in improved strength; this did not occur. The amyloid degeneration in victims of this disease may be similar to the amyloid degeneration seen in the brains of patients with Alzheimer-type dementia. Therefore, IBM may indeed be a primary degenerative disease, such as amyotrophic lateral sclerosis or Alzheimer-type senile dementia. If the theory is correct, this does not bode well for immunomodulating therapeutic trials.

Metabolic Myopathies

Metabolic myopathies are separated predominantly into glycogen-storage diseases and mitochondrial disorders; these genetically determined abnormalities of glucose and fatty acid metabolism result in myoglobinuria, severe muscle weakness, and disordered acid-base homeostasis. They are a cause of occult fatigue in numerous patients and can be difficult to

diagnose. The fact that the mitochondrion possesses its own genome complicates heritability further and underscores the relationship between cytoplasmic inheritance and nuclear inheritance in many of these disease states.

Glycolytic pathway pathology Abnormal metabolism of the stored form of glucose (glycogen) is determined by numerous heritable defects in glycolytic and glycogenolytic enzyme pathways. The lack of phosphofructokinase, phosphoglycerate kinase, lactate dehydrogenase, and phosphoglycerate mutase results in myoglobinuria with exercise. In these conditions, glucose cannot be used efficiently as fuel for muscular contraction, and cramps and myalgia occur with exertion. When the exertion is significant, muscle breakdown occurs and serious systemic consequences can result, such as acidosis and renal failure. An outmoded method of diagnosing glycogen-storage diseases involved exercising a muscle under ischemic conditions and failing to record an appropriate rise in venous lactate; the sophisticated enzymatic analysis of muscle biopsy specimens has supplanted this test. Exercise intolerance is generally the rule in these disorders, and hypoglycemia can occur as well. Therapy is largely related to altering the patient's level of activity and ratio of protein to complex carbohydrate intake.

Mitochondrial myopathies Mastery of the evolving field of mitochondrial myopathies is complicated by almost monthly discoveries and supplantation of previous ones. Abnormalities in mitochondrial function result in disordered fatty acid metabolism. Patients with fatigue and exercise intolerance have abnormal aggregations of mitochondria in their muscle tissue. Patients with weakness, mild or severe, are found to have elevated of lactic acid levels in the serum and cerebrospinal fluid (CSF) due to the reliance of the body on glycolytic function for the production of ATP rather than on the breakdown of the fatty acids. There is no effective treatment for any of the mitochondrial disorders. Attempts to supplement the diet with a variety of respiratory chain constituents, such as coenzyme Q, have met with only anecdotal success. The predominant importance of defining a syndrome of muscle weakness, exercise intolerance, and fatigue is to differentiate an organic from a psychiatric cause of impaired function in patients. Fatigue and myalgia are the predominant complaint of patients who present to the neurologist or their general physician with these diseases. It is likely that the vast majority of patients with these symptoms are initially diagnosed to have stress-related psychiatric conditions – a disservice foisted

on a small subgroup who may suffer from mitochondrial pathology.

Toxic Myopathy

Numerous toxins affect the muscle protein matrix. The most common exogenous muscle toxins are ethanol, hydroxymethylglutaryl coenzyme A (HMG-CoA) reductase inhibitors, and steroids.

Ethanol Ethanol causes a direct toxic effect, raising serum CPK and causing electromyographic changes of myopathy and proximal as well as distal muscle weakness. In severe cases, myoglobinuria occurs; cardiotoxicity is common. Alcohol itself may not be the lone offender; other electrolyte disturbances and vitamin deficiencies are contributory. Abstinence from ethanol is mandatory.

HMG-CoA reductase inhibitors HMG-CoA reductase inhibitors are used commonly for lipid- and cholesterol-lowering effects, having proved to decrease the incidence of fatal myocardial infarction in patients with elevated cholesterol. Unfortunately, these drugs may also interfere with fatty acid metabolism, and muscle damage occasionally occurs with severe elevations in serum CPK, renal failure, and muscle weakness. A specific pathology has not been clearly delineated; however, abnormal energy metabolism within muscles is blamed for muscle breakdown in susceptible patients.

Steroids A number of endocrinological conditions such as thyrotoxicosis and Cushing's syndrome, as well as hypoparathyroidism and vitamin D deficiency, cause myopathy. The most common endocrinologically induced myopathy represents the toxicity of large doses of endogenous glucocorticoids to the muscle milieu.

Acute steroid myopathy Acute steroid myopathy has only recently been distinguished from chronic steroid myopathy. Groups of asthmatic patients given high doses of steroids with or without the concomitant use of nondepolarizing muscle relaxants in intensive care units were found to develop static weakness and myosin-deficiency myopathy by muscle biopsy. Weakness occurring in these conditions is independent of prolonged neuromuscular blockade. It improves. At the present time, there is only supposition as to why the combined use of high doses of steroids and the critical illness state or neuromuscular blockade prompts myosin loss. The trophic influence of the electrochemical message sent through the neuromuscular junction to muscle combined with the

catabolic effects of glucocorticoids probably combine to create this pathology.

Chronic steroid myopathy Unlike acute steroid myopathy, patients on low to medium doses of steroids for any of a number of medical conditions can become progressively weak over weeks to months. The problem tends to be proximally predominant, and the pathology of the condition involves type II muscle fiber atrophy, a similar pathology seen in patients who are chronically bedridden and not exposed to steroids.

Differential Diagnosis of Myopathy

Central Nervous System Disease

It is distinctly uncommon for diseases of the brain, brain stem, or spinal cord to cause weakness bilaterally without other accompanying symptoms, whereas muscle weakness can occur in isolation without myalgia, sensory loss, or clear serum elevations of CPK. Bilateral brain disease resulting in bilateral limb weakness or bilateral cranial nerve abnormalities generally impairs the mental status. It is this preservation of mental status in patients with myopathy that makes distinguishing myopathy from brain or brain-stem diseases simple. It is only slightly more problematic to distinguish spinal cord causes of weakness from myopathy. Mental status is preserved in spinal cord disease unless respiratory muscle function is compromised severely and hypoxemia causes delirium. Sensory loss is the rule in myelopathic disease, as well as severe bowel and bladder dysfunction. The bowel and bladder dysfunction occurring in myopathy is due more to immobilization and the inability of the patient to physically get to a bathroom than it is to any primary pathology of the detrusor or anal sphincter muscles.

Neuropathy

Neuropathy can cause unilateral or bilateral muscle weakness generally with a distal predominance; however, sensory loss occurs in most neuropathies. Very rare neuropathies, such as multifocal motor conduction block neuropathy or primary motor neuropathy from motor neuron disease or lead intoxication, can mimic myopathy, but reflexes are absent or depressed. Patients with severe myopathy may have decreased or even absent reflexes, but this is extremely rare. In difficult cases, distinguishing neuropathic from myopathic weakness depends on electrodiagnosis and muscle biopsy. It is not unusual for an individual undergoing a workup for amyotrophic lateral sclerosis, which affects the lower motor neuron, to undergo

a muscle or nerve biopsy to help distinguish benign myopathic disease from the devastating diagnosis of anterior horn cell disease. However, the majority of neuropathies cause greater distal than proximal weakness, with profound sensory loss early in the course of the syndrome, thus allowing the clinician to distinguish neuropathy from myopathy in most cases.

Neuromuscular Junction Diseases

Patients with neuromuscular junction disease may have a predominant disease of the cranial innervated musculature and present with double vision, lid drooping, dysphagia or dysarthria, or problems with neck extension. Isolated respiratory muscle weakness has also been reported, but not commonly. Extracellular calcium, responding to the action potential reaching the end of the motor nerve, enters the presynaptic nerve through voltage-sensitive channels, causing the vesicles of acetylcholine to couple with the nerve membrane, releasing acetylcholine as the chemical message; this results in muscle contraction in the postjunctional area. Botulinum toxin envenomation creates a condition in which toxin enters the prejunctional membrane, preventing acetylcholine vesicles from coupling with the prejunctional nerve membrane, aborting the message for the muscle to contract. This condition results in the degeneration of this part of the nerve; in a sense, botulinum toxin envenomation not only causes a neuromuscular junction short circuit, but actually causes a distal motor neuropathy. Patients generally need months to regrow this part of the nerve.

Lambert-Eaton syndrome is the prototypic autoimmune disease in which antibodies are raised to voltage-gated calcium channels. Molecular mimicry among moieties on small-cell carcinomas and the prejunctional neuromuscular junction membrane creates a situation whereby the host raises antibodies both to small-cell cancer and the prejunctional membrane, resulting in an attack on the cancer and the fluctuating weakness seen as calcium inefficiently enters the prejunctional neuromuscular junction membrane due to competitive inhibition by these antibodies.

Most postjunctional neuromuscular diseases are the result of abnormalities induced in the coupling of acetylcholine to its receptor in the postjunctional membrane or in abnormalities of the acetylcholine receptor itself. Iatrogenic neuromuscular junction blockade occurs with the use of depolarizing or nondepolarizing muscle relaxants. The autoimmune disease myasthenia gravis results in inefficient neuromuscular transmission and weakness due to abnormalities in the acetylcholine receptor through competitive inhibition by autoantibodies. Similar to

the Lambert-Eaton syndrome, myasthenia gravis can be treated with a number of immunomodulating agents such as steroids or with the removal of exogenous antibody through plasma exchange or through other forms of nonspecific immunosuppression. Many patients have been thought to have conversion reaction because weakness can be subtle. Exogenous stressors such as the flu can stress the neuromuscular junction and cause even more inefficient neuromuscular communication, resulting in fluctuating weakness. Some victims are treated for psychiatric disease, delaying proper treatment.

Myopathy

Myopathy is generally characterized by mild to moderate elevations of the CPK level, greater proximal than distal muscle weakness, with the preservation of deep tendon reflexes and no sensory loss. The varieties of myopathy have already been discussed.

Electrodiagnosis

The proper designation of weakness as originating in the muscle can be made noninvasively through clinical examination and then confirmed through the electrodiagnostic capabilities of electromyography and nerve conduction studies. Using our knowledge of the anatomy of the muscle and recruitment of muscle fibers in the development of muscle contractions, we can evoke patterns of responses from muscles with a needle electrode for quick and efficient disease localization.

Using the principle of differential amplification, a needle electrode with two active recording surfaces separated by an insulator is introduced superficially into muscle. Recording the difference in voltage between the two active surfaces, a simple wave form is generated during contraction. The muscle cell group that contracts, innervated by one particular motor neuron, is termed the motor unit. As the muscle contraction is maintained voluntarily by the patient, a repeating waveform appears on the oscilloscope. Minimal contraction is necessary for this waveform to appear repetitively, up to 5–15 times s^{-1} . As the individual increases the force of the contraction, another motor unit is recruited to increase muscle tension, and another waveform appears on the screen, twitching between 5 and 15 Hz. Because there are hundreds of motor units in a typical muscle, hundreds of motor units appear and create a pattern of discharge on the oscilloscope. In patients with nerve disease, the number of motor neurons is decreased. Therefore, the number of motor units recorded on the screen is decreased. If the nerve disease has been occurring

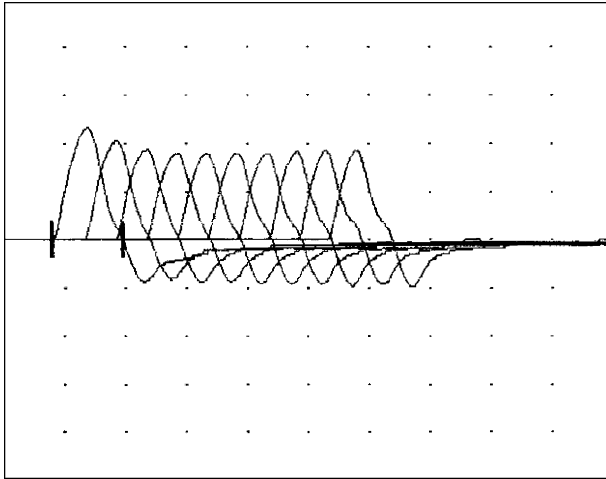


Figure 1 Successive decreases in waveform amplitude during repetitive stimulation testing in a patient with myasthenia gravis.

long enough, the muscle cells that were served by the damaged nerve in question will have been reinnervated by the remaining normal nerves; then each of the waveforms seen will be larger than generally expected. In addition, the rate at which each motor unit contracts is much faster than the normal 5–15 Hz. Neuropathy is then the likely diagnosis.

In contrast, in myopathy the number of motor units remains the same, but the number of muscle cells contracting effectively decreases. Assuming that a random distribution of muscle cells either have died or are damaged significantly, then each resulting motor unit creates a smaller waveform on the electromyographer's oscilloscope. In addition, because a contraction of the weak muscle, even to a moderate extent, necessitates the recruitment of motor units to maintain the same force as in a normal muscle, the contraction pattern appearing on the oscilloscope appears to be that of a maximum voluntary contraction, even though full voluntary contraction has not yet been achieved. Myopathy is then the likely diagnosis.

The repetitive stimulation method may distinguish myopathy from disorders of the neuromuscular junction. Repetitive nerve stimulation results in numerous recorded compound motor action potentials (Figure 1). In the case of neuromuscular junction disease, however, either too little or varying amounts of acetylcholine are released at the neuromuscular junction and fewer muscle cells contract. Less force is generated in the muscle, which expresses itself as weakness clinically and as a decrease in waveform amplitude, as shown in Figure 1.

Muscle Biopsy

It is through sophisticated staining methods of muscle tissue that our understanding of the myopathic

conditions has shifted since the mid-1980s. This has resulted in more accurate diagnoses and in the discovery of new forms of myopathy. Undoubtedly, additional forms of myopathy and a further sub-specification of disorders now lumped together under one designation will be discovered in the near future. One example of this is IBM, which was previously thought to be a steroid-unresponsive form of polymyositis and now enjoys its own designation as a separate disease entity, largely due to sophisticated histopathological techniques.

Therapy

Muscular dystrophies by definition are designated untreatable. Gene therapy trials will probably remove this untreatable designation. At the present time, physical rehabilitation is really our only means of keeping patients with muscular dystrophy as strong as possible. In the case of myotonic muscular dystrophy, in which weakness is generally not the worst problem, the avoidance of late-stage complications such as heart block and respiratory compromise from central neurogenic hypoventilation is the rule. In the case of glycogen-storage diseases, in which the avoidance of sudden exertion and dehydration helps best, genetic counseling is our only other means of prevention. The therapeutic armamentarium in myopathy, however, is greatly widened when we address inflammatory myopathy.

Corticosteroids

Corticosteroids prevent the mobilization of plasma cells and communication among a variety of white blood cells by decreasing interleukin production, minimizing inflammatory reactions and subsequent muscle destruction. This form of immunosuppression is still considered nonspecific in that no one particular process of the immune response is knowingly targeted. Prednisone is the mainstay of treatment for dermatomyositis, polymyositis, and the immune-mediated myopathies accompanying collagen vascular diseases such as systemic lupus erythematosus and rheumatoid arthritis.

However, the previous discussion of steroid-induced myopathy, whether acute or chronic, should not be taken lightly. The very treatment for these myopathies may precipitate a chronic form of weakness and, therefore, the steroid dosage necessary to prevent weakness should be titrated downward to the lowest effective dose. Alternate day therapy may slightly decrease the incidence of steroid myopathy and other abnormal side effects such as ulcer disease, cataract formation, glaucoma, hypertension,

diabetes, and aseptic necrosis of the hip, as well as severe weight gain from lipogenesis. Many inflammatory myopathies burn out and do not need to be treated for more than 2–3 years. Even this length of time may create enough of an opportunity for untoward side effects to occur.

Intravenous Immunoglobulin

Intravenous immunoglobulin (IgG) has failed to show significant results in polymyositis; however, some cases of dermatomyositis stabilize or improve with the use of pooled IgG. Perhaps it is because dermatomyositis is a B-cell-, plasma-cell-, and antibody-mediated disease that the application of large quantities of anti-idiotypic and nonspecific antibodies act as competitive and non-competitive inhibitors, blocking pathogenic host antibodies from destroying muscles. IgG is used more for other causes of weakness at other locations in the neuroaxis, such as the neuromuscular junction in the case of myasthenia gravis, and for Guillain-Barré syndrome, an inflammatory proximal neuropathy.

Plasmapheresis

The removal of plasma proteins, cytokines, and immunoglobulins and their replacement either with another individual's plasma or with an albumin and saline solution has been used in other immune-mediated conditions, both neurological and non-neurological. A controlled trial of plasma exchange in patients with polymyositis and dermatomyositis failed to show any benefit, however.

Conclusion

Current medications are helpful but are also toxic when used chronically, so although medical science can provide benefits for some clearly defined myopathic syndromes, we can expect more clinical trials of other pharmaceuticals. The variety of myopathies mirrors the variety of diseases in other parts of the

neuroaxis. Genetic influences, environmental influences, immunological influences, and toxic causes of weakness clearly illustrate the vulnerability of muscle tissue to a variety of stressors. However, psychic stress does not cause weakness. It can create fatigue, but clearly it does not create a consistently measurable abnormality of force generation due to muscle damage or neuromuscular junction disease. Therefore, weak patients need to be worked up for organic diseases. The complicated structure of nerve, neuromuscular junction, and muscle allows for breakdowns at any point along the way, and many diseases result, some of which have been clearly defined; all of which are measurable as weakness at the bedside, in the electrophysiology laboratory, and pathologically.

The physiological muscle response to numerous exogenous and endogenous stimuli may be normal contraction, paralysis, or anatomic disruption. The most explicit exposition of stress affecting muscle occurs in acute steroid myopathy, in which normal contraction may be affected to the point of paralysis without anatomic disruption of the muscle membrane and intracellular contents. In addition, acute steroid myopathy may also express its severity in myosin dissolution and the eventual anatomic disruption of the muscle membrane. This form of toxic myopathy best illustrates the effects of stress on muscle.

Further Reading

- Engel, A. and Franzini-Armstrong, C. (1994). *Myology* (2nd edn.). New York: McGraw-Hill.
- Griffin, J., Low, P. and Poduslo, J. (1993). *Peripheral neuropathy* (3rd edn.). Philadelphia: Saunders.
- Kimura, J. (1981). *Electrodiagnosis in diseases of nerve and muscle: principles and practice* (2nd edn.). Philadelphia: Davis.
- Rowland, L. (1995). *Merritt's textbook of neurology* (9th edn.). Baltimore, MD: Williams & Wilkins.
- Sethi, R. and Thompson, L. (1989). *The electromyographer's handbook* (2nd edn.). Boston: Little, Brown and Company.

N

Natural Disasters See: Community Studies; Disasters and Mass Violence, Public, Effects of.

Natural Killer (NK) Cells

T L Whiteside, M Boyiadzis and R B Herberman

University of Pittsburgh Cancer Institute and
University of Pittsburgh School of Medicine,
Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
T L Whiteside, A Baum, and R B Herberman, volume 3, pp 1–7,
© 2000, Elsevier Inc.

Characteristics of Natural Killer Cells
Biological Importance of NK Cells
Measurement of NK Cells
NK Cells in Human Disease
NK Cells and Aging
Stress and NK Cells

Effector cells
NK cells

NK cells that induce lysis of target cells. A subpopulation of lymphocytes characterized by the expression of surface markers that are distinct from those present on T or B lymphocytes. NK cells have a unique capability to recognize and eliminate a broad range of abnormal cell targets without prior sensitization. They are components of innate immunity responsible for immune surveillance against infectious agents and transformed cells.

NK-cell assay

An assay used to measure NK activity by coincubation of effector (E) with radioactively labeled target (T) cells at several different E : T ratios. The assay estimates the percentage of specific lysis attributable to NK cells based on levels of radioactivity released from a defined number of target cells.

Glossary

Activated
NK cells

A subset of NK cells responding to stimulatory signals that have upregulated expression of receptors for interleukin-2 as well as other surface molecules and have acquired new functional attributes that facilitate cellular interactions.

Antibody-
dependent
cellular
cytotoxicity

A form of cell death mediated by NK cells in the presence of an antibody that is specific for an antigen expressed on the surface of target cells and decorates the NK cell through binding to a receptor for immunoglobulins. This antibody forms a bridge between effector and target cells.

Cytotoxicity

One of the functions of NK cells commonly used for assessing the level of NK activity in the peripheral blood, other body fluids, or mononuclear cells isolated from tissues.

Characteristics of Natural Killer Cells

NK cells are effectors of the innate immune system that play an important role in host response against viruses and tumor cells through the production of cytokines and direct cytolytic activity. Unlike other effector cells, NK cells are able to kill these targets spontaneously, without having had prior encounters with them. NK cells also appear to be highly sensitive to stress, exhibiting dramatic changes in function in response to acute or chronic stress. Because they appear to be involved in important immunological defense against infectious agents and against cancer, and because they are affected so readily by stress or emotional state, NK cells have been a well-studied variable in the stress–illness relationship.

Natural killer cells were first described in the early 1970s based on their functional capability to lyse

tumor cells in the absence of prior stimulation. This characteristic of NK cells was considered to fit their role as mediators of immune surveillance. Since then, NK cells have been shown to mediate other biologically important functions. NK cells are now known to produce cytokines that regulate the development of acquired specific immunity, play a role in the rejection of bone-marrow grafts, be responsible for the selective elimination of abnormal cells, and be implicated in the development of autoimmune diseases and the graft versus host disease (GVHD).

NK cells make up from 10 to 15% of human peripheral blood lymphocytes, as determined by phenotyping using labeled monoclonal antibodies (MAbs). NK cells can home to secondary lymphoid organs. It is estimated that approximately 5% of mononuclear cells in uninflamed lymph nodes are NK cells, and they constitute 0.4–1% of the lymphoid cells found in inflamed tonsils and lymph nodes. Thus, the extravascular lymphoid tissues represent a large pool of innate effector cells because, in aggregate, lymph nodes harbor 40% of all lymphocytes, whereas peripheral blood contains only 2%.

NK Cell Development and Subsets

NK cells originate in the bone marrow from CD34+ hematopoietic progenitor stem cells and require the bone marrow microenvironment for complete maturation. In the presence of stromal cell-derived growth factors, c-kit ligand and flt-3 ligand, the CD34+ stem cells acquire the surface receptor for the cytokines interleukin (IL)-15/IL-2. The acquisition of the IL-15/IL-2 receptor renders the NK progenitors (NKP) responsive to IL-15, a critical cytokine for NK cell development. The next steps of development involve the acquisition of the unique NK cell receptor repertoire and proliferation into phenotypically and functionally mature NK cells.

Two subsets of NK cells are present in the peripheral blood of healthy adults. The majority of NK cells express low levels of CD56 (CD56^{dim}) and high levels of Fcγ receptor III (FcγRIII, CD16^{high}), whereas 10% of NK cells are CD56^{bright} CD16^{dim/−}. In the lymph nodes, the majority of NK cells belong to the NK-cell bright subset. These two NK subsets differ in receptor expression, trafficking, cytokine production, and the ability to mediate cytotoxicity. CD56^{dim} NK cells are more cytotoxic and mediate higher antibody-dependent target cell lysis but produce low levels of cytokines (Table 1). In contrast, CD56^{bright} NK cells can produce abundant cytokines, but exert low natural cytotoxicity and antibody-dependent cellular cytotoxicity (ADCC).

Table 1 Subsets of natural killer cells in humans^a

	CD56 ^{bright}	CD56 ^{dim}
CD56	++	+
CD16	−/+	++
NK receptors		
KIR	−/+	++
CD94	++	−/+
NG2A	+	−/+
Cytokine and chemokine receptors		
IL-2 Rαβγ	+	−
IL-2 R βγ	+	+
c-Kit	+	−
CCR7	++	−
CXCR1	−	++
CX ₃ CR1	−	++
Adhesion molecules		
CD2	++	+
CD44	++	+
CD62L	++	−/+

^a++ denotes high-density expression; + denotes low-density expression; −/+ denotes variable expression; − denotes no expression. CCR7, chemokine receptor C7; CX, chemokine receptor; IL, interleukin; KIR, killer immunoglobulin-like receptor.

Receptors

The triggering of NK cell cytotoxicity is determined by the balance of activating and inhibitory signals received via the broad repertoire of receptors found on individual NK cells. These receptors monitor the interactions of NK cells with other effector cells of the immune system and with cellular targets (Figure 1). In humans, there are three main classes of receptors: natural cytotoxicity receptors (NCRs), killer immunoglobulin-like receptors (KIRs), and killer C-type lectin receptors.

The NCRs, which bind to as yet unidentified tumor antigens, are all activating receptors and have been demonstrated to be involved in the NK-cell-mediated killing of various cancer cell lines. Three NCRs (NKp46, NKp44, and NKp30) have been identified. NKp46 and NKp30 are constitutively expressed by all peripheral blood NK cells and are not found on other immune cells. NKp44 is not expressed by resting NK cells, but is upregulated by IL-2 stimulation.

The KIRs belong to the immunoglobulin superfamily. There are two functionally distinct sets of KIRs: inhibitory and activating. KIRs specifically recognize major histocompatibility complex (MHC) class I alleles, including groups of HLA-A, -B, and -C molecules. Each specific set of inhibitory and activating KIRs has an identical extracellular domain and, consequently, each set binds to identical ligands. However, because of differences in their cytoplasmic domains, one set of KIRs signals an inhibitory response and the other set signals an activating response following their

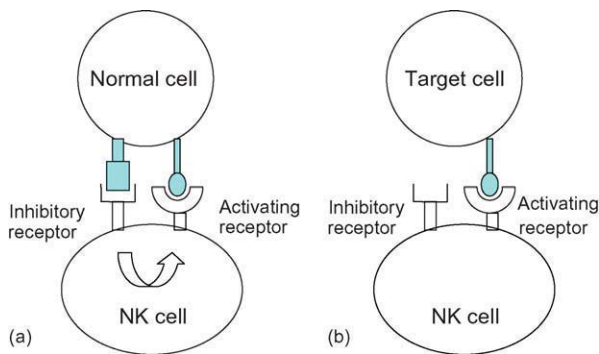


Figure 1 Natural killer cell cytotoxicity. a, No lysis; b, lysis. NK cell cytotoxicity depends on a delicate balance between activating and inhibitory receptors. The presence of activating receptors and inhibitory receptors on the surface of NK cells is necessary to protect normal cells from lysis. Inhibitory signals normally predominate over activating receptors, and a cell able to induce such signals (e.g., through expression of abundant MHC molecules) avoids lysis. This process occurs constantly as NK cells survey normal host tissues. However, when activating receptors engage their ligands on target cells in the absence of inhibitory receptor–ligand interactions, a net activating signal is generated, resulting in lysis of the target. Virally infected or tumor cells downregulate ligands that are recognized by the NK-cell inhibitory receptors (i.e., MHC molecules) and, thus, NK cells are able to recognize these targets as nonself and to kill them.

binding to identical MHC class I alleles. Inhibitory KIRs usually have a higher affinity for MHC-class I ligands and, therefore, coligation of both activating and inhibitory receptors results in a net negative signal. Presumably, this double set of receptors with opposing signaling properties exists to regulate NK activity and to protect normal cells from being eliminated.

The killer C-type lectin receptors include inhibitory heterodimers of CD94 complexed with NKG2A or NKG2B and activating heterodimers complexed with NKG2C or NKG2E. These receptors all recognize the overall expression of MHC class I, in that they bind to HLA-E, which is expressed only in association with MHC leader peptides. Another activating C-type lectin receptor, NKG2D, is expressed as a homodimer that binds to peptides associating with the stress-inducible molecules MIC-A, MIC-B, or ULBPs.

In addition to the various activating and inhibitory receptors, NK cells express several receptors that play an important role in enhancing their interaction with target cells and enhancing the NK-cell-mediated cytotoxicity (Table 2).

Mechanisms of Cytotoxicity

NK cells possess a unique capacity for recognizing and killing a broad range of abnormal targets such as virally infected or transformed cells and certain allogeneic targets. Although the molecular events involved in the process of target recognition and

Table 2 NK-cell receptors and ligand specificity^a

	Activating	Inhibitory	Ligand specificity
KIRs	+	+	MHC class I alleles including groups of HLA-A, -B, -C
NCRs	+		Unknown
C-type lectin receptors			
CD94/NKG2A/B		+	HLA-E
CD94/NKG2C	+		HLA-E
NKG2D	+		MICA, MICB, ULBP-1, -2, -3
Coreceptors			
CD16 (FcγRIII)			Fc of IgG
CD2			CD58
2B4			CD48
CD 40 ligand			CD40

^aKIRs, killer immunoglobulin-like receptors; MHC, major histocompatibility complex; NCRs, natural cytotoxicity receptors.

signaling are not fully known, the cellular alterations that result from effector (E) cell–target (T) cell interactions have been identified. NK cells can use two different mechanisms for killing, and several distinct sequential steps are involved in the process of target cell lysis.

- *Perforin-mediated cytotoxicity.* NK cells can lyse target cells through a calcium-dependent release of intracytoplasmic granules containing perforin or granzymes A and B. This secretory activity mediated by perforin/granzyme release results in necrosis and is directed primarily at NK-sensitive targets, such as K562 tumor cells. Perforin forms pores in target cells, whereas the granzyme release is necessary for target cell killing.

- *Induction of apoptosis through death receptors.* Activated NK cells express the Fas ligand (FasL) and can kill Fas+ (CD95+) targets through the activation of enzymes that fragment DNA. In addition, NK cells express tumor necrosis factor (TNF)-related apoptosis-induced ligand (TRAIL) and can thereby activate apoptosis in cells that express TRAIL receptors, including death receptor 4 and death receptor 5. Nonsecretory NK activity (surface receptor-mediated apoptosis) is preferred for NK-resistant targets such as fresh tumor or solid tissue cells.

Activated NK cells can rapidly kill a number of targets by combining these two mechanisms. NK cells that survive a lytic event are able to recirculate and can bind to another susceptible target and kill it. The efficiency of NK cell-mediated lysis (killing) is dependent on the ratio of NK cells to abnormal cells and on the level of NK-cell activation.

Biological Importance of NK Cells

NK cells participate either directly or indirectly in a variety of essential biological processes, ranging from reproduction to immunoregulation (Table 3).

Defense against Infections

The major function of NK cells appears to be defense against infection. They have been shown to mediate antiviral, antifungal, and antibacterial activities *in vitro*. NK cell importance in the control of infections *in vivo* is supported by studies of experimental viral infections in mice and by reports correlating low NK activity with increased sensitivity to frequent and severe viral infections in humans, including herpes simplex virus (HSV), Epstein–Barr virus (EBV), and cytomegalovirus (CMV) infections. *In vivo* responses of NK cells to a viral challenge involve increases in cytotoxicity, proliferation, migration of NK cells to infected tissues, and production of interferon (IFN)- γ by NK cells. Virus-induced interferons increase NK cell functions and also modulate IL-12 expression, contributing to the development of an antiviral milieu early on during infection and prior to the appearance of virus-specific T cells. Activated NK cells may reduce viral replication by killing infected host cells, and they are able to produce a variety of cytokines that can inhibit viral replication as well as facilitate antigen processing or antigen presentation by increasing the expression of MHC molecules.

Antitumor Effects

NK cells were first discovered because of their ability to kill *in vitro* certain tumor cell lines. Subsequently, NK cell antitumor activity has been convincingly demonstrated in several animal models of cancer and metastasis. These experimental studies suggest that activated NK cells play an important role in limiting both local tumor growth and dissemination of metastases. In particular, activated NK cells are thought to be responsible and ideally suited for antitumor surveillance because they appear to be able to enter solid tissues, migrate to sites of metastasis, and eliminate malignant targets while sparing normal tissue cells. The depletion of NK cells from animals enhances the *in vivo* growth of implanted tumor cells, whereas the

administration of NK cells to animals with metastatic cancer eliminates the tumor and prolongs survival. In a prospective study following 3500 Japanese over 11 years, the incidence of cancer was significantly increased in those with lower initial NK cytotoxic activity. In gastric and colon cancer patients, lower NK cytolytic activity at diagnosis correlated with higher tumor volume, metastases, and a worse prognosis. These data suggest that an adequate NK cell number and levels of NK activity are needed to mediate immune surveillance against malignancy.

Regulation of Immunity

NK cells can exert immunoregulatory activity through their interactions with other immune effector cells and by producing several cytokines, bridging innate immunity with adoptive immunity. Recently, dendritic cells (DC) have been implicated in the activation of NK cells. The activation of NK cells by DCs may play a role in shaping the emerging adaptive responses.

NK cells are constantly primed for activation by normal host cells, but their activation is negatively balanced by the inhibitory receptors that recognize self MHC class I antigens. A loss of this self-inhibitory mechanism could lead to both cytotoxicity and cytokine release by NK cells and may play a role in autoimmunity. Alterations in NK cell functions have been described in patients with autoimmune diseases. Such patients may have low levels of natural killing due to both reduced levels of circulating NK cells and defects in lytic activity.

Measurement of NK Cells

The measurement of NK cells in body fluids or cell suspensions consists of determinations of the NK cell number and NK activity. Both are necessary for an adequate evaluation of natural immunity or for following its changes during disease or in response to stress. In most instances, the level of NK activity per cell is the best estimate of the functional potential of NK cells.

NK cell numbers are generally measured using monoclonal antibodies specific for NK cell-associated markers (e.g., CD56 and CD16 in humans, NKR-P1 in rats, and NK1.1 in mice), which are labeled with a fluorescent dye. Following staining, the percentage of positive cells in a population is determined by flow cytometry; this percentage value is then converted into the absolute number of NK cells based on the simultaneously obtained differential lymphocyte count. Or, in a single-platform flow cytometry method, fluorosphere beads are used to allow for the direct calculation of the stained NK cells in a tube.

Natural killer activity has been traditionally measured in short-term (4-h) ^{51}Cr release assays, in

Table 3 Biological significance of NK cells

Elimination of intracellular pathogens: viruses, fungi, and bacteria
Surveillance against cancer and elimination of metastases
Regulation of immunity
Involvement in pregnancy and fertility
Regulation of liver growth and regeneration
Regulation of hematopoiesis
Interactions with the neuroendocrine system

which a standard number (5×10^3) of cultured K562 cells is used as radiolabeled targets (T) and peripheral blood mononuclear cells as effectors (E). The latter are adjusted in number to give at least four different E : T ratios in the assay. This allows for the generation of a lytic curve by plotting the percentage of specific lysis against the E : T ratio. More recently, flow cytometry methods have been introduced that simultaneously measure NK-cell cytotoxicity and NK-cell phenotype. These methods are gaining acceptance because they dispense with radioactivity, are cost effective, and offer an opportunity to evaluate several distinct characteristics of effector cells. However, these methods have not yet been validated against the Cr-release assay, which remains the gold standard for measuring NK activity. The results of NK-cell assays are usually expressed as lytic units (LU) calculated from the lytic curve and defined as the number of E required to kill 5×10^3 targets in a batch of tested cells. Computer programs are available to facilitate the transformation of these data.

NK Cells in Human Disease

NK cell abnormalities either in NK-cell number or activity have been described in human diseases. In either case, only chronic abnormality in NK cell activity or number is considered to be pathological because transient changes accompany a variety of physiologically normal events (e.g., circadian variations, daily stress, exercise, or a common cold). Although the absolute number of NK cells in the peripheral circulation may be significantly below or above a normal range, this does not mean that NK activity is also abnormal. Often, individuals with a low number of NK cells have normal NK activity because their NK cells are activated. Conversely, some individuals with a high NK cell number may have little NK activity due to a lack of or low function of these cells. Therefore, to evaluate the functional potential of NK cells appropriately, it is necessary to measure both NK-cell number and NK activity and to express NK-cell functions on a per cell basis.

Chronically low levels of NK cell activity are seen in cancer, acquired or congenital immunodeficiency diseases (e.g., AIDS), severe life-threatening viral infections (HSV, EBV, and CMV), a subset of patients with chronic fatigue syndrome (CFS), certain autoimmune diseases (e.g., connective tissue diseases), and behavioral disorders (e.g., depression). Persistently elevated NK cell activity is relatively rare, being reported in autoimmune liver disease (including viral hepatitis) and in NK-cell lymphoproliferations. The latter present as elevations in the number or activity of NK cells or both and are characterized by a benign course,

lymphocytosis of large granular lymphocytes (LGL), cytopenias, and splenomegaly. Rare acute proliferations of LGL have been reported and are consistent with aggressive leukemia/lymphoma.

Persistently low levels of the NK-cell number or activity in patients with advanced metastatic or other diseases may be associated with more severe symptoms or an increased risk of disease progression. Therefore, it may be advantageous for patients with chronically low NK cell activity to receive therapy designed to augment NK cell functions. Such therapies are available and consist of the administration of agents known to upregulate endogenous NK activity (e.g., interferons, IFNs; IL-2) or an adoptive transfer of autologous A-NK cells, either systemically or regionally. These therapeutic strategies have not been extensively evaluated, however. More recently, adoptive transfers of NK-92, an NK-cell line established from a patient with non-Hodgkin's lymphoma, have been used in the therapy of hematological malignancies and solid tumors with acceptable toxicity and the goal of providing the host with an excess of antitumor effector cells able to eliminate tumor cells.

NK Cells and Aging

The cells of the immune system are constantly renewed from hematopoietic stem cells. With age, a reduction in the overall capacity for the renewal of these stem cells as a whole has been observed. Unlike T and B cells, the absolute number of NK cells is increased in older individuals compared to young or middle-aged groups, and IFN- γ production and phagocytosis are also increased. Total NK-cell cytotoxicity is stable, however, so the NK-cell cytotoxicity on a per cell basis is impaired in older individuals, as is the response of NK cells to IL-2. The age-associated increases in the number of NK cells and in T cells expressing NK receptors might reflect an alteration in regulatory mechanisms that favors NK-cell functions associated with the regulation of immune responses at the expense of antitumor functions, for example.

Stress and NK Cells

Interactions between NK cells and neuroendocrine systems have received increasing attention. NK cells express receptors for neuropeptides, opioids, prolactin, and other neurohormones, which regulate interactions between the brain and the immune system. Therefore, NK cells are highly responsive to changes in the serum levels of these hormones. Stress often involves all these hormones, and the effects of stress on the immune response can be monitored by serially following changes in NK cell activity.

For example, major stressful life events in otherwise healthy individuals are generally associated with low NK cell activity. However, this effect of stress on NK cells is dependent on several other variables, including the duration of a stressful event, how long afterward responses are measured, and variables that affect the intensity of stress experienced.

Several studies have addressed effects of stress on NK cells and have begun to identify mechanisms underlying these effects. These studies have examined stress effects on overall NK activity, on the number of NK cells in circulation, and on the level of activity per NK cell. NK cell numbers tend to change as NK cells leave the bloodstream and migrate into tissue or lymph nodes. Stress-related changes in cell numbers and activity reflect this process, and thus NK activity per cell is the most reliable way of measuring the effects of stress on these cells.

Two primary systems are involved in the regulation of stress: the sympathetic nervous system (SNS) and the hypothalamic-pituitary-adrenocortical (HPA) axis. SNS activation is initially achieved through neural stimulation and is more or less immediate. Measurable increases in organ or system function tied directly to SNS activation (e.g., heart rate) can be detected within 1 min after the onset of a stressor. Endocrine activity, principally catecholaminergic, intensifies and extends the duration of neural stimulation, and the release of epinephrine and norepinephrine from the adrenal medullae and sympathetic neurons is reliably associated with stress.

The effects mediated by the HPA axis, which culminate in the increased release of glucocorticoids (i.e., cortisol) from the adrenal cortexes, are slower to act or to recover. Ordinarily, changes in HPA activity are not detectable for several minutes, but once initiated, they may persist for hours. This suggests that initial or acute stress responses are governed principally by the SNS, which may have a stimulatory effect on NK activity or may affect cell migration (more NK cells released into circulation). In general, acute stress lasting as little as a few minutes is associated with large, rapid increases in NK-cell number in peripheral blood circulation and a corresponding increase in overall cytotoxicity. Chronic stress lasting days, weeks, or months appears to be associated with the suppression of NK-cell number and cytotoxicity (Table 4). Differences in how acute and chronic stress affect NK cells provide some insight into the mechanisms, functions, and consequences of stress modulation of NK activity, as outlined briefly in the following section.

NK Cells and Acute Stress

Acute stress is almost always associated with a rapid release of catecholamines, leading to an immediate increase in NK-cell number in the peripheral

Table 4 Summary of the effects of acute and chronic stress in human NK cells^a

	<i>Acute stress</i>	<i>Chronic stress</i>
Cell number	++	--
Cell cytotoxicity	+	-
Overall cytotoxicity	++	-
Cytokine production	+	++ (pro-inflammatory)
Possible mechanisms	SNS, HPA axis	SNS, EOPs, HPA axis

^a++ and +, highest/increased effects; - and --, lowest/decreased effects; EOPs, endogenous opioid peptides; HPA, hypothalamic-pituitary-adrenal; SNS, sympathetic nervous system.

circulation, probably due to changes in NK-cell trafficking (i.e., increased movement from lymphoid or nonlymphoid tissue into the peripheral circulation as a result of a reduction in NK cell adherence to endothelial cells combined with an increase in blood flow). This increase in the number of NK cells is accompanied by increases in individual cell function, as observed in some studies but not in others. However, an acute elevation of glucocorticoids within 10–20 min after stress causes immune cells to marginate and subsequently to enter target tissues, probably as a result of the local release of cytokines, which act as chemoattractants. Hence, the enhancing effect of acute stress on immune cells is short-lived and is followed by a suppression of NK cell activity that may be unrelated to changes in NK-cell number. These effects have been observed in naturalistic studies of tandem parachute jumpers and in laboratory studies using far less severe stressors. In numerous studies, the time course of acute stress on NK-cell number appears to be biphasic, increasing on the initial encounter with stress and decreasing rapidly to the baseline resting level or below as soon as the stressor is terminated. Thus, NK-cell-enhancing effects of acute stress are reversible. Age, social support, time of day, and other variables appear to modulate the influence of acute stress on NK cells. The acute reactivity of NK cells during or shortly after an acute stressor may be correlated with levels of SNS activation (e.g., epinephrine and norepinephrine levels) and varies across the time of day. It has been observed that the most labile responses in NK cell activity occur late in the afternoon. A variety of measures thought to completely or partly reflect sympathetic arousal, including blood pressure and stress hormone levels, have been correlated with changes in NK-cell number and function.

The available studies underscore the complexity of acute stress responses and factors determining NK activity. They also suggest that the regulation of immune responses to acute stress depends on both catecholamine and cortisol phases of the response.

Furthermore, there may be several other pathways (e.g., those involving modulation by endogenous opioids) by which acute stress might affect the activity of NK cell.

NK Cells and Chronic Stress

Chronic stress, involving the repeated or sustained elevation of stress hormones, appears to suppress NK cell activity as well as the number of cells in circulation. For example, studies of medical students at the time of low and higher stress periods confirmed that NK cells were suppressed during prolonged stress. Similarly, studies of caregivers for Alzheimer's disease patients and of bereaved populations also provide evidence of chronic suppression of NK cell activity. In one study, bereaved women exhibited lower NK cell activity than nonbereaved participants, and in another study, caregivers were observed to have lower NK cell activity and increased plasma cortisol concentrations than controls. Studies of disasters, such as earthquakes or nuclear accidents, have also provided evidence of a lower number of NK cells in stressed groups, particularly if participants reported greater worry, intrusive thoughts about the disaster, or general persistent distress. Depression in young adults is associated with depressed NK activity. The effects of chronic stress on NK cells are clearly different from those of more short-lived acute stress, and it is possible that the suppression of cytotoxicity, cytokine release, and/or the NK-cell number observed in chronic stress is due to more slowly acting hormonal systems. It appears that long-term (i.e., chronic) responses to stress are also likely to be characterized by SNS and HPA activity, but the latter is thought to exert predominantly suppressive, nonreversible influences on the immune system. Lasting stress also appears to deplete the number of NK cells in blood, although it is not clear where these cells go when they have left the bloodstream or how these changes differ from the reversible effects of acute stress. The potential mechanisms responsible for these long-term effects are presently unknown. Nevertheless, the potential consequences of having persistently low NK cell activity related to chronic stress may represent a health hazard, although no direct relationship has been established so far between disease development and chronically low NK activity.

NK Cells and Cancer

Based on NK cell responsiveness to stress, many studies have used NK activity as a biomarker of stress, attempting to correlate it with health, disease, and the ability to cope with life. This is particularly evident in studies linking stress and depression with

development of cancer. By and large, these studies were performed with small cohorts of patients and yielded conflicting results. Although the literature dealing with the involvement of stress in cancer development, cancer progression, or response of cancer to treatment is extensive, few large-cohort investigations of immune surveillance and stress are available. In one such study of 6284 Jewish Israelis who had lost an adult son, there was an increased incidence of melanoma in the parents of accident victims and war casualties relative to nonbereaved members of the population. Unfortunately, no NK-cell numbers or activity was measured in this study. Because stress and depression have been linked with the altered cytokine profile (e.g., increased pro-inflammatory cytokine production), downregulation of MNC class I and class II molecules, and reduced NK activity in numerous studies and because the same biological effects accompany cancer progression, the tendency has been to seek correlations among stress, NK activity, and cancer and to speculate that NK activity may be a surrogate marker linking stress to cancer. However, inconsistent preliminary data and a lack of prospective long-term studies in large cohorts of humans do not at this time permit the conclusion that NK activity can be considered a biomarker of stress or a biomarker of cancer development progression.

See Also the Following Articles

Depression Models; Industrialized Societies; Lymph Nodes; Immune Surveillance – Cancer, Effects of Stress on; Immunity; Immune Response; Immune Cell Distribution, Effects of Stress on; Immune Function, Stress-Induced Enhancement.

Further Reading

- Bryceson, Y. T., March, M. E., Barber, D. F., et al. (2005). Cytolytic granule polarization and degranulation controlled by different receptors in resting NK cells. *Journal of Experimental Medicine* 202, 1001–1012.
- Hamerman, J. A., Ogasawara, K. and Lanier, L. L. (2005). NK cells in innate immunity. *Current Opinion in Immunology* 17, 29–35.
- Lanier, L. L. (2003). Natural killer cell receptor signaling. *Current Opinion in Immunology* 15, 308–314.
- Rajagopalan, S. and Long, E. O. (2005). Understanding how combinations of HLA and KIR genes influence disease. *Journal of Experimental Medicine* 201, 1025–1029.
- Reiche, E. M. V., Nunes, S. O. V. and Morimoto, H. K. (2004). Stress, depression, the immune system, and cancer. *Oncology* 5, 617–625.
- Whiteside, T. L. and Herberman, R. B. (1995). The role of natural killer cells in immune surveillance of cancer. *Current Opinion in Immunology* 7, 704–710.

Necrotic Neural Injury See: Brain Trauma; Glucocorticoids – Adverse Effects on the Nervous System.

Negative Affect

A A Stone and A A Gorin

State University of New York, Stony Brook, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 8–11, © 2000, Elsevier Inc.

Definition of Negative Affect
 Role of Negative Affect in Stress Research
 Conceptual Issues
 Measurement

Glossary

<i>Arousal</i>	A state of activation, surprise, and intensity.
<i>Circumplex model</i>	A dimensional approach to mood wherein two independent dimensions are crossed to form a two-dimensional space encompassing all mood states.
<i>Negative affectivity</i>	A stable personality characteristic associated with high levels of negative mood; sometimes referred to as trait negative affect or neuroticism.
<i>Positive affect</i>	Experiences high in the qualities of joyfulness, happiness, pleasantness, and cheerfulness.
<i>Specific affect model</i>	Position that there is a limited set of distinctive moods that do not fall along one or two dimensions.

Definition of Negative Affect

Negative affect refers to a set of states that fall into the broad class of phenomena known as mood, emotion, and affect. Emotions and affects sometimes refer to shorter fluctuations than mood (minutes versus hours), although these terms are often used interchangeably. These terms all refer to constructs that include subjective feelings, behavioral changes, and physiological states, and the majority of research on negative affect has focused on subjective states. Surprisingly, there are only weak to moderate associations

among subjective, behavioral, and physiological indicators of mood. Research in the field of psychophysiology, for instance, has found that facial expressions associated with an affective state (say anxiety) are only moderately correlated with the subjective experience of either anxiety or physiological activation. This suggests that the three domains of mood may be measuring somewhat different aspects of the construct.

The functions of moods and emotions have been debated since Darwin, but most theorists agree that emotions have a strong biological basis, assisting in the preparation, shifting, and maintenance of action. In everyday life, individuals often assign emotions a primary role when describing the causes of their behavior. In addition to action, emotions and mood can also influence attention and memory. Emotional arousal, for example, may narrow attentional focus, and the recall of certain memories may be mood-dependent. While there is considerable debate about the function and definition of mood, for this article, we will largely focus on the subjective component of negative affect because this is the most common and useful definition available. Also, using a multidimensional definition of mood results in a tremendous overlap with other components of the stress model (namely, the environmental and physiological aspects of stress).

Role of Negative Affect in Stress Research

In order to understand the importance of negative affect in stress research, we first place it in the context of a general model of stress. As earlier articles of this work have shown, there are different conceptual models for thinking about stress. One model states that environmental changes interact with individuals' predisposing psychological states to produce a subjective experience and a concurrent set of physiological changes. Some stress theorists focus entirely on alterations in the environment to define stress, for example, using life events to define stress. Other theorists focus on the resulting physiological states; a good example of this is the seminal work of Hans Selye. Still other theorists base their definition of stress on the subjective

perception of stress, which is the underlying conceptualization used in the Perceived Stress Scale.

Clearly, the latter definition of stress is closely linked to the construct of negative affect. It is difficult to imagine an individual reporting that they are feeling stressed, out of control, and unable to cope, yet not feeling high levels of negative affect at the same time. In a similar vein, it is difficult to imagine how the undesirable events that environmentalists use to define stress would not invoke negative affect. Death of a loved one is a major life event that has been used to define stress, and certainly negative affect is associated with mourning the death of a loved one. However, theorists who define stress as any alterations in one's environment (even positive ones) would not necessarily require associated alterations in negative moods.

Conceptual Issues

Behavioral scientists have devised several ways of measuring negative affect and the more general concept of mood. These range from single-item measures that are anchored with positive and negative descriptors at opposite ends of the scale to much more complicated, multidimensional adjective checklists. Several conceptual issues need to be discussed to place assessment methods in an appropriate context.

Structure of Negative Affect

The first issue pertains to the structure of mood. Two opposing views have emerged about how to best represent structure, namely, the specific affect approach and the dimensional approach. The specific affect approach views mood as being made up of a number of distinct states that, though related to one another, have distinguishing characteristics. According to this approach, different negative affects are associated with unique patterns of physiological activation, for example, the type of arousal associated with fear differs from that associated with anger. Both of these states are considered negative affect, yet each is associated with a distinctive pattern of subjective, physiological, and facial characteristics. In contrast, the dimensional approach rejects the idea of specific affects in favor of a limited number of core affective dimensions. One of the dimensions often discussed concerns the degree of negativity. The most developed dimensional model is based on two affective dimensions that are crossed to yield a two-dimensional space. Like primary colors mixed together in order to create a vast array of hues, the moods in this two-dimensional space are each composed of a mixture of positive and negative affect. Within the dimensional

approach to mood, there are differing viewpoints concerning how to label the dimensions and create the two-dimensional space. Watson and Tellegen advance the position that positive affect and negative affect are the two core dimensions, whereas Russell advances the position that a single bipolar dimension of positive/negative is crossed with a unipolar activation dimension. Both viewpoints have support in the literature and scales based on each approach are currently in use.

Regardless of the specific dimensional labels used, poles of the dimensions are typically anchored with representative adjectives. The pleasant pole is generally indicated with adjectives such as happy, delighted, glad, cheerful, warm-hearted, and pleased, whereas the unpleasant pole is often marked with adjectives such as unhappy, miserable, sad, gloomy, and blue. Proponents of the activation dimensional construct label the high activation end with words such as aroused, astonished, stimulated, surprised, active, and intense, whereas the low activation end is marked with words such as quiet, tranquil, still, inactive, idle, and passive.

Content of Negative Affect

The second conceptual issue concerns the way people experience and recall their moods. When asked to report about their moods over a distinct period of time, for example, the last day or week, do people report about the frequency of their moods or do they report about the intensity of their moods? This question was addressed in a study by Diener and colleagues. By varying instructions on a mood checklist, they found that reports of mood intensity were not related to reports of mood frequency, indicating that frequency and intensity are two different ways of conceptualizing mood. From an applied perspective, when people are asked in a more general way how their mood was over the last day or over the last week, it is not clear whether they are reporting about the frequency or intensity. This ambiguity requires future research in order to facilitate interpretation of mood assessments.

Recall of Negative Affect

A third conceptual issue concerns individuals' abilities to recall mood states and whether reports of mood over long periods of time are accurate. Research on autobiographical memory has shown that there are a number of biases that occur in retrospective recall. For example, comparisons of momentary (immediate) and retrospectively recalled moods suggest that rather than providing an accurate summary of moods, retrospective reports are heavily based on

more recent experiences. Several other biases in the recall of mood have been identified, leading some experts to advocate that a summary of momentary moods may be most appropriate for characterizing mood over longer periods of time.

Stability of Negative Affect

A fourth issue concerns the stability of negative affect over time, an issue related to the prior two points. A distinction has been drawn in the literature between state negative affect, which refers to transient fluctuations in negative emotions, and trait negative affect, which refers to enduring and stable levels of negative affect over extended periods of time. Trait negative affect, also known as negative affectivity and neuroticism, is associated with trait anxiety, poor self-esteem, introspective and ruminative behaviors, and negative views of the world and others. Important to the conceptualization of negative affect, recent studies have revealed that state negative affect and negative affectivity have differential influences on health. A study by Cohen and colleagues, for example, measured the impact of negative affect on symptoms of respiratory viral infection. Levels of state and trait negative affect were both associated with amount of subjective health complaints; however, only state negative affect was predictive of objective measures of illness (i.e., mucus weight). This study highlights the need to consider the stability of negative affect in conceptualizations of mood. As discussed below, decisions made about the stability of negative affect influence its measurement.

Measurement

There are numerous methods for measuring negative affect and mood. Among the single-item self-report measures of negative affect available, the use of a simple 0 to 7 point scale has often been used. The ends of such scales are often marked with not at all and extremely, and people simply rate where their current mood (state) and/or their general mood (trait) falls along the continuum. This measurement strategy has been used by proponents of both the specific affect and dimensional approach. Other single-item scales include bipolar scales marked with adjectives such as happy at one end of the scale and sad at the other end, and graphical approaches that use adjective labeled grids and circles based on a circumplex model of emotion. While single-item measures are available, by far the most commonly used approach to the self-report measurement of mood is the multi-item adjective checklist. This type of scale presents a participant with a number of adjectives

that are selected to represent either the specific or dimensional approach to mood. Each adjective is rated by the participants as to how well it represents either their current mood (state) or typical mood (trait). Linear combinations of adjectives are created (usually a simple summation of responses to selected groups of adjectives representing scales or dimensions) that index the level of mood. The psychometric properties (reliability, validity) of these questionnaires are excellent, and participants and patients have little difficulty completing them.

Examples of well-established mood-adjective checklists include the Nowlis Mood Adjective Checklist, which has a brief, 36-adjective version; the Profile of Mood States, which is composed of 65 adjectives and is rated for how the subjects felt over the past week; the Multiple Affect Adjective Checklist, which uses 132 adjectives to measure depression, anxiety, and hostility; and the Positive and Negative Affect Schedule, which taps the positive and negative dimensions of mood with 10 adjectives for each scale.

Observational methods for the measurement of mood are also available, although they are usually much more cumbersome and difficult to employ. Observational coding of facial expressions is a highly complex, labor-intensive, yet reliable method for measuring negative emotion. Theories of emotion involving facial feedback have been proposed, with some evidence supporting the utility of the observational measurement. While not often used in the stress literature, facial expression measurement of affect might be profitably employed in laboratory-based experiments.

In terms of negative affectivity, perhaps the most widely employed measure is the Neuroticism, Extra-version, and Openness to Experience Personality Inventory (NEO-PI). The NEO-PI is a self-report personality inventory that assesses negative affectivity and other dimensions of personality. In addition to the NEO-PI, several other measures are available that can be used to assess either state or trait negative affect, depending on the instructions given regarding the time frame of focus.

There is, then, an impressive assortment of measures for the investigator interested in measuring negative affect. Choice of an instrument for a given study will depend upon the nature of the task and the researcher's decision about the model of affect that best suits their needs.

See Also the Following Articles

Affective Disorders; Anger; Anxiety; Depression Models; Hostility; Stress Generation.

Further Reading

- Cohen, S., Doyle, W. J., Skoner, D. P., Fireman, P., Gwaltney, J. M. and Newsom, J. T. (1995). State and trait negative affect as predictors of objective and subjective symptoms of respiratory viral infections. *Journal of Personality and Social Psychology* 68, 159–169.
- Diener, E., Larsen, R. J., Levine, S. and Emmons, R. A. (1985). Intensity and frequency: dimensions underlying positive and negative affect. *Journal of Personality and Social Psychology* 48, 1253–1265.
- Leventhal, H. and Tomarken, A. J. (1986). Emotion: today's problems. *Annual Review of Psychology* 37, 565–610.
- Oatley, K. and Jenkins, J. M. (1992). Human emotions: function and dysfunction. *Annual Review of Psychology* 43, 55–85.
- Stone, A. A. (1995). Measurement of affective response. In: Cohen, S., Kessler, R. C. & Underwood, L. U. (eds.) *Measuring stress: a guide for health and social scientists*, pp. 148–171. New York: Oxford University Press.
- Watson, D. and Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin* 96, 465–490.
- Watson, D. and Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin* 98, 219–235.

Negative Life Events See: Community Studies; Life Events Scale; Life Events and Health.

Neighborhood Stress and Health

A V Diez Roux

University of Michigan, Ann Arbor, MI, USA

© 2007 Elsevier Inc. All rights reserved.

Neighborhoods and Health
Neighborhood Stressors and Their Measurement
Evidence Linking Neighborhood Stressors and Health

Neighborhoods and Health

Neighborhoods have recently received increasing attention as relevant to health. A large body of work has documented associations of neighborhood socioeconomic characteristics (or measures of neighborhood deprivation) with a variety of health-related outcomes, including health behaviors as well as physical and mental health outcomes. For example, neighborhood disadvantage has been associated with higher prevalence of smoking, unhealthy diets, and greater incidence and prevalence of cardiovascular disease. Although the strength of the associations is often small compared to those observed for individual-level socioeconomic characteristics, they are consistent across studies and tend to persist after statistical

controls for person-level measures of social position, suggesting an independent effect of neighborhood context on health.

The extent to which associations between neighborhood socioeconomic characteristics and health reported in observational studies reflect causal mechanisms remains a subject of research and debate. An important limitation of existing work pertains to the use of neighborhood deprivation or neighborhood disadvantage as a proxy for the specific neighborhood characteristics that may be relevant to the health outcome being studied. The use of aggregate measures of neighborhood disadvantage also raises questions regarding whether residual confounding by person-level socioeconomic position could explain some of the associations observed. An important consideration in interpreting results to date is that neighborhood constructs are grossly misspecified in existing research because of its reliance on very crude proxies for neighborhoods (usually administrative areas such as wards in the UK or census tracts in the United States) and limitations in the neighborhood-level measures used. Hence, neighborhood effects are likely to be underestimated compared to the effects of individual-level variables, for which measurement strategies are much better developed.

Several specific mechanisms through which neighborhood contexts may affect health have been proposed. Some of these mechanisms involve features of the physical environment of neighborhoods and others involve features of the social environment.

The physical environment includes not only traditional physical exposures such as environmental toxins and noise, but also features of the man-made environment such as street layout, transportation availability, land use, presence and design of public spaces, and other aspects of urban design. These man-made physical features of neighborhoods have recently been referred to as the built environment. The social environment includes social norms, social cohesion, and social capital and features such as violence or social disorder. The proximate individual-level pathways through which neighborhood physical and social environments may affect individual health include health behaviors and exposures to acute and chronic stressors. For example, it has been hypothesized that street connectivity and land use mix may promote walking in daily life and physically active lifestyles. It has also been proposed that certain features of neighborhoods may serve as chronic or acute stressors with consequent physiological responses that may impair health. An important focus of current research on neighborhoods and health is to empirically examine the presence and relative importance of these different pathways and the specific neighborhood characteristics linked to them.

Neighborhood Stressors and Their Measurement

Although exposure to neighborhood sources of acute and chronic stress is often hypothesized to play a role in the relationship between neighborhood context and health, empirical documentation of the health effects of neighborhood stressors remains limited. There is still scant information of what the most relevant stress-generating features of neighborhoods might be, or on the way in which these features should be measured in empirical investigations. The most common domains investigated as potential neighborhood stressors include neighborhood problems, neighborhood disorder, violence and safety, and physical ambient characteristics.

Neighborhood problems are generally assessed by asking residents to report on the extent to which they perceive that different issues are a problem in their neighborhood. Examples of the neighborhood domains assessed include violence, noise, traffic, litter, air quality, vandalism, drug use, and presence and quality of

resources and services. Responses are typically added to construct an index of neighborhood problems. Cumulative exposure to neighborhood problems is hypothesized to be a source of chronic stress.

Neighborhood disorder refers to conditions and activities in the neighborhood that indicate a breakdown of social order. Scales to measure neighborhood disorder typically assess neighborhood markers of social incivility, disruption, and physical decay. Some researchers distinguish physical and social disorder. Signs of physical disorder include the presence of abandoned buildings, noise, graffiti, vandalism, and disrepair. Signs of social disorder include the presence of crime, loitering, public drinking or drug use, conflicts, and indifference. It is hypothesized that these signs of disorder signify a breakdown of social control and result in the experience of a threatening environment for all residents, regardless of whether they themselves have been victimized. Long-term exposure to neighborhoods with high levels of social and physical disorder is hypothesized to result in chronic stress.

Violence and perceived safety are among the most common neighborhood sources of stress investigated in empirical studies. Measures of neighborhood violence are constructed using survey responses to neighborhood violence scales or using crime statistics obtained from governmental agencies. Perceived safety is also assessed through surveys of residents. It is hypothesized that exposure to violence or to an unsafe residential environment is a source of chronic stress.

Physical ambient characteristics such as noise, crowding, housing characteristics, and proximity to environmental toxins have also been hypothesized to be sources of stress that vary across neighborhoods. These ambient characteristics are usually measured using census data or by linking residence data to other data sources, such as noise measurements or proximity to airports or environmental toxins. In contrast to neighborhood problems, neighborhood disorder, and violence/safety, which are hypothesized to affect health through psychosocial mechanisms, the factors investigated under the general rubric of physical ambient characteristics may affect health through pathophysiological mechanisms that do not involve psychological stress. For example, crowding may affect the probability of contracting infectious diseases through increased disease transmission, and environmental toxins may have direct carcinogenic or cardiovascular effects. However, it has also been argued that the perception of the presence of these physical environmental characteristics may have psychosocial consequences analogous to those observed for other stressors.

Evidence Linking Neighborhood Stressors and Health

The evidence linking neighborhood stressors to health remains limited. Most research has focused on self-reported health outcomes and health behaviors. For example, perceived neighborhood disorder has been found to be associated with poor self-reported health, poor physical functioning, adverse mental health outcomes, the presence of chronic health conditions, and heavy drinking. Although it has long been hypothesized that neighborhood sources of stress may contribute to physical health outcomes such as blood pressure, very few studies have linked neighborhood stressors to objectively measured physical health outcomes. Two physical health outcomes examined in relation to potential neighborhood stressors are blood pressure (or blood pressure reactivity) and low birth weight.

Causal inference from existing studies of neighborhood stressors and health is rendered complex by their observational design and the possibility of confounding by individual-level factors associated with place of residence. Most observational studies have only limited ability to control for these individual-level factors. Another important set of confounders includes other neighborhood attributes such as features of the built environment, which often covary with the measures of neighborhood stressors used. Measures of the physical and built environment of neighborhoods are often part of the neighborhood stressors scales used, making it difficult to determine whether it is the physical feature itself through its effects on health behaviors or the psychosocial consequence of the factor that results in the health effects. There is also little consensus on the conceptualization or the measurement of the features of neighborhoods that may be sources of stress. Thus, the extent to which neighborhood contexts affect health through the stress pathway and the relative importance of this pathway compared to pathways involving environmental influences on health behaviors or direct effects of environmental exposures (such as toxins or allergens) remain largely undetermined.

See Also the Following Articles

Community Studies; Crowding Stress; Cushing's Syndrome, Neuropsychiatric Aspects; Environmental Factors; Health and Socioeconomic Status.

Further Reading

- Aneshensel, C. and Sucoff, C. (1996). The neighborhood context of adolescent mental health. *Journal of Health and Social Behavior* 37(4), 293–310.
- Diez Roux, A. V. (2001). Investigating area and neighborhood effects on health. *American Journal of Public Health* 91(11), 1783–1789.
- Diez Roux, A. V., SteinMerkin, S., Arnett, D., et al. (2001). Neighborhood of residence and incidence of coronary heart disease. *New England Journal of Medicine* 345, 99–106.
- Evans, G. W. (2001). Environmental stress and health. In: Baum, A., Revenson, T. A. & Singer, J. E. (eds.) *Handbook of health psychology*, pp. 365–385. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Ewart, C. and Suchday, S. (2002). Discovering how urban poverty and violence affect health: development and validation of a neighborhood stress index. *Health Psychology* 21, 254–262.
- Gee, G. and Payne-Sturges, D. C. (2004). Environmental health disparities: a framework integrating psychosocial and environmental concepts. *Environmental Health Perspectives* 112, 1645–1653.
- Harburg, E., Erfurt, J. C., Chape, C., et al. (1973). Socioecological stressor areas and black white blood pressure: Detroit. *Journal of Chronic Diseases* 26, 595–611.
- Hill, T. and Angel, R. (2005). Neighborhood disorder, psychological distress, and heavy drinking. *Social Science and Medicine* 61, 965–975.
- Latkin, C. A. and Curry, A. D. (2003). Stressful neighborhoods and depression: a prospective study of the impact of neighborhood disorder. *Journal of Health and Social Behavior* 44, 34–44.
- Morenoff, F. (2003). Neighborhood mechanisms and the spatial dynamics of low birth weight. *American Journal of Sociology* 108, 976–1017.
- Perkins, D. D. and Taylor, R. B. (1996). Ecological assessments of community disorder: their relationship to fear of crime and theoretical implications. *American Journal of Community Psychology* 24(1), 63–107.
- Ross, C. E. and Mirowsky, J. (2001). Neighborhood disadvantage, disorder, and health. *Journal of Health and Social Behavior* 42(3), 258–276.
- Steptoe, A. and Feldman, P. M. (2001). Neighborhood problems as sources of chronic stress: development of a measure of neighborhood problems and associations with socioeconomic status and health. *Annals of Behavioral Medicine* 23, 177–185.
- Wilson, K., Elliott, S., Law, M., et al. (2004). Linking perceptions of neighbourhood to health in Hamilton, Canada. *Journal of Epidemiology and Community Health* 58, 192–198.

Nelson's Syndrome

A Stathopoulou and K Dimitriou

General Hospital of Athens, Athens, Greece

G Kaltsas

University of Athens, Athens, Greece

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G Kaltsas and A Grossman, volume 3, pp 12–13, © 2000, Elsevier Inc.

Pathogenesis

Incidence

Pathophysiology – Molecular Pathology

Predictive Factors

Clinical Features

Diagnosis and Treatment

Glossary

Adrenocorticotropin (ACTH)

A 39-amino-acid peptide hormone produced from the pituitary, from a larger precursor molecule (pro-opiomelanocortin) that stimulates the secretion of glucocorticoids, mineralocorticoids, and androgenic steroids from the adrenal cortex.

Cushing's disease

The most common cause of Cushing's syndrome as a result of ACTH hypersecretion from a pituitary tumor (usually a microadenoma 3–10 mm in diameter). Chronic glucocorticoid excess that leads to skin, muscle, bone, neuropsychiatric, gonadal and metabolic abnormalities (glucose intolerance and hyperlipidemia), obesity, and hypertension.

Cushing's syndrome

Chronic glucocorticoid excess that leads to skin, muscle, bone, neuropsychiatric, gonadal and metabolic abnormalities (glucose intolerance and hyperlipidemia), obesity, and hypertension.

Gamma knife

A stereotactic radiosurgery device that allows well-defined, deep-seated brain tumors (<3 cm in diameter) to be treated in a single session under local anesthesia, using the overlapping of narrow collimated beams from 201 cobalt-60 sources.

Trans-sphenoidal surgery

Treatment of choice for pituitary tumors; the pituitary is approached from the nasal cavity through the sphenoid sinus, the anterior–inferior sella is removed, and the dura layer is incised. The adenoma is removed selectively while normal pituitary tissue is identified and preserved.

Nelson's syndrome defines patients with pituitary-dependent Cushing's syndrome (CS), i.e., Cushing's disease (CD), who have undergone bilateral adrenalectomy and have developed pituitary expansion with

increased plasma adrenocorticotropin (ACTH) levels and progressive pigmentation of the skin. Total adrenalectomy still offers the only definitive curative procedure to lower cortisol levels adequately in patients with persistent ACTH-dependent CS, in whom radiotherapy with or without medical adrenocortical enzyme-blocking therapy has been insufficient.

Pathogenesis

Plasma ACTH levels are either within the normal range or elevated modestly in patients with CD. After bilateral adrenalectomy, even during cortisol replacement therapy, plasma ACTH levels may rise as much as 20-fold, and this may be associated with tumor growth. Nelson's syndrome represents the clinical progression of a pre-existing ACTH-secreting adenoma following bilateral adrenalectomy and normalization of the hypercortisolism; the suppressive effect of excessive levels of cortisol is no longer present, ACTH secretion markedly increases, and the pituitary adenoma may thus progress. However, the exact underlying mechanism(s) leading to tumor development and excessive ACTH secretion is not well defined. Following corticotropin-releasing hormone (CRH) infusion, ACTH secretion is stimulated greater and for a longer period of time in patients with Nelson's syndrome than in those with CD. An increase in basal ACTH production and pulsatile secretion with significantly greater quantitative regularity of serial ACTH release patterns in Nelson's syndrome compared to untreated CD has been demonstrated.

Extremely high plasma ACTH levels and aggressive neoplastic growth might be explained by the lack of appropriate glucocorticoid negative feedback due to defective glucocorticoid signal transduction. At the transcriptional level, expression of pro-opiomelanocortin (POMC), the precursor molecule of ACTH, is under negative feedback control by glucocorticoids. Although mutations of the POMC promoter region that could lead to relative resistance to glucocorticoids have generally been excluded in patients with Nelson's syndrome, in one of four such tumors a mutation was found at the level of the glucocorticoid receptor, explaining the glucocorticoid resistance and probably contributing to the tumoral progression.

Incidence

Currently, bilateral adrenalectomy is rarely used to treat CD as primary surgery, with an initial cure rate of approximately 50% and an improvement rate of

around 80%. Although longer follow up after primary surgery presents an increase in recurrences, there still remain some patients with CD who will benefit from bilateral adrenalectomy. Furthermore, bilateral adrenalectomy is still used after the failure of repeated transsphenoidal surgery or pituitary radiotherapy to cure CD, in patients in whom the source of ACTH cannot be determined easily, and when the medical control of hypercortisolemia cannot be achieved adequately. Such patients are at risk of developing Nelson's syndrome, which ranges from 8% to 38% in adults, mainly in patients treated at an early age. Approximately 30% of patients develop classical Nelson's syndrome following bilateral adrenalectomy with progressive hyperpigmentation and an obvious ACTH-secreting tumor with a time interval between adrenalectomy and the diagnosis of Nelson's syndrome ranging from 0.5 to 24 years. Another 50% develop evidence of a microadenoma without marked progression, while around 20% never develop a progressive tumor. The reason for these differences in clinical behavior is not well defined and difficult to be predicted currently prior to adrenal surgery.

Pathophysiology – Molecular Pathology

Although several candidate genes that control the cell cycle, apoptosis, angiogenic, and growth factors as well as other signaling pathways have been studied, no clear information unraveling the pathophysiology of corticotroph adenomas and/or Nelson's tumors has been established. It is not yet clear whether corticotroph adenoma formation is the result of a unique molecular mechanism or a constellation of many different possibilities. However, glucocorticoids, and therefore their deprivation also, might exert variable effects on tumor growth.

Predictive Factors

Neither the duration of disease before adrenalectomy, the preadrenalectomy cortisol levels, and the suppressibility to dexamethasone administration nor the duration of postadrenalectomy follow up has been found to be an important predictor. High basal ACTH levels after adrenalectomy in excess of 600 ng l^{-1} , particularly if found less than 1 year postoperatively, were predictive of subsequent Nelson's syndrome in some, but not all, studies suggesting that no unique threshold value can be established. Although several studies have suggested that the preadrenalectomy age was the most important predictive factor for developing Nelson's syndrome, as the great majority of patients who developed Nelson's syndrome underwent an adrenalectomy before the age of 35 years,

this was not confirmed from subsequent studies. Pregnant women and children also seem to be at higher risk for the disease. Subnormal dose or noncontinuous glucocorticoid replacement therapy has also been thought to be associated with increased risk of development of the syndrome. Although a clustering of clinical, biochemical, and radiological factors can be used to try to predict the development of Nelson's syndrome there is currently no unique factor with a strong predictive value. It is hoped that a detailed molecular analysis between the slowly and rapidly growing corticotroph adenomas following adrenalectomy might provide some answers and can be used as a predictive marker.

Clinical Features

Pituitary tumors in patients with classic Nelson's syndrome are among the most aggressive and rapidly progressive tumors. All patients with Nelson's syndrome present with hyperpigmentation similar to that in Addison's disease and often an expanding intra- and/or extrasellar mass lesion. Symptoms are a consequence of extension of the tumor and compression of surrounding structures; headaches, visual field defects, and extraocular nerve palsies are among the most common clinical signs. Cases of diabetes insipidus and hypopituitarism are exceptional. Malignant transformations with central nervous system and distant metastases, as well as behavioral changes, have been described; interestingly, a number of ACTH-secreting pituitary carcinomas have developed in the context of Nelson's syndrome. Pituitary apoplexy and hydrocephalus may also complicate the course of these tumors. In addition, complications related to increased ACTH levels have been reported. Paratesticular and paraovarian tumors producing cortisol or androgens in an ACTH-dependent manner as well as testicular adrenal rests with hyperplasia have been described.

Diagnosis and Treatment

The diagnosis is suspected clinically when hyperpigmentation and signs of other mass effects on adjacent tissues develop in a patient with previous bilateral adrenalectomy and is confirmed by measuring plasma ACTH levels, which are extremely high, often ranging from 220 pmol l^{-1} (1000 ng l^{-1}) to 2200 pmol l^{-1} ($10\,000 \text{ ng l}^{-1}$). However, ACTH levels do not reflect accurately the size or aggressiveness of the tumor.

It is important to note the prevailing cortisol level when measuring ACTH, as patients with CD postadrenalectomy often have considerably elevated ACTH levels before their morning hydrocortisone

dose, whereas 1–2 h later the level may show a highly significant fall.

Thus, the diagnosis of Nelson's syndrome requires the presence of rising levels of ACTH in association with evidence of tumor growth and the presence of optimal hydrocortisone replacement therapy; pigmentation and high levels alone may simply indicate inadequate replacement therapy. The presence of tumor is confirmed by computerized tomography (CT) and magnetic resonance imaging (MRI) scans; in contrast to CD, these tumors are usually macroadenomas.

Nelson's tumors are recognized for their large size and aggressive and invasive growth. Once the diagnosis of Nelson's syndrome is made, pituitary surgery is the initial mode of treatment. Most neurosurgeons approach the tumor by the transsphenoidal route and resect as much as possible. Pituitary surgery of Nelson's macroadenomas should be performed before extrasellar expansion of the tumor occurs, as it is probably more successful when Nelson's adenomas are relatively small. Parasellar expansion, especially cavernous sinus invasion, by the pituitary tumor has been related to low surgical cure rate. Moreover, in large Nelson's tumors, spontaneous hemorrhage in the tumor can occur, which can lead to severe neurological complications. However, complete excision is not always possible and adjuvant treatment is usually required, especially if the tumor has not been irradiated previously.

The role of external beam radiation either before or postoperatively in preventing Nelson's syndrome in CD treated with bilateral adrenalectomy has not been defined clearly; adequate doses of pituitary radiotherapy, particularly immediately postoperatively, appear to decrease such an occurrence. Following the application of conventional radiotherapy, 50% of patients with Nelson's syndrome obtained normalization of plasma ACTH levels at a median of 9.5 years. However, no prospective studies comparing the incidence of Nelson's syndrome between patients who had been irradiated with either conventional or focal radiotherapy with patients who have not received such treatment are currently available.

The role of adjuvant focal beam radiotherapy in cases of patients who had previously received conventional pituitary radiotherapy, either with stereotactic multiaxial treatment (SMART) from a linear accelerator or the so-called gamma knife, is currently being investigated. A recent study failed to demonstrate any substantial response in tumor size in patients with Nelson's syndrome who received SMART.

The majority of these highly aggressive and infiltrating pituitary tumors in patients with Nelson's

syndrome frequently respond poorly to (repeated) surgical and radiotherapeutic procedures with progressive tumor growth and rising ACTH levels. An alternative approach for such patients is medical treatment to suppress ACTH hypersecretion; bromocriptine, cyproheptidine, and sodium valproate have all been used to transiently control ACTH secretion, but with minimal effect on pituitary tumor growth. Somatostatin analogs have also been used, obtaining some control of ACTH levels but not tumor size. The use of cytotoxic chemotherapy for highly aggressive tumors has also been advocated although this treatment modality is applied when no other therapeutic options are available. Recently it has been demonstrated that corticotroph adenomas express in higher density somatostatin receptor subtype 5 rather than subtype 2, which is the main target of currently available somatostatin analogs. It is believed that the newly developed somatostatin analogs that selectively bind to somatostatin receptor 5 will be more efficacious. Following recent *in vitro* data demonstrating that rosiglitazone (a peroxisomal proliferator-activated receptor (PPAR)- γ agonist) may lead to enhanced apoptosis and decreased cell division, PPAR- γ agonists show new potential for the prevention and treatment of Nelson's syndrome. Retinoids are also under investigation as it seems that they directly suppress ACTH secretion by glucocorticoid tumors.

Further Reading

- Andreassen, M. and Kristensen, L. O. (2005). Rosiglitazone for prevention or adjuvant treatment of Nelson's syndrome after bilateral adrenalectomy. *European Journal of Endocrinology* **153**, 503–505.
- Assie, G., Baharel, H., Bertherat, J., et al. (2004). The Nelson's syndrome ... revisited. *Pituitary* **7**, 209–215.
- Gaffey, T. A., Scheithauer, B. W., Lloyd, R. V., et al. (2002). Corticotroph carcinoma of the pituitary: a clinicopathological study. Report of four cases. *Journal of Neurosurgery* **96**, 352–360.
- Herman-Bonert, V. and Fagin, J. A. (1995). Molecular pathogenesis of pituitary tumors. *Baillieres Clinical Endocrinology and Metabolism* **9**, 203–223.
- Jenkins, P. J., Trainer, P. J., Plowman, P. N., et al. (1994). The long-term outcome after adrenalectomy and prophylactic pituitary radiotherapy in adrenocorticotropin-dependent Cushing's syndrome. *Journal of Clinical Endocrinology and Metabolism* **79**, 165–171.
- Karl, M., Von Wichert, G., Kempter, E., et al. (1996). Nelson's syndrome associated with somatic frame shift mutation in the glucocorticoid receptor gene. *Journal of Clinical Endocrinology and Metabolism* **81**, 124–129.
- Kasperlink-Zaluska, A. A. and Jeske, W. (2005). Letters to the editors. *Clinical Endocrinology* **63**, 232–237.

- Kasperlik-Zaluska, A. A., Nielubxicz, J., Wislawski, J., et al. (1983). Nelson's syndrome: incidence and prognosis. *Clinical Endocrinology* 19, 693–698.
- Kemink, L., Pieters, G., Hermus, A., et al. (1994). Patient's age is a simple predictive for the development of Nelson's syndrome after total adrenalectomy for Cushing's disease. *Journal of Clinical Endocrinology and Metabolism* 79, 887–889.
- Kemink, S. A., Grotenhuis, J. A., De Vries, J., et al. (2001). Management of Nelson's syndrome: Observations in fifteen patients. *Clinical Endocrinology* 54, 45–52.
- Kelly, P. A., Samandouras, G., Grossman, A. B., et al. (2002). Neurosurgical treatment of Nelson's syndrome. *Journal of Clinical Endocrinology and Metabolism* 87, 5465–5469.
- Nagesser, S. K., van Seters, A. P., Kievit, J., et al. (2000). Long-term results of total adrenalectomy for Cushing's disease. *World Journal of Surgery* 24, 108–113.
- Oldfield, E. H., Schulte, H. M., Chrousos, G. P., et al. (1986). Corticotropin-releasing hormone (CRH) stimulation in Nelson's syndrome: response of adrenocorticotropin secretion to pulse injection and continuous infusion of CRH. *Journal of Clinical Endocrinology and Metabolism* 62, 1020–1026.
- Pernicone, P. J., Scheithauer, B. W., Sebo, T. J., et al. (1997). Pituitary carcinoma: a clinicopathologic study of 15 cases. *Cancer* 79, 804–812.
- Van Aken, M. O., Pereira, A. M., van den Berg, G., et al. (2004). Profound amplification of secretory-burst mass and anomalous regularity of ACTH secretory process in patients with Nelson's syndrome compared with Cushing's disease. *Clinical Endocrinology* 60, 765–772.
- Van der Hoek, J., Lamberts, S. W. and Hofland, L. J. (2004). The role of somatostatin analogs in Cushing's disease. *Pituitary* 7, 257–264.

Neonatal Handling and Stress See: Stress Hyporesponsive Period; Eclampsia and preeclampsia; Fetal Stress; Pregnancy, Maternal and Perinatal Stress, Effects of; 11 β -Hydroxysteroid Dehydrogenases.

Neonatal Maternal Deprivation See: Stress Hyporesponsive Period; Eclampsia and preeclampsia; Fetal Stress; Pregnancy, Maternal and Perinatal Stress, Effects of; 11 β -Hydroxysteroid Dehydrogenases.

Nerve Growth Factors See: Neuroinflammation.

Nervous System See: Brain and Brain Regions; Hippocampal Neurons; Hippocampus, Overview; Neuroendocrine Systems.

Neural Stem Cells

**U S Sohur, J G Emsley, B D Mitchell and
J D Macklis**

Massachusetts General Hospital, Harvard Medical
School, and Harvard University, Boston, MA, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction

Definitions: Neural Stem Cells, Progenitors, and Precursors

Functional Adult Neurogenesis Occurs in Nonmammalian
Vertebrates

The Discovery of Constitutive Adult Mammalian
Neurogenesis

The Location of Neural Precursor Cells and Their
Isolation *in Vitro*

Neurogenic versus Nonneurogenic Regions in
the Adult Brain

Factors That Stimulate Neurogenesis

Induction of Neurogenesis in Nonneurogenic Regions
of the Adult Brain

Function of Adult Neurogenesis and Neural Precursor
Cells in the Adult Brain

Consequences of Disturbed Adult Neurogenesis and
the Reaction of Neurogenic Regions to Disease

Conclusions and Future Prospects

Glossary

Gliogenesis

The generation of neuroglia; glia were recognized about 150 years ago as connective elements in the nervous system. There are three main types of glial cells in the CNS: astrocytes, oligodendrocytes, and microglia. Astrocytes and oligodendrocytes both develop from neuroepithelial cells, while microglia develop from mesodermal hemopoietic precursors. Glial cells are estimated to occupy about half the brain and outnumber neurons by about 10 to 1. Glia are involved in most aspects of neural function. Thus, during development, radial glia guide neuronal migration. Microglia are immune cells and are involved in neuroinflammation. Oligodendrocytes myelinate axons. Astrocytes play a pivotal role in the blood-brain barrier.

Lineage-specific precursors or progenitors *Multipotent progenitors*

Cells with restriction to one distinct lineage (e.g., neuronal, astroglial, glial, oligodendroglial).

Proliferative cells of the adult brain with only limited self-renewal that can

differentiate into at least two different cell lineages.

The birth of new neurons.

A generic term encompassing both stem and progenitor cells.

Neurogenesis precursor cell

Stem cells

Stem cells of the mammalian CNS are (1) self-renewing, with the theoretically unlimited ability to produce progeny indistinguishable from themselves, (2) proliferative, continuing to undergo mitosis (though perhaps with quite long cell cycles), and (3) multipotent for the different neuroectodermal lineages of the CNS, including the multitude of different neuronal and glial subtypes.

Subgranular zone (SGZ) *Subventricular zone (SVZ)*

Zone of the dentate gyrus that is the origin of stem cells in the hippocampus. Sometimes also called the subependymal zone, this zone lies immediately adjacent to the cells (ependymal) that line the walls of the lateral cerebral ventricles. In the adult, this zone gives rise to neural precursors/stem cells, the progeny of which are immature neurons that normally migrate to the olfactory bulb

Introduction

Over most of the past century of modern neuroscience, it was thought that the adult mammalian brain was incapable of generating new neurons or having neurons added to its complex circuitry. However, recent findings show that neurogenesis, the birth of new neurons, constitutively occurs in two specific regions of the adult mammalian brain (olfactory bulb and hippocampal dentate gyrus) and that there are significant numbers of multipotent neural precursors, or stem cells, in many parts of the adult mammalian brain.

The rise of precursor cell biology has brought new life to neural transplantation and the consideration of cellular replacement strategies to treat diseases of the brain. The idea of making new neurons is appealing for neurodegenerative diseases or selective neuronal loss associated with chronic neurological or psychiatric disorders. One goal of neural precursor biology is to learn from this regionally limited, constitutive neurogenesis how to manipulate neural precursors toward therapeutic neuronal or glial repopulation. Elucidation of the relevant molecular controls might allow both control over transplanted precursor cells and the development of neuronal replacement therapies based on the recruitment of endogenous cells.

This article deals with adult neurogenesis, cellular repair of the adult mammalian central nervous system (CNS), and what is known about the location, behavior, and function of precursor cells in the adult brain. In the context of CNS regeneration, this information lies at the core of all attempts to guide neuronal or glial development from neural precursors in the adult CNS for therapeutic purposes. These topics are also important for at least two other reasons. First, adult neurogenesis and precursor cell function might play an important role in the function of the healthy and diseased brain; specifically, they may underlie particular aspects of neuronal plasticity in response to functional demands. Second, neurogenesis and precursor cell biology in the adult CNS also appear to recapitulate some of the molecular, cellular, and other requirements for neuronal development in the developing brain. Neural repair is of potential importance for neurodegenerative disorders such as amyotrophic lateral sclerosis (ALS), Parkinson's disease, and Huntington's disease, as well as acquired damage such as with spinal cord injury and stress-induced neurodegeneration.

Definitions: Neural Stem Cells, Progenitors, and Precursors

Rigorously defined, adult CNS stem cells exhibit three cardinal features: (1) they are self-renewing, with the theoretically unlimited ability to produce progeny indistinguishable from themselves; (2) they are proliferative, continuing to undergo mitosis (though perhaps with quite long cell cycles); and (3) they are multipotent for the different neuroectodermal lineages of the CNS, including the multitude of different neuronal and glial subtypes. Multipotent progenitors of the adult brain are proliferative cells with only limited self-renewal that can differentiate into at least two different cell lineages (multipotency). Lineage-specific precursors or progenitors are cells with restriction to one distinct lineage (e.g., neuronal, astroglial, glial, oligodendroglial). Together, CNS stem cells and all precursor/progenitor types are, broadly defined, precursors. It is preferred to use the term stem cell as rigorously defined previously, and to use precursor or progenitor for most forms of multipotent or lineage-restricted mitotic cells. This article uses precursor cell as a generic term encompassing both stem and progenitor cells.

Adult neurogenesis comprises the entire set of events of neuronal development, beginning with the division of a precursor cell and ending with the presence and survival of a mature, integrated, functioning new neuron. Neurogenesis is sometimes incorrectly used only in the sense of precursor cell proliferation,

but this definition clearly falls short. Precursor cell proliferation is not predictive of net neurogenesis. True neuronal integration depends on many complex variables and progressive events.

Even in the absence of a fully detailed molecular characterization, neural precursor cells of the adult brain clearly possess features that justify a view of these cells as extremely promising for the goals of CNS cellular repair. They are undifferentiated, often highly mobile cells, relatively resistant to hypoxia and other injury, proliferatively active, and able to produce mature neurons and glia. The most critical caveat, however, is that precursor cells in the adult mammalian brain are heterogeneous, and, while all of these criteria might apply in many cases, it is not clear how many of the sometimes site-specific and condition-specific observations in neural precursor biology can be broadly generalized, and how many reflect particulars of a certain brain region or a certain condition.

While the identity of true CNS stem cells is as yet unknown, the heterogeneity of neural precursor cells is also an important aspect of the promise of such cells. Their variable specialization and differentiation competence provides great promise for application in medicine. The idea that one cell fits all might turn out to be as unrealistic as it is unnecessary. Adult neural precursors with partial fate restriction may, in some cases, allow for more efficient production of desired cell types. The brain itself may provide the environmental determinants for the different forms and sequelae of adult neurogenesis and gliogenesis and may powerfully enable and direct specific lineage development.

Functional Adult Neurogenesis Occurs in Nonmammalian Vertebrates

Adult neurogenesis occurs in many nonmammalian vertebrates. The medial cerebral cortex of lizards, which resembles the mammalian dentate gyrus, undergoes postnatal neurogenesis and can regenerate in response to injury. Newts can regenerate their tails, limbs, jaws, and ocular tissues and the neurons that occupy these regions. Goldfish undergo retinal neurogenesis throughout life and, impressively, can regenerate surgically excised portions of their retina in adulthood. Although some nonmammalian animals can undergo quite dramatic regeneration of neural tissue, it is unclear how relevant this phenomenon is to mammals. Indeed, it has been proposed that selective evolutionary pressures have led mammals to lose such abilities.

Birds, whose brains are much closer to mammals in complexity, also undergo postnatal neurogenesis.

Lesioned postnatal avian retina undergoes some neurogenesis, with the new neurons most likely arising from Müller glia. In songbirds, new neurons are constantly added to the high vocal center, a portion of the brain necessary for the production of learned song, as well as to specific regions elsewhere in the brain (but not all neuronal populations). It is possible to manipulate experimentally the extent to which new neurons are produced in at least one songbird, the zebra finch, which does not normally seasonally replace high vocal center (HVC-RA) projection neurons in the song production network. Inducing cell death of HVC-RA neurons in zebra finches leads to deterioration in song. Neurogenesis increases following induced cell death, and birds variably recover their ability to produce song coincident with the formation of new projections from area HVC to area RA, suggesting that induced neuronal replacement can restore a learned behavior.

The Discovery of Constitutive Adult Mammalian Neurogenesis

Ramon y Cajal (1928, p. 750) has been widely quoted as writing that “in the adult centers the nerve paths are something fixed, ended and immutable. Everything may die, nothing may be regenerated.” The relative lack of recovery from CNS injury and neurodegenerative disease and the relatively subtle and extremely limited distribution of neurogenesis in the adult mammalian brain resulted in the entire field reaching the conclusion that neurogenesis does not occur in the adult mammalian brain. Joseph Altman was the first to use techniques sensitive enough to detect the ongoing cell division that occurs in adult brain. Using tritiated thymidine as a mitotic label, he published evidence that neurogenesis constitutively occurs in the hippocampus and olfactory bulb of the adult mammalian brain. These results were later replicated using tritiated thymidine labeling followed by electron microscopy. However, the absence of neuron-specific immunocytochemical markers at the time resulted in identification of putatively newborn neurons being made based on purely morphological criteria. These limitations led to a widespread lack of acceptance of these results and made research in the field difficult.

The field of adult neurogenesis was rekindled in 1992, when it was shown that precursor cells isolated from the forebrain can differentiate into neurons *in vitro*. These results, and technical advances such as the development of immunocytochemical reagents that could more easily and accurately identify the phenotype of various neural cells, led to an explosion of research in the field. Normally occurring neurogenesis

in the subventricular zone (SVZ)/olfactory bulb and the dentate gyrus subregion of the hippocampus have now been well characterized in the adult mammalian brain.

The Location of Neural Precursor Cells and Their Isolation *in Vitro*

If adult multipotent precursors were limited to the two neurogenic regions of the brain, the SVZ and dentate gyrus, it would severely limit the potential for neuronal replacement therapies based on *in situ* manipulation of endogenous precursors (Figure 1). However, *ex vivo* studies, in which precursors were isolated from the adult brain, provide evidence for cells with precursor cell properties in the adult brain. Such cells were first found embryonically, then in the adult hippocampus and the SVZ. Following these first reports, cells with similar properties have been isolated from many other adult mammalian brain regions as well. Neural precursor cells can be propagated in two main forms: in adherent cultures and under floating conditions, in which they aggregate to form heterogeneous ball-like structures, termed neurospheres. While it has become standard for some groups to use this so-called neurosphere-forming assay as a key criterion for neural stem cells, the ability to form a neurosphere, alone, is not a sufficiently reliable hallmark of stem cells and does not fully differentiate between various mitotic and precursor cell populations. Clonal analysis provides important additional information. By subcloning individual cells, one can test whether an individual cell from these spheres can again give rise to secondary spheres, which upon transfer into differentiation conditions can produce all neural lineages (see Figure 1).

In vitro, many such factors are likely to be present in the complex media used to culture neural precursors. Epidermal growth factor (EGF) and fibroblast growth factor 2 (FGF-2) are two key growth factors used in neural precursor cultures to maintain cells in their mitotic and undifferentiated state. The general effectiveness of both factors as mitogens for precursors has also been demonstrated *in vivo*, and they have been used more recently to enhance induced neurogenesis in the normally nonneurogenic CA1 region of the hippocampus after focal ischemia. However, their effects *in vivo* are not identical to those *in vitro*, and, consequently, their physiological relevance for normal neural precursor cell function of the adult brain is not known.

The efficacy of an individual growth factor *in vitro* does not prove that it plays a role in neurogenic permissiveness *in vivo*. Cell culture systems are highly artificial in many respects. *In vivo*, precursor cells

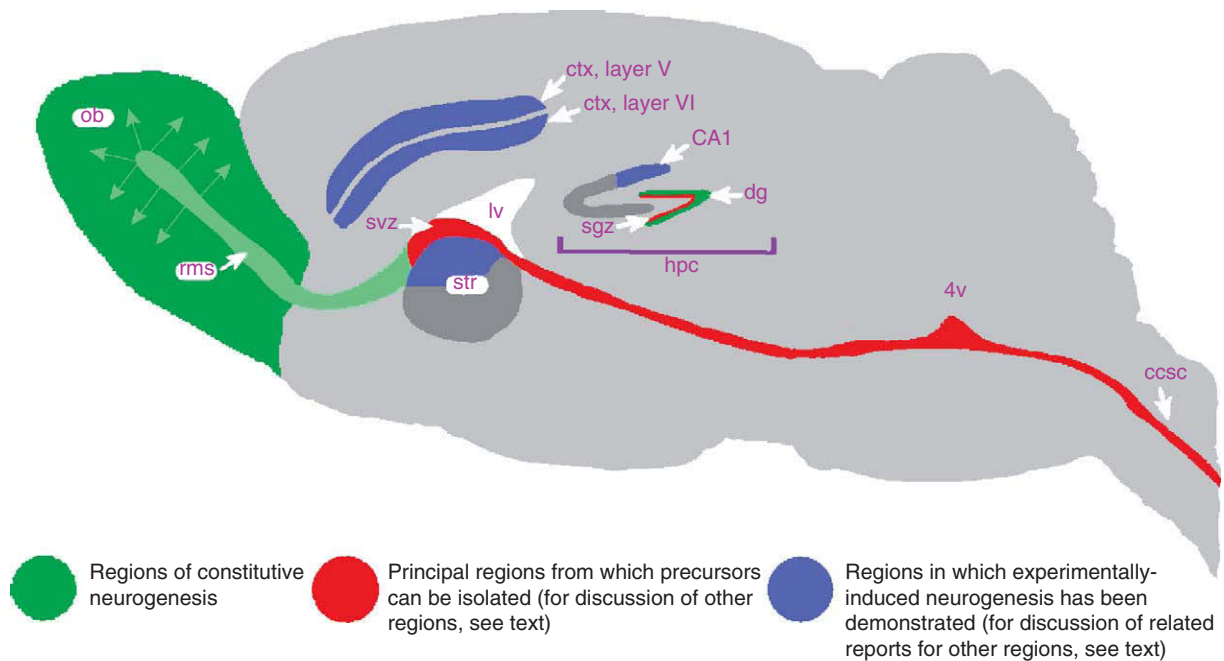


Figure 1 From Emsley, J. G., et al. (2005). *Progress in Neurobiology* 75, 321–341.

are never alone. Their relationship to a neurogenic microenvironment might be inseparable from their inherent properties. Regional differences *in vivo* reinforce the importance of the microenvironment for normal precursor cell function and neurogenesis. Nevertheless, *in vitro* studies are indispensable for many questions in neural precursor biology, including investigation of the transcriptional control of adult neurogenesis and neural differentiation.

Neurogenic versus Nonneurogenic Regions in the Adult Brain

In the adult mammalian brain, constitutive neurogenesis occurs only in the olfactory bulb and the dentate gyrus of the hippocampus. In the olfactory system, precursor cells reside in the anterior aspect of the SVZ in the walls of the lateral ventricles, migrate into the olfactory bulb, and differentiate into granule or periglomerular interneurons. In the dentate gyrus of the hippocampus, the precursor cell population is found in the subgranular zone (SGZ) of the dentate gyrus. Here, new hippocampal granule cell neurons are produced. These two brain regions are thus referred to as neurogenic regions of the brain. As is discussed in more detail later, reports of low-level and transient constitutive neurogenesis in the normal neocortex have been disputed, primarily for methodological reasons. Single reports have also appeared on constitutive adult neurogenesis in the amygdala, area CA1 of the hippocampus, the dorsal vagal complex of the adult brain stem, the spinal cord, and the substantia

nigra, although some of these results have been disputed and require further investigation.

Transplantation studies support the principle of defining neurogenic and nonneurogenic regions and provide evidence about the role of the microenvironment in influencing the potential of neural precursors. If precursor cells are transplanted into neurogenic regions, they can differentiate into neurons in a region-specific way. SVZ precursor cells generate hippocampal neurons when transplanted into the hippocampus, and SGZ precursor cells generate olfactory interneurons after transplantation into the rostral migratory stream of the olfactory ventricle (anterior extension of the lateral ventricle). When implanted outside the neurogenic regions, both types of precursor cells generate only glia. These data further support the interpretation that neurogenesis is dependent upon a permissive microenvironment, rather than on the regionally different properties of precursor cells. Local microenvironmental influences on the behavior of precursors and their ability to differentiate into neurons support the argument that it will be critical for the goal of neuronal repopulation to understand the cellular and molecular controls over cell type-specific precursor cell differentiation in the adult CNS.

It came as a surprise to many in the field that neurogenic regions are defined by the neurogenic permissiveness of the local microenvironment, rather than by the presence of neural precursor cells. Neural precursor cells have been found in a large number of brain regions, including white matter tracts. It is

possible that the entire brain contains neural precursor cells (albeit at very low density), just as it contains partially restricted glial precursors that act as progenitor cells. These widely distributed neural precursor cells do not seem to differ fundamentally from one another, although they are far from homogenous, with notably different growth kinetics and differentiation potential. Their function outside the classical neurogenic regions is not known, nor is their relation to the precursor cells of the neurogenic regions. For example, neural precursor cells from the spinal cord behave *in vitro* much like their counterparts from the SGZ and SVZ. That is, they are multipotent after implantation into the hippocampus, producing granule cell neurons, but generate only glial cells and no neurons *in situ* or when transplanted back to their original site in the spinal cord. *In vivo*, no adult neurogenesis can be found in the rodent spinal cord. In sum, the extent of neurogenesis in a region is a function of the local microenvironment on precursor cells with (varying) neurogenic potential.

There is, therefore, an important conceptual distinction between neurogenic and nonneurogenic regions, which might reflect complex molecular and functional states rather than a fixed cellular environment. Can nonneurogenic regions be molecularly modified to promote neurogenesis? Could this change in neurogenic permissiveness be induced under certain pathological conditions? There are strong natural precedents in other vertebrate systems, such as the avian telencephalon and the olfactory epithelium, for the concept that selective neuronal death can alter the differentiation of precursors, thereby inducing or increasing neurogenesis. While not yet proven, it is hypothesized that this change in neurogenesis occurs via alterations in the local molecular microenvironment that resemble conditions seen during nervous system development and in adult neurogenic regions.

Factors That Stimulate Neurogenesis

Seizure Activity

Seizure activity can increase hippocampal neurogenesis. However, it appears that seizure-induced neurogenesis may contribute to inappropriate plasticity, highlighting the fact that newly introduced neurons need to be appropriately integrated into the brain in order to have beneficial effects. Chemically or electrically induced seizures induce the proliferation of SGZ precursors, the majority of which differentiate into neurons in the granule cell layer. However, some newborn cells differentiate into granule cell neurons in ectopic locations in the hilus or molecular layers of the hippocampus and form aberrant connections to

the inner molecular layer of the dentate gyrus, in addition to the CA3 pyramidal cell region. It is hypothesized that these ectopic cells, with their aberrant connections, might contribute to the phenomenon of hippocampal kindling.

Behavioral Environment

Events occurring in the hippocampus dramatically demonstrate that behavior and environment can have a direct influence on the brain's microcircuitry. Animals living in an enriched environment containing toys and social stimulation have more newborn cells in their hippocampus than control mice living in standard cages. Experiments to further assess which aspects of the enriched environment contribute to increased neurogenesis revealed that part of the increase can be attributed simply to exercise via running. Associative learning tasks that involve the hippocampus also appear to increase neurogenesis. While this is still quite speculative, it is possible that the processes mediating these effects on neurogenesis may underlie some of the benefits that physical and social therapies provide for patients with stroke and brain injury. IGF-I appears to be one of the factors that might mediate behavioral influences on hippocampal neurogenesis.

Stress and Neurotransmitter Effects

Stress increases systemic adrenal corticosteroid levels and reduces hippocampal neurogenesis. Conversely, adrenalectomy, which virtually abolishes adrenal corticosteroids, increases hippocampal neurogenesis, suggesting that adrenal hormones at least partially mediate the effects of stress on hippocampal neurogenesis. Serotonin is a strong positive regulator of adult hippocampal neurogenesis. Drugs that increase serotonin action, such as antidepressants, increase neurogenesis. Together, these results demonstrate that adult neurogenesis can be modified by systemic signals, suggesting that modifying such systemic signals in addition to local signals may be useful in developing potential future neuronal replacement therapies involving manipulation of endogenous precursors.

Transcription Factors and Signaling Pathway Effects

Transcription factors and chromatin proteins, such as the methyl-CpG binding protein and HMGB1 chromatin protein, may regulate adult hippocampal neurogenesis, and Wnt signaling might also play a critical role in controlling the fate of hippocampal neural stem cells. The developmentally critical basic helix-loop-helix (bHLH) transcription factor *NeuroD* has

also been shown to be involved in adult hippocampal neurogenesis.

Depression and Neurogenesis

An interesting, but quite speculative and controversial, hypothesis concerning the role of hippocampal neurogenesis in human depression has been proposed and is outlined later in the article.

Induction of Neurogenesis in Nonneurogenic Regions of the Adult Brain

Induced neurogenesis in nonneurogenic regions can be envisioned in two basic forms: (1) local parenchymal precursor cells might be activated to generate new neurons upon a stimulus associated with a pathological event or with a set of directed molecular and/or genetic signals, or (2) precursor cells from normally neurogenic regions (e.g., the SVZ) might respond to either type of stimulus and migrate into the parenchyma. There are examples for both possibilities.

Several groups have recently demonstrated the induction of modest levels of neurogenesis in normally nonneurogenic regions in response to selective neuronal death or degeneration. As described later, Magavi et al. found that endogenous multipotent precursors normally located in the adult brain can be induced to differentiate into neurons in the normally nonneurogenic adult mammalian neocortex, a result recently extended to corticospinal motor neurons. More recently, other groups have reported similar and complementary results in a normally nonneurogenic region of the hippocampus and in the striatum, confirming and further supporting this direction of research. These reports of seemingly regenerative neurogenesis in the vicinity of targeted cortical neuron degeneration and cerebral ischemia suggest that the lack of normal neurogenic permissiveness in these regions is not unalterable. The controversial examples of reportedly extremely low-level constitutive neurogenesis outside the classical neurogenic regions might also support this possibility, potentially as leakage from normally tightly inhibited systems. Are precursor cells resident in nonneurogenic regions waiting for either instructive neurogenic signals or release from inhibition to awaken them to develop into new, function-restoring neurons?

Magavi et al in 2000 were the first to show that regenerative neurogenesis is possible in the neocortex. Induced synchronous apoptotic degeneration of corticothalamic neurons of layer VI of anterior cortex in adult mice led to the generation of new nerve cells in this region that projected to the thalamus. This work served as a proof of concept for the induction of adult mammalian neurogenesis in cortex

that has been extended to the striatum, hippocampus, and corticospinal system (see below). Thus, although the density of multipotent precursors in adult brain outside the anterior SVZ and the dentate gyrus is exceedingly low, these precursors under an appropriate *in vivo* environment (such as that generated by apoptotic neurons) can differentiate into astroglia, oligodendroglia, and neurons: the latter can receive afferents and extend axons to their appropriate targets. That is, cerebral cortex undergoing targeted apoptotic degeneration can direct multipotent precursors to integrate into adult cortical microcircuitry. The population specificity of neuronal death was found to be essential, allowing for maintenance of local circuitry and survival of surrounding cells that alter gene expression to direct precursors. Other groups have reported similar and complementary results in hippocampus and striatum, confirming and further supporting this direction of research. Nakatomi and colleagues, using an *in vivo* ischemia model, found that repopulation of neurons in the CA1 region of the hippocampus is possible following the elimination of CA1 neurons. Recently, Chen et al. demonstrated induction of low-level neurogenesis of corticospinal motor neurons, with progressive extension of axons into the cervical spinal cord over 3 to 4 months. The critical challenge will be to assess true functional integration of recruited adult-born neurons.

Taken together, these results demonstrate that endogenous neural precursors can be induced to differentiate into CNS neurons in a region-specific manner and at least partially replace neuronal populations that do not normally undergo neurogenesis. It appears that these results may be generalized to at least several classes of projection neurons, including the clinically relevant striatal medium spiny neurons and corticospinal motor neurons. Microenvironmental factors appear capable of supporting and instructing the neuronal differentiation and axon extension of adult-born neurons derived from endogenous precursors.

Function of Adult Neurogenesis and Neural Precursor Cells in the Adult Brain

Given the restricted regional occurrence of constitutive neurogenesis in the adult mammalian brain, it is reasonable to hypothesize that there are specific contributions of adult neurogenesis to the functions of the hippocampus and the olfactory system. Possibly, these functions differ fundamentally from functions found elsewhere in the brain.

For the most part, the adult brain seems to do just fine without the ability to generate or regenerate

neurons. Obviously, this characteristic of the mammalian CNS presents as a disadvantage in cases of neuronal loss, which cannot be replaced, but one theory is that this is the price paid during evolution for increasingly powerful and complex computing power. This reasoning posits that the cardinal requirement for stability in the system precludes the capability for neuronal replacement. Despite its appeal and clarity, this theory has for years been at odds with the fact that, even without considering actual cellular replacement or addition via neurogenesis, there is amazing structural and synaptic plasticity in the brain. Nonetheless, the fact remains that there is normally no (or extremely little) neurogenesis outside the two specific neurogenic regions of the adult brain. This striking restriction of neurogenesis to these two regions suggests that the question should be asked differently: What are the special functions of the dentate gyrus and olfactory bulb that require maintenance adult neurogenesis?

In simple terms, the hippocampus and olfactory system would appear to have little in common functionally. Whereas the hippocampus is one of the best-studied areas of the brain and has been named the gateway to memory because of its important and indispensable role in learning and memory, much less is known about the olfactory system. However, considering the evolutionary importance of olfactory perceptual learning, and the similar output modulatory roles of dentate gyrus granule neurons and olfactory bulb granule neurons that have been postulated, the functions of neurogenesis in each system might indeed be more closely related.

A gatekeeper theory has been proposed to help explain the potential function of adult-born dentate gyrus granule neurons. This theory hypothesizes that new granule cells allow for optimization of the mossy fiber connections between the dentate gyrus and CA3 in accordance with the level of complexity and novelty frequently encountered by an individual. Because adult neurogenesis occurs at a bottleneck connection, even a few new neurons could have a large impact. Selectively eliminating the adult-born neurons would be the method of choice to test this idea. However, both irradiation and treatments with cytostatic drugs as a means of killing dividing precursor cells are less selective than would be desired, with many structural and potentially functional side effects. Therefore, despite being highly intriguing, results of such studies have remained ambiguous and difficult to interpret.

Similarly, while the precise function of new neurons in the olfactory bulb is not yet known, recent evidence from Magavi, Macklis, and colleagues has identified an entirely novel function of adult-born olfactory

bulb granule neurons – they are especially responsive to novel odorants and undergo experience-dependent modification in response to familiarization with novel odorants presented as the neurons are integrating into functional circuitry. Selective circuit integration of activated neurons might underlie at least some of this change in population response. These results indicate that adult-born granule neurons provide a novel form of synaptic plasticity via insertion of entirely new cellular units into existing functional circuitry.

Consequences of Disturbed Adult Neurogenesis and the Reaction of Neurogenic Regions to Disease

Because so little is known about the physiological function of adult neurogenesis, it has been difficult to evaluate the possible consequences of disturbed neurogenesis. Nevertheless, altered adult hippocampal neurogenesis has been proposed to play a role in a number of disorders. The most robust data that support this role come from studies of temporal lobe epilepsy.

Experimental seizures are a robust inducer of cell proliferation and net neurogenesis. Because of this, dysregulated adult neurogenesis might provide an explanation for the development of the granule cell dispersion found in many cases of temporal lobe epilepsy. There is some evidence that ectopic neurons can form in seizure-induced neurogenesis and that they might contribute to seizure activity. In addition, there have been recent reports of increased neurogenesis in Alzheimer's disease, as well as in Huntington's disease. Further, experimental depletion of dopamine in a rodent model of Parkinson's disease has been reported to decrease precursor cell proliferation in both the SVZ and dentate gyrus. The reproducibility and functional significance of these phenomena in neurological disorders have yet to be evaluated.

Several experimental data suggest another quite speculative hypothesis, that a failure of adult hippocampal neurogenesis might underlie major depression. Patients with long-lasting depression have global hippocampal atrophy. Consistent with this hypothesis, stress-related glucocorticoids are associated with a decrease in neurogenesis, and increased serotonin levels are associated with an increase in neurogenesis. If adult neurogenesis is blocked by irradiation, the behavioral effects of fluoxetine treatment (a serotonin reuptake inhibitor) have been found to be abolished. However, the hippocampus is generally thought to be involved in memory consolidation and less involved in the generation of mood, suggesting that altered hippocampal neurogenesis may be secondary

to, rather than causative of, depression. Further, depression is not thought to be primarily a hippocampal disorder, and the presumed link to adult neurogenesis would need to depend strongly on the functional contribution of new dentate gyrus neurons to hippocampal function. The fact that neurogenesis occurs only in the dentate gyrus, while hippocampal atrophy in depression and positive effects of therapies are widespread throughout the hippocampus, argues that these phenomena are not directly linked. Thus a possible link between depression and neurogenesis remains speculative and in need of further rigorous analysis.

Conclusions and Future Prospects

Neurogenesis and neural precursor/stem cell biology suggest that the generation, insertion, and functional integration of new neurons is possible in the adult mammalian brain. From the point of view of fundamental neuroscience, it is important to learn about the normal role of precursor cells and the normal function of neurogenesis in the neurogenic regions of the adult CNS. This information might have immense implications for our understanding of brain function in development, normal adulthood, and disease states. Circuit plasticity at the level of individual cells, neuronal or glial, adds a new layer of complexity to our understanding of how brain structure and function interact. Better understanding of these issues may enable the prevention of disease and dysfunction, beyond the established goals of trying to repair neurons lost as a consequence of neurodegeneration and CNS damage. A better understanding of the cellular and molecular controls over differentiation of neural precursor cells along specific cellular lineages during development and in the adult CNS will be critical for the development of cellular therapeutic approaches for repopulating damaged or diseased areas of the nervous system. The future prospect of directing the development and integration of precursors in the adult mammalian brain toward the replacement of lost neurons or glia is exciting, and several recent lines of work provide remarkable progress toward this aim.

Acknowledgments

This article was updated and modified from a review article for a different readership (Sohur, U. S., Emsley, J. G., et al. (2006) Adult neurogenesis and cellular brain repair with neural progenitors, precursors, and stem cells. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 361(1473), 1477–1497).

See Also the Following Articles

Neurogenesis; Hippocampal Neurons; Hippocampus, Overview.

Further Reading

- Altman, J. (1969). Autoradiographic and histological studies of postnatal neurogenesis. IV. Cell proliferation and migration in the anterior forebrain, with special reference to persisting neurogenesis in the olfactory bulb. *Journal of Comparative Neurology* 137, 433–457.
- Altman, J. and Das, G. D. (1965). Autoradiographic and histological evidence of postnatal hippocampal neurogenesis in rats. *Journal of Comparative Neurology* 124, 319–335.
- Arlotta, P., Molyneaux, B. J., Chen, J., Inoue, J., Kominami, R. and Macklis, J. D. (2005). Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo. *Neuron* 20, 183–185.
- Chen, J., Magavi, S. S. and Macklis, J. D. (2004). Neurogenesis of corticospinal motor neurons extending spinal projections in adult mice. *Proceedings of the National Academy of Sciences USA* 101, 16357–16362.
- Gage, F. H. (2000). Mammalian neural stem cells. *Science* 287, 1433–1438.
- Jacobs, B. L., Praag, H. and Gage, F. H. (2000). Adult brain neurogenesis and psychiatry: a novel theory of depression. *Molecular Psychiatry* 5, 262–269.
- Kempermann, G., Wiskott, L. and Gage, F. H. (2004). Functional significance of adult neurogenesis. *Current Opinions in Neurobiology* 14, 186–191.
- Koketsu, D., Mikami, A., Miyamoto, Y. and Hisatsune, T. (2003). Nonrenewal of neurons in the cerebral neocortex of adult macaque monkeys. *Journal of Neuroscience* 23, 937–942.
- Magavi, S. S., Leavitt, B. R. and Macklis, J. D. (2000). Induction of neurogenesis in the neocortex of adult mice. *Nature* 405, 951–955.
- Magavi, S. S. P., Mitchell, B. D., Szentirmai, O., Carter, B. S. and Macklis, J. D. (2005). Adult-born and pre-existing olfactory granule neurons undergo distinct experience-dependent modifications of their olfactory responses in vivo. *Journal of Neuroscience* 25, 10729–10739.
- Molyneaux, B. J., Arlotta, P., Hirata, T., Hibi, M. and Macklis, J. D. (2005). Fezl is required for the birth and specification of corticospinal motor neurons. *Neuron* 47, 817–831.
- Nakatomi, H., Kuriu, T. and Okabe, S. (2002). Regeneration of hippocampal pyramidal neurons after ischemic brain injury by recruitment of endogenous neural progenitors. *Cell* 110, 429–441.
- Parent, J. M., Vexler, Z. S., Gong, C., Derugin, N. and Ferriero, D. M. (2002). Rat forebrain neurogenesis and striatal neuron replacement after focal stroke. *Annals of Neurology* 52, 802–813.
- Rakic, P. (2002). Neurogenesis in adult primate neocortex: an evaluation of the evidence. *Nature Reviews Neuroscience* 3, 65–71.

- Ramon y Cajal, S. (1928). *Degeneration and regeneration of the nervous system*. Volume 2. Haffner Publishing Co., New York. p. 750.
- Reynolds, B. A. and Weiss, S. (1992). Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. *Science* **255**, 1707–1710.
- Richards, L. J., Kilpatrick, T. J. and Bartlett, P. F. (1992). De novo generation of neuronal cells from the adult mouse brain. *Proceedings of the National Academy of Sciences USA* **89**, 8591–8595.
- Scharff, C., Kirn, J. R., Grossman, M., Macklis, J. D. and Nottebohm, F. (2000). Targeted neuronal death affects neuronal replacement and vocal behavior in adult songbirds. *Neuron* **25**, 481–492.
- Seaberg, R. M. and van der Kooy, D. (2003). Stem and progenitor cells: the premature desertion of rigorous definitions. *Trends in Neuroscience* **26**, 125–131.

Neurodegenerative Disorders

M F Mendez and A M McMurtray

West Los Angeles Veteran's Affairs Medical Center and University of California, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Overview of Neurodegenerative Disorders

Neurodegenerative Dementias: Alzheimer's Disease

Neurodegenerative Movement Disorders: Parkinson's Disease

Conclusions

Glossary

<i>Alzheimer's disease</i>	A progressive neurodegenerative dementia affecting memory and other cognitive functions.
<i>Hippocampus</i>	Area of the brain concerned with encoding and storage of new declarative memory.
<i>Neuron</i>	Single cell functional unit of the nervous system.
<i>Parkinson's disease</i>	Progressive neurodegenerative disorder affecting movement, balance, and postural stability.
<i>Substantia nigra</i>	Area of the brain principally affected in Parkinson's disease containing a high concentration of dopaminergic neurons.

Overview of Neurodegenerative Disorders

Neurodegenerative disorders are characterized by a gradual loss of neurons, often leading to death. The term encompasses a broad range of clinical diseases, including progressive dementing conditions, of which the most common are Alzheimer's disease (AD), movement disorders, exemplified by Parkinson's disease (PD), and a range of other neurological disorders. Neurodegenerative disorders have in common a

gradual loss of neurons and synaptic connections, usually occurring in later life. Specific diseases are distinguished by the presence of characteristic symptoms that depend on the location at which neuronal loss is occurring in the brain. Typically, the degree of neuronal loss correlates directly with the appearance and progression of the clinical symptoms. In AD, neuronal loss occurs early in the hippocampus, a brain region concerned with declarative episodic memory. In PD, the characteristic clinical triad of tremor, bradykinesia, and postural instability only become evident after 70–80% of dopaminergic neurons in the substantia nigra are lost.

The gradually progressive nature of these disorders, often combined with a long preclinical course before the development of symptoms, allows ample time and opportunity for environmental factors such as stress to contribute to disease development and progression. These features, however, make detection of environmental factors with disease-modifying effects by traditional epidemiological methods difficult. Prospective studies attempting to identify associations between stress beginning early in life and a clinical syndrome many years later in middle or even late life are hindered by the long length of time between exposure and disease onset, while retrospective studies are hampered by trying to quantitate environmental stress and account for the large fluctuations that occur over the course of a lifetime. Additionally, definitive diagnosis of most neurodegenerative disorders requires inspection of brain pathology after death, further complicating study of patients during life. Consequently, most knowledge concerning the relationship of stress to neurodegenerative disorders originates either from studies of animal models or from changes in stress-related hormones in those with neurodegenerative disorders compared to those without.

Despite the limitations, the scientific literature substantially supports an association of psychological

and physiological stress with the risk of neurodegenerative disease. This article describes the clinical and neuropathological findings of AD and PD, as illustrative of the most common neurodegenerative disorders. It describes the effects of stress on animal models of neurodegenerative disorders and evidence for alterations in stress hormones in the setting of neurodegenerative disease.

Neurodegenerative Dementias: Alzheimer's Disease

AD is the most common neurodegenerative disorder and a major health-care problem that will continue to increase in importance as the U.S. population ages. Currently, more than 4 million Americans have a diagnosis of AD, resulting in enormous economic and social costs for their care, with a total in the United States alone approaching \$100 billion per year.

AD occurs primarily but not exclusively in the elderly, with onset in most patients occurring after age 65. More rarely, onset may occur as early as the fourth decade of life, particularly in those with a familial form of the disease. Overall, prevalence for AD among the elderly (age 65 or greater) is about 6%, doubling approximately every 5 years after age 60, affecting between 26 and 45% of those age 85 or older. The annual incidence of new cases of AD per 100 000 people in the United States also rises with advancing age, one reason why investigators predict that AD will affect an increasingly large number of persons in the coming years, as the average age of the U.S. population increases. Some have predicted that in the United States the number of people with this disease may reach 14 million by the year 2050.

In the absence of a definitive clinical test or biomarker, the clinical diagnosis of AD depends on clinical criteria. The diagnosis of clinically probable AD by National Institute of Neurologic and Communicative Disorders and Stroke and the AD and Related Disorders Association Work Group (NINCDS-ADRDA) criteria requires impairment in at least two cognitive domains, progressive deterioration, absence of delirium, onset between 40 and 90 years of age, and lack of evidence for other dementing diseases after a standard work-up. Besides memory, other cognitive domains that may be affected include language, visuospatial perception or constructions, calculations, praxis, and executive functions. Patients meeting these criteria typically display progressive impairment in activities of daily living such as driving, buying groceries, preparing meals, doing laundry, and basic functions such as walking safely and maintaining personal hygiene. The diagnosis of clinically probable AD using NINCDS-ADRDA

criteria reportedly results in correct identification at autopsy in 85 in 95% of cases. Like other neurodegenerative disorders, gradual deterioration until death is typical, with a mean survival after symptom onset of 10.3 years.

The definitive diagnosis of AD requires clinico-pathological correlation. One hallmark of the disorder is the presence of intracellular neurofibrillary tangles (NFTs) containing the microtubule associated tau protein in an abnormal phosphorylated state. Another hallmark of AD is the presence of extracellular senile or neuritic plaques greater than expected for the patient's age. The central core of these plaques is composed of β -amyloid peptides such as A β 42. Around this amyloid core are degenerating dendrites and axons, some of which contain paired helical filaments like those found in NFTs. In addition to NFTs and neuritic plaques, the neuropathology of AD, like all neurodegenerative disorders, includes a progressive loss of neurons and their synaptic connections.

Both familial and environmental risk factors exist for development of AD. First, the development of AD at an early age (younger than age 65) often occurs as a result of autosomally dominantly inherited mutations in genes such as presenilin 1 and 2 and the amyloid precursor protein. Second, presence of an apolipoprotein E genotype with at least one ϵ 4 allele results in an earlier age of onset of the disorder. Apolipoproteins are cholesterol-bearing proteins that mediate neuronal protection and repair and may participate in A β 42 deposition. Third, women appear to be at slightly greater risk for AD, even when accounting for their greater longevity over men. Fourth, in addition to age, risk for AD increases with traumatic head injury, small head size, and low education, presumably on the basis of decreased neuronal reserve. Finally, later age of symptom onset, prominent psychiatric symptoms, and comorbidity such as cerebrovascular disease, heart failure, a history of heavy alcohol use, excessive weight loss, and institutionalization all may have a negative impact on clinical progression and eventually survival. In contrast, environmental and familial risk factors such as ethnic background and marital status are not proven to affect survival or rate of disease progression.

More recent evidence implicates psychological stress as an additional risk factor for the development of AD. Several studies show that patients with a high susceptibility to distress are more prone to AD in later life. When patients are followed to autopsy, those with a high susceptibility to distress as indicated by the personality trait of neuroticism are twice as likely as those without this susceptibility to have the NFTs and neuritic plaques of AD. Those with high distress proneness have a greater than 10-fold increase in

episodic memory decline compared to those with low distress proneness. Other studies indicate that everyday hassles and challenging life events may exacerbate age-related declines on episodic memory tests. These findings implicate accelerated damage to the hippocampi as the mediator of the effects of stress on memory.

Much data indicate that the hippocampal region is involved in the response to stress. Animal and human studies indicate memory deficits associated with hippocampal damage after chronic stressful experiences. Corticosteroid receptors are abundant in the hippocampus, and high, stress-induced concentrations of glucocorticoids become toxic to the hippocampal neurons. Studies have reported high concentrations of cortisol in association with hippocampal atrophy in patients with dementia, as well as in posttraumatic stress disorder and major depression.

Many studies suggest abnormal function of the hypothalamic-pituitary-adrenal (HPA) axis in patients with AD. There are reports of increased serum cortisol levels and reduced feedback inhibition of cortisol secretion in patients with AD. Corticosteroid hypersecretion may downregulate the hippocampal corticosteroid receptors, which in turn dampens the feedback inhibition of the HPA, leading to further hypersecretion and eventual hippocampal neuronal loss. Other studies have focused on abnormal neuroendocrine responses to challenge with a chemical agent. Several studies report that AD patients demonstrate elevated cortisol levels relative to nondemented controls after challenge by a dexamethasone suppression test. In these studies, cortisol levels after administration of dexamethasone correlated directly with dementia severity. Another study compared cortisol and prolactin responses to administration of naltrexone, an opiate antagonist, in AD patients and normal elderly controls. In this study, normal control patients demonstrated a significantly greater increase in cortisol levels after a single dose of naltrexone compared to the AD patients, with no difference in prolactin response between the groups. Finally, cortisol levels have been shown to directly correlate with severity of cognitive decline and degree of hippocampal atrophy in AD patients.

One further mechanism for stress-induced injury in AD is vascular. The neuropathology of AD and the cerebrovascular changes may be synergistic. Higher stress-induced systolic and diastolic blood pressure (BP) reactivity, including variability in BP on 24-h ambulatory monitoring, has been associated with decreased performance in verbal memory and executive tests. The relation of BP reactivity to decreased cognitive function may be mediated by

silent cerebrovascular changes from cerebral hypoperfusion. Repeated episodes of stress-induced BP reactivity can result in white matter changes or other silent cerebrovascular lesions on neuroimaging. Finally, life may also produce a hypercoagulable state that could contribute to cerebrovascular disease.

Neurodegenerative Movement Disorders: Parkinson's Disease

PD is an extrapyramidal movement disorder of unknown etiology, with onset between the ages of 50 and 65 years, and is more common in males than females with a ratio of 3:2. PD is fairly common, with a prevalence of approximately 100 to 180 per 100 000 population. Disease course is shorter on average than AD, with an 8-year mean duration of illness. Patients with PD have a two- to fivefold increased mortality compared to nondemented people in the same community. Risk of mortality directly relates to symptom severity, with death typically occurring due to aspiration pneumonia, urinary tract infection, or unrelated conditions of the elderly.

Clinically, PD has five primary characteristics: resting tremors, rigidity, bradykinesia, postural instability, and responsiveness to treatment with dopaminergic agents. In any particular patient some components may be evident and others not, and components are commonly asymmetric at onset. Of the primary clinical characteristics, presence of tremor is most suggestive of PD as opposed to other parkinsonian disorders. The tremor is classically present at rest, occurring at approximately four to eight cycles per second, and may develop in one upper extremity before spreading to involve all four limbs, face, and tongue. Bradykinesia contributes to the expressionless or masked face, sialorrhea, and micrographia. The gait is festinating, with abnormal righting reflexes and a slightly flexed posture. Eye movements may have subtle abnormalities in voluntary gaze, pursuit, and convergence. Many patients develop autonomic symptoms including postural hypotension, impotence, atony of the large bowel with constipation, and esophageal spasm.

Characteristic pathological changes of idiopathic PD typically occur in the pars compacta of the substantia nigra. Classically, there is marked depigmentation and neuronal loss, with presence of Lewy bodies in this region. As a consequence of these changes, dopaminergic neurons in the substantia nigra pars compacta are lost, affecting projections to the corpus striatum. Less pronounced changes are evident in a variety of other brain stem and diencephalic nuclei, including the locus ceruleus, dorsal

vagal nucleus, and sympathetic ganglia. Although depletion of dopamine is the classic neurotransmitter deficit in PD, levels of norepinephrine, glutamate decarboxylase, GABA, methionine-enkephalin cholestokinin, and serotonin are also decreased in the substantia nigra of PD patients. No evidence supporting or suggesting a link between stress and neurotransmitter levels has been reported in PD.

Lewy bodies are characteristic of PD. Certain sites show a predilection for development of Lewy bodies, including the entorhinal region, the CA2 sector of the hippocampal formation, limbic nuclei of the thalamus, anterior cingulate, agranular insular cortex, amygdale, ventromedial divisions both of the basal and accessory basal nuclei, and central nucleus. Lewy bodies are composed of aggregates of α -synuclein, suggesting a possible central role for this protein in the pathogenesis of PD. Further evidence in the form of dopaminergic loss and inclusion body formation in α -synuclein mice reinforces a causal role for α -synuclein in this disorder. The proposed pathological mechanism starts with a reduction in the solubility of α -synuclein, resulting in the formation of filaments, which then aggregate into the cytoplasmic inclusions called Lewy bodies that contribute to the dysfunction or death of the glial cells and neurons in which they form in PD patients. However, no evidence for a relationship between degree of pathological changes or presence of Lewy bodies and stress has been reported in PD.

Environmental risk factors for PD include age, family history, and rural living possibly related to pesticide exposure. A genetic component to risk for development of PD is suggested by some studies reporting an increased prevalence rate of PD in relatives of PD patients compared to the general population. This has not been confirmed in twin studies, which failed to demonstrate increased concordance among identical twin pairs. Exposure to environmental toxins may be a rare cause of PD, such as from exposures to pesticides or herbicides and well water consumption. In contrast to AD, most studies do not find that APOE e4 is a risk factor for PD or PD dementia; however, PD patients show a trend toward slightly higher APOE e4 allele frequencies (12.5 to 17%) compared to normals (10 to 12.5%). One study showed that the APOE e4 allele frequency was significantly higher for PD dementia compared to PD and may affect an earlier age of onset for PD. Additionally, while female gender is associated with an increase risk of developing AD, male gender is associated with increased risk of PD.

Stress is probably an additional risk factor for PD. When compared to levels in unaffected spouses and

siblings, cortisol levels in untreated patients with idiopathic PD are elevated. Other neuroendocrine abnormalities such as elevated levels of prolactin and decreased levels of growth hormone and adrenocorticotrophic hormone have been demonstrated in patients with PD compared to normal controls. The clinical understanding of HPA axis dysfunction in PD, however, is often complicated by medical treatment, since selegiline, a medication used in treating patients with PD, may lower cortisol levels. Investigators have studied neuroendocrine abnormalities in a dog model of PD, using 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)-treated animals. After MPTP treatment, which damages neurons in the substantia nigra, plasma ACTH and cortisol are elevated in the treated dogs compared to untreated controls. In these animals, elevated cortisol levels result in excessive stimulation of neuronal glucocorticoid receptors, contributing to excitotoxic cell death. Excitatory stimulation of postsynaptic glutamate (NMDA) receptors in nigrostriatal neurons increases intracellular concentrations of calcium, possibly leading to disruption of the cell cytoskeleton or activation of caspase pathways leading to programmed cell death. Increased oxidative stress at the cellular level is also important in the pathogenesis of PD. Oxidative stress, through the depletion of glutathione and other free radical detoxificant enzymes, impaired mitochondrial complex I activity, or enhanced iron-mediated free radical production may contribute directly to neuronal loss in PD.

Conclusions

In summary, stress and its physiological effects may play a role in the progressive loss of neurons that occurs in neurodegenerative disorders. Increased cortisol levels due to both acute and chronic stress have demonstrated associations with hippocampal degeneration and memory impairment. This may be particularly relevant in the study of neuronal loss in AD, a disorder in which hippocampal degeneration occurs early in the disease course. Stress-induced alterations in BP and consequent vascular injury may further predispose to AD. In PD, elevated cortisol levels result from abnormalities of the function of the HPA axis, possibly contributing to loss of dopaminergic neurons in the substantia nigra by a variety of mechanisms, including excitotoxic cell death, increased intracellular calcium levels, and increased oxidative stress. Much more research is vitally needed in order to more clearly characterize and understand the contributions of stress to AD, PD, and the many other neurodegenerative disorders.

See Also the Following Articles

Aging and Stress, Biology of; Alzheimer's Disease; Behavior, Overview; Brain and Brain Regions; Glucocorticoids – Adverse Effects on the Nervous System; Parkinson's Disease.

Further Reading

- Belanoff, J. K., Gross, K., Yager, A., et al. (2001). Corticosteroids and cognition. *Journal of Psychiatric Research* 35, 127–145.
- Bremner, J. D. (1999). Does stress damage the brain? *Biological Psychiatry* 45, 797–805.
- Deshmukh, V. D. and Deshmukh, S. V. (1990). Stress-adaptation failure hypothesis of Alzheimer's disease. *Medical Hypotheses* 32, 293–295.
- Fratiglioni, L., Paillard-Borg, S. and Winblad, B. (2004). An active and socially integrated lifestyle in late life might protect against dementia. *Lancet Neurology* 3, 343–353.
- Kirschbaum, C., Wolf, O. T., May, M., et al. (1996). Stress and treatment-induced elevations of cortisol levels associated with impaired declarative memory in healthy adults. *Life Science* 58, 1475–1483.
- McEwen, B. S. (2002). Sex, stress and the hippocampus: allostasis, allostatic load and the aging process. *Neurobiology of Aging* 23, 921–939.
- Meyer, J. S., Rauch, G., Rauch, R. A., et al. (2000). Risk factors for cerebral hypoperfusion, mild cognitive impairment, and dementia. *Neurobiology of Aging* 21, 161–169.
- Nasman, B., Olsson, T., Viitanen, M., et al. (1995). A subtle disturbance in the feedback regulation of the hypothalamic-pituitary-adrenal axis in the early phase of Alzheimer's disease. *Psychoneuroendocrinology* 20, 211–220.
- Sapolsky, R. M. (1999). Glucocorticoids, stress, and their adverse neurological effects: relevance to aging. *Experimental Gerontology* 34, 721–732.
- Sapolsky, R. M., Romero, L. M. and Munck, A. U. (2000). How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrinology Review* 21, 55–89.
- VonDras, D. D., Powless, M. R., Olson, A. D., et al. (2005). Differential effects of everyday stress on the episodic memory test performances of young, mid-life, and older adults. *Aging and Mental Health* 9, 60–70.
- Waldstein, S. R. and Katzel, L. I. (2005). Stress-induced blood pressure reactivity and cognitive function. *Neurology* 64, 1746–1749.
- Waldstein, S. R., Siegel, E. L., Lefkowitz, D., et al. (2004). Stress-induced blood pressure reactivity and silent cerebrovascular disease. *Stroke* 35, 1294–1298.
- Wilson, R. S., Evans, D. A., Bienias, J. L., et al. (2003). Proneness to psychological distress is associated with risk of Alzheimer's disease. *Neurology* 61, 1479–1485.
- Wilson, R. S., Barnes, L. L., Bennett, D. A., et al. (2005). Proneness to psychological distress and risk of Alzheimer disease in a biracial community. *Neurology* 64, 380–382.

Neurodevelopmental Disorders in Children

N J Rinehart and B J Tonge

Monash University, Clayton, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

Autism and Asperger's Disorder

Attention-Deficit/Hyperactivity Disorder (ADHD)

Stress Associated with Autism and ADHD

Conclusion

Glossary

Pervasive developmental disorders (PDD) A group of serious neuro-developmental disorders affecting communication and social skills and behaviour.

Autism (autistic disorder) A PDD that involves delayed and deviant language development, impaired social skills and ritualistic and repetitive

patterns of behaviour, usually associated with intellectual disability but a minority have normal IQ (high functioning).

A PDD characterised by normal intellectual ability and language development but the presence of impaired social skills and ritualistic and repetitive patterns of behaviour.

A term usually applied to represent autistic disorder, Asperger's disorder and atypical autism or PDD not otherwise specified (PDD-NOS) which share some common features of problems with communication and social skills and behaviour.

A neuro-developmental disorder presenting with inattention, impulsiveness and motor hyperactivity excessive for the developmental age and usually associated with learning problems.

Asperger's disorder

Autism spectrum disorder

Attention-deficit/hyperactivity disorder (ADHD)

Autism and Asperger's Disorder

Autistic disorder (referred to as autism here) and Asperger's disorder are classified under the umbrella term pervasive developmental disorders (PDDs). PDD applies to children who share the core features of severe and pervasive impairment in social and communication skills, together with the presence of restricted and repetitive patterns of behavior and interest. Other disorders that are categorized as PDDs include Rett's disorder, childhood disintegrative disorder, and pervasive developmental disorder, not otherwise specified (PDD-NOS). These disorders have their onset within the first 3 years of life. This section focuses on the two main PDDs, autism and Asperger's disorder. The term autistic spectrum disorders (ASDs) is frequently used to describe a group of individuals with either autism or Asperger's disorder and is used accordingly in this article.

Epidemiology

The prevalence of autism is around 10–12 per 10 000. There is no evidence that prevalence varies between countries or racial groups, and social class and level of parental education are not associated with autism. Autism is more common in boys than girls (4:1), and this gender distribution is even more marked in Asperger's disorder (10–13:1).

Clinical Features of Autism

Autism presents with delays and abnormalities in the development of language and social skills and the presence of rigid, repetitive, stereotyped play and behavior, often in association with intellectual disability (i.e., an IQ below 70) and a variety of neurological conditions such as epilepsy. A reliable diagnosis can be made from the age of 2 years. Parents usually first seek help because their child has language delay and a lack of nonverbal communication. About 50% of children with autism fail to develop functional speech and only slowly learn to compensate with gesture. Language development is often abnormal in the remainder, with echolalia, self-directed jargon, and the repetition of irrelevant phrases, for example, from a television show. The correct use of pronouns and the related development of a sense of self and others are delayed. Poor comprehension, problems expressing needs by words and gesture, and difficulty in social understanding are frequently the causes of stress, frustration, and disturbed behavior. Children who do develop functional language usually have difficulty in using language socially and in initiating or sustaining reciprocal conversations. In contrast to children with autism, young people with Asperger's disorder

have no delay in the development of normal expressive and receptive language, including the use of communicative phrases by the age of 3. However, children with Asperger's disorder have problems in their social use of language, for example, being verbose and preoccupied with a favorite topic. Their speech may appear odd due to the use of an unusual accent or the presence of abnormalities in pitch and volume, leading, for example, to a flat and monotonous delivery.

The play, behavior, and daily life of children with autism is usually rigid and repetitive. Their ritualistic play lacks imagination and social imitation. The older children may develop preoccupations with themes such as train timetables or dinosaurs, and this will be the focus of their play, drawing, and conversation. Change or unexpected events can be a significant source of stress: for example, the change of a school timetable.

Approximately 80% of children with autism also have intellectual disability, and a range of other emotional and behavioral disturbances is common. Children with autism who have intellectual (thinking) abilities within the normal range are referred to as high functioning. In contrast, all children with Asperger's disorder have normal intellectual abilities.

High levels of inattentiveness and behavioral hyperactivity are common problems in children with autism and Asperger's disorder; for example, at least 13% of children with autism also meet criteria for ADHD (although a diagnosis of ADHD cannot be made concurrently with a diagnosis of ASD). The overlap of these disorders suggests that ASDs and ADHD share areas of brain dysfunction.

The Neuropsychology of ASDs

There are three main brain theories of autism: theory of mind, the executive dysfunction theory, and the theory of weak central coherence. A poor theory of mind refers to problems with understanding that other people have their own perspectives and thoughts; it is believed to be a reason why people with ASDs have difficulties understanding and relating to other people. The executive dysfunction theory accounts for the problems that people with ASDs experience with timing and planning and producing appropriate and relevant behavior (e.g., coping with changes in the school timetable, planning homework). This breakdown in the ability to effectively direct one's behavior is also thought to be related to the repetitive, stereotyped, and restricted behavioral patterns of ASDs. Weak central coherence refers to a problem with being unable to see the forest for the trees and is thought to be responsible for why people with ASDs tend to be obsessed with details and often miss the bigger picture.

Standardized tests have demonstrated that the neuropsychological profile of autism is characterized primarily by deficient cognitive flexibility and planning, but without problems in the ability to sustain and direct attention (problems attributed to the prefrontal regions of the brain). Research suggests that the kinds of problems that individuals with Asperger's disorder have in directing their behavior may be different from those observed in autism, for example, their problem may be more one of initiating behavior than moving effectively from one behavior to another, perhaps suggesting that different parts of the prefrontal brain are involved in each disorder. Neuromotor testing has found that the walking and upper body postural alignment of people with ASDs is similar to that observed in patients who have Parkinson's disease and patients with cerebellar ataxia, indicating the role of the basal ganglia and cerebellum in these disorders. Again, autism and Asperger's disorder may be associated with different kinds of motor problems, which would suggest the differential involvement of these brain regions.

Neuroimaging Studies in ASDs

Neuroimaging techniques that enable us to look at the structure and function of the brain (e.g., computed tomography [CT], magnetic resonance imaging [MRI], and functional magnetic resonance imaging [fMRI]) have been unsuccessful in revealing a part or function of the brain that is consistently abnormal in all children diagnosed with ASDs. Structural changes in the brains of individuals with autism include a slightly increased average brain volume, a reduction in neuronal integrity in prefrontal areas (an area important for executive functioning – see preceding), decreased gray matter volume in the limbic system (an area important for social understanding), reduced neuron numbers in the vermis of the cerebellum, and gross structural changes in the cerebellum (an area important for motor and cognitive functioning) and the parietal lobes (an area important for efficient attention). fMRI studies have shown decreased activation in the highly interconnected cortical and subcortical frontal structures, suggesting a disconnection between thinking and motor processors in the brain. In summary, the brain imaging research data showing multiple brain region involvement and inconsistencies between people with ASDs fits with the clinical definitions of these disorders as being psychiatrically and neurologically complex and heterogeneous.

Genetics

The cause of the majority of cases of ASD remains unknown, but they are almost certain to have a

multifactorial and complex genetic basis. Autism is associated with defined environmental causes, such as rubella and cytomegalovirus, fetal infections, perinatal brain injury, toxins, and specific genetic abnormalities such as tuberous sclerosis and fragile X syndrome in less than 10% of cases. A higher incidence of prenatal stressors occurring between 21 and 32 weeks of gestation has been associated with the later development of autism. A suggested link between measles, mumps, and rubella vaccination and the use of thiomerosol in vaccines as a cause of the increased prevalence of autism has been discounted by several comprehensive studies.

Attention-Deficit/Hyperactivity Disorder (ADHD)

The symptoms of ADHD may vary in different settings, and diagnosis is often subdivided into the number of symptoms in each of the dimensions of inattention or hyperactivity/impulsiveness. Thus, a child may be diagnosed as ADHD predominately hyperactive type or predominantly inattentive type. Estimates of the prevalence of ADHD vary widely, ranging from 20 to 1–2% depending on the application of diagnostic criteria. The combined inattentive-hyperactive subtype is the most common presentation.

Epidemiology

The childhood prevalence of ADHD is approximately three times higher in males than in females. Symptoms usually reduce with maturation, but at least 30% of children with ADHD will continue to suffer from the disorder in adulthood.

Clinical Features of ADHD

The diagnosis of ADHD is based on a clinical judgment that there are sufficient symptoms of inattention and hyperactivity/impulsiveness, together with the decision that these symptoms cause significant impairment in daily functioning in at least two settings and are not consistent with the developmental level of the child. Therefore, diagnosis requires a careful and comprehensive history of the child's development and behavior from the parents and other informants such as the teacher, together with observation of the child's behavior.

Apart from showing high levels of distractibility and inattention, children with ADHD are disorganized and are usually unable to follow routine or complete tasks. They have difficulty monitoring their behavior and therefore often interrupt others, have difficulty following rules, and display inappropriate and impulsive behavior. Those who also suffer from

hyperactivity are constantly restless and fidgety, have difficulty remaining seated, and behave as if they are driven by a motor. These behaviors are influenced by aspects of the environment, such as the degree of external stimulation and sensory complexity. Therefore, observers may report differences in behavior depending upon the context. For example, a teacher in a busy, noisy classroom setting is more likely to observe inattention than a teacher's aide who meets with the child for individual teaching in a quiet library environment. However, the symptoms and impairments are usually observed, at least to some extent, in all aspects of the child's daily life.

The Neuropsychology of ADHD

In common with ASDs, neurocognitive testing has revealed that ADHD involves deficiencies within the prefrontal regions of the brain described under the umbrella term of executive dysfunction. However, the type of prefrontal executive dysfunction differs between ASDs and ADHD. The classic ADHD profile is characterized by inhibitory deficits and problems with sustained attention, compared to planning and cognitive flexibility deficits in the cognitive profile of children with autism.

Neuroimaging Studies of ADHD

As is the case with ASDs, no consistent structural abnormalities have been recorded for individuals affected by ADHD. ADHD has been associated in some but not all studies with smaller whole brain volumes, right prefrontal anomalies, and structural anomalies in the basal ganglia, cerebellum, and corpus callosum. As is the case with ASDs, fMRI studies have shown decreased activation in the highly interconnected cortical and subcortical frontal structures, particularly in neural regions that subserve key attentional and inhibitory functions.

Genetics and Stressors

The inherited basis for ADHD has been estimated in around 80% of cases, with perinatal brain injury (e.g., caused by stress, alcohol, tobacco, and other substance abuse during pregnancy) responsible for the remainder. The cause in the inherited cases is likely to be a complex interaction of multiple genes, while in some individuals there will also be an interaction with perinatal brain injury or other environmental stressors such as fetal alcohol exposure or postnatal malnutrition. As is the case with ASDs, there has been a large number of candidate gene and linkage studies of ADHD. Genetic screening for ASDs and ADHD is not yet possible.

Stress Associated with Autism and ADHD

Neurodevelopmental disorders present a considerable stress and burden for the affected individual, parents, siblings, and caregivers in the community. The particular stressors change over time and therefore need to be considered in the context of a human life span approach.

Stress Issues from Birth to Preschool

The trajectory of the stress and devastation that accompany a diagnosis of a neurodevelopmental disorder begins for the parents when their child is between 1 and 3 years of age, when parents realize, either through their own observations or through those of others (e.g., extended family, maternal health nurse, general practitioner), that something is not quite right with their child's developmental progress, for example, behavioral deficits in the form of language or social delays, inattentiveness, or behavioral excesses in the form of increased motor activity and hypervigilance. This realization usually triggers a string of subsequent referrals for pediatric, audiological, psychological, and speech and language assessment. Incumbent in each of these referrals is set of neurobiological and developmental test protocols to rule out various diagnoses. For example, a pediatrician may request a series of metabolic (e.g., phenylketonuria) and genetic tests (e.g., fragile X). The psychologist may undertake behavioral and cognitive standardized assessments, while the speech pathologist may request standardized assessments of receptive (i.e., understanding) and expressive (communicative) language. Investigation and testing may occur over a period of 3–12 months, at which time the family will be referred to a specialist (e.g., child psychiatrist, clinical psychologist) or specialist clinic for final diagnostic confirmation. The time length and uncertainty about their child's diagnostic status that parents experience during the assessment process is often stressful.

The formal diagnosis of a neurodevelopmental condition will often trigger a significant grief reaction in parents, who may mourn the loss of their perfect child and be required to reconsider future plans they may have had for their child, e.g., attending particular schools. In the case of ASDs, or when the child has additional severe intellectual disability, there may be a sense of loss about not having grandchildren, as many people affected by these disorders do not go on to partner and have their own children. Furthermore, research has shown that mothers of children with autism are more likely to suffer from depression than mothers of children with other physical conditions (e.g., cystic fibrosis) and mental disabilities.

Parents are typically provided with information about the genetic component of neurodevelopmental disorders at the time of diagnosis. For example, in the case of ASDs, if an older sibling has autism, the risk that a subsequent sibling will have autism is 2–8%. This information is stressful for parents because it has implications for younger siblings and for future family planning.

After a formal assessment has taken place, parents will play an important advocacy role throughout the child's life, a role that requires a significant emotional and financial commitment to ensure that a high quality of life is maintained for their child at each developmental stage. During the preschool years, parental advocacy typically involves organizing appropriate early interventions (e.g., speech and language, occupational therapy, behavioral management). The process of dealing with often lengthy waiting lists to access public services, or self-funding private interventions places a considerable emotional and financial burden on families and may place stress on the larger family unit by increasing work hours for caregivers and reducing time and resources for other family members, in particular, unaffected siblings. Surveys of maternal stress levels have estimated that approximately two-thirds of mothers of children with ASDs experience clinically significant levels of stress when their child is between 2 and 7 years of age. Stress levels are particularly high when a child with ASD has high levels of maladaptive and disruptive behaviors.

Children with ADHD often have hostile interactions with other children (e.g., hitting and biting) during the preschool years. This places a considerable stress not only on parents, but also on the broader preschool community (e.g., parents of other children and teachers). Compared to fathers, mothers have reported that they find these behaviors to be significantly more stressful.

Receiving education about the disorder and skills training in managing problem behaviors, learning relaxation and stress management techniques, and increased social support (e.g., respite care and community support groups) have been found to be useful for improving parental mental health and reducing parental stress associated with having a child with a neurodevelopmental condition.

Stress Associated with the Primary School Years

Interventions put in place during preschool years will typically continue into the primary school years depending on the availability of community and privately funded resources. During the school years, parents often become concerned about their child's

education and adverse school experiences such as being teased. They focus on advocating for appropriate educational modifications and placements to promote their child's ability to reach individual academic potential and to experience a safe and supportive school environment. This usually involves families again seeking out the services of psychologists and speech pathologists so that their child can have updated assessments of language and cognitive ability. This process is a burden for children and parents because it is time consuming and costly. In some cases, at the end of such assessment processes, children are deemed ineligible for education aide support or special consideration because of bureaucratic and funding reasons (e.g., the child's IQ is 72 points rather than 70 points). This issue sometimes requires higher levels of advocacy from involved health professionals to assist parents with appeal processes.

For children with autism who have good early intervention, focused on the parent (e.g., behavior management training, psychoeducation) as well as the child, disruptive behaviors are manageable, and the quality of life for the child and family is generally improved. For children who do not receive good early interventions, or when standard interventions are inadequate, perhaps due to additional neurobiological comorbidity and ongoing disruptive behaviors (e.g., stubbornness, self-injury, and aggression), there is an increased risk of failure in school and community activities, in some cases leading to more restrictive care. When psychological and behavioral interventions have not been successful in the management of disturbed emotions and behaviors, which may be underpinned by high levels of anxiety, treatment with neuroleptic medication (e.g., haloperidol, risperidone, imipramine, clomipramine, fluoxetine) that targets specific symptoms might prove helpful.

While the antisocial behavioral tendencies (e.g., biting and hitting) of children with ADHD may subside with good management during the preschool years, for older children poor management and associated comorbidity (e.g., conduct disorder and oppositional defiant disorder, learning disorders) increase the chances of school failure and peer rejection. Comorbidity with oppositional defiant disorder (e.g., clashing verbally with authorities) has been suggested at between 32 and 60%, while comorbidity with conduct disorder (e.g., engaging in theft) has been estimated at anywhere between 12 and 25%. Stimulant medications (e.g., methylphenidate and dexamphetamine) are effective treatments for the symptoms of ADHD. Side effects, which may include decreased appetite, insomnia, fatigue, irritability, and increase in pulse rate and blood pressure, may be a

source of stress for the individual and for parents. Tricyclic antidepressants (imipramine) clonidine, or atomoxetine (an inhibitor of the presynaptic norepinephrine transporter) are alternative treatments for children who have intolerable side effects or are unresponsive to stimulant medication. Parents may be stressed by the uncertainty of potential long-term effects of placing their child on medication. Currently there is a lack of good long-term evidence for the effectiveness of stimulant medication on learning, but they may decrease the risk of later substance abuse.

Stress Associated with the Secondary School Years

In the normal scheme of development, the transition from a relatively protected and highly structured primary school setting to secondary school is often a time of stress for children and families; these concerns and issues are magnified when a child has a neurodevelopmental disorder. In contrast to primary school, secondary school is often larger and requires the young person to be independent, navigate the school timetable, and engage in homework. All of these activities require good executive functions, e.g., planning, preparation, and anticipation, skills classically affected by neurodevelopmental disorders.

Adolescence is a developmentally complex time defined by social and sexual development and identity formation. Because social understanding and adaptability are innately reduced in ASDs it is an especially complex time for these young people. The affected young person may be more stressed than their nonaffected peers as they go through puberty and deal with changes in physical body shape, development in their thinking and insight, and emerging sexuality. Stress is also elevated for those individuals who develop comorbid psychiatric conditions, particularly depression, during this time. Depression may manifest as mood disturbance and irritability, sleep and appetite disturbance, and thoughts of suicide that may be enacted. The increased vulnerability to depression during adolescence may be associated with self-awareness of the disability, pubertal brain development, and a family history of depression. Stress and anxiety have been associated with the onset of tic disorder, or Tourette syndrome, in young people with ASDs. There is also an increased risk of developing psychosis during adolescence or early adult life, particularly for young people with Asperger's disorder. Individuals with associated intellectual disability are at greater risk for developing epilepsy during the adolescent and early adult years.

In addition to ongoing issues that first emerge in the primary school years (e.g., associated learning difficulties estimated in up to 50% of children with

ADHD) for young people affected by ADHD and their families, new hazards during secondary schooling revolve around ongoing and escalating child-parent conflicts, delinquent behavior, substance abuse (e.g., nicotine, alcohol, and illicit drug use), and an increased risk of being involved in car accidents, all of which may lead to school suspensions, community marginalization, and an increased risk of a breakdown of family relationships. It has been noted that increased marital discord when a child in the family has been diagnosed with ADHD may be influenced by the parents themselves having ADHD, given the associated genetic vulnerability. Young people who are treated with stimulant medication for ADHD are less likely to use substances compared with adolescents with ADHD who are not receiving treatment.

In contrast to individuals with ADHD, individuals with ASDs often go to lengths not to use drugs and alcohol, an issue that may paradoxically cause stress and social conflict because the individual with ASD may feel driven to impose their rigid views about abstinence onto unsuspecting others, for example, taking a cigarette or alcoholic beverage out of the hand of a friend or restaurant patron.

Stresses during Adulthood

For people affected by ASDs, the adulthood stressors vary widely according to the individual's intellectual ability, capacity for independent living, and provision of social, emotional, financial, and occupational support. It has been suggested that individuals with associated intellectual disability and epilepsy are at greater risk for deterioration during early adulthood, while there may be some improvement in functioning in normally intelligent young adults with ASDs. People with moderate to profound intellectual ability may be placed in special housing accommodation or residential care, depending on the family's emotional and financial ability to retain care within the family unit. This time of transition may be particularly stressful for the individual because it involves adaptation to a new environment and many changes in daily routine. Transition is stressful for family members involved in the decision-making process because they often feel a great sense of guilt and loss. Thus, the grieving process over the loss of the young person's potential, which begins for parents when they first come to terms with the diagnosis, is often revisited at this developmental point. Alternatively, the prospect of providing continuing care, without the hope of relief, is a source of anxiety for aging parents.

Even when young people with ASDs have relatively intact intellectual ability (e.g., ranging from mildly impaired to normally intelligent), they remain greatly

impaired in their ability to process social information. Like the transition from primary school to secondary school, the transition to adulthood involves another leap of independence, involving a greater need for executive abilities (e.g., budgeting money, planning meals) and social adjustment. Indeed, the social rules may become even more perplexing for an adult with ASDs, as the immediate supports and supervision from parents, teachers, and school aides reduce or disappear as the young person is expected to live independently in the community. The transition to adulthood is perhaps more complex for individuals who have poorer adaptive and life skills, with a lack of early intervention and skills training and poor management of associated comorbid psychiatric conditions such as anxiety during the formative years.

Reduced social ability and obsessional characteristics associated with ASDs may put an individual at greater risk for being unwittingly involved in inappropriate or illegal behavior, for example, sexual offending or stalking, which may occur in the context of naive and inept attempts to make social contact. Similarly, there have been reports in the Australian and European media about individuals with ASDs stealing trams or trains to pursue an obsessional interest in the rail system. Interventions to manage these types of forensic issues may be punitive and inappropriate, particularly when the individual's diagnosis may not be known, leading to additional feelings of being rejected and further emotional and behavioral problems.

Adults with ADHD are usually less disruptive and hyperactive than children with ADHD, but usually remain somewhat impulsive, disorganized, inattentive, and restless or lack motivation. Young adults with comorbid conduct disorder and an inability to control their behavior are at increased risk for being in trouble with the law, ranging from accumulating multiple speeding fines to committing more significant crimes, such as robbery. The disinhibited, poorly focused, and overactive symptoms of those with ADHD may make it difficult to maintain employment and social and stable intimate relationships.

Conclusion

ASDs and ADHD are lifelong neurodevelopmental conditions, possibly involving overlapping brain regions, that are associated with numerous and accumulating individual stressors across the life span. Despite the many differences between ASDs and ADHD, there are many commonalities in the factors that may protect against or lessen the impact of these stressors. At the forefront of these is the availability of early interventions, parent support and guidance, and ongoing management at key developmental stages from health, education, and welfare services (e.g., general practitioners, psychologists, psychiatrists). These may prevent the accumulation of stressful events and related complications that significantly reduce the quality of life of the affected individual and family and are costly to society with an increased prospect of unemployment, family and relationship breakdown, legal issues, and psychiatric care.

See Also the Following Articles

Attention-Deficit/Hyperactivity Disorder, Stress and; Parenting, Stress of.

Further Reading

- Barkley, R. A. (1998). *Attention deficit hyperactivity disorder: a handbook for diagnosis and treatment* (2nd edn.). New York: Guilford Press.
- Bradshaw, J. L. (2001). *Developmental disorders of the frontostriatal system: neuropsychological, neuropsychiatric and evolutionary perspectives*. Hove, East Sussex, UK: Psychology Press Ltd.
- Brereton, A. V. and Tonge, B. J. (2005). *Preschoolers with autism: an education and skills training program for parents*. London: Jessica Kingsley.
- Cohen, D. J. and Volkmar, F. R. (eds.) (1997). *Handbook of autism and pervasive developmental disorders*, (2nd edn.). New York: Wiley & Sons.
- Howlin, P. (2000). Outcome in adult life for more able individuals with autism or Asperger syndrome. *Autism* 4(1), 63–83.

Neuroendocrine Systems

G Fink

Mental Health Research Institute, Melbourne, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G Fink, volume 3, pp 14–30, © 2000, Elsevier Inc.

Introduction

Hypothalamic-Pituitary Axis

Neurohemal Junctions and Circumventricular Organs

The Hypothalamo-Adenohypophyseal System

The Hypothalamo-Neurohypophyseal System

The Intermediate Lobe of the Pituitary Gland

Pineal Gland: Melatonin and Photoperiodic Control of Circadian Rhythms

Hormonal Effects on the Nervous System

Psychoneuroendocrinology

Glossary

<i>Action potentials</i>	Electrical impulses that propagate signals along nerve fibers. The potentials are produced by the flux of ions, in particular sodium and potassium ions, through ion channels along the nerve cell membrane.
<i>Exocytosis</i>	The release of a secretory product, including neurotransmitters, in packets or quanta. The secretory products are packaged by the Golgi apparatus of the cell in membrane-bound vesicles. Exocytosis involves fusion of the vesicular membrane with the cell membrane, resulting in the release of the quantum of product inside the vesicle. This is thought to be the major mechanism of secretion in most secretory cells, including neurons.
<i>Limbic system</i>	The system comprising a number of structures, all interconnected, to form a circuit that is concerned primarily with emotions, memory, olfaction, and neuroendocrine control; gets its name from the fact that it forms a border between the new and old (in evolutionary terms) brains.
<i>Neurohemal junction</i>	A specialized junction between nerve terminals and a plexus of capillaries that facilitates the transport of neurohormones from the nervous system into the bloodstream or <i>vice versa</i> .
<i>Neurohormone</i>	A chemical substance released from nerve terminals and transported to its target cells by the bloodstream. Neurohormones

are often the same chemical substances as neurotransmitters.

Neurosecretion

The secretion of hormones by nerve cells.

Neurotransmitter

A chemical compound released by nerve cells (at nerve synapses or neuromuscular junctions) that either excites or inhibits a target cell. The target cell may be another nerve cell, glandular cell, or muscle fiber. Chemical neurotransmission is widely used in the nervous system and, because it is capable of analog functions, has a considerable advantage over electrotonic transmission, which is necessarily digital.

Portal vessels

The vessels that transport blood and solutes from one capillary bed to a second capillary bed before joining the systemic circulation. The hypophyseal portal vessel system transports neurohormones released from hypothalamic nerve terminals into a primary capillary plexus in the median eminence to a secondary capillary–sinusoidal plexus in the pituitary gland.

Steroid hormones

Molecules composed of a basic fused unit of three 6-membered (cyclohexane) and one 5-membered (cyclopentane) carbon rings with a variety of functional groups attached. The large number of carbon–hydrogen bonds makes steroids nonpolar.

Stimulus secretion coupling

The link between a stimulus, usually membrane depolarization (axon potentials in the case of nerve cells), and the release of a neurotransmitter or neurohormone.

Synapse

A specialized junction between two nerve cells where signaling most commonly occurs by way of the release of neurotransmitters.

Introduction

The human pituitary gland weighs no more than one gram, but it nonetheless controls all the major endocrine systems and is indispensable for life. Located at the base of the brain and closely surrounded by protective dense bone and fibrous membranes, the gland is composed of the neurohypophysis or neural lobe and the adenohypophysis. Derived embryologically from a neural downgrowth, the neural lobe is composed of axons that project from nerve cells in the hypothalamus and terminate on capillaries of the

inferior hypophysial artery. This is the site at which arginine vasopressin (AVP) (lysine vasopressin in pigs) and oxytocin are released into the systemic circulation. Synthesized in the supraoptic nucleus (SON) and paraventricular nucleus (PVN), AVP, also called the antidiuretic hormone, controls the volume of body water, whereas oxytocin is concerned mainly with stimulating milk ejection during lactation and contraction of the uterus (womb) during parturition.

The adenohypophysis is derived from an outgrowth of the roof of the mouth and, in subhuman species, is divided into the pars distalis and the pars intermedia. The pars distalis is more commonly called the anterior lobe, whereas the pars intermedia together with the neural lobe or pars nervosa forms the posterior lobe. There is no distinct pars intermedia in the human. The anterior pituitary gland secretes hormones that control the adrenal and thyroid glands, the gonads, body growth, and development of the breast and lactation. The hormones involved in this pituitary control are, respectively, adrenocorticotrophic hormone (ACTH), thyrotropin (TSH, thyroid stimulating hormone), the gonadotropins, growth hormone (GH), and prolactin. In addition to being growth promoting (trophic) and stimulating immediate hormonal or cellular events (tropic), these hormones all affect metabolism. The loss of ACTH with the consequent loss of adrenocortical hormone secretion, and perhaps to a lesser degree GH and TSH, makes removal of the gland (hypophysectomy) lethal.

Although ACTH is considered to be the main anterior pituitary stress hormone, GH and prolactin are also released in response to stress and so receive special attention in this article. Thus, for example, insulin-induced hypoglycemia is a potent metabolic stress that, in humans, results in a massive and rapid release of GH. Stress-induced prolactin release may, under certain circumstances, be causally related to infertility.

Anterior pituitary hormone secretion is under central nervous system (CNS) control and is modulated by the feedback of hormones secreted by the pituitary target organs: the adrenal and thyroid glands, the gonads, and, as shown recently, fat tissue by way of adipokines such as leptin and adiponectin. Neural control of the anterior pituitary hormones is mediated by the hypothalamic-pituitary regulatory neurohormones (formerly termed factors), which are released from the hypothalamus into the hypophysial portal vessels and transported by them to the anterior lobe, where they either stimulate or inhibit the release of the anterior pituitary hormones. The hypothalamic-pituitary regulatory factors are now termed neurohormones because, instead of being released at

synapses, they are released and transported to their target cells in the bloodstream.

The hypothalamic-pituitary system is the interface between the CNS and the endocrine system by which external factors, such as stress and day length, and internal factors, such as emotion, trigger endocrine responses. It was therefore termed the neuroendocrine system, and the pituitary gland is said to be under neuroendocrine control. In addition to the hypothalamic-pituitary system, the circumventricular organs also satisfy the criteria of neuroendocrine systems. The term neuroendocrine is also applied to interactions between nerve and endocrine cells in the periphery and especially the viscera, of which the gastrointestinal system and its appendages (especially the pancreas) are the most prominent. Of special relevance to stress is the sympathetic component of the autonomic nervous system and the adrenal medulla. Derived from the embryonic neural crest, the cells of the adrenal medulla are modified postganglionic neurons that secrete their neurotransmitters, mainly adrenaline (epinephrine) and opioids, directly into the bloodstream.

The present article focuses on the principles of neuroendocrine control of the pituitary gland and illustrates how the hypothalamic-pituitary system (1) was used to demonstrate that certain peptides satisfy the criteria of neurotransmitters or neurohormones; (2) played a significant role in our understanding of gene transcription, translation, and posttranslational processing; and (3) could be used as a window to study brain function in the conscious human (neuroendocrine window of the brain). Although only mentioned briefly in this article, neuroendocrinology is also concerned with the effects of pituitary target hormones on the brain-pituitary system, that is, on the way that hormones secreted by the anterior pituitary target glands play an important role in brain differentiation and plasticity; affect central neurotransmission and thereby mood, mental state, and memory; and exert feedback effects on the brain-pituitary system. These hormonal feedback systems constitute the afferent limb of homeostatic regulatory systems, which ensure that the output of pituitary hormones is maintained at a preset and functionally optimal level.

Hypothalamic-Pituitary Axis

Brief History

The central importance of the anterior pituitary gland as conductor of the endocrine orchestra was not understood until the early 1930s when P. E. Smith published his parapharyngeal method for removing

the gland (hypophysectomy). The effects of hypophysectomy proved to be so dramatic that for a short period most scientists in the field, including the distinguished neurosurgeon, Harvey Cushing, thought that the pituitary gland was autonomous. However, around the same time, William Rowan, working in Alberta on the annual migration of birds, showed that day length had a potent effect on the growth of the gonads. Rowan's experiments, together with those on seasonal breeding in animals, the effects of stressful stimuli on endocrine organs, and the effects of brain lesions on pituitary hormone secretion led to the concept that the anterior pituitary gland must be under CNS control, a view that Cushing soon adopted. The observational and experimental evidence that supported this concept was summarized by F. H. A. Marshall in his 1936 Croonian lecture.

It had long been known that the pituitary gland and brain were connected by the pituitary stalk, but several lines of evidence suggested that neural control of the anterior pituitary gland was mediated not by nerve fibers in the pituitary stalk but by chemical substances released into the hypophyseal portal vessels. These vessels, first described by G. T. Popa and U. Fielding in 1930, surround the pituitary stalk, linking a primary plexus of capillaries at the base of the hypothalamus with a second capillary plexus in the anterior pituitary gland. Throughout the 1930s and 1940s, a debate raged about the direction of blood flow in the hypophyseal portal vessels. Based on histological evidence, this debate could have been avoided had someone read the 1935 report by Bernado Houssay and associates that in the living toad blood flowed from the hypothalamus down to the pituitary gland. But Houssay's paper was published in French.

The neurohumoral hypothesis of the control of the anterior pituitary gland was first formally advanced by H. B. Friedgood in 1936 and J. C. Hinsey in 1937. However, it was the elegant pituitary graft experiments of G. W. Harris and Dora Jacobsohn that showed beyond doubt that the anterior pituitary gland was controlled by substances released at nerve terminals in the median eminence at the base of the hypothalamus and transported to the pituitary gland by the hypophyseal portal vessels. The characterization of the first three of these substances, thyrotropin-releasing factor (later thyrotropin-releasing hormone, TRH), luteinizing hormone-releasing factor (LRF; later, luteinizing hormone-releasing hormone, LHRH; or gonadotropin-releasing hormone, GnRH), and somatostatin, was to take a further 18–21 years of hard work in the laboratories of Roger Guillemin and Andrew Schally, for which they were awarded the 1977 Nobel Prize for Physiology and Medicine.

The sequence of corticotropin releasing factor-41 (CRF-41; later, corticotropin releasing hormone, CRH-41) was determined by Wylie Vale and associates in 1981, a breakthrough that had eluded many workers for more than 25 years.

Soon after its characterization as a decapeptide in 1971, G. Fink and M. Jamieson measured LHRH by radioimmunoassay in hypophyseal portal blood of the anesthetized rat and showed that its release could be increased fivefold by a small electrical stimulus applied to the medial preoptic area of the hypothalamus. The neurohumoral hypothesis of neural control of the anterior pituitary control had been proved.

General Anatomy and Development

The pituitary gland is anatomically linked to the hypothalamus at the base of the brain by way of the pituitary stalk ([Figure 1](#)). The hypothalamus comprises a medial part adjacent to the third cerebral ventricle, in which are located the major hypothalamic nuclei, and a lateral part made up mainly of a large cable of nerve fibers that carries reciprocal fiber tracts between midbrain and forebrain, the medial forebrain bundle ([Figure 2](#)), in which are embedded aggregations of nerve cell bodies. Axons from nerve cell bodies located in the hypothalamic nuclei project to the median eminence, where they either terminate on the loops of primary capillaries of the hypophyseal portal vessels in the external layer of the median eminence or form a cable that passes through the internal layer of the median eminence to form the bulk of the pituitary stalk and then the neural lobe ([Figure 3](#)). The median eminence, so called because it protrudes as a small dome in the midline from the base of the hypothalamus, forms the floor of the third ventricle and is delineated by the optic chiasm in front, the mamillary bodies behind, and a depression (hypothalamic sulcus) on either side. Arising from the median eminence is the neural stalk, which links the pituitary gland to the brain.

The pituitary gland is located in a fossa in the basisphenoid bone at the base of the skull, the sella turcica (so called because its shape resembles a Turkish saddle). The close proximity of the hypothalamus and pituitary gland to the optic chiasm means that tumors in either the hypothalamus or the pituitary gland may press on, or more rarely invade, the optic chiasm or tracts, thereby leading to visual defects.

The hypothalamic-pituitary axis is divided functionally into two systems ([Figure 4](#)): the hypothalamo-adenohypophyseal axis (the hypothalamus, hypophyseal portal vessels, and adenohypophysis) and the hypothalamo-neurohypophyseal axis (the hypothalamus, neural stalk, and neural lobe). The neural stalk is made up of numerous nerve fibers that project

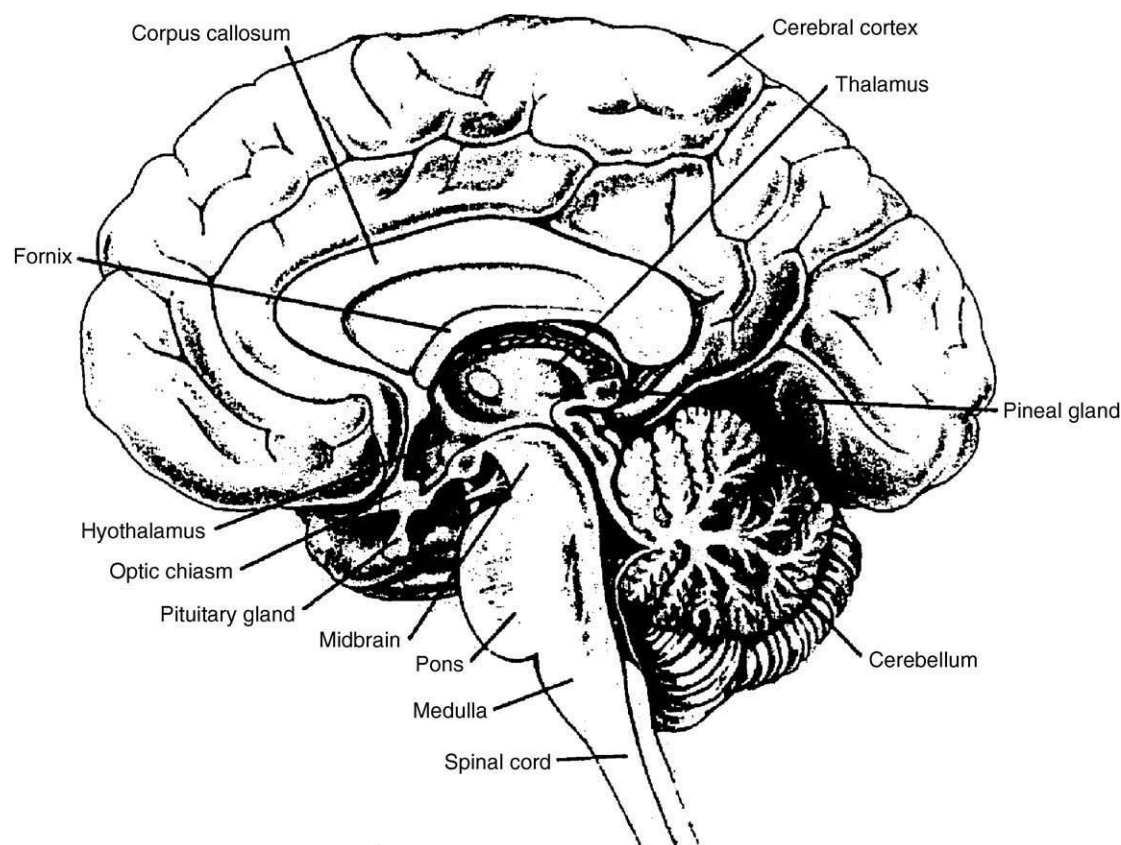


Figure 1 A midsagittal section of the human brain showing the inside (medial) surface. Note the pituitary gland attached by way of the pituitary stalk to the floor of the hypothalamus. The hypothalamus and the thalamus, which lies above it, form the wall of the third cerebral ventricle, at the posterior end of which is the pineal gland.

mainly from the PVN and SON of the hypothalamus to terminate on a capillary bed derived from the inferior hypophysial artery and located in the neural lobe of the pituitary gland. Nerve fibers in the neural lobe (or pars nervosa) are surrounded by pituicytes, the equivalent of glial cells. Also present in the neural stalk are nerve fibers of other types of chemical neurotransmitter (in addition to AVP and oxytocin), such as the endogenous opioids and dopamine, which are present in nerve fibers that project from the arcuate nucleus and innervate the pars intermedia. Because the stalk and median eminence are continuous, these dopaminergic neurons are a continuation of a dense palisade of dopaminergic fibers that also terminate on the primary plexus of the hypophysial portal vessels enmeshed with nerve fibers that contain other neurohormones. The pituitary stalk and median eminence are covered by a single layer of cells termed the pars tuberalis, which is continuous with the pars distalis.

The adenohypophysis develops from an outgrowth of the ectodermal placode, which forms the roof of

the embryonic mouth (or stomadeum). This ectodermal outgrowth forms Rathke's pouch and meets the neurohypophysis, which grows down from the floor of the embryonic third ventricle. Rathke's pouch closes and separates from the roof of the mouth. The caudal (rear) part of the pouch remains thin to form the pars intermedia, which becomes tightly juxtaposed to the rostral surface of the neurohypophysis ([Figure 5](#)). The rostral part of the pouch develops into the pars distalis ([Figure 5](#)). Vascularization of the median eminence and the pituitary gland begins on approximately day 15 of embryonic life (E15) in the rat, and the hypophysial portal vessels become defined by E18. Nerve terminals in the median eminence with granular vesicles (presumably neurohormones) are first evident on E16 and the first secretory granules appear in pars distalis cells on E17. This sequence of embryonic development suggests that the development of secretory cells in the pars distalis may depend in part on the development of nerve terminals in the median eminence and the anlage of the hypophysial portal vessels.

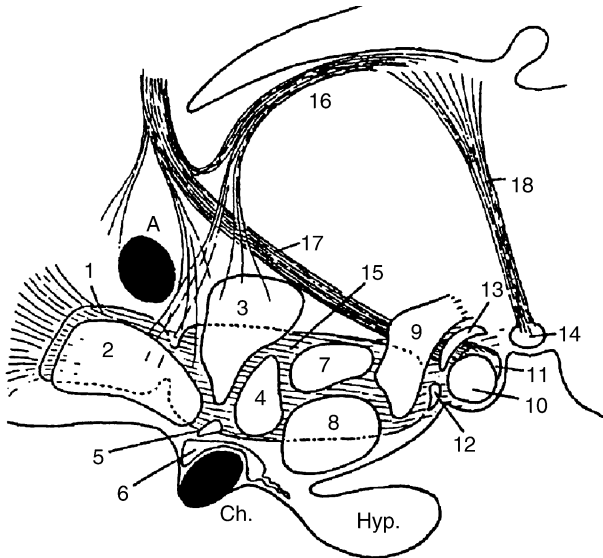


Figure 2 Diagram showing the relative positions in a sagittal plane of hypothalamic nuclei in a typical mammalian brain and their relation to the fornix, stria habenularis, and fasciculus retroflexus. (1) Lateral preoptic nucleus (permeated by the medial forebrain bundle). (2) Medial preoptic nucleus. (3) Paraventricular nucleus. (4) Anterior hypothalamic area. (5) Suprachiasmatic nucleus. (6) Supraoptic nucleus. (7) Dorsomedial hypothalamic nucleus. (8) Ventromedial hypothalamic nucleus. (9) Posterior hypothalamic nucleus. (10) Medial mamillary nucleus. (11) Lateral mamillary nucleus. (12) Premamillary nucleus. (13) Supramamillary nucleus. (14) Interpeduncular nucleus (a mesencephalic element in which the fasciculus retroflexus terminates). (15) Lateral hypothalamic nucleus (permeated by the medial forebrain bundle). (16) Stria habenularis. (17) Fornix. (18) Fasciculus retroflexus of Meynert (habenulopeduncular tract). A, anterior commissure; Ch., optic chiasma; Hyp., hypophysis (pituitary gland).

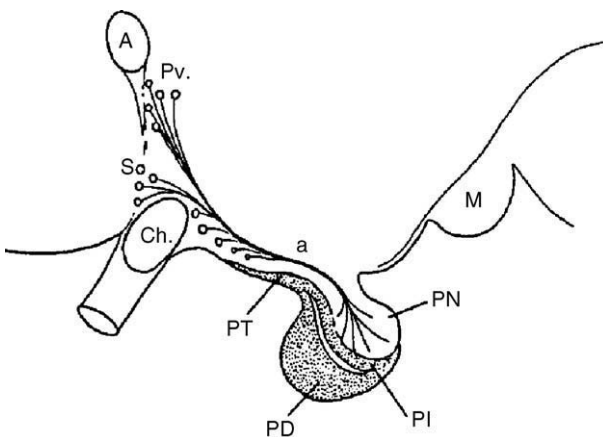


Figure 3 Schematic section of the mammalian hypothalamus and pituitary gland showing the neurohypophyseal tract (labeled a) comprised mainly of fibers derived from the paraventricular (Pv.) and supraoptic (So) nuclei. A, anterior commissure; Ch., optic chiasma; M, mamillary bodies; PD, pars distalis; PI, pars intermedia; PN, pars nervosa; PT, pars tuberalis.

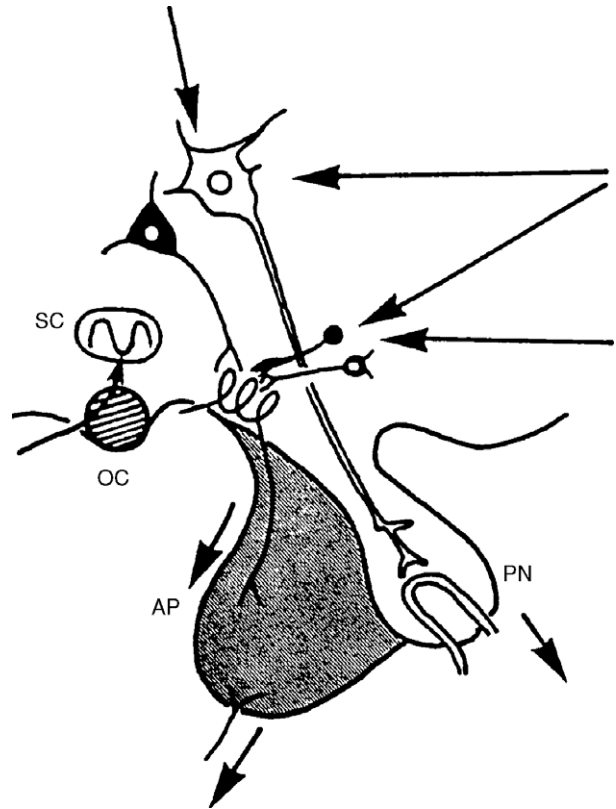


Figure 4 Schematic diagram of the hypothalamic-pituitary system showing the magnocellular (white) projections directly to the systemic capillaries of the pars nervosa (PN) and the parvocellular (black) projections to the primary plexus of the hypophyseal portal vessels, which convey neurohormones to the pars distalis of the anterior pituitary gland (AP). Dorsal to the optic chiasm (OC) are suprachiasmatic nuclei (SC), which receive direct projections from the retina and play a key role in the control of circadian rhythms (indicated by the sinusoidal curve). Activity of the intrinsic neurons of the hypothalamus is influenced greatly by projections (arrows) from numerous areas of the forebrain, midbrain, and hindbrain, particularly the limbic system, as well as by hormones, mainly estrogen and progesterone in the case of the hypothalamic-pituitary-gonadal system.

Neurohemal Junctions and Circumventricular Organs

Neurohemal junctions are the fundamental functional modules of the major central neuroendocrine system, the median eminence. They are composed of nerve terminals and capillaries that are closely juxtaposed and thereby facilitate the transport of chemical messengers from nerve terminals into the bloodstream and *vice versa* (Figure 6). Neurohemal junctions are also the fundamental units of the neurohypophysis and of the circumventricular organs, such as the organum vasculosum of the lamina terminalis, the subfornical organ, and the pineal gland, which are located at various sites around the third

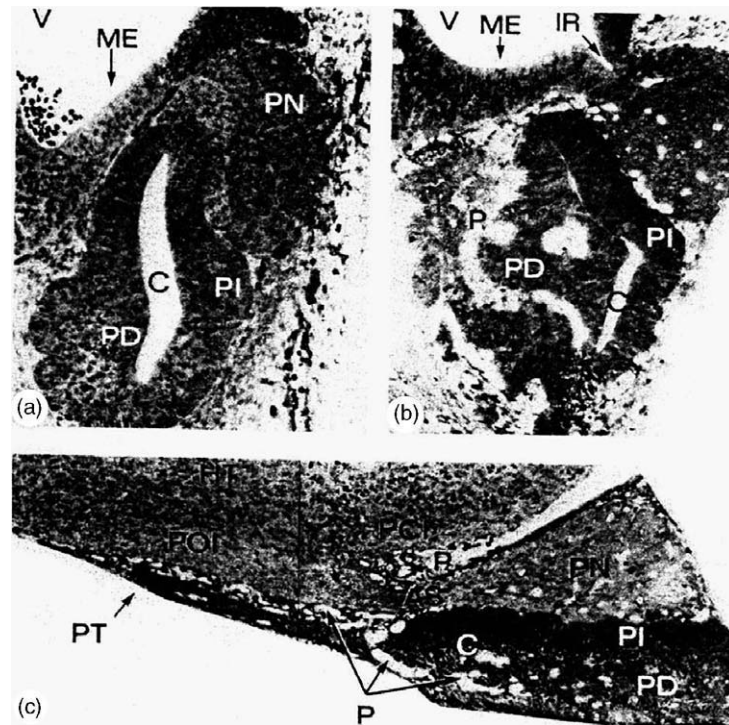


Figure 5 Photomicrographs of midline sagittal sections through the hypothalamic-pituitary complex showing the development of the pituitary gland in the rat. a, On embryonic day 15 (E15; $\times 120$); b, on E17 ($\times 120$); c, on E20 ($\times 80$). In a, the pituitary anlage shortly after closure of Rathke's pouch, which migrates dorsally to meet neurohypophysial downgrowth from the floor of the hypothalamus. Rotation of the pituitary gland caudally through 135° with respect to the base of the diencephalon (hypothalamus) is seen, as is the invasion of the pars distalis (PD) by the leash of portal vessels (P) on E17. AS, anatomical stem; C, hypophysial cleft; HT, hypothalamus; IR, infundibular recess; ME, median eminence; PCI, pars caudalis infundibuli; PI, pars intermedia; PN, pars nervosa; POI, pars oralis infundibuli; PT, pars tuberalis; V, third ventricle. (1- μ m Araldite sections, toluidine blue stain.) Reproduced with permission from Fink, G. and Smith, G. C. (1971), Ultrastructural features of the developing hypothalamo-hypophysial axis in the rat, *Zeitschrift für Zellforschung und mikroskopische Anatomie* **119**, 208–226.

cerebral ventricle, and the area postrema, located in the floor of the fourth cerebral ventricle. All the circumventricular organs are characterized by the fact that their vessels are fenestrated (Figure 6) and that the blood–brain barrier is inoperative at these sites. The neurohemal junctions in the median eminence, neurohypophysis, and pineal gland facilitate the transport of neurohormones from the nerve terminals or nerve cell derivatives (pineal) into the bloodstream, whereas at the other circumventricular organs, neurohemal junctions facilitate the transport of neurohormones from the blood to nerve cells. The latter mechanism has been implicated in the cross-talk between peripheral organs and the brain so that, for example, the hypertensive effect of the peptide, angiotensin, is due in part to its activation of neurons of the subfornical organ, which have a high density of angiotensin receptors. In addition to their importance as sites for the transfer of chemical messengers from nerve terminals into blood vessels, and *vice versa*, the neurohemal junctions of circumventricular organs are also sites at which drugs or toxins, which are

normally excluded by the blood–brain barrier, can penetrate the brain and affect brain function.

The Hypothalamo-Adenohypophysial System

Neurohormonal Control of Anterior Pituitary Hormone Secretion

The transmission of signals between the brain and the anterior pituitary gland is mediated by chemical messengers called neurohormones, which are transported by the hypophysial portal vessels from the hypothalamus to the anterior pituitary gland (Figures 3–5) where they either stimulate or inhibit the release of anterior pituitary hormones. Synthesized in nerve cells of the hypothalamic nuclei, the neurohormones are released from nerve terminals into the plexus of primary capillaries of the hypophysial portal vessel system. These capillaries are derived from the superior hypophysial arteries and coalesce to form the hypophysial portal veins, which run on the

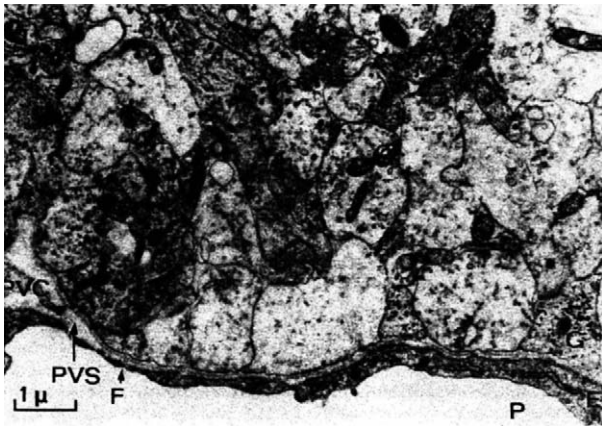


Figure 6 Electromicrograph of the external layer of the median eminence of a rat on the first postnatal day ($\times 13,200$). Note the high density of nerve terminals around part of a primary portal capillary vessel (P), which is fenestrated (F). Note also the large number of agranular and granular vesicles in the nerve terminals. These vesicles contain the packets (quanta) of neurohormone or neurotransmitter that are released on nerve depolarization as a consequence of nerve action potentials. The neurohormones are released into the perivascular space and from there move rapidly into portal vessel blood for transport to the pituitary gland. This arrangement is typical of neurohemal junctions found in the several circumventricular organs of the brain. E, endothelial cell; G, glial process; PVC, perivascular cell; PVS, perivascular space. Reproduced with permission from Fink, G. and Smith, G. C. (1971), Ultrastructural features of the developing hypothalamo-hypophysial axis in the rat, *Zeitschrift für Zellforschung und Mikroskopische Anatomie* **119**, 208–226.

surface or through the pituitary stalk to the anterior pituitary gland, where they form a secondary plexus of vessels called the pituitary sinusoids. Vessels on the surface of the stalk are the long portal vessels, whereas those within the substance of the stalk are the short portal vessels.

Hypophysial Portal Vessels

The portal vessels are so called because they transport chemical messengers from one capillary bed (primary capillaries) to a second capillary bed before entering the general circulation. In principle, this is identical to the hepatic portal system, which transports substances from the primary bed of capillaries in the intestine and its appendages (e.g., the pancreas) to a second bed of capillaries or sinusoids in the liver. Fenestration of the primary and the secondary (sinusoids) plexus of capillaries (Figure 6) facilitates the transport of substances across the capillary wall. Hormones released from anterior pituitary cells are transported by pituitary veins into the systemic circulation by which they are transported to their major target organs, the gonads and the adrenal and thyroid glands.

Hypothalamic Neurohormones

Outline Most of the neurohormones that mediate the neural control of anterior pituitary hormone secretion are peptides that are synthesized in discrete hypothalamic nuclei. Neurohormone-secreting neurons are the final common pathway neurons for neural control of the anterior pituitary gland, a term borrowed by G. W. Harris from Sherrington's description of the alpha-motor neurons of the spinal cord that innervate and control the contraction of skeletal muscles. Like the alpha-motor neurons, hypothalamic neurons are controlled by inputs to the hypothalamus from the brain stem, midbrain, and higher brain centers. The hypothalamic neuroendocrine neurons are thus connected with many other regions of the nervous system and, in particular, the components of the limbic system, which is involved in higher brain functions, including emotion, olfaction, and memory.

The neural control of all established anterior pituitary hormones is mediated by at least one or more neurohormones. In some cases, two neurohormones may act synergistically, as is the case of ACTH, the release of which is stimulated by both the 41-amino-acid-residue peptide CRH-41 and the nonapeptide AVP. The control of ACTH secretion is further complicated by the fact that urocortins, relatively new members of the CRH-41 family of peptides, act mainly by way of CRH type 2 receptors in the PVN to modulate AVP and CRH-41 synthesis. Hypophysial portal blood and immunoneutralization studies suggest that ACTH secretion may be inhibited by atrial natriuretic peptide (ANP).

In other cases two neurohormones may act antagonistically, as is the case for GH, whose release is stimulated by the 44-amino-acid-residue peptide GH-releasing hormone (GHRH-44) and inhibited by the 14- or 28-residue peptide somatostatin-14 or -28. The neural regulation of pituitary GH secretion is even more complex than originally thought; Bowers and associates discovered in 1981 that a hexapeptide, referred to as GH-releasing peptide-6 (GHRP-6), is a potent GH secretagogue. Subsequent studies showed that GHRP-6 binds to a receptor, GH secretagogue receptor (GHS-R), that is distinct from the GHRH-44 receptor. The endogenous ligand for the GHS-R is ghrelin, a 28-amino-acid-residue peptide secreted mainly by neuroendocrine cells of the gastric mucosa, which also stimulates GH release.

Prolactin seems to be the only anterior pituitary hormone that is predominantly under the inhibitory control of the brain. Evidence for this came first from the studies of Everett and Nikitowitch-Winer, who showed that prolactotrophs were the only cell type

that did not undergo atrophy in pituitary grafts under the kidney capsule, where they are far removed from central neural control. This histological observation was confirmed by the finding that prolactin concentrations in plasma are increased in animals bearing pituitary grafts under the kidney capsule. There is substantial evidence that dopamine, released into hypophyseal portal blood from tuberoinfundibular dopaminergic neurons, inhibits prolactin release. Indeed, dopamine agonists such as bromocriptine, cabergoline, and quinagolide are highly effective in treating hyperprolactinemia, a relatively common cause of infertility in women. Dopamine is now regarded as the prolactin inhibitory factor (PIF). Hyperprolactinemia is frequently due to benign tumors (adenomas) of the anterior pituitary gland, and these too can often be controlled by treatment with dopamine agonists. Stress is another common cause of hyperprolactinemia. Indeed, the prolactin response to stress in the human is as sensitive as that of ACTH and adrenal glucocorticoid response to stress.

The search for a prolactin-releasing factor (PRF) has continued for many years. Potential candidates have included TRH, vasoactive intestinal peptide, GnRH-associated peptide (GAP), and two novel, closely related hypothalamic peptides (termed PrRP31 and PrRP20) discovered in 1998 by Hinuma and associates when searching for a ligand for an orphan receptor (hGR3). Recent evidence suggests that the peptides intermedin and adrenomedullin may also release prolactin. However, although all these compounds are capable of stimulating prolactin release from pituitary cells *in vitro*, and in some cases under pharmacological conditions *in vivo*, robust proof that any of these compounds are the physiological PRF remains to be established.

The neural control of the gonadotropins, luteinizing hormone (LH) and follicle-stimulating hormone (FSH), is mediated by one and the same decapeptide, GnRH.

Criteria for neurohormones and neurotransmitters Werman, in 1972, summarized the criteria for classifying a candidate compound as a neurotransmitter or neurohormone thus: "If it can be shown that a substance is released into the extracellular space from presynaptic nerves in quantities consistent with the amount and rate of stimulation and the physiology of transmitter release at that junction, and if it can be shown that the material released acts on postsynaptic membranes by using molecular mechanisms identical with those used by the physiologically evoked transmitter, then that substance is a transmitter." Werman's two principal criteria – endogenous release (collectibility) and exogenous

mimicry (identity of action) – should be supplemented by a third criterion invoking the topochemistry of the mechanisms for the synthesis and inactivation of the transmitter.

The technique of collecting hypophyseal portal blood in the anesthetized rat (developed by Fink and Worthington in 1966) has made it possible to satisfy these criteria for all the hypothalamic-pituitary neurohormones outlined here, as well to clarify the mechanism and physiological significance of post-translational processing. The measurement of neurohormone release into the hypophyseal portal blood has also made it possible to ascertain whether newly discovered hypothalamic compounds could serve as hypothalamic-pituitary regulatory factors. This is exemplified by work in our laboratory, which showed that the concentrations of the cardiac peptide ANP were about four times greater in portal than in systemic blood, a finding that led to immunoneutralization studies that suggest that ANP is an ACTH-inhibiting factor. Studies of hypophyseal portal blood can also exclude the neurohormonal role of a candidate neurotransmitter. Thus, for example, although immunohistochemistry suggests that angiotensin and cholecystokinin are both present in the median eminence, their concentrations in portal blood are not greater than in peripheral blood, which makes it unlikely that either plays a role as a neurohormone.

All the hypothalamic-pituitary regulatory neurohormones are present in regions of the nervous system outside the hypothalamus, where Werman's criteria for a neurotransmitter need to be proved by determining whether the neurohormones are released (by push-pull cannulae or dialysis), activate cells (e.g., electrophysiologically), and are responsible for a behavioral effect. The functional significance of neurohormone action in areas of the CNS outside the hypothalamus is illustrated poignantly by CRH-41, which is believed to synchronize the endocrine, autonomic, immunological, and behavioral responses to stress.

Neurohormone receptors The hypothalamic neurohormones and their receptors, all members of the G-protein-coupled, seven transmembrane superfamily, underscore the adage that the evolution of hormones is less dependent on structural changes in the hormones than on the uses to which they are put. The latter depends entirely on the receptors and the signal transduction systems that mediate receptor activation, a point illustrated by the receptors that mediate the neurohormone control of ACTH and GH secretion.

The actions of CRH-41 in brain and periphery are mediated by two main receptor subtypes, CRH1 and

CRH2, with splice variants CRH2 α and CRH2 β . A third CRH receptor has recently been discovered in catfish pituitary. cAMP-linked CRH receptors also mediate the action of the urocortins, the mammalian homologs of urotensin in fish, and of sauvagine, a homolog of CRH-41 derived from the South American tree frog. In addition to mediating the neural control of ACTH release, CRH-41 and the urocortins exert actions on the brain and the cardiovascular system. CRH1 receptors (1) are localized predominantly in brain and pituitary, (2) have a high affinity for CRH-41, (3) are expressed by pituitary corticotrophs, and (4) are responsible for mediating the stress-induced CRH-41 stimulation of ACTH release. CRH2 receptors have a greater affinity for the urocortins than for CRH-41 and are present mainly in heart and skeletal muscle. However, the CRH2 α splice variant is also present in brain and pituitary gland, where it is expressed by pituitary gonadotrophs, but not corticotrophs, leading to the suggestion that CRH2 α may mediate stress effects on gonadotropin secretion. The ACTH-releasing action of CRH-41 is potentiated by AVP, which binds to the V $_{1\beta}$ receptor. The V $_{1\beta}$ receptor activates the inositol phospholipid cycle. Synergism between CRH-41 and AVP involves calcium-activated cAMP production. ACTH secretion is inhibited by ANP, an action mediated by two guanylyl cyclase-coupled receptor subtypes: A and B.

Human GHRH receptor cDNA was first cloned and sequenced in Mike Thorner's laboratory in 1993 and was found to be a member of a family of receptors that includes secretin, calcitonin, vasoactive intestinal peptide, and parathyroid hormone. The activity of the GHRH receptor (GHRH-R) promoter is enhanced by the pituitary-specific transcription factor POU1FI (formerly Pit-1). The importance of POU1FI, the prototypic POU-homeodomain transcription factor, for the development and activity of the GHRH-R is demonstrated by the fact that the GHRH-R is not present in the pituitary gland of dwarf mice (*dw/dw*) that lack POU1FI. Glucocorticoids stimulate, whereas estrogen inhibits, GHRH-R gene expression, actions that are likely to be mediated by the glucocorticoid and estrogen response elements present in the promoter and that explain in part the physiological effects of glucocorticoids and estrogen on GH secretion. The pituitary expression of GHRH-R mRNA in the rat increases with embryonic development to reach a peak at E19.5 (2 days before parturition), followed by a decline through to postnatal day 12 and then an increase to day 30, followed by a decline with age. If a similar pattern occurs at the corresponding times in humans, then that could be a major factor in the decline in plasma GH

concentrations with age in the elderly. The GHRH-R is distinct from the GHS-R whose ligand is the 28-amino-acid-residue gastric peptide ghrelin. Both types of receptor are members of the seven-transmembrane helix, G-protein-coupled receptor family.

There are six receptors for somatostatin (SSTR), encoded by five distinct genes. All six receptors are present in the anterior pituitary gland; SSTR-5 is the most prominent on the somatotrophs, followed by SSTR-2 (which has splice variants A and B). Hypothalamic arcuate neurons that secrete GHRH express SSTR-2 and SSTR-1, supporting the proposition that the action of somatostatin may involve an effect on the release of GHRH. Relatively large amounts of somatostatin are also present in the cerebral cortex and hippocampus, regions that express the SSTR-2A. In the periphery, somatostatin secreted by pancreatic islet cells inhibits the secretion of insulin, glucagon, and a range of gut peptides.

The complexity of the receptors involved in the control of ACTH and GH secretion contrasts with the apparent simplicity of the gonadotropin control system, in which GnRH mediates the neural control of both LH and FSH, and the apparent simplicity of the thyrotropin and prolactin control systems, in which there is evidence, so far, for only one neurohormone each, TRH and the PIF dopamine, respectively. Dopamine inhibition of prolactin release is mediated by the dopamine 2 (D2) receptor. Alternative splicing results in a long and short form of the D2 receptor, and this allows coupling to different types of G-protein. When activated, the D2 receptor inhibits adenylyl cyclase.

Whether the complexity of the neurohormones and receptors involved in the neural control of GH and ACTH release is a consequence of evolutionary convergence of different regulatory systems that are not deleterious and are therefore retained or whether complexity has developed as a consequence of selective pressure to provide flexibility or increase the number of variables that can influence the two systems remains to be determined.

Clinical relevance All the hypothalamic neurohormones have been used extensively as tests of anterior pituitary function, or pituitary responsiveness. In addition, neurohormone agonists and antagonists have been deployed for a variety of diagnostic tests and therapeutic regimes. Thus, analogs of CRH-41 and urocortin have been developed with a view to their possible use in the treatment of depression, anxiety, anorexia nervosa, and stroke. The use of dopamine agonists, such as bromocriptine, cabergoline, and quinagolide, for the treatment of hyperprolactinemia and prolactin-secreting pituitary adenomata has been

mentioned previously. GnRH superactive agonists and antagonists are used for chemical castration, as an adjunct treatment of cancer of the breast and prostate gland, for fertility control, and for the treatment of infertility and precocious puberty. Somatostatin, or its potent synthetic analogs such as Octreotide, has been used in the treatment of GH-secreting pituitary adenomas that lead to acromegaly and also ectopic GH-secreting tumors, most commonly small-cell carcinomas of the lung or tumors of the gastrointestinal system. The action of Octreotide in treating GH-secreting tumors can be potentiated by dopamine agonists such as bromocriptine and cabergoline.

Somatostatin and its agonists are also effective in the treatment of pancreatitis. Orally active, nonpeptide mimetics of GHRP-6, such as MK-677, have been developed and these may have therapeutic value in the treatment of children of short stature and adult GH deficiency. GHRH, GHRP-6, and its mimetics are also being explored as useful agents for the treatment of metabolic deficiencies that occur in aging.

Teleology of Neurohormonal Control

The hypothalamic-pituitary axis illustrates the remarkable economy of physiological systems. First, and perhaps most impressive, are the hypophyseal portal vessels, which, by transporting the neurohormones from the hypothalamus to the pituitary gland, undiluted by mixture in the systemic circulation, ensure that hypothalamic neurohormones released in very small amounts from the hypothalamus reach the pituitary gland at concentrations several orders of magnitude greater than in the systemic circulation and therefore sufficient to exert their effects. The corollary of this is that relatively little neurohormone needs to be released to exert its effect and that therefore only a small amount of new neurohormone needs to be synthesized. The metabolic economy of the hypophyseal portal system is brought into sharp relief by the fact that the hypothalamic concentration and total content of the hypothalamic-anterior pituitary regulatory neurohormones are three or more orders of magnitude lower than those of the neurohypophyseal nonapeptides AVP and oxytocin, which reach their peripheral targets by the systemic circulation.

Second, the transport of neurohormones at effective concentrations by the hypophyseal portal vessels also protects the body from the potential adverse effects of the high concentrations of neurohormones necessary to stimulate or inhibit pituitary hormone secretion. Thus, for example, the high portal blood concentrations of somatostatin may, in the systemic

circulation, have adverse effects on the gut and on insulin secretion by the pancreatic islets. Similarly, as shown by Fink and associates, the portal plasma concentrations of ANP that inhibit ACTH secretion would, in the systemic circulation, cause a lethal drop in blood pressure.

Third, neurohormones in the hypothalamo-adenohypophyseal system, by also serving as chemical messengers in several related systems or mechanisms, may serve to orchestrate appropriate physiological responses. Thus, for example, somatostatin, in addition to inhibiting GH secretion is also secreted by pancreatic islets and inhibits the secretion of insulin, glucagon, and a range of gut peptides. As mentioned earlier, CRH-41 is present in higher brain centers and in the periphery, and it has been implicated in stress-related behaviors, cardiovascular control, and the synchronization of the endocrine, autonomic, and immunological components of the stress response.

The Hypothalamo-Neurohypophyseal System

Principles and Overview

The neural lobe is the site of release of the nonapeptides oxytocin and AVP into the systemic circulation. Presumably because a considerable distance removes the target cells of these peptides from their site of release in the neural lobe and because there is massive dilution in the systemic circulation, the rate and the amount of synthesis of oxytocin and AVP and their content in the hypothalamus are about three orders of magnitude greater than that of neurohormones concerned in anterior pituitary control. This fact, together with their synthesis in discrete hypothalamic nuclei, their ease of assay, and that both nonapeptides contain disulfide bridges that allow the incorporation of the radioactive tracer [³⁵S]-cysteine, contributed to four important landmarks in neuroendocrinology. First, oxytocin and AVP were the first of the hypothalamic neurohormones to be sequenced (by Du Vigneaud and colleagues). Second, the glycoprotein components of their precursor proteins led to their conspicuous histochemical staining and thereby to the concept of neurosecretion (the Scharrers), a term applied to neurons whose main purpose seemed to be to secrete neurohormones rather than transmit signals by propagated action potentials. Third, the fact that both peptides have disulfide bridges and that synthesis involves the incorporation of cysteine, which can be labeled with the relatively high-energy radionuclide ³⁵S, made them excellent models for the first studies of neuropeptide synthesis and transport (first shown by Sachs and associates). Fourth,

oxytocin- and AVP-containing neurons were the first neuroendocrine neurons from which electrophysiological recordings were made (by J. Green and B. Cross in 1959).

The concept of neurosecretion was soon discarded because it was shown that the magnocellular neurons are no different from others – they propagate action potentials and release neurohormones or neurotransmitters. In fact, oxytocin- and AVP-containing neurons facilitate electrophysiological recording because they are large and because their axons terminate in the surgically accessible neural lobe. This makes it relatively easy to locate and verify by antidromic stimulation the identity of the neurons and record from them. Furthermore, the electrophysiological activity of the hypothalamic magnocellular neurons can be correlated with the secretion of AVP and oxytocin.

Finally, because the neural lobe is largely made up of nerve terminals, it proved to be a useful model for the study of exocytosis and stimulus-secretion coupling, the coupling between the cascade of ion fluxes through membrane channels (membrane depolarization) triggered by action potentials and the calcium-dependent release of neurohormones or neurotransmitters. Neurohormones, like all other known neurotransmitters and peptide and protein hormones, are packaged in secretory vesicles (Figure 6), which are released in packets or quanta.

Function

AVP (or the antidiuretic hormone) is concerned mainly with the control of body water. As its name implies, AVP also induces vasoconstriction and thereby can increase blood pressure, although this mechanism usually becomes operational only after a substantial loss of blood volume. AVP cells of the SON and PVN respond to osmotic stimuli. An increase in plasma osmolality triggers the release of AVP, which increases water reuptake in the nephron, which as a consequence causes an overall increase in body water with a fall in plasma osmolality. This is a perfect example of a homeostatic regulatory mechanism. In addition to its role in osmoregulation, AVP synthesized in the smaller (parvocellular) neurons of the PVN acts synergistically with CRH-41 to release ACTH.

Oxytocin, concerned mainly with milk ejection during lactation and parturition (the birth process), provides the perfect example of the role of neurohormones in neuroendocrine reflexes. Thus, oxytocin is the neurohormonal component of the milk-ejection reflex whereby suckling at the nipple of lactating mothers triggers volleys of impulses that travel through the mammary nerves to the spinal cord and,

by way of a multisynaptic pathway, reach the hypothalamus where they trigger the release of oxytocin. Oxytocin transported by the systemic circulation stimulates the contraction of the myoepithelial cells of the breast, resulting in milk ejection. During parturition, oxytocin coordinates and reinforces uterine contractions. That is, as uterine contractions force the head of the fetus against the cervix, volleys of impulses are triggered that ascend through multisynaptic pathways involving the pelvic nerves and the spinal cord to the hypothalamus to trigger the release of oxytocin, which acts on the smooth muscle cells of the uterus. This is a classical positive feedback system, which is terminated by expulsion of the fetus and placenta.

The Intermediate Lobe of the Pituitary Gland

Model for Posttranslational Processing

The major secretion of the intermediate lobe of the pituitary gland is α -melanocyte-stimulating hormone (α -MSH), a 13-amino-acid-residue peptide that, together with ACTH and β -endorphin, is derived from the precursor proopiomelanocortin (POMC). In addition to α -MSH, the pars intermedia also contains other derivatives of POMC – β -MSH, γ -MSH, corticotropin-like intermediate lobe peptide (CLIP), and β -endorphin – which together with the ACTH make up the melanocortins. The fact that in the pars distalis ACTH is the major hormone derived from posttranslational processing of POMC, whereas in the pars intermedia α -MSH is the major active hormonal product of POMC processing, reflects the presence of different enzymatic processing pathways in the two parts of the gland. The release of α -MSH is inhibited by dopaminergic neurons that originate in the arcuate nucleus and reach the intermediate lobe by way of the neural stalk. Humans are conspicuous among mammals in that the pars intermedia is not defined as a separate lobe of the human pituitary gland.

Melanocortins

Although they are all derived from POMC, melanocortins have diverse physiological functions. α -MSH, first thought to stimulate the growth of melanocytes and pigment formation (melanogenesis), is also involved in fever, inflammation, and the immune response. Indeed, several of the melanocortins, including ACTH, affect cytokine-induced effects on thymocytes, T and B lymphocytes, and neutrophils. The actions of the melanocortins are mediated by five melanocortin receptor subtypes (MC1–5). Although

all the melanocortin receptors are seven-transmembrane G-protein-coupled, the tissue distribution of the five receptors is quite distinct. Thus, the MC1 is located on melanocytes and mediates MSH-stimulated melanocyte proliferation and melanogenesis, while the MC2 receptor is located in the adrenal cortex, where it mediates the ACTH-induced secretion of glucocorticoids and mineralocorticoids. The MC3 and MC4 receptors are both located in the brain; the function of MC3 is unknown, but MC4 is involved, together with leptin (a protein produced by adipocytes) and neuropeptide Y, in the regulation of body weight. Studies on mutant *ob/ob* mice show that the action of leptin in regulating body weight is mediated, at least in part, by hypothalamic melanocortin activation, which results in decreased food intake. The importance of hypothalamic melanocortin activity for moderating food intake is illustrated by the agouti peptide, which is a high-affinity antagonist for the MC1 receptor and induces obesity in the mouse. The MC4 receptor is also involved in grooming behavior. The MC5 receptor, whose function remains unknown, is widely distributed in body organs, including the spleen, thymus, skin, and bone marrow. Cutaneous pigmentation is a function of the MC1 receptor, and MC1 gene variants may predispose to red hair and light skin color, which may make the individual susceptible to melanoma.

Pineal Gland: Melatonin and Photoperiodic Control of Circadian Rhythms

The pineal gland is a circumventricular organ (Figure 1) that secretes melatonin into the circulation and plays an important role in the photoperiodic control of reproduction in seasonal breeding animals. The gland deserves special mention here because it reinforces the principles of neuroendocrine control. Philosophers (e.g., Descartes, who called this the seat of the soul) and scientists have long been intrigued by the pineal. The secretion of melatonin is exquisitely sensitive to light. Pinealocytes in submammalian species are photoreceptors, and the gland offers an excellent experimental model for studies of the transduction of light into nerve impulses and neurohormone secretion.

The outer segment (sensory pole) of the pinealocyte in fish, amphibia, and reptiles has all the ultrastructural characteristics of a true photoreceptor, but these features are only vestigial in mammals and intermediate forms exist in birds. In fish, amphibia, and reptiles, the effector pole of the pinealocyte synapses with secondary pineal neurons that give rise to the pineal tract, which transmits signals to the CNS.

In birds and mammals, however, pinealocytes secrete melatonin directly into the circulation (or cerebrospinal fluid) in a neuroendocrine manner.

Melatonin, a derivative of serotonin, is synthesized within pinealocytes in two steps. First, serotonin is converted by *N*-acetyl-transferase (NAT) to *N*-acetyl serotonin, which is then converted to melatonin by hydroxyindole-*O*-methyltransferase. Because little if any melatonin is stored, the rate of melatonin secretion is tightly linked to its synthesis, which depends on NAT action, which in turn depends on norepinephrine release from the dense sympathetic innervation of the gland. In mammals, the control of melatonin secretion by light is mediated by a multisynaptic pathway that starts in the retina of the eye and successively involves synapses in the suprachiasmatic nuclei, the PVN, the intermediolateral column of the spinal cord, and the neurons of the superior cervical ganglion of the sympathetic nervous system. Sympathetic terminals in the pineal gland release norepinephrine, which stimulates melatonin secretion by an action on adrenoreceptors on pinealocytes. cAMP seems to be the main intracellular second messenger that mediates the action of norepinephrine in this system. The main point of regulation seems to be NAT, which, depending on the species, can be affected at the level of NAT gene expression as well by posttranslational modifications that alter the activity of the enzyme.

The secretion of melatonin starts with the onset of the dark period (night) and stops with the onset of the light period (day). The secretion of melatonin during the dark period is stopped abruptly by exposure to light. In blind people, the secretion of melatonin takes on the typical 25-h free-running period. Taken together, these and other data suggest that the secretion of melatonin is predominantly controlled by light exposure superimposed on the intrinsic rhythm of the major neural clock, the suprachiasmatic nuclei.

An interaction between melatonin and the hypothalamic-pituitary-adrenal stress response system has been suggested; indeed, melatonin has been implicated in glucocorticoid receptor-induced gene expression. However, the precise role of melatonin in the circadian rhythm of ACTH and adrenal corticosteroids, and in the stress response, has yet to be established. This applies equally to the long-researched but unanswered questions as to whether melatonin is a hypnotic or sedative and whether it can resynchronize circadian rhythms that have been disrupted by rapidly crossing time zones (jet lag). Melatonin is a powerful antioxidant, which may explain some of its alleged beneficial effects with respect to aging. So far, the most robust evidence regarding the function of melatonin in mammals relates to its inhibitory actions with respect to reproduction, especially in

seasonal breeding animals such as the wallaby, hamster, vole, and sheep.

What is well established is that the suprachiasmatic nuclei constitute the central generator of circadian rhythms of the body, with one notable exception – the daily anticipation of a meal. That is, rats in which the suprachiasmatic nuclei have been lesioned are still able to anticipate one meal per day. This anticipation is associated with an increase in activity, core body temperature, and plasma corticosterone concentrations that occur 2–4 h before access to the meal. The precise nature and location of the feeding-entrained oscillator that presumably determines the daily rhythm of feeding anticipation have yet to be established. However, studies with *c-Fos* (early immediate gene) activation have implicated the nucleus accumbens and the limbic system. The relative importance of the pineal gland and its precise role in the photoperiodic control of circadian rhythms also wait to be determined. With the discovery of genes that regulate circadian rhythms and melatonin synthesis, the prospects of defining what is likely to be a remarkably elegant and robust molecular control system seem excellent.

Hormonal Effects on the Nervous System

The steroid and thyroid hormones, the secretions of the three major pituitary target organs, exert powerful effects on the brain. These effects may be classified in terms of (1) feedback actions, (2) brain differentiation and neural plasticity, (3) neurotransmission, and (4) membranes and ion channels. Feedback actions of the corticosteroids on the brain and pituitary gland are discussed in detail elsewhere. Many entries also deal with the specific effects of corticosteroids on brain that are mediated by corticosteroid receptors, but that are not necessarily related to the feedback control of ACTH release. For information on the actions of steroids on brain differentiation, neurotransmission, cell membranes, and ion channels, the reader is referred to the Bibliography.

Psychoneuroendocrinology

The recognition that the secretion of pituitary hormones reflects the activity of hypothalamic neurons, which in turn reflects neurotransmitter activity in the brain, provides the basis for the use of pituitary hormone secretion as a marker of disordered central neurotransmission that may underlie or be associated with mental disorders. The development of this discipline, psychoneuroendocrinology, has been impeded by our lack of fundamental knowledge about the precise neurotransmitter regulation of hypothalamic

neurons that regulate the anterior pituitary gland. Nonetheless, several robust findings have been made and these provide important clues for further research that might optimally be carried out in conjunction with human genetics and sophisticated brain imaging.

For example, it has long been known that plasma cortisol concentrations are abnormally high in patients with psychosis, especially severe depression. Furthermore, as shown by the groundbreaking work of Carroll and Rubin and their associates, in many psychotic patients it is not possible to suppress these high cortisol concentrations by injecting a glucocorticoid agonist such as dexamethasone (the dexamethasone suppression test). Although it cannot accurately discriminate among the different types of psychoses, the failure of dexamethasone to suppress cortisol levels provides robust biological evidence that a patient is suffering from a severe psychosis. The inability of dexamethasone to suppress cortisol levels in many patients with depressive psychosis points to a significant abnormality in the regulatory mechanism for ACTH release. The elucidation of the precise nature of this abnormality may help explain the central neurotransmitter dysfunction that leads to severe depression or other psychosis.

The growth hormone and prolactin responses to pharmacological challenge are also abnormal in certain psychotic states. In anorexia nervosa, for example, a life-threatening condition that occurs predominantly in young women, the normal GH surge that can be induced by the infusion of L-tryptophan, the precursor for serotonin synthesis, is completely absent. This cannot simply be attributed to a lack of food intake because the GH response to L-tryptophan infusion is enhanced, rather than suppressed, in women of a similar age who have been placed on a low-calorie diet. Again the precise explanation for this phenomenon, first shown by Guy Goodwin and associates, is not established, but points to an abnormality in central serotonin function that may underlie or be associated with anorexia nervosa.

The advantages of the psychoneuroendocrine investigation of mental disorders are that it is relatively inexpensive, only mildly invasive, and can be carried out in severely ill patients. Its power will depend on further fundamental research that may enable us to interpret accurately the neuroendocrine responses to pharmacological challenge.

See Also the Following Articles

Adrenal Medulla; Circadian Rhythms, Genetics of; Feedback Systems; Food Shift Effect; Primate Hierarchies and Personality; Reproductive Dysfunction in Primates, Behaviorally Induced; Secretagogue; Urocortins; Corticotropin

Releasing Factor (CRF); Corticotropin Releasing Factor-Binding Protein; Corticotropin-Releasing Factor Receptors; Anti-CRF; CRH Anxiety Circuitry in Brain; Dex-CRH Test; Corticotropin Releasing Factor Receptor Deficiency in Mice; Corticotropin-releasing factor (CRF) Family of Neuropeptides – Role in Inflammation.

Further Reading

- Bale, T. L. and Vale, W. W. (2004). CRF and CRF receptors: role in stress responsivity and other behaviors. *Annual Reviews of Pharmacology and Toxicology* **44**, 525–557.
- Balthasar, N., Dalgaard, L. T., Lee, C. E., et al. (2005). Divergence of melanocortin pathways in the control of food intake and energy expenditure. *Cell* **123**(3), 493–505.
- Boorse, G. C. and Denver, R. J. (2006). Widespread tissue distribution and diverse functions of corticotropin-releasing factor and related peptides. *General Comparative Endocrinology* **146**(1), 9–18.
- Cone, R. D. (2005). Anatomy and regulation of the central melanocortin system. *Nature Neuroscience* **8**(5), 571–578.
- Ellacott, K. L., Halatchev, I. G. and Cone, R. D. (2006). Characterization of leptin responsive neurons in the caudal brainstem. *Endocrinology* **147**(7), 3190–3196.
- Fink, G. (1988). The G. W. Harris Lecture: Steroid control of brain and pituitary function. *Quarterly Journal of Experimental Physiology* **73**, 257–293.
- Fink, G. (1997). Mechanisms of negative and positive feedback of steroids in the hypothalamic-pituitary system. In: Bittar, E. & Bittar, N. (eds.) *Principles of medical biology*, pp. 29–100. Greenwich, CT: JAI.
- Fink, G., Dow, R. C., Casley, D., et al. (1992). Atrial natriuretic peptide is involved in the ACTH response to stress and glucocorticoid negative feedback in the rat. *Journal of Endocrinology* **135**, 37–43.
- Fink, G. and Smith, G. C. (1971). Ultrastructural features of the developing hypothalamo-hypophysial axis in the rat. *Zeitschrift für Zellforschung und Mikroskopische Anatomie* **119**, 208–226.
- Fink, G., Sumner, B., Rosie, R., et al. (1999). Androgen actions on central serotonin neurotransmission: relevance for mood, mental state and memory. *Behavioral Brain Research* **105**, 53–68.
- Ganguly, S., Weller, J. L., Ho, A., et al. (2005). Melatonin synthesis: 14-3-3-dependent activation and inhibition of arylalkylamine N-acetyltransferase mediated by phosphoserine-205. *Proceedings of the National Academy of Science USA* **102**(4), 1222–1227.
- Getting, S. J. (2006). Targeting melanocortin receptors as potential novel therapeutics. *Pharmacology and Therapeutics* **111**(1), 1–15.
- Goodwin, G. M., Shapiro, C. M., Bennie, J., et al. (1989). The neuroendocrine responses and psychological effects of infusion of L-tryptophan in anorexia nervosa. *Psychological Medicine* **19**, 857–864.
- Guillemet-Guibert, J., Lahlou, H., Cordelier, P., et al. (2005). Physiology of somatostatin receptors. *Journal of Endocrinological Investigation* **28**(11 supplement), 5–9.
- Hauger, R. L., Grigoriadis, D. E., Dallman, M. F., et al. (2003). International Union of Pharmacology. 36: Current status of the nomenclature for receptors for corticotropin-releasing factor and their ligands. *Pharmacological Reviews* **55**(1), 21–26.
- Larsen, P. R., Kronenberg, H. M., Melmed, S. and Polonsky, K. S. (eds.) (2003). *Williams textbook of endocrinology* (10th edn.) Philadelphia: Saunders.
- Luque, R. M., Kineman, R. D., Park, S., et al. (2004). Homologous and heterologous regulation of pituitary receptors for ghrelin and growth hormone-releasing hormone. *Endocrinology* **145**(7), 3182–3189.
- Mayo, K. E., Miller, L. J., Bataille, D., et al. (2003). International Union of Pharmacology. 35: The glucagon receptor family. *Pharmacological Reviews* **55**(1), 167–194.
- Mendoza, J., Angeles-Castellanos, M. and Escobar, C. (2005). Entrainment by a palatable meal induces food-anticipatory activity and c-Fos expression in reward-related areas of the brain. *Neuroscience* **133**(1), 293–303.
- Schumacher, M., Weill-Engerer, S., Liere, P., et al. (2003). Steroid hormones and neurosteroids in normal and pathological aging of the nervous system. *Progress in Neurobiology* **71**(1), 3–29.
- Turek, F. W., Joshi, C., Kohsaka, A., et al. (2005). Obesity and metabolic syndrome in circadian clock mutant mice. *Science* **308**(5724), 1043–1045.
- Werman, R. (1972) CNS cellular level: membranes. *Annual Review of Physiology* **34**, 337–374.
- Yoo, S. H., Ko, C. H., Lowrey, P. L., et al. (2005). A noncanonical E-box enhancer drives mouse Period2 circadian oscillations in vivo. *Proceedings of the National Academy of Science USA* **102**(7), 2608–2613.

Neurogenesis

P Tanapat and E Gould

Princeton University, Princeton, NJ, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 31–36, © 2000, Elsevier Inc.

Development of the Dentate Gyrus

Stress Effects on Neurogenesis

Adrenal Steroids Affect Neurogenesis through NMDA Receptors

Functional Implications of Stress Effects on Neurogenesis

Glossary

Bromo-deoxyuridine (BrdU) labeling

A nonradioactive method for labeling proliferating cells and their progeny. BrdU is a thymidine analog that is incorporated into cells that are in the DNA synthetic phase of the cell cycle. BrdU incorporation can be visualized using immunocytochemical methods and can be combined with stereological analyses to estimate the total numbers of new cells generated at a given time. Short survival times (e.g., 2 h) can be used to label proliferating cells. When combined with the use of immunohistological markers, longer survival times (e.g., several weeks) can be used to determine the cellular phenotype of the newly generated cells.

Dentate gyrus

A brain region that is part of the hippocampal formation. It consists of the following three major components: (1) the granule cell layer, densely packed with granule cell bodies; (2) the hilus, a region that contains the granule cell axons (mossy fibers), interneurons, and progenitor cells; and (3) the molecular layer, a region that contains the granule cell dendritic trees and interneurons.

Neurogenesis

The process of forming a new neuron. Neurogenesis involves the proliferation of neuronal progenitors and the subsequent differentiation of these cells into neurons.

[³H]Thymidine autoradiography

A method that labels proliferating cells and their progeny with radio-labeled thymidine, a molecule that is incorporated into cells that are in S phase. Cells that have incorporated [³H]thymidine can be visualized with autoradiography. However, because [³H]

thymidine autoradiography allows the detection only of labeled cells in the top 3 μm of a section, this technique cannot be used for stereological analyses.

Most neurons in the mammalian brain are produced during the embryonic period. In contrast, the granule cell population of the dentate gyrus of the hippocampal formation is produced during an extended period that begins during gestation and continues well into adulthood. Previously, it was shown that the production of new granule neurons is suppressed by stress. Taken together, these observations suggest that stressful experiences during postnatal development and adulthood have the potential to significantly alter both the structure and function of the hippocampal formation.

Development of the Dentate Gyrus

In most regions of the mammalian brain, the majority of neurons are produced during a discrete part of the embryonic period. In contrast, granule neurons in the dentate gyrus of the hippocampal formation are produced primarily during the postnatal period and throughout adulthood in mammalian species ranging from rodents to primates, including humans. During the embryonic period, the first granule neurons arise from progenitor cells located in the wall of the lateral ventricle that migrate across the hippocampal rudiment to reside in the developing granule cell layer (Figure 1a). A subpopulation of these progenitor cells does not undergo final division, is maintained in the hilus and subgranular zone of the dentate gyrus, and continues to give rise to new granule neurons. In the rodent, the majority of neurons in the dentate gyrus are produced during the first 2 postnatal weeks from this secondary proliferative zone; cells in the hilus and subgranular zone divide, migrate along radial glial fibers, and ultimately reside in the granule cell layer (Figure 1b). In addition, some granule neurons are also produced from cells within the granule cell layer during this period. Following the second postnatal week, cell proliferation and migration decrease and the granule cell layer is formed. In the primate, the formation of the granule cell layer occurs during the prenatal period, but the production of new granule neurons continues postnatally as well.

In both rodents and primates, new neurons are produced throughout adulthood from progenitor cells

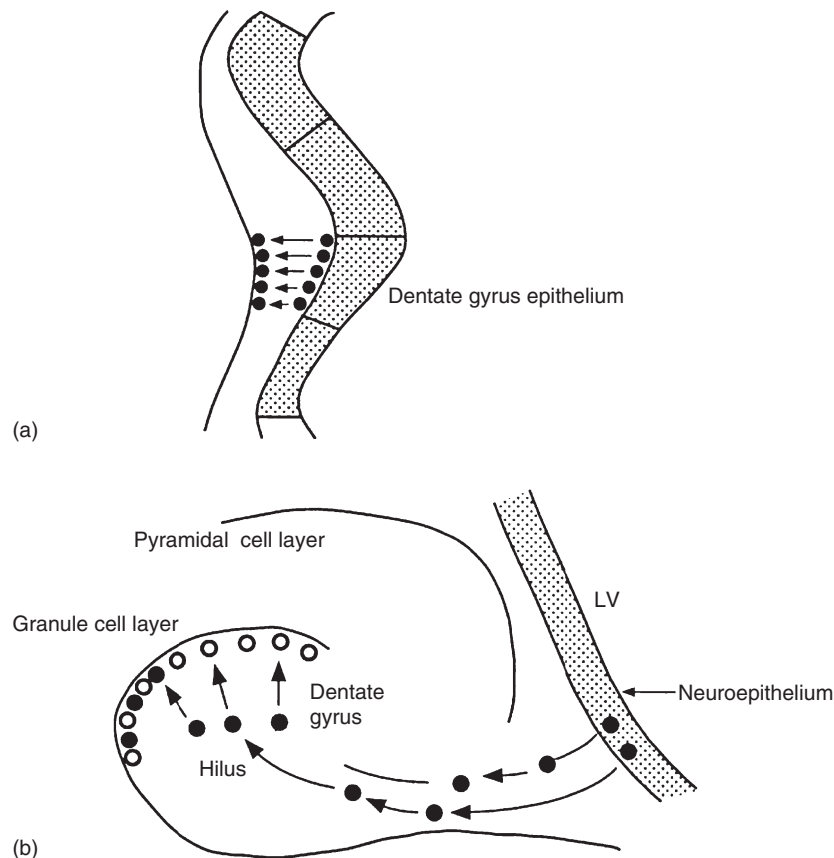


Figure 1 Development of granule cells. a, Embryonic period; b, late embryonic period. During the embryonic period, the first granule neurons arise from progenitor cells located in the wall of the lateral ventricle that migrate across the hippocampal rudiment to reside in the developing granule cell layer. During the late embryonic period and early postnatal development, granule cell precursors in the hilus divide and migrate to the granule cell layer. At the same time, precursor cells in the lateral ventricle continue to migrate into the dentate gyrus. The stippled area represents dentate gyrus neuroepithelium of the lateral ventricle (LV); solid circles represent migrating granule neuron precursors; open circles represent postmitotic immature granule neurons.

located in the hilus and subgranular zone (**Figure 2**). Studies using [^3H]thymidine or bromodeoxyuridine (BrdU) molecules, which are incorporated by cells during S phase, and various histological markers of cell phenotype have revealed that cells divide, migrate, extend axons, become incorporated into the granule cell layer, and ultimately express the morphological and biochemical characteristics of mature granule neurons. Although the numbers of new cells produced are much fewer than those observed during development, stereological studies have demonstrated that several thousand new cells are produced each day in the adult rodent brain. Because the majority of neurons in the dentate gyrus are produced postnatally, this brain region is particularly sensitive to changes in the environment during both postnatal development and adulthood compared to other brain regions. Thus, factors that regulate the production of granule neurons are capable of exerting significant effects on both the structure and the function of the hippocampal

formation, which may extend throughout the life of the organism. The remaining part of this article focuses on the regulation of granule neuron production by stress, the factors likely to mediate this regulation, and the possible functional implications of stress-induced changes in postnatal neuron production.

Stress Effects on Neurogenesis

During the first 2 postnatal weeks, termed the stress hyporesponsive period, rat pups exhibit a diminished response to most stressors that are known to increase circulating levels of adrenal steroids in adulthood. However, male rat pups do exhibit an increase in circulating corticosterone levels following acute exposure to the odor of an unfamiliar adult male (adult males are a natural predator of rat pups). Using this paradigm, it has been shown that a single exposure to the odors of an unfamiliar adult male during the stress hyporesponsive period results in a significant

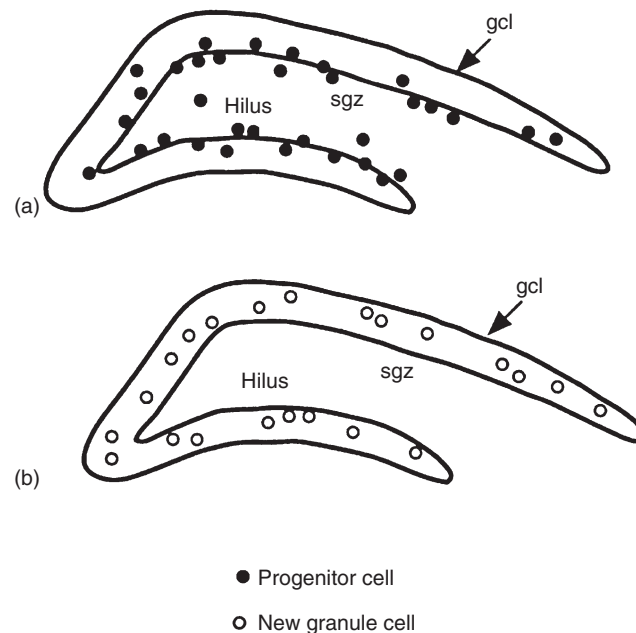


Figure 2 Production of new neurons. a, Progenitor cells; b, incorporation into granule cells. In both rodents and primates, new neurons are produced throughout adulthood from progenitor cells located in the hilus and subgranular zone (sgz). Studies using [^3H]thymidine or bromodeoxyuridine, both of which are incorporated by cells during S phase, and various histological markers of cell phenotype have revealed that these cells become incorporated into the granule cell layer (gcl) and ultimately express the morphological and biochemical characteristics of mature granule neurons.

decrease in the proliferation of granule cell precursors in the developing dentate gyrus 24 h following exposure.

Acute stressful experiences also decrease the number of new granule neurons produced during adulthood in a number of different species, including rats, tree shrews (animals considered to be phylogenetically between insectivores and primates), and marmosets. In adult rats, exposure to the odors of natural predators, such as foxes or weasels, elicits a stress response characterized by increased serum corticosterone levels and a characteristic electrophysiological response in the dentate gyrus. A single brief exposure to trimethyl thiazoline, a component of fox feces, rapidly suppresses the proliferation of cells in the dentate gyrus.

Similar observations have been made in studies that have examined the effect of social stress in nonrodent species. In tree shrews, the introduction of two adult same-sex conspecifics immediately results in the establishment of an enduring dominant-subordinate relationship. During this time, the subordinate exhibits a robust stress response characterized by elevated cortisol levels, increased heart rate, hypoactivity, and vigilance. Although able to recover physically on removal from the presence of the dominant, the subordinate exhibits an immediate stress response during subsequent encounters with the dominant. A single

brief exposure to a dominant animal, without physical contact, results in a rapid decrease in the number of new cells produced in the dentate gyrus of subordinate tree shrews compared to control animals. Likewise, studies of marmosets using a resident intruder paradigm, a model of psychosocial stress similar to the dominant-subordinate paradigm used with tree shrews, have yielded similar results. In this paradigm, a male marmoset is introduced into the cage of another male. Immediately after being placed in the cage, the intruder exhibits a stress response characterized by elevated mean arterial pressure and a dramatically increased heart rate. A single brief exposure to a psychosocial stress in the resident-intruder paradigm results in a rapid decrease in the number of new granule neurons produced in the dentate gyrus compared to naïve control animals.

In addition to exerting an acute impact on granule cell production, stress has also been shown to produce prolonged suppression of this process with repeated exposure. It has been demonstrated that 28 days of exposure to a dominant animal for 1 h each day results in a chronic increase in cortisol levels in adult tree shrews. Furthermore, exposure to chronic stress using this paradigm also results in a decrease in new cell production in the dentate gyrus, which is associated with a decrease in the total volume of the granule cell layer.

Adrenal Steroids Affect Neurogenesis through NMDA Receptors

Although the factors that mediate the stress-induced suppression of granule cell production remain unknown, adrenal steroids are likely to play a significant role in this effect. Several lines of evidence have indicated that adrenal steroids naturally regulate the production of hippocampal granule neurons. First, throughout life, a negative correlation between circulating adrenal steroid levels and the rate of granule neuron production exists. During the late embryonic period, adrenal steroid levels are high and few granule cells are produced. Following parturition, basal levels of glucocorticoids are low during the stress hyporesponsive period. This 2-week period is characterized by a diminished response to stress and maximal rates of neurogenesis. By the third postnatal week and into adulthood, circulating adrenal steroid levels are relatively high and granule production is diminished compared to the levels observed during development. With aging, the levels of glucocorticoids are elevated further and the rate of granule cell production becomes minimal.

Consistent with these natural observations, several studies have demonstrated that the production of neurons can be altered by the experimental manipulation of circulating adrenal steroid levels both during development and adulthood. During the stress hyporesponsive period, the administration of corticosterone results in a significant decrease in the number of granule neurons produced. Similar suppressive effects of adrenal steroids on the proliferation of granule cell precursors have been observed in adulthood as well. The treatment of adult rats with corticosterone results in a decrease in the production of new cells, whereas the removal of adrenal steroids by adrenalectomy results in a rapid increase in the number of new cells produced and, ultimately, in the number of new granule neurons. Despite the dramatic effects of hormone manipulations on cell proliferation in the dentate gyrus, very few granule cell precursors express adrenal steroid receptors, suggesting that adrenal steroids influence neurogenesis indirectly.

In addition to increasing circulating adrenal steroid levels, stressful experiences stimulate hippocampal glutamate release and alter the expression of the NMDA subtype of glutamate receptor in the hippocampal formation. Several lines of evidence suggest that adrenal steroids affect cell proliferation in the dentate gyrus via excitatory input. First, the timing of perforant path development and increased expression of NMDA receptors in the developing dentate gyrus coincides with an observed decrease in granule neuron production. Second, experimental manipulation

of NMDA receptor activation during development or adulthood results in rapid changes in the numbers of new granule neurons. The administration of NMDA receptor antagonists increases the proliferation of granule cell precursors and the production of new granule cells during the early postnatal period and in adulthood in rats and tree shrews. It is likely that these effects are mediated by the perforant pathway, inputs that can involve NMDA receptor activation, because lesion of the entorhinal cortex enhances cell proliferation in the dentate gyrus in adulthood.

The regulation of granule cell production by adrenal steroids and NMDA receptor activation occur via a common pathway. Treatment with corticosterone and MK801 during adulthood results in an increase in cell proliferation that is similar to that observed with MK801 treatment alone. Conversely, adrenalectomized rats that are treated with NMDA demonstrate numbers of proliferating cells that are similar to animals treated with NMDA alone; both groups demonstrate a significant decrease in the numbers of proliferating cells compared to control animals. Collectively, these results indicate that adrenal steroids and NMDA receptor activation decrease cell proliferation via a common pathway. In addition, these results indicate that NMDA receptor activation exerts a more proximal effect on cell proliferation, as evidenced by the fact that the experimental manipulation of NMDA receptor activation can block the effects of corticosterone.

Although the mechanism by which adrenal steroids and NMDA receptor activation suppress cell proliferation is not known, it is possible that these factors initiate a cascade of events that result in a lengthening of the cell cycle. *In vitro* studies in nonneuronal systems have shown that glucocorticoids are capable of lengthening the G1 phase or G2/M phases of the cell cycle. Alternatively, these factors may induce changes that prevent precursor cells from progressing through the cell cycle. *In vitro* studies have also demonstrated that glucocorticoids can reversibly halt the progression of the cell cycle in nonneuronal systems. If adrenal steroids do in fact inhibit progenitor cells from dividing, the removal of adrenal steroids may result in the disinhibition of cells that are normally prevented from progressing through the cell cycle. Such an effect could account for the speed and magnitude of the increases in the number of cells in S phase following the removal of adrenal steroids.

Functional Implications of Stress Effects on Neurogenesis

The functional implications of the stress-induced suppression of neurogenesis are not known. However, the

suppression of granule cell production during a time when the majority of granule neurons are produced is likely to have a significant effect on the structure of the adult hippocampal formation. The hippocampal formation has been implicated in learning and memory, raising the possibility that decreases in the production of granule neurons during development may have a negative impact on learning and memory in adulthood. Although there is some evidence to support the idea that stressful experience during development is correlated with impaired cognitive function, an involvement of stress-induced changes in granule neuron production has not yet been investigated.

Although the exact functional significance of adult-generated neurons is not known, evidence suggests a relationship between these new cells and hippocampal-dependent learning. Specifically, the survival of adult-generated cells appears to depend on learning hippocampal-dependent tasks at a critical time after their production. Taken together with the observation that stress suppresses the production of granule neurons, the finding that learning enhances the number of granule neurons suggests that stress-induced changes in neurogenesis may have a significant impact on learning. It is important to note that functional changes are not likely to be immediately evident. New cells require time to differentiate and become incorporated into functional circuitry. Thus, the period following stress-induced changes in cell proliferation is most likely to be functionally relevant. In addition, it is important to note that the functional changes attributable to acute changes in cell production are not likely to be of an observable magnitude. For this reason, the impact of stress-induced changes in neuron production is most likely to be of interest when considering conditions of chronic stress. Thus, although previous studies have demonstrated enhancements in learning as a result of conditions of acute stress, the mechanisms that underlie these effects are likely to involve processes other than changes in neuron production, such as changes in synaptic efficacy.

Several studies have demonstrated that chronic stress or corticosterone treatment exerts a significant impact on hippocampal-dependent learning. In addition, later studies have demonstrated that a stress-induced impairment of spatial learning on different tasks can be observed with both a shorter and a longer duration of stress; 8 days of psychosocial stress result in impaired performance on the hole board spatial discrimination task, and exposure to 6 months of psychosocial stress results in impaired performance on the Morris water maze spatial learning task. The stress-induced impairment observed is not permanent; rats that were tested on the radial arm maze 18 days

following the termination of stress performed comparably to control rats. This observation is consistent with a possible involvement of adult-generated cells. If adult-generated cells are indeed necessary for normal performance on this task, an impairment may only persist as long as the production of these cells is altered. Once the stress-induced decrease is no longer present, we expect to observe normal levels of performance. However, stress-induced impairments during the early postnatal period, when the granule cell layer is forming, may have lasting effects.

In general, the degree to which the structure of the adult brain can be altered by experience is limited by the fact that neurons in most brain regions are produced during a discrete period of early embryonic development. Thus, the postnatal production of neurons in the dentate gyrus presents an unusual situation in which the experience of an organism may exert dramatic and potentially long-lasting effects on the structure and function of this brain region.

See Also the Following Articles

Glucocorticoids – Adverse Effects on the Nervous System; Glucocorticoid Effects on Memory: the Positive and Negative; Hippocampal Neurons; Hippocampus, Corticosteroid Effects on.

Further Reading

- Altman, J. and Bayer, S. A. (1990). Migration and distribution of two populations of hippocampal granule cell precursors during the perinatal and postnatal periods. *Journal of Comparative Neurology* **301**, 365–381.
- Gould, E. and Cameron, H. A. (1996). Regulation of neuronal birth, migration and death in the rat dentate gyrus. *Developmental Neuroscience* **18**, 22–35.
- Gould, E., McEwen, B. S., Tanapat, P., et al. (1997). Neurogenesis in the dentate gyrus of the adult tree shrew is regulated by psychosocial stress and NMDA receptor activation. *Journal of Neuroscience* **17**, 2492–2498.
- Gould, E., Tanapat, P., Hastings, N. and Shors, T. J. (1999). Neurogenesis in adulthood: a possible role in learning. *Trends in Cognitive Sciences* **3**, 186–191.
- Gould, E., Tanapat, P., McEwen, B. S., et al. (1999). Proliferation of granule cell precursors in the dentate gyrus of adult monkeys is diminished by stress. *Proceedings of the National Academy of Sciences USA* **95**, 3168–3171.
- Sapolsky, R. M. and Meaney, M. J. (1986). Maturation of the adrenocortical stress response: neuroendocrine control mechanisms and the stress hyporesponsive period. *Brain Research* **396**, 64–76.
- Tanapat, P., Galea, L. A. and Gould, E. (1998). Stress inhibits the proliferation of granule cell precursors in the developing dentate gyrus. *International Journal of Developmental Neuroscience* **16**, 235–239.

Neuroimaging and Emotion

N A Harrison and H D Critchley

University College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

Emotion

Definition of Emotion

Emotion Perception and Attention

Memory and Learning

Interoception and Subjective Feeling States

Social Interaction

Emotion Regulation

Emotion Dysregulation

Glossary

<i>Amygdala</i>	An almond-shaped cluster of inter-related nuclei and associated cortex in the anterior medial temporal lobe.
<i>Arousal</i>	A dimension of emotion that varies from calm to excitement and predicts behavioral activity level.
<i>Blood-oxygen level-dependent (BOLD) signal</i>	T2*-weighted signal in functional magnetic resonance imaging (fMRI) that reflects hemodynamic changes in regional perfusion evoked by neural activity.
<i>Functional magnetic resonance imaging (fMRI)</i>	A non-invasive technique with good spatial resolution that detects changes in regional blood flow associated with neural activity.
<i>Interoception</i>	A representation of visceral activity; the afferent limb of homeostatic neural control of bodily organs.
<i>Magnetoencephalography (MEG)</i>	A non-invasive technique with high temporal resolution that detects the changing magnetic fields associated with neural activity.
<i>Valence</i>	A dimension of emotion that varies from unpleasant (negative) to pleasant (positive) to reflect motivational value.

Emotion

Emotion is central to our everyday human experience. Emotional events or objects within our environment are assigned value, bias our attention, and enhance our subsequent memory. Neuroimaging techniques, perhaps more than other approaches, have led to a growing awareness that emotion, unlike many other psychological functions, is relatively unencapsulated and interacts with and influences multiple

other areas of functioning. In addition to effects on attention, perception, memory, and learning, emotion also forms the backbone of our social relationships and is an important component of empathetic understanding of others. Emotional cues also act as powerful reinforcers, have a marked influence on our thoughts and reasoning, and serve to bias our ongoing behavior. Although emotions come in many flavors that may equally serve to bias multiple psychological functions, much of the contemporary research on emotional neuroscience has focused on the processing of fear. For the purposes of this article, we follow this focus, which also accords with an emphasis on stress responses.

Consistent with a broad role in human functioning are the wide-ranging manifestations associated with disorders of emotion regulation. Emotion disequilibrium underlies not just anxiety and stress-related illness but the entire range of mental disorders and acts as the common denominator in much of human unhappiness. This article reviews the contributions of neuroimaging studies to our developing understanding of the neurobiology of human emotion. One particular aim is to highlight how emotion interacts with and influences other areas of cognition and to illustrate how an understanding of the mechanisms underlying these interactions can enable a greater understanding of stress-related disorders.

Definition of Emotion

Differences in definition can lead to markedly different conceptualizations of emotion and underlie seemingly contradictory theories of emotion in the literature. A prevalent view conceptualizes emotion as a transient perturbation in an organism's functioning evoked by a triggering (i.e., emotive) stimulus (either external or internal). Most accounts of emotion also subscribe to a multicomponential description with physiological arousal, motor expression, and subjective feelings forming the reaction triad of emotion. However, the extent to which changes in these various components are necessary and sufficient to define an episode as emotional is more controversial. For the purposes of this article, we use the following definition: emotions are transient events, produced in response to external or internal events of significance to the individual; they are typically characterized by attention to the evoking stimulus and changes in neurophysiological arousal, motor behavior, and subjective feeling state that engender a subsequent biasing of behavior.

Emotion Perception and Attention

Implicit within this definition is that events of significance to the individual (emotive stimuli) are susceptible to preferential perceptual processing. One mechanism through which this is achieved is through the emotional biasing of attentional processes. This is illustrated in behavioral studies using visual search paradigms, in which target items are discriminated from an array containing a variable number of distracter items. Emotionally valenced items or objects are indeed identified more rapidly than nonemotional items. Furthermore, in spatial tasks, the presentation of an emotional cue on one side of the visual space increases the speed of identification of nonemotional stimuli subsequently presented on the same side (i.e., attention is drawn to the location of the preceding emotive stimulus). Neuroimaging studies investigating the neural basis of this emotional capture of attention in spatial orientation tasks show that it is correlated with activity in regions of the frontal and parietal cortices as well as the lateral orbitofrontal cortex.

Evidence from patient studies, neuroimaging, and electroencephalography suggests that even unattended emotional stimuli are more likely to enter awareness, suggesting further mechanisms through which emotional stimuli influence perception. Neuroimaging studies continue to delineate the brain mechanisms through which unattended emotional stimuli gain enhanced access to awareness. One influential hypothesis suggests that the emotional salience of a stimulus is rapidly detected by the amygdala after cursory representational processing. The amygdala subsequently facilitates attention and perception via projections to the sensory cortices to enable more detailed sensory processing. Convergent anatomical studies suggest that this is mediated by both direct and indirect amygdala influences on sensory cortices; reciprocal direct connections exist between the amygdala and sensory cortices, and a projection from the central nucleus of the amygdala to the cholinergic nucleus basalis of Meynert enables indirect ascending cholinergic neuromodulatory influences to be exerted on the same cortical regions.

Functional neuroimaging studies have tested this hypothesis using visual backward-masking paradigms in which briefly presented visual targets are rendered invisible by subsequently presented stimuli. Unseen emotional stimuli, presented in this subliminal manner, can nevertheless still evoke neural responses in the amygdala, compared to nonemotional (neutral) stimuli. Moreover, masked presentations of face stimuli portraying fearful (vs. neutral) facial expressions enhances activation within the fusiform cortex (an

extra-striate visual area containing a region implicated as being specialized for processing faces). This activation of the fusiform face area (FFA) is predicted by the magnitude of amygdala activation. Further evidence that emotional processing in the amygdala may enhance activity within sensory cortex comes from the observation that patients with amygdala damage fail to show this enhancement of activity within the face-processing visual cortex when tested on the same paradigm. The presentation of fearful face cues has also been shown to enhance even early perceptual functions such as contrast sensitivity, which is consistent with the suggestion that the amygdala may modulate even activity within the primary visual cortex.

Preattentive processing of emotional stimuli indicates early representational discrimination between emotional and nonemotional events. Studies in humans using magnetoencephalography (MEG), a neuroimaging technique with high temporal resolution, showed discriminatory responses to emotional faces as early as 100–120 ms. This compares to characteristic face-related responses that occurs much later, at 170 ms. Electrophysiological studies in humans using intracranial electrodes also showed an early, 120- to 160-ms response to aversive stimuli. The early recognition of emotionally valenced stimuli accords with animal data suggesting that the amygdala may be activated by a rapid subcortical pathway. Neuroimaging studies in both normal and brain injury subjects highlight subcortical amygdala activation in humans. One study exploited two separate findings: (1) that low- and high-spatial frequency information is processed by separate visual neural pathways and (2) that coarse emotional cues are carried in low-frequency components. When subjects were presented low-frequency (blurred) face stimuli fearful (vs. neutral face stimuli), enhanced activity was seen in the amygdala, pulvinar nucleus of the thalamus, and superior colliculus, components of a proposed subcortical circuit (Figure 1). In contrast, the presentation of the same faces at high spatial frequency activated the cortical regions, including the fusiform (face-identity-specific) cortex. Implicit processing of emotional valence can be demonstrated in patients with damage to the primary visual cortex, who are not consciously aware of stimuli presented in their blind hemifield (i.e., blindsight). These patients show amygdala activation to fearful faces presented in their blind hemifield, illustrating the presence of the subcortical visual processing of emotion to the level of the amygdala.

Together these lesion and functional imaging studies provide mechanistic insight into the role of amygdala, through interactions with the sensory cortices, in facilitating enhanced attention and perception of

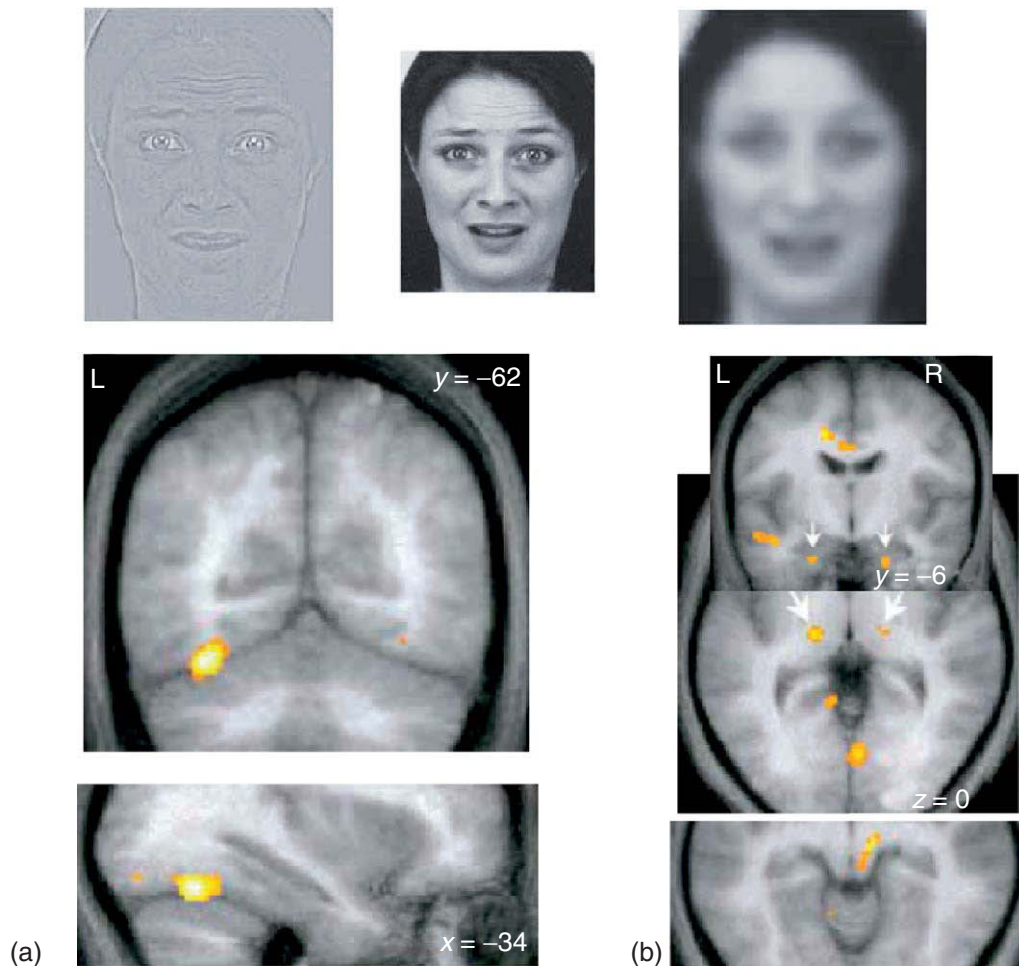


Figure 1 Neural activity in response to viewing facial expressions. a, High-spatial frequency information; b, low-spatial frequency information. High-spatial frequency images activated the fusiform cortex, specialized for the processing of facial identity. Fearful (vs. neutral) facial expressions activated the amygdala bilaterally (vertical arrows) with greater response to low-spatial frequency images. Low-spatial frequency fearful images also activated the posterior inferior thalamus bilaterally (diagonal arrows) and the superior colliculus on the right suggesting that low-spatial frequency images containing emotion information activate a subcortical pathway to the amygdala. Reprinted by permission from Macmillan Publishers Ltd: *Nature Neuroscience*; Vuilleumier, P., Armony, J. L., Driver, J., et al., Distinct spatial frequency sensitivities for processing faces and emotional expressions, **6**, 624–631, copyright (2003).

visually presented emotional stimuli. The preferential early processing of emotionally valenced material, illustrated most powerfully with fearful stimuli, sets up a primary bias in subsequent information processing that underpins the regulation of ongoing and prospective behavior.

Neuroscientific studies of human emotion have tended to focus on the visual processing of emotional facial expressions. Nevertheless, this visual communication of emotional state through changes in the facial musculature is only one part of a richer multi-modal vocabulary of affect. Information signaling emotional state is also conveyed by autonomic changes in the skin (leading to color change and sweating) and pupils, by posture and bodily movements, and by content and prosody of speech. Low-level auditory

processing of emotional signals are the primary source of emotional information in speech. The prioritized processing of emotional sounds, particularly non-speech intonations of emotional state (gasps, screams, etc.), is illustrated within the primary and accessory auditory cortices. As with the visual system, amygdalo-cortical interactions are implicated in enhanced auditory attention to emotive sounds. Moreover, specificity in emotional processing may cross sensory modalities, such that ventral insular regions implicated in disgust-processing may be activated by auditory and visual expressions of disgust as well as by disgusting stimuli.

A degree of regional specialization in human emotional processing is suggested from the convergence of lesion and neuroimaging studies. Acquired damage to the amygdala impairs the recognition and

behavioral experience of fear. Similarly, the neuroimaging literature emphasizes a predominant sensitivity of amygdala responses to threat and fear signals. As previously indicated, other studies implicate regions of insula and striatum in the selective processing of disgust. The impaired recognition of facial expressions of disgust was reported in a patient with damage to these regions, whereas the stimulation of the insula may evoke feelings of nausea or a perception of unpleasant tastes. These lines of evidence for specialized emotion-specific processing within different brain regions are generally circumscribed to fear/threat (amygdala), disgust (insula and striatum), and, to a lesser extent, sadness (subgenual cingulate activity). Many neuroimaging studies highlight a general sensitivity of amygdala responses to the emotional intensity of external emotional stimuli, for which threat is perhaps most arousing.

The insula cortex also appears to play a more complex role in emotion perception, apparently through relating perceptual information with interoceptive representations of bodily states. Thus, insula activity is engaged during the recognition and experience of disgust, sadness, happiness, and fear; during the learning and expression of threat; during the perception of noxious stimuli; and during the experience of phobic symptoms, hunger, and satiety; and even during explicit facial emotion categorization.

Memory and Learning

Emotion has a striking and well described impact on memory processes. Animal studies, particularly the work of LeDoux, highlight the central role of the amygdala in fear conditioning, a form of implicit memory. The same is true in humans, illustrated in many neuroimaging and lesion studies. Fear conditioning is a rapid form of learning that a stimulus is potentially harmful. If an adverse event (unconditioned stimulus, US; e.g., a painful shock) occurs after a relatively innocuous stimulus, the latter (the conditioned stimulus, CS) comes to predict the adverse event and will engender arousal responses associated with the presence of threat. Patients with bilateral amygdala damage do not acquire conditioned fear responses (autonomic or motor), despite having an intact explicit knowledge of the association between the CS and US. Functional neuroimaging studies also confirm the importance of the amygdala in fear conditioning (i.e., the learning of CS-US associations). However the role of the amygdala in the expression of threat responses is more time-limited and shows habituation. Consistent with animal data, neuroimaging studies also describe amygdala engagement in more general associative reinforcement learning,

including operant stimulus-reward learning and appetitive behavior.

In addition to simple associative responses, human memory is characterized by declarative richness in episodic and semantic memory. These domains of memory, typically supported both at encoding and recollection by activity within the hippocampus, are strongly modulated by emotional experience. Typically, emotional events and stimuli are remembered better than nonemotional occurrences. Animal studies highlight noradrenergic mechanisms enhancing amygdalo-hippocampal connectivity in preferential encoding of emotional objects. Brain-imaging studies in humans have gone much further, highlighting amygdala activation during encoding actually predicts the accuracy of subsequent declarative memory of positively or negatively valenced stimuli (Figure 2) and in delineating the functional architecture for mood-congruency effects (in which negative information is encoded and recalled to a greater extent when the subject is in a negative mood and, likewise, positive material is encoded and recalled to a greater extent in positive mood states). However, the role of the amygdala extends beyond encoding processes; it is also engaged during the retrieval of emotional items and contexts.

Interoception and Subjective Feeling States

Changes in bodily state, reflected in both physiological arousal and motor behavior, are obligatory defining characteristics of emotion. James and Lange, and more recently Damasio, argued that emotional feeling states must have their origin in the brain's representation of the bodily (arousal) state. Without changes in bodily arousal, there is no emotion. This peripheral account of emotion predicts that afferent feedback signals from muscles and viscera associated with physiological arousal are integral to emotional experience. These mechanisms have been illuminated in a number of neuroimaging experiments examining central representation and rerepresentation of autonomic arousal states and homeostatic sensations, such as temperature.

The role of the insula cortex in this context is particularly interesting. In a positron emission tomography (PET) study, activation within the viscerosensory mid-insula cortex correlated closely with the actual stimulus temperature, yet the subjective ratings of the perceived intensity or pleasantness of hot and cold stimuli correlated with activity in the right anterior insula and orbitofrontal cortex. This finding suggests a dissociation between a primary central representation of viscerosensory information, within

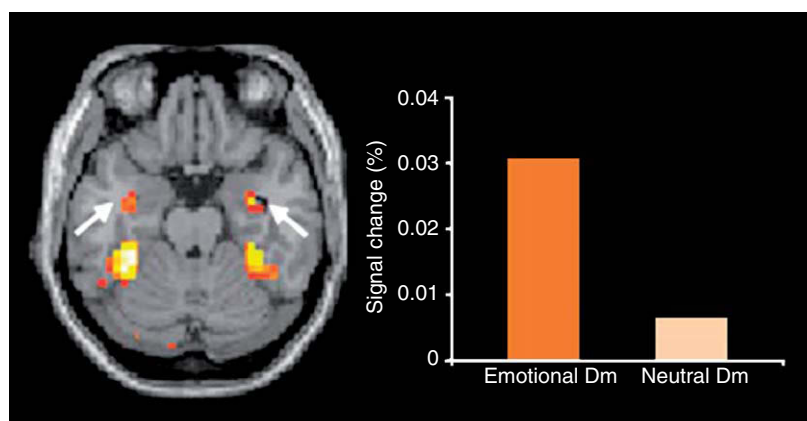


Figure 2 Level of activity within the amygdala (arrows) during the encoding of emotional images predicts whether a subject is subsequently able to recall emotional pictures (emotional Dm) but not neutral pictures (neutral Dm). Reprinted from *Neuron* 48, E. A. Phelps & J. E. LeDoux, Contributions of the amygdala to emotion processing: from animal models to human behavior, pp. 175–187, Copyright 2005, and adapted from *Neuron*, 42, F. Dolcos, K. S. LaBar & R. Cabeza, Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events, pp. 855–863, Copyright (2004), with permission from Elsevier.

mid-insula, and subjective feelings of warmth or cold, represented in the right anterior insula and orbitofrontal cortex. A general role for these regions in the representation of emotional feelings is also supported by their activation observed during generated states of sadness and anger, anticipatory anxiety and pain, panic, disgust, sexual arousal, trustworthiness, and even subjective responses to music.

Behavioral and neuroimaging studies in patients with selective peripheral autonomic denervation (pure autonomic failure, PAF) lend further support to the importance of viscerosensory feedback to emotional feelings. These patients, who do not have autonomic (e.g., heart-rate or skin-conductance) changes in response to emotional stimuli or physiological stressors, show a subtle blunting of subjective emotional experience. In a fear conditioning study of PAF subjects and controls, both groups showed amygdala activity during the learning of threat (face stimuli, CS, were paired with an aversive blast of white noise, US). However, compared to controls, the right insula cortex showed attenuated responses to threat stimuli in the PAF patients, suggesting this region is sensitive to the presence and absence of autonomic responses generated during emotional processing. Moreover, when the threat stimuli were presented unconsciously (using backward masking), the same insula regions (including the right anterior insula) were sensitive to both the conscious awareness of emotion events and induced bodily arousal reactions. This study, particularly, suggests that the right insula is a neural substrate for the contextual integration of external emotional events with feeling states arising from bodily reactions.

The notion that subjective emotional feelings arise from central representations of bodily arousal states

predicts that people who are more aware of their bodily responses may experience more intense emotions and feelings. Investigations examining individual differences in interoceptive awareness suggest that patients with anxiety disorders are more attuned to bodily reactions such as their own heartbeat or stomach motility. One such interoceptive task was translated into a neuroimaging study. Subjects were played a series of notes, presented in groups of 10, either synchronously or out of phase with their own heartbeats (see [Figure 3a](#)). In half the trials, subjects had to judge the timing of their heartbeat in relation to the notes (the control condition was a detection of a rogue note in the sequence). When subjects focused interoceptively to the timing of their own heartbeat, there was enhanced activity in the somatomotor, insula, and cingulate cortices. However, only one region, the right anterior insula cortex, reflected a conscious awareness of internal bodily responses (i.e., how accurately subjects performed the interoceptive task). Significantly, activity in this region, and even the relative amount of right anterior gray matter here ([Figure 3b](#)), predicted day-to-day awareness of bodily reaction and the subjective experience of negative emotions, particularly anxiety symptoms across individuals. Together these studies suggest that the right anterior insula is a principal neural substrate supporting conscious emotional feelings arising from interoceptive information concerning bodily arousal state ([Figure 4](#)).

Social Interaction

A characteristic of our social interactions is an intuitive ability to understand other people's mental and

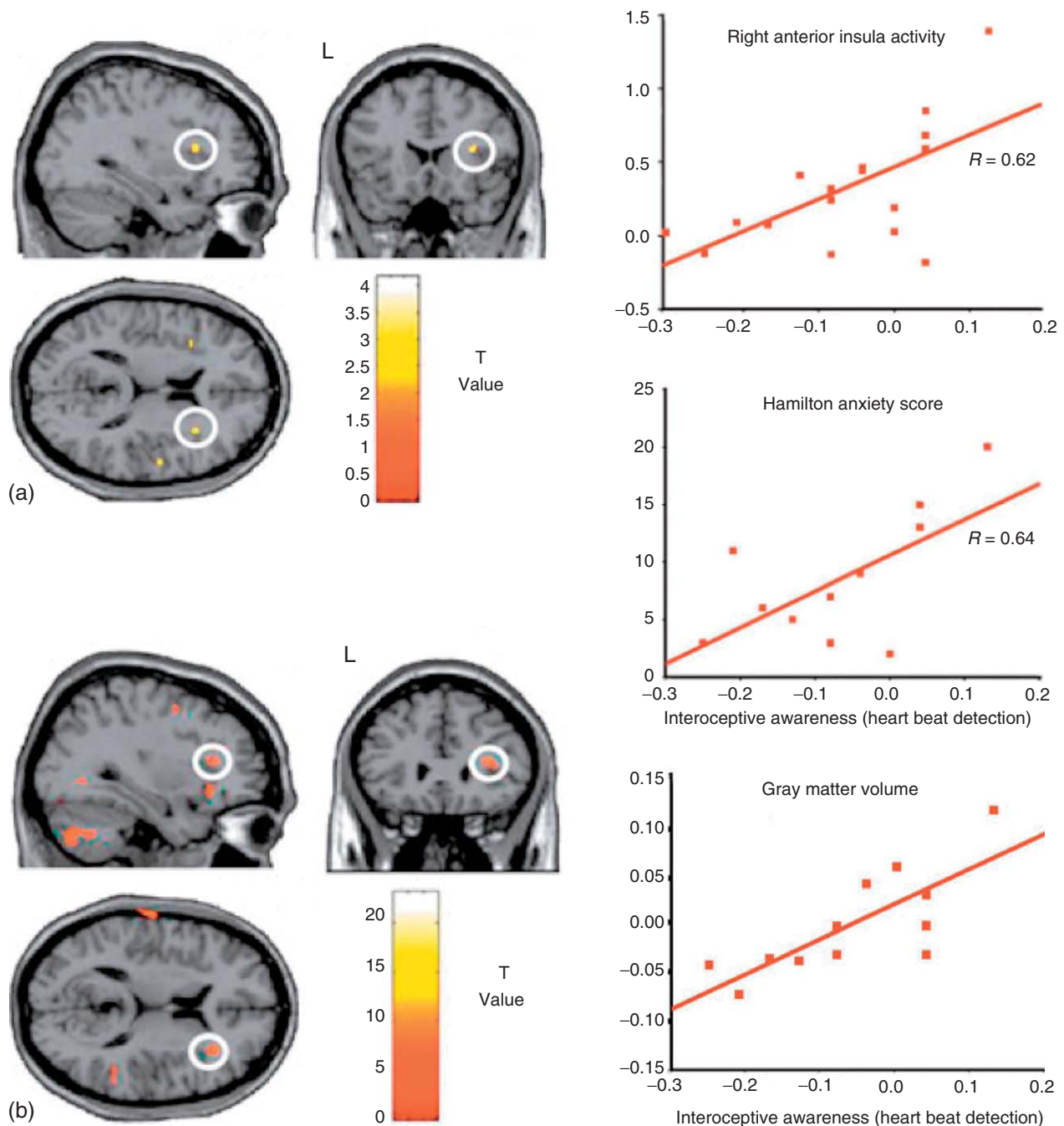


Figure 3 Interoceptive awareness (measured as accuracy on a heartbeat detection task). a, Association with increased neural activity in the right anterior insula/operculum; b, correlation with the gray matter volume of the right anterior insula region. Further subjects with greater interoceptive awareness also showed greater neural activity in this region during interoceptive trials. Activity in this region also correlated with anxiety scores, suggesting that emotional feeling states are supported by explicit interoceptive representations within right insula cortex. Reprinted by permission from Macmillan Publishers Ltd: *Nature Neuroscience*; Critchley, H. D., Wiens, S., Rotshtein, P., et al., Neural systems supporting interoceptive awareness, 7(2), 189–195, copyright (2004).

emotional states. Humans show a marked tendency to mimic one another's gesticulations, emotional facial expressions, and body postures, suggesting that this mirroring of activity mediates and facilitates emotional understanding among individuals. At a neural level, cells in monkey premotor cortex have been described

that respond to observing or performing the same motor actions. Corresponding human imaging studies of action observation also illustrate premotor cortical activity when observing the actions of others, and magnetoencephalography (MEG) studies report resynchronization in the primary motor cortex. Significantly,

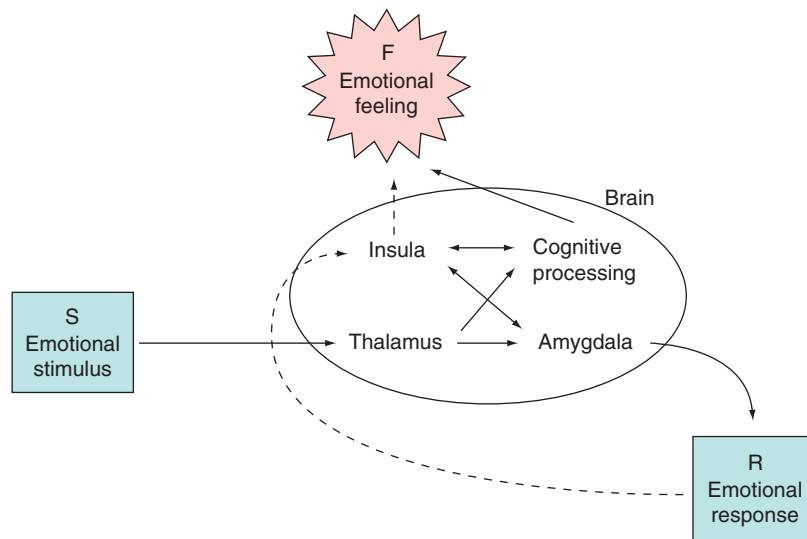


Figure 4 A functional neuroanatomical modification of William James's (1884) model of emotional responses. Emotional stimuli (S) elicit automatic emotional responses (R) via thalamus–amygdala pathways (solid line arrows). This low-level processing occurs independently of conscious cognitive processing. Peripheral autonomic responses are fed back to the mid-insula and then are remapped to the right anterior insula where they interact with higher-level cognitive processing. Emotional feeling states therefore result from peripheral-central interaction in the right anterior insula. Reprinted from *Trends in Cognitive Sciences*, 6(8), J. S. Morris, How do you feel?, pp. 317–319, Copyright (2002), with permission from Elsevier.

these mirror activations are somatotopic with respect to the body part performing the action.

Accumulating evidence supports the extension of these action-perception mechanisms to emotions and feeling states. A common neural representation is proposed for the perception of actions and feelings in others and their experience in self. Thus, viewing emotional facial expressions in others activates automatically mirrored expressions on one's own face (surprisingly, even when the observed facial expressions are masked, hence unavailable to conscious awareness). Significantly, this automatic mimicking of emotional expressions also extends to autonomic expressions of emotional states such as pupil size in expressions of sadness. Neuroimaging studies show shared neural activation in the somatosensory cortex when a subject experienced or observed another being touched, right insula activity when a subject experienced disgusting odors or observed expressions of disgust, and the activation of a common matrix of regions when a subject experienced or observed pain. The imitation or observation of emotional facial expressions also activates a largely similar network of brain regions including the premotor areas that are critical for action representation. There is also selectivity; compared to nonemotional facial movements, emotional expressions activate the right frontal operculum and anterior insula, and regions such as the amygdala, striatum, and subgenual cingulate are differentially responsive when imitating happy or sad emotions.

Emotion Regulation

Understanding the adaptive mechanisms through which we can control the expression of our emotions is clinically and socially important. Unlearning established emotional reactions is difficult. Many studies have examined the modulation of previously learned emotional responses with experimental extinction; the repeated presentation of a previously CS without its associated US rapidly leads to a diminished emotional response. Nevertheless, the relationship between the two stimuli is not forgotten but inhibited and can be rapidly reinstated. Neuroimaging of extinction learning (or the inhibition of emotional responses) shows increased amygdala activity in early extinction; however, retention after a day was associated with activity in the subgenual anterior cingulate and ventromedial prefrontal cortices, which suggests a role for this region in the longer-term recall of extinction learning (Figure 5).

Emotional responses may also be regulated volitionally, using cognitive control or reappraisal. For example, an image of a woman crying may be interpreted as one of sadness, but, if the subject is then told that the picture was taken after her daughter's wedding, it may be reappraised as one of great joy. Neuroimaging studies of emotional reappraisal demonstrate an attenuation of activity in the amygdala associated with enhanced activity in the left middle frontal gyrus, suggesting that activity within the lateral

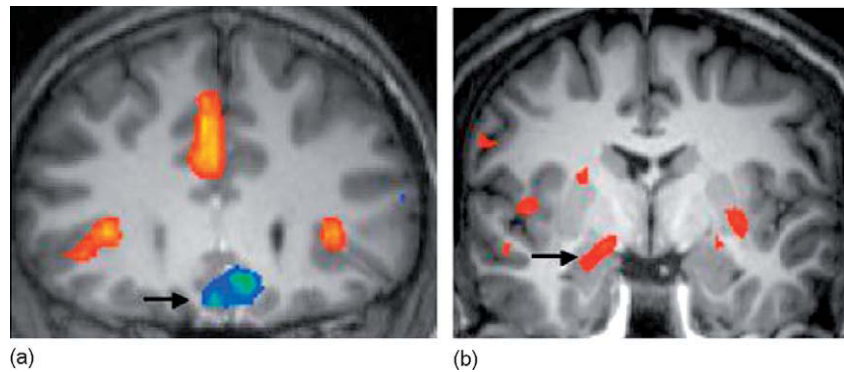


Figure 5 Regions activated in the extinction of conditioned fear. a, Ventromedial prefrontal cortex (arrow) showing less response to the conditioned stimulus (CS+) during acquisition; b, amygdala activation (arrow) in response to the CS+ during acquisition versus early extinction. During extinction training, ventromedial prefrontal cortex activity gradually increased, and the magnitude of this increase predicted the retention of extinction learning. Extinction training resulted in a reduction of amygdala activity in response to the CS+ and predicted the success of early extinction. Reprinted from *Neuron*, **48**, E. A. Phelps & J. E. LeDoux, Contributions of the amygdala to emotion processing: from animal models to human behavior, pp. 175–187, Copyright (2005), and adapted from *Neuron*, **43**, E. A. Phelps et al., Extinction learning in humans: role of the amygdala and vmPFC, pp. 897–905, Copyright (2004), with permission from Elsevier.

prefrontal cortex modulates emotional reactivity of the amygdala, possibly via connections to medial prefrontal cortex. Similar cognitive emotion-regulation techniques can powerfully diminish responses acquired through fear conditioning, and neuroimaging studies using this paradigm again show increased activity in the left middle frontal gyrus with cognitive control. Taken together, these studies suggest that cognitive emotion regulation and lower-level extinction learning recruit similar overlapping neural mechanisms for the regulation of the amygdala to control primary emotional responses.

Emotion Dysregulation

The wide-ranging influences of emotion on other psychological functions may also be seen in the variety of symptoms and presentations associated with emotional disorders. Posttraumatic stress disorder (PTSD), resulting from exposure to a highly traumatic emotional event and usually associated with injury or risk of injury to self or others, is frequently associated with symptoms in multiple psychological domains. Enhanced arousal and startle responses to previously innocuous stimuli are common, as are intrusive traumatic memories (flashbacks) of the triggering event, the numbing of emotional feelings, and impairments to social relationships. The identification of brain regions that show increased activation associated with the enhanced perception and memory for emotional events and associated feeling states enables the linkage of psychological mechanisms in conditions such as PTSD and phobias with their underlying neural substrates.

Abnormal cortisol regulation and adrenergic stress responses are also seen in PTSD, as is a reduction in hippocampal volume. Imaging studies showing a

similar reduction in hippocampal volume in endocrine disorders associated with chronically elevated cortisol levels, such as Cushing's disease, suggest an etiological role for stress hormones such as cortisol in this reduction in hippocampal volume. These conclusions are further supported by a recent study of healthy postmenopausal women, which showed a strong correlation between hippocampal volume and lifetime levels of perceived stress, suggesting an influence of stress on regional brain volumes even in the normal population.

Maladaptive cognitive interpretations of situations or events are central components of many emotional disorders such as depression and the anxiety disorders and are a feature focused on by many of the psychological therapies used in their treatment. Understanding the neural mechanisms mediating the modulation of emotional responses will be essential to developing a comprehensive neurobiological understanding of these disorders. Furthermore, a better understanding of these mechanisms offers the opportunity to develop more targeted approaches specifically focused on changing these maladaptive emotional reactions.

See Also the Following Articles

Amygdala; Brain and Brain Regions; Emotions: Structure and Adaptive Functions; Hippocampus, Overview.

Further Reading

- Calder, A. J., Keane, J., Manes, F., et al. (2001). Impaired recognition and experience of disgust following brain injury. *Nature Neuroscience* 3(11), 1077–1078.
- Craig, A. D. (2002). How do you feel? interoception: the sense of the physiological condition of the body. *Nature Reviews Neuroscience* 3, 655–667.

- Critchley, H. D., Mathias, C. J. and Dolan, R. J. (2001). Neuroanatomical basis for first and second order representations of bodily states. *Nature Neuroscience* 4(2), 207–212.
- Critchley, H. D., Mathias, C. J. and Dolan, R. J. (2002). Fear conditioning in humans: the influence of awareness and autonomic arousal on functional neuroanatomy. *Neuron* 33(4), 653–663.
- Critchley, H. D., Wiens, S., Rotshtein, P., et al. (2004). Neural systems supporting interoceptive awareness. *Nature Neuroscience* 7(2), 189–195.
- Damasio, A. R. (1994). *Descartes error: emotion, reason and the human brain*. New York: Grosset/Putnam.
- Damasio, A. R. (1999). *The feeling of what happens: body and emotion in the making of consciousness*. New York: Harcourt Brace.
- Darwin, C. (1872). *The expressions of emotion in man and animals*. London: John Murray. (Reprinted 1998 (3rd edn.), Ekman, P. (ed.). London: Harper Collins.)
- Dimberg, U., Thunberg, D. and Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science* 11(1), 86–89.
- Dolan, R. J. (2002). Emotion, cognition, and behavior. *Science* 298, 1191–1194.
- Dolcos, F., LaBar, K. S. and Cabeza, R. (2004). Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron* 42, 855–863.
- Gallese, V. (2003). The manifold nature of interpersonal relations: the quest for a common mechanism. *Philosophical Transactions of Royal Society London B* 358(1431), 517–528.
- Gianaros, P. J., Jennings, J. R., Sheu, L. K., et al. (2007). Prospective reports of chronic life stress predict decreased grey matter volume in the hippocampus. *NeuroImage* 35(1), 220–232.
- Harrison, N. A., Singer, T., Rotshtein, P., et al. (2006). Pupillary contagion: central mechanisms engaged in sadness processing. *Social and Affective Neuroscience* 1, 5–17.
- James, W. (1884). Physical basis of emotion. *Psychological Review* 1, 516–529.
- LaBar, K. S. and Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience* 7, 54–64.
- Lange, C. (1885/1922). *The emotions*. Haupt, I. A. (trans.). Baltimore, MD: Williams & Wilkins.
- LeDoux, J. E. (1996). *The emotional brain: the mysterious underpinnings of emotional life*. New York: Simon & Schuster.
- Morris, J. S. (2002). How do you feel? *Trends in Cognitive Sciences* 6(8), 317–319.
- Morris, J. S., DeGelder, B., Weiskrantz, L., et al. (2001). Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain* 124, 1241–1252.
- Morris, J. S., Friston, K. J., Buchel, C., et al. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain* 121, 47–57.
- Morris, J. S., Ohman, A. and Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature* 393, 467–470.
- Ochsner, K. N. and Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences* 9(5), 242–249.
- Penfield, W. and Faulk, M. E. (1955). The insula: further observations on its function. *Brain* 78(4), 445–470.
- Phelps, E. A. and LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* 48, 175–187.
- Phillips, M. L., Young, A. W., Senior, C., et al. (1997). A specific neural substrate for perception of facial expressions of disgust. *Nature* 389, 495–498.
- Scherer, K. R. (2000). Psychological models of emotion. In: Borod, J. C. (ed.) *The neuropsychology of emotion*, pp. 137–162. New York: Oxford University Press.
- Singer, T., Seymour, B., O'Doherty, J., et al. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science* 303(5661), 1157–1162.
- Vuilleumier, P., Armony, J. L., Driver, J., et al. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience* 6, 624–631.
- Vuilleumier, P., Richardson, M. P., Armony, J. L., et al. (2004). Distant influences of amygdala lesions on visual cortical activation during emotional face processing. *Nature Neuroscience* 7, 1271–1278.
- Wicker, B., Keysers, C., Plailly, J., et al. (2003). Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. *Neuron* 40, 655–664.

Neuroimmunomodulation

N R Rose

Johns Hopkins University, Baltimore, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 37–48, © 2000, Elsevier Inc., with revisions made by the Editor.

Communication between the Nervous System and the Immune System

Communication Channels between the Nervous System and the Immune System

Cytokine Interactions with the Nervous System

Clinical Implications of Neural-Immune Signaling

Glossary

<i>Cytokines</i>	Signal molecules released from immunocytes, supporting cells, or other cell types (endothelial cells) that can interact with other cells of the immune system and alter their intracellular capabilities.
<i>Immunocytes</i>	Cells of the immune system.
<i>Innervation</i>	The distribution of nerve fibers that end in association with a particular structure, region, or organ.
<i>Neurohormone</i>	A blood-borne signal of endocrine nature that is either derived from the brain directly or released from the pituitary gland due to neurally derived factors.
<i>Neuro-modulation</i>	The interaction of a neurally derived signal with a target cell that alters the intracellular capabilities of that cell through either genomic or nongenomic mechanisms.
<i>Neuropeptides</i>	Neural signal molecules derived from larger prohormones that consist of 3–100 or more amino acids; released from nerve endings, cells of neural origin, or cells of immune or endocrine origin.
<i>Neuro-transmitter</i>	A signal molecule released from nerve endings that can interact with specific receptors on target cells and alter either ion channels or second messengers following transduction.

Neuroimmunomodulation is the bidirectional signaling between the central nervous system (CNS) (the brain and spinal cord) and the cells and organs of the immune system. The CNS can send neuroendocrine signals that act on specific subsets of cells of the immune system and can also send direct neural

connections into lymphoid organs, where released neurotransmitters act on specific cell subsets of the immune system. These subsets of immunological cells possess a variety of receptors for neurohormones and neurotransmitters that permit signal transduction, influencing the activities and synthetic capacities of those cells. This direct neural-immune signaling modulates individual cellular functions as well as collective immune responses. Thus, the CNS can modulate immune responses, including innate immunity, cell-mediated immunity, autoimmunity, and complex collective immune reactivity to infections, cancer, and other challenges. This neural-immune circuitry permits behavioral influences, including those resulting from physical and psychological stressors, to modulate the activities of the immune system. The immune system, in turn, through the release of cytokines from specific subsets of cells, can signal neurons in the CNS and peripheral nervous system (PNS) directly and indirectly. This signaling influences the visceral and neuroendocrine functions of the hypothalamus, as well as the cognitive and affective functions of the forebrain. Some of the inflammatory cytokines, acting through the hypothalamus, can influence the activation of two highly significant components of the stress axes profoundly, the hypothalamic-pituitary-adrenal (HPA) axis and the sympathetic component of the autonomic nervous system (ANS). Therefore, the two great memory systems and communication systems of the body, the CNS and the immune system, are in constant communication with one another through biological signaling using endocrine hormones, neurotransmitters, and cytokines.

Communication between the Nervous System and the Immune System

Since the time of Galen, an association has been suspected between the mind or emotions and the clinical course or outcome of disease, as in Galen's description of melancholia and breast cancer. Early studies in humans pointed toward an association between psychosocial factors, particularly psychological stressors and the response of humans to various diseases thought to involve the immune system, such as bacterial and viral infections, allergic disorders, autoimmune diseases, and cancer. However, early correlational studies were unable to identify the communication channels through which such interactions between the nervous system and immune system could take place. Early studies in rodents

also demonstrated a link between psychological stressors and antibody responses, graft-versus-host responses, susceptibility to viral infections, and other reactions of the immune system, but, again, the neural communication channels were not elucidated in these early studies of the 1950s and 1960s.

Evidence for Neural-Immune Communication in Humans

Numerous studies of depressed individuals have reported diminished lymphocyte responses to proliferative stimuli (mitogens) or the diminution of other immune parameters such as natural killer (NK) cell activity or ratios of lymphocyte subsets. However, not all studies found such changes. Many variables influence the extent or presence of alterations in immunological parameters, including the type and severity of depression, time and treatment factors, and the age and gender of the patients. The most enduring findings reported that these diminished measures of immune response correlate best with the severity of depression and older age, particularly in men. The extent to which these diminished measures of immunological reactivity contribute to increased morbidity and mortality remains elusive.

Bereavement studies from the loss of a spouse also demonstrate diminished lymphocyte responses to mitogens and diminished NK cell activity, particularly in men. As in studies of other stressors in humans, the quality and extent of social support and the individuals' perception of control over their circumstances can act as buffers to ameliorate the stress-induced decline in measures of immunological reactivity.

The most extensive literature on neural-immune communications in humans is derived from studies of life stressors and psychosocial factors in selected groups of patients, particularly the studies of Janice Kiecolt-Glaser and Ronald Glaser. These investigators demonstrated diminished lymphocyte responses to mitogens, altered lymphocyte subsets, diminished NK cell activity, and increased antibody responses to viral epitopes on previously latent viruses in medical students during examination stress. Subsequent reports also demonstrated a diminished production of Th1 (cell-mediated) cytokines in medical students faced with examinations. Loneliness or lack of social support contributed to the presence and persistence of these alterations. Similar alterations in immune responsiveness (diminished mitogen responses and NK cell activity, and altered antibody or cytokine responses) have also been reported in a variety of other stressors, including marital separation and divorce, hostile behavior toward spouses in both recent and long-standing marriages, caregiving to a loved one with dementia (e.g., Alzheimer's disease),

and adaptation to a nursing home. In some of the marital studies, a significant impact of hostile behavior was noted on stress mediators, with the elevation of epinephrine and norepinephrine (NE), adrenocorticotrophic hormone (ACTH), growth hormone (GH), and prolactin. These important studies indicate direct links between behaviors or psychosocial factors, neuroendocrine and neurotransmitter mediators, and altered immune responses. It now appears that these links have direct implications for health through the modulation of wound repair, inflammation, NK cell response and tumor metastases, lymphocyte functions, and collective cell-mediated (vs. humoral) immune responses to infectious challenges by viruses and bacteria.

Positive emotions and positive factors also can modulate immune reactivity, often in a direction opposite to that of stressors. Social support interventions in nursing home subjects demonstrated an enhancement of some of the immunological parameters that were diminished previously. Lee Berk and colleagues have shown that mirthful laughter is accompanied by a diminution of stress mediators such as NE and cortisol and by enhanced NK cell activity and other immunological measures. Exercise that does not exceed 70–80% of capacity results in enhanced NK cell activity, enhanced lymphocyte mitogen responses, and enhanced measures of cell-mediated activity; exercise that exceeds these limits depresses many of these same parameters substantially. Due to the nature of the measurements of immune responsiveness in humans, limited mainly to blood samples, the extent to which these changes reflect the direct effect of mediators on individual cell functions, or the indirect effects of such parameters as altered cell trafficking, is not known. Some studies of various relaxation responses report the enhancement of some measures of immune response as well.

Studies in women with breast cancer by David Spiegel and colleagues and in patients with malignant melanoma by Fawzy Fawzy and colleagues have reported that structured psychological interventions involving counseling and peer support can extend longevity and enhance the quality of life. These provocative studies have evoked great interest in both scientific and lay communities and await further confirmation in multicenter trials and in studies in which neural-immune mediators and immunological measures are carried out in the same patients.

Caution is needed in studies of immune responses that accompany various stressors, positive behavioral and emotional factors, or structured interventions. Specific measures of immunological reactivity imply specific health consequences. For example, enhanced NK cell activity appears to aid in the innate response

to viral killing and in the killing of metastatic cancer cells, particularly when blood-borne. Enhanced cell-mediated immunity, production of Th1 cytokines such as interleukin (IL-2) and interferon (IFN- γ), boosts host responses to viral infections and appears to be protective in elderly individuals; human immunodeficiency virus (HIV)-positive individuals; and patients with a variety of cancers, including relatively nonimmunogenic cancers. However, enhanced cell-mediated T lymphocyte responses may be highly detrimental to an individual with an autoimmune disease such as multiple sclerosis; clinical trials with INF- γ showed an increased frequency of demyelinating attacks on the CNS. Thus, enhancing the immune system is a meaningless phrase unless we define (1) which parameters of immunological reactivity are enhanced, (2) which specific group of individuals is being studied, and (3) what specific immunological challenges the individuals are facing. The popular literature and lay understandings of this concept are virtually absent and need to be addressed.

Evidence for Neural-Immune Communication in Experimental Animals

A striking body of evidence for neural-immune communication in experimental animals comes from the conditioning studies of Robert Ader and Nicholas Cohen. Using a taste aversion model, they demonstrated that immune suppression from the administration of cyclophosphamide could be conditioned by a novel taste stimulus. This conditioned immune suppression was of sufficient functional magnitude that the conditioned rodents who were given the novel taste stimulus at a later time in their drinking water died of pneumonia; this agent had no effect in unconditioned animals. These investigators later studied conditioned immunosuppression in genetically autoimmune animals with a lupuslike syndrome that kills them early in life. These conditioned animals showed remarkably extended longevity without evidence of the lupus-related lymphadenopathy. As a further demonstration (one more pleasing to immunologists), Ader and Cohen demonstrated that an antigen itself (rather than a drug) could be paired with a novel taste stimulus to produce a conditioned enhancement of the magnitude of secondary antibody responses. Although the implications for application to humans facing difficult drug therapy with immune-reactive agents are highly promising, those studies are only now being pursued.

Studies in animals have investigated the immunological consequences of a wide range of physical and psychological stressors. In some circumstances, stressors such as extensive restraint result in suppressed NK cell activity, diminished lymphocyte proliferative

responses to mitogens, and even diminished cell-mediated and humoral immune responses. However, many complications arise in these stress and immunity studies. Sometimes, a stressor elicits enhanced immune responses rather than suppressed immune responses. Subtle differences in shock paradigms, related to frequency, intensity, and repetitions of the shocks, can produce the opposite effects, or no change at all, in immunological measurements. Even more perplexing is the finding that some stressors elicit the enhancement of one form of cell-mediated immune response and the suppression of a closely related but different cell-mediated immune response.

Thus, the impression expressed frequently in the lay literature that all stressors are bad for the immune system is simply incorrect. Some intriguing findings related to the exposure of rodents to stressors such as mild foot shocks involve the controllability of the stressors. Rodents who could not control the termination of the shocks showed a greater extent of immunological impairment than their counterparts who were able to control the termination of the shock and had less ability to respond to challenges from viruses, bacteria, or tumors. In general, acute laboratory stressors result in a wider and more complicated array of changes in immunological reactivity than chronic stressors. It is the chronic stressors that appear to be most likely to result in suppressed immune reactivity across a wide range of responses. Perhaps one of the most intriguing findings related to stressors and immune responses is their consequence for the progression of autoimmune diseases in experimental animals. Physical and psychological stressors, in general, protect against the progression of autoimmune symptoms and consequences; however, in the poststress interval, that benefit is lost or reversed, sometimes worsening the disease progression. A recent fascinating finding is the ability of IL-1 β to stimulate the corticotropin releasing hormone (CRH) system in the hypothalamus, thereby activating both the HPA axis (cortisol secretion) and the sympathetic nervous system (SNS; norepinephric, NE, neurotransmission). It appears that IL-1 β is one of the most potent activators of both of these stress axes compared with other physical and psychological stressors. Perhaps the immunological reactivity that results in the production of systemic inflammatory mediators should be considered at the top of the list of potent stressors in both experimental animals and humans.

One additional area of investigation that demonstrates a link between the nervous system and the immune system is the immunological consequences of direct lesions placed in the CNS. In general, lesion sites that result in altered immune responses are

found in the hypothalamus, the limbic forebrain (septum, hippocampal formation, amygdala, and some cortical regions), and brain-stem areas related to the regulation of autonomic outflow and processing or the reticular formation. These lesion studies suggest a complex array of circuitry that involves both excitatory and inhibitory neural connections that ultimately are able to alter the reactivity of specific portions of the immune system via neuroendocrine or autonomic outflow.

Communication Channels between the Nervous System and the Immune System

The CNS has two major routes of communication by which it can signal cells and tissues of the immune system. The first is the neuroendocrine channels that generally emanate from the hypothalamus and regulate the outflow of anterior and posterior pituitary hormones. Some of these hormones exert direct influences on immunocytes (cells of the immune system), whereas other anterior pituitary hormones act by stimulating additional hormones from target organs in the periphery. The second is the direct neural channels that provide anatomical nerve fiber connections between the CNS and lymphoid organs and tissues. These neural channels are either part of the ANS (and connect via a two classical two-neuron chain to the target structure) or part of the one-neuron sensory (afferent) connections from the periphery into the CNS. Some primary afferents are capable of releasing neurotransmitters both proximally, into the CNS, and distally, into the site of distribution of the sensory receptive elements. The autonomic nerve distribution to lymphoid organs and tissues derives principally from the sympathetic division of the ANS.

The immune system can signal neurons or supporting cells of the CNS and PNS principally via the release of cytokines. The cytokines can act on peripheral neuronal elements in the local paracrine environment and can act on CNS cells either through the actions of circulating cytokines or through the activation of primary afferents that activate central afferent autonomic circuitry.

Immunocytes Possess Receptors for Hormones and Neurotransmitters

Immunocytes, including lymphocytes (both B and T cells), macrophages/monocytes, granulocytes, thymocytes, stem cells, and other associated cell types, possess receptors for hormones and neurotransmitters. Some hormone receptors are present on the plasma membrane of immunocytes; their stimulation results in the altered activation of ion channels or the altered production of second messengers. Other

hormone receptors are found intracellularly and, when combined with a properly configured hormone (ligand), can exert genomic influences by acting as transcription factors or in combination with other transcription factors. The hormone–receptor complex can also alter intracellular transduction pathways. In addition, immunocytes possess neurotransmitter receptors on the plasma membrane. When activated by the appropriate ligand, these receptors on immunocytes can influence intracellular events through one or more second-messenger pathways. Some receptors for neurotransmitters (e.g., the β adrenoceptor) appear to be identical to their counterparts on neurons and peripheral target organs of the SNS, whereas other receptors (e.g., the substance P, SP, receptor) appear to be configured slightly differently from their neuronal counterparts. Immunocytes possessing this SP receptor respond more broadly to SP and substance K elements of the tachykinin family than does a neuron with its more narrowly responsive SP receptor.

Neurotransmitter and hormone receptors are not expressed by all immunocytes and, when expressed, are not distributed uniformly across all subsets or during all states of activation. For example, Virginia Sanders and others have demonstrated that β adrenoceptors, responsive to catecholamines such as NE and epinephrine, are expressed more abundantly on B lymphocytes than on T lymphocytes and more abundantly on T suppressor/cytotoxic (CD8) cells than on T helper/inducer (CD4) cells. Although β adrenoceptors are present on immature thymocytes, they increase in density during maturation. Activated T cells possess higher numbers of β adrenoceptors than resting T cells. Lymphocytes from aged animals possess higher numbers of β adrenoceptors than their younger counterparts. At first glance, this suggests an increasing sensitivity of these cells to catecholamines with age; however, these receptors on aged lymphocytes are poorly coupled with their G-protein and show a diminished transduction of cAMP via adenylate cyclase. Thus, many variables may influence the ability of neurotransmitters to influence immunocytes.

Immunocyte subsets possess receptors for a variety of neurotransmitters, including both α and β adrenoceptors, dopamine receptors, muscarinic cholinergic receptors, excitatory amino acid receptors, and peptidergic receptors (SP; calcitonin gene-related peptide, CGRP; somatostatin; vasoactive intestinal peptide, VIP; and opioid peptides, to name a few).

The signaling of immunocytes through neurotransmitter–receptor interactions may initiate direct intracellular changes that can be detected readily, or they may exert indirect effects that become apparent

mainly during signaling from a second ligand. For example, activation of the T-cell receptor with anti-CD3 antibody will result in the more robust production of the second messenger cAMP when β adrenoceptor activation is present than when it is absent. It is likely that neurotransmitters often act in synergistic (or countersynergistic) fashion with other neurotransmitters, hormones, cytokines, growth factors, and other signals; in this way, they resemble the dual-signaling and modulatory capacities of cytokines.

Neurohormonal Signaling of Immunocytes

A wide range of hormones, including those of the anterior pituitary, the posterior pituitary, and peripheral glands, can signal cells of the immune system and exert profound effects on the ability of the immune system to mount an appropriate defense response. The best-known hormonal system with regard to immune function, the HPA system (acting via CRH, ACTH, and cortisol), plays such an important and integral role in basic immunophysiology that Hugo Besedovsky proposed that the HPA axis is an integral component of the immune system itself. Although even a basic review of the hormonal effects on immune reactivity is far beyond the scope of this article, a few highlights are mentioned in selected areas.

The hypothalamic-pituitary-adrenal axis CRH, produced mainly in the paraventricular nucleus (PVN) of the hypothalamus, is released from axons into the private vascular channel of the hypophyseal-portal system. The resultant high concentration of CRH stimulates the release of ACTH from corticotrophs in the anterior pituitary. ACTH is released into the general circulation and activates cells in the adrenal cortex to release cortisol into the general circulation in humans (corticosterone in rodents). Both ACTH and cortisol exert effects on a variety of immunocytes. Interestingly, CRH is found in sites in addition to the PVN, such as macrophages, in a variety of peripheral organs. Thus, even CRH may exert hormonal influences on immunocytes directly, in addition to its contribution in activating ACTH and cortisol production.

CRH-producing neurons in the PVN are a major target of action for circulating inflammatory cytokines, especially IL-1 β , acting either directly or indirectly to enhance the activity of the CRH neurons. CRH neuronal activation resulting from either peripheral or central IL-1 β stimulation enhances the production of ACTH and cortisol via the HPA axis and, in addition, activates the hypothalamic-sympathetic axis, thus enhancing both of the major classical stress axes physiologically. Esther Sternberg

and colleagues showed that CRH hypothalamic neurons are important regulatory cells in autoimmune diseases. She studied genetically autoimmune-prone animals, such as the Lewis/N rat, to identify potential defects contributing to the susceptibility of these rats to autoimmune diseases. She found a lack of responsiveness of the hypothalamic CRH neurons following attempts to activate them by psychological stressors, neurotransmitter signaling, and the administration of cytokines. Sternberg concluded that the diminished capacity of these rats to activate the HPA axis following appropriate stimuli contributed to their inability to generate a cortisol-evoked suppression of unwanted cell proliferation during immune challenge, particularly by self-reactive clones of cells. When the hypothalamic CRH neurons in closely related autoimmune-resistant rat, the Fischer 344, were lesioned or blocked, these rats became as susceptible to autoimmune-provoking stimuli as Lewis/N rats. These important findings suggest that Besedovsky was correct in considering the HPA axis to be a vital physiological component of the immune system.

For many years, cortisol and other glucocorticoids were thought to be major immunosuppressive hormones. Dexamethasone and other such steroids were often given to patients for their anti-inflammatory or immune suppressing effects, including the treatment of autoimmune diseases. Studies of glucocorticoid effects on immunocytes have reported lymphocyte and thymocyte apoptosis; inhibition of T-cell proliferation; inhibition of cytokine production, such as IL-1 and IL-2; and suppression of a host of innate and acquired immune responses. Interestingly, work by Daynes and Araneo noted that physiological concentrations of glucocorticoids, in contrast to the massively suppressive effects of pharmacological doses of these agents, shifted the balance of cytokine production in T-helper lymphocytes from Th1 (diminished IL-2 and IFN- γ) to Th2 (enhanced IL-4 and IL-10), thus altering acquired immune responses away from cell-mediated immunity (Th1) and toward humoral immunity (Th2). Many researchers and clinicians have proposed that the maintenance of robust cell-mediated immunity is critical for protection against viral infections and some tumors and is necessary for maintaining a relative state of health in patients with HIV infection, some cancers (including some that are relatively nonimmunogenic), and immunosenescence in the elderly. ACTH has also been noted to exert a wide range of immune effects, including reduced IFN- γ production, increased NK cell activity, and complex effects on antibody production.

Dehydroepiandrosterone (DHEA) is an adrenal hormone that some investigators have reported as countering the effects of cortisol. DHEA can enhance

some cell-mediated immune responses, restore some immune responses in aged animals, and block some of the immune-suppressive effects of cortisol. The DHEA : cortisol ratio may be a good index of eustress activity (with positive effects on cell-mediated immune responses, such as moderate exercise), in contrast with distress from aversive stimuli and activities, such as maximal physical exertion in marathon running.

Growth hormone and prolactin GH is a hormone produced by cells of the anterior pituitary gland. A GH deficiency in many animals results in small, hypocellular primary and secondary lymphoid organs. The reconstitution of these deficient animals with GH restores cellularity to the lymphoid organs; restores T-cell development in the thymus; and restores antibody responses, cell-mediated responses, and NK cell activity. GH also potentiates cytotoxic T-lymphocyte activity. In animals with involution of the thymus, the administration of GH and prolactin restores thymic morphology and many T-lymphocyte functions. This reinforces the contention that GH exerts a particularly important influence on thymic function. GH also primes macrophages and neutrophils to produce superoxide anion, a functional component of phagocytosis. Surprisingly, GH is also produced by lymphocytes themselves and may act as a paracrine or autocrine signal for lymphocyte proliferation.

Prolactin is another hormone produced by cells of the anterior pituitary. The administration of prolactin contributes to the restoration of immune function in hypophysectomized animals and acts in a similar fashion with GH to help restore thymic cellularity and function in animals with involuted thymuses. Prolactin blockade decreases host resistance to some bacterial infections, such as *Listeria monocytogenes*. Mechanistic studies point toward a role for prolactin as a permissive factor or progression factor, necessary but not sufficient for IL-2-induced T-lymphocyte proliferation. A prolactin-like molecule has also been identified as a product of concanavalin A (Con A)-stimulated T lymphocytes and may be an important autocrine or paracrine lymphocyte growth factor, particularly in conjunction with a possible synergistic effect with estrogen.

Opioid peptides The opioid peptides are families of low-molecular-weight peptides, the endorphins, the enkephalins, and the dynorphins. They are found in the CNS, the pituitary gland, and peripheral structures such as the adrenal gland. In the brain, opioid peptide interactions with their receptors can induce

analgesia, exert profound behavioral changes, and alter outflow channels to the periphery that can modulate immune responses. Some of these changes can be blocked by naloxone and other opioid blockers, whereas others cannot. The opioid peptides are cleaved from larger prohormone peptides. For example, the precursor proopiomelanocortin (POMC) is cleaved to ACTH and β -lipotropin, the latter of which gives rise to β -endorphin. β -Endorphin levels increase in the serum in parallel with ACTH levels following some stressors, suggesting that β -endorphin may be an important stress-induced neuromediator with an influence on immunocytes or immune responses. Although a review of opioid peptide actions is beyond the scope of this article, we briefly mention of some immune effects of β -endorphin.

β -Endorphin reduces antibody production in mice through μ - and κ -opioid-receptor-mediated effects. β -Endorphin inhibits lymphocyte chemotactic factor in humans and may itself act as a chemotactic factor. Some additional β -endorphin effects are thought to be nonopioid-receptor-mediated, including enhanced Con A-induced T-lymphocyte proliferation in mice, reduced phytohemagglutinin (PHA)-induced T-lymphocyte proliferation in humans and the blockade of prostaglandin E1-induced suppression of lymphocyte proliferation in rats. The opioid peptides exert very complex and varied actions on immunocytes *in vivo* and *in vitro*; sometimes exert contradictory effects *in vitro* versus *in vivo*; and can act in synergistic fashion with other neural, endocrine, and immunological mediators.

Other hormones Many other hormones, in addition to those mentioned earlier, have been shown to influence immunocytes and immune function in a variety of *in vitro* and *in vivo* models. These hormones include thyroid-stimulating hormone (TSH), thyroid hormones, gonadal steroid hormones (estrogen and androgens), posterior pituitary hormones (oxytocin and vasopressin), and melatonin (produced by the pineal gland).

Nerve Fiber Connections with Lymphoid Organs: The Hard Wiring

Peripheral nerve fibers are found in primary lymphoid organs (the bone marrow and thymus), in secondary lymphoid organs (the spleen and lymph nodes), in mucosal-associated lymphoid tissue (the gastrointestinal tract and lung), and in the skin. Tracing studies have revealed the location of some (but not all) of these connections. The two major sources of connectivity are the ANS and the primary sensory system. In the ANS, the vast majority of connections derive

from the sympathetic division via a two-neuron chain, with the postganglionic component releasing NE as a major neurotransmitter. These SNS connections distribute adjacent to at least some components of all collections of lymphoid organs and tissues. Very few connections to lymphoid tissues derive from the parasympathetic division of the ANS. Classic two-neuron connections of the cholinergic parasympathetic system have not been identified in either the primary or secondary lymphoid organs; indeed, there is little evidence for cholinergic neurotransmission in these organs, despite the presence of muscarinic cholinergic receptors on some immunocytes found there. Parasympathetic connections that may influence immunocytes are found mainly in mucosal-associated lymphoid tissue and the skin.

A wide range of neuropeptides has been identified in nerve fibers in the lymphoid organs. Some of these neuropeptides (SP and CGRP) are thought to be associated mainly with sensory connections. However, very few tracing studies have demonstrated a primary sensory (afferent) origin for this neuropeptidergic innervation of lymphoid organs and tissues. Some peptides (neuropeptide Y, NPY) appear to be colocalized with other neurotransmitters in cells of the SNS. Some primary afferents, particularly the small SP fibers, have been shown to release their neurotransmitter from both proximal and distal arborizations of their axons. Thus, it is possible for a primary afferent cell to be activated and then release the neurotransmitter at the site in the periphery where the initial activating stimulus took place. It is clear that a much more thorough mapping of the cells of origin, distribution, and associations with target cells in the immune system needs to be accomplished for the nerve fiber connections that provide the innervation to cellular targets in the organs and tissues of the immune system.

Sympathetic noradrenergic nerve fibers Postganglionic sympathetic NE nerve fibers supply primary lymphoid organs, secondary lymphoid organs, and mucosal-associated lymphoid organs with innervation. These nerve fibers distribute along the vasculature, travel with smooth muscle structures such as the trabeculae, and arborize directly into the parenchyma where they form close junctional appositions (neuro-effector junctions) with cells of the immune system. The innervation that extends into the parenchyma is associated with specific compartments of lymphoid organs and maintains that compartmentation even when involution or loss of cellularity occurs. Some neuroimmunocyte appositions are as close as 6 nm (a gap junction is 2 nm, and a classical synapse in the

CNS is 20 nm). Felten and colleagues hypothesized that communication in the parenchyma of lymphoid organs occurs in both directions: (1) neurotransmitters released by sympathetic nerve endings can act on receptors on target cells of the immune system and (2) cytokines released by target cells of the immune system can modulate the extent of release of neurotransmitters such as NE. For the most part, sympathetic NE nerve terminals also colocalize NPY, although there are some terminals that contain only one and not the other. Extensive studies from the laboratory of David Felten and colleagues identified abundant innervation of lymphoid organs.

NE nerve fibers enter the bone marrow with the nutrient arteries, distribute along those arteries, and arborize extensively into the parenchyma, where they associate with stem cells and other constituents of the marrow. Most of these fibers appear to colocalize NPY. Many cells of the bone marrow possess adrenoceptors, especially β adrenoceptors. Studies using a three-dimensional bone marrow culture showed that catecholaminergic stimulation of bone marrow can augment granulocyte production. NE nerve fibers enter the thymus along the capsular surface and distribute through the septal system. Some NE fibers distribute into the thymic cortex and medulla and extend along the vasculature at the corticomedullary junction. Thymocytes possess adrenoceptors, with increasing numbers noted during thymocyte maturation. Some investigators have reported that catecholamines can suppress thymocyte proliferation through the activation of adenylate cyclase and generation of cAMP; however, this interaction may be more likely to occur during involution or during diminished cellularity, at which time the NE innervation is very dense.

NE nerve fibers enter the spleen with the splenic artery, travel along the trabecular system, and distribute extensively into the splenic white pulp. In the white pulp, NE nerve fibers follow the central artery of the white pulp, extend along the marginal sinus (where macrophages reside) and into the marginal zone, and distribute extensively into the periarteriolar lymphatic sheath. NE nerve terminals form close associations with T lymphocytes of both CD4 and CD8 subsets and with macrophages and NK cells. Very few NE fibers are found in the B-lymphocyte zones, the follicles. NE innervation appears to enhance cell-mediated immunity during the initiative phase of an immune response and has some inhibitory effect on humoral immunity in Th1-dominant animals.

NE nerve fibers enter lymph nodes with the vasculature in the hilar zone and with the capsular surface. These fibers distribute through the medullary cords,

where they associate with macrophages and other cells and arborize into the cortical and paracortical zone where T lymphocytes of both CD4 and CD8 subsets reside. As in the case of the spleen, very few NE terminals distribute into the B-lymphocyte zones, the nodules. In lymph nodes, NE fibers also enhance cell-mediated immunity and, in some cases, inhibit humoral immunity, particularly in the case of autoimmune-provoking stimuli. The denervation of these NE fibers in both the spleen and lymph nodes results in diminished cell-mediated immune responses, enhanced humoral responses in Th1-dominant animals, and increased severity of autoimmune-provoking stimuli. NE also appears to exert extensive influences on cell trafficking, particularly on NK cells and T lymphocytes. Even very low concentrations of catecholamines can result in the release of marginated and organ-localized NK cells into the circulation. Interestingly, during aging, the spleen and lymph nodes progressively lose their NE innervation, according to Denise Bellinger and colleagues, due to autodestruction by oxidative metabolites of catecholamines released into the paracrine environment by NE nerve terminals in these secondary lymphoid organs. When the NE innervation is restored in old animals, with the administration of growth factor-stimulating pharmacological agents (e.g., low-dose deprenyl), IL-2 and IFN- α production is enhanced markedly, NK cell activity is elevated, and T-cell proliferation is restored.

Noradrenergic innervation is also found in mucosal-associated lymphoid organs. In the gastrointestinal tract, these fibers distribute through T-lymphocyte zones and arborize in submucosal zones.

Neuropeptide-containing nerve fibers Extensive networks of neuropeptide-containing nerve fibers distribute into both the primary and secondary lymphoid organs. A few neuropeptides, such as NPY, are colocalized with NE sympathetic nerve terminals, but most neuropeptide-containing nerve fibers are independent of the known sympathetic innervation.

Neuropeptide Y-containing terminals appear to colocalize with NE as a component of the sympathetic innervation in the bone marrow, thymus, spleen, and lymph nodes, although some non-NE NPY+ nerve terminals are present in the lymphoid organs. NPY is also found in the platelets in the spleen and is found in the circulation, apparently from release following sympathetic activation. Some immunocytes possess receptors for NPY.

SP and CGRP are found frequently as colocalized neuropeptides in non-NE nerve terminals in all lymphoid organs and tissues, as well as the skin. SP nerve fibers have been identified along the

vasculature and in the parenchyma in the bone marrow, thymus, spleen, and lymph nodes. CGRP is often colocalized with SP; however, in some sites, CGRP-containing nerve fibers distribute independently (e.g., adjacent to antigen-presenting Langerhans cells in the skin).

SP receptors have been identified on both T and B lymphocytes as well as on granulocytes and mast cells. SP is a potent regulator of inflammation and can induce extravasation and edema, attract other pro-inflammatory cells and mediators, and exacerbate inflammation in sites such as joint spaces in rheumatoid arthritis. SP can also enhance polyclonal immunoglobulin (Ig)A and IgM production by gut-associated lymphoid tissue, enhance T-cell proliferation and IL-2 production in the secondary lymphoid organs, and increase some antibody responses and lymphocyte-proliferative responses. Many of these immunoenhancing effects of SP can be counteracted by somatostatin, but an extensive understanding of somatostatin innervation of lymphoid organs is not yet available. In the skin, CGRP acts as a potent inhibitory factor on antigen presentation by Langerhans cells.

VIP-containing nerve fibers are present in Peyer's patches in the gut, thymus, and spleen. VIP receptors are present on some T lymphocytes; Cliff Ottaway showed that their expression can enhance the migration of these T cells into the gut-associated lymphoid tissue and the mesenteric lymph nodes, whereas their absence prevents such migration. VIP can also inhibit T-lymphocyte proliferation.

Many other neuropeptides have been found in the primary and secondary lymphoid organs. These neuropeptides include the opioid peptides, cholecystokinin, neurotensin, CRH, somatostatin, and galanin. Little is known about the details of the functioning of these neuropeptide-containing nerve fibers. It is particularly difficult to discern the unique role of specifically compartmentalized neuropeptides, such as the opioid peptides, that appear in neurons, pituitary cells, and immunocytes.

The origin of cell bodies for most of the neuropeptide-containing nerve fibers that distribute into the lymphoid organs and tissues is not known. Thus, we do not even know if they are components of the primary afferent system, the sympathetic or parasympathetic divisions of the ANS, or the enteric nervous system. Neuropeptide receptors are present on many specific subsets of lymphocytes, macrophages/monocytes, and other lymphoid cells. We do not yet know the details of the distribution of these receptors, the details of the second-messenger pathways in specific subsets of immunocytes, or the functional role of these receptors in immunoregulation.

Immunocyte Production of Neuropeptides

Ed Blalock, Eric Smith, and colleagues made the remarkable discovery that cells of the immune system, particularly lymphocytes and macrophages/monocytes, can synthesize neuropeptides directly. They showed that some immunocytes can synthesize the POMC gene following stimulation by a virus (Newcastle's disease) or a releasing factor (CRH). The resultant neuropeptides produced and released by the immunocytes are, for the most part, identical to their counterparts made by neurons and pituicytes. Immunocytes have been reported to produce the POMC-derived products ACTH and β -endorphin, TSH, and CRH. In some instances, the same neuropeptide can be produced by neurons, pituicytes, and immunocytes in the periphery.

Cytokine Interactions with the Nervous System

Immunologically derived mediators, the cytokines, can act directly or indirectly to influence the activity of neurons in the CNS and PNS. In addition, some supporting cells of the nervous system (microglia and astrocytes), and even a small subset of neurons in the CNS and PNS, can produce cytokines directly, particularly IL-1 β . The most extensively studied cytokines with regard to effects on CNS neurons and CNS behavior are the pro-inflammatory cytokines, IL-1, IL-6, and tumor necrosis factor (TNF)- α . IL-1 β can act on neurons in the CNS, particularly on hypothalamic neurons, and can induce fever, slow-wave sleep, lethargy, loss of appetite, diminished reproductive behavior, and other classic illness behaviors. Interestingly, this illness behavior can be induced by IL-1 β injected by intracerebroventricular (ICV), intraperitoneal (IP), or intravenous (IV) routes; the ICV route requires lower concentrations of IL-1 β to induce the illness behavior than the IP or IV routes.

Many mechanisms have been proposed by which IL-1 β may exert its actions on the CNS:

1. IL-1 β may cross the blood-brain barrier directly, enter into the cerebrospinal fluid, and then act directly on neurons. Banks and colleagues reported such a transport system, although the greatest extent of crossing of IL-1 into the CNS occurs in the cortex, not the hypothalamus.
2. IL-1 β may act on the cells of the organum vasculosum of the lamina terminalis to release mediators such as prostaglandin E₂, which, in turn, activates preoptic neurons.
3. In the circulation, IL-1 β may act on endothelial cells in the brain vasculature, resulting in the

release of mediators such as nitric oxide and prostaglandins that, in turn, activate neurons.

4. IL-1 β may act on vagal afferents (via paraneurons) or on other afferents, according to the work of Watkins and Maier, stimulating cells in the nucleus solitarius complex, which, in turn, activate the hypothalamus.

It is likely that a combination of all of these routes operates to some extent in physiological circumstances.

IL-1 β also can activate the HPA axis through its actions on the CRH neuronal system of the PVN. To a lesser extent, so can IL-6 and other cytokines. This activation of the CRH system requires the input of the A1 and A2 brain-stem NE projections to the PVN, suggesting an action of IL-1 β on the central NE nerve terminals. IL-1 β also activates the hypothalamic-spinal cord system, which enhances sympathetic outflow from the intermediolateral cell system of the thoracolumbar spinal cord. Thus, cytokines that activate the HPA axis and the SNS can, in and of themselves, be considered as classical stressors. In addition, viral infections and immune responses that result in the production of pro-inflammatory cytokines, such as IL-1 β , also should be considered as classic stressors.

Inflammatory mediators also appear to be involved in some neurodegenerative diseases, such as Alzheimer's disease. IL-1 β and many other inflammatory mediators, complement components, and acute-phase proteins can be found at the site of plaque formation and neuronal degeneration in the brains of patients with Alzheimer's disease. The long-term administration of nonsteroidal anti-inflammatory medication or specific COX-2 inhibitors appears to diminish the inflammatory response and may protect against the progression of this disease.

In the PNS, both IL-1 and IL-2 appear to act on nerve terminals of the sympathetic NE system in secondary lymphoid organs and modulate the release of NE. These findings of cytokine modulation of central and peripheral neuronal activities indicate that the immune system and nervous system are in constant communication with one another on a continuing physiological basis.

Clinical Implications of Neural-Immune Signaling

At the most basic level, neurohormones and neurotransmitters can regulate the individual cellular functions of immunocytes, such as the proliferation, differentiation, maturation, expression of cell adhesion molecules (influencing trafficking); expression of specific cell markers; synthesis of specific cell

products (cytokines, immunoglobulins, and growth factors); or expression of specific portions of the genome. This can occur wherever the cell encounters an appropriate ligand, whether in an organ or in the blood.

At a more integrative level, neurohormones and neurotransmitters can also modulate immune responses, the collective interactions of multiple cell types that result in a reaction to challenge. These neural signals may either modulate aspects of innate immunity (e.g., location, number, and activity of NK cells) or modulate the acquired immune responses of both cell-mediated and humoral types. The final outcome depends on the collective impact of the specific signal molecules on antigen-presenting cells, accessory cells, lymphocytes, and any other cell types that contribute to the overall response.

The final clinical outcome is neural modulation of (1) the host response to viral, bacterial, or parasite infections, including both progression and latency; (2) wound healing; (3) inflammatory responses; (4) autoimmune reactivity; (5) immune senescence in old animals; and (6) the progression, metastatic spread, and recurrence of some specific cancers. In some instances, a single neurally derived signal may exert a predominant influence. In other instances, the collective interaction of glucocorticoids, other neurohormones, catecholamines, and neuropeptides all may contribute to the end effect. Studies demonstrate that many of these neural signals synergize with cytokines, growth factors, hormones, and other neural signals. This suggests that elucidating the integrated physiological influence of neural signaling on immune reactivity, and its potential to alter the outcome of disease processes, is a daunting but rewarding challenge.

See Also the Following Articles

Autoimmunity; Breast Cancer; Cytotoxic Lymphocytes; Immune Response; Immune Suppression; Natural Killer (NK) Cells; Opioids; Social Support.

Further Reading

- Ader, R., Cohen, N. and Felten, D. L. (1995). Psychoneuro-immunology: interaction between the nervous system and the immune system. *Lancet* **345**, 99–103.
- Ader, R., Felten, D. L. and Cohen, N. (1990). Interactions between the brain and the immune system. *Annual Review of Pharmacology and Toxicology* **30**, 561–5602.
- Ader, R., Felten, D. L. and Cohen, N. (1991). *Psychoneuro-immunology* (2nd edn.). San Diego: Academic Press.
- Bellinger, D. L., Madden, K. S., Felten, S. Y. and Felten, D. L. (1994). Neural and endocrine links between the brain and the immune system. In: Lewis, C. E., O'Sullivan, C. & Barraclough, J. (eds.) *The psychoimmunology of cancer*, pp. 55–106. New York: Oxford University Press.
- Felten, D. L., Felten, S. Y., Bellinger, D. L., et al. (1993). Fundamental aspects of neural-immune signaling. *Psychotherapy and Psychosomatics* **178**, 1–10.
- Kiecolt-Glaser, J. K. (1999). Stress, personal relationships, and immune function: health implications. *Brain, Behavior & Immunity* **13**, 61–72.
- Madden, K. S. and Felten, D. L. (1995). Experimental basis for neural-immune interactions. *Physiological Reviews* **75**, 77–106.
- Madden, K. S., Sanders, V. M. and Felten, D. L. (1995). Catecholamine influences and sympathetic neural modulation of immune responsiveness. *Annual Review of Pharmacology and Toxicology* **35**, 417–448.
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* **338**, 171–179.
- Reichlin, S. (1993). Neuroendocrine-immune interactions. *New England Journal of Medicine* **329**, 1246–1253.
- Sheridan, J. (1998). Stress-induced modulation of anti-viral immunity. *Brain, Behavior & Immunity* **12**, 1–6.
- Watkins, L., Maier, S. and Goehler, L. (1995). Cytokine-to-brain communication: a review and analysis of alternative mechanisms. *Life Sciences* **57**, 1011–1026.
- Wilder, R. L. (1995). Neuroendocrine-immune system interactions and autoimmunity. *Annual Review of Immunology* **13**, 307–338.

Neuroinflammation

G Z Feuerstein, R R Ruffolo and L J Rutkowski

Wyeth Research, Collegeville, PA, USA

© 2007 Elsevier Inc. All rights reserved.

Inflammation and Neuroinflammation: General Considerations

Neuroinflammation in Stroke

Neuroinflammation in Multiple Sclerosis

Neuroinflammation in Alzheimer's Disease

Summary

Glossary

<i>Alzheimer's disease</i>	A neurological disorder of progressive loss of cognitive functions occurring mostly in the elderly due to the accumulation of protein deposits that damage neurons and activate glial cells.
<i>Alzheimer's plaque</i>	Insoluble deposits of aggregated β -amyloid protein that accumulates in the brain and cerebral blood vessels of patients with Alzheimer's disease. Amyloid plaque triggers inflammatory responses that contribute to the death of neurons and progressive dementia.
<i>Cytokines</i>	A family of secreted proteins originally identified as products of leukocytes that serve in innate and specific immune and inflammatory reactions. Cytokine regulate diverse functions including cell growth, differentiation, function, and disease and are regarded as pleiotropic mediators.
<i>Inflammation</i>	A localized protective response that follows tissue injury and aims to destroy hostile agents, resolve tissue damage, and aid in repair and rehabilitation. Inflammation involves local vascular responses, leukocyte recruitment, and tissue cell activation stimulated by both local and systemic humoral factors.
<i>Interleukins</i>	A family of cytokines produced by leukocytes in natural immune reactions that serve to activate or modulate the function of other leukocytes as well as non-hematopoietic cells.
<i>Leukocytes</i>	White blood cells that are produced in the bone marrow and lymphatic tissue and circulate in the blood. White blood cells are responsible for immune

Lymphocytes

Magnetic resonance imaging (MRI)

Multiple sclerosis

Neurofibrillary tangles

Preconditioning

Prostaglandin

Stroke

surveillance, innate immunity, and inflammation as part of host defense and tissue repair. In certain conditions, certain leukocyte subsets can mediate toxic reactions that lead to autoimmune diseases and chronic inflammatory diseases. A subset of white blood cells that are responsible for innate and specific immune reactions. T-lymphocytes are the central arm of the acquired cellular immunity and carry receptors that largely recognize antigens of foreign origin. Following clonal expansion, T-cells acquire the ability to clear the foreign invaders. B lymphocytes are the central arm of the humoral immune response. They largely recognize soluble antigens via immunoglobulin on their surface, upon which they differentiate to memory B cells.

A noninvasive imaging method that deploys brief magnetic pulses that polarize hydrogen atoms and then monitors energy emission. High-resolution images of soft tissue, particularly brain and spinal cord, are generated by computer algorithms. Tissue structure imaged via computerized algorithms can be deduced by differences in emission densities.

A neurological disease characterized by discrete white matter lesions (demyelination) and typified by intermittent, progressive, yet variable loss of motor and sensory (speech and visual) function and sometimes death.

Paired helical filaments of tau protein that accumulate inside neurons. Neurofibrillary tangles are one of the pathological hallmarks used to confirm the diagnosis of Alzheimer's disease at autopsy.

Short and limited stress conditions that modulates tissue vulnerability to subsequent injurious events.

Lipid mediators that are produced from fatty acids and serve in local and systemic functions of diverse tissues and organs. Certain prostaglandins participate in immune and inflammatory reaction, including leukotriene B₄ and thromboxane A₂.

Abrupt ischemic or hemorrhagic event in the brain that leads to loss of motor, sensory, and cognitive functions or death.

Inflammation and Neuroinflammation: General Considerations

Inflammation is a fundamental, carefully orchestrated, and highly conserved reaction of the organism meant to protect, defend, and amend diverse forms of injury. Inflammation (as well as innate immune responses) is critical for the organism's survival, yet in specific occasions, these reactions could result in a detrimental outcome. While inflammation features prominently in injuries to peripheral tissue, the situation is quite different in the central nervous system (CNS). The brain has long been considered an immunologically privileged organ. This notion is based on observations that tissue transplants into the CNS are not well rejected by the immune system, and inflammatory cells and mediators are sparse compared to peripheral tissues. The anti-inflammatory and anti-apoptosis environment in the brain and the limited access of brain-derived antigens to lymphoid organs may contribute to its immune seclusion. However, recent studies in infectious and autoimmune models have challenged this view. Injuries to the CNS activate cells within the CNS, called microglia, that like the peripheral monocyte system are capable of releasing a plethora of inflammatory mediators as well as activation of immune permissive molecules such as MHC class I. Proteins that are released from brain cells during injury can provoke acute and chronic inflammatory reactions in which both brain cells and circulating immune inflammatory cells participate.

Neuroinflammation in Stroke

Stroke Prevalence and Epidemiology

Stroke is a major medical emergency due to the resulting high rate of mortality and morbidity. In the United States alone, about 750,000 strokes, resulting in about 150,000 deaths, are estimated annually, which makes stroke the third leading cause of death in the United States. In addition, stroke management incurs a high burden of health costs due to prolonged rehabilitation, permanent (~30%) residual deficits, and loss of productive professional capabilities. The incidence of stroke has been on the decline from 1980 to 1990, primarily due to advances in aggressive management of key risk factors such as hypertension, atherosclerosis, and lipid disorders; lifestyle modifications (smoking); and in recent years, improved time to admittance to health facilities that are better equipped and organized to diagnose and treat strokes in specialized stroke care units. However, the incidence of stroke has not decreased in recent years in most of the industrialized countries while the prevalence has

actually increased due to the aging population and increased rates of comorbidities such as diabetes, heart failure, and the metabolic syndrome. In spite of the medical severity of stroke, its prevalence, and health burden, little progress has been realized in the treatment of this condition. The only specific, pharmacological treatment registered for stroke is tissue plasminogen activator (tPA), which enhances thrombolysis. This agent, however, finds limited utility (2–5% of admitted patients) due to substantial limitation of time to administration (3 h) and associated toxicity. In addition, no mortality benefits have been proven in prospective clinical trials with thrombolysis agents.

The reasons for the difficulties in discovering and developing new drugs for treatment of stroke are complex. Likely causes for failure to discover new drugs include poor understanding of the mechanism of neuronal death in stroke, which results in development of inadequate therapeutics. Also, the lack of proper experimental models for this human condition and limited insightful translational medicine may contribute. Recently, recognition of the role of inflammation in stroke has emerged, which may provide new opportunities to discover and develop effective treatment for stroke.

Cells and Mediators in Neuroinflammation in Stroke

Cerebral ischemia is marked by rapid activation of microglia, the brain cells that are part of the innate immune and inflammation cellular elements along with infiltration of inflammatory cells from the circulation into the ischemic territory (see [Figure 1](#)). Experimental stroke models revealed that early accumulation of blood-borne inflammatory cells in the ischemic brain persists for days to weeks after the initial insult. The early infiltration of polymorphonuclear cells (also called neutrophils) into the capillaries of the ischemic zone parallels the early activation of the glia cells (microglia and astroglia). Monocytes arrive in the injury zone much later, around 12–24 h after the ischemic injury, and are rapidly transformed into tissue macrophages displaying aggressive phagocytosis. Other immune/inflammatory cells such as lymphocytes are also observed at later time points. The significance of the exogenous leukocyte infiltration into the ischemic brain is controversial. The presence of leukocytes in the capillaries was argued to contribute to low or no flow in the capillaries due to clogging that exacerbates the ischemic insult. Furthermore, the gamut of toxic mediators released by the activated leukocytes, e.g., reactive oxygen radicals, proteases, and cytokines, has also been suspected to cause

inhibiting the polymorphonuclear leukocyte adhesion to the endothelium) have failed to show efficacy.

Chemokines

Chemokines are a family of more than 50 structurally and functionally related proteins involved in inflammatory cell recruitment. These relatively small proteins are usually subdivided into four subfamilies on the basis of their structural conservation of cysteine residues. Chemokines and their receptors are constitutively expressed in normal brain cells such as microglia, astrocytes, and neurons throughout the CNS, albeit at very low levels. As shown in [Figure 1](#), rapid expression of many chemokines takes place in response to acute ischemic brain injury, which corresponds with cytokine and leukocyte adhesion molecule expression. Most prominent activation has been documented for MCP-1 (monocyte attractant protein-1), interleukin-8, interferon-inducible protein-1, macrophage inflammatory protein-1, and fractalkine. While these chemokines and others seem to be expressed in the location and at the time that could lend them a role in propagation, aggravation, or repair of the ischemic injury, no agent to date has been deployed in clinical trials in stroke or other brain injury conditions. Experiments with prototypic inhibitors of certain chemokines have shown the potential for better recovery following stroke, yet proof of the concept has not been obtained in clinical studies.

Proteases and Matrix Proteins

Matrix metalloproteinases (MMPs) and serine proteases including tissue and urokinase plasminogen activators (tPA and uPA) are two large and important neutral protease gene families that are involved in the neuroinflammatory response to brain ischemia. These proteases have been the subjects of extensive research in other human diseases such as cancer and arthritis, and diverse inhibitors for many of these proteases have already been deployed to assess their role in stroke. The major pathophysiological role of MMPs has been disruption of the blood–brain barrier (BBB) and hemorrhagic transformation following ischemia. These proteases are induced, processed, and released from diverse cellular elements such as activated microglia within the ischemic zone and invading neutrophils and macrophages. However, MMPs and other proteases have many more functions that can bear on both tissue injury and repair. Angiogenesis, matrix remodeling that enables cells migration, axonal navigation, and cell–cell communication are important, understudied functions. While many MMP inhibitors have been synthesized, and some shown efficacy in

mitigation of stroke outcome in experimental models, none has been tried in human stroke. Issues regarding specificity, potency, and brain access (crossing the BBB) pose significant limitation for human trials.

Other Inflammatory Mediators

Ischemic conditions to the brain result in activation of numerous biochemical pathways including lipid degradation, neurotransmitter release, and oxygen radical generation. Most of these mediators are released in large quantities in ischemia, yet mostly in a short period of time. Lipid products such as prostaglandins (PGs), thromboxane A₂ (TXA₂), and leukotrienes B₄ (LTB₄) are potent inflammatory agents that induce strong chemotactic stimuli that facilitate inflammatory cell migration into the brain. Phospholipid products such as platelet-activating factor (PAF), a potent proinflammatory mediator, may also be important. However, while potent antagonists to PGs, TXA₂, and LTB₄ have been discovered, none has been tried as a potential therapeutic in stroke. Reactive oxygen radicals (RORs) have been extensively reported as a major response to brain ischemia. RORs via activation of nuclear factor- κ B (NF κ B) elicit diverse transcriptional events that lead to synthesis and release of many cytokines, interleukins, and other inflammatory mediators. Experimental studies with numerous antioxidant compounds have been most promising, although none has demonstrated efficacy as yet in human clinical trials.

Neuroinflammation in Multiple Sclerosis

Multiple Sclerosis: Epidemiology and Pathology

Multiple sclerosis (MS) is the most common neurological disease affecting young adults in the Western world, affecting 0.05–0.15% of Caucasians. MS prevalence varies depending on the genetic background. It is much more prevalent in Caucasians than in Africans and Asians. The disease usually begins in young adulthood, and it affects women more frequently than men. In 80–90% of cases, MS starts as a relapsing–remitting course, but over time, the number of relapses decreases, with most patients experiencing progressive neurological deficits that occur independently of relapses. In 10–20% of patients, MS begins as a primary progressive course without acute relapses. In the relapsing–remitting form of MS, acute lesions in the white matter of the brain, identified by magnetic resonance imaging (MRI), can resolve spontaneously. Disruption of the BBB, demyelination, and edema are indicative of an inflammatory process. In progressive stages, global brain atrophy and disability evolve. Contemporary concepts define MS as a chronic immune and inflammatory

disease of the CNS that is focused in the myelin areas of the CNS and spinal cord resulting in axonal demyelination. The originating causes of MS are unknown, but it is widely believed to be an autoimmune and inflammatory disease process.

Inflammatory Cells and Mediators in Multiple Sclerosis

The hallmark of the brain lesion in MS is a plaque of demyelination in the white matter area of the brain where B cells, T cells, macrophages, and activated microglia are notable. The inflammatory changes are accompanied by neuronal and oligodendrocyte (myelin-producing cells) loss and astrocyte (a type of glia cell) activation. These fundamental observations were fortified by observations in an experimental model of MS, experimental allergic encephalitis (EAE), in which myelin proteins were identified as the cause of EAE. Furthermore, adoptive transfer studies suggested that CD4⁺ T cells could transmit the disease to naive animals. Further research identified cytokines released from T helper 1 (Th1) such as interferon- γ (IFN γ), TNF β , and IL-2 as well as the proinflammatory cytokine TNF α as the primary disease mediators. By contrast, T cells that secrete IL-4, IL-5, IL-10, and IL-13 were viewed as immune protective mechanisms. Thus, the Th1 cells have taken center stage in MS research based on information obtained in the EAE model that also shares many similarities to human MS.

However, this simplistic hypothesis has been challenged by studies that cast doubts on the distinction between Th1 and Th2 cytokines as mediators of MS and that also suggested a role for Th2-secreted cytokines as well as another subset of T cells, Th0, which secrete cytokines of both Th1 and Th2 cells. Furthermore, B cells are also likely to play a role in the mechanism of brain lesions in MS as well as the brain innate immunity and especially the microglia, which upon activation provide a proinflammatory milieu, which is permissive for T cell activation.

Immune and Inflammatory Modulators in Treatment of MS

There are currently three different immunomodulatory therapies that have been approved for the treatment of relapsing-remitting MS, as they have been shown to improve clinical or MRI symptoms and signs. These are recombinant IFN- β -1b (Betaseron and Betaferon); recombinant IFN1a (Avonex and Rebif), and glatiramer acetate (copolymer-1 or Copaxone). The discrete mechanism of action of any of these drugs is not entirely clear, but slowing disease progression toward sustained disability or annual exacerbation rate of the disease and reduction in MRI lesion burden has been demonstrated.

Neuroinflammation in Alzheimer's Disease

Alzheimer's Disease: Epidemiology and Pathology

Alzheimer's disease (AD) is the most common cause of dementia in the elderly population and the fourth leading cause of death in persons over 65 years old. AD is an incurable and progressive disease that affects millions of people worldwide. A prevalence of 10% has been cited in developed countries, with doubling in frequency every 5 years from the age of 65 to 50% at the age of 90. AD carries a substantial economic burden due to the need for intensive patient care. AD was originally described in 1906 by Dr. Alzheimer, a pathologist who noted prominent pathological features in brain of deceased patients that suffered from severe dementia. There are three pathological hallmarks in AD brain: (1) senile plaques, now known to be composed from deposits of β -amyloid protein, (2) neurofibrillary tangles, and (3) inflammation. AD etiology is heterogeneous, with about 10% of cases associated with a strong familial background and Mendelian inheritance pattern. Certain mutations in chromosome 1, 14, and 21 were identified in addition to or association with polymorphism of the ApoE gene. Metabolic and nutritional factors are often cited as confounding if not causal factors.

Neuroinflammation in Alzheimer Disease

The inflammatory condition in AD is quite different from that in stroke or MS. Most prominently, it lacks the classical T and B cell immune component so notable in MS and lacks the systemic inflammatory cell infiltration of neutrophils and macrophages seen in stroke. It is believed that the inflammation in AD is a result of microglia and astrocyte activation by the plaque material of aggregates of β -amyloid protein. In addition, advanced glycation end products (AGE) and their receptor, RAGE, have recently been implicated in the neurotoxicity of AD. Cytokines and interleukins including TNF α , IL-6, IL-1, and macrophage colony-stimulating factor (M-CSF) are the most likely mediators generated by the activated microglia and astrocytes. In addition, RORs and complement factors are often cited as players in the inflammation cascade of AD. Although the significance and contribution of inflammation to the progression of AD are a topic of debate, recent clinical (the Rotterdam Study) and epidemiological research on the use of anti-inflammatory medications in AD suggests a significant correlation between inflammation and neurodegeneration. An important set of inflammatory mediators of a potential source of therapeutics are the cyclooxygenase (COX) products of arachidonic

acid metabolism, especially those generated by the COX-2 enzyme complex.

Anti-inflammatory Therapy in Alzheimer's Disease

Current treatment of AD patients is based on drugs that are supposed to increase brain levels of acetylcholine (a neurotransmitter that is associated with cognitive function). This treatment has modest and partial efficacy. Nonsteroidal anti-inflammatory drugs (NSAIDs) are presumed to have beneficial effects in AD due to reduction in microglia COX-1 or COX-2 action. However, the fact that patients who are treated with NSAIDs for reasons other than AD, such as rheumatoid arthritis, are not particularly protected from AD cast doubt on the potential efficacy of these agents. Reports on the possible efficacy of aspirin in reducing AD-associated dementia may support the COX-1/2 hypothesis since aspirin inhibits COX-1 action. Reducing platelet reactivity might be a key part of aspirin benefits. The potential of other inflammatory mediators to serve as therapeutic targets for AD remains a subject of intensive research.

Summary

Immune and inflammatory cells and mediators play an important role in health and disease. The brain is unique in its ability to secure an immune privileged environment, yet brain cells can launch inflammation when activated by an aberrant biochemical milieu (amyloid accumulation) or a physiological condition (ischemia). Brain injuries also result in the exposure to systemic immune and inflammatory processes that could contribute to damage but also provide critical repair elements. Only multiple sclerosis has gained new therapeutic opportunities in the form of immune modulators, albeit at limited efficacy. It is hoped that better understanding of the role immune and inflammatory cells play in brain diseases and injuries could lead to novel and effective treatment.

See Also the Following Articles

Alzheimer's Disease; Inflammation; Cardiovascular Disease, Stress and; Multiple Sclerosis.

Further Reading

Baecher-Allen, C., Viglietta, V. and Hafler, D. A. (2005). The potential for targeting CD4+CD25+ regulatory T-cells in the treatment of multiple sclerosis in humans.

- In: Taams, L. S., Akbar, A. N. & Wauben, M. H. M. (eds.) *Regulatory T-cells in inflammation*, 133–152, Basel: Birkhauser.
- Bowirrat, A., Friedland, R. P. and Korczyn, A. D. (2002). Alzheimer's disease. In: Abramsky, O., Compston, A. S., Miller, A. & Said, G. (eds.) *Brain disease: therapeutic strategies and repair*. London, UK: Martin Dunitz.
- de Kleer, I. M., Wedderburn, L. R., Taams, L. S., et al. (2004). CD4+CD25+ (bright) regulatory T-cells actively regulate inflammation in the joints of patients with remitting form of juvenile idiopathic arthritis. *Journal of Immunology* 172, 6435–6443.
- Feuerstein, G. (2001). In: Feuerstein, G. (ed.) *Inflammation and stroke*. Basel: Birkhauser Verlag.
- Halliday, G., Robinson, S., Shepherd, C. and Kril, J. (2000). Alzheimer's disease and inflammation: a review of cellular and therapeutic mechanisms. *Clinical Experimental Pharmacology and Physiology* 27, 1–8.
- Hemmer, B., Archelos, J. and Hartung, H.-P. (2004). New concepts in the immunopathogenesis of multiple sclerosis. *Nature Reviews* 3, 291–301.
- Howard, G. and Howard, V. J. (2004). Distribution of stroke: heterogeneity of stroke by age, race and sex. In: Mohr, J. P., Choi, D. W., Grotta, J. C., Weir, B. & Wolf, P. A. (eds.) *Stroke, pathophysiology, diagnosis and management* (4th edn., pp. 3–12). Philadelphia, PA: Churchill Livingstone.
- Juurlink, B. H. J. (2001). Anti-oxidant strategy to treat stroke. In: Feuerstein, G. (ed.) *Inflammation and stroke*. Basel: Birkhauser Verlag.
- Lassman, H. and Ransohoff, R. M. (2004). The CD4-Th1 model for multiple sclerosis: a crucial re-appraisal. *Trends in Immunology* 25, 132–137.
- Lindsberg, P. (2001). Clinical evidence of inflammation as a risk factor in ischemic stroke. In: Feuerstein, G. (ed.) *Inflammation and stroke*. Basel: Birkhauser Verlag.
- Logan-Clubb, L. and Stacy, M. (1995). An open-labeled assessment of adverse effects associated with interferon 1-beta in the treatment of multiple sclerosis. *Journal of Neuroscience Nursing* 27, 344–347.
- Minagar, A., Gonzalez Toledo, E., Alexander, J. T. and Kelley, R. E. (2004). Pathogenesis of brain and spinal cord atrophy in multiple sclerosis. *Journal of Neuroimaging* 14, 5S–10S.
- Moalem, G., Leibowitz-Amit, R., Yoles, E., et al. (1999). Autoimmune T cells protect neurons from secondary degeneration after central nervous system axotomy. *Nature Medicine* 5, 49–55.
- Rosenberg, G., Navratil, M., Barone, F. C. and Feuerstein, G. (1996). Proteolytic cascade enzymes increase in focal cerebral ischemia in rat. *Journal of Cerebral Blood Flow and Metabolism* 16, 360–366.
- Stuchbury, G. and Munch, G. (2005). Alzheimer's associated inflammation, potential drug targets and future therapies. *Journal of Neural Transmission* 112, 429–453.

Neuron Death *See: Glucocorticoids – Adverse Effects on the Nervous System.*

Neuropeptide Y

C Carvajal, Y Dumont and R Quirion

McGill University, Montréal, Canada

© 2007 Elsevier Inc. All rights reserved.

The Effect of Stress on Central Neuropeptide Y Signaling

The Effect of Exogenous Neuropeptide Y on Anxiety-Related Behaviors

Anxiety-like Behavior in Neuropeptide Y Receptor Knockout and Transgenic Animals

Clinical Evidence for Neuropeptide Y in Anxiety Disorders

How Does Neuropeptide Y Regulate Emotionality?

Neuropeptide Y and Vulnerability to Disease

Glossary

<i>Anxiety</i>	A state of apprehension, uncertainty, or fear resulting from the anticipation of a realistic or fantasized threatening event or situation.
<i>Depression</i>	A mood disorder characterized by symptomatic criteria such as chronically depressed mood; low self-esteem; feelings of hopelessness, worthlessness, and guilt; and recurrent thoughts of death and suicide.
<i>Neuropeptide Y (NPY)</i>	A 36-amino-acid neuropeptide that is one of the most abundant within the central nervous system of all mammals, including humans.
<i>Neuropeptide Y receptors</i>	Receptor molecules for neuropeptide Y and the related peptides polypeptide YY (PYY) and pancreatic polypeptides (PP); they include the Y ₁ , Y ₂ , Y ₄ , Y ₅ , and Y ₆ subtypes, all of which couple to G-proteins and inhibit the production of cAMP.
<i>Sedation</i>	A reduction of anxiety, stress, irritability, or excitement, often by using a drug.

Neuropeptide Y (NPY) was isolated from porcine brain more than 2 decades ago and is a member of a large family of peptides that also include polypeptide YY (PYY) and the pancreatic polypeptides (PP). NPY is a 36-amino-acid peptide and is one of the most

abundant peptide expressed in the central nervous system (CNS) of all mammals, including humans. It is also one of the most conserved peptides in evolution, suggesting that it has an important role in the regulation of basic physiological functions. At present, five NPY receptor subtypes have been cloned and designated: Y₁, Y₂, Y₄, Y₅ and Y₆. All of them couple to G-proteins and inhibit the production of cAMP. In the rat brain, the Y₁ receptor has predominant expression in the cerebral cortex, thalamus, and brain-stem nuclei. The Y₂ receptor is abundantly expressed in the hippocampal formation and brain-stem nuclei. This subtype is considered to mostly act as presynaptic autoreceptors that provide negative feedback to NPY-ergic nerve terminals to modulate NPY release. There are only low levels of expression of the Y₄ receptor in the brain; *in situ* hybridization signals of the Y₅ receptor mRNA and its translated protein are seen in the olfactory bulb, hippocampus, and hypothalamic regions. The Y₆ receptor is not expressed in the rat and is nonfunctional in humans and primates. NPY has important modulatory functions in the immune and cardiovascular systems, circadian rhythms, food intake, and seizure activity. NPY is consistently involved in stress and anxiety-related behaviors and there is increasing support for the role of NPY in mood disorders such as depression and in various forms of addiction.

The Effect of Stress on Central Neuropeptide Y Signaling

NPY is considered an important neuromodulator in the regulation of anxiety-related behaviors. Several studies have shown that exposure to acute and chronic stress paradigms such as physical immobilization produces widespread changes in NPY expression throughout the CNS. For example, 1 h of restraint stress in Sprague–Dawley rats decreased NPY mRNA levels and NPY-like immunoreactivity (NPY-IR) by 30% in the amygdala; 2 h following restraint stress, NPY mRNA decreased by 35% in the neocortex and amygdala. However, 10 h after restraint stress, NPY-IR increased 23% in the hypothalamus. Another study

Table 1 Summary of the effects of exogenous NPY on anxiety-related behaviors^a

<i>Model</i>	<i>Treatment</i>	<i>Effect</i>	<i>Source^b</i>
Open field	ICV NPY	Dose-dependent anxiolytic effect	14
	NPY Y ₁ agonist: ICV [D-His ²⁶]NPY	Dose-dependent anxiolytic effect	14
	NPY Y ₂ agonist: ICV C2-NPY	No effect	14
	NPY Y ₅ agonist: ICV [cPP(1–7), NPY(19–23), Ala ³¹ , Aib ³² , Gln ³⁴]hPP	Anxiolytic-like effect	14
Elevated plus-maze	ICV NPY	↑ Time on open arms	7
		Dose-dependent anxiolytic effect	14
	ICV PYY	↑ Time on open arms	3
	ICV NPY(2–36)	↑ time on open arms	3
	NPY Y ₁ agonist: ICV [D-His ²⁶]NPY	Dose dependent anxiolytic effect	14
	NPY Y ₁ antagonist: ICV BIBP3226	Anxiogenic-like effect	9
	NPY Y ₁ antisense: ICV; CeA NPY Y ₁ antisense	Attenuated NPY-induced anxiolytic effect	4, 15
	NPY Y ₂ agonist: ICV C2-NPY	No effect	14
	NPY Y ₅ agonist: ICV [cPP(1–7), NPY(19–23), Ala ³¹ , Aib ³² , Gln ³⁴]hPP	Anxiolytic-like effect	14
	NPY Y ₁ /Y ₄ /Y ₅ agonist: ICV [Leu ³¹ , Pro ³⁴]NPY	↑ Time on open arms	3
	NPY Y ₂ /Y ₅ agonist: ICV	↓ Time on open arms	7
	NPY(3–36)	No effect	9
	ICV NPY	↑ Transitions (anxiolytic-like effect)	11
	BLA NPY	↑ Social behavior	13
	Septum NPY	↑ Social behavior	8
Light–dark box Social interaction	NPY Y ₁ antagonist: BLA BIBO3304	Blocks NPY-induced ↑ in social behavior	13
	NPY Y ₂ agonist: BLA C2-NPY	↓ Social behavior	12
	NPY Y ₅ agonist: BLA [cPP(1–7), NPY(19–23), Ala ³¹ , Aib ³² , Gln ³⁴]hPP	Anxiolytic-like effect	12
	NPY Y ₂ /Y ₅ agonist: BLA NPY(3–36)	Anxiolytic-like effect	12
	NPY Y ₅ antagonist: BLA Novartis 1	Blocked the anxiolytic-like effect of NPY(3–36)	12
	IP CGP71683A	No effect	10

Vogel conflict	ICV NPY	Anticonflict/anxiolytic-like effect	7
Geller–Seifter	ICV NPY	Anticonflict/anxiolytic-like effects	1, 2, 6
	NPY Y ₁ /Y ₄ /Y ₅ agonist: ICV [Leu ³¹ , Pro ³⁴]NPY	Anticonflict/anxiolytic-like effects	5
Fear-potentiated startle	ICV NPY, PYY, NPY(2–36)	↓ Startle response	3
	NPY Y ₁ /Y ₄ /Y ₅ agonist: ICV [Leu ³¹ , Pro ³⁴]NPY	↓ Startle response	3
	NPY Y ₂ antagonist: ICV NPY(13–36)	No effect	3

^a↑, increase, ↓, decrease, BLA, injected into basolateral nucleus of the amygdala; ICV, injected intracerebroventricularly; IP, injected intraperitoneally. Adapted from Carvajal, C., Dumont, Y. and Quirion, R. (2006), Neuropeptide Y: role in emotion and alcohol dependence, *CNS & Neurological Disorders – Drug Targets* **5**, 181–195, with permission.

^b 1. Britton, K. T., Akwa, Y., Spina, M. G., et al. (2000). Neuropeptide Y blocks anxiogenic-like behavioral action of corticotropin-releasing factor in an operant conflict test and elevated plus maze. *Peptides* **21**, 37–44.

2. Britton, K. T., Southerland, S., Van Uden, E., et al. (1997). Anxiolytic activity of NPY receptor agonists in the conflict test. *Psychopharmacology* (Berlin) **132**, 6–13.

3. Broqua, P., Wettstein, J. G., Rocher, M. N., et al. (1995). Behavioral effects of neuropeptide Y receptor agonists in the elevated plus-maze and fear-potentiated startle procedures. *Behavioral Pharmacology* **6**, 215–222.

4. Heilig, M. (1995). Antisense inhibition of neuropeptide Y (NPY)-Y1 receptor expression blocks the anxiolytic-like action of NPY in amygdala and paradoxically increases feeding. *Regulatory Peptides* **59**, 201–205.

5. Heilig, M., McLeod, S., Brot, M., et al. (1993). Anxiolytic-like action of neuropeptide Y: mediation by Y1 receptors in amygdala, and dissociation from food intake effects. *Neuropsychopharmacology* **8**, 357–363.

6. Heilig, M., McLeod, S., Koob, G. K., et al. (1992). Anxiolytic-like effect of neuropeptide Y (NPY), but not other peptides in an operant conflict test. *Regulatory Peptides* **41**, 61–69.

7. Heilig, M., Soderpalm, B., Engel, J. A., et al. (1989). Centrally administered neuropeptide Y (NPY) produces anxiolytic-like effects in animal anxiety models. *Psychopharmacology* (Berlin) **98**, 524–529.

8. Kask, A., Nguyen, H. P., Pabst, R., et al. (2001). Neuropeptide Y Y1 receptor-mediated anxiolysis in the dorsocaudal lateral septum: functional antagonism of corticotropin-releasing hormone-induced anxiety. *Neuroscience* **104**, 799–806.

9. Kask, A., Rago, L. and Harro, J. (1996). Anxiogenic-like effect of the neuropeptide Y Y1 receptor antagonist BIBP3226: antagonism with diazepam. *European Journal of Pharmacology* **317**, R3–R4.

10. Kask, A., Vasar, E., Heidmets, L. T., et al. (2001). Neuropeptide Y Y(5) receptor antagonist CGP71683A: the effects on food intake and anxiety-related behavior in the rat. *European Journal of Pharmacology* **414**, 215–224.

11. Pich, E. M., Agnati, L. F., Zini, I., et al. (1993). Neuropeptide Y produces anxiolytic effects in spontaneously hypertensive rats. *Peptides* **14**, 909–912.

12. Sajdyk, T. J., Schober, D. A. and Gehlert, D. R. (2002). Neuropeptide Y receptor subtypes in the basolateral nucleus of the amygdala modulate anxiogenic responses in rats. *Neuropharmacology* **43**, 1165–1172.

13. Sajdyk, T. J., Vandergriff, M. G. and Gehlert, D. R. (1999). Amygdalar neuropeptide Y Y1 receptors mediate the anxiolytic-like actions of neuropeptide Y in the social interaction test. *European Journal of Pharmacology* **368**, 143–147.

14. Sorensen, G., Lindberg, C., Wortwein, G., et al. (2004). Differential roles for neuropeptide Y Y1 and Y5 receptors in anxiety and sedation. *Journal of Neuroscience Research* **77**, 723–729.

15. Wahlestedt, C., Pich, E. M., Koob, G. F., et al. (1993). Modulation of anxiety and neuropeptide Y-Y1 receptors by antisense oligodeoxynucleotides. *Science* **259**, 528–531.

looked at repeated stress in Wistar–Kyoto rats. After 1 h of restraint stress, NPY mRNA increased in the arcuate nucleus of the hypothalamus by 81%, but after 3 days (1 h day⁻¹), NPY mRNA increased by 40% in the arcuate nucleus and 50% in the medial amygdala (MeA). After chronic restraint (1 h day⁻¹ for 9–10 days), there was an upregulation of prepro-NPY mRNA and NPY itself in the amygdala but not in hypothalamic or cortical regions. Consequently, acute physical restraint, which promotes experimental anxiety, primarily suppressed NPY expression, but chronic physical restraint generally enhanced NPY signaling, especially in the amygdala. These results are believed to support an early hypothesis that suggests NPY may act to buffer the behavioral effects of stress-promoting signals.

The Effect of Exogenous Neuropeptide Y on Anxiety-Related Behaviors

NPY is consistently implicated in the pathogenesis of anxiety disorders, based on a significant number of findings that show NPY-induced anxiolytic activity in animal models widely used for the screening of anxiolytic compounds. NPY is reported to elicit anxiolytic-like effects in models of anxiety, including exploratory behavior-based tests such as the open field, elevated plus-maze, and light–dark compartment test, social interaction, punished responding tests, and fear-potentiated startle (see [Table 1](#)).

The open field test is often used to measure behavioral changes induced by anxiolytic or anxiogenic-like compounds in animal models. Anxiolytic drugs such as benzodiazepines and barbiturates increase the number of entries and time spent in the central area of the arena. Total crossings are presented as a measure of general locomotor activity in the arena. At higher doses, traditional anxiolytic drugs have significant sedative effects and suppress locomotor activity. Similarly, NPY causes dose-dependent suppression in open field activity when the intracerebroventricular (ICV) dose of NPY exceeds 5 µg. The significant effect of NPY on sedation in the open field led investigators to examine the potential anxiolytic properties of this peptide in the elevated plus-maze.

The elevated plus-maze is one of the most widely used tests for anxiety in animal models. The test is based on the conflict between the natural aversion of rodents for open spaces and the drive to explore a novel environment. Typically, the time spent on the open arms and the numbers of entries onto the open and closed arms of the maze are recorded. The percentage of open arm entries relative to the number of total arm entries is considered to be the superlative measure reflecting innate fearfulness. NPY,

administered ICV, decreased the preference for closed arm entries and increased the time spent on open arms. Higher doses of NPY (exceeding 2 nmol) suppressed entries into both closed and open arms, consistent with the sedative action of NPY observed at high doses in the open field.

A central mechanism involving the NPY Y₁ receptor subtype in the anxiolytic effect of NPY is supported by several pharmacological studies. For example, ICV administration of an antisense oligonucleotide targeted at Y₁ receptor mRNA attenuated NPY-induced anxiolytic-like effects in the elevated plus-maze and blocked the anxiolytic-like effect of NPY in the central nucleus of the amygdala (CeA) in the same test. This was confirmed in a study that demonstrated anxiolytic-like activity of the NPY, PYY, and the Y₁/Y₄/Y₅ agonist [Leu³¹, Pro³⁴]NPY but not the Y₂ agonist NPY(13–36). Further support for the Y₁ receptor was revealed with the anxiogenic-like effect of the Y₁ receptor antagonist BIBP3226. The Y₂-type receptor agonist, NPY(13–36), was later shown to induce anxiogenic-like effects in this model, in mice, when administered ICV.

More recently, there was an attempt to confirm the receptor subtype(s) involved in the anxiolytic/anxiogenic action of NPY as well as to determine which receptor(s) are involved in the sedative action of NPY. This study examined ICV injection of NPY as well as the specific receptor agonists for the Y₁ receptor ([D-His²⁶]NPY), Y₂ receptor (C2-NPY), and Y₅ receptor ([cPP(1–7), NPY(19–23), Ala³¹, Aib³², Gln³⁴]hPP) in the elevated plus-maze and open field tests. The results revealed that NPY and the Y₁ agonists had a dose-dependent anxiolytic-like effect in both behavioral tests. However, in contrast to NPY, which caused significant sedation in the open field test, the Y₁ agonist was without sedative effect. The Y₂ agonist showed neither anxiolytic-like nor sedative effects, and the Y₅ agonist showed anxiolytic-like activity in both behavioral tests and caused sedation in the same dose range as NPY in the open field test. The authors conclude that the anxiolytic-like effects of ICV-administered NPY in rats are mediated via both Y₁ and Y₅ receptors, whereas sedation is mediated via Y₅ receptors.

Additional support for the anxiolytic-like effect of NPY has been confirmed in the light–dark compartment test in which ICV NPY increased the number of transitions between the two compartments, a validated measure of anxiolytic activity in this test. Social interaction has also been pharmacologically validated as an experimental model of anxiety. The time spent in active social behavior, as well as locomotor activity, is recorded. NPY, when microinjected into the basolateral nucleus of the amygdala (BLA) and into

the caudal dorsolateral septum increases social interaction. Thus, NPY-induced anxiolysis in this paradigm of anxiety appears to be mediated by several brain regions. As in the elevated plus-maze, the involvement of the Y_1 receptor in social interaction is supported by blocking the anxiolytic-like effect of NPY with the intra-amygdalar injection of the selective Y_1 receptor antagonist BIBO3304. The role of the Y_5 receptor was also shown using [cPP(1–7),NPY(19–23),Ala³¹,Aib³²,Gln³⁴]hPP, a selective Y_5 receptor agonist. This agonist and the mixed Y_5/Y_2 agonist NPY(3–36) caused anxiolytic-like effects when injected into the BLA. The effect of NPY(3–36) was blocked by pretreatment with a purported Y_5 antagonist, Novartis 1 (synthesized by Eli Lilly). However, the Y_5 antagonist CGP71683A did not influence anxiety in any exploratory model of anxiety.

It has also been suggested that the Y_2 NPY receptor may mediate anxiogenic-like behaviors in the amygdala in this task. One study found a dose-dependent decrease in the time spent in social interaction when the Y_2 receptor agonist C2-NPY was directly injected into the BLA. However, the authors did not attempt to reverse the anxiogenic-like effects of this molecule with a selective Y_2 receptor antagonist, such as BIIE0246. The Y_2 receptor agonists may have an anxiogenic effect because the activation of the pre-synaptic Y_2 receptor may lead to decreased release of NPY and a subsequent decrease in Y_1 receptor activation. Overall, in the exploratory tests of anxiety including the open field, elevated plus-maze, and social interaction, NPY produces robust anxiolytic-like activity that is mediated through the Y_1 (and possibly Y_5) receptor subtypes, whereas the anxiogenic effects of NPY are mediated through Y_2 receptors.

Several versions of operant behavioral analysis for anxiety have been developed such as the Vogel conflict test and Geller–Seifter test, which are based on conflict of motivations rather than unconditioned exploratory models. In these tests, the subject experiences opposing and concomitant tendencies of desire (e.g., to obtain a reward) and of fear (avoidance of a potentially aversive stimulus). In the Vogel conflict test, water-deprived rodents are exposed to mild and intermittent electric shock via the spout of a water bottle when attempting to drink. In this test, treatment with anxiolytic drugs increases the number of accepted shocks during the punished phase. It has been shown that ICV NPY markedly increased the number of electric shocks accepted in this test. At the doses employed, NPY was reported not to affect pain sensitivity in a shock threshold test or to affect thirst. Thus, the anticonflict effects are considered to be related to a reduction of anxiety. In the Geller–Seifter conflict test, rats are trained to respond for a food

reward and are exposed to a modest electric shock during the conflict component of the procedure. The ICV injection of NPY consistently produces dose-dependent anticonflict/anxiolytic-like effects in the Geller–Seifter test, an established animal model of anxiety especially suitable for detecting the effects of benzodiazepine-like anxiolytics. The Y_1 receptor subtype is implicated in this test as well. Intra-amygdalar injection of [Leu³¹, Pro³⁴]NPY was shown to have anxiolytic-like effects. As in the exploratory models of anxiety, NPY can also induce anxiolysis in the conflict-based tests.

Startle is an adaptive response to acoustic stimuli that enables the individual to avoid, or reduce, the risk of an injury by a predator. In fear-potentiated startle, an acoustic stimulus is paired with an aversive intervention such as foot shock or air-puff, and after training, the stimulus alone is capable of elevating startle amplitude. As in the elevated plus-maze, ICV injections of NPY, PYY, and the $Y_1/Y_4/Y_5$ agonist [Leu³¹Pro³⁴]NPY inhibited fear-potentiated startle, whereas the Y_2 agonist NPY(13–36) had no effect. The Y_1 receptor is consistently indicated across diverse animal models of anxiety to mediate the anxiolytic-like activity of NPY. However, the Y_5 receptor cannot be ignored because [Leu³¹Pro³⁴]NPY does have some efficacy for the Y_5 receptor subtype. In addition, the Y_5 – Y_2 differentiating receptor agonist NPY(2–36) has shown anxiolytic-like activity in the elevated plus-maze and fear-potentiated startle.

Anxiety-like Behavior in Neuropeptide Y Receptor Knockout and Transgenic Animals

The development of the NPY transgenic rat has provided a unique opportunity to study the effects of this peptide on anxiety-related behaviors. In this transgenic rat model, there is central overexpression of prepro-NPY mRNA and NPY peptide in the CA1 region of the hippocampus and decreased Y_1 binding sites within the hippocampus (CA1, CA2, and dentate gyrus). Recently, NPY protein levels were also shown to be significantly higher in the paraventricular, suprachiasmatic, and supraoptic nuclei of the hypothalamus and tended to be increased in the arcuate nucleus in these rats. The transgenic animals were generated using a 14.5-kb fragment of the rat NPY genomic sequence that includes the normal intronic sequence elements and is flanked by a 5-kb 5' sequence thought to contain the major regulatory elements that normally control NPY expression. Consequently, the regulation of the NPY transgene is predicted to be similar to the regulation of endogenous

NPY. These molecular and neurochemical events led to an altered anxiety profile in NPY transgenic rats that included an insensitivity to restraint stress in the elevated plus-maze and an increase in the number of punished drinking events in the Vogel conflict test. The anxiolytic-like profile of the NPY transgenic rat was replicated in a study that showed NPY transgenic animals were resistant to acute physical restraint stress measured by the elevated-plus maze and displayed anxiolytic-like activity in the open field. The behavioral profile observed in the NPY transgenic rats was not associated with any significant changes in corticosterone levels before or following a stress challenge. This suggests a mechanism other than hypothalamic-pituitary-adrenal axis modulation and supports the role of the limbic system in the modulation of anxiety-related behaviors by NPY. This may be critically important and could distinguish NPY from other substances modulating stress-responses (although there are links between NPY and corticotropin releasing hormone, CRH).

The central role of the limbic system is also supported by recent studies involving Y_2 receptor subtype knockout ($Y_2^{-/-}$) mice. Mice deficient in the Y_2 receptor subtype displayed an anxiolytic-like phenotype in the elevated plus-maze and open field test. More recently, NPY $Y_2^{-/-}$ mice displayed increase preference for the central area of the open field when compared to wild-type control ($Y_2^{+/+}$) animals without changes in locomotor activity. These findings have been confirmed by an independent laboratory using the open field, elevated plus-maze, and light-dark compartment tests. These receptors are considered to be autoreceptors that provide negative feedback to NPY-ergic nerve terminals to modulate NPY release. Consequently, $Y_2^{-/-}$ mice are predicted to have increased endogenous peptide expression, analogous to the phenotype of NPY transgenic rats, which may contribute to the underlying mechanism responsible for anxiolytic-like behaviors regulated by NPY.

There is also evidence that NPY exerts anxiolytic-like effects through Y_1 receptors within the amygdala. The predicted anxiogenic-like phenotype of NPY Y_1 knockout mice is supported by a study that found significantly higher levels of territorial aggression in Y_1 knockout mice compared to control animals using established aggression paradigms.

Interestingly, NPY knockout ($NPY^{-/-}$) mice did not show a dramatic change in anxiety-like behavior in the elevated plus-maze. $NPY^{-/-}$ mice were found to have a significant increase in the amplitude response in the acoustic startle test, and they were less active in the central part of the open field, suggesting that these animals were more anxious. Overall, the results from NPY transgenic and knockout animals

provide compelling evidence for the role of NPY in the modulation of anxiety.

Clinical Evidence for Neuropeptide Y in Anxiety Disorders

In addition to the extensive preclinical data supporting NPY in the regulation of anxiety-like behaviors, a number of clinical studies have been published on the subject. In human subjects, there is a positive association between acute, uncontrollable psychological stress and robust increases in plasma levels of NPY. The increase in plasma NPY was positively correlated with increased cortisol and norepinephrine concentrations. A correlation between higher levels of stress-related NPY release and lower levels of subjective psychological distress was also found, supporting the possibility that NPY exhibits anxiolytic activity during stress in human subjects.

How Does Neuropeptide Y Regulate Emotionality?

The mechanistic action of NPY in anxiety-related behaviors is strongly associated with the amygdala. The amygdala is a key structure to the regulation of anxiety and expression of emotional responses to stress. Attention has been focused particularly on the basolateral complex (lateral, basolateral, and basomedial nuclei) and CeA. The amygdala contains NPY and significant levels of Y_1 and Y_2 receptor subtypes. These areas may be potential neural substrates to the behavioral effects of NPY on emotional regulation. NPY is also produced by neurons in the hippocampus with considerable expression of the Y_1 and Y_2 receptor subtype. In addition to the substantial evidence for the involvement of the amygdala and hippocampus, there is also evidence for the periaqueductal gray, septum, and locus coeruleus. However, the specific contribution of these anatomical regions requires further elucidation.

The interaction between CRH and NPY has been proposed as a means by which emotional behavior is regulated. In contrast to the anxiolytic-like effects of NPY, there is evidence that CRH has significant anxiogenic-like effects in animal models. For example, the ICV administration of CRH increases anxiety-like behavior in the elevated plus-maze, and CRH antagonists can block the behavioral effects of a stress challenge. The anxiogenic-like role of CRH has been confirmed in studies using CRH-overexpressing and CRH knockout mice. It has been shown that NPY can counteract the anxiogenic-like effect of CRH in the hippocampus, hypothalamus, locus coeruleus, periaqueductal gray, and septal nucleus. Moreover,

deletion of Y_2 receptors causes a 60% reduction in CRH mRNA expression that might contribute to the anxiolytic-like behavior of NPY Y_2 receptor knock-out mice. Thyroxine-treated adult animals displayed reduced anxiety in the motility box and elevated plus-maze, a reduction in the number of CRH-IR neurons in the CeA, and an increase in the number of NPY-IR neurons in nuclei of the basolateral complex of the amygdala.

It has also been proposed that NPY increases γ -aminobutyric acid (GABA) signaling to produce anxiolysis, analogous to many classes of anxiolytic drugs such as benzodiazepines. NPY is localized in GABAergic neurons in the amygdala, and there is evidence that NPY may directly modulate the activity of GABAergic neurons by stimulating Y_1 receptors. It has been shown that diazepam counteracts the anxiogenic effect of the Y_1 receptor antagonist BIBP3226, suggesting that the equilibrium between GABAergic and NPY-ergic neurotransmission may be important for the regulation of the emotional state in animal models. This conclusion is supported by a recent study that found chronic treatment with positive (anxiolytic) or negative (anxiogenic) regulators of GABA_A receptors modulate Y_1 receptor-mediated transmission in the amygdala. The authors suggest that the NPY Y_1 -mediated transmission and the GABAergic system may act together on the same postsynaptic target sites to decrease anxiety and that the Y_1 receptors might also play a role in controlling both the NPY and GABA presynaptic release.

However, it was recently shown that the anxiolytic-like effect produced by inhibiting the metabotropic glutamate receptor 5 (mGlu5) with the mGlu5 receptor antagonist MPEP is mediated by the NPY receptor and not by GABA. In order to determine the mechanism that contributes to the anxiolytic action of MPEP, the authors injected either the benzodiazepine antagonist flumazenil or the Y_1 receptor antagonist BIBO3304 into the BLA. Flumazenil antagonizes the effect of the several different classes of anxiolytic agents including the benzodiazepine diazepam, serotonergic agents, and noradrenergic ligands in animal models of anxiety through a GABA-mediated mechanism. However, the anxiolytic effects of MPEP were not changed by flumazenil; but they were abolished by BIBO3304 administration. The authors also found a decrease in NPY-IR neurons after three doses of MPEP administration, which suggests a decrease in NPY levels in the amygdala. The authors propose that the decrease in NPY levels may result from enhanced release of the peptide and that this release is responsible for the anxiolytic action of MPEP. Interestingly, the administration of metabotropic glutamate receptor antagonists into the hippocampus also produces

anxiolytic-like effects. NPY has been shown to inhibit glutamate release in the hippocampus, leading to the speculation that the anxiolytic-like effect of hippocampal NPY overexpression in NPY transgenic rats resembles the anxiolytic-like action of the metabotropic glutamate receptor antagonists. The interaction between glutamate acting via mGlu5 and NPY may represent a novel mechanism of anxiolytic action in the brain, independent of the traditional GABA-mediated benzodiazepine signaling.

Overall, these studies demonstrate that NPY and its receptor subtypes are significantly involved in the regulation of anxiety-like behaviors in both animal models and human subjects.

Neuropeptide Y and Vulnerability to Disease

Vulnerability to Depression

Depression is traditionally viewed as the manifestation of an inability to cope with various lifetime stressors that is presumably determined by genetics. Given the significance of NPY in the CNS as a modulator of anxiety, NPY may have a significant role in the etiology of depression and/or the effective reversal of symptoms. This theory is supported by preclinical data that consistently indicate a role for NPY in depression. For example, maternal deprivation is an animal model of depression/vulnerability to stress that posits that early life stress may cause changes in the CNS (e.g., hypothalamic-pituitary-adrenal dysregulation) that are associated with an increased risk of adult life depressive psychopathology. Using this model in rats for 3 h day⁻¹ during postnatal days 2–14, NPY levels were shown to be reduced in the hippocampus and striatum and increased in the hypothalamus. However, if lithium treatment was employed on days 50–83, the changes in NPY-IR induced by maternal deprivation were not observed in the hippocampus and striatum, whereas NPY levels were further increased in the hypothalamus. Consequently, early life stress has long-term effects on NPY levels in the CNS and may be a factor in the development of depression; possibly through an increased vulnerability to stress.

Additional evidence for NPY in depression is through the antidepressant-like properties of NPY. The Porsolt forced-swim test is widely used for the screening of potential antidepressant drugs. It has recently been shown that NPY displayed antidepressant-like activity in the rat forced-swim test. These results have since been confirmed in the mouse version of this test. ICV NPY administration significantly reduced immobility time in a dose-dependent manner,

as did [Leu³¹Pro³⁴]PYY. Attempts to determine the specific receptor subtype(s) involved has shown that BIBP3226 and BIBO3304 (selective Y₁ antagonists) and NPY(13–36) (a preferential Y₂ agonist) did not display any activity at the doses tested. However, pretreatment with BIBP3226 and BIBO3304 significantly blocked the anti-immobility effects of NPY. It is predicted that potentiation of NPY signaling at central postsynaptic receptors through Y₁ receptors is essential to reverse the behavioral characteristics of depression. The upregulation of NPY expression through Y₁ receptor activation may be advantageous for behavioral adaptation to stress-induced changes within the CNS. Consequently, a maladaptive response to stress may be the result of an intrinsic abnormal function of central NPY transmission that is characteristic of depressive states. Thus, NPY may be involved in the vulnerability to mood disorders through its endogenous modulation of emotional behavior.

Vulnerability to Alcoholism

Recent evidence suggests that NPY has a significant role in the neurobiological response to alcohol, including alcohol consumption, dependence, and withdrawal. NPY is beginning to emerge as an important neuromodulator in the etiology of alcoholism that is independent from the addictive and reinforcing properties of the traditional system and, instead, is strongly associated with innate emotionality. It is proposed that individuals with low innate levels of NPY and high basal levels of anxiety, or with low levels of NPY and increased anxiety caused by withdrawal, may drink increased amounts of alcohol in an attempt to reduce anxiety. There is evidence that supports this hypothesis from genetic lines of alcohol-preferring rats. For example, alcohol-preferring (P) rats have altered levels of NPY in several brain regions, including low levels of NPY in the amygdala compared to non-alcohol-preferring (NP) rats. P rats also have high basal levels of anxiety, which may predispose them to excessive alcohol drinking. The central infusion of NPY may reverse high alcohol drinking in P rats by reducing their high levels of anxiety. Studies in Wistar rats also indicate that animals with high levels of anxiety drink more alcohol than those with lower levels. Furthermore, because NPY is expressed at low levels in the amygdala following ethanol withdrawal, it is predicted that preventing this decrease in NPY expression might prevent subsequent withdrawal symptoms such as increased anxiety. Consequently, the motivational aspect of alcohol intake may be strongly tied to innate anxiety.

In summary, much evidence suggests a key role for NPY and its receptors in anxious behaviors. NPY and

congeners could also play a role in more chronic stressful situations that can possibly result in the development on depressive states or addictive behaviors.

See Also the Following Articles

Alcohol, Alcoholism and Stress: A Psychobiological Perspective; Corticotropin Releasing Factor (CRF); Depression and Manic-Depressive Illness; Depression Models.

Further Reading

- Carvajal, C., Dumont, Y. and Quirion, R. (2006). Neuropeptide Y: role in emotion and alcohol dependence. *CNS & Neurological Disorders – Drug Targets* 5, 181–195.
- Dumont, Y., Chabot, J. G. and Quirion, R. (2004). Receptor autoradiography as mean to explore the possible functional relevance of neuropeptides: focus on new agonists and antagonists to study natriuretic peptides, neuropeptide Y and calcitonin gene-related peptides. *Peptides* 25, 365–391.
- Heilig, M. (2004). The NPY system in stress, anxiety and depression. *Neuropeptides* 38, 213–224.
- Kask, A., Harro, J., von Horsten, S., et al. (2002). The neurocircuitry and receptor subtypes mediating anxiolytic-like effects of neuropeptide Y. *Neuroscience and Biobehavioral Reviews* 26, 259–283.
- Lin, S., Boey, D. and Herzog, H. (2004). NPY and Y receptors: lessons from transgenic and knockout models. *Neuropeptides* 38, 189–200.
- Nestler, E. J., Barrot, M., DiLeone, R. J., et al. (2002). Neurobiology of depression. *Neuron* 34, 13–25.
- Pandey, S. C., Carr, L. G., Heilig, M., et al. (2003). Neuropeptide y and alcoholism: genetic, molecular, and pharmacological evidence. *Alcoholism: Clinical & Experimental Research* 27, 149–154.
- Redrobe, J. P., Carvajal, C., Kask, A., et al. (2004). Neuropeptide Y and its receptor subtypes in the central nervous system: emphasis on their role in animal models of psychiatric disorders. In: Michel, M. C. (ed.) *Handbook of experimental pharmacology. Vol.162: Neuropeptide Y and related peptides*, pp. 101–136. Heidelberg: Springer.
- Redrobe, J. P., Dumont, Y. and Quirion, R. (2002). Neuropeptide Y (NPY) and depression: from animal studies to the human condition. *Life Sciences* 71, 2921–2937.
- Sajdyk, T. J., Shekhar, A. and Gehlert, D. R. (2004). Interactions between NPY and CRF in the amygdala to regulate emotionality. *Neuropeptides* 38, 225–234.
- Thorsell, A. and Heilig, M. (2002). Diverse functions of neuropeptide Y revealed using genetically modified animals. *Neuropeptides* 36, 182–193.
- Thorsell, A., Michalkiewicz, M., Dumont, Y., et al. (2000). Behavioral insensitivity to restraint stress, absent fear suppression of behavior and impaired spatial learning in transgenic rats with hippocampal neuropeptide Y over-expression. *Proceedings of the National Academy of Sciences USA* 97, 12852–12857.

Valdez, G. R. and Koob, G. F. (2004). Allostasis and dysregulation of corticotropin-releasing factor and neuropeptide Y systems: implications for the development of alcoholism. *Pharmacology Biochemistry and Behavior* 79, 671–689.

Zukowska, Z. and Feuerstein, G. (eds.) (2005). *NPY family of peptides in inflammation, immune disorders, angiogenesis, and cancer*. Basel, Switzerland: Birkhauser Verlag.

Neuropeptides, Stress-Related

A J Tilbrook

Monash University, Melbourne, Australia

© 2007 Elsevier Inc. All rights reserved.

Stress-Related Functions of Corticotropin Releasing
Hormone and Arginine Vasopressin

Opioids

Oxytocin and Stress

Appetite-Regulating Neuropeptides and Stress

Glossary

*Hypophyseal
portal vessel
system*

A highly specialized blood vessel system that plays a vital role in the rapid transport of neuropeptides produced by the hypothalamic parvicellular neurons to target cells in the anterior pituitary gland. The hypophyseal portal system begins at the base of the hypothalamus with a group of capillaries (primary capillary plexus) that recombine into small portal veins, which pass down through the pituitary stalk into the anterior pituitary gland. Here they branch to form most of the anterior pituitary capillaries and sinusoids, which in turn drain into the systemic venous system. Almost all the blood supply to the anterior pituitary must first pass through the hypothalamo-hypophyseal portal system, arriving via the superior hypophyseal artery and leaving the anterior pituitary in the anterior hypophyseal vein.

*Hypothalamic-
pituitary-
adrenal (HPA)
axis*

One of the major physiological systems activated by stress; it consists of a hormonal cascade involving the hypothalamus of the brain, the anterior pituitary gland, and the cortex of the adrenal glands. The HPA axis is the humoral component of an integrated neural and endocrine system and results in the release of corticosteroid hormones from the adrenal glands. The major

*Median
eminence*

corticosteroids released during stress are the glucocorticoids, such as cortisol and corticosterone.

The median eminence is located at the base of the hypothalamus and is a region of confluence of neural and blood-borne messages where neuropeptides can be secreted from the terminals of neurons into the hypophyseal portal vessel system.

Neuropeptides

Peptides synthesized in, stored in, and released from neural tissue.

Stress

A complex physiological state that embodies a range of integrative physiological and behavioral processes that occur when there is a real or perceived threat to homeostasis.

There are numerous neuropeptides that are involved in the regulation of stress responses and released as effector molecules as a result of stress responses. Among the most widely researched stress-related neuropeptides are corticotropin releasing hormone (CRH) and arginine vasopressin (AVP), which regulate the hypothalamic-pituitary-adrenal (HPA) axis, and the opioid peptides, which have many functions. Oxytocin is a neuropeptide now recognized as a potentially important modulator of stress responses (**Figure 1**). Other neuropeptides, such as those involved in appetite regulation, can also influence the activity of the HPA axis during stress, although there is considerable variation in the scientific literature concerning the effects of these neuropeptides.

Stress-Related Functions of Corticotropin Releasing Hormone and Arginine Vasopressin

The key stress-related action of CRH and AVP is to act as hypophysiotropic hormones to stimulate the synthesis and secretion of adrenocorticotrophic hormone (ACTH) from the anterior pituitary gland (**Figure 1**).

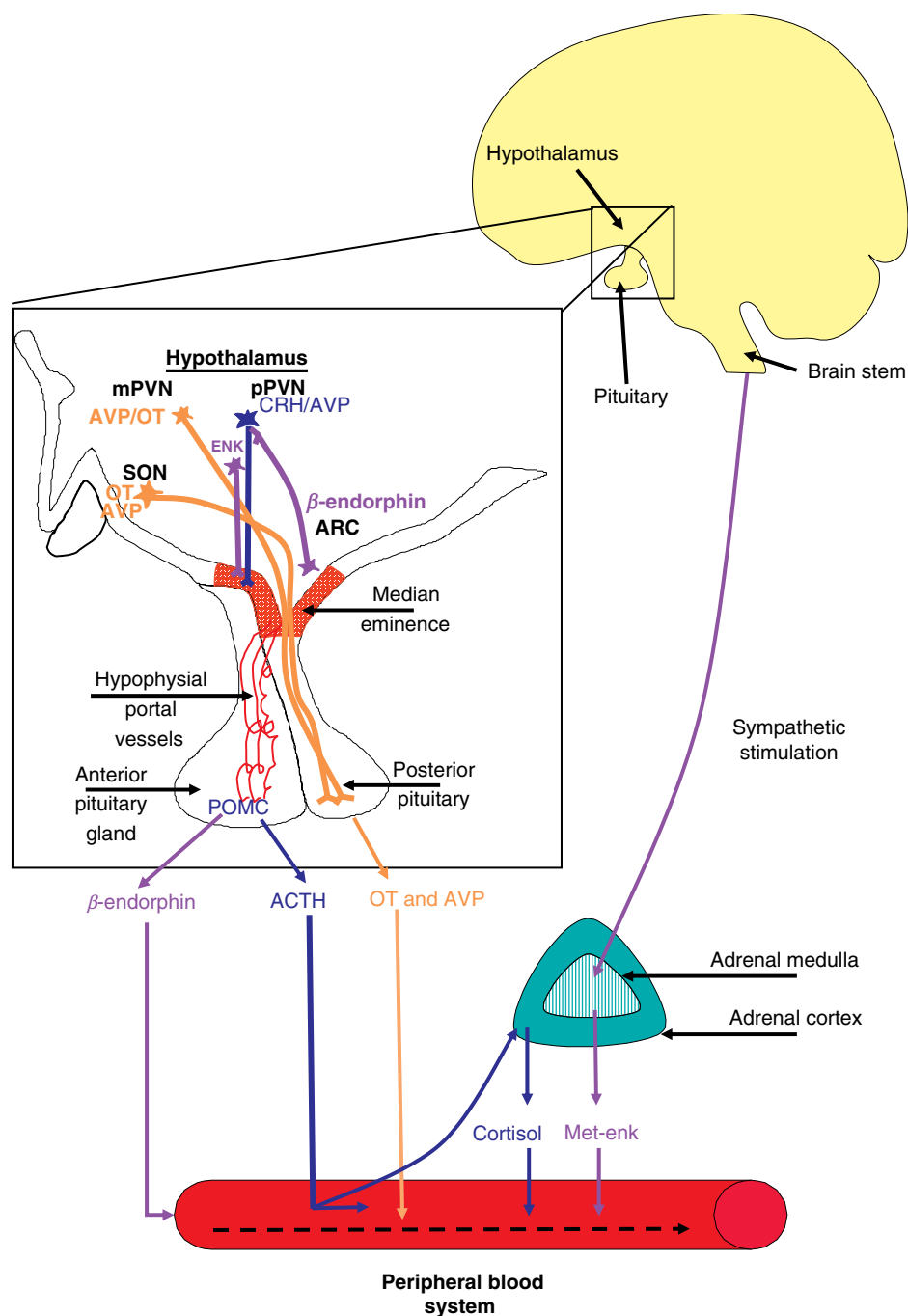


Figure 1 Diagram of some of the key neuropeptides involved in stress-related functions. Corticotropin releasing hormone (CRH) and arginine vasopressin (AVP) are synthesized and stored in parvocellular neurons in the paraventricular nucleus of the hypothalamus (pPVN) and are released into the hypophyseal portal blood to stimulate the synthesis of peptides derived from proopiomelanocortin (POMC), including adrenocorticotrophic hormone (ACTH) and the opioid peptide β -endorphin, which are released into the systemic circulation. ACTH stimulates the synthesis of glucocorticoids, such as cortisol, from the adrenal cortex. Opioids such as enkephalin (ENK) and β -endorphin are synthesized in neurons in the brain and have many actions, including the modulation of stress responses. Neurons in the arcuate nucleus of the hypothalamus (ARC) that produce β -endorphin have important interactions with neurons in the pPVN that produce CRH and AVP. The opioid met-enkephalin (Met-enk) is released from the adrenal medulla due to sympathetic stimulation during stress. Oxytocin (OT) is stored in and released from neurons in the posterior pituitary as well as in the brain and acts to attenuate stress responses and reduce anxiety. Magnocellular neurons that synthesize OT and AVP have cell bodies in the magnocellular region of the PVN (mPVN) and supraoptic nucleus (SON) of the hypothalamus. There is also evidence that some neuropeptides best known for their actions in regulating appetite also influence the activity of stress systems, especially the hypothalamic-pituitary-adrenal axis.

The target endocrine cells in the anterior pituitary gland on which CRH and AVP act are called corticotropes. ACTH is released from corticotropes into the systemic circulation and stimulates the synthesis of the glucocorticoids by the adrenal glands. CRH and AVP also have various other physiological actions.

Corticotropin Releasing Hormone

The principal source of CRH is neurons located in the parvicellular division of the paraventricular nucleus of the hypothalamus, which project to the median eminence. CRH is synthesized in the cell bodies of these neurons, stored in the terminals and released into the hypophyseal portal blood. It is often considered that CRH is the primary stimulator of the release of ACTH, although it is clear that AVP from parvicellular neurons located in the paraventricular nucleus also plays a role. There may be species differences in the relative importance of these neuropeptides in stimulating the release of ACTH. Various stressful stimuli (stressors) activate the HPA axis, with the consequent release of CRH and AVP into the hypophyseal portal blood in a pulsatile pattern. Although the majority of CRH neurons projecting to the median eminence arise from the paraventricular nucleus, some CRH neurons from other hypothalamic nuclei project to the median eminence. CRH neurons also project to the brain stem and are found elsewhere, such as in limbic structures that are involved in processing sensory information and the regulation of the autonomic nervous system.

In addition to CRH, there are other CRH-like peptides, including urocortin I, urocortin II (or stresscopin-like peptide), and urocortin III (or stresscopin); at least two receptors (CRH-R1 and CRH-R2), and a CRH-binding protein. The CRH system is complex, and, in addition to the classic hypophysiotropic actions already described, it is important in regulating behavioral responses to stress. The activation of the HPA axis by CRH is mediated via CRH-R1, which is present in corticotropes. The binding of CRH to CRH-R1 stimulates the secretion of peptides derived from proopiomelanocortin (POMC) which, in addition to ACTH, includes β -endorphin, α -melanocyte-stimulating hormone, β -melanocyte-stimulating hormone, corticotropin-like intermediate peptide, γ -lipotropic hormone, and met-enkephalin.

Arginine Vasopressin

With the exception of the pig, all mammals studied synthesize AVP in the hypothalamus. In the pig, lysine is substituted for arginine, giving lysine vasopressin in this species. Within the hypothalamus, AVP is

synthesized in parvicellular neurons that project to the median eminence and magnocellular neurons that project to the posterior pituitary (Figure 1). The parvicellular neurons are located in the paraventricular nucleus and synthesize AVP in the cell bodies, store it in the terminals, and release it into the hypophyseal portal blood. There is an increase in AVP secretion during stress, and it travels by way of the hypophyseal portal system to the anterior pituitary gland to act synergistically with CRH to stimulate the release of ACTH from corticotropes. Some parvicellular neurons synthesize and store both CRH and AVP. The cell bodies of the magnocellular neurons that produce AVP are located predominantly in the supraoptic nucleus of the hypothalamus, and AVP is stored in the terminals in the posterior pituitary from which it is released directly into the systemic circulation. AVP released from the posterior pituitary is also known as antidiuretic hormone. Although this source of AVP is generally considered to function mainly to control fluid homeostasis and, in extreme stress, effect vasoconstriction, recent studies suggest that AVP from magnocellular neurons may also stimulate ACTH secretion.

AVP acts on three different receptors (V1a, V1b, and V2), which are expressed differently in various tissues and exert different actions. The actions of AVP to stimulate the release of ACTH are mediated by the V1b receptor. Many authors consider CRH, not AVP, to be the major stimulator of ACTH synthesis and secretion. Indeed, textbooks and scientific publications often fail to recognize the role of AVP as an ACTH secretagogue. Nevertheless, it is clear that AVP plays an important role in stimulating the secretion of ACTH during stress, although its relative importance may differ between species. For example, the majority of research in rats suggests that CRH is the major ACTH secretagogue, with AVP playing more of a permissive role, perhaps potentiating the effects of CRH. In contrast, in sheep there is evidence that AVP plays a major role in stimulating ACTH secretion, although other work has suggested that CRH is more important. It also appears likely that different stressors may influence the extent to which CRH and AVP stimulate ACTH secretion.

Regulation of Corticotropin Releasing Hormone and Arginine Vasopressin Secretion

The synthesis and secretion of CRH and AVP are influenced by both excitatory and inhibitory neural inputs from various regions of the brain that process information resulting from different stressors, negative feedback by glucocorticoids, and circadian rhythms. CRH and AVP neurons in the paraventricular nucleus

receive direct input from the limbic system and brain stem and indirect input from the cerebral cortex. Major excitatory input to CRH and AVP neurons arises from noradrenergic afferents of the brain stem (A_1 , A_2 , and A_6 noradrenergic cell groups), and there are reciprocal projections between CRH and AVP neurons in the paraventricular nucleus and neurons in the arcuate nucleus that produce POMC products, including the opioid β -endorphin. Glucocorticoids exert negative feedback by acting directly on the neurons that produce CRH and AVP and through other brain regions that have afferent inputs to the paraventricular nucleus. Glucocorticoids also act directly on pituitary corticotropes. The negative feedback actions of glucocorticoids regulating CRH and AVP synthesis are exerted via high-affinity mineralocorticoid receptors (MRs) and low-affinity glucocorticoid receptors (GRs). MRs are localized in the hippocampus and other regions of the limbic brain such as the lateral septum and amygdala, as well as hypothalamic sites. The glucocorticoids act via MRs to maintain basal activity. GRs have a more extensive distribution within the brain, being most abundant in the hypothalamus and in pituitary corticotropes, and are involved in the feedback actions of both basal and stress-induced levels of glucocorticoids. The circadian regulation of the HPA axis is illustrated by the circadian rhythms of CRH and AVP levels, which peak in the early morning and fall during the day in diurnal species. The negative feedback effects of glucocorticoids also vary in a circadian fashion.

Opioids

The opioid peptides and their receptors are distributed widely in the body. This distribution includes the central, peripheral, and autonomic nervous systems as well as several endocrine tissues and various target organs. The opioids are involved in a broad range of functions, including acting as neuropeptides to regulate stress responses. There are three classes of opioids: β -endorphin, the enkephalins (met-enkephalin and leu-enkephalin), and the dynorphins. It is often considered that the affinity of β -endorphin, the enkephalins, and dynorphins is highest for the μ -, δ -, and κ -opioid receptor subtypes, respectively. However, it is also apparent that opioids can bind to more than one receptor subtype.

Peptides Derived from Proopiomelanocortin

As previously indicated, a large number of peptides are derived from processing POMC. This includes the opioid β -endorphin. Peptides derived from POMC

are found in the brain, anterior pituitary gland, gut, thyroid, thymus, adrenal medulla, pancreas, spleen, macrophages, placenta, testes, and ovaries. There is a wide distribution of β -endorphin within the brain, with the highest concentrations occurring in the medial hypothalamus and arcuate nucleus. Acting as a neuropeptide and neurotransmitter, β -endorphin is involved in a range of stress-related functions, including the modulation of CRH and AVP neurons in the paraventricular nucleus. Furthermore, during stress, the secretion of β -endorphin from the anterior pituitary gland increases in parallel with ACTH. The physiological significance of this is not known, although it is possible that this source of β -endorphin contributes to some of the known actions of β -endorphin during stress. These actions include analgesia and the regulation of blood glucose concentrations and the immune system.

Peptides Derived from Pro-Enkephalin

Pro-enkephalin (also referred to as pro-enkephalin A) is the major biological source of met-enkephalin and is a source of leu-enkephalin. Met-enkephalin-containing peptides are present in the central nervous system, adrenal medulla, gastrointestinal tract, and many other tissues. The neurons that synthesize met-enkephalin are extensively distributed throughout the central nervous system, making enkephalin the most widely distributed opioid in the brain. Pro-enkephalin neurons exist throughout the hypothalamus, including the paraventricular nucleus. The chromaffin cells of the adrenal medulla also synthesize pro-enkephalin. Met-enkephalin is often cosecreted with the catecholamines from the adrenal medulla into the peripheral circulation during stress. The biological significance of this is also unknown.

Peptides Derived from Pro-Dynorphin

Pro-dynorphin (also referred to as pro-enkephalin B) can be processed to produce dynorphin A, dynorphin B, α -neoendorphin, and β -neoendorphin. Pro-dynorphin is found in many parts of the central nervous system, including the hypothalamus, hippocampus, brain stem, amygdala, and gastrointestinal tract. The neurons that produce dynorphins in the hypothalamus are mainly located in the supraoptic nucleus and paraventricular nucleus. Dynorphins have greatest affinity for the κ -opioid receptor. Less is known about the dynorphins than the other classes of opioids, although it is apparent that they have many different physiological actions depending on their site of production. These actions include influencing the pattern of electrical activity

of magnocellular neurons that produce AVP and negative feedback regulation of oxytocin secretion.

Stress-Related Functions of Opioids

Opioids exert a diverse range of stress-related actions. In particular, opioids act to attenuate or terminate stress responses. This modulatory influence on stress responses may provide a mechanism for preventing the detrimental effects of sustained or chronic stress. Specifically, opioids can inhibit the activity of the HPA axis. At least part of this mechanism involves the inhibition of the synthesis of CRH and, possibly, AVP. The reciprocal interactions between CRH and AVP neurons in the paraventricular nucleus and β -endorphin neurons in the arcuate nucleus are clearly important for this modulatory effect (Figure 1). Further, opioids regulate sympathetic, cardiovascular, and respiratory neural control systems and are involved in the regulation of pain, reinforcement and reward, the release of neurotransmitters, and other autonomic and neuroendocrine functions.

Opioids are involved in the control of pain and play a major role in suppressing the perception of pain (i.e., analgesia). In particular, opioids are mediators of stress-induced analgesia. The phenomenon of stress-induced analgesia is well known. There are many accounts of people being unaware of pain during a stressful experience. The origin of pain is normally at the site of intense stimulation or tissue damage, causing the discharge of impulses in a specialized set of nerve endings (nociceptors). The intensity of the excitation is regulated by local chemical signals that are released when the tissue is damaged. These include the opioids. Information about painful stimulation is carried as impulses in afferent nerve fibers that synapse in the dorsal horn of the spinal cord. Postsynaptic neurons convey the nociceptive information to the brain. Transmission at these synapses is modulated by locally released neurotransmitters, including opioids. The perception of the pain can be altered by modulation in the dorsal horn. The brain influences the processing of nociceptive input in the spinal cord, partly by activating opioid interneurons in the dorsal horn. It seems that both opioids produced in the brain and circulating opioids are involved in the induction of analgesia, with the source of the opioids that induces the analgesia varying with the type of stressor.

Oxytocin and Stress

Oxytocin is synthesized in cell bodies of magnocellular neurons located principally in the paraventricular

nucleus of the hypothalamus, is stored in the terminals of these neurons in the posterior pituitary, and is released from there into the peripheral circulation (Figure 1). Oxytocin is also synthesized in neurons that are widely distributed within the central nervous system. The actions of oxytocin are mediated via specific high-affinity receptors that are distributed both peripherally and in many forebrain areas, including regions of the hippocampus and amygdala, bed nuclei of the stria terminalis, ventrolateral septum, and several hypothalamic nuclei. Thus, oxytocin has both peripheral and central actions, some of which concern the regulation of stress responses. The classic peripheral actions of oxytocin are to cause contraction of the myoepithelial cells of the mammary glands to stimulate milk let-down and to stimulate the myometrium of the uterus during parturition. Oxytocin also has stress-related effects, and circulating concentrations of oxytocin increase in response to some, but not all, types of stress. Oxytocin can act within the brain to reduce the responsiveness of the HPA axis to stress with reduced cortisol secretion due, at least in part, to a reduction in CRH synthesis. Oxytocin can also reduce blood pressure, increase tolerance to pain, and reduce anxiety. Oxytocin facilitates mother–infant interactions and tends to facilitate behaviors that oppose fight-or-flight behavioral responses to stress. Lactating females (in which oxytocin levels are elevated) generally display attenuated responses to stress, and although a number of different mechanisms may be involved, oxytocin is likely to play a role. Generally, oxytocin is not considered to readily cross the blood–brain barrier although this dogma has recently been challenged. It is difficult to know whether the actions of oxytocin to dampen stress responses are due to central and/or peripheral effects.

Appetite-Regulating Neuropeptides and Stress

Various neuropeptides that regulate appetite can affect the activity of the HPA axis in response to stress. These include orexin, neuropeptide Y (NPY), agouti-related protein, and cocaine and amphetamine-regulated transcript (CART).

The peptides orexin-A and orexin-B are produced in the hypothalamus and can modulate the activity of the HPA axis, both centrally and peripherally. Orexin receptors are widely distributed and are found in the paraventricular nucleus, median eminence, pituitary corticotropes, adrenal cortex, and adrenal medulla. These neuropeptides can stimulate the release of

CRH and AVP as well as having direct stimulatory effects on the adrenal cortex.

NPY may mediate the excitatory effects of orexins. Certainly NPY can influence the activity of the HPA axis, although it is not clear how because NPY has been found to both inhibit and stimulate the release of CRH. The synthesis of NPY increases in the arcuate nucleus and hilus of the dentate gyrus following restraint stress in rats.

Agouti-related protein is an appetite stimulant that is coexpressed in the arcuate nucleus with NPY, and there is differential regulation of this neuropeptide in relation to stress. For example, foot shocks in rats decreased agouti-related protein synthesis and increased NPY synthesis.

CART is synthesized in neurons in various hypothalamic structures, including the paraventricular and arcuate nuclei. Glucocorticoids can influence the synthesis of CART, and CART can act within the paraventricular nucleus to activate the HPA axis.

See Also the Following Articles

Adrenocorticotrophic Hormone (ACTH); Corticotropin Releasing Factor (CRF); Glucocorticoid Negative Feedback; Hypothalamic-Pituitary-Adrenal; Anatomy of the HPA Axis; Paraventricular Nucleus; Urocortins; Vasopressin; CRH Anxiety Circuitry in Brain; Neuropeptide Y; Orexin.

Further Reading

- Altemus, M. (1995). Neuropeptides in anxiety disorders: effects of lactation. *Annals of the New York Academy of Sciences* 771, 697–707.
- Charmandari, E., Tsigos, C. and Chrousos, G. (2005). Endocrinology of the stress response. *Annual Review of Physiology* 67, 259–284.
- Chrousos, G. P. (1998). Stressors, stress, and neuroendocrine integration of the adaptive response: the 1997 Hans Selye memorial lecture. *Annals of the New York Academy of Sciences* 851, 311–335.
- de Kloet, E. R. (2004). Hormones and the stressed brain. *Annals of the New York Academy of Sciences* 1018, 1–15.
- Herman, J. P., Figueiredo, H., Mueller, N. K., et al. (2003). Central mechanisms of stress integration: hierarchical circuitry controlling hypothalamo-pituitary-adrenocortical responsiveness. *Frontiers in Neuroendocrinology* 24, 51–180.
- Sapolsky, R. M. (1998). Why zebras don't get ulcers: an updated guide to stress, stress related diseases and coping. New York: W. H. Freeman.
- Tilbrook, A. J., Turner, A. I. and Clarke, I. J. (2002). Stress and reproduction: central mechanisms and sex differences in non-rodent species. *Stress* 5, 83–100.
- Tilbrook, A. J. and Clarke, I. J. (2006). Neuroendocrine mechanisms of innate states of attenuated responsiveness of the hypothalamo-pituitary adrenal axis to stress. *Frontiers in Neuroendocrinology* 27, 285–307.

Neurosis

D J Lynn

Thomas Jefferson University, Philadelphia, PA, USA

© 2007 Published by Elsevier Inc.

The term neurosis was used officially with confidence and authority in American psychiatry as recently as 1980, but, at least since the 1994 *Diagnostic and Statistical Manual of Mental Disorders* (4th edn.; DSM-IV), it has been officially abandoned. For official American psychiatry, it is now of primarily historical interest. To understand the evolution of the use of this term, and the related controversy, is to understand a great deal about the recent changes within American psychiatry.

The Scottish Physician William Cullen (1710–1790) introduced the term in his *First Lines of the Practice of Physic*, which first appeared in English in 1777. Cullen used the term neuroses to refer to one of four general categories of illnesses. The other three

were pyrexiae (febrile diseases), cachexiae (diseases resulting from bad habits), and locales (diseases consisting of local lesions). The category Cullen named neuroses included all conditions of the central nervous system not related to fever or focal lesions.

During the nineteenth century, the term evolved into psychoneuroses, which Richard von Kraft-Ebing defined in 1872 as mental illnesses in cases in which the brain was normal.

Neuroses were at the heart of the psychoanalytic writings of Sigmund Freud (1856–1939). He was most concerned with psychoneuroses, which he distinguished from actual neuroses by their etiologies. For Freud, psychoneuroses were caused by unconscious conflicts, generally of a sexual nature, with roots in childhood. Actual neuroses were caused by current factors within the patient's awareness. In American psychiatry in the mid-twentieth century, the influence of Freud's writings and the writings of his followers was extremely strong, and in DSM-II,

the term neurosis was used pervasively. In general in this era, the unconscious conflict believed to be causing the manifest neurotic symptoms attracted much more attention than the symptoms themselves. For any particular patient, the symptoms were seen as a product of, and in some sense a solution to, that patient's specific unconscious conflict. For psychoanalytically oriented psychiatrists, attention to the symptoms was important not for the purposes of arriving at a prognosis or conducting research but for delineating, and interpreting for the patient, the unconscious conflict.

DSM-III, published in 1980, was a systematic attempt to redirect attention to observable and quantifiable symptoms. Disorders were classified in a descriptive manner; the neuroses, so prominent in DSM-II, were discarded, and the term neurosis was relegated to parentheses. The authors stated that "at the present time, there is no consensus in our field as to how to define 'neurosis'" American Psychiatric Association (1980: 9). Some psychiatrists were using the term in a descriptive sense to denote a pattern of observable symptoms; others used it to designate a psychogenic etiology. DSM-III used an atheoretical approach concerning the disorders formerly termed neuroses for two stated reasons: (1) the etiology of these disorders was declared to be unknown and (2) the goal was to arrive at a system of classification that could be used by clinicians whether or not they agreed with particular theories of etiology. The editors suggested that the term neurotic disorder could be used, but only descriptively (to refer to an observable symptom pattern), and that, whenever psychiatrists wanted to convey an etiology based in unconscious conflict, the term neurotic process should be used. DSM-III-R (1987) basically continued the use of its predecessor, but DSM-IV (1994) contains no mention of the term neurosis.

For some psychoanalytic psychiatrists, these developments led to consequences that were seen as unfor-

tunate. As Glen Gabbard put it: "although this classification accommodates the recent biological research delineating different kinds of anxiety, it also encourages clinicians to think about anxiety as an *illness* rather than as an overdetermined *symptom* of unconscious conflict".

See Also the Following Articles

Anxiety; Depression Models; Psychoanalysis.

Further Reading

- American Psychiatric Association (1968). *Diagnostic and statistical manual of mental disorders* (2nd edn.). Washington, DC: APA Press.
- American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd edn.). Washington, DC: APA Press.
- American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd edn. rev.). Washington, DC: APA Press.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th edn.). Washington, DC: APA Press.
- Freud, S. (1953–1974). Anxiety, instinctual life (1933). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vol. 22), pp. 81–111. London: Hogarth Press.
- Freud, S. (1953–1974). Sexuality in the aetiology of the neuroses (1898). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vol. 3), pp. 263–285. London: Hogarth Press.
- Gabbard, G. O. (2000). *Psychodynamic psychiatry in clinical practice* (3rd edn.). Washington, DC: APA Press.
- Person, E. S., Cooper, A. M. and Gabbard, G. O. (eds.) (2005). *The American Psychiatric Publishing textbook of psychoanalysis*. Washington, DC: APA Press.
- Shorter, E. (2005). *A historical dictionary of psychiatry*. New York: Oxford University Press.

Neuroticism, Genetic Mapping of

M W Nash

King's College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

What Is Neuroticism, and Is It Genetic?

Which Regions of the Genome Are Linked to Neuroticism?

Have Association Studies Identified Any Genes Associated with Neuroticism?

Roles of Candidate Genes in Stress

Glossary

<i>Complex traits</i>	Phenotypes governed by variation in many genes; often continuously and/or normally distributed. The trait does not segregate in the population following any obvious patterns. Environmental influences often account for as much as, or more than, genetic influences.
<i>Genome</i>	The set of all genes for a particular organism. In humans this includes both nuclear and mitochondrial DNA sequences.
<i>5-HTTLPR</i>	Gene polymorphism in the 5'-promoter region of the gene encoding the serotonin (5-hydroxytryptamine; 5-HT) transporter (5-HTT) that has been implicated in susceptibility to life events/stressors.
<i>LOD score</i>	Logarithm of the odds score favoring linkage. The higher the score, the greater the evidence for linkage.
<i>Mendelian or simple traits</i>	Phenotypes governed by a single major locus or gene and hence the traits can be seen to segregate in the population following obvious additive, dominant, or recessive patterns.
<i>Phenotype</i>	Any observable and measurable characteristic of an organism produced by genetic and/or environmental influences.
<i>Quantitative trait locus (QTL)</i>	A locus implicated in genetic variation that causes a change in a continuous phenotype.

What Is Neuroticism, and Is It Genetic?

Neuroticism is a unique dimensional measure of personality thought to capture emotional stability and a temperamental sensitivity to negative stimuli. Indeed, it is a widely agreed upon higher order factor in many different personality constructs. Currently, three measures of personality are commonly used in genetic research: Eysenck's Personality Questionnaire, Cloninger's Temperament and Character Inventory, and

the Neuroticism Extraversion Openness (NEO) personality inventory. For Cloninger's measure of harm avoidance (HA) and Eysenck's measure of neuroticism (EPQ-N), there is some evidence that these different personality constructs are influenced by similar genetic and environmental variables. In the case of the EPQ-N, high scorers are conceived to be overly emotional, react strongly to stimuli, and find it difficult to recover after an emotionally stimulating experience, while low scorers are conceived to be calm, even tempered, and emotionally controlled. The genetic influence on neuroticism and harm avoidance has been widely assessed, and the estimates of heritability are 40–60% of the phenotypic variance. The EPQ-N is known to overlap genetically with many internalizing disorders, such as major depression, panic disorder, and generalized anxiety disorder, and because of this, neuroticism has been used in psychiatric genetic research to identify the genetic variation causing individual differences in the susceptibility to both depression and anxiety.

Which Regions of the Genome Are Linked to Neuroticism?

Traditionally, there have been two statistical approaches to relate variation of the genome to a phenotype of interest, known as linkage mapping and association mapping. Linkage mapping has played a huge historical role in the gene mapping of Mendelian traits. In essence, the method relies on the phenomenon of allele sharing among relatives. For example, monozygotic twins share 100% of their genome, while siblings and dizygotic twins should share only 50% – on average – of their genome. To illustrate this, if two siblings are both highly neurotic, then genetic variation at a region of the genome that is truly linked to neuroticism would be expected to be shared by the siblings. If such a phenomenon is observed in more siblings than would be expected by chance, this region of the genome is declared linked to neuroticism. Linkage is a useful approach for coarse mapping, as it identifies large regions of the genome as being linked to the phenotype, but this is also a disadvantage, as large regions can contain hundreds of candidate genes. As this analysis cannot distinguish which gene or genetic variant is responsible for the linkage, it is usual after the identification of a region to perform the finer association mapping approach to localize more precisely the disease gene (see next section).

To date, there have been seven linkage studies examining a neuroticism-like phenotype. The first

reported study, undertaken by Cloninger and colleagues, examined HA and observed significant linkage to the long (p) arm of chromosome 8. Additive and nonadditive interaction of the chromosome 8 peak with other loci suggested significant interaction with chromosomes 11, 18, 20, and 21. A subsequent study assessed selected regions of the genome for linkage to HA and also found linkage in an overlapping region of chromosome 8. It is not possible, however, to deduce that the same sequence variant(s) is responsible for the linkage in both these studies. Further fine mapping of the chromosome 8p region in the same sample refined the region to one around the neuroregulin gene (*NRG1*), which has been implicated in schizophrenia susceptibility. The *NRG1* protein is involved in signaling based on the glutamate neurotransmitter and plays a central role in neural development. Interestingly, neuroticism and HA have been reported as elevated in first-degree relatives of schizophrenics and schizotypal individuals. A re-analysis of the original data set, using a novel method, reproduced the evidence for linkage to the chromosome 8p markers. As these four studies are derived from only two samples, it is difficult to decide if a true quantitative trait locus (QTL) actually resides on chromosome 8p. The possibility that *NRG1* is a candidate for harm avoidance is intriguing, but the gene needs to be extensively examined in multiple samples before any robust conclusion can be drawn.

The first published linkage analysis of the EPQ-N was carried out by Fullerton and colleagues in a UK selected sibling sample. The main findings were QTLs on chromosomes 1, 4, 7, 12, and 13. Further evidence was presented that these loci for neuroticism might be gender specific: the QTLs on chromosomes 1, 12, and 13 appeared to be female specific, while a QTL on chromosome 8, which was not found to be genome-wide significant in the complete sample analysis, appeared to be male specific. It is interesting to note that the chromosome 8 locus is in close proximity to the chromosome 8p locus identified in HA. The second published study used a composite index of four self-report measures of anxiety and depression, including Eysenck's measure of neuroticism in a large UK-based sibship sample. As opposed to only examining Eysenck's neuroticism measure, Nash and colleagues used structural equation modeling on the four self-report measures to derive a phenotypic composite index of the shared genetic liability to both depression and anxiety. While no significant findings were obtained, linkage peaks closely approaching significance were obtained on chromosomes 1p and 6p for both the composite index and neuroticism, with a high correspondence of results between the two phenotypes. Splitting the sample by gender led to the identification of a linkage peak for females on

chromosome 12, while also indicating female specificity of the chromosome 1 peak and male specificity of the chromosome 6 peak. The most recent linkage publication on EPQ-N was performed in a New Zealand sample. While no peak reached statistical significance, a suggestive finding of linkage was obtained on chromosome 12, with nominal findings of linkages on chromosomes 1, 3, 6, 11, 15, and 17. Comparing the results from these three studies indicates a low correspondence between the results. There is possible overlap on chromosomes 1 and 12. The peak on chromosome 12 is potentially exciting, as this also overlaps with linkage peaks from both unipolar and bipolar depression. Indeed, as discussed later, variants of a gene within this peak have recently been observed to segregate with major depression and might also be associated with individual variation in neuroticism.

Similarly, comparisons between all the studies on neuroticism and harm avoidance show only slight correspondence. This is not unexpected, unless a major locus (responsible for >20% phenotypic variance) was involved. While linkage is a powerful method for simple traits with genes of major effect, its poor ability to reliably detect loci with a small phenotypic effect relegates the method to a preliminary tool providing genomic targets on which to focus fine mapping projects. Its contributions in this field at best provide suggestive evidence that chromosomes 1, 8, and 12 might harbor genetic variants that are associated with neuroticism/harm avoidance.

Have Association Studies Identified Any Genes Associated with Neuroticism?

Association mapping has historically been employed to evaluate the potential contribution of specified candidate genes, or as a follow-up, fine mapping tool to linkage analysis. Due to technological advances in speed and the reduction in expense in measuring differences in the genome, whole genome association mapping will eventually become the predominant mapping approach. The simplest version of association mapping is in case control studies of unrelated individuals and essentially examines whether or not there is a correlation between the genetic variation and the phenotype. This basic approach suffers from the fact that a much larger number of sequence variants need to be measured to capture adequately the effective genetic variation compared to linkage analysis. Nevertheless, the increased ability to detect small changes is the most attractive feature of this approach to gene mapping.

Function-Based Candidate Genes

Perhaps the biggest shift in thinking with respect to psychiatric and personality genetics has been the

reduction in the expected effect size of a single genetic variant. While initial expectations were that a single genetic variant might account for around 10% of the phenotypic variance, it is now thought that 1%, or much less, is a more realistic expectation for single variants. Commonly, sample sizes are inadequate to detect such effects reliably. To tackle this problem, meta-analytic techniques can be implemented that use the results of multiple studies to estimate reliably the effect sizes and significance of genetic variants. In the only published meta-analysis of genetic polymorphisms in personality, Munafo and colleagues examined commonly researched genetic variants in the serotonin transporter (*SLC6A4*) and various dopamine genes (*DRD4*, *DRD2*, *DRD3*, and *DAT1*). Of all the results, the only significant meta-analytic association was between a polymorphism in *SLC6A4* and neurotic/avoidant traits.

Current evidence indicates that the serotonin transporter gene (*SLC6A4*), located on the long arm of chromosome 17 (17q), is probably the only gene that can be considered associated with neuroticism, albeit only weakly. The mature protein of this gene transports serotonin from the synaptic cleft back into the presynaptic neuron for reuse and is responsible for the quantity of serotonin in the synaptic cleft. A sequence variant, known as 5-HTTLPR, is present in either a short or long form in a region of the gene called the promoter, which regulates the rate at which the mRNA of the gene is made. The first study of this variant identified an association of the number of short variants with an increase in neuroticism. Subsequently, there have been many attempts at replicating this finding with varying degrees of success. To address this issue, there have been several meta-analyses of the many findings, with a general consensus supporting a small but positive association. Recently, Caspi and colleagues observed that the relationship of stressful life events and depression was mediated by 5-HTTLPR, such that individuals with two long alleles were unchanged in their risk to depression as life events increased, while the risk of depression was increased in those carrying one or more short alleles. This finding has subsequently been replicated for depression, but comprehensive studies looking only at neuroticism and 5-HTTLPR have found no mediation of the variant on the effect that life events have on neuroticism. It is highly probable that 5-HTTLPR and neuroticism are weakly associated; what is unknown, and a source of current interest, is the developmental mechanisms that 5-HTTLPR affects, which lead to its association with individual differences in neuroticism.

A more recent line of research has been driven by theories of aberrant neuronal apoptosis and

neurogenesis in major depression. Based on such theories, some research has been carried out on genetic variants known to affect these mechanisms. The most studied of these is the brain-derived neurotrophic factor gene (*BDNF*), located on chromosome 11p. The BDNF protein has a well-documented role in neuronal development, survival, and plasticity, making it an attractive candidate for many psychiatric disorders. Indeed, variants in the gene have been associated with disorders such as obsessive-compulsive disorder, bipolar disorder, and major depressive disorder. Because neuroticism is genetically related to major depression and is a phenotypic marker for vulnerability to depression, genetic variants in the *BDNF* gene have been examined for their relationship with neuroticism. The original association with this gene examined a single nucleotide polymorphism that changes the 66th amino acid from a valine to a methionine and its association with neuroticism as measured by the NEO personality scale. This study found an apparently strong finding, with those individuals carrying the Met/Met genotype having substantially lower neuroticism scores than the rest of the sample. Subsequent follow-up studies using the NEO personality scale or HA have not found such strong evidence, and a recent comprehensive study examining the EPQ-N found no evidence for an association. It appears that although *BDNF* might be a plausible candidate gene for neuroticism, the data simply do not currently support this. As these studies have focused on only one variant in the gene, it is plausible that other variants in the gene are associated with neuroticism.

Position-Based Candidate Genes

Despite the promising leads from linkage studies, no positional candidates have emerged so far, but there have been some possible leads from related scientific fields. *RGS2*, or regulator of G-protein signaling 2, is a positional candidate gene from linkage studies that examined emotionality in mice. This gene is located on mouse chromosome 1 and corresponds to a region on human chromosome 1. Whether this gene is responsible for the linkage signal seen on this chromosome has not yet been established. Nevertheless, there have been reports that variants in the human version of this gene might be associated with agoraphobia. Another possible gene is *APAF1*, apoptosis protease activating factor 1, involved in the apoptosis pathway. This gene, located at 12q, was recently associated with major depression in Utah families. Functional studies indicated that only those variants that segregated with major depression had a functional effect on the apoptosis pathway. Given the genetic

relationship between the EPQ-N and major depression, it is possible that *APAF1* also contributes to individual differences in neuroticism. This gene is located within the chromosome 12 linkage peak identified in two studies of neuroticism, as mentioned previously, lending weight to the suggestion that genetic variation in this gene might be associated with neuroticism.

Roles of Candidate Genes in Stress

There are a number of possible mechanisms that might explain the relationship of the aforementioned candidate genes with stress. Upon being stressed, the body enters into various response patterns, including at the molecular level an increase in firing from serotonergic neurons. As the role of the serotonin transporter is the reuptake of serotonin from the synaptic cleft, the rate at which new transporter can be made might affect the rate at which the biochemical system regains homeostasis. Because the short, low-activity 5-HTTLPR allele carriers produce less mRNA, these individuals might return to their biochemical baseline slower than long allele homozygotes. This delay might then leave these individuals more susceptible to the detrimental effects of stress. Another aspect of these detrimental effects is the neuronal toxicity of glucocorticoids, which are released as a negative feedback mechanism during response to a stressor, and it is known that glucocorticoids are capable of damaging hippocampal neurons. The rate and impact of this damage might be mediated by both BDNF and *APAF1*. BDNF is involved in neuronal survival, while *APAF1* is involved in cell death. The functional variation that has been identified for these genes might cause some individuals to be more resilient to the toxic effects of stress, while making others more vulnerable. This model suggests that individuals who carry the S allele of 5-HTTLPR, who are homozygous for the BDNF Met allele, and who also have the apoptosis rate-increasing alleles of *APAF1* may be more vulnerable to the detrimental effects of stress.

See Also the Following Articles

Genetic Polymorphisms in Stress Response; Serotonin Transporter Genetic Modifications.

Further Reading

Caspi, A., Sugden, K., Moffitt, T. E., et al. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 301, 386–389.

Cloninger, C. R. (1986). A unified biosocial theory of personality and its role in the development of anxiety states. *Psychiatric Developments* 4, 167–226.

Dina, C., Nemanov, L., Gritsenko, I., et al. (2005). Fine mapping of a region on chromosome 8p gives evidence for a QTL contributing to individual differences in an anxiety-related personality trait: TPQ harm avoidance. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics* 132, 104–108.

Fullerton, J., Cubin, M., Tiwari, H., et al. (2003). Linkage analysis of extremely discordant and concordant sibling pairs identifies quantitative-trait loci that influence variation in the human personality trait neuroticism. *American Journal of Human Genetics* 72, 879–890.

Lake, R. I., Eaves, L. J., Maes, H. H., et al. (2000). Further evidence against the environmental transmission of individual differences in neuroticism from a collaborative study of 45,850 twins and relatives on two continents. *Behavior Genetics* 30, 223–233.

Lesch, K. P., Bengel, D., Heils, A., et al. (1996). Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* 274, 1527–1531.

Munafo, M. R., Clark, T. G., Moore, L. R., et al. (2003). Genetic polymorphisms and personality in healthy adults: a systematic review and meta-analysis. *Molecular Psychiatry* 8, 471–484.

Nash, M. W., Huezo-Diaz, P., Williamson, R. J., et al. (2004). Genome-wide linkage analysis of a composite index of neuroticism and mood-related scales in extreme selected sibships. *Human Molecular Genetics* 13, 2173–2182.

Neale, B. M., Sullivan, P. F. and Kendler, K. S. (2005). A genome scan of neuroticism in nicotine dependent smokers. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics* 132, 65–69.

Sen, S., Nesse, R. M., Stoltenberg, S. F., et al. (2003). A BDNF coding variant is associated with the NEO personality inventory domain neuroticism, a risk factor for depression. *Neuropsychopharmacology* 28, 397–401.

Willis-Owen, S. A., Fullerton, J., Surtees, P. G., et al. (2005). The Val66Met coding variant of the brain-derived neurotrophic factor (BDNF) gene does not contribute toward variation in the personality trait neuroticism. *Biological Psychiatry* 58, 738–742.

Willis-Owen, S. A., Turri, M. G., Munafo, M. R., et al. (2005). The serotonin transporter length polymorphism, neuroticism, and depression: a comprehensive assessment of association. *Biological Psychiatry* 58, 451–456.

Zohar, A. H., Dina, C., Rosolio, N., et al. (2003). Tridimensional personality questionnaire trait of harm avoidance (anxiety proneness) is linked to a locus on chromosome 8p21. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics* 117, 66–69.

Neuroticism Response to Stress, Genetic Mapping of Mice

M R Munafò

University of Bristol, Bristol, UK

J Flint

University of Oxford, Oxford, UK

© 2007 Elsevier Inc. All rights reserved.

Personality and Neuroticism

Animal Models of Emotionality

Genetic Mapping in Rodents

Comparisons between Rodents and Humans

Glossary

<i>Allele</i>	An alternative form of a genetic locus; a single allele for each locus is inherited separately from each parent.
<i>Backcross</i>	A cross between one animal type that is heterozygous for alleles obtained from two parental strains and a second animal type from one of those parental strains. The term is often used by itself to describe the two-generation breeding protocol of outcross followed by a backcross that is used frequently in linkage analysis.
<i>F1</i>	The first filial generation; the offspring resulting from the first experimental crossing of animals. The progeny produced by intercrossing F1 individuals are referred to as F2 individuals.
<i>Heterozygous</i>	Having two different alleles at a given locus.
<i>Homozygous</i>	Having two identical alleles at a given locus.
<i>Inbred strain</i>	Animals that result from the process of at least 20 sequential generations of brother–sister matings. This process is called inbreeding.
<i>Intercross</i>	A cross between two animals that have the same heterozygous genotype at designated loci, for example, between sibling F1 hybrids that were derived from an outcross between two inbred strains.
<i>Outcross</i>	A cross between genetically unrelated animals.
<i>Neuroticism</i>	A personality trait dimension that refers to emotionality, emotional instability, and stress reactivity.
<i>Quantitative trait locus (QTL)</i>	A locus that affects a quantitative trait.
<i>Syntetic genes</i>	Genes thought to reside on the same chromosome.

Personality and Neuroticism

Broad behavioral traits underlie variation in human personality, and there is general consensus that certain of these relate to individual differences in response to stress and in vulnerability to anxiety and depression in humans. Behavioral evidence of anxiety-related traits has been observed from amphibians to rodents and higher mammals. In humans, these traits have been described variously as neuroticism, trait anxiety, and harm avoidance, and there is now general agreement that any adequate trait model of personality must include a broad trait related to vulnerability to anxiety- and depression-related disorders, representing an evolutionarily conserved capacity for fear and anxiety.

The notion that a single underlying risk factor underlies vulnerability to both anxiety- and depression-related disorders is supported by evidence from twin studies in humans that these disorders share a common genetic basis. Moreover, numerous studies have arrived at similar estimates of the contribution of genetic variance to variation in trait neuroticism, with additive genetic variance estimated at approximately 30% and nonadditive variance at approximately 15%. This has led to attempts to better understand the molecular genetic architecture of trait neuroticism in humans likely to result from multiple low-magnitude genetic effects and their interactions, both with other genetic loci and environmental factors; however, this research has generally yielded disappointing and inconsistent results.

Better progress has been made in studies of animal models of emotionality, which offer several attributes useful in genetic research, including short gestation, early puberty, and large litters, as well as a greater degree of experimental control by means of directed mating and environmental control, all of which are valuable when attempting to map the genetic architecture of these traits.

Animal Models of Emotionality

Defining an animal's temperament is not easy, and operational definitions typically require the animal to show a consistent pattern of response to a behavioral task or series of tasks. For example, an animal model of emotionality might be operationally defined as the length of time it takes an animal to emerge into an aversive or threatening environment. This obviously requires relatively strong assumptions regarding the

Table 1 Animal tests of emotionality

<i>Unconditioned responses</i>	
Open-field arena	Activity, defecation, thigmotaxis
Elevated plus-maze	Ratio of entries to open and closed arms
Light-dark box	Emergence time
Hyponeophagia	Latency to start eating
Holeboard	Head dipping
Social interaction	Frequency of social interaction in a novel environment
<i>Conditioned responses</i>	
Two-way avoidance conditioning	Rate of acquisition of response
Acoustic startle	Contraction in response to loud noise
Foot shock-induced freezing	Contraction in response to conditioned stimulus
Fear-potentiated startle	Contraction in response to loud noise in conjunction with conditioned stimulus
Geller-Seifter	Frequency of conditioned response coincidental to an electrical shock
Vogel conflict	Frequency of conditioned licking coincidental to an electrical shock

extent to which this operational definition is an accurate analog of the same construct in humans. The majority of animal studies of emotionality have employed rodent models (i.e., mouse or rat). [Table 1](#) lists a number of tests of emotionality in animals, divided into measures that depend on unconditioned (ethological) and conditioned responses. Conditioned responses have the advantage of offering much more experimental control. These measures are founded on principles of avoidance, autonomic activation, and behavioral inhibition (i.e., the discontinuation of species-typical behaviors such as exploration).

The open-field apparatus is one widely used paradigm and consists of a circular, white, brightly lit, and fully enclosed arena that allows remote monitoring of behavior over a short period. Unfortunately, many of these tests rely on features unrelated to the trait of interest (i.e., emotionality). For example, many depend in part on the basal activity level of the animal, which may be a potential confound. A potential solution to this problem is outlined here.

Genetic Mapping in Rodents

The mouse genome is well characterized, comprising approximately 22 000 predicted genes, of which approximately 80% have an identifiable human ortholog. Inbred rodent strains exhibit substantial variability in emotionality, and these differences can be exploited for the purposes of quantitative trait mapping. Crosses between inbred strains of rodents are frequently used to map the quantitative trait loci (QTL) that give rise to the genetic component of variation in behavioral traits such as emotionality. The basic experimental design is the analysis of association between genotypic and phenotypic variation in a cross between two inbred strains of rodents (usually, but not necessarily, with contrasting emotionality

phenotypes). Several strategies exist, but in the most widely used either the offspring of the cross (the F1 generation) are mated to produce a F2 intercross or the offspring are backcrossed to either of the parental strains. Molecular genetic markers are then used to determine which chromosomal segments segregate with the trait (i.e., which chromosomal regions are shared by animals that are phenotypically similar for the trait of interest). Over the past 10 years, QTLs putatively related to emotionality have been identified on a number of chromosomes, although the molecular nature of these remains unclear.

One problem is that genetic mapping identifies a potential functional variant, rather than a gene, and functionally important variants that affect gene expression may lie some distance away from a gene or even in the location of an unrelated gene. Another problem is that this experimental design offers only a limited number of recombinants in a single generation (typically, between one and two), so that the resolution to map a specific locus is limited and may identify regions that can still contain hundreds or perhaps thousands of genes.

It is possible to remedy the problem of limited mapping resolution by mapping loci in genetically heterogeneous stocks of rodents that are generated from multiple (usually eight) inbred strains that have been successively intercrossed for maximal diversity and maintained over multiple generations through a program of pseudorandom mating. Because these animals are several generations removed from the original progenitor strains, this offers the potential of mapping a QTL to a limited number of genes, although this requires a substantial increase in marker density and the analysis is considerably more complex in this case. At present, this is most economical once broad chromosomal regions have been

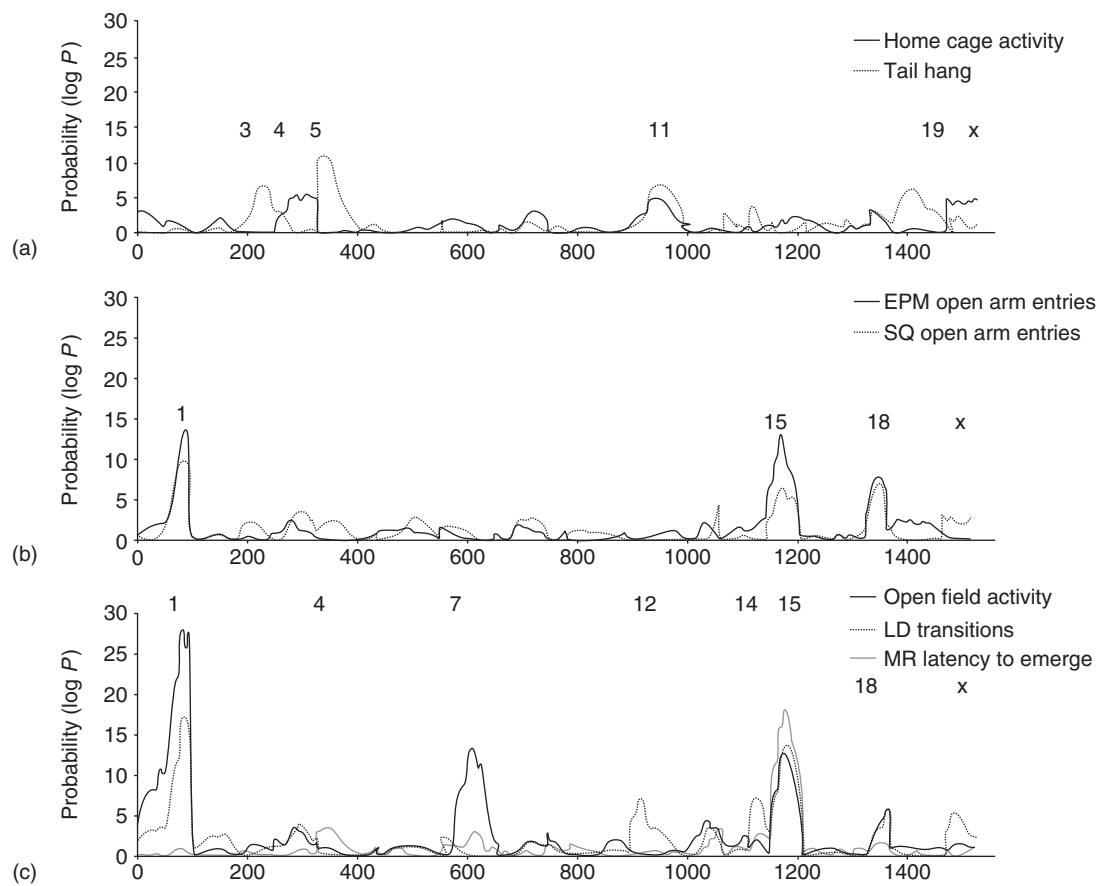


Figure 1 Quantitative trait loci that influence behavior in the De Fries strain of mice. a, Genome scan for two measures that control for putative confounds (home cage activity and tail hang); b, genome scan for two measures of emotionality (EPM open arm entries and SQ open arm entries); c, genome scan for three measures of emotionality (open field activity, LD transitions, and MR latency to emerge). Chromosomes that harbor significant genetic effects are shown as numbers on each panel. EPM, elevated plus-maze; SQ, square maze; LD, light–dark box; MR, mirror chamber. From Flint, J. (2004), The genetic basis of neuroticism, *Neuroscience Biobehavioral Review* 28, 307–316. © 2004 with permission from Elsevier.

identified, thereby limiting the number of markers required for QTL localization.

There also exists a potential solution to the problem of ensuring that the identification of loci related to emotionality in rodent models is specific to this trait. This is to attempt to identify chromosomal regions that influence multiple measures of emotionality, but not control measures, on the assumption that a single underlying trait should influence performance on multiple measures of emotionality to a comparable degree and should not influence performance on control measures.

The results of a recent attempt to do this are presented in Figure 1. Figure 1a identifies QTLs that influence control measures (i.e., potential confounds unrelated to emotionality), and Figures 1b–c identify QTLs that influence one or more of five tests of facets of emotionality related to emotionality (elevated plus-maze, square maze, open field arena, light–dark box, and mirror chamber).

Comparisons between Rodents and Humans

Recently, evidence for an overlap between the genetic basis for emotionality in rodent models and trait neuroticism in humans has emerged. Although a number of studies in humans have investigated individual genes, direct comparability with animal models is better achieved by a linkage analysis of human sibling pairs in which evidence of deviation from the expected value of allele sharing at a given point along the genome is evidence for the presence of a QTL at that locus. In this way, we can ask whether individuals who are behaviorally similar (i.e., have similar trait neuroticism scores) are also genetically similar (i.e., share two alleles). One locus on chromosome 1, identified in humans as potentially related to trait neuroticism, is potentially promising because it may be syntenic with loci identified in animal studies. The congruence of genetic mapping in human and

animal studies may help determine whether there is a common set of genes that contribute to trait neuroticism in humans and analog models of emotionality in animals.

Further Reading

Benjamin, J., Ebstein, R. P. and Belmaker, R. H. (2002). *Molecular genetics and the human personality*. Washington, DC: APA Press.

Flint, J. (2004). The genetic basis of neuroticism. *Neuroscience Biobehavioral Reviews* 28, 307–316.

Munafò, M. R., Clark, T. G., Moore, L. R., et al. (2004). Genetic polymorphisms in personality in healthy adults: a systematic review and meta-analysis. *Molecular Psychiatry* 9, 471–484.

Willis-Owen, S. A. and Flint, J. (2006). The genetic basis of emotional behavior in mice. *European Journal of Human Genetics* 14, 721–728.

Night Shiftwork

T Åkerstedt and G Lindbeck

Karolinska institutet, Stockholm, Sweden

© 2007 Elsevier Inc. All rights reserved.

Introduction
The Mechanism
Effects on Sleep
Alertness
Performance
Accidents
Stress Indices
Gastrointestinal Effects
Cardiovascular Effects
Mortality
Conclusion

Introduction

Shiftwork may not be the typical stressor in the classic sense, but it affects the systems of stress and restitution in a very pronounced way and causes cardiovascular disease, gastrointestinal disease, and insomnia, among other conditions. Here we review some of the connections. But before that we need to look at the phenomenon of shiftwork. This is an arrangement of work hours through which the daily duration of production/service extends to cover operation during 2–3 work days compressed into one 24-h period. Each day is called a shift. Normally, the first step of extension beyond a day shift is to introduce morning and evening shifts (6 a.m.–2 p.m. and 2 p.m.–10 p.m., respectively). When necessary, a night shift may also be used to cover the remaining 8 h. In Europe, workers often alternate among shifts, working the same shift two to four times in succession and then switch-

ing to a new shift. In transport, health care and the service sector, the shifts may be more variable in placement and duration, but the principle of alternating among different shifts remains. Even the permanent night worker alternates between a night shift and day life during his days off. Usually, the night shift is the most problematic, causing disturbed circadian rhythms, disturbed sleep, and excessive sleepiness and several diseases deriving from these disturbances.

The Mechanism

Shift work would not present a problem were it not for the biological clock that drives physiology and psychology through a 24-h cycle, with high and low periods of activity. The central oscillator is situated in the hypothalamus. It is a self-sustained oscillator and drives much of our physiology in an approximately 24-h cycle. It receives input from light and other stimuli that synchronize the pacemaker with the environmental light–dark cycle, called Zeitgebers, which entrain (influence) the biological clock to the light–dark changes of the normal environment. Essentially, light before the circadian trough (low) phase delays the biological clock (1–2 h) and light after the trough phase advances the clock.

The influence of the circadian cycle on sleep was demonstrated in monthlong isolation studies in which individuals could live according to their own preferred sleep–wake schedules. This led to the finding that some individuals developed a period of around 25 h, and that some individuals desynchronized, showing different period lengths for their sleep–wake rhythm and for their rhythm of metabolism (rectal temperature). However, the present estimate of the period length with control for light is 24.1 h. It was later shown that whenever the sleep–wake

rhythm placed sleep around the acrophase (circadian maximum) of rectal temperature, sleep was shortened, and when sleep was placed around the circadian trough phase, sleep was promoted. Thus, the biological clock interfered with sleep and wakefulness.

Effects on Sleep

The dominant health problem reported by shiftworkers is disturbed sleep and wakefulness. At least three-quarters of the shift-working population is affected. When comparing individuals with a very negative attitude to shiftwork with those with a very positive one, the strongest discriminator seems to be the ability to obtain sufficient quality of sleep during the daytime. Electroencephalograph (EEG) studies of rotating shiftworkers and similar groups showed that day sleep is 1–4 h shorter than night sleep. The shortening is due to the fact that sleep is terminated after only 4–6 h without the individual being able to return to sleep. The sleep loss is primarily taken out of stage 2 sleep (basic sleep) and stage rapid-eye movement (REM) sleep (dream sleep). Stages 3 and 4 (deep sleep) do not seem to be affected. Furthermore, the time taken to fall asleep (sleep latency) is usually shorter. Night sleep before a morning shift is also reduced, but the termination is through artificial means and the awakening is usually difficult and unpleasant. Interestingly, day sleep does not seem to improve much across series of night shifts. It appears, however, that night workers sleep slightly better (longer) than rotating workers on the night shift.

As indicated previously, shiftwork sleep duration is approximately 5.5–6 h after the night shift, and several laboratory studies have found that this amount of sleep curtailment may affect sleepiness moderately. The 2003 study by Van Dongen and colleagues indicates that the critical point for the accumulation of fatigue is approximately 7 h of sleep per night. The circadian trough and extended wakefulness (often up to 20 h) are other powerful contributors.

Alertness

Night-oriented shiftworkers complain as much about fatigue and sleepiness as they do about disturbed sleep. The sleepiness is particularly severe on the night shift, appears hardly at all on the afternoon shift, and is intermediate on the morning shift. The maximum is reached toward the early morning (5 a.m.–7 a.m.). Frequently, incidents of falling asleep occur during the night shift. At least two-thirds of the respondents report that they have experienced involuntary sleep during night work. Ambulatory EEG recordings verify that incidents of actual sleep occur during

night work in, for example, process operators. Other groups, such as train drivers or truck drivers, show clear signs of falling asleep while driving at night. This occurs toward the second half of the night and appears as repeated bursts of alpha and theta EEG activity, together with closed eyes and slow undulating eye movements. As a rule, the bursts are short (1–15 s) but frequent and seem to reflect let-downs in the effort to fend off sleep. Approximately one-quarter of the subjects recorded show the EEG/electrooculograph (EOG) patterns of fighting with sleep. This is clearly a larger proportion than is found in the subjective reports of episodes of falling asleep.

Performance

As may be expected, sleepiness on the night shift is reflected in performance. One of the classic studies in this area is by Bjerner and colleagues, who showed that errors in meter readings over a period of 20 years in a gas works had a pronounced peak on the night shift. There was also a secondary peak during the afternoon. Similar observations have been made for switchboard operators and other groups. Late-night shift performance capacity has been compared with performance following the ingestion of alcohol to a 0.08% blood alcohol level.

Accidents

If sleepiness is severe enough, interaction with the environment ceases, and if this coincides with a critical need for action, an accident may ensue. Most of the available accident data on night-shift sleepiness has been obtained from the area of transportation; the National Transportation Safety Board ranks fatigue as one of the major causes of heavy vehicle accidents.

Very little relevant data are available from conventional industrial operations, but fatal work accidents show a higher risk in shiftworkers and accidents in the automotive industry may exhibit night-shift effects. An interesting analysis has been put forward by the Association of Professional Sleep Societies Committee on Catastrophes, Sleep and Public Policy. Its consensus report notes that the nuclear plant meltdown at Chernobyl occurred at 1:35 a.m. and was due to human error (apparently related to work scheduling). Similarly, the Three Mile Island reactor accident occurred between 4 and 6 a.m. and was due not only to the stuck valve, which caused a loss of coolant water, but also, more important, to the workers' failure to recognize this event, which led to the near meltdown of the reactor. Similar incidents (although with the ultimate stage having been pre-

vented) occurred in 1985 at the David Beese reactor in Ohio and at the Rancho Seco reactor in California. Finally, the committee also stated that the NASA Challenger space shuttle disaster stemmed from errors in judgment made in the early morning hours by people who had had insufficient sleep (because of partial night work) for days prior to the launch. Still, there is very limited support for the notion that shiftwork outside the area of transportation actually carries a higher over all accident risk.

As with sleep, the two main factors behind sleepiness and performance impairment are circadian cycle homeostatic factors. Alertness falls rapidly after awakening, but gradually levels out as wakefulness is extended. The circadian influence appears as a sine-shaped superimposition on this exponential fall in alertness. Akerstedt describes this as a three-process model of alertness regulation. Space does not permit a discussion of the derivation of these functions here.

Stress Indices

A number of studies have looked at the effect of shiftwork on endocrine and other stress-related parameters. The circadian pattern of cortisol, growth hormone, and melatonin is clearly affected by night work, but the adjustment to night work is only partial. We should also consider the increased cortisol levels and reduced insulin sensitivity following the reduction of sleep to 4 h. On the other hand, there do not seem to be any data on the long-term consequences of night work on stress indices, except perhaps for a suppression of testosterone in male nightworkers with a negative overall attitude to night shifts. Also, testosterone seems linearly related to sleep duration, as demonstrated in experimental studies of shifted sleep.

Gastrointestinal Effects

Gastrointestinal complaints are more common among night shiftworkers than among day workers. In a review of a number of reports covering 34 047 people with day or shiftwork, ulcers were found to occur in 0.3–0.7% of day workers, 5% of morning and afternoon shiftworkers, 2.5–15% of rotating shiftworkers with night shifts, and 10–30% of ex-shiftworkers. Several other studies have come to similar conclusions. Other gastrointestinal disorders, including gastritis, duodenitis, and dysfunctions of digestion are more common in shiftworkers than in day workers.

The pathophysiological mechanism underlying gastrointestinal diseases in shiftworkers is unclear, but one possible explanation is that intestinal enzymes

and intestinal mobility are not synchronized with the sleep–wake pattern. Intestinal enzymes are secreted with circadian rhythmicity, and shiftworkers' intake of food is irregular compared with intestinal function. A high nightly intake of food may be related to increased lipid levels, and eating at the circadian low point may be associated with altered metabolic responses. In addition, reduced sleep affects lipid and glucose metabolism. Recently, an epidemiological study showed a large increase in ulcer risk in shiftworkers with complaints of sleep or excessive sleepiness.

Cardiovascular Effects

A number of studies have reported a higher incidence of cardiovascular disease, especially coronary heart disease, in shiftworkers than in those who work days. In a study of 504 paper-mill workers followed for 15 years a dose–response relationship was found between number of years of shiftwork and incidence of coronary heart disease in the exposure interval 1–20 years of shiftwork. A study of 79 000 female nurses in the United States gave similar results, as did studies of more than 1 million Danish men and of a cohort of Finnish workers. Again, disturbances of metabolic parameters such as lipids and glucose might contribute to this. The effects of shiftwork on metabolic changes do not seem as pronounced, however, as those of stress. Very few studies are available, but the prevalence of endocrine and metabolic diseases seems increased in shiftworkers, including increased insulin resistance.

Mortality

The mortality of shift- and day workers was studied by Taylor and Pocock, who studied 8603 male manual worker in England and Wales between 1956 and 1968. Day, shift-, and ex-shiftworkers were compared with national figures. The standardized mortality rate (SMR) can be calculated from observed and expected deaths reported in the paper. SMRs for deaths from all causes were 97, 101, and 119 for day, shift-, and ex-shiftworkers, respectively. Although the figures might indicate an increasing trend, the differences were not statistically significant. However, the reported SMR close to 100 is remarkable because the reference population was the general male population. Most mortality studies concerned with occupational cohorts reveal SMRs lower than 100, implying a healthy worker's effect. The same study showed a significantly increased incidence of neoplastic disease in shiftworkers (SMR 116).

Conclusion

Shiftwork (with night shifts) impairs sleep and alertness and increases the risk of cardiovascular and gastrointestinal disease. It also disturbs the circadian timing of the endocrine and other systems. Thus, it may serve as one form of stressor and may combine with traditional psychosocial stressors.

See Also the Following Articles

Sleep, Sleep Disorders, and Stress; Workplace Stress.

Further Reading

- Åkerstedt, T. (2003). Shift work and disturbed sleep/wakefulness. *Occupational Medicine* 53, 89–94.
- Bjerner, B., Holm, Å. and Swensson, Å. (1955). Diurnal variation of mental performance: a study of three-shift workers. *British Journal of Industrial Medicine* 12, 103–110.
- Dijk, D.-J. and Czeisler, C. A. (1995). Contribution of the circadian pacemaker and the sleep homeostat to sleep propensity, sleep structure, electroencephalographic slow waves, and sleep spindle activity in humans. *Journal of Neuroscience* 15, 3526–3538.
- Drake, C. L., Roehrs, T., Richardson, G., et al. (2004). Shift work sleep disorder: prevalence and consequences beyond that of symptomatic day workers. *Sleep* 27, 1453–1462.
- Folkard, S. and Tucker, P. (2003). Shift work, safety and productivity. *Occupational and Environmental Medicine* 53, 95–101.
- Harrington, J. M. (ed.) (1978). *Shift work and health: a critical review of the literature*. London: HMSO.
- Hastings, M. H., Reddy, A. B. and Maywood, E. S. (2003). A clockwork web: circadian timing in brain and periphery, in health and disease. *Nature Reviews Neuroscience* 4, 649–661.
- Knutsson, A. (2003). Health disorders of shift workers. *Occupational Medicine* 53, 103–108.
- Mitler, M. M., Carskadon, M. A., Czeisler, C. A., et al. (1988). Catastrophes, sleep and public policy: consensus report. *Sleep* 11, 100–109.
- Spiegel, K., Leproult, R. and Van Cauter, E. (1999). Impact of sleep debt on metabolic and endocrine function. *Lancet* 354, 1435–1439.
- Taylor, P. J. and Pocock, S. J. (1972). Mortality of shift and day workers 1956–68. *British Journal of Industrial Medicine* 29, 201–207.
- Tüchsen, F. (1993). Working hours and ischaemic heart disease in Danish men: a 4-year cohort study of hospitalization. *International Journal of Epidemiology* 22, 215–221.
- Van Dongen, H. P., Maislin, G., Mullington, J. M., et al. (2003). The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep* 26, 117–126.

Nightmares

M Hirshkowitz and A Sharafkhaneh

Baylor College of Medicine and Michael E. DeBakey
Veteran Affairs Medical Center, Houston, TX, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
M Hirshkowitz and C A Moore, volume 3, pp 49–52,
© 2000, Elsevier Inc.

Phenomenology of Frightening Awakenings
Nightmares and Posttraumatic Stress Disorder
Treatment Approaches

Glossary

Hypnagogic and hypnopompic hallucinations Vivid sensory images occurring at the transition between sleep and wakefulness. Hypnagogic refers to hallucinations occurring as a person falls asleep

while hypnopompic images occur as a person awakens. Hypnagogic imagery may be particularly vivid during rapid eye movement (REM) sleep occurring at sleep onset (as in narcolepsy). Similarly, dreamlike hypnopompic images may continue during wakefulness upon sudden or partial awakening from REM sleep. Hypnagogic and hypnopompic hallucinations may be accompanied by a complete inability to move (sleep paralysis).

Nachtmare

Nacht (night) + mare (spirit or monster). Mythologies throughout the world describe devils, demons, evil spirits, ogres, ghosts, vampires, old hags, and witches that attack individuals while they sleep. The *mare* usually sits on or crushes the sleeper's chest causing a smothering sensation and an inability to breathe. Additionally, the individual is usually unable

	to move. Sometimes there are sexual overtones as well with the spirit raping or assaulting the person while they lay paralyzed. This description suggests the occurrence of the parasomnia sleep paralysis accompanied by a hypnagogic or hypnopompic hallucination.
<i>Night terror (also known as sleep terror)</i>	A sleep disorder usually arising from slow-wave sleep that involves a sudden awakening with intense fear. The afflicted individual usually emits a sharp piercing scream. Autonomic nervous system activation accompanies the awakening and the individual will often have tachycardia, tachypnea, sweating, and increased muscle tone. The individual typically acts confused, disoriented, and inconsolable during the episode; however, there may be complete amnesia about the episode upon awakening the next morning.
<i>Parasomnia</i>	A disorder of arousal, partial arousal, or sleep stage transition. Parasomnias occur episodically during sleep and may interfere with sleep, sometimes producing insomnia or daytime sleepiness.
<i>Rapid eye movement (REM) sleep</i>	REM sleep is one of the five sleep stages defined in humans. REM sleep was discovered by Eugene Aserinsky who noted eye movements occurring every 90–100 minutes during sleep. In initial studies, dream recall was better correlated with awakenings from REM than other sleep stages. Consequently, REM sleep came to be considered the biological substrate of dreams.
<i>Sleep paralysis</i>	The inability to perform voluntary movements upon awakening from sleep or just before sleep onset. The individual is fully conscious, usually feels acutely anxious, typically is unable to vocalize, and may have hallucinatory activity accompanying the paralysis episode.

Nightmares, according to current terminology, are unpleasant, distressing, or frightening dreams that usually occur during rapid eye movement (REM) sleep. The International Classification of Sleep Disorders—Second Edition (ICSD-2) formally classifies the nightmare disorder as a parasomnia usually associated with REM sleep. Nightmares, as currently defined are “coherent dream sequences that seem real and become increasingly more disturbing as they unfold. Emotions usually involve anxiety, fear, or terror but frequently also anger, rage, embarrassment, disgust, and other negative feelings. Dream content most often focuses on imminent physical

danger to the individual but can also involve distressing themes.” The fears produced by the dream usually dissipate rapidly after awakening and clear sensorium returns.

Phenomenology of Frightening Awakenings

Although nightmares are currently conceptualized as frightening dreams that usually produce an awakening from REM sleep, the term derives from *nachtmare* which refers to a quite different phenomenon. The *nachtmare* (or night spirit) has taken many forms in folk lore and superstition, ranging from vampires (in Medieval Eastern Europe) to witch riders (among preAmerican Civil War slaves). Contemporary alien attacks are likely an offshoot of the same disorder. The *nachtmare* accompanies sleep paralysis with frightening hypnagogic or hypnopompic hallucinations and characteristically arise during the transition from wakefulness to sleep. The sleeper has difficulty breathing due to the creature sitting on his or her chest, there may also be a sense of sexual molestation, the sleeper may feel that their life force is being sucked out of them. To complicate things further, the mythical *nachtmares* Incubus and Succubus are commonly associated with yet as distinctly different phenomenon: the sleep terrors (*Pavor nocturnus*). Sleep terrors usually occur in slow-wave sleep, are not dreamlike, and are poorly recalled. The person may remember a frightening image; however, narrative and plot are rare.

By contrast, the nightmare in modern parlance explicitly arises from a dream, implying imagery, narrative, and plot. Individuals awakened from a dream typically recollect its content. Nightmares may afflict 10–50% of children and 50–85% of adults have at least an occasional nightmare. Overall estimates for the prevalence of nightmares range from 2–8% in the general population. Nightmares arising from either acute or posttraumatic stress disorders (PTSDs) may occur in stage 2 sleep as well as during REM sleep. However, there are other widely recognized types of awakenings associated with subjective reports of panic, dysphoria, or anxiety. These included sleep-related panic attacks, sleep terrors, and sleep paralysis with hypnagogic or hypnopompic hallucinations. These paroxysmal events arise from different stages of sleep. Like REM sleep nightmares, sleep-related panic attacks are characterized by awakening with fear. Polysomnography reveals that these episodes customarily occur during sleep stage 2, often during transition to slow-wave sleep. However, sleep-related panic attacks may occur, although rarely, at the wake-to-sleep transition.

Nightmares and Posttraumatic Stress Disorder

Nightmare frequency increases during periods of stress and after an individual experiences a distressing event (e.g., an automobile accident, burglary, rape, or death of a friend or relative). Stress and sleep deprivation are also thought to exacerbate sleep paralysis, sleep terrors, and sleep-related panic attacks. One approach to exploring the relationship between stress and these phenomena is to study the sleep in individuals with PTSD.

Silva and colleagues reported on symptoms in 131 physically or sexually abused women interviewed in a primary care setting. They found that 65% reported bad dreams, flashbacks, and/or terror attacks. These women had higher intrusion and avoidance scores. Nightmares are also reported in school age children after being involved as passengers in traffic accidents. Psychological consequences were common and persisted in 33% of children after 4–7 months. Eleven percent were thought to be severely afflicted on follow up with 17% reporting nightmares and other sleep difficulties. Nightmares related to trauma are known to persist for prolonged periods. A study of Dutch war veterans found nightmares and anxiety dreams 40 years after the traumatizing event. In a large survey of Vietnam War veterans, using a careful sampling technique, nightmare frequency strongly correlated with combat exposure. Insomnia also was associated with combat exposure but not to the same degree as nightmares. By contrast, neither chronic medical illnesses nor psychiatric disorders (panic disorder, major depression, mania, and alcohol abuse) could predict nightmare frequency.

Although recurrent distressing dreams are a key PTSD diagnostic feature according to DSM-IV, the nature of these nightmares is in question. Many clinicians assume that PTSD nightmares represent a REM sleep phenomenon; however, PTSD nightmares are poorly understood. In many respects their features conform to those present in REM sleep nightmares. They typically contain imagery, narrative, and plot. However, the repetitive, recurring nature of these nocturnal frightening phantasms makes them unusual. Another uncommon characteristic of chronic PTSD nightmares is that they narrate traumatic events, often without elaboration or symbolism. Several published descriptions of PTSD nightmares include the adrenergic concomitants tachycardia and sweating that are atypical of REM sleep. These features more commonly accompany sleep terrors. Thus, it is not surprising that some sleep experts contend that PTSD nightmares can arise from any stage of sleep. Hartmann in his extensive monograph on

nightmares flatly states that PTSD nightmares "... are not typical REM nightmares; they occur sometimes in stage 2 and sometimes in all stages of sleep and even at sleep onset..." Kramer and Kinney also emphasize that PTSD nightmares can arise from both REM and stage 2 sleep.

Ross and associates argue strongly that REM sleep disturbance is closely involved in the pathogenesis of PTSD. This disturbance may be either "inappropriate recruitment of essentially normal REM sleep processes" or involvement of "inherently dysfunctional REM sleep mechanisms". They correctly contend that the vividness and recallability of PTSD nightmare content has more in common with REM than non-REM mentation. REM sleep is comprised of an array of individual features, including: low amplitude, mixed frequency electroencephalograph (EEG) activity; rapid eye movements; muscle atonia; pontine geniculate occipital (PGO) discharges; small muscle twitches; periorbital integrated potentials (PIPs); and middle ear muscle activity (MEMA). Although these activities usually occur in concert, they sometimes dissociate, as in REM sleep behavior disorder (RBD). In RBD, the atonia that characteristically accompanies REM sleep is absent. In fact, Ross and colleagues suggest RBD as a possible model for the pathophysiological mechanism underlying PTSD nightmares. In their overnight polysomnographic studies, these researchers showed that combat veterans with PTSD had elevated REM sleep percentage, longer average REM cycle duration, and increased REM density (number of eye movements per minute of REM sleep) compared to controls. These same investigators also report more phasic leg muscle activity during REM sleep in veterans with PTSD than in those without PTSD. Admittedly, the sample was small; however, these studies provide a foundation for further research.

Neuropharmacology of REM sleep clearly demonstrates involvement of cholinergic mechanisms. It is also widely accepted that acetylcholine plays an important role in the formation of memories. It has also been postulated that REM sleep mediates memory consolidation. When high adrenergic arousal accompanies a remembered experience (as occurs with acute trauma or re-experiencing a distant trauma), it appears that the memory process functions differently. Memories can be triggered by adrenergic arousal. The relationship between this arousal, flashbacks, and nightmares underlie current neuropsychiatric conceptualization of the consequences of traumatic stress. REM sleep has been characterized as a parasympathetic state with bursts of sympathetic activity (associated with phasic events). Nightmares and

recurrent dreams may be considered a failure of dreamwork to handle emotionally laden, presumably adrenergically coupled experiences.

PTSD nightmares are reported in 92% of patients with flashbacks and 57% of patients without flashbacks, they seldom occur during laboratory sleep studies. This is the case even among patients who indicate having nightmares four or more times per week. Nonetheless, on several occasions during our 30 years of clinical polysomnography experience, patients had nightmares while sleeping in the laboratory. On those rare occasions, we noticed increased REM density just before awakenings associated with nightmares. Consultation with several sleep disorder medicine colleagues revealed that this is a widely held impression (e.g., Drs. Hartmann (Boston, MA), Roffwarg (Jackson, MS), Mendelson (Chicago, IL), and Mahowald (Minneapolis, MN)). To the best of our collective knowledge and according to thorough search using medical databases, little systematic data exists beyond description of the phenomenon. Ross and associates mention exceedingly high REM density preceding a nightmare that fortuitously occurred during their study. We also have observed vastly increased phasic activity in combat veterans with PTSD referred to our Sleep Diagnostic Clinic for evaluation of insomnia, nightmares, or daytime sleepiness.

A phasic activity threshold likely exists that is necessary but not sufficient to produce nightmares. The level of phasic activity is elevated in patients with PTSD compared to control subjects. This conceptualization is analogous to the relationship between hypersynchronous slow-wave activity and parasomnias of arousal (confused awakening, sleepwalking, and sleep terror). Another perspective is suggested by the well-established relationship between sleep and some forms of epilepsy. Abnormal spike activity is enhanced by sleep deprivation. Interestingly, some neurophysiologically oriented researchers have posited a kindling model for flashbacks and nightmares. More research is needed to illuminate the biological substrate underlying the dysphoric mentation and paroxysmal awakenings that we call nightmares.

Treatment Approaches

Nightmares are experienced acutely (at some period during the lifetime) in many individuals. Furthermore, they typically resolve without intervention. However, persistent nightmares, particularly if the dream content is recurrent, can be a marker for significant psychopathology. In such cases, treatment is advised especially when the nightmare produces serious distress, insomnia, or interferes with daily

activities. Similarly, nightmares in early childhood (preschool and early school years) are of less concern (assuming they are outgrown) than those beginning or persisting in late childhood and adolescence. In adults, the frequency of psychopathology is higher among patients suffering from chronic nightmares than in control subjects. Hartmann posits that nightmare sufferers have thin boundaries, are more vulnerable, and have weaker psychological defenses. He also suggests that these individuals are at risk for schizophrenia. In addition to this profile for patients with nightmares, traumatic events apparently can induce nightmares. The nightmares may be immediate or delayed. They may afflict the individual acutely or persist for many years. Finally, drug abuse, drug or alcohol withdrawal, and specific medications are associated with nightmares. Thus, treatment must start with a careful history emphasizing current stressors, possible traumatic experiences, psychopathology (in patient and family), alcohol and drug abuse, and medication regimen.

Several medications have been associated with increased chance of nightmares. Some of them act by potentiating REM sleep and others by producing REM sleep rebound on withdrawal. Such drugs include L-DOPA, reserpine, thiothixene, alpha methyl dopa, metoprolol, propranolol, simvastatin, and atorvastatin. Interestingly, in addition to several statins appearing to induce nightmares, there is also a report by Ararun and associates (2005: 361–364) that nightmares are associated with low serum lipid levels.

Altering dose or substituting other pharmacological agents would be the first step in such cases. Alcohol or stimulant withdrawal also potentiates nightmares via presumed REM sleep mechanisms. In such cases, nightmares should be time-limited and resolve when sleep normalizes. If nightmares persist, intervention should be considered.

When psychopathology is found, it should be treated directly. Psychotherapy is thought to be beneficial. It is worth emphasizing that insomnia resulting from fear of sleep can pose a serious treatment obstacle. Sleep deprivation exacerbates nightmares, sleep terrors, and sleep paralysis attacks. Thus, the nightmare may perpetuate insomnia and thereby produce a vicious cycle. Behavioral interventions for insomnia are safe and effective in both the short and long term. Improved sleep hygiene, stimulus control therapy, sleep restriction therapy, and/or cognitive therapy are recommended. Krakow and Zadra (2006: 45–70) directly target the nightmare with imagery rehearsal therapy (IRT) with good results using four, 2-hour sessions. Another reportedly successful behavioral approach is lucid dream therapy. A lucid dream is one in which the dreamer becomes aware that he or

she is dreaming. In an uncontrolled case series of five individuals that could be trained to recognize their dream state, nightmare frequency declined. Whether the decline related to increased lucidity or volitional alteration of dream content is unknown. In some cases, medications may be needed. Several medications reportedly provide some relief from nightmares when administered to patients with PTSD. Davidson and colleagues report decreased nightmare and general sleep disturbance scores in response to the atypical antidepressant nefazodone in a group of 17 patients treated with up to 600 mg day⁻¹ for 12 weeks. Other antidepressants have been tried. Cavaljuga and co-workers (2003: 12–16) found fluoxetine more effective at reducing nightmares than amitriptyline; however, amitriptyline was more effective overall in acute PTSD. In another study by Gupta and co-workers, cyproheptadine (4–12 mg at bedtime) was found to be beneficial on retrospective review of nine patient charts. Response to this histamine-1 blocking, serotonin antagonist ranged from decreases in nightmare frequency (or intensity) to complete remission.

At present, the most systematic work on pharmacological treatment of nightmares is with prazosin. Peskind and colleagues (2003: 165–171) treated nine older men with posttraumatic nightmares with 2–4 mg prazosin 1 hour before bedtime. The drug substantially reduced nightmares in eight of the patients and was well tolerated. Furthermore, a 20-week, double-blind, crossover trial was conducted by Raskind and associates (2003: 371–373) in 10 Vietnam theatre combat veterans. A mean dose of 9.5 mg of prazosin was given at bedtime. Drug, compared to placebo, significantly reduced scores on recurrent distressing dreams.

In a review article concerning nightmare treatments, van Liempt (2006: 193–202) concludes that open-label reports suggest efficacy for some antidepressants, anticonvulsants, and atypical antipsychotic drugs. Placebo-controlled studies show possible efficacy with olanzapine and prazosin. We find these preliminary results are encouraging; however, more double-blind, placebo-controlled trials are needed. Finally, assorted benzodiazepines can provide relief from dream anxiety and help improve sleep. However, few randomized, placebo-controlled, clinical trials have been conducted and they had small sample sizes. Furthermore, the withdrawal effects from pharmacotherapeutic interventions have not been systematically studied.

Further Reading

Arargun, M. Y., Gulec, M., Cilli, A. S., et al. (2005). Nightmares and serum cholesterol level: a preliminary report. *Canadian Journal of Psychiatry* 50, 361–364.

- Bell, C. C., Shakoor, B., Thompson, B., et al. (1984). Prevalence of isolated sleep paralysis in black subjects. *Journal National Medical Association* 76, 501–508.
- Boriani, G., Biffi, M., Strocchi, E., et al. (2001). Nightmares and sleep disturbances with simvastatin and metoprolol. *Annals of Pharmacotherapy* 35, 1292.
- Cavaljuga, S., Licanin, I., Mulabegovic, N., et al. (2003). Therapeutic effects of two antidepressant agents in the treatment of posttraumatic stress disorder (PTSD). *Bosnian Journal of Basic Medical Sciences* 3, 12–16.
- Davidson, J. R., Weisler, R. H., Malik, M. L., et al. Treatment of posttraumatic stress disorder with nefazodone. *International Clinical Psychopharmacology* 13, 111–113.
- Diagnostic Classification Steering Committee. (1990). *International classification of sleep disorders – 2nd edition: diagnostic and coding manual*. Westchester, IL, USA: American Academy of Sleep Medicine.
- Ermin, M. K. (1987). Dream anxiety attacks (nightmares). *Psychiatric Clinics of North America* 10, 667–674.
- Firestone, M. (1985). The “old hag” sleep paralysis in Newfoundland. *Journal of Psychoanalytic Anthropology* 8, 47–66.
- Gupta, S., Popli, A., Bathurst, E., et al. (1998). Efficacy of cyproheptadine for nightmares associated with post-traumatic stress disorder. *Comprehensive Psychiatry* 39, 160–164.
- Hartmann, E. (1984). *The nightmare: the psychology and biology of terrifying dreams*. New York: Basic Books.
- Hauri, P. J., Friedman, M. and Ravaris, C. L. (1989). Sleep in patients with spontaneous panic attack. *Sleep* 12, 323–337.
- Kales, A., Soldatos, C., Caldwell, A., et al. (1980). Nightmares: clinical characteristics and personality patterns. *American Journal of Psychiatry* 137, 1197–1201.
- Krakow, B. and Zadra, A. (2006). Clinical management of chronic nightmares: imagery rehearsal therapy. *Behavioral Sleep Medicine* 4, 45–70.
- Kramer, M. and Kinney, L. (1988). Sleep patterns in trauma victims with disturbed dreaming. *Psychiatric Journal of the University of Ottawa* 13, 12–16.
- Morin, C. M., Colecchi, C., Stone, J., et al. (1999). Behavioral and pharmacological therapies for late-life insomnia: a randomized controlled trial. *JAMA* 17, 991–999.
- Neylan, T. C., Marmar, C. R., Metzler, T. J., et al. (1998). Sleep disturbance in the Vietnam generation: findings from a nationally representative sample of male Vietnam veterans. *American Journal of Psychiatry* 155, 929–933.
- Peskind, E. R., Bonner, L. T., Hoff, D. J., et al. (2003). Prazosin reduces trauma-related nightmares in older men with chronic posttraumatic stress disorder. *Journal of Geriatric Psychiatry and Neurology* 16, 165–171.
- Raskind, M. A., Peskind, E. R., Kanter, E. D., et al. (2003). Reduction of nightmares and other PTSD symptoms in combat veterans by prazosin: a placebo-controlled study. *American Journal of Psychiatry* 160, 371–373.
- Ross, R. J., Ball, W. A., Sullivan, K. A., et al. (1989). Sleep disturbance as the hallmark of PTSD. *American Journal of Psychiatry* 146, 697–707.

- Silva, C., McFarlane, J., Soeken, K., et al. (1997). Symptoms of post-traumatic stress disorder in abused women in a primary care setting. *Journal of Womens Health* **6**, 543–552.
- van Liempt, S., Vermetten, E., Geuze, E., et al. (2006). Pharmacotherapy for disordered sleep in post-traumatic

- stress disorder: a systematic review. *International Clinical Psychopharmacology* **21**, 193–202.
- Woodward, S. H., Arsenault, N. J., Murray, C., et al. (2000). Laboratory sleep correlates of nightmare complaint in PTSD inpatients. *Biological Psychiatry* **48**, 1081–1087.

Nitric Oxide

S M McCann

Louisiana State University, Baton Rouge, LA, USA

© 2007 Published by Elsevier Inc.

Role of Nitric Oxide in Physiological Release of Hypothalamic Neuropeptides and Transmitters
 Role of Nitric Oxide in Stress-Induced Alterations in Hypothalamic-Pituitary Function
 Effect of Infection on Cytokine and Nitric Oxide Formation in the Pineal Gland and in Other Organ Systems
 The Role of Nitric Oxide in Reproduction
 The Nitric Oxide Hypothesis of Aging
 Conclusion

Glossary

Adrenocorticotrophic hormone (ACTH) A polypeptide hormone secreted from the corticotropes of the anterior pituitary that stimulates the secretion of glucocorticoids (cortisol in human, corticosterone in rat) from the adrenal cortex.

Cytokine A polypeptide or small protein secreted classically by cells of the immune system, such as lymphocytes, monocytes, and macrophages, and now known to be secreted by other cells, such as endothelial cells and adipocytes. In many instances, cytokines have autocrine, paracrine, and endocrine effects to alter growth and function of their target cells.

Follicle-stimulating hormone (FSH) A heterodimeric glycoprotein secreted by gonadotropes of the anterior pituitary that promotes follicular growth in the ovaries and spermatogenesis in the testes and interacts with the luteinizing hormone to produce estrogen secretion by the ovarian follicles.

Growth hormone (GH) A protein hormone secreted by the somatotropes of the anterior pituitary that is responsible for growth by anabolic

Interferon γ

Interleukin (IL)

Luteinizing hormone (LH)

Prolactin (PRL)

Thymosin α_1

Thyroid-stimulating hormone (TSH)

effects in many tissues often mediated by insulin-like growth factor I, secreted principally by the liver.

One of several interferons that particularly has antiviral activity in contrast to the antibacterial activity of many of the interleukins.

A group of cytokines. IL-1 is a pro-inflammatory cytokine and was the first cytokine to be characterized; two closely related proteins, IL-1 α and β , have similar actions on closely related receptors in many tissues. Subsequently characterized cytokines were named in the order of their discovery. IL-2 is a pro-inflammatory cytokine. IL-10, IL-13, and IL-1 receptor antagonist are anti-inflammatory cytokines.

A heterodimeric glycoprotein secreted by gonadotropes of the anterior pituitary that induces ovulation and corpus luteum formation, estradiol secretion from the follicle, and progesterone secretion from the corpus luteum of female mammals and stimulates the interstitial tissue of the testis to secrete testosterone in males.

A protein hormone, related to growth hormone, secreted by the lactotropes of the anterior pituitary. Its best known function is to stimulate the growth and secretion of milk from the mammary gland. In males, it affects male sex accessories. It also has actions on the cells of the immune system that enhance immune responses.

A polypeptide cytokine secreted by the thymus that has paracrine and hormonal actions.

A heterodimeric glycoprotein secreted by the thyrotrope of the anterior pituitary that stimulates the secretion of thyroid hormones.

Tumor necrosis factor α A pro-inflammatory cytokine mediating local inflammatory and systemic responses to infection. It was named according to its originally described action; also known as cachectin.

Stress may be defined as any stimulus that distorts homeostasis in the organism. Therefore, stress can be physical, as in the case of excessive heat or cold, trauma, surgical operations, or infection, or it can be emotional and psychological. Selye, who coined the term, even talked about the stress of life itself. It is the stresses of life that contribute to degenerative disease and old age.

Stress can affect every organ system in the body, but the principal mediators of the response to stress are the sympathetic nervous system and the hypothalamic-pituitary-adrenal axis. Various hypothalamic-releasing and -inhibiting hormones are released from their axon terminals in the median eminence (ME) into the capillaries of the hypophyseal portal veins that transport them to the anterior lobe of the pituitary gland. There they stimulate or inhibit the release of the various pituitary hormones that act either directly on body tissues or on their target glands, which in turn release hormones that act on various tissues of the body. Examples of the latter are adrenocorticotrophic hormone (ACTH), follicle-stimulating hormone (FSH), luteinizing hormone (LH), and thyroid-stimulating hormone (TSH); examples of the former are prolactin (PRL) and growth hormone (GH).

In the dozen years since the discovery of nitric oxide (NO) in the body and the enzymes forming it, it has become evident that NO plays a pivotal role in the function of every organ system of the body. This article first examines the role of NO under physiological conditions and then indicates how this role is affected by stress. In this connection, most attention has focused on the role of NO in the induction of the responses to infection. Little attention has been given so far to its role in psychological and other types of stress.

Role of Nitric Oxide in Physiological Release of Hypothalamic Neuropeptides and Transmitters

Corticotropin Releasing Hormone

The release of corticotropin releasing hormone (CRH) from hypothalami incubated *in vitro* is controlled by muscarinic cholinergic receptors because it can be blocked by atropine. The acetylcholine-producing interneurons in the hypothalamus release

acetylcholine, which stimulates a muscarinic-type receptor, which in turn stimulates CRH release from the CRH neurons. Neural NO synthase (nNOS) has been located in neurons in the paraventricular nucleus (PVN) of the hypothalamus. Stimulated CRH release can be blocked by N^G-monomethyl-L-arginine (NMMA), a competitive inhibitor of all forms of NOS. Consequently, CRH release from the neurons in the PVN is stimulated by cholinergic neurons that synapse on these NOergic neurons to activate NOS. NOS synthesizes NO, which diffuses into the CRH neurons and activates CRH release by activating cyclooxygenase I (COX I), leading to the generation of prostaglandin E₂ (PGE₂) from arachidonate (AA). PGE₂ activates CRH via the activation of adenyl cyclase and the generation of cyclic adenosine monophosphate (cAMP). cAMP activates protein kinase A, which induces the exocytosis of CRH secretory granules into the hypophyseal portal vessels that then activate ACTH release from the corticotrophs of the anterior pituitary gland. NO activates not only COX, but also lipoxygenase (LOX), which also plays a role in the activation of CRH release. NO also activates guanylyl cyclase (GC), which converts guanosine triphosphate into cyclic guanosine monophosphate (cGMP). cGMP is postulated to increase intracellular [Ca²⁺], which is required to activate phospholipase A₂, which converts membrane phospholipids into AA, the substrate for COX and LOX, permitting the generation of prostaglandins (PGs) and leukotrienes, respectively. The activation of CRH release can be blocked by the synthetic glucocorticoid dexamethasone and also by blockers of the three pathways of AA metabolism (e.g., clotrimazol, which blocks epoxygenase, which converts AA into epoxides, and indomethacin, which inhibits COX) and by 5' 8' 11-eicisotrienoic acid, which blocks LOX. Thus, CRH release is activated by the AA cascade. α -Melanotropic-stimulating hormone (α -MSH) also inhibits CRH release. Cyclosporin inhibits CRH release as well, probably by inhibiting calcineurin. Calcineurin dephosphorylates NOS, rendering it active.

Luteinizing Hormone-Releasing Hormone

The role of NO in luteinizing hormone-releasing hormone (LHRH) release has been elucidated. Norepinephrine (NE), glutamic acid (GA), and oxytocin stimulate LHRH release by activation of nNOS. The pathway is as follows. Oxytocin and/or GA activate NE neurons in the medial basal hypothalamus (MBH), which activate NOergic neurons by α_1 -adrenergic receptors. The NO released diffuses into LHRH terminals and induces LHRH release by the activation of GC and COX. NO controls not only the release of

LHRH bound for the pituitary, but also that which induces mating activity by actions in the brain stem. The adipocyte hormone, leptin, also increases the release of LHRH from medial basal hypothalamic explants by stimulating the release of NO. However, β -endorphin and γ -aminobutyric acid (GABA) block the activation of LHRH release by NO. In contrast to the clear evidence for the stimulation of LHRH release by NO, it apparently has no effect on FSH release mediated by the FSH releasing factor.

Other Hypothalamic Peptides

Evidence from *in vivo* experiments suggests that NO also stimulates the release of GH releasing hormone (GHRH), and *in vitro* experiments clearly show that it also stimulates the release and synthesis of somatostatin. The release of oxytocin and vasopressin from the hypothalamus is suppressed by NO. NO stimulates the release of the PRL-inhibiting hormone, dopamine.

Effects of Nitric Oxide on the Release of Pituitary Hormones

nNOS is present in folliculostellate (FS) cells, a cell type of the anterior pituitary not previously thought to produce hormones but that is now known to produce several cytokines. nNOS has also been found in LH gonadotrophs. NO produced by this nNOS in the pituitary has important functions in altering pituitary hormone secretion. Intravenous injection of nitro-arginine methyl ester (NAME) augments vasopressin-induced ACTH release, so NO probably blocks ACTH secretion induced by vasopressin or CRH, but this has not been studied directly. NO inhibits PRL secretion. Apparently, NO from FS cells diffuses to adjacent lactotrophs where it induces the production of cGMP by activating GC. cGMP then inhibits PRL release by blocking the exocytosis of PRL secretory granules.

The principal PRL-inhibiting hormone, dopamine, apparently acts on cell surface D2 receptors on FS cells to activate NOS followed by NO production. However, the action of FSH-releasing factor (FSHRF), LHRH, and leptin to stimulate not only FSH but also LH is brought about by their action on specific receptors on gonadotrophs to activate nNOS, thereby generating cGMP, which, instead of inhibiting, as in the case of lactotropes, promotes hormone secretion. Specificity of action of these compounds on gonadotropin release is accomplished by specific receptors for each of these ligands that increase intracellular $[Ca^{2+}]$, thereby activating nNOS. Both LHRH and leptin preferentially stimulate LH with a lesser potency to release FSH, whereas FSHRF selectively releases FSH and has a lower potency to release

LH. In this way, it is possible to have differential effects of various stimuli on FSH and LH secretion. This mechanism accounts for the pulses of FSH alone generated by FSHRF and of LH alone generated by LHRH.

Role of Nitric Oxide in Stress-Induced Alterations in Hypothalamic-Pituitary Function

Nearly all types of stress elicit a stereotyped pattern of pituitary hormone secretion in humans that consists of a rapid increase in ACTH, PRL, and GH secretion (in rats, GH secretion is inhibited) and inhibition of the secretion of LH, TSH, and, to a lesser extent, FSH. The pattern is caused primarily by alterations in the secretion of hypothalamic-releasing and -inhibiting hormones.

Immediate Response to Infection

The role of NO has been studied extensively in a model of the stress from infections that consists of injecting bacterial lipopolysaccharide (LPS) intravenously (IV) or intraperitoneally (IP). This induces an identical pattern of pituitary hormone secretion as that seen in infection. There is a very rapid increase in plasma ACTH and PRL within a few minutes following the IV injection of LPS. The response is dose-related and is accompanied by a rapid inhibition of LH and TSH (but not FSH) secretion. GH secretion is stimulated in humans but suppressed in rats. This initial response to LPS is mediated by the constitutive nNOS present in the brain. There is no participation of the NO synthesized by inducible NOS (iNOS) in this initial response. Indeed, the initial response must be due to action on receptors for LPS on the endings of vagal afferents and also in areas where the blood-brain barrier is not present, such as the choroid plexus, ME, organum vasculosum lamina terminalis (OVLT), area postrema, and other circumventricular organs. LPS alters the release of the various hypophysiotropic hormones, increasing the release of CRH and vasopressin, GHRH, and somatostatin, particularly in the rat, which overpowers the effect on the release of GHRH. LPS decreases the release of dopamine, the PRL-inhibiting hormone; LHRH; and TRH (but not FSHRF). All of these pathways involve the participation of nNOS as described earlier. Input to the hypothalamus from LPS by vagal afferents occurs, at least in part, by activation of the locus ceruleus, which sends noradrenergic axons to the hypothalamus to activate CRH release.

In the authors' studies, LPS itself had no acute effect on ACTH release from hemianterior pituitaries *in vitro*. However, LPS induces cytokine production

in the pituitary. Cytokine production would be increased in a few hours and undoubtedly would modify the responses of the pituitary to the continued altered secretion of releasing and inhibiting hormones.

Delayed Response

Central nervous system (CNS) infection is a powerful inducer of cytokine production in glia and neurons of the brain, which causes the induction of iNOS and the production of potentially toxic quantities of NO. Following IV injection of an intermediate dose of LPS, there was an induction of IL-1 α immunoreactive neurons in the preoptic-hypothalamic region. These cells were shown to be neurons by the fact that double staining revealed the presence of neuron-specific enolase. The neurons were found in saline-injected control animals, suggesting that they are normally present, but they increased in number by a factor of two within 2 h following injection of LPS. They are located in a region that also contains thermosensitive neurons. They may be the neurons that are stimulated to induce fever following the injection of LPS. They have short axons that did not project clearly to the areas containing the various hypothalamic-releasing and -inhibiting hormones, but they could also be involved in the stimulation or inhibition of their release, which occurs following infection.

This study led to further research that demonstrated that the IP injection of a moderate dose of LPS induced IL-1 β and iNOS mRNA in the brain, anterior pituitary, and pineal glands. The results were very exciting because an induction of IL-1 β and iNOS mRNA occurred with the same time course as found in the periphery following the injection of LPS, namely, clear induction of IL-1 β followed by iNOS mRNA within 2 h, reaching a peak in 4–6 h, followed by a decline to near basal levels at the next measurement by 24 h after the injection. The induction of both mRNAs occurred in the meninges; choroid plexus; circumventricular organs, such as the subfornical organ and ME; ependymal cells lining the ventricular system; and, very surprisingly, parvocellular neurons of the PVN and arcuate nucleus (AN), areas of particular interest because they contain hypothalamic-releasing and -inhibiting hormone-producing neurons and also other neurotransmitters controlled by NO. The greatest induction occurred in the anterior lobe of the pituitary, where the iNOS mRNA was increased at 2 h by a factor of 45 and in the pineal where the activity was increased at 6 h by a factor of 7, whereas the increase in the PVN was fivefold. At 6 h, the medial basal hypothalamus was found to have an increased content of NOS measured *in vitro* and the collected cerebrospinal fluid (CSF) had

increased concentrations of the NO metabolite nitrate. These results indicate that the increase in iNOS mRNA was followed by the *de novo* synthesis of iNOS that liberated NO into the tissue and also into the CSF. Presumably, LPS was bound to its receptors in the circumventricular organs and in the choroid plexus. These receptors, as in macrophages, activated DNA-directed IL-1 β mRNA synthesis, which, in turn, caused the synthesis of IL-1 β . IL-1 β then activated iNOS mRNA and synthesis.

How can neurons in the AN and PVN be activated inside the blood–brain barrier? In the case of the AN, the neurons may have axons that project to the ME. These neurons may have LPS receptors on their cell surface, which then induce IL- β mRNA and IL-1 β synthesis. IL-1 β then induces iNOS mRNA followed by NO synthesis. Alternatively, LPS, acting on its receptors, may induce IL- β mRNA and iNOS mRNA simultaneously.

Active transport mechanisms for IL-1 and other cytokines, and perhaps LPS, are present in the choroid plexus. On the basis of the authors' results, cells of the choroid plexus must have LPS receptors on them. LPS then stimulates IL-1 β and iNOS mRNA, followed by the synthesis of IL-1 β and iNOS in the choroid plexus. LPS and IL-1 β are then transported into the CSF. LPS is carried by CSF flow to the third ventricle, where it either crosses the ependyma or acts on terminals of PVN neurons in the ependyma to induce IL-1 β and iNOS mRNA. This massive delayed increased NO production should further increase the effects of NO to maintain the pattern of hypothalamic hormone secretion already induced by LPS. Unfortunately, the effect of inhibitors of NOS on these later stages in the response to LPS or infection has not yet been studied. Interestingly enough, in studies on LHRH release induced by NO, it has been shown that increasing concentrations of NO provided by the release from sodium nitroprusside (NP) produce a bell-shaped dose–response curve in terms of LHRH release, with values reaching a peak and then declining as the concentration of NO increases. Therefore, the massive increase in NO produced by iNOS, several hours after injection of LPS, might actually reduce the effects of NO on releasing hormone discharge below the peaks achieved earlier.

In addition to inducing the production of pro-inflammatory cytokines such as IL-1, IL-2, IL-6, and TNF- α , LPS also induces the production of anti-inflammatory cytokines such as IL-10 and L-13 and the IL-1 receptor antagonist in the brain, pituitary, and pineal gland. In the periphery, these inhibit the inflammatory response induced by the pro-inflammatory cytokines. Limited studies indicate

that these anti-inflammatory cytokines antagonize the actions of the pro-inflammatory cytokines in the brain as well as the hypothalamic-pituitary response to infection.

Direct Actions of Lipopolysaccharide, Cytokines, and Nitric Oxide on the Anterior Pituitary Gland during Infection

As indicated earlier, LPS had no acute direct effect to alter ACTH release from incubated pituitaries. Thus, the initial LPS-induced changes in pituitary hormone secretion are mediated by the altered release of hypothysiotropic hormones; however, LPS induces a massive induction of cytokine and, consequently, iNOS production in the anterior pituitary. Therefore, after the first few hours these cytokines and the massive release of NO will certainly modify the response of the gland to the hypothalamic factors. Unfortunately, these later times after LPS have not been studied experimentally.

Interleukin 1 The effects of IL-1 β and several other cytokines released by LPS have been studied extensively *in vitro*. It was reported that IL-1 releases ACTH and other pituitary hormones from perfused pieces of pituitary glands. In the case of the other cytokines, the authors have found actions on the pituitary gland; however, it required a period of 1–2 h for these actions to become manifest.

Tumor necrosis factor α After at least 1 h of incubation, TNF- α stimulated the release of ACTH, GH, TSH, and minimally PRL from either overnight-cultured dispersed pituitary cells or hemipituitaries; however, the minimal effective dose (MED) for these actions of TNF- α was 100-fold greater with dispersed cells than with hemipituitaries. We speculate that this decreased sensitivity in the dispersed cell preparation is due either to the loss of receptors or to paracrine actions of various pituitary cells to augment the effects on other pituitary cells in the hemipituitary preparation. There was a bell-shaped dose–response curve of response of the pituitaries, and the MED of 10^{-12} M for the hemipituitaries is within the concentrations that might be encountered *in vivo* in infection. Consequently, the authors concluded that cachectin may play an important role in altering pituitary hormone release by direct actions on the gland in infection; however, because of the relatively slow onset of these actions, it is likely that the acute response to infection is brought about by alterations in the release of hypothalamic-releasing and -inhibiting hormones, which then affect the gland. During more chronic infection, the actions at the pituitary

modulate the effects on the release of the hypothalamic peptides.

Interleukin 6 The incubation of IL-6 with hemipituitaries *in vitro* for at least 2 h increased the release of ACTH and GH only into the culture medium at the single concentration of 10^{-13} M. Concentrations of 10^{-14} or 10^{-12} M were ineffective. Again, there was a bell-shaped dose–response curve with the abolition of the actions with higher doses. Using 4-day monolayer-cultured cells, Spangelo and co-workers found not only an IL-6-induced increase in GH release but also an increase in PRL release. These differences may be accounted for by the different preparations of IL-6 used or, perhaps more likely, by their use of the dispersed cell system. This system may result in up- or downregulation of receptors for IL-6 in the relative absence of endogenous ligand and loss of paracrine actions because of separation of the cells. Because IL-6 is clearly produced by the pituitary gland itself and the production is increased by LPS (see earlier discussion), it is quite likely that IL-6 has important direct actions on the pituitary to modify pituitary hormone release. These actions probably modulate the effects of the cytokine on the release of hypothalamic hormones mentioned previously.

Therefore, the effect of IL-6 released from the hypothalamus and pituitary and also reaching the gland after release in the periphery during infection is the result of its actions at both the hypothalamic and pituitary levels, which induces a stimulation of ACTH and GH, and a decrease in TSH release in acute situations *in vivo*.

The authors have not investigated chronic exposure to IL-6. It is quite likely that there might be other effects, perhaps similar to those that occur in infection, with chronic exposure to this IL. Indeed, the acute effects of most of the cytokines at hypothalamic and pituitary sites can be characterized over a short period of an hour or two; however, much more research should be done to characterize the effects of more prolonged exposure of both the hypothalamus and the pituitary gland to these substances.

Interleukin 2 IL-2 is the most potent agent known to act directly on the pituitary to alter pituitary hormone release. At a concentration of 10^{-15} M, it elevated PRL and TSH release and inhibited the release of FSH, LH, and GH from hemipituitaries *in vitro*. ACTH release was induced with a higher concentration (10^{-12} – 10^{-10} M). After reaching the minimal effective stimulatory or inhibitory concentration, the dose–response curve was flat, but at much higher concentrations (10^{-9} or 10^{-8} M), responses tended

to diminish to insignificant values. Again, a bell-shaped dose–response curve was obtained similar to those observed with other cytokines.

Interferon- γ The incubation of INF- γ with hemipituitaries *in vitro* revealed no effect on the release of PRL, TSH, and GH, and it revealed a stimulation of ACTH release only at the relatively high concentration of 10^{-8} M. Later studies revealed that, at a concentration of 10^{-12} M, INF- γ also decreased the release of GH from pituitaries incubated *in vitro*. At a much higher concentration (10^{-8} M), it potentiated GH release induced by growth releasing hormone (GRH) but had no direct effect by itself. Because this is a 10,000-fold increase in dose over that which produced the inhibition of release, the authors question whether it is of pathophysiological significance. This inhibitory action on GH release vanished with higher doses, again giving the bell-shaped dose–response curve indicated earlier.

Thymosin α_1 Thymosin α_1 evoked a dose-dependent release of TSH and ACTH, although there was no effect on the release of PRL and GH from hemipituitaries incubated *in vitro*. There was a remarkable ability of the peptide to stimulate LH in a dose-related manner at a dose as low as 10^{-12} M, whereas FSH release was unaltered. Because thymosin α_1 has been localized to both the hypothalamus and the pituitary, it may have physiological significance in neuroendocrine immunology. As in the case of the other cytokines, both hypothalamic and pituitary sites of action may be of importance for the induction of changes in pituitary hormone release.

Therefore, although LPS does not alter pituitary hormone secretion directly, it induces cytokine release that in turn stimulates iNOS synthesis. The NO generated, together with the direct actions of the cytokines, produces modifications in the pituitary response to the altered pattern of hypothalamic peptides reaching the gland. Unfortunately, again, there have been no studies of these later events in terms of pituitary hormone release. From what is known, the general pattern that occurs immediately after the injection of LPS is maintained. Further studies of these later times are essential.

Role of Nitric Oxide in the Pattern of Pituitary Hormone Secretion in Other Physical and Psychological Stresses

In contrast to the extensive research carried out on the effects of LPS to mimic the effects of infection, little attention has been paid to the role of NO in mediating the effects of psychological stresses on the hypothalamic-pituitary axis. It has been shown that

restraint stress increases the expression of nNOS, but not iNOS, with the same time course and in the same regions in which LPS induces iNOS. Furthermore, the injection of nitroarginine methyl ester (L-NAME), an inhibitor of NOS, blocked ACTH release induced by foot shock and other physical-emotional stresses.

Effect of Infection on Cytokine and Nitric Oxide Formation in the Pineal Gland and in Other Organ Systems

Space limitations limit this article mostly to the effect of NO on the hypothalamic-pituitary axis; however, as described earlier, LPS causes the induction of IL-1 β mRNA followed by iNOS mRNA in the pineal as well, and experiments indicate that a similar response occurs in the heart and kidney. It is probably a uniform response throughout all organ systems of the body. Under normal conditions, NO exercises an important role in practically every, if not every, organ system of the body. The massive increase in NO production following infection is already known to be responsible for the production of toxic shock, and its toxicity probably contributes to the malaise that occurs during infections.

The Role of Nitric Oxide in Reproduction

Because LHRH induces ovulation by releasing LH and because animals go into heat around the time of ovulation, it occurred to me that LHRH could induce mating behaviour. Indeed, this turns out to be the case. Mating behaviour in female rats is brought about by LHRH neural axons that extend to the brain stem and release NO, which activates descending pathways and induces the lordosis reflex. Efferents from the pelvic spinal cord activate the pelvic ganglia, which send axons that extend to the corpora cavernosa of the penis and release Ach, which, in turn, acts on the muscle that releases NO, which activates GC to generate cGMP, which dilates the corpora leading to erection. In the presence of stress or infection, these pathways are inactive because stress blocks LHRH release. Viagra (sildenafil) promotes penile erection by inhibiting the phosphodiesterase 5 in the corpora cavernosa, thus slowing the inactivation of cGMP and prolonging the erection. The role of NO in mating behaviour is established, but its action in the external genitalia of women is not certain, although we predict that it would induce clitoral erection to facilitate orgasm.

The Nitric Oxide Hypothesis of Aging

As indicated earlier, NO generated by endothelial NOS (eNOS) and nNOS plays a ubiquitous role in

the body in controlling the function of almost every, if not every, organ system. Bacterial and viral products, such as bacterial LPS, induce iNOS synthesis, which produces massive amounts of NO toxic to the invading viruses and bacteria but also to host cells by the inactivation of enzymes leading to cell death. The actions of all forms of NOS are mediated not only by the free-radical oxidant properties of this soluble gas but also by its activation of GC, leading to the production of cGMP, which mediates many of its physiological actions. In addition, NO activates COX and LOX, leading to the production of physiologically relevant quantities of PGE₂ and leukotrienes.

The authors' hypothesis is that recurrent infections over the life span play a significant role in producing aging changes in all systems of the body via the release of toxic quantities of NO. NO may be a major factor in the development of coronary heart disease (CHD). Considerable evidence has accrued indicating

a role for infections in the induction of CHD and, indeed, patients treated with a tetracycline derivative had 10 times less complications of CHD than their controls. Stress, inflammation, and infection have all been shown to cause the induction of NOS in rats, and it is likely that this triad of events is very important in the progression of coronary arteriosclerosis leading to coronary occlusion. The aging of the anterior pituitary and pineal, with the resultant decreased secretion of pituitary hormones and the pineal hormone melatonin, respectively, may be caused by NO.

Inside the blood-brain barrier, the induction of iNOS in the temperature-regulating centers by infections may cause the decreased febrile response in the aged by loss of thermosensitive neurons. iNOS induction in the PVN may cause the decreased nocturnal secretion of GH and PRL that occurs with age, and its induction in the AN may destroy LHRH neurons, thereby leading to a decreased release of gonadotropins. Recurrent infections may play a role in the aging

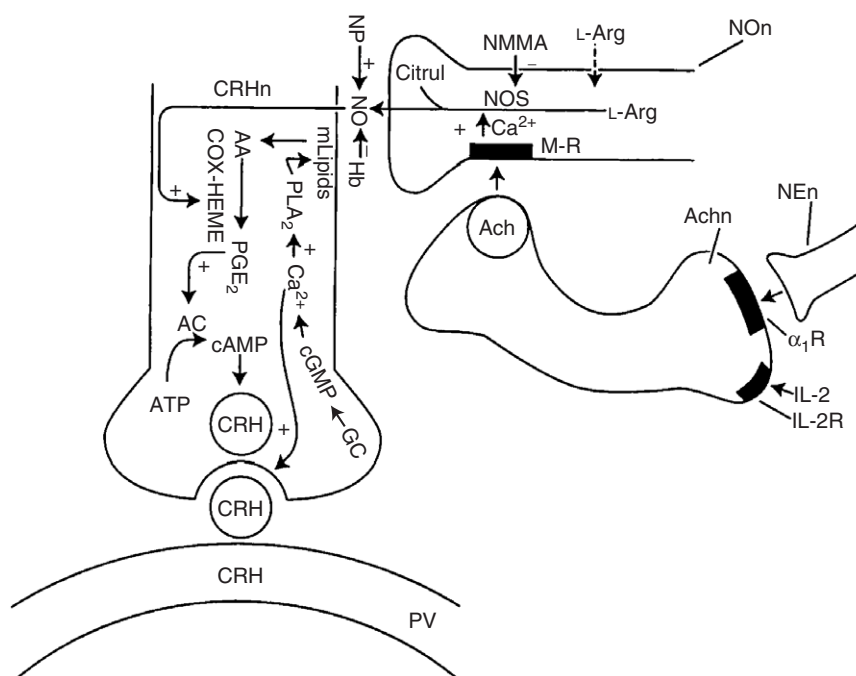


Figure 1 Diagram of the mechanism of action of cytokines such as IL-2 to release CRH by nitric oxide. In the case of systemic administration of inflammatory cytokines such as IL-2 or pain stimuli, afferent input occurs by vagal afferents that project to the locus ceruleus, a small brain-stem nucleus that contains many norepinephrine neurons (NEn) that project axons to the PVN, where they are visualized to act on α_1 receptors (α_1R) on acetylcholine interneurons (Achn) there. When IL-2 is in the circulation, it can also reach these neurons and act on its receptors there (IL-2R) to activate the Achn. This releases acetylcholine (ACh) from synaptic vesicle ACh from the cholinergic neuron, which acts on muscarinic receptors (M-R) on the nitric oxide neuron (NON), increasing intracellular free calcium, which combines with calmodulin to activate nitric oxide synthase (NOS). This converts L-arginine within the neuron, which is also pumped into the neuron by an active transport mechanism in the presence of oxygen and various cofactors into citrulline (Citrul) and nitric oxide (NO) that diffuses over to the CRH neuron to activate guanylyl cyclase (GC), which causes the production of cGMP, which increases intracellular free calcium. The increased intracellular calcium not only plays a role in exocytosis of CRH granules, but also activates phospholipase A₂, which converts membrane phospholipids to arachidonic acid (AA). NO activates cyclooxygenase by interacting with the heme group on the molecule, leading to the conversion of AA to prostaglandin E₂ (PGE₂), which activates adenylyl cyclase (AC), which generates cAMP from ATP, which also leads to the extrusion of CRH synaptic vesicles into the hypophyseal portal veins (PV), which carry it to the anterior pituitary to release ACTH. CRHn, corticotropin-releasing hormone (CRH) neuron; mLipids, membrane lipids; NMDA, N-methyl-D-aspartic acid; PLA₂, phospholipase A₂.

of other parts of the brain because there are increased numbers of astrocytes expressing IL-1 β throughout the brain in aged patients. IL-1 and products of NO activity accumulate around the plaques of Alzheimer's disease and may play a role in the progression of the disease. Flu encephalitis during World War I possibly induced iNOS in cells adjacent to substantia nigra dopaminergic neurons, leading to the death of these cells, which, coupled with ordinary aging fallout, led to early-onset Parkinsonism. The CNS pathology in acquired immunodeficiency syndrome (AIDS) patients bears a striking resemblance to aging changes and may also be largely caused by the action of iNOS. Antioxidants, such as melatonin, vitamin C, and vitamin E, probably play an important acute and chronic role in reducing or eliminating the oxidant damage produced by NO.

Conclusion

NO clearly plays a very important role in mediating the response to stress of every organ system in the body. This article concentrates on its actions mediating the response to stress and infection of the hypothalamic-pituitary axis. Afferent input (e.g., by LPS acting on receptors on vagal afferents) reaches the brain stem, where it activates the locus ceruleus, an important nucleus containing neuroadrenergic neurons that project to the hypothalamus (Figure 1). In the case of the CRH-ACTH-adrenocortical axis, the NE released in the vicinity of the PVN activates CRH release by stimulating cholinergic interneurons that release Ach, which in turn activates NOergic neurons by muscarinic receptors. The NO released then triggers CRH release. LPS also acts on receptors in circumventricular neurons, such as the ME and organum vasculosum lamina terminalis, to activate the NOergic control of CRH release. CRH passes down the hypophyseal portal vessels and activates ACTH release. CRH releases ACTH from the corticotropes of the anterior pituitary gland, which circulates to the adrenal cortex to stimulate the release of cortisol. Cortisol exerts negative feedback to suppress CRH release and decrease the pituitary response to the CRH released. Continued stimulation by stress, such as LPS, maintains the stimulatory

drive, maintaining the elevated ACTH and cortisol concentrations.

Acknowledgments

This work was supported by NIH Grants DK 43900 and MH51853. We thank Judy Scott and Jill Lucius for their excellent secretarial assistance.

See Also the Following Articles

Adrenocorticotrophic Hormone (ACTH); Aging and Stress, Biology of; Cytokines; Neuroimmunomodulation.

Further Reading

- Banks, W. A., Kastin, A. J., Huang, W., et al. (1996). Leptin enters the brain by a saturable system independent of insulin. *Peptides* 17, 305–311.
- McCann, S. M., Karanth, S., Kamat, A., et al. (1994). Induction by cytokines of the pattern of pituitary hormone secretion in infection. *Neuroimmunomodulation* 1, 2–13.
- McCann, S. M., Licinio, J., Wong, M.-L., et al. (1998). The nitric oxide hypothesis of aging. *Experimental Gerontology* 33, 813–826.
- McCann, S. M., Mastronardi, C., Walczewska, A., et al. (2003). The role of nitric oxide in control of LHRH release that mediates gonadotropin release and sexual behavior. *Current Pharmaceutical Design* 9, 381–390.
- Rettori, V., Dees, W. L., Hiney, J. K., et al. (1994). An interleukin-1-alpha-like neuronal system in the preoptic-hypothalamic region and its induction by bacterial lipopolysaccharide in concentrations which alter pituitary hormone release. *Neuroimmunomodulation* 1, 251–258.
- Rivier, C. (1998). Role of NO and carbon monoxide in modulating the ACTH response to immune and non-immune signals. *Neuroimmunomodulation* 5, 203–213.
- Wong, M.-L., Bongiorno, P. B., Rettori, V., et al. (1997). Interleukin (IL) 1 β , IL-1 receptor antagonist, IL-10, and IL-13 gene expression in the central nervous system and anterior pituitary during systemic inflammation: pathophysiological implications. *Proceedings of the National Academy of Sciences USA* 94, 227–232.
- Wong, M.-L., Rettori, V., Al-Shekhlee, A., et al. (1996). Inducible nitric oxide synthase gene expression in the brain during systemic inflammation. *Nature Medicine* 2, 581–584.

Northern Ireland, Post Traumatic Stress Disorder in

P Bell

South and East Belfast Health and Social Services
Trust, Belfast, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
P Bell, volume 3, pp 62–65, © 2000, Elsevier Inc.

Early Survey of Survivors of Violence
The Enniskillen Bombing
The Kegworth Air Disaster
Theoretical Model Used in Debriefing

Glossary

Posttraumatic stress disorder (PTSD) A psychological injury after violence, characterized by reexperiencing trauma, hyperarousal, and withdrawal.

Psychiatrists in Northern Ireland had been treating the survivors of violence for more than 10 years, since PTSD was defined in the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III; 3rd edn.). In general psychiatry, most of the patients encountered are suffering from a psychiatric illness. This results when an individual who is vulnerable is subjected to a stressor. In the 1970s, the concept of psychological injury was not well developed, and so people who suffered from psychological problems after exposure to violence were viewed according to the psychiatric illness model. Either these patients were assumed to have some constitutional vulnerability to anxiety or depression and that a neurotic illness had been precipitated by an adverse life event or else the psychological injuries were dismissed as a normal reaction. Previous attempts to define psychological injuries, such as shell shock and combat neurosis, had been discredited because these diagnoses carried with them some secondary gain. In the early 1980s, a group of psychiatrists based at the Mater Hospital decided to examine the diagnosis of PTSD in order to see whether it applied to the victims of violence in Northern Ireland. The diagnosis had considerable appeal because of the clear operational definition involving a pattern of symptoms. However, local psychiatrists were skeptical about the diagnosis for three main reasons. First, community surveys in Northern Ireland had indicated that the number of cases of psychiatric morbidity in the province was no higher

than in England or Scotland. Second, experience with survivors of violence indicated that they were a very different population from the archetypal Vietnam veteran. In particular, there seemed little evidence of impulsive violence or substance abuse. And the final reason for skepticism about the diagnosis of PTSD was the fact that this diagnosis had already begun to fall into disrepute in the United States, where veterans of the Vietnam War were able to obtain a higher pension if they were given this diagnosis.

Early Survey of Survivors of Violence

This first study conducted in Northern Ireland had a simple methodology. The case notes of almost 700 patients were examined retrospectively, and for each case a total of 80 items of information were recorded. These 80 items of information included the symptoms of PTSD, social and environmental factors, and a number of other items that might be expected to predispose to psychological problems. The results of the survey were valuable because of the large number of patients involved and because the retrospective nature of this study (normally a disadvantage) meant that both the patients and interviewers were naïve to the diagnosis of PTSD and, thus, that the symptoms could not have been feigned. Eighty items of information were analyzed by the computer department of Queen's University, Belfast, where a mainframe computer applied the diagnostic criteria defined in DSM-III to the symptom checklists. The research team was surprised by the results.

Despite the skepticism of psychiatrists, when the diagnosis of PTSD was applied to this group of patients, 23% fulfilled the diagnostic criteria. This percentage is very high, especially considering that the symptoms of PTSD were not specifically being looked for at the time when the patients were interviewed. A number of factors were found to be significantly associated with the diagnosis. These findings were very much in keeping with literature on PTSD from other parts of the world, in that female, elderly, and unmarried individuals were more likely to develop PTSD. However, the negative findings were perhaps more striking than the positive findings. Despite the large numbers involved, there was no significant association of PTSD with a past history of psychiatric illness, family history of psychiatric illness, or any of the other putative vulnerability factors that were examined in this study.

It was also striking that 75% of the population had recovered from the psychological difficulties within 18 months. This indicated that most people had recovered before litigation issues had been settled. At that time, the conventional wisdom was the view of Millar, that psychological injuries do not resolve until compensation issues have been resolved.

The results of this study forced psychiatrists in Northern Ireland to reappraise their views on the psychological effects of violence. For the first time, PTSD began to be seen as a true psychological injury. Until then, it was not uncommon for victims of violence to be told by their doctors, "Of course, you're upset. It is normal to be upset when you have been involved in an explosion." Psychiatrists began to realize that saying this to patients was rather like saying to the victim of a road-traffic accident, "Of course, you've got a broken leg. It is normal to have a broken leg when you have been run down by a bus."

It was becoming clear that PTSD is an injury analogous to a grief reaction, in that it is a psychological reaction that happens to normal people when placed under extreme stress, and that there is a natural course of healing after such an injury which takes about 1.5 years. The analogy with grief reaction goes further in that certain circumstances at the time of the trauma might produce a delayed injury and that, if the natural course of the healing process is disrupted, a chronic injury may result. People in Northern Ireland were well aware that a grief reaction is normal and that allowances have to be made for normal people who suffer psychological problems for a year or so after a bereavement. Through a process of education, the public has become aware that psychological injuries after trauma do not imply any psychological weakness and that people with these injuries usually recover with the passage of time.

This study contained a large number of individuals with a wide spectrum of traumatic experiences, ranging from being a casual witness to a bombing to surviving a specific assassination attempt (e.g., one individual had survived an attempted sectarian murder; a gun was put to his head but jammed when the trigger was pulled). In addition, there were a number of individuals who experienced traumatic events unrelated to the Troubles, including common assault and sexual assault. When these different types of traumatic experience were grouped together, the researchers were surprised to find that there was no significant difference in the rate of PTSD among these different groups. It seemed that the seriousness of the traumatic event and the objective level of danger were less important in producing PTSD than the individual's subjective experience of the event.

The study raised a number of questions, which were addressed in a later study. A number of speculative suggestions have been put forward to explain the paradox that individuals affected by violence are psychologically ill and yet no increase in morbidity is seen in the community as a whole. The nature of violence in Northern Ireland changed during the 1970s from widespread civil unrest to attacks on specific targets. A much smaller number of individuals were involved in these violent incidents, and so community surveys might not be sensitive enough to pick up the small numbers involved. Anecdotally, psychiatrists noticed that the level of functioning of certain marginal individuals improved in an atmosphere of violence. It was believed that this might be because sectarian polarization had produced tightly knit ghettos with increased community spirit and support. Durkheim pointed out the importance of these social factors in determining the mental health of a community, and it might be that the detrimental psychological effects of violence for the individuals directly involved might be partly compensated for by the increased community spirit in tightly knit sectarian ghettos. The hypothesis was proposed that a blitz spirit in these communities might be masking the psychological ill effects on those directly effected by violence. If this is the case, it was proposed that, if the Troubles were to end and Northern Ireland returned to more normal social functioning, we might see an upsurge in psychological morbidity.

During the 1980s, a number of scientific papers were published describing survivors of various civilian disasters from all parts of the world. These papers emphasized the validity of the concept of PTSD. It was striking that the psychological sequelae seen after trauma were remarkably similar in different cultures and with different types of trauma. It became clear from the growing literature and from clinical experience that there is a core group of symptoms, involving hyperarousal, emotional constriction, and reliving the event, that are experienced after all types of trauma and that, in addition, to these symptoms there are a variety of other symptoms that depend on the type of trauma experienced. For example, passive victims of violence in Northern Ireland had low rates of substance abuse, survivor guilt, and acting as if the incident were reoccurring compared to active participants in violence such as Vietnam veterans. The survivors of hostage situations seemed to develop symptoms of learned helplessness in addition to the core syndrome of PTSD, and victims of sexual violence often suffered the additional symptoms of post-rape syndrome, including obsessional cleanliness, mistrust of members of the opposite sex, and an aversion to sexual stimuli. The definition of PTSD became more focused in

subsequent revisions of DSM-III, and the difference between the symptoms experienced by adults and children were more clearly delineated.

The Enniskillen Bombing

A terrorist bomb exploded in the town square at Enniskillen on Remembrance Sunday 1987 as the town gathered to remember the dead in two world wars. A group of survivors were followed up at 6 months and 12 months in order to determine the relationship between their physical and psychological injuries. Millar had said that there is an inverse relationship between physical injuries and psychological injuries after a traumatic event. This was found to be true when the survivors were followed up at 6 months, but at 12 months it seemed that those individuals who suffered more severe physical injuries tended to develop delayed PTSD after 6 months. This group of survivors were particularly unfortunate because their psychological injuries were only beginning to emerge at a time when other victims of the trauma were beginning to put the incident behind them. In working with survivors of the Enniskillen bombing, researchers emphasized the importance of local supports, debriefing, and education. Many survivors and their families were relieved to be told that the psychological reactions that they were seeing were, in fact, normal under the circumstances and were likely to improve within 18 months or so. Families were also relieved to hear that the withdrawal and irritability, which they were finding so hard to live with, were in fact predictable parts of a recognized psychological injury that tended to improve with time. In this way, education was able to prevent one of the main complications of PTSD, namely marital disharmony.

The Kegworth Air Disaster

A passenger jet carrying over 90 people traveling between London and Belfast in 1989 crashed on to a stretch of motorway as it was diverted to East Midlands Airport. Research teams followed up the survivors on both sides of the Irish Sea. The relationship between physical and psychological injuries was once more explored in some detail. One of the most striking findings from this piece of research, which was not reported in the original paper, was that individuals from England and Scotland (68%) were significantly more likely to develop PTSD than residents of Northern Ireland (43%) (Table 1). A number of possible reasons have been put forward to explain this result. It may be that even when applying the research criteria psychiatrists in Northern Ireland

Table 1 Rates of PTSD according to home address (%)^a

	<i>England/Scotland</i>	<i>Northern Ireland</i>
No PTSD	7	24
PTSD	15	19

^aPTSD, posttraumatic stress disorder.

tend to be more reluctant to make the diagnosis of PTSD, perhaps applying the research criteria more strictly than our colleagues in mainland United Kingdom. Another possible reason is that individuals living in Northern Ireland have been in some way inured to violence and were, therefore, less traumatized than individuals from mainland United Kingdom, who had less direct experience of traumatic situations. Anecdotal experience indicates that individuals may suffer a series of traumatic events without any psychological ill effects, only to break down very severely after a relatively minor trauma. This suggests that repeated exposure to trauma can strengthen defense mechanisms up to a certain point, but once these defense mechanisms have been overwhelmed the psychological injury seems more severe than otherwise would be the case.

Theoretical Model Used in Debriefing

Successful debriefing seems to depend on making PTSD understandable to the individual involved. Models that are too technical or difficult to believe are of little value. The following model has been developed from various sources, as well as from direct experience of PTSD in Northern Ireland.

Three groups of symptoms are involved in PTSD: hyperarousal, reliving the incident, and emotional constriction. The hyperarousal is best explained as part of the fight-or-flight response to stress. This is one of the survival mechanisms by which nature has equipped us to deal with dangerous or threatening situations. Symptoms such as increased aggression, increased startle reaction, sleep disturbance, and hypervigilance are seen. These symptoms may improve an individual's chances of survival in a dangerous situation and arguably improve the functioning of a combat soldier.

The symptoms of reexperiencing the incident in the form of flashbacks and nightmares is best understood according to Janet's model of traumatic memory. Under normal circumstances, long-term memories are laid down in a process of digesting experience. Long-term memory is not a factual record of events but a story that we tell ourselves to explain the past. It seems that traumatic memories are portions of experience so threatening to the individual that they

cannot be digested by this process. In this way, the traumatic event remains recorded as a factual record, almost like a videotape. Individuals can remember what they had for breakfast on the day of a traumatic event 15 years after the incident. Perhaps this is why listening to the accounts of survivors is so distressing – the events are recorded so literally and usually accurately. These traumatic memories are so distressing that they are phobically avoided. They are suppressed, and the conscious mind attempts to keep them in the subconscious. However, they reemerge when the individual is off-guard (e.g., when asleep). It may be that it is this phobic element that is improved when individuals undergo eye movement desensitization.

Emotional constriction and estrangement from others can be viewed as a loss of innocence. This assumes a state of normal denial, in which most people believe that the world is a safe place and that death and mutilation are things that happen to other people. However, traumatic events force us to confront our own mortality and destroy this normal denial. Thus, people with PTSD are living in a different – more dangerous – world than those individuals around them, who continue to convince themselves that the world is a safe place. In order to be normal again, individuals with PTSD must convince themselves once more that the world is a safe place, and they attempt to do this by suppressing the memory, thus leading to the development of traumatic memories that reemerge in the form of nightmares and flashbacks.

This model has been found to be very useful in talking to survivors of violence. People with PTSD often think of themselves as members of an elite club who are able to see the truth when those around them remain naïve. It is perhaps for this reason that group therapy can be successful in PTSD, although there is a strong danger of scapegoating in groups because of the high levels of free-floating anger that survivors suffer as part of PTSD.

The Clinical Resource and Efficacy Support Team for Northern Ireland reviewed the Northern Ireland literature and produced guidelines for management of PTSD in June 2003. This found strong evidence that a single-session individual debriefing was not helpful and may well be harmful. The team also found evidence that brief programs of cognitive-behavior therapy were effective, as well as pharmacological treatment with selective serotonin reuptake inhibitors (SSRIs), monoamine oxidase inhibitors (MAOIs), and tricyclic antidepressants (TCAs). Eye movement desensitization and supportive group therapy were also found to be effective, and there was evidence that nonprofessional and practical assistance reduced the frequency of avoidance and intrusion symptoms.

Because individuals with PTSD are normal, they often do not respond well to mental health professionals. Normal supports are more useful in supporting these individuals (e.g., family, friends, employment, and church), but debriefing and education are useful in reducing the suffering for survivors and their families. Symptomatic relief can be offered, but, if PTSD persists beyond 1.5 years, the prognosis becomes less positive and the response to treatment is usually incomplete, leaving lasting emotional scars.

Further Reading

- Bell, P., Kee, M., Loughrey, G. L., et al. (1988). Post traumatic stress disorder in N. Ireland. *Acta Psychiatrica Scandinavica* 77, 166–170.
- Cairns, E. and Wilson, R. (1985). Psychiatric aspects of stress in N. Ireland. *Stress Medicine* 1, 193–196.
- Loughrey, G. L., Curran, P. S. and Bell, P. (1993). Post traumatic stress disorder and civil violence in N. Ireland. In: Wilson, J. P. & Raphael, B. (eds.) *The international handbook of traumatic stress syndromes*, pp. 377–380. New York: Plenum.
- Lyons, H. A. (1979). Civil violence – the psychological aspects. *Journal of Psychosomatic Research* 23, 373–376.

Nuclear Warfare, Threat of

J Thompson

University College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J Thompson, volume 3, pp 66–70, © 2000, Elsevier Inc.

Threat of Nuclear Warfare

Evaluating the Threat

Questionnaires: The Chief Method of Studying
Nuclear Anxiety

Individual Studies

Summary

Glossary

<i>Anxiety</i>	The fearful anticipation of potential harm.
<i>Denial</i>	A concept of psychoanalytic origin implying an unwillingness to face unpleasant facts.
<i>Questionnaire</i>	A French word, meaning a list of questions, usually on a particular subject of psychological interest, with a common restricted format for answers. If properly standardised and validated on relevant populations they can be useful in estimating psychological characteristics and states of mind.

Threat of Nuclear Warfare

One thing can be said for certain about nuclear anxiety – it is not an innate fear. On the contrary, no other anxiety depends so much on an appreciation of a wide range of political, cultural, and military threats. Anxiety about the threat of nuclear warfare is caused by the belief that a nuclear war may happen and that it will bring widespread destruction to civilization. This anxiety is linked to powerful images of the explosive power of nuclear weapons and to the dreadful consequences for individuals, society, and the environment if they should come to be used. If all anxiety is based on an assessment of threat, even though the assessment may be faulty, then anxiety about nuclear war is threat assessment at its most difficult.

Evaluating the Threat

A balanced, rational assessment of the threat of nuclear war includes assessing the following issues:

the destructive power of the weapons, the reliability of the weapons-delivery systems, the protection afforded by buildings and shelters, the viability of countermeasures, the extent of the damage caused by the weapons to national infrastructure, the effects of radiation, the effects of major fires on the climate, the possibilities of reconstruction, the nature of enemies, the international political scene, the personalities of the nuclear weapons operators and political leaders, the mood of people in countries that have nuclear weapons, the reliability of warning systems, the possibilities of false alarms, the likelihood of terrorist actions, and the spread of nuclear weapons capability to other nations and to armed groups. By the very nature of the threat, such a list can never be exhaustive. To complicate matters, nuclear weapons are now often categorized as weapons of mass destruction, so the risks are conflated with chemical and biological warfare risks. A less rational assessment of the likelihood of nuclear war is subject to the distorting effects of personal depression, selective monitoring of news, paranoid ideas about other nations, and generalized sociopolitical anxiety.

As a consequence, any proper evaluation of the topic must contrast the objective threat with the perceived threat and then explain any identified differences. In fact, it has been extremely hard for the average citizen to evaluate this threat, which has been with us in various forms for over 50 years. Much of this uncertainty relates to the secrecy that has surrounded nuclear warfare and to the difficulty of determining the intentions of opposed nations.

One approach to getting a more objective assessment is to look at the Doomsday Clock, which was established by the *Bulletin of the Atomic Scientists* in 1947 as a way of showing the consensus view of scientists of the likelihood of nuclear war. Assuming for the moment that these scientists know more about the risks than the average person, the highest danger point, at 2 min to midnight, came in 1953 with the hydrogen bomb tests. The next closest, at 3 min to midnight, came in 1949 when the Soviet Union exploded its first atomic bomb and in 1984 when the arms race accelerated. By 1991 the clock had achieved a safety high point, with a leisurely 17 min to midnight, as arms reductions took place and the Soviet Union contracted. It currently stands at 7 min to midnight, unchanged since February 2002 as a consequence of political extremism and the proliferation of weapons.

There is little doubt that the destructive power of nuclear weapons is very great and has increased

greatly over successive decades. However, many citizens find it difficult to comprehend the destructive power of these weapons, which is so far outside normal human experience, and at the same time may incorrectly believe that everything will be destroyed during a nuclear war. Although it is very hard to be sure about the consequences, there are a range of possibilities.

A full-scale nuclear war might destroy some nations immediately by so damaging their cities and infrastructures that it renders them incapable of providing basic services to their citizens, in the same way that conventional wars have damaged many parts of the world and left the local population destitute, without shelter and food, and dependent on outside help for survival. Such a general world war with nuclear weapons could well leave the survivors incapable of maintaining themselves at their former level of living, in the way that the Black Death disrupted the medieval world. Strategists argue as to whether it would be possible to have a limited nuclear war, many feeling that a nuclear war would quickly spread out of control.

However, from the perspective of the beginning of a new century, the threat of the immediate destruction of civilization may seem remote. The end of the cold war has brought about a change of mood as well as a change in the real level of threat of war between superpowers. Nuclear weapons have come to be seen as an ancient terror, last year's music. Yet the number of nuclear nations has grown, and the threat of regional nuclear conflicts has increased. It would indeed be a cruel fate for our global society to perish from an unfashionable cause.

Questionnaires: The Chief Method of Studying Nuclear Anxiety

Where would the study of anxiety be without the questionnaire? No other method in the social sciences has borne so heavy a load and has been pressed into service in so many campaigns and in so many different theatres of research. The procedure is as follows. The authors make a list of questions to which they would like answers, sometimes drawn from their own interests and imaginations, sometimes drawn from matters of public debate, and even sometimes drawn from opposed points of view. The answers given by willing subjects are often reported in raw form by giving the percentages of respondents who agree with a particular proposition. Sometimes the question list is subjected to item analysis to detect faulty questions and to factor analysis in order to search for an underlying structure to the answers. This generally

leads to revisions of the question list and the methods for scoring the answers to produce subscores with greater explanatory power. The questionnaire is now judged fit for wider use, and comparative studies follow with different populations. Done well, this can be a considerable help in understanding the answers to the questions and can produce interesting comparisons between different groups. Done badly, it can produce many ineffective lists of questions leading to a confusing mass of contradictory findings.

A more profound problem with questionnaires is trying to understand what questionnaire responses really mean. Do they tap real opinions that may lead to real actions, or are they little more than party chatter, the surface moods and immediate responses of the disengaged? In plain terms, is a person who says he or she is anxious about nuclear war really anxious about nuclear war, and what does the person's anxiety amount to? Does it restrict the person in the way that a phobia might? Does the person experience the agony of someone who fears flying but must travel by air for some urgent reason?

To answer those questions, the questionnaire approach must be supplemented with appropriate observations and suitable cross-checks. So, if someone is anxious that there will be a nuclear war, how will he or she behave? If the person believes that war is imminent, some of his or her behaviors ought to be obvious, but if the person cannot be sure how soon the war will start, then it is harder to make predictions. We might surmise that the person would be less likely to make long-term investments and preparations, that he or she would have some sort of emergency nuclear war plan, and that he or she would frequently refer to possible nuclear wars when contemplating all other plans. Such behavioral indicators of nuclear anxiety have been studied infrequently.

The typical subjects of questionnaire studies are captive student populations, although occasionally a proper epidemiological sample is obtained. The ages of the subjects are generally 11 and above, and usually the subjects are tested on one occasion, without much reference to other measures such as intelligence, personality, medical history, school records, spare-time activities, and other interests. Although some questionnaires are used in comparative studies, question lists tend to proliferate and to tap different domains, and often they advertise their purpose in a way that may influence the answers that respondents give. For example, people may be moderately optimistic until faced with a list of dangers and then they become more fearful as they contemplate

all the threats that face them. The demand characteristics of the questionnaire may lead to self-serving distortions.

Individual Studies

Nuclear Warfare Anxiety

The category of nuclear warfare anxiety may contain a number of related fears. These include fears about death, from whatever cause; worries about the dangers of technology and the pace of change in industrial societies; despair at not being able to bring about political change on the nuclear weapons issue; and a general sense of foreboding about the future. Nonetheless, when questionnaires are appropriately refined, nuclear weapons can be shown to be a cause of anxiety in their own right.

How Common Is Nuclear Anxiety?

In the 1980s and early 1990s, nuclear anxiety was common. It was generally cited as being in the top three worries of most young people, although the rate declined as the cold war diminished. The death of parents, doing very badly at school, getting cancer, and being unemployed were the other most common worries. The mid-1980s seem to have been the peak, with as high as 54% of U.S. and 93% of USSR adolescents responding that they were worried about nuclear war. A few years later, the rates had diminished by approximately 10%. Large-scale epidemiological studies show big variations in the rates, with as high as 80% of Finnish adolescents and as (relatively) low as 30% of English and Austrian adolescents reporting such fears. Interestingly, most of the literature on this subject dates from before 1993; after that year, there are far fewer published findings.

Who Is Affected by Nuclear Anxiety?

Most people are affected by nuclear anxiety to some extent. Despite some early anxieties about biased questionnaire samples, nuclear anxiety is not restricted to wealthy, politicized families but is found in all social strata, including profoundly disadvantaged teenagers living in violent inner-city neighborhoods (who might be thought to be too much at present risk to worry about future dangers). In this last group, despite not having activist parents and despite personally despairing about being able to do anything about the nuclear threat, teenagers ranked nuclear war (in 1991) as their second worry – after fear of being murdered – and this fear was higher in those who had actually experienced violence.

Those with nuclear anxiety tend to be more aware of global issues; have good self-esteem; believe that they, and not outside forces, are responsible for their success in life; be not unusually anxious in general; be not religious; have a healthy optimism about the future; and have a tendency toward political activism. Women generally are more likely than men to experience nuclear anxiety.

Does Nuclear Anxiety Run in Families?

The few studies that matched children with their parents found an understandable similarity of attitudes. Nuclear concern is higher in higher-income families; scholastically able children are more knowledgeable and more fearful about nuclear war; and children's hopelessness about the arms race increases as parents' worry about nuclear war increases. Girls express more worry than boys. Interestingly, children with a high degree of general anxiety do not indicate a high degree of nuclear fears. However, parents are often poor at predicting nuclear anxiety in their children and are more likely to deny their own anxiety than are their children.

How Do People Cope with Nuclear Anxiety?

One coping mechanism is to downplay a real anxiety. Keilin sought to go beyond questionnaire studies by subjecting those who reported themselves to be high or low on nuclear anxiety to a laboratory investigation of their emotional responses to nuclear images. People with low nuclear anxiety, although reporting lower levels of emotionality, were more emotional when shown nuclear war stimuli, as reflected by several physiological and behavioral measures. People with high nuclear anxiety were also anxious about relevant social issues, but did not have a greater disposition to be anxious. Keilin interpreted these findings as supporting the existence of nuclear war-related denial processes on the part of the low-anxiety-reported group but only when faced with highly salient and dramatic nuclear stimuli. So, some people repress their nuclear anxieties and others sensitize themselves to the emotional threat, and questionnaires may pick up their preferred coping styles rather than their real emotions.

Another approach is to become an activist. Peace activists sometimes find that their initial anxiety about nuclear war subsides after they have been active for a period of time. They see activism as intrinsically valuable, something they turned to at periods of flux in their lives, believing that they should act to prevent current problems caused by the arms race and not just the anticipated damage from a nuclear war.

How Do the Nuclear Anxious Differ from Others?

Anti-nuclear activists tend to be cooperation-oriented individuals who express spontaneous nuclear concern, have high nuclear anxiety, and also have feelings of empowerment on personal and nuclear issues. They possess feminine traits and are generally tolerant of ambiguity, are open-minded, and have low manipulative tendencies in interpersonal relationships. They are likely to believe that chance does not determine nuclear outcomes and that nuclear war cannot be survived. The most common contrasting group has been those who are pro-defense, that is to say, those who believe that the best protection against nuclear war is to hold nuclear weapons as a deterrent. Pro-defense activists tend to be strength-oriented individuals who express low nuclear anxiety and tend to be masculine, dogmatic, persuasive, and intolerant of ambiguity. They are likely to believe that nuclear issues are controlled by powerful others and chance. This is a very clear differentiation between attitudes and suggests a profoundly different worldview, leading to widely differing rates of nuclear anxiety.

Does Nuclear Anxiety Cause Psychological Disorders?

Specific anxiety about nuclear war does not appear to lead to psychological disorders, although despair about the future (which may show up as nuclear anxiety on general questionnaires) may be linked with adolescent depression. Those with nuclear anxiety followed up 8 years later were very slightly more likely to have psychosomatic health complaints and

emotional distress but not drug, relationship, or work problems.

Summary

Nuclear anxiety appears to be somewhat different from more trait-based stress conditions in that it is a reaction to real-world events and it seems to have subsided with the passing of the cold war. In terms of the usual imperfect measures, it appears that children survived their upbringing in the nuclear age. Although it is hard to prove identifiable major effects of nuclear anxiety, it would be surprising if so widespread a perceived danger had no effect at all. Perhaps the main effect has been to frighten many, to reduce their confidence in the future, and to goad some into activism, leading to considerable relief for all when the arms race ended in a truce.

See Also the Following Article

Hiroshima Bombing, Stress Effects of.

Further Reading

- Keilin, W. J. (1987). Reactions to the threat of nuclear war: An investigation of emotional and defensive processes. PhD thesis, Colorado State University.
- National Academy of Sciences (1986). *The medical implications of nuclear war*. Washington, DC: Institute of Medicine.
- Schatz, R. and Fiske, S. T. (1992). International reactions to the threat of nuclear war: the rise and fall of concern in the eighties. *Political Psychology* 13, 1–30.

Nutrition

J E Morley

Saint Louis University Medical School, St. Louis, MO, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 71–74, © 2000, Elsevier Inc.

Introduction

Stress-Induced Eating: the Tail-Pinch Model

Stress-Induced Anorexia

Anorexia of Aging: a Response to a Physiological Stressor

Stress, Chewing, and Feeding in Humans

Conclusion

Glossary

Corticotropin-releasing hormone (CRH)

A neurotransmitter that plays a role in co-ordinating the response to stress. Administration of CRH causes anorexia, and it appears to be the major neurotransmitter responsible for stress-induced anorexia.

Stress-induced anorexia

Decreased food ingestion secondary to a severe stressor that can be either exteroceptive or interoceptive.

Stress-induced eating

Eating induced in response to an external stressor.

Tail-pinch eating

Eating induced in rodents or cats by gently pinching their tail. It is mainly due to activation of an endogenous opioid, dynorphin.

Introduction

There are numerous interactions between stress and nutrition. Nutrition insecurity is one of the most important stressors of the human race. The need to maintain nutritional intake is the primal driving force of human beings. Simple organisms, such as the *Pleurobranchia*, spend the majority of their life eating and only when adequately nourished (their state of eustress) do they release a satiety hormone that stops eating and then results in egg laying. Severely malnourished women lose the ability to produce luteinizing hormone surges and become amenorrheic. Malnutrition leads to immune dysfunction, including a marked decline in CD4+ T lymphocytes producing a similar picture of immune function to that seen in patients with AIDS. Malnutrition increases the propensity of organisms to develop a variety of infections, which, with associated cytokine release, lead to activation of the stress response. Thus, deficient nutritional intake can be associated with both psychological and physical stressors.

Much of the long-term injury produced by repetitive stress is due to damage produced by free radicals.

A number of nutrients such as vitamins C and E and the trace mineral selenium play an important role as free radical scavengers and as such may limit stress-induced damage.

In addition to the effects of nutrition as a stressor and stress alleviator, an important component of the stress response is alterations in food ingestion. Dependent on the kind and duration, stress can either increase or decrease food intake. A number of psychiatric diseases, such as depression and anorexia nervosa, produce alterations in food intake secondary to activation of the classic stress response. Obesity, which has become a national epidemic in the United States, has been associated with a variety of psychological stressors in relationship to the alteration in body image. Release of stress hormones, such as cortisol, can alter fat deposition, resulting in an increase in visceral obesity. Visceral obesity is associated with an increase in diabetes mellitus, hypertension, myocardial infarction, stroke, and mortality. Aging is a physiological stressor that is associated with anorexia.

Figure 1 summarizes the potential interactions between nutrition and stress.

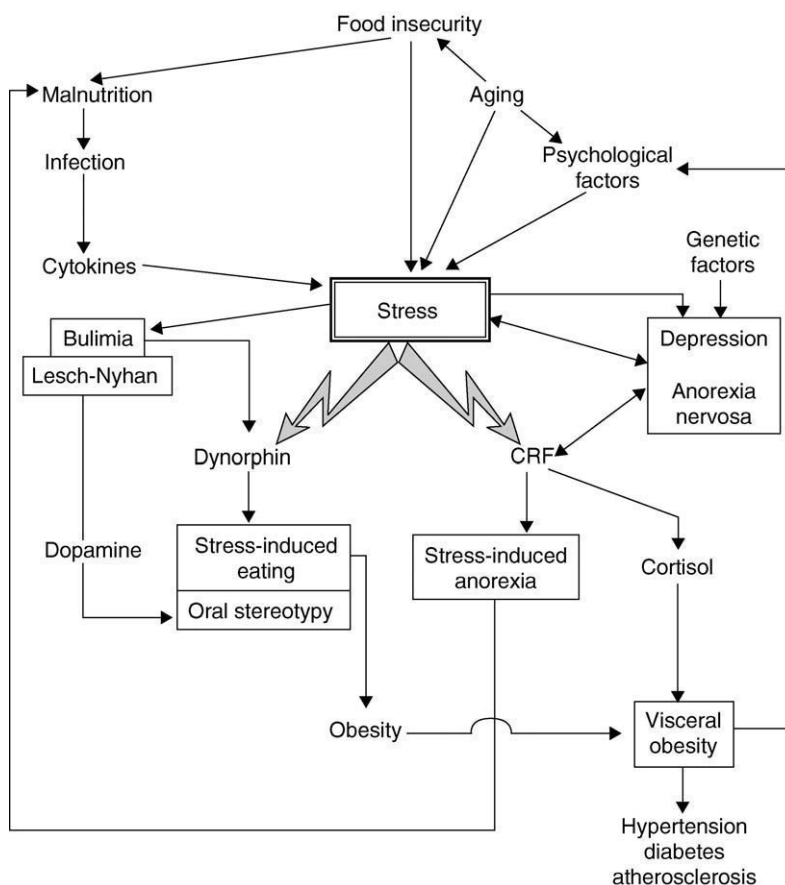


Figure 1 Nutrition and stress interrelate at multiple levels. Nutritional factors can produce stress and stress can lead to under- or overnutrition. Eventually a vicious cycle develops. CRF, corticotropin releasing factor, synonymous with CRH, corticotropin-releasing hormone.

Stress-Induced Eating: the Tail-Pinch Model

Stress-induced ingestion has been classically described in the ethology literature. Eating is commonly observed during boundary disputes in a variety of birds. Pecking at the ground is common, and the frequency can be increased experimentally by replacing the substratum with grain. Cichlid fish bite at the substratum during intervals between fighting. Electric shocks given to sticklebacks result in feeding similar to that seen with food deprivation. Handling rats may increase food intake. Electric shocks induce chewing in monkeys and rabbits. These findings led researchers to develop a stress-induced eating model, namely mild tail pinch in cats, rats, and mice. The stressor appears to be the pinch as a mild pinch to other parts of the body also induced feeding.

The tail-pinch model represents a stimulus-bound behavior in that the rodent stops eating as soon as the pinch is removed. The phenomenon generalizes to sweetened milk but not to water ingestion. Tail pinch appears to drive chewing to a greater degree than it drives ingestive behaviors. Tail pinch may also induce other behaviors such as copulation and maternal behaviors. Tail pinch demonstrates a number of similarities to the behavior seen following electrical stimulation of the lateral hypothalamus.

The neurochemistry involved in tail pinch-induced eating has been explored in some detail with pharmacological approaches. Both dopamine and endogenous opioids appear to play a central role in the modulation of tail-pinch eating. Dopamine appears to be involved predominantly in the production of chewing behaviors, whereas opioids appear to be more involved in ingestive behaviors.

A number of neurotransmitters and peptide hormones have been shown to decrease tail pinch-induced ingestive behaviors. Centrally, these include serotonin, prostaglandin E₂, and thyrotropin-releasing hormone. Peripherally, both cholecystokinin and bombesin decrease food intake in the tail-pinch model.

Prolonged chronic tail pinch eventually results in weight loss and naloxone-perceptible withdrawal symptoms. These animals show similar changes to those seen in morphine-addicted animals.

In summary, tail pinch-induced eating appears to be an excellent model of stress-induced eating, which is predominantly dependent on the activation of dopaminergic and opioid neurotransmitters.

Stress-Induced Anorexia

Severe stressors, such as restraint stress, in rodents result in anorexia and weight loss. Corticotropin-releasing hormone (CRH) has been demonstrated

to be a potent anorectic agent when administered intracerebroventricularly to rodents. CRH levels increase in the hypothalamus during restraint stress. Inhibition of CRH with antibodies or antagonists reverses the anorexia associated with restraint stress.

Anorexia produced by CRH can be attenuated by tail pinch, suggesting a delicate balance between stress-induced eating and stress-induced anorexia.

Overall, considerable evidence shows that stress-induced anorexia is due to the activation of CRH. CRH also appears to play an important role in stress-induced anorexia in humans (*vide infra*).

Anorexia of Aging: a Response to a Physiological Stressor

The aging process can be conceived to be a physiological stressor that produces its effects over the life span. Over the life span there is approximately a 30% decrease in food intake. This is predominantly due to a decrease in the central opioid-feeding drive. There does not appear to be an alteration in the effects of CRH with aging. In addition to the central neurotransmitter changes with aging, there are also peripheral alterations in adaptive relaxation of the fundus of the stomach due to decreased nitric oxide synthase activity. This leads to early antral filling and a greater degree of satiation for a given amount of food ingested.

The anorexia of aging makes the older individual at greater risk for developing severe anorexia and weight loss when exposed to stressors. The release of cytokines from varying sources has a potent anorectic effect in older persons secondarily to activation of the CRH stress cascade.

Stress, Chewing, and Feeding in Humans

Stress has long been recognized as a cause of masticatory muscle contraction. As pointed out by St. Matthew, "and there shall be a weeping and gnashing of teeth." Biting into a hard object was used by early surgeons to decrease the pain of surgery in the days antedating anesthesia. Anxiety increases bruxism (tooth grinding). In Lesch-Nyhan syndrome (hypoxanthine-guanine phosphoribosyltransferase deficiency), oral stereotypy is increased markedly during periods of emotional distress.

In humans, approximately two-thirds become anorectic when stressed, whereas one-third overeat. In stressed overeaters the act of chewing appears to be as important as the actual ingestion of food. When the stressor is boredom, overeating is the most common associated behavior. Bulimia (binge eating, which may be associated with subsequent vomiting) is often precipitated by a traumatic event.

Anorexia nervosa represents a classic pathological example of intrapsychic stress resulting in anorexia and severe weight loss. Patients with anorexia nervosa have been demonstrated to have elevated CRH levels in the cerebrospinal fluid.

Depression represents a major intrapsychic stressor. Two-thirds of depressed persons tend to be anorectic, and in these individuals, CRH levels in the cerebrospinal fluid are increased.

Conclusion

There is now ample evidence that not only does nutritional insufficiency result in stress, but stressors can lead to either over- or undereating in human beings. Both the endogenous opioid system (dynorphin) and CRH have been demonstrated to be neurotransmitters playing a role in the alterations in ingestive behavior produced by stress. A number of diseases that have a predominant component of altered food ingestion, such as anorexia nervosa and bulimia, appear to be precipitated by stress. Nutrient intake remains the most important behavior for the survival of all organisms. It is not surprising, therefore, that nutrition and stress are intertwined so intimately.

See Also the Following Articles

Diet and Stress, Non-Psychiatric; Eating Disorders and Stress; Obesity, Stress and; Diet and Stress, Psychiatric.

Further Reading

- Greeno, C. G. and Wing, R. R. (1994). Stress-induced eating. *Psychological Bulletin* **115**, 444–464.
- Grunberg, N. E. and Straub, R. O. (1992). The role of gender and taste class in the effects of stress on eating. *Health Psychology* **11**, 97–100.
- Morley, J. E. and Levine, A. S. (1980). Stress-induced eating is mediated through endogenous opiates. *Science* **209**, 1259–1261.
- Morley, J. E., Levine, A. S. and Rowland, N. E. (1983). Stress induced eating. *Life Sciences* **32**, 2169–2182.
- Nemeroff, C. B. (1988). The role of corticotropin-releasing factor in the pathogenesis of major depression. *Pharmacopsychiatry* **21**, 76–82.
- Troop, N. A., Holbrey, A., Trowler, R., et al. (1994). Ways of coping in women with eating disorders. *Journal of Nervous and Mental Disease* **182**, 535–540.
- Weinstein, S. E., Shide, D. J. and Rolls, B. J. (1997). Changes in food intake in response to stress in men and women: psychological factors. *Appetite* **28**, 7–18.

VOLUME ONE

ENCYCLOPEDIA OF STRESS

SECOND EDITION

EDITOR-IN-CHIEF
GEORGE FINK



DEDICATION

For Ann Elizabeth

EDITOR-IN-CHIEF

George Fink

Professorial Research Fellow (formerly Director)
Mental Health Research Institute of Victoria
Parkville, Melbourne, Victoria
Australia
Formerly Director, MRC Brain Metabolism Unit
Edinburgh, Scotland, UK.

ASSOCIATE EDITORS

Bruce McEwen

The Rockefeller University
Laboratory of Neuroendocrinology
1230 York Avenue, Box 165
New York NY 10021

E. Ronald de Kloet

Universiteit Leiden
Natuurwetenschappen Wiskunde en
LACDR Medical Pharmacology
Einsteinweg 55, kamer HB 902
Leiden 2333 CC
The Netherlands

Robert Rubin

Department of Psychiatry (116A)
VA Greater LA Healthcare System
11301 Wilshire Blvd.
Los Angeles, CA 90073

George Chrousos

First Department of Pediatrics,
Athens University Medical School,
Aghia Sophia Children's Hospital,
115 27 Athens, Greece

Andrew Steptoe

University College London
Department of Epidemiology and Public Health
Psychology
1-19 Torrington Place London WC1E 6BT
UK

Noel Rose

Johns Hopkins University
Center for Autoimmune Disease Research
Molecular Microbiology and Immunology
615 North Wolfe Street, Rm E5014
Baltimore MD 21205

Ian Craig

SGDP Research Centre
Institute of Psychiatry, Molecular Genetics Group
De Crespigny Park, Box P082
Denmark Hill
London SE5 8AF
UK

Giora Feuerstein

Merck Research Laboratories 42-209
Department of Cardiovascular Diseases
770 Sumneytown Pike
West Point PA 19486

CONTENTS

<i>Contents by Subject area</i>	<i>xxiii</i>
<i>Preface to First Edition</i>	<i>xxxi</i>
<i>Preface to the Second Edition</i>	<i>xxxiii</i>
<i>Guide to Encyclopedia</i>	<i>xxxv</i>
<i>Foreword</i>	<i>xxxvii</i>

VOLUME 1

A

Acute Stress Disorder and Posttraumatic Stress Disorder	1
R. Yehuda and C. M. Wong	
Acute Stress Response: Experimental	7
K. Pacak and R. McCarty	
Acute Trauma Response	15
W. C. Chiu, D. E. Carlson and M. P. Lilly	
Adenylyl Cyclases and Stress Response	21
F. A. Antoni	
Adjustment Disorders	24
M. Dascalu and D. Svrakic	
Adolescence	28
G. N. Swanson	
Adolescent Suicide	36
M. Berk, R. Suddath and M. Devich-Navarro	
Adrenal Cortex	38
G. P. Vinson, B. J. Whitehouse and J. P. Hinson	
Adrenal Insufficiency	47
H. S. Willenberg, S. R. Bornstein and G. P. Chrousos	

Adrenal Medulla	52
R. Kvetnansky and R. McCarty	
Adrenaline	60
T. M. Pollard	
Adrenocortical Function, Factors Controlling Development Thereof	64
T. Else and G. D. Hammer	
Adrenocorticotrophic Hormone (ACTH)	69
M. E. Rhodes	
Aerobic Exercise and Stress Reduction	73
E. J. C. de Geus and J. H. Stubbe	
Affective Disorders	78
D. F. MacKinnon	
Aggression	84
E. F. Coccaro and E. C. Manning	
Aggressive Behavior	89
J. M. Koolhaas	
Aging and Adrenocortical Factors	92
C. Lord and J. C. Pruessner	
Aging and Psychological Stress	96
B. W. J. Penninx	
Aging and Stress, Biology of	102
M. A. Horan, R. N. Barton and G. J. Lithgow	
AIDS	108
M. H. Antoni and D. G. Cruess	
Airline Accidents	114
G. Li	
Alarm Phase and General Adaptation Syndrome	119
R. McCarty and K. Pacak	

Bereavement	317	Caregivers, Stress and	416
P. J. Clayton		S. H. Zarit, K. Bottigi and J. E. Gaugler	
Beta-Adrenergic Blockers	323	Catecholamines	419
M. B. Hamner and G. W. Arana		U. Lundberg	
Beta-Endorphin	332	Central Stress Neurocircuits	424
M. Lee and S. L. Wardlaw		S. Kollack-Walker, H. E. W. Day and H. Akil	
Blood Pressure	335	Cerebral Metabolism, Brain Imaging	432
A. Sherwood and R. A. Carels		K. P. Ebmeier, C. L. Donaghey and N. J. Dougall	
Blood-Brain Barrier, Stress and	342	Chaperone Proteins and	
S. N. Malaeb and B. S. Stonestreet		Chaperonopathies	438
Borderline Personality Disorder	348	A. J. L. Macario and E. Conway de Macario	
H. W. Koenigsberg and L. J. Siever		Chaperonopathies	444
Brain and Brain Regions	351	A. J. L. Macario and E. Conway de Macario	
A. G. Watts		Chemical Warfare	449
Brain Natriuretic Peptide (BNP)	357	J. Berberich	
P. Pervanidou and G. P. Chrousos		Chernobyl, Stress Effects of	452
Brain Trauma	360	A. Tønnessen and L. Weisæth	
B. Pentland		Child Abuse	457
Breast Cancer	364	C. C. Swenson and L. Saldana	
A. Moyer		Child Physical Abuse	460
Burnout	368	C. C. Swenson	
C. Maslach and M. P. Leiter		Child Sexual Abuse	463
		J. A. Cohen	
		Childbirth and Stress	467
		S. Ayers and E. Ford	
		Childhood Stress	472
		S. Sandberg	
		Cholesterol and Lipoproteins	478
		C. M. Stoney	
		Chronic Fatigue Syndrome	484
		A. J. Cleare and S. Wessely	
		Chronic Social Stress: GR Sensitivity in	
		Leukocytes	493
		A. Weizman and B. Rotberg	
		Circadian Clock Genes as Modulators of	
		Sensitivity to Genotoxic Stress	496
		M. P. Antoch and R. V. Kondratov	
		Circadian Rhythm Effects on	
		Cardiovascular and Other Stress-Related	
		Events	500
		R. Manfredini, B. Boari, R. Salmi, A. M. Malagoni and F. Manfredini	
Cardiovascular System and Stress	396		
P. Hjelm Dahl			
Cardiovascular Disease, Stress and	410		
G. P. Chrousos and G. Kaltsas			

C

Calbindin	373
F. A. Antoni	
Calcium, Role of	374
F. A. Antoni	
Calcium-Dependent Neurotoxicity	375
J. S. Kelly	
Cancer	378
D. Spiegel	
Cancer Treatment	384
F. I. Fawzy, A. L. Canada and N. W. Fawzy	
Captivity, Adaptation to	388
R. H. Rahe	
Captivity, Recovery from	392
R. H. Rahe	
Cardiovascular System and Stress	396
P. Hjelm Dahl	
Cardiovascular Disease, Stress and	410
G. P. Chrousos and G. Kaltsas	

Circadian Rhythms, Effects of Prenatal Stress in Rodents	505	Corticosteroids and Stress	613
S. Maccari and O. Van Reeth		A. Munck	
Circadian Rhythms, Genetics of	508	Corticotropin Releasing Hormone (CRH)	620
F. W. Turek and M. H. Vitaterna		A. T. Lim	
Cognition and Stress	513	Corticotropin Releasing Factor Receptor Deficiency in Mice	623
M. W. Eysenck		N. J. Justice and K.-F. Lee	
Cognitive-Behavioral Therapy	515	Corticotropin Releasing Factor-Binding Protein	627
G. A. Fava		P. J. Lowry, C. F. Kemp and R. J. Woods	
Combat Reaction, Chronic	518	Corticotropin-Releasing Factor Circuitry in the Brain – Relevance for Affective Disorders and Anxiety	630
R. H. Rahe		D. A. Gutman and C. B. Nemeroff	
Combat Stress Reaction	524	Corticotropin-Releasing Factor (CRF) Family of Neuropeptides – Role in Inflammation	635
M. Dobson		A. Gravanis and A. N. Margioris	
Combat, Acute Reactions to	529	Corticotropin-Releasing Factor Receptors	641
R. H. Rahe		D. E. Grigoriadis	
Common Cold and Stress	533	Cortisol Awakening Response	649
A. Smith		A. Steptoe	
Community Studies	536	C-Reactive Protein	653
C. J. Holahan, R. H. Moos and L. M. Groesz		W. J. Kop and A. A. Weinstein	
Comorbid Disorders and Stress	542	Crime Victims	659
S. Sundram and Avril Pereira		I. Robbins	
Comparative Anatomy and Physiology	549	Crisis Intervention	662
A. G. Watts		D. Hamaoka, D. Benedek, T. Grieger and R. J. Ursano	
Concentration Camp Survivors	553	Critical Thermal Limits	667
J. D. Kinzie		J. Roth	
Congenital Adrenal Hyperplasia (CAH)	556	Crowding Stress	669
A. Solomon and P.-M. G. Bouloux		L. Kovács and P. Csermely	
Conservation of Resources Theory	562	Cultural Factors in Stress	672
S. E. Hobfoll and J. S. Ford		J. W. Berry and B. Ataca	
Control and Stress	568	Cultural Transition	678
A. Steptoe		M. S. Kopp	
Coping and Stress: A Lens and Filter Model	574	Cushing's Syndrome, Medical Aspects	682
R. H. Rahe		S. R. Bornstein, M. Gruber, H. S. Willenberg, C. A. Stratakis and G. P. Chrousos	
Coping Skills	578	Cushing's Syndrome, Neuropsychiatric Aspects	688
A. DeLongis and E. Puterman		M. N. Starkman	
Corticosteroid Receptor Genes: Functional Dissection in Mice	584	Cytokines	692
F. Tronche		G. D. Marshall	
Corticosteroid Receptors	594		
O. C. Meijer, E. R. de Kloet and B. S. McEwen			
Corticosteroid-Binding Globulin (Transcortin)	605		
B. E. P. Murphy			

Cytokines, Chronic Stress, and Fatigue S. Jain and P. J. Mills	698	Diet and Stress, Non-Psychiatric J. Wardle and E. L. Gibson	797
Cytokines, Stress, and Depression B. E. Leonard and C. Song	705	Diet and Stress, Psychiatric V. March and M. H. Fernstrom	806
Cytotoxic Lymphocytes M. A. Fletcher and N. G. Klimas	711	Disaster Syndrome P. Valent	811
D		Disasters and Mass Violence, Public, Effects of G. Stevens, B. Raphael and M. Dobson	814
Death Anxiety R. Kastenbaum	717	Disease, Stress Induced H. S. Willenberg, S. R. Bornstein and G. P. Chrousos	824
Defensive Behaviors D. C. Blanchard, M. Yang, M. Hebert and R. J. Blanchard	722	Dissociation J. R. Maldonado	828
Demand–Control Model T. Theorell	727	Distress G. Matthews	838
Dental Stress T. K. Fábán, G. Fábán and P. Fejérdy	733	Divorce, Children of K. N. Hipke, S. A. Wolchik and I. N. Sandler	844
Depersonalization: Systematic Assessment M. Steinberg	736	Domestic Violence B. Donohue, H. Hill and T. Maier-Paarlberg	848
Depression and Coronary Heart Disease F. Lespérance and N. Frasure-Smith	741	Dopamine, Central G. D. Stanwood	852
Depression and Manic–Depressive Illness R. T. Rubin and B. J. Carroll	744	<i>Drosophila</i> Genes and Anoxia G. G. Haddad	859
Depression and Stress, Role of n-3 and n-6 Fatty Acids C. Song and B. E. Leonard	754	<i>Drosophila</i> Studies M. Allikian and J. Tower	864
Depression Models K. Matthews and C. Stewart	760	Drug Use and Abuse J. R. Mantsch	866
Depression, Immunological Aspects M. R. Irwin	766	E	
Dermatological Conditions M. A. Gupta	773	Earthquakes, Stress Effects of M. Livanou and M. Başoğlu	871
Desensitization F. A. Antoni	778	Eating Disorders and Stress D. C. Jimerson	876
Dexamethasone Suppression Test (DST) R. T. Rubin and B. J. Carroll	780	Eclampsia and Pre-Eclampsia A. Makrigiannakis, G. Petsas and G. P. Chrousos	880
DEX-CRH Test N. C. Schommer and I. Heuser	784	Economic Factors and Stress R. A. Catalano	884
DHEA J. Herbert	788	Education Levels and Stress J. Mirowsky and C. E. Ross	888
Diabetes, Type 1 A. Riazi and C. Bradley	792		

Food Shift Effect F. K. Stephan	88	Glucocorticoid Receptor Mutations and Polymorphisms J. W. Koper	183
Freud, Sigmund D. J. Lynn	90	Glucocorticoids – Adverse Effects on the Nervous System R. M. Sapolsky	185
G		Glucocorticoids, Effects of Stress on J. C. Buckingham	190
GABA (Gamma Aminobutyric Acid) J. D. C. Lambert	97	Glucocorticoids, Overview B. E. Pearson Murphy	198
Gastrointestinal Effects R. Murison and A. M. Milde	109	Glucocorticoids, Role in Stress J. C. Buckingham	210
Gender and Stress R. J. Handa and W. C. J. Chung	115	Glucose Transport A. L. McCall	217
Gene–Environment Interactions in Early Development I. S. P. Davis and R. Plomin	122	Glycobiology of Stress G. Lauc and M. Flögel	222
Genetic Factors and Stress C. A. Koch and C. A. Stratakis	128	Gonadotropin Secretion, Effects of Stress on M. Ferin	228
Genetic Polymorphisms in Stress Response I. W. Craig	135	Graves' Disease (Thyrotoxicosis) W. M. Wiersinga	234
Genetic Predispositions to Stressful Conditions D. Blackwood and H. Knight	141	Grieving A. Ray and H. Prigerson	238
Genetic Testing and Stress V. Senior and M. Cropley	146	Group Therapy N. Wong	242
Genetic Variation of HPA Axis Activity and Function in Farm Animals P. Mormède	150	Gulf War Syndrome, Psychological and Chemical Stressors H. Soreq	248
Ghrelin and Stress Protection T. Brzozowski, M. Pawlik, D. Drozdowicz, Z. Sliwowski, S. J. Konturek, W. W. Pawlik and P. C. Konturek	153	H	
Glia or Neuroglia G. W. Bennett and D. E. Ray	161	Health and Socioeconomic Status T. Chandola and M. Marmot	255
Glucocorticoid Effects on Memory: the Positive and the Negative O. T. Wolf	166	Health Behavior and Stress A. Steptoe	262
Glucocorticoid Negative Feedback M. F. Dallman	172	Heart Disease/Attack G. J. Baker, S. Suchday and D. S. Krantz	266
Glucocorticoid Receptor Mutant Mice as Models for Stress-Induced Affective Disorders P. Gass	176	Heart Failure, Stress Effects M. Alevizaki and G. P. Chrousos	272
		Heart Rate J. R. Jennings	274
		Heat Resistance A. J. L. Macario and E. Conway de Macario	278

Heat Shock Genes, Human	284	Hyperthermia	381
A. J. L. Macario and E. Conway de Macario		J. Roth	
Heat Shock Proteins: HSP60 Family Genes	288	Hyperthyroidism	388
H. Kubota		I. M. Lesser and D. L. Flores	
Heat Shock Response, Overview	292	Hyperventilation	390
A. J. L. Macario and E. Conway de Macario		C. Bass	
Hemostasis and Stress	300	Hypnosis	394
Roland von Känel		W. G. Whitehouse, E. C. Orne and M. T. Orne	
Herpesviruses	305	Hypocortisolism and Stress	400
D. A. Padgett, M. T. Bailey and J. F. Sheridan		C. M. Heim and U. M. Nater	
Hippocampal Neurons	311	Hypoglycemia	408
R. L. Spencer and S. T. Bland		B. M. Frier	
Hippocampus, Corticosteroid Effects on	321	Hypotension, Hypovolemia, and Septic Shock	413
M. Joëls and H. Karst		A. Beishuizen, A. B. Johan Groeneveld and I. Vermes	
Hippocampus, Overview	327	Hypothalamic-Pituitary-Adrenal Axis	421
M. J. Meaney and S. J. Lupien		M. F. Dallman, S. Bhatnagar and V. Viau	
Hiroshima Bombing, Stress Effects of	332	Hypothermia	428
R. J. Lifton		M. J. Taylor	
HIV Infection/AIDS	336	Hypothyroidism	439
B. W. Dixon		R. T. Joffe	
Holocaust Survivors, Experiences of	339	Hysteria	442
P. Valent		K. Pajer	
Holocaust, Stress Effects of	342		
P. Valent			
Homeostasis	347		
B. S. McEwen			
Homosexuality, Stress and	348	Immobilization Stress	445
J. Drescher		R. Kvetnansky and R. McCarty	
Hostility	354	Immune Cell Distribution, Effects of Stress on	449
L. H. Powell and K. Williams		F. S. Dhabhar	
HPA Alterations in PTSD	359	Immune Function, Stress-Induced Enhancement	455
R. Yehuda		F. S. Dhabhar	
Hurricane Katrina Disaster, Stress Effects of	364	Immune Response	462
C. Piotrowski		P. J. Delves	
11β-Hydroxysteroid Dehydrogenases	368	Immune Suppression	470
J. R. Seckl		P. Prolo and F. Chiappelli	
Hyperreactivity (Cardiovascular)	372	Immune Surveillance – Cancer, Effects of Stress on	477
A. Georgiades		D. Spiegel and F. S. Dhabhar	
Hypertension	376	Immune System, Aging	481
A. Steptoe		M. A. Horan	

Immunity	485	L	
F. Chiappelli			
Impact of Terrorism on the Development of Mental Health Symptoms	493	Learned Helplessness	567
R. Yehuda		D. M. Isaacowitz and M. E. P. Seligman	
Impotence, Stress and	497	Learning and Memory, Effects of Stress on	571
M. R. Gignac, G. M. Rooker and J. K. Cohen		M. Lindau, O. Almkvist and A. H. Mohammed	
Impulse Control	500	Left Ventricular Mass	577
I.-M. Blackburn		T. G. Pickering	
Incest	502	<i>Leishmania</i>, Stress Response in	579
A. P. Mannarino		M. Shapira	
Income Levels and Stress	506	Leptin, Adiponectin, Resistin, Ghrelin	584
S. V. Subramanian and I. Kawachi		A. Gavrilu, D. Barb and C. S. Mantzoros	
Indigenous Societies	511	Leukocyte Trafficking and Stress	592
W. W. Dressler		S. Hong, M. U. Goebel and P. J. Mills	
Industrialized Societies	517	Life Events and Health	599
J. Siegrist		P. Surtees and N. Wainwright	
Infection	521	Life Events Scale	603
H. Friedman and S. H. Pross		E. Wethington	
Inflammation	530	Lockerbie Air Crash, Stress Effects of	608
G. Z. Feuerstein, R. R. Ruffolo, C. Coughlin, J. Wang and D. Miller		H. Livingston and M. Livingston	
Instinct Theory	535	Loss Trauma	612
R. Gardner Jr.		A. Bifulco	
Integrative Medicine (Complementary and Alternative Medicine)	537	Lymph Nodes	616
D. Spiegel		T. L. Whiteside	
Interactions Between Stress and Drugs of Abuse	540	Lymphocytes	622
P. V. Piazza and M. Le Moal		N. R. Rose	
J		M	
Job Insecurity: The Health Effects of a Psychosocial Work Stressor	549	Macrophage Antimycobacterial Activity, Effects of Stress on	629
J. E. Ferrie and P. Martikainen		C. S. Boomershine and B. S. Zwilling	
K		Macrophages	634
Kidney Function	557	W. P. Fehder, F. Tuluc, W.-Z. Ho and S. D. Douglas	
W. J. Welch		Major Depressive Disorder	640
Korean Conflict, Stress Effects of	563	A. B. Negrão and P. W. Gold	
C. A. Goguen and M. J. Friedman		Male Partner Violence	645
		M. Ingram, N. P. Yuan and M. P. Koss	
		Marital Conflict	651
		P. T. McFarland and A. Christensen	

Marital Status and Health Problems	653	Mitochondria	754
I. M. A. Joung		I. Manoli, S. Alesci and G. P. Chrousos	
Marriage	660	Monoamine Oxidase	761
A. C. Yoneda and J. Davila		P. Huezo-Diaz and I. W. Craig	
Maternal Deprivation	667	Motor Vehicle Accidents, Stress Effects of	764
R. L. Huot, C. O. Ladd and		T. C. Buckley and E. B. Blanchard	
P. M. Plotsky		Mucosal Secretory Immunity, Stress and	768
Medical Profession and Stress	674	J. A. Bosch and D. Carroll	
K. G. Power and V. Swanson		Multi Drug Resistance P Glycoprotein	
Meditation and Stress	678	and other Transporters	774
J. L. Kristeller		E. C. M. de Lange	
Membrane Glucocorticoid Receptors	686	Multiple Personality Disorder	783
P. J. Gasser and M. Orchinik		R. P. Kluft	
Memory and Stress	693	Multiple Sclerosis	790
S. J. Lupien and F. S. Maheu		A. T. Reder	
Memory Impairment	699	Multiple Trauma	795
A. J. Parkin [†]		P. N. Soucacos and E. O. Johnson	
Menopause and Stress	703	Musculoskeletal Problems and Stress	800
N. E. Avis		S. Svebak	
Menstrual Cycles and Stress	706	Myopathy	807
R. Suri and L. Altshuler		G. A. Small	
Mental Stress Testing	712		
P. G. Saab, K. A. Kline and			
J. R. McCalla			
Metabolic Syndrome	717	Natural Killer (NK) Cells	815
L. Keltikangas-Järvinen		T. L. Whiteside, M. Boyiadzis and	
Metabolic Syndrome and Stress	721	R. B. Herberman	
R. Rosmond		Negative Affect	822
Metals, Oxidative Stress, and Brain		A. A. Stone and A. A. Gorin	
Biology	724	Neighborhood Stress and Health	825
D. I. Finkelstein, T. Lynch, S. Wilkins,		A. V. Diez Roux	
R. A. Cherny and A. I. Bush		Nelson's Syndrome	828
Metastasis	726	A. Stathopoulou, K. Dimitriou and	
M. K. Demetrikopoulos		G. Kaltsas	
Metyrapone: Basic and Clinical Studies	730	Neural Stem Cells	832
E. A. Young		U. S. Sohur, J. G. Emsley,	
Migraine	733	B. D. Mitchell and J. D. Macklis	
N. M. Ramadan		Neurodegenerative Disorders	840
Mineralocorticoid Receptor		M. F. Mendez and A. M. McMurtry	
Polymorphisms	744	Neurodevelopmental Disorders	844
R. H. DeRijk and E. R. de Kloet		in Children	
Minorities and Stress	748	N. J. Rinehart and B. J. Tonge	
I. Mino, W. E. Profit and C. M. Pierce		Neuroendocrine Systems	851
		G. Fink	
		Neurogenesis	865
		P. Tanapat and E. Gould	

[†]Deceased.

Pheromones	119	Prison	217
A. Kumar, C. A. Dudley, S. Chakravarty and R. L. Moss		D. L. Whitehead and A. Steptoe	
Pituitary Regulation, Role of	127	Prisoners of War	223
F. A. Antoni		C. Tennant	
Police, Stress in	131	Problem-Solving Skills Training	227
R. J. Burke		A. M. Nezu, C. M. Nezu and T. J. D’Zurilla	
Posttraumatic Stress Disorder – Clinical	135	Prolactin and Stress	231
N. C. Feeny, L. R. Stines and E. B. Foa		G. Tolis, G. Rombopoulos, D. Kaltsas, E. Katounda, V. Kaltzidou and N. Angelopoulos	
Posttraumatic Stress Disorder – Neurobiological basis for	140	Pro-opiomelanocortin (POMC)	233
M. Barad		A. B. Bicknell	
Posttraumatic Stress Disorder in Children	145	Prostaglandins	240
A. M. La Greca		S. Moshonov, U. Zor and Z. Naor	
Posttraumatic Stress Disorder, Delayed	150	Proteases in Prokaryotes and Eukaryotic Cell Organelles	247
A. Holen		E. Conway de Macario and A. J. L. Macario	
Posttraumatic Stress Disorder, Neurobiology of	152	Proteases in the Eukaryotic Cell Cytosol	252
J. D. Bremner		E. Conway de Macario and A. J. L. Macario	
Posttraumatic Therapy	157	Protein Synthesis	258
A. M. Rasmusson, C. M. Monson and P. A. Resick		M. A. Brostrom and C. O. Brostrom	
Pregnancy – Maternal and Perinatal Stress – Effects of	165	Proteosome	266
P. Smirnaki and M.-A. Magiakou		M. Maldonado and J. Wang	
Premenstrual Dysphoric Disorder	173	Psoriasis	271
S. Nowakowski, P. Haynes and B. L. Parry		A. B-. Kirschbaum	
Pre-pulse Inhibition	180	Psychoanalysis	274
M. van den Buuse		P. Roazen	
Pressure, Effects of Extreme High and Low	184	Psychological Stressors, Overview	278
R. J. Værnes		S. M. Monroe and G. M. Slavich	
Primate Hierarchies and Personality	194	Psychoneuroimmunology	284
R. M. Sapolsky		R. Dantzer	
Primate Models, Behavioral– Immunological Interactions	199	Psychosocial Factors and Stress	288
M. L. Laudenslager and S. Tiefenbacher		J. Siegrist	
Primate Models, Cardiovascular Disease	204	Psychosomatic Heart Disease: Role of Sympathetic and Sympathoadrenal Processes	292
J. R. Kaplan		G. W. Lambert and T. Dawood	
Primate Models, Overview	211	Psychosomatic Medicine	296
D. M. Lyons		T. Theorell	
Primates: Rearing and Effects of Stress on Primate CNS Function	214	Psychotherapy	302
T. K. Newman and C. S. Barr		G. C. Smith	

Psychotic Disorders J. Ventura	308	Revenge Fantasies M. Horowitz and S. Meffert	391
Q		S	
Quality of Life S. M. Skevington	317		
R			
Racial Harassment/Discrimination D. R. Williams and S. A. Mohammed	321	Salivary Cortisol C. Kirschbaum and D. H. Hellhammer	405
Recovery from Stress S. Sonnentag and C. Fritz	327	Salt Appetite R. S. Weisinger and N. Chen	409
Reductive Stress J. P. Kehrer	331	Schizophrenia M. van den Buuse and D. Copolov	418
Reenactment Techniques N. Wong	336	School Stress and School Refusal Behavior C. A. Kearney, L. C. Cook and G. Chapman	422
Refugees, Stress in J. D. Kinzie	338	School Violence and Bullying J. Juvonen	425
Regional Blood Flow, Stress Effects W. W. Blessing	341	Seasonal Changes in Stress Responses R. J. Nelson and L. B. Martin II	427
Relaxation Techniques W. G. Whitehouse, E. C. Orne and M. T. Orne	345	Seasonal Rhythms L. M. Romero	432
Religion and Stress S. Packer	351	Secretagogue G. Fink	435
9/11, Religion and Stress S. Packer	357	Selective Serotonin Reuptake Inhibitors (SSRIs) M. Gitlin	440
Remodelling of Neuronal Networks by Stress E. Fuchs and G. Flügge	364	Self-Esteem, Stress, and Emotion D. Roger	443
Renal and Adrenocortical Actions of Dopamine B. C. Williams, Y.-C. Lo and S. W. Walker	371	Selye, Hans B. Tuchweber and P. Bois	448
Reproduction, Effects of Social Stress On C. A. Shively	374	Sepsis, Acute Respiratory Distress Syndrome, and Glucocorticoid Resistance G. U. Meduri	450
Reproductive Dysfunction in Primates, Behaviorally Induced J. L. Cameron	380	Serotonin B. C. Williams and Y.-C. Lo	457
Resistance L. M. Zabarenko	386	Serotonin in Stress S. Kusljic and M. van den Buuse	461
Restraint Stress R. J. Servatius, G. Salameh, K. M. Coyle and W. P. Paré	389	Serotonin Transporter Genetic Modifications K. Sugden	465
		Sex Differences in Human Stress Response B. M. Kudielka, D. H. Hellhammer and C. Kirschbaum	469

Sex Steroids, Response to Stress and Susceptibility to Depression	474	Startle Response	561
M. V. Seeman and L. E. Ross		C. O. Ladd, P. M. Plotsky and M. Davis	
Sex-Specific Effects of Early Social Stress in Mammals: A Study in Guinea Pigs	479	Steroid Hormone Receptors	568
S. Kaiser and N. Sachser		M. Beato and J. Klug	
Sexual Assault	484	Steroid Hydroxylases	581
N. C. Feeny, T. J. Linares and E. B. Foa		F. H. de Jong	
Sexual Dysfunction	490	Strain Differences in Stress Response in Rodents	585
W. T. O'Donohue and L. Woodward Tolle		P. Mormède	
Sexual Offenders	494	Stress and Anxiety: Treatment with 2nd Generation Antipsychotic Drugs	587
F. M. Saleh and H. M. Malin		J. S. Ballon, J. A. Boyd and D. A. Wirshing	
Sickle Cell Disease and Stress	502	Stress and CNS Arousal: Genomic Contributions	591
K. Midence		A. C. Ribeiro and D. W. Pfaff	
Sleep Loss, Jet Lag, and Shift Work	504	Stress Effect of Assisted Reproduction	596
E. Van Cauter		F. M. Helmerhorst, R. Sibug and E. R. de Kloet	
Sleep, Sleep Disorders, and Stress	506	Stress Effects, Overview	599
A. N. Vgontzas, S. Pejovic and M. Karataraki		A. Steptoe	
Smoking and Stress	515	Stress Generation	601
F. J. McClernon and D. G. Gilbert		J. E. Roberts and J. A. Ciesla	
Social Capital	520	Stress Hyporesponsive Period	606
I. Kawachi		S. Levine, E. R. de Kloet, G. Dent and M. S. Schmidt	
Social Networks and Social Isolation	523	Stress in University Students	612
L. F. Berkman		B. Andrews and J. Hejdenberg	
Social Status and Stress	528	Stress Induced Anovulation	615
D. de Ridder		S. L. Berga and T. L. Loucks	
Social Stress, Animal Models of	533	Stress Management and Cardiovascular Disease	631
J. Bugajski, A. Gądek-Michalska and A. J. Bugajski		D. Lane and D. Carroll	
Social Support	539	Stress Management, CAM Approach	636
T. C. Antonucci, J. E. Lansford and K. J. Ajrouch		K. Krebs	
Social Support in Trauma	542	Stress of Self Esteem	640
K. O'Donnell and A. Steptoe		J. C. Pruessner, S. Wuethrich and M. W. Baldwin	
Somatic Disorders	545	Stress System Balance Hypothesis	646
F. Creed		E. R. de Kloet	
Space, Health Risks of	548	Stress, Beneficial Effects of	650
V. A. Convertino		S. Joseph and P. A. Linley	
Spinal Cord Injury, Physical Stress of	554	Stress, Definitions and Concepts of	653
S. Bhatia and M. Quigley		B. S. McEwen	
Spinal Cord Injury, Psychological Stress of	557		
S. D. Schnakenberg-Ott and M. R. Lovell			

Stress, Insulin Resistance, and Type II Diabetes	654	Torture	749
C. Tsigos and I. Kyrou		I. Genefke, H. Marcussen and O. V. Rasmussen	
Stress, NPY, and Cardiovascular Diseases	660	Transport-Related Stress	756
Z. Zukowska		R. G. Smart	
Suicide Terrorism, Genesis of	667	Trans-sexualism	760
A. Speckhard		R. A. Allison	
Suicide, Biology of	677	Trauma and Memory	765
M. A. Oquendo, L. Giner and J. J. Mann		B. A. van der Kolk	
Suicide, Psychology of	684	Trauma Group Therapy	767
D. Lester and R. L. Walker		J. Dwyer and L. G. Martin	
Suicide, Sociology of	689	Traumatic Stress and Posttraumatic Stress Disorder, the Israeli Experience	771
D. Lester and R. M. Fernquist		E. Klein and J. Zohar	
Surgery and Stress	693	Trier Social Stress Test	776
C. Vögele		B. M. Kudielka, S. Wüst, C. Kirschbaum and D. H. Hellhammer	
Survivor Guilt	695	Type A Personality, Type B Personality	782
P. Valent		W. S. Shaw and J. E. Dimsdale	
Sympathetic Nervous System	697		
D. S. Goldstein			
Synthetic Glucocorticoids	704		
A. M. Karssen and E. R. de Kloet			
Systemic Lupus Erythematosus	708		
D. J. Wallace			
		U	
		Ulceration, Gastric	787
		R. Murison and A. M. Milde	
		Ultradian Rhythms	791
		S. L. Lightman	
		Understimulation/Boredom	794
		V. J. Sutherland	
		Unemployment, Stress, and Health	797
		M. Bartley	
		Urocortins	804
		É. M. Fekete and E. P. Zorrilla	
		V	
		Vaccination	813
		V. E. Burns, A. C. Phillips and K. M. Edwards	
		Vasoactive Peptides	817
		W. K. Samson and M. M. White	
		Vasopressin	824
		L. P. Renaud	
		Vietnam Veterans, Postwar Experiences and Health Outcomes	830
		J. A. Boscarino	
Teaching and Stress	713		
E. R. Greenglass			
Temperature Effects	717		
J. Roth			
Terrorism	719		
P. Bell			
Thermal Stress	723		
C. M. Blatteis			
Thermotolerance, Thermoresistance, and Thermosensitivity	726		
S. Gatti McArthur, D. Alberati and T. Bartfai			
Three Mile Island, Stress Effects of	735		
A. L. Dougall and A. Baum			
Thymus	738		
M. S. Vacchio			
Thyroid Hormones	743		
T. J. Visser and E. Fliers			

Violence	838	War, Suicide and Sacrifice	860
E. K. Englander		V. Hazboun	
Viral Virulence and Stress	842	War-Related Posttraumatic Stress Disorder, Treatment of	865
A. J. L. Macario and E. Conway de Macario		L. B. Slone and M. J. Friedman	
Viruses and Stress	850	Work-Family Balance	868
G. Z. Feuerstein, R. R. Ruffolo and B.-N. David		J. G. Grzywacz and A. B. Butler	
		Workplace Stress	871
		U. Lundberg	

W

Waist-Hip Ratio	853
R. Rosmond	
War Stress in the Former Yugoslavia	855
M. Flögel, S. Šupraha Goreta and G. Lauc	

VOLUME 4

<i>Contributors</i>	1
<i>Subject Index</i>	21

CONTENTS BY SUBJECT AREA

ANIMAL STUDIES AND MODELS

- Animal Models (Nonprimate) for Human Stress
- Circadian Rhythms, Effects of Prenatal Stress in Rodents
- Fish, Stress in
- Immobilization Stress
- Pre-pulse inhibition
- Primate Hierarchies and Personality
- Primate Models, Behavioral-Immunological Interactions
- Primate Models, Cardiovascular Disease
- Primate Models, Overview
- Primates: Rearing and Effects of Stress on Primate CNS Function
- Sex-Specific Effects of Early Social Stress in Mammals: a Study in Guinea Pigs

CONFLICT, WAR, TERRORISM

- Chemical Warfare
- Gulf War Syndrome, Psychological and Chemical Stressors
- Hiroshima Bombing, Stress Effects of
- Holocaust, Stress Effects of
- Impact of Terrorism on the Development of Mental Health Symptoms
- Korean Conflict, Stress Effects of
- Lockerbie Air Crash, Stress Effects of
- Northern Ireland, Post Traumatic Stress Disorder in
- Nuclear Warfare, Threat of
- Oklahoma City Bombing, Stress Effects of
- Persian Gulf War, Stress Effects of
- Prisoners of War
- 9/11, Religion and Stress
- Suicide Terrorism, Genesis of
- Terrorism
- Torture

- Traumatic Stress and Posttraumatic Stress Disorder; the Israeli Experience
- Vietnam Veterans, Postwar Experiences and Health Outcomes
- War Stress in the Former Yugoslavia
- War, Suicide and Sacrifice

DISASTERS

- Airline Accidents
- Chernobyl, Stress Effects of
- Disasters and Mass Violence, Public, Effects of
- Disaster Syndrome
- Earthquakes, Stress Effects of
- Floods, Stress Effects of
- Hurricane Katrina Disaster, Stress Effects of
- Motor Vehicle Accidents, Stress Effects of
- Three Mile Island, Stress Effects of

DIURNAL, SEASONAL AND ULTRADIAN RHYTHMS

- Circadian rhythm effects on cardiovascular and other stress-related events
- Circadian Rhythms, Genetics of
- Night Shiftwork
- Seasonal Changes in Stress Responses
- Seasonal Rhythms
- Sleep Loss, Jet Lag, and Shift Work
- Sleep, Sleep Disorders, and Stress
- Ultradian Rhythms

DRUGS (EFFECTS)

- Alcohol, Alcoholism and Stress:
A Psychobiological Perspective
- Drug Use and Abuse
- Interactions between Stress and Drugs of Abuse
- Opioids

DRUGS (TREATMENT)

- Antipsychotic Drugs and Stress
- Anxiolytics
- Benzodiazepines
- Beta-Adrenergic Blockers
- Pharmacological Treatments of Stress
- Selective Serotonin Reuptake Inhibitors (SSRIs)
- Stress and Anxiety: treatment with 2nd generation antipsychotic drugs

GENERAL CONCEPTS AND MODELS

- Alarm Phase and General Adaptation Syndrome
- Allostasis and Allostatic Load
- Behavior, Overview
- Conservation of Resources Theory
- Control and Stress
- Demand–Control Model
- Effort–Reward Imbalance Model
- Environmental Factors
- Evolutionary Origins and Functions of the Stress Response
- Fight-or-Flight Response
- Instinct Theory
- Life Events Scale
- Psychological Stressors, Overview
- Remodelling of neuronal networks by stress
- Selye, Hans
- Stress, Definitions and Concepts of
- Stress Effects, Overview
- Stress System Balance Hypothesis
- Stress, Beneficial Effects of

GENETICS AND GENOMICS

- Circadian clock genes as modulators of sensitivity to genotoxic stress
- Corticosteroid Receptor Genes: Functional Dissection in Mice
- Corticotropin Releasing Factor Receptor Deficiency in Mice
- Drosophila Genes and Anoxia
- Expression Profiling of Stress Responsive Gene Patterns
- Gene Environment Interactions in Early Development
- Genetic Factors and Stress
- Genetic Polymorphisms in Stress Response
- Genetic Predispositions to Stressful Conditions
- Genetic Testing and Stress
- Genetic variation of HPA axis activity and function in farm animals
- Glucocorticoid Receptor Mutant Mice as Models for Stress-Induced Affective Disorders

- Glucocorticoid Receptor Mutations and Polymorphisms
- Mineralocorticoid Receptor Polymorphisms
- Neuroticism Genetic Mapping of
- Neuroticism Response to Stress, Genetic Mapping of Mice
- Serotonin Transporter Genetic Modifications
- Strain Differences in Stress Response in Rodents
- Stress and CNS Arousal; Genomic Contributions

HUMAN COGNITION, EMOTION AND BEHAVIOR

- Adolescence
- Aggression
- Aggressive Behavior
- Aging and Psychological Stress
- Anger
- Anxiety
- Bereavement
- Burnout
- Captivity, Adaptation to
- Captivity, Recovery from
- Caregivers, Stress and
- Childbirth and Stress
- Cognition and Stress
- Combat, Acute Reactions to
- Combat Reaction, Chronic
- Combat Stress Reaction
- Concentration Camp Survivors
- Coping Skills
- Coping and Stress: a Lens and Filter Model
- Cortisol Awakening Response
- Death Anxiety
- Dental Stress
- Depersonalization: Systematic Assessment
- Diet and Stress, Non-Psychiatric
- Dissociation
- Distress
- Emergency Personnel, Stress in
- Emotional Inhibition
- Emotions: Structure and Adaptive Functions
- Fatigue and Stress
- Fear
- Fear and the Amygdala
- Firefighters, Stress in
- Grieving
- Holocaust Survivors, Experiences of
- Homosexuality, Stress and
- Hostility
- Impotence, Stress and
- Impulse Control
- Industrialized Societies
- Learning and Memory, Effects of Stress on
- Marriage

Medical Profession and Stress
 Memory Impairment
 Menopause and Stress
 Neuroimaging and Emotion
 Nightmares
 Optimism, Pessimism and Stress
 Pain
 Parenting, Stress of
 Peacekeeping
 Police, Stress in
 Prison
 Psychosomatic Medicine
 Recovery from Stress
 Refugees, Stress in
 Religion and Stress
 Self-Esteem, Stress, and Emotion
 Somatic Disorders
 Stress Generation
 Stress in University Students
 Suicide, Psychology of
 Suicide, Sociology of
 Surgery and Stress
 Survivor Guilt
 Teaching and Stress
 Trauma and Memory
 Type A Personality, Type B Personality
 Understimulation/Boredom

HUMAN HEALTH AND PHYSICAL ILLNESS

Amnesia
 Adrenal Insufficiency
 AIDS
 Alzheimer's Disease
 Arthritis
 Asthma
 Atherosclerosis
 Attention-Deficit/Hyperactivity Disorder, Stress and
 Breast Cancer
 Cancer
 Cardiovascular Disease, stress and
 Chaperonopathies
 Chronic Fatigue Syndrome
 Common Cold and Stress
 C-Reactive Protein
 Cytokines, Chronic Stress, and Fatigue
 Cushing's Syndrome, Medical Aspects
 Cushing's Syndrome, Neuropsychiatric Aspects
 Dermatological Conditions
 DHEA
 Diabetes, Type I
 Disease, Stress Induced
 Eclampsia and preeclampsia
 Endometriosis
 Enuresis

Epilepsy
 Fibromyalgia
 Graves' Disease (Thyrotoxicosis)
 Health Behavior and Stress
 Heart Disease/Attack
 Heart Failure, Stress Effects
 Hemostasis and Stress
 Hyperreactivity (Cardiovascular)
 Hypertension
 Hyperthermia
 Hyperthyroidism
 Hyperventilation
 Hypoglycemia
 Hypotension, Hypovolemia, and Septic Shock
 Hypothermia
 Hypothyroidism
 Hysteria
 Metastasis
 Metastasis
 Migraine
 Multiple Trauma
 Nelson's Syndrome
 Neurodegenerative Disorders
 Organ Transplantation, Stress of
 Parkinson's Disease
 Perinatal Dexamethasone
 Pregnancy, Maternal and Perinatal Stress, Effects of
 Premenstrual Dysphoric Syndrome
 Prolactin and stress
 Psychosomatic heart disease: role of sympathetic
 and sympathoadrenal processes
 Rheumatic Disorders
 Sickle Cell Disease and Stress
 Space, Health Risks of
 Spinal Cord Injury, Physical Stress of
 Spinal Cord Injury, Psychological Stress of
 Stress Effect of Assisted Reproduction
 Stress Induced Anovulation
 Stress, Insulin Resistance and Type II Diabetes
 Stress, NPY and Cardiovascular Diseases
 Ulceration, Gastric

HUMAN MENTAL HEALTH AND PSYCHOPATHOLOGY

Acute Stress Disorder and Posttraumatic Stress
 Disorder
 Adjustment Disorders
 Adolescent suicide
 Affective Disorders
 Antisocial Disorders
 Anxiety
 Arthritis – Psychological
 Borderline Personality Disorder
 Child Abuse
 Child Sexual Abuse

Comorbid Disorders and Stress
Corticotropin-Releasing Factor Circuitry in the
Brain – Relevance for Affective Disorders
and Anxiety
Defensive Behaviors
Depression and Coronary Heart Disease
Depression and Manic-Depressive illness
Depression and Stress, Role of n-3 and n-6
Fatty Acids
Depression, Immunological Aspects
Depression Models
Diet and Stress, Psychiatric
Disaster Syndrome
Dissociation
Eating Disorders and Stress
Elder Abuse
Ethnicity, Mental Health
HPA alterations in PTSD
Hypocortisolism and Stress
Incest
Learned Helplessness
Life Events and Health
Loss Trauma
Male Partner Violence
Major Depressive Disorder
Multiple Personality Disorder
Negative Affect
Neurodevelopmental Disorders in Children
Neurosis
Obsessive–Compulsive Disorder
Panic Disorder and Agoraphobia
Paranoia
Posttraumatic Stress Disorder – Clinical
Posttraumatic Stress Disorder, Delayed
Posttraumatic Stress Disorder in Children
Posttraumatic Stress Disorder – neurobiological
basis for
Posttraumatic Stress Disorder,
Neurobiology of
Psychotic Disorders
Revenge Fantasies
Schizophrenia
Sexual Assault
Sexual Offenders
Social support in trauma
Stress of Self Esteem
Suicide, Biology of
Trans-sexualism
Violence

IMMUNOLOGY, INFECTION AND INFLAMMATION

Antibody Response
Autoimmunity

Corticotropin-releasing factor (CRF) family of
neuropeptides – role in inflammation
Cytokines
Cytokines, Stress, and Depression
Cytotoxic Lymphocytes
Herpesviruses
HIV Infection/AIDS
Immune Cell Distribution, Effects of Stress on
Immune Function, Stress-Induced Enhancement of
Immune Response
Immune Suppression
Immune Surveillance – Cancer, Effects of Stress on
Immune System, Aging
Immunity
Infection
Inflammation
Leishmania, Stress Response in
Leukocyte Trafficking and Stress
Lymphocytes
Mucosal Secretory Immunity, Stress and
Multiple Sclerosis
Natural Killer (NK) Cells
Neuroimmunomodulation
Neuroinflammation
Psoriasis
Psychoneuroimmunology
Sepsis, Acute Respiratory Distress Syndrome and
Glucocorticoid Resistance
Systemic Lupus Erythematosus
Vaccination
Viral Virulence and Stress
Viruses and Stress

LABORATORY STUDIES AND TESTS

Antimineralocorticoid Challenge
Dex-CRH Test
Dexamethasone Suppression Test (DST)
Mental Stress Testing
Metyrapone: Basic and Clinical Studies
Trier Social Stress Test
Waist–Hip Ratio

THERAPIES

Aerobics, Exercise and Stress Reduction
Behavior Therapy
Cancer Treatment
Desensitization
Exercise
Family Therapy
Group Therapy
Hypnosis
Integrative Medicine (Complementary and
Alternative Medicine)

Meditation and Stress
 Problem-Solving Skills Training
 Psychosomatic Medicine
 Reenactment Techniques
 Relaxation Techniques
 Stress Management and Cardiovascular Disease
 Stress Management, CAM Approach
 Trauma Group Therapy

PHYSIOLOGICAL, PHARMACOLOGICAL AND BIOCHEMICAL ASPECTS

Acute Stress Response: Experimental
 Acute Trauma Response
 Adenylyl Cyclases and Stress Response
 Adrenal Cortex
 Adrenaline
 Adrenal Medulla
 Adrenocorticotrophic Hormone (ACTH)
 Adrenocortical Function, factors controlling development thereof
 Aging and Stress, Biology of
 Aging and Adrenocortical Factors
 Alcohol and Stress: Social and Psychological Aspects
 Aldosterone and Mineralocorticoid Receptors
 Ambulatory Blood Pressure Monitoring
 Amenorrhea
 Amygdala
 Anatomy of the HPA Axis
 Androgen Action
 Angiotensin
 Angiotensin receptors
 Annexin
 Anti-CRF
 Antidepressant Actions on Glucocorticoid Receptors
 Apoptosis
 Arterial Baroreflex
 Autonomic Nervous System
 Autotolerance
 Avoidance
 Beta-Endorphin
 Blood–Brain Barrier, Stress and
 Blood Pressure
 Brain and Brain Regions
 Brain Natriuretic Peptide (BNP)
 Brain Trauma
 Calbindin
 Calcium-Dependent Neurotoxicity
 Calcium, Role of
 Cardiovascular System and Stress
 Catecholamines
 Central Stress Neurocircuits
 Cerebral Metabolism, Brain Imaging

Chaperone Proteins and Chaperonopathies
 Cholesterol and Lipoproteins
 Comparative Anatomy and Physiology
 Congenital Adrenal Hyperplasia (CAH)
 Corticosteroid-Binding Globulin (Transcortin)
 Corticosteroid Receptors
 Corticosteroids and Stress
 Corticotropin Releasing Factor (CRF)
 Corticotropin Releasing Factor-Binding Protein
 Corticotropin Releasing Factor Receptors
 Critical Thermal Limits
 Dopamine, Central
 Drosophila Studies
 Electrodermal Activity
 Endocrine Systems
 Estrogen
 Ethanol and Endogenous Opioids
 Excitatory Amino Acids
 Excitotoxins
 Familial Patterns of Stress
 Febrile Response
 Feedback Systems
 Fibrinogen and Clotting Factors
 Feeding Circuitry (and Neurochemistry)
 Fetal Stress
 Food Intake and Stress, Human
 Food Intake and Stress, Non-Human
 Food Shift Effect
 GABA (Gamma Aminobutyric Acid)
 Gastrointestinal Effects
 Gender and Stress
 Ghrelin and Stress Protection
 Glia or Neuroglia
 Glucocorticoid Negative Feedback
 Glucocorticoid Effects on Memory: the Positive and Negative
 Glucocorticoids – Adverse Effects on the Nervous System
 Glucocorticoids, Effects of Stress on
 Glucocorticoids, Overview
 Glucocorticoids, Role in Stress
 Glucose Transport
 Glycobiology of Stress
 Gonadotropin Secretion, Effects of Stress on
 Heart Rate
 Heat Resistance
 Heat Shock Genes, Human
 Heat Shock Proteins: HSP60 Family Genes
 Heat Shock Response, Overview
 Hippocampal Neurons
 Hippocampus, Corticosteroid Effects on
 Hippocampus, Overview
 Homeostasis
 11 β -Hydroxysteroid Dehydrogenases
 Hypothalamic-Pituitary-Adrenal Axis

Instinct Theory
Kidney Function
Left Ventricular Mass
Leptin, Adiponectin, Resistin, Ghrelin
Lymph Nodes
Macrophage Antimycobacterial Activity, Effects of
Stress on
Macrophages
Maternal Deprivation
Membrane Glucocorticoid Receptors
Memory and Stress
Menstrual Cycles and Stress
Metabolic Syndrome
Metals, Oxidative Stress and Brain Biology
Mitochondria
Monoamine Oxidase
Multi Drug Resistance P Glycoprotein and other
Transporters
Musculoskeletal Problems and Stress
Myopathy
Neural Stem Cells
Neuroendocrine Systems
Neurogenesis
Neuropeptide Y
Neuropeptides, Stress-Related
Nitric Oxide
Nutrition
Obesity, Stress and
Orexin
Oxidative Stress
Oxidative Stress and Acidosis, Molecular
Responses to
Oxidative Stress and Aging
Oxytocin
Paraventricular Nucleus
Peptides
Personality Processes
Pheromones
Pituitary Regulation, Role of
Pressure, Effects of Extreme High and Low
Pro-opiomelanocortin (POMC)
Prostaglandins
Proteases in Prokaryotes and Eukaryotic Cell
Organelles
Proteases in the Eukaryotic Cell Cytosol
Protein Synthesis
Proteosome
Reductive Stress
Regional Blood Flow, Stress Effects
Renal and Adrenocortical Effects of
Dopamine
Reproduction, Effects of Social Stress on
Reproductive Dysfunction in Primates,
Behaviorally Induced

Resistance
Restraint Stress
Salivary Cortisol
Salt Appetite
Secretagogue
Serotonin
Serotonin in Stress
Sex Differences in Human Stress Response
Sex Steroids, Response to Stress and Susceptibility
to Depression
Sexual Dysfunction
Smoking and Stress
Startle Response
Steroid Hormone Receptors
Steroid Hydroxylases
Stress Hyporesponsive Period
Sympathetic Nervous System
Synthetic Glucocorticoids
Temperature Effects
Thermal Stress
Thermotolerance, Thermoresistance, and
Thermosensitivity
Thymus
Thyroid Hormones
Urocortin
Vasoactive Peptides
Vasopressin

PSYCHOLOGICAL THERAPY

Cognitive Behavioral Therapy
Freud, Sigmund
Posttraumatic Therapy
Psychoanalysis
Psychotherapy
War-Related Posttraumatic Stress Disorder,
Treatment of

PSYCHOSOCIAL AND SOCIOECONOMIC ASPECTS

Childhood Stress
Child Physical Abuse
Chronic Social Stress: GR Sensitivity
in Leukocytes
Community Studies
Crowding Stress
Crime Victims
Crisis Intervention
Cultural Factors in Stress
Cultural Transition
Divorce, Children of
Domestic Violence
Economic Factors and Stress
Education Levels and Stress

Employee Assistance and Counseling
 Environmental Stress, Effects on Human
 Performance
 Health and Socioeconomic Status
 Income Levels and Stress
 Indigenous Societies
 Job Insecurity; the health effects of a psychosocial
 work stressor
 Marital Conflict
 Marital Status and Health Problems
 Minorities and Stress
 Neighborhood Stress and Health
 Psychosocial Factors and Stress

Quality of Life
 Racial Harassment/Discrimination
 School Violence and Bullying
 School Stress and School Refusal Behavior
 Social Capital
 Social Networks and Social Isolation
 Social Status and Stress
 Social Stress, Animal Models of
 Social Support
 Transport-Related Stress
 Unemployment, Stress and Health
 Work–Family Balance
 Workplace Stress

PREFACE TO FIRST EDITION

“Stress” remains one of the most frequently used but ill-defined words in the English language. Stress is a phenomenon that has quite different meanings for the politician, social scientist, physician, nurse, psychotherapist, physiologist, molecular biologist, and perhaps you and me. This diversity of meanings was one impetus for creating the *Encyclopedia of Stress*, the aim being to derive a definition of stress from a variety of expert descriptions. The second impetus was the obvious need for an up-to-date compendium on one of the most important social, medical, and psychological phenomena of our age. We were fortunate in attracting stars for our Editorial Board and a set of most distinguished contributors for the 400 or so entries—indeed, the list of contributors is a Who’s Who in stress research.

We anticipate that the diversity of our readers will equal the diversity of the topics covered. They will find that the coverage of the Encyclopedia extends well beyond the general adaptation theory of Hans (Janos) Selye and the fight-or-flight response of Walter Cannon. Nevertheless, the general principles annunciated by these two great pioneers in the field still underpin our understanding of the biology of the stress phenomenon. That is, stress is a real or perceived challenge, either endogenous or exogenous, that perturbs body equilibrium or “homeostasis.” The stressor may range from overcrowding, traffic congestion, violence, bereavement, redundancy, or unemployment to physical, chemical, biological, or psychological insults. Whether the person can adapt to or cope with the stress will depend on the nature and severity of the stressor and the person’s physical and mental state, which in turn depends on genetic, experiential, social, and environmental factors. These issues are discussed in depth in the Encyclopedia, as are the mechanisms of coping and the impact of stress on health and predisposition to diseases such as

cancer, infection, rheumatoid arthritis, heart disease, high blood pressure, and mental disorder.

Aggression remains a hallmark of human behavior, even as we move into the third millennium, and so the Encyclopedia covers several topical areas that have only recently been analyzed systematically. These include war and specific wars, posttraumatic stress disorder (PTSD; formerly thought of vaguely as “shell shock”), rape, torture, marital discord and spousal abuse, and the Holocaust. In tackling these topics we accept that our entries may not include all the nuances that are necessary for a full understanding of what these phenomena are all about and described so graphically and sensitively in Tolstoy’s *War and Peace* or Pat Barker’s monumental *Regeneration* trilogy on the horrific psychological traumas of the First World War. Nevertheless, an important start has been made in that we now accept that PTSD is not just a lack of “bottle” (courage or “guts”), but rather a syndrome that needs to be and can be understood within the framework of medicine and psychology.

Biologically, the stress response reflects a set of integrated cascades in the nervous, endocrine, and immune defense systems. As in most areas of biology, molecular genetics has made a significant difference in the precision with which we now understand the physiopathological processes of the stress response. And so the adage formerly applied to diabetes mellitus may now apply equally to stress: “Understand stress and you will understand medicine.”

In summary, we hope that this first *Encyclopedia of Stress* will indeed define the phenomenon and at the same time provide a valuable source of information on a phenomenon that affects us all. In setting out on this adventure we were aware that there is nothing new under the sun and that stress has been around since the first biological particles, bacteria, or even viruses competed for the same mechanisms of replication.

There is a tendency for each generation to imagine that stress and its untoward effects are uniquely harsh for them, but it is not this misconception that underlies this work. Rather, the stress of “stress” itself—the massive accumulation of knowledge—made it seem propitious to bring the information together in a systematic manner that allows ready access to all who need or wish to understand the phenomenon.

The idea of producing this Encyclopedia was conceived at an Academic Press reception in San Diego held in conjunction with the annual meeting of the Society for Neuroscience in 1996. I am deeply indebted to Erika Conner for enabling conversion of the idea to a concept and then a project and to

Jennifer Wrenn, Christopher Morris, and Carolan Gladden, all of Academic Press, for their enthusiasm, encouragement, and herculean efforts that converted the concept into a reality. For giving generously of their intellect, expertise, sound advice, and unstinting work, I am greatly indebted to my friends and colleagues on the Editorial Board, who made the project such a satisfying experience. To all our contributors go our profound thanks for taking time out from wall-to-wall schedules to produce their entries, which together have made this Encyclopedia.

George Fink
August 1999

PREFACE TO THE SECOND EDITION

The popularity of the first edition of the Encyclopedia of Stress, the several major stressful conflicts that have occurred since 1999, and the significant advances in our knowledge of stress prompted the production of this second edition. In addition to more than 140 new articles, most of the original 400 first edition articles have been updated for the second edition to reflect, for example, advances in our understanding of corticotropin releasing factor (CRF) and urocortin receptors, ideas about the role of central CRF in the overall control of the psychological/mental as well as neuroendocrine response to stress, and the novel discoveries of the complex neuropeptide regulation of hunger and satiety. The effects of maternal and perinatal stress on subsequent body development and function in the adult, the adverse effects of stress on brain (and particularly hippocampal memory function), the role of the amygdala in fear and aggression, and the role of stress in the etiology of obesity and the metabolic syndrome all feature prominently in this new edition. 9/11 and the many new conflicts since 1999 have engendered new articles on terrorism, suicide bombers and their effects, as well as a retrospective on previous human apocalypses such as Hiroshima. Coverage of the whole field of post traumatic stress disorder (PTSD) has been expanded, reflecting the greater recognition and understanding of PTSD, and possibly related disorders such as *combat fatigue* and *burnout*. Stress and depression and anxiety, already touched on in the first edition, now feature prominently. On the conceptual side, *allostasis* (modified from Hans Selye's *heterostasis*), or stability through centrally regulated change in physiological set points, receives greater emphasis than it did in the first edition.

Overall, our understanding of stress mechanisms in the human has benefited from two major quantum leaps in technology. First, the advances in molecular genetics and genomics and sequencing of the human

genome (published after the first edition had appeared) have increased the rigor and precision of our understanding of the molecular neurobiology of stress. Thus, new to the second edition are a series of articles on the genetic factors that play a role in susceptibility to stress and in various components of the stress response. Secondly, functional brain imaging has enhanced our understanding of the neurobiology of stress in the human. Many of the articles in the second edition reflect the positive impact of these two advances.

Accessibility to the second edition of the Encyclopedia of Stress will be facilitated by virtue of the fact that it is published online as well as in hard copy. My profound thanks for their selfless work go to our team of distinguished authors and to my friends and colleagues on the Editorial Board. Production of the Encyclopedia would not have been possible without our colleagues at Elsevier, and in particular Bob Donaldson and Hilary Rowe whose dedication, commitment and excellent work I am pleased to acknowledge.

George Fink
November 2006

Note on Terminology of Corticotropin Releasing Factor/Hormone and the Catecholamines

The central nervous regulation of the anterior pituitary gland is mediated by substances, mainly peptides, which are synthesized in the hypothalamus and transported to the gland by the hypophysial portal vessels. Because these compounds are transported by the blood, the term "hormone" gained acceptance in the neuroendocrine literature. The major hypothalamic peptide involved in the stress response is the

41-amino acid corticotropin releasing factor (CRF). The Endocrine Society (USA), following convention, adopted the term corticotropin releasing hormone (CRH). However, this nomenclature has been challenged. Hauger et al. (*Pharmacol Rev* 55: 21–26, 2003), in liaison with the International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification, argued that the CRF's function extends well beyond the biology of a hormone, and that therefore it should be termed corticotropin releasing factor (CRF) rather than hormone. Since the terminology of CRF versus CRH has yet to be resolved, the two terms and abbreviations are here used interchangeably, depending on author preference.

Adrenaline and *noradrenaline* are catecholamines that play a pivotal role in the stress response. These

terms are synonymous with *epinephrine* and *norepinephrine*, respectively. Both sets of terms are used interchangeably in the endocrine, neuroendocrine and stress literature, and this principle has been adopted for the Encyclopedia. Style has depended on author preference, but wherever possible within-article consistency has been ensured.

George Fink
November 2006

Hauger, R. L., Grigoriadis, D. E., Dallman, M. F., Plotsky, P. M., Vale, W. W., and Dautzenberg, F. M. (2003). International Union of Pharmacology. XXXVI. Current status of the nomenclature for receptors for corticotropin-releasing factor and their ligands. *Pharmacological Reviews* 55(1), 21–26.

FOREWORD

This is the second edition of the *Encyclopedia of Stress*, the three heavy volumes that first appeared in 2000 under the overall editorial guidance of George Fink. From its original coining in reference to rather unusual and unpleasant situations, the word *stress* is now quasi synonymous with *life*, i.e. modern life! This second edition of the *Encyclopedia of Stress* presents several hundreds of articles written by experts in that particular field, especially for this series, discussing, clarifying, explaining to the average-reader as well as to anyone more educated in one field or another, the extraordinary involvement of stress and our normal or abnormal reactions to it, both in our physical body and its accompanying mind. The table of contents is truly amazing, from the stress of being born to the stress of dying and so many circumstances in between, with clear presentations of the pertinent, current knowledge about them. This is truly an encyclopedia where, surely, you will find some explanation and answer to your own stresses (and those of others).



Roger Guillemin, MD, PhD
Nobel laureate in Physiology & Medicine



Obesity, Stress and

R Rosmond

Björndammsterrassen 41, Partille, Sweden

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Björntorp, R Rosmond and J Uddén, volume 3, pp 75–76, © 2000, Elsevier Inc.

Stress-Eating
Stress-Eating and Obesity
Mechanism of Stress-Eating

Glossary

<i>Hypothalamic-pituitary-adrenal axis</i>	The hormonostatic endocrine axis that regulates cortisol production and secretion.
<i>Leptin</i>	An adipocyte-produced hormone that is involved in the regulation of satiety and energy expenditure.

The current understanding of the systems of the body and brain related to the control of energy intake, energy balance, and body weight has increased significantly since 2001. In the periphery, a variety of substances are believed to be involved in integrating physiological and behavioral systems directed towards energy homeostasis. The hormones insulin, leptin (secreted by adipocytes), and ghrelin (mainly secreted by the stomach) link the control of peripheral energy metabolism to the central nervous system (CNS) feeding behavior integrating unit by modulating short-term signals that determine meal initiation and termination as well as energy balance. The adrenal glucocorticoids (cortisol) predominantly influence the ingestion and metabolism of fat and carbohydrates. In addition to these hormones, the process of integrating metabolic information from the periphery in the CNS requires specialized functions of multiple brain areas. These include the lower brainstem, which integrates neural information between peripheral autonomic-endocrine organs and forebrain structures,

the hypothalamus, which is closely linked to circulating metabolites and hormones, and the forebrain structures, such as the nucleus accumbens and amygdala, which integrate incoming information with various cognitive aspects of food.

The role of the hypothalamus in the process of relating hormones and metabolism to eating behavior has received considerable attention. A number of neuroendocrine systems have been identified in this structure and are believed to be involved in controlling appetite for the macronutrients, carbohydrate, fat, and protein. They also modulate metabolism and contribute ultimately to weight gain and obesity.

Stress-Eating

Many overweight people, particularly those with abdominal obesity, find that eating increases in situation of discomfort and also soothes their feelings of stress and anxiety. This phenomenon is usually called stress-eating or emotional-eating. While obese individuals may realize that in general they eat in response to stress/emotions, they are typically unable in the moment to stop themselves from eating. As will be discussed below, this type of stress-eating might be due to disruptions of central regulatory systems involved in the control of energy intake. In obese binge eaters, stress-eating occurs, and it is up to 100% specifically triggered by stress and negative moods (e.g., depression, anger, anxiety, etc.).

Stress-Eating and Obesity

It is not clear if stress-eating is secondary to being obese or vice versa, stress causing overeating, resulting in obesity. It is, however, clear that obese subjects have more body image distortion and distress about their weight than normal-weight subjects. Indeed, in our society being obese is so stigmatized that few individuals who are obese are able to feel good about their size. Negative body image and feelings of stress and anxiety about being obese may thus

serve as a trigger for stress-eating. In addition, stress and anxiety about being obese may work indirectly, increasing negative emotions such as depression and thus rendering the individual more susceptible to eating in response to stress. However, no studies have been reported indicating that obese subjects overeat by stress mechanisms leading to a maintained or escalated obese state.

Our knowledge of human obesity has progressed beyond the simple generalization that obesity is fully explained by inappropriate eating. Still, in more than 80% of the cross-sectional studies linking intake data to degree of obesity, dietary fat intake was associated with obesity. It is, however, important to recognize that human obesity is not a single entity, and subgroups are clearly more or less afflicted by stress. Realizing this distinction, it may be possible to examine how stress may be directly associated with an increased energy intake, and eventually with obesity.

Obesity with an increased number of adipocytes (i.e., hypercellular obesity) typically begins in early or middle childhood. In contrast, obesity with enlarged adipocytes (i.e., hypertrophic obesity) tends to correlate with truncal fat distribution, which in turn is an important predictor of the health hazards of obesity. Certainly, differences in clinical symptomatology favor a separation of obesity into two broad categories: visceral and nonvisceral obesity. These two forms seem to have different associations to stress. Central, visceral obesity is often associated with psychosocial and socioeconomic handicaps and excess use of alcohol and tobacco, as well as traits of depression and anxiety. Psychosocial stressors might induce eating as comfort food that may directly result in reduction of the negative effects of the chronic stressor in the nucleus accumbens. In addition, socioeconomic difficulties may make it hard to purchase nutritionally healthy food.

Mechanism of Stress-Eating

The expression of appetite is underpinned by a complex set of peripherally generated signals. Key hormonal appetite factors include insulin, cortisol, and leptin. Leptin is a protein that is synthesized exclusively in adipose tissue and moderates the drive to consume food. Within the neural circuitry numerous neuropeptides, both orexogenic (e.g., neuropeptide Y and melanocortin), and anorexigenic (e.g., pro-opiomelanocortin and corticotropin-releasing hormone) have been identified which stimulate and inhibit food intake.

Resistance to the appetite-suppressing effects of leptin is associated with common forms of obesity. Is

it possible that stress factors associated with central, visceral obesity might be interfering with the leptin system? The human body holds many effector systems including the hypothalamic-pituitary-adrenal (HPA) axis for maintaining homeostasis. The secretory end product of the HPA axis, cortisol, is kept within an optimal range through the feedback action of cortisol interacting with neural control mechanisms. Distressing events or situations evokes prominent HPA system activation, and after long-term exposure, the HPA axis will eventually become dis-habituated, resulting in a disruption of central regulatory systems. Together with insulin, the cortisol stimulate drive for and ingestion of comfort foods that may directly result in reduction of the negative effects of the chronic stressor in the nucleus accumbens, through stimulation of the anterior, more pleasure-associated part of this cell group, thus reducing the weight of the stress-stimulated posterior, more defensive part. Studies in rats have indicated that cortisol is directly involved in the sensitivity of the leptin regulation of satiety. When rats are adrenalectomized, their leptin sensitivity becomes high and feeding is inhibited. When corticosterone is replaced in a dose-dependent manner, food intake and body weight follow in parallel, and with excess corticosterone, obesity due to overeating appears. Furthermore, most animal models of obesity have enlarged adrenals and overproduction of glucocorticoids. Upon adrenalectomy, the somatic feature of obesity disappears. In normal-weight subjects, cortisol induces a dose-dependent increase in plasma leptin concentrations. In obese humans, food intake is clearly elevated after glucocorticoid administration in spite of increased levels of circulating leptin, indicating the presence of leptin resistance. Taken together it seems reasonable to assume that stress-eating is primarily caused by interference of stress-related cortisol secretion, with the normal signaling of satiety by the leptin system. The resulting deficient satiety signal causes overeating and finally the appearance of overweight and obesity.

Further Reading

- Björntorp, P. and Rosmond, R. (2000). Obesity and cortisol. *Nutrition* **16**, 924–936.
- Dallman, M. F., Pecoraro, N. C. and la Fleur, S. E. (2005). Chronic stress and comfort foods: self-medication and abdominal obesity. *Brain, behavior, and immunity* **19**, 275–280.
- Gautier, J. F., Chen, K., Salbe, A. D., et al. (2000). Differential brain responses to satiation in obese and lean men. *Diabetes* **49**, 838–846.
- Mastorakos, G. and Zapanti, E. (2004). The hypothalamic-pituitary-adrenal axis in the neuroendocrine regulation

of food intake and obesity: the role of corticotropin releasing hormone. *Nutritional Neuroscience* 7, 271–280.

Rosmond, R., Dallman, M. F. and Björntorp, P. (1998). Stress-related cortisol secretion in men: relationships with abdominal obesity and endocrine, metabolic and hemodynamic abnormalities. *Journal of Clinical Endocrinology and Metabolism* 83, 1853–1859.

Rosmond, R. (2005). Role of stress in the pathogenesis of the metabolic syndrome. *Psychoneuroendocrinology* 30, 1–10.

Udden, J., Björntorp, P., Arner, P., et al. (2003). Effects of glucocorticoids on leptin levels and eating behaviour in women. *Journal of Internal Medicine* 253, 225–231.

Zakrzewska, K. E., Cusin, I., Sainsbury, A., et al. (1997). Glucocorticoids as counterregulatory hormones of leptin: toward an understanding of leptin resistance. *Diabetes* 46, 717–719.

Obsessive–Compulsive Disorder

R T Rubin

VA Greater Los Angeles Healthcare System,
Los Angeles, CA, USA

B J Carroll

Pacific Behavioral Research Foundation, Carmel,
CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition
article by R T Rubin, volume 3, pp 77–82,

© 2000, Elsevier Inc.

Introduction

Clinical Characteristics of OCD

Biological Characteristics of OCD

Brain Circuit Theories of OCD

Neuropsychological Testing in OCD

Neuroimaging Studies in OCD

Conclusions

Glossary

<i>Anti-depressants</i>	Drugs that are used in the treatment of depression and other psychiatric disorders, including obsessive–compulsive disorder (OCD).
<i>Basal ganglia</i>	Groups of neurons deep in the brain that serve as integrative way stations for brain circuits.
<i>Cerebral cortex</i>	The outer layer of the brain, composed mainly of nerve cells (neurons).
<i>Compulsions</i>	Repetitive behaviors or mental acts that reduce anxiety or distress.
<i>Neuro-transmitters</i>	Chemicals that carry messages from one nerve cell to another.
<i>Obsessions</i>	Intrusive and inappropriate recurrent and persistent thoughts.

Introduction

Psychiatric disorders are considered functional when there is no clearly understood central nervous system (CNS) pathophysiology underlying them. Thus, schizophrenia, the major affective illnesses, and anxiety disorders have been and are still considered functional psychiatric illnesses. This is an evolving nomenclature, however – as our knowledge about the pathophysiology of the brain in these disorders increases, the adjective functional conveys less and less meaning.

The brain substrates of the major psychiatric disorders are being examined in several ways. Complex genetic underpinnings are slowly being defined. Structural and functional imaging studies of the brain are providing some anatomical localization, limited by the resolution of the imaging techniques and the particular methodology of each study. Specific classes of drugs are known to be therapeutic for specific disorders, but the balance among chemical systems in the brain makes inferences about disturbance of a specific system in a particular psychiatric disorder difficult. Furthermore, no drug is curative of any psychiatric disorder. Obsessive–compulsive disorder (OCD) is one of the functional psychiatric illnesses that is giving way to an understanding of its pathophysiology in the brain. Stress can play a major role in the severity of OCD, and the illness itself can be very stressful to afflicted persons.

Clinical Characteristics of OCD

OCD is classified among the anxiety disorders. Its hallmarks are recurrent obsessions and/or compulsions severe enough to be time consuming (more than 1 h/day) or to interfere with occupational or social functioning. At some point, the individual must

recognize that these symptoms are excessive or unreasonable (this does not apply to children), and, if another major psychiatric disorder is present, the content of the obsessions or compulsions must be unrelated to it. Finally, the OCD must not be a direct result of a drug or an underlying medical condition.

Obsessions are recurrent and persistent thoughts, impulses, or mental images that are experienced as intrusive and inappropriate, that cause marked anxiety and distress, and that are not simply excessive worries about real-life problems. The individual recognizes that the mental intrusions are a product of his or her own mind (this distinguishes them from psychotic symptoms such as thought insertion) and attempts to neutralize them with other thoughts or actions (for example, compulsive rituals). The content of obsessions is most often violent (for example, the murder of one's child), sexual, contaminative (for example, touching unclean objects), or doubting (for example, worrying repeatedly about having performed a particular act).

Compulsions are repetitive behaviors or mental acts engaged in to reduce anxiety or distress. Common compulsive behaviors are hand washing, ordering of objects, and repetitive checking (for example, a locked door); common mental acts are repetitive counting, praying, and saying words silently. The frequency and severity of these obsessions and compulsions as necessary for a diagnosis are a minimum for these activities; not infrequently, many hours each day are spent in the throes of OCD, and compulsions may be physically damaging (for example, washing one's hands until the skin is raw). As mentioned, the individual knows these are alien thoughts and behaviors and is distressed by them, but is powerless to moderate them.

There appear to be different OCD symptom profiles across subjects. An analysis of scores on the Yale-Brown Obsessive-Compulsive Scale in more than 300 OCD patients showed that there were four relatively independent clusters of symptoms: obsessions and checking, symmetry and order, cleaning and washing, and hoarding. Other recognized clinical variants include obsessional slowness (for example, taking several hours to perform routine bathing, grooming, and dressing each day) and obsessional rumination, commonly with a hypochondriacal theme. The background personality of individuals with OCD tends to be rigid and perfectionistic. The end result of severe OCD can be impressively bizarre behavior: the housewife who must purchase new china whenever guests are invited; the adult son who cannot join his parents for the weekend because he spends 12 hours checking that his house has been properly locked; the hoarder

whose house is stuffed with trash and newspapers to the point that it is declared a fire hazard. At present, there are no laboratory tests that are diagnostic of OCD.

The lifetime prevalence of OCD is 2–3%; thus, it is not rare. OCD occurs equally in men and women and usually begins in early adulthood, although childhood OCD is not uncommon. It may co-occur with other psychiatric illnesses such as major depression, other anxiety disorders, and eating disorders. There is a high incidence of OCD (30–50%) in patients with Gilles de la Tourette syndrome, which consists of severe tics (sudden, involuntary muscle movements) and involuntary utterances, occasionally vulgar in nature. Tourette syndrome is quite rare, but approximately 5% of OCD patients may have some form of it, and 20–30% of OCD patients have a history of tics. Indeed, one formulation considers that an obsession is a psychic tic. OCD affects both twins more often in monozygotic (single-egg) twins than in dizygotic twins, and there is a higher incidence of OCD in first-degree relatives (parents, siblings, and children) of both OCD and Tourette syndrome patients than in the general population. OCD patients in stressful life situations often have a worsening of their symptoms, for example, more frequent intrusive obsessional thoughts or more frequent compulsive behaviors such as repetitive hand washing.

Biological Characteristics of OCD

Neurotransmitters (for example, serotonin, norepinephrine, and dopamine) are chemicals in the brain that are released by neurons and carry a chemical message across a short space (synapse) to activate receptors on downstream neurons. An interesting treatment finding is that OCD responds best to drugs that are serotonin uptake inhibitors (SUIs) in the brain. These drugs (for example, clomipramine, fluoxetine, paroxetine, sertraline, and fluvoxamine), block the serotonin transporter, which recycles the neurotransmitter serotonin from the synapse back into the nerve cell that released it. By blocking transporter uptake, these drugs increase synaptic serotonin concentrations, at least in the short term. This implies that reduced serotonin neurotransmission may be a neurochemical basis for OCD, but clear evidence of this, by measurement of serotonin or its metabolites, transport mechanisms, or receptors, has not yet been forthcoming.

Complicating this perspective is recent evidence that chronic treatment with an SUI completely blocks the rise of extracellular brain serotonin that occurs in response to stress. Serotonin is mainly an

inhibitory neurotransmitter in the brain, especially in areas that are implicated in OCD, such as orbital prefrontal cortex and hippocampus. In these sites, serotonin reduces the firing of cortical pyramidal cells that are essential components of the cortico-striato-thalamo-cortical circuits now implicated in OCD and major depression. SUIs are most often used as antidepressants, and it is interesting that OCD and major depression, two illnesses with different symptom patterns (although occurring together in some patients), respond to treatment with the same drugs. The common factor is now considered to be that these disorders have a similar neuropathology that is expressed in functionally segregated neuronal circuits in the brain. Relief of OCD symptoms by SUIs is not complete (about 35% on average), and only about 50% of patients respond, suggesting that more than one brain neurotransmitter system may be involved. The addition of antipsychotic drugs (for example, haloperidol) that block dopamine receptors on neurons, thereby reducing dopamine neurotransmission, can be effective, certainly when there is a tic disorder associated with the OCD, and perhaps in cases refractory to SUI treatment. Other drugs useful in selected cases as adjuncts to treatment with SUIs are lithium (used most commonly to treat manic-depressive illness) and fenfluramine, a direct stimulator of serotonin receptors. Interestingly, dextroamphetamine, a dopamine and norepinephrine neurotransmission-potentiating drug, can reduce OCD symptoms in some patients.

Another aspect of OCD that supports an underlying brain abnormality is the presence of OCD-like disorders in other mammalian species, including sub-human primates, dogs, cats, horses, pigs, cows, sheep, bears, and parrots. Compulsive behaviors in these species include repetitive grooming to the point of skin damage; hair and feather pulling; tail chasing, pacing, whirling, bouncing, and somersaulting; sucking, bar biting, and lip flapping; compulsive sexual behaviors; aggression; checking; and territorial marking. Some equine and canine compulsive behaviors are familial and breed-specific, providing support for genetic factors in these conditions.

It cannot be determined which of these behaviors animals engage in to reduce anxiety or distress vs. which are distressing to the animal, but the animal is powerless to control them, the latter being a necessary criterion for the diagnosis of OCD in humans, as indicated previously. Some behaviors, such as pacing, whirling, and compulsive sexual activity, may be engaged in because a captive animal is deprived of necessary physical exercise and environmental stimulation, and these behaviors provide some degree of

pleasure and relief from boredom. Nevertheless, some compulsive behaviors are pathological, such as acral lick in dogs, resulting in dermatitis, and avian compulsive feather picking. Several of these animal behaviors have been treated successfully with SUIs as well as with dopamine-blocking drugs.

The pharmacology of drugs useful in treating OCD points to which brain areas might be involved. Both the cerebral cortex of the frontal lobes (frontal cortex), which receives prominent serotonin connections from brain stem nuclei, and the basal ganglia, which receive dopamine connections from the substantia nigra, have been considered in OCD pathology. However, it should be noted that psychological treatments such as behavior therapy also can have a significant therapeutic effect, and brain imaging studies have shown that behavior therapy can induce changes in brain metabolism in OCD similar to those induced by drug therapy. Whether psychological therapy induces changes in neurotransmitter activity similar to those induced by pharmacotherapy is as yet unknown. In some instances of severe, incapacitating OCD that was unresponsive to all other treatments, stereotaxic cingulotomy (surgery that interrupts connections between the thalamus and the prefrontal cortex) has been effective. A significant minority of such patients experience gratifying clinical remission, but a partial response is equally common. Some 30–50% of patients, however, have little improvement from surgery.

Brain Circuit Theories of OCD

Brain circuit theories of OCD represent variations on the postulated involvement of certain areas of the frontal parts of the brain (frontal lobes) and their functional connections with deeper parts of the brain – the basal ganglia, thalamus, and other structures, including hippocampus and amygdala. There also are connections from these structures back to the cortex itself. In a general sense, the frontal lobes of the brain mediate executive functions, that is, an individual's decision-making and behavioral actions based on detailed evaluation of environmental demands, along with an appreciation of their historical context and a coordinated affective/emotional component. Particular areas of the frontal cortex mediate executive behavior, social behavior, and motivation.

The underlying theme of the brain circuit theories of OCD is that the lower, or orbital, frontal cortex is a generator of afferent impulses to the caudate nucleus, which is a part of the basal ganglia, and the caudate in turn serves as a filter or gating area. The caudate nucleus is considered to normally suppress extraneous thoughts, sensations, and actions, with little need

for control by the frontal cortex. Some investigators have hypothesized a primary defect in caudate function in OCD, which leads to an increased requirement for the frontal cortex to suppress and emotionally neutralize intrusive neuronal activity. Other investigators have viewed overactivity of the orbital frontal-thalamic circuit as the cause of compulsions in OCD, since compulsions represent unmodulated behavioral drive. The functional result is that the caudate nucleus cannot reasonably filter all the neuronal impulses sent to it from frontal cortical areas; some inappropriate or otherwise innocuous stimuli therefore are passed on to other structures, where they acquire excessive emotional meaning that results in inappropriate thoughts and actions.

Electroencephalographic (brain wave) studies of OCD subjects undergoing cognitive neuropsychological testing also support involvement of the frontal lobes, and functional neuroimaging studies suggest increased blood flow and metabolic overactivity of the orbital frontal cortex. Imaging studies of the caudate nucleus are less consistent in their findings. Thus, it is not possible at present to determine from imaging studies where in the frontal cortical/caudate nucleus/thalamus/frontal cortical circuit the primary defect resides. Structural lesions (for example, strokes and tumors) of both the frontal cortex and the basal ganglia have been associated with obsessions and stereotyped behaviors resembling compulsions. And, as mentioned, drug treatments of OCD do not particularly clarify the matter, since alteration of cortical serotonergic systems with SUIs often results in only partial relief of symptoms, and dopamine-enhancing drugs, which act on caudate nucleus dopamine systems, have some role in the management of OCD symptoms, especially when there is associated tic disorder.

Current theories emphasize hyperactivation of the lateral orbitofrontal loop in OCD, with overactivity of the orbital frontal cortex, the anterior cingulate cortex (ACC), and the caudate nucleus. As indicated previously, these brain structures also have been implicated in major depression and anxiety disorders. The ACC has been further implicated in OCD through correlative functional magnetic resonance imaging (MRI) studies of ACC activity and error checking. Compared to normal control subjects, patients with OCD have excessive error-related activation of the rostral ACC, which is correlated with the severity of OCD symptoms. One of the primary executive functions of the ACC is error detection or mismatch recognition; dysfunction in this brain region might well be associated with the inability of OCD patients to achieve cognitive closure. Failure to achieve cognitive closure is precisely what they demonstrate in their inability to terminate rituals and obsessions

despite intellectual awareness that these behaviors are not justified. For the OCD patient, it is never okay: there is always the possibility that he or she has failed to rule out some contingency. This cognitive style is also apparent in formal testing of, for example, deductive and inductive reasoning: Patients with OCD have no problem with deductive reasoning, but they are significantly impaired on inductive reasoning tasks.

Neuropsychological Testing in OCD

Neuropsychological tests can tap some of the components of frontal lobe function, but it is difficult to implicate the involvement of a particular area of the frontal lobes in OCD from impaired performance on a specific neuropsychological test. Studies indicate deficits in OCD patients' ability to shift task set. This is consistent with theories of disrupted response feedback, in which repeated compulsive acts are required to determine their adequacy. An inherent need for high accuracy in a given individual may aggravate the tendency to lose response feedback, setting up the repetitive behavioral pattern.

Another issue is whether neuropsychological deficits or changes in brain metabolic activity appears first. Altered metabolic activity may result in the behavioral and neuropsychological test changes found in OCD, but it also may be true that the anxiety and need for absolute mastery caused by the illness itself underlie the increased cerebral metabolic activity noted on electroencephalogram (EEG) and functional neuroimaging studies.

Neuroimaging Studies in OCD

With the advent of neuroimaging techniques, it now is possible to probe the structure and function of the living human brain. Brain structure can be viewed with x-ray computed tomography (CT) and MRI. Nuclear medicine techniques such as single-photon and positron-emission computed tomography (SPECT and PET) allow the visualization and quantitation of regional cerebral blood flow, glucose metabolic rate, and neurotransmitter receptor occupancy, which are indirect probes of regional neuronal function. SPECT and PET use radiolabeled compounds as tracers; as these compounds decay, high-energy photons are emitted that are counted by external detectors. The distribution of the tracer molecules is computed, and cross-sectional images of the brain are created in which image brightness is proportional to the underlying physiological process being measured.

Studies of brain structure with CT and MRI in OCD patients have not demonstrated consistent

abnormalities. Most of these studies, however, focused on the caudate nucleus or lateral cerebral ventricles; less attention has been paid to the frontal lobes. Taken together, the findings suggest that structural abnormalities may be more associated with subtle neurological deficits accompanying OCD in some patients than with the disorder itself.

In contrast to the contradictory data from structural imaging studies, a consistent finding in SPECT and PET functional neuroimaging studies of OCD has been increased frontal cortical blood flow and glucose metabolism (hyperfrontality), although the precise cortical location has varied. The hyperfrontality does not seem to be a reflection of heightened anxiety in OCD patients. On the other hand, a relationship may exist between OCD hyperfrontality and serotonin neurotransmitter dysfunction. SUIs not only reduce OCD symptoms, but also alleviate the hyperfrontality, as shown by SPECT and PET.

A tendency toward lateralized right frontal hyperactivity in OCD also has been noted. Right frontal hypermetabolism in OCD is consistent with theories of frontal emotional lateralization. The right frontal cortex is considered to play a greater role in perception of and reaction to negative primary emotions and stimuli, whereas the left frontal cortex may be more active in response to positive stimuli and positive social emotions. The inability of OCD patients to suppress negative primary emotions and thoughts occurs along with their right frontal hyperactivity. A similar formulation of predominantly right prefrontal activity has been advanced for major depression. Suppression of right hemisphere cerebral activity, either with the Wada test (transient suppression of a cerebral hemisphere by barbiturate injection into the same-sided carotid artery) or as a result of lesions, is associated with decreased apprehension, euphoric mood, and minimization of negative emotions, while left hemisphere suppression produces feelings of depression, guilt, and worries about the future. With regard to basal ganglia function, neither structural nor functional neuroimaging studies strongly substantiate the hypothesis of a caudate abnormality in OCD. They do not rule out this possibility, but rather indicate that whatever abnormality might exist has not been consistently seen with current functional neuroimaging technology.

Conclusions

There is considerable variation in obsessive and compulsive symptoms, ranging from normal time-limited worry, through obsessive-compulsive personality disorder, to self-consuming OCD. The criteria for OCD represent a consensus statement among mental health

professionals, based not only on the nature of the OCD symptoms but also on their frequency and interference with social and occupational functioning. OCD thus remains an arbitrarily defined syndrome, no matter how much expertise underlies the definition, so long as there are insufficient empirical data to validate its qualitative difference from obsessive-compulsive personality disorder, as well as from other major anxiety syndromes.

While animal models of OCD provide an interesting theoretical perspective, as noted earlier, this analogy also suffers from a certain nonspecificity. Repetitive behaviors such as pacing, whirling, hitting oneself, toe walking, and rocking, as well as increased anxiety, also are common symptoms of autism, a pervasive developmental disorder that begins in early childhood and persists into adult life. In autistic adults, dietary tryptophan depletion, which depletes brain serotonin, increases these repetitive behaviors, and treatment with SUIs decreases them. This suggests an involvement of brain serotonin neurotransmitter systems in autism similar to that in OCD.

It is interesting to speculate on the opposite of OCD, that is, excessive neglect in conducting one's activities of daily living. A large segment of society fails to adequately check their surroundings and to protect themselves from even known threats. Preventable accidents are a leading cause of mortality and morbidity, and epidemics of avoidable diseases such as AIDS are only too well known. This can be considered inappropriate denial of dangerous situations that have the potential to cause real harm to the individual. However, we approach inappropriate denial and risk taking only from a cognitive, educational standpoint, for example, trying to convince persons at risk to just say no. Perhaps our relative failure to make headway with such an approach suggests that we should take a fresh, psychobiological look at this aspect of behavior. For example, how is compulsive risk taking psychobiologically similar to or different from compulsive checking and hand washing? Psychologically they could be viewed as opposite ends of a spectrum, the former representing pathological self-neglect and the latter pathological self-protection. These and other issues remain the exciting areas of future research into OCD.

Acknowledgment

This work was supported by NIH grant MH28380.

See Also the Following Articles

Behavior Therapy; Depression and Manic-Depressive Illness; Selective Serotonin Reuptake Inhibitors (SSRIs).

Further Reading

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th edn.). (DSM-IV). Washington, D.C.: American Psychiatric Association.
- Barr, L. C., Goodman, W. K., Price, L. H., McDougle, C. J. and Charney, D. S. (1992). The serotonin hypothesis of obsessive compulsive disorder: implications of pharmacologic challenge studies. *Journal of Clinical Psychiatry* 53(Supplement), 17–28.
- Chamberlain, S. R., Blackwell, A. D., Fineberg, N. A., Robbins, T. W. and Sahakian, B. J. (2005). The neuropsychology of obsessive compulsive disorder: the importance of failures in cognitive and behavioral inhibition as candidate endophenotypic markers. *Neuroscience and Biobehavioral Reviews* 29, 399–419.
- Cummings, J. L. (1995). Anatomic and behavioral aspects of frontal-subcortical circuits. *Annals of the New York Academy of Science* 769, 1–13.
- Denys, D., Zohar, J. and Westenberg, H. G. M. (2004). The role of dopamine in obsessive-compulsive disorder: preclinical and clinical evidence. *Journal of Clinical Psychiatry* (Supplement 14), 11–17.
- Dodman, N. H., Moon-Fanelli, A., Mertens, P. A., Pflueger, S. and Stein, D. J. (1997). Veterinary models of OCD. In: Hollander, E. & Stein, D. J. (eds.) *Obsessive-compulsive disorder*, pp. 99–144. New York: Marcel Dekker.
- Fitzgerald, K. D., Welsh, R. C., Gehring, W. J., et al. (2005). Error-related hyperactivity in the anterior cingulate cortex in obsessive-compulsive disorder. *Biological Psychiatry* 57, 287–294.
- Hoehn-Saric, R. and Benkelfat, C. (1994). Structural and functional brain imaging in obsessive-compulsive disorder. In: Hollander, E., Zohar, J., Marazziti, D. & Olivier, B. (eds.) *Current insights in obsessive-compulsive disorder*, pp. 183–211. New York: Wiley.
- Jenike, M. A. (1995). Obsessive-compulsive disorder. In: Kaplan, H. I. & Sadock, B. J. (eds.) *Comprehensive textbook of psychiatry* (6th edn., pp. 1218–1227). Williams & Wilkins: Baltimore, MD.
- Jenike, M. A. (2004). Obsessive-compulsive disorder. *New England Journal of Medicine* 350, 259–265.
- McGuire, P. K., Bench, C. J., Frith, C. D., et al. (1994). Functional anatomy of obsessive-compulsive phenomena. *British Journal of Psychiatry* 164, 459–468.
- Otto, M. W. (1992). Normal and abnormal information processing: a neuropsychological perspective on obsessive-compulsive disorder. *Psychiatry Clinics of North America* 15, 825–848.
- Pelissier, M. C. and O'Connor, K. P. (2002). Deductive and inductive reasoning in obsessive-compulsive disorder. *British Journal of Clinical Psychology* 41(Part 1), 15–27.
- Rubin, R. T. and Harris, G. J. (1999). Obsessive-compulsive disorder and the frontal lobes. In: Miller, B. L. & Cummings, J. L. (eds.) *The human frontal lobes: functions and disorders*, pp. 522–536. New York: Guilford.
- Rubin, R. T., Villanueva-Meyer, J., Ananth, J., Trajmar, P. G. and Mena, I. (1992). Regional xenon 133 cerebral blood flow and cerebral technetium 99m HMPAO uptake in unmedicated patients with obsessive-compulsive disorder and matched control subjects. *Archives of General Psychiatry* 49, 695–702.
- Saxena, S., Brody, A. L., Schwartz, J. M. and Baxter, L. R. (1998). Neuroimaging and frontal-subcortical circuitry in obsessive-compulsive disorder. *British Journal of Psychiatry* 173(Supplement 35), 26–37.
- Zielinski, C. M., Taylor, M. A. and Juzwin, K. R. (1991). Neuropsychological deficits in obsessive-compulsive disorder. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology* 4, 110–126.

Oklahoma City Bombing, Stress Effects of

C Piotrowski and S J Vodanovich

University of West Florida, Pensacola, FL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by C Piotrowski, volume 3, pp 83–85, © 2000, Elsevier Inc.

Site of the Bombing
Effects on Survivors
Impact on Children
Emergency Personnel and Disaster Mental Health

Glossary

<i>Access victims</i>	Individuals who offered support and human assistance to survivors and families of victims, such as mental health professionals, nurses, and clergy.
<i>Direct-contact victims</i>	Emergency and rescue workers who searched the rubble for survivors and dead bodies, in addition to morgue personnel.
<i>Direct-impact victims</i>	Survivors who witnessed the bombing, death, and maiming of immediate victims firsthand.

<i>Disaster mental health</i>	A competency-based intervention model, based on disaster-related psychopathology, that addresses a community's mental health needs.
<i>Intraself risk factors</i>	A victim's idiosyncratic perception of the meaning and appraisal of a catastrophe and any premorbid factors that impact postdisaster reactions.
<i>Mass-casualty event</i>	The multidimensional stressors that may result in differential patterns of posttraumatic reactions, both short-term and long-term.
<i>Self-efficacy adaptational coping</i>	Survivors' struggles to maintain psychosomatic and social equilibrium.
<i>Socioenvironmental risk factors</i>	The degree of ecological and interpersonal destruction and intensity of exposure to morbidity.

Site of the Bombing

The bombing of the Alfred P. Murrah federal office building in Oklahoma City, Oklahoma, on April 19, 1995, remains the most devastating domestic terrorist act on U.S. soil. This tragic event impacted a close-knit community in America's heartland and tore at the fabric of our national sense of security. The bomb blast had a devastating impact; 168 lives were lost and nearly 1000 people were injured. Sadly, 19 children were killed and 57 were injured, some quite seriously. A majority (60%) of the fatalities were government employees. Because the event occurred in the morning, thousands of commuters were in the downtown area, in close proximity of the blast, heard the shattering explosion, and witnessed the calamity unfold. The rescue response was immediate, with hundreds of emergency workers, medical personnel, and law enforcement officials attending to the scene. During the first 48 h, the city's hospitals were overwhelmed with the critically wounded and exasperated family members while the bombing scene was amass with rescue teams in search of the survivors and victims. In a matter of minutes, Oklahoma City experienced a calamity of monumental proportions. Although the Oklahoma City bombing has many features in common with most disasters, the fact that the perpetrators were U.S. citizens added an additional stress factor to the readjustment of victims and the recovering community. The resolution of intense emotional concerns has had a central role in the healing process of not only the residents of Oklahoma City but also the nation as a whole. Perhaps this national consciousness is most evident in the poignant annual commemoration for all the victims of this tragedy.

Effects on Survivors

Disasters are usually classified as natural (e.g., hurricanes) or human-made (Three-Mile Island), but with the bombing in Oklahoma City, the American psyche was introduced to the ominous reality of disaster by intentional human strategy. In 1995, Parson coined the term Oklahoma City Bombing Stress Response (OKC-BSR), defined as (1) traumatic stress responses and (2) traumatically shattered meaning elements. In understanding victims' collective psychological responses and adaptation, Parson argued that ORC-BSR involves psychophysiological reactions to dissociation, anxiety symptoms, and acute stress disorder with a concomitant need to find meaning from a terrorist act. To address the overwhelming need for mental health services in the Oklahoma City area, a special program for individuals reporting psychological distress from the bombing (entitled Project Heartland) was established. Some 9000 residents received services such as crisis intervention and support groups. This program offered a model for investigating risk factors for posttraumatic stress disorder (PTSD) symptoms among the survivors of the trauma-related community stress.

Research studies on mental health symptomology and stress syndromes following the Oklahoma City bombing have been widespread in recent years. A study by Sprang, based on the premise that post-disaster reaction is a function of degree of exposure, found Oklahoma City residents reported high levels of PTSD and victimization symptomatology. Tucker and colleagues reported on 86 adults who sought assistance for stress related to the bombing 6 months after the tragedy. Reactions of nervousness to how others reacted at the time of the bombing accounted for approximately one-third of the variance in post-traumatic stress scores. These findings suggest that people who experience high anxiety during a traumatic event are subject to prolonged stress reactions over time. Pfefferbaum and colleagues reported in a series of studies on the long-term psychopathology and PTSD symptoms evident in both the survivors and the local population after the bombing. Survivors ($n = 37$) in the direct path of the explosion were assessed 17 months postdisaster. The results showed that 31% of the sample was designated as chronic PTSD. In another survivor sample ($n = 182$), 45% was diagnosed with postdisaster psychiatric disorders and 45% exhibited PTSD symptoms, particularly intrusive memories and hyperarousal, 6 months after the bombing. Furthermore, Pfefferbaum and colleagues, in a group of 84 individuals from the Oklahoma City community seeking support and mental health services, found a strong association between

PTSD and substance use behaviors such as smoking and alcohol intake. Such coping mechanisms have been found to exacerbate functional impairment in trauma/disaster victims. These studies on the long-term effects of the bombing confirm not only the immediacy of PTSD onset but also lingering stress response syndromes in populations with proximity to the disaster impact area. This calls for early intervention with long-term treatment options.

Because victims of a disaster may have had previous trauma, PTSD, or premorbid psychiatric conditions, they may be vulnerable to reacting pathologically to subsequent stressful events. In this regard, Moyers reported on the exacerbation of PTSD symptoms following the Oklahoma City bombing in veterans who were attending support groups.

Impact on Children

In recent years, the effects of major disasters on the mental health and emotional functioning of children has become a specialized area of research endeavor. Children are a particularly vulnerable group with the added stress of adapting to developmental milestones and dealing with emotional turmoil. In a sample of over 3000 middle school and high school students in the Oklahoma City school system, researchers found that children who knew someone who was killed by the blast (vicarious exposure) reported higher PTSD severity than did children who were not highly exposed to the traumatic event. Bereaved youths were more likely than nonbereaved peers to experience symptoms of anxiety and fear, changes in both school and home environments, and PTSD-type symptoms. Interestingly, girls reported significantly higher PTSD severity scores than boys. In addition, students who viewed extensive television coverage of the bombing were most negatively affected. These findings support prior research on the functional role that the defense mechanism of denial plays in the adaptation of children to the onerous impact of psychic trauma and disasters.

Emergency Personnel and Disaster Mental Health

A review of the literature on trauma provides conflicting evidence on the prevalence of stress reactions in emergency services personnel and rescue workers in the aftermath of disasters. Studies of firefighters and rescue personnel who responded to the bombing reported an unexpectedly low incidence rate (13%) for PTSD, even 3 years postdisaster. Body

handlers were found to have negligible symptoms of depression or PTSD, and their overall mental health status improved over time. The resilience factors that aided in adaptation for these emergency personnel were identified as professional hardiness, job experience, and postdisaster debriefing.

Conversely, mental health workers in Oklahoma City reported chronic stress due to the increased workload for an extensive time period after the bombing. In fact, disaster mental health personnel were most prone to compassion fatigue and burnout. These findings illustrate the critical importance of professional education and psychosocial support initiatives. The implications for occupational stress, mental health, and organizational efficacy for emergency service workers cannot be ignored. Future research must focus on competency in conducting both disaster planning and postdisaster intervention.

See Also the Following Articles

Alarm Phase and General Adaptation Syndrome; Disaster Syndrome; Disasters and Mass Violence, Public, Effects of; Terrorism; Workplace Stress; Hurricane Katrina Disaster, Stress Effects of.

Further Reading

- Allen, S. F., Dlugokinski, E. L., Cohen, L. A., et al. (1999). Assessing the impact of a traumatic community event on children and assisting with their healing. *Psychiatric Annals* 29(2), 93–98.
- Figley, C. R. (ed.) (2002). *Treating compassion fatigue*. New York: Brunner/Routledge.
- Gordon, N. S., Farberow, N. L. and Maida, C. A. (1999). *Children and disasters*. Philadelphia: Brunner/Mazel.
- Moyers, F. (1996). Oklahoma City bombing: exacerbation of symptoms in veterans with PTSD. *Archives of Psychiatric Nursing* 10, 55–59.
- North, C. S., Nixon, S. J., Shariat, S., et al. (1999). Psychiatric disorders among survivors of the Oklahoma City bombing. *Journal of the American Medical Association* 282, 755–762.
- North, C. S., Pfefferbaum, B., Tivis, L., et al. (2004). The course of posttraumatic stress disorder in a follow-up study of survivors of the Oklahoma City bombing. *Annals of Clinical Psychiatry* 16, 209–215.
- North, C. S., Tivis, L., McMillen, J. C., et al. (2002). Psychiatric disorders in rescue workers after the Oklahoma City bombing. *American Journal of Psychiatry* 159, 857–859.
- Parson, E. R. (1995). Mass traumatic terror in Oklahoma City and phases of adaptational coping. Part I: Possible effects on intentional injury/harm on victims' post-traumatic responses. *Journal of Contemporary Psychotherapy* 25, 155–184.

- Pfefferbaum, B. (2001). The impact of the Oklahoma City bombing on children in the community. *Military Medicine* 166 (supplement 2), 49–50.
- Pfefferbaum, B. and Doughty, D. E. (2001). Increased alcohol use in a treatment sample of Oklahoma City bombing victims. *Psychiatry* 64, 296–303.
- Piotrowski, C., Armstrong, T. and Stopp, H. (1997). Stress factors in the aftermath of Hurricanes Erin and Opal: data from small business owners. *Psychological Reports* 80, 1387–1391.
- Piotrowski, C. and Dunham, F. (1983). Locus of control orientation and perception of “Hurricane” in fifth graders. *Journal of General Psychology* 109, 119–127.
- Reyes, G. and Elhai, J. D. (2004). Psychosocial interventions in the early phases of disasters. *Psychotherapy: Theory, Research, Practice, Training* 41, 399–411.
- Sprang, G. (1999). Post-disaster stress following the Oklahoma City bombing – an examination of three community groups. *Journal of Interpersonal Violence* 14, 169–183.
- Tucker, P., Dickson, W., Pfefferbaum, R., et al. (1997). Traumatic reactions as predictors of posttraumatic stress six months after the Oklahoma City bombing. *Psychiatric Services* 48, 1191–1194.

Opioid System *See: Opioids.*

Opioids

J A Russell

University of Edinburgh, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J A Russell and A J Douglas, volume 3, pp 86–98, © 2000, Elsevier Inc.

Overview

Neuroendocrine Responses

Opioids and Stressor Processing

Opioids, Stress, and Reproduction

Conclusion

Glossary

β-Endorphin

A 31-amino-acid endogenous opioid peptide that acts mainly on δ - and μ -opioid receptors, synthesized from the proopiomelanocortin (POMC) gene. Cells expressing POMC are found in the arcuate nucleus of the hypothalamus, the corticotroph cells of the anterior pituitary, and the cells of the intermediate pituitary. Nerve fibers containing β -endorphin innervate the hypothalamus and project to other brain regions.

Agonist

Drug or exogenous or endogenous substance that binds to a receptor in a cell to induce a response.

Analgesia

The relief from pain; it may be local, spinal cord-mediated, or central.

Antagonist

A drug or endogenous substance that binds to a receptor in a cell-displacing bound agonist to prevent or reverse the effect of the endogenous ligand or agonist.

Dependence

The result of chronic exposure to an opiate drug that causes withdrawal signs to occur when the opiate exposure is interrupted; also called addiction.

Dynorphins

Endogenous opioid peptides that are 8 or 17 amino acids long and act mainly on κ -opioid receptors but also partly on δ - and μ -opioid receptors; synthesized from the preprodynorphin gene (also known as preproenkephalin B).

Endomorphins

Endogenous opioid peptides (discovered in 1997) that are 4 amino acids long and act selectively on the μ -opioid receptor; the gene or precursor is not known.

Enkephalins

Endogenous opioid peptides that are five amino acids long (Met⁵ and Leu⁵ enkephalins) and act mainly on δ - but also on κ - and μ -opioid receptors; synthesized from the preproenkephalin A gene. They are present in neurons widely distributed in brain, spinal cord, and adrenal medulla.

Euphoria

The conscious sensation of pleasure or feeling of well-being.

<i>Limbic system</i>	Part of the brain comprising the amygdala, bed nucleus of the stria terminalis, hippocampus, septum, and nucleus accumbens that is concerned with memory, emotionality, and the processing of stressors.
<i>Morphine</i>	An opiate alkaloid produced by the opium poppy that acts in the brain and peripherally to cause analgesia, euphoria, and reduced alarm; it is a relatively selective μ -opioid receptor agonist. Opium and heroin, a modified form of morphine, are widely abused for their pleasurable effects, but they induce tolerance and dependence (addiction).
<i>Naloxone</i>	A general opioid receptor antagonist with greatest activity at μ -opioid receptors; it enters the brain from the circulation readily and reverses opioid agonist effects rapidly. It precipitates withdrawal symptoms in opiate addiction.
<i>Nociceptin</i>	A peptide related to dynorphin that binds to ORL1 (OP4) receptors, which are related to κ -opioid receptors. It lacks the classic action of opioids, is not antagonized by naloxone, and has central pronociceptive properties; also called orphanin FQ.
<i>Nociception</i>	Processes involved in perceiving pain following the stimulation of nociceptive nerve endings (nociceptors).
<i>Opiate</i>	A drug derived from exogenous sources, originally the opium poppy, producing opium-like effects.
<i>Opioid</i>	A drug or endogenously produced compound (e.g., opioid peptide) that acts through opioid receptors to produce effects that are antagonized by naloxone.
<i>Opioid receptors</i>	Proteins in the cell membrane that confer sensitivity to opioids and opiates in the cells that express them; there are three, μ , δ , and κ . They have seven transmembrane domains, ligand-binding sequences projecting extracellularly, and a G-protein binding region intracellularly; a $G_{i/o}$ -protein mediates the intracellular effects.
<i>Orexigenic</i>	Being appetite stimulating; applied to an endogenous substance.
<i>Pain threshold</i>	Often defined by the time a subject takes to respond to an applied painful stimulus of set intensity (e.g., in the tail flick test, which is applying radiant heat to the tail).
<i>Plasticity</i>	Adaptation of an endogenous system so that its response to a particular stimulus or set of stimuli changes with repeated exposure or with a change in the physiological state of the individual.

Tolerance The effect of an opiate drug on receptor and postreceptor mechanisms in a cell so that the same dosage of a drug has a reduced effect and the dosage must be increased to maintain the same effects.

Overview

The soporific and analgesic properties of opium have been known to humans for thousands of years. Morphine, an alkaloid produced by the opium poppy, *Papaver orientale*, is the prototypic opioid. It and related synthetic opiates, such as heroin, have been widely used to control pain while reducing the state of alarm in those with severe injury or suffering the pain of acute myocardial infarction, advanced cancer, surgery, or childbirth. These drugs are widely abused because of their euphoric properties, but they induce tolerance and dependence. Sexual activity is reduced in opiate-dependent individuals. An overdose of morphine results in respiratory depression that can lead to death, but the potent antagonist naloxone acts rapidly to reverse the depression. Experimentally, these drugs have provided the tools to characterize opioid mechanisms. Pharmacological studies with various opiate drugs and antagonists identified three principal types of opiate receptor (μ , δ , and κ , according to their affinities for different agonists). These were found in the brain, spinal cord, posterior pituitary gland, gut, and other tissues. They are proteins, synthesized within the opioid-sensitive cells and inserted into the lipid plasma membrane so that the amino acid sequence that tightly binds the opioid ligand is exposed on the cell surface. The hypothesis that the expression of these high-affinity receptors for opiate drugs indicated the presence of endogenously produced ligands led to the discovery of the endogenous opioid peptides in the early 1970s. Three families were discovered: the enkephalins, β -endorphin, and the dynorphins. Subsequently, a peptide related to the dynorphins, nociceptin (or orphanin FQ), was discovered, but this acts on a different receptor (ORL1) that the other opioids do not act on. Naloxone is not effective at this receptor, and nociceptin lacks the classical actions of opioids. Indeed, in some paradigms it enhances pain perception. The three types of opioid receptors, although defined initially on the basis of their distinctive affinities for different opiate drugs and, in the early 1990s, by the characterization of their genes, in general do not selectively bind the different types of endogenous opioids. The enkephalins bind to μ - and δ -, β -endorphin binds to μ - and κ -, and the dynorphins bind to κ -receptors. However, opioid peptides (endomorphins 1 and 2)

selective for the μ -receptor have been identified in the brain. The enkephalins, β -endorphin, the dynorphins, and nociceptin are produced by the processing of higher-molecular-weight precursors, each translated from a distinct gene. However, each of the opioid precursor proteins contains opioid peptide sequences also present in the other two, indicating their common origin from an ancestral gene. Similarly, the opioid receptors are all members of the seven-transmembrane-domain G-protein-linked family, with considerable homology in their sequences. The G-protein binding part of the receptor is in the cytoplasmic loops. The intracellular G-protein links the receptor to the intracellular process that the opioid affects; this link may be direct (e.g., to an ion channel) or involve the inhibition of second-messenger (cAMP) generation. The G-protein is the $G_{i/o}$ type; it is inactivated by the pertussis toxin, and this is used as an experimental tool to determine whether a $G_{i/o}$ -protein is involved in a particular opioid action.

Two principal types of opioid actions on neurons have been identified: presynaptic and postsynaptic. When opioids act presynaptically on nerve terminals, they inhibit the entry of Ca^{2+} ; because such Ca^{2+} entry is essential for the release of a neurotransmitter into the synaptic cleft, opioids inhibit release, reducing the effectiveness of the synapse. If the opioid-inhibited presynaptic neuron produces an excitatory transmitter, then the postsynaptic neuron is inhibited; conversely the presynaptic opioid inhibition of an inhibitory neuron excites the postsynaptic neuron. Opioids can also act postsynaptically on the dendrites or cell bodies of neurons, inhibiting the processes that generate action potentials and, thus, indirectly inhibiting the release of the neurotransmitter from the terminals of these neurons. The mechanism of this type of opioid inhibition involves the activation of K^+ channels in the plasma membrane; the outflow of K^+ then hyperpolarizes the neuron, making it less excitable. This mechanism also contributes to opioid actions on nerve terminals. Nociceptin has similar modes of inhibiting neural activity. The actions of opioids reverse when the opioid diffuses away or is broken down by local proteases into peptides that have little affinity for the opioid receptors. The combination of an opioid receptor with an endogenous peptide leads to the internalization and activation of other intracellular regulatory mechanisms that blunt further opioid action. It seems that these mechanisms do not follow the occupancy of the μ -receptor with morphine, and this may underlie the addictive properties of morphine.

Plasticity is a feature of opioid systems. This is evident in the increased or decreased expression of

opioid receptors or opioid peptides, which may be a result of the altered activity of their genes. The expression of the preproenkephalin and preprodynorphin genes is regulated by promoter sequences responding to intracellular messengers. The pro-opiomelanocortin (POMC) gene (which encodes β -endorphin) has a glucocorticoid receptor-response element that mediates the inhibition of expression in the anterior pituitary but mediates its stimulation in β -endorphin neurons. Opioid receptor expression can change because of redistribution within a neuron or altered turnover at the plasma membrane.

The principal actions of opiate drugs on mental state and pain perception suggest that the activation of endogenous opioid mechanisms during stress moderate the processing of stressors and the responses to them. There are several lines of evidence for this. First, the opioid peptides and their receptors are found, at appropriate sites, to influence reactions to stress. Second, the actions of opioid antagonists reveal activity in endogenous mechanisms; the effects of exogenous opioids, and the consequences of absence of opioid peptides or their receptors in mutants, all point to important roles for endogenous opioids in stress. Third, the loss of responses suspected to be due to endogenous opioid action after chronic morphine treatment, which produces tolerance to opioid actions at μ -receptors, indicates a cross-tolerance to an endogenous opioid action. It is clear that endogenous opioid mechanisms are involved in modulating reactions to stress at many levels. For physically painful stimuli, endogenous opioids modulate processing in the nociceptive pathways. Opioids are present in the brain-stem nuclei concerned with stressor processing and, similarly, in the limbic system. Opioids act in parts of the brain concerned with pleasure and reward, and here they affect appetite. Opioid peptides are coexpressed in the neuroendocrine neurons that mediate hormonal stress responses, and these neurons can be targets for opioid peptides. In the periphery, endogenous opioids modify the activity of immune cells and modulate inflammatory pain by actions on nociceptive nerve terminals. Opioids impinge on the neuroendocrine neurons regulating gonadal activity.

The major sources of endogenous opioids are the anterior pituitary gland and the medullae of the adrenal glands, which can both secrete opioid peptides (β -endorphin and enkephalins, respectively) into the circulation when stimulated; the brain and spinal cord, where opioids are produced by specific neurons; and the cells of the immune system, when activated. Opioids produced by neurons (β -endorphin, enkephalins, or dynorphins) act locally in the immediate vicinity of the site of release from the terminals of

the processes of the neurons. Enkephalins and dynorphins are commonly produced in neurons also producing other transmitters, whereas β -endorphin neurons also produce other products of the POMC gene. Opioids from neurons generally inhibit the activity of other neurons or act back on the neuron terminals from which they have been released, inhibiting further the release of transmitters. The opioids (principally β -endorphin) produced by immune cells also act locally at the site of inflammation. For opioids secreted into the circulation, there is a problem about their site(s) of action. The concentrations of opioid peptides in the circulation may not reach levels sufficient to activate the receptors in distant tissues, and blood-borne opioids may not penetrate the tissues to reach cells bearing opioid receptors. In particular, the tight junctions between the lining cells of the blood capillaries in most regions of the brain (the blood-brain barrier) prevent the free diffusion of peptides into the brain. There are transport mechanisms in the choroid plexus for selectively transporting certain peptide messengers (e.g., leptin) from blood into the cerebroventricular system. However, there is little transfer of β -endorphin from the blood into cerebrospinal fluid, and it does not then permeate the brain tissue. Enkephalins produced in the brain are removed partly by being transported into the blood, which seems to preclude their being transported in the opposite direction. Dynorphins are not transported into the brain. Furthermore, in experimental animals, after the removal of the pituitary and adrenal glands, thus eliminating these sources of opioid peptides, pain control during stress can persist; this shows the importance of opioids produced within the brain or spinal cord. Thus, it is unlikely that the activity of neurons in the brain can be directly affected by opioid peptides circulating in the blood. This is not true of the opiate drugs, such as morphine, which enter the brain and have major actions on receptors and on neurons inside the blood-brain barrier, or of antagonists, particularly naloxone, which enters the brain readily and reverses the effects of an opiate drug overdose very rapidly. A modified form of naloxone (naloxone methylbromide), with an altered charge, does not cross the blood-brain barrier; this is used experimentally to distinguish the activity of opioids within the brain from the activity of those outside.

In general, the activation of endogenous opioid mechanisms by stress reduces the perception of pain; modulates autonomic (sympathetic nervous system) and neuroendocrine mechanisms that bring about metabolic adjustments to meet the stress; stimulates eating, especially of highly palatable foods during stress; and inhibits energetically costly activities that

are not essential for immediate survival, especially reproduction. This article reviews the mechanisms.

Neuroendocrine Responses

Opioids and the Hypothalamic-Pituitary-Adrenal Axis

Endogenous opioid and opioid receptor expression in the hypothalamic-pituitary-adrenal system The defining neuroendocrine outcome of stress is the release of corticotropin releasing hormone (CRH) and vasopressin from the terminals of parvocellular paraventricular nucleus (PVN) neurons in the median eminence, causing increased secretion from the anterior pituitary gland of adrenocorticotrophic hormone (ACTH), which stimulates the secretion of glucocorticoids from the adrenal cortex. Endogenous opioids are cosynthesized and/or have actions at all these locations to modulate responses to stress ([Figure 1, i, ii](#)). In the paraventricular nucleus, the parvocellular CRH neurons also synthesize enkephalins from the preproenkephalin A gene and express opioid (at least, κ -) receptors. Preproenkephalin A gene expression is rapidly stimulated in these neurons in response to stress. Endomorphins 1 and 2 have also been described in neurons and fibers in the paraventricular nucleus. There are opioid receptors in the median eminence that mediate local regulation at the level of the terminals secreting releasing factors into the hypothalamo-hypophyseal portal system. Opioids acting here could be enkephalins coreleased with CRH or secreted by the adrenal medulla or by β -endorphin from the anterior pituitary because the blood-brain barrier is deficient in this region. Corticotrophs in the anterior pituitary synthesize both ACTH and β -endorphin from the POMC gene, and they are cosecreted; however, much of the β -endorphin may be acetylated and, consequently, inactive. This changes in pregnancy (the placenta produces POMC in women). ACTH release from the corticotrophs is not directly influenced by any opioid released locally or carried in the portal blood from neuroendocrine neurons. The adrenal cortex does not produce opioid peptides in the rat, guinea pig, human, or bovine, but in the rat it does show specific receptors for β -endorphin and dynorphin in the zona fasciculata-reticularis. However, the adrenal medulla in many species (rat, mouse, human, guinea pig, horse, rabbit, and bovine) expresses preproenkephalin A and POMC-derived opioids. These appear to influence the adjacent cortex locally; enkephalins and dynorphins are generally stimulatory in the rat and endorphins are generally inhibitory to glucocorticoid secretion in both the rat and human. Adrenal

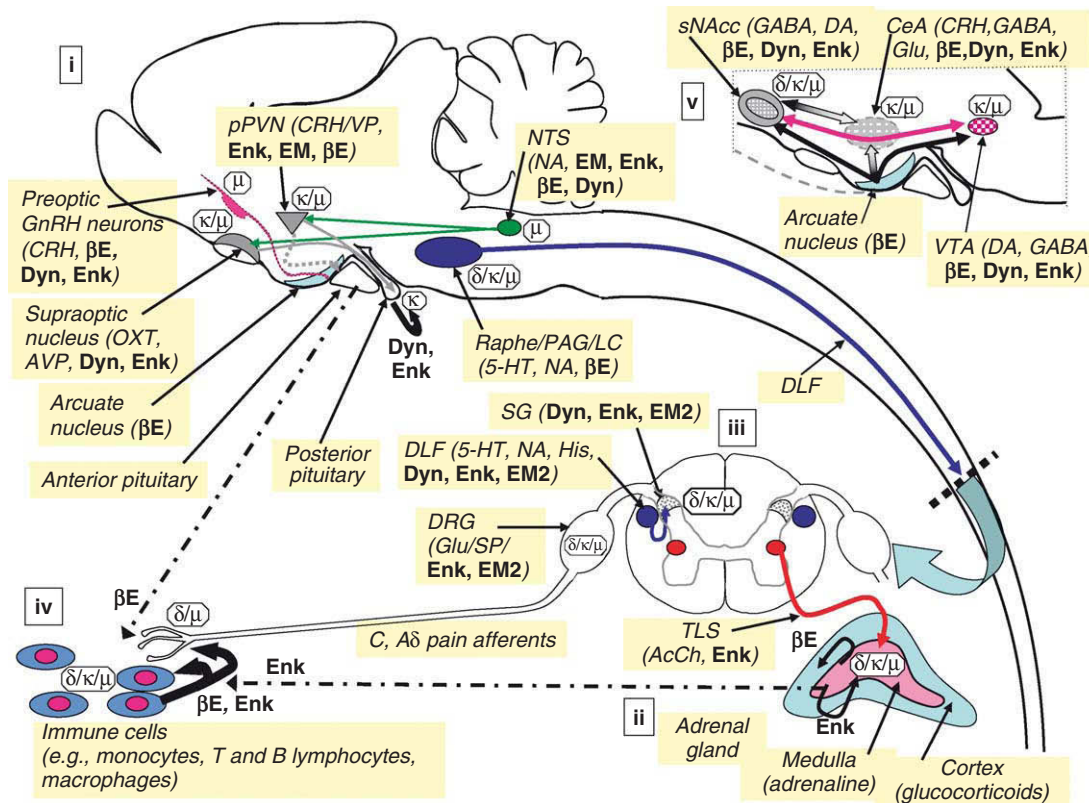


Figure 1 The distribution of opioid peptides and opioid receptors in key elements of central and peripheral structures that organize stress responses. The elements illustrated (not to scale) are (i) neuroendocrine brain and pituitary gland (sagittal section), (ii) adrenal gland, (iii) spinal cord pain pathways (transverse section), (iv) immune cells, and (v) mesolimbic reward circuit. For each element, the predominant types of opioid receptor and the type of opioid peptide (in parentheses in boldface) present or produced are indicated. Note that β -endorphin-producing neurons in the brain are located in the arcuate nucleus (with a few in the nucleus tractus solitarius) and project to many brain regions (as in (v)). Endomorphin-producing neurons are localized in the hypothalamus and nucleus tractus solitarius. Enkephalin- and dynorphin-producing neurons are more widely distributed. Other transmitters in the key central nervous system areas are indicated (in parentheses in italics). Neural pathways modulated by opioids are indicated by straight or long curved arrows; short curved arrows indicate local opioid action; dot-dash lines indicate opioids carried in blood. 5-HT, serotonin; β E, β -endorphin; δ , κ , and μ , δ -, κ -, and μ -opioid receptors; AcCh, acetylcholine; C, A δ , C- and A δ -fiber; CeA, central nucleus of amygdala; CRH, corticotropin releasing hormone; DA, dopamine; DLF, dorsolateral funiculus; DRG, dorsal root ganglion; Dyn, dynorphin; EM, endomorphin; Enk, leu-, met-, or extended enkephalin; GABA, γ -aminobutyric acid; Glu, glutamate; GnRH, gonadotropin releasing hormone; His, histamine; LC, locus coeruleus; NE, norepinephrine; NTS, nucleus tractus solitarius; OXT, oxytocin; PAG, periaqueductal gray; pPVN, parvocellular paraventricular nucleus; Raphe, raphe nucleus; sNAcc, shell of nucleus accumbens; SP, substance P; TLS, thoracolumbar sympathetic outflow; VP, vasopressin; VTA, ventral tegmentum.

corticosteroid secretion is increased after naloxone in the human independently of a change in ACTH secretion.

Endogenous opioid and opioid receptor expression in inputs to the hypothalamic-pituitary-adrenal axis Neural inputs to the hypothalamic-pituitary-adrenal (HPA) axis that contain opioids include β -endorphin nerve terminals of arcuate nucleus neurons (adjacent to the median eminence), which traverse the median eminence and innervate the hypothalamus, including the paraventricular nucleus. CRH neurons in the paraventricular nucleus are contacted by β -endorphin terminals. Thus, β -endorphin can act on both the cell bodies of CRH neurons

and their terminals in the median eminence to influence CRH and vasopressin and/or opioid release at these sites. Endomorphin 2-expressing cell bodies have been located in the hypothalamus, near the third ventricle. In addition, opioids (enkephalins and dynorphins) are synthesized in the neurons of the nucleus of the tractus solitarius (NTS) and ventral medulla in the brain stem, colocalized in noradrenergic neurons, among others; these project to the hypothalamus, including to the paraventricular nucleus and median eminence (Figure 1, i). NTS neurons coexpressing prodynorphin or proenkephalin A mRNA are activated during seizures and other stressors, increasing their expression of prodynorphin mRNA after seizures. The limbic brain regions (the

amygdala; septum; bed nucleus of stria terminalis, BNST; and hippocampus) express opioid receptors and receive opioidergic inputs. Locus coeruleus noradrenergic neurons, important in arousal and attention in response to stressors, receive stimulatory CRH and inhibitory enkephalin-containing terminals, giving the basis for the activation and then the suppression of the activity of these noradrenergic neurons with an acute stressor. After chronic morphine exposure, the effect of CRH is enhanced, and this may exemplify how opiate addiction sensitizes to stressors and how stress predisposes to opiate craving. The amygdala has μ -opioid receptors in both rat and human, and it expresses endomorphin 2 in the rat; μ -, δ -, and κ -opioids all stimulate neurons in the amygdala in the guinea pig. Naloxone prevents stress-induced changes in long-term potentiation in the hippocampus. The septum expresses μ -opioid receptors. Enkephalin-containing neurons in the BNST are contacted by noradrenergic axon terminals and are activated by stressors. Nociceptin and its receptor (ORL1) have been described in both the amygdala and the hippocampus. δ -Receptors are not found in the paraventricular nucleus, pituitary, or adrenal cortex but may weakly contribute to regulating basal HPA secretion, perhaps via distant inputs. Thus, both brain-stem and limbic system neurons involved in mediating the activation of paraventricular nucleus CRH neurons by stressors either produce opioids or can be affected by them.

Endogenous opioids and hypothalamic-pituitary-adrenal stress responses The major effects of opioids on HPA axis stress responses are in the brain. Acutely administered opioids generally act in the hypothalamus to inhibit CRH secretion and hence ACTH secretion in most species. In the sheep, pig, and human, endogenous opioids inhibit basal and stress-stimulated HPA axis activity; and in the horse, opioids are inhibitory but naloxone increases ACTH secretion without altering CRH or vasopressin release. However, in the rat, opioids are stimulatory to the HPA axis. Thus, the release of CRH is increased following the central application of an opioid peptide (e.g. β -endorphin) *in vivo* or *in vitro*, and the effect is blocked by naloxone. The activation of endogenous opioid peptide mechanisms impinging on CRH neurons during stress (including emotional and physical stressors) is indicated by a reduced ACTH response after treatment with the antagonist naloxone. However, there are conflicting reports of tonic opioid-mediated inhibition. The subtle change of a single nucleotide polymorphism (SNP) in the rhesus monkey μ -opioid receptor gene, which increases the affinity of the receptor for β -endorphin, is associated with

reduced basal cortisol secretion. Surprisingly, whereas the central administration of morphine, a μ -receptor agonist, increases corticosterone secretion in the rat, a selective endogenous μ -receptor agonist (endomorphin-1 or -2) does not. In humans, naloxone increases basal ACTH secretion in most subjects and has a greater effect in patients with depression than in patients with Cushing's disease. A naloxone test has been suggested as an aid to distinguishing these conditions, on the basis that naloxone acting in the brain will further enhance CRH release, and hence ACTH secretion, in depressed patients with hypercortisolism but not in patients with Cushing's disease. The different opioid effects on the HPA axis may be the result of different opioids acting through different types of opioid receptor, such that μ -opioids activate and κ -opioids inhibit, and at different sites, depending on the stimulus conditions. The balance of these different actions may even change the evident direction of action from stimulation to inhibition (or vice versa) and account for species differences. Such a change has been described in late pregnancy in the rat, when endogenous opioid inhibits the activation of CRH neurons in response to several stressors by inhibiting the release of norepinephrine in the paraventricular nuclei.

The opioid enhancement of HPA axis secretory responses in rats shows plasticity during chronic stress, aging, and pregnancy/parturition. Because chronic stress increases the expression of the preproenkephalin A gene in parvocellular PVN neurons, their release of enkephalins is expected to increase. Furthermore, the neurons in the parvocellular PVN that produce enkephalin and CRH continue to be activated by an acute stressor after chronic stress. This suggests an important role for the coproduced enkephalin in mediating or modulating the HPA stress response after chronic stress. The opioid regulation of stress responses seems to decline with age because naloxone has a reduced effect, both on HPA axis responses and on stress-induced analgesia. During parturition in the rat, the usual opioid enhancement of secretory responses switches to inhibition, restraining HPA axis secretion. This follows an attenuation of HPA axis stress responses in late pregnancy, partly due to a lack of endogenous opioid enhancement. However, arcuate nucleus POMC expression and hypothalamic β -endorphin content increase in pregnancy in the rat (β -endorphin concentration in the blood also increases in both rat and human during pregnancy). Furthermore, β -endorphin can reduce CRH expression in the PVN and reduce POMC expression and ACTH secretion in the pituitary after stress. Thus, although it is possible to define the importance and characteristics of opioid actions at multiple sites in

the neural networks regulating the HPA axis, the net effect on the output is not easy to predict. The experimental finding that chronic prenatal exposure to morphine can dampen the ACTH response of the offspring to stress when they become adult indicates that the baby born to an opiate-addicted mother may be programmed to respond less to stress.

In summary, endogenous opioid mechanisms are activated by stress and modify HPA axis responses to stress at several levels in the axis. The activity of these opioid mechanisms is not fixed, and this contributes to altered HPA responses to stress, for example, as a result of reproductive activity or chronic disease. β -Endorphin is secreted into the circulation when ACTH is secreted and may have distant actions through opioid receptors outside the brain.

Opioids in the Hypothalamic-Neurohypophyseal System

Oxytocin is secreted into the blood from the nerve terminals in the posterior pituitary gland of magnocellular neurons that have their cell bodies in the supraoptic nucleus and PVN in the hypothalamus. These neurons, and hence oxytocin secretion, are stimulated by emotional and physical stressors in the rat (but generally not in other species), in which vasopressin is the posterior pituitary stress hormone. Although oxytocin may act directly on the adrenal cortex to enhance basal corticosteroid secretion, its role in the periphery in coping with stress remains unclear. Oxytocin secretion is inhibited by opioids. In the rat, vasopressin is secreted into the blood (from magnocellular neurons adjacent to oxytocin cells) in response to noxious stimuli and probably mediates cardiovascular compensation as well as conserving water by its renal actions; endogenous opioids can restrain the vasopressin response. Other stressors, such as conditioned fear, inhibit vasopressin secretion, but opioids are not responsible for this.

Opioid peptides and opioid receptor expression in the hypothalamic-neurohypophyseal system Oxytocin neurons coexpress the endogenous opioids, enkephalins and dynorphins, in both the supraoptic nucleus and PVN. These opioids are costored in the nerve terminals in the posterior pituitary and are coreleased with oxytocin on stimulation, and thus they can autoregulate oxytocin secretion into the blood (Figure 1, i). Opioids originating from the nearby anterior and intermediate lobes of the pituitary are evidently not important. Endogenous opioids also come from the adjacent vasopressin neurons, which coexpress and secrete dynorphins from their nerve terminals in the posterior pituitary and from inputs

to the oxytocin neuron cell bodies from the arcuate nucleus (β -endorphin) or brain stem (enkephalins). The magnocellular cell bodies of oxytocin neurons in the supraoptic nucleus and PVN have κ - and μ -opioid receptors, but their nerve terminals in the posterior pituitary have only κ -opioid receptors; vasopressin neurons are sensitive to κ -opioids only. Although Met⁵- and Leu⁵-enkephalins are not active at κ -receptors, the extended forms Met-enkephalin-Arg⁶Phe⁷ and Met-enkephalin-Arg⁶Gly⁷Leu⁸ are; these are produced by oxytocin neurons, and like dynorphins they can inhibit stimulated oxytocin secretion from the posterior pituitary. δ -Opioid receptors are not evident in the hypothalamic-neurohypophyseal system (HNS) and δ -acting opioids are not important in its regulation. Inputs to the magnocellular supraoptic nucleus and PVN arise from other areas of the hypothalamus, the brain stem, the circumventricular organs, the BNST, and the olfactory bulb. Although most of these regions express opioid peptides and opioid receptor binding, and opioids could affect the activity of these inputs, their role in oxytocin and vasopressin responses to stress is not known.

Endogenous opioids in hypothalamic-neurohypophyseal system responses to stress Oxytocin neuron activity and secretion are inhibited by exogenous or endogenous opioids under stimulated conditions in males and females. Thus naloxone increases and κ - or μ -opioids inhibit oxytocin secretion in response to parturition, suckling, hyperosmotic stimuli, and stress (e.g., forced swim, restraint, surgery, or white noise), the major physiological conditions that increase oxytocin secretion. In the supraoptic nucleus, opioids can act directly on the oxytocin neurons, inhibiting their discharge of action potentials via κ - or μ -opioid receptors, and act presynaptically on μ -opioid receptors to inhibit the release of neurotransmitters from brain-stem inputs. However, the inhibition of oxytocin secretion by endogenous opioids is so far known to occur only at the oxytocin nerve terminals in the posterior pituitary, except in pregnancy. The mechanisms of opioid action here are similar to those on other nerve terminals. Opioid antagonists such as naloxone increase oxytocin secretion into the blood in response to stressors in the rat. Researchers believe that κ - and μ -opioid receptors are differentially involved in regulating responses to different stressors in the rat; thus, endogenous μ -opioids regulate oxytocin secretory responses to immobilization, whereas endogenous κ -opioids regulate responses to intraperitoneal hypertonic saline. The μ -opioid receptor actions are likely to be in the brain rather than the posterior pituitary, but the

sites of action of endogenous opioids during stress are unknown.

There is plasticity in the opioid regulation of the HNS. Acute hyperosmotic stress induces the secretion of oxytocin and vasopressin and upregulates proenkephalin (weakly) and prodynorphin gene expression in the supraoptic nucleus. However, chronic stress in rats more strongly upregulates preproenkephalin A gene expression in supraoptic neurons. In pregnancy, oxytocin secretory responses to stress are attenuated more strongly by endogenous opioids. A central site of opioid inhibition is likely at this time because there is a downregulation of κ -opioid receptors and a decreased sensitivity to κ -opioids in the posterior pituitary. The upregulated endogenous opioid mechanism in pregnancy involves increased proenkephalin A and μ -opioid receptor mRNA expression in the brain-stem input and, possibly, increased β -endorphin action from arcuate nucleus neurons. The upregulation of endogenous opioid inhibition of oxytocin neurons in pregnancy is driven by neuroactive steroid products of progesterone metabolism in the brain. The more powerful opioid restraint of oxytocin neurons that emerges in pregnancy acts to restrain the responses of oxytocin neurons to a range of stressors, including interleukin-1 β . This opioid-induced resistance of the oxytocin system to stressors, and especially immune signals, in pregnancy may be important in preventing premature activation of parturition. Paradoxically, stress slows parturition by decreasing oxytocin secretion in a naloxone-reversible way. Stress may still have a stimulatory effect on oxytocin secretion in parturition, but may disrupt the pulsatile pattern of secretion that is essential for parturition (see **Oxytocin**).

In the rat, vasopressin secretion from the posterior pituitary into the blood is restricted to selected stressors – for example, to electrically applied foot shock, surgical stress, and noxious stimuli but not to restraint, forced swimming, noise, and novel environment exposure. Opioids generally inhibit magnocellular vasopressin secretion via central κ -opioid receptor mechanisms in rats and humans. Partly, this reflects the autoregulatory action of dynorphin, which is evidently released by the dendrites of vasopressin neurons during periods of electrical activity (phasic bursts) and, acting through κ -opioid receptors on the neurons, eventually terminates each burst. Otherwise, the effects of endogenous opioids on stress-induced vasopressin release are dependent on the stressor. For instance, naloxone augments the vasopressin secretory response to foot shock. Endogenous opioids are not responsible for the lack of vasopressin secretion in response to stressors such as forced swimming or restraint in the rat and this does

not change in altered physiological conditions such as pregnancy.

In summary, for posterior pituitary hormone secretion, and on stimulated oxytocin secretion in particular, endogenous opioids have powerful moderating actions at several levels under stressful conditions. These actions can be seen as contributing to a successful outcome of pregnancy.

Opioids and the Adrenal Medulla

The stimulation of the sympathetic nervous outflow from the spinal cord drives the secretion of epinephrine from the adrenal medulla chromaffin cells as part of the flight-or-fight response to stress. The adrenal medulla is an important site of opioid modulation of stress responses.

Endogenous opioid and opioid receptor expression in the adrenal medulla Soon after the discovery of the endogenous opioids, it was recognized that the adrenal medulla produces abundant opioids. In particular, proenkephalin A is expressed by chromaffin cells, which produce the catecholamines, epinephrine, and norepinephrine. These are secreted into blood when chromaffin cells are excited through nicotinic receptors by their cholinergic preganglionic sympathetic innervation from the spinal cord, that is, in acute stress (**Figure 1, ii**). Thus, the phenylethanolamine *N*-methyltransferase (the enzyme synthesizing epinephrine) and preproenkephalin A genes are expressed in the same cells. The preproenkephalin A protein is processed by a serine protease to yield Met⁵- and Leu⁵-enkephalins but also the extended peptides Met-enkephalin-Arg⁶Phe⁷ and Met-enkephalin-Arg⁶-Gly⁷Leu⁸. Enkephalins and epinephrine are packaged in the same secretory vesicles, and the stimulation of the chromaffin cells with acetylcholine or nicotine increases the secretion of both epinephrine and enkephalins into the blood, with the possibility of distant opioid actions. Such excitation increases the expression of the preproenkephalin A gene and the subsequent synthesis of enkephalins, so the secretion of opioid can be sustained during continual stimulation. Pro-inflammatory cytokines stimulate or enhance the release of enkephalins from the chromaffin cells, indicating that production is increased with infection. The capacity of adrenal chromaffin cells to produce opioids is demonstrated by the analgesia caused by fragments of adrenal medulla implanted adjacent to the dorsal horns of the spinal cord or in the periaqueductal gray (PAG) in experimental models of chronic pain. POMC is also expressed. Only a few adrenal medullary cells express the prodynorphin gene, except when these cells become tumorous (pheochromocytoma), when dynorphin is

abundantly produced. Nociceptin is also produced. There are opioid receptors in the adrenal medulla, κ -opioid receptors are present on the nerve fiber tracts and in the epinephrine-producing chromaffin cells, δ -opioid receptors are present on the norepinephrine-producing cells, and μ -opioid receptors are present at a low level. The adrenal medulla thus provides a microcosm to investigate the pre- and post-synaptic actions of opioids on neurons.

Endogenous opioids and adrenal medulla stress responses

A presynaptic action of opioids has been shown with *in vivo* microdialysis of the adrenal medulla, revealing the inhibition by μ - and δ -opioid agonists of stimulated acetylcholine release. The cholinergic nerve terminals probably release enkephalin when stimulated, causing autoinhibition at these terminals as well as inhibition by enkephalin released from the chromaffin cells (Figure 1, ii). Exogenous opioids inhibit catecholamine (norepinephrine) release from chromaffin cells induced by splanchnic nerve stimulation or by nicotine, but this is not reversed by naloxone, and indeed opioid antagonists have a similar effect, so classic opioid receptors do not seem to mediate this action on the strong stimulation of the chromaffin cells. These actions of both opioid agonists and antagonists involve the inhibition of nicotine-stimulated Ca^{2+} entry, essential for exocytosis of the contents of the secretory granules. However, other studies at the level of individual cells show clear stereospecific, opioid receptor-mediated actions. Opioids released by chromaffin cells act through the δ - or μ -opioid receptors on these cells to inhibit secretion by reducing the opening of voltage-activated non-L type (N or P/Q) Ca^{2+} channels through a pertussis-toxin-sensitive G-protein but also by inhibiting L-type channels, thus reducing the Ca^{2+} signal for exocytosis. The effects on N or P/Q channels are voltage dependent, that is, are evident when the cells are activated by depolarization. These effects last for several seconds after the secretion of enkephalin. In addition, μ -opioids activate large conductance Ca^{2+} -dependent K^+ channels (BK), which reduce Ca^{2+} influx by decreasing the action potential duration when the chromaffin cells are excited by sympathetic nerve stimulation. In this case, there is a tight coupling between the μ -opioid receptor and channel, not involving the G-protein or phosphorylation. Together, these opioid-mediated negative feedback mechanisms (in conjunction with similar ATP-mediated actions) tonically brake the secretory activity of the chromaffin cells. However, with strong stimulation, the voltage-dependent effects of the opioids are lost; thus, the secretion of catecholamines (and enkephalins) is facilitated. The

coproduction of enkephalins may thus ensure that there is minimal secretion of catecholamine from the chromaffin cells under basal conditions, but may permit a sharp increase when the animal responds to stress by increasing the frequency of impulses in the sympathetic nerves controlling the adrenal medulla.

In summary, enkephalins cosecreted into the circulation with epinephrine when the adrenal medulla is stimulated can have distant actions. Within the adrenal medulla, enkephalins automatically moderate or inhibit epinephrine release, except when the chromaffin cells are strongly stimulated.

Opioids and Stressor Processing

Pain

The origin of pain is normally at the site of intense stimulation or tissue damage, causing the discharge of impulses in a specialized set of nerve endings (nociceptors). The extent of excitation of these nerve endings is related primarily to the intensity of the painful stimulus, but this is regulated by local chemical signals, including opioid peptides, released when tissue is damaged. Information about painful stimulation is carried as impulses in afferent nerve fibers (especially the small diameter, unmyelinated, and slowly conducting C fibers and the myelinated A δ fibers, responsible for dull and sharp pain, respectively). These synapse in the dorsal horn, in the substantia gelatinosa (lamina II) and lamina I, respectively, of the spinal cord. Through their long axonal projections in the spinal cord, postsynaptic neurons convey the nociceptive information to the brain. Transmission at these synapses is subject to modulation by locally released neurotransmitters, including opioid peptides. Modulation in the dorsal horn alters the perception of the pain and its impact as a stressor, enhancing (hyperalgesia) or decreasing (analgesia) its effects. The application of morphine to the spinal cord (intrathecal injection or infusion) is highly effective in producing analgesia.

Endogenous opioid and opioid receptor expression in pain-mediating cells

There is a high density of all three types of opioid receptors on the nerve fibers in laminae I and II of the dorsal horn. The C (or A δ) fiber neurons have their cell bodies in the dorsal root ganglia and express μ -opioid receptors in particular, but also δ -receptors, transporting the receptors into both their peripheral and central processes (Figure 1, iii). Postsynaptic neurons also have opioid receptors, including κ -opioid receptors. Opioids (enkephalins) in the substantia gelatinosa are produced by the C fiber neurons themselves and by local neurons in the dorsal

horns. Opioids in the dorsal horns include enkephalins, dynorphins (especially in the substantia gelatinosa), and endomorphin 2 (in the nerve fibers of the afferent dorsal root ganglion neurons). Presynaptic opioid actions in the spinal cord involve hyperpolarizing the terminals by activating K^+ channels so that there is less transmitter released when the impulses triggered in the periphery arrive. The main excitatory transmitter released from the C (or A δ) fiber neurons is glutamate (acting through NMDA receptors), but its effects are enhanced by the coreleased peptide substance P. Substance P and δ -opioid receptors are found together in the secretory vesicles in the nerve terminals of dorsal root ganglion neurons in the dorsal horn. Thus, the opioid receptors are exposed for opioid binding as the nerve terminals release their excitatory neuropeptide. The inhibition by opioids of the release of these excitatory transmitters consequently brakes the activation of relay neurons in the spinal cord that would otherwise have been excited further. Opioids also inhibit the postsynaptic neurons directly, which are probably excitatory interneurons. These oppose the excitatory actions of glutamate through NMDA receptors and, for κ -agonists, also oppose glutamate actions via AMPA receptors.

Opioids in inputs to pain-processing systems The brain influences the processing of nociceptive input in the spinal cord, partly by activating the opioid interneurons in the dorsal horns. There are descending fibers, in the dorsolateral funiculus of the spinal cord, from neurons in the brain stem to the dorsal horns (Figure 1, i, iii). These neurons are in the raphe nuclei and locus coeruleus, and they receive input from the PAG and, in turn, from the reticular formation, limbic system, cerebral cortex, and thalamus. They produce monoamines as their transmitters (serotonin, 5-HT; or norepinephrine), and opioids are important in regulating their activity as well as their actions in the spinal cord. When these areas are stimulated, especially the PAG, opioids are released in the dorsal horns, which act via δ -, κ -, or especially μ -opioid receptors and inhibit nociceptive neurons; analgesia results, which is reversed by naloxone. Opioid application in these brain-stem regions has the same effect as stimulation. Hence, endogenous opioids acting on these groups of neurons in the brain, via δ -, κ -, or μ -opioid receptors, are important in controlling or filtering indirectly the flow of nerve impulses from the spinal cord that cause the sensation of pain. Opioids may act directly on these neurons or on the terminals of their afferent neurons; for example, met⁵-enkephalin acts like this after its release from presynaptic terminals on locus coeruleus neurons. Appropriately, the descending mechanism is activated

by stressors that are likely to lead to pain, giving rise to stress-induced analgesia. Pain pathways continue to ascend to the thalamus and onward to the limbic system and cerebral cortex, affecting mood and bringing the sensation of pain into consciousness. The analgesic actions of morphine include actions at the level of the limbic system and cerebral cortex, which modify further the impact of a painful stressor on the activity of the HPA axis and other stress response mechanisms.

Stress-Induced Analgesia

It has long been recognized that in the heat of battle the perception of pain from an injury can be suppressed. Experimental studies on rats, first reported in 1976, showed that inescapable foot shock increases the pain threshold and that this is partly reversed by naloxone, indicating the involvement of endogenous opioids. The increase in pain threshold is generally tested with a different stimulus from that inducing analgesia, such as, the measurement of tail-flick latency. The mechanism of this stress-induced analgesia involves the pathway previously described, from the brain to the spinal cord; thus, other neurotransmitters involved include monoamines (norepinephrine, 5-HT, and histamine). Microinjections of antagonists have shown that opioid mechanisms (δ - and κ -, as well as μ -, opioid receptors) are activated in the brain stem and spinal cord (Figure 1, i, iii). There are differences between individuals or strains of rats or mice in the analgesic response to a stressor, and these are related to genetic factors and the density of μ -opioid receptors in the brain. Also, transgenic mice that have a disruption of the prodynorphin gene, or that do not produce β -endorphin, do not show swim-stress-induced analgesia, which is also suppressed in normal mice by the administration of a κ -opioid receptor antagonist. Stress-induced analgesia can be conditioned, and the conditioned response (analgesia) is prevented by naloxone; thus, the expectation of a painful or stressful stimulus evidently activates brain opioid mechanisms to reduce the anticipated pain. In particular, δ - and μ -opioid receptors (rather than κ -opioid receptors) are involved. An analogous opioid-mediated analgesic effect of anticipated pain has been shown in humans. Such conditioning involves higher brain levels. In contrast, lesion studies show that the PAG and structures more rostral are not necessary for analgesia induced by physical pain (foot shock), which is mediated via the brain stem, particularly the nucleus raphe alatus, and by the spinal cord. Presumably, the ascending information about the foot shock pain stimulates the opioid mechanisms in the brain stem.

Because a stress response (adrenocortical activation) is not essential for stress-induced analgesia,

although basal corticosteroid secretion has a permissive role, it has also been termed adaptive pain suppression. The involvement of opioids in this analgesia varies according to the pattern of stressful stimulation (naloxone reverses analgesia following intermittent but not continuous foot shocks) and its nature (analgesia with tepid- but not cold-water swimming is naloxone reversible and shows morphine cross-tolerance). Clearly, endogenous opioids are important in the analgesia produced by some, generally less persistent and less intense, stressors but not others. In contrast, nociceptin can block supraspinal analgesia, and the potentiation of mild stress-induced analgesia by an ORL-1 receptor antagonist indicates that endogenous nociceptin limits this analgesia. Nonopioid analgesia is a consequence of activation by other transmitters of the descending pathway that opioids can also activate; evidently, endocannabinoids act in the PAG to mediate nonopioid stress-induced analgesia. Lesion of the arcuate nuclei (the location of β -endorphin neurons) reduces opioid-mediated stress-induced analgesia, and transgenic mice with an inactivation of the POMC (β -endorphin) gene do not show normal swim-stress analgesia. In contrast, mice with an inactivation of the preproenkephalin A gene still show opioid-mediated stress-induced analgesia. Thus, β -endorphin, rather than an enkephalin, is the important opioid peptide in the brain.

In mice, social defeat induces analgesia, which is due to endogenous opioid action because it is prevented by naloxone and shows cross-tolerance to morphine. Because adrenalectomy or hypophysectomy does not affect this stress induction of analgesia, it is caused by opioids within the brain rather than by circulating opioids. In contrast, the removal of the pituitary gland or adrenal medulla (sources of β -endorphin and enkephalins, respectively) reduces the analgesia induced by foot shock or immobilization. Evidently, circulating opioids as well as opioids released within the brain can be involved in this analgesia.

An inescapable stressor can evoke learned helplessness (effects caused by the uncontrollability of events that are beyond the organism's control rather than by the events *per se*) as well as analgesia, and this too is opioid-mediated. The inhibition of motor activity, together with a reduced sensitivity to pain, may help the individual to cope or survive. Posttraumatic stress disorder may be modeled by the consequences of exposing an animal to inescapable electric shocks, with opioid-mediated stress-induced analgesia as a common feature.

In summary, pain control by opioids acting within the brain and spinal cord involves automatic feedback control at synapses in pain pathways *pari*

passu with noxious stimulation and activation of opioid neurons according to previous experience. The consequence is that these opioid mechanisms contribute to reducing pain sensitivity temporarily, for instance, during acute stress. Morphine mimics these actions to suppress pain.

Opioids and Immune Mechanisms: Inflammatory Pain

The functions of immune cells are modified significantly by the actions of glucocorticoids, and these actions are a major feature of the body's response to stress. However, the chemical messengers, including opioids and CRH, involved in the hypothalamus in regulating ACTH secretion also have important actions directly on immune cells.

Endogenous opioid and opioid receptor expression in immune cells The immune cells themselves produce opioids. T and B lymphocytes, monocytes, and macrophages contain the mRNAs for β -endorphin and met-enkephalin, and produce the peptides; expression is increased at sites of inflammation, and, significantly, immune cells contain pro-protein convertases, which process the opioid precursors. Furthermore, immune cells express the CRH1 receptor and CRH that is released locally from nerve endings stimulates inflammatory responses. One effect of CRH is to stimulate the release, from immune cells, of opioid peptides (especially β -endorphin, a product of the POMC precursor), particularly from lymphocytes migrating to the site of injury; some cytokines act similarly, as does lipopolysaccharide (bacterial endotoxin). The released opioid can then act locally to regulate the inflammatory process (Figure 1, iv).

Opioid effects on immune responses Morphine affects both cell-mediated and humoral immune mechanisms. Opiate (heroin) addicts have impaired immune mechanisms. This has implications for the progression of human immunodeficiency virus (HIV) infection in such individuals. Similarly, morphine has immunosuppressive effects in mice. *In vitro* studies showed that opioids (e.g., met-enkephalin acting through μ - or δ -opioid receptors) have direct inhibitory effects on a range of functions of immune cells, including the inhibition of killer activity (NO production) in macrophages and leukocytes. Lymphocytes express δ -, κ -, and μ -opioid receptors, and responses are mediated by $G_{i/o}$ -proteins and effects on calcium mobilization. Cytokines are signaling peptides important in the local regulation of immune cell function, and opioids inhibit the cytokine-induced activation and chemotaxis of granulocytes, acting to

desensitize cytokine receptors. However, the rapid effects of opioids can promote immune cell activity, and morphine or opioid peptides (β -endorphin or met-enkephalin) act as chemoattractants to leukocytes. Also, enkephalins stimulate the proliferation of B and T lymphocytes.

In addition, opioids released by immune cells, especially by polymorphonuclear leukocytes, at the site of inflammation have local analgesic actions, hence modulating stressful stimulation. Opioids, including morphine and enkephalins, applied to sites of injury, act through μ -opioid receptors (δ -opioid receptors are also expressed) on the nerve terminals, responding to painful stimuli (nociceptors) and to the chemicals released by local inflammatory cells, and reduce pain (Figure 1, iv). These antinociceptive actions are more effective, especially via μ -opioid receptors, if there is also local inflammation. This increased effectiveness is a result of increased opioid receptor expression in the C and A δ nerve-fiber terminals. It is partly a result of the disruption of the nerve sheath so that opioids have easier access. Opioids act on the nerve terminals to reduce their excitability, probably by opening K^+ channels, hyperpolarizing the terminals. This reduces both the initiation of nerve impulses and the release from these nerve terminals of chemical messengers (e.g., substance P), which promote the inflammatory process. The inhibitory action of opioids on immune cells will reduce their release of chemical pain messengers. Stress, such as swimming in cold water, activates the opioid antinociceptive mechanism in inflamed tissue. This perhaps indicates that circulating opioids, released during stress, may act peripherally at such sites (Figure 1, iv). The duration of action of peripherally released opioid peptides is limited by their inactivation by extracellular metalloproteinases, produced by immune cells themselves. Thus, it seems that the stimulated local release of opioids in the initial stages of inflammation may first promote activity of the immune cells and then dampen the response. At the same time, these local opioids reduce inflammatory pain. The nervous and immune systems interact in these processes. The development of drugs to target the peripheral opioid-mediated analgesic mechanisms may afford a new approach to controlling inflammatory pain.

Chronic inflammation leads to changes in the opioid mechanisms that modulate the processing of nociceptive input in the spinal cord. The expression of opioid receptors and opioid peptides in the dorsal horn afferents and neurons shows plasticity. In experimental polyarthritis (a stressful, painful chronic inflammatory condition in which pain sensitivity is locally increased, hyperalgesia), preproenkephalin and prodynorphin mRNAs in the dorsal horn are

strongly upregulated and the release of dynorphin in the spinal cord is increased. However, preproenkephalin mRNA expression in dorsal root ganglion neurons is reduced, as is enkephalin release in the spinal cord. Hence, the modulation of pain input in chronic inflammation changes from enkephalin to dynorphin mediation at this level. Nociceptin expression in different dorsal root ganglion cells is increased, and the proposed pronociceptive actions of this peptide may contribute to the hyperalgesia. There is an increase in the number of μ -opioid receptors (and a decrease in δ - and κ -opioid receptors) in the dorsal root ganglion neurons, but this may not result from increased gene expression. Opioid receptor upregulation in the dorsal horn lamina II does involve increased μ - and δ -opioid receptor gene expression in neurons in laminae I and II. These changes may explain the rapid increase in the effectiveness of morphine in inhibiting nociceptive dorsal horn neurons after the induction of focal peripheral inflammation. Clearly, because the chronic inflammatory condition is painful, the changes described do not succeed in preventing such pain.

In summary, the peripheral and central actions of opioids in the brain and spinal cord are complementary in controlling the transmission of information through pain pathways from a site of injury or inflammation. Opioids released at the site of inflammation, or arriving from the anterior pituitary or adrenal medulla, reduce the activity of immune cells and have local analgesic actions. Endogenous opioids reduce the perceived strength of the damaging stressor and thus dampen the HPA axis response to stress; with chronic inflammation, this mechanism may not function.

Opioids and Hedonic Appetite in Stress

It has been known for a long time that opiates and endogenous opioids stimulate the desire for sweets and that this craving is also a feature of stress. Naloxone reduces this desire in humans and in objective tests in laboratory animals. Hence, it has been proposed that the activation of endogenous opioid mechanisms is involved, when an individual is under stress, in enhancing the incentive value of rewarding stimuli, including the pleasant sensations from eating sweet and fatty foods. Consequently, endogenous opioids and their receptors are implicated as a cause of obesity due to hedonically driven overeating in chronically stressed individuals. A consequence of such overeating in animal models is the reduction of the activation of the HPA axis, evidently by a signal from the enlarged abdominal fat store.

Evidence that endogenous opioids are released as a result of consuming sweet foods includes the finding

that this results in analgesia reversible by naltrexone, a long-acting version of naloxone. Studies on mice with transgenic manipulations of the POMC and proenkephalin-A genes so they do not produce β -endorphin or enkephalins showed a reduced incentive to obtain sweet nutrient over and above their need for calories; these findings implicate both of these endogenous opioids in the stimulation of hedonic appetite. In contrast, although nociceptin can also stimulate feeding, it does not stimulate hedonic appetite, so is unlikely to be involved in increasing the desire for sweets and fatty food in stress.

The causal relationship between the activation of brain endogenous opioid mechanisms and the palatability of food or drink extends to the desire to drink alcohol in susceptible individuals; accordingly, naltrexone can be prescribed for the treatment of alcoholism. Further evidence that some actions of nociceptin are functionally opposed to those of opioids is that nociceptin given in the brain in rats acts like naloxone to reduce the stress-induced reinstatement of alcohol-seeking behavior in rats with a previously established preference for alcohol. This has prompted research to develop a nonpeptide ORL-1 agonist to treat alcoholism.

The involvement of endogenous opioids in the normal regulation of appetite and food intake without stress, studied for over 40 years, is complex. Multiple opioid peptides and multiple receptors are involved at several sites in the brain. For instance, male mice genetically engineered so they do not produce β -endorphin eat more and become obese, indicating an anorexigenic role for this opioid in normal appetite regulation and contrasting with its orexigenic actions on hedonic appetite. Because naloxone still reduces food intake in these mice, it is clear that opioids other than β -endorphin usually mediate the orexigenic actions that have previously been suspected to be due to β -endorphin. Sites in the brain where opioids have been shown to stimulate eating have been identified by seeking evidence of opioid modulation of the activity of pools of neurons in discrete brain areas known to be involved in appetite regulation (Figure 1, v). These studies involved the microinjection of opioid receptor agonists and antagonists into these brain regions, sometimes into different sites in the same animal, to investigate opioid-mediated interconnections. The effects of such local opioid interventions on appetite, measured in a range of tests with food intake as an end point, and on the activities of neurons in the treated area have then been studied. Neuronal activity has been evaluated principally by the postmortem identification of stimulated immediate-early gene expression, using Fos immunocytochemistry. Of particular interest in the

context of stress-induced eating is that fat or sugar intake is stimulated by the microinjection of μ - or δ -, but not κ -, opioid agonists into the shell of the nucleus accumbens. This nucleus receives a dopaminergic input from the ventral tegmental area and is a key part of the mesocorticolimbic system, the reward circuitry that is implicated in drug dependence, including alcoholism. Opioids act in the shell of the nucleus accumbens, via μ - and δ -opioid receptors to increase the release of dopamine, which underlies hedonic sensation. Here the local injection of a μ -opioid receptor agonist stimulates the ingestion of a sweet or fatty meal, and especially a fatty one – increasing almost threefold the amount of fat ingested. The double-injection technique (opioid agonist microinjected into one site, and opioid antagonist microinjected into another) has shown there are opioid-mediated interactions between the nucleus accumbens and the central nucleus of the amygdala, with a μ -opioid agonist stimulating feeding and reciprocal opioid connections between the nucleus accumbens shell and the ventral tegmental area. This reciprocal opioid connection promotes feeding via μ - and δ -opioid receptors. Naltrexone, given systemically, activates neurons in the nucleus accumbens shell, as indicated by Fos expression, and prevents the activation of presumably different neurons in this structure that otherwise follows the ingestion of a palatable meal. Appropriately, the consumption of a sweet, fatty meal activates central β -endorphin neurons. These opioid-driven mechanisms for individuals preferring to ingest energy-rich sweet and fatty foods, notably when they are stressed, can be seen as having a survival advantage in that energy storage is favored. However, with the superabundance of such foodstuffs in advanced societies, these opioid-based reward mechanisms evidently contribute to obesity as a consequence of stress.

The involvement of multiple opioids and their multiple receptors in appetite control, and the multiple other roles of endogenous opioid mechanisms, means that opioid antagonists are insufficiently precise for use in obesity treatment. Instead, recent discoveries have revealed the roles of endocannabinoids and several nonopioid neuropeptides in the regulation of appetite, including α -melanocyte stimulating hormone (MSH), which, like β -endorphin, is a POMC derivative produced by arcuate nucleus neurons, and these have become the focus of attention as novel drug targets.

In summary, endogenous opioids regulate activity in the reward circuitry in the brain, increasing appetite for sweet and fatty foods and the desire for alcohol. The seeking of such hedonic rewards is increased by stress, and opioids drive this behavior.

Opioids, Stress, and Reproduction

Opioids and the Hypothalamic-Pituitary-Gonadal Axis

Reproductive capability depends primarily on the intermittent, approximately hourly, synchronized discharge of gonadotropin-releasing hormone (GnRH) neurons in the anterior hypothalamus/preoptic area. The GnRH so released into the hypothalamo-hypophysial portal system drives gonadotropin secretion from the anterior pituitary gland, which in turn stimulates the sex steroid- and gamete-producing activities of the ovaries or testes. The suppression of this GnRH pulse generator leads to failing gonadal function and to diminishing reproductive capacity or sexual desire. The striking consequence in females of many species is the failure of ovulation and in women the accompanying cessation of the menstrual cycle. Because these effects are a consequence of the suppression of the GnRH pulse generator, this is hypothalamic amenorrhea. Stress depresses reproductive activity in this way, although actions of stress hormones on the anterior pituitary gonadotrophs or gonads may also be involved. Stress inhibits the activity of GnRH neurons through the indirect actions of CRH within the brain via an opioid mechanism in both males and females.

Endogenous opioid effects on gonadotropin secretion during stress: males and females There are many β -endorphin terminals, along with CRH nerve terminals, in contact with, or in the vicinity of, the GnRH neurons (Figure 1, i). It is important to note that endogenous opioids also play an important role in the normal regulation of GnRH neurons in active reproductive life, being involved in the interplay between the sex steroid hormones and the neurons. Furthermore, it does not follow that because opioids have been involved in suppressing reproductive function that simply antagonizing opioid actions will reactivate GnRH neurons. Different processes are involved in reactivation. Hence, in women with amenorrhea associated with intensive physical training, naloxone may have no effect. The suppression by restraint stress of the luteinizing hormone (LH) surge in rats that occurs every 4–5 days and triggers ovulation is evidently not caused by endogenous opioids because naloxone does not prevent this.

An acute stressor in intact males can increase LH and testosterone secretion transiently, perhaps aiding aggressive responses. In castrated male rats, foot shock stress decreases LH secretion rapidly, which is reversed by the injection of antiserum to β -endorphin or dynorphin or an irreversible μ -opioid receptor

antagonist into the cerebral ventricles, with no effects of δ -opioid receptor ligands. β -Endorphin or CRH injected centrally inhibits LH secretion. Furthermore, CRH is not effective after β -endorphin antiserum, so the inhibition by CRH evidently requires β -endorphin neurons. However, ablation of the PVN, and hence of the parvocellular CRH neurons, does not prevent the foot-shock-induced inhibition of the GnRH pulse generator in male rats and CRH neurons in the PVN evidently do not project to the GnRH neurons, so CRH neurons elsewhere may exert the inhibitory effects of stress on reproductive function. Also, calcitonin gene-related peptide (CGRP), given centrally, suppresses LH pulses, and this action, which is partly via the CRH neurons, is prevented or reversed by naloxone, indicating the mediation of CGRP actions by opioids. In male rats, the inhibition of LH secretion by acute (but not chronic) restraint is reversible by naloxone and hence involves endogenous opioids. Similarly, in female marmosets, aggression by a conspecific causes the naloxone-reversible inhibition of the GnRH pulse generator.

Fasting and insulin-induced hypoglycemia, produced experimentally, are stressors, and they reduce the frequency of GnRH pulses. This reflects a generally inhibitory effect of reduced energy availability on the GnRH pulse generator, so that reproductive activity is suppressed when it would strain further the individual's ability to support its own needs. Leptin is an important signal from adipose tissue in this regard, and the suppression of central kisspeptin production is involved; kisspeptin is a hypothalamic peptide essential for the reactivation of GnRH neurons at puberty. The inhibition of the GnRH pulse generator by fasting or hypoglycemia, which is evident only in females or males with sufficient estrogen or testosterone, respectively, is mediated through μ -opioid receptors; a similar suppression in rhesus monkeys does not involve opioids. The requirement for sex steroids may relate to their stimulatory effects on POMC neurons in the arcuate nucleus. CRH can, with vasopressin, stimulate the release of β -endorphin from these POMC neurons, and this could be the link for the hypoglycemic inhibition of the GnRH pulse generator. Opioids from the adrenal medulla may be involved because the removal of the adrenal medullae abolishes this inhibition. Glucose deprivation stimulates PVN CRH neurons by increasing norepinephrine release, inhibiting GnRH neurons indirectly. The inhibitory actions of CRH on the cell bodies of the GnRH neurons are partly blocked by local μ -opioid receptor antagonists, and thus endogenous opioids released in the vicinity may mediate these CRH actions. However, GnRH neurons do not express

opioid receptor genes, so opioid actions are indirect. Specifically, opioid action in the medial preoptic area seems important, where β -endorphin-containing terminals are found, with noradrenergic and neuropeptide Y (NPY)-containing terminals, on neurons projecting to the PVN. Dynorphin and enkephalin-containing terminals are also found in the vicinity of GnRH neurons. In addition, CRH can act by stimulating endogenous opioid release (β -endorphin or met-enkephalin) in the median eminence, with consequent inhibition, via μ -opioid receptors, at the level of the GnRH nerve terminals. It is not known which endogenous opioid neurons are activated by stress to inhibit GnRH neurons.

Endogenous opioid effects on gonadotropin secretion during stress: males In males, stress decreases the circulating testosterone levels, as we expect from the inhibition of LH secretion. Leydig (interstitial) cells of the testis are the principal source of the male sex steroid hormone, testosterone, but also express the POMC gene, producing β -endorphin under stimulation by gonadotropin along with dynorphins. Sertoli cells (nurse cells for spermatogenesis in the seminiferous tubules) have opioid receptors and produce enkephalins. Opioids affect spermatogenesis through Sertoli cells and testosterone production via any Sertoli cell effects on the nearby Leydig cells. Locally administered naloxone can increase testosterone production. These mechanisms are evidently activated during stress so that immobilization for several hours decreases the circulating level of testosterone by approximately half, due to an action of opioids within the testes involving reduced sensitivity to gonadotropins. It is not clear how the stressor can activate intratesticular opioid mechanisms, but this seems to be a way in which the sex drive may be indirectly reduced during stress.

In summary, endogenous opioids, in particular β -endorphin, mediate the suppression of the pulsatile activity of GnRH neurons and hence mediate the inhibition of gonadotropin secretion and gonadal function that may occur with stress.

Conclusion

In sum, endogenous opioids act as censors, forbidding excessive stimulation by noxious inputs, opposing the actions of chemical messengers acting on neural circuits that favor greater responses to stressors. Endogenous opioids can make the outside world seem less hostile, albeit less sexually interesting, and increase the appetite for rich food.

See Also the Following Articles

Hypothalamic-Pituitary-Adrenal; Anatomy of the HPA Axis; Oxytocin; Pain.

Further Reading

- Amit, Z. and Galina, Z. H. (1986). Stress-induced analgesia: adaptive pain suppression. *Physiological Reviews* **66**, 1091–1120.
- Appleyard, S. M., Hayward, M., Young, J. I., et al. (2003). A role for the endogenous opioid beta-endorphin in energy homeostasis. *Endocrinology* **144**, 1753–1760.
- Bodnar, R. J., Lamonte, N., Israel, Y., et al. (2005). Reciprocal opioid-opioid interactions between the ventral tegmental area and nucleus accumbens regions in mediating mu agonist-induced feeding in rats. *Peptides* **26**, 621–629.
- Brunton, P. J., Meddle, S. L., Ma, S., et al. (2005). Endogenous opioids and attenuated hypothalamic-pituitary-adrenal axis responses to immune challenge in pregnant rats. *Journal of Neuroscience* **25**, 5117–5126.
- Condren, R. M. and Thakore, J. H. (2001). Cushing's disease and melancholia. *Stress* **4**, 91–119.
- Coggeshall, R. E. and Carlton, S. M. (1997). Receptor localization in the mammalian dorsal horn and primary afferent neurons. *Brain Research Reviews* **24**, 28–66.
- Dickenson, A. H. (1995). Spinal cord pharmacology of pain. *British Journal of Anaesthesia* **75**, 193–200.
- Drolet, G., Dumont, E. C., Gosselin, I., et al. (2001). Role of endogenous opioid system in the regulation of the stress response. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **25**, 729–741.
- Fabbri, A. (1990). The role and regulation of testicular opioids. *Trends in Endocrinology and Metabolism* **1**, 117–120.
- Fields, H. L. and Basbaum, A. I. (1978). Brainstem control of spinal pain-transmission neurons. *Annual Review of Physiology* **40**, 217–248.
- Grossman, A. (1988). Opioids and stress in man. *Journal of Endocrinology* **119**, 377–381.
- Levine, A. S. and Billington, C. J. (2004). Opioids as agents of reward-related feeding: a consideration of the evidence. *Physiology & Behavior* **82**, 57–61.
- Mansour, A., Fox, C. A., Burke, S., et al. (1994). Mu, delta, and kappa opioid receptor mRNA expression in the rat CNS: an in situ hybridization study. *Journal of Comparative Neurology* **350**, 412–438.
- Mercer, M. E. and Holder, M. D. (1997). Food cravings, endogenous opioid peptides, and food intake: a review. *Appetite* **29**, 325–352.
- Meunier, J-C. (1997). Nociceptin/orphanin FQ and the opioid receptor-like ORL1 receptor. *European Journal of Pharmacology* **340**, 1–15.
- Olszewski, P. K. and Levine, A. S. (2004). Minireview: characterization of influence of central nociceptin/orphanin FQ on consummatory behavior. *Endocrinology* **145**, 2627–2632.

- Pasternak, G. W. (2004). Multiple opiate receptors: déjà vu all over again. *Neuropharmacology* 47(supplement 1), 312–323.
- Pechnik, R. N. (1993). Effects of opioids on the hypothalamo-pituitary-adrenal axis. *Annual Review of Pharmacology and Toxicology* 33, 353–382.
- Pecoraro, N., Reyes, F., Gomez, F., et al. (2004). Chronic stress promotes palatable feeding, which reduces signs of stress: feedforward and feedback effects of chronic stress. *Endocrinology* 145, 3754–3762.
- Rivier, C. and Rivest, S. (1991). Effect of stress on the activity of the hypothalamic-pituitary-gonadal axis: peripheral and central mechanisms. *Biology of Reproduction* 45, 523–532.
- Russell, J. A., Brown, C. H. and Caron, R. W. (1998). Endogenous opioids. In: Neill, J. D. & Knobil, E. (eds.) *Encyclopedia of reproduction* (vol. 1), pp. 1043–1060. San Diego, CA: Academic Press.
- Russell, J. A. and Brunton, P. J. (2005). Neuroactive steroids attenuate oxytocin stress responses in pregnancy. *Neuroscience* 138, 879–889.
- Salzet, M., Vieau, D. and Day, R. (2000). Crosstalk between nervous and immune systems through the animal kingdom: focus on opioids. *Trends in Neuroscience* 23, 550–555.
- Schafer, M., Mousa, S. A. and Stein, C. (1997). Corticotropin-releasing factor in antinociception and inflammation. *European Journal of Pharmacology* 323, 1–10.
- Stamford, J. A. (1995). Descending control of pain. *British Journal of Anaesthesia* 75, 217–227.
- Stefano, G. B., Scharrer, B., Smith, E. M., et al. (1996). Opioid and opiate immunoregulatory processes. *Critical Reviews in Immunology* 16, 109–144.
- Tilbrook, A. J., Turner, A. I. and Clarke, I. J. (2002). Stress and reproduction; central mechanisms and sex differences in non-rodent species. *Stress* 5, 83–100.
- Twitchell, W. A. and Rane, S. G. (1993). Opioid peptide modulation of Ca^{2+} -dependent K^{+} and voltage-activated Ca^{2+} currents in bovine adrenal chromaffin cells. *Neuron* 10, 701–709.
- Yeomans, M. R. and Gray, R. W. (2002). Opioid peptides and the control of human ingestive behaviour. *Neuroscience and Biobehavioral Reviews* 26, 713–728.

Optimism, Pessimism, and Stress

M F Scheier

Carnegie Mellon University, Pittsburgh, PA USA

C S Carver

University of Miami, Coral Gables, FL USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M F Scheier and C S Carver, volume 3, pp 99–101, © 2000, Elsevier Inc.

Models of Motivation and Optimism

Goals, Stress, and Optimism

Effects of Optimism on Well-Being

Future Directions

Glossary

<i>Coping</i>	Attempts to deal with obstacles to goal attainment and with losses.
<i>Cross-sectional study</i>	A study in which the predictor and outcome variables are assessed at same point in time.
<i>Expectancy-value theories</i>	Models of motivation in which expectancies and values (goals) interact to determine behavior.

Explanatory style

A person's characteristic way of explaining prior negative (and positive) outcomes.

Optimists

People who generally expect to obtain favorable personal outcomes in the future.

Prospective study

A study in which the predictor and outcome variables are assessed at some appropriate baseline and then the outcome variables are reassessed at a later time.

Stress

Experience that arises when people have difficulty moving toward their goals.

Models of Motivation and Optimism

Optimists are people who hold generalized positive expectations for the future. They expect good things to happen to them, not bad things. Pessimists are just the reverse. They expect to miss out on the good things in life and to experience things that are bad. Research suggests that individual differences in optimism and pessimism may play an important role in the manner in which people react to stressful circumstances. To understand the nature of the effects, it is important to understand the broader theoretical context in which research on optimism and pessimism is

embedded. It is also important to understand how goals and the pursuit of goals are involved in creating stress.

Expectancies are pivotal in most contemporary theories of optimism. Defining optimism and pessimism in terms of expectancies creates a link to expectancy-value theories of motivation (some of which date back to the first few decades of the twentieth century). Expectancy-value models assume that behavior occurs in order to attain desired values or goals. If engagement of effort is to occur, there must be a goal that matters enough (has enough value) to try to reach it. The other element in the equation is expectancy: Confidence or doubt that the goal will be obtained. If the person lacks confidence, again there will be no action. If the person loses confidence along the way, action will stop. Only when confidence is sufficiently high will the person act and remain engaged in goal-directed efforts.

Expectancies exist at many levels of generality. A patient lying in a hospital bed recovering from an automobile accident can have a variety of different expectancies. She can have an expectancy about being able to move her index finger one more centimeter forward, an expectancy about being able to get out of bed unaided by next Saturday, and an expectancy about fully recovering from the trauma in the long run. Presumably, the principles embodied by expectancy-value theories pertain equally well to expectancies that are specific in nature and expectancies that are more general. They should even apply to the most general kinds of expectancies, those that characterize optimists and pessimists. The confidence that is at issue is simply broader in scope, relevant to a greater class of situations and behaviors.

Goals, Stress, and Optimism

Expectancies are important facets of expectancy-value theories, but so are goals and values. The emphasis on goals meshes well with aspects of contemporary theorizing in personality and social psychology. Recent years have seen a reemergence of interest in goal constructs. One prevailing view is that human behavior is organized around the pursuit of goals – trying to attain things that are desired and avoid things that are not. Because of differences in emphasis, theorists use different terms to refer to goals, including current concerns, personal strivings, personal projects, and life tasks. Goals that are even broader in nature carry labels such as possible selves and self-guides.

Although these various goal constructs certainly differ among themselves in ways that are far from

trivial, they also share an underlying commonality. All include the idea that goals energize and direct activities. All include the sense that goals give meaning to people's lives and that understanding the person means understanding the person's goals and the organization of those goals. As such, goals are seen as providing the structure that define people's lives, imbuing lives with meaning, both in the short run and the long run.

Goal constructs provide an interesting window on the experience of stress and coping. From this vantage point, stress occurs when people experience difficulty moving toward desired end states or difficulty keeping away from end states that are unwanted. Stated somewhat differently, stress occurs whenever impediments to goal attainment are encountered. From this perspective, coping involves efforts to create conditions that permit a person to continue to move in the right direction (toward the desired goals or away from the undesired goals) or to disengage from goals that are seen as no longer attainable. Thus, stress and coping are not seen as distinct classes of phenomena but, rather, are seen as arising within the dynamics of normal goal-related self-regulatory processes.

What determines how a person reacts to the blockage of goal-directed activities? Expectancy-value theories suggest that the key determinant is the person's confidence or doubt that goal attainment will eventually occur. Thus, when confronting minor challenges, optimists take a posture of confidence and continue to persist in their goal-directed efforts, even if progress is difficult or slow. Pessimists are more doubtful and reserved in their efforts. This divergence should also be displayed as the threat to goal attainment becomes more extreme. Optimists are likely to believe that the adversity can be handled successfully; pessimists are likely to anticipate disaster. To the extent that continued engagement in goal-directed efforts is adaptive, optimists should experience a coping advantage over pessimists.

Effects of Optimism on Well-Being

Measuring Optimism

Research on the beneficial effects of optimism has flourished over the past 20 years. This research has tended to take two different routes to assessing optimism, leading to two somewhat distinct literatures. One approach measures expectancies directly, asking people to respond to statements such as "I'm optimistic about my future." The second approach is more indirect. It assesses expectancies by examining attributional style, the characteristic manner in which a

person explains prior events. Bad outcomes that are explained in terms of causes that persist into the future, influence a broad range of events, and involve an aspect of the self are thought to be pessimistic in nature in that they carry the implication that negative outcomes will continue to occur. The opposite attributional style – explaining negative events in terms of causes that are more time limited, narrow in their effects, and external to the self – are thought to reflect a more optimistic orientation. Although in principle attributional style can be assessed with respect to both negative and positive prior events, in practice assessment is usually limited to negative outcomes only.

Psychological Well-Being

Dozens of studies have been conducted examining the relationships between optimism, pessimism, and distress among groups of people undergoing adversity of one type or another. The range of stressors embodied by this research has been extremely varied. Studies have examined the experiences of students entering college, employees of businesses, and survivors of missile attacks. Studies have measured the reactions of people caring for cancer patients and people caring for patients suffering from Alzheimer's disease. Research has examined the experiences of people dealing with medical procedures such as childbirth, abortion, coronary artery bypass surgery, and attempts at *in vitro* fertilization, as well as heart and bone marrow transplantation. Yet other studies have looked at how people deal with a diagnosis of cancer, the pain of arthritis, and the progression of AIDS.

The results of these various studies all point in the same direction: optimists experience less distress during times of adversity than do pessimists. This conclusion holds for studies that are cross-sectional in nature – that is, those that assessed optimism and distress at the same point in time. This conclusion also holds true for studies that are prospective in nature – that is, those that assessed optimism and distress at some appropriate baseline and then reassessed distress at a later time. The results of the prospective studies are particularly noteworthy in that they suggest that optimism is associated with beneficial changes in distress over time. Thus, these prospective studies also help to get around the potential problem of confounding optimism with subjective well-being, which is inherent in the cross-sectional research.

Physical Well-Being

Much more is known about the effects of optimism on psychological well-being than is known about the

effects of optimism on physical well-being. Still, a number of studies have explored the links between optimism and physical health and between optimism and parameters of physiological functioning. The findings from this body of research mirror those just presented with respect to psychological functioning: optimists typically show signs of better physical health or signs of more adaptive physiological responses when under adversity than do pessimists.

For example, compared to pessimists, optimists report fewer physiological symptoms during times of duress and maintain a higher health status across their lives. Other studies show that optimists are less likely to suffer negative side-effects from major surgery or to be rehospitalized within the first few months after major surgery. In a similar vein, less pessimistic people have been found to outlive people who are more pessimistic in outlook when diagnosed with a life-threatening illness such as recurrent cancer. They also exhibit less extreme cardiovascular reactivity during the course of their daily lives. On a somewhat different note, optimists tend to show signs of more adaptive immune functioning than do pessimists, but the evidence on this point is more mixed than on the other physical-health outcomes that have been examined.

Future Directions

At present, the evidence strongly indicates that optimism confers important benefits. Still, a number of questions remain for future research to resolve. First, there is a question concerning the structure of optimism and pessimism. Researchers tend to treat optimism and pessimism as though they were opposite poles of a single continuum. However, some evidence suggests that optimism and pessimism may represent related but independent dimensions. This is one question that researchers have been trying to sort out for a while and will continue to sort out into the future.

Second, little research has been directed toward understanding the developmental antecedents of optimism and pessimism. We do know from several studies that optimism and pessimism have a genetic basis. Less is known about the environmental influences on optimism and pessimism. What accounts for the nongenetic variation in optimism and pessimism? What kinds of life experiences cause a person to become more or less optimistic? These will also be important questions for future research to address.

Third, what can be done to turn pessimists into optimists? To date, relatively little research has concerned itself with generating and evaluating interventions designed to reduce pessimism. Given the toxic effects

of a pessimistic orientation, the future will no doubt see a significant increase in the number of studies that are designed to develop effective interventions to counteract the negative effects of pessimism.

Finally, it now seems well established that coping differences underlie the differences in psychological adjustment that emerge in the manner in which optimists and pessimists adjust to stress. Less clear are the biological and physiological mechanisms that underlie the differences in physical health outcomes that have been observed. The search for the biological pathways by which optimism gets translated into health outcomes will also occupy a substantial portion of the research landscape in the future.

Acknowledgments

The preparation of this article was facilitated by NIH grants HL065111, HL065112, HL076858, HL076852, CA64710, and CA84944.

See Also the Following Article

Personality Processes.

Further Reading

Carver, C. S. and Scheier, M. F. (1998). *On the self-regulation of behavior*. New York: Cambridge University Press.
 Carver, C. S. and Scheier, M. F. (1999). Stress, coping, and self-regulatory processes. In: Pervin, L. A. & John, O. P. (eds.) *Handbook of personality* (2nd edn., pp. 553–575). New York: Guilford Press.

Carver, C. S. and Scheier, M. F. (2002). Optimism. In: Snyder, C. R. & Lopez, S. J. (eds.) *Handbook of positive psychology*, pp. 231–243. New York: Oxford University Press.
 Carver, C. S. and Scheier, M. F. (2003). A self-regulatory perspective on personality. In: Millon, T. & Lerner, M. J. (eds.) *Comprehensive handbook of psychology*. Vol. 5: *Personality and social psychology*, pp. 185–208. New York: John Wiley.
 Pervin, L. (ed.) (1989). *The goal concept in personality and social psychology*. Hillsdale, NJ: Lawrence Erlbaum.
 Peterson, C. and Bossio, L. M. (1991). *Health and optimism: new research on the relationship between positive thinking and well-being*. New York: Free Press.
 Scheier, M. F. and Carver, C. S. (2003). Goals and confidence as self-regulatory elements underlying health and illness behavior. In: Cameron, L. D. & Leventhal, H. (eds.) *The self-regulation of health and illness behavior*, pp. 17–41. Reading, UK: Harwood.
 Scheier, M. F., Carver, C. S. and Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a re-evaluation of the Life Orientation Test. *Journal of Personality and Social Psychology* 67, 1063–1078.
 Scheier, M. F., Carver, C. S. and Bridges, M. W. (2001). Optimism, pessimism, and psychological well-being. In: Chang, E. C. (ed.) *Optimism and pessimism: implications for theory, research, and practice*, pp. 189–216. Washington, DC: American Psychological Association.
 Scheier, M. F., Matthews, K. A., Owens, J. F., et al. (1999). Optimism and rehospitalization following coronary artery bypass graft surgery. *Archives of Internal Medicine* 159, 829–835.
 Seligman, M. E. P. (1991). *Learned optimism*. New York: Knopf.

Orexin

W K Samson and M M White

St. Louis University, St. Louis, MO, USA

A V Ferguson

Queen's University, Kingston, ON, Canada

© 2007 Elsevier Inc. All rights reserved.

Discovery of the Orexins

Orexins and the Stress Response

Orexin and Stress-Related Behaviors

Summary and Conclusion

Glossary

<i>Adrenocorticotrophic hormone (ACTH)</i>	A stress hormone that activates hormone secretion from the adrenal gland.
<i>Cerebro-ventricles</i>	Fluid-filled cavities in the brain where test compounds often are injected <i>in vivo</i> .
<i>Circumventricular organs</i>	Sites in the brain where there is no blood–brain barrier.
<i>Corticosterone</i>	A steroid hormone released from the adrenal gland during stress that facilitates energy delivery to the nervous system and stimulates the production of new glucose.

<i>Hypothalamus</i>	A brain region responsible for coordinating the appropriate endocrine, cardiovascular, and behavioral responses to stress.
<i>Prolactin (PRL)</i>	A hormone released from the anterior pituitary gland in stress that exerts multiple tissue-specific effects.
<i>Rostral ventrolateral medulla</i>	An important site in the hindbrain where neurons organize and control sympathetic efferent activity.
<i>Stress hormones</i>	Chemical messengers released from the anterior pituitary gland during physical or emotional stimuli.

Discovery of the Orexins

The lateral hypothalamic area (LHA) has long been considered to be an important site for the regulation of feeding behavior. In 1998, two groups independently identified the production of novel peptides in this region. deLecea and colleagues named these peptides hypocretins because of their apparent production in the hypothalamus and their structural homology to the members of the secretin family of peptides. Sakurai and collaborators named these peptides orexins on the basis of their ability when administered intracerebroventricularly (ICV) to stimulate food intake. It is now recognized that the peptides exert many physiologically relevant actions in brain, the majority of which are seemingly related to behavioral, endocrine, and cardiovascular responses to stress.

For convenience we use here the nomenclature advanced by Sakurai. Orexin A and orexin B are posttranslational products of the same prohormone (Figure 1). Two receptor proteins have been identified, the orexin type 1 receptor (OX₁R), which binds both peptides but prefers orexin A as a ligand, and the orexin type 2 receptor (OX₂R), which binds both peptides with approximately equal affinity. The two receptors are differentially distributed, as listed in Table 1.

The distribution of the orexin receptors predicted more biological actions than simply those related to food intake. Indeed, receptor localization in the dorsal raphe, tuberomammillary nuclei, and locus ceruleus (LC), areas known to be important in arousal state and sleep/wakefulness, predicted the eventual finding that the orexins are important determinants of sleep/wakefulness. The absence of the receptors in dogs and degeneration of orexin-expressing neurons in animal models and humans results in the syndrome narcolepsy/cataplexy. Thus, the major role recognized at this time for these peptides is one of arousal. Aspects of the narcolepsy/cataplexy syndrome highlight other reported biological actions of the orexins, particularly those related to stress.

Orexins and the Stress Response

Orexins and Cardiovascular Function

Soon after their discovery, the orexins were demonstrated by several groups to act in the brain to elevate

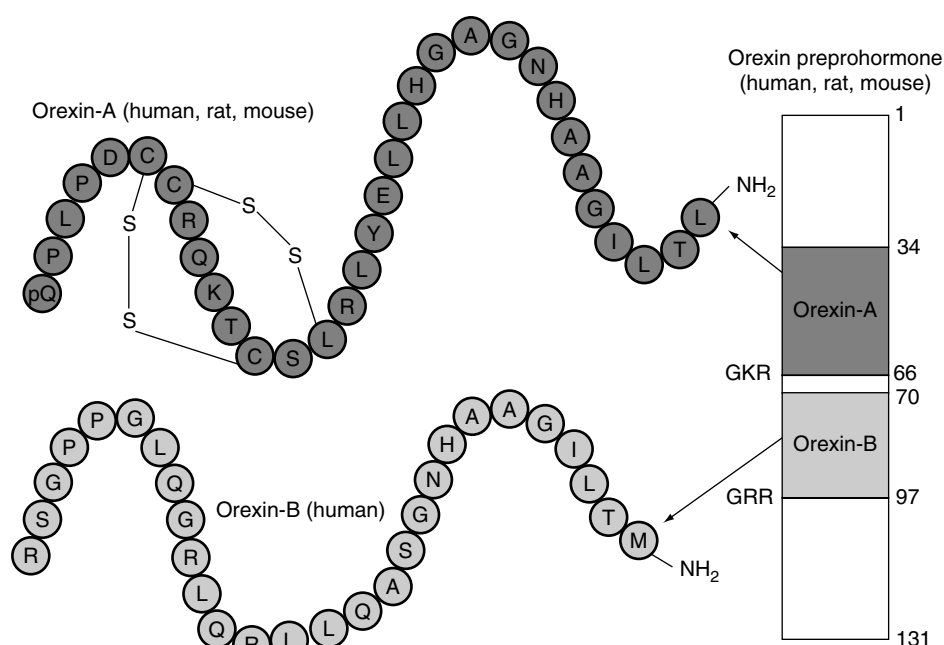


Figure 1 Structure of the orexins. From *Trends in Endocrinology and Metabolism* volume 11, 2000.

Table 1 Distribution of orexin receptor subtypes in brain regions important to the stress response^a

Brain site	OX ₁ R mRNA	OX ₂ R mRNA
Bed nucleus, stria terminalis	High	Moderate
Medial amygdala (nucleus)	Moderate	Low
Cortical amygdala (nucleus)	Low	Moderate
Periventricular thalamus (nucleus)	High	Moderate
Medial preoptic hypothalamus (area)	Absent	Moderate
Medial preoptic hypothalamus (nucleus)	Moderate	Moderate
Organum vasculosum lamina terminalis	Low	Moderate
Median preoptic area	Low	Moderate
Anterior hypothalamic nucleus	High	Low
Hypothalamic arcuate nucleus	Absent	High
Lateral hypothalamus (area)	Low	High
Medial paraventricular nucleus (parvocellular)	Absent	High
Lateral paraventricular nucleus (parvocellular)	Absent	Low
Posterior paraventricular nucleus (magnocellular)	Absent	Absent
Tuberomammillary nuclei	Absent	High
Ventromedial hypothalamus (nucleus)	High	Moderate
Dorsomedial hypothalamus (nucleus)	Absent/low	Moderate
Midbrain A4, A5, and A7 cell groups	High	Absent
Dorsal motor nucleus of vagus	Moderate	Moderate
Dorsal raphe nucleus	Moderate	Moderate
Locus coeruleus	High	Low
Nucleus of the solitary tract	Moderate	Low
Parabrachial nucleus	Absent	Low
Periaqueductal grey	Moderate	Moderate
Sensory and spinal trigeminal nuclei	Absent	Moderate
Substantia nigra (pars compacta)	Moderate	Moderate
Ventral tegmental area	Moderate	Moderate
Zona incerta	Moderate	Moderate

^aAs determined by *in situ* hybridization histochemistry. Density of the histochemical signal detected is indicated as high, moderate, low, or absent (at background).

blood pressure and heart rate. The effects are mediated via the activation of the sympathetic nervous system, as demonstrated by the use of α -adrenergic antagonist blockade (phentolamine) and the demonstration of increased sympathetic nerve activity. Increased cardiovascular function was observed following the ICV or direct tissue-specific administration of the peptides into the area postrema, rostral ventrolateral medulla, and medial nucleus tractus solitarius; the direct effects on preganglionic sympathetic neurons in the intermediolateral cell column (IMDL); and the ability of intrathecal administration of orexin to activate these neurons. In addition, orexin has been shown to influence the excitability of paraventricular nucleus (PVN) parvocellular and nucleus tractus solitarius (NTS) neurons. This does not rule out the potential contribution of vagal withdrawal to the observed increased heart rate following central orexin administration. In fact, direct effects of orexin on neurons in the dorsal motor nucleus of the vagus also have been reported.

Cellular mechanisms of action In an attempt to determine the cellular mechanism of action of orexin

to raise blood pressure and heart rate by an action in the NTS, experiments were conducted in medullary slice preparations. Whole-cell recordings from NTS neurons in current clamp mode revealed that the depolarizing effect of orexin was maintained in the presence of tetrodotoxin, demonstrating direct postsynaptic actions of the peptide. In voltage clamp mode, orexin was observed to inhibit a specific potassium conductance (the sustained K⁺ current) and to activate a nonselective cation conductance, leading to depolarization. Later a G-protein signaling mechanism was demonstrated to activate phospholipase C and protein kinase C but not protein kinase A (PKA).

The cellular effects of orexin also have been characterized in the area postrema (AP), a circumventricular organ located adjacent to the NTS. In dissociated AP neurons, orexin similarly activated a nonselective cation conductance, leading to depolarization. In these cells, however, no effects on potassium currents were observed. The activation of a nonselective cation conductance also has been demonstrated in parvocellular PVN neurons, which may in fact have been preautonomic cells projecting

to sympathetic centers in the rostral ventrolateral medulla. As in the AP, there did not appear to be any action of orexin on potassium currents in neurons in parvocellular PVN.

In sympathetic preganglionic neurons in spinal cord, orexin's depolarizing effects are blocked by pertussis toxin, indicating a G-protein signaling mechanism. In these neurons, a PKA-mediated inhibition of potassium channels has been demonstrated. The direct effects of orexin to depolarize noradrenergic neurons in the LC occur due to decreased potassium conductance and the activation of an inward current. The exact receptor subtype responsible for these depolarizing actions of orexin in the NTS, AP, PVN, LC, or IMDL has not been identified; however, the similar (and in some cases diverse) signaling mechanisms described predict heterogeneity of receptor subtypes depending on tissue localization.

Physiological relevance of the pharmacological effects of orexin on cardiovascular function Mice in which the gene for orexin has been deleted (embryonic knockouts) and mice and rats engineered to suffer age-related loss of orexin neurons (orexin-ataxin 3 transgenes) display lower mean arterial blood pressures than their wild-type littermates under resting conditions and fail to mount a sympathetic response (increased heart rate or increased blood pressure) in the resident/intruder stress model. The reason for the lower resting blood pressures in these transgenic animals was established to be due to a loss of vasoconstrictor sympathetic tone because prazosin (α -adrenergic blockade) or hexamethonium (ganglionic blockade) failed to alter the already low pressures in these animals but brought pressures in wild-type littermates to levels similar to the transgenic animals. This was not accomplished by treatment with captopril (angiotensin-converting enzyme inhibitor) or a V_1 (vasopressin antagonist).

Does the syndrome narcolepsy/cataplexy in humans demonstrate compromised sympathetic function? Autonomic dysfunction has been reported in these patients, including the attenuation of cardiovascular reflexes to muscle contraction and the Valsalva maneuver. Frequent occurrences of orthostatic hypotension have been reported as well.

Neuroendocrine Actions of the Orexins

Orexin and the hypothalamic-pituitary-adrenal axis The presence of orexin receptors (Table 1) and broad distribution of orexin-positive nerve fibers in the hypothalamus suggested actions of the peptides on neuroendocrine centers in the diencephalon or the pituitary gland itself. Initial attention was drawn to potential actions in anterior pituitary gland because

orexin receptors are abundant in the gland and the peptide appears to be delivered to the median eminence, where it could gain access to the hypophyseal portal circulation. Although no significant effect of the orexins on prolactin (PRL) release were observed in dispersed anterior pituitary cell cultures, the release of the other major anterior lobe stress hormone, adrenocorticotrophic hormone (ACTH), was significantly affected by the concentrations of peptide in the nanomolar range. Basal ACTH release was not affected by orexin; however, corticotropin releasing hormone (CRH)-stimulated ACTH release was inhibited in a concentration-related fashion. The effect appeared to be mediated by the OX_1R because the median effective concentration (EC_{50}) for orexin A was at least 1 log lower than that for orexin B and because the inhibitory action of orexin was absent when the production of the OX_1R (but not the OX_2R) was compromised by antisense oligonucleotide treatment. Recently, the selective OX_1R antagonist has been demonstrated to also block the action of orexin A on CRH-induced ACTH release *in vitro*. Furthermore, the relatively selective OX_2R agonist (Ala¹¹, Leu¹⁴) orexin B did not exert actions similar to orexin A or orexin B. Orexin's action on the corticotroph appear to require the activation of a non-pertussis toxin-(non-PTX-) sensitive G-protein-coupled receptor to phospholipase C activation and the formation of inositol trisphosphate. In addition, the activation of PKC seems to be a necessary step in transduction of the inhibitory signal.

In the hypothalamus, orexin acts to stimulate the hypothalamic-pituitary-adrenal (HPA) axis by activating CRH-producing neurons. This was demonstrated by several groups using pretreatment with CRH antiserum or antagonists prior to the ICV administration of orexin. Whether this is a direct or indirect action of orexin on CRH-producing neurons in the parvocellular PVN is currently unknown, although single-cell the reverse transcription polymerase chain reaction (RT-PCR) approaches now employed in the authors' laboratories promise a future definitive answer to this question. Neurons in the PVN express both receptor subtypes, although the OX_2R appears more abundant. In addition, orexin-producing neurons that project to the PVN express glucocorticoid receptors, and thus the closed feedback loop of the HPA axis appears to include orexinergic neurons in the LHA (direct effects of orexin, of unknown origin, on adrenal steroid production and release have also been reported). As previously mentioned, direct postsynaptic effects of orexin on PVN neurons have been reported. The orexin effects on parvocellular PVN neurons, those projecting to median eminence or other brain sites (e.g., sympathetic

centers in the brain stem) are direct (i.e., postsynaptic) and are mediated by the activation of nonselective cation conductances. In magnocellular PVN, the depolarizing effects of orexin are indirect (i.e., tetrodotoxin blockable), involving a glutamatergic interneuron. Because high concentrations of arginine vasopressin (AVP) are capable of stimulating ACTH release from cultured pituitary cells *in vitro* and under at least some *in vivo* conditions the compromise of AVP action reduces stress-induced ACTH release, it is possible that the ACTH response to the ICV administration of orexin may be mediated by AVP release into the median eminence. The central administration of orexin has been reported to increase both CRH and AVP mRNA levels in the PVN; however, in hypothalamic explant cultures, orexin did not stimulate AVP release. Thus, it appears that the hypothalamic action of orexin to stimulate the HPA axis is mediated primarily by CRH release into the median eminence.

Orexin and prolactin secretion Controversy exists over the potential action of orexin on PRL secretion. Orexin, in log molar concentrations ranging from 1.0 pM to 100 nM, did not alter basal or thyrotropin-releasing hormone (TRH)-stimulated PRL release from dispersed anterior pituitary cells *in vitro*. However, when it was administered ICV some groups observed stimulatory and others inhibitory effects on PRL levels in plasma. The suppressive effects of orexin on *in vivo* PRL secretion were reported to be the result of increased dopamine turnover in male, but not female, rats. Evidence for a stimulatory role for orexin on PRL and luteinizing hormone (LH) releases come from fasting studies in which the pre-ovulatory PRL and LH surges, absent during fasting, were restored by central orexin administration. Even more support for the stimulatory actions of orexins comes from steroid-priming studies in ovariectomized rats, in which the PRL and LH surges were blocked by pretreatment with anti-orexin antibodies. Whether any of these pharmacological effects of orexin on PRL or LH secretion are physiologically relevant is unclear and may not be likely because orexin knockout mice are fertile and can nurse their pups. It is possible that physiologically relevant actions of the orexins on at least PRL secretion may not be related to reproductive function but instead to stress-induced hormone release. Thus, the actions of the peptide on tuberoinfundibular dopamine neurons of the arcuate nucleus, where the OX₂R is abundantly expressed, may be observed only in models of acute stress such as restraint (physical) or novel environment (anticipatory). The expression of early response genes in neurons of the LHA has been described in several experimental models of stress.

Orexin and Stress-Related Behaviors

The presence of orexin receptors and orexin-positive nerve fibers in central nervous system (CNS) sites known to be important in the arousal state, such as the LC and raphe nuclei, forms the anatomical basis for the effects of pharmacological doses of orexin on motivated and spontaneous behaviors. Several groups have described effects of orexin on locomotor behavior and general arousal state, and the obligatory role for endogenous orexin in sleep/wakefulness is well established. In our hands, the effects of orexin on general locomotor behavior and stereotypic behaviors parallel in dose relationship and time course the actions on sympathetic activity (i.e., blood pressure elevation). In addition, some of the feeding responses to orexin may be related to the peptide's ability to stimulate arousal, as well as exploratory and reward behaviors. These behaviors, observed following central orexin administration, mirror behaviors observed during stress; however, they differ from those observed following CRH administration, which are more anxiogenic in nature. There may be an important interplay between orexin neurons in the LHA and CRH neurons in the PVN. As already detailed, orexin appears to activate CRH neurons (resulting in ACTH release), and it has been reported that CRH receptors are present on orexin neurons in the LHA, where CRH itself exerts depolarizing effects. Thus, at least some of the behavioral arousal stimulated by CRH may be mediated by the effects of the peptide on endogenous orexin release.

The pharmacological actions of orexin on locomotor activity and arousal state are mirrored by determinations of orexin neuronal activity in these behaviors. Indeed, c-FOS expression correlates directly with wakefulness and locomotor activity, and the spontaneous activity of these neurons during such states has been characterized as well. One site of orexin action related to spontaneous locomotor activity is the PVN. In addition to stimulating feeding in the early light phase, orexin injection into the PVN increased spontaneous locomotor activity at all the times examined, and this was accompanied by increased oxygen consumption. The effect of orexin on spontaneous locomotor activity was antagonized by pretreatment with SB-334867 (an OX₁R antagonist). Those results suggest that physical activity independent of feeding is due to the actions of orexin in the hypothalamus. Indeed in recent studies, we observed that many, but not all, of the behavioral actions of exogenous orexin could be antagonized by pretreatment with the OX₁R antagonist, whereas some of the similar behavioral actions of the selective OX₂R agonist remain unaffected by OX₁R

antagonist pretreatment. Thus, both receptors may be responsible for differing behavioral modalities elicited by orexin administration. It will be important to determine the receptor subtype selectivity of the individual behaviors so that the roles played by each receptor in not only locomotor but also stress-related behaviors can be elucidated.

Non-Central Nervous System Sites of Orexin Action Related to Stress

Orexin receptors have been identified in non-CNS tissues, and significantly, the adrenal gland expresses both receptors but the OX₂R more abundantly. Orexin acts both *in vitro* and *in vivo* to stimulate corticosterone release, with orexin A being more potent than orexin B. This suggests a role for the OX₁R in the control of adrenal cortical function. The effect is selective (aldosterone secretion was unaffected) and is blocked by H-89 pretreatment *in vitro*, suggesting a signaling pathway involving adenylyl cyclase. Because the major site of orexin production was the LHA and plasma levels of the peptide are extremely low, it remained unclear that these pharmacological effects of orexin in the adrenal gland had any physiological correlate. However, it is now known that the orexins are also produced in the adrenal gland and thus may exert autocrine or paracrine actions in the gland. It should be remembered that the direction of blood flow in the adrenal gland is from the outer cortex to the inner medulla, and thus orexin produced in the adrenal cortex might bathe the medulla in relatively high concentrations, similar to the situation with corticosterone. At least in human pheochromocytomas, orexins act via the OX₂R to stimulate epinephrine and norepinephrine release, suggesting an action of orexin to regulate catecholamine release from the normal adrenal gland. Whether orexin production and/or release from cortex increases during stress is at this point unknown; however, it is tempting to speculate that the adrenal actions of orexin complement the peptide's actions in brain related to metabolism and stress.

Summary and Conclusion

Abundant evidence support a role for endogenous orexins in the regulation of sleep/wakefulness and metabolic state. Central to all the actions of the orexins is the consistent finding that multiple indicators of stress are revealed following orexin administration and some are absent when orexin production or action is compromised. Remarkable progress has been made in understanding the importance of the endogenous orexin system, largely due to the development of transgenic models of orexin peptide or receptor

compromise. In particular, the development of the orexin-ataxin-3 transgenic mice and rats has been a landmark in our ability to understand the broad impact of these peptides on multiple behavioral, endocrine, and cardiovascular systems. Now, with the availability of selective OX₁R antagonists and at least a relatively selective OX₂R agonist, additional elucidation of the sites and mechanisms of orexin actions in the brain can be accomplished. Because the replacement of the peptide restores normal sleep/wakefulness patterns in transgenic animals, the use of orexin or orexin analogs for the treatment of human narcolepsy/cataplexy holds much promise. Critical to the development of those potential therapeutic strategies is the recognition of the receptor subtypes responsible for the effects of orexin on cardiovascular and neuroendocrine function. Clearly, potential side effects related to cardiovascular control and hormone release are possible and should be monitored if orexin is to be tested for its therapeutic effect in humans.

See Also the Following Articles

Food Intake and Stress, Human; Feeding Circuitry (and Neurochemistry).

Further Reading

- de Lecea, L., Kilduff, T. S., Peyron, C., et al. (1998). The hypocretins: hypothalamus-specific peptides with neuroexcitatory activity. *Proceedings of the National Academy of Sciences USA* **95**, 322–327.
- de Lecea, L. and Sutcliffe, J. G. (eds.) (2005). *Hypocretins: integrators of physiological functions*. New York: Springer.
- Ferguson, A. V. and Samson, W. K. (2003). The orexin/hypocretin system: a critical regulator of neuroendocrine and autonomic function. *Frontiers in Neuroendocrinology* **24**, 141–150.
- Hara, J., Beuckmann, C. T., Nambu, T., et al. (2001). Genetic ablation of orexin neurons in mice results in narcolepsy, hypophagia, and obesity. *Neuron* **30**, 345–354.
- Kayaba, Y., Nakamura, A., Kasuya, Y., et al. (2003). Attenuated defense response and low blood pressure in orexin knockout mice. *American Journal of Physiology* **285**, R581–R593.
- Kilduff, T. S. (2005). Hypocretin/orexin: maintenance of wakefulness and a multiplicity of other roles. *Sleep Medicine Reviews* **9**, 227–230.
- Kiwaki, K., Kotz, C. M., Wang, C., et al. (2004). Orexin A (hypocretin 1) injected into hypothalamic paraventricular nucleus and spontaneous activity in acts. *American Journal of Physiology* **286**, E551–E559.
- Marcus, J. N., Aschkenasi, C. J., Lee, C. E., et al. (2001). Differential expression of orexin receptors 1 and 2 in rat brain. *Journal of Comparative Neurology* **435**, 6–25.

- Mieda, M., Willie, J. T., Hara, J., et al. (2004). Orexin peptides prevent cataplexy and improve wakefulness in an orexin-ablated model of narcolepsy in mice. *Proceedings of the National Academy of Sciences USA* 101, 4649–4654.
- Mileykovskiy, B. Y., Kiyaschenko, L. I. and Siegel, J. M. (2005). Behavioral correlates of activity in identified hypocretin/orexin neurons. *Neuron* 46, 787–798.
- Peyron, C., Tighe, D. K., van den Pol, A. N., et al. (1998). Neurons containing hypocretin (orexin) project to multiple neuronal systems. *Journal of Neurosciences* 18, 9996–10015.
- Sakurai, T., Amemiya, A., Ishii, A., et al. (1998). Orexins and orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. *Cell* 92, 573–585.
- Samson, W. K., Taylor, M. M. and Ferguson, A. V. (2005). Non-sleep effects of hypocretin/orexin. *Sleep Medicine Reviews* 9, 243–252.
- Willie, J. T., Chemelli, R. M., Sinton, C. M., et al. (2001). To eat or sleep?: orexin in the regulation of feeding and wakefulness. *Annual Reviews of Neuroscience* 24, 429–458.

Organ Transplantation, Immune Response See: Organ Transplantation, Stress of.

Organ Transplantation, Stress of

M A Dew, A F DiMartini and R L Kormos

University of Pittsburgh School of Medicine and Medical Center, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

Prevalence of Organ Transplantation

The Transplantation Process for the Individual Patient

Interventions to Minimize Stressor Effects during the Transplant Process

Glossary

<i>Graft rejection</i>	The process by which the body's immune response system attacks the foreign organ or tissue.
<i>Immuno-suppression</i>	The suppression of the immune response by an agent such as a chemical or drug in order to enhance the body's acceptance of foreign organs or tissue.
<i>Organ transplantation</i>	The process of removing an organ or tissue from one person's body and implanting it in another person's body.

Organ transplantation is increasingly often the procedure of choice for patients with end-stage diseases of the kidney, liver, heart, lung, pancreas, or intestines. Organ transplantation can significantly extend life, as well as dramatically improve the quality of

individuals' lives. Yet each step in the transplantation process – encompassing the initial evaluation and wait for an organ, the transplant surgery, postsurgical recovery, and living permanently with the transplant – is also associated with significant stressors. These stressors include both acute events and chronic, ongoing health and psychosocial strains. Across the entire process, transplant patients face and must adapt to changing health and functional capacity, altered social relationships, new perceptions of self, and revised life goals and plans. An understanding of both the nature of such stressors and individuals' reactions to them is critical for three reasons. First, as transplantation becomes more common, it is imperative that patients and their families know what to expect during the transplant process and thereafter. Education regarding likely stressors as well as potential benefits of transplantation will increase the likelihood that organ recipients will achieve the best possible quality of life (QOL) in the months and years after transplant. Second, it is only through knowledge about the nature of transplant-related stressors that we will be able to develop and test appropriate stress reduction interventions for transplant recipients and their families. Third, knowledge regarding transplant-related stressors, and reactions to them, also has value for understanding similar reactions to the experience of other life-threatening chronic physical illnesses, their treatment, and their long-term outcomes.

Prevalence of Organ Transplantation

Although the potential benefits of solid organ transplantation were first realized in the 1950s, it was not until the advent of modern immunosuppression in the 1970s and 1980s that organ transplantation fully evolved from an experimental procedure to the standard of care for many end-stage diseases. Today, the greatest obstacle to receiving a transplant is the shortage of donor organs. As illustrated in [Figure 1](#), the total number of individuals registered on the wait list for a transplant (including kidney, liver, heart, lung, heart-lung, pancreas, or intestines) far exceeds the number of transplants performed, and the gap continues to widen, despite the success of some social and governmental programs in increasing rates of donation. The lives of kidney transplant candidates can be maintained by dialysis during the wait for an organ. But for other solid organ candidates, for whom few or no such long-term alternative treatments exist, the consequence of the organ shortage is that a large proportion of them will die before a suitable organ becomes available.

In some areas of transplantation, living donor organs have helped to reduce waiting times and allow more candidates to receive transplants. For example, in the United States, living kidney donation – whether from a donor who is biologically or emotionally related to the recipient or from an anonymous,

unrelated donor – is almost as prevalent as deceased donor transplantation. Living liver donation is also becoming a viable option for some patients. [Figure 2](#) shows the total numbers of deceased donor transplants (and living donor transplants, where applicable) that were performed for each type of organ transplant in 2004 in the United States. Overall, kidney transplantation (whether from deceased or living donors) is the most common type of transplant performed, followed by liver, heart, pancreas, and lung transplantation. Intestine and heart-lung transplants are the most rare. Similar distributions are reported in other countries.

As shown in [Figure 3](#), survival rates following organ transplantation vary across organ type, but are good to excellent for many types of organs, especially in the early years after transplant. Kidney recipients, especially those with living donors, show the highest rates of survival. Even at 10 years posttransplant, 76% remain alive. Liver and heart recipients show slightly lower survival rates, but, even so, almost 75% survive at least 5 years posttransplant, and about 50% remain alive at 10 years posttransplant. Survival rates for intestine and lung recipients remain substantially poorer, especially in the long-term. But even these rates have shown marked improvements in recent years. Given that annual mortality rates range up to 30% or more among candidates awaiting transplant, it is clear that transplantation provides a significant extension of life for many patients.

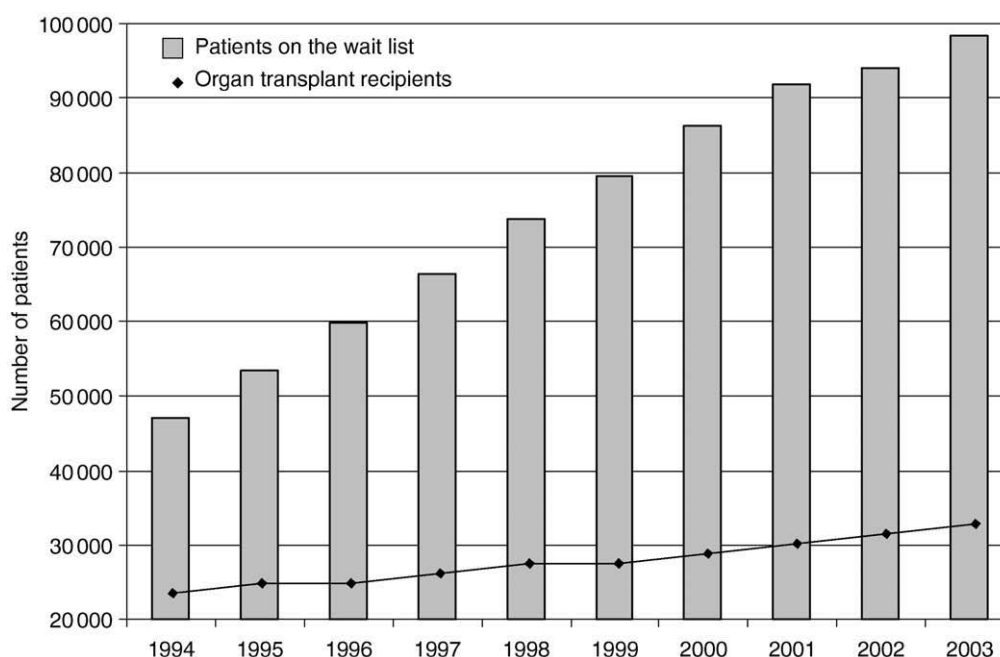


Figure 1 Combined United States and Eurotransplant statistics. From 2004 Annual Reports from United Network for Organ Sharing, Organ Procurement and Transplantation Network, which includes all United States data (www.optn.org), and Eurotransplant, which includes data from Austria, Belgium, Germany, Luxembourg, Netherlands, and Slovenia (www.eurotransplant.org).

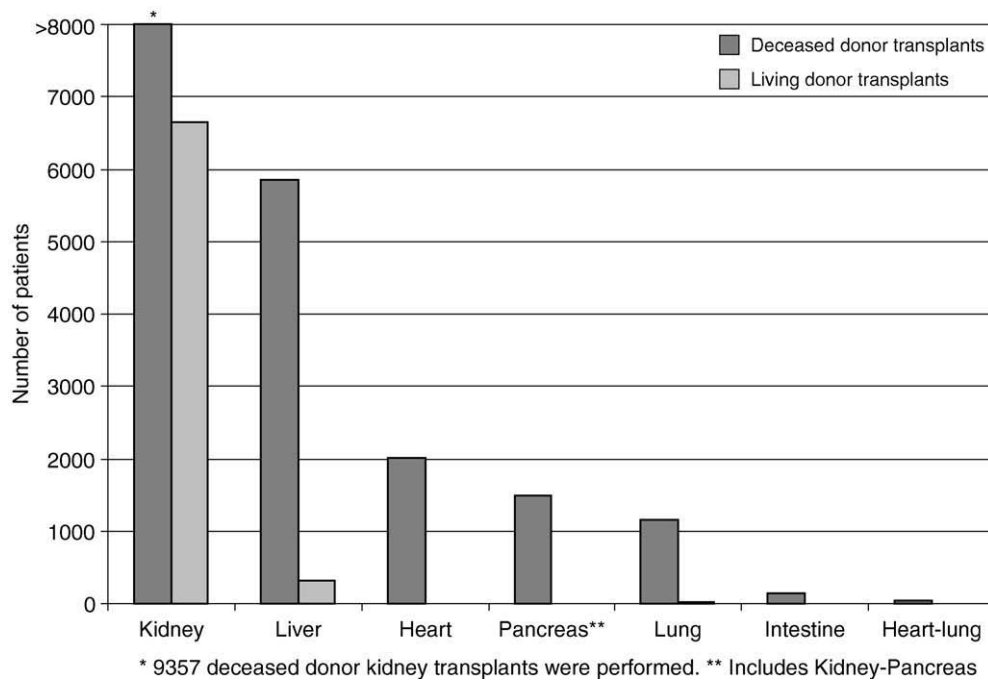


Figure 2 Numbers of deceased donor and living donor transplants of each organ type (including kidney-pancreas) in the United States in 2004; 9357 deceased donor kidney transplants were performed. From United Network for Organ Sharing, Organ Procurement and Transplantation Network (www.optn.org).

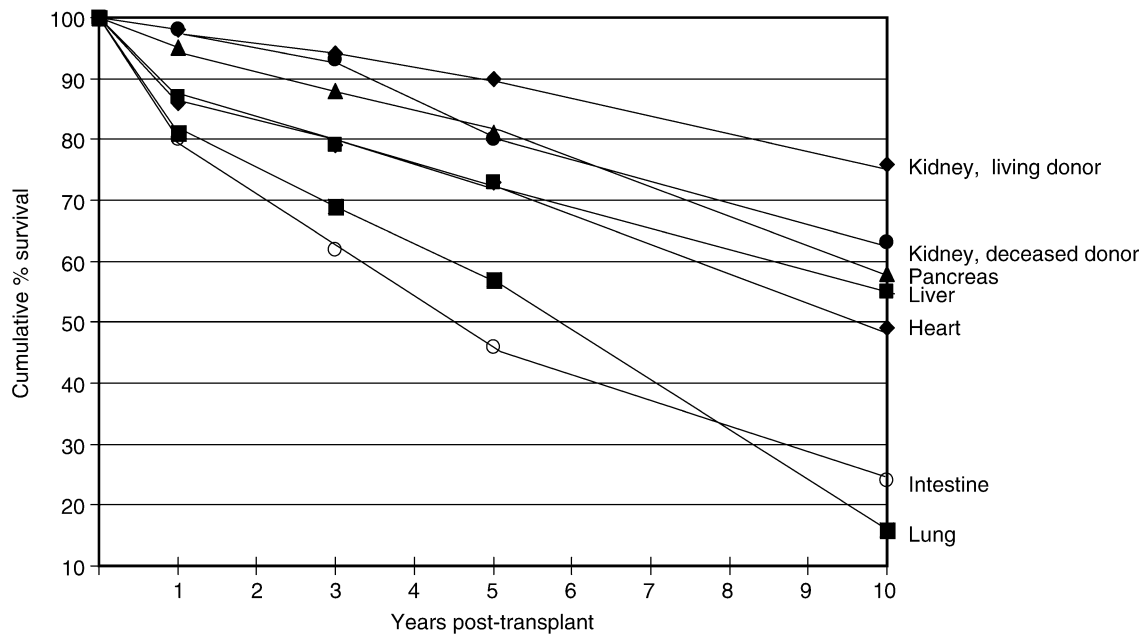


Figure 3 Patient survival, U.S. transplants, 1996–2003. From 2004 Annual Report, United Network for Organ Sharing, Organ Procurement and Transplantation Network.

The Transplantation Process for the Individual Patient

From the point of their initial contact and evaluation by the transplant team, patients and their families move through a relatively standard series of events

and time periods associated with the transplant process. The typical time line is shown in [Figure 4](#). The time line is punctuated by specific events that occur during the transplant process, and these events mark the onset and offset of important periods or stages, including the wait for a donor organ, the

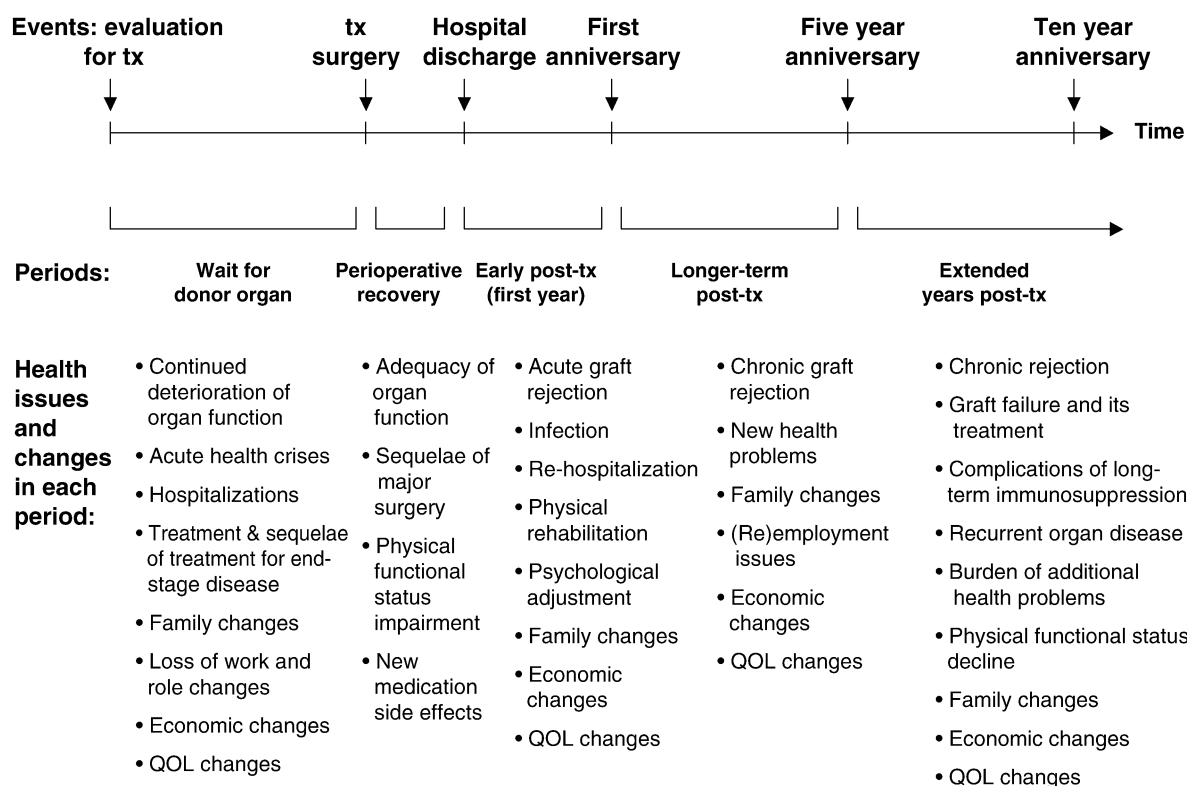


Figure 4 Organ transplant timeline: critical events, time periods, and health and psychosocial issues. Abbreviations: tx, transplantation; QOL, quality of life. Adapted from Dew, M. A. et al. (2002). Psychosocial aspects of transplantation. Reprinted with permission from Medscape Transplantation. © 2002, WebMD, Inc.

perioperative recovery period, the early posttransplant months, and the longer-term posttransplant years. The duration of each of these periods can vary considerably, depending on such factors as the rapidity with which end-stage organ disease develops in a given patient, the availability of a living donor (when living donation is possible) or a suitable deceased donor organ, and the medical complications that may occur perioperatively or later posttransplant. However, the essential ordering of the events and time periods is similar for all patients.

Figure 4 also lists the typical health and psychosocial issues that are associated with each time period. Some of these issues are unique to specific time periods (e.g., continued deterioration of the native organ's function during the waiting period, or early posttransplant physical rehabilitation). Others are present across almost all of the periods (e.g., changes in QOL), although the central elements of concern to patients and their families in these areas vary across time. Each of the sections below considers the stressors and related health and psychosocial issues that are relevant to the major events and the time periods in Figure 4.

The Initial Evaluation for Transplant

The formal medical evaluation for transplant candidacy is often the first occasion during which patients and their families seriously consider transplantation as a therapeutic option. For both patients and their family members, the evaluation may evoke conflicting feelings, including relief and the hope of having a healthy future, fears about surgical risks and adapting to life with an organ from another person, worry about whether the patient will be judged to be a suitable transplant candidate, worry about the financial costs of transplantation, concerns that an organ will not become available soon enough to save the patient's life, and eagerness to try a new treatment option combined with reservations as to whether there might still exist other seemingly less drastic medical interventions. For all of these reasons, studies of psychological distress among patients and family members conducted at the time of the evaluation typically find that many respondents' distress levels are high. Symptoms of anxiety and depression have been noted most frequently, and distress in these areas may reach the level of diagnosable disorder in some individuals.

The evaluation for transplant is itself usually lengthy and complex because it is designed to assess all aspects of patients' medical and psychosocial status. Patients may be hospitalized during the several days of the evaluation in order to complete all required tests and procedures. An extensive array of medical and laboratory tests is required, and virtually all transplant programs consider patients' psychosocial and psychological histories. Although there is considerable variation across programs in the depth with which psychosocial history is probed, important elements that are usually evaluated to at least some degree include lifetime history and treatment of psychiatric disorder; lifetime history and current use of alcohol, nicotine, and other substances; current and past level of adherence to medical regimens; cognitive status and ability to understand what will be required for posttransplant care; social history including employment, financial circumstances, marital status, and availability of social supports from family and friends; personal views and expectations about the prospect of receiving a transplant; and personal strategies for coping with functional limitations due to health.

These psychosocial components are routinely assessed because patients' history and current status in these areas may influence their psychological adaptation to the wait for the transplant and to life after the transplant. However, these assessments often serve as important additional stressors to patients and families. They are stressful because of the personal nature of the questions as well as because the results may be used for decision making by the transplant team. Typically, these evaluations will contribute to judgments as to whether patients should be immediately listed as transplant candidates or whether they need to undergo additional interventions to improve their health status and adherence to the medical requirements set by the transplant team. For example, patients may be required to enroll in alcohol or smoking cessation programs before the transplant team will accept them onto the waiting list.

The Waiting Period

Many patients and their families perceive the waiting period to be the most psychologically stressful part of the transplant experience, due largely to patients' continued physical health deterioration and the inherent uncertainties about whether and when a suitable donor organ will become available. The duration of this period can vary from a few days (although brief waits are now rare) to several years, depending on the patient's medical status and other factors such as blood type and the patient's physical size. The geographical location of the transplant center can be a major

determinant of waiting time as well, since organs are allocated in many countries through a regional system that depends on the location from which the donor organ originated. Given the uncertainties of obtaining a donor organ, plus the risks of the transplant surgery itself (especially as the patient's medical condition deteriorates), patients and their families are faced with the mutually opposing prospects of preparing to live and preparing to die.

In the face of these stressors, the risk of significant psychological distress and cognitive impairment rises. Thus, as for many chronic disease populations, transplant candidates are much more likely to experience elevated distress and diagnosable depressive and anxiety-related disorders than nonpatient, community-based populations. Whether a particular transplant candidate is at greater risk for depression or anxiety appears to depend in part on their end-stage organ disease. For example, depression appears to be more common in kidney, liver, and heart candidates, while anxiety disorders – panic disorder in particular – are more often observed in patients with end-stage lung disease.

In general, the longer a patient waits for a transplant, the greater the psychological toll on the patient and family. However, it is noteworthy that an extremely brief wait (e.g., in the case of patients who have very rapidly declining organ function and are fortuitously able to undergo transplantation very quickly) may have negative psychological consequences as well. For example, it has been found that levels of anxiety and rates of anxiety disorders in the first year after transplantation are higher in patients who wait for very brief periods of a few months or less, compared to those who wait for longer periods. It is thus possible that the waiting period – despite its central elements of uncertainty and its associated risk for psychological distress when the wait is prolonged – can provide important time for patients to adapt to the idea that they need a transplant. If they do not have time before the surgery to adjust to the prospect of transplantation, the immediate consequence may be greatly heightened distress as they attempt to adapt to the transplant experience in the early postsurgical period.

Neurocognitive impairments are common in transplant candidates. For example, hepatic encephalopathy is prevalent in liver candidates and ranges from mild cognitive deficits to delirium and coma. Even patients with chronic liver disease who do not have hepatic encephalopathy may have cognitive and behavioral abnormalities as results of more subtle derangements of metabolic and detoxification functions. Cognitive deficits in heart candidates are extremely common once cardiac ejection fraction (a measure of the heart's ability to pump) drops below

30%; when brain perfusion is substantially inadequate, delirium results. Although kidney patients' conditions are generally more stable compared to candidates for other types of solid organs, subtle cognitive deficits often occur when patients are not well dialyzed.

Continued elevations in psychological distress and/or neurocognitive changes may have their own negative effects on other areas of patient well-being. For example, impaired patients may have more difficulty adhering to the medical regimen requirements set by the transplant team: they may have trouble following prescribed diets or abstaining from nicotine, alcohol, or other substances. These difficulties may arise not only due to impaired motivation (e.g., as a result of depression) but as a result of confusion and memory loss.

Finally, as patients' organ function declines, they may require increasingly complicated medical technologies and treatment regimens in order to maintain their lives. These treatments also serve as stressors that further tax patients' strained emotional and cognitive resources. The strain may extend to the family as well. The patient's marriage and relationships with primary family caregivers (who are most often spouses of adult patients and parents of pediatric patients) are particularly vulnerable due to role changes within the family and changes in daily living activities and schedules. The financial ramifications of these changes – arising from medical expenses, the inability of the patient to work, and/or the need for key family members to take substantial time off from work – can further add to strain within the family.

The Surgery and the Perioperative Recovery Period

This phase of the transplant process is characterized by major, but often positive, physical and emotional transitions. With improvements in procedures and medical care in the immediate days posttransplant, hospital stays have been dramatically reduced. For example, patients who receive the most common types of transplant (kidney, liver, or heart) are frequently able to be discharged in 7 to 10 days. However, as with all major surgery, the possibility for medical complications exists, and this, plus the patient's physical status immediately prior to surgery, will affect the speed of perioperative recovery.

During the first few weeks after surgery, most patients and their families voice high levels of optimism about the patient's prospects for recovery. Their central concerns usually involve how well the new organ is functioning and the probable speed of the patient's physical rehabilitation. These concerns are sometimes heightened by the brief postsurgical hospital stays. Moreover, the short duration of the

hospitalization is often perceived as stressful because patients and families typically receive many educational materials about life posttransplant during this time. (This information may have been provided before the transplant as well, but patients may have had difficulty absorbing it due to other stressors and/or cognitive impairment.) The positive feelings and optimism that often dominate after the surgery can interfere with educational efforts during the perioperative period because these feelings may lead patients and their families to be less able to accept and focus on information concerning potential complications, financial issues, and family difficulties that can arise posttransplant.

An additional source of stress arises from medication changes made in conjunction with the transplant. The immunosuppressants that transplant recipients will take for the remainder of their lives are introduced with the transplant, and they bring an extensive range of side effects. While some side effects abate with time, transplant patients – just like any other chronic disease population – will find that others appear to be permanent and unpleasant (e.g., tremor, altered body appearance, changes in sexual functioning, appetite changes). The new medications can also precipitate mood changes. For example, the introduction of corticosteroids can precipitate the onset or recurrences of mood or anxiety disorders. In addition, it is not uncommon for patients who were stabilized on psychotropic medications before transplant to have those medications abruptly discontinued after surgery. This can lead to rapid destabilization for patients, whose mood and/or anxiety disorders may quickly recur.

The First Year after the Transplant

For most patients, the first year posttransplant is a period of readjustment and rehabilitation, with gradual improvement in all domains of QOL, including physical, emotional, and social well-being. A frequent source of frustration for both patients and their families is the fact that the recovery process is generally considerably slower than they had expected. Moreover, important psychosocial elements of the transplant process may further prolong recovery and may also affect family members' well-being. These elements pertain to patients' need to (1) cope with common complications such as acute graft rejection episodes and infections, (2) alter their self-image from that of a critically ill or dying patient and resume a less illness-focused lifestyle, and (3) psychologically accept the fact that they have an organ from another person. Aside from these personal issues, patients may feel daunted by continuing financial issues related to the costs of the transplant, follow-up care,

and medications. In addition, in terms of interpersonal relationships, both patients and family members can be dismayed that the same marital, parent-child, or other family-related difficulties that existed before (and/or during) the period of critical illness and transplant surgery continue after the patient returns home. In fact, the stress of the waiting period may have held some relationships together temporarily, and these may then dramatically decline after the transplant.

An additional area of difficulty for many patients is adherence to the complex posttransplant treatment regimen. Not only must patients take multiple medications, but they also must complete routine medical follow-up evaluations and laboratory tests, monitor vital signs at home (e.g., blood pressure, temperature), follow exercise and dietary requirements, and adhere to restrictions on smoking, alcohol, and other substance use. Just as for most chronic disease patients who must adhere to a complex regimen, transplant recipients' adherence tends to worsen with time. This worsening itself can be a source of frustration (as well as a health risk) for patients, who may become quite discouraged at their inability to refrain from long-standing habits (e.g., smoking) or maintain new ones (e.g., exercise).

A likely result of these many concerns is that the likelihood of experiencing high levels of emotional distress, as well as clinically significant episodes of depressive and anxiety-related disorders, is higher during the first year posttransplant than during subsequent years. For most patients, distress levels gradually abate over the course of the year, and some psychiatric disorders – for example, posttraumatic stress disorder (PTSD) related to the transplant – appear to rarely if ever recur in later years. Individuals with PTSD related to the transplant appear to be unable to come to terms with the transplant experience and may experience flashbacks, nightmares, and extreme distress when thinking about one or more of the stressors that occurred during the waiting period, the surgery, and/or the perioperative recovery period.

Not only can the elements that serve as posttransplant stressors exacerbate psychological distress, but also such distress can itself serve as a major source of strain that, in turn, contributes to the worsening of these other stressors. Thus, for example, episodes of depression and anxiety can lead to increased difficulties in adhering to the medical regimen, as well as affect the patient's ability to cope with new medical complications or the need for rehospitalizations. There is evidence that, perhaps via these behavioral mechanisms, psychological distress can lead to poorer

clinical outcomes and reduced survival rates after transplant.

The Longer-Term Years after the Transplant

By the end of the first year after transplant, most patients have achieved their maximal level of physical rehabilitation and have psychologically adjusted to the transplant experience. Most then embark on a multiyear period in which graft functioning remains high and the rate of serious or life-threatening medical complications is low. Acute graft rejection episodes may be rare or relatively infrequent, compared to the first year after transplant. Patients' QOL in the areas of physical, emotional, and social well-being is often high, and rates of rehospitalization for rejection, infections, etc., are low. Some patients will develop new health problems, including diabetes, hypertension, and kidney disease, often as a result of long-term immunosuppression use. Patients are also at greater risk for cancers due to these drugs. Generally, however, patients' level of functioning remains quite high during these years.

Focal issues for patients and their families during these years pertain to maintaining as high a level of QOL as possible and postponing the effects of any decline in graft function or other complications for as long as possible. Economic and financial concerns related to costs of medications and continuing health care can become more important as the years go by. Nevertheless, psychological distress related to the transplant experience and its medical or psychosocial sequelae is uncommon during these years. Instead, patients' mental health is much more likely to be influenced by the same sorts of acute life events and chronic life strains experienced by other community populations. Indeed, the fact that nontransplant life stressors assume increasing prominence as predictors of emotional difficulties during these years attests to the fact that the vast majority of patients – even those who had great psychological difficulty immediately posttransplant – eventually are able to incorporate the experience into their lives and go on to focus on other things.

The Very Extended Years after the Transplant

We know the least about the stressors and psychosocial outcomes experienced by transplant recipients and their families beyond the 5-year anniversary of the transplant. Even in the 1980s and early 1990s, survival rates were sufficiently poor beyond this point that most clinical and psychosocial research did not consider these later-term years. Currently, however, with large proportions of patients surviving even well

Table 1 Examples of interventions that have been evaluated for their ability to improve patient and family psychosocial outcomes during the organ transplantation process

<i>Intervention</i>	<i>Sample and study groups</i>	<i>Design</i>	<i>Key results</i>	<i>Reference</i>
Education classes (3x/week for 6 weeks) plus exercise and weight training sessions (2x/week for 6 weeks), compared to education alone. Education focused on lung disease, home health care, stress reduction techniques	Nine lung transplant candidates randomized to education plus exercise (n = 5) or education alone (n = 4)	Randomized controlled trial; assessments pre- and postintervention, no blinding	No significant between-group changes, but global QOL and physical functional status significantly improved in both groups. No other improvements in physical, emotional, or social well-being	Manzetti, J. D., Hoffman, L. A., Sereika, S. M., et al. (1994). Exercise, education, and quality of life in lung transplant candidates. <i>Journal of Heart and Lung Transplantation</i> 13 , 297–305
Telephone-based sessions (1x/week for 8 weeks) compared to usual care (medical management by transplant team). Telephone sessions were supportive counseling with cognitive-behavioral techniques to address stress, health, and coping	71 lung transplant candidates randomized to telephone (n = 36) or usual care (n = 35)	Randomized controlled trial; assessments pre- and postintervention; transplant team blinded to treatment assignment; assessors not blinded	Telephone group had lower somatic, anxiety, and depression symptoms and better global QOL, global mental health, and perceived social support postintervention, controlling for baseline levels. No group differences in physical functional QOL or perceived stress	Napolitano, M. A., Babyak, M. A., Palmer, S., et al. (2002). Effects of a telephone-based psychosocial intervention for patients awaiting lung transplantation. <i>Chest</i> 122 , 1176–1184
Individual compared to group psychotherapy (1x/week for 12 weeks). Each condition used systemic integrative psychotherapy, which focuses on patient understanding of problems and means to address them	89 kidney recipients randomized to individual (n = 49) or group therapy (n = 40). A control group receiving no treatment (n = 37) was also enrolled	Partially randomized trial, with the control group enrolled after all other subjects; assessment pre- and postintervention (or at similar intervals in controls), with 3-, 6-, 9-, and 12-month follow-up. No blinding	Both intervention groups showed significant reductions in depressive symptoms that were maintained through 1 year postintervention. Control group showed considerably lower depressive symptoms than intervention groups at baseline, and symptom levels rose over time. By the final assessment, symptom levels in the three groups were similar	Baines, L. S., Joseph, J. T. and Jindal, R. M. (2004). Prospective randomized study of individual and group psychotherapy versus controls in recipients of renal transplants. <i>Kidney International</i> 65 , 1937–1942
Psychosocial intervention (4 months exposure to Internet-based program) compared to usual care (medical management by transplant team). Internet intervention included components to address health education, stress management, psychological well-being, and medical compliance	60 heart recipients receiving either the intervention (n = 20) or usual care (n = 40). Each recipient's primary family caregiver was also enrolled	Controlled trial with historical comparison group (usual care); assessments pre- and postintervention (or at similar intervals in controls); transplant team blinded to intervention participation; assessors not blinded to intervention participation but blinded to study hypotheses	Intervention group patients had significantly improved depression and anxiety symptoms, and caregivers had significantly improved anxiety and anger symptoms, relative to the comparison group. Social functioning QOL significantly improved in intervention patients and caregivers (not assessed in comparison group). Patient medical compliance did not change, but the subgroup who used the component of the intervention focused on the medical regimen showed significant improvement in keeping clinic appointments, completing blood work, and following diet	Dew, M. A., Goycoolea, J. M., Harris, R. C., et al. (2004). An internet-based intervention to improve psychosocial outcomes in heart transplant recipients and family caregivers: development and evaluation. <i>Journal of Heart and Lung Transplantation</i> , 23 , 745–758

Mindfulness-based stress reduction program (1x/week for 8 weeks). Program consisted of training sessions and daily practice of meditation and gentle yoga techniques	20 transplant recipients (2 lung, 12 kidney, 5 kidney-pancreas, 1 pancreas)	Single-group trial; assessments pre- and postintervention and at 3 months follow-up; no blinding	Sleep, depressive symptoms, and overall mental health significantly improved pre- to postintervention. At 3-month follow-up, sleep improvement was maintained, but depressive symptoms had worsened. At 3-month follow-up, anxiety symptoms were improved from baseline. No changes in global QOL, physical functional QOL, or medical adherence	Gross, C. R., Kreitzer, M. J., Russas, V. et al. (2004). Mindfulness meditation to reduce symptoms after organ transplant: a pilot study. <i>Advances in Mind-body Medicine</i> 20 , 29–29
Telephone-based QOL therapy compared to telephone-based support therapy (1x/week for 8 to 12 weeks). QOL therapy used cognitive-behavioral strategies to address patient-identified problems in various QOL domains. Supportive therapy involved listening and providing information about the transplant experience	35 lung transplant candidates randomized to QOL therapy (n = 17) or supportive therapy (n = 18)	Randomized controlled trial; assessments pre- and 1-month, 3-month postintervention; transplant team blinded to treatment assignment	While the groups were equivalent at baseline, the QOL therapy group showed significantly better global QOL, lower mood disturbance, and a better relationship with spouse/partner at 1 month postintervention. QOL and mood improvements were maintained at 3-month postintervention	Rodrigue, J. R., Baz, M. A., Widows, M. R, et al. (2005). A randomized evaluation of quality-of-life therapy with patients awaiting lung transplantation. <i>American Journal of Transplantation</i> 5 , 2425–2432

beyond 10 years posttransplant, it is critical that the stressors, as well as the continued benefits, of transplantation be more carefully delineated.

Many of the complications that began but remained at low levels of severity in earlier years become symptomatic and can lead to major functional limitations beyond 5 years posttransplant. For example, chronic graft rejection, kidney disease that progresses to kidney failure, complications of diabetes, and cancers all increase in prevalence. In all types of organ transplant recipients, kidney failure, due to long-term immunosuppressant use, can precipitate the need for a kidney transplant. Many individuals who had originally received nonrenal transplants (i.e., liver, heart, lung) do not expect to eventually need another type of transplant, and learning that they now need a kidney transplant can reawaken the depression and anxiety that they experienced while awaiting their earlier transplants. Individuals who originally received kidney transplants can experience a profound sense of bereavement as their renal graft fails and they return to dialysis.

Recipients of nonrenal transplants may develop chronic rejection that is sufficiently severe that they require a retransplant with a new heart, liver, or pancreas. Retransplantation of nonrenal organs remains considerably less successful than first-time transplantation. Therefore, progressive graft failure for these individuals is likely to be associated with increasing psychological distress because treatment alternatives and the potential for good outcomes are poor.

In short, the accumulating medical burden experienced by all types of transplant recipients in the extended years posttransplant is likely to lead to decrements in all areas of well-being and QOL. Some of the health-related stressors and their likely sequelae will be similar to those experienced during the end-stage organ disease that necessitated the original transplant. Other stressors may be more similar to those experienced by other chronic disease and late-life populations as they approach the upper end of their life expectancies.

Interventions to Minimize Stressor Effects during the Transplant Process

Transplant programs have attempted a variety of strategies to help patients and their families cope with transplant-related stressors and/or to minimize the stressors' impact on patients' emotional well-being and QOL. These interventions have included educational, psychotherapeutic, and self-help support

groups, patient-to-patient mentoring programs, and the provision of educational and referral materials using a variety of audiovisual formats. The effectiveness of the majority of these programs has not been empirically evaluated. A very small literature to date has tested psychosocial interventions that could be employed either face to face when organ candidates and recipients return to the medical center for care or by telephone or via the Internet. Examples of these interventions and study findings regarding improved patient and family psychosocial outcomes are listed in [Table 1](#). More extensive intervention development and testing for effectiveness are critical needs in this field.

Acknowledgments

This article was supported in part by grant MH072718 from the National Institute of Mental Health, Rockville, MD, USA.

See Also the Following Articles

Disease, Stress Induced; Psychosomatic Medicine.

Further Reading

- Cupples, S. A. and Ohler, L. (eds.) (2002). *Solid organ transplantation: a handbook for primary health care providers*. New York: Springer Publishing Co.
- Dew, M. A., Switzer, G. E., DiMartini, A. F., et al. (2000). Psychosocial assessments and outcomes in organ transplantation. *Progress in Transplantation* 10, 239–259.
- Dew, M. A., Manzetti, J., Goycoolea, J. M., et al. (2002). Psychosocial aspects of transplantation. In: Smith, S. L. & Ohler, L. (eds.) *Organ transplantation: concepts, issues, practice and outcomes*, chap. 8. New York: Medscape Transplantation, WebMD, Inc. (www.medscape.com/transplantationhome).
- DiMartini, A. F., Dew, M. A. and Trzepacz, P. T. (2005). Organ transplantation. In: Levenson, J. L. (ed.) *The American Psychiatric Publishing textbook of psychosomatic medicine*, pp. 675–700. Washington, D.C.: The American Psychiatric Press, Inc.
- Olbrisch, M. E., Benedict, S. M., Ashe, K. and Levenson, J. L. (2002). Psychological assessment and care of organ transplant patients. *Journal of Consulting and Clinical Psychology* 70, 771–783.
- Rodrigue, J. R. (ed.) (2001). *Biopsychosocial perspectives on transplantation*. New York: Kluwer.
- Trzepacz, P. T. and DiMartini, A. F. (eds.) (2000). *The transplant patient: biological, psychiatric and ethical issues in organ transplantation*. Cambridge, UK: Cambridge University Press.

Oxidative Stress

H Sies

Heinrich-Heine-Universität, Düsseldorf, Germany

D Jones

Emory University, Atlanta, GA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by H Sies, volume 3, pp 102–104, © 2000, Elsevier Inc.

Oxidative Stress Definitions

Related Terms

Reductive Stress

Nitrosative Stress

Concluding Remarks

Glossary

Oxidative stress An imbalance between oxidants and antioxidants in favor of the oxidants, leading to a disruption of redox signaling and control and/or molecular damage.

Oxidative Stress Definitions

The imbalance between oxidants and antioxidants in favor of the oxidants, potentially leading to damage, forms the core of the definition of oxidative stress. In the biological-biomedical field oxidative stress denotes a disturbance in the pro-oxidant–antioxidant balance in favor of the former. Much experimental evidence now supports the thesis that the favorable aspects of aerobic life are also linked to potentially dangerous oxygen-linked oxidation processes. These are thought to form the basis of a number of physiological and pathophysiological phenomena and to participate in processes as diverse as inflammation, atherosclerosis, aging, carcinogenesis, drug action and drug toxicity, and defense against protozoa. Research on oxidative stress has intensified in recent years, as documented by the close to 42,000 entries at the ISI Web of Knowledge at the beginning of 2005, mainly about the role of oxidants in cellular redox signaling.

In view of the proliferation of publications using the term, a few cautionary words might be appropriate. How can oxidative stress be defined operationally? In the normal healthy state, oxidative challenge occurs in many cell types, but this alone does not constitute an oxidative stress. Likewise, a simple loss of antioxidants resulting from a limited nutritional

supply is not sufficient. However, when there is an increased formation of pro-oxidants such as hydrogen peroxide, which is accompanied by a loss of glutathione caused by the formation of glutathione disulfide (GSSG), we approach a definition. Even a severe loss of antioxidants, however, may still mean that there is no resulting damage. Novel roles of oxidants and antioxidants as part of signaling cascades and as mediators of adaptive responses have been unraveled, so it is possible that the time course and even the final outcome of a variety of diseases can be critically modulated by strengthening the antioxidant side of the balance.

This concept can be further refined by a better understanding of the normal biological uses of the reactive species that cause oxidative stress. The biological functions of reactive species that cause oxidative stress include fundamental processes such as the signaling of cell division and controlled cell elimination in the renewal of tissues. Furthermore, the balance of oxidants and antioxidants cannot be defined by a single entity. This has led to the view that multiple antioxidants are needed to protect cells and organs (the antioxidant network) and that different antioxidants are part of a protection system in specific disease processes. Following this line of reasoning, a refinement in the definition of oxidative stress has been suggested: “a disruption of redox signaling and control”. Such a definition focuses attention on abnormal routes of electron flow, which both alter normal physiological signaling and damage the macromolecular machinery.

As these definitions indicate, oxidative stress is a concept that is becoming clearer with further scientific development, each ramification and extension providing valuable further input. Redox chemistry involves reduction–oxidation reactions, so obviously tipping the balance in favor of the reductants may also cause an imbalance (reductive stress). A landmark in the development of oxidative stress is the 1954 work of Gerschman, who viewed oxygen toxicity along with x-ray irradiation. A summary of the accumulated knowledge of the field appears in a 1979 review of hydroperoxide metabolism in mammalian organs by Chance et al. In a 1970 study of glutathione in red blood cells, the effect of glutathione reductase deficiency on the stimulation of a hexose monophosphate shunt under oxidative stress was examined by Beutler and colleagues; this may have been the first mention of the term, at least in the title of a publication that is readily retrievable.

Related Terms

Dietary Oxidative Stress

The term dietary oxidative stress was coined by Levander in 1995. The relationship between oxidative stress and viral infections has been scrutinized in regard to this concept. This area has received much attention from food science and in functional food considerations. In a definition of dietary antioxidants published by the Food and Nutrition Board of the Institute of Medicine, National Academy of Sciences, in 2000, the term was defined as “as a substance in foods that significantly decreases the adverse effects of reactive species, such as reactive oxygen species, reactive nitrogen species, on normal physiological function in humans.” The ability of dietary anti- and pro-oxidants to affect oxidative stress has been widely discussed; Ames et al. have emphasized the presence of carcinogens and anticarcinogens in food.

Nutritional, or dietary, oxidative stress denotes a disturbance of the redox state resulting from excess oxidative load or from an inadequate nutrient supply that favors pro-oxidant reactions. The low intake of dietary antioxidants, including vitamins E and C, carotenoids, polyphenols, and other micronutrients (e.g., selenium), or the impaired availability of these antioxidants weakens the antioxidant network.

Postprandial oxidative stress, subform of nutritional oxidative stress, ensues from sustained postprandial hyperlipidemia and/or hyperglycemia and is associated with a higher risk for atherosclerosis, diabetes, and obesity. In Western societies, a significant part of the day is spent in the postprandial state. Unsaturated fatty acids incorporated into low-density lipoproteins (LDLs) and oxidized LDLs are an atherogenic factor. Lipid hydroperoxides present in the diet are absorbed and contribute to the pro-oxidant load. In hyperlipidemic and hyperglycemic subjects, endothelium-dependent vasodilation is impaired in the postprandial state, making postprandial oxidative stress an important factor modulating cardiovascular risk. Postprandial oxidative stress is attenuated when dietary antioxidants are supplied together with a meal rich in oxidized or oxidizable lipids. The ingestion of dietary polyphenols (e.g., from wine, cocoa, or tea) improves endothelial dysfunction and lowers the susceptibility of LDL lipids to oxidation. Polyphenols affect endothelial function not solely as antioxidants but also as modulatory signaling molecules.

Physiological Oxidative Stress

The fact that oxygen and its metabolites are not evenly distributed in organs and in subcellular organelles has led to the consideration that some sites are

under physiological oxidative stress. The mitochondria are considered to be a site of physiological oxidative stress; in fact, it is possible that basically all cells, in the quasi-normal situation, experience this form of normal challenge. In view of the growing interest in adaptive responses, it is of interest to identify the need for and the extent of the physiological exposure to oxidative stress. The basal level of macrophage and neutrophil activation, for example, may fall into this category. Likewise, the physiological activity of glutamate (NMDA) receptors could be included here. Thus, in specific locations in the physiological system there is a routine deployment of oxidants for normal functioning. A difficulty arises, of course, in defining the level at which the threshold for the pathological oxidative stress is transcended. It may be that the accumulation of damage products, as part of the general definition of oxidative stress, could be a useful parameter in this regard. The adaptation to oxidative stress, initially studied predominantly with bacteria (e.g., in the *oxyR*-response) has also been demonstrated in eukaryotes and may have interesting further functions in the mammalian organism. The field of cell and tissues responses in terms of gene expression in mammalian systems is developing rapidly. The term exercise-induced oxidative stress may also fall under the present heading.

Spontaneous mutations and aging can be considered to be the result of the exposure of the organism to oxidative stress, be it physiological or pathophysiological. The loss of the *oxyR* regulon, which controls the expression of a battery of oxidant-defense genes, leads to an enormous increase in the number of spontaneous mutations. The relationship of oxidative stress to the aging process has been studied in many ways.

Photooxidative Stress

The generation of oxidants by electronic excitation through impinging light represents a major part of photobiology. The photoexcitation of endogenous or exogenous sensitizer molecules, leading to the formation of reactive oxygen species, notably singlet molecular oxygen, can be the cause of molecular and structural damage. The wavelength ranges of biological importance are ultraviolet B (UVB) and ultraviolet A (UVA), but visible light and even infrared (IR) are known to generate photobiological responses. In human health, exposed tissues such as skin and the eye are the most studied. There is nutritional protection against this damage from sunlight in dietary micronutrients such as carotenoids and polyphenols (flavonoids). Photodynamic therapy (PDT) uses the biological effects of photooxidative stress.

Radiation-Induced Oxidative Stress

The chemical basis of radiation biology involves the generation of free radicals. This is a large research field, ranging from chemistry to the treatment of cancer.

Oxidant Stress

Oxidant stress has been used much less frequently. This term (although it sounds slightly ambiguous) may be considered to be more or less synonymous with oxidative stress.

Pro-oxidant Stress

Pro-oxidant stress was used by Levander in his 1979 work on metal toxicity, again using red blood cells as a model, in describing the effects of lead. Cerutti used the term pro-oxidant states in his analysis of the role of oxidants in tumor promotion.

Oxidative Stress Status

The problem of assessing the biological status of intact systems has been tantalizing. Noninvasive methods are preferable, and stable end points are good candidates for useful applications. Our early work was on using GSSG efflux as an indicator. Many efforts were directed toward this (e.g., the use of 8-oxodG or certain isoprostanes in urinary samples). The study of biomarkers of oxidative stress is being actively pursued.

Approaches are evolving to evaluate oxidative stress and its association with the risk of disease in humans. Most methods are currently available only on an experimental basis. A key approach involves genetic analyses of variations in key antioxidant systems that have been associated with disease risk. Clinical measures of the status of the nutritionally required antioxidants vitamins C and E are available, but there is no consensus on the spectrum of additional antioxidants that should be measured for an effective evaluation of oxidative stress-related risk. The glutathione redox state and the determination of reactive oxygen metabolites provide more dynamic measures of the ongoing balance of pro-oxidant and antioxidant processes. In the blood plasma, the amino acid cysteine and its disulfide cystine are much more abundant than the glutathione system and provide an additional indicator of the balance of pro-oxidant and antioxidant systems. The cysteine/cystine balance in human plasma has been correlated with oxidative stress in aging, cigarette smoking, and cardiovascular disease. Although no consensus exists on ways to measure the damage from oxidative stress, measurement of isoprostanes, the free radical

products of lipid oxidation found in the urine, by mass spectrometry is becoming a standardized and more generally accepted procedure. For the measurement of damage from reactive nitrogen species, the mass spectrometry measurement of nitrotyrosine in plasma proteins is being used as well.

Reductive Stress

In analogy to hyperoxia's being linked to oxidative stress, hypoxia or anoxia may be linked to reductive stress. The term reductive stress was introduced in 1987 by Wendel to denote the metabolic initiation of iron redox cycling following the overproduction of nicotinamide adenine dinucleotide (NADH). From a toxicological point of view, therefore, nicotinamide adenine dinucleotide phosphate (NADPH), or NADH, can be considered to be a detoxicant and a toxicant.

The effects caused by respiratory inhibition (chemical hypoxia), which favors the formation of toxic oxygen species, have been called reductive. The role of vascular endothelial growth factor in mediating the vascular dysfunction induced by hypoxia-like cytosolic metabolic imbalances, such as reductive stress and increased superoxide and nitric oxide production, has been studied. A gene battery responsive to reductive stress has been analyzed.

Nitrosative Stress

While working with nitrosothiols, Hausladen et al. observed a transcriptional activation of the *oxyR* regulon, which they called nitrosative stress. This area, which is of particular interest to vascular physiology and pathophysiology, has been treated in some detail.

Concluding Remarks

It is obvious that the term oxidative stress evokes different associations to researchers coming from different fields of interest. Thus, sometimes it has been overused, and caution is advised when the molecular species are not specified in the transition under study. The various concepts presented here may, in fact, not be at variance with one another; rather, the various terms emphasize a common phenomenon. Likewise, there are many parallelisms between the reactions distinguished as nitrosative versus oxidative. This shows that protection can be afforded against oxidative stress and nitrosative stress simultaneously, underlining the close relationship between the two.

Further Reading

- Ames, B. N., Shigenaga, M. K. and Hagen, T. M. (1993). Oxidants, antioxidants, and the degenerative diseases of aging. *Proceedings of the National Academy of Sciences USA* **90**, 7915–7922.
- Cerutti, P. A. (1985). Prooxidant states and tumor promotion. *Science* **227**, 375–381.
- Chance, B., Sies, H. and Boveris, A. (1979). Hydroperoxide metabolism in mammalian organs. *Physiological Reviews* **59**, 527–605.
- Cutler, R. G. and Rodriguez, H. (eds.) (2003). *Oxidative stress and aging*. Singapore: World Scientific Publishing.
- Dalle-Donne, I., Scaloni, A., Giustarini, D., et al. (2005). Proteins as biomarkers of oxidative/nitrosative stress in diseases: the contribution of redox proteomics. *Mass Spectrometry Reviews* **24**, 55–99.
- Food and Nutrition Board, Institute of Medicine (2000). Dietary reference intakes for vitamin C, vitamin E, selenium, and carotenoids, p. 17. National Academy Press: Washington.
- Griffiths, H. R., Moller, L., Bartosz, G., et al. (2002). Biomarkers. *Molecular Aspects of Medicine* **23**, 101–208.
- Halliwell, B. (1996). Oxidative stress, nutrition and health: experimental strategies for optimization of nutritional antioxidant intake in humans. *Free Radical Research* **25**, 57–74.
- Hausladen, A., Privalle, C. T., Keng, T., et al. (1996). Nitrosative stress: activation of the transcription factor OxyR. *Cell* **86**, 719–729.
- Hermes-Lima, M. (2004). Oxidative stress and medical sciences. In: Story, K. (ed.) *Functional metabolism: regulation and adaptation*, pp. 369–382. New York: John Wiley.
- Jones, D. P., Mody, V. C. Jr., Carlson, J. L., et al. (2002). Redox analysis of human plasma allows separation of prooxidant events of aging from decline in antioxidant defenses. *Free Radical Biology & Medicine* **33**, 1290–1300.
- Levander, O. A., Fontela, R., Morris, V. C., et al. (1995). Protection against murine cerebral malaria by dietary-induced oxidative stress. *Journal of Parasitology* **81**, 99–103.
- Paniker, N. V., Srivastava, S. K. and Beutler, E. (1970). Glutathione metabolism of the red cells: effect of glutathione reductase deficiency on the stimulation of hexose monophosphate shunt under oxidative stress. *Biochimica et Biophysica Acta*, **215**, 456–460.
- Pryor, W. A. and Godber, S. S. (1991). Noninvasive measures of oxidative stress status in humans. *Free Radical Biology & Medicine* **10**, 177–184.
- Schwarz, K. B. (1996). Oxidative stress during viral infection: a review. *Free Radical Biology & Medicine* **21**, 641–649.
- Sies, H. (1985). Oxidative stress: introductory remarks. In: Sies, H. (ed.) *Oxidative stress*, pp. 1–8. London: Academic Press.
- Sies, H. (1986). Biochemistry of oxidative stress. *Angewandte Chemie International Edition in English* **25**, 1058–1071.
- Sies, H. (ed.) (1997). *Antioxidants in disease mechanisms and therapy*. San Diego: Academic Press.
- Sies, H. (2000). What is oxidative stress? In: Keaney, J. F. (ed.) *Oxidative stress and vascular diseases*, pp. 1–8. Boston: Kluwer Academic.
- Sies, H. and Stahl, W. (2004). Nutritional protection against skin damage from sunlight. *Annual Review of Nutrition* **24**, 173–200.
- Sies, H., Stahl, W. and Sevanian, A. (2005). Nutritional, dietary and postprandial oxidative stress. *Journal of Nutrition* **135**, 969–972.
- Sohal, R. S. and Weindruch, R. (1996). Oxidative stress, caloric restriction, and aging. *Science* **273**, 59–63.
- Storz, G., Christman, M. F., Sies, H., et al. (1987). Spontaneous mutagenesis and oxidative damage to DNA in *Salmonella typhimurium*. *Proceedings of the National Academy of Sciences USA* **84**, 8917–8921.
- von Sonntag, C. (1987). *The chemical basis of radiation biology*. London: Taylor & Francis.
- Wendel, A. (1987). Measurement of *in vivo* lipid peroxidation and toxicological significance. *Free Radical Biology & Medicine* **3**, 355–358.
- Wink, D. A., Cook, J. A., Kim, S. Y., et al. (1997). Super-oxide modulates the oxidation and nitrosation of thiols by nitric oxide-derived reactive intermediates: chemical aspects involved in the balance between oxidative and nitrosative stress. *Journal of Biological Chemistry* **272**, 11147–11151.

Oxidative Stress and Acidosis, Molecular Responses to

C Mobbs

Mt. Sinai School of Medicine, New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by C Mobbs, volume 3, pp 105–108, © 2000, Elsevier Inc.

Oxidative Stress and Acidosis

Molecular Responses to Oxidative Stress and Acidosis

Mechanisms Mediating Molecular Responses to Oxidative Stress and Acidosis

Summary

Glossary

<i>Acidosis</i>	An increase in the acidity, and thus a decrease in pH.
<i>Cysteine</i>	An amino acid (subunit of protein) particularly sensitive to oxidation.
<i>Cytochrome B</i>	An essential component of mitochondrial metabolism, which functions in electron transport chain complex III.
<i>Cytochrome oxidase</i>	An essential component of mitochondrial metabolism, which functions to facilitate electron transport.
<i>Free radical</i>	A molecular species such as O_2^- , which by virtue of a free electron is energetically favored to reactively oxidize many compounds and cellular constituents.
<i>Heme oxidase 1</i>	An enzyme that protects cells from oxidative damage.
<i>Metallothionein</i>	A protein that protects cells from oxidative damage.
<i>Mitochondrial rRNAs</i>	Ribosomal RNAs used to translate proteins in the mitochondria.
<i>Oxidation</i>	A chemical reaction in which a compound or molecular constituent loses electrons.
<i>Oxidative state</i>	A reflection of the likelihood that atoms in the environment will gain or lose net electrons or hydrogen atoms; in a more highly oxidative state sulfur atoms in the side chains of the amino acid cysteine will tend to lose net electrons and hydrogen atoms, thus enhancing the production of disulfide bridges.
<i>Oxidative stress</i>	Potentially toxic increase in oxidative state of the cell, leading to characteristic molecular responses.
<i>OxyR protein</i>	A protein that regulates gene expression which, when activated by oxidative stress, stimulates the expression of anti-oxidant genes.

Reduction

A chemical reaction in which a compound or molecular constituent gains electrons.

Superoxide dismutase

An enzyme that protects against oxidative stress by inactivating the highly oxidant free radical O_2^- .

Oxidative Stress and Acidosis

The internal milieu of the cell is highly regulated, and any displacement from this milieu can have fatal consequences. Therefore molecular mechanisms have evolved to counter environmental influences which threaten to disrupt this milieu. These molecular mechanisms entail responses which are common (stereotypic) in a wide range of cells and in response to a wide variety potentially toxic influences. Thus, these molecular responses may usefully be considered to be molecular stress responses. However, these general stress responses are also accompanied by more specific changes which depend on the cell type and specific stressor. Perhaps the best-studied of these molecular stress responses are the heat-shock proteins, which were first observed to be induced by a potentially lethal increase in temperature. The heat-shock proteins are discussed in detail elsewhere in this encyclopedia. As discussed there, the heat-shock proteins are not only induced by elevated temperature, but also by a wide variety other stressors, among which are changes in the oxidative state and the pH of the cell. However, oxidative stress and acidosis also induce other molecular responses which are not prominent responses to heat shock. These other responses are the focus of this present article.

Oxidative state can best be thought of as the relative attraction of electrons to atoms or compounds. Oxygen is one of the strongest attractors of electrons, so it is relatively oxidizing comparing to almost any other atom. In comparison, hydrogen has relatively little attraction for electrons. Thus when hydrogen and oxygen combine to form water, the hydrogen is oxidized by (loses electrons to) oxygen, and, conversely, the oxygen is reduced by (gains electrons from) hydrogen. In a more general sense, it may be said that the atmosphere, consisting of a relatively high level of oxygen due to ongoing photosynthesis, is characterized by a relatively high oxidative state compared, for example, to the internal state of the cell, in which free oxygen is at much lower concentrations. This disparity partly reflects the fact that life originally evolved in

atmospheric conditions which were much less oxidative than exists today. The physiological significance of this disparity is that the proper function of most proteins requires an oxidative state lower than is found in outside the cell. Even within the cell, some compartments, such as the endoplasmic reticulum, require even more highly reduced conditions than other parts of the cell. However, most cells require oxygen to produce energy, and in the course of producing energy molecular species called free radicals are produced which potentially greatly enhance the local oxidative state inside the cell. Furthermore, other cellular functions, such as certain immunological functions, may entail enhanced production of free radicals. When free radicals or oxidative chemicals increase the oxidative state inside the cell, the result can be the oxidation of proteins, especially at disulfide bonds that are highly sensitive to oxidative status. Because such an increase in oxidative state may be toxic and induces a common set of molecular responses, this change is called oxidative stress. In principle cellular function could also become impaired if oxidative state becomes too low (that is, the internal milieu could be too reduced), but this circumstance rarely occurs in most cells and is not further considered in this article.

In addition to the highly regulated oxidative state, the acidity, or pH, inside the cell is also highly regulated. Acidity, which reflects the concentration of free hydrogen ions, is critically important because the ionic status of most amino acid side chains in proteins is dependent on pH. Therefore, agents that alter internal pH can, like agents that alter oxidative status, also be toxic, and thus not surprisingly cellular responses to acidosis have also evolved. Much less is known about these responses, but molecular responses to acidosis and oxidative stress entail induction of some common species.

Molecular Responses to Oxidative Stress and Acidosis

Molecular responses to oxidative stress entail some common responses and some responses that depend on the cell type and nature of the oxidative stressor. For example, as with many potentially toxic insults, oxidative stress may generally entail induction of heat-shock proteins. However, other responses are different from those induced by heat shock. The range of molecular responses to oxidative stress is indicated by comparing three papers that used the differential display technique to assess molecular responses of three different cell types to either hydrogen peroxide or diethylmaleate. Hydrogen peroxide directly increases oxidative state since it is a highly

oxidative molecule, whereas diethylmaleate increases the oxidative state by depleting the reduced form of glutathione. In a Chinese hamster fibroblast cell line, hydrogen peroxide was observed to cause a decrease in a several transcripts derived from mitochondria (including both 16S and 12S mitochondrial rRNAs, nicotinamide adenine dinucleotide (NADH) dehydrogenase subunit 6; ATPase subunit 6; and cytochrome oxidase subunits I and III). In contrast, most nucleus-derived ribonucleic acid (RNA) species were either not influenced or increased, including messenger RNAs (mRNAs) for several growth-inhibiting proteins. In a human lens epithelial cell line, hydrogen peroxide was observed to influence transcripts derived from both mitochondrial and nuclear genomes. The mitochondrial transcripts (including NADH dehydrogenase subunit 4, cytochrome-b, cytochrome oxidase III, and 16S RNA) were decreased. In contrast, all but one of the nucleus-derived transcripts that were regulated (cytokine-inducible nuclear protein, alternative splicing factor 2, and beta-hydroxyisobutyryl-coenzyme A hydrolase) were increased. Among the nucleus-derived transcripts, only glutamine cyclotransferase was decreased. In the human-derived Hela cell line, diethylmaleate was observed to decrease the levels of RNA species derived from mitochondria, whereas nuclear-derived species, including the ribosomal protein 4, c-fos, and previously uncharacterized mRNAs, were induced. Thus the differential display technique, which in effects surveys across all RNA species, clearly indicates that a common response to oxidative stress is a decrease in mitochondrial RNA species and an induction of a small number of nuclear-derived mRNAs. Since oxidative stress would normally be produced by mitochondrial oxidative metabolism, the decrease in mitochondrial species can be understood as a homeostatic response to decrease the production of oxidative stress. This common oxidative stress response, however, is not observed with heat shock. In contrast, other proteins, including antioxidant proteins such as glucose-6-phosphate dehydrogenase (which catalyzes the rate-limiting step in the production of nicotinamide adenine dinucleotide phosphate hydrogen (NADPH), the source of reducing potential for the antioxidant glutathione), catalase, superoxide dismutase, heme oxidase 1, and metallothionein, are induced by many forms of oxidative stress in many cell types. While these gene products may be induced by thermal or other stressors under some circumstances, their induction by oxidative stress is much more robust. Interestingly, other than heat-shock proteins, the most robust reported response to acidosis is also metallothionein. Both metallothionein and heme oxidase 1 probably serve as antioxidants

by sequestering potentially oxidizing metals, but this remains to be established.

The molecular responses to oxidative stress depend on the severity of the stress, in relation to whether the stress is lethal, produces programmed cell death (apoptosis), or is relatively mild and not only allows the cell to survive but become relatively resistant to subsequent cell death. DNA microarray studies suggest that relatively mild oxidative stress produce molecular responses that could lead to apoptosis, but these potentially lethal responses are inhibited, interestingly, by a concomitant reduction in glycolysis. This reduction in glycolysis actually appears to produce a protective effect, in part by inhibiting apoptosis, but through other mechanisms as well. However, if the oxidative stress becomes severe enough, the apoptotic effects overwhelm the blockade produced by glycolysis inhibition. Even more severe oxidative stress produces cell death directly, without apoptosis (necrosis).

Mechanisms Mediating Molecular Responses to Oxidative Stress and Acidosis

In eukaryotic systems, oxidative stress appears to lead to degradation of mitochondrial RNA species through unknown mechanisms. Oxidative stress increases nuclear-derived mRNAs through several mechanisms, including increased transcription and, as for heme oxidase 1, increased mRNA stability. The molecular details of these effects of oxidative stress in eukaryotes is not well-understood. However, oxidative stress responses in prokaryotes have been examined in some detail.

In bacteria, hydrogen peroxide induces about 30–40 proteins, whereas agents which lead to free production induce an approximately 40 additional proteins. Hydrogen peroxide induces gene expression in part by activating the OxyR protein, which activates transcription of about 10 antioxidant proteins including catalase (the enzyme which converts hydrogen peroxide to water), and glutathione reductase. OxyR is normally inactive in the relatively reduced normal cellular milieu, but hydrogen peroxide leads to an oxidation of two cysteine residues on OxyR which activates the protein. Free radical production induces gene expression in part through the oxidation of SoxS, a transcriptional activator which stimulates the transcription of several antioxidant genes including superoxide dismutase and glucose-6-phosphate dehydrogenase. The oxidation of SoxS takes place at an iron–sulfur center, so requires a higher energy form of oxidation than is produced by hydrogen peroxide alone. Thus, molecular responses to different

forms of oxidative stress can be appropriately generated through effector molecules with sensitivities to different forms of oxidation. Recent studies suggest that a similar process occurs in eukaryotes, entailing the activation of the transcription factor Nrf2 via the mitogen-activated protein (MAP) kinase signaling pathway.

Summary

Life initially evolved under conditions of much lower oxygen than is the case at present and most cellular processes therefore require a lower state of oxidation than is found in air. Therefore specific and highly regulated processes have evolved to maintain this lower state of oxygen even in the face of metabolic insults. Several levels of antioxidant defenses exist, to accommodate precise levels of oxidative challenges. Some of these antioxidative proteins, including metallothioneins, are also induced to protect against acidosis.

Further Reading

- Ammendola, R., Fiore, F., Esposito, F., et al. (1995). Differentially expressed mRNAs as a consequence of oxidative stress in intact cells. *FEBS Letters* 371, 209–213.
- Colussi, C., Albertini, M. C., Coppola, S., et al. (2000). H₂O₂-induced block of glycolysis as an active ADP-ribosylation reaction protecting cells from apoptosis. *Faseb Journal* 14, 2266–2276.
- Carper, D. A., Sun, J. K., Iwata, T., et al. (1999). Oxidative stress induces differential gene expression in a human lens epithelial cell line. *Investigative Ophthalmology & Visual Science* 40, 400–406.
- Choi, A. M. and Alam, J. (1996). Heme oxygenase-1: function, regulation, and implication of a novel stress-inducible protein in oxidant-induced lung injury. *American Journal of Respiratory Cell and Molecular Biology* 15, 9–19.
- Conklin, D. R., Tan, K. H. and Aschner, M. (1998). Dimethyl sulfoxide, but not acidosis-induced metallothionein mRNA expression in neonatal rat primary astrocyte cultures is inhibited by the bioflavonoid, quercetin. *Brain Research* 794, 304–308.
- Crawford, D. R., Wang, Y., Schools, G. P., et al. (1997). Down-regulation of mammalian mitochondrial RNAs during oxidative stress. *Free Radical Biology & Medicine* 22, 551–559.
- Demple, B. (1999). Radical ideas: genetic responses to oxidative stress. *Clinical and Experimental Pharmacology and Physiology* 26, 64–68.
- Narasimhan, P., Swanson, R. A., Sagar, S. M., et al. (1996). Astrocyte survival and HSP70 heat shock protein induction following heat shock and acidosis. *Glia* 17, 147–159.
- Nebert, D. W., Petersen, D. D. and Fornace, A. J., Jr. (1990). Cellular responses to oxidative stress: the [Ah] gene

- battery as a paradigm. *Environmental Health Perspectives* 88, 13–25.
- Nishimura, R. N., Dwyer, B. E., Cole, R., et al. (1989). Induction of the major inducible 68-kDa heat-shock protein after rapid changes of extracellular pH in cultured rat astrocytes. *Experimental Cell Research* 180, 276–280.
- Pahl, H. L. and Baeuerle, P. A. (1994). Oxygen and the control of gene expression. *Bioessays* 16, 497–502.
- Piette, J., Piret, B., Bonizzi, G., et al. (1997). Multiple redox regulation in NF-kappaB transcription factor activation. *Biological Chemistry* 378, 1237–1245.
- Ryter, S. W. and Choi, A. M. (2005). Heme oxygenase-1: redox regulation of a stress protein in lung and cell culture models. *Antioxidants & Redox Signaling* 7, 80–91.
- Schreck, R., Albermann, K. and Baeuerle, P. A. (1992). Nuclear factor kappa B: an oxidative stress-responsive transcription factor of eukaryotic cells (a review). *Free Radical Research Communications* 17, 221–237.
- Tawe, W. N., Eschbach, M. L., Walter, R. D., et al. (1998). Identification of stress-responsive genes in *Caenorhabditis elegans* using RT-PCR differential display. *Nucleic Acids Research* 26, 1621–1627.
- Zhang, Y., Fong, C. C., Wong, M. S., et al. (2005). Molecular mechanisms of survival and apoptosis in RAW 264.7 macrophages under oxidative stress. *Apoptosis* 10, 545–556.

Oxidative Stress and Aging

C Mobbs

Mt. Sinai School of Medicine, New York, USA

© 2007 Elsevier Inc. All rights reserved.

Reduction

A chemical reaction in which a compound or molecular constituent gains electrons.

Metabolism Drives Aging: The Free Radical Hypothesis
Oxidative Stress Increases with Age
Functional Evidence That Oxidative Stress Limits Life Span

Glossary

<i>Free radical</i>	A molecular species such as O_2^- , which by virtue of a free electron is energetically favored to reactively oxidize many compounds and cellular constituents.
<i>Mitochondria</i>	The powerhouse of the cell, where ATP, the main source of cellular energy, is produced.
<i>Mitochondrial electron transfer chain</i>	A specialized process in the mitochondria whereby electrons are passed down a chain of five complexes, allowing for the production of ATP.
<i>Oxidation</i>	A chemical reaction in which a compound or molecular constituent loses electrons.
<i>Oxidative state</i>	A reflection of the likelihood that atoms in the environment will gain or lose net electrons or hydrogen atoms; in a more highly oxidative state, sulfur atoms in the side chains of the amino acid cysteine tend to lose net electrons and hydrogen atoms, thus enhancing the production of disulfide bridges.
<i>Oxidative stress</i>	The potentially toxic increase in oxidative state of the cell, leading to characteristic molecular responses.

Metabolism Drives Aging: The Free Radical Hypothesis

The causes of aging have been the subject of speculation since before the dawn of recorded history, and in the era of science such speculations have become even more numerous. At present, to the extent that a unitary mechanism of sorts is seriously entertained by gerontologists, the hypothesis that aging is caused by oxidative damage is the most widely accepted. This hypothesis had its origins in the observation that life span seems inversely correlated with metabolic rate, which gave rise, early in the twentieth century, to the rate of living theory. This hypothesis, inspired by the fantastically successful conservation laws of classical physics, hypothesized that the total amount of metabolic activity an individual could sustain over its life span is a constant. This hypothesis accorded well with the basic observation with inanimate objects that use leads to deterioration and with the intuition that, in human life at least, those who live fast, die young; it also accorded well with a more objective analysis that, for example, heart rate is much higher in short-lived mammals than in long-lived mammals. Indeed, this observation gave rise to the hypothesis that the lifetime number of heartbeats, thought to be a proxy of metabolic rate, is conserved. The hypothesis was strengthened by the incontrovertible demonstration that reducing the temperature in

poikilotherms such as flies and nematodes increased life span. Because reducing the temperature *in vitro* reduces the rate of metabolic and chemical reactions, these observations were highly suggestive that metabolic processes drive aging and limit life span.

Nevertheless, precisely which aspects of metabolism limits life span has continued to be elusive. Recent studies have suggested that total metabolic activity is in fact not conserved; for example, caloric restriction can increase life span without producing a concomitant decrease in tissue-specific metabolic rate. In 1956, by analogy with radiation-induced cellular damage, Dennis Harman suggested that the metabolic process that limits life span is the production of free radicals whose main damaging effect is mediated through the oxidation of macromolecules such as protein and DNA. This free-radical hypothesis of aging emphasized the role of oxidative stress, although in principle oxidative stress could occur through other mechanisms in addition to free radicals. This hypothesis has continued to stimulate much research and continues to be among the most attractive mechanisms to mediate the aging process.

Oxidative stress, defined as evidence of a relatively oxidized state of the cell and its constituents, has been implicated in the aging process by several lines of evidence. Originally, this hypothesis was suggested by the observation that radiation, which can certainly produce oxidative stress, can in some ways mimic the aging process, especially in reducing life span and increasing the incidence of tumor burden. On the other hand, it is not certain that oxidative damage necessarily mediates the pathological effects of radiation, and it has been similarly difficult to prove that oxidative stress mediates the pathologies of aging. Nevertheless, evidence implicating oxidative stress as mediating the pathological effects of aging has continued to grow over the last 50 years.

Oxidative Stress Increases with Age

Perhaps the most obvious evidence is that oxidative stress incontrovertibly increases with age. For example total protein oxidation, as indicated by protein carbonyl levels and nitrotyrosine immunoreactivity, increases by approximately 50% in the brain from maturity to old age. Conversely, total protein reduction, as indicated by sulfhydryl groups, decreased with age. For some proteins, such as glyceraldehyde phosphate dehydrogenase, the age-related oxidation has been located at specific cysteine residues, and the oxidation has been shown to account for the robust decrease in specific activity characteristic of many enzymes during aging. Furthermore, protein oxidation in specific brain regions of aged mice has been

shown to correlate with impairments in functions regulated by those specific brain regions. Thus, for example, the oxidation of hippocampal proteins correlates with impairments in memory function, whereas the oxidation of cerebellar proteins correlates with impairments in tasks requiring fine motor coordination. Oxidative damage to DNA similarly increases with age, especially in mitochondrial DNA. In mitochondria, DNA oxidative damage is mainly manifested by characteristic deletions that can entail the loss of several thousand bases and whose occurrences increases robustly with age in many tissues. In nuclei, DNA oxidation is mainly manifest as the modification of specific bases, especially the conversion of guanine to 8-oxoguanine, and specific oxidized bases in the nucleus, which are most likely due to oxidative damage. Finally, the oxidation of lipids, mainly manifest in the form of 4-hydroxynonenal (HNE) and malondialdehyde, also increases with age in several tissues.

The increased oxidative damage that occurs during aging appears to be due primarily to the increased production of reactive oxygen species during aging, although reduced turnover of proteins and lipids may also play a role. Most reactive oxygen species are produced by the mitochondria, especially in electron transport chain complexes I and III. It is not clear why the production of reactive oxygen species increases with age, but in some tissues at least this may be due to a shift toward the greater use of complex I and away from the use of complex II.

Functional Evidence That Oxidative Stress Limits Life Span

Four main lines of evidence suggest that the increase in oxidative damage that occurs during aging plays an important role in determining life span. First, age increases the lethality of oxidative stress, for example in *Caenorhabditis elegans*. Second, dietary restriction, which increases life span, also increases resistance to oxidative stress. Third, genetic modifications that increase life span (e.g., mutations in the insulin-like receptor) also increase resistance to oxidative stress. Fourth, genomewide unbiased screening revealed that the most prominent class of genes whose ablation increases life span are genes coding for mitochondrial electron transport chain complexes I, III, IV, and V. As previously indicated, these complexes, especially I and III, are the major sources of reactive oxygen species in most cells. Notably missing from these screens were genes coding for proteins in mitochondrial electron transport chain complex II. Indeed, it had already been shown by classic genetic screens that mutations in complex II reduced life span rather than

increasing it. The significance of these observations is that reactive oxygen species are produced at much greater levels in complex I than in complex II. Thus, mutations that reduce complex II activity presumably entail elevated compensatory activity of complex I, increasing the production of reactive oxygen species and reducing life span, whereas genetic manipulations that inhibit complex I have the opposite effect. Precisely such a mechanism seems to mediate the age-related pathologies (and reduced life span) caused by Huntington's disease.

A final line of evidence supporting a role for oxidative stress in life span involve studies that genetically manipulate genes coding for the classic anti-oxidative defense enzymes superoxide dismutase and catalase. Although early studies in flies suggested that transgenic overexpression of these genes might increase life span, later studies failed to replicate this effect. Furthermore, in at least one study genetic reduction in superoxide dismutase activity had no effect on life span, even though this manipulation increased oxidative stress as well as tumor burden. Therefore, the effect of manipulating the activity of classic anti-oxidant defenses on life span must be considered an open question. On the other hand, these classic anti-oxidant enzymes are not the sole determinant of oxidative stress, so the best interpretation of the current evidence is that oxidative stress, but not necessarily the classic anti-oxidant enzymes, is a major determinant of age-related pathologies and mortality.

Further Reading

Butterfield, D. A., Drake, J., Pocernich, C., et al. (2001). Evidence of oxidative damage in Alzheimer's disease

brain: central role for amyloid beta-peptide. *Trends in Molecular Medicine* 7(12), 548–554.

Ishii, N., Goto, S. and Hartman, P. S. (2002). Protein oxidation during aging of the nematode *Caenorhabditis elegans*. *Free Radical Biology & Medicine* 33(8), 1021–1025.

Lee, S. S., Lee, R. Y., Fraser, A. G., et al. (2003). A systematic RNAi screen identifies a critical role for mitochondria in *C. elegans* longevity. *Nature Genetics* 33(1), 40–48.

Lin, Y. J., Seroude, L. and Benzer, S. (1998). Extended life-span and stress resistance in the *Drosophila* mutant methuselah. *Science* 282(5390), 943–946.

Lopez-Torres, M., Gredilla, R., Sanz, A., et al. (2002). Influence of aging and long-term caloric restriction on oxygen radical generation and oxidative DNA damage in rat liver mitochondria. *Free Radical Biology & Medical* 32(9), 882–889.

Nicolle, M. M., Gonzalez, J., Sugaya, K., et al. (2001). Signatures of hippocampal oxidative stress in aged spatial learning-impaired rodents. *Neuroscience* 107(3), 415–431.

Perez-Campo, R., Lopez-Torres, M., Cadenas, S., et al. (1998). The rate of free radical production as a determinant of the rate of aging: evidence from the comparative approach. *Journal of Comparative Physiology B* 168(3), 149–158.

Sohal, R. S., Mockett, R. J. and Orr, W. C. (2002). Mechanisms of aging: an appraisal of the oxidative stress hypothesis. *Free Radical Biology & Medicine* 33(5), 575–586.

Wallace, D. C. (2005). A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a dawn for evolutionary medicine. *Annual Review of Genetics* 39, 359–407.

Yuh, K. C. and Gafni, A. (1987). Reversal of age-related effects in rat muscle phosphoglycerate kinase. *Proceedings of the National Academy of Sciences USA* 84(21), 7458–7462.

Oxytocin

G Leng and N Sabatier

University of Edinburgh, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G Leng, volume 3, pp 109–114, © 2000, Elsevier Inc.

Effects of Stress on the Milk-Ejection Reflex
Stress and Parturition
Stress-Induced Stimulation of Oxytocin Secretion
Oxytocin Release within the Brain
Conclusion

Glossary

Dendrites

Parts of a cell where information is received and processed; thick, specialized processes arising from the cell body (soma) of a nerve cell (neuron). Most inputs to a nerve cell from other nerve cells arrive at the dendrites. However, for many cells, the transfer of information at the dendrites is bidirectional; neurotransmitters and neurohormones secreted from the dendrites can affect the activity of neighboring cells and can also affect the excitability of the afferent nerve endings.

<i>Natriuresis</i>	Sodium excretion in urine. To keep the plasma concentration of sodium stable, the intake of salt in the diet has to be balanced with appropriately regulated excretion. Natriuresis is regulated by a variety of neural and hormonal mechanisms.
<i>Posterior pituitary gland</i>	The site at which oxytocin and vasopressin are secreted into the general circulation; also called the neurohypophysis. The gland contains only the axons and neurosecretory nerve terminals of magnocellular vasopressin and oxytocin cells; the cell bodies are within the hypothalamus.

Oxytocin, a peptide hormone, is synthesized by large (magnocellular) neurons in the supraoptic and paraventricular nuclei of the hypothalamus. The axons of these neurons project to the posterior pituitary gland, where oxytocin is stored in large vesicles within nerve endings and axon swellings. Electrical activity in the cell bodies is propagated down the axons in the form of action potentials, resulting in the secretion of oxytocin, which then enters the circulation. Thus, the concentration of oxytocin in the blood reflects the electrical activity of oxytocin neurons in the hypothalamus. There are also some peripheral sources of oxytocin, but these are likely to play local, autocrine roles and do not make a substantial contribution to concentrations in the circulation.

Effects of Stress on the Milk-Ejection Reflex

Oxytocin is essential for the transfer of milk from mother to infant. The sucking of a baby at a mother's breast activates sensory receptors in and around the nipple, leading to the activation of spinal nerves. These nerves project indirectly to the hypothalamus, leading to the activation of oxytocin neurons and secretion of oxytocin. The oxytocin is carried in the systemic circulation to the mammary gland, where it acts on myoepithelial cells, causing milk to be let down from these cells into the collecting duct of the mammary gland, from which the baby can draw the milk into his or her mouth. Without oxytocin secretion, there is no milk let-down and a suckling baby will go hungry.

This reflex – the milk-ejection reflex – is abrupt and intermittent. In response to suckling, oxytocin is secreted in pulses every few minutes. These pulses are the result of brief, intense, coordinated bursts of electrical activity in the oxytocin cells, separated

by periods of relative electrical quiescence. Although the reflex is normally activated by suckling, it can also be conditioned – and the sight, sound, or smell of a baby can, in some circumstances, elicit milk let-down.

This reflex is very susceptible to disruption by stress. In humans, as in many mammals, a quiet, familiar, and safe environment and a calm state of mind are conducive to the reflex. Indeed, in rats the milk-ejection reflex appears to require that cortical activity be in a state of slow-wave sleep. However, although stress might thus be described as inhibiting the oxytocin-mediated milk-ejection reflex, it does not follow that stress inhibits oxytocin secretion. The pulsatile pattern of oxytocin secretion in response to suckling means that the mammary gland is exposed to very high circulating concentrations of oxytocin – but for only a few seconds at a time. The gland responds only to high concentrations of oxytocin, and continuous exposure to high concentrations soon leads to desensitization, so anything that disrupts the appropriate patterning of oxytocin secretion can block the milk-ejection reflex. The reflex requires a coordinated, abrupt activation of the oxytocin cells, and many stimuli that increase the basal activity of oxytocin cells also impair their ability to display coordinated, abrupt episodes of activation. Hence, perversely, many stimuli that lead to increased concentrations of oxytocin in the blood at the same time block the pulsatile pattern of secretion that is essential for the physiological efficacy of oxytocin at the mammary gland.

Stress and Parturition

Oxytocin plays an important role during parturition in most, if not all, mammals. In all mammals studied, the uterus becomes highly sensitive to oxytocin toward the end of pregnancy, as a result of the increased expression of oxytocin receptors in both the myometrium and endometrium. Oxytocin acts on the myometrium of the pregnant uterus to induce contractions directly and on the endometrium to stimulate the production of prostaglandins, which also induce uterine contractions. During delivery, the stimulation of the cervix initiates the Ferguson reflex, during which large amounts of oxytocin are secreted from the pituitary gland into the circulation and uterine contractions and abdominal compression also contribute to the activation of oxytocin secretion.

In the rat, oxytocin cells behave during parturition in much the same way as during suckling – they display intermittent intense bursts of activity leading to pulsatile secretion of hormone. Again, this pattern is important for the physiological efficacy of

oxytocin, although less so than in the case of the milk-ejection reflex; the uterus is sensitive to lower concentrations of oxytocin in the circulation and is less susceptible to desensitization.

Like many mammals, including the rabbit and the sow, the pregnant rat builds a nest in which to deliver; some other mammals, such as the cat, ewe, and bitch, do not build a nest but instead seek out a sheltered place. Mammals that are normally most active at night tend to give birth during the day, and those most active during the day tend to give birth at night; in women, the peak incidence for delivery is between 3:00 and 4:00 a.m. Thus, labor generally begins at a time when the animal is normally asleep or resting and in a place that offers relative seclusion and apparent safety. Disturbance – the sight or sound of an intruder or physical movement from the place where labor has begun to a new environment – can disrupt parturition markedly. If a mouse or a rat is taken from her nest once birth has started and is placed in an empty glass chamber or if a sow is moved from the farrowing nest that she has made to an empty pen, subsequent deliveries are delayed. In bitches, ferrets, rabbits, and ewes, even the noticed presence of a human observer can result in prolonged labor associated with the inhibition of uterine contractions. In women too, anxiety is associated with prolonged labor, and although this has not been the subject of controlled observations, uterine contractions often decrease temporarily when a woman in labor is moved from home to the hospital or from the labor room to the delivery room.

Thus, the progress of parturition in mammals is very susceptible to disruption by acute stress, and many authors have suggested that this disruption might be a consequence of a disruption of oxytocin secretion. The acute administration of oxytocin antagonists prolongs parturition in the rat, and the pharmacological inhibition of oxytocin secretion, particularly by opioid agonists, leads to a prolongation of parturition that can be reversed by the administration of physiological doses of oxytocin. Thus, if stress impairs the activation of oxytocin neurons in parturition, then this would be likely to impair the progress of parturition.

When a rat is moved in mid-parturition from her nesting cage to an empty glass bowl, the oxytocin concentrations in the blood fall, but whether this is a primary consequence of stress or is secondary to reduced uterine activity is not clear. In the rat and in the sow, after the administration of the opioid antagonist naloxone, parturition progresses normally despite environmental disturbance, suggesting that stress might activate endogenous opioid pathways

that inhibit oxytocin secretion; oxytocin neurons are known to be very sensitive to inhibition by exogenous opioids. However, in the rat, naloxone administered during parturition potentiates oxytocin secretion even in the absence of overt stress, so it is possible that the normal progress of parturition in the naloxone-treated disturbed rat merely reflects the ability of hypersecretion of oxytocin to overcome the influence of mild stress.

Thus, mild stress can subvert the physiological actions of oxytocin, preventing milk let-down in response to suckling and prolonging parturition. In lactation, stress disrupts the pattern of oxytocin secretion, preventing the coordinated intermittent burst discharge of oxytocin cells that is necessary for pulsatile oxytocin secretion. During parturition, it is less clear how stress interferes with the progress of labor, but it seems plausible that a similar disruption of the patterning of oxytocin secretion leads to a less effective hormonal signal at the uterus. In the case of the milk-ejection reflex, the disruption of the appropriate patterning of secretion might be accompanied by (and might even be caused by) an increased but uncoordinated activation of the oxytocin cells, giving rise to higher average concentrations of oxytocin in the circulation.

Stress-Induced Stimulation of Oxytocin Secretion

In the rat, on which most studies have been performed, most acute stressors that have been tested markedly increase oxytocin secretion from the pituitary gland. Plasma concentrations of oxytocin are increased, for instance, following forcible restraint, noxious stimuli, cold exposure, forced swimming, aversive foot shocks, administration of nausea-inducing agents, and ether inhalation. At least some of the neural pathways that are involved are also involved in mediating the release of corticotropin releasing hormone (CRH) and hence the secretion of adrenocorticotrophic hormone (ACTH). However, stressors described as emotional only, such as social defeat, do not induce oxytocin secretion. Oxytocin secretion also depends on the species; plasma concentrations of oxytocin do not change in humans exposed to stress or in horses exposed to a novel environment. Stress-induced oxytocin secretion in rats is partly mediated by A2 and A1 noradrenergic cells of the brain stem because the oxytocin response to fear-related stimuli or noxious stimuli, for instance, is impaired after central norepinephrine depletion by neurotoxin treatment or by central injection of an α_1 -adrenergic antagonist. However, not all stressors are mediated by

noradrenergic pathways; for example, the disruption of noradrenergic pathways does not affect oxytocin secretion in rats exposed to a novel environment.

These observations raise several questions. Is the stress-induced secretion of oxytocin into the circulation of any physiological significance? Or is it an epiphenomenon, reflecting a disruption of the normal pattern of activity in oxytocin cells that is essential for reflex milk ejection? The disruption of the milk-ejection reflex is undoubtedly a common consequence of stress. Does this disruption have any adaptive benefits, or does it reflect pathological malfunction?

In fact, much less oxytocin is secreted in response to stress in lactating rats than in virgin rats. This might reflect a general hyporesponsiveness to stress during lactation because ACTH secretion in response to many stresses is also reduced. At least in part, the reduced secretion of oxytocin in lactation reflects depleted stores of oxytocin at the posterior pituitary gland because the amount of oxytocin that is secreted in response to a given activation of hypothalamic neurons is proportional to the amount of oxytocin stored in the neurosecretory nerve endings; in lactation, large amounts of oxytocin are secreted in pulses in response to suckling, and the stores are thus normally 40% or more below the levels in virgin rats. However, the pituitary stores of oxytocin increase through pregnancy, so the reduced oxytocin secretory responses to stress seen in pregnant rats requires a different explanation. The hyporesponsiveness of pregnancy seems to suggest that a specific mechanism is activated to reduce the impact of stressors on oxytocin secretion. This might be part of the mechanisms that allow pituitary stores of oxytocin to accumulate during pregnancy, or it might be part of a generalized hyporesponsiveness of stress-related signaling that is reflected by a depression of hypothalamic-pituitary-adrenal (HPA) axis responsiveness to stressors.

Much effort has been expended attempting to find some physiological rationale for the stress-induced secretion of oxytocin. These efforts have addressed three possibilities:

1. Oxytocin secreted in response to stress acts at other targets in the periphery, where its actions are appropriate and adaptive.
2. Oxytocin secreted in response to stress participates in the regulation of ACTH secretion from the anterior pituitary gland
3. The activation of oxytocin neurons by stress releases oxytocin within the brain as well as from the pituitary gland, and this release within the brain mediates the appropriate adaptive responses to stress.

Possible Peripheral Targets for Oxytocin Released during Stress

From this, it should be clear that neither the uterus nor the mammary gland seems a promising candidate as the physiological target of oxytocin that is secreted during stress. However, receptors for oxytocin are also expressed at other peripheral sites, including the heart, the thymus, and, in the rat, the kidney. Although stress induces oxytocin secretion, it is not clear which peripheral target oxytocin acts on or what its physiological effect is.

At the rat kidney, oxytocin promotes natriuresis; this provides an apparent explanation for the elevated secretion of oxytocin observed in response to increased plasma osmotic pressure. In the rat, oxytocin is also secreted in response to gastric distention, possibly indicating that it mediates a reflex natriuresis following food intake. Similar secretion of oxytocin can be evoked by peripheral injections of cholecystokinin; cholecystokinin is a gut peptide released during feeding and acts on afferent gastric vagus nerve endings. It is difficult to envisage any adaptive value to reflex natriuresis in response to generalized stress, but it seems possible that the activation of this pathway is an incidental consequence of vagal activation. This pathway seems to be a characteristic of the rat; in humans, cholecystokinin increases the secretion of vasopressin but not oxytocin.

In the rat heart, oxytocin receptors are expressed in atrial myocytes, and oxytocin acting at these receptors might participate in the release of atrial natriuretic peptide (ANP). Despite its name, the major physiological role of ANP is probably not to regulate natriuresis, although this is one of its actions; instead, ANP appears to regulate the rate and force of cardiac contractions. The actions of oxytocin reduce blood pressure and reduce cardiac output – not an obviously helpful response to an acute stress, but it might be protective against excessive cardiovascular effects of sympathoadrenal activation. These actions of oxytocin seem more consistent with a role for oxytocin in body-fluid homeostasis through coordinated actions at the heart and kidney.

The thymus is responsible for the selection of the peripheral T-cell repertoire. Oxytocin, like many neuropeptides, has been identified in the human thymus by immunoreactivity and at the transcriptional level; it is present in the thymus in surprisingly large amounts in subtypes of thymic epithelial cells. Oxytocin is co-localized with the cytokeratin network of thymic epithelial cells rather than in secretory granules, so it is not secreted, but behaves like an antigen presented at the cell surface. Thymic oxytocin

also behaves as a cryptocrine signal, interacting with receptors expressed by pre-T cells. Oxytocin receptors are expressed by cytotoxic CD8⁺ lymphocytes, and oxytocin induces phosphorylation of focal adhesion kinase in pre-T cells. Thus oxytocin might intervene in T-cell differentiation as a neuroendocrine self-antigen and as a promoter of T-cell focal adhesion. There is little evidence that circulating oxytocin has important effects at this site, and the possible adaptive value of actions of circulating oxytocin at the thymus is unclear. No convincing case has been made so far that oxytocin secreted into the peripheral circulation in response to stress has any adaptive physiological consequences at the thymus, but there has been very little investigation of this possibility.

Possible Actions of Oxytocin at the Anterior Pituitary Gland

The secretion of oxytocin in response to acute stress has prompted speculation that it might play a role in regulating ACTH secretion from the anterior pituitary gland, and pharmacological studies have indicated that oxytocin can indeed induce ACTH secretion from dispersed rat anterior pituitary cells in synergy with CRH. The ACTH-secreting activity is apparently not mediated by oxytocin receptors because oxytocin receptor mRNA is not expressed in corticotrophs and because specific oxytocin agonists have little effect on ACTH secretion; this effect might be mediated by vasopressin receptors because it can be inhibited by a vasopressin V₁ receptor antagonist. The action of oxytocin on corticotrophs suggested by *in vitro* studies has yet to be confirmed to be physiologically relevant *in vivo* in conditions of stress.

ACTH is secreted in response to the synergistic releasing activities of CRH and vasopressin; these two peptides are co-expressed in parvocellular neurosecretory neurons of the paraventricular nucleus that project to the median eminence, where these peptides are released into the hypothalamic-hypophyseal portal circulation. Oxytocin secreted from the posterior pituitary gland does not appear to have direct access to the anterior pituitary gland because few blood vessels connect the separate lobes of the pituitary gland. However, some collaterals of the axons of magnocellular neurons might terminate in the median eminence, and there are also separate parvocellular populations of oxytocin neurons in the paraventricular nucleus, some of which project to the median eminence. Physiological circumstances associated with elevated secretion of oxytocin into the circulation are not consistently associated with high rates of secretion of ACTH, but the conjunction of high CRH release with high oxytocin secretion might lead to a potentiation of ACTH secretion under some

circumstances because oxytocin at high concentrations is capable of triggering ACTH secretion from corticotrophs and oxytocin appears to act synergistically with CRH. It is not clear whether these actions of oxytocin are physiological rather than purely pharmacological; oxytocin receptors appear to be expressed at relatively low levels in the anterior pituitary gland, but oxytocin at high concentrations can also act at vasopressin receptors and vasopressin receptors on corticotrophs mediate the synergistic ACTH-releasing effects of vasopressin co-secreted with CRH.

A fully convincing case has yet to be made that oxytocin, released either from magnocellular neurosecretory neurons into the circulation or from parvocellular oxytocin neurons projecting to the median eminence, is important for the physiological regulation of ACTH secretion at the level of the pituitary gland, but oxytocin released centrally does appear to be involved in regulating HPA axis activity.

Oxytocin Release within the Brain

Oxytocin does not cross the blood-brain barrier in significant quantities, and oxytocin receptors do not appear to be prominently expressed at brain sites that lack an effective blood-brain barrier. Thus, oxytocin secreted from the posterior pituitary gland into the systemic circulation does not act in the brain. However, oxytocin is present at high concentrations in the cerebrospinal fluid (CSF) and at a number of sites within the brain; this oxytocin does not derive from the blood but from oxytocin released within the brain. A lot of oxytocin is released from the soma and dendrites of magnocellular neurosecretory neurons; this release, within the supraoptic and paraventricular nuclei, generally accompanies the electrical activation of the magnocellular oxytocin cells. However, there are differences in the time course of oxytocin release centrally and peripherally, and different stimuli have differential effects on the central and peripheral release of oxytocin; these observations suggest that there are important differences in the regulation of dendritic oxytocin release and release from nerve terminals. In the rat supraoptic and paraventricular nuclei, forced swimming, shaker stress, and social defeat all induce dendritic release of oxytocin from magnocellular neurons. The role for dendritic oxytocin during stress is unclear, but endogenous oxytocin modulates the activity of supraoptic oxytocin cells and their noradrenergic inputs in response to noxious stimuli.

Many stressors induce the central release of oxytocin, producing very high extracellular concentrations in several regions of the brain including the hypothalamus (especially the supraoptic and paraventricular

nuclei), the septum, and the amygdala. Central oxytocin is thought to be an antistress and anti-anxiety factor; in rats, the central injection of an oxytocin antagonist increases basal plasma concentrations of ACTH and facilitates ACTH secretion in response to various stressors and the central injection of oxytocin attenuates glucocorticoid secretion in response to stress. Rats previously exposed to the central or peripheral administration of oxytocin are less anxious when challenged in an elevated-plus maze, a behavioral test of anxiety. Central oxytocin also attenuates stress-induced CRH mRNA expression in the paraventricular nucleus and attenuates the expression of *c-fos* mRNA in stress-activated neural circuits, including the paraventricular nucleus, the septum, and the hippocampus.

The significance of the presence of oxytocin in the CSF is not clear. Although it has been speculated that the CSF might be a medium for disseminating oxytocin, mediating brain hormone-like actions, it might be that the CSF is simply a sewer, clearing oxytocin from its localized sites of activity.

The oxytocin neurons that project centrally are mainly located in the paraventricular nucleus. There are several functionally and anatomically distinct populations of parvocellular oxytocin cells in this nucleus. One population projects to the brain stem and appears to be implicated in gastric reflexes; these neurons, like magnocellular oxytocin neurons, are activated following systemic injections of cholecystokinin, and they might be part of a reflex loop that regulates gastric motility. Another population projects to the spinal cord and, in males, appears to be involved in penile erection. Oxytocin-containing fibers are also present in many other brain areas, including the preoptic area, the ventromedial nucleus of the hypothalamus, the bed nucleus of the stria terminalis, and the hippocampus. The origin of these diffusely scattered fibers is generally thought to be a separate population or populations of parvocellular oxytocin neurons in the paraventricular nucleus, whose functions are largely unknown.

Experiments involving central injections of oxytocin or of oxytocin antagonists have led to suggestions that oxytocin influences many functions and behaviors. The strongest evidence concerns ingestive behaviors (central oxytocin inhibits food intake and sodium appetite) and reproductive behaviors. Centrally acting oxytocin has been implicated in sexual behavior in both male and female rats, in the initiation of maternal behavior after parturition in the rat, and in pair bonding, notably in the prairie vole. The central release of oxytocin is undoubtedly involved in the milk-ejection reflex: intracerebroventricular injections of 1 ng oxytocin in a lactating rat lead

to a striking increase in the frequency of suckling-induced milk ejections over the following 30 min, whereas intracerebroventricular injections of selective oxytocin antagonists block this reflex. These effects probably reflect the dendritic release of oxytocin from magnocellular neurons that occurs in response to suckling. Intracerebroventricular injections of much larger amounts of oxytocin (1 μ g) have many consequences – including an increased activity of the HPA axis and of the sympathoadrenal system. Such doses greatly exceed the total oxytocin content of the brain and, indeed, are close to the total amount of oxytocin stored in the pituitary gland. It is not clear where in the brain the administered oxytocin acts to produce these effects, that these effects are all mediated by specific oxytocin receptors, or that they mirror any physiological role of endogenous oxytocin. Daily repeated intracerebroventricular injections of high doses of oxytocin produce contrasting effects, including lowered blood pressure, behavioral calm, and lowered activity of the HPA axis.

Suckling itself has a calming influence on the mother and, indeed as mentioned previously, the state of slow-wave sleep observed in rats to be a prerequisite for the milk-ejection reflex might be facilitated by suckling. Whether endogenous oxytocin is involved in this is not clear.

Thus, although centrally projecting oxytocin cells are activated by a variety of stresses, there is no convincing evidence yet that this activation plays any useful or adaptive role or has any important consequences. However, oxytocin is involved in a wide range of behaviors, particularly reproductive behaviors and social behaviors, that are strongly influenced by stressors. Although it has been generally assumed that the source of the oxytocin that is involved in behaviors is the centrally projecting parvocellular oxytocin neurons, there is growing evidence that oxytocin release from the dendrites of magnocellular oxytocin neurons might be important in many behaviors, including behaviors related to stress. In the final section, therefore, we indicate the key experimental evidence that has led to a current focus of attention on dendritic oxytocin release.

Dendritic Oxytocin Release and Priming

The dendrites of magnocellular neurons are large processes, packed with neurosecretory vesicles that contain oxytocin or vasopressin, and they form a dense plexus at the ventral surface of the brain below the supraoptic nucleus. We now know that the dendrites of many different neurons are electrically active, they can conduct action potentials, and they can release many neuroactive substances. Probably only the classical prejudice that dendrites are not

electrically active but are merely passive receivers of information caused us to be slow to recognize the ventral dendritic plexus of the supraoptic nucleus for what we now know it to be – a neurosecretory gland within the brain itself that is the source of very high release levels of oxytocin and vasopressin that can then diffuse to act at distant sites within the brain.

Stimuli that result in oxytocin secretion from the pituitary gland generally also result in dendritic oxytocin release, but in many circumstances dendritic release is very delayed compared to peripheral secretion, although in a few circumstances central release can precede systemic secretion. The regulation of peripheral secretion is now very well characterized – all secretion from this site is thought to be regulated by action potential activity, depolarization of neurosecretory terminals in the pituitary gland result in calcium entry into the terminals via voltage-gated calcium channels, and this calcium entry is the key step in triggering exocytosis of the large neurosecretory vesicles in which oxytocin is packaged. However, secretion from the dendrites is not all activity-dependent in the same way. Some agents, such as the peptide α -melanocyte-stimulating hormone (α -MSH), can trigger dendritic secretion even though they cause no increase at all in the action potential activity of oxytocin cells; on the contrary, α -MSH inhibits the electrical activity of oxytocin cells and thus depresses the secretion of oxytocin into the blood. α -MSH-induced oxytocin release from dendrites appears to be the result of calcium mobilization from intracellular stores rather than the result of calcium entry into the cell through membrane channels, and other peptides, including oxytocin itself, can act similarly.

Interestingly some agents that mobilize intracellular calcium stores also induce an important change in the behavior of oxytocin cells, called priming. In addition to directly causing dendritic oxytocin release, these agents also prime a readily releasable pool of oxytocin, making it available for subsequent activity-dependent release. Thus it appears that the dendrites might contain two functional pools of neurosecretory vesicles, a readily releasable pool that is available for release in response to depolarization-induced calcium entry, as at the neurosecretory terminals, and a much larger pool of vesicles that are not released in response to action potential activity but can be released by agents that mobilize calcium from intracellular stores. Such agents can augment the readily releasable pool, making more oxytocin available for subsequent activity-dependent release, and can thus exert complex and very long-lasting effects on oxytocin cell behavior.

Oxytocin neurons express oxytocin receptors; oxytocin released from dendrites acts back on oxytocin

cells to modulate their activity, including by priming activity-dependent oxytocin release, and this local feedback plays an important part in the patterning of discharge of oxytocin cells. Whether oxytocin has physiological actions on other neurons in the supraoptic or paraventricular nuclei has not been established yet; but it does not act directly on magnocellular vasopressin cells. The dendritic release of oxytocin is very substantial, and spillover of this release into the CSF probably accounts for some of the oxytocin detected there. However, other populations of oxytocin cells in the paraventricular nucleus project centrally, and these other sources probably also contribute to oxytocin concentrations in the CSF.

Conclusion

In the rat, the release of oxytocin both peripherally and centrally is influenced profoundly by stress. Stresses generally increase oxytocin secretion in the rat, but they disrupt the normal pattern of secretion, leading to a hormonal message that is less effective. Thus, the interruption of milk let-down and parturition that accompanies mild stress might be physiologically appropriate; the enhanced secretion of oxytocin *per se* has no clearly established benefit. However, there is currently speculation, based mainly on behavioral effects seen after the central administration of high doses of oxytocin, that the central release of oxytocin has a generally calming influence.

Further Reading

- Amico, J. A., Mantella, R. C., Vollmer, R. R., et al. (2004). Anxiety and stress responses in female oxytocin deficient mice. *Journal of Neuroendocrinology* 16, 319–324.
- Douglas, A. J. (2005). Central noradrenergic mechanisms underlying acute stress responses of the hypothalamo-pituitary-adrenal axis: adaptations through pregnancy and lactation. *Stress* 8, 5–18.
- Geenen, V., Keeha, O. and Martens, H. (1998). Thymic expression of neuroendocrine self-peptide precursors: role in T cell survival and self-tolerance. *Journal of Neuroendocrinology* 10, 811–822.
- Jankowski, M., Hajjar, F., AlKawas, S., et al. (1998). Rat heart: a site of oxytocin production and action. *Proceedings of the National Academy of Sciences USA* 95, 14558–14563.
- Leng, G. and Brown, D. (1997). The origins and significance of pulsatility in hormone secretion from the pituitary. *Journal of Neuroendocrinology* 9, 493–513.
- Ludwig, M. (1998). Dendritic release of vasopressin and oxytocin. *Journal of Neuroendocrinology* 10, 881–895.
- Ludwig, M., Sabatier, N., Bull, P. M., et al. (2002). Intracellular calcium stores regulate activity-dependent neuropeptide release from dendrites. *Nature* 418, 85–89.

- Onaka, T. (2004). Neural pathways controlling central and peripheral oxytocin release during stress. *Journal of Neuroendocrinology* **16**, 308–312.
- Russell, J. A., Leng, G. and Douglas, A. J. (2003). The magnocellular oxytocin system, the fount of maternity: adaptations in pregnancy. *Frontiers in Neuroendocrinology* **24**, 27–61.
- Sabatier, N., Caquineau, C., Bull, P., et al. (2003). Alpha-melanocyte-stimulating hormone stimulates oxytocin release from the dendrites of hypothalamic neurons while inhibiting oxytocin release from their terminals in the neurohypophysis. *Journal of Neuroscience* **23**, 10351–10358.
- Stricker, E. M. and Verbalis, J. G. (1996). Central inhibition of salt appetite by oxytocin in rats. *Regulatory Peptides* **66**, 83–85.
- Uvnas-Moberg, K. (1998). Antistress pattern induced by oxytocin. *News in Physiological Sciences* **13**, 22–26.

P

Pain

H J Strausbaugh and J D Levine

University of California, San Francisco, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 115–118, © 2000, Elsevier Inc.

Introduction

Stress-Induced Analgesia

Stress-Induced Hyperalgesia

Glossary

<i>Analgesia</i>	A decreased perception of pain.
<i>Hyperalgesia</i>	An increased perception of pain.
<i>Nociception</i>	A perception of pain.
<i>Opioid</i>	Any of three families of endogenous peptides (i.e., as endorphin, enkephalin, and dynorphin) that induce analgesia by binding to μ , δ , or κ opioid receptors.

Introduction

Definition and Measurement of Pain

Pain has been defined as an aversive sensation originating from a noxious stimulus in a defined area of the body. Such intense, potentially tissue-injurious stimuli activate nociceptive sensory neurons (nociceptors), which transmit signals via the spinal cord to the brain that are perceived by the organism as pain. The central nociceptive processing areas include the thalamus, hypothalamus, mesencephalic tegmental area, mesolimbic areas, parabrachial nuclei, amygdala, bulbopontine reticular formation, and the somatosensory cortex. Pain has two components: a sensory or discriminate component and an affective or emotional component. Because pain involves both of these components, simply recording from peripheral nociceptors does not provide a sufficient measurement of pain because the affective component of pain is not

assessed in this manner. To circumvent this problem, behavioral measures of pain are used. If animals display characteristic behaviors after application of a noxious stimulus, including withdrawal of the area to which the stimulus is applied, licking of the affected area, vocalization, or escape behavior, it is inferred that the animal is experiencing pain. A commonly used behavioral test for the measurement of pain in animals and the test upon which much of the current information on stress effects on pain is based is the tail flick test. In this test, an animal is lightly restrained and its tail is exposed to a heat source intense enough to produce pain in humans. The time it takes for the animal to remove its tail from the heat source is recorded as tail flick latency; the more painful the stimulus, the shorter the latency. An additional tool used to identify pain is the administration of opioid analgesics such as morphine. Morphine activates an endogenous pain control circuit to block the transmission of nociceptive signals. Therefore, if a noxious stimulus-induced behavior is blocked by morphine, it is considered to be painful.

Definition and Measurement of Stress

Stress is generally defined as any stimulus to which the organism is not adapted. Like pain, stress involves a perceptual component and is therefore difficult to assess in laboratory animals. To circumvent this problem, hormonal measurements are commonly used. Stressful stimuli activate three neuroendocrine circuits: the hypothalamic-pituitary-adrenocortical (HPA) axis, the sympathoadrenal axis, and the sympathetic neural axis. Although different stressors may activate these axes to varying degrees, all axes are activated by most stressful stimuli. Therefore, the hormonal measurement of activation of one of these axes, generally the HPA axis, is used as an operational definition of stress. Both adrenocorticotrophic hormone (ACTH) and corticosterone are secreted rapidly into the blood upon activation of the HPA axis. Stimuli that evoke increases beyond normal circadian

levels in either of these hormones are considered stressful.

How do these broad, somewhat overlapping concepts, pain and stress, relate to each other? One of the first written observations of the relationship between pain and stress comes from observations made by field doctors in Europe during World War II. These doctors observed that up to 70% of soldiers presenting with severe battle wounds would not report significant levels of pain. This apparent insensitivity to pain was short-lived; 24 h later, all soldiers reported significant levels of pain. Reports such as these suggested that the perception of pain could be modified cognitively. Because stress seemed to be a common factor in all of these reports, the question arose as to whether stress could affect pain perception. In 1976, a decreased perception of pain after exposure to a stressor was shown in laboratory experiments for the first time and was termed stress-induced analgesia. Although less well studied, it has also been shown that stress can enhance pain perception in some circumstances. This phenomenon has been termed stress hyperalgesia and is generally referred to as stress-induced hyperalgesia.

Stress-Induced Analgesia

Stress-induced analgesia refers to the well-established phenomenon that exposing humans or animals to a wide variety of stressors induces the suppression of pain perception. Although electric foot shock and cold water swim have been the most well-studied eliciting stressors, others include rotation, immobilization, and food deprivation. The decrease in pain sensitivity in these studies is usually measured as an increase in tail flick latency to noxious thermal stimuli. However, other pain sensitivity tests such as paw withdrawal, paw licking, and escape behavior in response to exposure to electrified grids or hot plates are also used.

Although the mechanisms of stress-induced analgesia are not completely understood, two subtypes of stress-induced analgesia – endogenous opioid dependent (opioid mediated) and opioid independent (nonopioid mediated) – can be distinguished. Opioid-mediated stress-induced analgesia is defined as stress-induced analgesia that either is blocked by the opiate receptor antagonist naloxone or is cross-tolerant to opioid agonists such as morphine. Cross-tolerance is determined by treating an animal for several days with morphine until the drug no longer produces analgesia. A stressful stimulus is then applied, and if the analgesia is blocked, then it is considered cross-tolerant to morphine. The nonopioid-mediated type of stress-induced

analgesia is unaffected by these treatments. Characteristics of the stressor, including the temporal pattern of administration and severity, seem to predict whether opioid or nonopioid analgesia will be induced. In general, stressors that are intermittent and less severe produce opioid-mediated analgesia, whereas stressors that are continuous and of greater severity produce nonopioid-mediated analgesia. Both forms can be classically conditioned.

The distinct nature of opioid- and nonopioid-mediated subtypes of stress-induced analgesia is supported by the fact that they are not cross-tolerant. In fact, a collateral inhibition model has been proposed by Bodnar and colleagues in which activation of one form of stress-induced analgesia would actively inhibit the other form. These investigators have shown that, if administered closely in time, non-opioid-mediated analgesia inhibits opioid-mediated analgesia. This is in contrast to the well-established phenomenon that distinct elicitors of opioid-mediated stress-induced analgesia synergize to produce greater analgesia. Bodnar and colleagues propose that collateral inhibition allows the animal to use its most adaptive system in response to potentially widely varying stressful situations presented by the environment.

The mechanism of opioid-mediated stress-induced analgesia is thought to involve both central and peripheral sources of opioids as well as both central and peripheral sites of action. Stress activates endogenous opioidergic neural circuits in the midbrain, which then transmit signals through the dorsolateral funiculus to inhibit pain transmission circuits in the dorsal horn of the spinal cord. Stress also induces the release of opioids from peripheral sources, including the pituitary gland and the adrenal medulla, and, during inflammation, from leukocytes. These opioids may act by binding to receptors on terminals of nociceptive peripheral sensory nerves and inhibiting noxious stimulation-induced activation or sensitization of these neurons. It is important to note that while these opioidergic pathways are a key element in mediating this type of analgesia, other neurotransmitters (e.g., serotonin, norepinephrine, acetylcholine), neuropeptides (e.g., cholecystokinin), and hormones (e.g., corticosterone) may also participate in this pathway.

Although it is clear that distinct pathways mediate opioid and nonopioid forms of stress-induced analgesia, there is some evidence that there may also be some pathways in common. Similar to opioid-mediated forms of stress-induced analgesia, both midbrain structures and the dorsolateral funiculus of the spinal cord are involved in mediating nonopioid stress-induced forms of analgesia. Additionally, mouse strains have been developed that express high or low levels of

stress-induced analgesia. This phenotype is present whether the animals are exposed to stressors that characteristically induce opioid- or nonopioid-mediated analgesia. However, in the absence of specific gene products, this is not strong evidence for an overlap in the mechanism but does offer a promising avenue for future studies. In contrast to opioid-mediated stress-induced analgesia, vasopressin, thyrotropin-releasing hormone, and histaminergic systems are involved in mediating nonopioid-mediated stress-induced analgesia. Some forms of nonopioid-mediated stress-induced analgesia (e.g., continuous cold water swim) are mediated by the HPA axis and, in particular, by the adrenocortical system, whereas other forms (e.g., continuous foot shock of short duration) are not. Additionally, dopaminergic and cholinergic systems may play a role in mediating continuous cold water swim analgesia. Clearly, further research is required to elucidate the pathways that mediate nonopioid-mediated stress-induced analgesia.

Stress-Induced Hyperalgesia

Stress-induced hyperalgesia is defined as an increased sensitivity to noxious stimuli or a decrease in analgesia after exposure to stress. In the laboratory, an increased sensitivity to noxious stimuli is usually indicated by a decreased tail flick latency to noxious heat stimulation. Stress-induced hyperalgesia has been observed in animals after exposure to a variety of stressors, including ether, mechanical oscillation, vibration, a novel environment, and repeated exposure to cold temperatures. Stress-induced hyperalgesia has been much less well studied than stress-induced analgesia, and consequently the mechanisms of stress-induced hyperalgesia are still poorly understood. However, a general concept is emerging that an induction of stress-induced hyperalgesia is due to changes in the central, possibly perceptual, processing of nociceptive information by the organism and not due to changes in the processing of noxious stimuli at the level of the peripheral nociceptive pathways as may contribute to stress-induced analgesia. This emerging concept is based on several lines of evidence. First, electrophysiological recordings from nociceptive sensory neurons in the peripheral nerves indicate no change in sensitivity to the noxious stimulus before and after stress-induced hyperalgesia is elicited. Second, stress-induced hyperalgesia can be classically conditioned such that even when a noxious stimulus is not applied, a neutral stimulus that had been paired previously with the noxious stimulus can elicit hyperalgesia. Third, emotional characteristics of the rat can determine whether

stress-induced hyperalgesia is induced. Although it is difficult to assess the emotional state of a rat, studies have shown that rats displaying behaviors thought to indicate anxiety (e.g., vocalization, agitation) during exposure to stress and the noxious stimulus display stress-induced hyperalgesia, whereas rats not displaying these behaviors and exposed to identical stimuli do not show stress-induced hyperalgesia. This concept is supported by the fact that administering antianxiety drugs (e.g., clonidine, diazepam) to these rats blocks stress-induced hyperalgesia, whereas administering anxiety-producing drugs (e.g., yohimbine) enhances and prolongs stress-induced hyperalgesia. Finally, clinical studies indicate that anxious patients display an increased sensitivity to pain. Of course, in these studies it is difficult to assess whether patients are more sensitive to pain because they are anxious or whether they are anxious because they are perceiving more pain. Further investigation will be required to elucidate the factors that elicit and the mechanisms that produce stress-induced hyperalgesia. Additionally, it has been shown that stress may inhibit the analgesic effects of drugs as well as enhance pain perception; however, the relationship between these two effects is currently unknown. Further investigation will be required to distinguish these two phenomena.

See Also the Following Articles

Hypothalamic-Pituitary-Adrenal; Opioids.

Further Reading

- Amit, Z. and Galina, Z. H. (1986). Stress-induced analgesia: adaptive pain suppression. *Physiological Reviews* **66**, 1091–1120.
- Bardiani, A. and Pavone, F. (1991). Reduction of oxotremorine-induced analgesia after chronic but not acute restraint stress. *Psychopharmacology (Berlin)* **104**, 57–61.
- Bodnar, R. J. (1990). Effects of opioid peptides on peripheral stimulation and “stress”-induced analgesia in animals. *Critical Reviews in Neurobiology* **6**, 39–49.
- Foa, E. B., Zinbarg, R. and Rothbaum, B. O. (1992). Uncontrollability and unpredictability in post-traumatic stress disorder: an initial model. *Psychological Bulletin* **112**, 218–238.
- Gamaro, G. D., Xavier, M. H., Denardin, J. D., et al. (1998). The effects of acute and repeated restraint stress on the nociceptive response in rats. *Physiology and Behavior* **63**, 693–697.
- Gue, M., Del Rio-Lacheze, C., Eutamene, H., Theodorou, V., Fioramonti, J. and Bueno, L. (1997). Stress-induced visceral hypersensitivity to rectal distension in rats: role

- of CRF and mast cells. *Neurogastroenterology and Motility* 9, 271–279.
- Kelly, D. D. (ed.) (1986). *Stress-induced analgesia*. Annals of the New York Academy of Sciences (Vol. 467). New York: New York Academy of Sciences.
- King, T. E., Crown, E. D., Sieve, A. N., Joynes, R. L., Grau, J. W. and Meagher, M. W. (1999). Shock-induced hyperalgesia: evidence forebrain systems play an essential role. *Behavioural Brain Research* 100, 33–42.
- Madden, J., Akil, H., Patrick, R. L. and Barchas, J. D. (1977). Stress-induced parallel changes in central opioid levels and pain responsiveness in the rat. *Nature* 265, 358–360.
- Mogil, J. S., Sternberg, W. F., Marek, P., Sadowski, B., Belknap, J. K. and Liebeskind, J. C. (1996). The genetics of pain and pain inhibition. *Proceedings of the National Academy of Sciences USA* 93, 3048–3055.
- Murison, R. and Overmier, J. B. (1993). Parallelism among stress effects on ulcer, immunosuppression and analgesia: commonality of mechanisms? *Journal of Physiology (Paris)* 87, 253–259.
- Nijenhuis, E. R., Vanderlinden, J. and Spinhoven, P. (1998). Animal defensive reactions as a model for trauma-induced dissociative reactions. *Journal of Trauma and Stress* 11, 243–260.
- Simone, D. A. (1992). Neural mechanisms of hyperalgesia. *Current Opinion in Neurobiology* 2, 479–483.
- Vidal, C. and Jacob, J. J. (1982). Stress hyperalgesia in rats: an experimental animal model of anxiogenic hyperalgesia in human. *Life Sciences* 31, 1241–1244.
- Watkins, L. R. and Mayer, D. J. (1982). Organization of endogenous opiate and nonopiate pain control systems. *Science* 216, 1185–1192.

Panic Disorder and Agoraphobia

J C Ballenger

Medical University of South Carolina, Charleston, SC, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J C Ballenger, volume 3, pp 119–123, © 2000, Elsevier Inc.

Symptoms

Diagnosis

PD in the General Medical Setting

Epidemiology

Risk Factors

Course and Prognosis

Treatment

Psychological Treatments:

Psychodynamic Psychotherapy

Behavioral Treatments

Selection of Treatments

Glossary

<i>Panic attacks</i> (PA)	The name used currently for unexpected attacks of fear.
<i>Panic disorder</i> (PD)	Originates from the Greek god of flocks, Pan, who would frighten animals and humans out of the blue.
<i>Spontaneous panic attacks</i>	The main characteristic of PD that is central to its recognition and diagnosis.

Symptoms

One of the main characteristics of panic disorder (PD) is spontaneous out of the blue panic attacks (PAs). PAs usually occur immediately upon exposure or in anticipation of exposure to particular situations, often ones in which panic attacks have previously occurred.

Symptoms of PAs in order of their frequency are palpitations, pounding heart, tachycardia, sweating, trembling or shaking, shortness of breath or smothering, feeling of choking, chest pain or discomfort, nausea or abdominal distress, feeling dizzy, unsteady, lightheaded or faint, derealization or depersonalization, fear of losing control or going crazy, fear of dying, chills, and hot flashes.

Panic attacks usually last for several minutes, but they can also last for hours. In one of the largest modern PD clinical trials, patients averaged one to two PAs per week. However, the frequency and severity of PAs vary greatly between individuals and, at times, in individuals. Some may have only one to three PAs per year, whereas others may have multiple PAs on a daily basis. Some individuals will even have bursts of PAs and then not experience attacks for a long period of time.

Because PAs are extremely frightening, patients will often develop a logical fear that they will reoccur. This anticipatory anxiety increases prior to exposure to situations in which patients have experienced previous PAs. Because of this, PD patients often develop

agoraphobia, which is avoidance of certain places or situations. Agoraphobics often avoid situations in which they have had PAs because they are afraid that escape would be difficult or embarrassing or that help might not be available during these attacks. In community samples, one-third to one-half of patients who meet criteria for PD also have significant agoraphobia avoidance.

There are typical situations that are avoided most frequently. These include taking buses, trains, or planes; riding in or driving a car; being in large crowds; standing in lines; shopping; and being in spaces where they might feel trapped (e.g., bridges, tunnels, elevators). Patients experiencing a PA will often have an overwhelming need to escape or return to a place of safety, like home. These patients do not actually fear the situation itself but are afraid of the feelings of panic that might happen while in that situation.

Many patients even have PAs that begin during sleep. These nocturnal panic attacks are quite common, and the majority of PD patients experience them. These PAs occur during the beginning states of the sleep cycle and are essentially identical to daytime PAs.

There also is a close relationship between depression and PD. In various samples, comorbid or secondary depression ranged from 22 to 68%. Proper recognition of comorbid depression is especially important because of the marked increase in suicide attempts (23.6–50%) in patients with both PD and depression.

Diagnosis

The diagnosis of PD has several requirements.

1. Recurrent, unexpected PAs (situational PAs could also occur, but there would need to be at least two unexpected PAs by history).
2. Panic attacks need to be followed by at least 1 month of (1) persistent anxiety about the potential recurrence of further PAs, (2) persistent anxiety about the implications of these attacks (e.g., going crazy, something wrong medically), or (3) a significant behavioral change because of these attacks.

PD in the General Medical Setting

There is poor recognition of PD in the general medical setting. It is often undiagnosed, because physical symptoms of PD can be mistakenly linked to other medical conditions. However, panic-like symptoms can occur in a few medical conditions (e.g., hyperthyroidism, pheochromocytoma, seizures, cardiac arrhythmias, and chronic obstructive pulmonary disease).

Conservative estimates of PD in primary care have ranged from 3 to 8%, with at least 50% of patients with PD being unrecognized. The average length of time for a patient in the general medicine setting to be diagnosed correctly with PD is 10 years, with an escalation of the use of health-care services over this 10-year period. Overall, PD patients are three times more likely to utilize general medical services than the traditional patient in a primary care setting. The percentage of PD patients is also markedly higher in certain medical groups (e.g., cardiology and gastroenterology).

A study sponsored by the World Health Organization (WHO) studied primary care patients in 14 different countries. Ninety-nine percent of patients with one PA in the previous month had an anxiety disorder or depression or a subthreshold anxiety disorder or depressive disorder. If replicated, it would appear that the occurrence of even a single PA should serve as a signal for increased scrutiny for anxiety and depressive syndromes.

Epidemiology

There is a striking uniformity worldwide for the observed prevalence of PD, generally averaging between 1 and 2% in multiple studies. In the WHO Primary Care Survey of 14 countries mentioned previously, the prevalence for PD ranged from a low of 1.4% to a high of 16.5% for PAs and 0–03.5% for PD itself. The average for PD was 1.1% (current) and 3.5% (lifetime), which was similar to community samples. As mentioned, rates are much higher in certain medical settings, ranging from 15% in dizziness clinics to 16–65% in cardiology practices and 35% in hyperventilation clinics.

Risk Factors

In treatment samples, PD has been observed to be at least two times more prevalent in females than in males. Males tend to have a longer duration of illness but are less likely to seek help or develop agoraphobia and depression. The onset of PD for both genders usually occurs between 15 and 24 years of age, with a second peak at 45–54 years of age. The onset of PD after the age of 65 is rare (0.1%).

As mentioned previously, agoraphobia is also more prevalent in females, with the highest rates of PD and agoraphobia occurring in widowed, divorced, or separated individuals living in cities. Early parental loss, limited education, and physical or sexual abuse are also risk factors for PD.

Genetic factors can even be associated with PD. Heightened anxiety, known as behavioral inhibition,

is higher in children of anxiety-disordered parents and in children of adults with PD. As these children have aged, they have been observed to have higher rates of anxiety and phobic disorders.

Precipitating events have been reported in 60–96% of PD cases. Multiple studies also suggest that traumatic early events may figure into the vulnerability leading to PD.

Course and Prognosis

Available evidence suggests that PD is a chronic condition, once criteria for the disorder are met. It is common for patients to report symptoms that they have had for 10–15 years prior to diagnosis. As mentioned previously, the Klein model suggests that spontaneous PAs are the first manifestation of the illness, followed by anticipatory anxiety and then agoraphobia in some individuals.

In fact, most PD patients do not have PAs as their first symptoms. Over 90% have had mild phobic or hypochondriac symptoms prior to the onset of their first PA. The first PA is usually experienced in a phobogenic situation, such as a public place, street, public transportation, crowd, elevator, tunnel, bridge, or open space. Despite this, the Klein model remains largely correct and helpful in conceptualizing the development of the condition and its treatments.

After diagnosis and some sort of treatment, functional recovery occurs in the majority of patients. Poor responses to treatment were associated most consistently with initial high symptom severity and high agoraphobic avoidance at baseline. A poor response is also associated with a low socioeconomic status, less education, longer duration, limited social networks, death of a parent, divorce or unmarried status, and personality disorders.

Treatment

Since the early 1960s, effective psychological and pharmacological treatments for PD have been developed.

Tricyclic Antidepressants (TCAs)

In the 1960s, the TCA imipramine was one of the first effective drug treatments. Important advantages of imipramine include its once-a-day dose, its low cost, and the extensive research that has been conducted on its effectiveness. Its principal difficulties are initial hyperstimulation reactions, significant weight gain with long-term use, and the dangers involved with an overdose. The highly serotonergic clomipramine became the most popular treatment in Europe during the 1970s, 1980s, and 1990s.

Monoamine Oxidase Inhibitors (MAOIs)

During the 1960s, MAOIs were also proven to be effective in the treatment of PD. The most definitive trial was an early comparison of phenelzine and imipramine in the mid-1970s. This was the first trial to suggest that MAOIs might be the most effective medication class, especially if the PD patient is depressed. The biggest drawbacks of MAOIs are the need for a patient to be on a restricted diet and the danger of hypertensive crisis if the diet is not followed. There are also dangers if certain other medications are used, especially certain analgesics and other antidepressants. The principal side effects are insomnia, weight gain, postural hypotension, anticholinergic side effects, and sexual dysfunction.

The reversible MAOIs, brofaramine and meclobemide, have stimulated interest because they require no restrictive diet. However, trials with meclobemide were unable to distinguish medication from placebo, and despite promising early results with brofaramine, its development has been discontinued. There is still active research in this area.

Benzodiazepines (BZs)

High potency BZs (e.g., alprazolam, clonazepam) were developed in the 1980s and are also used in the treatment of PD. In fact, BZs remain the most frequently utilized treatment for PD worldwide, probably because of their quick onset of action, increased efficacy against anticipatory anxiety, low cost, and tolerability. They are frequently utilized in conjunction with the antidepressants. The principal drawback to BZs is the general concern with their abuse potential, although it is alcoholics and opiate or pill addicts that abuse the drug. Abuse in uncomplicated PD patients is extremely rare. Other side effects include the high rate of withdrawal symptoms and relapse that can occur after discontinuing BZ treatment.

Selective Serotonin Reuptake Inhibitors (SSRIs)

More recent studies in the 1990s have demonstrated that the SSRIs are effective, and they are now frequently recommended as the first treatment of choice. This is largely based on the many controlled trials documenting the high tolerability of SSRIs and the decreased rate of jitteriness and hyperstimulation responses often associated with tricyclic antidepressants. They are preferred over benzodiazepines in patients with significant difficulty with depression. Most (70–85%) patients retain their benefit when SSRIs are tapered and discontinued after 12 months of successful treatment.

Psychological Treatments:

Psychodynamic Psychotherapy

Although psychodynamic psychotherapy remains a popular treatment for many psychiatric disorders, including PD, there is limited research demonstrating its efficacy in PD. An emotion-focused treatment for PD has been developed that explores typical fears of being abandoned or trapped as stimuli for panic attacks. This often involves a 12-session acute treatment with 6 sessions of monthly maintenance. Patients are encouraged to identify, reflect upon, and attempt to change problematic feelings and their responses.

Behavioral Treatments

There is a rapidly increasing body of evidence for combinations of exposure-based treatments and cognitive-behavioral therapy (CBT) that are effective in the treatment of PD.

Exposure Treatments

Behavioral treatments, involving *in vivo* exposure, require patients to gradually systematically re-expose themselves to their phobic situations. A large number of studies have documented that 58–83% of patients improve after exposure treatments.

One consistent observation of exposure-based treatments is its long-lasting effects. In one trial with 110 individuals treated with 12 weeks of exposure-based treatment, 74% of the sample achieved panic-free status, and this improvement was maintained in almost all patients throughout a 7-year follow-up. Studies that have compared the use of exposure to CBT have found them to be essentially equal in efficacy. There appears to be little benefit from the combination, although the dropout rate was lower.

Evidence supporting the importance of exposure-based treatments is very clear, but many patients remain symptomatic. In addition, many studies report that between 10 and 25% of patients drop out of the therapy because it is an anxiety-provoking treatment method.

Cognitive-Behavioral Therapy

Cognitive-behavioral therapy is based on the theory that patients with PD interpret physical symptoms in a catastrophic way and that these cognitive distortions need to be challenged and corrected. CBT of PD evolved from early work of Aaron Beck but has been applied to PD primarily by Barlow and colleagues and is called panic control therapy (PCT).

Treatment generally involves an initial education component, which is followed by the identification of critical misinterpretations of panic symptoms. Patients are then taught ways in which they can challenge and correct these misinterpretations. This includes evaluating how likely (or unlikely) the imagined consequences are and how catastrophic it would be even if their worse fears did happen. Most CBT is also combined with interoceptive exposure to the physical symptoms that frighten them. For instance, a patient fearful of dizziness might be spun in a chair.

Although applied relaxation has been shown to be somewhat effective in PD, other studies have demonstrated that CBT, especially combined with interoceptive exposure (i.e., PCT), is more effective than medication.

Most experts would agree that the treatment of PD with pharmacotherapy and a CBT approach – employing a combination of exposure and cognitive therapy – are roughly comparable in efficacy. The most carefully performed trial demonstrating this was a multicenter, 11-week acute trial comparing imipramine to CBT. Both CBT and imipramine performed equally well, but the combination of CBT and imipramine was not proven to be any more effective in the treatment of PD. This trial provided evidence that CBT applied early in treatment did not offer protective effects against imipramine relapse. However, other data suggest that it is protective if given during the discontinuation period.

Selection of Treatments

Pharmacotherapy (BZs or antidepressants) and CBT/exposure treatments appear to be equally effective, although response is perhaps faster with medication treatment and longer term response is better in some patients with CBT/exposure-based treatments. Most often, selection of treatment is determined by patient choice or individual characteristics.

One important issue in the treatment of PD is that the majority of patients do not fully recover with either pharmacotherapy or psychotherapy. For both types of treatment, only 25–45% of patients could be considered fully recovered, despite the excellent progress that has been made since the 1980s. This has led to ongoing research attempts to find more effective treatment methods for PD.

See Also the Following Articles

Anxiety; Cognitive Behavioral Therapy; Pharmacological Treatments of Stress.

Further Reading

- Ballenger, J. C. (1998). Treatment of panic disorder in the general medical setting. *Journal of Psychosomatic Research* 44, 5–15.
- Ballenger, J. C. (1998). Comorbidity of panic and depression: Implications for clinical management. *International Journal of Clinical Psychopharmacology* 13(Supplement), 513–517.
- Ballenger, J. C. (2003). Selective serotonin reuptake inhibitors in the treatment of the anxiety disorders. In: Nutt, D. & Ballenger, J. C. (eds.) *Anxiety disorders*, pp. 339–361. Oxford, UK: Blackwell Science.
- Ballenger, J. C., Burrows, G. and DuPont, R. L., Jr., et al. (1988). Alprazolam in panic disorder and agoraphobia. Results from a multicenter trial. I. Efficacy in short-term treatment. *Archives of General Psychiatry* 45, 413–422.
- Ballenger, J. C., Lydiard, R. B. and Turner, S. M. (1997). Panic disorder and agoraphobia. In: Gabbard, G. O. (ed.) *Treatments of psychiatric disorders* (2nd edn., vol. 2), pp. 1421–1452. Washington, D.C.: American Psychiatric Press.
- Ballenger, J. C., Davidson, J. R. T., Lecrubier, Y. and Nutt, D. J. (International Consensus Group on Depression and Anxiety), et al. (1998). Consensus statement on panic disorder from the International Consensus Group on Depression and Anxiety. *Journal of Clinical Psychiatry* 59, 47–54.
- Klein, D. F. (1964). Delineation of two drug responsive anxiety syndromes. *Psychopharmacology* 5, 397–408.
- Klerman, G. L., Weissman, M., Ovellette, R., Johnson, J. and Greenwald, S. (1991). Panic attacks in the community: social morbidity and health care utilization. *Journal of the American Medical Association* 265, 742–746.
- Margraf, G., Barlow, D. H., Clark, D. M. and Telch, M. J. (1993). Psychological treatment of panic: work in progress in outcome, active ingredients, and follow-up. *Behavioral Research Therapy* 31, 1–8.
- Marks, I. M. (1969). *Fears and phobias*. London: Heinemann.
- Mavissakalian, M. R. and Prien, R. F. (eds.) (1996). *Long-term treatments of anxiety disorders*. Washington, D.C.: American Psychiatric Press.
- Rosenbaum, J. F., Biederman, J., Gersten, M., et al. (1988). Behavioral inhibition in children of parents with panic disorder and agoraphobia: a controlled study. *Archives of General Psychiatry* 45, 463–470.
- Roy-Byrne, P. P. and Cowley, D. S. (1995). Course and outcome in panic disorder. A review of recent follow-up studies. *Anxiety* 1, 151–160.
- Sartorius, N., Ustun, T. B., Jorges-Alberto, C., et al. (1993). An international study of psychological problems in primary care. *Archives of General Psychiatry* 50, 819–824.

Paranoia

P Kinderman

University of Liverpool, Liverpool, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Kinderman, volume 3, pp 124–129,

© 2000, Elsevier Inc.

Paranoia

An unreasonable belief held by a person that other people are victimizing him or her; an unreasonable fear of threat.

Paranoid personality disorder

A disorder characterized by a pervasive distrust and suspiciousness of others, such that their motives are interpreted as malevolent.

Definition and Description

Significance and Social Impact of Paranoia

Assessment

Theories of Paranoia

Treatment of Paranoia

Glossary

Causal attribution

An explanation for the occurrence of events.

Delusion

A false personal belief sustained in spite of apparently incontrovertible and obvious proof or evidence to the contrary.

Definition and Description

The meaning of paranoia and paranoid has fluctuated over the centuries. The literal translation of the Greek paranoid is “by the side of the mind” and was used in a general manner to mean crazy or mad. The use of the term paranoid subsided until its resurgence in Germany in the first half of the nineteenth century to refer to a distinct delusional disorder. In the English-speaking world, a distinct confusion of use emerged, with the United Kingdom tending to mirror continental Europe and using the word paranoia to describe systematized delusional insanity. In the United States, the acceptance of Freud’s 1911 psychodynamic model

of paranoia led to the widespread use of the term and to a diminution of any distinctions between delusional and nondelusional paranoia.

The links among paranoia, persecutory delusions, and paranoid schizophrenia remain close but confused. In ordinary language, however, the word paranoid has become general, applied to people who either have unreasonable or unwarranted complaints of victimization unreasonable fears of threat.

At its extreme, such concerns may lead to the label of paranoid personality disorder, defined as a pervasive distrust and suspiciousness of others such that their motives are interpreted as malevolent. People with this disorder are described as being suspicious, being cynical, and tending to bear grudges and take criticism extremely badly. Paranoid personality disorder appears to be relatively common in the general population and has been placed on a continuum in which paranoid schizophrenia is seen as a more severe form of paranoia, which is in turn seen as a more severe form of paranoid personality disorder.

Significance and Social Impact of Paranoia

The terms paranoid or paranoia are used repeatedly in political and social contexts. Paranoia and persecutory delusions are also commonly associated with violence. Paranoia has been described as a characteristic of a vast variety of political systems. The societies in Plato's *Republic* and Hobbes's *Leviathan* have been described as characteristically paranoid as have a great number of political systems from Tudor England between the fifteenth and seventeenth centuries, the United Kingdom in the eighteenth to nineteenth centuries, and both the U.S. and Soviet Cold War political systems. Political leaders as individuals (particularly Hitler, Stalin, and Nixon) have also been described as paranoid.

Paranoid delusions have been associated with many violent acts. The M'Naughten rules, the legal basis for an insanity defense, were established after Daniel M'Naughten attempted to kill British Prime Minister Sir Robert Peel in 1843. M'Naughten mistook the prime minister's aide, Edward Drummond, for Sir Robert and shot him instead. M'Naughten was found to be clearly and delusionally paranoid. According to an analysis of the so-called White House cases (individuals who were hospitalized because of their psychotic preoccupation with U.S. politicians), paranoid delusions were common, but the best predictor of arrest for a violent crime after hospital discharge was an arrest for a violent crime before hospitalization. This finding has been generally supported; the likelihood of acting on delusions is generally unrelated to

phenomenology of the delusion but is predicted by the perceived evidential support for the belief, negative emotions, and social functioning.

Assessment

Paranoid ideation can be assessed in its own right as delusional beliefs or as a symptom of schizophrenia.

Delusional classification is complex, and it is often difficult to determine the difference between unusual beliefs and full delusions, between brief episodes and persistent delusions, or between bizarre versus non-bizarre delusions. Thus, continua rather than dichotomous classifications may be appropriate. Delusions are also usually assessed on a variety of parameters: conviction, cultural or stimulus determinants, preoccupation, implausibility, extension (the degree to which the delusional belief involves various areas of the patient's life), bizarreness, disorganization, and emotional commitment.

General psychiatric rating scales include brief assessments of the presence or absence of the symptoms of schizophrenia, including delusional beliefs. A more precise assessment can be obtained using specific research devices. Within such scales, however, delusional ideation is viewed as a part of a whole syndrome rather than as a phenomenon worth examining alone. More specific rating scales and assessment devices have been developed for the assessment of delusions alone.

Despite doubts about the validity and reliability of scales for the assessment of paranoia, the fact that they seem incestuous in their development, and their reliance on doubtful theories of personality structure, diagnosed patients with delusions of persecution score high on these measures, whereas nonpatient samples show predictable associations with other measures of cognitive functioning. However, paranoid ideation is complex. It may be the case that measures applicable for normal and clinical populations are different.

Theories of Paranoia

Theories of paranoia are occasionally subsumed under theories of the development of delusions or schizophrenia. The presumed causes of schizophrenia are not dealt with here; note, however, that nearly all variables believed to influence human behavior have been invoked. Specific theories concerning the etiology of delusions are equally numerous.

Biological Approaches

Medical and psychiatric conditions Paranoid delusions occur in association with at least 105 general

medical conditions. In many nonpsychiatric conditions associated with delusions, delirium (a state of disordered or impaired consciousness combined with an alteration of perceptual functioning or affect) seems important. In such medical conditions, paranoid ideas are presumed to be secondary to the physical pathology. Although it seems difficult to understand the etiology of the abnormal beliefs, it does seem that when the underlying illness is treated the delusions usually disappear. Paranoid delusions are far more common in psychiatric conditions, in which they are often difficult to treat and tend to persist.

Delusions occur in 10–20% of depressed inpatients, with mood-congruent delusions being more common. Delusions occur in up to 70% of manic patients and are usually grandiose, concerning wealth, importance, status, or special powers. As in depression, separating mood-congruent, exaggerated, and abnormal (but nondelusional) beliefs from true delusions in mania is difficult and perhaps of little clinical value.

The presence of delusions is highly likely to lead to a diagnosis of paranoid schizophrenia. Delusions of persecution have been reported in over 60% of patients with a diagnosis of paranoid schizophrenia. Such delusions in paranoid schizophrenia are usually systematized and persistent, as opposed to the fleeting and haphazard beliefs of people with organic diseases.

Genetics Genetic studies of paranoid schizophrenia seem to imply that inherited factors are important but not directly causal. Genetic influences on delusional ideation have received little specific study. When links between genetic factors and delusional beliefs have been made, these usually emerge as post hoc findings from research protocols designed to study paranoid schizophrenia.

Dopamine Dopamine is no longer viewed as a likely contender for the status of single schizophrenogenic factor. However, dopamine-releasing illicit drugs are thought to trigger delusions, especially delusions of persecution.

There is some evidence of dopaminergic abnormalities specific to delusional ideation. However, biological researchers seem to have unwittingly hampered their own investigations in a potentially important area by studying broad constructs such as paranoid schizophrenia. It may be the case, for instance, that disturbance of mental health in general is too broad a concept to be related to specific neurotransmitter mechanisms even if individual psychotic phenomena, such as hallucinations and delusions, may be. Dopamine functioning remains a possible avenue for the investigation of paranoia.

Psychological Approaches

In 1983, Winters and Neale distinguished between motivational and deficit psychological theories of delusional thinking. Deficit theories generally assume that cognitive or attentional abnormalities lead to delusions, whereas motivational theories, in general, assume that individuals develop delusions to explain either unusual perceptual experiences or aversive psychological states.

Psychological deficit theories

Deficits (or excesses) in emotional expression or processing A number of theorists have suggested that delusions stem from excessive or exaggerated emotion. There is some evidence of disruption (or facilitation) of emotional processing in delusions, particularly paranoia. Generally, research examining cognitive processes in persecutory delusions has examined the processing of information related to emotions and the self-concept.

Deficits in logical thinking Because delusions are illogical, many authors have suggested that delusional beliefs occur as a result of deficits in an individual's ability to think logically. In the presence of schizophrenic thought disorder, normal associative links may be disrupted. In the absence of coherent logical associations, affective influences may become dominant and hence delusional ideas may be generated. Although originally advocated by influential authors, there remain reservations about this very general model.

There have been a number of more specific models of delusions involving deficits in logical processes. In delusions, errors or biases in inductive reasoning have been investigated by a number of researchers. These studies have often been cited as evidence of rapid, overconfident, less evidence-based patterns of inductive reasoning in deluded patients.

Perceptual deficits and excesses It has been suggested that paranoid delusions are the products of healthy rational attempts to explain anomalous or disturbed perceptions. There is some evidence in favor of this model, although some investigations have been somewhat unconventional. It has also been suggested that hallucinations may lead to rational delusional beliefs. It might seem quite rational to develop delusional explanations for inexplicable perceptual experiences. There is some clinical evidence of a relationship between deafness and delusions of persecution in elderly patients, for whom the misperception of the conversations of others may serve as fuel for paranoia. Despite this, and the logic of this

approach, there is little experimental support for such a perceptual deficit model in younger patients. One group of perceptual excesses known to lead to delusions is drug-induced hallucinations. However, the delusions seen in drug use seem qualitatively different from functional psychoses.

Theory-of-mind deficits In 1994, Frith suggested that a deficit or abnormality in what is termed the Theory of Mind may explain a number of schizophrenic symptoms. We do not, as a rule, appraise other people's actions and conversation as simple behavioral or verbal patterns. Rather, according to the Theory of Mind, we assume that we can use behavior to divine wishes, hopes, beliefs, intentions, and so forth. Such an understanding is believed to be vital for successful social intercourse. In Frith's model, paranoid delusions are seen as stemming from a lack of understanding of the actions of other people. Although support for this model comes mainly from phenomenological discussions, a number of recent empirical studies have been reported to support it.

Attentional deficits In 1994, Hemsley suggested that delusions stem from a process involving abnormalities in selective attention in which attention is focused on trivialities. This leads to a lack of influence from learned regularities. Clinically, abnormalities in selective attention seem common. However, certain details of the causal, psychological, linking processes seem vague and poorly related to clinical details of delusional beliefs.

Psychological motivational theories

Learning theory The development of learning theory as a major psychological discipline has, of course, led to behavioral theories of delusions. One general approach is to assume that psychotic phenomena are maintained by positive reinforcement. Indeed, behavioral approaches to the treatment of delusions have frequently proved successful, although this tells us little about the origins of delusional beliefs.

Another behavioral approach to delusions is to assume that delusions are reinforcing through avoidance. Patients with schizophrenia may well be sensitive to anxiety and stress, and therefore that avoidance through delusional ideation may be welcome. Overall, however, these models are unsatisfying. The fact that psychopathology gets worse in conditions of stress is not a good test of any model. Such models seem to invoke, but not address, cognitive variables.

Psychoanalytical theories

The paranoid pseudocommunity Some psychoanalytic theorists have suggested that delusions result

from weaknesses in an individual's ego boundaries. If a person's distinction between self and world and between self and fantasy is weak, the distinction between reality and falsehood is assumed to be similarly weak. In this context, it has been suggested that a person frustrated by a lack of social skills seeks refuge in fantasy and daydreams. However, these primitive fantasies, born from frustration, also threaten the ego. Because the defenses (fantasy) are immature, these threats are denied and repressed. Because the individual is not in contact with a normal social network, paranoia develops in a context of imagined and delusional others – the paranoid pseudocommunity.

Although this model appears overly complex, with many assumptions, many researchers have found it persuasive. Empirical data, however, undermine the central thesis that paranoia is born from a lack of social skills. In fact, people diagnosed as having schizophrenia have a higher premorbid social adjustment if they have paranoid symptoms.

Freud More commonly, psychoanalytic theories have invoked defense mechanisms, particularly projection. Freud (in 1911) and many others suggested that delusions represent the externalization of desires, fears, or conflicts. Sigmund Freud's theory of the causes of paranoia has been extremely influential. In essence, it suggests that paranoia develops as a defensive support for a denial of homosexual love. Broad acceptance of this model seems to have shaped American psychiatry's approach to paranoia.

Freud proposed a psychoanalytic theory of persecutory delusions. Paranoid delusions, Freud claimed, are parts of a process protecting the conscious ego from awareness of conflict with unacceptable homosexual impulses stemming from the id. Such homosexual urges are, Freud argued, denied or contradicted and then countered by the defenses of rationalization and projection. A male patient is essentially confronted with the idea that "I (a man) love him (a man)." This is unacceptable, leading to reaction formation and the idea "I do not love him. I hate him." Such inhuman hatred is still unacceptable, and is rationalized as "I hate him because he hates and persecutes me."

Freud's account is colorful, but has been widely challenged. Freud's model suggests that the delusional persecutor is a person for whom the patient expresses secret homosexual love. However, paranoid people frequently do not know their persecutors or even have a specific idea of who they might be. Moreover, many patients fear persecution by members of the opposite sex, who are unlikely to be fueled by repressed homosexual feelings.

It is notable that Freud's model of paranoia includes two elements: a defensive projection or

externalization of threatening material and latent homosexuality. It is possible that the defensive component is more valid than the homosexual part. Psychoanalytic writers after Freud suggested that persecutory delusions serve a defensive function, without necessarily stressing latent homosexuality.

In an more recent variant of this model, Colby suggested that paranoia stems from a tendency of a person to perceive or generate threats to his or her self-esteem combined with a protective mechanism of projection and externalization of the threat to others. In an interesting twist, Colby described a computer simulation model of his theory. A test was conducted comparing the output of his computer model with a real paranoid patient. Because clinical judges were reported to be unable to distinguish the two, the computer model was considered a success. One of the main benefits of Colby's theory is that it is testable. People with delusions of persecution should readily perceive potential threats to self-esteem, and they should also locate the source of such threats as external. Many more recent cognitive investigations of paranoia are compatible with Colby's model.

Camouflaged depression Zigler and Glick suggested that paranoia is a form of camouflaged depression. This suggestion was originally formulated as a statement regarding classification, with paranoia being properly seen as a form of affective disorder. However, the claim is also a causal one. Taking as sources the psychological theories of paranoia previously briefly reviewed, Zigler and Glick claimed that paranoia consists of low self-esteem or feelings of inadequacy (the depression part) covered with defensive projections of these feelings onto others.

Two types of paranoia A subdivision of paranoia has recently been proposed into the rather pejoratively labeled poor-me and bad-me paranoias. In the first form, paranoia is seen as stemming from a low self-esteem, so low that persecution is expected. In the second, paranoia is seen as stemming from defensive processes similar to those suggested by Kinderman and Bentan. Despite weak evidence supporting the subdivision, this categorization has the benefit of being based on theoretical distinctions between different cognitive processes believed to be etiologically important.

Attributional defenses against threats to the self-concept In the attributional model of paranoia, proposed by Kinderman and Bentan, events that might activate an implicit negative self-representations are hypothesized to evoke externalizing (self-protective yet paranoid) causal attributions: "I am not responsible for the bad things that are happening to me, other people are." This reduces the accessibility of self-actual:

self-ideal discrepancies, but it leads to the belief that others have negative views of the self. Such an explanation is consistent with psychodynamic accounts of paranoia and with the suggestion that paranoia represents a form of camouflaged depression. This model may have relevance to "normal" psychology beyond the clinical arena.

Normal emotional processing More recently, psychologists Garety and Freeman have suggested that paranoid thought emerges from normal, rather than excessive or exaggerated, emotional processes. These commentators point out that relatively normal emotional problems (anxiety and depression) are common in paranoia and contend that it is possible that these processes may lead to delusional beliefs.

Although it is an attractive argument, this approach fails to account for why some people with these problems remain simply anxious and depressed, whereas other people become deluded.

Treatment of Paranoia

A range of drug treatments are commonly used in the treatment of psychosis, including paranoia. The most important are antipsychotic drugs (sometimes called neuroleptics), which are now regarded as the preferred medical treatment for psychotic conditions. Antipsychotic drugs may be used in the acute phase of psychosis or prophylactically and have been shown to result in acute symptom management and the prevention of relapse. However, neuroleptics are not perfect treatments. Some patients fail to respond or show only a partial response, and neuroleptic treatment can lead to serious and distressing side effects.

Modern developments in clinical psychology mean that schizophrenia may now be seen as amenable to psychological treatment. In particular, delusional ideas can be addressed successfully using the techniques of reality testing and cognitive restructuring. Relatively straightforward cognitive therapy strategies may be beneficial to deluded patients. A common feature of the cognitive-behavioral therapy (CBT) strategies that have been advocated for use with deluded patients is the avoidance of direct confrontation of the patient.

Kinderman and Bentall described one final therapeutic approach, directly driven by theory. In this approach, the causal attributions made by a severely deluded, paranoid patient are addressed through the use of cognitive therapeutic techniques. External personal attributions for minor negative events (blaming other people for bad events) are thought to be characteristic of paranoia. The use of the techniques of CBT may address this tendency.

It has been suggested, with respect to the teleological and religious quality of many delusions, that a

patient's recovery is heralded by the acceptance of the accidental in life. If the situational-external locus of causal attribution is meaningful, it is essentially an acceptance of the accidental. Such an acceptance may be important in many areas of life. Perhaps paranoid individuals need to learn that, sometimes, strange things just happen. An acceptance of the accidental in life may be psychologically important.

See Also the Following Article

Schizophrenia.

Further Reading

- Bentall, R. P., Corcoran, R., Howard, R., et al. (2001). Persecutory delusions: a review and theoretical integration. *Clinical Psychology Review* 21, 1143–1192.
- Cole, J., Kleberman, G. L. and Goldberg, S. C. (1964). Phenothiazine treatment of acute schizophrenia. *Archives of General Psychiatry* 10, 246–261.
- Freud, S. (1953–1974). Psycho-analytic notes upon an autobiographical account of a case of paranoia (dementia

- paranoides) (1911). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vol. 12), London: Hogarth Press and the Institute of Psychoanalysis.
- Frith, C. (1994). Theory of mind in schizophrenia. In: David, A. S. & Cutting, J. C. (eds.) *The neuropsychology of schizophrenia*, pp. 147–161. Hove, UK: Lawrence Erlbaum.
- Garety, P. A. and Freeman, D. (1999). Cognitive approaches to delusions: a critical review of theories and evidence. *British Journal of Clinical Psychology* 38, 113–154.
- Garety, P. A., Hemsley, D. R. and Wessely, S. (1991). Reasoning in deluded schizophrenic and paranoid patients. *Journal of Nervous and Mental Disease* 179, 194–201.
- Kinderman, P. and Bentall, R. P. (1999). The clinical implications of a psychological model of paranoia. In: Sanavio, E. (ed.) *Behavior and cognitive therapy today: essays in honour of Hans J. Eysenck*, pp. 131–162. Oxford: Elsevier Press.
- Winters, K. C. and Neale, J. M. (1983). Delusions and delusional thinking: a review of the literature. *Clinical Psychology Review* 3, 227–253.

Paraventricular Nucleus

L W Swanson

Neuroscience Research Institute, University of Southern California, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by L W Swanson, volume 3, pp 130–133, © 2000, Elsevier Inc.

Location and Basic Divisions

Output Control

CRH Neurons

Stressor-Specific Patterns of Gene Expression

Regulation of Homeostasis and Ingestive Behaviors

Glossary

- Corticotrope** Cell type in the anterior pituitary that synthesizes and secretes adrenocorticotrophic hormone (ACTH), which in turn stimulates glucocorticoid secretion from the adrenal cortex.
- Corticotropin-releasing hormone (CRH)** Peptide hormone that increases the secretion of adrenocorticotrophic hormone (ACTH) from corticotropes in the anterior pituitary.

Magnocellular neurosecretory system

Large neurons in and around the hypothalamus that send their axons directly to the posterior pituitary, where they release the peptide hormones vasopressin and oxytocin into the general circulation.

Median eminence

Functionally, the initial part of the stalk of the pituitary gland along the base of the hypothalamus; its inner layer contains the axons of magnocellular neurosecretory neurons and its outer layer contains the axon terminals of parvocellular neurosecretory neurons.

Paracrine

Chemical messengers that act locally in a general way; in contrast to hormonal messengers that act via the bloodstream and to synaptic messengers that act as specialized intracellular junctions.

Parvocellular neurosecretory system

Small neurons in and around the hypothalamus that send their axons to the outer layer of the median eminence, where they are taken up and delivered to the anterior pituitary by a portal system.

Portal system

A system of veins with a capillary network at either end; used to deliver substances from one capillary bed (as in the median eminence) to another (as in the anterior pituitary).

The paraventricular nucleus of the hypothalamus (PVH) is the most important part of the brain for controlling the stress response because its neurons control the release from the anterior pituitary gland of adrenocorticotrophic hormone (ACTH), which in turn controls glucocorticoid release from the adrenal cortex. In addition to constituting the final common pathway for all centrally mediated stress responses, the PVH (1) controls the release of other hormones from the anterior and posterior parts of the pituitary gland, (2) influences the size of autonomic reflexes mediated by circuits in the brain stem and spinal cord, and (3) modulates eating and drinking behaviors. Furthermore, steroid hormone and neural inputs can dramatically alter the expression of neuropeptide and many other genes in the PVH, with different stressors producing different patterns of altered expression.

Location and Basic Divisions

In all mammals, the PVH lies adjacent to the third ventricle, in the rostral half of the hypothalamus, rostral to the infundibulum or stalk of the pituitary gland. The PVH, together with the supraoptic nucleus, was originally known as the origin of the magnocellular neurosecretory system. That is, under most conditions separate groups of neurosecretory neurons in each of these two nuclei synthesize significant amounts of either vasopressin or oxytocin and release them into the general circulation at the level of the posterior pituitary gland. These peptide hormones (which are nine amino acids long) play especially important roles in controlling diuresis and blood pressure (vasopressin), and in parturition, lactation, and associated social behaviors (oxytocin).

In the 1970s, it became clear that the PVH also has two other divisions. First, it contains neurons that project to the median eminence and that are thus involved in controlling hormone release from the anterior pituitary (the parvocellular neurosecretory system) and, second, it contains neurons that send descending projections to the brain stem and spinal cord – especially to regions involved in autonomic reflexes, the relay of nociceptive information, and somatic motor behavioral responses. Neuroanatomical studies with multiple retrograde tracers demonstrate that there are three essentially separate cell groups in the PVH of the rat: the magnocellular neurosecretory, parvocellular neurosecretory, and descending divisions.

The parvocellular neurosecretory division appears to contain three major cell types. Most importantly regarding stress, it contains virtually all of the neurons that synthesize corticotropin-releasing hormone (CRH) and release it into the pituitary portal system

to increase ACTH secretion from the anterior pituitary. Second, the PVH contains virtually all of the neurons that synthesize thyrotropin-releasing hormone (TRH) and release it into the pituitary portal system to increase thyroid-stimulating hormone secretion from the anterior pituitary. And third, the periventricular zone of the PVH contains a large fraction of the neurons that release somatostatin into the pituitary portal system to inhibit growth hormone secretion from the anterior pituitary.

Neurons in the descending division of the PVH innervate many parts of the brain stem and spinal cord. Best understood, however, are inputs to preganglionic neurons of both the parasympathetic and sympathetic parts of the autonomic nervous system to visceral sensory regions including the nucleus of the solitary tract and parabrachial nucleus, to the periaqueductal gray, and to various parts of the reticular formation. As discussed in the following, these descending projections appear to coordinate autonomic and behavioral responses with neuroendocrine responses elicited from the PVH.

Output Control

As the final common pathway for all types of stress response mediated by the central nervous system, it stands to reason that the PVH receives pathways for all of the relevant stimuli. Thus, it should come as no surprise that the afferent control of the PVH is exceedingly complex. In fact, on the order of 50 different sources of axonal inputs to the PVH have been identified with modern neuroanatomical pathway tracing methods, and in many instances associated neurotransmitters have been identified as well. Each neural input appears to end in relation to more than one cell type in the PVH, and a majority of the neural inputs that have been identified to date end in the region of CRH neurons that control ACTH release.

Neural inputs to parvocellular neuroendocrine CRH neurons fall into a number of functional categories. First, there are inputs that relay sensory information. The best-known inputs of this type arise in three hindbrain regions – the nucleus of the solitary tract, ventrolateral medulla, and parabrachial nucleus – and are known to relay viscerosensitive, and probably nociceptive, information to CRH neurons. Second, there is a direct input from a part of the basal ganglia (of the cerebral hemispheres) known as the bed nuclei of the stria terminalis. The bed nuclei in turn receive inputs from other parts of the cerebral hemispheres – including the ventral hippocampus, amygdala, and prefrontal cortex – that may exert cognitive influences on the PVH. Third, there are direct inputs from certain behavioral state control

nuclei, such as the suprachiasmatic nucleus (which generates the circadian rhythm) and the midbrain raphe nuclei (which may be involved in certain aspects of behavioral state or arousal). And fourth, parvocellular neurosecretory CRH neurons in the PVH receive inputs from many cell groups in the hypothalamus that have complex, as yet poorly understood functions related to homeostasis and behavior.

In addition to neural inputs, the PVH appears to be influenced by humoral inputs, that is, by feedback influences of steroid and other hormones. For example, specific sets of neurons in the PVH express glucocorticoid, mineralocorticoid, estrogen, and thyroid hormone receptors. These hormone receptors may change the plasma membrane electrical properties of certain PVH neurons, but it is becoming increasingly clear that they may also change – either directly or indirectly – the pattern of gene expression in PVH neurons in profound ways.

CRH Neurons

CRH neurons that control ACTH secretion are quite remarkable from a functional point of view. In the first place, they have the traditional role of releasing a factor or hormone (CRH) into the proximal capillary plexus of the hypothalamo-hypophysial portal system – at the level of the external layer of the median eminence. The axon of these neurons initially extends laterally from the cell body in the PVH and then arches ventrally through the region of the fornix before curving medially around the ventrolateral border of the hypothalamic ventromedial nucleus to enter the median eminence, where it ends in a rich proliferation of terminal boutons. Peptides (such as CRH) released from these terminals are of course transported via the portal system to act as hormones in the anterior pituitary (for example, on corticotropes).

In addition, however, there is evidence that released substances may also act in a paracrine way within the median eminence itself. For example, it would appear that CRH can act presynaptically to inhibit the release of gonadotropin-releasing hormone (GnRH) at the level of the median eminence. This may account in part for the inhibition of gonadal steroid hormone release observed after some types of stress.

Furthermore, the results of lucifer yellow dye cell-filling experiments indicate that the axon of parvocellular neurosecretory neurons generates terminal boutons of passage in the hypothalamus, on its way to the median eminence. This raises the distinct possibility that these neurons also form classical chemical synapses within the hypothalamus itself. Thus, the evidence suggests that the axon of CRH parvocellular

neurosecretory neurons in the PVH can have spatially segregated hormonal, paracrine, and synaptic effects.

In addition to the fact that ACTH-controlling CRH neurons receive more than 50 sources of neural inputs, and that their axon can exert hormonal, paracrine, and synaptic effects, it is now clear that they can also express a bewildering complement of neurotransmitter and/or neuromodulator genes – on the order of 10 to 20. In addition to CRH, they include vasopressin and angiotensin, all of which act synergistically to secrete ACTH, as well as enkephalin, neurotensin, and cholecystokinin, which do not appear to act on corticotropes but may instead interact with receptors at the level of the median eminence and/or hypothalamus.

Stressor-Specific Patterns of Gene Expression

Shortly after CRH was characterized and antibodies to it were raised by Vale and his colleagues, it became clear from multiple labeling immunohistochemical studies that the ratio of putative neurotransmitters in individual PVH neurons can vary dramatically, depending on the endocrine status and stress history of the animal. This was first observed in adrenalectomized rats, in which high levels of previously undetected vasopressin were found in CRH neurons, indicating that vasopressin mRNA levels are much more sensitive to the negative feedback effects of glucocorticoids than CRH mRNA levels.

Since then, a vast amount of experimental evidence has suggested that many forms of stress alter patterns of neuropeptide gene expression in the PVH, as well as in numerous other parts of the brain, and that each form of stress may be accompanied by a specific and distinct pattern of altered gene expression in the brain. For any particular gene, the occurrence, magnitude, and time course of altered expression depend on the type, magnitude, and time course of a particular stressor or pattern of stressors. Thus, it is clear that stressors do not produce synchronous changes in brain gene expression.

Interestingly, a number of neurotransmitter or neurotransmitter-related genes are expressed across cell types in the PVH. For example, under appropriate conditions, CRH, vasopressin, angiotensin, dynorphin, and tyrosine hydroxylase (the enzyme that synthesizes dopamine) are expressed in parvocellular neurosecretory, magnocellular neurosecretory, and descending neurons of the PVH, and a particular stressor or condition may influence gene expression differentially in different cell types. In the rat, corticosterone (the principal endogenous glucocorticoid) decreases CRH and vasopressin mRNAs in parvocellular neurosecretory

neurons, increases CRH mRNA but has no effect on vasopressin mRNA in magnocellular neurosecretory neurons, and increases CRH mRNA but has no effect on vasopressin mRNA in descending neurons.

Regulation of Homeostasis and Ingestive Behaviors

A wide range of experimental work has shown that the PVH is part of the circuit or system that controls eating and drinking behaviors, that is, ingestive behaviors, and that descending projections from the PVH (probably to the periaqueductal gray or nearby areas) are involved in these responses. This involvement in behaviors that play a critical role in maintaining homeostasis for metabolic fuels and for body water is perhaps not surprising in view of the neuroendocrine and autonomic projections of the nucleus. Regarding metabolic control, recall that virtually all parvocellular neurosecretory CRH and TRH neurons (controlling adrenal glucocorticoid and thyroid hormone secretion, respectively) are found in the PVH, as are a majority of the somatostatin neurons that inhibit growth hormone release. And regarding body water regulation, recall that the PVH contains a major population of magnocellular neurosecretory vasopressin (antidiuretic hormone) neurons. In addition, the PVH establishes major bidirectional connections with brain stem and spinal cord circuits that mediate both sympathetic and parasympathetic reflexes, some of which are undoubtedly involved in metabolic and body water regulation.

In summary, the PVH consists of three functionally distinct divisions with essentially motor functions: parvocellular neurosecretory, magnocellular neurosecretory, and descending, with inputs to autonomic and somatomotor control systems. As such, the PVH is in an ideal position to integrate neuroendocrine, autonomic, and behavioral responses to particular stressors, which is a hallmark of hypothalamic function in general. Because the PVH contains few if any interneurons, and recurrent axon collaterals have

been difficult to identify unequivocally, this functional integration appears to be mediated by neural inputs, virtually all of which innervate more than one functional division, and perhaps by humoral inputs as well. In addition to containing the CRH neuroendocrine motoneurons that control ACTH release, the PVH is critical for the integrated behavioral, autonomic, and neuroendocrine responses that assure a relatively constant supply of nutrients and body water.

See Also the Following Articles

Homeostasis; Hypothalamic-Pituitary-Adrenal; Pituitary Regulation, Role of.

Further Reading

- Herman, J. P., Ostrander, M. M., Mueller, N. K. and Figueiredo, H. (2005). Limbic system mechanisms of stress regulation: hypothalamo-pituitary-adrenocortical axis. *Progress in Neuropsychopharmacology and Biological Psychiatry* **29**, 1201–1213.
- Leibowitz, S. F. (1992). Neurochemical-neuroendocrine systems in the brain controlling macronutrient intake and metabolism. *Trends in Neuroscience* **15**, 491–497.
- Reyes, T. M., Walker, J. R., Decino, C., Hegenesch, J. B. and Sawchenko, P. E. (2003). Categorically distinct acute stressors elicit dissimilar transcriptional profiles in the paraventricular nucleus of the hypothalamus. *Journal of Neuroscience* **23**, 5607–5616.
- Swanson, L. W. (1991). Biochemical switching in hypothalamic circuits mediating responses to stress. *Progress in Brain Research* **87**, 181–200.
- Swanson, L. W. and Sawchenko, P. E. (1980). Paraventricular nucleus: a site for the integration of neuroendocrine and autonomic mechanisms. *Progress in Neuroendocrinology* **31**, 410–417.
- Watts, A. G. (2005). Glucocorticoid regulation of peptide genes in neuroendocrine CRH neurons: a complexity beyond negative feedback. *Neuroendocrinology* **26**, 109–130.
- Young, L. J., Murphy Young, A. Z. and Hammock, E. A. (2005). Anatomy and neurochemistry of the pair bond. *Journal of Computational Neurology* **493**, 51–57.

Parenting, Stress of

K D Jennings

University of Pittsburgh, Pittsburgh, PA, USA

LJ Dietz

University of Pittsburgh Medical Center, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

Why Parenting Can Be Stressful

Types of Parenting Stress

Effects of Stress on Parents

Effects of Parenting Stress on Children

Interventions for Parents and Children

Glossary

<i>Child temperament</i>	An enduring pattern of biological and physiological responses to the environment.
<i>Coping</i>	Psychological processes that allow an individual to regulate negative emotions and behavior to adapt to stressful situations.
<i>Daily hassles</i>	Normal or typical life experiences that increase negative affect and interpersonal distress by virtue of their high demands and low satisfaction.
<i>Major life events</i>	Exceptional but time-limited life circumstances that produce stress in an individual by acutely increasing negative affect, interfering with daily functioning, and creating change in relationships and routines.
<i>Parenting stress</i>	Pattern of adverse psychological and physiological reactions resulting from the demands of being a parent.
<i>Self-efficacy</i>	An individual's perception of his or her competence in completing the tasks that define a particular domain.
<i>Social support</i>	Interpersonal relationships that meet the emotional and instrumental needs of individuals experiencing distress.

Raising children is often a rewarding and positive experience, but it is also a challenging task that is inevitably stressful at times. Parenting requires a high level of effort and emotional involvement that can be overwhelming. It also involves a wide range of skills that must be used in a flexible manner. Parenting involves meeting the basic needs of children, socializing them to culturally appropriate rules of behavior, and promoting their self-esteem and emotional well-being. In the past 25 years, increasing attention has

been paid to the stress of raising children and the effects of parenting stress on caregivers and families. Parenting stress has been defined as a pattern of adverse psychological and physiological reactions resulting from attempts to adapt to the demands of being a parent.

Why Parenting Can Be Stressful

Parenting Tasks

On a daily basis, parents must meet children's survival needs by providing them with shelter, food, clothing, and other necessities. They must ensure their physical safety by carefully monitoring their activities in different environments and by ensuring that proper medical attention is received. Parents must also meet their children's emotional needs by providing attention and affection in order to promote a healthy self concept and socioemotional development. Parents are also responsible for socializing children. They teach them how to adapt to their sociocultural environment by monitoring their behavior and reinforcing culturally appropriate responses. Parents also act as the primary models for regulating negative emotions and coping with distress in culturally specific ways. Parents actively help younger children cope with anger and disappointment by emotional scaffolding and redirecting attention. With older children, parents assist in problem solving and anticipating consequences prior to responding to a situation. Parents provide negative consequences to dissuade inappropriate anger responses such as aggression and prolonged displays of negative affect; in contrast, they praise and reward behaviors that are appropriate or polite. Finally, parents are responsible for introducing children to the larger community outside the family by providing opportunities for peer interaction, formal learning in school, and group activities. In all, parents are expected to carry out a wide variety of tasks that enable their children to successfully adapt to the demands of their sociocultural group and ultimately to become contributing members of their group.

Parenting Must Change as Children Develop

As children grow and develop, their needs and reliance on parents continually change. Being a parent is thus not a singular transition into a static role but rather a set of transitions within a dynamic role that is continually evolving to meet the changing needs of children. Furthermore, the role of being a parent

requires synchrony between parents and children, and this synchrony falls largely on parents to achieve, especially when children are younger. Thus, children's developmental stages influence the amount of stress experienced by parents. For example, caring for a newborn is a time of high stress, particularly for the first-born child. In addition to the change in responsibilities and life style that occurs with the first child, newborns demand round-the-clock intensive care-taking. The preschool age has also been identified as stressful because parents must provide constant care and supervision while negotiating limits with their children as they begin to establish their independence. Other developmental stages provide their own set of challenges: entry into school, with exposure to and judgment by the larger community; adolescence, with increased demands for independence and a renegotiation of the parental role; and finally launching, facilitating the children's assumption of adult roles. Each developmental stage demands different characteristics of the parent; thus, parents differ on which stages they find easier or more difficult.

Parenting Must be Adapted to Children's Temperament and Special Needs

In addition to development, parents must adapt their interactions and interventions according to their child's individual temperament. There is not a single style of parenting that is ideal for every child. Instead, for example, parents must adapt their style to be more soothing for children with difficult temperaments who have problems regulating negative affect and more firm for children with temperaments that are high in novelty seeking. Children with more extreme temperaments or temperaments that differ from their parents can be experienced as more stressful. In addition, disabilities, developmental delays, chronic illnesses, or other special needs of children influence the level of stress experienced by parents. Such children typically require more physical care and more supervision; they also require that parents adjust to the loss of their perfect child and learn to cope with the medical and educational systems. Similarly, children with behavioral or emotional problems cause more stress for their parents.

Balancing Parenting with Other Roles

Parenting stress can also result from difficulty in balancing parenting responsibilities with other adult roles, particularly the roles of spouse and worker. For example, the transition to parenthood has been associated repeatedly with decreased marital satisfaction. The overwhelming, new responsibilities that come from the arrival of a newborn, combined with

fatigue and the intense demands of caring for an infant, decrease the time and attention available for the marital relationship, which in turn may decrease marital satisfaction. Over time, parents must learn to balance the needs of all in the family, including the needs of their spouse and their own needs. Parents' difficulty in finding time and energy to meet their own needs is a major contributor to parenting stress.

Balancing work with parenting is another source of stress. Parents must make provisions for routine child care and also accommodate nonroutine events such as children's illnesses and school snow days. The rising demographic trend in first-world nations for dual-wage-earner families indicates that a high number of new mothers will continue to work outside of the home after a child is born. Thus, many families will be faced with the difficulties of balancing multiple roles.

Circumstances of Child Rearing

Specific circumstances of childrearing can exacerbate the level of stress experienced by parents. These circumstances include poverty, number of children, amount of social support, marital status, and physical or mental illness in parents. These stressors impede parents' efforts to care for their children. Thus, for example, single parents raising a number of children in poverty are likely to experience more stress than parents raising children in more favorable circumstances.

Types of Parenting Stress

Major Life Events

Early research on parenting stress focused on the consequences of major life events that increased emotional distress in parents and complicated normal parenting tasks. These major life events, such as death of a spouse, loss of employment, or divorce, were viewed as atypical of the everyday experience of parenting and were thought to increase parental distress and dysfunction and to impair effectiveness of meeting children's needs. Likewise, parenting stress was examined in relation to parenting children with a developmental disability, chronic physical illness, or an emotional disorder. This research did indeed indicate a link between major life events and various indices of parent, child, and family functioning. Parents who were coping with out-of-the-ordinary difficult circumstances reported reduced emotional well-being and increased psychopathology. They were also observed to have less optimal parenting skills. In turn, their children showed more psychopathology and lower social competence. However, a

shortcoming of the major life events model is that it focuses on low-frequency events that cluster in high-risk populations with many problems. The life events model fails to capture the minor stress incurred by all parents in meeting the daily demands of childrearing.

Chronic Minor Stressors

Current research on stress in adults focuses on both major life events and the more mundane chronic stress stemming from meeting the demands of daily living. For parents, minor stressors are the daily hassles of meeting children's needs, including monitoring children's behavior, physically caring for children (i.e., feeding and cleaning), and balancing conflicting schedules of work and family life. Thus, the construct of minor stressors, or parenting hassles, refers to commonly occurring events in all families. Crnic and others argue that the chronic nature of these daily hassles of parenting may make them more likely to adversely affect family functioning than are major life events that occur infrequently. Furthermore, they argue that minor stress, or daily hassles, can provide a more universal model of the effects of stress on family functioning.

Perceived Stress

Stress resulting from daily hassles is largely a subjective experience. Some parents can cope with challenging aspects of parenting better than others. Accordingly, all parents do not experience the daily hassles of parenting in the same way. For example, coping with children's misbehavior is so stressful for some parents that changes occur in the parents' mood and level of anxiety; whereas for other parents these are merely transient annoyances of the day. Individual differences in perceptions of parenting stress reflect both biological and environmental influences. Adults bring to parenting different personality styles, different emotional resources, different coping skills, and different childhood experiences of being parented. Indeed, the best estimate of emotional well-being and functioning in parents is the level of well-being and functioning prior to becoming a parent, that is, prior to the birth of the first child. Personality styles most susceptible to experiencing stress are those marked by the prevalence of negative affect (i.e., moodiness and irritability) and trait anxiety (i.e., intolerance of uncertainty and intolerance of arousal). In addition, parents who as children experienced less positive relationships with their own parents and insecure attachments are more likely to perceive their experiences with their own children as less positive.

Critics of the minor stressors, or daily hassles, approach argue that the construct of daily hassles is inherently subjective, whereas major life events can be assessed more objectively. These critics argue that perceptions of minor hassles are confounded with the outcome measures of family functioning; for example, there is an overlap between perceiving parenting as burdensome and perceiving children's behavior as problematic. On the other hand, individual differences clearly exist in parents' ability to cope with the everyday challenges of parenting. And a large literature verifies that perceptions of events influence functioning.

Effects of Stress on Parents

Parental stress from both major life events and from daily hassles has been shown to negatively impact parents' emotional well-being, marital satisfaction, and parenting behavior.

Parental Mood and Emotional Well-Being

Parenting stress is associated with lower emotional well-being in parents. That is, parents who report more daily hassles in parenting and/or more major life events describe themselves as having less life satisfaction and more negative mood and emotional distress. There is some evidence that stress from daily hassles relates more strongly than does stress from major life events, but stress from each of these sources contributes to lower emotional well-being. In addition to the negative impact on their own emotional well-being, parents with high stress also report a less positive outlook on parenting and less satisfaction in the parental role. Furthermore, they tend to experience less pleasure in and enjoyment of their children. Not surprisingly, they report lower feelings of self-efficacy in the parenting role; that is, they feel less competent in carrying out their parental responsibilities and less confident that their efforts will have a positive impact on their children. For some parents, high levels of parenting stress contribute to psychological disorders, such as depression and anxiety. For example, mothers with higher parenting stress from low-birth-weight or medically ill infants are at higher risk for developing postpartum depression.

A bidirectional relationship exists between parenting stress and negative parental mood. Parenting stress can increase negative parental mood, but, conversely, negative mood may also increase perceptions of parenting stress. Parents with negative mood are more likely to attend to their children's negative behaviors, attribute negative intent to ambiguous behaviors, and have a decreased threshold for aversive behaviors.

Marital Satisfaction and Social Support

The stress of parenting can also affect the marital relationship. When parents feel competent in their abilities to care for their children and cope effectively with daily hassles of parenting, the marital relationship can be strengthened and serve as a solid foundation of the family. However, when high levels of parenting stress are experienced, parents often report lower levels of marital satisfaction. One reason for the decrease in marital satisfaction may be that parents who experience high levels of parenting stress have less time and emotional resources to devote to the marital relationship and to meeting their spouse's needs. Another reason may be that high levels of parenting stress increases the frequency of marital conflict and demonstrations of negative affect toward the spouse. Because the primary social support in parenting is typically the spouse, the decrease in marital satisfaction typically decreases the support a parent experiences in the role of parenting. However, others (such as grandparents) may increase the support that they provide so that a parent does not perceive a decrease in the overall support that he or she receive in the parental role.

Again bidirectional relationships occur. Parenting stress can impair marital satisfaction, but conversely, marital satisfaction can mitigate parenting stress. If parents work together, divide responsibilities, and support one another's efforts, the daily hassles of parenting are less stressful.

Parental Interactions with Children

Parenting stress can also have an adverse effect on parents' behavior toward their children. Parents who experience high levels of stress report more negative and coercive interactions with their children; similarly, they report using more authoritarian and power-assertive discipline strategies. They also describe themselves as less involved with their children. These disparate self-reports are suggestive of inconsistent management techniques. Indeed, high levels of stress have been associated with parents' reports of inconsistent disciplinary styles, that is, alternating between setting harsh and rigid limits for their children to being overly lax and permissive. Although much research has established that stressed parents report negatively on their parenting skills, it is less clear that their actual parenting skills are impaired. Research on observed parenting behaviors has been equivocal, although there is some evidence that mothers are more irritable with their children on days when they have experienced more minor stressors. Thus, parents who report higher levels of stress in caring for children may be less likely to initiate

pleasurable interactions with the children and may be less able to use proactive disciplinary strategies.

Effects of Parenting Stress on Children

Parenting stress has repeatedly been found to relate to children's social competence and to behavior problems. Indeed, it is thought to be a primary contributor to the development of psychopathology in children. Again, however, much of this research is based on parental reports. Parents who report high stress describe their children as less socially competent and as having more behavior problems, both externalizing and internalizing problems. These parents also describe their children as exhibiting behaviors consistent with insecure attachments. Limited research using teacher reports and children's self-reports provides some support that parents' negative views may be accurate. Teachers rated the children of highly stressed parents as less socially competent, and children reported more adjustment problems when their parents were highly stressed.

How Might Parenting Stress Affect Children's Outcome?

Parenting stress negatively affects parents' interactions with their children, and the negative interactions, in turn, negatively impact child adjustment. In other words, parenting stress may lead to critical and negative parent-child interactions as well as inconsistent discipline; then these changes in parenting may have negative effects on the children's attachment, self-esteem, and behavior. Another way that parenting stress may affect children's outcome is through modeling affect and behavior regulation. Children observe how parents cope with frustrating situations and how they regulate negative emotions and impulsive behaviors. Parents who experience high parenting stress may model poor regulation of negative affect and more reactive behavioral responses to frustration. Finally, children may be directly experiencing some of the same factors that are causing stress in their parents, such as overcrowding.

Bidirectional effects occur in this arena as well. Parents are crucial in shaping children's behavior, but children also shape adults' parenting behavior. The bidirectional nature of parent-child interactions can create coercive relationships. In these relationships, parents' irritability and less-sensitive parenting fuel behavior problems in children, which evoke further punitive and negative parental behaviors. The cyclical interaction of parents' reactions to stressful circumstances and children's aversive behavior can reinforce a maladaptive pattern of interactions that can lead to

chronic stress in the family environment and can place children at higher risk for delinquency.

Interventions for Parents and Children

With increased attention to the potentially harmful effects of stress on parents, the study of how to promote resiliency in parents has gained increasing attention. Social support and active coping skills have emerged as two primary areas of intervention for relieving parental stress. Social support has been related to decreases in self-reported parental stress. Being connected to others, sharing difficult experiences, and receiving help with the day-to-day care of children reduces parenting stress. Teaching parents strategies for coping with the daily hassles of parenting has also been shown to reduce stress in the parental role. Successful coping involves adapting to the daily hassles of parenting by anticipating problems and having a plan for actively addressing them. For some parents, coping may involve monitoring their children's behavior and environment in order to circumvent the escalation of problem behaviors; for others, it may involve establishing predictable daily routines for children and scheduling periods for taking time off, when another caregiver (e.g., spouse, grandparent, or babysitter) can provide respite care. Specific programs that teach effective parenting skills seem to reduce stress by increasing parents' confidence and efficacy in managing children's problem behaviors. Likewise, interventions that teach young children social problem-solving skills can also decrease parents' stress and increase positive parent-child exchanges. Finally, interventions to decrease parenting stress at the larger community level are also needed to support families with children. Systematic efforts to provide high-quality, affordable child care for working parents, job flexibility for balancing

the demands of full-time employment and parenting, and community recreation programs for families will help reduce the burden of parenting and enhance parents' ability to care for their children.

Further Reading

- Anthony, L. G., Anthony, B. J., Glanville, D. N., et al. (2005). The relationships between parenting stress, parenting behavior and preschoolers' social competence and behavior problems in the classroom. *Infant and Child Development* 14, 133–154.
- Belsky, J. (1984). The determinants of parenting. *Child Development* 55, 83–96.
- Crnic, K. A. and Acevedo, M. (1995). Everyday stresses and parenting. In: Bornstein, M. H. (ed.) *Handbook of parenting Vol. 4: Applied and practical parenting*, pp. 277–297. Mahwah, NJ: Lawrence Erlbaum.
- Crnic, K. A., Gaze, C. and Hoffman, C. (2005). Cumulative parenting stress across the preschool period: relations to maternal parenting and child behavior at age 5. *Infant and Child Development* 14, 117–132.
- Crnic, K. A. and Greenberg, M. T. (1990). Minor parenting stresses with young children. *Child Development* 61, 1628–1637.
- Deater-Deckard, K. (1998). Parenting stress and child adjustment: some old hypotheses and new questions. *Clinical Psychology: Science and Practice* 5, 314–332.
- Deater-Deckard, K. (2004). *Parenting stress*. New Haven, CT: Yale University Press.
- Deater-Deckard, K. (2005). Parenting stress and children's development: introduction to the special issue. *Infant and Child Development* 14, 117–132.
- Rodgers, A. Y. (1998). Multiple sources of stress and parenting behavior. *Children and Youth Services Review* 20, 525–546.
- Sepa, A., Frodi, A. and Ludvigsson, J. (2004). Psychosocial correlates of parenting stress, lack of support and lack of confidence/security. *Scandinavian Journal of Psychology* 45, 169–179.

Parkinson's Disease

S Baser

The Movement Disorder and Spasticity Center,
Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

Clinical Features of Parkinson's Disease

Clues to the Etiology of Parkinson's Disease

Oxidative Stress

Environmental Stresses and Parkinson's Disease

Interactions between Genes and Environment

Medications Used for Parkinson's Disease

Summary

Glossary

Dopamine

A catecholamine neurotransmitter that plays a pivotal role in several key functions of the brain. Normal dopamine neurotransmission in brain is vital for normal motor function. The abnormal movements and deficiency in motor control that characterize Parkinson's Disease are due to a deficiency in central dopamine consequent on the degeneration of dopaminergic neurons.

L-DOPA (Levodopa)

3,4-dihydroxy-L-Phenylalanine is the immediate precursor of dopamine in its biochemical synthesis. Since L-DOPA can cross the blood brain barrier, it is used clinically in the management of Parkinson's Disease to drive increased dopamine synthesis.

Parkin

An E3 ligase protein that normally tags alpha synuclein protein molecules with ubiquitin. This Parkin-induced tagging (ubiquitination) of alpha-synuclein, triggers alpha-synuclein degradation by the cell proteasome. In many cases of inherited Parkinson's disease, however, Parkin is abnormal and fails to ubiquitinate alpha synuclein. As a consequence, synuclein does not undergo degradation. Rather, it accumulates in Lewy bodies, which leads to dopaminergic cell dysfunction and death.

Parkinson's disease (PD)

The second most prevalent neurodegenerative disorder of the Western world that leads to abnormal movements and loss of motor control and function. Specifically, the clinical features include slowness of movement (bradykinesia), muscular rigidity, resting tremor and postural instability. PD is characterized

Substantia nigra

by the appearance of intracytoplasmic inclusions called Lewy bodies in dopamine neurons in the substantia nigra and the progressive loss of these neurons. Lewy bodies are enriched in filamentous α -synuclein and other proteins that are often highly ubiquitinated. The majority of PD cases are sporadic and of unknown origin, but several genes have been identified that, when mutated, give rise to rare, familial forms of the disease.

A midbrain nucleus of the basal ganglia complex that controls motor function. The pars compacta of the substantia nigra contains the cell bodies of the dopaminergic neurons that project to the striatum in the forebrain. Degeneration of these neurons is the most prominent neuropathological feature of Parkinson's Disease and is thought to be the prime cause of the disorder.

Synucleins

Small, soluble proteins expressed primarily in neural tissue and in certain tumors. The family includes three known proteins: alpha-synuclein, beta-synuclein, and gamma-synuclein. Mutations in alpha-synuclein are associated with rare familial cases of early-onset Parkinson's disease, and the protein accumulates abnormally in Parkinson's Disease, Alzheimer's Disease, and several other neurodegenerative illnesses. Alpha synuclein has also been implicated in non-familial (sporadic) PD, but the mechanism has yet to be determined.

Ubiquitin proteasome system

Molecular system by which misfolded proteins are tagged with polyubiquitin for subsequent degradation by the proteasome of the cell.

Clinical Features of Parkinson's Disease

Parkinson's disease (PD) is a slowly progressive, degenerative, central nervous system disorder. First described by James Parkinson in 1817, it is a movement disorder that results from the death of dopamine neurons in the substantia nigra. The substantia nigra is located in the midbrain region of the brain. It consists of two parts: the pars compacta (SNc) and the pars reticulata (SNr). The SNc produces dopamine and is the part that degenerates in PD. After 50% of the dopamine neurons and 75–80% of striatal dopamine is lost, patients start to show symptoms

of PD. Associated with the loss of dopamine is the presence of protein-containing deposits known as Lewy bodies. It is not known what actually initiates the development of PD.

Approximately 1 million patients worldwide have PD, with approximately 60 000 new patients diagnosed per year. The average age of onset is 60 years old; 5–10% show symptoms prior to age 40; 3% show symptoms after age 65, and 10% show symptoms after age 80.

To be diagnosed with PD, the patient must have at least three of four cardinal signs: rest tremor, bradykinesia (slowness of movement), rigidity, and loss of balance. Pharmacologically, PD is differentiated from clinically similar disorders by the absence of secondary causes and by the presence of a good response to the medication L-dopa. Definitive diagnosis of PD can only be made by autopsy.

Parkinsonism

Parkinsonism is a clinical syndrome that may share some features with PD, secondary to a variety of etiologies. These etiologies include infectious, vascular, pharmacological, toxic, metabolic, structural, and various degenerative disorders.

Clues Suggesting Atypical Parkinsonism

- Early onset of, or rapidly progressing, dementia
- Rapidly progressive course
- Supranuclear gaze palsy
- Upper motor neuron signs
- Cerebellar signs: dysmetria or ataxia
- Loss of bladder/bowel control
- Early symptomatic low blood pressure
- Early loss of balance

Clues to the Etiology of Parkinson's Disease

For decades scientists knew that the substantia nigra in the midbrain of Parkinson's patients show loss of dopamine nerve cells, but what caused this loss remained elusive. Von Economo's encephalitis, which started as an epidemic in Europe in 1918 and then spread worldwide was associated with symptoms similar to PD in 80% of survivors. This condition was termed postencephalitic parkinsonism.

An important breakthrough in the pathophysiology of PD occurred in the late 1950s when researchers gave rats powerful antipsychotic drugs that blocked dopamine nerve cells. In the 1960s, scientists used the toxin 6-hydroxy-dopamine in animals to destroy these cells and produce a rat model of PD.

Finding effective treatments for PD proved more difficult because dopamine does not pass directly into the brain. As depicted in the movie *Awakenings*, doctors in the late 1960s began to give patients the dopamine precursor L-dopa, which crosses the blood–brain barrier and is converted into dopamine.

In the 1980s, a sudden outbreak of severe PD appeared in young Californian drug addicts. These addicts had used a drug contaminated with 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP). This chemical is converted to MPP⁺ in the brain, damaging and killing dopamine cells in the substantia nigra. This inadvertent discovery created a second animal model with which to explore the disease.

Oxidative Stress

Damage from oxygen may be one cause of PD. Oxygen free radicals are oxygen molecules that carry an extra electron. They are destructive because in excessive amounts they chemically attack the components of the cell, including proteins, mitochondria, DNA, and lipids in cell membranes. One of the major problems of normal aging is an increased level of these free-radical-damaged proteins, along with damaged DNA and lipids.

Recent research by Bennett at the University of Virginia isolated complex I from the mitochondria of the brains of PD patients. They discovered that the complex I assembly in Parkinson's had 50% more damage from oxygen. The complex I in the brains of PD patients also had evidence of not being properly assembled and had reduced electron flow.

Environmental Stresses and Parkinson's Disease

Toxins and Parkinson's disease

Rural living and pesticide exposure Although studies in the past pointed to an association between rural living and PD, the results of the most recent studies are quite inconsistent. Because there are many specific exposures associated with living in rural areas, recent studies have attempted to measure these exposures in order to explain the association between PD and areas of low population density.

Three recent studies have found an association between PD and agricultural work. A German study reported an elevated and significant relative risk associated with mushroom harvesting during childhood and adolescence, but no association was found with previous farm activity or employment in agricultural work, living on or near a farm, having contact with farm animals, or being involved in slaughtering

animals. Well water has also been implicated in PD; recent studies in Italy and Spain found a positive association between drinking well water and the risk of PD.

The similarity between the structures of the MPP⁺ and the herbicide paraquat encouraged the investigation of a possible relation between pesticide exposure and PD. A study in Taiwan indicated that the incidence of PD was two times greater in people who had used both paraquat and herbicides/pesticides than in those exposed to pesticides and herbicides alone.

Previous studies found a substantially increased rate of PD among city dwellers who gardened as a hobby. Recent studies have consistently shown an increased risk of PD with pesticide exposure, and in some, this achieved statistical significance. A 2000 California-based study showed that individuals with high-level herbicide exposure had a 70% increased risk of developing PD, compared with those who were not exposed. People who used insecticides in the garden showed a 50% increased risk of developing PD compared to those who had never been exposed to home pesticides of any type. In-home use of insect-killing chemicals was associated with a 70% increased risk of developing PD compared with no use of pesticide.

Metal/industrial exposure Several early case reports described Parkinsonian symptoms in association with manganese mining fumes and accidental industrial poisoning. A recent study of 15 welders with Parkinsonian symptoms showed that the welders had an earlier onset of symptoms. Another study reported that counties in Michigan with iron and copper industries had higher PD mortality rates.

Higher mortality rates from PD are reported in regions with paper and chemical industries. More PD patients than controls reported that they had wood paneling in their homes; this may implicate wood preservatives in the etiology of PD.

Head Injury

Head injury has been implicated in the etiology of PD, possibly through mechanisms that are involved in the inflammatory process. Several recent studies found significant risks associated with head trauma, and several well-known boxers are known to have PD.

Smoking

Many epidemiological studies have shown smoking to be protective for PD. A recent prospective study involving the Honolulu Heart Study supported the idea that nicotine may be protective for PD. Other case-control studies, however, do not support the

claim that smoking is protective for PD; the previous negative association was explained by the fact that smokers die sooner than nonsmokers.

Diet and Antioxidants

A low-calorie diet may help to reduce the risk of PD. A team from the U.S. National Institute on Aging found a long-term reduction in caloric intake protects rhesus monkeys from developing the disease; they also had higher levels of a growth factor, glial cell line-derived neurotrophic factor (GDNF), which may protect brain cells from destruction.

The role of dietary antioxidants has been examined in the etiology of PD. According to the oxidative stress model, an increase in antioxidants prevents damage and death to the dopaminergic cells by scavenging more free radicals. Epidemiological studies examining the association between antioxidant consumption, including vitamin E, C, and A, have not shown a consistent role.

Vitamin B₆ is essential for the metabolism of protein and for brain function. In a study of more than 5000 people, Dutch researchers found those who reported taking in the most vitamin B₆ were approximately half as likely to develop PD as those who consumed the least of this vitamin.

High levels of the amino acid homocysteine can damage brain cells. To see whether a higher intake of folate and vitamins B₆ and B₁₂, which can reduce homocysteine levels, also reduced the risk of PD, researchers followed 5289 men and women ages 55 and older. The researchers found no association between the consumption of B₁₂ or folic acid and the risk of developing the disease; however, the more B₆ people consumed, the lower their risk of PD.

Interactions between Genes and Environment

Genes, Cell Death, and α -Synuclein

No one single underlying factor has been shown to cause PD. Most current research shows that it is a complex multifactorial disease resulting from interaction between one or more genes and the environment. The genes that cause inherited PD can involve the proteins α -synuclein, parkin, and ubiquitin. α -Synuclein forms a major component of Lewy bodies, a classic pathological finding in PD patients.

The discovery of the dominantly inherited gene (α -synuclein, SCNA gene, PARK1) was first described in the Contursi region of Italy. This discovery developed the recognition of a direct role for genetics in PD. Abnormally high levels of α -synuclein, which is produced in dopaminergic cells, may play an

Table 1 Medications^a

<i>Medication</i>	<i>Mechanism</i>	<i>Examples</i>
L-dopa	Crosses the blood–brain barrier and is converted into dopamine	
MAOIs	Block dopamine metabolism	Selegiline, rasagaline
Dopamine agonists	Act directly on dopamine receptors (D1 and D2) to take the place of dopamine	Bromocriptine (Parlodel), pergolide (Permax), pramipexole (Mirapex), ropinirole (Requip), apomorphine
NMDA antagonists	Block excitatory receptors	Amantadine and flumadine
COMT inhibitors	Block the breakdown of L-dopa in the bloodstream, allowing more L-dopa to cross the blood–brain barrier	Tolcapone (Tasmar), entacapone (Comtan)
Anticholinergics	Decrease the activity of acetylcholine (a balancing neurotransmitter)	Trihexyphenidyl (Artane), benztropine mesylate (Cogentin)

^aCOMT, catechol-O-methyltransferase; MAOIs, Monoamine oxidase inhibitors; NMDA, N-methyl-D-aspartic acid.

important role in the death of these cells. In patients with the defective gene, synuclein accumulates, causing dopamine neurons to become more vulnerable to neurotoxic insults. Normally, two other molecules, parkin and ubiquitin, are involved in the natural self-destruction of synuclein in the natural process of programmed cell death called apoptosis.

Genes, Cell Death, and Ubiquitin

Recessively inherited juvenile parkinsonism (AR-JP, PARKIN, PARK2) is one of the most common forms of familial PD. AR-JP is characterized by the selective and massive loss of dopamine neurons in the substantia nigra of the midbrain and the absence of Lewy bodies (the pathological hallmark of PD). Parkin mutations in this form of PD cause the failure of proteolysis mediated by the ubiquitin system and the accumulation of protein, causing dopamine nerve cell death.

Emotional Stress and Parkinson's Disease

Emotional stress increases the symptoms of PD, including the worsening of tremor. Depression in Parkinson patients is common, with some estimates reaching 50% of those affected. Biochemical changes in the brain involving serotonin, catecholamines, and dopamine and emotional responses to chronic disease play a role.

Encouraging patients to perform as many daily activities as possible and to follow a program of regular exercise can help people with PD maintain mobility. Recent research suggests additional benefits of exercise, including neuroprotective effects. Physical and occupational therapy can help patients maintain or regain muscle tone, maintain range of motion, and learn adaptive strategies. Music therapy has been shown to improve slowness in PD.

Caring for a spouse with PD is associated with emotional and social stress, underlining the importance of also assessing the needs of caregivers. One

study showed that spouses of PD patients had more severe depression and that a higher proportion reported tiredness, sadness, and less satisfaction with life than healthy elderly subjects. Caregivers can benefit from learning about the physical and psychological effects of PD and about ways to enable patients to function as well as possible. Because such care is tiring and stressful, caregivers may benefit from support groups.

Medications Used for Parkinson's Disease

The medications used for PD patients compensate for the loss of the dopamine-producing cells by either increasing dopamine levels or simulating the effect of dopamine in the brain (Table 1). Dopamine cannot be given directly because it does not cross the blood–brain barrier.

Summary

Many stressors have been associated with the development of PD. Complex interactions among the environment, oxidants, toxins, genes, and emotions play a role in the etiology of PD.

See Also the Following Articles

Brain Trauma; Diet and Stress, Non-Psychiatric; Dopamine, Central; Environmental Factors; Environmental Stress, Effects on Human Performance; Oxidative Stress; Renal and Adrenocortical Actions of Dopamine; Smoking and Stress; Diet and Stress, Psychiatric; Early Environment and Adult Stress; Neurodegenerative Disorders.

Further Reading

Aarsland, D., Larsen, J. P., Karlsen, K., et al. (1999). Mental symptoms in Parkinson's disease are important contributors to caregiver distress. *International Journal of Geriatric Psychiatry* 14(10), 866–874.

- Fearnley, J. M. and Lees, A. J. (1991). Ageing and Parkinson's disease: substantia nigra regional selectivity. *Brain* 114(5), 2283–2301.
- Gosal, D., Ross, O. A. and Toft, M. (2006). Parkinson's disease: the genetics of a heterogeneous disorder. *European Journal of Neurology* 13(6), 616–627.
- Kontakos, N. and Stokes, J. (2000). Monograph series on aging-related diseases. XII: Parkinson's disease-recent developments and new directions. *Chronic Diseases in Canada* 20(2), 58–77.
- Lang, A. E. and Lozano, A. M. (1998). Parkinson's disease, I. *New England Journal of Medicine* 339(15), 1044–1053.
- Lang, A. E. and Lozano, A. M. (1998). Parkinson's disease, II. *New England Journal of Medicine* 339(16), 1130–1143.
- Marsden, C. D. (1994). Clinical Neuropharmacology. *Journal of Neurol Neurosurg Psychiatry* 57(6), 672–681.
- Olanow, C. W. and Tatton, W. G. (1999). Etiology and pathogenesis of Parkinson's disease. *Annual Review of Neuroscience* 22, 123–144.
- Pacchetti, C., Mancini, F., Aglieri, R., et al. (2000). Active music therapy in Parkinson's disease: an integrative method for motor and emotional rehabilitation. *Psychosomatic Medicine* 62, 386–393.
- Polymeropoulos, M. H. (1998). Autosomal dominant Parkinson's disease and alpha-synuclein. *Annals of Neurology* 44(3 supplement 1), S63–S4.
- Riess, O. and Kruger, R. (1999). Parkinson's disease – a multifactorial neurodegenerative disorder. *Journal of Neural Transmission* 56(supplement), 113–125.
- Tanaka, K., Suzuki, T., Chiba, T., et al. (2001). Parkin is linked to the ubiquitin pathway. *Journal of Molecular Medicine* 79(9), 482–494.
- Trimmer, P. A., Swerdlow, R. H., Parks, J. K., et al. (2000). Abnormal mitochondrial morphology in sporadic Parkinson's and Alzheimer's disease cybrid cell lines. *Experimental Neurology* 162(1), 37–50.
- Wersinger, C. and Sidhu, A. (2006). An inflammatory pathomechanism for Parkinson's disease? *Current Medicinal Chemistry* 13(5), 591–602.

Peacekeeping

B Litz and S Maguen

VA Boston Healthcare System, Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC), and Boston University School of Medicine, Boston, MA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by B Litz and E Bolton, volume 3, pp 134–137, © 2000, Elsevier Inc.

Types of Peacekeeping Missions
Unique Stressors of Peacekeeping
Psychological Impact of Peacekeeping
Conclusions and Recommendations

Glossary

<i>Peace enforcement</i>	A process by which military personnel necessitate a cessation of hostilities and facilitate a peace process.
<i>Peacekeeping</i>	Maintaining a military presence that increases the likelihood of peace and reduces the likelihood of fighting among two or more formerly warring parties.
<i>Peacekeeping roles</i>	Duties such as guarding and patrolling vulnerable areas, helping maintain crowd control, operating checkpoints, providing humanitarian assistance, and facilitating the rebuilding of infrastructure, all of which can expose peacekeepers to

potentially dangerous and life-threatening situations.

Posttraumatic stress disorder (PTSD)

A psychological disorder that may result from exposure to severe stress and that involves re-experiencing of the traumatic event, avoidance of cues associated with the event, emotional numbing, and hyperarousal symptoms.

Role conflict in peacekeepers

Conflict resulting from exposure to competing demands of neutrality and restraint in the face of life threat or danger, which may lead to feelings of helplessness or frustration.

Rules of engagement

Rules prohibiting use of weapons except in extreme circumstances, which is quite different from ways in which soldiers are traditionally trained to handle provocation and life threat.

Types of Peacekeeping Missions

Peacekeeping has changed considerably following the inception of the United Nations in 1945. There have been 60 peacekeeping missions since 1948, and currently there are 17 active operations throughout the world. Historically, the role of the military in peacekeeping operations has been that of maintaining a strictly neutral presence by overseeing peace accords between formerly warring parties. However, post-Cold War, peacekeepers have shifted to enforcing peace in the midst of active hostilities.

These modern missions have placed peacekeepers into more dangerous and conflict-laden environments. Modern missions require soldiers to be combat trained, yet able to model restraint and maintain neutrality. As a result, peace enforcement is one of the most stressful duties a soldier can be exposed to in the modern military, and the possibility of exposure to traumatic events and death has become even more likely in modern missions. For example, Seet and Burnham found that more deaths have occurred among UN peacekeepers in the past decade than in the previous 40 years, and that deaths from hostile acts have significantly increased following the Cold War.

This article provides brief summaries of the stressors associated with peacekeeping in three different peacekeeping missions, briefly reviews recent peacekeeping missions, summarizes the unique stressors of peacekeeping, and highlights the psychological impact of these missions.

The Multinational Force and Observers in the Sinai

Since the establishment of the state of Israel in 1948, the UN and other multinational organizations have engaged in an ongoing struggle to create and maintain peace between Israel and its neighbors. An example of a very successful attempt to maintain peace in this region is the Multinational Force and Observers (MFO) in the Sinai. Due to the stability of the peace between Israel and Egypt and the receptiveness of both parties to the presence of the MFO, peacekeepers in the Sinai have been exposed to few potentially traumatizing events directly related to their service as peacekeepers. The MFO is an example of a successful peacekeeping mission in which the duty of peacekeepers is strictly to observe, monitor, and report transgressions.

Operation Restore Hope and Operation Continue Hope in Somalia

As a result of the civil war that erupted in Somalia in 1991, humanitarian relief efforts aimed at curbing famine and the spread of disease were interrupted and sabotaged. The UN, with extensive support from the United States, decided to guarantee the provision of humanitarian aid as well as to enforce the peace in Somalia. During this mission, peacekeepers became peace enforcers, and the risk of exposure to potentially traumatic events grew exponentially. Although Operation Restore Hope (ORH) was a great success with regard to the provision of medical and food supplies, the mission is often considered to be a failure because Somalia continues to be at risk for the devastating effects of famine, political instability, and violence.

While in Somalia, U.S. military personnel were assigned to a variety of tasks, ranging from policing

duty to combat-like duty (e.g., patrols, disarming civilians). Peacekeepers were exposed to a fairly well armed civilian population who were actively engaged in interclan war. Peacekeepers in Somalia were also subjected to acts of aggression by some Somalis, yet the strict rules of engagement sharply restricted options for protection or retaliation, creating a general sense of threat.

Operation Joint Endeavor and Operation Joint Guard

In 1991, following the declaration of independence of Slovenia and Croatia, a civil war erupted in the former Yugoslavia. The disillusion of the state gave rise to power struggles over governing control of the remaining states and to old race-based hatreds. Horrible atrocities were perpetrated, including genocidal acts. In response, approximately 60 000 North Atlantic Treaty Organisation (NATO) and U.S. military personnel were deployed as peacekeepers. Peacekeepers in Bosnia-Herzegovina have been subjected to shelling and sniper attacks, and a minority of peacekeepers have also been taken hostage. Some UN peacekeepers have had to stand by helplessly while atrocities were taking place.

Recently Established Peacekeeping Missions

Locations of recent UN peacekeeping missions include (1) Sudan (UNMIS), which began in March 2005 to support implementation of the Comprehensive Peace Agreement, (2) Burundi (ONUB) since May 2004, (3) Haiti (MINUSTAH) since April 2004, (4) Cote D'Ivoire (UNOCI) since February 2004, (5) Liberia (UNMIL) since September 2003, (6) East Timor (UNMISSET) since May 2002, and (7) Ethiopia and Eritrea (UNMEE) since September 2000.

Unique Stressors of Peacekeeping

Peacekeepers are at risk for exposure to stressors that are typical of a war zone (artillery fire, land mines, small arms fire). Even when peacekeepers are tasked with maintaining a firm peace, they are faced with the possibility of life threat. They may be fired upon as a result of a misunderstanding, accidentally in cross-fire between two armed feuding parties, or during firing close, which occurs when the opponent wishes to intimidate the peacekeepers in order to keep them away from a certain area. Peacekeepers may also witness atrocities committed against fellow peacekeepers and civilians and the malicious destruction of property. Other types of stressful experiences for peacekeepers are boredom, isolation, family separation, exhaustion, unfavorable climatic conditions, and demoralization about the mission's efficacy.

Role conflicts may also provoke additional stress in peacekeepers. For combat-trained soldiers from larger nations, peacekeeping duty may feel incongruent with their training. For example, peacekeepers, unlike traditional soldiers, are often restrained from taking offensive action in conditions of life threat. In addition, the types of defensive military structures that are commonplace in war are often not as available on peacekeeping missions due to the proximity that is required in order to provide humanitarian assistance and protection. The emphasis on proximity rather than protection creates considerable hypervigilance and arousal in peacekeepers and contributes to a general sense of fear. Peacekeepers who are unclear about how to respond to threats and/or experience repeated threats of injury, with little or no opportunity for recourse, are likely to experience great anxiety. Furthermore, peacekeepers who are forced to suppress their frustration, fear, and anger are at risk for acting out their feelings during a mission and/or upon their return home.

Psychological Impact of Peacekeeping

Although there are many potentially traumatizing experiences and psychological conflicts associated with being a peacekeeper, the great majority of soldiers appear to adapt well. However, a sizeable number of peacekeepers are at risk for the development of psychopathology related to their deployment. The most frequent mental health problems associated with peacekeeping are PTSD, depression, anger and hostility problems, and alcohol abuse. Mehlum and Weisaeth reported that 7 years following service, 16% of prematurely repatriated soldiers from southern Lebanon met criteria for PTSD. Passey found that more than 20% of the soldiers deployed as peacekeepers to Bosnia endorsed symptoms of PTSD and depression. In Litz and colleagues' examination of U.S. military personnel who deployed to Somalia, 25% of Somalia veterans reported clinically significant psychological distress, particularly hostility and anger problems, and 8% reported clinically significant PTSD; Dirkzwager and colleagues found similar rates of PTSD among peacekeepers who served in Yugoslavia. Among Kosovo peacekeepers, Maguen and colleagues found that 4% endorsed clinically significant PTSD symptoms; Dirkzwager and colleagues found similar rates of PTSD among peacekeepers who served in Cambodia. Gray and colleagues found that although most follow a standard pattern of PTSD development, some peacekeepers exhibit delayed-onset PTSD after a period of minimal distress.

Adler and colleagues examined factors that predicted PTSD in U.S. military personnel deployed to

the Bosnia peacekeeping mission and found that longer deployments and first-time deployments were associated with an increase in distress scores. Interestingly, the association between deployment length and distress was significant only for male soldiers. In a prospective study of Kosovo peacekeepers, Maguen and colleagues found that pre-existing stress symptoms and exposure to stressful war zone events (e.g., going on dangerous patrols) were the most robust predictors of PTSD symptoms. Similarly, in Litz and colleagues' study of Somalia peacekeepers, the extent of exposure to stressful war zone events and frustrations with aspects of the peace enforcement mission (e.g., restrictive rules of engagement) predicted the severity of PTSD. The relationship between war zone exposure and PTSD was strongest for Somalia veterans who had high levels of frustration with the negative aspects of peacekeeping duty. Additional variables associated with PTSD symptoms reported by Dirkzwager and colleagues were feelings of powerlessness and threat, belief that the mission had become meaningless, and lack of control during deployment. Litz and colleagues also found that cohesion and morale during deployment were negatively associated with symptoms of psychological distress and PTSD. On the other hand, Michel and colleagues found that postdeployment adversities most strongly contributed to poor mental health among Swedish peacekeepers that deployed to Bosnia.

Peacekeeping duty can also promote positive outcomes. Dirkzwager and colleagues found that the majority of peacekeepers reported positive consequences of their deployment, with 82% reporting a broadening of their horizon and 52% reporting increased self-confidence. Britt and Adler, and Bartone documented that following deployment to Bosnia, many peacekeepers reported an increase in political understanding, stress tolerance, and professional qualifications. They also found that soldiers who identified more closely with the role of peacekeeper and who believed their assignment to be important were more likely to report a perceived benefit from their peacekeeping experience and were less likely to report burnout and adverse psychological consequences. Finally, the researchers noted that soldiers who reported benefits as a result of the deployment had at least some exposure to the physical damage caused to the local civilians and soldiers from other nations.

Conclusions and Recommendations

Even in the context of conflicts that arise as a result of counterterrorism, peacekeeping and peace enforcement missions will thankfully outnumber wars internationally. Although the psychological casualties

resulting from peacekeeping stress do not compare to combat, it is important to appreciate the mental health needs of soldiers who return from peacekeeping missions, particularly ones that entail unforeseen escalations in hostilities. Peacekeepers may experience a wide range of potentially traumatic events and numerous uniquely stressful experiences. Peacekeeping, especially peace enforcement, missions should include provisions for secondary prevention interventions for those soldiers most at risk for chronic mental health problems following homecoming.

See Also the Following Article

War-Related Posttraumatic Stress Disorder, Treatment of.

Further Reading

- Adler, A. B., Huffman, A. H., Bliese, P. D. and Castro, C. A. (2005). The impact of deployment length and experience on the well-being of male and female soldiers. *Journal of Occupational Health Psychology* 10, 121–137.
- Britt, T. W., Adler, A. B. and Bartone, P. T. (2001). Deriving benefits from stressful events: the role of engagement in meaningful work and hardiness. *Journal of Occupational Health Psychology* 6, 53–63.
- Dirkzwager, A. J., Bramsen, I. and van der Ploeg, H. M. (2005). Factors associated with posttraumatic stress among peacekeeping soldiers. *Anxiety, Stress, and Coping* 18, 37–51.
- Gray, M. J., Bolton, E. E. and Litz, B. T. (2004). A longitudinal analysis of PTSD symptom course: delayed-onset PTSD in Somalia peacekeepers. *Journal of Consulting and Clinical Psychology* 72, 909–913.
- Litz, B. T., King, L. A., King, D. W., Orsillo, S. M. and Friedman, M. J. (1997). Warriors as peacekeepers: features of the somalia experience and PTSD. *Journal of Consulting and Clinical Psychology* 65(6), 1001–1010.
- Litz, B. T., Orsillo, S. M., Friedman, M., Ehlich, P. and Batres, A. (1997). Post-traumatic stress disorder associated with peacekeeping duty in Somalia for U.S. military personnel. *American Journal of Psychiatry* 154, 178–184.
- Maguen, S. and Litz, B. T. (2006). Predictors of morale in U.S. peacekeepers. *Journal of Applied Social Psychology* 36, 820–836.
- Maguen, S., Litz, B. T., Wang, J. L. and Cook, M. (2004). The stressors and demands of peacekeeping in Kosovo: predictors of mental health response. *Military Medicine* 169, 198–206.
- Mehlum, L. and Weisaeth, L. (2002). Predictors of posttraumatic stress reactions in Norwegian U.N. peacekeepers seven years after service. *Journal of Traumatic Stress* 15, 17–26.
- Michel, P. O., Lundin, T. and Larsson, G. (2003). Stress reactions among Swedish peacekeeping soldiers serving in Bosnia: a longitudinal study. *Journal of Traumatic Stress* 16, 589–593.
- Passey, G. and Crocket, D. (1995). *Psychological consequences of Canadian UN peacekeeping in Croatia and Bosnia*. Paper presented at the meeting of the International Society for Traumatic Stress Studies, Boston, MA.
- Seet, B. and Burnham, G. M. (2000). Fatality trends in United Nations Peacekeeping Operations, 1948–1998. *Journal of the American Medical Association* 284, 598–603.

Peptides

E P Zorrilla and G F Koob

The Scripps Research Institute, La Jolla, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by E P Zorrilla and G F Koob, volume 3, pp 138–141, © 2000, Elsevier Inc.

Definition and Classification of Peptides

Biosynthesis and Processing of Peptides: The Proopiomelanocortin Model

Peptide Release

Peptide Receptors

Modulators of Peptide Action

Glossary

Hormone

A peptide, polypeptide, or protein that is released into the circulation via endocrine or exocrine routes to signal to distant tissues.

Neurotransmitter

A compound synthesized and stored within a neuron that, on physiological release into the synapse via a calcium-dependent mechanism, elicits a postsynaptic, receptor-mediated response. Neurotransmitters are also subject to degradation.

Preprohormone

A multipotential precursor protein that contains the amino acid sequence for several different peptides that can be

	liberated with appropriate enzymatic processing; typically, it is inert.
<i>Receptor desensitization</i>	The process by which the repeated stimulation of a receptor with an agonist elicits a smaller response or requires larger doses to achieve a similar response.
<i>Ribosome</i>	The structure outside the nucleus of the cell in which mature mRNA is translated into the coded protein.

Definition and Classification of Peptides

Peptides are generally defined as any member of a class of low-molecular-weight compounds that yield two or more amino acids on hydrolysis. Conversely, peptides also can be defined as any compound produced by amide formation between a carboxyl group (COOH) of one amino acid and an amino group (NH₂) of another, thereby resulting in the loss of water from the adjacent amino acids. Although the word peptide is most often used to refer to compounds in which the amide bonds are formed between C-1 of one amino acid and N-2 of another (or eupeptide bonds), it also applies to compounds with residues linked by other amide bonds (or isopeptide bonds). Because the elements of water are removed from

amino acids in forming peptide bonds, what remains of each amino acid is referred to as an amino acid residue (e.g., glutamic residue instead of glutamic acid). The 21 amino acids known to be encoded genomically and their common abbreviations are listed in [Table 1](#).

Peptides act as signaling agents to transfer information between tissues. Naturally occurring peptides consist of 100 or fewer residues (longer sequences generally are referred to as proteins and have a more complex tertiary structure), making them up to 50 times larger than classic neurotransmitters. Although peptides are found in much lower concentrations than conventional neurotransmitters, their larger size offers a greater capacity for encoding receptor-appropriate structural information, such that peptides can bind with 1000-fold higher affinity and greater selectivity than their smaller transmitter cousins. Peptides generally colocalize with other peptides and transmitting agents, such as classical neurotransmitters and cytokines. The diversity of available chemical messengers greatly increases the permutations of information transfer from a given cell. Peptides are thought to represent an evolutionarily older class of transmitter than the biogenic amines, with peptides having their roots in hormonal function. More than 70 peptides

Table 1 Amino acids and their symbols

<i>Amino acid</i>	<i>One-letter symbol</i>	<i>Three-letter symbol</i>	<i>DNA codons</i>
Alanine	A	Ala	GCT, GCC, GCA, GCG
Aspartic acid or asparagine ^b	B	Asx	See specific entries
Cysteine	C	Cys	TGT, TGC
Aspartic acid	D	Asp	GAT, GAC
Glutamic acid	E	Glu	GAA, GAG
Phenylalanine	F	Phe	TTT, TTC
Glycine	G	Gly	GGT, GGC, GGA, GGG
Histidine	H	His	CAT, CAC
Isoleucine	I	Ile	ATT, ATC, ATA
Lysine	K	Lys	AAA, AAG
Leucine	L	Leu	CTT, CTC, CTA, CTG, TTA, TTG
Methionine	M	Met	ATG
Asparagine	N	Asn	AAT, AAC
Proline	P	Pro	CCT, CCC, CCA, CCG
Glutamine	Q	Gln	CAA, CAG
Arginine	R	Arg	CGT, CGC, CGA, CGG, AGA, AGG
Serine	S	Ser	TCT, TCC, TCA, TCG, AGT, AGC
Threonine	T	Thr	ACT, ACC, ACA, ACG
Selenocysteine	U	Sec	Specific in-frame TGA codons ^a
Valine	V	Val	GTT, GTC, GTA, GTG
Tryptophan	W	Trp	TGG
Tyrosine	Y	Tyr	TAT, TAC
Glutamic acid or glutamine ^b	Z	Glx	See specific entries
Unknown/Other	X	Xaa	—

^a Although TGA also functions as a stop codon, it is now recognized that certain in-frame TGA nucleotide sequences conditionally direct the incorporation of selenium as selenocysteine, the twenty-first amino acid.

^b Used when identity is ambiguous. Once identity is confirmed, the appropriate amino acid-specific symbol is used. For glutamic acid/ glutamine (Z, Glx), the symbols also refer to any substance that yields glutamic acid on acid hydrolysis of the peptide.

have been identified in neural tissue, many in extra-hypothalamic tissue, where they act locally as neurotransmitters, neuromodulators, or trophic factors. Possibly related to their evolutionary history, many peptides have a longer duration of action than their amine counterparts. Thus, some have been described as modulating the tone of the information circuits over long time intervals. In other instances, the neuromodulatory action of peptides is to tune the gain of circuits in response to fast-acting, classic synaptic transmitters such as glutamate or γ -aminobutyric acid. Nevertheless, some peptides that meet the conventional criteria for neurotransmitters also signal information more dynamically. Although there is no consensus regarding a classification system for peptides, one approach has been to divide them according to their major anatomical location and effector mechanism. According to this scheme, biologically active peptides can be divided into the following classes: (1) gut peptides and central nervous system (CNS) neuropeptides, (2) hypothalamic neuropeptides that control the release of anterior pituitary hormones via a neurohormonal route, and (3) pituitary peptides that act primarily as hormones. Although this classification served a heuristic purpose, most peptides or homologous peptide gene products (e.g., gonadotropin-releasing hormone (GnRH) I and II and urocortins 1, 2, and 3) and their putative cognate receptors were subsequently identified in tissues other than those in which they were originally identified, in many cases acting in autocrine or paracrine, rather than endocrine, fashion. These tissues include the muscle; lungs; heart; adipose tissue; skin; and glial, immune, and hematopoietic cells and indicate the functional diversity of peptide action.

Biosynthesis and Processing of Peptides: The Proopiomelanocortin Model

Unlike biogenic amines, which are produced nonribosomally by biosynthetic enzymes, neuropeptides and hormones are derived from larger, mostly inactive precursor molecules that are gene products. Bioactive peptides are enzymatically cleaved and stored from these multipotential precursors. The biosynthesis and processing of one such precursor – proopiomelanocortin (POMC), the prohormone for adrenocorticotrophic hormone (ACTH), melanocyte-stimulating

hormone (MSH), lipotrophic hormone (LPH), and several opioids – has been well characterized and serves as a useful model for understanding the production of peptides *in vivo*. The POMC gene was the first neuropeptide precursor gene to be cloned and, like most peptide genes, has been found to be organized similarly across mammalian species. Spanning 6 kb, the POMC gene comprises three exons and two introns (see Figure 1). Upon transport from the nucleus to the cytoplasm, functional, mature POMC mRNA transcripts span 1100–1400 nt due to tissue-specific differences in the degree of 3' polyadenylation of the mRNA. Transcripts subsequently bind to free ribosomes, where the translation of the N-terminus of the preprohormone begins. Exon 2 of the POMC gene codes for a 26-amino-acid signal (pre)peptide characteristic of all known secretory proteins. The binding of this α -helical signal peptide to a signal-recognition ribonucleoprotein and the association of this complex to a docking protein on the rough endoplasmic reticulum (RER) is required for further translation of peptide precursors. Once in the RER, the signal peptide is removed and the translation of the prohormone is completed. As can be seen in Figure 2, the translated POMC prohormone contains a multipotential N-terminal domain that contains the sequence for γ -MSH, a joining peptide, and domains that contain the full sequences for ACTH and β -lipotropin. From this point on, the prohormone undergoes posttranslational processing in a biologically regulated and tissue-specific manner to yield tissue- and state-specific peptide cocktails ready to be released from secretory vesicles. As can be seen in Figure 2, a multitude of peptides could be cleaved from this hormone, depending on physiological demands. Adding yet further flexibility to the system, certain peptide sequences are contained in multiple prohormones (e.g., met-enkephalin is present in both POMC and preproenkephalin). Much of this processing involves proteolytic cleavage by convertases at sites distal to dibasic residues (Arg-Lys, Lys-Arg, Lys-Lys) that flank the desired peptide sequence. For some peptides, a single basic residue (Arg) serves as the cleavage signal. Subsequently, the flanking dibasic peptides are snipped from the target peptide by carboxypeptidases. En route from the RER to the smooth endoplasmic reticulum, the Golgi apparatus and, ultimately, the secretory granules, the prohormone, and its peptide derivatives are also subject

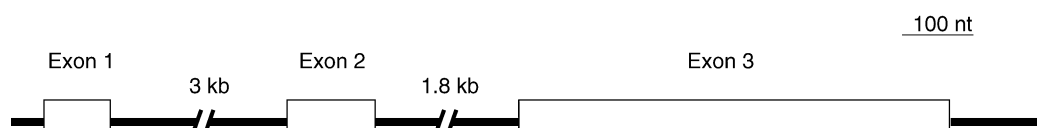


Figure 1 Representation of the structure of the rat proopiomelanocortin (POMC) gene.

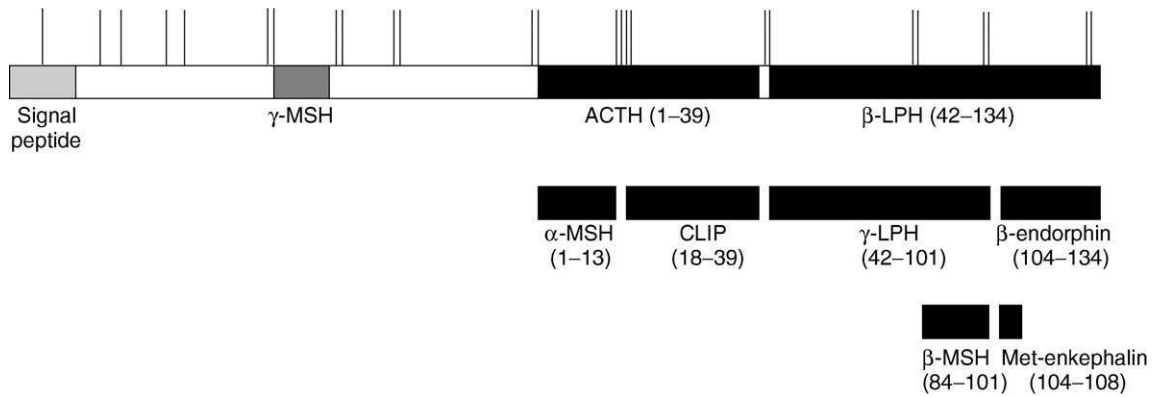


Figure 2 Representation of the structure of the bovine preproopiomelanocortin prohormone. The locations of known peptide components are shown in closed boxes with the amino acid numbers in parentheses. Paired basic amino acid residues (Arg-Lys, Lys-Arg, and Lys-Lys) and single cysteine residues, which represent possible cleavage sites, are marked with vertical bars. ACTH, adrenocorticotropic hormone; CLIP, peptide fragment; LPH, lipotrophic hormone; MSH, melanocyte-stimulating hormone.

to several enzymatic nonproteolytic modifications, including glycosylation, deglycosylation, acetylation, sulfation, phosphorylation, carboxymethylation, and C-terminal amidation. These modifications facilitate transport or storage, increase metabolic stability, and alter the product's bioactivity (e.g., the importance of C-terminal amidation for many peptides or of Ser3-acetylation for ghrelin). Thus, the process by which a particular peptide or peptide cocktail is localized in a secretory granule is extremely complex, with regulation (or dysregulation) possible at the level of prohormone synthesis, enzymatic processing, transport, and storage.

Peptide Release

Some peptides, notably growth factors, follow a constitutive route of release whereby they proceed directly from the Golgi network to the cell surface, bypassing storage in secretory granules. More typical of stress-related peptides found in endocrine cells and neurons, however, is a regulated route of release, as described in the POMC model, whereby the peptide is packaged in large, dense, core secretory vesicles until liberated by the appropriate stimulation of the cell. Neuropeptide release, like that of biogenic amine neurotransmitters, is induced by cell depolarization in a calcium-dependent manner. However, unlike bioamines, which are released from smaller clear synaptic vesicles during low-frequency activity, neuropeptides are generally released from large, dense core vesicles only on high-frequency burst firing. Because peptides lack reuptake sites to replenish secretory stores, *de novo* synthesis of new peptide is the major mechanism for replacement after release, such that peptide release can often be inferred indirectly from a compensatory increase in peptide mRNA expression.

Peptide Receptors

Once released, peptide ligands must interact with their receptors to transmit information. Neuropeptide receptors are generally expressed at low densities (1–10 pmol/mg protein at highest concentrations). Prior to the advent of molecular biological techniques, this made them difficult to isolate, clone, and characterize, although much recent progress has been made. They have been observed both pre- and postsynaptically on the soma, axons, and dendrites of neurons, indicating a complex mode of communication.

G-protein-coupled receptors

With few exceptions, peptide receptors are guanine nucleotide-binding (G)-protein-coupled receptors (GPCRs). The G-protein-coupled superfamily of receptors, which represents approximately 80% of all known receptors, has a highly conserved domain structure, with seven predicted transmembrane α -helices. The hydrophobic transmembrane core regions show the greatest homology across peptide receptors, whereas variation in the intra- and extracellular domains yields the diversity of receptor families and subtypes. A conserved, paired acid–base residue at the base of the third transmembrane appears to play a role in G-protein activation. The full range of nonphotorceptor G-protein-mediated signal transduction mechanisms has been described for these receptors, including the positive and negative regulation of adenylyl cyclase, stimulation of phospholipase C, and positive and negative regulation of ion channels (Ca^{2+} , K^{+} , and nonselective $\text{Na}^{+}/\text{K}^{+}$). Of approximately 360 GPCRs coded in the human genome and identified in a database maintained by the International Union of Pharmacology Basic and Clinical Pharmacology Committee on Receptor Nomenclature and

Drug Classification, approximately one-fourth are known to have peptide ligands. Of peptide GPCRs, approximately 70 belong to the largest family, the class A or rhodopsinlike receptors; 15 belong to the class B (secretinlike) family of receptors for large polypeptidelike hormones; and none belong to the class C receptor family. A further 70 orphan GPCRs are estimated to have peptide ligands.

A few peptide receptors belong to the guanylate cyclase (rather than GPCR) family of membrane receptors. Ligands for these receptors, which are single-transmembrane proteins that regulate intracellular cyclic guanine monophosphate (cGMP) accumulation, include atrial natriuretic peptide, C-type natriuretic peptide, uroguanylin, and guanylin. Finally, a very limited number of peptides bind to other receptor types, including an ionotropic receptor for an invertebrate tetrapeptide FMFRamide and sortilin, a large, single-transmembrane receptor for neurotensin that has a long extracellular domain.

Theories of Peptide–Receptor Interaction

Peptides' activation of receptors cannot be explained by a simple lock-and-key model of amine–receptor interaction. Rather, two heuristic models, although theoretical, provide a framework for understanding peptide–receptor interactions: a zipper model and an address–message model. The former emphasizes that the association of a single amino acid on the peptide to a complementary receptor binding site initiates a series of interactions that results in a receptor-bound activating confirmation. The latter notes that certain portions of the peptide (the address) are more important for binding the ligand with the receptor, whereas other portions (the message) are more important for initiating the conformational changes in the receptor that lead to signal transduction. Many, if not most, peptide receptors also exhibit a state of constitutive activation in the absence of ligand. Like their ligands, peptide receptors undergo posttranslational processing, including alternative splicing and differential N-glycosylation and phosphorylation. These modifications, which alter receptor binding affinities and signal-transduction efficiency, form the basis of several receptor subtypes.

Desensitization and Regulation of Peptide Receptors

Many peptide receptors, like their amine counterparts, undergo rapid-onset (<30 min) homologous receptor desensitization. The C-terminal tail of the receptor appears to participate in this process, perhaps by modifying intracellular protein interactions. Some

receptors also exhibit heterologous augmentation of homologous desensitization; for example, arginine vasopressin synergistically enhances corticotropin-releasing factor (CRF)-induced desensitization of CRF receptors. However, the mechanism for this process is unclear. Finally, peptide receptor synthesis and expression also are subject to homologous and heterologous regulation.

Modulators of Peptide Action

Unlike biogenic amines, whose actions are terminated in part by cellular reuptake, peptide action is modulated by two specific largely extracellular mechanisms – peptidases and peptide-binding proteins.

Peptidases

Proteases are enzymes encoded by 2% of genes in all living organisms that hydrolyze peptide bonds, potentially with ensuing functional consequences. Proteases can be classified as the somewhat more common endopeptidases, which cleave at internal residues, or as exopeptidases, which break terminal peptide bonds. Exopeptidases are further subclassified into aminopeptidases and carboxypeptidases, according to their localized site of action. Proteases also can be described according to their chemical site of action (e.g., cysteine, serine, or threonine proteases). Peptidases exhibit varying degrees of substrate specificity, with few exhibiting high selectivity (e.g., the thyrotropin-releasing hormone-degrading ectoenzyme specifically inactivates thyrotropin-releasing hormone). Many peptidases, such as endopeptidase 24.11, angiotensin-converting enzyme, and pyroglutamate aminopeptidase II, are synaptosomally located, poised to limit neurotransmission by inactivating peptides released into the synaptic cleft. Indeed, peptidase release appears to be induced by neuronal depolarization. Other peptidases, however, such as aminopeptidases A and M, dipeptidyl aminopeptidase IV, and carboxypeptidase, appear to be located in the microvasculature and to be situated to limit volume transmission, or signaling occurring over a great distance, as well as to limit passage across the blood–brain barrier. A third class of membrane-associated cytoplasmic peptidases in the neural and glial tissue, for example, pyroglutamyl peptidase I, also may help to inactivate the released peptide. Finally, receptor-mediated endocytosis, with subsequent lysosomal degradation, is an additional route of limiting peptide activity. Each peptide is subject to inactivation by a variety of peptidases, with each reaction yielding different metabolic products. Because many of these metabolic products can act as agonists

or antagonists (in a manner sometimes functionally distinct from those of the parent peptide – for example, the conversion of peptide YY to peptide YY_{3–36}), the physiological control of bioavailable peptidases is an additional important determinant of peptide action. Protease activity is regulated *in vivo* by endogenous inhibitors, zymogen activation, and the modulation of their synthesis or degradation. Because of peptidases, natural peptides generally have very short half-lives (on the order of minutes) *in vivo*. Peptides with deletions, unnatural amino acids, or substitutions that eliminate a natural cleavage site can result in compounds with greater metabolic stability. Depending on the specific intrinsic activity of the analog, it could then function as a long-acting antagonist (e.g., astressin-B compared to corticotropin-releasing factor), an agonist (exendin-4 compared to glucagon-like peptide-1) or an inhibitory superagonist ([D-Trp⁶, Pro⁹-NET]-GnRH compared to GnRH). Similarly, inhibitors of peptidase action can augment or prolong peptide action. The MEROPS database maintains an integrated list of known peptidases and peptidase inhibitors classified according to their structural similarity or catalytic site of action.

Peptide-Binding Proteins

Peptide-binding proteins are soluble or membrane-bound proteins to which peptides bind in lieu of their cognate receptor. Often, they act to sequester peptide ligands, thereby limiting their bioavailability (e.g., CRF-binding protein). Some binding proteins (e.g., insulin-like growth factor-binding protein), however, potentiate peptide action by protecting peptides from extracellular peptidases and forming a bound complex that associates with the membrane, thereby facilitating peptide–receptor interaction. Binding proteins are physiologically regulated through enzymatic modifications of their affinity for the peptide ligand or through changes in synthesis and expression, with consequent effects on peptidergic activity.

Acknowledgments

Supported by NIH grants DK64871 and DK26741 from the National Institute of Diabetes and Digestive and Kidney Diseases. This is publication number 17220-NP from The Scripps Research Institute.

See Also the Following Articles

Beta-Endorphin; Corticotropin Releasing Factor (CRF); Endocrine Systems; Gastrointestinal Effects; Hypothalamic-Pituitary-Adrenal; Anatomy of the HPA Axis; Annexin; Opioids; Oxytocin; Pro-opiomelanocortin (POMC); Pituitary Regulation, Role of; Protein Synthesis; Secretagogue; Urocortins; Vasoactive Peptides; Vasopressin.

Further Reading

- Crawley, J. N. and McLean, S. (eds.) (1996). Preface. In: *Special issue on neuropeptides: basic and clinical advances, Annals of the New York Academy of Sciences* 780, ix–xi.
- Guillemin, R. (2005). Hypothalamic hormones a.k.a. hypothalamic releasing factors. *Journal of Endocrinology* 184, 11–28.
- Harmar, A. J. (2004). Receptors for gut peptides. *Best Practice & Research: Clinical Endocrinology & Metabolism* 18, 463–475.
- Holly, J. (2004). Physiology of the IGF system. *Novartis Foundation Symposiums* 262, 19–26.
- Hökfelt, T., Broberger, C., Xu, Z.-Q., et al. (2000). Neuropeptides – an overview. *Neuropharmacology* 39, 1337–1356.
- Rawlings, N. D., Tolle, D. P. and Barrett, A. J. (2004). MEROPS: the peptidase database. *Nucleic Acids Research* 32, D160–D164.

Relevant Websites

- International Union of Pharmacology Basic and Clinical Pharmacology Committee on Receptor Nomenclature and Drug Classification database. <http://www.iuphar-db.org>.
- MEROPS database. <http://merops.sanger.ac.uk>.

Perinatal (Fetal) Stress/Barker's Hypothesis See: Fetal Stress.

Perinatal Dexamethasone

E Fuchs

German Primate Center and Medical School, University of Göttingen, Göttingen, Germany

© 2007 Elsevier Inc. All rights reserved.

Early Life Programming

Dexamethasone: Actions and Potencies

Placental Metabolism of Glucocorticoids

Dexamethasone and Fetal Development

Metabolic Programming

Effects on Brain, Endocrine, and Immune Functioning and Behavior

Intergenerational Effects of Dexamethasone

The Underlying Mechanism: A Permanent Alteration of Gene Expression?

Conclusions

Glossary

Dexamethasone
Early life programming

A synthetic fluorinated glucocorticoid hormone structurally related to cortisol. Programming reflects the action of a factor during a sensitive developmental period, or time window, to affect the development and organization of specific tissues that are concurrently vulnerable, producing effects that persist throughout life. Different cells and tissues are sensitive at different times, so the effects of environmental challenges will have distinct effects, depending not only on the challenge involved but also on its timing.

Early Life Programming

Lifestyle and genetic predisposition are considered to be the main determinants of noncommunicable diseases in humans. However, increasing evidence from studies in several distinct human populations worldwide indicates that many disorders occurring in adult life date back to influences of the intrauterine environment during the prenatal period. Examples of illnesses for which there is evidence of fetal origin are hypertension, coronary heart disease, diabetes, and neurological disorders. Behavioral disorders during childhood and adolescence may also be programmed during the prenatal period. As it is generally accepted that a multitude of pre- and postnatal influences are involved in pathophysiological processes that predispose individuals to such diseases, it is plausible to propose that there are factors triggering fetal

development, thus amplifying or minimizing certain long-term processes. One identified trigger is hormonal disturbance such as high glucocorticoid exposure during pregnancy. However, most of the studies in humans have been restricted to either epidemiological or some physiological measurements that in sum cannot explain the cause and basic mechanisms underlying the pathophysiological processes.

A high percentage of pregnant women with an increased risk of preterm delivery are treated with glucocorticoids for enhancement of fetal lung maturation before term. The underlying mechanism for why this treatment is used in fetal pulmonary development is the ability of glucocorticoids to turn off cell proliferation and initiate mature cell function. In addition to being used in cases of preterm labor, synthetic glucocorticoids are administered to pregnant women in other clinical situations, including congenital adrenal hyperplasia and a wide spectrum of autoimmune disorders such as multiple sclerosis, asthma, rheumatic diseases, or chronic inflammatory intestinal diseases (M. Crohn, colitis ulcerosa). Often these treatments occur in early pregnancy, when it has not yet been diagnosed, and in some cases throughout gestation. There is currently little information about the impact of such treatment regimes on organ development. Epidemiological data indicate that fetuses exposed to synthetic glucocorticoids such as dexamethasone display significant increases in emotionality, unsociability, and behavioral problems in later life as children. However, further resolution of the safety of these treatment regimes is urgently needed. As a result, two National Institutes of Health (NIH) consensus reports published in 1995 and 2001 suggested that repeated or prolonged treatment should be avoided during pregnancy and highlighted the requirement for further research on the effects of prenatal glucocorticoids on organ development and function after birth.

Dexamethasone: Actions and Potencies

Synthetic glucocorticoids were developed by modifying various sites on the cortisol molecule in order to, for example, enhance the anti-inflammatory and reduce the Na^+ -retaining mineralocorticoid activity. Among synthetic agents, the fluorinated steroid dexamethasone (DEX) has the lowest mineralocorticoid effect; it is characterized by a very long half-life, with the consequence of a greater potential to cause glucocorticoid-related side effects. DEX can be administered orally and, as with other steroids, is rapidly absorbed from the jejunum with high bioavailability,

Table 1 Comparative potencies of cortisol and dexamethasone

	Equivalent dose (mg)	Plasma half-life (min)	Duration of action (h)	Mineralocorticoid potency	Glucocorticoid potency	Relative placental inactivation (% of drug inactivated)
Cortisol	20	90	8–12	1	1	68
Dexamethasone	0.5	240	36–72	0	30	2

From Lin and Paget (2002).

meaning that up to 90% of an orally administered dose appears in the circulation, although bioavailability after oral administration is less than after parenteral administration because of substantial first-pass liver metabolism (Table 1).

DEX is used in the differential diagnosis of hypercortisolemia as it occurs, for example, in Cushing's disease, atypical pituitary adenomas, or depression (see **Hypocortisolism and Stress**). In this context, it is interesting to note that at least the adult rodent brain appears to be protected against high amounts of DEX by a multidrug-resistant P-glycoprotein located in the blood–brain barrier.

Placental Metabolism of Glucocorticoids

Transplacental passage of the different corticosteroid molecules varies, so different substances have to be used for maternal than for fetal targets. Dexamethasone and betamethasone, a second member of the clinically used fluorinated glucocorticoids, cross the placenta more readily than the natural glucocorticoid cortisol does because they are not bound to transporter proteins in the maternal blood, nor are they as easily converted to cortisone by placental enzymes such as 11 β -hydroxysteroid dehydrogenase (11 β -HSD). The relative inactivation of various corticosteroids to their 11-ketometabolites, through action of placental 11 β -HSD, ranges from 2% (DEX) to 68% (cortisol; see Table 1). Within the placenta, two isoforms of 11 β -HSD exist: a low-affinity NADP-dependent dehydrogenase/oxoreductase and a high affinity NAD-dependent dehydrogenase. Expression and distribution of these isoforms within the placenta most likely have an impact on levels of active glucocorticoids in fetal blood so that a complex interplay of factors determines the degree of fetal exposition.

Dexamethasone and Fetal Development

The assumption that excessive fetal exposure to glucocorticoid hormones constitutes an origin of adult diseases is based on studies in experimental animals. Findings in laboratory rodents (mice, rats, guinea pigs) concur with observations in ruminants and monkeys, mainly rhesus monkeys, documenting many lasting endocrine, neural, and behavioral effects when

immature animals are exposed to steroidal hormones early in life. To model perinatal hyperexposure to glucocorticoids under controlled conditions, many research groups have studied the *in utero* effects of DEX, whose cellular effects are mediated via the glucocorticoid receptors (GRs) that are present in nearly all cells of the body (see **Glucocorticoids, Effects of Stress on**). It is known that synergistic and/or antagonistic interactions between activated GRs and several transcription factors regulate gene expression. Many organs, such as the brain, and the cardiovascular, skeletal, immune, reproductive, and urogenital systems contain cells that are vulnerable to changes in glucocorticoid homeostasis, and it is very likely that excess antenatal glucocorticoids such as DEX have multiple pathophysiological effects. The developmental consequences of prenatal DEX treatment are summarized in Figure 1.

Animal studies have revealed that, via the effects on gene transcription, glucocorticoids exert well-documented prenatal programming actions in a variety of target tissues to induce effects that may persist throughout the life span. For example, the adult rat offspring of pregnancies treated with DEX exhibited lifelong elevated blood pressure and high levels of plasma glucose, insulin, and corticosterone, as well as altered behavior. Variations in the effectiveness of the fetoplacental enzymatic barrier to maternal glucocorticoids because of changes in 11 β -hydroxysteroid dehydrogenase type 2 (11 β -HSD2), which catalyzes the rapid inactivation of cortisol (or corticosterone in rats) to inert cortisone (11-dehydrocorticosterone), correlate directly with birth weight in rats and humans. Thus, a deficiency of the placental barrier with respect to inactivation of maternal glucocorticoids, or overexposure to the hormones, may lead to increased effects of the hormones on gene expression in the fetal cells, and thus to prenatal programming. This idea is supported by the finding that, comparable to the effects of DEX, inhibition of 11 β -HSD in pregnant rats by treating them with carbenoxolone causes reduced birth weight and hypertension later in life. Interestingly, the effects of carbenoxolone require the presence of endogenous maternal glucocorticoids: the offspring of adrenalectomized pregnant rats were found to be protected from the effects of carbenoxolone with respect to birth weight and adult physiology. However, it has

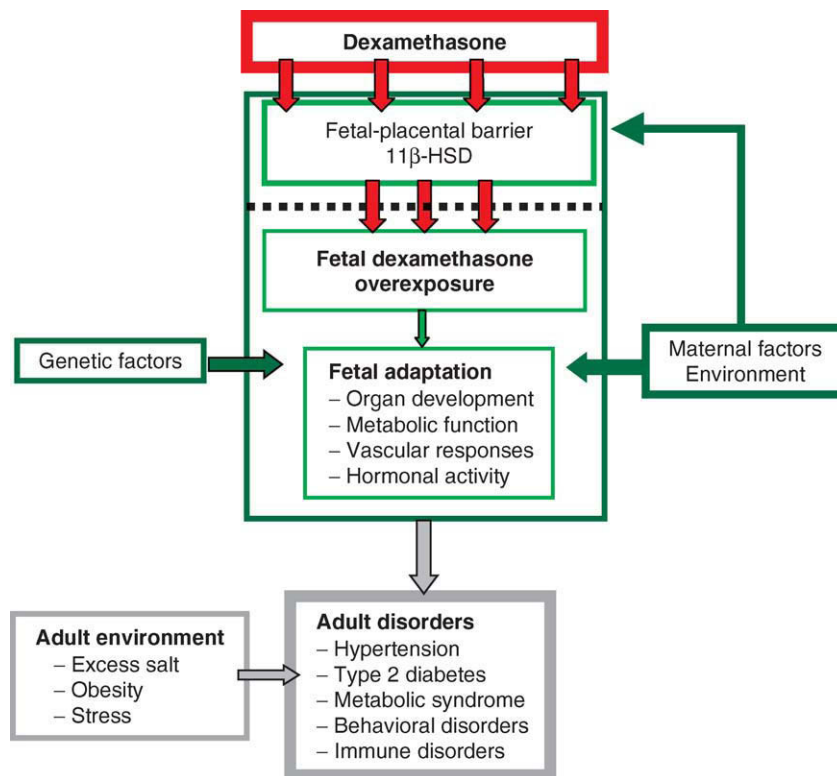


Figure 1 Developmental consequences of prenatal dexamethasone treatment. Fetal dexamethasone overexposure in combination with genetic and environmental factors will influence the development of many organ systems and programs tissue responses in the developing offspring, resulting in an increased risk for pathophysiological conditions later in life.

to be considered that carbenoxolone is a nonselective compound that inhibits both 11β -HSD isoforms as well as related dehydrogenases.

The timing of antenatal glucocorticoid exposure is important, in part because of the complex species-specific patterns of expression of the 11β -HSD isoenzymes that regulate maternal glucocorticoid transfer to the fetus. In rats, exposure to glucocorticoids during the last week of pregnancy is sufficient to produce permanent adult hypertension, whereas in sheep, the time window of sensitivity for such effects is earlier in gestation.

Mice and rats have been regarded as corticoid-sensitive species because their newborn pups are extremely immature at birth. Therefore, the relevance of rodent models of prenatal glucocorticoid overexposure has sometimes been questioned. However, similar types of adverse effects have also been observed in sheep, including an influence of physical growth and suppression of maternal and fetal adrenal activity. These glucocorticoid effects in sheep are surprising because high maternal cortisol levels play an important role in the normal physiology of their pregnancy and parturition. It is possible that in certain species the fetus is typically protected from excessive transplacental diffusion of cortisol/corticosterone through the enzyme 11β -HSD.

In monkeys, the natural ligand cortisol is metabolized even more rapidly in the fetus than in the full-term neonate. Conversely, DEX takes considerably longer to clear from the circulation. This extended half-life may be important when DEX is administered to the mother while the baby is still *in utero* and thus depends on the transfer of steroid hormones back across the placenta, rather than on its own renal excretion after birth.

Unfortunately, not enough animal studies have been conducted with betamethasone to permit a systematic comparison with DEX. In general, the research in humans tends to find more side effects with DEX than with betamethasone.

Metabolic Programming

The preceding findings imply that glucocorticoids exert organizational effects on the immature system during defined developmental stages, leading to changes in physiology that have persistent consequences throughout the lifetime. Because reduced prenatal growth and low birth weight are determined largely by nongenetic factors, these findings have led to the fetal origin hypothesis. It proposes that the physiology of the fetus adapts to the adverse intrauterine environment, which in turn reduces fetal growth

and programs the lifelong changes. This programming predisposes an individual to the metabolic syndrome that manifests itself only in adulthood. It emanates from a general disturbance of the body's metabolic processes and leads to heart disease, diabetes, and kidney failure. These illnesses have long been observed as being somehow linked to modern civilization, but the fact that they constitute an entire syndrome has been recognized only recently.

Blood Pressure

The impact of prenatal glucocorticoid exposure during development on arterial blood pressure has been investigated in the offspring of several species. Treatment of pregnant rats with DEX results in persistent elevations of arterial blood pressure in the adult offspring. Maternal administration of the 11 β -HSD inhibitor carbenoxolone leads to elevated blood pressure in the adult offspring. As previously mentioned, this effect requires the presence of maternal glucocorticoids, as the offspring of adrenalectomized pregnant rats are protected from the adverse actions of carbenoxolone.

The mechanisms of glucocorticoid-programmed adult hypertension are not clearly understood but probably involve a variety of processes. Potential factors may involve primary effects on the heart by interfering with the development of cardiac noradrenergic innervation, modifications in vascular responsiveness, alterations in renal structure (irreversible reductions in nephron number), and enhanced activity of the renin-angiotensin system.

Glucose Homeostasis

Prenatal overexposure to glucocorticoids also causes permanent hyperglycemia and hyperinsulinemia in the adult offspring of rats and sheep. As with blood pressure, the window of sensitivity for programming hyperglycemia in the rat occurs in the last third of gestation. Earlier DEX exposure or *post partum* treatments do not program hyperglycemia/hyperinsulinemia, suggesting that there is a narrow window for this effect. The mechanisms through which prenatal superphysiological levels of glucocorticoids program hyperglycemia have not been fully determined but may involve derangements in several target organs such as the liver, the pancreas, fat, and skeletal muscle.

Glucocorticoids regulate the expression of several hepatic metabolic enzymes, including phosphoenolpyruvate carboxykinase (PEPCK). This enzyme develops in late gestation and catalyzes a rate-limiting step in gluconeogenesis. In rats, exposure to excess DEX *in utero* leads to offspring with permanent elevations in PEPCK mRNA and enzyme activity. Interestingly the PEPCK activity is increased in patients with type 2 diabetes.

Relatively little is known about the effects of prenatal exposure to glucocorticoids on the development of the endocrine functions of the pancreas. However, indirect evidence suggests that excessive prenatal glucocorticoid levels may cause long-term beta cell dysfunction.

Prenatal DEX exposure in rats programs fat metabolism. Adult offspring of dams exposed to DEX in pregnancy have increased intra-abdominal fat depots, with a parallel increase in leptin levels. The same antenatal treatment may also affect skeletal muscle glucose metabolism.

Effects on Brain, Endocrine, and Immune Functioning and Behavior

In addition to their metabolic role, glucocorticoids are essential for normal brain development. Glucocorticoids are important for normal maturation in most regions of the developing central nervous system, initiating terminal maturation and remodeling axons and dendrites, and for cell survival.

Exposure to glucocorticoids *in utero* has widespread acute effects on neuronal structure and synapse formation and may permanently alter brain structure. Prenatal glucocorticoid administration retards brain weight at birth in sheep, delaying maturation of neurons, myelination, glia, and vasculature. Rhesus monkeys that were prenatally exposed to DEX exhibited marked atrophy of hippocampal cells at birth, and their hippocampi failed to recover to normal size, even at 2 years *post partum*. Human and animal studies have demonstrated that altered hippocampal structure may be associated with memory impairment and behavioral alterations (see **Hippocampus, Corticosteroid Effects on**).

Overall, the development of the hippocampus in rhesus monkeys is comparable to that in humans, although it is difficult to specifically compare the rate of fetal growth and stage of maturity at birth. However, in humans, important phases in maturation of the brain, including brain growth, take place during the postpartum period. At birth, the brain of a rhesus monkey is 60% of its adult size, whereas the human brain is only 24% of its adult size. Thus, the impact of prenatal DEX overexposure on the developing nervous system may be of a greater magnitude in the monkey.

The hypothalamic-pituitary-adrenocortical (HPA) circuit, which is central for the integration of an individual's endocrine and behavioral responses and for the whole body's homeostasis, appears highly sensitive to excess glucocorticoids during development. A number of animal studies have shown that exposure to synthetic glucocorticoids *in utero* results in disturbed hypothalamic-pituitary circuit functioning.

Studies in rats provided evidence that perinatal DEX treatment enhances susceptibility to experimental autoimmune encephalitis (EAE) in adult life. EAE is an experimental autoimmune inflammatory disease that serves as an animal model for multiple sclerosis. Low reactivity of the HPA axis and increased Th1-type cytokine production are discussed as potential mechanisms for this increased susceptibility in adult rats after DEX treatment. Furthermore, perinatally DEX-treated rats have a long-term decreased production of cytokines by macrophages. This may lead to increased susceptibility to bacterial infections in later life. Apparently, perinatal DEX treatment has permanent programming effects on endocrine as well as on immune functioning in adult life.

As revealed mainly by studies in rats, overexposure to glucocorticoids *in utero* leads to alterations in adult behavior. Late gestational DEX in rats apparently impairs coping in aversive situations later in life. Behavioral changes in adult rats exposed prenatally to glucocorticoids may be associated with altered functioning of the amygdala, a key brain structure in the regulation of fear and anxiety. Interestingly, in rats, adult offspring of DEX-treated dams show increased corticotropin-releasing hormone levels, specifically in the central nucleus of the amygdala.

Intergenerational Effects of Dexamethasone

The effects of fetal programming may not be limited to the first-generation offspring. A recent study investigated epigenetic effects in the DEX-programmed rat, a model in which fetal exposure to excess corticosteroids results in low birth weight, adult glucose intolerance, and increased activity of the key hepatic gluconeogenic enzyme PEPCK. Male offspring of female rats that had been exposed prenatally to DEX, but that were not treated during their pregnancy, also had reduced birth weight, glucose intolerance, and elevated hepatic PEPCK activity. These effects resolved in a third generation. Similar lifelong programming was observed in the offspring of male rats exposed prenatally to DEX mated with control females. The persistence of such programming effects across several generations, transmitted by either maternal or paternal lines, indicates – at least in rodents – the potential importance of non-genomic factors in the intergenerational inheritance of the programming phenotype.

The Underlying Mechanism: A Permanent Alteration of Gene Expression?

The GR is critical for cell function, and GR gene expression shows marked tissue-specific regulation. In the peripheral tissues, such as the liver and visceral

adipose tissue, prenatal glucocorticoid exposure causes a permanent increase in GR expression; this upregulation is paralleled by increased glucocorticoid sensitivity in these tissues.

Until recently it was unclear how perinatal glucocorticoid overexposure could permanently alter gene expression. Recent evidence suggests that this priming may involve selective methylation/demethylation processes of specific promoters of the GR gene. Changes in the GR promoter DNA methylation pattern were associated with altered histone acetylation and transcription factor (NGFI-A) binding to the GR promoter. Treatment of the adult rat offspring with a histone deacetylase inhibitor removed the epigenetic differences in histone acetylation and DNA methylation and changes in NGFI-A binding. During development, such target promoter demethylation occurs before birth and may tune the promoter to remember regulatory events occurring during development. This novel mechanism of gene control by environmental events early in life, which then persists throughout the life span, needs to be confirmed in species other than rodents.

Conclusions

Balanced levels of glucocorticoids are of central importance for normal development, and it can be assumed that glucocorticoid overexposure during critical periods of fetal development may have long-term effects that could be deleterious later in life. To date, most of the studies in humans have been restricted to either epidemiological or some physiological measurements that in sum cannot explain the cause and basic mechanisms underlying the pathophysiological processes. In contrast, animal studies have indicated a link between glucocorticoid overexposure and an increased risk for pathophysiological conditions later in life. Until now, experimental studies on the mechanisms induced by prenatal hyperexposure to glucocorticoids such as DEX have been conducted predominantly in rodents. However, rodents are far from ideal models of primate pregnancy due to their multiple offspring and short gestation lengths and life span. Clearly, mechanistic studies need to be extended to species more closely related to humans.

See Also the Following Articles

Hypocortisolism and Stress; Hippocampus, Corticosteroid Effects on; Glucocorticoids, Effects of Stress on.

Further Reading

Aghajafari, F., Murphy, K., Matthews, S., et al. (2002). Repeated doses of antenatal corticosteroids in animals:

- a systematic review. *American Journal of Obstetrics and Gynecology* 186, 843–849.
- Bakker, J. M., Kavelaars, A., Kamphuis, P. J., et al. (2000). Neonatal dexamethasone treatment increases susceptibility to experimental autoimmune disease in adult rats. *Journal of Immunology* 165, 5932–5937.
- Bakker, J. M., van Bel, F. and Heijnen, C. J. (2001). Neonatal glucocorticoids and the developing brain: short-term treatment with life-long consequences? *Trends in Neuroscience* 24, 649–653.
- Berger, B., Alvarez, C. and Goldman-Rakic, P. S. (1993). Neurochemical development of the hippocampal region in the fetal rhesus monkey. I. Early appearance of peptides, calcium-binding proteins, DARPP-32, and monoamine innervation in the entorhinal cortex during the first half of gestation (E47–90). *Hippocampus* 3, 279–305.
- Coe, C. L. and Lubach, G. R. (2005). Developmental consequences of antenatal dexamethasone treatment in nonhuman primates. *Neuroscience and Biobehavioral Reviews* 29, 227–235.
- Drake, A. J., Walker, B. R. and Seckl, J. R. (2005). Inter-generational consequences of fetal programming by *in utero* exposure to glucocorticoids in rats. *American Journal of Physiology: Regulatory, Integrative, and Comparative Physiology* 288, R34–R38.
- Kay, H. H., Bird, I. M., Coe, C. L. and Dudley, D. J. (2000). Antenatal steroid treatment and adverse fetal effects. What is the evidence? *Journal of Society for Gynecologic Investigation* 7, 269–278.
- Krozowski, Z., Li, K. X., Koyama, K., et al. (1999). The type I and type II 11 β -hydroxysteroid dehydrogenase enzymes. *Journal of Steroid Biochemistry and Molecular Biology* 69, 391–401.
- Lin, A. N. and Paget, S. A. (eds.) (2002). *Principles of corticosteroid therapy*. London: Arnold.
- Matthews, S. G. (2000). Antenatal glucocorticoids and programming of the developing CNS. *Pediatric Research* 47, 291–300.
- Meijer, O. C., de Lange, E. C., Breimer, D. D., et al. (1998). Penetration of dexamethasone into brain glucocorticoid targets is enhanced in *mdr1A* P-glycoprotein knockout mice. *Endocrinology* 139, 1789–1793.
- National Institutes of Health Consensus Development Conference Statement. August 17–18, 2000 (2001). Antenatal corticosteroids revisited: repeat courses. *Obstetrics and Gynecology* 98, 144–150.
- NIH Consensus Developmental Panel on the Effect of Corticosteroids for Fetal Maturation on Perinatal Outcomes. (1995). Effect of corticosteroids for fetal maturation on perinatal outcomes. *Journal of the American Medical Association* 273, 413–418.
- Seckl, J. R. (2004). Prenatal glucocorticoids and long-term programming. *European Journal of Endocrinology* 151, 1–16.
- Uno, H., Eisele, S., Sakai, A., et al. (1994). Neurotoxicity of glucocorticoids in the primate brain. *Hormones and Behavior* 28, 336–348.

Perinatal Stress Animal Studies See: Fetal Stress.

Persian Gulf War, Stress Effects of

**S M Southwick, D Vojvoda,
C A Morgan III and D Lipschitz**

Yale University School of Medicine, New Haven,
CT, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3,
pp 142–148, © 2000 Elsevier Inc.

Health Risks of the Gulf War

Mental Health

Predictors of Posttraumatic Stress Disorder

Physical Symptoms in Gulf War Veterans

Relationship between Physical and Psychological
Complaints

Summary

Glossary

*Posttraumatic
stress disorder
(PTSD)*

This disorder results from situations in which an individual believes that he or she or someone else is in danger of being seriously injured or killed. Symptoms of PTSD include repetitive re-experiencing of the trauma in the form of intrusive memories, nightmares, and flashbacks; persistent avoidance of stimuli reminiscent of the trauma; a general numbing of emotions; and increased arousal as expressed by hypervigilance, exaggerated startle, irritability, and insomnia.

Health Risks of the Gulf War

In 1990 the United States and allied countries responded to the Iraqi invasion of Kuwait by deploying

a large number of soldiers to the Persian Gulf. Operation Desert Storm began in 1991 when Allied forces launched a 5-week intensive air attack on Iraq. This was followed by a 4-day ground war. The number of deaths to American and Allied forces during the Gulf war was limited compared to World War I and II, the Korean war, and the Vietnam war. Nevertheless, since the end of the Gulf war, a substantial number of Desert Storm veterans have experienced a host of physical and psychological symptoms that, in some cases, have had a pronounced impact on psychosocial functioning.

Military personnel stationed in the Gulf were exposed to a wide variety of potential health risks. These included a hostile environment (e.g., extreme temperatures, humidity and rainfall, blowing sand, crowded living conditions), infectious agents (e.g., leishmaniasis, sandfly fever), medical prophylaxis (e.g., pyridostigmine bromide, anthrax, and botulinum toxin vaccines), occupational exposures (e.g., chemical exposures associated with routine maintenance and repair of equipment, fuels), environmental hazards (e.g., pesticides, smoke from oil well fires), depleted uranium (used in some artillery shells), the possibility of chemical and biological warfare agents (e.g., organophosphate nerve agents, sulfur mustard, anthrax, botulinus toxin), and repetitive physiological and psychological stress.

Soldiers serving in the Persian Gulf were exposed to numerous potential psychosocial stressors. Stressors associated with deployment included uncertainty about parameters of deployment (e.g., the period of waiting for deployment, unknown length of tour and return date), overestimation of Iraqi military forces, media anticipation of high casualty rates, and separation from home and family. As in all wars, soldiers described the anticipation of combat as being highly stressful. In a study of 1524 deployed military personnel, 80% of subjects reported feeling stressed by the persistent threat of SCUD missile attacks, terrorist attacks, and biological and chemical warfare. Many experienced strong emotional responses to repeated missile warnings and the rapid donning of protective gear that they did not fully trust would be effective against all possible toxic agents.

Reports of directly experienced traumatic events vary depending on the units studied. A series of scientific investigations commissioned by the U.S. Department of Defense found that 40% of soldiers in the units studied experienced incoming indirect fire, between 40 and 70% reported exposure to mines and booby traps, 40% in two units recalled firing rounds at the enemy, and 80% in one unit and 60% in another claimed to see enemy killed or wounded in action. In two units more than 20% of soldiers believed they were in imminent danger of being killed at some point during their tour in the Persian Gulf.

Reports of seeing an American wounded by friendly fire had great emotional impact, as did work in grave registration units. Sadly, some traumas experienced in the Gulf, such as sexual harassment and assault, were inflicted by Americans on their own fellow soldiers.

Other frequently cited Gulf war stressors include austere physical environment and severe climatic conditions, heavy workload, sleep deprivation, inadequate supplies and equipment, lack of information, poor leadership and unit morale, and unsatisfactory living conditions.

Mental Health

The psychological consequences of war have been described for centuries and have been referred to by a variety of different terms, including shell shock, irritable heart of soldiers, physioneurosis, and, most recently, posttraumatic stress disorder (PTSD). PTSD results from situations in which the individual believes that he or she or someone else is in danger of being seriously injured or killed. Symptoms of PTSD include repetitive re-experiencing of the trauma in the form of intrusive memories, nightmares, and flashbacks; persistent avoidance of stimuli reminiscent of the trauma; a general numbing of emotions; and increased arousal as expressed by hypervigilance, exaggerated startle, irritability, and insomnia. Depression, anxiety, substance use, and changes in personality traits are other frequently cited consequences of trauma.

Psychological Symptoms in Persian Gulf Veterans

Prior to the ground war it was estimated that acute psychiatric casualties might reach as high as 50 000. This estimate was based on experience from previous wars. However, during the Gulf war there appeared to be very few psychiatric casualties. Between August 1990 and August 1991, 476 soldiers were evacuated from southwest Asia for psychiatric reasons. This number constituted only 6.5% of all medical evacuations from the Persian Gulf. Similarly, only 158 soldiers from the U.S. Army 7th Corps received mental health treatment while stationed in the Gulf, and 99% were returned to duty after brief treatment. These low rates led to the belief that psychiatric problems in Gulf veterans would be lower than expected. Explanations for low rates included the brief nature and success of the war, limited exposure to traditional combat stressors, and high levels of perceived social support.

Since the end of the Gulf war, however, a number of different research groups have reported rates of psychiatric symptoms and disorders in Persian Gulf veterans that are high in light of the earlier evacuation

figures. The Fort Devons Operation Desert Storm Reunion Survey was conducted within 5 days of return to the United States before soldiers rejoined their families. The most commonly endorsed war zone stressors were formal alert for chemical or biological attack (approximately 75%), receiving incoming fire from large arms (approximately 73%), and witnessing death or disfigurement of enemy troops (approximately 48%). Four percent of 2136 men and 9% of 208 females met a presumptive diagnosis of PTSD. Symptoms of hyperarousal and exaggerated startle were the most commonly endorsed PTSD symptoms. Twenty-eight to 31% had general severity index (GSI) scores on the brief symptom inventory (BSI), which reflects general psychological stress, in the clinically significant range.

Approximately 6 months after Operation Desert Storm, Labatte and Snow administered a self-report questionnaire to 57 soldiers stationed in Germany who were members of a single mechanized infantry unit that suffered four casualties during the brief ground war. Fourteen soldiers (25%) reported traumatic nightmares within 2 months of the war. Nightmares began 2 to 6 months after the war for an additional nine soldiers (16%).

Two other studies assessed PTSD symptoms and general psychopathology in a large number of reservists within the first year of their return from the Persian Gulf. In a study of 591 Army, Navy, and Marine reservists, Perconte and co-workers found that those who had been deployed to the Gulf war combat zone had significantly higher Mississippi PTSD scores, higher Beck depression scores, and higher levels of global distress as measured by the symptom check list-90 (SCL-90) than reservists who had been activated but not deployed to combat zones. Over 15% of deployed reservists met Mississippi cutoff scores for PTSD compared to less than 4% of nondeployed reservists. Sutker and colleagues reported low levels of psychological distress in their overall group of 775 combat-deployed Reserve and National Guard troops. However, 12.5% of the sample were diagnosed with PTSD.

Approximately 1 year after the war, Holmes and colleagues evaluated 179 members of an Air National Guard unit who were never deployed to the Persian Gulf and compared them to 296 members of the same unit who were deployed. Most members of the deployed group did not experience any traditional combat stressors, although they were exposed to non-traditional combat stressors such as anticipation of missile attacks and biological warfare. Compared to the nondeployed, the deployed group had statistically higher mean scores on the Mississippi PTSD scale and on the GSI as well as on the obsessive-compulsive, depression, hostility, phobic anxiety, and paranoia

subscales of the BSI; 6.8% of deployed and 1.7% of nondeployed were identified as having increased risk for PTSD. The findings of this study suggest that multiple noncombat stressors were distressing enough to cause both PTSD and non-PTSD psychological symptoms in some deployed veterans.

In a follow-up study of the Fort Devons sample 18–24 months after the war, nearly half of the sample reported no physical or psychological symptoms. Twenty-seven percent believed that their psychological health had deteriorated since the war. PTSD symptomatology was significantly associated with high health symptom endorsement. Twenty-two percent of soldiers with high combat exposure met a presumptive diagnosis of PTSD compared to 8.1% with low combat exposure. Data were consistent with a number of recent studies suggesting that an apparent association between combat exposure and health is significantly mediated by a diagnosis of PTSD.

In a survey administered approximately 2 years after the war to 1524 deployed and 2727 nondeployed military personnel from Hawaii and Pennsylvania, Stretch and co-workers reported possible PTSD in 8.0% of active duty veterans and in 9.35% of reserve veterans who had been deployed to the Persian Gulf. By comparison, 1.3% of active duty nondeployed and 2.1% of reservist nondeployed met possible criteria for PTSD. The sample was also compared to a large group of deployed active duty Army veterans from the XVIIIth and VIIth Corps. The possible risk for PTSD in these groups was 15.2 and 12.9%, respectively. The authors speculate that these units had higher rates of possible PTSD because their duty was more dangerous and because they were evaluated 1 year earlier. Deployed personnel also scored significantly higher on the somatization, obsessive-compulsive, and hostility subscales of the BSI and the GSI.

As part of a follow along study of Gulf veterans from two National Guard units, Southwick and colleagues reported a rate of presumptive PTSD in 8.3% of 84 subjects 6 months after the war and in 10% of 64 subjects 2 years after the war. The greatest increase in overall PTSD symptoms occurred between 1 month and 6 months after the war, and symptoms of hyperarousal were elevated above re-experiencing and avoidance symptoms at 1-month, 6-month, and 2-year evaluations. Morgan and colleagues found evidence for anniversary reactions in this same group of veterans.

Approximately 5 years after the war, the Iowa Persian Gulf study group conducted telephone interviews with 1896 Gulf war veterans and 1799 personnel who served in the military during the war but who were not stationed in the Persian Gulf. Persian Gulf war personnel reported higher rates of PTSD (1.9%

vs. 0.8%), depression (17.0% vs. 10.9%), cognitive dysfunction (18.7% vs. 7.6%), asthma (7.2% vs. 4.1%), fibromyalgia (19.2% vs. 9.6%), alcohol abuse (17.4% vs. 12.6%), chronic fatigue (1.3% vs. 0.3%), and sexual dysfunction compared to non-Persian Gulf war military personnel. Symptom levels were generally higher in Persian Gulf war National Guard reserve personnel than in Persian Gulf war regular military personnel.

Finally, in a report of 18 495 Gulf war veterans who presented for evaluation to the Department of Defense Comprehensive Clinical Evaluation Program, 36% met ICD-9-CM diagnoses for psychological conditions. In most cases, symptoms did not develop until after the war. There was no apparent association between individual physical or psychological symptoms and specific exposures. Compared with a referent group of patients having no symptoms, an excess prevalence of disorders was highest for psychological conditions compared to a variety of physical conditions such as digestive and skin conditions.

Psychological Symptoms in Civilians Affected by the Persian Gulf War

Clearly the Gulf war had profound stress-related effects on Iraqi military personnel and civilians living in Iraq and surrounding Middle Eastern countries. In general, the rates of death, physical injury, and psychological injury have not been reported accurately in the scientific literature. However, there have been a number of publications related to the effects of the Gulf war on American and Israeli civilians as well as Vietnam veterans and Holocaust survivors.

In American civilians, it has been reported that the Gulf war was associated with mild to moderate symptoms of anxiety and depression, particularly in those who had a relationship with a deployed soldier. In children of deployed American soldiers, a major predictor of psychological symptoms was level of symptoms in other household members. Although deployment rarely provoked pathological levels of symptoms in healthy children, it was associated with increases in symptoms of depression, especially among younger children. In a study of Israeli civilians that was conducted during the war, Solomon and co-workers found that 80% of individuals whose homes had been destroyed by Iraqi missiles described symptoms consistent with DSM-III-R criteria for PTSD. Similarly, Laor and colleagues reported elevated levels of Gulf war stress-related symptoms in children and mothers who were displaced from their homes due to damage from SCUD missiles and in young children living in families characterized by inadequate cohesion.

In Vietnam veterans, no consistent response to the Gulf war has been described. Reported responses have included support for the war, anger at the U.S.

government for starting another war, irritability, intrusive thoughts, and increased depression and suicidality. In a study of 76 female Vietnam veterans with PTSD, Wolfe and colleagues reported some exacerbation of symptoms in most subjects, with the greatest increases in those who had high levels of preexisting PTSD symptoms. In a study of Holocaust survivors living in Israel during the Gulf war, Robinson and colleagues reported that many were still vulnerable 50 years after WWII and reported a revival of feelings and memories associated with the Holocaust.

Predictors of Posttraumatic Stress Disorder

Numerous investigations of WWII, Korean, and Vietnam veterans have shown that not all individuals develop PTSD and/or other psychological disorders after life-threatening traumas. Given an equal degree of traumatic exposure, some individuals are more likely than others to develop subsequent trauma-related psychopathology. In order to explain such individual differences in veterans of the Gulf war, researchers have studied a number of potential risk factors or predictors for the development of PTSD. These have included characteristics of the traumatic event, as well as available personal and environmental resources.

The most commonly cited risk factor for the development of PTSD is the nature and severity of the trauma itself. Many war-related investigations, including a number of Desert Storm studies, have demonstrated a clear dose-response relationship between the severity of psychological trauma and subsequent psychopathology (i.e., the more severe the trauma, the greater the likelihood of developing PTSD and/or other mental disorders). When traumatic exposure involves actual physical injury, the risk for subsequent psychopathology is even greater. In a hospital chart review study of medical and surgical patients who were evacuated from the Gulf to the Walter Reed Hospital, Brandt and co-workers reported a significantly higher rate of axis I disorders (and psychiatric symptoms of concern) in the hospital records of soldiers with traumatic injuries (e.g., shrapnel, gunshot wounds, motor vehicle accidents, SCUD missile attack) compared to soldiers without traumatic injuries (e.g., gastrointestinal bleeding, chest pain, malignancies). Other trauma-related risk factors identified in Gulf war veterans include exposure to dead or dying bodies, having a buddy wounded or killed in action, exposure to American soldiers killed or wounded by friendly fire, being in the reserve (e.g., lack of psychological preparation for deployment), and handling of human remains.

Personal variables that have been studied in Gulf war veterans include gender, race, intelligence, personality, hardiness, coping style, and perceived social

and family support. In a study of 653 Persian Gulf war zone-exposed veterans, Sutker and co-workers reported significantly greater psychological distress and PTSD symptoms in minority men compared to white men but no differences between war zone males and females. In a separate report, Sutker and colleagues found more avoidance, wishful thinking, and self-blame coping strategies; fewer problem-oriented coping strategies; lower scores on hardiness dimensions of commitment, control, and challenge; and less perceived social and family support among Gulf war veterans with PTSD ($n = 97$) compared to those without psychopathology ($n = 484$). No differences in intellectual sophistication as measured by the Shipley Institute of Living Scale were found between the two groups.

Physical Symptoms in Gulf War Veterans

Since the end of the Gulf war, a large number of veterans have reported an array of troubling physical symptoms, including fatigue, joint pain, gastrointestinal problems, skin rash, headache, memory loss, insomnia, and shortness of breath. In a relatively small percentage of veterans, specific etiologies for physical complaints have been linked directly to military service in the Gulf (e.g., leishmaniasis, malaria, and embedded fragments of uranium). However, for the majority of veterans with physical complaints, a specific relationship between physical symptoms and biological, chemical, or other environmental exposures has not been reported in the literature. Further, no specific medical diagnoses appear to explain most of the complaints. As a result, some have postulated a Gulf war syndrome to account for these unexplained physical symptoms.

While some studies clearly report a greater number of physical complaints among military personnel who served in the Gulf compared to those who did not serve in the Gulf, most have not found high rates of any one medical disorder or conclusive evidence for a new syndrome. To date, a number of expert panels (Defense Science Board Task Force, National Institutes of Health Technology, Institute of Medicine, Presidential Advisory Committee on Gulf War Veteran's Illnesses) have failed to identify a new Gulf war syndrome. For example, in 1996 the Institute of Medicine reported that "the committee has not identified scientific evidence to date demonstrating adverse health consequences linked with PGW service other than the documented incidents of leishmaniasis, combat-related or injury-related mortality or morbidity, and increased risk of psychiatric sequelae of deployment." Similarly, in 1996 Gray and associates reported no increased risk of prewar or postwar

hospitalization for any cause among 547 076 Persian Gulf war veterans compared to 618 335 nondeployed. Persian Gulf war veterans did have higher rates of psychiatric hospitalization that appeared to be related to the use of drugs and alcohol. Two large-scale mortality studies have reported no increased mortality as a result of medical diseases in Persian Gulf war veterans. However, an excess number of deaths due to accidents and unintentional injury have been reported.

Relationship between Physical and Psychological Complaints

In a comprehensive literature review on the relationship among psychological trauma, posttraumatic stress disorder, and physical health in diverse trauma populations, Friedman and Schnurr reported that exposure to catastrophic events (independent of actual physical injury) was frequently associated with adverse health reports and increased medical utilization, morbidity, and mortality. Based on work by Wolfe and others, Friedman and Schnurr's review suggested that PTSD – the reaction to a trauma – may serve as an important mediator in the association between traumatic exposure and perceived physical health. Baker and co-workers found support for this hypothesis in a sample of 188 Desert Storm veterans where fatigue, nausea, muscle aches, dizziness, back pain, stomach ache, and numbness were much more likely to be reported by Persian Gulf war veterans with PTSD compared to those without PTSD. Similarly, Stretch and colleagues reported a probable link between Gulf war syndrome and PTSD, even though they found the relationship between stress and physical health to be relatively weak in a large group of Desert Storm veterans. Approximately 30% of veterans with Gulf war syndrome also met criteria for PTSD. In a study of female veterans conducted within 1 year of their return from the Gulf, Wolfe and colleagues found anxiety and PTSD, but not combat exposure, to be significant predictors of health symptoms.

Neurobiological, psychological, and behavioral hypotheses have been offered to explain the relationship between PTSD and physical health. As noted by Friedman and Schnurr, medical problems in trauma survivors may result from neurobiological alterations characteristic of PTSD, such as exaggerated cardiovascular reactivity, adrenergic dysregulation, disturbed sleep physiology, altered hypothalamic-pituitary-adrenal axis activity, and enhanced thyroid function. It is also possible that these neurobiological alterations increase susceptibility to immunologic disorders and infections. Psychological and behavioral factors that may contribute to health problems and that are commonly seen in survivors with PTSD include hostility,

depression, alcohol and drug use, smoking, poor social support, and avoidant and emotion-focused coping.

Summary

Soldiers serving in the Gulf war were exposed to a wide variety of potential health risks, including traumatic stress. Across multiple studies it appears that between 6 and 10% of Gulf war veterans developed PTSD as a result of their service in the Middle East. A host of other psychological and physical symptoms have been reported, and some researchers have proposed a Gulf war syndrome. In Gulf war veterans the precise relationship among stress, the neurobiological alterations that accompany stress, and toxic or infectious exposure currently is poorly understood and the subject of intense scientific investigation.

Acknowledgments

This work was partially supported by grant number DAMD17-96-1-6125, U.S. Army, and by the National Center for PTSD, West Haven, CT.

See Also the Following Articles

Combat Stress Reaction; Gulf War Syndrome, Psychological and Chemical Stressors; Korean Conflict, Stress Effects of; Posttraumatic Therapy; Vietnam Veterans, Postwar Experiences and Health Outcomes; War-Related Posttraumatic Stress Disorder, Treatment of.

Further Reading

- Baker, D. G., Mendenhall, C. L., Simbartl, M. S., Magan, L. K. and Steinberg, J. L. (1997). Relationship between posttraumatic stress disorder and self-reported physical symptoms in Persian Gulf War veterans. *Archives of Internal Medicine* 157, 2076–2078.
- Brandt, G. T., Norwood, A. E., Ursano, R. J., et al. (1997). Psychiatric morbidity in medical and surgical patients evacuated from the Persian Gulf war. *Psychiatric Services* 48, 102–104.
- Friedman, M. J. and Schnurr, P. P. (1995). The relationship between trauma, posttraumatic stress disorder and physical health. In: Friedman, M. J., Charney, D. S. & Deutch, A. (eds.) *Neurobiological and clinical consequences of stress: from normal adaptation to PTSD*. Philadelphia, PA: Lippincott-Raven.
- Gray, G. C., Coate, B. D., Anderson, C. M., et al. (1996). The postwar hospitalization experience of U.S. veterans of the Persian Gulf War. *New England Journal of Medicine* 335, 1505–1513.
- Holmes, D. T., Taniel, P. N. and Cox, C. (1998). Preliminary evidence of psychological distress among reservists in the Persian Gulf war. *Journal of Nervous Mental Disorders* 186, 166–173.

- Institute of Medicine (1995). *Health consequences of service during the Persian Gulf War: initial findings and recommendations for immediate action*. Washington, D.C.: National Academy Press.
- Iowa Persian Gulf Study Group. (1997). Self-reported illness and health status among Gulf war veterans. *Journal of the American Medical Association* 277, 238–245.
- Kroenke, K., Koslowe, P. and Roy, M. (1998). Symptoms in 18 495 Persian Gulf war veterans: Latency of onset and lack of association with self-reported exposures. *Journal of Occupational and Environmental Medicine* 40, 520–528.
- Labatte, L. A. and Snow, M. P. (1992). Posttraumatic stress symptoms among soldiers exposed to combat in the Persian Gulf. *Hospital and Community Psychiatry* 43, 831–833.
- Landrigan, P. (1997). Illness in Gulf war veterans: causes and consequences. *Journal of the American Medical Association* 277, 259–261.
- Laor, N., Wolmer, L., Mayes, L. C., et al. (1996). Israel preschoolers under SCUD missile attacks. *Archives of General Psychiatry* 53, 416–423.
- Morgan, C. A., III, Hill, S., Fox, P., Isingham, P. and Southwick, S. M. (1999). Anniversary reactions in Gulf war veterans: A follow-up inquiry 6 years after the war. *American Journal of Psychiatry* 156, 1075–1077.
- Perconte, S. T., Wilson, A. T., Pontius, E., Dietrick, A. L. and Spiro, K. J. (1993). Psychological and war stress symptoms among deployed and non-deployed reservists following the Gulf war. *Military Medicine* 158, 516–521.
- Presidential Advisory Committee on Gulf War Veteran's Illnesses (1996). *Final report*. Washington, D.C.: U.S. Government Printing Office.
- Robinson, S., Hemmendinger, J., Netanel, M., et al. (1994). Retraumatization of holocaust survivors during the Gulf War and SCUD missile attacks on Israel. *British Journal of Medical Psychology* 67, 352–362.
- Solomon, Z., Laor, N., Weiler, D., et al. (1993). The psychological impact of the Gulf war: a study of acute stress in Israel evacuees. *Archives of General Psychiatry* 50, 320–321.
- Southwick, S. M., Morgan, A., Nagy, L. M., et al. (1993). Trauma-related symptoms in veterans of Operation Desert Storm: a preliminary report. *American Journal of Psychiatry* 150, 1524–1528.
- Stretch, R. H., Marlowe, D. H., Wright, K. M., et al. (1996). Post-traumatic stress disorder symptoms among Gulf war veterans. *Military Medicine* 161, 407–410.
- Sutker, P. B., Uddo, M., Davis, M. J. and Ditta, S. R. (1995). War zone stress, personal resources, and PTSD in Persian Gulf war returnees. *Journal of Abnormal Psychology* 104, 444–452.
- Wolfe, J., Brown, P. J. and Kelley, J. M. (1993). Reassessing war stress: exposure and the Persian Gulf war. *Journal of Social Issues* 49, 15–31.
- Wolfe, J., Proctor, S. P., Davis, J. D., Borgos, M. S. and Friedman, M. J. (1998). Health symptoms reported by Persian Gulf war veterans 2 years after return. *American Journal of Industrial Medicine* 33, 104–113.

Personality Processes

R J Larsen

Washington University, St. Louis, MO, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R J Larsen, volume 3, pp 149–154, © 2000, Elsevier Inc.

Transactional Model of Stress

Summary

Glossary

<i>Appraisal</i>	The part of the transactional model related to an individual's cognitively assessing the demands of the event and determining his or her resources for coping with those demands.
<i>Attributional style</i>	The way an individual explains the causes of bad events.
<i>Coping</i>	The part of the transactional model related to strategies for enduring and overcoming stress.
<i>Expectation</i>	The part of optimism related to the anticipation of positive outcomes.
<i>Internal/external style</i>	The degree to which people blame the cause of bad events on themselves or on things outside of their control.
<i>Neuroticism</i>	A personality characteristic associated with such characteristics as emotional instability, physiological hyperreactivity, and a lowered threshold for the experience of a variety of negative emotions, such as anxiety, anger, irritation, depression, and distress.
<i>Optimistic bias</i>	The part of optimism related to an individual's perception that he or she is at a lower risk for negative events than the average person.
<i>Optimistic explanatory style</i>	The combination of having external, unstable, and specific explanatory styles with respect to bad events.
<i>Primary appraisal</i>	The estimate that an individual makes about the degree to which a particular event is a threat to his or her personal goals.
<i>Secondary appraisal</i>	The estimate that an individual makes about the degree to which he or she has the personal resources to cope with the demands of some particular event.
<i>Self-disclosure</i>	A coping strategy related to the expression of thoughts and feelings about upsetting events.
<i>Self-efficacy</i>	The part of optimism related to an individual's belief in his or her ability to do the behaviors necessary to achieve some desired outcome.

Specific/global style

The degree to which people blame the cause of bad events on factors that are particular to a given situation or on factors that are much more general and applicable to a large number of life situations.

Stable/unstable style

The degree to which people blame the cause of bad events on factors that are long lasting and relatively permanent or on factors that are more temporary and changeable.

Stress

The subjective feeling that is produced by events that an individual perceives as overwhelming and beyond his or her control.

Stress exposure

The part of the transactional model related to being in situations that have a high probability for evoking stress in the average person.

Stressor

An event capable of eliciting stress.

Transactional model of stress

A model that suggests that stress occurs when (1) an individual is exposed to a challenging event, (2) the person appraises the demands of the event and appraises his or her own resources for adjusting to those demands, and (3) the person initiates a strategy for coping.

Stress is the subjective feeling that is produced by events that individuals perceive as overwhelming and beyond their control. Events that typically elicit stress are called stressors. Not all people respond to stressors in the same way, however. Two people may undergo the same event, say a painful loss such as losing their job. One person may be devastated and completely overwhelmed, whereas the other may accept the event as a challenge and mobilize to take positive action. Differences between people in how they react to the same objective event highlight the fact that stress is not out there in the environment. Instead, stress really lies in the transaction between the individual and the characteristics of the environment. Personality processes may moderate this transaction; whether a particular event is experienced as stressful is determined, in part, by the personality characteristics of the person exposed to the event. This article reviews briefly the transactional model of stress and what is known about several stress-moderating personality processes.

Transactional Model of Stress

In general terms, the transactional model of stress, first formulated by psychologist Richard Lazarus,

holds that stress unfolds as follows: (1) an individual is exposed to a challenging event, (2) the person appraises the demands of the event (primary appraisal) and appraises his or her own resources for coping with those demands (secondary appraisal), and (3) the person initiates a strategy for coping. Outcomes then follow, such as a change in the situation or in the health, mood, or behavior of the individual. Personality processes can affect each of the three steps in this transactional model.

Stress Exposure and Neuroticism

Some personality variables are related to higher exposure levels to stress-initiating events. Many studies have documented the role of neuroticism in the exposure to events that elicit stress. Neuroticism is a broad and inclusive personality characteristic associated with such characteristics as emotional instability, physiological hyperreactivity, and a lowered threshold for the experience of a variety of negative emotions, such as anxiety, anger, irritation, depression, and distress. Individuals high in neuroticism (high-N) are frequently found to report greater rates of unpleasant events in their daily lives than low-N subjects, particularly social events such as family difficulties, problems with co-workers, and personal relationships that create conflict and distress. In general, high-N subjects report much more interpersonal stress in their lives than low-N individuals.

The process whereby high-N people come to be exposed to more stressful events is less well understood. One hypothesis is that high-N subjects evoke unfavorable reactions in others. That is, the general characteristics associated with neuroticism lead those around high-N people to react in ways that create stress for them. A second hypothesis is that high-N subjects simply elect to be in situations with higher conflict potential, seeking out those stress-provoking situations that are avoided successfully by low-N subjects. A third hypothesis is that high-N subjects have a lower threshold for reacting to potential stressors; that is, given the same event, high-N subjects exhibit a larger stress response than low-N subjects. Although a dozen or more studies report positive correlations between N and stress responses, more recent controlled laboratory studies using standardized stressors have found that high-N subjects react to identical stressors (e.g., an unpleasant mood induction) with more negative emotions than low-N subjects. For example, studies have shown that high-N people have more cortisol (a biological by-product of stress) in their systems following standardized stress tests than do low-N people. High-N subjects appear to react to challenging events in a way that amplifies their biological stress response. Other studies have docu-

mented that high-N subjects are more reactive to naturally occurring interpersonal and noninterpersonal events (e.g., financial stressors). Stress reactivity blends into the next step in the transactional model – how people cognitively process or appraise events can determine their degree of reactivity to that event.

Appraisal, Attributional Style, and Optimism

According to the transactional model of stress, in order for stress to be evoked for individuals, two cognitive events have to occur. In the first cognitive event, called primary appraisal, people perceive that the event is a threat to their personal goals; in the second necessary cognitive event, called secondary appraisal, people conclude that they do not have the resources to cope with the demands of the event. If either of these appraisals is absent – if people do not perceive the event as threatening or if they feel they have plenty of resources for coping with the threat – then stress is not evoked. If some event is perceived as being threatening to individuals' goals, yet people feels they have the resources necessary to meet that event, then they may experience the event more as a challenge than as stress. Alternatively, individuals may feel that they do not have the resources demanded by the event (secondary appraisal) and that the event is not important to their long-term goals (primary appraisal) and so may not respond with stress.

The appraisal step in the transitional model of stress highlights the importance of cognition in understanding how an event becomes stressful. Individuals may differ in how they cognitively process events such that some individuals experience more stress than others. Attributional style refers to individual differences in the way people explain the causes of bad events. For example, I might ask you why you just had a minor automobile accident. You might attribute your accident to your lack of driving ability, which is an internal attribution, due to you. Or you might attribute your accident to the car's being particularly difficult to drive because it is so unresponsive, which is an external attribution, due to something outside of your control.

Although people might explain any specific event with either internal or external attributions, most people tend to consistently lean in one direction or the other. That is, some people typically give internal explanations and others typically give external explanations. One way to think about attributional style is in terms of the question: Where does the person typically place the blame when things go wrong? For example, some people tend to blame themselves when something goes wrong (e.g., I crashed the car

because I am a clumsy driver) and others never blame themselves for any bad event (e.g., I crashed the car because the road was not built correctly).

In addition to the internal versus external style distinction, attributions can also be characterized as referring to stable versus unstable causes and to whether those causes are specific versus global. For example, say that you are asked why you got a low grade on an exam. An internal and stable attribution might be that you did poorly because you have low ability or are low in intelligence. This is a stable attribution cause because intelligence is thought of as a stable characteristic, as difficult to change. An internal but unstable attribution might be that you did poorly because you were tired that day. Being tired is an unstable attribution because you can get over that relatively easily. Global versus specific causes are whether the cause of the event is due to some global state of affairs or is specific. For example, in explaining a low grade in your statistics course, a global explanation may be that you are low in general intelligence; a more specific explanation may be that you just have trouble with numbers.

The three attributional style dimensions tend to be correlated in a particular pattern. Consequently, many psychologists prefer to view this pattern as a single personality dimension, which they call the optimistic versus pessimistic explanatory style. Optimists believe that they are in charge, that they are the master of their fates, and that their actions actually influence their outcomes in life. Pessimists, however, believe that they are fairly helpless when it comes to bad events and that nothing they can do will make much of a difference in preventing bad events. Consequently, optimists have the expectation that what they do will influence what happens to them, whereas pessimists expect that their behavior is not related to the outcome.

In addition to being related to an internal, unstable, and specific explanatory style, optimism is also related to other personality concepts. For example, optimism also includes the people's expectations that good events will be plentiful, and bad events rare, in the future; for example, optimistic people expect that they will achieve success in most areas of life.

Another related psychological concept, developed by psychologist Albert Bandura, is self-efficacy. This is the people's belief that they can do whatever is necessary to achieve some desired outcome. Self-efficacy also refers to the confidence people have in their ability to perform the actions needed to achieve some specific outcome. For example, if you lose the job you have had for many years, self-efficacy is your subjective belief and confidence that you can find a new job.

A third concept related to optimism concerns perceptions of risk. For example, imagine you are asked to estimate the probability of various unpleasant events happening to you using a scale from 0 to 100, where 0 means "it will never happen to me" and 100 means "it is certain to happen to me." The events you are asked to estimate include dying in a plane crash, being diagnosed with cancer, having a heart attack, and being hit by lightning. Optimists perceive that they are at lower risk for such negative events than the average person is. This is referred to as the optimistic bias, and it may actually lead people in general to ignore or minimize the risks inherent in life or to perceive the world as a fairly benign place.

Psychologists have theorized about the possible mechanisms whereby the optimistic explanatory style buffers stress. These mechanisms should not be viewed as competing hypotheses; instead, two or more of these mechanisms may be operating to buffer stress for optimists.

- **Immune system.** Researchers have shown that the immune systems of optimists respond to challenges better and with more strength than the immune systems of pessimists. Other researchers have examined how optimism relates to the progression of HIV. Although results are mixed, there are suggestions in the literature that optimists have a longer survival time after the development of AIDS symptoms, and that this may be due to stronger immune function among optimists.

- **Emotional mechanism.** There is a very large literature showing that optimists are resistant to depression. It could be that optimism is related to stress buffering indirectly because it protects people from depression and the debilitating consequences of this emotional condition.

- **Cognitive process.** Optimism may be related to an extensive set of beliefs that people have about themselves and the world, and these beliefs may, in part, promote a more positive or adaptive perspective toward potential stressors. For example, researchers have shown that optimism is related to the belief that people can maintain and promote their own positive life circumstances as well as the belief that they can reduce various risks. Other studies have shown that optimists see the world as less filled with hassles and therefore feel less stress than pessimists.

- **Social contact.** Pessimists tend to be loners, and social isolation is a reliable predictor of general distress and dissatisfaction with life. Friends and family may provide the earliest medical feedback when things start going bad. For example, a friend might say, "You look stressed today," or a roommate might remark, "You are looking really tired these days; is

something wrong?" Because pessimists are isolated and avoid social contact, they may have less of this social feedback and hence not have this source of information about their health status.

- **Direct behavior.** Optimism may set in motion certain behaviors that lead directly to benefits for coping. The answer to how optimism and coping are linked may be as simple as that optimists act differently and take better care of themselves and their circumstances than pessimists do. Behavior may prove to be the most critical link in the connection between optimism and coping, simply because optimists do more of the right things (e.g., exercising, sleeping enough, and getting advice from others) and fewer of the wrong things (e.g., drinking and distracting themselves from problems).

As a caveat, most of the research on the association of optimism and stress is correlational, and therefore psychologists cannot pinpoint exactly which factors are responsible for the correlations; they can only speculate. Future research will help determine what are the most effective components of the optimistic personality style in producing better health and a longer life.

Coping and Self-Disclosure

What are some of the strategies that people use to overcome stress and the accompanying feelings of anxiety, being overwhelmed, and being afraid? One strategy studied by psychologists is self-disclosure, the tendency for people to express their thoughts and feelings to others.

Many theorists have suggested that keeping things to ourselves (not opening up to other people) may be a source of stress and ultimately may lead to psychological distress and physical disease. These theorists have further argued that being open with others about our feelings may promote coping. For example, talk therapy may work, in part, because through it patients uncover secrets and reveal what they have been keeping to themselves.

Psychologist James Pennebaker is a leader in the field of self-disclosure. In some early studies, he asked subjects to think of something upsetting or traumatic that had happened to them, something they had not discussed with anyone. He asked his subjects to write these secrets down. People wrote about embarrassing moments, sexual indiscretions, illegal or immoral behaviors, humiliations, and so on. It is interesting that all subjects came up fairly quickly with a secret that they had been keeping.

Not discussing some upsetting event can lead to problems because it takes physical energy to inhibit the thoughts and feelings associated with such events.

It is not easy to keep a secret, and keeping something in, especially if it is a major trauma or an especially significant or upsetting event, takes a lot of energy. Over time the cost of keeping the secret builds and builds and can increase the likelihood of stress-related problems such as trouble sleeping, irritability, physical symptoms such as stomach aches and headaches, and even illness.

According to Pennebaker's early theorizing, telling the secret relieves the compounded stress of keeping it in. In other words, getting something off your chest is a relief from the cost of keeping it inside. Confronting the traumatic memory by telling someone or even writing about it releases the person from the work of keeping a secret. Letting the secret out by talking or writing about it lowers stress and improves the chances for coping and positive adjustment.

In his recent writings, Pennebaker has proposed a different mechanism of action for self-disclosure. By telling the story of some stressful event, people reinterpret the event, often finding a new way to think about it that lessens its impact. For example, a college student talking about a painful breakup may reinterpret the event to actually have some positive consequences (e.g., He really was not my type anyway). Telling others about misfortunes gives people an opportunity to reframe the event and to look for the silver lining in even the worst events.

Many researchers have conducted studies on the self-disclosure of stressful events or telling secrets. In some of these studies, researchers contacted subjects who had undergone a severe event, such as the loss of a spouse through accident or suicide. Subjects were asked how much they had discussed the event with friends, family, or other helping professionals, such as a priest or minister or therapist. Researchers also typically assessed the subjects' health during the time following the stressful event. Results typically showed that the more people had talked about the tragedy, the better their subsequent health. In other words, those who kept the trauma to themselves tended to suffer more problems subsequent to the event than those who disclosed their feelings to others.

In related research, experiments were conducted in which subjects were randomly assigned to conditions. In one condition, subjects were typically asked to recall and write about a distressing event. In the other condition, subjects were typically asked to write about some trivial topic, such as what they normally eat for breakfast. Results frequently showed that subjects writing about the traumatic event reported feeling more distress and discomfort while writing. However, follow-up surveys showed that subjects who wrote about a trauma typically had fewer illnesses compared to subjects who wrote

about trivial topics. Moreover, objective records (e.g., number of health-service visits) also suggest that subjects who self-disclosed about trauma did indeed have fewer illnesses than the subjects who wrote about trivial topics.

Other studies show that people who keep unpleasant information about themselves secret are more likely to develop anxiety or depression than those who tell someone. Often psychotherapists will ask their clients, especially those who have experienced a trauma or some extreme event, to talk or write about that trauma. Some psychologists even recommend that their patients keep a diary of the events in their life and how they are reacting to those events. Such a daily self-disclosure or private secret-telling helps people put their feelings into perspective and make some meaning out of the events in their life.

In summary, research on self-disclosure suggests that keeping traumatic events and the feelings about those events unexpressed can be stressful. Expressing emotions in words can, it appears, produce stress-reducing effects. Moreover, it appears that it does not matter how the feelings are put into words; it may not matter whether we talk to a trusted friend or relative, go to a caring psychotherapist, confess at our church, talk with our minister or rabbi, have a discussion with our husband or wife, or write it in a diary. Whatever form it takes, self-disclosure appears to be a personality process that can help people cope with stress.

Summary

The differences among people in how they respond to events exist because stress is not out there in the environment; instead, stress is in the subjective reaction of individuals to events. This is worth emphasizing because many people often refer to certain events as stressful, as if stress were an inherent characteristic of certain events. Instead, stress lies in how individuals respond to events.

Personality processes may have their effects at one or more points in the stress response. This article presents the transactional model of the stress response, which emphasizes three steps: exposure, appraisal, and coping. Certain personality characteristics have been associated with each of these steps; for example,

exposure has been associated with neuroticism, appraisal with attributional style and optimism, and coping with individual differences in self-disclosure. In reality, the steps of the stress response blend into one another, and so the personality processes presented here certainly overlap; for example, neuroticism may not only be associated with stress exposure, but may also be associated with coping efficacy. In addition, because of space constraints, there are other personality processes in the literature that were not discussed, such as the personality variables hardiness, locus of control, and emotional intelligence. Moreover, there are many individual differences related to health behavior (e.g., smoking, exercise, and social support) that may provide buffers for stress. Identifying and understanding individual differences in the stress response are fascinating and important goals for psychological research.

See Also the Following Articles

Coping Skills; Depression Models; Emotions: Structure and Adaptive Functions; Immune Response; Optimism, Pessimism, and Stress; Social Support.

Further Reading

- Cooper, C. L. and Dewe, P. (2004). *Stress: a brief history*. Boston: Blackwell.
- Dohrenwend, B. P. (1998). *Adversity, stress, and psychopathology*. New York: Oxford University Press.
- Horowitz, M. (2001). *Stress response syndromes: personality styles and interventions*. Northvale, NJ: Jason Aronson.
- Lazarus, R. S. (2006). *Stress and emotion: a new synthesis*. New York: Springer.
- Lazarus, R. S. and Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.
- Moller, T. W. (1997). *Clinical disorders and stressful life events*. Madison, CT: International Universities Press.
- Pennebaker, J. W. (1997). *Opening up: the healing powers of expressing emotion*. New York: Guilford.
- Pennebaker, J. W. (2004). *Writing to heal: a guided journal for recovering from trauma and emotional upheaval*. Oakland, CA: New Harbinger.
- Seligman, M. E. (2006). *Learned optimism: how to change your mind and your life*. New York: Vintage.

Phagocytes See: Macrophages; Glia or Neuroglia.

Pharmacological Treatments of Stress

O G Cameron

University of Michigan, Ann Arbor, MI, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
O G Cameron, volume 3, pp 155–161, © 2000, Elsevier Inc.

stressor); usually, but not invariably, has the connotation that the environmental event is detrimental or experienced as negative or unpleasant.

Stress

General Clinical Psychopharmacology

Psychopharmacology of Stress

Conclusion

Glossary

<i>Acute stress disorder (ASD)</i>	The disordered human stress response that occurs in the first month after exposure to a trauma (stressor).
<i>Fight or flight</i>	Types of reactions in which organisms engage as adaptive responses to environmental demands.
<i>General adaptation syndrome (GAS)</i>	A name given to the organism's reaction to a stressor as described by Hans Selye. The GAS is comprised of three main features – thymic and lymphatic involution, gastric ulceration, adrenal hypertrophy.
<i>Homeostasis</i>	The hypothetical idealized state of the body (<i>milieu interieur</i>) in which all physiological functions are in their optimal basal or most unperturbed conditions.
<i>Neuro-transmitter</i>	Any chemical in the brain or other part of the nervous system that carries information among nerve cells.
<i>Posttraumatic stress disorder (PTSD)</i>	The disordered human stress response that occurs subsequent to the first month after exposure to a trauma (stressor).
<i>Psychopharmacology</i>	The study of the pharmacological mechanisms of, or the therapeutic use of, drugs that affect the functions of the nervous system related to behavior and complex brain functions.
<i>Stress</i>	The complex of biological, psychological, and behavioral changes (that is, the stress response) that occurs as an attempt by an organism to respond adaptively (not necessarily successfully) to exposure to an environmental event (that is, the

Stress

Stress is a widely used hypothetical concept that refers both to environmental events that evoke organismic responses and to the responses themselves. There is little systematic data on either the efficacy or appropriateness of using drugs to prevent or modify the normal stress response. In contrast, a persistent disorder of the stress response (posttraumatic stress disorder [PTSD]) can be at least partially relieved by medication. The most efficacious drugs are the antidepressants, especially the selective serotonin reuptake inhibitors (SSRIs). Furthermore, there are medications that can both relieve the acute stress response and diminish the likelihood of developing PTSD (benzodiazepines and β -adrenergic blockers).

The Definition of Stress

Before reviewing information concerning the psychopharmacological treatment of stress, it is essential to attempt to define stress. This is difficult, and not all individuals involved in stress research or the treatment of people with stress disorders are likely to agree with any particular definition. It is important, nonetheless, to identify basic issues about the definition of this concept as it has been used over the past half century.

Stress has generally been defined in one of two fundamental ways. In the first, stress is an environmental event or circumstance that produces an effect of a particular kind on the organism. That effect has been called the stress response (see **Evolutionary Origins and Functions of the Stress Response, Genetic Polymorphisms in Stress Response, Seasonal Changes in Stress Responses, Sex Differences in Human Stress Response**).

Alternatively, stress has been used to refer to the effects that occur in the organism. In this case, the

environmental event is typically called the stressor. Whenever stress is discussed, one must know in which way the speaker or writer is using the term. For the psychopharmacological treatment of stress, the word is typically used in the second sense, that is, to modify the effect of a stressor on the organism. If one wished to treat stress in the other sense, that would involve not changing the organism's response with treatment, but rather changing the environment itself. In other words, it would involve developing strategies to diminish the occurrence of stressors (a difficult thing to do in real life, and unlikely to involve drugs).

Circularity in the Definition of Stress

An ambiguity or circularity exists in the definitions of stress. In a situation in which an environmental event is quite different from normal circumstances (in other words, unambiguously traumatic), and especially if the event is not avoidable or escapable by the organism (assuming it is a negative event or circumstance, which is usually – but not always – considered to be the case in stress), most people would agree that a stress or stressor has occurred. These are the kinds of events often used in stress research, especially with animals. However, in more typical real-life circumstances, it is frequently not as clear whether a stressor has occurred. For example, while some people find sky diving exhilarating and enjoyable, others find it very frightening. Is it a stressor? Is it a stressor only for those who report a negative subjective response to it? Is it a stressor for anyone who reports a strong subjective reaction, positive or negative? Is it a stressor for everyone, but some people cope with it more effectively? How can stress be unambiguously defined and recognized?

In an attempt to avoid this dilemma, some have advocated using the organism's response instead of the environmental event as the defining characteristic. Classic markers have included physiological reactions such as changes in heart rate, blood pressure, and sudomotor response (sweating), as well as changes in hormones, especially glucocorticoids (such as cortisol) and epinephrine (adrenaline). In this situation, there is a risk that the physiological (or other) changes will be equated with stress rather than being a marker of it. Unless an external event or circumstance is so extraordinary that all observers agree to its divergence from normal circumstances, this definitional problem appears to be almost inevitable. The issue of normal versus abnormal stress is important in understanding the treatment of stress, as described next.

Stress Normal and Abnormal, and Homeostasis

In the middle of the nineteenth century Claude Bernard defined the concept of the *milieu interieur*, his point being that the organism regulates all of the physiological processes going on in the body (under the skin). Early in the twentieth century, Walter Cannon hypothesized that the organism seeks homeostasis, that is, there is an optimal status for the *milieu interieur*, and the organism attempts to maintain that set of conditions. While there is much truth to this concept, it has been found to be an oversimplification. For example, most physiological processes have been shown to fluctuate normally over predetermined time cycles, for example, the circadian (around 24 h) rhythm and the monthly menstrual cycle.

In the context of homeostasis, ignoring the definitional problem, stress can be understood as a perturbation of this set of psychological, physiological, biochemical, and even gene expression states. A fundamental question thus arises: is the stress response – the organism's response to the stressor – a normal reaction? And, if so, what is its function?

Cannon gave an answer to this question, although he was not thinking exactly in terms of the stress response. He viewed the various physiological changes that occur as preparing the organism for fight or flight, in other words, for action. Extensive study of the multitude of reactions that occur during the response to the stressor has generally been consistent with this hypothesis. In short, Cannon was largely correct.

Perhaps the scientist most closely associated with the concept of stress was Hans Selye. He called it the general adaptation syndrome (GAS). He viewed the stress response somewhat differently, not clearly as an adaptive mechanism but as evidence of dysfunction, at least in its more extreme form. In short (admittedly an oversimplification), both Cannon and Selye viewed the changes that are now called the stress response as normative, that is, always or almost always occurring when a stressor is present. However, Cannon viewed the response as adaptive and healthy, while Selye viewed it, at least in its more severe form, as destructive.

Disorders of Stress

Whether or not one views the stress response as adaptive, it does appear that it is normal; in other words, it does always or almost always occur when a stressor is present, at least if the stressor is intense enough. It is also true, however, that a minority of individuals experience stress responses that are abnormal in character, intensity, and/or duration.

Clinical descriptions of detrimental psychological changes occurring in response to traumatic events go back at least as far as the Civil War, and probably much earlier. Phenomena previously identified by such terms as soldier's heart and shell shock refer to such changes occurring in association with traumatic events in war, but these negative psychological reactions can occur to a wide variety of severely traumatic events. The current psychiatric diagnosis nomenclature, as sanctioned by the American Psychiatric Association, is called the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (DSM-IV). In it, these detrimental changes are defined as disorders, and are classified as anxiety disorders. The two specifically defined disorders are acute stress disorder (ASD) and PTSD. The diagnostic criteria for these two disorders are listed in Table 1.

The major differences between the two types of stress disorders are (1) the duration criterion, (2) the dissociation criterion, which is present only in ASD, and (3) the exclusion criterion for substance use, etc., which also is only in ASD. However, for the purpose of this article, the main issue is the fact that both stress disorders are just that, disorders. In other words, these are not descriptions of how the normal or modal person responds to a traumatic event. These are descriptions of how responses occur when they go wrong in some way, definitions of abnormal responses. Research has shown that the majority of people who experience most kinds of traumatic events do not develop stress disorders. Predictors of who will develop these syndromes after experiencing

a trauma are beginning to be identified. Treatment is usually recommended only for the individuals who develop a stress disorder; the question of whether or not (and under what circumstances) to treat the stress response – that is, to treat any individual exposed to a severe stressor – is revisited at the end of this article.

General Clinical Psychopharmacology

Before discussing the treatment of stress disorders with psychopharmacological agents, a brief introduction to psychotropic drugs is in order. Major classes of these drugs are listed in Table 2. The two classes that are most useful in the treatment of PTSD – antidepressants and anti-anxiety drugs – are listed separately. The remaining drugs are listed only as the major classes rather than the individual drug groups.

In many branches of medicine, drugs are listed in classes related to the first or most well defined use for which they have been shown to be effective. For example, a number of drugs with various chemical structures – heterocyclics, monoamine oxidase inhibitors (MAOIs), SSRIs – are all called antidepressants because (1) they all demonstrate an antidepressant effect, (2) this is a very important effect, and (3) this is often the effect for which each drug was first demonstrated to have benefit. Despite this, drugs in the group of antidepressants are beneficial for a variety of disorders. For example, the heterocyclics are also used to treat enuresis (bed wetting), some kinds of pain, obsessive-compulsive disorder, and panic anxiety disorder. MAOIs are useful, for example, for panic disorder and social phobia. SSRIs treat a wide

Table 1 DSM-IV diagnostic criteria for acute stress disorder (ASD) and posttraumatic stress disorder (PTSD)

Acute Stress Disorder (ASD)

- A. Exposure to a traumatic event
- B. Three or more of five dissociative symptoms
- C. Persistent re-experiencing of the event
- D. Marked avoidance of stimuli that arouse recollections
- E. Marked symptoms of anxiety or arousal
- F. Clinically significant distress or impairment
- G. Occurs within 4 weeks of the event and lasts between 2 and 28 days
- H. Not due to a substance, medical condition, or other psychiatric disorder

Posttraumatic Stress Disorder (PTSD)

- A. Exposure to a traumatic event
- B. Persistent re-experiencing of the event in one or more of five ways
- C. Persistent avoidance and/or numbing in three or more of seven ways
- D. Persistent increased arousal in two or more of five ways
- E. Duration is more than 1 month
- F. Clinically significant distress or impairment

Table 2 Classes and examples of psychotropic drugs

Antidepressants

1. Heterocyclic
2. Monoamine oxidase inhibitors (MAOIs)
3. Selective serotonin reuptake inhibitors (SSRIs)
4. Other agents

Anxi-anxiety drugs

1. Benzodiazepines
2. Other

Other classes

1. Mood stabilizers
2. β -adrenergic antagonists
3. α 2-adrenergic agonists
4. Antipsychotics
5. Stimulants
6. Cognition enhancers
7. Narcotic analgesics
8. Drugs for substance abuse prevention

variety of disorders such as panic disorder and the eating disorder bulimia nervosa. The antidepressant drugs, especially the SSRIs, are the drugs that have been demonstrated most clearly to have benefit for people with stress disorders.

The other class of drugs for which there is controlled-trial research concerning treatment of the stress disorders is the benzodiazepines. These drugs are used most widely as anti-anxiety agents. They are also used as sedatives, muscles relaxants, hypnotics (sleeping pills), and anticonvulsants (anti-seizure). This class includes such drugs as alprazolam, diazepam, and clonazepam.

As can be seen in Table 2, there are several other groups of drugs classified as psychopharmacological agents. Either these other groups have not been tested for the treatment of stress disorders, or the studies have not been controlled trials. Table 2 does not contain all drugs or drug groups with psychotropic effects that are potentially useful, however; undoubtedly new drugs will be added in the future.

Psychopharmacology of Stress

In addressing the question of how to treat stress with medications (or otherwise), the first issue that arises is which meaning of stress is under consideration. In addition to PTSD and ASD, many psychiatric syndromes such as schizophrenia, depression, and various anxiety disorders, as well as various medical disorders, are often considered as involving stress. The concept of stress is used very broadly, while research specifically related to DSM-IV defined disorders of stress is much more restricted. Only a small number of controlled studies have been published.

Acute Stress Disorder (ASD)

There has been less research on the treatment of ASD than on that of PTSD, but two issues have emerged. First, medications with rapid onset of effect that can reduce anxiety and/or arousal-like symptoms appear to be appropriate sometimes. Second, a number of acute risk factors for development of PTSD have been identified, including the occurrence of acute stress symptoms and especially adrenergic arousal-like symptoms (mainly, elevated heart rate soon after the stressor, and also elevated urinary epinephrine). (History of childhood trauma, severity and persistence of physical injury, and abnormalities of cortisol secretion are among the other implicated predictors.) Specific medications potentially useful to treat acute stress symptoms include the benzodiazepines and the β -adrenergic blocking drugs (which also can reduce the likelihood of developing PTSD).

Posttraumatic Stress Disorder (PTSD)

Even though the relationship between the experience of a highly disruptive event and the development of post-event PTSD symptoms has been known for many years, systematic research into the use of psychopharmacological agents to treat this type of symptoms has been ongoing for only about 20 years. The number of studies is relatively small, and the number of controlled studies is smaller still. One reviewer has summarized these studies by concluding that drug treatment has modest but meaningful effects.

Specific drugs for PTSD Controlled studies of drug treatment of PTSD have mainly evaluated drugs in the antidepressant class. Among the heterocyclic antidepressants, there is evidence for benefit from imipramine, amitriptyline, and probably chlorimipramine. It appeared that approximately two-thirds of the treated individuals experienced meaningful improvement. The symptoms most likely to improve were intrusive symptoms such as distressing recollections or dreams. Avoidance symptoms were typically less responsive. On the other hand, desipramine did not appear to be as helpful. It has been observed that the drugs that were most helpful have a more prominent effect on the neurotransmitter serotonin in the brain, while less beneficial drugs have a more prominent effect on the brain neurotransmitter norepinephrine.

Antidepressants from two other groups, MAOIs and SSRIs, have also been analyzed in controlled studies. Results with MAOIs have been mixed, but have also supported their usefulness for intrusive symptoms. SSRI drugs have shown robust benefit, not only for the intrusive symptoms but also for avoidance and numbing. For all of these medications, early improvement was associated with a better ultimate outcome, but significant benefit nonetheless typically required 5–8 weeks or more, and further improvement was sometimes observed well beyond 8 weeks.

The benzodiazepine alprazolam was assessed in a controlled trial. It was less beneficial than the antidepressants. Hyperarousal symptoms did, however, show some improvement. Other benzodiazepines are also likely to show benefit for the hyperarousal symptoms.

In addition to the controlled studies, a number of uncontrolled treatment trials have been reported, and a number of different drugs and drug classes have been assessed. Although these results must be viewed as preliminary, promising drugs include the anticonvulsant mood stabilizers carbamazepine and valproic acid. Other drugs also probably have a role as second-tier agents to be used either in combination

with the first-tier drugs or when the first-tier drugs fail. These include the mood stabilizer lithium, the β -adrenergic antagonist drugs, the α_2 agonists, the nonbenzodiazepine anti-anxiety agent buspirone, the sedating antidepressant trazodone, and opiate antagonists (possibly beneficial for numbing symptoms). Antipsychotic drugs are not indicated for PTSD unless psychotic symptoms are present.

Issues and factors in drug treatment of PTSD The definition of PTSD includes three groupings of symptom types: (1) intrusive, (2) avoidance and numbing, and (3) hyperarousal. The heterocyclic antidepressant drugs (primarily those that predominantly affect serotonin) and the MAOIs have been shown to be most beneficial for the intrusive symptoms. SSRIs are also beneficial for the intrusive symptoms, as appears to be the mood stabilizer anticonvulsant carbamazepine. Considering all three types of symptoms, SSRIs appear to be the most beneficial. In general, the intrusive symptoms are probably the ones most responsive to drug treatment.

The avoidance and numbing symptoms appear to be the most difficult to control. They are least likely to improve, and they require the longest treatment – many weeks – before improvement occurs, if it occurs at all. The antidepressant drugs, mainly the SSRIs, are most likely to produce improvement for these symptoms.

Hyperarousal symptoms seem intermediate in their likelihood of response. Benzodiazepines, buspirone, and the β -adrenergic antagonists are most specifically efficacious for the hyperarousal symptoms.

Several other issues in the drug treatment of PTSD are noteworthy. Not only do these drugs take longer to provide benefit for PTSD than they usually do for depression (2 months or more versus approximately 1 month), but they often require higher doses. Often only partial symptom relief is attained with the drugs now available. A total treatment program, including nonpharmacological treatments, is almost always appropriate. The presence of comorbid disorders such as depression, other anxiety disorders, and personality disorders modulates treatment outcome, although the direction of effect is variable. For example, some research suggests that the presence of depression improves prognosis, while other research suggests prognosis is worse when depression is present, possibly because of the increased overall severity due to two co-existing (comorbid) disorders. (Medication benefit does not depend on the presence of a comorbid disorder.) The presence of a personality disorder, especially antisocial personality, makes prognosis worse. The source of the trauma (for example, combat

Table 3 Factors and goals of drug treatment of PTSD^a

1. Improve intrusive symptoms
2. Control avoidance and numbing symptoms
3. Reduce hyperarousal symptoms
4. Provide distance between the individual and intensity of the emotional experience
5. Relieve depressive symptoms
6. Minimize irritability and aggressiveness
7. Gain control of dissociative and psychotic symptoms if they occur
8. Facilitate efficacy of psychotherapy

^aModified from Vargas and Davidson (1993).

versus non-combat related) can influence outcome. Finally, it is likely that the earlier that treatment is provided after the PTSD syndrome develops (or even in the first month after the trauma – ASD), the more improvement can be anticipated.

It has been suggested that the phasic (intermittent) PTSD symptoms are more responsive to treatment than the tonic (persistent) symptoms. A list of the ways in which drug treatment can be helpful in the management of PTSD appears in [Table 3](#).

Stressful Circumstances: Should Normal Stress Be Treated?

The research reviewed previously demonstrates that medications are beneficial, but not usually curative, for PTSD. Older research had suggested the use of sedative drugs for abreaction for ASD, a treatment now not generally recommended. Some clinicians have suggested that sedation, possibly including sedative effects from benzodiazepines, might be detrimental, although the anti-arousal, anxiolytic effects can be acutely beneficial. These data lead to the question, can the normal stress response itself be diminished or prevented effectively (with pharmacological and/or other means)? Should it be?

If one views the stress response as Cannon conceived of it (that is, the fight-or-flight reaction), it would seem that prevention of this normal physiological and behavioral response to environmental stressors would be detrimental rather than beneficial. On the other hand, if one views the stress response as Selye did, that is, as a reaction that at least in its more severe form is harmful to the organism, preventing it would be appropriate. As pointed out by Yehuda and McFarlane, there is considerable confusion about the relationships among stressors, potential detrimental effects of the stress response itself, and the development of PTSD. For example, although it is widely believed that stress has many negative consequences on health, the relationship between stress and the development, maintenance, and/or

exacerbation of specific medical symptoms and syndromes remains at best only partially understood. Provocative preliminary hypotheses exist, such as the apparent detrimental effects of chronic elevations of corticosteroid stress hormone substances on specific areas of the brain (especially the hippocampus, a structure very important for normal memory function), but much more needs to be known to clarify these linkages. It seems that the answer to the question of whether to treat normal stress must await not only additional data, but also improved clarity concerning the definition(s) and use of the word and concept of stress as applied to humans in clinical situations. What is clear is that abnormal stress responses call for intervention.

While investigators and clinicians struggle to obtain greater conceptual clarity in the field of stress research and treatment, new data have begun to address the question of early pharmacological intervention for abnormal stress responses. Some early risk factors, predictive of increased likelihood of the development of PTSD, can now be used to lower that risk. Specifically, the presence of an elevated heart rate or other evidence of significant adrenergic activation soon after the trauma appears to call for measures, pharmacological and/or nonpharmacological, to promptly lessen this activation.

Conclusion

Disordered chronic responses to significant stressors (PTSD) are treatable with medications, although most people with the disorder do not experience a full remission from drug treatment, and some symptoms (avoidance and numbing) are harder to treat than others (intrusive, hyperarousal). Most individuals need a treatment program including pharmacological and nonpharmacological methods to derive maximal benefit. Prophylactic drug treatment given soon after trauma exposure to reduce arousal can diminish the risk of subsequent development of PTSD. The larger question of whether or not stress per se is treatable (that is, the normal stress response is diminished or prevented), and whether or not that would be advantageous or disadvantageous to the organism, depends in part on which definition of stress is under consideration.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Alarm Phase and General Adaptation Syndrome;

Evolutionary Origins and Functions of the Stress Response; Fight-or-Flight Response; Genetic Polymorphisms in Stress Response; Homeostasis; Seasonal Changes in Stress Responses; Sex Differences in Human Stress Response; Selye, Hans; Stress, Definitions and Concepts of; Sympathetic Nervous System.

Further Reading

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th edn.) (DSM-IV). Washington, D.C.: American Psychiatric Press.
- Brady, L. S. (1994). Stress, antidepressant drugs, and the locus coeruleus. *Brain Research Bulletin* 35, 545–556.
- Cannon, W. B. (1939). *The wisdom of the body*. Philadelphia: W.W. Norton.
- Davidson, J. (1992). Drug therapy of post-traumatic stress disorder. *British Journal of Psychiatry* 160, 309–314.
- Goldstein, D. S. (1994). Stress and science. In: Cameron, O. G. (ed.) *Adrenergic dysfunction and psychobiology*, pp. 179–236. Washington, D.C.: American Psychiatric Press.
- Katz, L., Fleischer, W., Kjernisted, K., et al. (1996). A review of the psychobiology and pharmacotherapy of posttraumatic stress disorder. *Canadian Journal of Psychiatry* 41, 233–238.
- Marmar, C. R., Neylan, T. C. and Schoenfeld, F. B. (2002). New directions in the pharmacotherapy of posttraumatic stress disorder. *Psychiatric Quarterly* 73, 259–270.
- Pitman, R. K., Sanders, K. M., Zusman, R. M., et al. (2002). Pilot study of secondary prevention of post-traumatic stress disorder with propranolol. *Biological Psychiatry* 51, 189–192.
- Selye, H. (1956). *The stress of life*. New York: McGraw Hill.
- Tomb, D. A. (ed.) (1994). *Psychiatric clinics of North America, vol. 17: Post-traumatic stress disorder*. Philadelphia: W.B. Saunders Company.
- Turnbull, G. J. (1998). A review of post-traumatic stress disorder, part II: treatment. *Injury* 29, 169–175.
- Vargas, M. A. and Davidson, J. (1993). Post-traumatic stress disorder. In: Dunner, D. L. (ed.) *Psychiatric clinics of North America, vol. 16: psychopharmacology II*. Philadelphia: W.B. Saunders Company.
- Weiner, H. (1992). *Perturbing the organism*. Chicago: University of Chicago Press.
- Yehuda, R. (ed.) (2002). *Psychiatric clinics of North America, vol. 25: Recent advances in the study of biological alterations in post-traumatic stress disorder*. Philadelphia: W.B. Saunders Company.
- Yehuda, R. and McFarlane, A. (1995). Conflict between current knowledge about posttraumatic stress disorder and its original conceptual basis. *American Journal of Psychiatry* 152, 1705–1713.

Pheromones

A Kumar, C A Dudley, S Chakravarty and R L Moss

University of Texas Southwestern Medical Center,
Dallas, TX, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
C A Dudley, A Kumar and R L Moss, volume 3, pp 162–168,
© 2000, Elsevier Inc.

Chemical Communication and Social Odors

The Vomeronasal Organ

Vomeronasal Organ Response to Olfactory Stimuli and
Olfactory System Response to Pheromonal Stimuli

Stress and Reproduction

Stress Signaling Mediated by Pheromones

Glossary

<i>Accessory olfactory bulb (AOB)</i>	A neural structure located at the dorso-caudal end of the main olfactory bulb but distinguished from the main olfactory bulb by the types of signals detected and by its central nervous system projection sites.
<i>c-Fos immunocytochemistry</i>	A method for measuring cellular activation by detecting the expression of the c-fos oncogene.
<i>Chemo-architecture</i>	The relationship between cellular topography and the cellular machinery used in signal detection and processing.
<i>Chemosensory stimulation</i>	An environmental stimulus detected by a sensory system via ligand–receptor binding.
<i>Gonadotropin-releasing hormone (GnRH)</i>	A decapeptide produced in the hypothalamus, responsible for stimulating the release of luteinizing hormone from the pituitary gland; also referred to as LHRH.
<i>Sociogenic stress</i>	A set of stresses caused by conspecific interactions.
<i>Vomeronasal organ (VNO)</i>	a specialized chemosensory structure for pheromone signaling.

Chemical Communication and Social Odors

Interorganism communication is an indispensable feature of species survival, and the signals employed for this purpose are quite diverse. Among the various senses employed, chemosensation (the detection and transduction of chemical environmental stimuli) is

perhaps the most primitive. In addition to aiding in locating nutrition, chemosensory transduction allows an animal to find suitable mating partners and avoid dangerous predators. Pheromones are compounds used for intraspecies chemical communication that are secreted or excreted by one individual and elicit behavioral or physiological responses in the recipient. However, the great diversity of signals used in chemical communication by vertebrates suggests that this communication is not mediated exclusively by pheromones. A multiplicity of signals may be present in the same secretion, and it is unlikely that the response to this complex signal is mediated exclusively by a single sensory system. Johnston has recently suggested the term chemical signals for the wide array of odors regulating vertebrate behavior and physiology. Indeed, various chemosensory structures have been found in vertebrates that transduce chemosensory cues. Pheromones act on the recipient through olfaction, ingestion, or absorption; however, all mammalian pheromones studied so far act through olfactory pathways. Olfactory-mediated pheromonal communication in mammals is used to organize a wide range of social behaviors, such as mating, dominance, aggression, and maternal bonding, and hence these cues are also frequently referred to as social odors.

Mammalian pheromones elicit both long-lasting effects that bring about changes in the endocrine state of the recipient animal and short-term rapid effects on the recipient's behavior. For example, male mouse pheromones have been shown to induce advanced onset puberty in female counterparts (the Vandenberg effect), the onset of estrus (the Whitten effect), and the termination of pregnancy (the Bruce effect) by an increase in luteinizing hormone (LH) and a decrease in prolactin. Female mouse pheromones cause a delay in the onset of puberty and a suppression of estrus in other female mice by increased release of prolactin. These delayed neuroendocrine effects on reproductive physiology and behavior have been termed the primer effect of mammalian pheromones.

The faster-acting releaser effects are elicited immediately, and these pheromones have been implicated in the regulation of intermale aggression, initiation of male ultrasonic vocalization and copulatory behavior, induction of lordosis (the receptive posture adopted by female rodents anticipating sexual intercourse), and parental care. Releaser effects are also

modulated by the hormonal status of the animal, which is itself under pheromonal regulation. Gonadotropin-releasing hormone (GnRH) and LH play crucial roles in regulating these pheromone-induced fast behavioral responses.

The Vomeronasal Organ

Pheromone molecules are chemically unrelated and are contained in body fluids such as urine, sweat, and sebum. Pheromones are detected by peripheral sensory neurons sequestered in a tiny cigar-shaped bony capsule called the vomeronasal organ (VNO) or Jacobson's organ (Figure 1a). Located bilaterally on the posterodorsal aspect of the nasal septum, this tubular structure opens into the nasal or oral cavity via the vomeronasal duct, allowing chemical stimuli to reach VNO neurons from both the nose and mouth. In some mammals such as ungulates and felines, the males display a behavior called flehmen, which consists of curling of the upper lip and throwing the head up after investigating the female's genital area. This helps in facilitating the transport of the odorous compounds in female urine into the VNO of male, thus allowing the detection of the female's receptive state. In rodents, however, a pumping activity of the VNO associated with the autonomic afferent system appears to actively draw pheromone compounds into the lumen of the VNO, where they bind to receptors. These receptors lie on microvillous extensions of the dendrites of the vomeronasal

receptor neurons (VNRNs). Patch clamp recordings from dissociated VNRNs reveal that they possess voltage-gated inward and outward currents consistent with the ability to generate action potentials. VNRNs from the mouse respond to putative pheromones from mouse urine with the generation of an outward current accompanied by a decrease in membrane conductance (Figure 2). In a current clamp mode, pheromone application causes hyperpolarization of the membrane. The binding of pheromone molecules to their specific membrane receptors activates an intracellular signaling cascade involving G-proteins ($G_{\alpha_{i2}}$ in the apical neurons, G_{α_o} in basal neurons), phospholipase C- β (PLC), and diacylglycerol (DAG). DAG, the second messenger, is the gating factor for a cation channel known as transient receptor potential channel 2 (TRPC2), whose opening leads to an influx of $\text{Na}^+/\text{Ca}^{2+}$ that causes membrane depolarization. The spread of membrane depolarization along the dendrite leads to the activation of voltage-gated Na^+ and K^+ channels, with the consequent generation of an action potential. The propagation of impulse along the nerve fibers allows VNO neurons to communicate with the first central processing unit, the accessory olfactory bulb (AOB).

Much insight into the molecular and anatomical components of VNO-signal transduction has been gained from the study of mutant mice. Mice that lack a functional TRPC2 channel show normal olfactory function but fail to respond to urinary vomeronasal stimuli. These mice do not display male-male

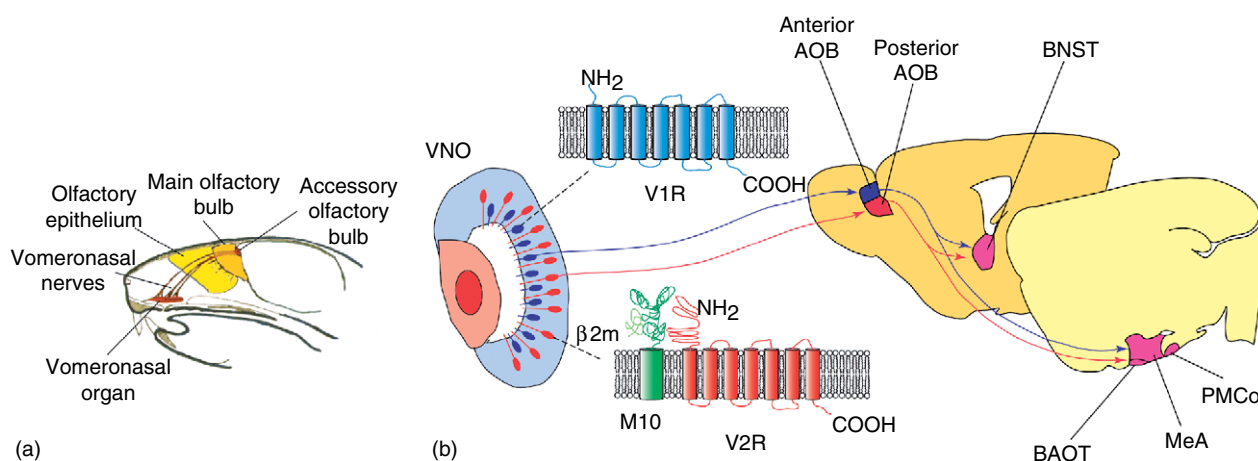


Figure 1 Rodent vomeronasal organ (VNO). a, Sagittal view showing the location in relation to the main olfactory epithelium; b, anatomy and central neural projections, with the sensory vomeronasal receptor neurons (VNRNs) shown in a coronal cross section of the VNO. Two discrete classes of VNRNs (V1Rs and V2Rs), G-protein-coupled receptors, project separately to the anterior and posterior accessory olfactory bulb (AOB), respectively. The V2Rs differ markedly from the V1Rs in that they have a large extracellular N-terminus and form a receptor complex with members of the M10 family of major histocompatibility complex (MHC) class Ib molecules and β_2 -microglobulin ($\beta_2\text{m}$). The information from the V1Rs and V2Rs are processed in the AOB and then converges in overlapping projections to the bed nucleus of the stria terminalis (BNST), bed nucleus of the accessory olfactory tract (BAOT), medial amygdala (MeA) and the posteromedial cortical nucleus (PMCo) of the amygdala. COOH, From Brennan, P. A. and Keverne, E. B. (2004), Something in the air?: new insights into mammalian pheromones, *Current Biology* 14, R81–R89, with permission.

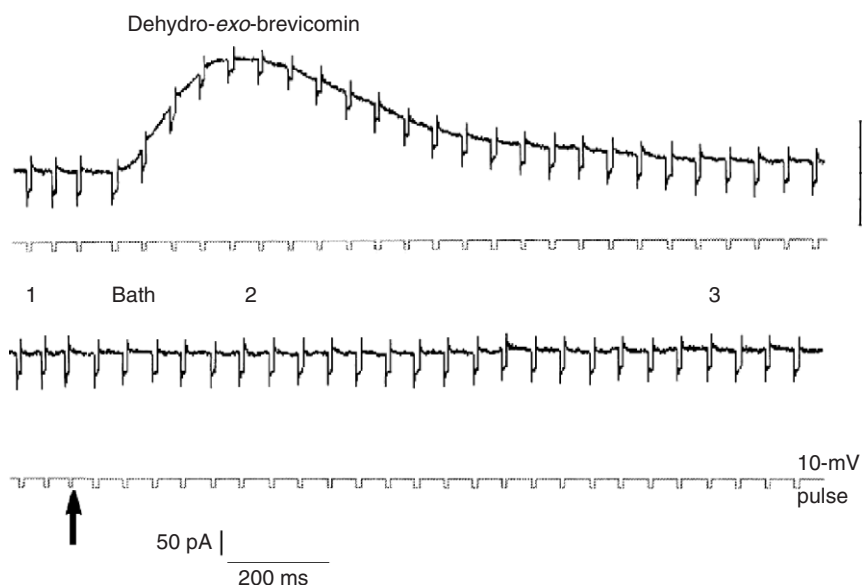


Figure 2 Application of the putative pheromone dehydro-exo-brevicomin (DHB) to a VNRN evoked an outward current accompanied by a decrease in conductance. Depolarizing pulses of 10 mV were applied continuously before (1), during (2), and after (3) DHB or bath. In the presence of DHB, the amplitude of the current pulses was decreased. From Moss, R. L., Flynn, R. E., Shen, X. M., et al. (1997), Urine-derived compound evokes membrane responses in mouse vomeronasal receptor neurons, *Journal of Neurophysiology* **77**, 2856–2862, with permission.

aggression and engage in sexual and courtship behaviors toward both males and females. They also have deficits in maternal aggression toward male intruders. This lack of aggressiveness in these mutant mice is in accordance with similar findings in mice following VNO lesions. The mouse VNO has two broad receptor gene families, V1Rs and V2Rs, comprising over 250 receptors genes. When the genes for 16 of the V1R family clustered together were deleted using chromosomal engineering (which amounted to 12% of the V1R population), male mice were impaired in their sexual behavior toward females, with fewer males initiating sexual activity toward females than the genetically intact controls. These mice, too, fail to initiate maternal aggression toward intruder males, as was found in TrpC2-deficient mice, although there were no differences in the aggressive behavior of these V1R cluster-deficient males. These and other studies have yielded important insights into the functioning of the vomeronasal system and have helped researchers appreciate the complexities of this chemosensory system.

Chemoarchitecture of the Vomeronasal Organ to Accessory Olfactory Bulb Projection

The diversity in electrical responses of VN receptor neurons to putative pheromones, as well as the vast array of behaviors subjected to pheromone regulation, may be related to the fact that VNRNs appear to be divided into at least two chemoarchitecturally

distinct populations: VNRNs located in layers close to the lumen of the VNO express $G_{\alpha_{i2}}$ and project to the anterior portion of the AOB; VNRNs located more basally express G_{α_o} and project to the posterior part of the AOB. The significance of this division has become more apparent with studies identifying two families of VNO receptor genes, with V1Rs and V2Rs expressed in $G_{\alpha_{i2}}$ and G_{α_o} VNRNs, respectively. This expression pattern suggests that there exist two functionally distinct processing units responsible for relaying signals from the VNO to the AOB.

Until now, the physiological basis for chemoarchitecturally segregated populations of VNO neurons was largely unexplored. Presuming that both populations of VNO cells are involved in pheromone detection, they may recognize separate categories of pheromones (volatile vs. nonvolatile, or primer vs. releasing) or may be involved in the mediation of separate behavioral responses. In the male mouse, the VNO is involved in at least two distinct behaviors: reproduction and aggression. Experiments have demonstrated that neurons in the anterior portion of the AOB of the male mouse are preferentially activated by exposure to female-soiled bedding, as opposed to familiar male-soiled bedding. The differential activation was dependent on an intact VNO and was not observed in the main olfactory bulb (MOB). These data indicate that the position of a vomeronasal receptor cell within the VNO is reflected at the level of the AOB by a functional division along the anterior–posterior axis.

Recently, Katz's group performed the first recordings of pheromonal responses in the AOB of awake-behaving mice that demonstrated that responses of the AOB mitral cells are highly selective to strain and the sex of conspecifics. It is also possible that nuclei downstream of the AOB, such as the medial amygdala (mAMG), may also be preferentially activated. The nature of the information processing by V1Rs and V2Rs is yet to be understood, but the anterior and posterior AOB projections completely overlap at the level of amygdala, the bed nucleus of the accessory olfactory tract (nAOT), and the bed nucleus of the stria terminalis (BNST).

Central Neural Pathways of the Vomeronasal System

The anatomical pathway linking pheromone detection in the VNO to the hypothalamic centers controlling sexual and reproductive events has been well defined by studies using transsynaptic tracers. VNRNs project directly to the AOB, where they synapse onto mitral cell dendrites. The mitral cell axons project to the mAMYG, BNST, and nAOT. Connections between the mAMYG and the preoptic area (POA) and between the mAMYG and the ventromedial hypothalamus (VMH) complete the circuit by which pheromones may modulate the hypothalamic and/or limbic centers involved in reproductive behavior. Lesions of the VNO, AOB, mAMYG, or VMH significantly impair the ability of the female rat to display increases in sexual receptivity in response to male sexual contacts. The possibility of direct projections from the VNO to the POA and VMH has received some support and may provide an additional pathway for pheromonal control over GnRH-secreting neurons located in the hypothalamus. The surgical removal of the VNO adversely affects the copulatory behavior in males and lordosis behavior in female rodents. GnRH has been shown to facilitate mating behavior in the female rat, and GnRH has also been demonstrated to restore mating behavior in the female rat and male hamster after VNO removal.

Activation of Vomeronasal Neural Pathways

Studies employing c-fos-immunocytochemistry as a marker of cellular activation have revealed that vomeronasal neurons are activated by chemosensory stimulation. Following mating with a male, female rats exhibited a VNO-dependent increase in c-fos expression in all vomeronasal system nuclei. A particularly dense concentration of activated cells was observed in the mitral and granule layers of the AOB (Figure 3a) and in the mAMYG (Figure 3b), with increases in cellular activation also noted in the

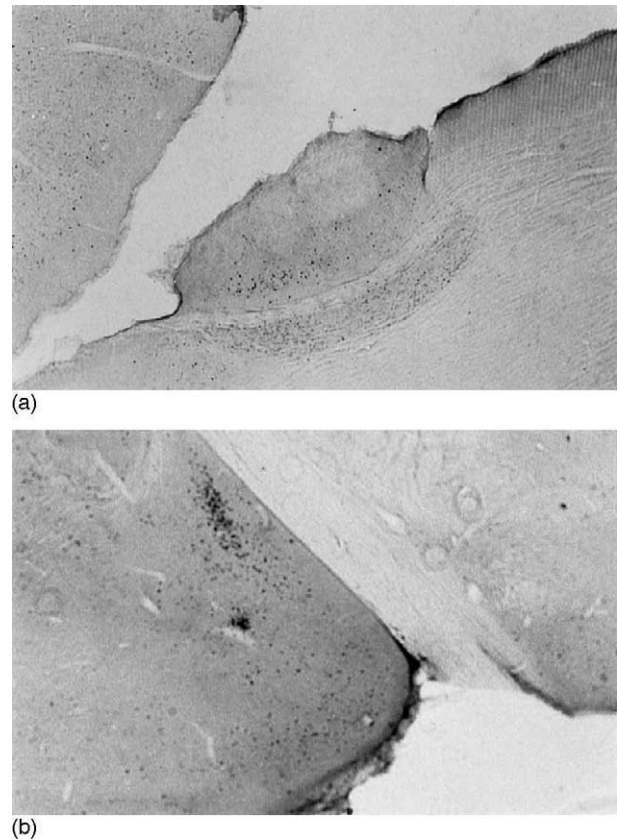


Figure 3 Samples from female rats exposed to 3 h of repetitive mating. a, the accessory olfactory bulb; b, the medial amygdala (mAMYG). c-Fos-immunopositive cells are indicated by small black dots. Note the cluster of c-fos-immunopositive cells parallel to the optic tract in the mAMYG. From Dudley, C. A., Sudderth, S. B. and Moss, R. L. (1992), LHRH neurons in the medial septal-diagonal band-preoptic area do not project directly to the hippocampus: a double-labeling immunohistochemical study, *Synapse* 12, 139–146, with permission.

BNST, the nAOT, and the ventrolateral VMH. Female cervical stimulation or exposure of male rats to female-soiled bedding evoked a similar pattern of c-fos activation. In the male hamster, mating or exposure to vaginal discharge also produced activation of the vomeronasal system.

Activation of Gonadotropin-Releasing Hormone Neurons

During copulation, the female rat becomes more receptive over time and exhibits a release of LH (Figure 4). Because LH release is stimulated by GnRH, the chemosensory stimulation obtained during copulation is likely to activate GnRH neurons. Indeed, double-labeling immunocytochemical studies demonstrated that mating activated approximately 80% of the GnRH neurons, whereas after VNO removal only approximately 38% of the GnRH neurons were activated (see Figure 5). Similar studies

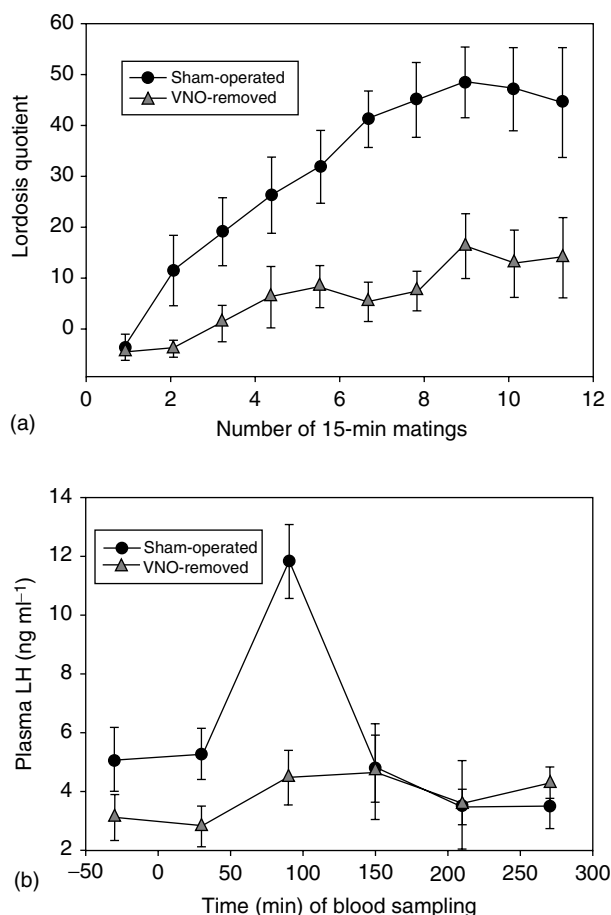


Figure 4 The effect of VNO removal in female rats. a, On lordotic responsiveness; b, on mating-induced luteinizing-hormone (LH) release. From Rajendren, G., Dudley, C. A. and Moss, R. L. (1990), Role of the vomeronasal organ in the male-induced enhancement of sexual receptivity in female rats, *Neuroendocrinology* **52**, 368–372, with permission.

from other laboratories also revealed an increase in the activation of GnRH neurons in the female mouse during mating and in the female rat during vagino-cervical stimulation. These observations provide firm evidence that chemosensory stimulation at the level of the VNO influences endocrine events at the hypothalamic level.

Vomeronasal Organ Response to Olfactory Stimuli and Olfactory System Response to Pheromonal Stimuli

Recent reports have indicated that cross-talk seems to exist between the vomeronasal system (VNS) and the main olfactory system. Although the VNO is mainly dedicated to sensing pheromones and the main olfactory epithelium that lines the nasal cavity senses common odors, recent findings make us rethink how pheromones control hormonal responses. This is particularly relevant for humans, in which adults lack a

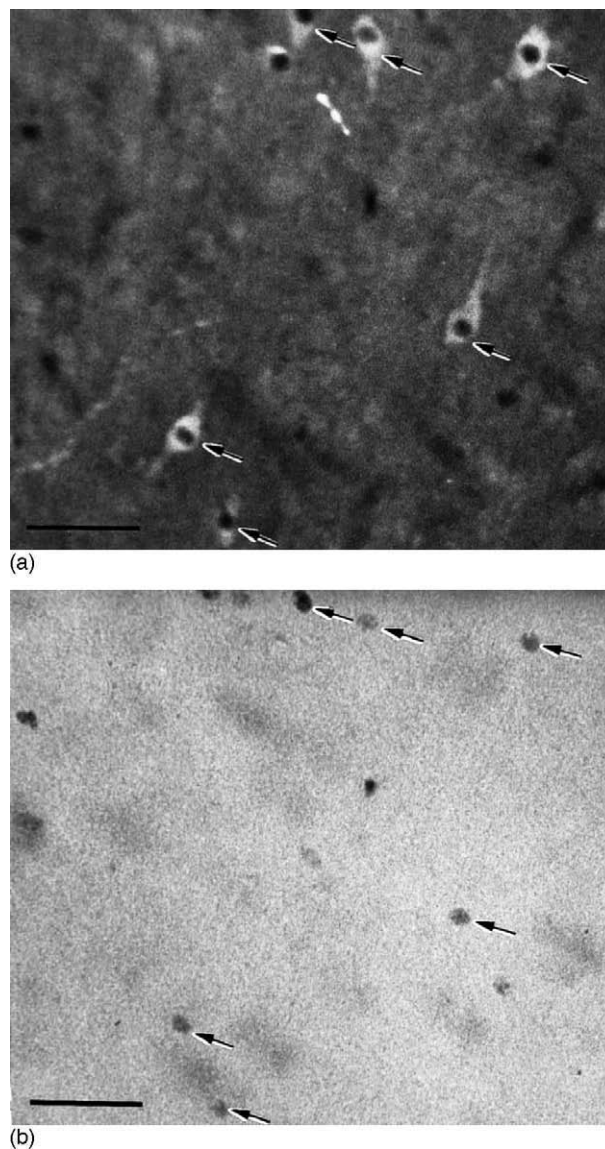


Figure 5 GnRH-containing neurons that were activated by 3 h of repetitive mating in the female rat. a, Arrows point to GnRH neurons labeled immunofluorescently with Texas Red; b, arrows point to c-fos-immunopositive nuclei. In a, the five GnRH-containing cells also contained the c-fos protein. From Rajendren, G., Dudley, C. A. and Moss, R. L. (1993), Influence of male rats on the luteinizing hormone-releasing hormone neuronal system in female rats: role of the vomeronasal organ, *Neuroendocrinology* **57**, 898–906, with permission.

functional VNO. In mammals, mating and the reproductive behavior of females are controlled by GnRH neurons in the hypothalamus. GnRH is secreted into the hypophysial portal system, through which it regulates the action of gonadotropin-secreting pituitary neurons. Recent findings have brought the main olfactory system into the limelight in the pheromone and behavior field, not only by showing its crucial role in sexual and aggressive behavior but also by revealing a direct anatomical projection from the

main olfactory epithelium to luteinizing hormone-releasing hormone (LHRH) neurons in the hypothalamus, which regulate the gonadal hormone release. There is also evidence that the VNO in the sexually naïve male hamster is crucial for the female pheromone-induced mating behavior in male, but the main olfactory system takes over the functional role of the VNO in the context of sexual behavior. Thus, the traditional distinction that common odors are perceived through the olfactory pathway and pheromones by the vomeronasal pathway does not hold true; rather, the main olfactory system and vomeronasal system complement one another in chemosensory communication.

In the human fetus, a VNO develops and appears to guide the migration of GnRH neurons to the hypothalamus. However, the lack of a functional VNO in human adults does not imply that pheromone signaling does not occur in humans. Most of the vomeronasal receptor genes in the human genome are pseudogenes; however, four of these seem to be functional, of which one, hV1RL1, is expressed in olfactory mucosa and is likely to have a transducing function. At present, the only reproducible pheromonal effect in humans is the effectiveness of axillary secretions from women in different stages of their menstrual cycles to alter the cycle lengths of other women living in close proximity, thus causing menstrual synchronization. The recent findings in mice that the main olfactory pathway is required for reproductive behavior and that it projects to and activates hypothalamic GnRH/LHRH neurons might explain how organisms without a functional VNO, such as humans, use olfactory inputs in order to activate subcortical hard-wired circuits controlling innate sexual and reproductive behavior. Advances in neuroimaging have allowed us to visualize brain activity patterns in awake humans, and these techniques have shown that sex pheromones activate wide regions of the human brain, including areas traditionally thought to mediate cognition. These studies suggest that, as in the rodent, the pheromone-activated endocrine system is much more extensive than previously thought. Thus, it appears that we need to learn more about how intimately neuroendocrine functions, controlled by pheromones acting through our noses, might interact with other operations within the brain to control human behavior and cognition.

Stress and Reproduction

Social Interactions

Stresses of a social nature have the potential to exert negative effects on reproduction. The most common

social stress in animals of the same species is caused by high population density. Crowded conditions give rise to complex emotional states that may interfere with reproductive function. In addition, pheromonal cues emitted and detected in crowded conditions contribute to a decline in endocrine function. Although the exact mechanisms are uncertain, a wide array of noxious physical stimuli and distressed emotional states, including those of social origin, can enhance the activity of the hypothalamic-pituitary-adrenal (HPA) axis. The enhanced activity of this system is correlated with the suppression of reproductive processes at many levels in the reproductive axis. This suppression appears to be brought about by several specific actions that affect the GnRH system of neurons. First, corticotropin releasing hormone (CRH) and the proopiomelanocortin (POMC)-derived peptides adrenocorticotrophic hormone (ACTH) and β -endorphin can act in the brain to inhibit the release of GnRH. Second, corticosteroids themselves can decrease the responsiveness of the pituitary to GnRH. All these actions can interfere with the stimulatory effect of LH on steroidogenesis in the gonad. However, the extent of reproductive inhibition varies with social rank, sex, maturity, season, and some other factors.

The negative effect of emotional stress on reproduction is well illustrated in studies performed on cages of mice that are densely populated. Typically, a large cage is populated by a few pairs and the population is allowed to grow until it regulates itself. Maintaining this regulated population level often involves a marked inhibition of sexual maturation among young animals, as well as a cessation of reproduction by adults. In crowded deer mice, the effects of high population density on reproductive function are profound. When young deer mice are allowed to grow in pairs in small cages with adults of the opposite sex, they become fertile at 5–8 weeks of age. On rearing in dense populations, however, by 12–13 weeks of age, as many as 95% of these deer mice still are not fertile. According to Terman and colleagues, many die at well over 1 year of age without ever having achieved fertility. To some extent, pheromones can alleviate the negative effects of environmental stress on estrous cycling. Female mice housed in grouped, crowded conditions experience a lack of estrous cycling, which can be overcome by urinary cues from a conspecific male or by putative pheromones isolated from male mouse urine.

Agonistic Interactions

Intermale aggression in the form of fighting is common among densely populated animals. Losers of these

fight as well as other social subordinates and most of the onlookers seem physiologically stressed. Male mice fight to establish dominance even when only two are brought together. Agonistic interactions have been shown to have a profound impact on the endocrine system. The defeated animal experiences an almost immediate inhibition of LH secretion, lowered testosterone levels, and a dramatic rise in the secretion of corticosterone. Moreover, the increase in corticosterone can be conditioned to occur in response to the mere presence of a dominant male. Thus, even without the occurrence of overt fighting, subordinate animals may have depressed LH secretion. Such alterations in hormone levels may have negative consequences for health and fertility because the long-term hyperactivation of the HPA axis is related to a decline in immune system activity and reduced gonadal activity may result in sterility.

In higher vertebrates, behavioral responses to unfamiliar conspecifics are dependent on the actual environmental situation as well as social experiences made during behavioral development. Sapolsky's studies of a population of wild African olive baboons found that high- and low-ranking baboons differ in levels of cortisol and testosterone during periods of social instability, when overt aggression is quite common. Dominant males tended to have the lowest levels of cortisol. During periods of relative stability in the colony, differences in cortisol and testosterone between dominant and subordinate males were no longer detectable. Differences in cortisol levels appear to be extremely important, not only in terms of reproductive success but also in terms of survival. Males with the lowest cortisol levels survived the longest. A strong relationship exists among environmental factors (the rearing conditions as well as the actual situation), aggressive behavior, endocrine responses, and health.

Recently, two synthetic pheromones, 2,3-dehydro-*exo*-brevicomin (DHB) and 2-*sec*-butyl-dihydrothiazole (SBT), the urinary constituents in male mice, have been shown to elicit aggression in mice. The level of two other urinary constituents from male mouse, α - and β -farnesene, were elevated in the excreted urine of dominant male mice even several days after establishing dominance. The farnesene, absent in bladder urine but abundant in the preputial gland, when mixed with water or bladder urine significantly reduced the amount of time subordinate males investigated the urine deposits of dominant males. Thus, these four different urinary constituents have been shown to contribute to male-male agonistic interactions in mice.

Nonsocial Interactions

Reproductive events are also affected by stress encountered in nonsocial situations. The endocrine system appears to be more susceptible to emotional stress than to physical stress. For example, immobilized female rats display an immediate loss of LH pulsatile activity, but no change in pulsing was observed after an experimenter administered limb fracture. The stressful nature of immobilization has also been observed in cattle restricted for the first time by a stanchion, haltered, and prepared surgically for sequential blood collection via jugular cannulae. In these conditions, greatly elevated plasma cortisol levels and no pulsatile release of LH were noticed. However, repeated exposure to this procedure resulted in low levels of plasma cortisol and a normal episodic pattern of LH release.

Stress Signaling Mediated by Pheromones

Pheromone detection is an important component in the animal's ability to respond appropriately to stressful cues in social and agonistic interactions. It is well documented in several mammalian species that animals experiencing stress and fear release chemical signals that induce behavioral, endocrinological, and immunological changes in the recipient animals of the same species. Aggressive behavior is displayed in many species in day-to-day social interactions. This behavior is elicited by the appropriate recognition of an adversary as a potential threat or sexual competitor, and this recognition involves the detection of social odors by the VNS and/or olfactory system. The involvement of the VNS in such recognition is clearly demonstrated by several studies. Genetic ablation of the TrpC2 channel results in the failure of the male mice to display aggression toward other males, similar to behavior shown by male mice after vomeronasectomy, and also eliminates aggressive behavior of lactating female mice. However, in contrast, vomeronasectomy does not reduce the maternal aggression of lactating female rats. Vomeronasectomy also has been shown to decrease aggressive behavior in the male lesser mouse lemur, a prosimian.

In socially stressed crowded female rats, the removal of the VNO eliminates the resumption of estrous cycling induced by male rat urine. Mechanisms underlying the VNO-dependent responses to sociogenic stress are largely unknown at this time. Finding neural sites in which the stress-activated HPA axis interacts with the vomeronasal system seems a good strategy for determining where to begin research from a mechanistic point of view.

See Also the Following Articles

Aggressive Behavior; Crowding Stress; Reproduction, Effects of Social Stress on.

Further Reading

- Boehm, U., Zou, Z. and Buck, L. B. (2005). Feedback loops link odor and pheromone signaling with reproduction. *Cell* **123**, 683–695.
- Brennan, P. A. and Keverne, E. B. (2004). Something in the air?: new insights into mammalian pheromones. *Current Biology* **14**, R81–R89.
- Bronson, F. H. (1988). Seasonal regulation of reproduction in mammals. In: Knobil, E. & Neill, J. D. (eds.) *The physiology of reproduction* (vol. 2), pp. 1831–1871. New York: Raven Press.
- Bronson, F. H. (1995). Seasonal variation in human reproduction: environmental factors. *Quarterly Review of Biology* **70**, 141–164.
- Calegro, A. E., Bagdy, G. and D'Agata, R. (1998). Mechanisms of stress on reproduction: evidence for complex intra-hypothalamic circuit. *Annals of the New York Academy of Sciences* **851**, 364–370.
- Dudley, C. A., Rajendren, G. and Moss, R. L. (1996). Signal processing in the vomeronasal system: modulation of sexual behavior in the female rat. *Critical Review of Neurobiology* **10**, 265–290.
- Dudley, C. A., Sudderth, S. B. and Moss, R. L. (1992). LHRH neurons in the medial septal-diagonal band-preoptic area do not project directly to the hippocampus: a double-labeling immunohistochemical study. *Synapse* **12**, 139–146.
- Dulac, C. and Torello, T. A. (2003). Molecular detection of pheromone signals in mammals: from genes to behaviour. *Nature Reviews Neuroscience* **4**, 551–562.
- Halpern, M. and Martinez-Marcos, A. (2003). Structure and function of the vomeronasal system: an update. *Progress in Neurobiology* **70**, 245–318.
- Kumar, A., Dudley, C. A. and Moss, R. L. (1999). Functional dichotomy within the vomeronasal system: distinct zones of neuronal activity in the accessory olfactory bulb correlate with sex-specific behaviors. *Journal of Neuroscience* **19**(20), RC32.
- Luo, M., Fee, M. S. and Katz, L. C. (2003). Encoding pheromonal signals in the accessory olfactory bulb of behaving mice. *Science* **299**, 1196–1201.
- Marchlewska-Koj, A. (1997). Sociogenic stress and rodent reproduction. *Neuroscience and Biobehavioral Review* **21**, 669–703.
- Meredith, M. and Fernandez-Fewell, G. (1994). Vomeronasal system, LHRH, and sex behaviour. *Psychoneuroendocrinology* **19**, 657–672.
- Mombaerts, P. (2004). Genes and ligands for odorant, vomeronasal and taste receptors. *Nature Reviews Neuroscience* **5**, 263–278.
- Moss, R. L., Flynn, R. E., Shen, X. M., et al. (1997). Urine-derived compound evokes membrane responses in mouse vomeronasal receptor neurons. *Journal of Neurophysiology* **77**, 2856–2862.
- Przekop, F. and Tomaszewska, D. (1996). Responses in the hypothalamic monoaminergic system activity in ewes to beta-endorphin, CRF, and their antagonists. *Acta Neurobiologiae Experimentalis* **56**, 807–817.
- Rajendren, G., Dudley, C. A. and Moss, R. L. (1990). Role of the vomeronasal organ in the male-induced enhancement of sexual receptivity in female rats. *Neuroendocrinology* **52**, 368–372.
- Rajendren, G., Dudley, C. A. and Moss, R. L. (1993). Influence of male rats on the luteinizing hormone-releasing hormone neuronal system in female rats: role of the vomeronasal organ. *Neuroendocrinology* **57**, 898–906.
- Yoon, H., Enquist, L. W. and Dulac, C. (2005). Olfactory inputs to hypothalamic neurons controlling reproduction and fertility. *Cell* **123**, 669–682.

Phobias See: Anxiety.

Photoperiodism See: Seasonal Changes in Stress Responses.

Physical Injury See: Combat, Acute Reactions to; Loss Trauma; Multiple Trauma.

Pituitary Regulation, Role of

F A Antoni

Centre for Integrative Physiology, University of
Edinburgh, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
F A Antoni, volume 3, pp 169–174, © 2000, Elsevier Inc.

The Endocrine Response to Stress

General Overview of the Control of the Pituitary Gland

Neurohypophyseal Responses to Stress

Adenohypophyseal Hormones in Stress

Intermediate Lobe

Glossary

Neuroendocrine motoneurons

Hypothalamic nerve cells that produce and secrete neurohormones into the hypophyseal portal circulation and thus regulate the synthesis and release of anterior pituitary hormones.

Pituitary gland

The master gland of the endocrine system; as such, it mounts the primary endocrine response to stress, alongside the cells of the adrenal medulla, which are controlled by the autonomic nervous system.

The Endocrine Response to Stress

The stress response is characterized by marked changes of the levels of several hormones in blood. In the short term, the hormonal response to stress facilitates the mobilization of nutrients, e.g., breakdown of glycogen, promotes cardiovascular adaptation, e.g., increase of heart rate, the preservation of body fluid and salt, e.g., increase of salt reabsorption, as well as behavioral responses, e.g., explorative behavior. Exactly which hormone is responsible for a particular function may vary from species to species. For instance, the pituitary hormone prolactin is a hormone of salt-water balance in fish, while in mammals it is involved in the development of the breast in pregnancy and milk production during lactation. Of course, this latter process is a major strain on the osmotic balance of the mammalian organism, hence the phylogenetic traces of osmoregulation are evident, but only during breastfeeding and only in the female. Remarkably, prolactin is acutely released during stress in virtually all species, including male mammals; however, the functional significance of this response is still not understood.

If the stressor is chronic, i.e., recurs repeatedly or is never removed, the stress response may undergo habituation. However, in many cases this is not sufficient to prevent long-term detrimental changes of physical condition, such as reduced resistance to infectious agents, hypertension, and behavioral anomalies.

General Overview of the Control of the Pituitary Gland

The pituitary gland consists of three functionally and structurally distinct compartments ([Figure 1](#)).

The neurohypophysis or the posterior lobe is a large array of axon terminals and their supporting glial cells called the pituicytes. The neuronal perikarya that give rise to the terminal axons in the neurohypophysis lie in the hypothalamic paraventricular and supraoptic nuclei and form the magnocellular neurosecretory system ([Figure 1](#)). Separate populations of these cells produce the small peptide hormones vasopressin or oxytocin, which are synthesized in the cell bodies, transported along the axons, and released into the systemic circulation by the nerve endings in the neurohypophysis.

The adenohypophysis or anterior lobe contains various types of cells that are derived from the neural placode epithelium and produce polypeptide hormones for export into the systemic circulation. These hormones are fundamentally important for the control of the endocrine system. The cell types and their secreted hormones are as follows: corticotrophs, adrenocorticotropin (ACTH); mammothrophs, prolactin (PRL); somatotrophs, growth hormone (GH); somatomammothrophs, GH and PRL; thyrotrophs, thyroid-stimulating hormone (TSH); and gonadotrophs, luteinizing hormone (LH) and follicle-stimulating hormone (FSH). In addition, a further type of cell, the folliculo-stellate cells that are of mesenchymal origin, produce various cytokines and peptides that may be important for local paracrine control of the adenohypophysis.

The functional activity of the cells in the adenohypophysis is controlled by neurohormones (hypothalamic-releasing and -inhibiting factors) of hypothalamic origin. These are synthesized by neurons of the medial hypothalamus and secreted by axons that terminate in the external zone of the median eminence, which is a specialized neurovascular contact site at the base of the brain. This is called the parvicellular hypophyseotropic neurosecretory system of the hypothalamus ([Figure 1](#)).

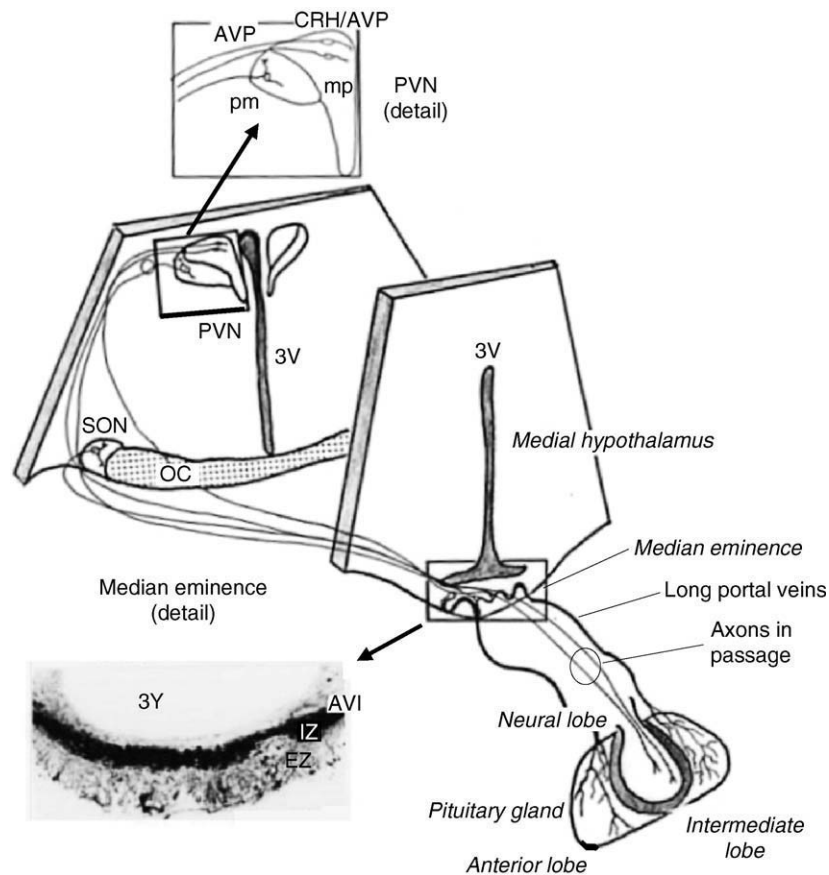


Figure 1 Schematic representation of the hypothalamo-hypophyseal system that controls the secretion of ACTH. Cell bodies of parvocellular hypophysiotropic neurons synthesize corticotropin releasing hormone (CRH) and vasopressin (AVP) in the hypothalamic paraventricular nucleus (PVN). The axons reach the median eminence along an archlike pathway. The fibers first pass laterally above the fornix (F) and then turn ventral and caudal to reach the rostral part of the median eminence through the lateral retrochiasmatic area at the base of the brain, behind the optic chiasm (OC). Once in the median eminence, the fibers form terminal varicosities abutting on the pericapillary space in the external zone (EZ) of the median eminence. The effluent blood collects in the long portal veins and irrigates the anterior lobe of the pituitary gland. Magnocellular hypothalamic neurons in the PVN and the supraoptic nucleus (SON) produce AVP. The axons of these cells take a similar course in the hypothalamus as the parvocellular axons to reach the median eminence. In the median eminence, the magnocellular axons remain largely in the internal zone and appear to release AVP through conventional exocytotic mechanisms. AVP released in this manner is also found in hypophyseal portal blood. The coronal section of the median eminence shown here (median eminence detail) was stained with an antiserum against vasopressin-associated neurophysin and was prepared from the brain of an adrenalectomized rat. Adrenalectomy markedly increases the AVP content in the external zone. 3V, third cerebral ventricle; pm, magnocellular part of the PVN; mp, medial parvocellular compartment of the PVN.

The capillary blood vessels in the median eminence lie immediately adjacent to the parvocellular nerve terminals and have special large pores to allow diffusion of neurohormones released by the nerve fibers into the blood irrigating the median eminence. The venous effluent of the median eminence carries the hypothalamic hormones to the adenohypophysis, where the capillaries again branch out into smaller vessels, forming the pituitary portal system that irrigates the adenohypophysis. The transit time in this system from the nerve endings to the hormone producing cells is approximately 5 s.

The internal zone of the median eminence contains the varicose nerve fibers of the magnocellular neurons

running to the neurohypophysis. Functional evidence indicates that vasopressin and oxytocin are also secreted in the median eminence by these fibers in passage, thus neurohypophyseal secretions may also control anterior pituitary function.

In addition to hypothalamic and paracrine pituitary-derived factors, the function of anterior pituitary cells may also be influenced by factors reaching the adenohypophysis via the systemic circulation. For instance, steroid hormones produced and secreted by the adrenal cortex influence the synthesis and the release of ACTH. Cytokines derived from the immune system may also target these cells. The actions in the adenohypophysis of factors derived

from the endocrine target glands of anterior pituitary hormones are important for the formation of neuroendocrine feedback loops.

The intermediate lobe of the pituitary gland is composed almost entirely of cells that produce melanocyte-stimulating hormone, which is important in species that change their skin color rapidly as part of their adaptive response. The intermediate lobe is innervated directly by nerve fibers originating from the medial basal hypothalamus. It has no obvious functional anatomical communication with the other two lobes.

Neurohypophyseal Responses to Stress

Both vasopressin and oxytocin, the neurohormones released from the nerve terminals in neurohypophysis, have multiple functions in the body.

The principal function of neurohypophyseal oxytocin is the contraction of the uterus and the ejection of milk during breastfeeding. Actions on the kidney are controversial. Oxytocin may also contribute to the regulation of prolactin secretion by the adenohypophysis, as receptors and actions on prolactin release have been described. Some stressors, e.g., immobilization, increase the plasma levels of oxytocin in experimental animals; however, the importance of this in the stress response is not understood.

Vasopressin derived from the neurohypophysis increases blood pressure by contracting blood vessels, increases water reabsorption in the kidney (hence its other name, antidiuretic hormone [ADH]), and promotes glycogenolysis in the liver. The relative importance of these effects shows considerable species variation. In humans only the actions on the kidney appear relevant.

A further important action of vasopressin in the body is the stimulation of ACTH secretion. At physiological concentrations of vasopressin, this action appears to require a cohormone called corticotropin-releasing hormone (CRH). Vasopressin derived from the neurohypophysis is distributed in the systemic circulation and reaches levels required to influence ACTH release (0.1–2 nM) only in larger mammals (sheep, cat) where peripheral levels of vasopressin may reach 1–2 nmol/L. In rodents and humans, plasma levels rarely exceed 10 pmol/L. However, vasopressin released by the fibers of magnocellular neurons in passage through the median eminence and by their dendrites in the paraventricular nucleus (Figure 1) significantly contributes to the control of ACTH secretion.

The magnocellular neurons that produce vasopressin are often referred to as the osmosensitive neurosecretory system. It is important to bear in mind that

these neurons are osmosensitive in the widest sense of the word. On the one hand, changes of plasma osmolality will directly enhance the firing rate and secretory activity of these cells. On the other, conditions that potentially threaten osmotic homeostasis and body fluid volume such as strenuous exercise, pain, and hemorrhage all stimulate the release of vasopressin through neural afferent pathways.

Adenohypophyseal Hormones in Stress

The two main anterior pituitary hormones released by all types of stress are ACTH and PRL. While the physiological significance of ACTH in the stress response is firmly established, the role of PRL is not well understood. The stress-induced increase of plasma PRL is, however, important in the clinic as it poses a potential problem for the differential diagnosis of hyperprolactinemia of other causes, e.g., pituitary tumors.

Hypothalamic Control of ACTH

Like most midsize (39 amino acid residues) polypeptide hormones, ACTH is initially synthesized as a larger precursor molecule called proopiomelanocortin. In the anterior pituitary gland the main hormones derived from this precursor are ACTH and β -lipotropin. In target tissues such as the adrenal cortex extracellular proteases may cleave these hormones further to other biologically active fragments such as β - and γ -melanocyte-stimulating hormone.

The main mediators of the hypothalamic control of ACTH are CRH and vasopressin. CRH is produced by parvicellular hypophysiotropic neurons located in the medial parvicellular part of the hypothalamic paraventricular nucleus (Figure 1). These parvicellular neurons also produce vasopressin, which is co-packaged with CRH and co-released from the same secretory vesicles. Significantly, the relative amount of CRH and vasopressin synthesized and released by parvicellular neurons shows wide variation. In humans the level of vasopressin in parvicellular neurons increases with age. In experimental animals, vasopressin synthesis in parvicellular hypophysiotropic neurons may be enhanced from one- to twofold that of CRH to eight- to 10-fold CRH by various forms of chronic or repeated stress as well as adrenalectomy. This is somewhat paradoxical, as chronic stressors cause persistent increases of plasma corticosteroids whereas adrenalectomy brings about a drastic reduction of these hormones. The likely explanation for these findings is that the actions of corticosteroids on hypothalamic neurons are context dependent, i.e., are a function of the state of activation of other signaling pathways that regulate

the expression of vasopressin in parvicellular hypophysiotropic neurons.

As outlined in the previous section, the magnocellular vasopressinergic system is also involved in the control of anterior pituitary function. The innervation and regulation of these cells by afferent neural pathways are distinct from that of the parvicellular CRH neurons. However, the dendrites of vasopressinergic magnocellular neurons form synaptic contacts with CRH-containing parvicellular neurons, and vasopressin appears to influence the level of CRH in hypophysial portal blood. Thus, there is intrahypothalamic interaction between the magnocellular and parvicellular systems controlling the adeno-hypophysial secretion of ACTH. In addition, vasopressin derived from magnocellular axons in the median eminence is found in hypophysial portal blood.

A prominent functional interaction between CRH and vasopressin is evident at the level of the pituitary gland. CRH stimulates the release of ACTH through the intracellular messenger cyclic AMP. Vasopressin alone has little or no effect on ACTH release or cAMP synthesis, respectively, but synergistically enhances the effect of CRH on both ACTH release and cellular cAMP content. The basis of this is the augmentation of CRH-induced cAMP synthesis by vasopressin through protein kinase C. Thus, the adenylyl cyclase system of the adeno-hypophysial corticotroph cell appears to behave as a coincidence detector; it responds to vasopressin only in the presence of CRH. It is therefore likely that corticotroph cells express specialized adenylyl cyclase proteins (see **Adenylyl Cyclases and Stress Response**) that are activated by protein kinase C in the presence of the stimulatory G protein G_s . Indeed, two species of protein kinase C-stimulated adenylyl cyclase (isoforms 2 and 7) have been found in the adeno-hypophysis, and evidence for different pathways of cAMP biosynthesis induced by CRH alone and CRH and vasopressin in combination, respectively, has been reported. There are some species variations in the potency of CRH and vasopressin to stimulate ACTH. In the sheep, vasopressin appears to be the more potent ACTH secretagogue.

The significance of vasopressin in the control of ACTH release lies in the fact that in the presence of this neurohormone corticotroph cells become relatively resistant to the negative feedback action (see below) of adrenal corticosteroids. The opposition of glucocorticoid feedback is inversely correlated to the levels of cAMP, indicating that the intracellular levels of cAMP are major determinants of the efficiency of corticosteroid feedback inhibition at the pituitary level. The concentration of vasopressin required for inducing corticosteroid resistance is in the range of 1–2 nM, which requires either vasopressin derived

from magnocellular neurons or an increase of vasopressin expression and biosynthesis by parvicellular neurons such as that seen in chronic stress.

Peripheral and Paracrine Factors

The release of cytokines during inflammation appears to influence anterior pituitary ACTH release. For example, interleukin-1 induces ACTH secretion by an action in the brain and possibly directly on corticotroph cells. Of further interest is the production of leukocyte-inhibiting factor (LIF) and migration inhibitory factor (MIF) in corticotroph cells. LIF appears to be important for the maintenance of corticotroph functions, while MIF may attenuate corticosteroid feedback.

Corticosteroid Feedback Inhibition

Generally the plasma levels of adrenal corticosteroids are tightly controlled by a neuroendocrine feedback loop. Identified targets of corticosteroid action are adeno-hypophysial corticotrophs and parvicellular hypophysiotropic neurons that produce CRH and vasopressin. These are clearly sites of physiological relevance that control the final common elements of the neuroendocrine pathway that leads to the stimulation of ACTH. Under certain conditions, some of which are clearly pathological, glucocorticoid feedback inhibition is reduced. A well-characterized physiological example in humans is strenuous exercise. Cushing's disease and Nelson's syndrome are characterized by compromised glucocorticoid feedback at the pituitary level, which may be due to the tumoral transformation of corticotrophs. Furthermore, depressed patients have abnormally high levels of plasma corticosteroids and reduced feedback inhibition. In animals, chronic autoimmune inflammation produces high circulating levels of ACTH and corticosteroids. At the hypothalamic level this condition is characterized by a marked increase of vasopressin expression in parvicellular CRH neurons of the paraventricular nucleus. Furthermore, surgical stress and hemorrhage are also characterized by reduced corticosteroid feedback at the pituitary level, and it is known that under these conditions the magnocellular vasopressinergic neurons become active. Overall, this evidence indicates that vasopressin is mobilized in the short term as well as in the long term to modulate the efficiency of corticosteroid feedback inhibition of pituitary ACTH secretion.

Intermediate Lobe

The intermediate lobe is stimulated by epinephrine and norepinephrine derived from the adrenal medulla as well as by hypothalamic CRH. It is normally inhibited

by dopaminergic input from the arcuate nucleus. The physiological role of the intermediate lobe in adult humans is unclear. In rodents, stress induces the secretion of alpha melanocyte stimulating hormone (α -MSH), which is not susceptible to corticosteroid feedback inhibition.

See Also the Following Articles

Adenylyl Cyclases and Stress Response; Vasopressin.

Further Reading

- Antoni, F. A. (1986). Hypothalamic regulation of adrenocorticotropin secretion: advances since the discovery of 41-residue corticotropin-releasing factor. *Endocrine Reviews* 7, 351–378.
- Antoni, F. A. (1993). Vasopressinergic control of anterior pituitary adrenocorticotropin secretion comes of age. *Frontiers in Neuroendocrinology* 14, 76–122.
- Antoni, F. A. (1996). Calcium checks cyclic AMP – mechanism of corticosteroid feedback in adeno-hypophysial corticotrophs. *Journal of Neuroendocrinology* 8, 659–672.
- Antoni, F. A., Sosunov, A. A., Haunso, A., Paterson, J. M. and Simpson, J. (2003). Short-term plasticity of cyclic adenosine 3',5'-monophosphate signaling in anterior pituitary corticotrope cells: the role of adenylyl cyclase isoforms. *Molecular Endocrinology* 17, 692–703.
- Calandra, T., Bernhagen, J., Metz, C. N., et al. (1995). MIF as a glucocorticoid-induced modulator of cytokine production. *Nature* 377, 68–71.
- Chowdrey, H. S., Larsen, P. J., Harbuz, M. S., et al. (1995). Evidence for arginine vasopressin as the primary activator of the HPA axis during adjuvant-induced arthritis. *British Journal of Pharmacology* 116, 2417–2424.
- Dallman, M. F., Akana, S. F., Scribner, K., et al. (1992). Stress, feedback and facilitation on the hypothalamo-pituitary-adrenal axis. *Journal of Neuroendocrinology* 4, 517–526.
- de Kloet, E. R., Vreugdenhil, E., Oitzl, M. S. and Joëls, M. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Reviews* 19, 269–301.
- Engelmann, M., Landgraf, R. and Wotjak, C. T. (2004). The hypothalamic-neurohypophysial system regulates the hypothalamic-pituitary-adrenal axis under stress: an old concept revisited. *Frontiers in Neuroendocrinology* 25, 132–149.
- Jones, M. T. and Gillham, B. (1988). Factors involved in the regulation of adrenocorticotrophic hormone/ β -lipotropic hormone. *Physiology Reviews* 68, 743–818.
- Petrides, J. S., Mueller, G. P., Kalogeras, K. T., et al. (1994). Exercise-induced activation of the hypothalamic-pituitary-adrenal axis: marked differences in the sensitivity to glucocorticoid suppression. *Journal of Clinical Endocrinology and Metabolism* 79, 377–383.
- Plotsky, P. M. (1991). Pathways to the secretion of adrenocorticotropin: a view from the portal. *Journal of Neuroendocrinology* 3, 1–9.
- Tilders, F. J. H., Berkenbosch, F., Vermes, I., Linton, E. A. and Smelik, P. G. (1985). Role of epinephrine and vasopressin in the control of the pituitary-adrenal response to stress. *FASEB Journal* 44, 155–160.

PMS See: Premenstrual Dysphoric Disorder.

Police, Stress in

R J Burke

York University, Toronto, Ontario, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R J Burke, volume 3, pp 175–178, © 2000, Elsevier Inc.

Glossary

<i>Burnout</i>	Response to chronic work stress characterized by exhaustion, cynicism, and low feelings of accomplishment.
<i>Coping</i>	Responses by individuals to manage work stressors.

Work Stressors

In comparison with other occupations, police work has been identified as a particularly stressful occupation. Negative aspects of the job such as boredom,

Work Stressors

Psychological Burnout

Organizational Initiatives

Individual Initiatives

lack of respect from members of the public, excessive paperwork, contacts with the public that are sometimes negative and confrontational, shift work, threats of violence, and the militaristic nature of the bureaucratic structure of policing are among the work stressors that confront police officers. As a result of these and other stressful aspects of policing, a variety of symptoms and reactions may occur. These include deteriorating work performance (absenteeism, low morale), negative psychological states (emotional burnout, frustration, depression, anger), and psychosomatic and physical conditions (headaches, ulcers). The stresses of policing have been offered as one reason why capable police officers leave the profession for other careers.

Figure 1 shows the wide variety of potential stressors in police work. These include external demands and factors in the broader social context, aspects of the criminal justice system, the police organization and the police role itself, aspects of individual differences, and work–family issues. In addition, individual characteristics of police officers (e.g., authoritarian personality) may increase levels of work stress. Some of these stressors are common to a number of occupations while others are unique to police work.

The police organization and the police job pose particular demands. Police officers, particularly early in their careers, work varying shifts. In addition, they often are on-call. Overtime work is an inevitable part of their job. These factors result in heavy workloads in terms of hours per week spent working.

The police role also can be a concern. Police officers are required to be suspicious and hypervigilant. They are trained to exercise emotional control. Finally, their role conveys considerable authoritativeness. Police officers are always right and have considerable power and authority over citizens.

The peer culture and socialization influences in policing are also somewhat unique. These include the presence of masculine attitudes and values (e.g., control, dominance, authority), an extreme in-group/out-group mentality, and intense solidarity with peers. Police are likely to socialize with other police officers. Since many contacts with citizens are negative in nature, this fosters a mentality of us versus them.

Cynicism among police officers has received considerable attention as well. Officers begin their careers with positive attitudes toward their jobs and their role. These attitudes have been shown to steadily deteriorate as officers approach mid-career, then

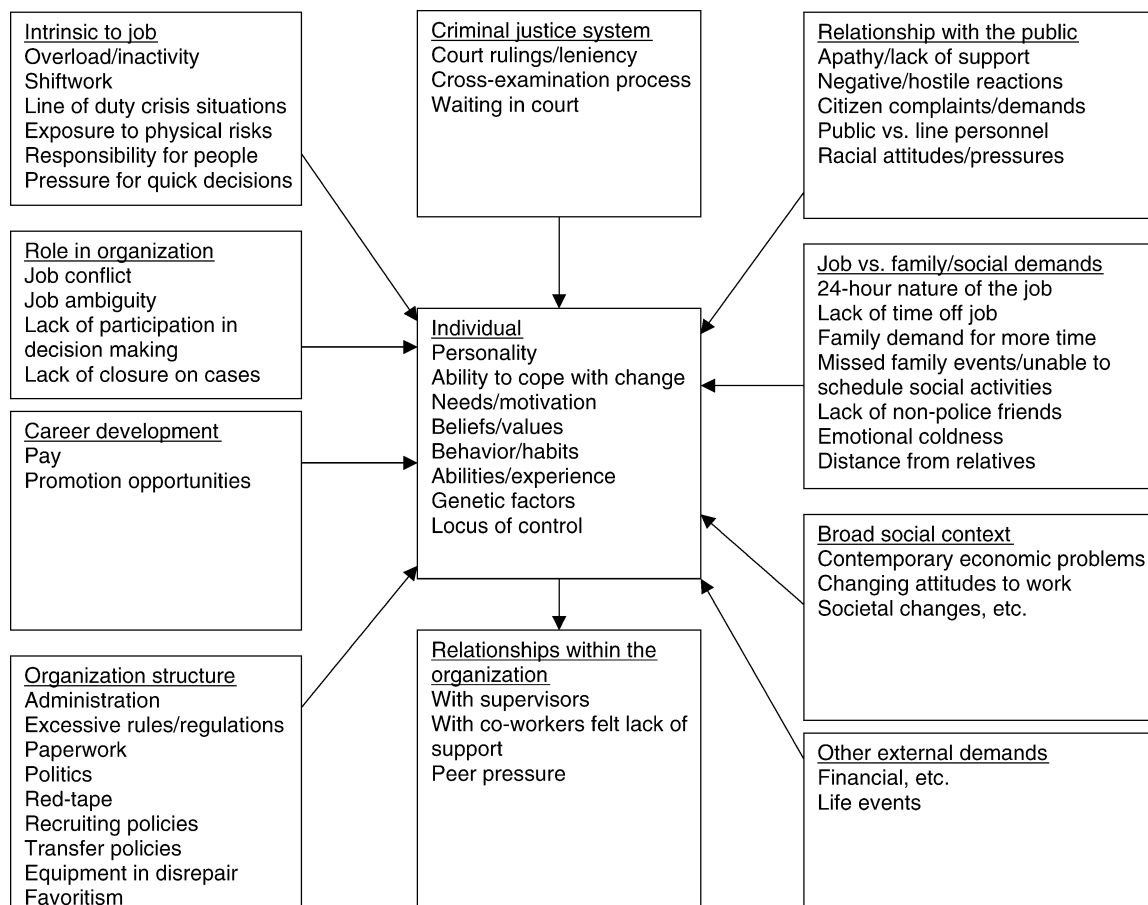


Figure 1 Understanding stress in policing.

improve modestly with increasing police tenure as officers approach retirement. These attitude changes are important to both understand and manage since police officers are vulnerable to career plateauing given the steep organizational hierarchy in police management. The challenge for senior police management is to maintain positive attitudes among officers who will never be promoted.

These factors produce some common problem areas. Police officers become overinvolved with their force and underinvolved with their spouses or partners. The cynicism police have about their profession and the broader justice system spills over to the home front. Police officers make authoritative demands at home, as they do in their work. The irregular schedule and the leave time demands results in limited time with their spouses and families. Spouses or partners often have fear or anxiety about the safety of the officer. The spouses or partners are often jealous of the time that police officers spend with co-workers. Police families are sometimes isolated from the broader community. Finally, there are weapons in the home.

Work stressors in policing have been found to be related to work–family conflict, which in turn has been related to psychological burnout. Jackson and Maslach collected data from men and women in police work and their spouses or partners. They found that police officers' levels of burnout affected their interactions with their spouses and children. Burke examined antecedents and consequences of work–family conflict among 828 men and women in police work. Ratings of work–family conflict were weakly correlated with demographic antecedents but strongly correlated with work setting characteristics, levels of social support, levels of work and non-work stress, and several outcome measures (including burnout).

The combination of the types of individuals attracted to police work and the nature of the job has been suggested to increase particular problematic behaviors. These behaviors include anger, verbal outbursts, physical violence, alcoholism, and infidelity. Forty percent of police officers reported at least one episode of physical aggression during a conflict with their spouse or partner in the previous year.

Some published literature has downplayed the negative effects on officers of participating in risky calls and has pointed to the organizational structure as the greatest source of stress and pressure, adversely affecting job satisfaction and the perceived quality of life.

Although most police officers are men, an increasing number of women have become officers over the past decade. About 15 to 20% of officers in the developed world are women. Women and men in police work generally report similar levels of work

stressors, job satisfactions, burnout, and psychological well-being. Interestingly, the biggest source of job stress reported by women officers comes primarily from their male colleagues and to a lesser extent from male members of the general public. Women officers report prejudice, discrimination, and sexual harassment from both of these sources.

Psychological Burnout

Research attention has increasingly focused on psychological burnout, a common reaction to stress among police officers. Maslach and Leiter defined burnout as a syndrome of emotional exhaustion, depersonalization, and low personal accomplishment. Jackson and Maslach studied the effects of levels of self-reported psychological burnout on police officers' interactions with their spouses and children. Police officers reporting higher levels of psychological burnout were more likely to display anger, spend time off away from the family, be uninvolved in family matters, and have unsatisfactory marriages.

Empirical studies of psychological burnout among police officers have attempted to identify both individual difference characteristics and job conditions associated with higher burnout levels. Relationships of burnout with individual demographic characteristics are sometimes found but tend to be both inconsistent and weak. However, there is considerable evidence that work setting characteristics, particularly chronic work stressors, influence levels of psychological burnout. Such characteristics have included features of the job itself, quality of supervision, unmet expectations, and constraints in one's organizational environment.

Police forces are concerned when their officers use excessive force toward citizens. Recent research has shown a relationship between officer burnout levels and both attitudes toward the use of force and the actual use of force. Burned-out officers were more likely to advocate and use force and violence. Cynicism toward the general public lowers the threshold for using force or violence. Exhaustion and low professional efficacy reduce problem-solving skills while increasing use of force and violence.

Organizational Initiatives

Recruitment and Selection Interventions

Police work is a relatively high-paying occupation. As a result, a large number of individuals are interested in developing careers in policing. It seems desirable for police forces to use realistic job previews so that applicants have a greater understanding of the nature

and demands of the occupation, make greater use of personality assessment to weed out applicants likely to have a difficult time performing effectively, and use the training period to both socialize desired behaviors and identify applicants that may not perform well.

Counseling Interventions

It is critical to offer trauma counseling to police officers in cases of acute stress (using a weapon, shooting a suspect, death of a citizen in an accident). In addition, the provision of counseling to police officers via Employee Assistance Programs (EAPs) to handle alcohol abuse, work–family difficulties, and family problems is useful. The macho nature of the police culture and the real (or imagined) stigma of seeking help must be addressed.

Training Interventions

Two areas of training are evident. One specifically addresses work stressors and coping, particularly the use of active coping responses. Although such training has not been found to produce lasting benefits, it may still be justified. A second aspect of training, aimed at the organizational level, involves management development. It is unrealistic to expect that the nature and culture of policing will change dramatically or quickly, but efforts in this direction must be made.

Work–Family Initiatives

As noted earlier, the work–family interface can be particularly problematic in policing. Several work–family initiatives have been undertaken by innovative focus. These include a police spouse/partner support group, a couple workshop that both educates and addresses concrete issues, the inclusion of work–family content in training and development activities, and the use of peer counseling efforts.

Individual Initiatives

Although coping is a critical individual response to perceived work stressors due to its complexity, it is still not well understood. A simple taxonomy is to classify individual coping responses as active (problem solving, talk with others, anticipate upcoming

demands) or passive (do nothing, avoidance, use alcohol). Hurrell and Kroes reported that 25% of all police officers are dependent on alcohol as a stress reliever. In a large sample of police officers, Burke found that use of active coping responses was associated with greater job satisfaction, and use of passive coping responses was associated with greater psychological distress. Use of active and passive coping responses was, not surprisingly, negatively correlated. Police officers making greater use of active coping made lesser use of passive coping responses.

See Also the Following Articles

Work–Family Balance; Workplace Stress.

Further Reading

- Abdollahi, K. M. (2002). Understanding police stress research. *Journal of Forensic Psychology Practice* 2, 1–24.
- Brown, J. and Heidenshoen, F. (2000). *Gender and policing: comparative perspectives*. Basingstoke: MacMillan.
- Burke, R. J. (1997). Work and non-work stressors and well-being among police officers: the role of coping. *Anxiety, Stress & Coping* 11, 345–362.
- Cooper, W. H. (1982). Police officers over career stages. *Canadian Police College Journal* 6, 93–112.
- Hart, P. M., Wearing, A. J. and Headey, B. (1993). Assessing police work experiences: Development of the police daily hassles and uplift scales. *Journal of Criminal Justice* 21, 553–572.
- Jackson, S. E. and Maslach, C. (1982). After-effects of job-related stress: families as victims. *Journal of Occupational Behavior* 3, 63–77.
- Kop, N. and Euwema, M. C. (2001). Occupational stress and the use of force by Dutch police officers. *Criminal Justice and Behavior* 28, 631–652.
- Kroes, W. H. (1976). *Society's victim – the policeman. An analysis of job stress in policing*. Springfield, IL.: Charles C. Thomas.
- Kroes, W. H. and Hurrell, J. J. (1975). *Job stress and the police officer*. Washington, D.C.: U.S. Government Printing Office.
- Maslach, C. and Leiter, M. P. (1997). *The truth about burnout*. San Francisco: Jossey-Bass.
- Neidig, P., Russell, H. and Seng, A. (1992). Interpersonal aggression in law enforcement families: a preliminary investigation. *Police Studies* Spring 30–38.

Posttraumatic Stress Disorder – Clinical

N C Feeny and L R Stines

Case Western Reserve University, Cleveland, OH, USA

E B Foa

University of Pennsylvania, Philadelphia, PA, USA

© 2007 Elsevier Inc. All rights reserved.

Overview of Posttraumatic Stress Disorder
Cognitive-Behavioral Treatments for Posttraumatic
Stress Disorder
Real-World Treatment Considerations
Childhood Sexual Abuse-Related Posttraumatic
Stress Disorder
Conclusion

Glossary

<i>Arousal</i>	A cluster of PTSD symptoms that includes difficulty concentrating, sleep disturbance, irritability and anger, exaggerated startle response, and overalertness.
<i>Avoidance</i>	A cluster of PTSD symptoms that includes behavioral and cognitive avoidance (i.e., avoidance of thoughts, feelings, and reminders of the trauma), emotional numbing, loss of interest in activities, feeling disconnected from others, psychogenic amnesia, and a sense of foreshortened future.
<i>Cognitive-behavioral treatments</i>	A class of psychosocial treatment approaches focused on the interrelationships among thoughts, behaviors, and emotions.
<i>Posttraumatic stress disorder (PTSD)</i>	An anxiety disorder that develops in response to a traumatic event and is characterized by three clusters of symptoms: reexperiencing, avoidance, and physiological hyperarousal. PTSD is diagnosed after a 1-month period of sustained problems following the traumatic event.
<i>Reexperiencing</i>	A cluster of PTSD symptoms that includes reexperiencing the traumatic event through intrusive distressing thoughts, nightmares, flashbacks, and intense emotional upset and physiological arousal on exposure to reminders of the trauma.
<i>Trauma</i>	An event that a person experiences, witnesses, or learns about that involves actual or threatened physical injury or death and to which the person responds with feelings of fear, helplessness, or horror.

Overview of Posttraumatic Stress Disorder

Prevalence of Traumatic Events

According to the *Diagnostic and Statistical Manual of Mental Disorders* (4th edn.; DSM-IV) a trauma is an event that involves actual or perceived physical threat or injury, to which a person responds with fear, helplessness, or horror. Although in the past, trauma was conceptualized as unusual and outside the realm of human experience, large epidemiological studies have demonstrated high rates of trauma exposure among adults. In a nationally representative sample of 6000 U.S. adults, Kessler and colleagues found that 60% of men and 51% of women reported experiencing at least one traumatic event in their lifetime. Similar findings emerged from Norris's 1992 study using a large, racially diverse sample – 74% of men and 65% of women reported experiencing at least one traumatic event. A lower prevalence of trauma was found in a sample of members of a health maintenance organization – 43% of men and 39% of women experienced a trauma in their lifetime. Taken together, these studies suggest that trauma is very common among adults in the United States.

Diagnostic Criteria of Posttraumatic Stress Disorder

The constellation of psychological difficulties that is observed most often following a trauma is called posttraumatic stress disorder (PTSD). PTSD is an anxiety disorder, and its symptoms fall into three clusters: reexperiencing the traumatic event (e.g., intrusive and distressing thoughts, flashbacks, and nightmares), avoidance of trauma-related reminders (e.g., avoiding thoughts or situations related to the trauma, emotional numbing, sense of foreshortened future), and hyperarousal (e.g., sleep disturbance, difficulty concentrating, irritability, and hypervigilance). In order to meet diagnostic criteria for PTSD, these difficulties must last for more than 1 month and cause substantial impairment in social or occupational functioning.

Prevalence and Course of Posttraumatic Stress Disorder

In the general population, the lifetime prevalence of PTSD is estimated to be around 9%. Estimates suggest that women are two times more likely to develop PTSD than men (e.g., 10% of women vs. 5% of men

in the general population). Because a diagnosis of PTSD requires an exposure to a traumatic event, rates of PTSD are inherently higher among those who have experienced such an event, with lifetime prevalence estimated at 25%. Among trauma survivors, the prevalence of current PTSD varies quite widely. For example, 15% of Vietnam combat veterans, 12–65% of female assault survivors, and up to 40% of survivors of serious motor vehicle accidents have been reported to meet diagnostic criteria for PTSD. Thus, although trauma exposure is quite common, persistent psychological difficulties in the aftermath of a traumatic event are less so and vary in frequency, in part, as a function of the traumatic experience.

In the initial days and weeks following a trauma, it is common to experience difficulties such as fear, disrupted sleep, and difficulty concentrating. Although most people experience such difficulties shortly after a traumatic event, the majority also experience a natural reduction in these difficulties over the following several months. Some trauma survivors, however, continue to experience significant PTSD symptoms for months and even years after exposure to the traumatic event. After 1 year, chronic PTSD is unlikely to remit without treatment; thus there is a critical need for effective treatments.

Cognitive-Behavioral Treatments for Posttraumatic Stress Disorder

Several cognitive-behavioral therapies for PTSD have been developed and tested in prospective randomized studies. These treatments include prolonged exposure therapy (PE), stress inoculation training (SIT), cognitive therapy (CT), cognitive processing therapy (CPT), and eye movement desensitization and reprocessing (EMDR). In this section, these treatments are reviewed and selected outcome studies that evaluate the efficacy of these treatments are presented.

Prolonged Exposure Therapy

Exposure therapy has been extensively evaluated as a treatment for persistent, pathological anxiety. The chief aim of this form of treatment is to help patients confront feared and avoided objects, situations, memories, and images until anxiety declines. PE for trauma survivors with PTSD is a treatment program that is based on the notion that avoidance of trauma-related memories and external reminders interferes with recovery from the traumatic event. Specifically, avoidance prevents opportunities to emotionally process the trauma memory, to learn to distinguish safe from unsafe situations, and to disconfirm erroneous trauma-related cognitions, thus maintaining the

unrealistic cognitions underlying PTSD that the world is utterly dangerous and that the trauma survivor is extremely incompetent in coping with stress. This view of the mechanisms involved in exposure therapy is central to emotional processing theory.

PE includes two forms of exposure: imaginal exposure and *in vivo* exposure. Imaginal exposure involves repeatedly recounting the traumatic memory. Specifically, the patient is instructed to vividly imagine the traumatic event and to describe it in detail aloud, along with the thoughts and feelings that occurred during the event. Imaginal exposure is typically conducted during several treatment sessions and via homework of listening to audiotapes of the in-session recounting of the memory. *In vivo*, or real-life, exposure involves systematic and repeated confrontation with safe or low-risk trauma-related reminders that evoke excessive or unrealistic anxiety. *In vivo* exposure is typically conducted outside of the treatment sessions as homework, during which the patient confronts situations, places, objects, or activities that are realistically safe but that trigger trauma-related fear and anxiety because they are related to the trauma. With *in vivo* exposure, the patient is encouraged to remain in the anxiety-provoking situation until his or her fear decreases by a significant amount. Other components of PE include psychoeducation about common reactions to traumatic events and breathing retraining.

Across several studies in different centers, PE has consistently been found to be highly effective. PE has demonstrated outcomes superior to supportive counseling (SC) and a wait-list control (WL). Most patients who receive PE experience a significant reduction in PTSD, depression, anxiety, anger, and guilt and experience improvements in overall functioning. In fact, in the treatment guidelines developed by experts in the field in conjunction with International Society for Traumatic Stress Studies (ISTSS), PE was identified as the most empirically supported intervention for PTSD.

Stress Inoculation Training

SIT was among the first of the cognitive-behavioral treatments to be applied to PTSD. According to the theory underlying SIT, stress occurs when people experience environmental events as exceeding their coping resources and thereby as threatening to their safety or welfare. Anxiety is a normal response to stress and signals people to increase efforts to cope with or somehow manage the stress-eliciting situation. In the case of individuals with PTSD, the anxiety evoked by trauma-related reminders is excessive and disruptive. Thus, SIT involves the learning and

repeated practice of specific coping skills, with the goal of helping patients to develop or enhance their ability to manage stress effectively and to reduce anxiety. Coping skills typically include tools such as breathing and relaxation training, guided self-dialog, assertiveness training, role-playing, covert modeling, and cognitive restructuring.

The efficacy of PE, SIT, SC, and WL were compared for female sexual-assault survivors with chronic PTSD. PE and SIT showed significant pre- to posttreatment reductions in reexperiencing and avoidance symptoms, whereas SC and WL did not. Also, at the end of treatment, 50% of patients who had SIT and 40% who had PE no longer met the criteria for PTSD; in contrast, only 10% of patients who had SC and none who had WL lost their diagnosis. At follow-up, there was a tendency for patients in the PE group to show further improvement in PTSD symptoms, whereas patients in the SIT and SC groups did not.

In a second study, the efficacy of PE and SIT was compared to the efficacy of a combination of both treatments (PE/SIT) and WL. All three active treatments resulted in considerable symptom reduction and were more effective than WL. However, PE appeared superior to the other treatments, leading to greater reductions of anxiety and depression, with fewer dropouts than both SIT and PE/SIT. At the end of treatment, 70% of patients who had PE, 58% who had SIT, and 54% who had PE/SIT lost their PTSD diagnosis compared to none in the WL group. At follow-up, women who received PE exhibited better social functioning than did those who received SIT or PE/SIT.

Cognitive Therapy

The aims of CT or cognitive restructuring (CR) for PTSD patients are to help patients to understand the role of their beliefs and appraisals of situations in influencing their emotional reactions, to identify trauma-related irrational thoughts or beliefs that trigger avoidance and/or excessive negative emotions (e.g., fear, shame, and rage), and to learn to challenge unrealistic beliefs and expectations in a rational evidence-based manner. In challenging these beliefs, evidence is weighed and alternative ways of viewing the situation are evaluated. In treatment sessions, and in daily life, patients practice responding to automatic thoughts and interpretations by reviewing the facts, considering alternative explanations, and sometimes experimenting with different ways of behaving in response to situations or events that elicit anxiety and other negative emotions. As a consequence, patients learn to evaluate whether their trauma-related beliefs and expectations accurately reflect

reality and are appropriate or helpful and modify them accordingly.

In a 2005 randomized trial, PE was compared with a program that included PE and CR (PE/CR) in female victims of sexual and nonsexual assault. The results suggest that both PE and PE/CR produce substantial improvements on multiple outcomes, including evaluator ratings of PTSD severity, depression, and social functioning. Similar to the PE/SIT findings, combining PE with CR did not improve the efficacy of PE. In an examination of PE in a sample of mixed trauma victims with chronic PTSD, Marks et al. compared PE to CR, PE/CR, and a relaxation control condition (R). The results suggest that PE, CR, and PE/CR are all quite effective and superior to R. Again, both PE and CR were effective, and combining them did not improve efficacy.

Cognitive Processing Therapy

Another cognitive behavioral therapy program for PTSD is CPT. CPT was originally developed for use with sexual assault victims and comprises elements of both CT and PE. Specifically, CPT includes CR around beliefs in five core areas – safety, power and control, esteem, trust, and intimacy – written exposure (i.e., writing about the traumatic memories and reading the account aloud to the therapist).

Resick et al. compared the efficacy of 12 sessions of CPT, 9 sessions of PE, and WL in rape victims with chronic PTSD. Results indicated that after 9 weeks of PE and 12 weeks of CPT, both treatments were highly effective in reducing PTSD; posttreatment only 19.5% of completers of CPT and 17.5% of PE met the criteria for PTSD. Among the completers, 76% of patients who had CPT and 58% of patients who had PE met the criteria for good end-state functioning (i.e., low PTSD and depression). Gains were maintained over time as well. At a 3-month follow-up, only 16.2% of those who received CPT and 29.7% of those who received PE were PTSD positive. Finally, at a 9-month follow-up, 80.8% of CPT and 84.6% of PE patients remained in sustained remission from PTSD symptoms. In sum, CPT appears to be an efficacious intervention for the short- and long-term treatment of PTSD. This study also provides further evidence for the efficacy of PE.

Eye Movement Desensitization and Reprocessing

EMDR emerged as a treatment for PTSD in the early 1990s, and since that time, a number of studies have investigated its efficacy. A core component of this intervention is theorized to be the therapist's repeated elicitation of rapid, saccadic eye movements (or other bilateral stimulation) from the patient during

the processing of traumatic memories. Shapiro theorized that rapid eye movements in some way override or reverse neural blockage or obstruction induced by the traumatic event. During EMDR treatment sessions, patients are asked to generate an image of the trauma and to focus on trauma-related thoughts, feelings, and/or sensations. Simultaneously, the therapist elicits the saccadic eye movements by having patients visually track a finger rapidly waved back and forth in front of their face. Other forms of laterally alternating stimuli (e.g., tapping or alternating sounds) are sometimes used rather than the original finger tracking. Patients are asked to evaluate the cognitions associated with these images and experiences and to generate alternative cognitive appraisals of the trauma or their behavior during the event. As patients focus on the distressing images and thoughts, and later focus on the alternative cognition, the saccadic eye movement (or another form of bilateral stimulation) is intermittently generated.

EMDR has been the focus of considerable controversy due to early claims regarding remarkable success in only a single session. Of the studies evaluating the efficacy of EMDR, more recent studies have typically used well-controlled designs and thus yielded clearly interpretable results. One well-controlled 1997 study evaluated the efficacy of EMDR relative to WL for PTSD in female rape victims. Three sessions of EMDR resulted in greater improvement of PTSD symptoms (57% reduction in PTSD at posttreatment) than WL (10% reduction at posttreatment). However, in contrast to other controlled studies, one therapist conducted all the treatments, and therefore the contribution of therapist and treatment effects are confounded. A direct comparison between EMDR and a combined treatment of PE plus stress inoculation training (called trauma treatment protocol, TTP) was conducted by DeVilly and Spence. Both TTP and EMDR reduced PTSD severity (63 and 46%, respectively), but TTP patients maintained their gains at follow-up, whereas EMDR patients showed higher rates of relapse (symptom reduction at follow-up of 61 and 12%, respectively). Rothbaum, Astin, and Marsteller compared PE, EMDR, and WL in female sexual assault victims. Compared to WL, both treatments resulted in a significant reduction in PTSD severity and related psychopathology. Immediately after treatment, only 5% of PE participants and 25% of EMDR participants continued to meet PTSD diagnostic criteria, compared to 90% of WL participants. The active treatments were significantly different from WL but did not differ from one another. At follow-up, however, significantly more participants receiving PE (78%) achieved good end-state functioning than did participants receiving EMDR

(35%). Taylor and colleagues compared PE, EMDR, and R and found that PE produced significantly greater reductions in reexperiencing and avoidance symptoms than EMDR and R, which did not differ from one another. Indeed, a 2001 meta-analysis found EMDR to be no more effective than exposure therapy programs and, moreover, suggested that the eye movements integral to the treatment and its underlying theory are unnecessary.

In summary, several cognitive-behavioral therapy programs have received empirical support for their efficacy in reducing PTSD severity and diagnosis in controlled studies: PE, SIT, CT, CPT, and EMDR. Studies that have attempted to combine these empirically supported approaches into new treatment packages to improve outcome have failed to demonstrate superior outcomes compared to the single-treatment protocols in these clinical trials.

Real-World Treatment Considerations

Treatment Tolerability and Preference

It also important to explore the tolerability and acceptability of empirically supported treatments for PTSD. Notably, although concerns have been raised by some clinicians and researchers that PE is too difficult to tolerate and may lead to symptom worsening and treatment drop-out, there is evidence that PE is tolerable. For example, Foa and colleagues examined self-reported symptom exacerbation in a sample of women with chronic assault-related PTSD who were receiving PE. After the first imaginal exposure, only a minority of participants in the study exhibited reliable symptom exacerbation during treatment: 10.5% reported an increase in PTSD symptoms, 21.1% reported an increase in anxiety, and 9.2% reported an increase in depression. Significantly, patients who reported such symptom exacerbation benefited from PE as much as those who did not report such exacerbation and were not more likely to drop out of treatment. Thus, although some patients experienced a brief period of exacerbation of symptoms during PE, this exacerbation decreased within 2 weeks and was unrelated to treatment outcome or completion. With respect to the completion of treatment, additional evidence supports that PE is tolerable. A 2003 review of dropout rates from various cognitive-behavioral therapy treatments for PTSD across 25 controlled studies indicated no significant differences in dropout rates among trials of PE, CT, SIT, EMDR, and combined treatments. PE did not produce a higher rate of treatment dropouts than other PTSD treatments; thus, concerns about excessive treatment dropout from PE appear to be unwarranted.

In fact, recent data demonstrate that treatments including PE are actually preferred over other psychosocial and psychopharmacological treatments. In a forced-choice comparison of treatment preference between PE and sertraline (an antidepressant medication) for PTSD, approximately 87% of female college students chose PE compared to only 7% who chose medication (6% chose no treatment). Consistent with these data, Tarrier and colleagues found that adults strongly prefer PE for PTSD compared to other psychosocial interventions, including EMDR, group therapy, and psychodynamic psychotherapy.

Treatment Dissemination

Disseminating empirically supported treatments such as PE to community-based clinicians is an important step in bridging the critical gap between research and practice for trauma survivors with PTSD. In one of the first studies to evaluate such dissemination, Foa and colleagues trained clinicians at a community-based rape crisis center in the delivery of PE for assault survivors with PTSD. These community clinicians provided treatment to a subset of participants in a randomized controlled trial, and their outcomes were compared to participants treated by those with expertise in PE at an academic medical center. After an initial training with follow-up consultation and supervision, results showed that the community mental health counselors obtained outcomes comparable to the university-based providers. This provides good evidence that empirically supported treatments such as PE can be successfully disseminated.

Childhood Sexual Abuse-Related Posttraumatic Stress Disorder

Despite the documented efficacy of PE for PTSD, there is some concern reported in the clinical literature that adults who have PTSD secondary to maltreatment during childhood, specifically child sexual abuse, may be a particularly vulnerable population for whom PE may be too difficult to tolerate. In response to this concern, a new treatment protocol, skills training affect and interpersonal regulation plus modified prolonged exposure (STAIR/MPE), was designed as a two-phase treatment, with phase 1 consisting of skills training in affect and interpersonal regulation and phase 2 consisting of modified PE (imaginal exposure plus coping skills training, with no *in vivo* exposure). Data suggested that STAIR/MPE resulted in significant reductions in PTSD and depression compared to WL. However, a critical weakness of this study is that this combination treatment was not compared to the usual PE. It is of note

that two recent clinical trials demonstrated that cognitive-behavioral treatments are effective for reducing PTSD among child sexual abuse survivors without a skills-building component. Moreover, a subanalysis of the data from Foa et al.'s 2005 study demonstrated that patients whose PTSD was related to child sexual abuse benefited from PE and PE/CR as much as did those whose PTSD was related to adult trauma. Thus, it seems that the skills-building component, which extends the number of treatment sessions and may delay recovery time, is unnecessary above the brief treatments that have empirical support.

Conclusion

Although exposure to potentially traumatic events is quite common, only a small minority of people develop persistent psychological difficulties such as PTSD in response to a trauma. The most validated types of treatment for PTSD are cognitive-behavioral therapies, and among the cognitive-behavioral therapy approaches available, PE has the strongest empirical support. Further, recent data suggest that PE is a preferred treatment among those who are given various treatment descriptions for PTSD. Concerns over the tolerability of PE, specifically that PE may lead to symptom exacerbation or premature dropout, are not supported by data. STAIR/MPE, a modified version of PE, has been shown to be helpful for childhood sexual abuse survivors with PTSD. However, this modified treatment has not been compared directly to already validated treatments for PTSD such as PE. Thus, at this point, we cannot determine whether it offers incremental benefit above such treatments. Future research should continue to evaluate ways to enhance treatment outcome and to disseminate efficacious treatments to a larger number of those who suffer with PTSD.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Child Sexual Abuse; Cognitive Behavioral Therapy; Posttraumatic Stress Disorder in Children; Posttraumatic Stress Disorder, Delayed; Posttraumatic Stress Disorder, Neurobiology of; War-Related Posttraumatic Stress Disorder, Treatment of; Posttraumatic Stress Disorder – Neurobiological basis for.

Further Reading

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental health disorders* (4th edn.). Washington, DC: American Psychiatric Association.
Cloitre, M., Koenen, K. C., Cohen, L. R., et al. (2002). Skills training in affective and interpersonal regulation

followed by exposure: a phase-based treatment for PTSD related to childhood abuse. *Journal of Consulting and Clinical Psychology* 70(5), 1067–1074.

Foa, E. B., Dancu, C. V., Hembree, E. A., et al. (1999). A comparison of exposure therapy, stress inoculation training, and their combination for reducing posttraumatic stress disorder in female assault victims. *Journal of Consulting and Clinical Psychology* 67(2), 194–200.

Foa, E. B. and Rothbaum, B. O. (1998). *Treating the trauma of rape*. New York: Guildford Press.

Kessler, R. C., Sonnega, A., Bromet, E., et al. (1995). Post-traumatic stress disorder in the national comorbidity survey. *Archives of General Psychiatry* 52, 1048–1060.

Zoellner, L. A., Feeny, N. C., Cochran, B., et al. (2003). Treatment choice for PTSD. *Behaviour Research & Therapy* 41(8), 879–886.

Posttraumatic Stress Disorder – Neurobiological basis for

M Barad

University of Los Angeles, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

General Considerations

Physiological Measures

Brain Imaging

Glossary

α_2 -Adrenergic receptor A receptor for the neurotransmitter norepinephrine; largely expressed presynaptically and acts to inhibit further release of norepinephrine.

Adrenergic system The cells and pathways in the brain that use the neurotransmitter norepinephrine, either by releasing it or by expressing receptors that respond to it.

Corticotropin releasing hormone (CRH) A peptide neurotransmitter released by the median eminence of the hypothalamus to stimulate the hypothalamic-pituitary-adrenal axis. It also acts widely in the brain as a neurotransmitter.

Cortisol The primary glucocorticoid hormone.
Dendritic pruning Reductions in the branching dendrites of neurons, the part of the neuron that receives signals from other neurons.

Functional imaging The examination of the brain by magnetic resonance imaging or positron emission tomography, yielding cross-sectional images of the brain and highlighting areas of increased or decreased blood flow that reflect localized brain activity.

Glucocorticoids Steroid hormones, especially cortisol, that regulate metabolism and the stress response.

Hippocampus A brain structure crucially involved in complex memories (those available to consciousness).

Hypothalamic-pituitary-adrenal (HPA) axis

Longitudinal studies

The regulatory pathway from the brain to the adrenal cortex that regulates the release of cortisol.

Studies that follow the development of a disease in an individual over time, along with its biological correlates. By comparing variables before and after the development of a disease, these studies allow a stronger identification of the pathological effects of the disease process than other types of studies.

A serotonin agonist.

Metachlorophenylpiperazine (m-CPP)
Naltrexone
Opioid system

An antagonist of opioids receptors.

The cells and pathways in the brain that use endorphins, or endogenous opioids, as neurotransmitters, either by releasing them or expressing receptors that respond to them. These are the same receptors that bind opium and other opioid drugs.

Provocation studies

Studies that use drugs or environmental stimuli (e.g., stories, images, or tasks) to provoke a specific brain response.

Serotonin system

The cells and pathways in the brain that use the neurotransmitter serotonin, either by releasing it or by expressing receptors that respond to it.

Structural imaging

The examination of the brain using high-resolution magnetic resonance imaging (or, in the past, computed tomography) to carefully quantify the size of identifiable structures.

Twin studies

Research comparing identical twins, one of whom suffers a pathology, such as posttraumatic stress disorder. Because the healthy twin is genetically identically and shares much the same environment as the patient, the similarities in his

brain structure or function to that of the patient, which differ from the average population and are similar to the patient population, suggest a preexisting condition or vulnerability to disease.

Yohimbine A pharmacological antagonist of α_2 -adrenergic receptor.

General Considerations

In general, two approaches have been used to examine the neurobiological basis of posttraumatic stress disorder (PTSD): imaging and biological measures. Imaging has been used to examine the anatomy of the brain and its activity both at rest and during a number of symptom provocation and cognitive challenges. Biological measures have been used to measure a number of candidate hormonal and neurotransmitter systems, again, both at rest and after symptom provocation. We consider some of the findings of these studies here.

It is important to remember that most studies of PTSD compare PTSD patients to normal controls. Such studies cannot differentiate between a finding in the PTSD patient that is the result of the disease process of PTSD and follows the traumatic event and a finding that preceded the trauma and thus may indicate a source of vulnerability to PTSD. Only longitudinal and twin or family studies can effectively differentiate between those possibilities, and there have been few such studies to date.

It is likely that preexisting factors, whether genetic or acquired from previous experience, are crucial to the development of PTSD. Exposure to trauma is not enough, no matter how severe. Only a small proportion of people exposed to trauma, even horrific trauma, develop PTSD. In the United States, approximately 90% of the population is exposed to at least one traumatic event serious enough to satisfy the trauma criterion for PTSD. However, only approximately 8% of the population develops PTSD. This means that it is particularly important to understand the neurobiological factors that contribute to vulnerability to PTSD or to its inverse, resiliency.

One obvious biological factor that seems to play a role in vulnerability to PTSD is sex. Women are twice as likely to develop PTSD after a trauma than men, although the interaction between biology and type of trauma is a complicated one. Women are much more likely to develop PTSD after an assault and men are much more likely to develop PTSD after being raped (even though their risk of rape is lower); both sexes have an equal risk of PTSD after natural disaster. These patterns suggest that the differences in rates of PTSD between men and women overall have

something to do with the meaning of the type of trauma to each sex. Both sexes suffer equally from natural disasters, which affect everyone in the area, whereas physical or sexual assault emphasizes the relative weakness of the victim.

PTSD is, by definition, a chronic condition in which symptoms persist for at least 1 month. However, in the immediate aftermath of a trauma, almost everyone experiences symptoms of PTSD. Thus, the symptoms of PTSD are almost certainly part of the normal human response to stress. It is the failure of recovery that is characteristic of PTSD, so it may be that the right place to look for vulnerability is in the mechanisms of recovery. Such mechanisms may include biological mechanisms, such as the glucocorticoid response to stress, or psychological mechanisms, such as the ability of people to suppress fearful responding in the short term through inhibition by prefrontal cortex or in the long term through behavioral extinction or desensitization of fearful responses to the reminders of the trauma.

Physiological Measures

Frightening experiences elicit a characteristic stress response. The faster response to a frightening stimulus is the adrenergic stress response, which orchestrates the readiness of an animal or a human to protect itself, the fight-or-flight response originally described by Walter Cannon. Simultaneously, a second, slower stress system is activated, the hypothalamic-pituitary-adrenal (HPA) axis, which results in the production of the effector glucocorticoid cortisol, as first described by Hans Selye. One important action of the slower cortisol response is to shut down or limit the adrenergic response.

Disruptions in stress hormone regulation stand out prominently in the neurobiology of PTSD. Some of the first studies of PTSD patients indicated not only that PTSD patients show increased excretion of epinephrine and norepinephrine at baseline but also that adrenergic system is hyperreactive. The administration of yohimbine to PTSD patients provided the most striking demonstration of the increase of adrenergic tone and its relation to PTSD symptoms. Yohimbine blocks the α_2 receptors for norepinephrine. These receptors are presynaptic on adrenergic terminals and act like a thermostat to inhibit the release of norepinephrine into the synapse. The administration of yohimbine is mildly anxiogenic in normal controls, but in PTSD patients it caused great distress, precipitating flashbacks in 40% and panic attacks in 70%, whereas none of the controls experienced either. These data indicate a profound dysregulation of the adrenergic system. Interestingly, yohimbine

also provokes panic attacks in patients with panic disorder but not in patients with schizophrenia, depression, obsessive-compulsive disorder, or generalized anxiety disorder and also not in normals. Thus, PTSD may share some part of its pathophysiology with panic disorder.

Increases of adrenergic tone and reactivity may be of great importance in the symptomatology of PTSD. It is clear from the experiment that increases of adrenergic neurotransmission may be precipitating some of the most troubling symptoms of PTSD, including not only flashbacks and the physiological hyperreactivity characteristic of panic attacks but also intrusive thoughts, emotional numbing, difficulty concentrating, and increased startle. However, it has also been extensively demonstrated in preclinical studies that the adrenergic system is crucial to memory strength, and particularly to the establishment of fearful or distressing memories. This explains in part why memories of frightening events are so vivid. Thus, high levels of norepinephrine during traumatic experiences may well strengthen the formation of distressing memories and make them more vivid for PTSD patients.

When challenged with yohimbine, the family members of PTSD patients do not have more panic attacks than normal subjects, whereas the healthy family members of panic disorder patients do have high rates of panic attacks upon yohimbine challenge. This suggests that a hyperreactive adrenergic system may be a predisposing factor for panic disorder, whereas the adrenergic hyperreactivity of PTSD patients may be acquired.

What might account for that hyperreactivity? One hypothesis is an insufficient regulation by an inadequate glucocorticoid response. We might expect from the strong response to trauma in PTSD patients that cortisol levels, as part of the stress response, should be high. In fact, most studies have found that glucocorticoid levels are low at baseline in PTSD patients and that the feedback inhibition of glucocorticoid synthesis is overly strong. Thus, when injected with an exogenous glucocorticoid, dexamethasone, patients with PTSD suppress the production of endogenous cortisol for a significantly longer time than normal controls. This is true even though the levels of the hypothalamic hormone that drives cortisol synthesis and release, corticotropin releasing hormone (CRH), are high in PTSD patients.

By contrast, in depression, high CRH levels drive high levels of circulating cortisol. The contrasting pattern in PTSD suggests that feedback inhibition of the HPA axis is dominant over CRH drive. Low circulating levels of cortisol may well allow the adrenergic response to stress to run out of control. But

are these low levels in PTSD patients the result of a burnout of the cortisol system by sustained high cortisol during and after the trauma?

Two longitudinal studies of motor vehicle-accident victims have addressed this question. In one study, 15-h urinary cortisol excretion was measured after accidents; in the other, cortisol levels were measured in blood samples drawn on the way to the emergency room. In both studies, the victims who went on to develop PTSD had significantly lower cortisol after their accidents than those who did not. Thus, even just after their trauma, the cortisol levels of those who later went on to PTSD were lower. This begins to hint that lower cortisol in PTSD patients may not be the result of burnout after prolonged elevations around the time of trauma but rather a preexisting state that contributes to a vulnerability to PTSD.

This preexisting state is almost certainly not entirely genetic. For example, in rape victims, cortisol levels after the rape were lower in those who had sustained a previous assault, and the victims of previous assaults were three times as likely to go on to develop PTSD. Women who had no previous trauma had cortisol levels that averaged three times higher than those who did and had higher cortisol levels when the index rape was more severe. Women with previous trauma showed no increase in cortisol levels in response to increased severity of rape.

However, the question remains whether low cortisol levels existed even before the rape. By comparing the children of Holocaust survivors to children of European Jews who escaped the Holocaust, Rachel Yehuda and her colleagues have examined this question. The children of Holocaust survivors are almost four times as likely to develop PTSD as the controls. The children of Holocaust survivors show lower cortisol excretion than do the controls, which argues for a preexisting defect in the cortisol system. However, this group was not homogeneous. When neither parent nor child had PTSD, cortisol excretion was the same as that of controls. However, when the parent had PTSD, the cortisol excretion was lower in the offspring; it was lowest when both parent and child had PTSD. Thus, it is clear that low PTSD levels are a preexisting risk factor that may not only be due to previous trauma. It is tempting to think that this risk factor is due to the transmitted anxiety of parents with PTSD, acting like a previous assault in rape victims to lower cortisol. However, there remains a potential for a genetic explanation. It might be that the Holocaust survivors with PTSD also had low cortisol as a risk factor for their PTSD and transmitted that risk genetically rather than behaviorally.

All of these data on endocrine stress responses outline one potential chain of causation in generating

the clinical picture of PTSD. A hypoactive cortisol response may lead to the dysregulation of the adrenergic stress response. An ungoverned adrenergic system may increase the intensity of traumatic memories, the strength of association of those memories with a wide range of stimuli in the environment, and the physiological and psychological arousal when those memories are recalled.

Beyond the alterations in the endocrine stress responses, other physiological abnormalities have been identified in PTSD patients. Endogenous opioids are low in PTSD sufferers, who also show a reduced pain threshold. However, when exposed to traumatic reminders, Vietnam veterans with PTSD showed a higher pain threshold that could be reversed by naltrexone (which blocks opioids receptors). Thus, PTSD patients appear to have both decreased resting levels of endogenous opioids and a more reactive opioid system. This may lead to cycles of trauma reminders inducing numbing, followed by opioid withdrawal (and PTSD) symptoms such as irritability and disturbed sleep. Because these hyperarousal symptoms of opioid withdrawal may be mediated by adrenergic activity, low tonic opioid levels may act in concert with low cortisol to increase adrenergic reactivity.

One study examined the effect of a serotonin agonist metachlorophenylpiperazine (m-CPP) on PTSD patients compared with the effect of yohimbine and placebo. As in the studies of yohimbine, m-CPP caused panic attacks in a substantial number of the patients, along with exacerbations of other PTSD symptoms. A similar number of patients experienced panic attacks and worsened PTSD symptoms after m-CPP injections. Interestingly, different subsets of patients tended to get panic attacks with the two agents. Another study showed that PTSD patients had a higher frequency of being homozygous for the short allele of the serotonin promoter, a genotype associated with depression after life stressors. Thus, although fewer data indicate a role for serotonergic than for noradrenergic mechanisms, serotonergic mechanisms may make a second and possibly independent contribution to PTSD vulnerability or pathophysiology.

Brain Imaging

Three areas of the brain have received particular attention in imaging studies of PTSD, because there are *a priori* reasons to suspect their involvement. The hippocampus has been examined in detail because a wealth of preclinical studies indicate that chronic stress has profound effects on this structure. The amygdala has been examined because this brain structure has an orchestrating role in all fear learning. Finally, the

prefrontal cortex has been examined because of its role in suppressing amygdaloid activation.

Structural Imaging

A variety of studies in animals indicate that elevated levels of the stress hormone cortisol is damaging to the hippocampus. Long-term exposure to exogenous glucocorticoid steroids and prolonged stress (which presumably elevates endogenous levels of these steroids) both damage the CA3 region of the hippocampus and cause dendritic pruning of neurons there, with the net effect of decreasing hippocampal volume. Although lower levels of cortisol are usually observed in PTSD patients, it is reasonable to hypothesize that cortisol levels might have been extremely high at the time of traumatic exposure.

Consistent with this hypothesis, hippocampal volumes in PTSD patients are decreased compared to controls in a number of structural imaging studies of combat veterans and of adult victims of childhood sexual abuse that compared PTSD patients to others exposed to combat or sexual abuse without PTSD. However, results have been less consistent with other patient populations and have not been observed in Holocaust survivors or patients with personality disorders who suffer from PTSD.

These studies are cross-sectional, comparing PTSD patients to controls, and do not address whether the hippocampal volume deficits might predate the trauma and development of PTSD. One study of identical twins in which one twin was a Vietnam combat veteran addressed this question indirectly. This study found that smaller hippocampal volume in the combat-exposed twin correlated to the severity of his PTSD symptoms. However, the hippocampal volume of the non-combat-exposed twin correlated just as closely with the severity of his brother's symptoms. These data suggest strongly that small hippocampi are a risk factor for, rather than a result of, combat exposure. Two longitudinal studies have also addressed this question, and neither found decreases of hippocampal volume associated with the development of PTSD or with ongoing PTSD pathology. Similarly, no changes in amygdala volume have been seen in PTSD patients.

Functional Imaging

Brain structure may be of less importance than brain function in psychiatric disease, including in PTSD. Functional imaging, looking at brain activity as reflected by blood flow, may thus be a better indicator of pathological brain function, especially when performed using biological or psychological probes designed to stimulate functional systems. Functional

studies have focused attention on the amygdala, a brain region important for orchestrating the fear response, as well as in related paralimbic and frontal structures. This approach remains in its early stages, and a clear anatomical pattern has not yet emerged, perhaps because of the wide variation in patient populations and provocation protocols in use.

One interesting approach has been to use pharmacological provocation with yohimbine, the antagonist of inhibitory presynaptic α_2 autoreceptors. As previously described, yohimbine, when injected into PTSD patients, exacerbates symptoms and causes panic attacks and flashbacks. In a positron emission tomography (PET) imaging study, yohimbine injections tended to increase metabolism in a variety of cortical regions including the prefrontal, temporal, parietal, and orbitofrontal cortices in control subjects, but tended to decrease metabolism in those regions in PTSD patients. These data indicate that the adrenergic dysregulation observed in PTSD patients profoundly changes brain activation patterns and suggests a role for those brain structures with changed activity in suppressing the set of symptoms evoked by yohimbine, specifically panic attacks and flashbacks.

Psychologically driven provocation studies have generally shown differences in brain activation between PTSD patients and controls, although these have not always been consistent. Left amygdala activity was increased by combat sounds, compared to white noise, in Vietnam veterans with PTSD but not in combat-exposed veterans without PTSD or in normal controls. In another similar study with combat sound, veterans with PTSD showed activation of the right amygdala as well as in other regions. In another comparison of veterans with and without PTSD who were asked to visualize combat, the veterans showed activation in the right amygdala as well as in the ventral anterior cingulate gyrus. However, when actually viewing combat images, no such activation was seen; however, PTSD patients showed decreased activity in Broca's area, an area important for language generation.

Trauma scripts and viewing of traumatic images have generally not yielded amygdala activation, whereas visualization and simple stimuli, such as combat sounds, which bypass the cortex, effectively activate that structure. Interestingly, masked presentations of fearful faces, too brief to register consciously, caused exaggerated amygdala responses in veterans with PTSD compared to healthy veterans.

On the other hand, script-driven provocation has effectively activated the posterior cingulate and motor cortex in women with a history of childhood

sexual abuse with PTSD compared to those without PTSD. Women with PTSD also failed to activate Broca's area and the anterior cingulate compared to controls, who did.

In summary, the functional imaging data so far do support changes in the function of the amygdala and of the limbic and prefrontal cortices in PTSD. However, the exact nature of those changes and their significance in the symptomatology and pathophysiology of PTSD remain to be defined.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Catecholamines; Corticotropin Releasing Factor (CRF); Corticotropin-Releasing Factor Receptors; Endocrine Systems; Fight-or-Flight Response; Glucocorticoid Negative Feedback; Glucocorticoids, Effects of Stress on; Glucocorticoids – Adverse Effects on the Nervous System; Glucocorticoids, Overview; Glucocorticoids, Role in Stress; Hippocampus, Corticosteroid Effects on; Hypothalamic-Pituitary-Adrenal; Posttraumatic Stress Disorder in Children; Posttraumatic Stress Disorder, Delayed; Posttraumatic Stress Disorder, Neurobiology of; Posttraumatic Therapy; Sex Steroids, Response to Stress and Susceptibility to Depression; Fear and the Amygdala; Neuroimaging and Emotion; Opioids; Selye, Hans; War-Related Posttraumatic Stress Disorder, Treatment of; Posttraumatic Stress Disorder – Clinical; Serotonin in Stress.

Further Reading

- Bremner, J. D. (2002). Neuroimaging studies in post-traumatic stress disorder. *Current Psychiatry Reports* 4, 254–263.
- Bremner, J. D. (2005). Effects of traumatic stress on brain structure and function: relevance to early responses to trauma. *Journal of Trauma and Dissociation* 6, 51–68.
- Foa, E. B. and Street, G. P. (2001). Women and traumatic events. *Journal of Clinical Psychiatry* 62(supplement 17), 29–34.
- Gilbertson, M. W., Shenton, M. E., Ciszewski, A., et al. (2002). Smaller hippocampal volume predicts pathologic vulnerability to psychological trauma. *Nature Neuroscience* 5, 1242–1247.
- Grossman, R., Buchsbaum, M. S. and Yehuda, R. (2002). Neuroimaging studies in post-traumatic stress disorder. *Psychiatric Clinics of North America* 25, 317–340.
- Resnick, H. S., Yehuda, R., Pitman, R. K., et al. (1995). Effect of previous trauma on acute plasma cortisol level following rape. *American Journal of Psychiatry* 152, 1675–1677.
- Rinne, T., Westenberg, H. G., den Boer, J. A., et al. (2000). Serotonergic blunting to meta-chlorophenylpiperazine (m-CPP) highly correlates with sustained childhood

abuse in impulsive and autoaggressive female borderline patients. *Biological Psychiatry* 47, 548–556.

Southwick, S. M., Bremner, J. D., Rasmusson, A., et al. (1999). Role of norepinephrine in the pathophysiology and treatment of posttraumatic stress disorder. *Biological Psychiatry* 46, 1192–1204.

Vermetten, E. and Bremner, J. D. (2002). Circuits and systems in stress. II: Applications to neurobiology and treatment in posttraumatic stress disorder. *Depression and Anxiety* 16, 14–38.

Yehuda, R. (2002). Current status of cortisol findings in post-traumatic stress disorder. *Psychiatric Clinics of North America* 25, 341–368.

Yehuda, R. (2004). Risk and resilience in posttraumatic stress disorder. *Journal of Clinical Psychiatry* 65(supplement 1), 29–36.

Yehuda, R., Schmeidler, J., Wainberg, M., et al. (1998). Vulnerability to posttraumatic stress disorder in adult offspring of Holocaust survivors. *American Journal of Psychiatry* 155, 1163–1171.

Posttraumatic Stress Disorder in Children

A M La Greca

University of Miami, Coral Gables, FL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A M La Greca, volume 3, pp 181–185, © 2000, Elsevier Inc.

Traumatic Events

Posttraumatic Stress Disorder in Children: Clinical Picture
Prevalence and Developmental Course

Demographic Trends

Factors that Precipitate Posttraumatic Stress Disorder in
Children

Assessing Posttraumatic Stress Disorder Symptoms in
Children

Glossary

*Avoidance or
numbing*

Refers to symptoms of posttraumatic stress syndrome (PTSD), such as avoiding thoughts, feelings, or conversations about the traumatic event, avoiding reminders of the event, diminished interest in normal activities, or feeling detached or removed from others.

Hyperarousal

Refers to symptoms of PTSD, such as difficulty sleeping or concentrating, irritability, angry outbursts, hypervigilance, and an exaggerated startle response.

Life threat

The perception that one's life is in danger; it is a key aspect of exposure to traumatic events, and is believed to be critical for the emergence of PTSD and related symptoms.

*Loss and
disruption*

Loss (of family and friends, of personal property and possessions) and disruption of everyday life (displacement from home, school, or community); it is also an aspect of exposure to traumatic events and can contribute to PTSD in children.

Re-experiencing

Refers to symptoms of PTSD, such as recurrent or intrusive thoughts or dreams about the event, or intense distress at cues or reminders of the event.

Posttraumatic stress disorder (PTSD) refers to a set of symptoms that develop following exposure to an unusually severe stressor or event. Typically, the event is one that causes or is capable of causing death, injury, or threat to the physical integrity of oneself or another person. In order to meet criteria for a diagnosis of PTSD, a child's reaction to the traumatic event must include intense fear, helplessness, or disorganized behavior. In addition, specific criteria for three additional symptom clusters must be met: re-experiencing, avoidance/numbing, and hyperarousal. Re-experiencing includes symptoms such as recurrent or intrusive thoughts or dreams about the event and intense distress at cues or reminders of the event; for young children, re-experiencing also may be reflected in repetitive play with traumatic themes or by a re-enactment of traumatic events in play, drawings, or verbalizations. Avoidance or numbing includes symptoms such as efforts to avoid thoughts, feelings, or conversations about the traumatic event, avoiding reminders of the event, diminished interest in normal activities, and feeling detached or removed from others. Hyperarousal symptoms include difficulty

sleeping or concentrating, irritability, angry outbursts, hypervigilance, and an exaggerated startle response; these behaviors must be newly occurring since the event. For a diagnosis of PTSD, the above symptoms must be manifest for at least 1 month, and be accompanied by significant impairment in the child's functioning (e.g., problems in school, friendships, or family relations). In addition, acute PTSD is specified if the duration of the symptoms is less than 3 months; chronic PTSD is specified if the duration of symptoms is 3 months or longer.

Traumatic Events

In recent years, media coverage of devastating natural disasters, bombings, and terrorist activities has drawn attention to the significant trauma that children and adults experience following such events. It has become increasingly apparent that exposure of children to such traumatic events may lead to reactions that interfere substantially with their day-to-day functioning and cause them significant distress. Specifically, exposure to natural disasters (e.g., hurricanes, tsunamis, tornadoes, fires, earthquakes, or floods) and to human-made disasters and acts of violence (e.g., plane crashes, sinking of ferries, war, terrorist attacks, bombings, sniper shootings) represent traumatic events that may result in the emergence of PTSD. Although considerable attention has been devoted to the effects of terrorism and community-wide disasters, it is also the case that children who experience motor vehicle accidents, life-threatening illnesses or medical procedures, or violence of a personal nature (e.g., physical and sexual abuse, rape, kidnapping, and community violence) also are vulnerable to developing PTSD. This article describes common symptoms of PTSD in children and factors that contribute to the development and course of this disorder.

Posttraumatic Stress Disorder in Children: Clinical Picture

The diagnostic category of PTSD was introduced in the third edition of the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III). At that time, PTSD was primarily considered to be an adult disorder. More recently, there has been a growing awareness that children and adolescents also experience PTSD, which is reflected in the current edition of the DSM (DSM-IV). (See previous description of the criteria for the disorder.)

Community studies of children following traumatic events suggest that the most commonly reported PTSD symptoms are those of re-experiencing the

event. For example, up to 90% of children exposed to a catastrophic hurricane reported symptoms of re-experiencing 3 months after the disaster. In contrast, symptoms of avoidance and numbing are far less common in children and, in fact, their presence may be a good indicator of PTSD.

Due to their limited verbal capacity, the diagnosis of PTSD is especially difficult in very young children, such as infants, toddlers, and early preschoolers. In such cases, generalized anxiety, fears, avoidance of certain situations that may be linked to the traumatic event, and sleep disturbances may be useful indicators of PTSD.

PTSD may also be present (i.e., co-morbid) with other psychological difficulties. In particular, PTSD often is co-morbid with anxiety disorders (i.e., separation anxiety disorder, specific phobias, generalized anxiety disorder, panic, and agoraphobia) and with affective disorders (e.g., major depressive disorder and dysthymia). Traumatic events that involve the loss of loved ones (e.g., earthquakes and bombings) are especially likely to lead to depressive reactions in children.

Prevalence and Developmental Course

Prevalence

It is difficult to estimate the prevalence of PTSD in children and adolescents because studies have been extremely diverse with respect to the type of trauma evaluated, assessment methods used, and the length of time that passed since the traumatic event occurred. Community studies suggest that approximately 24–39% of children and adolescents exposed to trauma (e.g., community violence or a natural disaster) meet criteria for PTSD. Rates of PTSD have been reported to be even higher among children and adolescents who witness death or physical injury in conjunction with acts of violence or following disasters associated with mass casualties.

When subclinical levels of PTSD are considered, up to 55% of the children in large community samples have reported at least moderate levels of PTSD symptoms during the first 3 months following a traumatic event, such as a hurricane. Thus, symptoms of PTSD appear to be common among children exposed to trauma, although fewer children and youth will meet criteria for a PTSD diagnosis.

Developmental Course

Little is known about the course of PTSD symptoms in children over time. However, it does appear that PTSD symptoms may emerge in the days or weeks

following a traumatic event, and can take months or years to dissipate in some children. In the absence of re-exposure to trauma or of the occurrence of other traumatic events, the typical developmental course of symptoms appears to be one of lessening frequency and intensity over time. For example, 3 months after a highly destructive natural disaster (Hurricane Andrew: a Category 5 hurricane, that struck South Florida in 1992), 39% of the children informally met criteria for PTSD disorder, but this was reduced to 24% at 7-months postdisaster, and to 18% by 10-months postdisaster. A subgroup of children who reported moderate-to-severe PTSD symptoms was followed 42 months postdisaster, revealing that 40% continued to report moderate-to-severe PTSD symptoms as well as impairment in their functioning; yet, almost none of the children who had mild or no symptoms at 10-months postdisaster reported any symptoms later on. Similarly, 51.5% of the adolescents who survived the sinking of the cruise ship *Jupiter* in 1988 (with 400 British school children on board) developed PTSD after the disaster, but 5–8 years later, only 17.5% still had PTSD.

These data suggest a steady reduction in the frequency and severity of PTSD and its symptoms over time (with no further exposure to similar trauma), although a significant minority of children and adolescents do not recover and continue to report substantial difficulties years later. The findings also indicate that it is highly unusual for children and adolescents to report significant symptoms of PTSD 1 year or more after a traumatic event, if they did not experience such symptoms closer to the event. Although there may be a brief period of shock or numbing, or sometimes elation and relief at still being alive, most children who develop PTSD will evidence stress symptoms within the first few months of the traumatic event.

Persistence of Symptoms

Overall, findings confirm that children's reactions to disasters and other traumatic events are not merely transitory feelings that dissipate quickly. On the contrary, they appear to linger and persist and are likely to cause distress to children and their families for some time. One factor that contributes to the persistence of PTSD symptoms in children over time is the occurrence of other significant life stressors. Children who encounter major life stressors (e.g., a death or illness in the family or parental divorce) in the months following a traumatic event, do not recover quickly, and report persistently high levels of PTSD symptoms over time.

Demographic Trends

Gender and Age

Some studies indicate that PTSD appears more frequently among girls than boys, but this has not consistently been the case. Preadolescent children may be more vulnerable to PTSD than older youths, although it is difficult to draw generalizations regarding children's vulnerability to PTSD at different ages, as findings on age-related differences have been inconsistent and may be influenced by diverse developmental manifestations of PTSD.

Ethnicity

PTSD appears to occur across diverse ethnic and cultural groups. Community studies suggests that minority youths exposed to severe natural disasters report more symptoms of PTSD and have a more difficult time recovering from such events than nonminority youths. It is possible that socioeconomic factors might account for such findings, in that children and families from minority backgrounds may have less financial resources or less adequate insurance to deal effectively and efficiently with the postdisaster rebuilding process. This could prolong the period of life disruption and loss of possessions that ensues after destructive natural disasters.

Factors that Precipitate Posttraumatic Stress Disorder in Children

Exposure to Trauma

A wide range of traumatic events have been linked to the emergence of PTSD in children and adolescents. However, two key aspects of exposure to these traumatic events are thought to contribute to children's reactions: the presence or perception of life threat, and personal loss and disruption of everyday events.

The presence or perception of life threat is considered to be essential for the emergence of PTSD. It is easy to understand why children who witness or are exposed to acts of violence, such as sniper shootings or severe physical abuse by a caretaker, feel that their life is in danger. It is also the case, however, that catastrophic natural disasters or personal disasters (such as residential fires) can elicit perceptions of life threat in children, even if no one is injured or hurt. For example, although very few lives were lost in South Florida as a result of Hurricane Andrew, the extensive destruction of homes and property that occurred during the storm was terrifying to many children and adults. In one study, 60% of the children interviewed "thought that they were going to die"

during the storm. Thus, perceptions of life threat can occur in the absence of actual loss of life or serious injury.

Evidence also underscores the importance of loss (of family and friends or personal property and possessions) and disruption of everyday life (displacement from home, school, or community) as contributing to PTSD in children. The life changes that result from this aspect of exposure to a traumatic event also predicts PTSD symptoms in children and adolescents.

The duration and intensity of life-threatening events are additional aspects of traumatic exposure associated with children's PTSD symptom severity. The prolonged nature of certain traumatic events (e.g., floods or child abuse) is very distressing to children when no immediate relief is in sight.

Role of Prior Psychological Functioning

Children with pre-existing psychological problems, especially anxiety, may be more vulnerable to developing PTSD following traumatic events. For example, children's anxiety levels 15 months before a traumatic event have been found to predict their levels of PTSD symptoms 3 and 7 months postdisaster, even when controlling for children's exposure to the event. This suggests that anxious children may have a vulnerability to PTSD, even if their degree of exposure to a trauma is relatively low. In addition, children who had greater levels of exposure to the disaster (i.e., more life threat or more loss/disruption) showed an increase in anxiety symptoms following the disaster relative to children with less disaster exposure. Other studies have shown predisaster levels of depression predicted postdisaster PTSD among adolescents exposed to a destructive natural disaster (earthquake).

Aspects of the Recovery Environment

Certain aspects of the posttrauma recovery environment may either contribute to the development and persistent of symptoms of PTSD, or may help to alleviate or mitigate such symptoms. Intervening life events/stressors and re-exposure to the traumatic event have been linked with greater persistence of PTSD and its symptoms among children and adolescents. In addition, parents' psychosocial functioning, including their levels of psychopathology and their own reactions to the disaster, are likely to affect children's postdisaster functioning. For example, mothers' distress in the aftermath of Hurricane Hugo, which struck Charleston South Carolina in 1989, was associated with the persistence of their children's postdisaster emotional difficulties. Other research has supported a linkage between children's symptoms

of PTSD and parents' trauma-related symptoms. Finally, children with more negative coping strategies for dealing with stress (e.g., anger or blaming others) show greater persistence in their symptoms of PTSD over time.

In contrast, the availability of social support has been found to mitigate the impact of trauma on children. Following traumatic events, such as severe abuse or catastrophic disasters, children with higher levels of social support from significant others (parents, teachers, friends, and classmates) report fewer symptoms of PTSD than those with low levels of social support. Because of these findings, efforts to enhance the social support of children exposed to trauma, and to encourage adaptive coping skills, may be useful for strategies for interventions with children following traumatic events.

Assessing Posttraumatic Stress Disorder Symptoms in Children

Evaluating the presence of PTSD symptoms in children and adolescents can be challenging and will depend, to some extent, on the age or developmental level of the child. For children of elementary school age and older (i.e., 7 years and above), the child or adolescent is likely to be the best informant; parents often underestimate symptoms of PTSD in their children.

The use of child-oriented structured interviews that contain items pertinent to the PTSD diagnosis, such as the Anxiety Disorders Interview Schedule for Children or the Diagnostic Interview Schedule for Children represent the most desirable way to assess such reactions. For screening purposes, when a large number of children are being evaluated, or when personnel constraints preclude the use of individual interviews, child-oriented self-report measures, such as the PTSD Reaction Index, have been found to be extremely useful.

Assessing symptoms of PTSD in children of pre-school age and younger typically requires the input of the parent or primary carer. Parent-versions of structured interviews would be the desired format to use, or parent-competed questionnaires reporting on children's PTSD symptoms. For young children, behavioral observations can be especially important in evaluating their reactions. Behaviors to evaluate include signs of distress, arousal, or fear, and behavioral avoidance of trauma-related objects, events, or situations.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Posttraumatic Stress Disorder, Delayed;

Posttraumatic Stress Disorder, Neurobiology of; War-Related Posttraumatic Stress Disorder, Treatment of; Posttraumatic Stress Disorder – Clinical; Posttraumatic Stress Disorder – Neurobiological basis for.

Further Reading

- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th edn.). (DSM-IV) Washington, DC: American Psychiatric Association.
- American Academy of Child Adolescent Psychiatry (1998). AACAP Official Action: Practice parameters for the assessment and treatment of children and adolescents with posttraumatic stress disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* 37, 4S–26S.
- Applied Research & Consulting LLC and the Columbia University Mailman School of Public Health. (2002). *Effects of the World Trade Center attack on NYC public school students: initial report to the New York City Board of Education*. Available online at: http://eric.ed.gov/ERICDocs/data/ericdocs2/content_storage_01/0000000b/80/27/ce/00.pdf. Last accessed 11 October 2006.
- Asarnow, J. (1999). When the earth stops shaking: earthquake sequelae among children diagnosed for pre-earthquake psychopathology. *Journal of the American Academy of Child and Adolescent Psychiatry* 38, 1016–1023.
- Drell, M. J., Siegel, C. H. and Gaensbauer, T. J. (1993). Posttraumatic stress disorder. In: Zeanah, C. H. (ed.) *Handbook of infant mental health*, pp. 369–381. New York: Guilford Press.
- Goenjian, A. K., Pynoos, R. S., Steinberg, A. M., et al. (1995). Psychiatric comorbidity in children after the 1988 earthquake in Armenia. *Journal of the American Academy of Child and Adolescent Psychiatry* 34, 1174–1184.
- Green, B. L., Korol, M. S., Grace, M. C., et al. (1991). Children and disaster: gender and parental effects on PTSD symptoms. *Journal of the American Academy of Child and Adolescent Psychiatry* 30, 945–951.
- Gurwitch, R. H., Sitterle, K. A., Young, B. H., et al. (2002). The aftermath of terrorism. In: La Greca, A. M., Silverman, W. K., Vernberg, E. M., et al. (eds.) *Helping children cope with disasters and terrorism*, pp. 327–358. Washington, DC: American Psychological Association.
- La Greca, A. M., Silverman, W. K., Vernberg, E. M., et al. (1996). Symptoms of posttraumatic stress after Hurricane Andrew: a prospective study. *Journal of Consulting and Clinical Psychology* 64, 712–723.
- La Greca, A. M., Silverman, W. K., Vernberg, E. M., et al. (eds.) *Helping children cope with disasters and terrorism*. Washington, DC: American Psychological Association.
- La Greca, A. M., Silverman, W. K. and Wasserstein, S. B. (1998). Children's predisaster functioning as a predictor of posttraumatic stress following Hurricane Andrew. *Journal of Consulting and Clinical Psychology* 66, 883–892.
- Nolen-Hoeksema, S. and Morrow, J. (1991). A prospective study of depression and posttraumatic stress symptoms after a natural disaster: the 1989 Loma Prieta Earthquake. *Journal of Personality and Social Psychology* 61, 115–121.
- Vernberg, E. M., La Greca, A. M., Silverman, W. K., et al. (1996). Predictors of children's post-disaster functioning following Hurricane Andrew. *Journal of Abnormal Psychology* 105, 237–248.
- Vogel, J. and Vernberg, E. M. (1993). Children's psychological responses to disaster. *Journal of Clinical Child Psychology* 22, 464–484.
- Yule, W., Udwin, O. and Bolton, D. (2002). Mass transportation disasters. In: La Greca, A. M., Silverman, W. K., Vernberg, E. M., et al. (eds.) *Helping children cope with disasters and terrorism*, pp. 327–358. Washington, DC: American Psychological Association.

Posttraumatic Stress Disorder Overview, General See: Posttraumatic Stress Disorder – Clinical; Posttraumatic Stress Disorder, Neurobiology of; Posttraumatic Therapy; Acute Stress Disorder and Posttraumatic Stress Disorder.

Posttraumatic Stress Disorder, Delayed

A Holen

Norwegian University of Science and Technology,
Trondheim, Norway

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by A Holen, volume 3, pp 179–180, © 2000, Elsevier Inc.

Diagnostic Systems and Delayed Posttraumatic Stress
Disorder

Nature of the Delay Period

Discussing Delayed Posttraumatic Stress Disorder

Glossary

<i>Delayed posttraumatic stress disorder (D-PTSD)</i>	Posttraumatic stress disorder when the onset of symptoms occurs at least 6 months after the stressor.
<i>Stressor</i>	A traumatic event that gives rise to major stress responses.

Diagnostic Systems and Delayed Posttraumatic Stress Disorder

Posttraumatic stress disorder (PTSD) can be found in both the *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (DSM-IV) of the American Psychiatric Association and the *International Classification of Diseases* (ICD-10) of the World Health Organization. The stressor criterion and the three symptom cluster criteria (intrusion, avoidance, and psychophysiological arousal), as well as the time frames, are fairly similar, though some differences exist.

In DSM-III through DSM-IV, PTSD is subsumed under anxiety disorders. When the disorder lasts at least 1 month and persists for less than 3 months after the stressful event, PTSD is considered acute. If the duration is more than 3 months, then PTSD is regarded as chronic.

In ICD-10, PTSD is placed under the heading F43, "Reaction to severe stress and adjustment disorders." The ICD-10 defines "Post-traumatic stress disorder" (F43.1) when the diagnostic criteria are met within 6 months of a major stressful event or of a period of stress. The ICD-10 gives yet another option: when patients have a chronic course, which involves major and lasting personality changes, the diagnosis of "Enduring personality changes, not attributable to brain damage and disease" (F62) may be used.

So far, there has been limited research into long-term personality changes.

Both DSM-IV and ICD-10 concur that delayed PTSD (D-PTSD) is diagnosed when at least 6 months have passed between the exposure of a major stressor and the full onset of PTSD.

Nature of the Delay Period

D-PTSD occurs when a person is reasonably well adapted for the initial 6 months or more after experiencing a major stressor but develops PTSD symptoms, such as intrusion, avoidance, and increased arousal related to the initial event, at some later stage. None of the diagnostic systems, however, characterize the nature of the delay period. Still, there is a limited body of research available to shed light on the subject. North et al., in a study published in 1997, found no cases of delayed-onset PTSD in 136 survivors 1 year after a mass shooting incident. Considerable discrepancy of identified PTSD cases was apparent between index and follow-up. The authors argued that inconsistency in reporting rather than true delayed onset was responsible for PTSD cases identified 1 year after the trauma. After the Oklahoma City bombing, survivors were studied 6 and 17 months postdisaster by the same author. Again, no delayed-onset cases were verified. At a 5-year follow-up, Holen found only one case of D-PTSD in the 72 survivors from the North Sea oil rig disaster.

A prospective study by Bryant et al. of 103 motor vehicle accident survivors studied within 1 month of the stressor and again at 6 months and 2 years post-accident identified five patients with PTSD at 2 years; they had not met the PTSD criteria at 6 months. Additionally, the authors reported that delayed-onset cases suffered subsyndromal levels of PTSD, with elevated resting heart rate levels prior to diagnosis. Buckley et al. also found delayed-onset survivors from motor vehicle accidents to be more symptomatic initially than those who did not develop PTSD. In addition, the delayed-onset survivors tended to have poorer social support systems than controls both prior to and after the stressor. Furthermore, the delayed-onset survivors had lower global assessment of function (GAF) scores than controls in the month prior to the motor vehicle accident. Schnurr et al. found that women were more likely to be in the delayed-onset cluster. In a prospective study of traffic accident victims by Koren et al., only 2 of 39 survivors without PTSD at 1 year posttrauma demonstrated delayed-onset PTSD 4 years later. In discussing

compensation issues, Eitinger and Holen argued that bridge symptoms, i.e., subclinical symptoms that persist from the time of the trauma to the manifestations of full-blown PTSD, may serve as precursors to delayed responses. In follow-up studies of the population from the collapse of the Buffalo Creek dam, Green et al. reported that 11% of the cases not identified with PTSD in 1974 met the criteria in 1986.

In a study published in 1989, Solomon et al. reviewed 150 randomly selected files of veterans who sought treatment for combat-related disorders; all the veterans were from the 1982 Lebanon war. Four reasons were given for why they asked for health services. In 10% of the sample, the delayed-onset PTSD was preceded by an initial asymptomatic period. After the war, some other veterans uninterruptedly suffered mild PTSD-associated symptoms until accumulated tension or exposure to subsequent adversity, either military or civilian in nature, resulted in full-blown PTSD. Such exacerbations of PTSD were seen in 33% of the veterans. Reactivation of an earlier combat stress reaction was found in 13% of the cases. The reactivation started only after an asymptomatic period and in connection with threatening military stimuli. Some veterans had no single trigger; they seemed to build up tensions over time with various exposures. Finally, one group of veterans was characterized by delayed treatment seeking for chronic PTSD; this accounted for as much as 40% of the cases. In the absence of proper definitions, the first three groups, i.e., those with an initial asymptomatic period, those with exacerbations, and those with reactivated responses, may qualify as varieties of D-PTSD or they may represent different kinds of posttraumatic processing. The veterans from the 1982 Lebanon war with delayed-onset PTSD used significantly more emotion-focused and distraction coping on the Ways of Coping Checklist than control subjects. Furthermore, they used coping skills significantly less than those with immediate-onset PTSD.

From 1968 to 1973, decades before PTSD was invented, Eitinger and Stroem studied the health of Norwegian concentration camp survivors. They reported a high occurrence of late-onset somatic and psychiatric manifestations associated with concentration camp syndrome, or the KZ syndrome.

In a study of aging Resistance veterans from Holland, Aarts et al. found that the period between the end of the war and when the first symptoms of PTSD appeared varied considerably. First onset of PTSD symptoms appeared in 26% of the cases during the first 5 years after World War II. However, in about 50% of the veterans, PTSD manifested after more than 20 years.

Discussing Delayed Posttraumatic Stress Disorder

There are many uncertainties and controversies in the current conceptualizations of D-PTSD. Some authors report no cases of D-PTSD and raise doubts about the legitimacy of the concept. Others bring up the question of whether D-PTSD represents a latent disorder or simply is unrecognized PTSD. In recent years, a growing number of studies involving both motor vehicle accidents and combat-related stressors have demonstrated occurrence rates of D-PTSD around 10% when manifestation of the disorder is preceded by an asymptomatic period. Moreover, several studies point toward vulnerability factors in both the survivors and their environments. Persons who tend to have subsyndromal symptom levels in the delay period are more likely to have poorer social networks, poorer coping skills, and lower GAF scores.

Some researchers, such as Horowitz, argue that posttraumatic symptoms may wax and wane with triggers or life events. According to the diagnostic criteria of PTSD in the DSM-IV, the triggers are internal or external cues that symbolize or resemble aspects of the primary stressor. When a person is exposed to triggers, his or her state may change from asymptomatic or subthreshold to symptomatic, or from prodromal mild to severe, and thus meet the diagnostic criteria of D-PTSD. Most researchers and clinicians would probably regard triggers as idiosyncratic, low-grade events with no or only a minimal impact on most people, which, in contrast, lead to major responses in survivors. On the contrary, events leading to reactivation may be regarded as having more impact and being traumatic in a general sense. The notion of reactivation implies that past traumas lower the threshold for certain stressors that may symbolize or resemble the primary stressor, and entail far stronger responses than would be seen in most people. The boundaries between episodes of reactivation as part of D-PTSD and episodes of a new PTSD are rather blurred.

In discussing D-PTSD and concepts such as triggers, reactivating events, and primary stressors, the need for clearer definitions becomes quite apparent. Furthermore, gradients and/or classifications of events may be relevant. More studies are needed to shed light on these issues as well as on the course of the posttraumatic processing. What roles do the primary stressor, the subsequent triggers and aversive events, personality, and the social environment play when full-blown PTSD occurs after a completely asymptomatic period, or after a time dominated by subthreshold levels of PTSD symptoms? If the subthreshold symptoms affect the life of the person to a maladaptive degree yet do not meet the criteria of

PTSD, the person might within certain time limits be diagnosed as having an adjustment disorder.

Many central questions regarding D-PTSD are still unanswered. We do not know much about possible links to aging. More research is needed to expand our current knowledge about the vicissitudes of delayed posttraumatic stress.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Anxiety; Depression and Manic-Depressive Illness.

Further Reading

- Arts, P. G. H., Op den Velde, W., Falger, P. R. J., et al. (1996). Late onset of posttraumatic stress disorder in aging resistance veterans in the Netherlands. In: Ruskin, P. E. & Talbott, J. A. (eds.) *Aging and the posttraumatic stress disorder*, pp. 53–76. Washington, D.C.: American Psychiatric Press.
- Bryant, R. A. and Harvey, A. G. (2002). Delayed-onset posttraumatic stress disorder: a prospective evaluation. *Australian and New Zealand Journal of Psychiatry* 36, 2005–2009.
- Buckley, T. C., Blanchard, E. B. and Hickling, E. J. (1996). A prospective examination of delayed onset PTSD secondary to motor vehicle accidents. *Journal of Abnormal Psychology* 105, 617–625.
- Eitinger, L. and Strøm, A. (1973). *Mortality and morbidity after excessive stress*. Norway: Universitetsforlaget.
- Green, B. L., Lindy, J. D., Grace, M. C., et al. (1990). Buffalo Creek survivors in a second decade: stability of stress symptoms. *American Journal of Orthopsychiatry* 60, 43–54.
- Holen, A. (1990). *A long-term outcome study of survivors from a disaster: the Alexander L. Kielland disaster in perspective*. Oslo, Norway: University of Oslo.
- Horowitz, M. J. (1986). *Stress response syndromes*. Northvale: Aronson.
- Koren, D., Arnon, I. and Klein, E. (2001). Long term course of chronic posttraumatic stress disorder in traffic accident victims: a three-year prospective follow-up study. *Behavior Research and Therapy* 39, 1449–1458.
- North, C. S., Smith, E. M. and Spitznagel, E. L. (1997). One year follow-up of survivors of a mass shooting. *American Journal of Psychiatry* 154, 1696–1702.
- North, C. S., Pfefferbaum, B., Tivis, L., et al. (2004). The course of posttraumatic stress disorder in a follow-up study of survivors of the Oklahoma City bombing. *Annals of Clinical Psychiatry* 16, 209–215.
- Pomerantz, A. (1991). Delayed onset of PTSD: delayed recognition or latent disorder? *American Journal of Psychiatry* 148, 1609.
- Schnurr, P. P., Lunney, C. A., Sengupta, A. and Waelde, L. C. (2003). A descriptive analysis of PTSD chronicity in Vietnam veterans. *Journal of Trauma and Stress* 16, 545–553.
- Solomon, Z., Kotler, M., Shalev, A. and Lin, R. (1989). Delayed onset PTSD among Israeli veterans of the 1892 Lebanon war. *Psychiatry* 52, 428–436.
- Solomon, Z., Mikulincer, M. and Waysman, M. (1991). Delayed and immediate onset posttraumatic stress disorder. II. The role of battle experiences and personal resources. *Social Psychiatry and Psychiatric Epidemiology* 26, 8–13.
- Stroem, A. (ed.) (1968). *Norwegian concentration camp survivors*. Norway: Universitetsforlaget.

Posttraumatic Stress Disorder, Neurobiology of

J D Bremner

Emory University School of Medicine, Atlanta, GA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J D Bremner, volume 3, pp 186–191, © 2000, Elsevier Inc.

Posttraumatic Stress Disorder: Biology and Phenomenology
Animal Models for Posttraumatic Stress Disorder
Neurotransmitter and Neuropeptide Systems Involved in Stress and Posttraumatic Stress Disorder

Effects of Stress on the Neurobiology of Learning and Memory
Concluding Remarks

Glossary

<i>Amygdala</i>	A small, walnut-shaped brain area in the center of the brain, involved in fear responses.
<i>Cortisol</i>	The stress hormone in humans (glucocorticoid is a more general term, referring to the stress hormone in all animal species).
<i>Extinction</i>	The inhibition of fear responses; felt to be mediated by the medial prefrontal cortex.

<i>Hippocampus</i>	The subcortical brain area involved in learning and memory.
<i>Medial prefrontal cortex</i>	Part of the frontal cortex (brain area just above the eyes) that is involved in emotion and social behaviors; may be involved in posttraumatic stress disorder symptoms.
<i>Magnetic resonance imaging</i>	An imaging technique that can provide detailed maps of brain structure as well as function.
<i>Norepinephrine</i>	A stress neurotransmitter that mediates fight-or-flight responses.
<i>Positron emission tomography (PET)</i>	An imaging technique that uses radiolabeled substances to provide information about brain function.
<i>Posttraumatic stress disorder (PTSD)</i>	A psychiatric disorder, listed in the <i>Diagnostic and Statistical Manual of Psychiatry</i> , that follows exposure to an extreme stress that was a threat to life; it is accompanied by intrusive thoughts, startle, sleep disturbance, hyperarousal, and other symptoms.

Posttraumatic Stress Disorder: Biology and Phenomenology

PTSD affects 8% of people at some time during their lives. PTSD is characterized by specific symptoms,

including intrusive thoughts, hyperarousal, flashbacks, nightmares, sleep disturbance, changes in memory and concentration, and startle responses, that develop following exposure to extreme stress. The symptoms of PTSD are a behavioral manifestation of stress-induced changes in brain structure and function. Stress results in acute and chronic changes in neurochemical systems and specific brain regions, which result in long-term changes in the brain circuits involved in the stress response. Brain regions felt to play an important role in PTSD include the hippocampus, amygdala, and medial prefrontal cortex. Cortisol and norepinephrine are two neurochemical systems that are critical in the stress response. When we refer to the neurobiology of PTSD, we are referring to the aggregate of these stress-induced changes in the brain (Figure 1).

Animal Models for Posttraumatic Stress Disorder

Animal models of stress have been used in the study of stress-related psychiatric disorders, including PTSD. Animals exposed to repeated stress from which they cannot escape develop specific fear-related behaviors that are associated with lasting changes in stress-related neurochemical systems. Behavioral changes related to stress include weight loss and decreases in food intake, active behavior in novel environments,

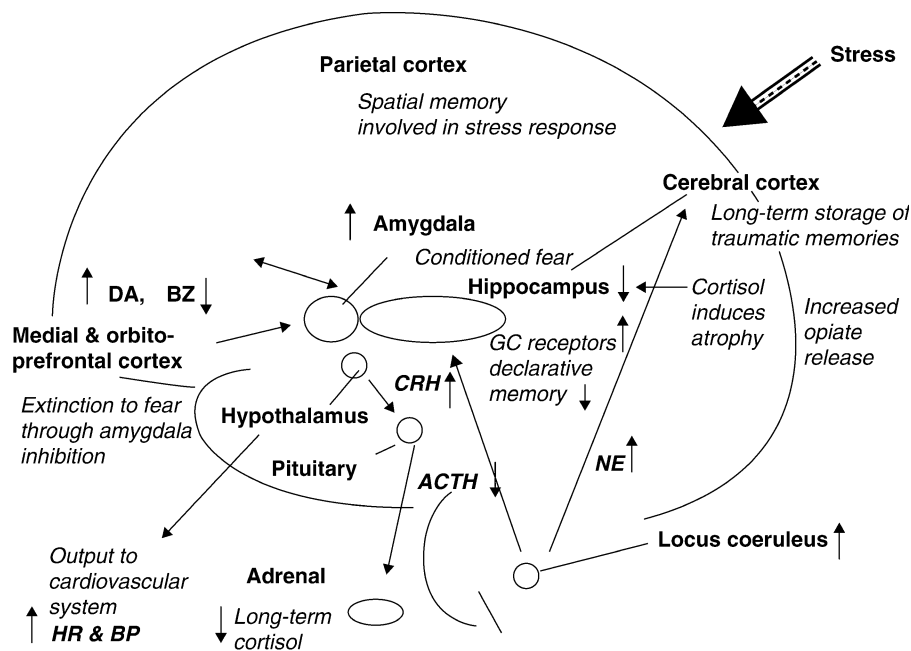


Figure 1 A model for the neurobiology of PTSD. Stress has an effect on multiple brain areas that are responsible for mediating the stress response. These include the amygdala, hippocampus, medial and orbitofrontal cortex, cerebral cortex, hypothalamus, and brain stem. These regions are functionally interrelated and represent important substrates for neurochemical systems involved in the stress response. Long-term changes in the function and structure of these regions lead to symptoms of PTSD. ACTH, adrenocorticotropic hormone; BP, blood pressure; BZ, benzodiazepine; CRH, corticotropin-releasing hormone; DA, dopamine; GC, glucocorticoid; HR, heart rate; NE, norepinephrine.

competitiveness and normal aggressiveness, grooming, play activity, responding for rewards, and self-stimulation of brain reward centers, as well as deficits in memory and attention, sleep disturbances, and increased defecation. These behaviors are similar to stress-related behaviors seen in patients with PTSD.

Neurotransmitter and Neuropeptide Systems Involved in Stress and Posttraumatic Stress Disorder

The locus coeruleus-noradrenergic system plays an important role in the stress response. The majority of norepinephrine cell bodies in the brain are located in a brain-stem site called the locus coeruleus (dorsal pons), with long neurons that project to multiple cortical and subcortical brain sites (e.g., the hippocampus, amygdala, and hypothalamus) for the direct release of the norepinephrine neurotransmitter. A variety of stressors result in an increase in locus coeruleus firing, with increased norepinephrine release in the brain and increased stress-related behaviors (e.g., increased vigilance and fight or flight) that are adaptive in survival. Animals with a history of chronic stress show a potentiated norepinephrine response with reexposure to subsequent stressors. The sensitization of noradrenergic responses may partially underlie the clinical phenomenon of stress sensitization seen in PTSD (e.g., childhood abuse increased the risk of PTSD in veterans who served in Vietnam).

A variety of studies showed long-term dysregulation of the noradrenergic system in PTSD. Kardiner noted in 1941 that veterans with psychiatric conditions related to their war experiences exhibited what appeared to be a hyperresponsiveness of the sympathetic system, manifested by increases in heart rate, blood pressure, sweatiness, irritability, palpitations, vertigo, dizziness, nausea, and syncope. Psychophysiology studies demonstrated an increase in sympathetic responses (heart rate, blood pressure, and galvanic skin response) to traumatic reminders, such as traumatic slides and sounds or traumatic scripts, in PTSD. An increase in norepinephrine and epinephrine was found in 24-h urine samples of patients with combat-related PTSD in some studies but not in others. Plasma levels of norepinephrine were not elevated at baseline in PTSD; however, a potentiated release of norepinephrine was found following exposure to traumatic reminders. Studies of peripheral adrenergic platelets (used as a marker of circulating levels of norepinephrine and epinephrine) have had mixed results. The administration of the adrenergic antagonist yohimbine, which results in an increase in the release of norepinephrine in the brain, resulted in an increase in PTSD-specific

symptomatology, as well as a greater release of norepinephrine metabolites in plasma in PTSD patients. Alterations in central metabolic responses to yohimbine were also found in PTSD patients as measured with PET. Yohimbine was associated with a differential effect on brain metabolism in neocortical areas, including the orbitofrontal, temporal, prefrontal, and parietal cortex, with metabolism having a tendency to increase in the controls and decrease in the patients. Given the dose-response effects of norepinephrine on cortical metabolism, these findings are consistent with a potentiated release of norepinephrine in PTSD. In summary, studies are consistent with an increased responsiveness of the norepinephrine system in PTSD.

The mesocortical dopaminergic system, which involves projections from the midbrain to the medial prefrontal cortex, is the most sensitive neural system to mild stressors. Both acute and chronic stressors result in a selective increase in dopamine release and metabolism in the medial prefrontal cortex. The few studies performed in PTSD showed elevations of dopamine and its metabolites in 24-h urine samples. Future work is needed in this important area.

The corticotropin-releasing hormone (CRH)-hypothalamic-pituitary-adrenal (HPA) axis system plays an important role in the stress response. Exposure to stressful situations is associated with a marked increase in cortisol release from the adrenal. Glucocorticoids are important in affecting many of the expressions of the stress response, such as increased gluconeogenesis, inhibition of growth and reproductive systems, and containment of inflammatory responses. Glucocorticoid release from the adrenal is regulated by adrenocorticotrophic hormone (ACTH) release from the pituitary, which in turn is regulated primarily by CRH release from the paraventricular nucleus (PVN) of the hypothalamus. In addition to the PVN, CRH is distributed in several brain areas that have been implicated in the behavioral and physiological responses to stress, including the central nucleus of the amygdala, hippocampus, prefrontal and cingulate cortices, locus coeruleus, thalamus, periaqueductal gray, and cerebellum. Intraventricular injection of CRH results in stress-related behaviors (e.g., fearfulness). Stressors early in life may have long-term effects on the CRH-HPA axis, including an increased glucocorticoid response to subsequent stressors. Glucocorticoids have profound effects on neuronal development, which means that early stressors could have lasting effects on development of the brain.

Studies in human subjects have validated the important role of glucocorticoids (cortisol in the human) for the stress response. Stress results in an acute increase in cortisol release, with psychological factors (such as

personal coping, appraisal, and previous stress exposure) modulating cortisol responsiveness. Long-term dysregulation of the HPA axis is associated with PTSD. Baseline levels of urinary cortisol were either decreased or unchanged in chronic PTSD, whereas decreased levels were found in 24-h samples of plasma cortisol levels. Cortisol may be elevated in the more acute phase of PTSD, although further research is needed in this area. A replicated finding has been a suppression of the cortisol response after lower doses of dexamethasone suppression test (DST) (0.5 mg), a finding that is the opposite of patients with major depression, who are nonsuppressors with the standard 1-mg DST test. PTSD patients had elevated levels of CRH in the cerebrospinal fluid. Other studies showed an increase in lymphocyte glucocorticoid receptors in PTSD and enhanced cortisol response to stress. One possible explanation of findings to date is an increase in neuronal CRH release, increased central glucocorticoid receptor responsiveness, and resultant low levels of peripheral cortisol due to enhanced negative feedback.

Other neurochemical systems may play a role in symptoms of PTSD. Benzodiazepinelike substances occur naturally in the brain and have been hypothesized to play a role in anxiety and stress. Animal models showed decreased benzodiazepine binding in the frontal cortex with acute and chronic stressors. Consistent with this, neuroimaging studies involving the quantitation of benzodiazepine receptor binding in the human brain showed decreased binding in frontal cortex of patients with PTSD. No difference was found in the anxiety response to the benzodiazepine receptor antagonist flumazenil between patients with combat-related PTSD and controls. Exposure to stress results in an increased release of opiate peptides in the brain associated with the development of an analgesia to pain known as stress-induced analgesia. Opiates cause a decrease in firing from the locus coeruleus; this provides an explanation for the favorable response of hyperarousal symptoms of PTSD to opiates such as heroin. Vietnam veterans with PTSD had a reduced sensitivity to pain during exposure to traumatic reminders in the form of a videotape with combat-related scenes; this was reversible with the opiate antagonist naloxone. Studies showed decreased levels of β -endorphin in the cerebrospinal fluid of patients with PTSD. Stressors result in increased serotonin (5-HT) turnover in the medial prefrontal cortex. Chronic restraint stress results in a decrease in 5-HT_{1A} binding in the hippocampus and in a decrease in 5-HT₂ binding in the parietal cortex. Patients with combat-related PTSD had reduced paroxetine-binding sites in platelets (evidence of reduced serotonin reuptake site function). Other studies found no difference in platelet

serotonin levels or uptake or in response to probes of 5-HT_{1A} function. Thyroid hormone (thyroxine, T₃) was found to be increased, with a pattern of increased thyrotropin-stimulating hormone response to thyrotropin-releasing hormone in PTSD. Other studies showed increased somatostatin levels in cerebrospinal fluid and decreased levels of neuropeptide Y (NPY) in plasma in PTSD.

Effects of Stress on the Neurobiology of Learning and Memory

There has been considerable interest in alterations in memory function in PTSD. Memory alterations in PTSD take the form of deficits in explicit verbal memory as well as in potentiation of recall of traumatic experiences, dissociative flashbacks, and changes in implicit memory (e.g., fear conditioning). Brain areas, including the hippocampus, amygdala, and medial prefrontal cortex, may mediate these alterations in memory. The hippocampus, a brain area involved in verbal declarative memory, is very sensitive to the effects of stress. Stress in animals was associated with damage to neurons in the CA3 region of the hippocampus with associated deficits in learning and memory. Studies in a variety of animal species suggested that direct glucocorticoid exposure results in decreased dendritic branching and a loss of neurons in the hippocampus. Glucocorticoids appear to exert their effect by increasing the vulnerability of hippocampal neurons to insults such as excitatory amino acids. Other factors, including alterations in serotonin function and decreased levels of brain-derived neurotrophic factor, also probably contribute to stress-related hippocampal damage. Recently, stress has been shown to inhibit neurogenesis, an effect that is reversed by antidepressant treatment.

Studies in PTSD are consistent with abnormalities of the hippocampus. PTSD patients show specific deficits in verbal declarative memory function based on measures that have been validated as sensitive to the loss of neurons in the CA3 region of the hippocampus. Patients with PTSD secondary to combat or childhood abuse also had a smaller volume of the hippocampus based on measurements with magnetic resonance imaging. Deficits in memory were correlated with reduced hippocampal volume in patients with combat-related PTSD. Declarative memory function improved and hippocampal volume increased with antidepressant treatments, including paroxetine and phenytoin. We have hypothesized that stress-induced hippocampal dysfunction may mediate many of the symptoms of PTSD that are related to memory dysregulation, including explicit memory deficits and

fragmentation of memory in abuse survivors. The hippocampus and adjacent cortex plays an important role in binding together information from multiple sensory cortices into a single memory at the time of retrieval. Abnormalities of hippocampal function in PTSD may impair the ability of the hippocampus to recreate a normal memory. Traumatic memories may not be placeable in space and time, leading to fragmentation or amnesia. Traumatic memories may intrude into consciousness without any discernible reason, leading to the symptom of intrusive memories and flashbacks.

The amygdala is involved in memory for the emotional valence of events. The paradigm of conditioned fear has been used as an animal model for stress-induced abnormalities of emotional memory. When an aversive stimulus (e.g., electric shock) is paired repeatedly with a neutral stimulus (bright light), exposure to the conditioned stimulus alone results in an increase in fear responsiveness and an increased startle response. This phenomenon, known as fear conditioning, is mediated by the amygdala. The startle response has been used as a probe of amygdala function in humans. In PTSD patients, the startle response has been shown to be either equal to or greater than that seen in normal subjects. Imaging studies have shown increased amygdala response to external threats, such as fear conditioning or exposure to fearful faces, but not during exposure to internally generated traumatic reminders.

The medial prefrontal cortex has been hypothesized to mediate symptoms of PTSD. The medial prefrontal cortex includes the anterior cingulate gyrus (Brodmann's area 32) and subcallosal gyms (area 25), as well as orbitofrontal cortex, and is functionally and anatomically distinct from the dorsolateral prefrontal cortex. Lesion studies demonstrated that the medial prefrontal cortex is involved in emotion and social behavior. Medial prefrontal cortical areas modulate emotional responsiveness through the inhibition of amygdala function, and dysfunction in these regions may underlie pathological emotional responses in patients with PTSD. Conditioned fear responses are extinguished following repeated exposure to the conditioned stimulus in the absence of the unconditioned (aversive, e.g., electric shock) stimulus. The extinguished memory is rapidly reversible following reexposure to the conditioned-unconditioned stimulus pairing even up to 1 year after the original period of fear conditioning, suggesting that the fear response did not disappear but was merely inhibited. This inhibition appears to be mediated by the medial prefrontal cortical inhibition of amygdala responsiveness. Studies in animals showed that early stress results in decreased brain neurons in the medial prefrontal cortex. Imaging studies in PTSD show a smaller volume of the anterior

cingulate. Imaging studies of brain function in PTSD are consistent with dysfunction of the medial prefrontal cortex during the presentation of traumatic cues. Stimulation of PTSD symptoms with the noradrenergic agent yohimbine resulted in a relative failure of activation of metabolism in parts of the medial prefrontal cortex, as well as decreased function in the hippocampus, as measured with the PET assessment of metabolism in PTSD. Exposure to traumatic reminders in the form of traumatic slides and sounds or traumatic scripts was associated with an increase in PTSD symptoms, decreased blood flow in the medial prefrontal cortex and anterior cingulate, and a failure of activation in the anterior cingulate (areas 32 and 24), as measured with PET. Traumatic reminders also resulted in increased activation in the posterior cingulate and motor cortex in two or more studies in PTSD. These areas are involved in visuospatial processing and preparation for action that may be critical to the fight-or-flight response. A decreased function in the hippocampus was also seen in two studies. In summary, dysfunction of the medial prefrontal cortex during exposure to traumatic reminders may underlie the symptoms of PTSD.

Concluding Remarks

Multiple neurotransmitter and neuropeptide systems are involved in the stress response, including CRH, norepinephrine, serotonin, dopamine, endogenous benzodiazepines, and endogenous opiates. These transmitters mutually regulate one another in the execution of the stress response. Chronic stress is associated with long-term alterations in these transmitters and peptides, which translate into changes in neuronal function and structure. Preliminary clinical studies in patients with PTSD are beginning to replicate many of the findings from animal studies. Specific brain regions, including the hippocampus and adjacent cortex, amygdala, and medial prefrontal cortex, are hypothesized to play a role in symptoms of PTSD. Neurotransmitter and neuropeptid systems involved in stress modulate behavior and memory function through direct actions on these brain regions. Chronic stress, probably through long-term alterations in these systems, can lead to long-term behavioral and memory disturbances. This may represent the mechanism for many of the symptoms of PTSD.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Amygdala; Learning and Memory, Effects of Stress on; Posttraumatic Stress Disorder in Children; Posttraumatic Stress Disorder, Delayed.

Further Reading

- Bremner, J. D. (2002). *Does stress damage the brain? Understanding trauma-related disorders from a mind-body perspective*. New York: W. W. Norton.
- Bremner, J. D. (2005). *Brain imaging handbook*. New York: W. W. Norton.
- Bremner, J. D., Randall, P. R., Scott, T. M., et al. (1995). MRI-based measurement of hippocampal volume in patients with combat-related posttraumatic stress disorder. *American Journal of Psychiatry* 152, 973–981.
- Bremner, J. D., Southwick, S. M. and Charney, D. S. (1999). The neurobiology of posttraumatic stress disorder: an integration of animal and human research. In: Saigh, P. A. & Bremner, J. D. (eds.) *Posttraumatic stress disorder: a comprehensive text*, pp. 103–143. New York: Allyn & Bacon.
- Elzinga, B. M., Schmah, C. S., Vermetten, E., et al. (2003). Higher cortisol levels following exposure to traumatic reminders in abuse-related PTSD. *Neuropsychopharmacology* 28(9), 1656–1665.
- Rauch, S. L., Whalen, P. J., Shin, L. M., et al. (2000). Exaggerated amygdala response to masked facial stimuli in posttraumatic stress disorder: a functional MRI study. *Biological Psychiatry* 47(9), 769–776.
- Sapolsky, R. M. (1996). Why stress is bad for your brain. *Science* 273, 749–750.
- Vermetten, E. and Bremner, J. D. (2002). Circuits and systems in stress. I: Preclinical studies. *Depression & Anxiety* 15, 126–147.
- Vermetten, E. and Bremner, J. D. (2002). Circuits and systems in stress. II: Applications to neurobiology and treatment of PTSD. *Depression & Anxiety* 16, 14–38.
- Vermetten, E., Vythilingam, M., Southwick, S. M., et al. (2003). Long-term treatment with paroxetine increases verbal declarative memory and hippocampal volume in posttraumatic stress disorder. *Biological Psychiatry* 54(7), 693–702.
- Yehuda, R., Teicher, M. H., Trestman, R. L., et al. (1996). Cortisol regulation in posttraumatic stress disorder and major depression: a chronobiological analysis. *Biological Psychiatry* 40(2), 79–88.

Posttraumatic Therapy

A M Rasmusson

Veterans Administration National Center for PTSD,
West Haven, and Yale University of Medicine,
New Haven, CT, USA

C M Monson and P A Resick

Veterans Administration National Center for PTSD and
Boston University School of Medicine, Boston, MA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
A M Rasmusson and D S Charney, volume 3, pp 192–200,
© 2000, Elsevier Inc.

Background: Diagnosis and Symptoms of Posttraumatic
Stress Disorder (PTSD) and Acute Stress Disorder (ASD)

Early Interventions

Treatment of PTSD

The Pharmacological Treatment of PTSD

Summary

Glossary

*Double-blind
study*

A research study in which both the research patient and the research clinician are blind or unaware of the type of medication or other intervention being used to treat the patient. This prevents the expectations of the patient and researcher

from biasing the research results. In such studies, a research pharmacist or another medical clinician not in direct contact with the patient knows which medication the patient is taking so that appropriate steps can be taken if the patient develops any side effects.

Extinction

The process of learning that a stimulus strongly paired with a traumatic event (e.g., gunshots) is no longer associated with the traumatic event (e.g., combat). This process requires that the person be re-exposed to trauma-related stimuli, and can cause high levels of distress. However, as the person learns that the traumatic event is no longer going to occur in association with that stimulus, anxiety, fear, defensive, and avoidance behaviors subside.

Flashback

Acting or feeling like one is re-experiencing a traumatic event after it is over. A flashback may seem to occur spontaneously or may occur in response to a sight, sound, smell, or other sensation associated with the initial trauma (e.g., in response to hearing a car backfire, seeing a person that resembles the perpetrator of an assault, feeling the ground vibrate, smelling fire). The person experiencing the flashback may retain a partial

Neuro-transmitter

awareness of reality or completely lose touch with reality.

A chemical released by neurons (nerve cells) in the brain or peripheral nervous system to communicate with other neurons. Neurotransmitters interact with receptors on nearby neurons, or on the neuron from which the neurotransmitters were released, to increase or decrease their activity. Medications used to treat PTSD are known to influence levels or efficacy of the following neurotransmitters or neurotransmitter receptors: serotonin, norepinephrine (noradrenalin), epinephrine (adrenaline), acetylcholine, dopamine, and γ -aminobutyric acid (GABA).

Placebo

An inactive pill given to persons participating in research studies to examine the effectiveness of new medications. Some psychiatric symptoms can improve in response to a placebo substance because individuals getting that substance think that the new medication will be helpful. This is called a placebo response. The symptom responses of persons getting a placebo medication are then compared to the responses of persons getting the real medication. Since real medications usually have some undesirable side effects, a real medication must be able to improve symptoms more than a placebo in order for the real medication to be considered effective and worthwhile.

Background: Diagnosis and Symptoms of Posttraumatic Stress Disorder (PTSD) and Acute Stress Disorder (ASD)

Diagnosis of PTSD

Posttraumatic stress disorder (PTSD) is diagnosed when criteria A, B, C, and D set forth in the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition, text revision (DSM-IV-TR) are met.

To meet criterion A, a person must have experienced or witnessed, or have been confronted with, an event perceived as extremely threatening (criterion A1) in association with an intense feeling of helplessness, horror, or fear (criterion A2). Examples of traumatic events include being physically assaulted, being raped, participating in combat, being a victim of a natural disaster such as an earthquake or tornado, being in a car accident, witnessing a murder, or hearing about the sudden and unexpected homicide of a loved one. Symptoms of PTSD that may develop

in response to a traumatic event comprise the B, C, and D criteria. At least one re-experiencing symptom (criterion B), three avoidance symptoms (criterion C), two hyperarousal symptoms (criterion D), and evidence of functional impairment must be present for a diagnosis of PTSD to be made.

Symptoms of PTSD

Re-experiencing symptoms Re-experiencing symptoms of PTSD include bothersome intrusive memories, thoughts, or images of the trauma, nightmares, flashbacks, and emotional or physiological reactions to stimuli associated with the initial trauma.

Avoidance/numbing symptoms Avoidance/numbing symptoms of PTSD include going out of one's way to avoid stimuli associated with the trauma (e.g., not driving after a car accident, avoiding 4th of July fireworks celebrations after experiencing combat trauma, using alcohol or other drugs of abuse to avoid feelings or memories of the trauma), not discussing the traumatic event because of feelings that get stirred up, partial or complete amnesia for the event, disinterest in everyday activities, emotional numbing, and the sense that one's future is dismal or will be cut short.

Hyperarousal symptoms Hyperarousal symptoms of PTSD include hypervigilance, irritability, anger or rage, difficulty concentrating, startle reactions, and sleep disturbances. The latter may include difficulty falling asleep, middle of the night awakening (often in reaction to nightmares), and early morning awakening. It is not uncommon for persons with untreated PTSD to regularly obtain less than 6 h of sleep per night.

Time Course Depending on the duration of PTSD symptoms and the onset of PTSD symptoms relative to the experience of trauma, three PTSD subtypes may be diagnosed. Treatment approaches may vary depending on the timing and course of symptoms. Acute PTSD is diagnosed when symptoms of PTSD remain or develop more than 1 month after the trauma. Symptoms of acute PTSD may last up to 3 months in duration. Chronic PTSD is diagnosed when PTSD symptoms last longer than 3 months, and delayed-onset PTSD is diagnosed when PTSD symptoms develop more than 6 months after the trauma.

Diagnosis of ASD

A diagnosis of acute stress disorder (ASD) requires the presence of at least three symptoms of dissociation (subjective numbing, detachment, or an absence

of emotional responsiveness, an unawareness of surroundings, derealization, depersonalization, or dissociative amnesia), along with one re-experiencing, one hyperarousal, and one avoidance symptom (discussed previously). By definition, these symptoms may last up to 4 weeks from the trauma. The development of ASD predicts the development of PTSD, though not all persons that develop ASD go on to develop PTSD, and not all persons that develop PTSD have experienced a readily diagnosable ASD.

Comorbid Disorders Associated with PTSD

Other psychiatric disorders such as depression, panic disorder, substance abuse, somatization disorder, phobias, and dissociative disorders are commonly associated with PTSD. It is thought that depression and panic disorder when associated with PTSD are neurobiologically distinct from these disorders diagnosed in the absence of trauma exposure. Alcohol, nicotine, or drug abuse may constitute attempts by the patient to maladaptively cope with PTSD symptoms. Thus, the presence of a substance use disorder in a person suffering from PTSD can, in many cases, be construed as an avoidance symptom that has significant secondary adverse effects on the person's adaptation. It is also possible that individuals with comorbid PTSD and substance use disorders have risk factors that predispose them to both. Interestingly, the risk for nicotine dependence is increased by having PTSD, while the risk for the development of PTSD, after trauma exposure, is increased by prior nicotine dependence. Chronic pain and stress-exacerbated medical disorders such as gastric ulcers or irritable bowel syndrome also occur frequently in persons with PTSD. The presence of these comorbid psychiatric and medical disorders can significantly complicate the treatment of PTSD (see later).

Trauma exposure can lead to an intensification of central neurochemical and peripheral neurohormonal responses to perceived threat. Associated state changes facilitate fight or flight reactions at the expense of higher-order cognitive efforts to appraise the actual degree of threat and enact appropriate, considered responses. Patients with PTSD therefore often have problems with impulse control and defensive aggression. Traumatic experiences in childhood and even early adulthood often lead to the development of affect dysregulation, cognitive misappraisals, interpersonal problems, and behaviors that cause self-harm. These symptoms are sometimes considered features of borderline, narcissistic, and antisocial personality disorders, while others consider them symptoms of complex PTSD. Trauma experienced during adulthood

may sometimes precipitate the emergence of similar problems along with the development of PTSD even in previously well-functioning individuals. Therapists engaging in the treatment of PTSD therefore should have familiarity with, comfort with, and experience in the management of these defensive, maladaptive patterns of function in order to maintain a therapeutic relationship within which treatment can proceed.

Early Interventions

The risk for developing PTSD after exposure to a traumatic event depends on the type and severity of the trauma, as well as the age, gender, and other premorbid characteristics of the person suffering the trauma. For example, PTSD is thought to occur at higher rates in response to unpredictable or uncontrollable events and in response to interpersonal trauma resulting from the breakdown of expectable psychosocial norms (e.g., rape, genocide, war atrocities, intrafamilial abuse). Research also suggests that a history of childhood trauma may predispose to the development of PTSD in response to adult trauma. Genetic factors may also predispose to the development of PTSD. Finally, the degree of psychosocial support experienced in the aftermath of trauma affects the progression of ASD to PTSD (see later).

Understanding the mechanisms by which the preceding factors contribute to the risk for PTSD will help lead to new methods of PTSD prevention. In the meantime, significant progress has been made in evaluating the efficacy of early interventions aimed at preventing the progression of acute trauma reactions to PTSD. Single-session debriefing therapy was the first such treatment widely employed. During debriefing therapy, trauma victims are educated about typical stress reactions and ways to cope with them. Victims are also encouraged to express feelings and thoughts about the trauma and are informed about other treatments that may be accessed later should the need arise. Controlled trials of debriefing therapy, however, have not only failed to show its efficacy, but also suggest that it might increase the risk for PTSD. Since these studies, research methodologists have emphasized the importance of comparing outcomes of experimental therapies to the natural evolution of trauma symptoms.

Much better outcomes have been found with cognitive-behavioral therapies that employ exposure techniques and/or cognitive restructuring interventions. These interventions have been introduced a few months after trauma exposure and are significantly more effective in preventing PTSD than supportive psychotherapies, the simple repetition of PTSD

symptom assessments, and self-help programs. Importantly, rates of PTSD remittance in response to early cognitive-behavioral interventions are also much higher than natural rates of PTSD remission.

Further research is needed to determine whether even earlier psychological or pharmacological interventions may protect against the development of PTSD, as suggested by preliminary controlled studies of short-term propranolol administration to trauma victims seen initially in an emergency room setting.

Treatment of PTSD

Exposure Therapy

Exposure therapies used to treat PTSD are derived from research into the mechanisms underlying classical and operant conditioning. For instance, it has been fairly well established that traumatic memories are indelibly encoded within the amygdala, a part of the brain that coordinates the hormonal, cardiovascular, behavioral, and cognitive responses to threat in mammals. Within the amygdala, memories of the naturally threatening (i.e., unconditioned) stimulus and the initially nonthreatening stimuli associated with the threatening stimulus become paired at a neuronal level, so that in the future, the previously neutral but now fear-conditioned stimuli alone can elicit fearful defensive responses. Over time, if the individual repeatedly reencounters conditioned stimuli in the absence of the original threatening stimulus, new information about the safety of the conditioned stimuli restrains activation of the amygdala, and the expression of conditioned fear eventually extinguishes. Indeed, it is believed that the success of exposure therapy in PTSD relies in large part on the engagement and enhancement of this natural extinction process.

There are several types of exposure therapy used in the treatment of PTSD. These treatments vary in three ways: (1) the exposure to trauma-related stimuli may be imaginal or involve actual confrontation with feared persons, places, or things; (2) the exposure may be short or long; and (3) the level of emotional arousal stimulated by the exposure may be low or high. In systematic desensitization, patients are instructed to relax, and then are asked to imagine a minimally threatening stimulus for a short period of time. This is repeated several times until the stimulus ceases to cause anxiety. The patient is then asked to imagine a more fearful stimulus for several repetitions until it ceases to elicit anxiety. This process continues until the most threatening situation can be comfortably confronted by the patient. In flooding, the patient starts by confronting the most highly

feared stimuli either in the imagination or *in vivo*. The exposure is prolonged and elicits very high levels of anxiety.

Perhaps the most rigorously studied type of exposure therapy is called prolonged exposure (PE). PE is typically conducted in 7 to 10 sessions and involves a combination of imaginal and *in vivo* exposure. After education regarding PTSD, rationale building, and breathing retraining, the therapist and client develop a hierarchy of feared situations for *in vivo* exposure practice between sessions. Then during sessions, which are 90 min in length, imaginal exposure is conducted in which the client is asked to relate the worst traumatic event, in present tense, including as many sensory details as possible. The account is repeated for 45 to 60 min during the session, which is audiotaped. The client listens to the tape between sessions until the next exposure session is taped.

In well-designed studies, PE has been found to result in remission of PTSD in approximately 50% of patients treated. Further studies have suggested that the degree of arousal experienced during exposure therapy influences its outcome. Thus, future research will be needed to identify factors that interfere with achievement of optimal levels of arousal necessary for extinction to occur. Interventions then could be developed to mitigate the impact of such factors. For example, Foa and Meadows have described several anxiety management techniques that could be used by patients with PTSD to manage disabling levels of anxiety induced by exposure to trauma-related cues. Other techniques may help patients who willfully or automatically avoid experiencing necessary levels of arousal and distress during exposure therapy. For patients with neurobiologically mediated tendencies toward over- or underarousal, medications might be used in combination with exposure techniques.

Cognitive Therapy

Cognitive therapy aims to identify, challenge, and restructure automatic systems of meaning that have become attached to traumatic experiences and that interfere with the ability to accurately interpret current events. Cognitive theories of PTSD focus on the effect that various beliefs and schemas (automatic rules) have on emotions and how the memory of the traumatic event may be discrepant with a person's understanding of him- or herself and his or her world view. Cognitive therapy assists the client in examining beliefs about the traumatic event, his or her self, and the world and helps the client challenge possible erroneous and maladaptive assumptions. In several studies, cognitive therapy has been found to be as effective as exposure therapy.

Combination Strategies

Cognitive processing therapy (CPT) Cognitive processing therapy (CPT) utilizes a combination of exposure and cognitive therapy over the course of 12 1-h sessions. CPT was first shown to be efficacious in women with assault-related chronic PTSD, inducing remission of PTSD in about 50% of those who initiated treatment and 80% of those who completed treatment. Since then, CPT has been successfully extended to the treatment of male Vietnam veterans with chronic combat-related PTSD.

During the exposure component of CPT, the patient writes an account of his or her worst traumatic event, reads the account daily, and then reads it to the therapist in session. The patient is encouraged to experience the full range of emotions associated with that event. Most people with PTSD re-experience the traumatic event as only brief sensory images rather than the complete story. Writing the account in detail restores the full event within its greater context. In addition, reviewing the details of the traumatic experience allows the therapist and patient to consider possible assumptions that were made at the time of the trauma and distortions in the patient's retrospective evaluation of the traumatic event that have interfered with recovery. For example, victims of trauma often overestimate the amount of control they had over the event. Self-blame by patients is frequent even when it is obvious to everyone else that the event was not preventable. These cognitive distortions, believed to play a role in the maintenance of PTSD, are referred to as stuck points.

During the subsequent cognitive therapy sessions, patients are taught to systematically challenge their counterproductive thoughts by asking and answering a series of questions and then generating more balanced alternative thoughts. During CPT, new reality-based thoughts begin to supplant the conditioned and overgeneralized automatic beliefs, leading to a marked reduction in re-experiencing, avoidance, and hyperarousal symptoms. Systematically addressing deleterious effects of the traumatic experiences on the victim's self-concept and relationships with others also results in significant reductions in depression and guilt. Indeed, CPT has been found to be modestly more effective than prolonged exposure therapy in reducing depression. It is also more effective in reducing guilt and somatic symptoms comorbid with PTSD.

The therapeutic effects of these evidence-based cognitive-behavioral therapies can be profound and have been shown, in a long-term follow-up study, to persist up to 5 years after treatment.

Eye movement desensitization reprocessing (EMDR) During eye movement desensitization reprocessing

(EMDR), a patient tracks the movement of a therapist's finger, while visualizing a traumatic scene and paying attention to any thoughts or feelings provoked by the scene. The therapist takes note of whether the patient's appraisal of the circumstances or meaning of the trauma may be distorted and thereby contribute to negative feelings associated with the incident. This process is repeated until the anxiety generated by the traumatic scene subsides. At this point, the therapist interjects an alternate, and ideally more realistic or constructive, thought about the trauma scene for the patient to consider. This positive association is thought to compete with negative associations and contribute to the extinction of conditioned fear.

It is important to note that recent research has suggested that having the patient follow the movement of the therapist's finger may lessen the intensity of the emotional response to the visualized traumatic material. Thus, having the patient engage in other mildly distracting repetitive movements also may be useful in optimizing the degree of arousal attained during this imaginative exposure exercise and thus may facilitate extinction.

A recent study has found that prolonged exposure therapy results in greater improvements in PTSD symptoms than EMDR, while EMDR did not differ significantly in efficacy from relaxation therapy. A head-to-head comparison between CPT and EMDR has not been conducted.

Treatment of Complicated PTSD

As noted, the presence of comorbid medical and psychiatric conditions can significantly complicate the treatment of PTSD. Although it has been generally believed that substance use disorders need to be addressed before the treatment of PTSD, there have been recent efforts to treat comorbid substance abuse and PTSD simultaneously, with positive effects. States of intoxication as well as withdrawal intensify PTSD symptoms and may compromise the natural biological processes involved in extinction and the development of more reality-based cognitions. In addition, substance dependence results in secondary psychosocial problems (legal, vocational, interpersonal) that further stress and destabilize individuals struggling with PTSD. Comorbid depression also may complicate recovery from PTSD. While both PE and CPT have been shown to significantly decrease depressive symptoms in patients with PTSD, women with PTSD and comorbid major depression have significantly higher self-reported PTSD and depression symptoms at baseline and after treatment with CPT and PE compared to women with PTSD alone. Severe problems with affect regulation frequently

seen in victims of childhood abuse can also present a challenge during PTSD treatment. In addition, chronic pain syndromes associated with PTSD need to be addressed because chronic pain appears to increase PTSD symptoms and vice versa.

Therefore, treatments for PTSD in persons with comorbid psychiatric and medical conditions deserve attention in future research studies. For example, severe dysregulation of affect has been specifically targeted in one therapy format prior to undertaking prolonged exposure therapy in victims of childhood abuse. It has been suggested that comorbid major depression or panic disorder should be addressed pharmacologically prior to undertaking formal treatment of core PTSD symptoms. On the other hand, it is possible that substantial numbers of such patients may be refractory to antidepressants or other medications currently available (see later), or that therapeutic techniques as well as medication administration may be helpful. The treatment of chronic pain syndromes associated with PTSD can also be difficult. Standard pain medications such as nonsteroidal anti-inflammatory agents may be ineffective, while tolerance to more effective agents such as opiates can occur; in addition, the possibility that drugs such as opiates may interfere with exposure and cognitive techniques needs to be evaluated. Antidepressants may effectively treat both chronic pain and PTSD in some individuals, but do not benefit all.

The Pharmacological Treatment of PTSD

Over the past several years, significant progress has been made in evaluating the efficacy of psychotropic medications used in the treatment of PTSD. A number of placebo-controlled trials of medication for the treatment of PTSD have been conducted, which has resulted in an increase in the number of agents that can be prescribed with some confidence. In addition, open label studies and case reports continue to suggest other promising medications that likely will be tested in future controlled trials.

Another challenge in the future will be to determine whether psychotropic medications might enhance the efficacy of evidence-based psychotherapies for PTSD. While these psychotherapies are relatively brief and appear to effectively cure PTSD in many individuals, many individuals remain symptomatic after treatment. In addition, otherwise well-controlled trials of these therapies have not controlled for medication use. Thus, we cannot be sure whether medication use might have enhanced or hindered the effects of these psychological treatments in some individuals. Comparisons between the psychological and medication treatments are also difficult because active

psychological treatments are generally compared to no treatment, while active medications are compared to nonactive placebo medications that typically lead to some improvement in PTSD symptoms themselves. This could contribute to the greater apparent efficacy of psychotherapies compared to medications.

Antidepressants Used in the Treatment of PTSD

Selective serotonin reuptake inhibitors (SSRIs) Selective serotonin reuptake inhibitors (SSRIs) interfere with the reuptake of serotonin into neurons from which serotonin is released. In addition, SSRIs appear to increase brain levels of neuroactive steroids that reduce anxiety, arousal, and pain. A number of preliminary open label studies, as well as placebo-controlled double-blind studies, have shown that SSRIs such as fluoxetine, sertraline, paroxetine, and citalopram are helpful in treating core symptoms of PTSD. For example, studies of sertraline show approximately 30% reductions in PTSD symptoms. There is some evidence that the efficacy of these agents is greater either in women or in victims of civilian trauma compared to male victims of combat trauma.

Tricyclic antidepressants Tricyclic antidepressants increase brain levels of neurotransmitters such as norepinephrine and serotonin by interfering with their reuptake into neurons after their release. Tricyclic antidepressants may be helpful in some patients with PTSD in the treatment of depression, as well as in the treatment of re-experiencing and hyperarousal symptoms. However, these agents are generally less effective than either monoamine oxidase inhibitors (MAOIs) or SSRIs in the treatment of core symptoms of PTSD. The anticholinergic side effects of tricyclics appear to be less well tolerated in patients with PTSD compared to other patient groups.

Monoamine oxidase (MAO) inhibitors MAO inhibitors are medications that inhibit enzymes that break down monoamine neurotransmitters such as dopamine, serotonin, and norepinephrine. The levels of these neurotransmitters in the brain and peripheral nervous system therefore increase. MAOIs have been shown in double-blind, placebo-controlled studies to be helpful in treating re-experiencing symptoms in PTSD. The use of MAO inhibitors is contraindicated in patients with comorbid substance use because of potential life-threatening interactions with illicit drugs and alcohol. Patients on MAOIs also must adhere to a special diet devoid of foods and beverages with tyramine (an amino acid that interacts with neurotransmitters increased by MAOIs) in order to avoid potentially dangerous side effects. Many

patients are not willing or reliably able to restrict their diet in this manner.

Sleep Medications

A number of medications can be useful in promoting sleep in PTSD. These include the antidepressants trazodone, mirtazapine, and amitriptyline, atypical neuroleptics (see next section), antiadrenergic agents (see later), and drugs that target the γ -aminobutyric acid type A (GABA_A) receptor such as the benzodiazepines. Problems with tolerance frequently complicate use of the latter, however.

Atypical Neuroleptics Used in the Treatment of PTSD

Neuroleptics act by blocking the effects of neurotransmitters such as dopamine, serotonin, and norepinephrine at their receptors. Traditional neuroleptics primarily block dopamine type 2 (D₂) receptors while newer atypical neuroleptics primarily block serotonin type 2 receptors (5-HT₂) and sometimes noradrenergic α 1 or α 2 receptors. Before PTSD was formally recognized as a distinct psychiatric disorder, veterans displaying symptoms of hypervigilance, paranoia, agitation, dissociation, trauma-related auditory hallucinations, and behavioral dyscontrol were sometimes treated with traditional neuroleptics. Neuroleptic use in PTSD was thereafter discouraged due to their poor efficacy, intolerable short-term side effects, and potential for inducing long-term movement disorders. More recently, the newer atypical neuroleptics have been found to be both tolerable and efficacious over the short term in treating severe combat-related PTSD, particularly impulsive aggression, irritability, and sleep disturbance. The long-term effectiveness, tolerability, and long-term potential for side effects associated with these agents have not yet been studied, however.

Other Medications of Potential Use in the Treatment of PTSD

A variety of other medications also may be useful in treating PTSD. However, double blind, placebo-controlled studies of the efficacy of these agents either alone or in combination with other medications have not yet been completed.

Antiadrenergic agents Antiadrenergic medications such as propranolol (which blocks the effects of epinephrine at β -receptors) or clonidine (which inhibits the release of norepinephrine) have been reported to effectively treat hyperarousal and re-experiencing symptoms in PTSD. Large controlled trials of propranolol administered in the immediate aftermath of trauma for the prevention of PTSD are currently

under way. Prazosin, a noradrenergic α 1 receptor antagonist, has shown some promise in reducing nightmares in PTSD.

Anticonvulsants Preliminary studies suggest that anticonvulsants such as carbamazepine or valproic acid have variable effects on hyperarousal, re-experiencing, and emotional numbing in PTSD. Benzodiazepines (tranquilizers) may be used for the nonspecific symptoms of irritability, anxiety, and insomnia, though tolerance and dependence are often a problem. The use of benzodiazepines in patients with past alcohol or drug abuse problems may be unwise since it may serve to trigger relapse; if prescribed, careful monitoring is necessary. A recent open label trial of topiramate, an anticonvulsant that antagonizes certain excitatory glutamate receptors while enhancing the effects of GABA at a novel subtype of the GABA_A receptor, suggests that it may be beneficial and well tolerated in PTSD. Controlled trials of this agent will need to be conducted to confirm its efficacy.

Pharmacological agents under development As our understanding of the underlying neurobiology of PTSD increases, the potential for the development of pharmacological agents that counteract deleterious biological changes associated with PTSD increases. For example, studies have demonstrated an increase in the central nervous system level of the peptide (protein) neurotransmitter called corticotropin-releasing hormone (CRH) in chronically stressed animals and in combat veterans with PTSD. CRH has been shown to confer an increase in stress sensitivity and to increase the expression of defensive behaviors in many animal models. Thus, drugs that interfere with the effects of CRH at its receptors (CRH antagonists) are under development for the possible treatment of PTSD.

Similarly, recent research suggests that blood levels of a substance called neuropeptide Y (NPY) may be decreased in combat veterans with PTSD. Centrally, NPY has been found to counteract the anxiogenic effects of CRH and to be generally anxiolytic. In addition, NPY acts to inhibit the release of norepinephrine in response to stress; therefore, a decrease in blood NPY levels may contribute to the hyperactivity of the peripheral sympathetic nervous system and the related fight-or-flight behaviors seen in PTSD. Consistent with a role for NPY in modulating hyperactive stress responses, low plasma NPY levels have been found in combat veterans with PTSD, while higher plasma NPY levels are associated with better performance and decreased dissociation symptoms in military personnel undergoing intense, stressful mock interrogation trials during training. Drugs that target

the NPY system are therefore under development and are being considered for the treatment of PTSD.

Synthetic neuroactive steroids that facilitate action of the inhibitory neurotransmitter GABA at GABA_A receptors are also in development. These steroids have been shown to exert potent anxiolytic, anesthetic, and sedative actions. Recent data suggest that cerebrospinal fluid levels of such neuroactive steroids are low in women with PTSD. Finally, medications that target excitatory, glutamatergic neurotransmitter system receptors such as the N-methyl-D-aspartate (NMDA) receptor are being considered for treatment of PTSD. D-cycloserine, for instance, has been shown to enhance the rate of recovery from specific phobias and thus may facilitate extinction of conditioned fear in patients with PTSD.

Summary

Substantial progress has been made in evaluating the efficacy of a variety of psychological and medication treatments for PTSD. Given the demonstrated efficacy of combinations of psychotherapeutic interventions, it is even beginning to appear that chronic PTSD need not be considered chronic. However, there is further work to be done. Many patients with PTSD remain relatively treatment resistant. Biological as well as psychological characteristics of these trauma victims may be important determinants of treatment resistance. Future research needs to better define such factors so that current treatment techniques can be appropriately modified and new treatment approaches developed. Combining pharmacological treatments targeting specific underlying neurobiological abnormalities with psychotherapeutic approaches is likely to be one such new approach. In addition, research aimed at understanding the pathogenesis of PTSD must be continued so that better methods of PTSD prevention and intervention can be developed.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Anxiety; Posttraumatic Stress Disorder, Delayed; Posttraumatic Stress Disorder, Neurobiology of.

Further Reading

- Albucher, R. C. and Liberzon, I. (2002). Psychopharmacological treatment in PTSD: a critical review. *Journal of Psychiatric Research* 36, 355–367.
- Armony, J. L. and LeDoux, J. E. (1997). How the brain processes emotional information. *Annals of the New York Academy of Science* 281, 259–270.

- Cloitre, M., Koenen, K. C., Cohen, L. R. and Han, H. (2002). Skills training in affective and interpersonal regulation followed by exposure: a phase-based treatment for PTSD related to childhood abuse. *Journal of Consulting and Clinical Psychology* 70, 1067–1074.
- Ehlers, A. and Clark, D. (2003). Early psychological interventions for adult survivors of trauma: a review. *Biological Psychiatry* 53, 817–826.
- Foa, E. B. and Meadows, E. A. (1997). Psychosocial treatments for posttraumatic stress disorder. A critical review. *Annual Reviews of Psychology* 48, 448–480.
- Hien, D. A., Cohen, L. R., Miele, G. M., Litt, L. C. and Capstick, C. (2004). Promising treatments for women with comorbid PTSD and substance use disorders. *American Journal of Psychiatry* 16, 1426–1432.
- Marmar, C. R., Neylan, T. C. and Schoenfeld, F. B. (2002). New directions in the pharmacotherapy of posttraumatic stress disorder. *Psychiatric Quarterly* 73, 259–270.
- Morgan, C. A., III, Rasmusson, A. M., Wang, S., et al. (2002). Neuropeptide Y and subjective distress in humans exposed to acute stress: replication and extension of a previous report. *Biological Psychiatry* 52, 136–142.
- Najavits, L. M., Weiss, R. D., Shaw, S. R. and Muenz, L. R. (1998). “Seeking safety”: outcome of a new cognitive-behavioral psychotherapy for women with posttraumatic stress disorder and substance dependence. *Journal of Traumatic Stress* 11, 437–456.
- Nishith, P., Nixon, R. D. and Resick, P. A. (2005). Resolution of trauma-related guilt following treatment of PTSD in female rape victims: a result of cognitive processing therapy targeting comorbid depression? *Journal of Affective Disorders* 86, 259–265.
- Pitman, R. K., Sanders, K. M., Zusman, R. M., et al. (2002). Pilot study of secondary prevention of posttraumatic stress disorder with propranolol. *Biological Psychiatry* 51, 189–192.
- Rasmusson, A., Pinna, G., Weisman, D., et al. (2005). Decreased CSF allopregnanolone levels in women with PTSD correlate negatively with reexperiencing symptoms. Society of Biological Psychiatry, poster presentation. May 19–21, 2005, Atlanta, GA.
- Resick, P. A., Nishith, P., Weaver, T. I., Astin, M. C. and Feuer, C. A. (2002). A comparison of cognitive processing therapy with prolonged exposure and a waiting condition for the treatment of chronic posttraumatic stress disorder in female rape victims. *Journal of Consulting and Clinical Psychology* 70, 867–879.
- Ressler, K., J., Rothbaum, B. O., Tannenbaum, L., et al. (2004). Cognitive enhancers as adjuncts to psychotherapy: use of D-cycloserine in phobic individuals to facilitate extinction of fear. *Archives of General Psychiatry* 61, 1136–1144.
- Taylor, S., Thordarson, D. S., Maxfield, L., et al. (2003). Comparative efficacy, speed, and adverse effects of three PTSD treatments: exposure therapy, EMDR, and relaxation training. *Journal of Consulting and Clinical Psychology* 71, 330–338.

Preeclampsia See: Eclampsia and preeclampsia.

Pregnancy – Maternal and Perinatal Stress – Effects of

P Smirnaki and M-A Magiakou

University of Athens School of Medicine, Athens, Greece

© 2007 Elsevier Inc. All rights reserved.

Stress, the Stress System and the Hypothalamic-Pituitary-Adrenal Axis in the Nonpregnant State

The Maternal Hypothalamic-Pituitary-Adrenal Axis during Pregnancy

The Hypothalamic-Pituitary-Adrenal Axis of the Fetus

Maternal-Placental-Fetal Stress Physiology

Effects of Prenatal and Perinatal Stress

Glossary

<i>Corticotropin releasing hormone (CRH)</i>	The key hormone of the stress system secreted by the hypothalamus; it stimulates corticotropin production and secretion by the pituitary, which, in turn stimulates cortisol production by the adrenal cortex. During pregnancy, CRH is also produced by the placenta.
<i>Early-life programming</i>	Structural and functional changes in prenatal development caused by early-life environmental factors influence; these persist in the later life or the entire life span of the offspring.
<i>Glucocorticoids</i>	Hormones produced and secreted by the adrenal glands in a baseline circadian fashion and in response to stress; the main endogenous glucocorticoid in humans is cortisol.
<i>Hypothalamic-pituitary-adrenal (HPA) axis</i>	The brain-controlled activation pathway for the secretion of corticosteroid hormones during stress through a hormonal cascade.
<i>Stress</i>	A state caused by the real or perceived challenge to homeostasis, accompanied by various hormonal responses.
<i>Stress system</i>	A complex system of brain circuits and peripheral effectors that are activated during stress. Its main components are the CRH and locus ceruleus–norepinephrine (LC–NE) arousal and autonomic systems and their peripheral effectors, the

pituitary-adrenal axis and the limbs of the autonomic nervous system.

Preeclampsia/eclampsia

Pregnancy associated with disturbances in maternal placental circulation, arterial hypertension, and activation of the acute-phase reaction and blood hypercoagulability. It is a potentially lethal state, frequently associated with spontaneous premature labor and delivery; artificial induction of delivery may be necessary for its control.

Stress, the Stress System and the Hypothalamic-Pituitary-Adrenal Axis in the Nonpregnant State

Living organisms maintain a complex dynamic equilibrium (homeostasis) that is constantly challenged by intrinsic and extrinsic adverse forces or stressors. In humans, the stress response (adaptive responses of the organism to various stressors) results in the activation of a complex repertoire of behavioral and physiological responses. The main components of the stress system are the corticotropin releasing hormone (CRH) and locus ceruleus–norepinephrine (LC–NE) autonomic systems and their peripheral effectors, the pituitary-adrenal axis and the limbs of the autonomic system ([Figure 1](#)).

During stress, the hypothalamic-pituitary-adrenal (HPA) axis is activated in response to a variety of stressors, and CRH appears to mediate the metabolic, circulatory, and behavioral changes that are necessary for an effective response. CRH, produced and secreted by the hypothalamus, stimulates adrenocorticotrophic hormone (ACTH) production and secretion by the pituitary. In turn, ACTH stimulates the production and secretion of cortisol by the adrenal cortex. The regulation of these hormones is achieved by complex negative feedback mechanisms.

CRH plays an important role in inhibiting gonadotropin-releasing hormone (GnRH) secretion during stress, and via somatostatin it also inhibits growth hormone (GH), thyrotropin-releasing hormone (TRH), and thyrotropin (TSH) secretion, thus suppressing

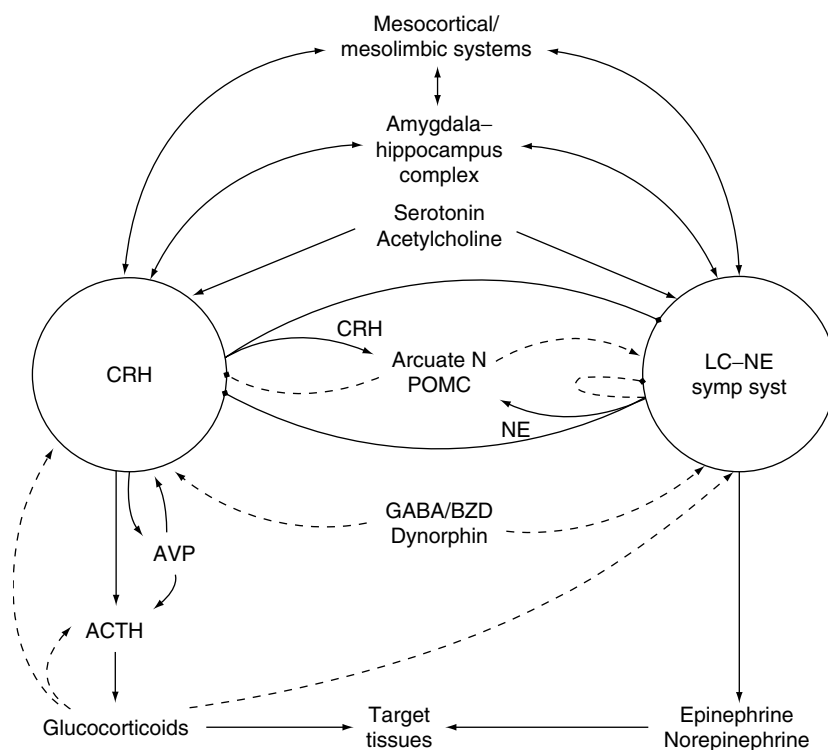


Figure 1 Brain circuits of the stress system. ACTH, adrenocorticotrophic hormone; Arcuate N, arcuate nucleus; AVP, arginine vasopressin; BZD, benzodiazepine; CRH, corticotropin releasing hormone; GABA, γ -aminobutyric acid; LC–NE symp syst, locus ceruleus–norepinephrine sympathetic system; NE, norepinephrine; POMC, proopiomelanocortin. From Chrousos, G. P. and Gold, P. W. (1992), The concepts of stress and stress system disorders: overview of physical and behavioral homeostasis, *Journal of the American Medical Association* **267**, 1244–1252.

Table 1 Glucocorticoid functions^a

Organ system	Functions
All fetal organ systems	Maturation (especially of the lung)
CNS	Suppression of pituitary gonadotropins, GH, and TSH Inhibition of the CRH, LC–NE, and β -endorphin systems Stimulation of the mesocorticolimbic dopaminergic system, the hippocampus, and the CRH perididergic central nucleus of the amygdala
Adipose tissue	Promotion of visceral adiposity, insulin resistance, dyslipidemia
Cardiovascular system	Enhancement of endothelin-induced vasoconstriction Attenuation of endothelium-dependent vasorelaxation Affecting renin–angiotensin system receptor density Alternation prenatally of the development of cardiac noradrenergic and sympathetic processes
Liver	Regulation of expression of critical hepatic metabolic enzymes, which regulate gluconeogenesis
Bone	Hepatic insulin resistance Low-turnover osteoporosis
Autoimmune system	Immunosuppression

^aCNS, central nervous system; CRH, corticotropin releasing hormone; GH, growth hormone; LC–NE, locus ceruleus–norepinephrine; TSH, thyrotropin.

reproductive, growth, and thyroid functions. Interestingly, these three functions depend on positive catecholaminergic input. On the other hand, the end hormones of the HPA axis, the corticosteroids, play multiple adaptive roles (Table 1). They simultaneously inhibit the CRH, LC–NE and β -endorphin systems

and stimulate the mesocorticolimbic dopaminergic system and the CRH-peptidergic central nucleus of the amygdala. In addition, they directly inhibit pituitary gonadotropin, GH, and TSH secretion. They also have direct and insulin-mediated effects on adipose tissue, ultimately promoting visceral adiposity, insulin

resistance, dyslipidemia, and hypertension (metabolic syndrome X) and direct effects on the bone, causing low-turnover osteoporosis. Central CRH via glucocorticoids and catecholamines inhibits the inflammatory reaction, whereas CRH directly secreted by peripheral nerves stimulates local inflammation.

The stress response is meant to be acute or at least of a limited duration. The time-limited nature of this process renders its accompanying antianabolic, catabolic, and immunosuppressive effects temporally beneficial and of no adverse consequences. The chronic, excessive, or deficient activation of the stress system, on the other hand, may lead to abnormal reactivity of the HPA axis, mostly in the form of increased or decreased cortisol secretion. In both cases, pathological adaptation may have detrimental effects on the organism.

The Maternal Hypothalamic-Pituitary-Adrenal Axis during Pregnancy

Pregnancy is a period of hypercortisolism. From weeks 8–10 of gestation on, CRH is also produced by the placenta and secreted primarily into the maternal and also into the fetal circulation. Placental CRH (pCRH) has the same biological activity as hypothalamic CRH (Table 2). pCRH has been postulated to stimulate maternal pituitary ACTH release and cortisol secretion, thereby contributing to pregnancy-associated hypercortisolemia. Paradoxically, cortisol stimulates the synthesis and release of pCRH in a positive feedback regulation fashion. pCRH secretion is also stimulated by inflammatory cytokines, mechanical stretch, and anoxic conditions and therefore by inflammatory and noninflammatory stress, including the stress of preeclampsia or eclampsia. The alterations mentioned result in a shift from normal negative feedback regulation of the maternal HPA axis to a feed-forward mechanism of peripherally produced pCRH, which stimulates adrenal cortisol secretion (Figure 2).

Generally, under normal conditions, circulation pCRH is inactivated to a large extent by a CRH

binding protein (CRH-BP). During the second half of pregnancy, plasma CRH levels increase exponentially, whereas the CRH-BP concentration starts decreasing at about weeks 33–34 of gestation. The outcome is even higher levels of bioactive CRH, which reaches a peak during labor and delivery and then rapidly decreases postpartum. The increase in CRH levels in the mother's plasma toward the end of pregnancy is normal. HPA axis hormones (CRH; ACTH; corticosteroids; and dehydroepiandrosterone sulfate, DHEAS) are involved in the timing of the onset of the birth process in humans. pCRH causes vasodilation of the fetoplacental circulation by activating nitric oxide synthase and participates in labor by stimulating the secretion of prostaglandin E2a and prostaglandin E2 by the fetal membranes and by promoting the constriction of myometrial cells. At the end of normal pregnancy, the simultaneous stimulation of organ maturation and initiation of parturition constitute a positive effect of a feed-forward mechanism. However, the premature activation of this system may result in preterm labor and delivery. This is supported by evidence of an earlier association between prematurely raised blood pCRH (and decreased CRH-BP) levels in pregnant women who had a preterm delivery. Both preeclampsia and eclampsia are characterized by prematurely elevated levels of pCRH and spontaneously premature labor and delivery.

At the end of parturition, the placental source of CRH is lost; thus, the hypercortisolism of pregnancy is followed by a transient suppression of hypothalamic CRH secretion in the postpartum period. The readjustment of the HPA axis and the postpartum recovery of hypothalamic CRH activity remains poorly characterized and varies among women. The hypothalamic-pituitary-ovarian axis is suppressed

Table 2 Placental corticotropin releasing hormone and its potential physiological functions

Sites of production	Functions
Cytotrophoblast, syncytiotrophoblast, amnion, chorion	Maternal hypercortisolism Fetoplacental circulation Timing of labor Fetal adrenal zone stimulation Labor and delivery

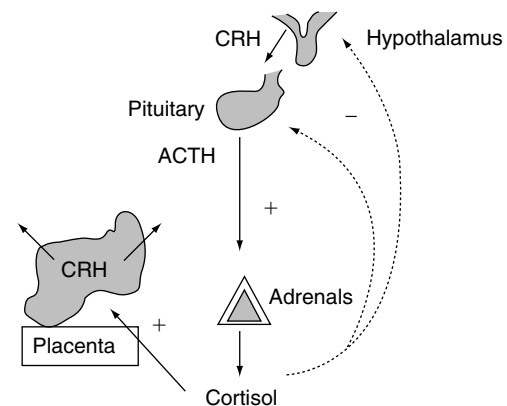


Figure 2 The hypothalamic-pituitary-adrenal axis and placental CRH in human pregnancy. ACTH, adrenocorticotrophic hormone; CRH, corticotropin releasing hormone. From Chrousos, G. P., Torpy, D. and Gold P. W. (1998), Interactions between the hypothalamic-pituitary-adrenal axis and the female reproductive system: clinical implications, *Annals of Internal Medicine* 129, 229–240.

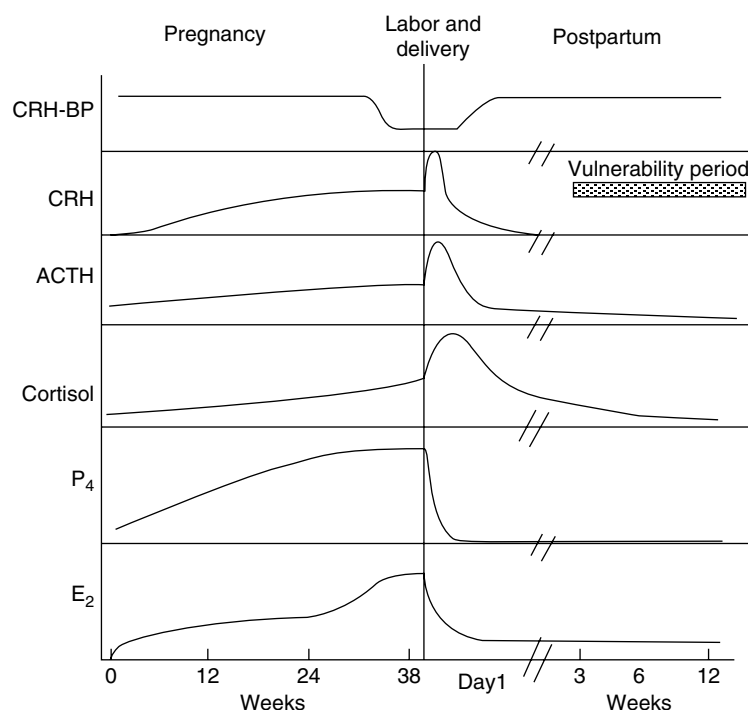


Figure 3 Hormonal changes and period of increased vulnerability to mood disorders and autoimmune phenomena during pregnancy and postpartum. ACTH, adrenocorticotrophic hormone; CRH, corticotropin releasing hormone; CRH-BP, CRH binding protein; E₂, 17 β -estradiol; P₄, progesterone. From Chrousos, G. P., Torpy, D. and Gold P. W. (1998), Interactions between the hypothalamic-pituitary-adrenal axis and the female reproductive system: clinical implications, *Annals of Internal Medicine* **129**, 229–240.

postpartum and does not recover for approximately 4–6 weeks. Both hypothalamic CRH hyposecretion and abrupt estrogen withdrawal may explain the postpartum blues or depression and autoimmune phenomena frequently seen during this period (Figure 3). Estrogen administration has been shown to ameliorate postpartum depression, probably by returning to normal the secretion of the suppressed CRH neuron.

During pregnancy, ACTH and cortisol continue to be secreted in a pulsatile and circadian fashion, whereas the secretion of pCRH is pulsatile but not circadian. The lack of circadian correlation between CRH and ACTH/cortisol is not surprising because the major part of CRH comes from the placenta, is bound to CRH-BP, and has limited action on the maternal corticotrophs. Moreover, another secretagogue, possibly arginine vasopressin (AVP), released at the median eminence from terminals of the parvocellular neurons, may be important in regulating pulsatile and circadian ACTH secretion during this period.

The Hypothalamic-Pituitary-Adrenal Axis of the Fetus

Fetal cortisol is important for the maturation of virtually all fetal organ systems. The importance of the increase in fetal cortisol secretion for the maturation

of the fetal lung and other systems is well established. The fetal HPA axis is regulated through negative feedback starting early in gestation. In a later stage of pregnancy, pCRH enters the fetal circulation via the umbilical vein. Maternal HPA axis hormones also stimulate the placenta to produce CRH. pCRH stimulates the fetal HPA axis to produce and secrete ACTH, cortisol, and androgens (DHEAS, which at the placenta is converted to estriol). Subsequently, fetal cortisol enters the placental circulation via the umbilical vessels and further stimulates the production of pCRH. This way, the fetal HPA axis is also regulated by a feed-forward mechanism at the end of pregnancy. This results in a large increase in cortisol through which maturation of the fetal organs is enhanced.

Maternal-Placental-Fetal Stress Physiology

It is most likely that maternal stress influences the fetus through (stress) hormones. The following three mechanisms have been proposed, which may operate simultaneously and may amplify one another's effects.

1. Reduced uteroplacental blood flow. Corticosteroids and catecholamines exert strong effects on peripheral blood vessels, and the placenta is replete

with receptors for these hormones. The activation of the sympathetic nervous system by stress may lead to reduced blood flow to the uterus and fetus, and it may contribute to fetal growth restriction.

2. Transplacental transport of maternal stress hormones. The human fetus is relatively protected against direct exposure to high cortisol concentrations. In the placenta, 50–90% of maternal cortisol is converted to the biologically inactive cortisone by the enzyme 11 β -hydroxysteroid dehydrogenase (11 β -HSD2). This enzyme protects the fetus for most of the pregnancy. However, under specific conditions (e.g., if the maternal cortisol concentration is very high, if the activity of 11 β -HSD2 is reduced or impaired due to interindividual differences, or when the placenta is immature in early pregnancy or poorly functioning), maternal cortisol is not completely inactivated in the placenta and the fetus is directly exposed to its detrimental actions. Interestingly, in normal pregnancy, a few days before normal labor the 11 β -HSD2 spontaneously decreases to further prepare the fetus for the extra-uterine life.

3. Secretion of CRH to the fetus. A diminished supply of nutrients and oxygen (hypoxemia) for any reason, including infection, inflammation, preeclampsia, eclampsia, starvation, or placental insufficiency, may cause a stress response in the fetus. This involves increased secretion of pCRH that, in turn, contributes to the feed-forward mechanisms at either side of the placenta and accelerates labor and delivery.

Effects of Prenatal and Perinatal Stress

The stress-induced activation of the HPA axis during pregnancy can be expected to have serious negative implications for the mother (pregnancy complications, preterm delivery, problems in lactation, and increased vulnerability to mood disorders and autoimmune phenomena postpartum), the fetus's development (intrauterine growth retardation, immaturity, neurological damage, etc.), and subsequent sequelae in the adult life of the newborn infant.

Maternal Stress: Pregnancy Complications and Preterm Delivery

Recent well-controlled human studies indicated that pregnant women with high stress and anxiety levels are at increased risk for spontaneous abortion, preeclampsia, and preterm labor. The HPA axis hormones (CRH, ACTH, corticosteroids, and DHEAS) are involved in the timing of the onset of birth in humans (rapid increase in free pCRH during the last 2–4 weeks of pregnancy). pCRH initiates, through an increase in DHEAS (the precursor of estrogens), a

cascade of events that may lead to increased uterine activity and eventually labor and delivery. The fetal DHEAS reaches the placenta, where it is converted to estrone (E_1) and estradiol (E_2) and, after 16-hydroxylation in the fetus's liver, also to estriol (E_3). The increase in E_2 and E_3 contributes to the onset of term and preterm delivery (through the stimulation of oxytocin, oxytocin receptor, gap junction, and prostaglandin synthesis).

Maternal stress and fetal stress or hypoxia increase maternal and/or fetal plasma cortisol levels and, because of its positive stimulatory effect, the latter causes an increase in pCRH. Maternal and fetal stresses are also associated with an increase in NE, angiotensin II, and AVP, all compounds that stimulate CRH release in the hypothalamus as well as in the placenta. Fetal distress and preterm delivery may also be driven by amniochorionic and maternal systemic inflammation. Infections and high levels of interleukins may also stimulate CRH secretion and the stress system. Women with preterm delivery, especially those with infections, show an increase in cytokines such as interleukin (IL-)1, tumor necrosis factor (TNF-) α , and IL-6, which is also secreted by the placenta in response to IL-1 and TNF- α . The activation of the cytokine network may lead to increased placental apoptosis and preterm delivery.

Under abnormal circumstances (preeclampsia, eclampsia, threatened preterm delivery, or maternal stress), all these alterations may be initiated prematurely. In mothers who deliver preterm, the hormone increases might be noticeable even in the second trimester, indicating that the trajectory toward preterm delivery may start early in pregnancy and may be recognizable, potentially leading to the development of preventive interventions. This issue is still in need of further investigation. Potent CRH antagonists might be used as labor retardants.

Maternal Stress and Lactation

Lactating women interact more positively with their babies, and nursing mothers (vs. bottle-feeding mothers) are more likely to describe positive mood states and less anxiety. Successful lactation involves the oxytocin-mediated milk ejection reflex and neuroendocrine adaptations of the HPA and gonadal axes. A recent study conducted in breast-feeding women reported decreased plasma ACTH, cortisol, AVP, epinephrine, and glucose responses to treadmill stress tests, suggesting that, under normal conditions, during lactation there is a selective inhibition of the hypothalamic stress response (stress hyporesponsive period). Central actions of oxytocin have been implicated as a cause of the downregulation of HPA

axis reactivity. On the other hand, it has long been known that stress and hormonal products of the HPA axis can influence the release and actions of oxytocin. Milk ejection is inhibited by stressful experiences, suggesting that oxytocin secretion in women could also be inhibited by stress.

Maternal Stress and Increased Vulnerability to Mood Disorders and Autoimmune Phenomena Postpartum

Epidemiological studies indicate that childbearing represents a period of vulnerability to depressive symptomatology in women. Postnatal depression is a significant and serious illness with a reported incidence of approximately 14%, and a further 30% of women experience an adjustment disorder and anxiety following childbirth. The impact of postpartum depression and psychosis on family life is pervasive, and its capacity for causing long-term harm in children is disturbing. The *Diagnostic and Statistical Manual of Mental Disorders* (4th edn.; DSM IV) indicates that postpartum disorders are distinguished not by their phenomena but by their timing. Any major depressive, manic, or mixed episode; bipolar I disorder or bipolar II disorder; or brief psychotic disorder should get a postpartum specifier if it occurs within the postpartum period.

The abrupt CRH withdrawal once the placenta is delivered and its impact on the mother's HPA axis equilibrium may be a major contributing factor to the development of postpartum blues or depression. It is possible that the development of postpartum blues or depression is related to the prolonged or exaggerated suppression of hypothalamic CRH during the postpartum period because the occurrence of postpartum depression has been found to be related to fluctuations in the levels of glucocorticoids. The resetting of the HPA axis and postpartum recovery of the CRH activity remains poorly characterized and varies among women. Moreover, stress affects glucocorticoid levels and can influence estradiol levels in nonpregnant and pregnant females. The sudden withdrawal of estradiol following parturition is also considered an important pathophysiological mechanism of postpartum disorders. In support of this hypothesis, estrogen is an effective treatment for postnatal depression. The newborns of mothers with prepartum and postpartum depressive symptoms have elevated cortisol and norepinephrine levels and lower dopamine levels.

The transient suppression of hypothalamic CRH secretion in the postpartum period might also explain the increased vulnerability to autoimmune phenomena at this period. A defective HPA axis might lead

to relative resistance to infections and neoplastic disease and increased susceptibility to autoimmune and inflammatory disease, such as Hashimoto's thyroiditis, rheumatoid arthritis, and multiple sclerosis.

Prenatal Stress and Its Effects on Fetal Development

Fetal cortisol levels are increased in human fetuses showing intrauterine growth retardation or in pregnancies complicated by preeclampsia, implicating endogenous cortisol in retarded fetal growth. Chronic maternal stress is associated with reduced fetal heart rate variability, which is an indicator of fetal maturation. Most studies assessing the impact of both acute and chronic stress in the fetus have measured changes in heart rate and fetal movements. Experimental studies have demonstrated a relation between specific episodes of maternal psychological stress and exacerbation of fetal asphyxia *in utero*. Experiencing stress during pregnancy leads to the birth of babies with low birth weight, congenital malformations, or growth retardation (reduced head circumference, in particular).

Stress also increases the risk of the infant's developing diseases in subsequent periods of life. Antenatal glucocorticoid administration has been linked with higher blood pressure in adolescence and subtle effects on neurological function, including reduced visual closure and visual memory. Multiple doses of antenatal glucocorticoids given to women at risk of preterm delivery were associated with reduced head circumference and an increased risk of externalizing behavior problems, distractibility and inattention in the offspring.

The Long-Term Programming Hypothesis

The importance of a good quality of psychoaffective communication between mother and child during pregnancy has been shown to be decisive for fetal growth, the perinatal period, and further development of the child. Early-life environmental factors influence prenatal development and may cause structural and functional changes, which persist for the life span. This organizational phenomenon is termed early-life programming or fetal programming of adult disease. This hypothesis proposes that a stimulus or insult acting during critical periods of growth and development may permanently alter tissue structure and function. Genes and environment are no longer considered to exert separate influences. Development is viewed not as a gradual elaboration of an architectural plan preconfigured in the genes but, rather, as a dynamic interdependency of genes and

environment, characterized by a continuous process of interactions in a space- and time-dependent manner. Evidence from both human and animal studies suggests that manipulating the environmental experience of the fetus can induce many diseases of adult life. Programming agents seem to include nutrients, growth factors, cytokines, and hormones, all of which can be altered by stress.

Numerous studies, initially in the United Kingdom and then worldwide, revealed an association between lower birth weight and the subsequent development of the common cardiovascular and metabolic disorders of adult life, notably hypertension, insulin resistance, type 2 diabetes, hyperlipidemia, and atherosclerotic cardiovascular diseases. Thus, early-life events altering birth weight are important predictors of adult morbidity. In addition, postnatal catch-up growth also appears to be predictive of the risk of adult cardiovascular disease, suggesting that it is the restriction of intrauterine growth (caused by overexposure to glucocorticoids) that is pathophysiologically important.

The human HPA axis also appears to be programmed by the early-life environment. Higher plasma and urinary glucocorticoid levels are found in children and adults who were of low birth weight. Two major environmental hypotheses have been proposed to explain the mechanism by which low birth weight is associated with adult disease: fetal nutrition and overexposure of the fetus to glucocorticoids. There is no doubt that in developing countries maternal diet can affect birth weight and may be extremely important in mediating intergenerational effects on birth weight and adult size. Glucocorticoid administration during pregnancy is well documented to both reduce offspring birth weight and alter the maturation of organs (hence, their use to accelerate fetal lung maturation in premature labor). Moreover, antenatal exposure to endogenous or exogenous glucocorticoids produces permanent hypertension, hyperglycemia, hyperinsulinemia, altered behavior, and neuroendocrine responses throughout the life span (in animal experiments).

Fetal undernutrition and fetal glucocorticoid overexposure may be linked by a common mechanism. Processes underlying fetal programming include the determination of the set point of the HPA axis and of tissue glucocorticoid receptor (GR) expression. The molecular mechanisms may reflect permanent changes in the expression of specific transcription factors, key among which is the GR itself. The differential programming of the GR in different tissues reflects effects on one or more of the multiple tissue-specific alternate first exons/promoters of the GR gene. The

proposed mechanisms by which overexposure to glucocorticoids prenatally lead to cardiovascular and metabolic disorders of adult life are as follows:

- **Hypertension and cardiovascular disease:** Prenatal glucocorticoid exposure leads to irreversible reductions in nephron number, fetal and adult vascular responses to vasoconstrictors, enhancing endothelin-induced vasoconstriction and attenuating endothelium-dependent vasorelaxation, indicating microvascular dysfunction. Renin–angiotensin system receptor density and tissue synthesis are also affected. Prenatal glucocorticoid exposure also alters the development of cardiac noradrenergic and sympathetic processes.
- **Liver and pancreas:** Prenatal overexposure to exogenous or endogenous glucocorticoids programs permanent hyperglycemia – particularly hyperinsulinemia – in the adult offspring, effects that are confined to the last third of gestation. Glucocorticoids regulate the expression of critical hepatic metabolic enzymes, notably phosphoenolpyruvate carboxykinase (PEPCK), which catalyses a rate-limiting step in gluconeogenesis. Prenatal undernutrition impairs pancreatic β -cell development, reducing β -cell mass and causing glucose intolerance. Fetal pancreatic insulin content correlates inversely with fetal corticosterone levels.
- **Adipose tissue:** Antenatal corticosteroid exposure seems to program fat metabolism, causing a marked increase in GR expression selectively in visceral adipose tissue, which contributes to both adipose and hepatic insulin resistance. Leptin concentrations in human fetal cord blood correlate directly with body weight and adiposity at birth.

Altered Regulation of Hypothalamic-Pituitary-Adrenal Axis after Prenatal Stress and Its Influence on the Fetal Central Nervous System

There is abundant evidence of an overactive and dysregulated HPA axis in the offspring of prenatally stressed animals. Prenatally stressed animals have higher basal blood glucocorticoid levels, and when exposed to a stressful stimulus, their hormonal response is more rapid, stronger, and more prolonged, indicating that the normal negative feedback mechanism of the HPA axis is heavily impaired. In addition, alterations have been found in almost all known regulatory and neurotransmitter systems in the brain, including the opioid, cholinergic, serotonergic, dopaminergic, GABA-ergic, and noradrenergic systems. These alternations could underlie some of the behavioral and physiological effects observed after prenatal stress and may further affect later stress responses.

Three mechanisms can explain the overactivity and altered feedback regulation of the HPA axis (increased levels of HPA axis hormones affect fetal central nervous system development and, especially, brain GR development): (1) the downregulation of receptors in the hippocampus, hypothalamus, pituitary, or adrenal glands; (2) the decreased sensitivity of the receptors at any of these levels; and (3) the alteration in CRH levels or the altered levels or affinity of plasma-binding proteins. The maternal HPA axis is involved in early neuroendocrine programming of the offspring.

The hippocampus appears to be particularly vulnerable to insult during early development because its pyramidal neurons contain a high concentration of GRs that are highly sensitive to either hypercortisolemia caused by severe stress or to exposure to exogenous glucocorticoids. Another neural structure that may contribute to increased responsiveness of the HPA axis following prenatal stress is the amygdala, a key locus for its effects on fear and anxiety. The content of CRH appeared to be substantially increased in the amygdala of prenatally stressed rats compared to levels in control animals. These increased levels of CRH are associated with increased emotionality, fear, and anxiety-like behaviors and facilitate stress-induced behavior and autonomic activation.

Moreover, prenatal glucocorticoid exposure contributes to schizoaffective, attention deficit hyperactivity, and extrapyramidal disorders. Stressful events in the second trimester of human pregnancy have been associated with increased incidence of offspring schizophrenia.

See Also the Following Articles

11 β -Hydroxysteroid Dehydrogenases; Corticotropin Releasing Factor (CRF); Hypertension; Metabolic Syndrome.

Further Reading

- Austin, M. P., Leader, L. R. and Reilly, N. (2005). Prenatal stress, the hypothalamic-pituitary-adrenal axis and fetal and infant neurobehaviour. *Early Human Development* **81**, 917–926.
- Buitelaar, J. K., Huizink, A. C., Mulder, E. J., et al. (2003). Prenatal stress and cognitive development and temperament in infants. *Neurobiology of Aging* **24**, S53–S60.
- Carter, C. S., Altemus, M. and Chrousos, G. P. (2001). Neuroendocrine and emotional changes in the postpartum period. *Progress in Brain Research* **133**, 241–249.
- Chrousos, G. P. (1999). Reproductive placental corticotropin-releasing hormone and its clinical implications.

- American Journal of Obstetrics and Gynecology* **180**, S249–S250.
- Chrousos, G. P. and Gold, P. W. (1992). The concepts of stress and stress system disorders: overview of physical and behavioral homeostasis. *Journal of the American Medical Association* **267**, 1244–1252.
- Chrousos, G. P., Torpy, D. and Gold, P. W. (1998). Interactions between the hypothalamic-pituitary-adrenal axis and the female reproductive system: clinical implications. *Annals of Internal Medicine* **129**, 229–240.
- de Weerth, C. and Buitelaar, J. K. (2005). Physiological stress reactivity in human pregnancy – a review. *Neuroscience and Biobehavioral Reviews* **29**, 295–312.
- Drake, A. J. and Walker, B. R. (2004). The intergenerational effects of fetal programming: non-genomic mechanisms for the inheritance of low birth weight and the cardiovascular risk. *Journal of Endocrinology* **180**, 1–16.
- Fowden, A. L., Giussani, D. A. and Forhead, A. J. (2005). Endocrine and metabolic programming during intrauterine development. *Early Human Development* **81**, 723–734.
- Halbreich, U. (2005). The association between pregnancy processes, preterm delivery, low birth weight, and postpartum depressions – the need for interdisciplinary integration. *American Journal of Obstetrics and Gynecology* **193**, 1312–1322.
- Kaiser, S. and Sachser, N. (2005). The effects of prenatal social stress on behaviour: mechanisms and function. *Neuroscience and Biobehavioral Reviews* **29**, 283–294.
- Magiakou, M. A. and Chrousos, G. P. (2005). Biological basis of stress-related diseases. In: Antoniou, A. S. G. & Cooper, C. L. (eds.) *Research companion to organizational health psychology*, pp. 70–86. Cheltenham, UK: Edward Elgar.
- Magiakou, M. A., Mastorakos, G., Rabin, D., et al. (1996). Hypothalamic corticotropin releasing hormone suppression during the postpartum period: implications for the increase of psychiatric manifestations during this time. *Journal of Clinical Endocrinology and Metabolism* **81**, 1912–1917.
- Magiakou, M. A., Mastorakos, G., Rabin, D., et al. (1996). The maternal hypothalamic pituitary-adrenal axis in third trimester human pregnancy. *Clinical Endocrinology* **44**, 419–428.
- Majzoub, J. A. and Karalis, K. P. (1999). Placental corticotropin-releasing hormone: function and regulation. *American Journal of Obstetrics and Gynecology* **180**, S242–S246.
- Mulder, E. J. H., Robles de Medina, P. G., Huizink, A. C., et al. (2002). Prenatal maternal stress: effects on pregnancy and the (unborn) child. *Early Human Development* **70**, 3–14.
- O'Regan, D., Welberg, L. L. A. M., Holmes, M. C., et al. (2001). Glucocorticoid programming of pituitary-adrenal function: mechanisms and physiological consequences. *Seminars in Neonatology* **6**, 319–329.
- Seckl, J. R. (2001). Glucocorticoid programming of the fetus; adult phenotypes and molecular mechanisms. *Molecular and Cellular Endocrinology* **185**, 61–71.

- Seckl, J. R. (2004). Prenatal glucocorticoids and long-term programming. *European Journal of Endocrinology* 151, U49–U62.
- Van den Bergh, B. R. H., Mulder, E. J. H., Mennes, M., et al. (2005). Antenatal maternal anxiety and stress and the neurobehavioural development of the fetus and child: links and possible mechanisms: a review. *Neuroscience and Biobehavioural Reviews* 29, 237–258.
- Weinstock, M. (2005). The potential influence of maternal stress hormones on development and mental health of the offspring. *Brain, Behavior and Immunity* 19, 296–308.
- Welberg, L. A. M. and Seckl, J. R. (2001). Prenatal stress, glucocorticoids and the programming of the brain. *Journal of Neuroendocrinology* 13, 113–128.

Premenstrual Dysphoric Disorder

S Nowakowski

San Diego State University/University of California,
San Diego, CA, USA

P Haynes

University of Arizona, Tucson, AZ, USA

B L Parry

University of California, San Diego, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
P Haynes and B L Parry, volume 3, pp 201–207, © 2000,
Elsevier Inc.

Diagnostic Issues

Etiology

Glossary

<i>Circadian oscillator</i>	A pacemaker located in the superchiasmatic nucleus of the hypothalamus with an output that repeats about once per day.
<i>Dysphoric mood</i>	A symptom of depression; feeling sad, despondent, discouraged, unhappy, anxious, or irritable.
<i>Entrainment</i>	The process by which the circadian oscillator becomes synchronized with the external environment.
<i>Follicular phase</i>	Also called the proliferative phase or estrogenic phase, the follicular phase of the menstrual cycle lasts from the first day of menstruation until ovulation. As assessed by the secretion of luteinizing hormone (LH), gonadotropin-releasing hormone (GnRH) is secreted from the hypothalamus in a pulsatile manner at a frequency of once per 90 min during the early follicular phase and increasing to once per 60–70 min. Follicle-stimulating hormone (FSH) is secreted by the anterior pituitary, and its concentrations in plasma increase from the first day of the cycle.

Luteal phase

The rise in FSH levels can be attributed to a decrease in progesterone and estrogen levels at the end of the previous cycle and the subsequent removal of inhibition of FSH by these ovarian hormones. FSH stimulates the development of 15–20 follicles each month and stimulates ovarian follicular secretion of estradiol. Estrogen levels peak toward the end of the follicular phase of the menstrual cycle, at which point estrogen exerts positive feedback on the hypothalamic-pituitary-gonadotrope system, thereby triggering the preovulatory LH surge.

Approximately the second half of the menstrual cycle occurring after ovulation during which there is a decrease in frequency and increase in amplitude of time, ovarian steroid, and pulsatile release LH (and presumably GnRH) secretion. The luteal cells of the corpus luteum produce progesterone, the plasma levels of which peak by the middle of the luteal phase (days 20–22). By day 28, the corpus luteum has regressed with a consequent decrease in the output of estrogen and progesterone. The latter causes local ischemia and subsequent necrosis of the endometrium (i.e., menstrual flow).

Zeitgeber

German for time giver; a regular, external cue that synchronizes the internal clock to environmental cues; light is the most powerful zeitgebers in humans.

Premenstrual dysphoric disorder (PMDD) is a clearly defined syndrome that is often referred to as premenstrual syndrome (PMS) among the nonmedical population. As delineated in the *Diagnostic and Statistical Manual*, 4th edition (DSM-IV-TR), five symptoms are required to make the diagnosis of

Table 1 Diagnostic criteria for premenstrual dysphoric disorder

A. In most menstrual cycles during the past year, at least five of the following symptoms must have been present most of the time during the last week of the luteal phase and disappear within a week after the onset of menstruation. At least one of the symptoms must be one of the first four.

1. Feeling sad, hopeless, or self-deprecating
2. Feeling tense, anxious or on edge
3. Marked lability of mood interspersed with frequent tearfulness or increased sensitivity to rejection
4. Persistent irritability, anger, and increased interpersonal conflicts
5. Decreased interest in usual activities
6. Difficulty concentrating
7. Feeling fatigued, lethargic, or lacking in energy
8. Marked changes in appetite, which may be associated with binge eating or craving certain foods
9. Hypersomnia or insomnia
10. A subjective feeling of being overwhelmed or out of control

B. The symptoms must cause an obvious and marked impairment in the ability to function socially or occupationally.

C. The symptoms may be superimposed on another disorder but are not merely an exacerbation of the symptoms of another disorder, such as major depression, panic disorder, dysthymia, or a personality disorder.

D. Criteria A, B, C, and D must be confirmed by prospective daily self-ratings during at least two consecutive symptomatic cycles. (The diagnosis may be made provisionally prior to this confirmation.)

From the American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition, text revision. Washington, D.C.: American Psychiatric Association, with permission.

PMDD, and at least one symptom must be a moderate to severe mood symptom (see [Table 1](#)). The criteria for PMDD are more stringent and delineate increased specificity and severity of premenstrual symptoms with an emphasis on affective symptoms (see [Table 1](#)). The criteria for PMDD are stringent and delineate increased specificity and severity of premenstrual symptoms with an emphasis on affective symptoms (see [Table 1](#)). Symptoms typically begin in the late luteal (premenstrual) phase of the menstrual cycle, which is associated with declining estrogen and progesterone levels from the degenerating corpus luteum. Premenstrual symptoms typically remit within the first few days of the follicular phase, when estrogen levels increase as a result of follicular growth. The diagnosis of PMDD requires documentation of luteal phase timing of symptoms and a symptom-free interval from approximately day 4 of menses to the onset of ovulation. The National Institute of Mental Health Premenstrual Syndrome Workshop (1983) specified that a 30% change in symptom severity during the luteal phase is required for a PMDD diagnosis. In addition, a diagnostic requirement is that symptoms must cause functional impairment.

Diagnostic Issues

R. T. Frank first identified physical, psychological, and behavioral changes corresponding to monthly changes in reproductive hormones in 1931. However, public awareness of the concurrent change in mood with phases of the menstrual cycle has increased

dramatically over the past two decades. The media has disseminated a wide body of information shaping a societal notion of what is popularly termed premenstrual syndrome (PMS). Definitions for PMS have varied widely. Typically, PMS refers to a large number of symptoms, including irritability, tension, fatigue, dysphoria, distractibility, impaired motor coordination, changes in eating and sleeping, and libido changes, which occur in the late luteal phase and remit after the beginning of menstruation. Up to 90% of women experience some premenstrual symptoms, and 50% of women may experience many symptoms together, thus comprising a syndrome.

Many women experience mood disruptions or physical complaints during the late luteal phase. For this reason, the PMDD diagnosis has been controversial among psychiatrists, psychologists, gynecologists, and sociologists. One reason for the ambiguity may be the failure to distinguish normal from pathological premenstrual mood disturbance. Unlike other psychiatric syndromes, the remission and relapse of premenstrual symptoms is connected to a physiological process, the menstrual cycle. Every woman does not experience a pathological premenstrual mood disorder, however. Furthermore, there are discrepant views in defining depression as a categorical versus spectrum illness. Depression can be defined as a discrete, all-or-none variable or on a continuum of symptom severity. Thus, the differentiation between the normal and pathological is important in terms of investigative research, interdisciplinary discourse, as well as treatment.

Premenstrual criteria were included in the appendix of DSM-III-R as late luteal phase dysphoric disorder (LLPDD), and they are currently included as PMDD in the appendix of the DSM-IV-TR as a depressive disorder, not otherwise specified. Some researchers have postulated that moving PMDD into the depressive disorders not otherwise specified section of the DSM-IV has reinforced the view that PMDD is depression, rather than a unique disorder. Moreover, the treatment of PMDD with selective serotonin reuptake inhibitors (SSRIs) may be treating symptoms of depression rather than the etiological disturbance; thus, treatment may provide little information on the nature of the disorder and raises the problem of comorbidity with other mood disturbances (e.g., major depressive disorder).

Besides emotional distress, the disturbance must cause impairment in occupational or social functioning. Also, the mood disturbances must begin during the late luteal phase and remit at the follicular phase; this trend of mood fluctuation must occur over at least two consecutive cycles. Because retrospective ratings may not be accurate due to an expectation bias, premenstrual trends must be documented by daily mood ratings. In the United States, approximately 3–5% of women are thought to experience symptoms that meet criteria for PMDD. Even with the DSM-IV-established criteria, debate over defining criteria of PMDD is ongoing. The requirement of absence of symptoms postmenstrually may be unreasonably conservative, and the criterion for premenstrual interference in daily activities contributes little to the diagnosis of PMDD. Poor classification proves to be particularly problematic because, over time, untreated PMDD becomes progressively more severe, and episodes of dysphoria extend in duration. For this reason, women diagnosed with PMDD are vulnerable to developing major depressive disorder at a later date.

Etiology

Various disciplines conceptualize the etiological basis of premenstrual mood disturbance. Biological theorists maintain that premenstrual symptoms originate from the biological reactivity or sensitivity to reproductive hormones. Psychological theorists suggest that premenstrual symptoms arise from the maladaptive response of the individual to the environment. Social theorists assert that the social, patriarchal environment elicits symptoms. Regardless of the etiological model, hormonal changes associated with the menstrual cycle are likely to serve as contributory factors to mood disruption.

Biomedical Model

Investigators and clinicians have recognized the spectrum of premenstrual mood: symptoms ranging from mild to severe. Those with severe symptoms have a decreased ability to function effectively and meet criteria for PMDD. Rather than psychosomatic origins, researchers have investigated physiological underpinnings to the disorder that are expressed in mood and behavior, that is, conceptualized as a somato-psychic illness.

Estrogen and progesterone Research in this area began with the hypothesis that the absolute amount of estrogen (E2) and progesterone (P4), e.g., P4 deficiency, E2 excess, were responsible for mood disturbances. Inconsistencies in research findings have called this hypothesis into question. The rate of hormonal change, however, may play an important role in the pathogenesis of the disorder. Schmidt and colleagues examined the role of E2 and P4 by testing the effects of ovarian suppression with leuprolide, an agonist analog of gonadotropin-releasing hormone, versus placebo in women with and without PMS. The authors found that leuprolide caused a significant decrease in symptom scores for women with PMS compared with placebo treatment. In a subsequent phase of the study, Schmidt and colleagues administered a conjunct hormone therapy (leuprolide plus estradiol or progesterone) to the women with PMS whose symptoms responded to ovarian suppression, which resulted in a significant recurrence of symptoms. Thus, baseline levels of gonadal steroids or changes in hormone levels triggered a change in mood in women sensitive to mood instability. This study suggests that the occurrence of PMS symptoms may represent an abnormal response to normal hormonal changes. In addition, it suggests that E2 and P4 variations likely exert their influence on mood indirectly rather than directly. The central nervous system (CNS) is thus a nontraditional target for sex steroids.

Other neuroendocrine variables Research also suggests that abnormal fluctuations in thyroid, cortisol, prolactin, mineralocorticoids, prostaglandins (PGs), and endogenous opiates correspond to mood changes during the menstrual cycle. In addition, women with PMDD have a brief, transient, heightened cortisol response to ovine (o) corticotropin-releasing hormone (CRH) stimulation, blunted cortisol responses to serotonergic agonists in the late luteal phase, and lower evening plasma cortisol levels. Alterations in phase but not amplitude measures of cortisol have been found in women with PMDD during the

menstrual cycle. These findings suggest a physiological dysfunction between the hypothalamic-pituitary-adrenocortical (HPA) and the hypothalamic-pituitary-ovarian (HPO) axis for women with PMDD. This relationship may be complex and include a variety of other systems in its interaction.

Neurotransmitters Current research suggests that neurotransmitter abnormalities in PMDD are similar to those implicated in major depressive disorder. Specifically, PMDD patients have been differentiated from normal control patients by deficiencies in the serotonergic system. Investigators have found that women with PMDD have reduced platelet uptake of serotonin 1 week before menstruation, low whole blood serotonin during the last 10 days of the menstrual cycle, and an abnormal response to tryptophan loading in the late luteal phase. SSRIs (e.g., fluoxetine, sertraline, paroxetine, fluvoxamine) have been found to be efficacious in the treatment of PMDD. The efficacy of this pharmacological class is further evidenced by the FDA approval of fluoxetine and sertraline for PMDD. Recent work has shown that fluoxetine, citalopram, clomipramine, and sertraline can be effective if administered intermittently during the luteal phase. Further work is necessary to corroborate these findings. This medication regimen may be preferable to some women who do not wish to take chronic medication for a periodic condition. Women on a luteal dosing schedule generally reach response rates of 65%, with some women experiencing no significant benefits. This response rate may be reduced by nonadherence with a luteal phase dosing regimen.

In addition to serotonin, norepinephrine has been examined as another neurotransmitter that might contribute to PMDD. However, trials comparing fluoxetine to bupropion, sertraline to desipramine, and paroxetine to maprotiline demonstrate that augmenting noradrenergic activity alone is not effective for the treatment of PMDD.

Circadian rhythms Besides antidepressant medication, sleep and light therapies also have been found to be beneficial in reducing symptoms. The efficacy of these treatments is based on a chronobiological hypothesis of depression. Chronobiological investigations indicate that people with major depressive disorder may have disruptions in the internal regulation of rapid eye movement (REM) sleep propensity, temperature, cortisol, and melatonin, which occur throughout the 24-h daily cycle. Furthermore, the oscillator regulating these circadian rhythms may be desynchronized with the sleep-wake cycle. External

cues, such as bright light, play an important role in the modulation and synchronization of these circadian rhythms. For example, exposure to 500 lux of light acutely suppresses melatonin secretion, a marker for circadian phase position. Light also shifts circadian rhythms and melatonin. Research indicates that patients with affective illness may be supersensitive to the suppressive effects of light on melatonin secretion, possibly due to inadequate exposure to daytime light; they may have altered amplitude of melatonin secretion and body temperature.

There may be differences between men and women in the regulation or expression of circadian rhythms. Compared with men, menstruating women have a shorter free-running period (i.e., length of time of a rhythm in an environment free of external cues, such as light), longer sleep duration, and lower amplitude in body temperature that varies with phases of the menstrual cycle. These findings suggest that reproductive hormones (including testosterone) affect circadian rhythm amplitude and synchronization. Indeed, comparative studies in rodents show that estrogen advances and progesterone delays circadian rhythms. Estrogen also appears to enhance synchronization between the oscillators.

Thus, fluctuations of menstrual cycle hormones may alter circadian rhythmicity. Instability of circadian rhythms resulting from changing reproductive hormonal levels may increase the risk for depression in some women. Chronobiological hypotheses contend that disrupted circadian rhythmicity is a factor contributing to depression. Women with PMDD have been reported to have decreased melatonin amplitudes, higher nocturnal body temperatures, and disturbances in the timing and amplitude of prolactin and thyroid-stimulating hormone (TSH) secretion. Also, the offset and duration of the melatonin rhythm may be disturbed and show an abnormal response to light. Light and sleep therapies may relieve PMDD symptoms by helping to realign underlying circadian clocks.

Psychological Model

In this model, women experiencing pathological premenstrual symptoms are distinguished from healthy women on the basis of the way they perceive or react to the environment. Therefore, treatment of PMDD is often directed toward physical activity, psychotherapy, relaxation training, support groups, and stress management.

One psychological model has studied PMDD from a state-dependent learning perspective, while biological markers tend to be trait dependent. Compared

with healthy subjects, women with menstrually related mood disorders report more negative life events and greater distress associated with those events only in the late luteal phase of the cycle. Thus, symptoms are conceptualized as occurring in an experiential state, such as the late luteal phase, which is characterized by affective and cognitive components. In support of the state-dependent model, women with severe premenstrual symptoms were more likely to use emotionally based coping techniques (i.e., catharsis) in the premenstrual phase than other strategies, such as cognitive reappraisal. Also during this phase, women with premenstrual dysphoria may be more likely than healthy controls to seek social support when perceiving daily stressors as uncontrollable.

Other findings fail to support the state-dependent hypothesis. Woods and colleagues suggested that a general stressful life context predicts premenstrual symptoms more than stress experienced during a particular menstrual cycle phase. A stressful milieu was a large predictor of perimenstrual symptoms and especially negative affect. Major life stressors play an important role in the development of depressive symptoms; however, the cumulative effect of daily stressors may be more influential in the generation of global premenstrual symptoms.

Type of stressor may also affect the manifestation of symptoms. For example, some research indicates that women with premenstrual dysphoria experience negative interpersonal interactions. Kuczmierczyk, Labrum, and Johnson found that women with PMDD perceive their work and especially their family lives as more stressful. Increased conflict and less self-sufficiency and assertiveness were noted in family environments of PMDD patients, which suggest that enmeshed and rigid family structures may have a role in the development of PMDD. These types of family structures have been previously implicated in other somatization processes.

In addition to disturbances in mood, women with PMDD may experience physical and cognitive changes. Cognitive deficits in major depression have been well documented and often include disturbances in attention, forgetfulness, psychomotor retardation, and proneness to confusion. In some cases, these cognitive disturbances may be secondary to mood disturbances. However, in the case of PMDD, the role reproductive hormone fluctuation may have on memory and cognitive functioning recently has been the focus of several studies. In healthy women without psychiatric illness, performance on neuropsychological tests has been found to vary with the monthly fluctuation in reproductive hormones. High levels of E2, as experienced in the late follicular phase, may

elevate performance on automatization abilities, perceptual motor speed, mental arithmetic, and verbal memory. Lower levels of E2, as experienced during the menstrual phase, may be associated with higher performance on visuospatial tasks. The majority of these effects have been subtle and were not able to be replicated; this effect may be due to difficulties in documentation and measurement of ovulation, plasma hormone levels, mood symptoms, and individual variation in cycle length. Furthermore, cyclic performance changes are minor and are unlikely to interfere with normal functioning.

Social Model

From the earliest times, various facets of women's personalities, capabilities, and moods have been attributed to menstruation. Instability, resulting from women's reproductive cycles, has been used to justify denying women equal access to education and employment. One feminist model maintains that premenstrual syndrome pathologizes a naturally occurring phenomenon. Theorists state that high prevalence rates of premenstrual symptoms support this assumption and that the distinction of a separate pathological disorder is a spurious dichotomy. Based on historical comparisons of PMS and hysteria, writers have suggested that the use of a biological explanation for this disorder may legitimize societal prejudices. Furthermore, social theorists state that the medical model may pathologize normal biological processes and oversimplify women's feelings, problems, and conflicts through the diagnosis of PMDD. In fact, medical investigations have revealed that increases in cyclicity can be viewed as adaptive because it leads to homeostasis, stabilization, and longevity in life. Thus, it is the diagnosis of PMDD that leads to the medicalization of women's experiences of the menstrual cycle and not the cycle-related changes themselves. It is suggested that PMDD is a socially constructed disorder rather than a psychiatric disorder.

Since Western society is a contributory factor in this model, research on the prevalence of PMDD in other cultures is of importance in understanding the universality of symptoms. A limited number of studies have examined symptoms of PMDD in non-Western societies. This research has been plagued by a number of important methodological concerns and questionable validity of instruments that have been translated for use in populations other than those in which they were developed. In addition, the sociopolitical position of women in other cultures and their societal beliefs and expectations about menstruation are difficult to quantify and thus have not been taken

into account. Despite these difficulties, cross-cultural studies are of interest in considering the relative importance of biological versus sociological factors in the etiology of PMDD. Investigative studies have indicated a widespread variance in the type and severity of premenstrual symptoms experienced, which supports the interpretation of normal cyclical change. Eriksen found that women in developing countries tended to report shifts in mood associated with the onset of menstruation or throughout the menstrual cycle. Several studies have found reports of higher rates of somatic than affective symptoms in populations outside the United States, suggesting that culture may influence the symptoms that women notice, as well as the symptoms they determine to be problematic during the premenstrual period. As a result of these observed cross-cultural differences, menstrual socialization has been proposed as a determinant of premenstrual complaints. These findings suggest that PMDD is a culture-bound syndrome specific to Western cultures in which most women have been socialized to have negative expectations about menstruation. More specifically, it is argued that North American culture and media perpetuate the idea that the premenstrual period will be associated with negative affect and mood instability, causing women to interpret normal physiological changes that are essentially neutral in nature to have negative connotations.

Emergence of a Biopsychosocial Model

Since the relatively recent invention of radioimmunoassays, biomedical research has clearly dominated the area of PMDD. Because PMDD symptoms are associated with hormonal fluctuations, cultural and psychological phenomena associated with the disorder have often been ignored in scientific research, which has resulted in a poor integration of the findings. In the field of depression and other mood disorders, multidimensional theoretical models have emerged and contributed greatly to understanding the experience of depression for an individual. In addition, exploring the interaction of biological, psychological, and social processes has expanded the range of treatment options and provided individuals who suffer from depression effective alternatives to medication. For similar reasons, an integrative approach could greatly benefit research and clinical work in the area of premenstrual mood disturbances. A two-phase approach to encompass a comprehensive biopsychosocial perspective is recommended, with the first phase emphasizing self-monitoring through charting, stress management, and dietary and

exercise programs. In addition, cognitive-behavioral coping strategies and reattribution training may help alleviate symptoms. The second phase consists of a pharmacological approach to mitigate biological aspects contributing to PMDD.

In the area of affective illness, a social zeitgeber theory of mood disturbance has been one unifying hypothesis linking biological and psychosocial models. Social zeitgebers are personal relationships, social demands, or tasks that entrain biological rhythms. Similar to the effect of light, social events are external regulators that synchronize circadian rhythms. For instance, marriage, birth of a child, divorce, and loss of a job are social events that disrupt natural meal-times, sleeping times, and times of activity, which affect the stability of biological rhythms. In individuals predisposed to affective disruption, disturbances in the biological clock may develop into a state of ongoing desynchronization as observed in major depression. Social rhythm disruptions also have been implicated in the onset of manic episodes. Ehlers and colleagues suggested that both interpersonal psychotherapy and cognitive-behavioral therapy, two empirically supported treatments for depression, address the regularity of social routine. In addition, they state that the model is likely to be influenced by personality factors, gender, social support, coping, genetic/familial loading, and past treatment experience.

Adoption of the social zeitgeber model for premenstrual mood disturbances may potentially elucidate the interaction of important psychosocial and biological etiological variables. Socially, women may be particularly prone to respond more dramatically to interpersonal conflicts because they may rely on social zeitgebers, or social relationships, more than men. Some feminist models may explain this perspective. For instance, Carol Gilligan proposed that women may view relationships in terms of interconnections and caring, and men may view relationships in terms of hierarchy and power. Developmental research has documented that girls may be more likely to engage in prosocial behavior than boys. On self-report measures, women are often more empathetic and nurturing than men. The emphasis that women may place on connection with others possibly serves as a vulnerability to a mood disturbance; interpersonal disruptions may have a greater impact on women than on men. Animal and human studies have also shown that responses to social interactions can lead to fluctuations in whole blood serotonin levels. Therefore, women may be able to regulate serotonin function by seeking appropriately rewarding social interactions; the inability to seek or receive the social interactions could lead

to further physiological dysregulation via circadian rhythm disruption and/or reduction of serotonin, manifesting as symptoms of PMDD.

At the same time, this increased need for social support may be more adaptive in the long term; women may cope with stress, such as job loss or a death, better than men. Therefore, the increased variation in mood may be more adaptive by enhancing the mechanism of homeostasis, thus protecting women against other types of long-standing illness and increasing longevity.

These correlations between behavior and neuroendocrinology have also been explored within a framework of functional hypothalamic amenorrhea. For some time, researchers have noted that psychosocial variables, such as exercise, personality traits, and environmental stress, have the potential to induce ovarian acyclicity by inhibiting the release of gonadotropin-releasing hormone (GnRH) through mediating variables, such as CRH, endogenous opioids, dopamine, and TSH. Further work in this area potentially could elucidate important somato-psychic interactions, which are key to understanding the etiology of PMDD.

The biological, psychological, and social approaches to PMDD taken independently do not take into account the more complex relationships that formulate the disorder. Practically, it may be useful to examine PMDD through a single pathway that focuses on main effects and a linear relationship. However, lack of attention to the cluster of influences and their dynamic relations may oversimplify pathways as to how individuals develop severe premenstrual disturbances. Increasingly, the biopsychosocial model of PMDD is being examined to generate a more comprehensive model. Although great strides have been made in the past decade in increasing our understanding of PMDD, we must continue to supplement current approaches with research that focuses on more complex interactions among various factors. This future direction can only be accomplished through the continued interdisciplinary effort on a multivariate, biopsychosocial approach to PMDD.

See Also the Following Articles

Anxiety; Depression and Manic-Depressive Illness; Menstrual Cycles and Stress; Selective Serotonin Reuptake Inhibitors (SSRIs).

Further Reading

- Blak, F., Salkovskis, P., Gath, D., et al. (1998). Cognitive therapy for premenstrual syndrome: a controlled study. *Journal of Psychosomatic Research* 45(4), 307–318.
- Dimmock, P. W., Wyatt, K. M., Jones, P. W., et al. (2000). Efficacy of selective serotonin-reuptake inhibitors in premenstrual syndrome: a systematic review. *Lancet (North American Edition)* 356, 1131–1136.
- Endicott, J. (2000). History, evaluation and diagnosis of premenstrual dysphoric disorder. *Journal of Clinical Psychiatry* 61(12), 5–8.
- Eriksson, E., Andersch, B., Ho, H. P., et al. (2002). Diagnosis and treatment of premenstrual dysphoria. *Journal of Clinical Psychiatry* 63(7), 16–23.
- Halbreich, U., Bergeron, R., Yonkers, K. A., et al. (2002). Efficacy of intermittent, luteal phase sertraline treatment of premenstrual dysphoric disorder. *Obstetrics & Gynecology* 100(6), 1219–1229.
- Parry, B. L. (2001). The role of central serotonergic dysfunction in the aetiology of premenstrual dysphoric disorder: therapeutic implications. *CNS Drugs* 15(4), 277–285.
- Parry, B. L. and Berga, S. L. (2002). Premenstrual dysphoric disorder. In: Arnold, A. P., Etgen, A. M., Fehrbach, S. E. & Rubin, R. T. (eds.) *Hormones, brain, and behavior* (vol. 5, chap. 99) pp. 531–552. New York: Academic Press/Elsevier Science.
- Parry, B. L., Berga, S. L., Mostofi, N., et al. (1989). Morning vs. evening light treatment of late luteal phase dysphoric disorder. *American Journal of Psychiatry* 146, 1215–1217.
- Ross, L. E. and Steiner, M. (2003). A biopsychosocial approach to premenstrual dysphoric disorder. *Psychiatric Clinics of North America* 26, 259–546.
- Schmidt, P. J., Nieman, L. K., Danaceau, M. A., et al. (1998). Differential behavioral effects of gonadal steroids in women with and in those without premenstrual syndrome. *The New England Journal of Medicine* 338, 209–216.

Pre-pulse Inhibition

M van den Buuse

Mental Health Research Institute, Melbourne, Australia

© 2007 Elsevier Inc. All rights reserved.

What Is Prepulse Inhibition?

Clinical Correlates

Effect of Stress on Prepulse Inhibition in Rats and Mice

Mechanism of Stress Effect on Prepulse Inhibition

Conclusion

Glossary

Prepulse inhibition (PPI)

A measure of sensorimotor gating, usually the reduction of the startle response to an intense stimulus if this is preceded by a low-intensity prestimulus.

Sensorimotor gating

A normal protective mechanism in the brain that functions to gate or filter irrelevant sensory stimulation, allowing for coherent thought.

What Is Prepulse Inhibition?

PPI is a measure of sensorimotor gating and is widely used to study information processing and attention dysfunctions. Sensorimotor gating is a normal protective mechanism in the brain that functions to gate or filter irrelevant sensory stimulation, allowing for coherent thought. Deficits in sensorimotor gating may result in stimulus overload and the misinterpretation of sensory information.

PPI of the acoustic startle response can be easily measured in several species, including rats, mice, and humans. The startle response is elicited by the presentation of short, loud acoustic stimuli (pulses or clicks) that usually induce whole-body startle responses in rats and mice and reflex eyeblink responses in humans. When the acoustic startle stimulus is preceded by a lower-intensity prestimulus (a prepulse), the subsequent behavioral response is reduced (see [Figure 1](#)). This mechanism is mediated by a well-defined neural circuit, including a short brain-stem pathway mediating the startle response and several modulating brain regions, such as the hippocampus, nucleus accumbens, and frontal cortex. PPI is deficient in a number of mental illnesses, including schizophrenia, obsessive-compulsive disorder, and attention-deficit disorder. In experimental animals, deficits in PPI can be induced by treatment with dopamine

receptor agonists, serotonin 1A and 2A receptor agonists, or NMDA receptor antagonists. Several studies have shown that deficits in humans or treated animals can be reversed by treatment with antipsychotic drugs, such as clozapine. Because of the importance of stress as a risk factor in mental illnesses, studies have addressed the effect of stress on PPI.

Clinical Correlates

A clear example of a link between stress and PPI comes from observations in patients with posttraumatic stress disorder (PTSD). There was a significant reduction of PPI when 21 Vietnam veterans with PTSD were compared to a civilian control group of 17 subjects. However, when the 21-subject group was compared to an 11-subject Vietnam veteran control group without PTSD, the difference reached only trend level. The magnitude of PPI disruption was correlated with the extent of PTSD symptoms (Mississippi scale), specifically the extent of intrusive memories and the level of arousal. Subsequent studies in other patient groups by the same authors have only partially replicated these findings. Similarly, whereas in a groups of 8- to 13-year-old children with PTSD PPI was clearly diminished, in another study of a group of 28 adolescent girls with PTSD, there were no significant differences in PPI compared to a group of 23 healthy nontraumatized girls. These discrepancies between studies could be related to the differences in the causes of PTSD in these subjects and also to parametric and methodological differences between the studies.

Many PPI studies have focused on schizophrenia. The majority of these studies have found significant disruptions of PPI in these patients, which was reversed by treatment with antipsychotic drugs. However, although stress is recognized as a risk factor in this illness, few studies have investigated stress modulation of PPI in patients with schizophrenia (see [Schizophrenia](#)).

Effect of Stress on Prepulse Inhibition in Rats and Mice

In contrast to the paucity of clinical studies on stress and PPI, several experimental animal studies have addressed this relationship. These studies can be divided according to the timing of the stress: perinatal, chronic postweaning, or acute stress.

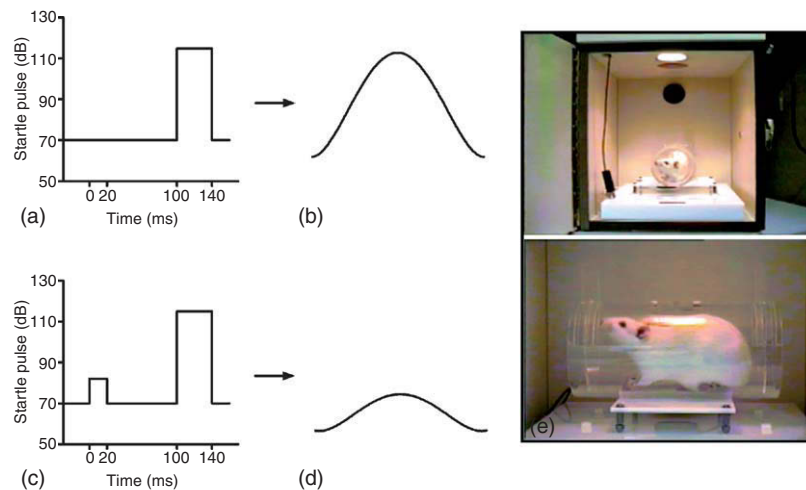


Figure 1 Prepulse inhibition of acoustic startle in rats. a, 40-ms 115-dB startle pulse; b, response in pulse-alone trials; c, startle pulse preceded 100 ms before by a 20-ms prepulse of 12 dB above baseline; d, reduced startle responses in a prepulse–pulse trial; e, rats in the automated startle apparatus used for PPI experiments. If PPI is disrupted because of pharmacological treatment or disease, the presentation of the prepulse–pulse trial evokes responses closer to or similar to those after pulse-alone trials.

Perinatal Stress and Prepulse Inhibition

The neurodevelopmental origin of schizophrenia has led to several studies on the effect of early stress on PPI. The rationale for these studies is that an early stressful event elicits subtle changes in brain development that result in changes in behavior comparable to symptoms of schizophrenia. When rats are born, they are much less developed than humans. Thus, neonatal maternal deprivation in rats has been used to study the effects of prenatal stress factors in humans. Several neuroendocrine studies have shown prolonged alterations in responsiveness of the hypothalamic–pituitary–adrenal axis in maternally deprived rats. A single 24-h separation of rat pups from their mother also induced a disruption of PPI later in life. This disruption developed only in adulthood, and the effect was strongly dependent on the timing of the stress. Treatment with typical or atypical antipsychotic drugs could reverse the disruption of PPI in these animals. Maternal deprivation did not cause PPI disruption in all rat strains tested, suggesting an important genetic modulation of stress sensitivity and PPI. Although baseline dopamine release in the forebrain nucleus accumbens was not altered, the stimulation of dopamine release by K^+ infusion or amphetamine treatment resulted in much greater increases in dopamine levels in maternally deprived rats than in controls.

The role of maternal care has also been identified. Rat dams may be grouped according to their maternal behavior into high- and low-licking/grooming dams. The offspring of low-licking/grooming dams, which are therefore subjected to relatively low levels of maternal care, showed lower PPI than the offspring

of high-licking/grooming dams, which are subjected to relatively high levels of maternal care. This result again suggests that stress very early on in development can have long-lasting effects on behavior in general and on PPI in particular.

There are, furthermore, other studies using early stressors, such as neonatal endotoxin treatment in rats to simulate bacterial infection, or chronic stress in the pregnant dam. In many of these studies, the offspring develop deficits in PPI when reaching adulthood.

Postweaning Stress and Prepulse Inhibition

Isolation rearing in rats consists of housing the animals in single cages for a prolonged period of time. This procedure induces a significant disruption of PPI, although a comparison of different rat strains showed that this was not seen in all strains. The timing of isolation housing was also important, with the greatest disruption of PPI seen in animals that had been singly housed from weaning. Isolation rearing, and its associated behavioral perturbations, in rats has been used as a developmental animal model of schizophrenia. Indeed, disruptions of PPI in isolation-reared rats could be reversed by treatment with various antipsychotic drugs. Social isolation was associated with alterations in the release and turnover of dopamine and other neurotransmitters in different parts of the brain, which could explain the effect of antipsychotic drugs. For example, similar to maternally deprived rats, in isolation-housed rats amphetamine-induced dopamine release in the nucleus accumbens was markedly enhanced when compared to group-housed controls. A similar finding has been made in

patients with schizophrenia when compared to normal controls.

Acute Stress and Prepulse Inhibition

Only few studies have specifically addressed the effect of acute stress on PPI. It is important to note that the PPI procedure itself is mildly stressful in rodents. Acute treatment with antipsychotic drugs has been shown to result in a subtle but significant increase in PPI. This may be explained by these drugs blocking enhanced activity of dopaminergic projections and/or other neurotransmitter pathways in response to the stress of the PPI procedure. This activation can be exacerbated by acute stress shortly before the procedure, although such effects are seen only under certain conditions. For example, we have observed that the mild stress of a saline injection caused a disruption of PPI in some mice strains but not others. Similarly, injection stress caused a disruption of PPI in rats that had undergone neonatal depletion of dopamine but not in controls. Interestingly, in healthy human volunteers, mental arithmetic stress was shown to cause a disruption of P50 gating, another measure of sensory information processing. The disruption took P50 gating to the reduced level seen in patients with schizophrenia.

Mechanism of Stress Effect on Prepulse Inhibition

Peripheral Effects

Two of the hallmarks of the stress response are enhanced release from the adrenals of adrenaline and cortisol (corticosterone in rodents). In a number of studies, it has been shown that adrenalectomy does not significantly affect PPI. However, we found that the 25% disruption of PPI caused by treatment with the dopamine-releasing drug amphetamine was absent in adrenalectomized rats (Figure 2). The disruption of PPI caused by treatment with the dopamine receptor agonist apomorphine, the glutamate receptor antagonist MK-801 (Figure 2), and the serotonin-1A receptor agonist 8-OH-DPAT was not altered in adrenalectomized rats. This suggests a modulatory action of circulating factors from the adrenal, presumably corticosterone, on the brain mechanisms involved in the effect of amphetamine on PPI. These brain mechanisms probably include dopaminergic activity. This was also seen in other studies. Thus, in adrenalectomized mice, haloperidol treatment did not cause the increase in PPI seen in intact mice or adrenalectomized mice supplemented with moderate doses of corticosterone. Interestingly, the haloperidol-induced increase of PPI was also absent in mice treated with

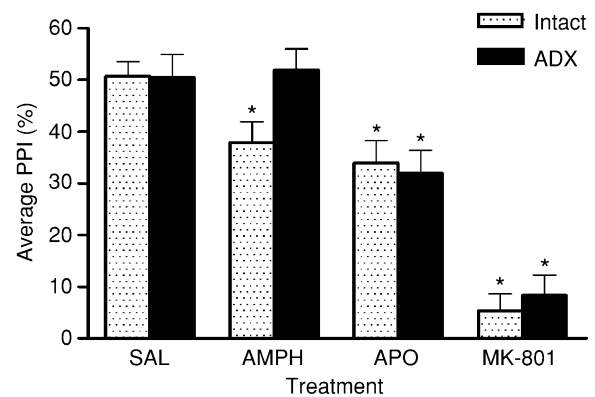


Figure 2 Average PPI of intact and adrenalectomized (ADX) rats after treatment with saline (SAL), 0.5 mg/kg of the dopamine-releasing drug amphetamine (AMPH), 0.1 mg/kg of the dopamine receptor agonist apomorphine (APO), and 0.1 mg/kg of the N-methyl-D-aspartate receptor antagonist MK-801. All drug treatments caused a disruption of PPI compared to saline treatment, except amphetamine treatment in adrenalectomized rats, which was ineffective. * $p < 0.05$ for difference with saline treatment in the same group. Based on unpublished data from Adams and van den Buuse; for methodological details, see van den Buuse et al. (2004).

high doses of corticosterone. Similar results were seen in rats. Chronic treatment with relatively high doses of corticosterone caused a decrease of baseline PPI in rats. Apomorphine-induced disruption of PPI was markedly greater in these animals when compared to either intact controls or adrenalectomized rats (Figure 3). Taken together, these results could indicate that the dopaminergic modulation of PPI, here exposed by the action of haloperidol and of apomorphine, is dependent on certain optimal circulating levels of corticosterone. Any deviation from these optimal levels, either by adrenalectomy or treatment with high doses of the steroid, affected this mechanism.

Central Nervous System Effects

In the hypothalamus, corticotropin releasing hormone (CRH) and vasopressin are involved in the regulation of adrenocorticotrophic hormone (ACTH) release from the pituitary. The central administration of CRH caused a disruption of PPI in rats. Similarly, in transgenic mice overexpressing CRH, there was a decrease of PPI that was reversed by treatment with haloperidol and risperidone. CRH 1 receptor (CRHR1) and CRHR2 appear to play opposing roles in mediating the effect of CRH on PPI. Thus, treatment with CRHR1 and CRHR2 antagonists caused an increase or a decrease of PPI, respectively. Furthermore, the disruption of PPI caused by CRH treatment could be blocked by pretreatment with a CRHR1 antagonist. In CRHR1 knockout mice, CRH did not disrupt PPI but, rather, enhanced it. Overall, these results could indicate that CRH in the brain plays an

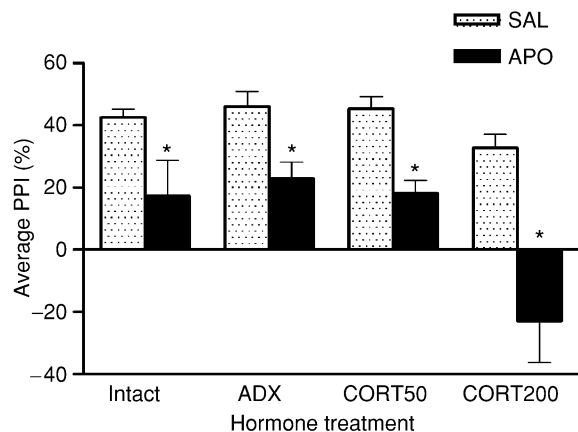


Figure 3 Average PPI of intact rats and of adrenalectomized rats that were implanted with either a cholesterol control pellet (ADX), a 50-mg corticosterone pellet (CORT50) or a 200-mg corticosterone pellet (CORT200). Two weeks after pellet implantation, the rats were injected with saline (SAL) or 0.1 mg/kg of the dopamine receptor agonist apomorphine (APO) and were tested for PPI. The 50–60% disruption of PPI observed in all groups was markedly enhanced (170%) in rats treated with a high dose of corticosterone. * $p < 0.05$ for difference with saline-treatment in the same group. Based on unpublished results from Adams and van den Buuse; for methodological details, see van den Buuse et al. (2004).

important role in the effects of stress on sensorimotor gating. Conversely, drugs acting on CRH receptors could be relevant in reversing PPI deficits in patients with PTSD or schizophrenia.

There is limited information on the role of vasopressin in PPI. Vasopressin-deficient Brattleboro rats display disruptions of PPI that can be reversed by chronic, but not acute, treatment with haloperidol. Also vasopressin V1b receptor knockout mice show a disruption of PPI. Chronic treatment with the atypical antipsychotic drugs, risperidone and clozapine (but not the typical antipsychotic, haloperidol) reversed this PPI disruption. The authors noted decreased extracellular levels of dopamine in the frontal cortex of these animals; however, the activity of the hypothalamic-pituitary-adrenal axis was not investigated.

Conclusion

Studies on the modulation of PPI by stress and peptides involved in the stress response provide

important insights into the role of stress in the symptoms of illnesses such as PTSD and schizophrenia. The effect of stress on PPI is mediated by circulating factors, such as corticosteroids, or central factors, such as CRH.

See Also the Following Articles

Corticosteroids and Stress; Corticotropin Releasing Factor (CRF); Dopamine, Central; Schizophrenia; Serotonin.

Further Reading

- Ellenbroek, B. A. (2003). Animal models in the genomic era: possibilities and limitations with special emphasis on schizophrenia. *Behavioural Pharmacology* **14**, 409–417.
- Geyer, M. A. and Markou, A. (1995). Animal models of psychiatric disorders. In: Bloom, F. E. & Kupfer, D. J. (eds.) *Psychopharmacology: the fourth generation of progress*, pp. 787–798. New York: Raven Press.
- Koch, M. and Fendt, M. (2003). Startle response modulation as a behavioral tool in neuropharmacology. *Current Neuropharmacology* **1**, 175–185.
- Swerdlow, N. R., Geyer, M. A. and Braff, D. (2001). Neural circuit regulation of prepulse inhibition of startle in the rat: current knowledge and future challenges. *Psychopharmacology* **156**, 194–215.
- van den Buuse, M., Garner, B., Gogos, A., et al. (2005). Importance of animal models in schizophrenia research. *Australian and New Zealand Journal of Psychiatry* **39**, 550–557.
- van den Buuse, M., Garner, B. and Koch, M. (2003). Neurodevelopmental animal models of schizophrenia: effects on prepulse inhibition. *Current Molecular Medicine* **3**, 459–471.
- van den Buuse, M., Morris, M., Chavez, C., et al. (2004). Effect of adrenalectomy and corticosterone replacement on prepulse inhibition and locomotor activity in mice. *British Journal of Pharmacology* **142**, 543–550.
- Zhang, T. Y., Chretien, P., Meaney, M. J., et al. (2005). Influence of naturally occurring variations in maternal care on prepulse inhibition of acoustic startle and the medial prefrontal cortical dopamine response to stress in adult rats. *Journal of Neuroscience* **25**, 1493–1502.

Pressure, Effects of Extreme High and Low

R J Værnes

NUTEC Crisis Management, Ytre Lakevaag, and
University of Bergen, Bergen, Norway

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
R J Værnes, volume 3, pp 208–220, © 2000, Elsevier Inc.

effects consist of neuropsychological complaints regarding cognitive functioning and fine visuomotor tempo and coordination.

Introduction

Humans Working under High Pressure

Humans Working under Low Pressure

Concluding Remarks

Glossary

<i>Acute mountain sickness (AMS)</i>	A high-altitude illness, which can be classified as normal or malignant AMS. Malignant AMS includes features of both pulmonary and cerebral edema.
<i>Chronic mountain sickness (CMS)</i>	A high-altitude illness whose symptoms are vague neuropsychological complaints, including headache, dizziness, fatigue, and difficulty concentrating.
<i>High-pressure nervous syndrome</i>	Behavioral disturbances after being at depths greater than 150 m. Central nervous system manifestations include tremor, dizziness, nausea, and electroencephalogram changes. It becomes more severe with increasing depth and during fast rates of compression.
<i>Hypoxia</i>	Lack of oxygen that results from, at normal pressure, a reduction of pO_2 and, at high altitude, the reduction in atmospheric pressure. Central nervous system manifestations differ between acute hypoxia (seconds to 1–2 h) and chronic hypoxia (hours to years).
<i>Nitrogen narcosis</i>	An illness from breathing compressed air at depths deeper than 20 m; characterized by euphoria, impaired cognitive functioning, and impaired neuromuscular coordination.
<i>Oxygen toxicity</i>	An illness that increases progressively with the increase of inspired pO_2 and with greater duration of exposure; central nervous system manifestations range from localized muscle twitching to grand mal seizures.
<i>Residual effects</i>	Residual impairment of central nervous system functioning following the return to normal pressure from deep depths and from high altitudes. For both high- and low-pressure conditions, the residual

Introduction

What Constitutes Extreme High- and Low-Pressure Conditions for Humans?

As we have throughout the centuries, today humans go into the sea for food, military operations, sport, treasure-hunting, and exploration. The earliest indications of breath-hold diving and the first techniques used by humans at work in the sea are from an archaeological dig in Mesopotamia dating from 4500 BC and from Aristotle, who in the fourth century BC mentioned a type of diving machine. The historical stages in the development of diving are:

1. Breath-hold diving (also called free diving)
2. Diving bells
3. Surface-supported or helmet diving (also called hard-hat diving)
4. Self-contained underwater breathing apparatus (scuba diving)
5. Saturation diving

Diving physiologists had, by the late 1950s, made great strides toward understanding and managing some of the principal hazards that beset divers in their progress toward the depths. These hazards included oxygen toxicity, the narcotic effects of inert gases used as oxygen diluents in high-pressure environments, and the mechanical effects of dense atmospheres on respiratory work and the effectiveness of ventilation.

When oil and gas exploration started in the North Sea in the 1960s, diving became important for the maintenance, inspection, and repair of subsea installations. Therefore, the saturation diving technique was developed further, and diving to depths between 50 and 180 msw (meters of seawater) became a standard procedure. In parallel with this offshore diving, experimental studies on diving down to a maximum of 850 msw were conducted in the United States, Europe, and Japan. This research led to the observation of the high-pressure nervous syndrome (HPNS), a syndrome consisting of tremors, dizziness and nausea, electrophysiological changes, and impaired performance.

Humans, however, expose themselves to more than just extremely high pressures. For centuries, humans

have attempted to climb higher and higher, and scientists have been astonished repeatedly by mountaineers' ability to ascend higher than their confident predictions. Whereas the central nervous system (CNS) problems from diving are due to oxygen toxicity, inert gas narcosis, and HPNS, the CNS problems of extreme altitude are due to hypoxia (because there is reduction in atmospheric, and hence oxygen, pressure with increases in altitude) and cold (because of the temperature decrease with altitude and wind chill).

One of the earliest European reports of the problems of high altitude was given by the Greeks who, during their yearly ceremonial visit to the summit of Mountain Olympus (2911 m), were not able to survive unless they applied moist sponges to their noses (possibly soaked in vinegar). The first documented description of mountain sickness comes from Chinese sources in 32–37 BC. It has been suggested that the occurrence of mountain sickness and its complications may have been taken by the Chinese as a sign for them not to transgress their natural boundaries.

In the late nineteenth and early twentieth centuries, there was increasing scientific and medical interest in both the oxygen transport system and altitude adaptation, and many discoveries were made. Viault recorded the best-known physiological response to high altitude – an increase in the number of red blood cells. The tempo of Himalayan exploration accelerated during this period, and in 1892 while on Pioneer Peak (7000 m), Conway described the debilitating effects of long periods at high altitude. In 1953, the first ascent of Mt. Everest (8848 m) was conducted by Hillary and Tensing. However, they used supplementary oxygen. In 1978, Messner and Habeler made the first ascent of Mt. Everest without supplementary oxygen, and in 1987 the same was done in the winter (December) by Sherpa Ang Rita.

What Constitutes the Physical Characteristics of These Conditions?

Diving Awareness of the natural laws covering diving, and the results of disobedience, has fostered the development of the systems and techniques that allow humans to dive safely to increasingly greater depths. The laws are the ideal gas law (behavior of gases at low pressures), Boyle's law (volume is inversely proportional to pressure), Charles's law (volume is directly proportional to temperature), real gas law (behavior of gases at high pressures), Dalton's law (law of additive partial pressures), Graham's law (diffusion rate is proportional to molecular rate), Amagat's law (law of additive partial volumes), thermal laws (specific heat, conduction and convection, and heat transfer), acoustical laws (conduction, diffraction, and the speed

of sound), density laws (Archimedes's principle), and optical laws (curvature of light and absorption of color). A direct line can be drawn from any existing diving technique to at least one of these fundamental laws.

When people dive down to a certain depth, a compression profile is followed. For each 10 msw of depth, the pressure increases 1 atm absolute (ATA). Especially in the 1970s, surface-oriented diving with a wet diving bell was common practice. Two men descended, for example, in the bell to approximately 100 msw (11 ATA). The compression rate was normally in the range of 30 msw min⁻¹ (3 ATA min⁻¹). The average time at maximum pressure was less than 30 min, and the tasks involved light or moderate effort. Only one diver exited to the work site while the other monitored his or her activity and stood by to assist.

Because the time required to come back to surface (decompress) is often longer than the time spent on the bottom, the ratio of diver work time to decompression time is not always a favorable one (if a diver decompresses too fast, he may get the bends, that is, gas bubbles in the skin, joints, spinal cord, or bloodstream).

The price that a diver must pay for staying deeper and longer led to the development of saturation diving. First accomplished in 1938, this diving technique was developed further in parallel with the oil and gas exploration in the North Sea. If a team of divers are to work at 150 msw (16 ATA), they compress to 16 ATA in a living chamber at a diving platform or vessel and stay at that pressure for 10 days. Pairs of divers then enter a dry diving bell, which is lowered to the work site at 150 msw. The bell stays at that depth for up to 8 h, and one diver at a time works in the water for half the period. After the 8-h period, the bell is taken up to surface and connected to the living chamber where the divers eat, relax, and sleep (all the time under the same pressure). After 10 days, they start decompression to surface pressure, which takes several days.

High altitudes In 1878, it was known already that most of the deleterious effects of high altitude in humans were caused by hypoxia, a direct result of the reduction in atmospheric pressure. However, there is still confusion about the relationship between barometric pressure and altitude, particularly at extreme heights. For example, the barometric pressure at the summit of Mt. Everest is considerably higher than that predicted by the standard pressure–altitude tables used in aviation.

Although most of the high-altitude effects are due to hypoxia, additional deterioration results from cold, dehydration, and solar and ionizing radiation.

Most of these additional hazards, however, can be avoided by using proper clothing or shelter. Only hypoxia is unavoidable unless supplementary oxygen is available. Low barometric pressure itself has no physiological consequences unless the decompression is rapid. For example, when a window fails in an aircraft, an explosive decompression can occur; and, just as in diving, rapid decompression causes barotraumas as a result of the very rapid enlargement of airspaces within the body, especially in the lungs and middle ear cavity.

Humans Working under High Pressure

Working Conditions (Types of Diving)

Helmet diving The first major breakthrough in helmet diving was the invention of an open diving dress in 1819 by Siebe, a waist-length jacket, a metal helmet, and air under pressure supplied from the surface to the helmet. Siebe later modified his open dress to the closed type, retaining the helmet and force pump and enclosing divers in a continuous airtight dress. Air was vented by a valve, the basis for modern hard-hat diving gear.

As advances were made in helmet diving, divers gained more air, more protection, and more mobility, but it was not until the development of self-contained underwater breathing apparatus (scuba) gear that they were afforded the freedom that breath-hold divers had to move about, coupled with an air supply that enabled them to stay down longer than free divers.

Scuba diving The Cousteau–Gagnan valve (aqua lung, presented in 1943) is the basis for modern open-circuit scuba gear. Open-circuit breathing gear usually contains compressed air as a breathing mix. It has a demand regulator, which provides air when the diver inhales, and the air is vented into the water on exhalation. To conserve the breathing mix and extend the time at depth, closed-circuit and semiclosed rebreathers have been developed.

A closed-circuit apparatus can employ pure oxygen or a mixed gas with oxygen and a diluent gas such as nitrogen or helium. The risks inherent in breathing pure oxygen, primarily oxygen toxicity, limit the operating depths to approximately 10 msw or 18 msw resting.

When divers go deeper than 30 msw on air, dangerous signs and symptoms of euphoria, intoxication, and narcosis occur. The first inference that the nitrogen constituent of compressed air is responsible for this narcosis was made by Behnke and colleagues in 1935. They concluded that compressed air at depths

deeper than 20 msw exerts a narcosis characterized by “euphoria, retardation of the higher mental processes, and impaired neuromuscular coordination” (Behnke et al. 1935). The signs and symptoms of the narcosis are similar to those due to drinking alcohol, hypoxia, and the early stages of anesthesia. Exposure to depths greater than 90 msw may also result in loss of consciousness.

Saturation diving The use of saturation diving in a commercial application was, in fact, suggested by Behnke as early as 1942. He noted that frequent decompressions were “not only potentially dangerous, but highly uneconomical. It would appear advisable therefore to keep men at work on a job continually under pressure” (Behnke 1942).

Although the first saturation dives were made using air, this method has never been used for diving at a depth less than 40 msw. For diving deeper, however, this has been the main diving technique since the early 1980s. What they initially had to do, however, was remove the nitrogen (which would lead to narcosis) and use helium instead. The breathing gas in saturation diving is therefore called heliox (helium and oxygen).

Modern-day saturation diving platforms in the North Sea have greatly extended diving capabilities. Up to 28 people can be in saturation, making excursions to different depths with two bells. Compression to depths greater than 150 msw, however, produces the signs and symptoms of high-pressure nervous syndrome (HPNS). The syndrome includes tremor in the hands and arms, electroencephalograph (EEG) changes, dizziness, nausea, and vomiting. At depths greater than 300 msw, lapses of consciousness may occur. The symptoms of HPNS become more severe with increasing depth and during fast rates of compression.

Effects on the Central Nervous System

As already indicated, several CNS reactions may occur when divers are exposed to this high-pressure environment. The stress reaction in the individual diver, however, is a product of a combination of factors, psychological and physiological, and is often a synergistic effect. This section focuses on the following six factors, which both alone and in combination can lead to severe stress and, in some cases, a fatal outcome in diving: (1) anxiety and panic, (2) oxygen toxicity, (3) nitrogen narcosis, (4) HPNS, (5) isolation and confinement, and (6) residual effects.

Anxiety and panic Stress during diving is probably the central problem causing the accidents and resulting injuries and fatalities that occur in sport divers.

Most diving-accident researchers implicate panic, as a response to stress, as the major cause of diving fatalities. The skills learned in training, such as the careful use of equipment and the techniques of coping with accidents, appear to be lost when divers panic. Therefore, it is important for divers, especially sports divers, to understand stress and to learn to handle stress.

Divers' use of defensive reactions to a threat also lead to inadequate performance under stress, and decrements in both defense mechanisms and performance relate to endocrine stress reaction. The first study demonstrating that perceptual defense mechanisms, as tested by the defense mechanism test (DMT), predicted individual reactions to dangerous situations was on attack divers. Three groups of divers ($N = 29$, 20, and 29) were tested with DMT, and the protocol was coded, rated, and ranked independently by three raters. All three raters' predictions, based exclusively on the DMT results, correlated significantly with the pass/fail criteria at the school.

In later studies in which more specific performance criteria were used, the same relationships were found. In a study of 45 navy divers, the individual score on defense mechanisms correlated significantly with the reduction in reasoning capacity at 60 msw. In a study of 29 navy divers, the same relationship between defense mechanisms and mental performance was found. In addition, the factor pattern of the endocrine postdive measures showed cortisol, catecholamine, and testosterone factors. There was a correlation between the cortisol factor and both defense pattern and performance decrement, confirming previous studies from other high-risk environments (parachutists).

Oxygen toxicity The greater application of hyperoxygenation both in therapy and in diving since the early 1970s has led to a better understanding of the universal dependence of vital biological processes on the cellular antioxidant defense mechanisms that have evolved as an adaptation to Earth's atmospheric oxygen tension. It is now apparent that the same oxygen pressures required to sustain life cause lethal oxygen poisoning in the absence of these mechanisms.

The severity of oxygen poisoning increases progressively with the increase in the inspired partial pressure of oxygen and greater duration of exposure. At sufficient pressure and exposure duration (i.e., an attack diver breathing pure oxygen at too great a depth), oxygen initially causes functional impairment and ultimately causes the chemical destruction of any living cell.

Donald exposed divers to oxygen pressures of 3 ATA or higher until CNS effects were experienced.

Unfortunately, the sequence of progression did not consistently include minor symptoms before the onset of convulsions. Even when preconvulsive aura was experienced, it was sometimes followed so rapidly by seizures that it had little practical value as a warning. An enormous variation in oxygen tolerance was found among divers exposed to the same conditions. Attempts to correlate CNS oxygen tolerance with factors such as age, weight, physical fitness, smoking, alcohol ingestion, psychological stability, or personality traits were all unsuccessful. Even in the same diver, there was considerable variability in oxygen tolerance from day to day. The conclusion was that variability in oxygen tolerance from day to day within the same subject was as great as between subjects.

The seizures caused by oxygen toxicity are essentially the same as those occurring in grand mal epilepsy. They can be preceded by an aura or by a sequence of premonitory sensations, or they can occur suddenly. In contrast to the cerebral hypoxia that occurs during an epileptic seizure, brain hyperoxygenation is maintained during an oxygen convulsion by a high alveolar partial pressure of oxygen in conjunction with marked hypercapnia and increased cerebral blood flow. Biochemical studies on free radicals under hyperoxic conditions have shown, however, that extracellular hyperoxia leads to cellular hypoxia through free-radical mechanisms.

However, it is not only the use of pure oxygen in diving that can lead to serious CNS problems. Ongoing research has shown that there is probably a synergistic effect between oxygen (in heliox with a partial pressure of O_2 of only 0.6 ATA) and high pressure when people dive to extreme depths (>300 msw), contributing to the explanation of CNS-related symptoms and signs (e.g., HPNS) that are observed.

In a series of experiments, the effects of various partial pressures of oxygen (0.2, 0.4, and 0.6 ATA) in the heliox breathing gas on HPNS were studied in rats. Rats were exposed to 91 ATA, equivalent to approximately 45 ATA for humans. The results showed that HPNS symptoms were reduced significantly with a lowered oxygen partial pressure. Follow-up studies on the effects of hyperbaric oxygen ($>95\%$) on neuronal transmission in the perforant path/dentate area of hippocampal formation showed that epileptic seizure activity produces long-term changes in synaptic excitability in granular cells of the dentate area.

The inhibition of nitric oxide (NO) production has been reported to delay the onset of hyperbaric oxygen-induced epileptic seizures in mice and in rats. The mechanism(s) behind this protective effect is still equivocal. It has been suggested that NO is

involved in the regulation of blood flow and arterial pressure, as a transmitter in peripheral nerves, and as an intercellular or, rather, transcellular transmitter between CNS neurons. The protective effect of inhibiting NO production could, therefore, be due to changes in cerebral blood flow, neuronal activity, or both.

Nitrogen narcosis All divers know that diving deeper than 30 msw on air can lead to dangerous symptoms and signs of euphoria, intoxication, and narcosis. One of the first reports of nitrogen narcosis was given by a Frenchman in 1835. Around the turn of the twentieth century, it was reported by several others in tunnel workers breathing air at 46 m. However, as mentioned previously, Behnke and co-workers were the first to claim that nitrogen was the agent responsible for the narcosis (Behnke 1942).

As might be expected, there is a wide variation in individual susceptibility. Whereas one diver becomes narcotic at pressures as high as 50 msw, another is affected quite severely at only 30 msw. In this connection, Cousteau claimed that emotionally stable individuals are usually affected to a smaller extent than less stable subjects. This was confirmed later in experimental studies by Værnes in which divers sat in a pressure chamber taking cognitive tests on paper and performing no other physical activity.

In both studies, a significant reduction in long-term memory (LTM) was found. When information input was done predive and recall at 60 msw, the mean reduction in LTM was 31.2%. When input was done at 60 msw and recall later at the surface, LTM impairment was even worse (45.6%). (The control situation was both input and recall at the surface in the same chamber with the same duration between input and recall.) Memory impairment was obviously a major effect of nitrogen narcosis. As reported earlier, some divers also had a reduction in reasoning, although the average reduction was not significant. However, for reasoning impairment, in both studies there was a significant correlation with the degree of defensive reaction pattern as defined by DMT, and in the second study, it also correlated with the degree of cortisol activation.

High-pressure nervous syndrome Compression in a heliox atmosphere induces a number of behavioral disturbances at depths greater than 150 msw (16 ATA). These disturbances were described by Bennett as nausea, vertigo, and tremor during rapid compression to 180 msw, and he attributed them to the helium (i.e., helium tremor). Brauer and colleagues and Fructus and colleagues described a number of symptoms occurring during simulated dives between 300 and

365 msw. The symptoms have been termed the high-pressure nervous syndrome (HPNS) and include tremor in the hands and arms, increased slow-wave activity (2–7 Hz) and depression of α waves (8–13 Hz) in the EEG, dizziness, nausea, and vomiting. The syndrome was replicated by Bennett and Towse during a dive to 457 msw, and it has been demonstrated that lapses of consciousness may occur at depths greater than 300 msw. The symptoms of HPNS become more severe with increasing depth and during fast rates of compression.

Various methods have been used to ameliorate the signs and symptoms of HPNS, including the selection of less sensitive divers, the use of excursions from a shallow saturation exposure, slow exponential compression with stops for adaptation, and the use of narcotics to antagonize the effects of HPNS. Studies of these methods have shown that the speed of compression correlates with the extent and intensity of the various symptoms. In particular, a slow compression interrupted by holding stages has been shown to reduce HPNS considerably.

A major problem in evaluating the effects of HPNS by manipulating the compression conditions or selecting less sensitive divers is the individual variation in the reaction pattern. Whereas one diver may show a normal EEG but significant tremor followed by impaired motor performance, another diver may show normal motor functions but significant EEG changes accompanied by impaired cognitive performance. To study these different responses, it is recommended that a relatively comprehensive battery of tests be used for evaluating a compression profile. The validity of such a battery of tests can be related not only to fitness at depth but also to the temporal aftereffects that have been observed in deep diving.

Therefore, a compression profile with holding stages and a slower compression rate at deeper depth has been developed for operational diving to 180 msw and deeper in the North Sea. In a series of simulated dives to 360 msw, including a total of 18 divers, the proposed profile was evaluated. HPNS testing revealed only mild effects of the compression. Only three divers had impairments of more than 2 standard deviations in peripheral motor function. Memory was impaired periodically in two divers. Their tremor had minimal effects on their performance. In only 1 of the 18 divers had a marked EEG change followed by cognitive impairment. The results confirmed that using a compression profile with rates decreasing progressively with increasing depth, and with several intermediate stops, provides fit divers at depth.

This profile was later used in the tie-in operation at 220 msw at, by then, the world's largest oil platform (Gullfaks C) in 1989. A total of 52 divers took part in

the operation and all were tested with a field battery of HPNS tests at pressure before being allowed to start work in the water. No dives were aborted due to HPNS, and only 1 out of the 52 divers had to have an extended observation period in the living chamber (6 h) before being allowed to start the in-water operation. Although made even more conservative through two revisions, this original profile is the standard profile today for all diving in the Norwegian sector of the North Sea deeper than 180 msw.

Several studies have suggested that cerebral monoamines may be involved in the occurrence of some HPNS symptoms. Some studies suggested that dopamine (DA) counters the convulsive activity induced by pressure. Nevertheless, others did not find significant changes in the concentration or kinetics of DA analyzed in the brain homogenates of animals after exposure to pressure.

More recent studies in free-moving rats using microdialysis confirmed that the pressure induced striatal increase and implicated D1, D2, and NMDA receptors in this change. Recent works performed on GABA_A or GABA_B receptors using agonists (muscimol or baclofen) injected focally in the substantia nigra pars reticulata (SNr) or substantia nigra pars compacta (SNc) implicated these receptors in the striatal DA release and showed that helium pressure inhibits the activity of GABA neurotransmission at the level of substantia nigra, which induces a DA change.

Microdialysis studies indicated that the helium pressure induces an increase of serotonin (5-HT) in the striatum, which can explain the activation of these receptors. At the level of nucleus accumbens, the use of 5-HT_{2C} receptor agonists indicates that this receptor plays a crucial role in the development of locomotor activity (LMA) and in the pressure-induced increase in DA accumbens release. In contrast, the administration of 5-HT_{2A} receptor antagonists had no effect on LMA and DA increase, but it reduced myoclonia. These results suggest that helium pressure may simultaneously induce an increase in 5-HT transmission at the level of 5-HT_{2A} receptors and a decrease in 5-HT transmission at the level of 5-HT_{2C} receptors.

These studies are some examples of the several neuropharmacological and neurochemical findings that have increased our information on the origin and mechanisms of HPNS. In summary, it can be stated that helium pressure acts specifically on some neurotransmissions, on some neurochemical interactions between the systems and the structures, and on some receptors or on their synthesis. It is evident, however, that more work remains to be done to increase our understanding of the effect of pressure on the nervous system.

Isolation and confinement In addition to these pressure- and gas-related effects of deep diving on the CNS, studies have found that isolation and confinement themselves can have an effect. Although the stimulus-deprivation studies performed in the 1950s showed that the CNS needs stimulus variation, these extreme experiments seemed to be a long way from living in a small chamber (be it a saturation diving complex or a space station) for several weeks.

How has it been possible to evaluate this effect? Did someone pay to lock up test subjects inside a pressure chamber at normal atmospheric pressure for weeks? American, Russian, and European space agencies have used diving saturation chambers to evaluate the effects of long-term isolation and confinement in conjunction with preparing astronauts for space station Mir and the future international space station Alpha. Using these space-related experiments, it is possible to analyze the effects of isolation that a diver is exposed to during deep, long-duration (>30 days) dives. Three such experiments performed by the European Space Agency (ESA) are described next.

The Long-Term Programme Office (LTPO) of the Directorate of Space Station and Microgravity (D/SSM) at ESA carried out studies on the psychological and physiological aspects of long-duration human spaceflight in multinational crews under conditions of isolation and confinement in diving saturation chambers. LTPO first performed an experimental test with a crew of six men for 28 days (ISEMSI-90) at the Norwegian Underwater Technology Center (NUTEC) in Bergen, Norway, and then performed a second one with a crew of four (three men, one woman) for 60 days (EXEMSI-92) at Deutsche Zentrum für Luft-und Raumfahrt (DLR) in Cologne, Germany. In 1994, ESA was planning the EUROMIR-95 long-duration flight (135 days) of a European astronaut on board the Russian orbital station Mir. In order to perform detailed investigations on the human-related aspects in such a long-duration isolation, ESA organized a ground simulation, the Human Behavior Study (HUBES), for which the chamber facility at the Institute for Biomedical Problems (IBMP) in Moscow was set up like a small space station.

Cognitive fatigue and subjective health symptoms occurred in all three studies. On performance, an increase in decision time and an increase in the number of errors were found in several of the experiments. For the physiological studies, HUBES confirmed the earlier findings of the ISEMSI-90 and EXEMSI-92 tests, showing weekly rhythms, individual rhythms, and overall phases as in the group psychology studies. It was therefore concluded that for long-duration

dives (and spaceflights) experimenters will have to take into account the individual biological patterns of the subjects. If this is done, the outcome of the experiments will be more predictable and the comparison of data between the subjects will become easier.

These three ESA experiments have therefore shown, among other things, that the timing of individual experiments is critical when parameters are involved that are strongly under endocrinological control, such as those of the salt-water homeostasis, and therefore display rhythmic patterns.

Residual effects As in high-altitude medicine, researchers in diving medicine have vigorously debated the possibility of residual impairment of CNS function following the subjects' return to normal pressure. Several deep saturation dives performed between 1968 and 1989 showed temporary CNS changes immediately after the dives. Given these CNS changes, observed both immediately postdive and during deep saturation dives, an important question was whether chronic CNS effects could develop after single or repeated exposures, as in the previous findings on diving accidents.

The neuropsychological status of saturation divers was assessed before and after 300–500 msw dives (deep saturation diving, DSD group, $N = 64$) and before and after 3.5 years of ordinary saturation diving (SD group, $N = 32$). The average baseline results showed divers to be slightly superior to the nondiving controls. Mild-to-moderate neuropsychological changes ($>10\%$ impairment) were found in measures of tremor, spatial memory, vigilance, and autonomic reactivity in 20% of the divers after deep dives (DSD group). One year postdive, no recovery was observed except for a vigilance test. In the SD group, 20% of the divers showed $>10\%$ impairment after 3.5 years of ordinary saturation diving. A significant reduction in autonomic reactivity was also found, and there was a relationship between low autonomic reactivity before saturation diving and the number of impairments that were $>10\%$. For the whole group (DSD + SD divers), negative correlations were found between saturation experience and results on memory and complex visuomotor tests. The number of years of diving from the first to last examination correlated positively with the number of impairments that were $>10\%$ and with a reduction in autonomic reactivity. No similar correlations were found at dive variables after approximately 3 years of air diving.

The mild-to-moderate changes seen in some divers, therefore, seem to be the effects of saturation diving. Because one deep dive may cause an effect similar to

the effects of 3.5 years of ordinary saturation diving, there is reason to believe that repeated deep diving may lead to more pronounced neuropsychological impairment.

Such mild-to-moderate changes in some divers indicate the start of a pathological process on the subcortical level (mild encephalopathy) such as a decrease in memory and vigilance, low autonomic reactivity, and increased tremor, but the underlying mechanisms are unclear. The degree of individual reaction and the variation in symptomatology indicate that the mechanisms for both the temporal and the subtle permanent changes after saturation diving are very complex. Individual factors, additive factors, and synergistic mechanisms may be involved.

Several international studies followed the publication of the Norwegian study and were reviewed by Dutka. In his review, Dutka states that divers who have had decompression sickness can develop long-term health effects, although there are conflicting results among deep-sea divers with no history of decompression sickness. His conclusion is in conflict with the follow-up study on North Sea divers presented in 2003 by Troland. In this study, the results seem even more convincing, supporting the statement from an international consensus conference that said: "There is evidence that changes in bone, the central nervous system and the lung can be demonstrated in some divers who have not experienced a diving accident or other established environmental hazards. The changes are in most cases minor and do not influence the divers quality of health. However, the changes are of a nature that may influence the diver's future health (Hope et al. 1994)."

Humans Working under Low Pressure

Conditions

Human responses to hypoxia are complex and depend, among other things, on the severity of the hypoxia and the rate at which it is imposed. Severe sudden hypoxia experienced by aviators when cabin pressure is lost at a high altitude produces an entirely different result than that experienced by a climber who reaches the same altitude after weeks of acclimatization. Furthermore, people with a lifelong exposure to high altitudes develop further changes in response to hypoxia compared with even acclimatized lowlanders. Adaptations may also occur after a species has lived at high altitudes for generations due to the selection of genes advantageous to life in the hypoxic environment.

Generally, hypoxic conditions can be divided into four categories: (1) acute hypoxia (from seconds to

1–2 h), (2) chronic hypoxia (from hours to years, i.e., subjects in chamber experiments, mountaineers, workers in high-altitude mining towns, sea-level residents who go to high altitudes), (3) lifelong hypoxia (i.e., high-altitude residents who were born, bred, and live at high altitudes), and (4) species that have lived at altitude for generations.

Effects on the Central Nervous System

People who go to high altitudes often have changes in neuropsychological function, including memory, mood change, and special senses such as vision. This has been observed both in lowlanders staying at high altitudes for some time and in high-altitude natives. There is also mounting evidence that there may be persistent CNS defects on return to sea level after periods of severe hypoxia at high altitude. This is analogous to the indications of permanent CNS after-effects in some divers after deep diving.

Because an increasing number of climbers choose to climb at great altitudes without supplementary oxygen, these findings are of special interest. The increase in morbidity and mortality during such expeditions and the occurrence of irrational decisions made by severely hypoxic climbers have also led to serious concern from medical and operational points of view.

Acute hypoxia At modest altitude (>1500 m), breathlessness may be felt on exertion and some rise in heart rate may be noticed, but the main effect is on the CNS. Night vision is impaired at altitudes as low as 1500 m. The tingling of the fingers and mouth may be noticed at 4000–5000 m. However, although now definitely hypoxic, climbers have very little subjective prewarning. Some people may become unconscious above 5000 m, whereas most become unconscious at above 7000 m. In response to an acute exposure to Mt. Everest's altitude (8848 m), unacclimatized subjects remain conscious for only approximately 2 min.

Within the brain, the hippocampus, white matter, superior colliculus, and lateral geniculates appear particularly sensitive to levels of oxygen. Brain lactate levels increase during the early stages. The brain tissue concentrations of ATP, ADP, and AMP apparently remain close to normal even during severe hypoxia. In animal studies in which the arterial partial pressure of oxygen is reduced gradually (from 80 to 20 mmHg), the subjects' EEG amplitude initially increases slightly and then slow waves and spikes appear. Subsequently, the slow waves decrease in amplitude and then disappear; the small spikes become sporadic, and finally the EEG flattens. This initial activation followed by depression may be due to hypoxic effects

on the reticular activating system. Visually evoked potentials are increased initially and are then abolished as the level of oxygen is reduced. The brain-stem auditory response is abolished by hypoxia.

The effects of hypoxemia and hypocapnia at high altitude have opposing effects on the cerebral circulation. There have not, however, been systematic studies of the cerebral blood flow at various altitudes, partly because of the difficulties of measurement.

Chronic hypoxia at moderate altitudes It has been important to study the effects of mild hypoxia on the performance of aircrews, who are required to maintain peak performance for hours while performing monotonous tasks under hypobaric conditions. There is a general agreement that CNS function is impaired at altitudes over approximately 4500 m; however, what is the lowest altitude at which minor, but operational, significant alterations in CNS occur? This question arises frequently in the aviation industry because it is important in selecting the cabin pressure of commercial aircraft.

Earlier studies showed that performance declines as the time devoted to vigilance tasks increases. Studies on the effects of hypoxia on vigilance performance have produced contradictory evidence. Impairment in performance was found at altitudes as low as 4000 m and carboxyhemoglobin (HbCO) levels of 1.8%. Other studies found no similar effects at 3500 m and at HbCO as high as 8.4%.

Further studies of people performing tasks at high terrestrial elevations have, however, demonstrated that impairment of a wide range of tasks occurs and that this correlates with the degree of hypoxia. In addition, animal performance in tests of maze performance, visual discrimination, or nondiscriminated avoidance are affected adversely by exposure to low oxygen tension. These changes are qualitatively similar to the behavior noted in rats when levels of brain biogenic amines, norepinephrine and DA, are reduced by pharmacological agents. Hypoxia also affects brain norepinephrine in rats. Exposure to 5 or 10% O₂ for 1 h resulted in a 23% decrease of this amine.

The variability in the response to hypobaric exposures, both within and between individuals, has been emphasized in relation to the symptoms and signs of hypoxia. Some proportion of the variation may be due to the differences in the respiratory responses to the hypoxia, causing significant temporal and individual differences in the tensions of oxygen and carbon dioxide in the arterial blood. Furthermore, the degree of cerebral vasoconstriction also causes a proportion of this variation.

Performing learned psychomotor tasks is unimpaired to at least 3000 m. Long- and short-term

memory, arithmetic, and conceptual reasoning tests are also unaffected below 3000 m. Previously, it was assumed that this degree of hypoxia did not produce any effects on human performance. However, breathing air at 2500 m causes a significant increase in the time required to achieve optimum performance on a novel task. The delayed learning is probably due to the retardation of some oxidative processes within the brain rather than to interference with oxidative phosphorylation, which hardly can be affected because the oxygen tension of the cerebral venous blood falls 2–4 mmHg during an ascent from ground level to 2500 m. The impairment of learning increases markedly at altitudes above 3704 m (12 000 feet).

A study by Værnes and colleagues was designed to examine the effects of a prolonged exposure to mild hypoxia on performance and endocrine reactivity. An altitude of 3048 m (10 000 feet) was simulated in an altitude chamber for 6.5 h. The study was initiated after reports of some cases of headache, dizziness, visual disturbances, and subjective feeling of performance decrement by an aircrew on a maritime surveillance aircraft during a trans-Atlantic flight with a cabin altitude of 3048 m. The aim of the study was to determine whether these symptoms were the effects of general stress, mild hypoxia, or a combination of both.

Performance tests indicated significant effects of hypoxia. In contrast to earlier studies on the relationship of hypoxia and performance, no relationship between impaired performance and duration of exposure to hypoxia was found. Repeated testing throughout exposure indicated stable individual reactions. The endocrine variables did not support the hypothesis that activation or stress caused the impairment observed. In addition to impaired neuropsychological test performance and impaired task performance, subjects reported headache, weakness, and some dizziness. Comparisons of the various tests confirmed the previous results showing that mild hypoxia yields varying degrees of impairment on different cognitive functions.

Changes that take place over a period of hours to months at altitude, termed acclimatization, occur in different systems of the body and have a different time course. In some cases, they involve a biphasic response (i.e., the heart rate shows a rise within minutes, followed by a fall over weeks at altitude). The best-known response is the increase in hemoglobin concentration due to the rapid decrease in plasma volume followed by a much slower increase in red cell mass.

Adaptation takes place over decades or is the result of being born and bred in a mountain environment. No definite examples of CNS-related adaptation

have, however, been demonstrated in humans, but examples include an increase in the number of branches leaving the main coronary trunks and the blunting of the hypoxic ventilatory response.

The observed deterioration in the mental and physical conditions of humans as a result of prolonged stay at altitude is termed high-altitude deterioration. The term was first used by members of the early Everest expeditions. High-altitude deterioration, in the absence of dehydration, hypothermia, exhaustion, and/or starvation, is characterized by poor appetite, weight loss, lethargy, slow recovery from fatigue, an increasing lack of will power to start new tasks, and irritability. The mechanisms underlying this deterioration are unknown, but it is quite obvious that the effect of chronic severe hypoxia on the CNS plays an important part.

Chronic hypoxia at high altitudes Both simple and more complex psychological functions are impaired significantly at high altitudes (>6000 m), including arithmetical tests, writing ability, and the appearance and disappearance of afterimages following exposure of the eye to bright light. Increased memory errors, errors in perseverance, and reductions in auditory threshold and words apprehended are also observed. In a series of now-classic studies, measurements of sensory and motor functions were made at several altitudes. It was found that impairment was not significant below 5300 m. Subjects with longer periods of acclimatization appeared to suffer less deterioration.

The general findings from several such performance studies showed consistently that there is an increased distractibility and lethargy that tend to reduce the subjects' ability to concentrate. These results are consistent with the impression reported by many people who have worked at high altitude, namely that thought processes are slowed but that, if one concentrates hard enough, accurate procedures can be carried out.

Changes in the human EEG have been observed at high altitudes (5500 m). Ryn reported both an α -wave frequency decrease and paroxysmal and focal pathology. In general, there were abnormalities in 11 out of 30 climbers. Partly the same pattern was later found by Zhongyuan and colleagues at altitudes above 5000 m. However, the changes were less than those observed during acute hypoxia at the same altitude simulated in a pressure chamber prior to the expedition. A general finding was that members of the expedition who tolerated the acute hypoxia well tended to show fewer EEG changes on the mountain.

Recent studies on altitude exposure above 6000 m showed an increased susceptibility to hallucinatory

experiences in mountaineers who refused to use supplementary oxygen. The autonomic nervous system seems to play a crucial role in the short-term adjustment to hypoxia. There is increased sympatho-adrenal activity associated with decreased parasympathetic tone, and vasomotor depression due to central hypoxia may occur in nonacclimatized subjects with reduced altitude tolerance at and above 4000 m.

Findings of possible impairment of higher cognitive functions at high altitude are still controversial. Recent studies have more specifically studied what is sometimes called mountaineer's high, which in some ways resembles the syndrome associated with lesions in the frontolimbic loops. The similarity in the patterns of cognitive and emotional deficits after frontal lobe impairment and in the high-altitude syndrome is best accounted for by the fact that the prefrontal and orbitofrontal areas and their connections to the limbic centers are among those most sensitive to acute cerebral hypoxia. When three groups of seven men were exposed to simulated altitudes of 3000 and 4500 m, short-term adaptation mechanisms led to a preservation of selected frontal lobe cognitive and emotional functions. Functional hemispheric superiorities for word recognition and affective processing were likewise unaffected, despite a drop in diastolic blood pressure at 4500 m, indicating a beginning central hypoxia in terms of functional impairment of the vasomotor center.

Residual effects As in diving medicine, those in high-altitude medicine express great interest in the possibility of residual impairment of CNS function following a return to normal pressure. Studies were done of 21 members of an American expedition to Mt. Everest in 1981 prior to the expedition, immediately afterward, and 1 year later using a comprehensive battery of neuropsychological tests. Changes were found for verbal learning and memory, aphasic errors, and fine-motor tempo. Other authors reported similar or consistent findings. Results from a second American expedition in 1985 confirmed the changes found in the 1981 expedition. Furthermore, a simulated altitude exposure similar to the Mt. Everest conditions produced the same impairments in motor speed and persistence, memory, and verbal expression afterward. There was also a significant negative correlation between hypoxic ventilatory response and neuropsychological function measured after the expedition. In other words, subjects with the largest hypoxic ventilatory response showed the greatest neuropsychological decrement. Townes explained this finding by the following hypothesis. Subjects with the highest hypoxic ventilatory response would reduce their arterial pressure of CO_2 the most and

therefore develop the most cerebral vasoconstriction. As a result, this would cause the most severe cerebral hypoxia, even though their arterial pressure of O_2 would actually be higher than in subjects with the smaller ventilatory response to hypoxia. Ward and colleagues found that this correlation between hypoxic ventilatory response and residual impairment of CNS function led to a paradox: the climber who is endowed by nature to go the highest is likely to suffer the most severe CNS damage.

Concluding Remarks

Both diving and high-altitude climbing are potentially dangerous. It requires time and patience to master all the skills necessary to move safely at deep depths and in mountain country. Vigilance and emotional stability are important, and extreme sensation seekers are not as likely to function effectively as more self-sufficient people. Those who are obliged to live harmoniously in close proximity for long periods of time in a high- or low-pressure environment should be stable, loyal, and have both a social and an intellectual tolerance for their companions.

Further Reading

- Behnke, A. R. (1942). Effects of high pressure prevention and treatment of compressed air illness. *Medical Clinics of North America* 26, 1228–1237.
- Behnke, A. R., Thomson, R. M. and Motley, E. P. (1935). The psychologic effect from breathing air at 4 atmospheres pressure. *American Journal of Physiology* 161, 417–425.
- Bennett, P. B. (1982). The high pressure syndrome in man. In: Bennett, P. B. & Elliott, D. H. (eds.) *The physiology and medicine of diving*. London: Baillière Tindall, 262–296.
- Brauer, R. W., Dimov, S., Fructus, X., et al. (1969). Syndrome neurologique et électrographique des hautes pressions. *Revue Neurologique* 121, 264–265.
- Brugger, P., Regard, M., Landis, T., et al. (1999). Hallucinatory experiences in extreme-altitude climbers. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology* 12, 67–71.
- Burnett, C. S. F. (1996). High altitude mountaineering 1600 years ago. *Alpine Journal* 88, 127.
- Cavaletti, G., Morom, R., Garavaglia, P., et al. (1987). Brain damage after high-altitude climbs without oxygen. *Lancet* 1, 101.
- Clark, J. M. (1982). Oxygen toxicity. In: Bennett, P. B. & Elliott, D. H. (eds.) *The physiology and medicine of diving*. London: Baillière Tindall, 200–238.
- Cousteau, J. Y. (1953). *The silent world*. London: Reprint Society.
- Darbin, O., Risso, J. J. and Rostain, J. C. (1999). The full expression of loco motor and motor hyperactivities

- induced by pressure requires both striatal dopaminergic and NMDA receptors activities in rat. *Neuroscience Letters* 267, 149–152.
- Donald, K. W. (1947). Oxygen poisoning in man. *British Medical Journal* 1, 667–672.
- Dutka, A. J. (2003). Long term effects on the central nervous system. In: Brubakk, A. O. & Neuman, T. S. (eds.) *Bennett & Elliott's physiology and medicine of diving*. Edinburgh: Saunders, 680–699.
- Hope, A., Lund, T., Elliott, D. H., Halsey, M. J. and Wiig, H. (eds) (1994). *Longterm health effects of diving: an international consensus conference*. Bergen: Norwegian Underwater Technology Center, 391.
- Værnes, R. J., Hammerborg, D., Ellertsen, B., et al. (1983). Central nervous system reactions during heliox and trimix dives to 51 ATA, Deep Ex 81. *Undersea Biomedical Research* 10, 169–192.
- Værnes, R. J., Kløve, H. and Ellertsen, B. (1989). Neuropsychologic effects of saturation diving. *Undersea Biomedical Research* 16, 233–251.
- Værnes, R. J. and Sandal, G. (2003). Human reactions to deep-water conditions. *Lancet* 362, 10–12.
- Ward, M. P., Milledge, J. S. and West, J. B. (1989). *High altitude medicine and physiology*. London: Chapman and Hall.

Primate Hierarchies and Personality

R M Sapolsky

Stanford University, Stanford, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R M Sapolsky, volume 3, pp 221–225,

© 2000, Elsevier Inc.

Background: Primate Sociality
Social Rank and Stress Physiology
Modifiers of the Rank–Physiology Relationship
Hierarchies, Humans, and Stress-Related Disease

Glossary

<i>Primate personality</i>	Stable patterns of behavioral style, emotional responsiveness and social roles.
<i>Dominance hierarchies</i>	The collectivity of an array of reasonably stable asymmetries in dyads of animals is competing for limited resources.
<i>Psychosocial stressors</i>	Stressors, in a social context, in which there is either anticipation of a challenge to homeostasis (without it having yet occurred) or in which there is no threat of a challenge whatsoever.

Whether one is a professional zoologist or a casual observer of animals in a zoo or on a public television special, few can resist the pull of non-humans primates. Part of the fascination with them is due to their intelligence, and part to their unnervingly close relationship to us. Perhaps the strongest pull, especially when considering the apes and the Old World monkeys, is the force of their individuality – these are

long-lived organisms with strong, pungent personalities and temperaments. For the psychobiologist, an irresistible question is whether there are physiological correlates of these individual behavioral differences and, if such physiological differences exist, whether they are causes or consequences of the behavior.

Stress physiologists have examined this question for a long time, framing it in the context of the stress response and stress-related disease. Do individual differences in primate behaviors, in social rank within the dominance hierarchy, or in personality predict differences in the quality of the stress response? Do they predict differences in patterns of stress-related disease? The answers to these questions are complex, and the very complexity is what offers the greatest insight into the human condition. To best appreciate the complexity, a brief overview of primate social systems will be helpful.

Background: Primate Sociality

As befits our close phylogenetic relatives, the primate order is arguably the most flexible, varied, and catholic of any in the animal kingdom. Primates can be found in ecosystems ranging from the most severe of deserts in Africa to snowy regions of north India and Japan, from thick rainforests to open savanna, from flat plains to steep mountainous regions. They are equally varied in their diet, ranging from strict vegetarians to organized hunters. Given this heterogeneity, it is not surprising that there is an enormous range of primate social systems, ranging from solitary, to small related groups built around male–female pair bonds, to fusion/fission societies involving hundreds of individuals with complex patterns of

polygamy and polyandry. Nonetheless, there are some themes common to most primate species.

First, with the exception of the few rare solitary primate species (such as the orangutan), these are profoundly social animals; the behavior of most primates outside the context of their group makes little sense. As but one measure of the importance of sociality, across the more than 150 primate species, the size of the cortex (relative to total brain weight) positively correlates with the size of social groups in that species. Second, most primate societies have exogamy of one sex, that is, the adolescents of one sex leave the group at puberty, transferring to another group as an inbreeding-avoidance mechanism. Among chimpanzees, for example, it is females who transfer, whereas it is males among the baboons and macaques. As a result, the gender that does not transfer spends its entire life in its natal group, surrounded by same-sex relatives and its mother. Thus, within any such primate society, it is impossible to understand the behavior of one of the sexes without considering the complexities of interactions among kin – cooperative coalitions among chimpanzee brothers or baboon sisters, aunting behavior by female macaques that increases the likelihood of a sister's child surviving, and so on. Finally, even in the most bountiful of ecosystems, resources are not infinite and, in most primate societies, are typically divided up unevenly along lines of conventionalized asymmetry referred to as dominance hierarchies. In a stable dominance hierarchy, limited resources are not contested with bloody tooth and claw. Instead, animals are smart enough to understand the status quo of the dominance hierarchy, allowing it to substitute for overt competition. Accordingly, the social rank of an animal affects its quality of life to an enormous extent.

The phenomenon of dominance hierarchies has long fascinated primatologists, as it can be the most immediate and dramatic version of individual differences in the behavior of these animals. Thus, when stress physiologists became interested in the psychology of individual differences among primates, they first focused on social rank.

Social Rank and Stress Physiology

The earliest theorizing in the field was badly derailed by the influence of the executive stress syndrome studies of the 1950s. This work seemed to suggest that monkeys with executive control over resources were more prone toward ulcers than monkeys without control, and was widely interpreted as evidence of the increased stressfulness of dominance. These findings were subsequently shown to be spurious, due to

nonrandom selection of subjects. Specifically, in order to facilitate the learning of the operancy performed as the executive task, animals were prescreened, and the most emotionally reactive ones were made executive animals.

Finally, freed from these erroneous findings, primatologists interested in the rank–health relationship turned their attention to rodent studies, which were pioneered by James Henry. These produced a fairly consistent picture. Subordinate animals (as typically detected by the outcome of paired fights) showed indices of chronic stress, including elevated circulating glucocorticoid levels, hyperplastic adrenals, immune suppression, reproductive problems, and hypertension. As studies began with non-human primates, it was expected that the same would be seen. This was reinforced by insights into the psychological variables that modulate the stress response. Specifically, classic studies by John Mason, Jay Weiss, Seymour Levine, Martin Seligman, and colleagues showed that the same physical stressor would trigger a larger or more prolonged stress response if the organism lacked a sense of control, lacked a sense of outlets, lacked predictive information about the stressor, or interpreted events as worsening.

Studies of an array of primate species demonstrated that lower-ranking individuals in stable dominance hierarchies disproportionately suffer those psychological stressors. They lack control and predictability, for example, potentially spending minutes obtaining a food item that can be seized from them, finding a grooming partner only to have the bout disrupted by someone else, being subject to unpredictable displacement aggression if someone else was having a bad day. Moreover, they typically lack the means to release frustration by displacing aggression onto someone else.

This generated the expectation that subordinate primates would have a maladaptive profile that would include stress responses that were hyperactive basally or sluggish in recovering from a stressor, and greater vulnerability to stress-related disease. This was precisely what was reported, during the 1980s and 1990s. Work by, among others, Jay Kaplan, Carole Shively, Deborah Gust and colleagues with macaques, Robert Sapolsky with baboons (a wild population, in this case), Barry Keverne and colleagues with talapins, Kirk Manogue and colleagues with squirrel monkeys, and Eberhart Fuchs and colleagues with tree shrews demonstrated that subordinate individuals displayed elevated basal glucocorticoid levels and negative feedback insensitivity, elevated resting blood pressure, a sluggish cardiovascular response to a stressor and a sluggish recovery, suppressed levels of good (i.e., HDL) cholesterol, impaired fertility and a

reproductive axis more readily suppressed by stress, fewer circulating white blood cells, an enhanced risk of atherosclerosis, increased endogenous benzodiazepine signaling, and less genesis of new neurons in the brain. Furthermore, when group membership was being manipulated among captive animals, and physiological measures were taken before and after formation of the group (and thus of the dominance hierarchy), it was clear that the physiological differences arose as a consequence of the rank differences. Finally, an array of subtle studies showed that these stress-related indices were far more a function of the psychosocial aspects of subordination, rather than one of the overt physical insults of rank.

These findings were readily interpreted as indices of the stressfulness of social subordination. A number of individuals carried out detailed studies uncovering the reductive mechanisms explaining these patterns. For example, Kaplan and colleagues have done detailed studies demonstrating that the stress-induced suppression of estrogen levels in subordinate female macaques helps contribute to the increased risk of atherosclerosis. Sapolsky and colleagues have shown that basal hypercortisolism of the subordinate baboons arises for many of the same reasons that hypercortisolism does in human depression (i.e., hypersecretion of corticotropin-releasing hormone [CRH] at the hypothalamic level, a partial compensation at the level of the pituitary via blunted sensitivity to CRH, and feedback resistance in the form of insensitivity to dexamethasone).

Despite this seeming clarity, major contradictions to this picture were emerging. David Abbott, working with marmoset and tamarin monkeys, and Cavigelli, working with feral lemurs (as well as Scott Creel, studying an array of non-primate species, including African wild dogs and dwarf mongooses), were reporting diametrically opposite findings, namely, that it is dominant individuals that show the indices of overactivated stress responses.

By the mid-1990s, a widespread impression was emerging that there was, in fact, little consistent relationship between primate social status and stress-related disease. The most recent work in the field has shown that this is not the case. Instead, the rank–health relationship can be modified dramatically by an array of additional factors.

Modifiers of the Rank–Physiology Relationship

What Is the Meaning of Rank in a Species?

Among the macaques and baboons, subordination is a state forcibly imposed from above by way of

aggression or harassment by dominant individuals; resources are divided with marked inequity along the lines of rank. Such despotic dominance systems generate considerable stress for subordinates. The situation for marmosets and tamarins, in contrast, is quite different. Among these New World monkeys, social groups of six to eight animals consist of a pair-bonded dominant pair and a number of subordinate individuals. Critically, these social groups are typically extended families of relatives, and successful reproduction involves cooperativity among them. A subordinate animal is usually a younger relative waiting her turn and helping older relatives with child-care in the interim, and is rarely subject to displacement aggression. This goes far to explain Abbott's reports of dominant animals showing the preponderance of stress-related physiology. One meta-analysis examined all studies of rank–basal glucocorticoid relationships in non-human primates (seven species, both genders), ranging from cases in which subordinate animals have far lower basal levels than dominant animals (marmosets and tamarins) to those in which subordinates have far higher (talapoin and baboons). Across these species, a significant predictor of such relative glucocorticoid levels was the degree to which subordinates were subject to the harassment and displacement aggression typical of despotic dominance systems.

How Is High Rank Maintained in the Particular Primate System?

In many primate social systems, rank can change over time, raising the question of how a high-ranking individual holds on to that rank as long as possible. In some cases, this involves frequent and overt aggression aimed at subordinate individuals, with the result that dominant animals are involved in the highest rates of fighting. In contrast, in other cases, well-entrenched dominant animals have the lowest rates of aggression, maintaining their dominance through little more than subtle psychological intimidation. A number of studies now demonstrate that in the former case, subordination is not associated with elevated basal glucocorticoid levels, whereas in the latter it is.

How Stable Is the Dominance Hierarchy?

In many primate social systems, rank is inherited at birth, remaining static throughout life. However, even in systems in which rank shifts over time, there can be long periods of hierarchical stability (i.e., where the dominant individual in any given dyad wins the vast majority of dominance interactions, thereby reinforcing the status quo). In such cases,

dominant individuals have the abundant physical and psychological advantages of their lofty rank. Occasionally, hierarchies will become unstable (due to some key individual entering or leaving the troop, a death or injury, or the formation or dissolution of a key coalition), and there will ensue an unstable period of dramatic, frequent shifts in rank until the hierarchy restabilizes. This period will produce high rates of aggression and competitive interactions among individuals, unpredictable formation and disintegration of coalitions, and low rates of socially affiliative behaviors. Under such circumstances, it is high-ranking individuals, at the center of the competition, who have the least control and predictability over events around them. As shown by numerous investigators, during such unstable periods, it is no longer the subordinate individuals who have the highest basal glucocorticoid levels or the greatest risk of developing atherosclerosis.

What Coping Strategies Are Available to Subordinate Animals?

The magnitude of the stress response is a function of the severity and frequency of stressors to which an individual is exposed and, just as importantly, the availability of coping outlets. In primates, some coping outlets are positive (grooming, being groomed, physical contact with another animal) and others are negative (displacing aggression on to another animal). Across the primate species studied, the more such outlets are available (both positive and negative), the lower the basal glucocorticoid levels. There are not sufficient data, however, to determine whether positive or negative outlets are more physiologically protective.

How Readily Can Subordinate Animals Avoid Dominant Ones?

In a rather simple and intuitive way, subordinate animals who can readily evade the attention of dominant individuals are likely to be subject to fewer stressors. This point helps explain some seeming contradictions in the literature. For example, studies of wolves have shown that in captivity (where subordinate individuals are markedly constrained in their ability to avoid dominant animals), it is subordinate animals who have the highest basal glucocorticoid levels. In contrast, in feral wolf populations, this is not the case.

What Is the Social Milieu of a Particular Primate Group?

Primate social systems not only vary between species, but also vary among different populations of the

same species. In some cases, this reflects ecological or genetic issues, but often, it reflects a fluke of demographics of the group and the personalities of its members. As such, the social milieu of a particular group can modify the rank–health relationship. Work by Gust and by Sapolsky has focused on social systems (of female macaques and male baboons, respectively) in which subordination is typically associated with elevated basal glucocorticoid levels. However, this is not observed in troops with atypically low levels of displacement aggression on subordinates and/or high levels of social affiliation.

What Is an Individual Animal's Experience of Life in Its Social Group?

While the generic features of a primate's social group (e.g., rates of reconciliation, extent of hierarchical stability) may be a significant predictor of stress-related physiology, so is the personal experience of those variables. For example, in a pair of studies of macaques, the severity of basal hypercortisolism varied as a function of how often individual animals were subject to dominance or aggressive interactions, or how often they were given affiliative support. Similarly, in a study of wild baboons during a period when the hierarchy was destabilized by a highly aggressive transfer male, the frequency with which that male attacked females was highly predictive of the extent to which they were immunosuppressed. Strikingly, in that study, females who had not been subject to such displacement aggression showed no evidence of immune suppression. Thus, these stress-related physiological markers do not respond to the abstract features of life in a particular troop (i.e., this is a troop with high rates of reconciliation/displaced aggression/coalitional fighting ...), but to the individual's very concrete personal experience of events in the troop.

What Is the Personality of the Animal in Question?

One of the most striking factors influencing individual differences in stress-related physiology among primates is personality. The notion of personalities among non-human primates was long ignored in many circles of primatology, being viewed as unacceptably anecdotal and anthropomorphic. Recent and rigorous work has shown there to be consistent and stable differences among individuals in their temperament, patterns of impulsivity, tendencies toward sociality, the social roles they will assume within the group, and so on. With those findings has come an interest in whether those differences predict different profiles of stress-related physiology.

As one example, Stephen Suomi and colleagues, Ned Kalin, as well as Kaplan, Steven Manuck, and colleagues reported stable personality differences in macaque monkeys in their extent of reactivity to novelty. Such traits are apparent very early in life, and the high reactor animals whose behaviors are more disrupted by novelty are hypercortisolemic throughout life and more prone toward atherosclerosis.

In another series of studies, Sapolsky and colleagues examined personality differences among wild baboons, after controlling for rank. Two clusters of traits predicted elevated basal cortisol levels. The first related to patterns of male–male competition; hypercortisolism was observed among individuals who were least capable of differentiating between threatening and neutral interactions with rivals (i.e., they responded to both as if they were provocative), were least likely to control the initiation of a fight, and were least capable of differentiating between winning and losing fights (i.e., both outcomes equally disrupted their social behavior thereafter). A second, independent cluster demonstrated hypercortisolism among the males with the lowest rates of social grooming or contact with females, or play with infants. These findings were interpreted as showing greater endocrine indices of stress among animals with the least amounts of control and predictability about stressors and the fewest sources of social coping outlets.

Hierarchies, Humans, and Stress-Related Disease

A number of investigators have examined human hierarchies in the hopes of correlating rank differences with differences in stress physiology. Outcomes of athletic competitions, military ranks, and position within a corporate workplace have all been considered as examples of human hierarchies, and differences in physiology have been reported. Other investigators have criticized these studies due to artificiality of these rankings (e.g., how much can rankings in a college wrestling tournament tell us about cardiovascular disease susceptibility?), the confounding ability of humans to be in more than one hierarchy (i.e., the low-ranking individual working in the mailroom of a corporation might, at the same time, be captain of the company softball team and consider that to be the more relevant hierarchy), and the human capacity for rationalizing and externalizing what would be perceived as being a low rank.

Some investigators have suggested that low socioeconomic status (SES) can be viewed as the most valid and pervasive example of low rank in a human, carrying with it many of the psychological stressors seen

among many low-ranking non-human primates (lack of control, predictability, or outlets). Low SES carries with it an enormously increased risk of a broad range of diseases, and careful study by pioneers such as Michael Marmot, Richard Wilkinson, Nancy Adler, and Ichiro Kawachi has shown that this SES gradient can not be explained by differences in health-care access and can only be minimally explained by differences in education or in exposure to risk factors and protective factors. In light of those findings, some researchers have suggested that part of the SES gradient of disease can be viewed as reflecting the psychosocial stressors of poverty. In support of this is not only the fact that some of the alternative explanations can be ruled out to some extent (e.g., a striking bit of evidence against the idea that the SES gradient is due solely to differential health-care access is the robust presence of such gradients in countries with universal health care), but also the fact that diseases that are probably the most stress sensitive (psychiatric disorders and cardiovascular disease) show the steepest gradients.

See Also the Following Articles

Health and Socioeconomic Status; Primate Models, Overview; Social Networks and Social Isolation; Social Status and Stress; Social Stress, Animal Models of.

Further Reading

- Abbott, D., Keverne, E., Bercovith, F., et al. (2003). Are subordinates always stressed? A comparative analysis of rank differences in cortisol levels among primates. *Hormones and Behavior* **43**, 67.
- Clark, A. and Boinski, S. (1995). Temperament in non-human primates. *American Journal of Primatology* **37**, 103.
- Creel, S. (2001). Social dominance and stress hormones. *Trends in Ecology and Evolution* **16**, 491.
- Gould, E. and Gross, C. (2002). Neurogenesis in adult mammals: some progress and problems. *Journal of Neuroscience* **22**, 619–623.
- McEwen, B. (2002). Sex, stress and hippocampus: allostasis, allostatic load and the aging process. *Neurobiology of Aging* **23**, 921.
- Sapolsky, R. (1999). Stress, glucocorticoids and their adverse neurological effects: relevance to aging. *Experimental Gerontology* **34**, 721.
- Sapolsky, R. (2000). Glucocorticoids and hippocampal atrophy in neuropsychiatric disorders. *Archives of General Psychiatry* **57**, 925.
- Sapolsky, R. (2005). The influence of social hierarchy on primate health. *Science* **308**, 648.
- Suomi, S. (1997). Early determinants of behaviour: evidence from primate studies. *British Medical Bulletin* **53**, 170.

Primate Models, Behavioral-Immunological Interactions

M L Laudenslager and S Tiefenbacher

University of Colorado Health Sciences Center Denver,
CO, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition
article by M L Laudenslager, volume 3, pp 226–229,
© 2000, Elsevier Inc.

Animal Models
The Immune System
Prenatal Experiences Modify Postnatal Immune
Competence of the Offspring
Postnatal Experiences Have Acute and Long-Term
Consequences
Social Relationships Affect Immunoregulation
Individual Differences in Immune Regulation
Environmental Factors Influencing Immunoregulation
Mechanisms of Brain-Immune Interactions
Immune Senescence in Nonhuman Primates
Summary

Glossary

<i>Acquired immunity</i>	The components of the immune system that require time (several days) to respond to antigens and are highly specific for an antigen.
<i>Antibody</i>	A protein secreted by B cells that binds to antigens and leads to their destruction and removal, also called immunoglobulin.
<i>Antigen</i>	Any substance that is recognized by the immune system as not belonging to the body (not self); includes viruses, bacteria, tumors, and tissue grafts.
<i>Cytokines</i>	Protein messengers made by a variety of cells including lymphocytes and phagocytes that regulate the immune response and affect central nervous system activity.
<i>Innate immunity</i>	The components of the immune system that respond immediately in the presence of antigens as a rapid nonspecific response, a first line of defense; includes neutrophils, natural killer T cells, and macrophages.
<i>Lymphocytes</i>	The cells of the immune system, including B cells (make antibodies), T cells, (regulate the immune response and destroy foreign cells), and natural killer

(NK) cells (first line of defense against tumors and virally infected cells; also regulate the immune response).

Phagocytes

The cells of the immune system that process antigens in preparation for activating the immune response; include neutrophils and macrophages.

Psychoneuro-immunology

An interdisciplinary science that examines bidirectional interrelationships among immune, neuroendocrine, and behavioral processes

Simian immunodeficiency virus (SIV)

A virus found only in monkeys that is the best model for human immunodeficiency virus (HIV) available to scientists.

In this article, the utility of animal models is described with a focus on nonhuman primates, followed by a simple overview of the immune system. The area of immunology is growing by leaps and bounds, and as such it is elusive with regard to any review that serves the entire discipline with any justice. The growing area of psychoneuroimmunology (PNI) suggests that there are important regulatory relationships between the brain and the immune system and furthermore that these relationships are bidirectional. The study of these relationships in nonhuman primates is of particular importance for improved understanding brain immune relations in humans due to the great physiological and behavioral similarity of the two groups. The following discussion focuses on several areas of investigation that have implications for immune and behavior relationships in nonhuman primates: prenatal and postnatal experiences, social relationships, and aging.

Animal Models

Animal models are used in a variety of biomedical studies because they permit the control of a number of factors not possible with human subjects. An adequate animal model of any human condition must fulfill several criteria, including a common phenomenology, etiology, underlying pathophysiology, and effective treatment. The evolutionary proximity of humans and nonhuman primates (e.g., monkeys and apes) makes the nonhuman primate a frequently studied model for both human health and disease. The most obvious similarity to humans resides in the

social nature and behavioral characteristics of non-human primates. In addition, nonhuman primates offer a unique opportunity for studying brain-immune relationships. It is likely that the phenotypic traits shared by humans and nonhuman primates are homologous, that is, that they are related to a common ancestral origin. Thus, they may share similar design features. These similarities are important for disentangling complex biomedical problems.

The Immune System

The immune system is a highly regulated system with numerous redundancies and backups. Simply put, the responsibilities of the immune system can be characterized by the three Rs: recognition, removal, and regulation. The immune system must recognize or differentiate self from not self (antigens or foreign substances) and respond appropriately. The antigen might be a simple microorganism located in a cut to a tumor developing in the lungs. Once an antigen is recognized as not self, it must be removed or destroyed by host defense mechanisms. This very complicated process is regulated by immune cells that communicate via complex proteins called cytokines, endocrine hormones, and neurotransmitters derived from neural structure. There are innate processes that respond immediately to the antigen until the acquired system can respond to the foreign intruder. The ability of natural killer (NK) cells to immediately destroy tumors or virally infected cells illustrates the actions of the innate immune system. The presence of specific antibodies protecting the individual against flu, following the administration of a flu vaccine in the fall, exemplify the actions of the acquired immune response. Acquired immunity is a slower and highly specific process; flu protection after vaccination does not occur until approximately 14 or more days after a person receives the vaccine. This response is also a highly integrated response of the immune system.

These systems are not independent but act together closely to control infection and disease. The innate system forms a first line of defense until the acquired system can develop responses specific for the pathogen. It might take from several days to weeks before the acquired response is at full capacity. The innate and the acquired response are regulated by the actions of cells of the immune system, including lymphocytes (T cells, B cells, and NK cells) and phagocytes (neutrophils, macrophages, and eosinophils). T cells, NK cells, and macrophages secrete a number of protein messengers (cytokines) that are critical in the regulation of the response. The capabilities of the immune system can be measured in either tissue culture

systems (*in vitro*) or in the intact organism (*in vivo*). *In vivo* measurements are preferred over *in vitro* assessment because the system is studied under naturally occurring conditions. For example, specific antibodies seen after immunization with a foreign protein represent the final output of a highly integrated response of the immune system. The antibodies in the blood reflect the initial recognition and processing of the antigen by phagocytic cells; release of regulatory cytokines by T cells, NK cells, and macrophages; and interactions with neural and hormonal factors resulting in the final production of antibody by the B cells. One of the first reports on stress in nonhuman primates described the suppressive effects of unpredictable and generally poor housing conditions of monkeys on their antibody levels after immunization.

Prenatal Experiences Modify Postnatal Immune Competence of the Offspring

The prenatal environment contributes significantly to postnatal neurobehavioral developmental and immune competence and as such is a significant contributor to individual differences in the overall regulation of the hypothalamic-pituitary-adrenal (HPA) axis and immune system. The mother and the developing fetus form a highly regulated amalgamation that has received considerable attention with regard to the fetal origins of later medical conditions in the offspring. Stress in the mother is synonymous with stress in the developing fetus. More specifically, prenatal stressors (behavioral and/or pharmacological, e.g., dexamethasone or adrenocorticotrophic hormone, ACTH) of pregnant rhesus monkeys are associated with a cascade of postnatal effects in the fetus. The power of prenatal stressors is indicated by the fact that they are sufficiently pervasive to redirect the gender-specific development of the offspring. Although the offspring are healthy and normal at birth, the effects of prenatal stress are reflected in a dysregulation of immune responses and a more responsive HPA axis in infants born to these prenatally challenged mothers. Immune effects include suppressed cytotoxic-like activity of lymphocytes, depressed proliferative responses, and reduced cytokine production by lymphocytes. The timing of the stressor during gestation (e.g., early vs. late) affects the direction of the specific impact on immunoregulation for some of these measures. That is, early gestational stressors are associated with enhanced mixed lymphocyte responses (MLR) responses toward syngeneic cells from the mother or father, whereas late gestational stressors are associated with reduced MLR to the same targets. Thus, the nature of prenatal stressors is complicated but nonetheless powerful. Prenatal stressors impact placental transfer from

mother to fetus, affecting the transfer of maternal immunoglobulins and iron. These effects lead to reduced levels of beneficial bacteria in the gut of the offspring.

Postnatal Experiences Have Acute and Long-Term Consequences

Early experience effects are not limited to events during the prenatal period but extend to the postnatal period as well. Behavioral challenges are similar in their impacts on later immune and HPA regulation. For example, significant life challenges such as brief maternal separations that might occur during the breeding season when mothers leave their offspring to mate with males in the troop, matrilineal dominance status, or the quality of maternal care have significant effects on the immune regulation in nonhuman primates, both in the field and in laboratory settings. Wild-born offspring of low-ranking matriline, for example, show greater immune and HPA perturbations to acute challenges as adolescents. In the laboratory, social-separation experiences such as the removal of the mother from the infant for brief periods of time produce a number of disturbances in both *in vitro* and *in vivo* immune parameters. Acutely, *in vitro* immune measures such as NK activity and activity of macrophages are elevated, whereas other markers are suppressed. Indeed, specific antibody levels following immunization at the time of separation are reduced and the magnitude of this reduction is greatest in the most distressed monkeys. These early separation experiences have also been shown to impact immune regulation over an extended time period, such as reduced *in vitro* markers of the T-cell function and an elevation in markers of the innate system, namely NK activity and the activity of the macrophage system.

Postnatal pharmacological challenges such as antidepressant drugs administered to young monkeys can have long-lasting effects on immune competence, resulting in lower specific antibody responses to antigenic challenge and increased reactivity of the HPA axis. Altered HPA regulation is one of many contributing factors to dysregulated immune function. The overarching observation is that, if something stressful happens to the mother-infant unit prenatally or the infant postnally, it is very likely to result in a disruption of the immune system, sometimes in subtle but important ways. These observations also indicate the utility of nonhuman primate models, with their shortened life cycle and the researchers' ability to control environmental exposure to pathogens that will surely impact the developmental trajectory. This makes the nonhuman primate model ideal for studies of aging.

A critical question facing the field of psychoneuro-immunology is the interpretation of the functional significance of these *in vitro* measures. As addressed earlier, *in vivo* measures are most likely to reveal important consequences mediated by immune activity that might not be indicated by *in vitro* measures. The implications of early challenges and stressor experiences are best exemplified by studies of monkeys infected with SIV, a nonhuman primate model for acquired immunodeficiency syndrome (AIDS). There is considerable variability in the disease course of AIDS after infection with human immunodeficiency virus (HIV). A significant question for understanding the pathophysiology of this virus is: Why do some individuals infected with HIV progress rapidly and become quite ill and others infected with the virus remain asymptomatic? In retrospective studies of rhesus monkeys infected with SIV, researchers found that the early developmental history of the monkeys contributed significantly to disease course. Thus, monkeys that experienced repeated housing changes early in development or were reared with peers rather than their mother showed a more rapid progression in indicators of virus activity. Furthermore, social reorganization stress at the time of viral inoculation was also associated with reduced survival. Moreover, recent studies from our group also observed that monkeys who progressed most rapidly following inoculation with SIV had reduced activity of cytotoxic T-cell (CD16+ lymphocytes) function before exposure to the virus and independent of early experiences. Although, many potential influences of early experiences on both immunoregulation and disease resistance in adult monkeys have been demonstrated, recent studies also indicate that the risk for rapid progression may rest in genetically determined aspects of immunoregulation as well.

Social Relationships Affect Immunoregulation

We have learned the most regarding the relation between behavior and immunity from observations of the negative consequences of stressors. Social psychology has shown for many years that the presence of social support in humans has positive consequences for health and well-being. There is growing evidence in nonhuman primates that a number of aspects of social relationships can also have positive consequences for markers of health and immunity. Most nonhuman primates are social creatures with their lives organized around social interactions with other members of their group. During times of social challenge, the presence of another monkey is important for reducing the negative endocrine and immune

consequences of the stressor. The negative consequences of maternal absence on behavior and physiology are significantly reduced in the presence of other monkeys who interact with the separated monkey. These social interactions may include affiliative behaviors, such as sitting nearby, play, and/or grooming. Being groomed by another monkey is quite effective in reducing heart rate to levels comparable to those noted during sleep. Grooming in monkeys strengthens social ties and facilitates reconciliation following fights between monkeys. Heart rate recovers rapidly following an aggressive interaction if one of the monkeys is groomed by the other involved in the fight. The immediate immune consequences of these behaviors have not been determined.

The relationship between dominance status in nonhuman primates and immune regulation is not as clear as we might expect. A high social rank does not necessarily imply enhanced immune capacity. Immune status assessed by *in vitro* methods in monkeys with high social status does not differ from low-ranking monkeys under stable social conditions nor do acute short-term stressors modulate these measures in a manner predicted by social status in the group. However, if the specific antibody responses to an antigen, an *in vivo* measure, are assessed under stable conditions, low-ranking monkeys mount a larger response than high-ranking monkeys. High rank may carry with it some subtle negative consequences. It appears that specific affiliative behaviors are most critical in modulating the impact of stressful social conditions in nonhuman primates, not unlike the role of social support for humans. Furthermore, the assumption that stress suppresses immune regulation is no longer entirely correct because acute stressors and challenges can enhance aspects of the inflammatory response.

Individual Differences in Immune Regulation

Monkeys differ along a number of behavioral and physiological dimensions. The concept of individual differences has received more attention from researchers in PNI. In the face of considerable variability in outcome following seemingly similar challenges, we may observe a wide range in the levels of immune measures even under resting conditions. Epidemiological research has tended to focus on demographic characteristics such as gender, age, socioeconomic status (SES), race, and so on in partitioning outcome variability. The long-term consequences of these factors have received considerable attention in a concept called allostatic load, the cumulative effects of long-term stressors that ultimately push the adaptational

capacity of the stress response system beyond its capacity. However, there are individual factors that contribute to the extent to which these stressors affect immunoregulation.

Recent observations in monkeys have suggested that behavioral characteristics or traits unique to each monkey are influential in determining these outcomes. For example, sociability, or the tendency to engage in affiliative interactions with other monkeys, is associated with more rapid clearance of the virus in SIV-inoculated monkeys. These findings support the observations that social affiliation reduces the negative consequences of social challenge. Under free-ranging conditions, monkeys characterized as irritable (e.g., active, irritable, and aggressive) tended to have higher active infection with cytomegalovirus (CMV), a virus common in these animals. Thus, in addition to demographic characteristics, individual differences in behavioral traits also contribute to immunological variance. Other aspects of traitlike behaviors on nonhuman primates (e.g., personality) affect immune regulation to the extent that animals' relative sociality in social groups affect their specific antibody response following relocation. Monkeys rated as low in sociability (e.g., motivation to seek social interactions with others) had reduced specific antibody responses during this challenge.

Environmental Factors Influencing Immunoregulation

The complex nature of nonhuman primates presents the researcher with a number of challenges to provide for their psychological well-being. These manipulations are strictly monitored by federal (U.S. Department of Agriculture, USDA) and private (Association for Assessment of Laboratory Animal Care, AALAC) agencies. The simplest way of achieving this goal is through social housing and avoiding single-individual housing of nonhuman primates. Only a limited number of studies have actually investigated the impact of environmental enrichment on immune responses *per se*. A large number of studies, however, have investigated the impact of environmental enrichment on social behaviors of macaque monkeys. The provision of foraging enrichments (e.g., foraging boards and small preferred food treats mixed into the bedding) and manipulative devices such as simple toys significantly reduces the number of stereotypical behaviors and increase positive social interactions in nonhuman primates. As mentioned earlier, unpredictable housing experiences are associated with immune dysregulation as reflected in an attenuated specific antibody responses. Observations in SIV-infected macaques also indicate that the number of cage changes

occurring in advance of inoculation also impact the response to viral inoculation; that is, numerous cage moves, which represent a significant social stressor, are associated with more rapid progression and fatality to the virus. Features of housing for research animals that were neglected in the past have also been shown to impact experimental outcomes in significant ways.

Mechanisms of Brain-Immune Interactions

If behavior (i.e., the central nervous system) is likely to affect the regulation of the immune response, there must be potential pathways for these interactions. Indeed, there are two major means of brain and immune communication: direct neural innervation of immune organs by the autonomic nervous system and soluble blood-borne factors (i.e., cytokines and endocrine hormones). Nerve fibers of the autonomic nervous system come in close proximity to immune cells in lymphoid organs such as the spleen, thymus, and lymph nodes. Stress hormones such as cortisol are present in the blood during times of environmental and social challenge. Both of these routes, activated during stress, provide the means for neural-immune interactions. In addition, the cells of the immune system possess receptors for the neural and hormonal systems that permit the various neurotransmitters and hormones to affect immune regulation.

This is not a one-way street; the actions and products of the immune system itself also influence the central nervous system. A major component of illness is the associated changes that include sickness behaviors such as fatigue, sleep changes, changes in cognition, and loss of appetite. In many respects, these behaviors associated with cytokine release parallel major depressive disorders, and it has been argued that sickness behavior is one of many factors that underlie depression in clinical populations. Only a few studies have approached sickness behavior in primate models, but cytokines such as interleukin (IL)-1 and interferon- α have been observed to produce sleeplike behavior, advance rapid eye movement (REM) onset, and activate the HPA axis. These effects are consistent with observations in rodent models of sickness behavior and suggest that similar relationships to behavior exist in nonhuman primates as well.

Immune Senescence in Nonhuman Primates

It is generally assumed that aging is associated with immune downregulation. However, observations in successfully aging seniors have indicated that this

may not universally be the case. As a group, older monkeys demonstrate reduced *in vitro* immune responses, such as lymphocyte proliferation and natural cytotoxicity. Functional measures indicate that, compared to a younger monkey, the older monkey's response to challenge with a trivalent flu vaccine is impaired. The extent to which cytotoxic responses are maintained with aging is an important predictor of long-term survival. It has been pointed out that the selection of old animals from a colony creates an inherent bias because they are really only members of a survivor cohort. In general, evidence supports the natural cytotoxic response as a potential proxy marker for successful aging and viral suppression in both old and younger animals.

Interestingly, feeding and exercise habits contribute significantly to the relationship between aging and immune senescence. For almost a century, it has been recognized that caloric restriction (not to be confused with malnutrition) is associated with a number of health benefits, including reduced morbidity and mortality in rodent models. A study that began in the late 1980s fed two groups of rhesus monkeys nutritionally fortified but either calorically restricted or *ad libitum* diets. As these monkeys now approach their senior years, a number of studies are underway to evaluate the impact of this restricted diet on brain structure, neuroendocrine regulation, and immune competence. Immune assessments have indicated attenuated age-related changes among the calorically restricted monkeys in many parameters, including *in vitro* cellular proliferation, cytokine levels, *in vitro* cytokine production, and cytokine gene expression, most notably in interferon- γ . These initial results suggest that some of the untoward aspects of aging are potentially attenuated in association with dietary restriction.

Summary

Nonhuman primates provide excellent models for research in PNI. The complexities of their behavioral interactions serve as an outstanding model for behavior and immune interactions relevant to humans. Early experiences affect immune regulation beyond the developmental period. Negative events experienced early in development contribute to increased risk for disease; however, social factors can reduce the impact of many of these negative experiences. Individual differences in behavioral characteristics also contribute to differences in immune regulation. There are abundant means whereby the immune system and the brain are able to interact. Understanding of these interactions poses a significant challenge for PNI in the twenty-first century.

Further Reading

- Ader, R., Felten, D. L. and Cohen, N. (eds.) (2001). *Psychoneuroimmunology* (3rd edn.) San Diego, CA: Academic Press.
- Capitanio, J. P. and Lerche, N. W. (1998). Social separation, housing relocation, and survival in simian AIDS: a retrospective analysis. *Psychosomatic Medicine* **60**, 235–244.
- Coe, C. L. and Lubach, G. R. (2005). Prenatal origins in individual variation in behavior and immunity. *Neuroscience and Biobehavioral Reviews* **29**, 39–49.
- Fleshner, M. and Laudenslager, M. L. (2004). Psychoneuroimmunology: then and now. *Behavioral and Cognitive Neuroscience Reviews* **3**, 114–130.
- Glaser, R. and Kiecolt-Glaser, J. K. (2005). Stress-induced immune dysfunction: implications for health. *Nature Reviews: Immunology* **5**, 243–251.
- Johnson, R. W. (2002). The concept of sickness behavior: a brief chronological account of four key discoveries. *Veterinary Immunology and Immunopathology* **8**, 443–450.
- Kemeny, M. and Laudenslager, M. L. (eds.) (1999). *Special issue on individual differences. Brain, Behavior, and Immunity* **13**(2).
- McEwen, B. S. (1998). Stress, adaptation, and disease: allostasis and allostatic load. *Annals of the New York Academy of Sciences* **840**, 33–44.
- Roitt, I., Brostoff, J. and Male, D. (2001). *Immunology* (6th edn.). Philadelphia: Mosby.
- Roth, G. S., Mattison, J. A., Ottinger, M. A., et al. (2004). Aging in rhesus monkeys: relevance to health intervention. *Science* **205**, 1423–1426.
- Schapiro, S. J. (2002). Effects of social manipulations and environmental enrichment on behavior and cell-mediated immune responses in rhesus macaques. *Pharmacology, Biochemistry, and Behavior* **73**, 271–278.
- Thierry, B., Singh, M. and Kaumanns, W. (eds.) (2004). *Macaque societies: a model for the study of social organization*. Cambridge, UK: Cambridge University Press.

Primate Models, Cardiovascular Disease**J R Kaplan**

Wake Forest University School of Medicine, Winston
Salem, NC, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition
article by J R Kaplan, volume 3, pp 230–235,

© 2000, Elsevier Inc.

Nonhuman Primates as Animal Models

Behavioral Influences on Risk Factors for Cardiovascular
Disease

Behavioral Influences on Atherosclerosis

Behaviorally Induced Abnormalities in Vascular Reactivity

Summary

Glossary

Abnormal coronary vascular reactivity The tendency for coronary arteries to lose their ability to dilate rather than constrict as demand for blood is increased.

Atherosclerosis A gruel-like accumulation of fatty substance in the intima (inner layer) of the large muscular and elastic arteries of the body.

Beta-blocker A class of drugs that inhibits the activation of the sympathetic nervous system (i.e., heart rate, blood pressure, and

cardiac output) by blocking the response of β -adrenergic receptors in the arteries and heart.

Coronary heart disease (CHD) A disease spectrum subsuming myocardial infarction (heart attack), angina pectoris (ischemic chest pain), and disruptions to the rhythm and electrical activity of the heart.

Dominant/subordinate social status The position that an animal occupies in a pecking order. High-status animals (dominants) predictably defeat low-status animals (subordinates) in competitive encounters. Subordinate animals generally exhibit signs of stress (e.g., exaggerated release of cortisol from the adrenal cortex and behavioral withdrawal).

Endothelium The inner wall of the artery; a normally functioning endothelium causes the artery to dilate and constrict appropriately in response to blood flow.

Heart rate reactivity The degree to which an individual's heart rate increases in response to stressful situations. High heart rate reactivity is probably due to excessive activation of the sympathetic nervous system, and it is associated with the increased development of atherosclerosis in people and monkeys.

Hypertension A sustained increase in blood pressure.

<i>Ischemia</i>	A reduction in blood flow to an organ or tissue, due to obstruction (e.g., a blood clot), arterial constriction, atherosclerosis, or a combination of factors.
<i>Plasma lipids</i>	Fats or fatlike substances (insoluble in water) that include total plasma cholesterol (TPC), high-density lipoprotein cholesterol (HDL), and low-density lipoprotein cholesterol (LDL). In general, $TPC = HDL + LDL$. Elevations in plasma levels of LDL (and the resulting increases in TPC) are highly atherogenic. High plasma concentrations of HDL (the good cholesterol) are protective.
<i>Quantitative coronary angiography</i>	An imaging procedure in which an opaque dye is infused into the coronary arteries, allowing them to be visualized radiographically and measured.
<i>Sympathetic nervous system (SNS)</i>	That part of the involuntary or self-regulating nervous system that controls responses (e.g., heart rate and blood pressure increases) to stressful situations.

Cardiovascular disease is manifested most prominently in human beings as coronary heart disease (CHD) and stroke. Atherosclerosis in the coronary arteries is a usual cause of CHD, whereas atherosclerosis in the cerebral arteries contributes to both transient ischemic attacks and ischemic stroke. Epidemiological and clinical studies have provided considerable evidence that behavioral factors influence the development of cardiovascular disease and the underlying atherosclerosis. However, logistical and ethical considerations prevent many questions regarding risk factors and mechanisms from being addressed directly in people. Animal models offer an alternative means of answering such questions.

Nonhuman Primates as Animal Models

The ideal animal for evaluating behavioral influences on cardiovascular disease would respond to a fat- and cholesterol-containing diet by exhibiting both an intimal accumulation of atherosclerosis in the large proximal coronary arteries and a susceptibility to infarction and ischemia. Furthermore, this animal would resemble human beings in cardiovascular and reproductive physiology and in neurobiology. Finally, the animal's behavioral responses and characteristics would vary along dimensions (e.g., aggressiveness and competitiveness, social support) analogous to those thought important in the epidemiology of atherosclerosis and CHD in human beings. Of all animal models, the nonhuman primates best approximate the foregoing characteristics.

Numerous investigators have used primates to explore one facet or another of the hypothesis that behaviorally evoked, excessive perturbations of the body's principal axes of neuroendocrine response (i.e., the adrenal cortex or the SNS) may produce pathophysiological consequences. Some of these studies provide data regarding the influence of stress on the risk factors for cardiovascular disease (e.g., elevations in plasma lipids and blood pressure) or on mediating factors (circulating concentrations of immune or inflammatory markers), whereas others have focused primarily on the characteristics of atherosclerotic lesions or abnormalities in vascular reactivity to changes in blood flow. Taken together, existing studies present a clear and convincing case that behavioral factors (either stressful environments or particular characteristics of temperament) influence both atherosclerosis and clinical events in ways that have relevance for human beings.

Behavioral Influences on Risk Factors for Cardiovascular Disease

Plasma Lipids

Elevations in plasma lipids (e.g., total plasma cholesterol, TPC; low-density lipoprotein cholesterol, LDL) are associated with the exacerbation of atherosclerosis and increased risk of CHD and ischemic stroke. Thus, behavioral factors may contribute to cardiovascular disease through a mechanism involving adverse changes in plasma lipids. Among primates, the lipids of squirrel monkeys (*Saimiri sciurius*, a New World primate) seem particularly susceptible to behavioral influences. In one study, for example, there was an increase in plasma lipids as animals were transported from South America to a laboratory in the United States. Interestingly, both TPC and LDL declined in these animals after they were put in a secluded environment, but increased again following exposure to a busy laboratory. All of these changes occurred in the absence of any cholesterol in the diet and are attributed most reasonably to the influence of psychic stress. Cholesterol-fed squirrel monkeys are similarly responsive to experimental manipulations; exposure to either a shock-avoidance paradigm (in which the animal presses a lever to avoid a shock) or a restraint box leads to a significant and sustained increase in TPC and LDL.

Among the Old World monkeys, adverse serum lipid profiles have been observed in relation to individual differences in social status and individual differences in heart rate reactivity to stress. In both baboons (*Papio hamadrayas*) and cynomolgus macaques

(*Macaca fascicularis*), for example, subordinate social status is associated with suppressed plasma concentrations of high-density lipoprotein cholesterol (HDL), which is a risk factor for atherosclerosis. Notably, cynomolgus macaques that reliably exhibit exaggerated heart rate responses to psychological stress (e.g., exposure to a threatening object) also tend to have suppressed plasma concentrations of HDL.

Data from Old and New World monkeys suggest that stressful environments, low social status, and exaggerated SNS responsivity to challenge are all associated with potentially harmful plasma lipid profiles. Among human beings, there are also numerous observations relating either elevations in LDL or reductions in HDL to stressful circumstances or particular personality characteristics.

Blood Pressure

Elevated blood pressure (hypertension) is associated with increased risk of CHD and stroke. High blood pressure also exacerbates atherosclerosis in the coronary and cerebral arteries. Numerous studies indicate that behavioral stress can lead to sustained, potentially pathogenic increases in both systolic and diastolic blood pressure. For example, squirrel monkeys trained to avoid a shock by pressing a lever ultimately exhibit increases in blood pressure that persist across experimental sessions. Rhesus monkeys (*M. mulatta*) exposed to a similar manipulation also develop a steady elevation in blood pressure. Further, rhesus monkeys as well as baboons experience sustained increases in blood pressure when light signals and shock are paired and combined with a bar press for the delivery of food (called conflict manipulation). Chronic exposure to loud noise also produces elevated blood pressure in cynomolgus and rhesus macaques. Notably, adult male baboons that typically live together with a number of females in a harem develop hypertension when they are separated from their harems and forced to watch the females feed first. Males also exhibit a hypertensive response when they are separated from their harems and are forced to observe a new male being introduced to these females. These latter manipulations are purely social and more closely model the kind of stress most likely to affect human beings.

Behavioral Influences on Atherosclerosis

Atherosclerosis in Males

Primates typically do not develop atherosclerosis to any significant degree unless they are fed a diet

relatively high in saturated fat and cholesterol, similar to the diet often consumed by people living in industrialized nations. Early investigations used physical restraint or exposure to a shock-avoidance manipulation to provide evidence that psychic stress, in combination with an atherogenic diet, could potentiate atherosclerosis in monkeys. However, social manipulations also influence the development of atherosclerosis in monkeys. For example, when cynomolgus males are housed in groups that either remain stable in composition or that are perturbed (by distributing members across groups), dominant and subordinate animals tend to retain their relative rankings irrespective of social condition (i.e., stable or unstable groups). A quantitative evaluation of the coronary arteries of such animals revealed that monkeys housed in an unstable condition had significantly more coronary artery atherosclerosis than animals in a stable environment but only if they were dominant in their groups. This finding suggests that the behavioral demands of retaining dominant social status in an unstable social setting potentiate atherosclerosis, at least in males of this species consuming an atherogenic diet. A similar pattern was obtained (although with markedly smaller lesions), among animals that consumed a diet modeled on the recommendations of the American Heart Association. Thus, behavioral factors might, even in the absence of dietary provocation, cause damage that could initiate the earliest lesions of atherosclerosis. Subsequent investigations have shown that exposure to the stress of social perturbation for as little as 72 h can cause the degeneration and death of endothelial cells, thus marking the first stage of atherosclerosis progression.

Mediating Role of Sympathetic Nervous System Activation in the Atherosclerosis of Males

Significantly, behavioral effects observed in the social manipulation experiments were independent of variation in serum lipid concentrations and resting blood pressure. Numerous investigators have proposed that recurrent SNS activation in response to behavioral stimulation could promote the development of atherosclerosis and its clinical sequelae. Testing the hypothesis that excessive SNS activation might have potentiated atherogenesis among dominant animals in the foregoing studies, male monkeys in a subsequent investigation were all housed in unstable social groups and fed an atherogenic diet; here, half of the monkeys also received a beta-blocker, propranolol, throughout the 2-year study. Untreated dominant monkeys again had twice the amount of atherosclerosis of their subordinate counterparts,

replicating the results from the unstable treatment condition of earlier experiments. Dominant monkeys treated with propranolol, however, developed less than half the coronary atherosclerosis of untreated dominants and had roughly the same extent of lesion as both treated and untreated subordinate animals. Hence, the more pronounced atherosclerosis characteristic of untreated (i.e., autonomically intact) dominant monkeys, when housed in groups of rotating membership and fed a cholesterol containing diet, can be prevented by the long-term administration of a beta-blocker and is therefore dependent on concomitant SNS activation. Furthermore, the monkey studies also provide evidence that behaviorally induced SNS activation initiates atherogenesis independently of the presence of a high-fat diet. Hence, the endothelial cell death associated with the short-term (72-h) perturbation previously described occurs in the absence of elevated plasma lipids and can be inhibited by pretreatment with a beta-blocker.

Other data from this species further support the speculation that a heightened SNS responsivity to behavioral stimuli increases the risk for atherosclerosis. Specifically, cynomolgus macaques display reliable and significant individual differences in heart rate response (assessed via radiotelemetry) to a standardized social stressor (exposure to an unknown human). Animals can thus be characterized as either high or low in heart rate reactivity to stress. Notably, the atherosclerotic lesions of high heart rate reactors are twice as large as those of their low-reactive counterparts. Taken together, the results of the macaque experiments indicate that some individuals (e.g., high heart rate reactors) may be susceptible to the development of atherosclerosis because of an intrinsic SNS hyperresponsivity to behavioral stimuli, whereas others (dominant individuals in unstable social environments) may be at increased risk because of frequent elicitation of such responses in stressful environments. SNS activation, provoked in either of two ways, may thus provide the common pathway by which pathogenesis is mediated in male monkeys.

Summary for Male Monkeys

The studies of male monkeys relate a behavioral predisposition (social dominance, i.e., successful aggression) and excessive SNS stimulation to the exacerbation of atherosclerosis. Both domains appear to have analogs among human beings. There are other behavioral risk factors, however, about which the studies in male monkeys provide little data or insight. Such factors include depression, sleep disturbances, and a propensity to anxiety or panic, all of which

have been associated reliably with elevated risk of CHD in human beings. Moreover, social support is often believed to be cardioprotective. Although several measures of social support have been evaluated in the foregoing studies of male monkeys, there is no evidence that animals differing by these measures vary in their extent of atherosclerosis.

Sex Differences in Atherosclerosis

The relative sparing of premenopausal women compared to men of similar ages is a prominent feature of CHD, ischemic stroke, and atherosclerosis. Although this phenomenon is sometimes referred to as female protection, it is characterized more accurately as a delay in disease onset, with the incidence curve for women lagging behind that of men by approximately 10 years. The various effects of estrogen are believed to account for most of this sex difference, at least with respect to CHD incidence and underlying atherosclerosis. Not only are premenopausal women protected from CHD, but the provision of estrogen replacement to women immediately prior to or immediately following menopause is associated with a significant reduction in CHD risk (it must be noted that such protection is not apparent when hormone therapy is initiated in women with preexisting disease or who are 10 or more years postmenopausal).

Because atherosclerosis progresses over decades, it is likely that the clinical events occurring in postmenopausal women have their beginnings in the premenopausal years. This conclusion is supported by autopsy assessments showing that by 34 years of age, at least one-third of all women (regardless of race) have raised lesions in their coronary arteries. *In vivo* arterial imaging also demonstrates the presence of relatively extensive focal atherosclerosis in premenopausal women. The observations that atherosclerosis begins premenopausally raises the question of whether there are particular risk factors that place some women on a higher-risk trajectory than others. Considerable effort has been focused on this question in studies employing premenopausal monkeys as a model of reproductive-age women. The major results of these studies are described next.

Social Stress and Female Protection from Atherosclerosis in Monkeys

Investigations using socially housed female cynomolgus monkeys provide evidence that stress-induced ovarian dysfunction eliminates premenopausal protection from atherosclerosis. For example, in one 2-year experiment contrasting the atherosclerosis of socially housed males and females, the females were divided

into groups of five. Each of these groups contained a vasectomized male; additional males were placed together in groups of five. All animals consumed an atherogenic diet. Quantitative evaluation of the coronary arteries at the end of this experiment demonstrated that female protection with regard to atherosclerosis extended only to the dominant animals (i.e., those ranking first or second in their social groups); subordinate females (ranking third or below) were indistinguishable from males. Subordinate females also had five times as many anovulatory cycles and three times as many cycles characterized by luteal-phase progesterone deficiencies as did their dominant counterparts. Notably, females with the most extensive atherosclerosis were all subordinate, and all had marked ovarian endocrine dysfunction. These data suggest that the stress of social subordination may increase the risk of atherosclerosis in females by inducing a relative ovarian endocrine deficiency state similar to that observed in postmenopausal women. Interestingly, investigations in premenopausal women also have linked stress-induced estrogen deficiency to severity of angiographically documented CHD.

In a further study, ovariectomized females were placed in social groups for the same amount of time and fed the same diet as the reproductively intact females just described. The ovariectomized females as a group had significantly more extensive coronary artery atherosclerosis than the intact females. This effect was due entirely to the loss of protection experienced by the ovariectomized dominant animals compared with their intact counterparts. The subordinates, intact and ovariectomized, were affected equally with atherosclerosis. Additional studies have extended these findings by showing that the treatment of premenopausal females with either an oral contraceptive or unopposed estrogen treatment eliminates the difference between dominant and subordinate individuals. Parenthetically, it should be noted that subordinate females tended to be hypercortisolemic compared to dominants, leading some scientists to suggest that the atherosclerosis distinctions between such animals may be mediated in part by the excessive activation of the hypothalamic-pituitary-adrenocortical axis. The fact that dominant and subordinate animals continue to differ in adrenocortical activation – but not atherosclerosis extent – following ovariectomy argues against this hypothesis.

Finally, a recent life-course study demonstrated that premenopausal behavioral and hormonal status predict postmenopausal atherosclerosis outcomes independently of any postmenopausal hormone exposure. Specifically, females that were subordinate in their

social groups – and thus were ovarian-suppressed and estrogen-deficient – developed more atherosclerosis premenopausally and postmenopausally than did their dominant, ovarian-normal counterparts. Significantly, the worsened atherosclerosis of the subordinate, socially stressed individuals could be mitigated by premenopausal treatment with exogenous estrogen in the form of an oral contraceptive. This important set of observations implies for women that premenopausal behavioral and associated hormonal (estrogenic) characteristics establish the trajectory of atherosclerosis progression that ultimately culminates in postmenopausal CHD events. The monkey research thus emphasizes the importance of early pathogenesis and early intervention, a perspective supported by recent reevaluations of clinical trial data.

The different patterns in which social status influences atherosclerosis in male and female monkeys deserves comment. Subordinate female primates, like subordinate and stressed females of many other mammalian species, tend to experience impaired ovarian function. This impairment eliminates female protection from atherosclerosis in monkeys and, by implication, women. Subordinate male monkeys also experience stress, possibly including impaired reproductive activity; however, the physiological sequelae of social subordination among males do not seem to affect the coronary arteries. Instead, dominant males in unstable settings are predisposed to atherosclerosis because they experience marked elevations in artery-damaging SNS activity as they attempt to maintain behavioral preeminence in the face of challenge.

Social Isolation, Depression, and Atherosclerosis in Females

The exacerbation of atherosclerosis in female monkeys is also observed in conjunction with single-cage housing (i.e., social separation). Specifically, females housed in single cages have significantly more coronary artery atherosclerosis than comparably treated animals housed in social groups. Single-caged animals are most similar in atherosclerosis extent to socially housed subordinates. Interestingly, this worsening of atherosclerosis among single-caged females is not related to ovarian impairment; rather, such animals exhibit elevations in heart rate. This latter observation suggests that, at least in female monkeys, social separation exacerbates coronary artery atherosclerosis via exaggerated SNS activity. No similar studies have been done in male monkeys. This outcome in females is, however, consistent with extensive literature documenting the cardioprotective

effects of social support and the detrimental impact on CHD of social isolation in human beings.

Epidemiological studies have linked depression to CHD, and it is well known that women suffer from a higher incidence of depression than men. The role of depression in the coronary artery atherosclerosis of female monkeys has not been systematically evaluated, in part because primate depression usually has been studied with respect to developmental abnormalities caused by early separation of infants from their mothers. Recently, however, it has been shown that a proportion of socially housed, premenopausal cynomolgus monkeys exhibited behavioral and physiological characteristics that model adult human depression. This model offers for the first time the possibility of determining experimentally whether depression is causally associated with the development of atherosclerosis, the primary underlying cause of CHD.

Behaviorally Induced Abnormalities in Vascular Reactivity

Evaluation of Abnormalities in Vascular Reactivity

Numerous investigators have proposed that behavioral stress contributes to the occurrence of myocardial ischemia by causing the arteries to constrict inappropriately, thus cutting off the flow of blood to the heart. The potential for ischemia is evaluated by assessing the responses of coronary arteries to the infusion of vasoactive substances in the context of behavioral investigations. The normal endothelium, for example, dilates in response to acetylcholine. In contrast, endothelial damage causes a paradoxical constriction in response to acetylcholine. Nitroglycerine, an endothelium-independent dilator, reliably causes arterial dilation irrespective of the condition of the endothelium. Using such vasoactive substances and appropriate imaging techniques, it is now possible to visualize and quantitate the responsiveness of the coronary arteries of monkeys exposed to various kinds of stress. Such functional assessments have markedly extended the usefulness of monkeys in cardiovascular research.

Studies in Male Monkeys

In one such study, quantitative coronary angiography was used to evaluate arterial responses to acetylcholine and nitroglycerine in male cynomolgus monkeys that either were or were not exposed to the stress of repeated reorganization of social-group membership (i.e., social instability). The coronary arteries of animals living in unstable groups constricted in response

to acetylcholine, whereas those of monkeys from stable groups dilated. Significantly, the degree of constriction observed in this experiment was independent of the extent of the underlying atherosclerosis; animals with little atherosclerosis but living in unstable groups had a degree of constriction equivalent to unstable monkeys that had extensive atherosclerosis. In contrast, monkeys living in stable groups dilated, irrespective of the amount of atherosclerosis.

The previous study, and others like it, demonstrate that chronic social stress impairs endothelium-dependent vascular responses, an effect that is not necessarily dependent on the extent of atherosclerosis. It should be noted that none of these investigations revealed an influence of dominance status on vascular reactivity. Hence, the pattern of behavioral influences on blood flow (and thus on ischemia) differs somewhat from that observed with respect to atherosclerosis, at least among socially housed cynomolgus males.

Studies in Female Monkeys

As with atherosclerosis, behavioral influences on vascular reactivity among females are apparently mediated by social subordination and ovarian dysfunction. Hence, the coronary arteries of socially subordinate female cynomolgus monkeys, compared to dominants, constrict in response to acetylcholine. In contrast, nitroglycerine dilates the arteries of all monkeys, irrespective of status. As in the other studies, subordinates here were characterized by ovarian dysfunction (low plasma estradiol and luteal-phase abnormalities in progesterone). Notably, the extent of vascular constriction in response to acetylcholine was directly related to the degree of ovarian impairment as indicated by plasma estradiol concentrations on the day of testing. These observations not only demonstrate the role of behavioral factors on coronary vascular reactivity among females, they emphasize again the cardioprotective significance of normal ovarian function.

Summary

Systematic studies in nonhuman primates have shown that stress, both social and nonsocial, consistently affects the risk factors for CHD and the underlying atherosclerosis. Studies involving socially housed individuals are probably of most relevance to human beings; such investigations have focused primarily on the development of coronary artery atherosclerosis in macaque monkeys. Stress clearly exacerbates atherosclerosis in these animals, albeit differently in males and females.

Among males, dominant individuals in unstable settings and those who exhibit exaggerated heart rate reactivity to challenge are at greatest risk. These outcomes are most likely mediated by excessive SNS activation. These findings are reminiscent of observations in men linking excessive hostility and high heart rate reactivity to increased CHD risk. The monkey results imply that the pathogenicity of hostility (approximated in the animals by social dominance) is most prominent under stressful disrupted social conditions. Notably, indices of social support have not been found to predict patterns of atherosclerosis among male monkeys.

The situation among females is different. Normal ovarian function seems critical for the premenopausal inhibition of atherosclerosis among animals consuming a diet relatively high in fat and cholesterol. Notably, among groups of females, animals occupying the lower status positions are stressed and experience ovarian impairment and relative estrogen deficiency. Such individuals are characterized by a rapid acceleration of premenopausal atherosclerosis, which can be inhibited by exogenous estrogen. A novel observation is that the atherosclerosis trajectory established by premenopausal behavioral and hormonal conditions predicts the postmenopausal atherosclerosis extent. These findings suggest that primary prevention of CHD in women should begin prior to or immediately at menopause. The studies in female monkeys also identify social isolation and social support as potential factors affecting disease risk. Systematic studies of the role played by depression in the development of atherosclerosis have not yet been conducted, but they hold promise for the future.

Further Reading

- Bairey Merz, C. N., Johnson, B. D., Sharaf, B. L., et al. (2003). Hypoestrogenemia of hypothalamic origin and coronary artery disease in premenopausal women: a report from the NHLBI-sponsored WISE study. *Journal of the American College of Cardiology* **41**, 413–419.
- Cohen, S., Kaplan, J. R. and Manuck, S. B. (1994). Social support and coronary heart disease: underlying psychologic and biologic mechanisms. In: Schumaker, S. & Czajkowski, S. M. (eds.) *Social support and cardiovascular disease*, pp. 195–221. New York: Plenum Press.
- Henry, J. P. and Stephens, P. M. (1977). *Stress, health and the social environment: a sociobiologic approach to medicine*. New York: Springer-Verlag.
- Kaplan, J. R. and Manuck, S. B. (1999). Status, stress, and atherosclerosis: the role of environment and individual behavior. *Annals of the New York Academy of Sciences* **896**, 145–161.
- Kaplan, J. R. and Manuck, S. B. (2004). Ovarian dysfunction, stress, and disease: a primate continuum. *ILAR Journal* **45**, 89–115.
- Kaplan, J. R., Manuck, S. B., Anthony, M. S., et al. (2002). Premenopausal social status and hormone exposure predict postmenopausal atherosclerosis in female monkeys. *Obstetrics and Gynecology* **99**(3), 381–388.
- Kaplan, J. R., Manuck, S. B., Clarkson, T. B., et al. (1985). Animal models of behavioral influences on atherogenesis. *Advances in Behavioral Medicine* **1**, 115–163.
- Klaiber, E. L., Vogel, W. and Rako, S. (2005). A critique of the Women's Health Initiative hormone therapy study. *Fertility and Sterility* **84**, 1589–1601.
- Mikkola, T. S. and Clarkson, T. B. (2006). Coronary heart disease and postmenopausal therapy: conundrum explained by timing? *Journal of Women's Health* **15**, 51–53.
- Rozanski, A., Blumenthal, J. A. and Kaplan, J. R. (1999). Impact of psychological factors on the pathogenesis of cardiovascular disease and implications for therapy. *Circulation* **99**, 2192–2217.
- Schneiderman, N. (1987). Psychophysiologic factors in atherogenesis and coronary artery disease. *Circulation* **76**(supplement I), I41–I47.
- Shively, C. S., Register, T. C., Friedman, D. P., et al. (2005). Social stress-associated depression in adult female cynomolgus monkeys (*Macaca fascicularis*). *Biological Psychology* **69**, 67–84.
- Skantze, H. B., Kaplan, J., Pettersson, K., et al. (1998). Psychosocial stress causes endothelial injury in cynomolgus monkeys via α_1 -adrenoceptor activation. *Atherosclerosis* **136**, 153–161.

Primate Models, Overview

D M Lyons

Stanford University Medical School, Stanford, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D M Lyons, volume 3, pp 236–240, © 2000, Elsevier Inc.

Social Psychological Sources of Stress
Psychobiological Consequences of Stress
Environmental and Genetic Risk Factors

Glossary

<i>Model</i>	Simplified experimental system used as a testing device.
<i>Primates</i>	Mammalian vertebrates with prehensile hands.
<i>Social psychobiology</i>	The study of relationships between social, psychological, and biological processes.

Practical limitations and ethical concerns restrict opportunities for prospective controlled studies of stress in healthy humans and patient populations. Animal models that use rodents are limited because corticolimbic brain substrates of social, emotional, and cognitive control of behavior differ significantly in rats and mice compared to humans and nonhuman primates. Studies of monkeys, baboons, and apes have clearly advanced our understanding of the social psychological sources of stress, psychobiological consequences of stress, and individual differences in vulnerability and resilience related to environmental and genetic risk factors.

Social Psychological Sources of Stress

Models of stress in rats and mice often rely on exposure to cold, electric shock, or physical restraint, whereas primate models highlight the significance of seemingly subtle sources of social psychological stress.

Maternal Separation and Loss

Primate infants generally respond to maternal separation with increased secretion of hypothalamic-pituitary-adrenal axis stress hormones. The hypothalamic release of corticotropin releasing hormone (CRH) stimulates pituitary secretion of adrenocorticotrophic hormone (ACTH), which triggers the secretion of glucocorticoids (i.e., cortisol) from the adrenal cortex. An informative exception to the rule

that maternal separation elicits a robust stress response is found in one of the few primate species in which fathers provide parental care. In titi monkeys, maternal separation does not elicit an adrenocortical stress response unless the father is also removed from the developing infant. Separation from the father invariably elicits a robust adrenocortical stress response even when infant titi monkeys are left in the company of their mother. These findings concur with behavioral evidence that fathers are the primary attachment figure for titi monkey infants.

Relationship Formation and Social Reunions

Among adult primates, the formation of new relationships is often stressful, as demonstrated by measures of social behavior, elevated plasma levels of cortisol, and evidence of impaired immunity. Even reunions with well-known relatives may elicit surprising effects. When previously separated adolescent mangabey monkeys are returned to their natal social group, these monkeys and their mothers engage in frequent affiliative behavior. Nevertheless, these reunions evoke in adolescent offspring increased plasma levels of cortisol and diminished cell-mediated immunity.

Crowding

High-density living conditions are especially stressful for certain individuals, as illustrated in rhesus macaques. A statistically significant interaction between crowding and behavioral inhibition (i.e., shyness) was found to predict individual differences in the incidence of physical injury. Compared to uninhibited individuals, highly inhibited cage-mates experienced higher rates of injury during (but not before or after) crowding in the presence of ample food. Evidently, inhibited rhesus macaques are targeted for violence in high-density conditions by uninhibited peers.

Competitive Conflicts and Foraging Demands

Stressful competitive conflicts occur during foraging and feeding in primates. Even in species in which overt aggression is uncommon, there is a general tendency toward mutual avoidance during foraging and feeding. Mutual avoidance minimizes fighting when food sources are dispersed, but when food is confined to limited locations stressful conflicts often emerge. Foraging demands likewise affect maternal behavior in vervet monkeys, squirrel monkeys, baboons, and macaques. Mother-reared bonnet macaque monkeys exposed as infants to stressful foraging demands

show increased emotional reactivity when challenged later in life by maternal separation or acute exposure to novelty. As adults, these bonnet macaque monkeys are also less social, more subordinate, and present with increased cerebrospinal fluid CRH levels.

Dominance Hierarchies and Social Conflicts

Primates that live in large groups can generally be ranked within a pecking order or social dominance hierarchy. The stressful aspects of hierarchies in primate societies are not necessarily related to a particular rank *per se*, but may reflect the style by which rank is expressed and the overall context in which social hierarchies emerge. In stable hierarchies, high-ranking primates tend to exhibit metabolic, immune, and cardiovascular profiles that appear to be less pathogenic than those that are seen in low-ranking subordinates. These rank-related differences are diminished during periods of recurring conflict and social instability. High-ranking primates in stable hierarchies also tend to have lower plasma levels of cortisol, but the levels maintained in dominant baboons that lack rudimentary social skills are as high as those in subordinates.

Psychobiological Consequences of Stress

Various psychobiological consequences of stress have been characterized in primate models at multiple levels of analysis.

Hippocampal Neuropathology

Cellular and molecular markers of hippocampal neuropathology have been identified in adult monkeys following exposure to chronic stress. Similar neuropathological damage is induced when sustained-release cortisol pellets are implanted in the hippocampus. Stress-induced reductions in glucocorticoid receptor binding sites within the hippocampus impair negative feedback sensitivity on dexamethasone suppression tests in cynomolgus macaques. Recent findings in New and Old World primates indicate that the hippocampal dentate gyrus continues to generate new neurons throughout adulthood and that neurogenesis is suppressed by exposure to psychosocial stress.

Neurochemistry and Cognitive Control

Squirrel monkeys that respond to social separation with greater increases in cortisol exhibit greater increases in cerebrospinal fluid levels of the dopamine metabolite homovanillic acid. Stress increases prefrontal dopamine release, and drugs that block dopamine receptors ameliorate stress-induced prefrontal dependent cognitive deficits in rhesus macaques. Stress also decreases serotonin in prefrontal brain

regions and reduces cerebrospinal fluid levels of the serotonin metabolite 5-hydroxyindoleacetic acid (5-HIAA). Low levels of cerebrospinal fluid 5-HIAA are associated with a higher incidence of impulsive risk-taking behavior and excessive consumption of alcohol in adolescent macaques.

Reproductive Biology

Stress inhibits the release of hypothalamic-pituitary-gonadal axis hormones, alters the onset of puberty, impairs reproductive behavior, suppresses lactation, and adversely affects diverse aspects of parental care. The specific mechanisms that produce these outcomes are not well understood. A role for endogenous opioids is indicated by the observation that the stress-induced suppression of hypothalamic-pituitary-gonadal axis hormones is blocked by the administration of the opiate receptor antagonist naloxone. Other studies suggest that the stress-induced inhibition of reproduction occurs through nonopioidergic mechanisms and that hypothalamic-pituitary-adrenal axis hormones mediate these antireproductive effects.

Immunology

Stressful conflicts, the disruption of relationships, and the formation of new social groups impair immune responses in primates. Glucocorticoid levels within the range normally achieved in response to stress suppress numerous aspects of immunity, including lymphocyte numbers, lymphocyte function, and *in vivo* antibody responses to antigenic challenges. An alternative pathway for these immunological effects is through the catecholaminergic innervation of the lymph nodes, thymus, and spleen.

Health Disorders and Disease

Primate models provide evidence for adverse stress effects in coronary heart disease, susceptibility to upper respiratory infections, and vulnerability to simian acquired immune deficiency syndrome (the primate homolog of AIDS). Sustained exposure to stress induces metabolic abnormalities that parallel those found in humans who develop abdominal obesity and early-onset diabetes. Stress exacerbates the effects of high dietary salt intake on hypertension in baboons. Sustained exposure to stress elicits changes in primate behavior and neurobiology that resemble features of human anxiety disorders, alcoholism, and major depression.

Environmental and Genetic Risk Factors

Stress is an inescapable aspect of life that cannot be eliminated nor avoided for long. Equally obvious is that certain individuals are particularly vulnerable to

stress, whereas others seem unscathed. Experimental studies of primates indicate that environmental and genetic risk factors contribute to individual differences in vulnerability and resilience.

Prenatal Development

Pregnant squirrel monkeys and rhesus macaques exposed to chronic sources of stress produce infants with deficits in muscle tone, impaired neuromotor coordination, and shortened attention spans. Increased secretion of cortisol induced in mothers by the administration of exogenous ACTH during mid-gestation results in neonatal outcomes similar to those produced by prenatal stress. Later in life, prenatally stressed monkeys, and monkeys derived from ACTH-treated mothers, exhibit exaggerated neuroendocrine responses to stress and deficits in immune physiology.

Postnatal Development

Rhesus macaque infants raised without mothers show prolonged and inappropriate responses to stress, self-directed emotional behavior, diminished norepinephrine neurotransmission, altered autonomic reactivity, and impaired immune responses. Neuronal changes in dendritic branching have been characterized in motherless monkeys, and selective alterations in tyrosine hydroxylase activity are evident in the striatum, substantia nigra, and ventral tegmental area. Attempts to rehabilitate motherless monkeys may partially ameliorate anomalies in behavior but allegedly fail to normalize neurobiological effects.

Early-Life Stress Inoculation

Severely stressful early experiences increase the risk for the development of adult psychiatric conditions, including depression and anxiety disorders. Far less researched, but of equal importance, is the suggestion that mild forms of stress, instead of increasing vulnerability, result in subsequent stress resistance. Various descriptions of inoculating, immunizing, steeling, toughening, or thriving, the notion that mild early-life stress reduces vulnerability to subsequent stressors has important implications for understanding the neurobiology and prevention of mental health disorders. Recent studies of monkeys suggest that early exposure to mild stress leads subsequently to diminished measures of anxiety, pro-social tendencies, enhanced cognitive control, larger prefrontal brain volumes, and attenuated activation of the hypothalamic-pituitary-adrenal axis.

Social Buffering

The deleterious effects of stress on immune physiology are often buffered or reduced by the presence of familiar companions in primates with social proclivities.

Under certain conditions, familiar companions likewise buffer or reduce the stress-induced secretion of glucocorticoids. Although glucocorticoids are known to regulate immune processes involved in susceptibility to disease, there is limited evidence in primates to implicate glucocorticoids as a direct mediator of social stress-buffering effects.

Phylogenetic Constraints

Species differences in neuroendocrine, autonomic, and behavioral responses to stress are important concomitants of species-typical modes of interacting with social and nonsocial environments. Interspecies differences in stress physiology and emotional vulnerability are well documented for two New World primates (squirrel monkeys and titi monkeys) and for at least five different species of macaques.

Inherited Variation

Coping with stress is increasingly recognized as key feature of health. Stimulated by the concept of primary prevention, health professionals are developing programs aimed at teaching people how to cope effectively with stress. That modes of coping are learned and not inherited from others is implicit in this approach. Yet human twin studies clearly suggest robust hereditary influences on aspects of coping with stress. Individual differences in autonomic reactivity to stress and cerebrospinal fluid levels of neurotransmitter metabolites are likewise heritable in rhesus macaques. High heritabilities for total brain size have been reported in humans and monkeys, but little is known about the genetics of inherited variation in stress-related corticolimbic brain circuits.

See Also the Following Articles

Animal Models (Nonprimate) for Human Stress; Crowding Stress; Hippocampus, Overview; Hypothalamic-Pituitary-Adrenal; Primate Hierarchies and Personality; Primate Models, Behavioral-Immunological Interactions; Primate Models, Cardiovascular Disease; Psychosocial Factors and Stress; Reproductive Dysfunction in Primates, Behaviorally Induced; Social Stress, Animal Models of.

Further Reading

- Abbott, D. H., Keverne, E. B., Bercovitch, F. B., et al. (2003). Are subordinates always stressed?: a comparative analysis of rank differences in cortisol levels among primates. *Hormones and Behavior* 43, 67–82.
- Arnsten, A. F. (2000). Stress impairs prefrontal cortical function in rats and monkeys: role of dopamine D1 and norepinephrine alpha-1 receptor mechanisms. *Progress in Brain Research* 126, 183–192.

- Bailey, M. T., Lubach, G. R. and Coe, C. L. (2004). Prenatal stress alters bacterial colonization of the gut in infant monkeys. *Journal of Pediatric Gastroenterology and Nutrition* 38, 414–421.
- Barr, C. S., Schwandt, M. L., Newman, T. K., et al. (2004). The use of adolescent nonhuman primates to model human alcohol intake: neurobiological, genetic, and psychological variables. *Annals of the New York Academy of Sciences* 1021, 221–233.
- Boyce, W. T., O'Neill-Wagner, P., Price, C. S., et al. (1998). Crowding stress and violent injuries among behaviorally inhibited rhesus macaques. *Health Psychology* 17, 285–289.
- Gorman, J. M., Mathew, S. and Coplan, J. (2002). Neurobiology of early life stress: nonhuman primate models. *Seminars in Clinical Neuropsychiatry* 7, 96–103.
- Gould, E. and Gross, C. G. (2002). Neurogenesis in adult mammals: some progress and problems. *Journal of Neuroscience* 22, 619–623.
- Hoffman, K. A., Mendoza, S. P., Hennessy, M. B., et al. (1995). Responses of infant titi monkeys. *Callicebus moloch*, to removal of one or both parents: evidence for paternal attachment. *Developmental Psychobiology* 28, 399–407.
- Kalin, N. H. (2004). Studying non-human primates: a gateway to understanding anxiety disorders. *Psychopharmacology Bulletin* 38, 8–13.
- Kaplan, J. R. and Manuck, S. B. (2004). Ovarian dysfunction, stress, and disease: a primate continuum. *ILAR Journal* 45, 89–115.
- Lyons, D. M. (2002). Stress, depression, and inherited variation in primate hippocampal and prefrontal brain development. *Psychopharmacology Bulletin* 36, 27–43.
- Parker, K. J., Buckmaster, C. L., Schatzberg, A. F., et al. (2004). Prospective investigation of stress inoculation in young monkeys. *Archives of General Psychiatry* 61, 933–941.
- Schneider, M. L., Moore, C. F., Kraemer, G. W., et al. (2002). The impact of prenatal stress, fetal alcohol exposure, or both on development: perspectives from a primate model. *Psychoneuroendocrinology* 27, 285–298.
- Shively, C. A., Register, T. C., Friedman, D. P., et al. (2005). Social stress-associated depression in adult female cynomolgus monkeys (*Macaca fascicularis*). *Biological Psychology* 69, 67–84.
- Tilbrook, A. J., Turner, A. I. and Clarke, I. J. (2002). Stress and reproduction: central mechanisms and sex differences in non-rodent species. *Stress* 5, 83–100.

Primates: Rearing and Effects of Stress on Primate CNS Function

T K Newman

University of Cape Town, Cape Town, South Africa

C S Barr

National Institutes of Health, Bethesda, Maryland, USA

© 2007 Elsevier Inc. All rights reserved.

Nonhuman Primate Models of Early Stress
Gene by Environment Interaction

Glossary

Gene by environment interaction

Describes how a phenotype (e.g., impulsive behavior) is the result of both genetic and non-genetic (environmental) factors such that the influence of an environmental effect is conditioned by the individual's genotype.

Hypothalamic-pituitary-adrenal (HPA) axis

The hypothalamus, pituitary gland and adrenal cortices form a linked neuroendocrine system (axis) that controls an organism's reaction to stress through the secretion and regulation of a hormonal cascade. Activation of the HPA axis

Rearing

induces the release of CRH (corticotropin-releasing hormone) from the hypothalamus, ACTH (adrenocorticotrophic hormone) from the pituitary gland and glucocorticoid hormones (in primates, cortisol) from the adrenal cortex.

Refers to the practice of raising captive laboratory animals under specific experimental conditions including (1) "mother-reared", or (2) "nursery-reared" in which subjects are permanently removed from their mothers within days of birth, after which the infant may be placed with age-matched peers (peer-rearing), or raised on an inanimate surrogate mother. Rhesus macaques (*Macaca mulatta*) are a species of Old World Monkey with broad indigenous distribution from eastern Afghanistan to Eastern China. They live in gregarious, hierarchical social groups centered on a core of related lines of females. Rhesus are the most common laboratory primate in the United States.

Rhesus macaque

On the basis of the available evidence from human studies, it is well established that stress early in life, particularly during infancy and early childhood, is a significant predictor and risk factor for a range of developmental, neurological, and behavioral deficits, often including autism-like behaviors. Some deficits manifest early in life, resulting in children who are fearful and anxious, socially withdrawn, or exhibit inappropriate and aggressive play behavior. These adverse experiences and impoverished environments in early human childhood are also known to greatly increase the risk of long-term neurobiological effects on systems essential for normal emotional growth, social competence, and appropriate development of neuroendocrine responsivity. It is clear that experience-based brain development in the very early years of life sets biological pathways that affect cognition, behavior (e.g., anxiety and aggression), capacity to learn, memory, and physical and mental health throughout the life cycle. The hypothalamic-pituitary-adrenal (HPA) axis is exceptionally vulnerable to the effects of early life stress, and such perturbations are suspected of playing a major role in the etiology of affective disorders, mood disorders, and other psychopathologies in adulthood. Delineating the proximal causal links among early experience, brain development, and normative and dysfunctional social behavior is a major goal. Although many questions remain, these biological pathways are becoming increasingly better understood, due largely to evidence that early stressful experience reduces neural plasticity and may lead to the permanent silencing of genes critical to the regulation of the stress response to stress experienced later in life.

Nonhuman Primate Models of Early Stress

Nonhuman primates are a remarkably important model for exploring and ultimately attempting to understand the impact of early stress and trauma on the developing brain, its effects on individual differences in temperament and psychopathology, and the potential interaction of these factors with underlying genetic differences that may impart resilience or vulnerability to early stress. In most primate species, particularly those with gregarious and hierarchical social organization such as rhesus macaques (*Macaca mulatta*), infants are nurtured and protected by their mothers and extended female relatives during their first years. They spend most of their first few months in direct physical contact with their mothers, and even a brief separation can elicit robust behavioral, neuroendocrine, and neurobehavioral activation in both mother and infant.

Studies using animal models such as rodents and nonhuman primates have enabled researchers to address experimentally the effects that early traumatic experiences and impoverished environments can have on a multitude of behavioral and physiological parameters. Early stress and trauma can be defined in a number of ways, but principally they include abuse and neglect of the infant and also may include stress or illness in the mother both pre- and postnatally. In the laboratory, early stress and trauma are modeled most commonly by manipulating the conditions under which infants are reared. The first studies investigating the impact of early adverse experiences were focused on the effects of permanent separation of the infant from its mother – a paradigm initiated by Harry Harlow in the 1960s. Variations of the maternal separation paradigm grew to examine the effects of surrogates on maternally deprived infants and the effects of varying exposure to like-aged peers, but all models emphasized the permanent separation of an infant from its mother.

Although these experiments revealed the devastating consequences of maternal deprivation, they nevertheless represent intense manipulations of early social and emotional experience. As such, they may not appropriately capture the less tangible effects on brain development resulting from the kinds of parental neglect and abuse reported in human clinical populations that increase the risk for adult onset psychopathologies. Alternatively, or in addition to, permanent maternal deprivation are studies investigating the effects of acute, periodic, or repeated mother–infant separations followed by reunion. Compared to the severe social and emotional deficits associated with permanent maternal deprivation, infants that experience short-term maternal separation exhibit more subtle impacts on neuronal and behavioral development. Regardless, long-term physiological and behavioral effects result, including alterations to the HPA axis and in stress reactivity. Indeed, infant monkeys exposed to repeated separations from their mothers show pronounced and persistent behavioral and physiological activation. These experiments in nonhuman primates are more likely to reflect the forms of stressful, unpredictable, and traumatic experiences that neglected and abused children may experience and suggest how such early trauma can affect the normal development of neuronal pathways and neural structures, as well as neuroendocrine responsivity.

Gene by Environment Interaction

The nonhuman primate model is also well suited for studying the interplay between heritable genetic

differences between individuals and their response to nongenetic, environmental insults, that is, gene by environment interaction. Compared to rodents, nonhuman primates such as rhesus macaques are closely related to humans, and this is evident in the degree of similarity between human and rhesus gene sequences of common behavioral candidate genes. At the same time, captive laboratory primates share tightly controlled housing conditions that help to limit the number of intangible, nongenetic variables, which are difficult to isolate in human studies. Finally, most psychopathology involves deficits in appropriate emotional and social functioning, and the complex social organization and behavioral repertoires of nonhuman primates approximate far more closely those of human subjects than do other laboratory animals.

Recent efforts to understand the interaction between genetic and nongenetic (environmental) factors that underlie interindividual differences in response to stress and trauma have focused on functional polymorphisms in genes that are thought to play an important role in mediating response to stress or to be influenced by the effects of early stress and trauma. One of the better-known examples comes from the serotonin system. Serotonin is an important neurotransmitter involved in the regulation of mood and affect, and perturbations in serotonergic neurotransmission, whether naturally occurring or induced, are thought to result in a host of emotional and behavioral disorders in humans. Serotonin is critical for the development of the central nervous system (CNS) during the early postnatal period when mother–infant bonds are cemented. Moreover, serotonergic neurotransmission is involved in the activation and feedback control of the neuroendocrine stress axis, and the activation of stress hormones is known to regulate gene expression, including the serotonin transporter gene (5-HTT).

Both the rhesus macaque and human versions of the serotonin transporter gene contain a complex length polymorphism in the gene's promoter region in which the shorter (s) of the two common alleles reduces transcriptional efficiency. The s allele in humans is associated with numerous mood and psychiatric disorders, including anxiety and depression, harm avoidance, and aggression-related traits. In a recent human longitudinal study, serotonin transporter gene variation increased the risk of depressive symptoms in individuals who had stressful, impoverished childhood experiences. In rhesus macaques, several studies have demonstrated an interaction between the s allele of the serotonin transporter and permanent and acute maternal separation on (1) cerebrospinal fluid measures of 5-hydroxy-3-indoleacetic acid (5-HIAA), a marker of central serotonergic

neurotransmission; (2) neonatal measures of motor coordination, reflex, and orienting ability and temperament (thought to be sensitive measures of CNS growth and maturity in the developing brain); and (3) HPA axis activation. In all cases, the effects of serotonin genotype were evident in maternally separated monkeys (but not in maternally raised monkeys), with persistent, lasting effects into adulthood on affective behaviors.

The monoamine oxidase A gene (MAOA), a common candidate gene in human psychiatric genetics involved in metabolizing serotonin, dopamine, and norepinephrine, also contains a functional length polymorphism in the regulatory region that is conserved across humans and rhesus macaques. As with the serotonin transporter gene, the association between high- and low-activity alleles and aggressive behavior was dependent on whether subjects had been raised by their mothers. That is, the effect of functional variation in MAOA was conditioned by early rearing experience.

Although early traumatic and stressful experiences in infancy and childhood have demonstrated effects on a wide range of downstream emotional, behavioral, and neurobiological deficits, it is also clear that not all individuals exposed to early life stress undergo such deficits. The nonhuman primate model for early stress and trauma provides an experimental vehicle for controlled hypothesis-driven analyses that have begun to delineate the complex interactions between early stress and normal gene variation and their effects on brain and CNS function. Yet the specific and causal links between genes and environment, and their interactions on neuronal growth and development in the infant brain, remain poorly understood. As such, there is a continued need for clinical and population studies in humans and in nonhuman animal model systems such as primates to elucidate the complex interplay of heritable predisposition and experience on brain development and function.

See Also the Following Articles

Aggression; Homosexuality, Stress and; Genetic Polymorphisms in Stress Response; Monoamine Oxidase; Serotonin Transporter Genetic Modifications.

Further Reading

- Barr, C. S., Newman, T. K., Becker, M. L., et al. (2003). The utility of the non-human primate model for studying gene by environment interactions in behavioral research. *Genes, Brain and Behavior* 2, 336–340.
- Bennett, A. J., Lesch, K. P., Heils, A., et al. (2002). Early experience and serotonin transporter gene variation

- interact to influence primate CNS function. *Molecular Psychiatry* 7, 118–122.
- Champoux, M., Bennett, A., Shannon, C., et al. (2002). Serotonin transporter gene polymorphism, differential early rearing and behavior in rhesus monkey neonates. *Molecular Psychiatry* 7, 1058–1063.
- Higley, J. D., Suomi, S. J. and Linnoila, M. (1991). CSF monoamine metabolite concentrations vary according to age, rearing and sex, and are influenced by the stressor of social separation in rhesus monkeys. *Psychopharmacology* 103, 551–556.
- Newman, T. K., Syagailo, Y., Barr, C. S., et al. (2005). Monoamine oxidase A gene promoter variation and rearing experience influence aggressive behavior in rhesus monkeys. *Biological Psychiatry* 57, 167–172.
- Sánchez, M. M., Ladd, C. O. and Plotsky, P. M. (2001). Early adverse experience as a developmental factor for later psychopathology: evidence from rodent and primate models. *Development and Psychopathology* 13, 419–449.
- Suomi, S. J. (1987). Genetic and maternal contributions to individual differences in rhesus monkey biobehavioral development. In: Krasnegor, N. A., Blass, E. M., Hofer, M. A. & Smotherman, W. P. (eds.) *Perinatal development: a psychobiological perspective*, pp. 397–420. New York: Academic Press.

Prison

D L Whitehead and A Steptoe

University College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J Griffith and A Steptoe, volume 3, pp 241–246, © 2000, Elsevier Inc.

Scope of the Problem

Prison Inmates

Prison Staff

Limitations and Conclusion

Glossary

<i>Bullying</i>	The act of repeatedly intimidating a weaker person by the real or threatened infliction of physical, verbal, or emotional abuse.
<i>Role conflict</i>	A form of social conflict that takes place when the person is forced to take on two different and incompatible roles at the same time.
<i>Depersonalization</i>	A feeling of detachment or estrangement from the self, in which the individual senses a loss of personal identity.
<i>Punitive approach</i>	The perspective on incarceration which reflects the belief that inflicting punishment on an individual will act as a deterrent and decrease the likelihood of the crime being repeated.
<i>Rehabilitative approach</i>	The perspective on incarceration which reflects the belief that an individual can be retrained to function appropriately in society through supportive educational programs in prison.

Scope of the Problem

In 2005, 714/100 000 U.S. residents were incarcerated in prisons and jails, a rate that had risen by 5.0% over the previous 5 years. Bureau of Justice statistics in the United States indicate that if current incarceration rates remain unchanged, an estimated 1 out of every 20 people (5.1%) will serve time in prison during his or her lifetime. U.S. incarceration rates are very much higher than in Europe, where there are wide variations between countries. Southern European countries have a relatively low rate of 80/100 000, compared with 184/100 000 in eastern and central Europe; rates in the United Kingdom fall in the middle, at 139/100 000 in 2005.

There is a very much higher incarceration rate among men than women, and more than 90% of convicts are men. However, one important change over recent years has been the rate of growth of the number of female inmates, which has outstripped that of men. The U.S. female prison population grew by 5.0% annually from 1999 to 2004, compared with a growth rate of 3.3% in the male population. Research has not kept pace with this change, and until recently most studies were conducted on the male inmate population. Overall, inmates are more likely to be young, single, and less well educated than the population at large. Ethnic minority groups are substantially overrepresented in the prison population. In 2004, 12.6% of Black American males in their late twenties were incarcerated, compared with 3.6% of Hispanic males and 1.7% of White males.

Imprisonment is one of the most stressful human experiences and figures high in all comparisons of negative life events. A combination of environmental

and personal deprivation in unpleasant and often frightening circumstances can lead to the perception of extreme stress. Factors such as adherence to a strict daily routine, close living proximity, loss of control over life course, and loss of freedom have all been identified as causing psychological distress. Interpersonal relationships with prison officers and other inmates and the strain on relationships with friends and family are also causes of psychological disturbance. Associated factors such as financial hardship, coping with guilt, and dealing with aggressive fellow inmates can contribute to stress-related problems in prisoners.

Prison staff also report a range of stressors associated with the job. The main stressors are the maintenance of order and the risk of loss of control over hostile inmates, combined with the lack of support from supervisors and society. The difficulty in balancing the responsibility for the welfare of prisoners with the responsibility for security leads to conflicting work demands.

Prison Inmates

Sources of Stress

A number of studies have identified a range of stressors reported among inmates. These fall into several categories.

Personal relationships The separation from partner and family has been cited as one of the main stressors of serving time in prison. For female inmates, the separation from their children is particularly stressful. The majority of women in prison are mothers and many are also heads of households. The separation from their children has been found to be associated with a wide range of feelings such as guilt, anxiety about welfare of the children, and fear of losing close attachments to the children. This may be one reason why women have been found to suffer more emotional distress in prison than men.

The effects of separation are also evident from the finding that married men show higher levels of distress than single men. Being deprived of intimacy and worrying about the welfare of their family are also significant. The lack of confiding relationships and being denied meaningful contact with the opposite gender may create new anxieties regarding sexual identity. Coming to terms with such a mixture of emotions can pose a risk to mental health.

Establishing friendships within prison is an important aspect of adaptation. One study of young inmates found that social support and the ability to confide in

friends in prison were major predictors of psychological well-being. The fear of rejection by other inmates and social isolation can cause heightened anxiety, especially in the initial stages of incarceration. However, in many cases such concerns are ill-founded because one study found that prisoners tend to receive more social support from other prisoners than they had anticipated.

Economic factors The loss of income associated with incarceration can be a major problem for prisoners and their families. Men who report financial hardship experience greater distress than those who are financially secure.

Prison environment Lack of privacy and crowding are serious problems within the prison environment. In many countries, prisons are filled beyond capacity. In the United Kingdom, for example, the Home Office released figures in 2005 that indicated a 106% occupancy rate across all prison service establishments. When personal space is threatened, angry or defensive reactions may result. Research has shown that men living in single cells report liking their accommodation more and perceive themselves having greater control than inmates living in multiple occupancy cells or dormitories. Violence is a serious concern for many inmates, and many have long-lasting fears of attack that may be associated with sustained neuroendocrine and autonomic activation. The rates of sexual and physical assault are difficult to estimate because much abuse goes unreported. Bullying and favoritism by staff are also found in many institutions.

Other aspects of the prison environment that can be sources of stress are noise levels, lack of control over lighting, physical discomfort arising from extremes of heat and cold, lack of contact with the natural environment, absence of intellectual stimulation, and constraints on physical activity.

One issue that is problematic is the pressure not to show evidence of distress. Often emotions associated with stress such as fear, grief, and depression are not admitted and are hidden from staff and other inmates. This hinders the analysis and solution of stress-related problems.

Time Course of Imprisonment

Certain sources of stress are more evident at different stages of imprisonment. Three phases of a prisoner's career have been identified.

Initial phase This is the most stressful period when the inmate first encounters prison life. First-time

offenders are particularly vulnerable due to the loss of liberty and separation from friends and family. This stage is often characterized by feelings of denial followed by anxiety and depression, although other emotions such as shock, fear, isolation, grief, and anger are also reported frequently. The adjustment period typically lasts for the first 4–6 weeks of confinement. It is at this stage that many suicide attempts are made.

Middle phase During the middle stage of the sentence, the focus is on the prison community. Friendships are established, and some inmates experiment with drugs and gay sex. Mandatory drug testing has been introduced in many countries; despite efforts to prevent the importation and use of illegal drugs into prisons, the problem is common. In 2001, a UK prison survey found that 19% of inmates reported using cannabis in the previous month, and 13% reported using opiates. Although drug use is prevalent, a smaller proportion of inmates use drugs while in prison than did prior to incarceration.

Final phase Some convicts become preoccupied with their release in the final few weeks of their prison terms. Apprehension about cultural changes after long sentences is coupled with worry about the attitudes of friends and neighbors and the need to reestablish family relationships and reenter employment. Health and economic disparities that existed before incarceration are intensified following release, because access to jobs, health insurance, public housing and health benefits may be limited to ex-prisoners, particularly in the United States, and disproportionately in ethnic minority groups. Thus incarceration continues to exert effects leading to increased stress and poor health even after release.

Responses to the Stress of Prison Life

The stress experienced by prisoners has been measured in several ways: by self-report measures of mood, well-being, and mental health; by the incidence and prevalence of physical health outcomes, mental health outcomes, and physiological indicators of stress; and by the most powerful indicator of hopelessness and despair, suicide.

The investigation of the psychological state of inmates and stress-related effects in prisons faces a number of methodological challenges. First, assessing the impact of incarceration is complicated by limited knowledge about the prior health and well-being of inmates. Many studies of stress in prisons that use outcomes such as hospitalization for mental illness

during incarceration are problematic in that incident psychopathology, that is, new onset mental illness, may be difficult to distinguish from prevalent psychopathology, which has been consistently demonstrated as having higher rates in prisoners. Longitudinal studies following prisoners from the first days of incarceration are the only way of studying the emergence of stress resulting from the prison environment rather than ongoing stress from psychosocial disadvantage or existing psychopathology that entails greater risk of conviction in the first place.

Second, the general population is a poor comparison group when assessing the relative prevalence of stress-related disorder in prisoners. Many individuals sent to prison have histories of psychological problems and drug abuse prior to sentencing. Inmates also tend to come from low-income families, to be unemployed, and to have limited education and therefore may have previously encountered negative life events, inadequate health care, and poor living conditions. Drug and alcohol problems frequently precede incarceration; a survey of convicts in state prisons in the United States indicated that 50% of those convicted of violent crimes were under the influence of drugs or alcohol at the time of the offense, and in 1993 drug offenders accounted for 61% of inmates in federal prisons. The high prevalence of some problems among inmates cannot, therefore, be attributed to the prison experience and is confounded by considerable preexisting psychosocial disadvantage in this group.

Nevertheless, anxiety, depression, and complaints of fatigue, sleeplessness, headache, and backache are common. In one study of female convicts, nearly half met the diagnostic criteria for posttraumatic stress disorder (PTSD), and major depression was more prevalent than in the general population. A study of a representative sample of imprisoned juvenile delinquents in California found that half experienced PTSD, with many having witnessed or been involved in traumatic violent episodes. The prevalence of infectious disease, sexually transmitted disease, and tuberculosis is also high compared with the general population. Substantial weight gain has been recorded among female inmates, resulting from the combination of the sedentary lifestyle and the availability of energy-dense high-fat foods. A recent British study investigating subjective health status in female inmates addressed the confounding by lower socioeconomic status in prisoners. Using a comparison group of women from lower-socioeconomic-status sectors of the community, a comparable level of physical health complaints and a greater level of mental health complaints were found in prisoners.

Blood pressure has been found to be elevated in prison inmates compared with matched controls and to increase with the duration of confinement and the degree of crowding in prison. One particularly impressive study used a quasi-experimental approach by measuring the blood pressure of male convicts assigned to individual cells, four-bed cells, and larger dormitories in a medium-security prison. Blood pressure increased to a greater extent in men who were transferred to multiple-occupancy cells and dormitories compared with those who were transferred to other individual cells.

However, a third methodological problem with studies of physiological stress markers is that prison studies have not kept pace with recent advances in measuring such markers as ambulatory blood pressure, cortisol, catecholamines, and inflammation. One notable exception is the Copenhagen Solitary Confinement study, which assessed changes in mood, orientation, and serum cortisol during the first 3 months and compared prisoners kept in solitary confinement (SC) with prisoners in shared cells. Although anxiety, depression, and morning cortisol levels were reduced over time in the non-SC group, such reductions did not occur in the SC group, indicating a lack of adaptation to the prison environment. Many physiological markers of stress are good prognostic indicators of health problems such as autoimmune and cardiovascular diseases, and therefore it may be fruitful to pursue this avenue.

International surveys have found that the suicide rate in prison is between 3 and 11 times greater than that of the general population, with a reported fivefold increase compared with the UK general male population from 1978 to 2003. Most suicides take place within the first few weeks of incarceration and are more common in crowded institutions than other prisons. High rates have been recorded among prisoners on remand and those convicted of murder and serving life sentences. Although suicide is most frequent in young males, self-harm is highly prevalent in female inmates, and recent UK Home Office statistics report a rate of 58.7%. Unlike the pattern in the general population, depression is not a common diagnosis at the time of death, although about one-third of male suicides had some form of psychiatric treatment in the past. Drug or alcohol dependence can be found in a substantial proportion of these cases.

There is also evidence for a European trend toward reinstitutionalization of those with mental health problems in recent years – but in prisons rather than in psychiatric hospitals. A systematic review of prevalence studies in Western countries in 2002 found an overall prevalence of 3.7% for psychotic disorders (10% for major depression), representing a two- to

fourfold increase in mental illness compared with the general population. It is likely that U.S. prisons and jails hold up to twice as many individuals with mental illness than do state psychiatric hospitals; the prison system is not well equipped to support these individuals.

Characteristics of the Individual Convict

Some individuals adapt to incarceration better than others and survive even lengthy prison terms without experiencing high levels of distress. Older convicts tend to have more friends in prison, to have more contact with family and friends outside, and to report greater psychological well-being than younger individuals. A recent UK Home Office report, however, drew attention to the often overlooked but growing population of prisoners aged 60 and over. Although physical health problems in this cohort are often well provided for, psychiatric problems, in particular major depression, are not well treated. Political demand for longer sentencing for some crimes means that this older cohort will continue to provide an ongoing challenge for prison services. As noted earlier, higher levels of anxiety and depression are found in female than in male prisoners, although it is not clear whether this is a specific response to imprisonment or reflects the more general difference in distress found in the adult population at large. Members of ethnic and religious minorities suffer more than members of the majority because prejudice and cultural isolation are added to the other sources of stress present in prison.

In the United States, it has been noted that White prisoners are more likely to attempt and complete suicide than African American prisoners, but the reasons are poorly understood. One suggestion is that being raised in the tough conditions experienced by many African Americans makes adults more resilient to the pressures of prison life. Alternatively, because proportionately more Black than White Americans are incarcerated, they may come from a broader spectrum of society; White inmates may include a higher proportion of marginalized, psychologically vulnerable individuals.

Methods of Reducing Stress among Inmates

The management of stress in prison inmates involves several different levels of intervention. Medical and psychiatric services have an important role in preventing suicide and self-harm, in managing drug dependence, and in treating other psychiatric and neurological problems. Support is often hampered by the reluctance of inmates to divulge personal information to authorities or to reveal weakness. Training prison staff more generally in suicide prevention

is also important. A recent public health initiative in Texas prisons led to improvements in health outcomes, and Telemedicine, allowing rapid access to telephone consultations for physical and mental health problems, proved to be a novel and effective way of delivering health care.

At a more general level, education, training, and other initiatives help build prisoners' self-esteem and confidence in their ability to adapt to civilian life after release. Increasing physical activity and exercise may also be beneficial. Studies have found that inmates who have a job detail or are on work-release programs show less evidence of stress than those who are not active.

Finally, changes at the institutional level may be warranted. It is likely that modifications of prison regimes in relation to aspects such as staff attitudes, privileges, training opportunities, and relocation of inmates to serve their terms near their families could play a key role in alleviating stress among inmates. However, the reality is that stress does not figure highly among the concerns of prison administrators or politicians, some of whom would argue that prison should be stressful.

Prison Staff

Sources of Stress

Research on prison staff members finds that many of the issues cited in the general literature on occupational stress are relevant for them as well. As with other jobs, factors such as work demands, role conflict, low control, and poor social support emerge as relevant. Thus, symptoms of psychological distress and job dissatisfaction are greater among correctional officers who report high work demands coupled with low control and low support from management and colleagues.

However, these problems are exacerbated by the peculiar requirements of working in prison. Prison officers cite the conflict of responsibilities in the job as an underlying source of stress and dissatisfaction. On the one hand, prison staff members are responsible for the welfare of inmates and for the maintenance of their physical and psychological well-being. On the other hand, they are required to maintain security and protect the public from the dangerous and sometimes violent convicts. The prison environment is often volatile, with a high risk of aggressive behavior from inmates. Confrontations with inmates and the threat of physical danger are cited as major sources of stress, irrespective of the age and gender of staff members. Ethnic minority prison officers have been found to experience lower stress levels, especially in the

United States, where ethnic minorities are grossly overrepresented among inmates.

Some differences have been observed between new and long-serving staff. Newly appointed staff members report a lack of support from supervisors and a concern with lack of control over the working environment as particularly problematic. Those who have been working for longer periods are more troubled by the absence of support and respect from society and the lack of variation in the job. The combination of tedium associated with mundane day-to-day activities and the pressure to maintain vigilance in order to react quickly to high-risk incidents appears to be especially stressful.

Many prison officers believe that the responsibilities for security and the welfare of inmates are incompatible obligations. A recent meta-analysis of factors predicting work-related stress revealed an interesting effect of punitive versus rehabilitative attitudes – officers in the United States who endorsed rehabilitative attitudes experienced greater stress, whereas those in Canada with stronger punitive attitudes were more stressed. This suggests that when individuals' beliefs about their role are at odds with the ethos of the prison service in a particular country, stress and conflict are the result.

Outcomes of Stress

High rates of sick leave for both minor complaints (fatigue, headaches, and indigestion) and physical health problems (high blood pressure and peptic ulcers) occur among prison staff. In one study, the rates of coronary heart disease, hypertension, and ulcers were found to be greater among correctional officers than a comparable sample of police officers. The divorce rate in correctional officers is reported as being twice that of blue- and white-collar workers in general. Job dissatisfaction is coupled with the increased prevalence of divorce in this occupational group. The rates of sick leave are greater among staff members in maximum security prisons and in prisons with many drug users than in other correctional institutions.

There is a greater prevalence of psychological distress in prison officers. Much of the research has focused on burnout, with male and female staff members reporting similar high levels of depersonalization and emotional exhaustion. Self-reported anxiety and depression are also elevated compared with the general working population. Female officers are more likely to take sick leave than male officers, but this pattern may result from family demands rather than different levels of stress at work.

Methods of Reducing Stress in Prison Staff

It has been noted in the research literature that many prison officers do not perceive themselves to be stressed but, rather, see their colleagues as being stressed. Pressure not to show the effects of stress may be particularly evident in a male-dominated environment where there is fear that any sign of weakness will be exploited by the inmates. Admitting to feeling stressed may be interpreted as an inability to cope and therefore as not being strong enough for the job at hand. Alternatively, some prison officers may not be able to identify the personal effects of stress in themselves. This suggests that education about stress and how it can be managed may both benefit prison staff members directly and help them to identify the repercussions of stress in inmates.

A number of studies have found that exercise and fitness programs have positive effects. One investigation involving a 6-week strength and aerobic program showed reductions in body weight, skinfold thickness, and cholesterol levels, together with beneficial changes in smoking, alcohol consumption, sleep, nutritional habits, and stress tolerance.

Consultation and efforts to enhance social support at work have been found to counteract the effects of stress. Factors such as goal consensus among staff, a proactive management style, and increasing of decision latitude can be incorporated into the working environment.

Limitations and Conclusion

There are a number of limitations in the studies that have been carried out thus far about stress in prisons. Most have been cross-sectional in nature, identifying associations between aspects of prison experience and diminished well-being, and the causal sequence is uncertain. Longitudinal designs and data on inmates prior to imprisonment are important in the assessment of the impact of stressors because some apparent effects of incarceration may be the direct consequences of prior experience. The nature of the offense, family situation, and previous physical and mental health are all important factors that should be taken into account. Also, the civilian groups with which prisoners are compared in many studies may not be matched appropriately in terms of background characteristics.

The generalizability of results across populations is uncertain. The majority of studies have been conducted on White male adults and may not be relevant to women, ethnic minorities, and juveniles. This is

not mere tokenism because it is apparent from the limited information available that the experience of prison is very different for people of different backgrounds and cultures. The applicability of findings to different prison systems must also be questioned. Justice systems in the United Kingdom and other European countries differ greatly from the U.S. system, and countries also vary in the extent to which individuals with serious psychiatric histories are found in prison. The application of management suggestions derived from one system must be reviewed carefully before their implementation elsewhere.

Nevertheless, the data do indicate that imprisonment is severely stressful for many inmates and that work within the prison system is also associated with higher rates of stress-related dysfunction than comparable jobs. Education and the implementation of specific stress-management programs may be beneficial to both groups and may help ensure that stress in prison is confined within tolerable limits.

See Also the Following Articles

Burnout; Crowding Stress; Prisoners of War; Racial Harassment/Discrimination; Workplace Stress.

Further Reading

- Andersen, H. S. (2004). Mental health in prison populations: a review. *Acta Psychiatrica Scandinavica* 110(supplement 424), 5–59.
- Andersen, H. S., Sestoft, D., Lillebaek, T., et al. (2003). A longitudinal study of prisoners on remand: repeated measures of psychopathology in the initial phase of solitary versus nonsolitary confinement. *International Journal of Law and Psychiatry* 26, 165–177.
- Cooke, D. J., Baldwin, P. J. and Howison, J. (1993). *Psychology in prisons*. London: Routledge.
- DeRosia, V. R. (1998). *Living inside prison walls: adjustment behavior*. Westport, CT: Praeger.
- Hayes, L. M. and Blaauw, E. (eds.) (1997). *Special issue on prison suicide. Crisis: The Journal of Crisis Intervention and Suicide Prevention* 18(4).
- Dowden, C. and Tellier, C. (2004). Predicting work-related stress in correctional officers: a meta-analysis. *Journal of Criminal Justice* 32, 31–47.
- Lindquist, C. H. and Linquist, C. A. (1997). Gender differences in distress: mental health consequences of environmental stress among jail inmates. *Behavioral Sciences & the Law* 15, 503–523.
- Ostfeld, A. M., Kasl, S. V., D'Atri, D. A. and Fitzgerald, E. F. (1987). *Stress, crowding and blood pressure in prison*. Hillsdale, NJ: Lawrence Erlbaum.
- Paulus, P. B. (1988). *Prison crowding: a psychological perspective*. New York: Springer-Verlag.

Prisoners of War

C Tennant

University of Sydney, St. Leonards, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 247–251, © 2000, Elsevier Inc.

Mortality, Physical Morbidity, and Cognitive Impairment
Psychological Morbidity
Etiological Factors
Summary

Glossary

<i>Physical morbidity</i>	Illness and hospital admissions.
<i>Psychological morbidity</i>	Anxiety, depression, posttraumatic stress disorder, and psychiatric admissions.

Mortality, Physical Morbidity, and Cognitive Impairment

Mortality

There have been relatively few studies of physical health and mortality in prisoners of war (POWs) and controls; studies from different countries have produced different findings. The largest series of studies is based on a cohort of U.S. World War II Pacific, European, and Korean POWs and controls assessed by a series of researchers. In the early postwar years, there was increased mortality in all POW groups, due largely to tuberculosis (TB) and postwar trauma. By 1965, no increased mortality appeared in European POWs. Mortality was still increased in Pacific POWs, due to cirrhosis, tuberculosis, and trauma, and in Korean veterans, due to trauma. This increased mortality in Pacific and Korean POWs had persisted for 9 and 13 years, respectively, postwar, whereas death from cirrhosis of the liver had increased from the tenth year of follow-up. In an Australian study of POWs of the Japanese in the 1950s, more deaths from suicide, trauma, cirrhosis, and TB were initially reported. In an Australian controlled cohort study (40 years postwar), no increase in mortality in POWs was found.

Physical Health and Hospital Admissions

In U.S. POWs from all three theaters, general hospital admission rates to 1965 were increased significantly. However, in a 40-year follow-up of a subsample,

no increased rates of hypertension, diabetes, or myocardial infarction were found. A full medical examination of the Australian cohort revealed no differences in current medical disorders. Higher rates of peptic ulcer were, however, generally found in World War II POWs and in U.S. naval POWs in Vietnam; the latter also experienced excessive disorders of the peripheral nervous system.

Thus, postwar peptic ulcer, trauma, TB, and cirrhosis appear more common in World War II POWs than in combatant controls. Although mortality appears to be increased in the U.S. studies, there may have been sampling biases; findings from a well-matched Australian cohort study suggested no difference.

Cognitive Impairment

U.S. World War II and Korean prisoners of war High- and low-weight-loss POWs have been compared with combatant controls. Cognitive impairment (memory and learning) was greatest in POWs with the greater weight loss; there was, however, little difference between low-weight-loss POWs and their controls.

Australian prisoners of war of the Japanese No significant differences in the Australian cohort follow-up (50 years postwar) were found between POWs and combatant controls using detailed psychometric tests; high- and low-weight-loss groups similarly did not differ on any tests or on prevalence of clinically diagnosed dementia. These similarities occurred despite POWs having a greater incidence of depression that could have impaired their psychometric performance, a factor that may have accounted for the positive findings in U.S. studies, which relied on volunteers. Finally, in the same cohort, POWs and controls did not differ on any neuropsychological tests or on brain width (using cerebral X-ray computed tomography, CT, scan); however, POWs, interestingly, had a shorter mean anteroposterior brain diameter. Taking into account methodological problems of some studies, cognitive impairment is not confirmed in POWs postwar.

Psychological Morbidity

Minnesota Multiphasic Personality Inventory Questionnaire Studies (U.S. World War II and Korean Prisoners of War)

U.S. World War II prisoners of war Minnesota Multiphasic Personality Inventory (MMPI) symptom profiles (depression, anxiety, and somatic concern)

were elevated (from highest to lowest) in POWs with chronic posttraumatic stress disorder (PTSD), POWs recovered from PTSD, POWs with other psychiatric diagnoses, POWs with no psychiatric disorder, and normal men. The personality profiles of denial and suppression were also elevated.

Korean prisoners of war and combatant controls Hypochondriasis, depression, hysteria, psychasthenia, paranoia, and ego strengths distinguished the POWs and the controls. The MMPI assesses both personality and psychological symptoms. It shows some unexpected personality differences in POWs and controls, presumably because the personality subscales are sensitive to change in mood. The Eysenck Personality Questionnaire (a more stable measure) showed no differences when used in other Australian POWs and controls.

Psychiatric Admissions and General Psychological Morbidity Data

U.S. prisoners of war (World War II and Korea) Hospital admission rates were significantly higher in the Pacific POWs for 27 of the 33 somatic conditions studied and for all nine of the psychiatric diagnoses, whereas in Korean POWs, admissions were elevated in only nine somatic conditions and three psychiatric conditions. Admission rates in the European POWs were not increased. Bias was possible due to a response rate of only 59% for the initial medical questionnaire. Again, 45 years postwar these POWs had moderately elevated rates of hospitalization and depressive disorders and greatly elevated PTSD; no difference was found in bipolar disorder, schizophrenia, alcoholism, or physical health.

Australian World War II prisoners of war of the Japanese These POWs (88% response) were no more likely to have had psychiatric admissions than combatant controls and indeed had fewer postwar multiple psychiatric admissions. Diagnosed clinical anxiety and major depression in POWs were twice that of controls; alcohol abuse and psychosis were no different. These findings were consistent with their significantly higher scores on questionnaire measures of anxiety and depression. Surprisingly, 10 years later there had been a significant decline in both diagnosed clinical anxiety and depression.

General psychological morbidity is elevated significantly for many years postwar; however, admission data findings vary between U.S. and Australian POWs.

Posttraumatic Stress Disorder

World War II and Korea

U.S. prisoners of war This group has been studied extensively by Engdahl and colleagues and by Sutker

and colleagues. Initial questionnaires were returned by only 35% of the sample surveyed; some 500 of these were then studied 40 years postwar with a 75% response. In the first series of 426 POWs, the rate of PTSD was increased greatly and lifetime prevalence of depression was increased moderately, but bipolar disorder, schizophrenia, and alcoholism were not. In 262 of this sample, 53% met the criteria for lifetime PTSD and 29% for current PTSD; in Japanese POWs, the rates of PTSD were higher, 84% lifetime and 59% current PTSD prevalence. In 62 of the World War II POWs, 50% had PTSD within 1 year of release and 29% has PTSD 40 years later. A lifetime diagnosis of generalized anxiety was found in over 50%. In another study of 915 of the original World War II cohort, 53% of POWs initially responded and 188 of these were interviewed. Sixty-seven percent of this sample had PTSD; 50% had shown no recovery, 24% had some residual symptoms, 39% had mild symptoms, and 29% had fully recovered. It is important to note the possibility of considerable sample selection bias in this series of studies (due to poor response rates), which may have inflated the rates of psychiatric morbidity.

The second U.S. series of studies are those of Sutker and colleagues. Korean and Pacific veterans ($n = 326$) had experienced more extreme trauma and had the highest prevalence rates for lifetime and current mental disorders, including PTSD. In a small subsample, 77% of the Pacific POWs had current PTSD and 78% a lifetime diagnosis, compared with 18 and 29%, respectively, in controls. In Korean POWs, PTSD symptoms were found in 90–100%. Finally, World War II aviators were less likely to have PTSD than nonaviators. The sample bias was also probable in these studies because the subjects were selected from attenders at a Veteran's Administration (VA) medical center. Other studies have used questionnaires to assess PTSD with similar high prevalence rates in military and civilian POWs.

Canadian prisoners of war Higher rates of PTSD and other symptoms were found in POWs from Europe 50 years postwar, whereas in former Far East POWs, 30% had current PTSD 50 years postwar and 90% had at least one PTSD symptom.

Australian prisoners of war In the Australian cohort, no subjects appeared to meet full *Diagnostic and Statistical Manual of Mental Disorders*, 3rd ed. (DSM-III) criteria for PTSD 40 years postwar, although a majority of the sample had some PTSD symptoms and 50% had a clinical affective disorder or anxiety. At a further follow-up 10 years later, 4% of subjects had been diagnosed as having PTSD using DSM-IV criteria; there was, however,

a significant decline in psychiatric disorders overall, from 50 to 36%.

French and German prisoners of war of the Russians Forty years postwar, 71% of a selected sample (response 40%) had symptoms consistent with PTSD, whereas in a randomly selected subsample, the frequency of individual PTSD symptoms ranged from 84% (for recurring dreams) to 18% (for panic symptoms).

Dutch Resistance Veterans Sample veterans were studied 40 years postwar (76% response); half had been imprisoned and half had not. Interestingly, 54% of both groups had PTSD and a similar rate of other psychological symptoms. Paradoxically, this study followed an earlier study showing that, 15 years postwar, imprisoned Dutch resistance fighters had less incidence of psychiatric disorders than those not imprisoned.

There is little doubt that significant rates of PTSD are found long after the war, although not in all samples; the very high rates of PTSD found, particularly in U.S. samples, may be due in part to various forms of sampling bias.

Other Conflicts

Israeli prisoners of war of the Yom Kippur War POWs, combatants (who had experienced combat stress reactions), and normal combat veterans were compared; the prevalence of PTSD more than 18 years postwar in the three groups was 23, 37, and 14%, respectively, whereas current PTSD prevalence was 13, 13, and 3%. Psychophysiological complaints were also more frequent among POWs, as were trauma-related and general psychiatric symptoms. In terms of coping, those with high sensation-seeking qualities adjusted better to the stress of captivity and had fewer PTSD or other psychiatric symptoms; active coping strategies were used by high sensation seekers, whereas low sensation seekers used detachment and denial. In general, poor psychological adjustment was associated with more helplessness, more suffering, and less active coping, as well as hostility and ambivalence.

British prisoners of war from the Persian Gulf War In POWs studied both 6 and 18 months postwar, psychological symptoms (using the General Health Questionnaire, GHQ) reduced significantly; a poorer psychological outcome was associated with POWs witnessing physical violence and self-perceived deterioration in physical and mental health during captivity.

Balkan prisoners of war In Croatian POWs, 36% had postwar psychiatric disorder; camp conditions contributed to psychiatric symptoms (particularly threats against family members); however, group cohesion diminished the level of disorder. In another sample, only 4.5% met the full criteria for PTSD, 28% had no PTSD symptom cluster, and the remainder had some PTSD symptoms.

Comorbidity of Psychiatric Disorders in Prisoners of War

In one selected sample of Korean POWs, PTSD was found the most; depressive comorbidity occurred in 75%, anxiety in 45%, and alcohol abuse in 20%.

In the other U.S. series, of 262 World War II and Korean POWs (who had sought mental health treatment), 53% met the criteria for lifetime PTSD; 29% of those with PTSD had other axis I disorders. In a subsample, depression and anxiety coexisted in 61% of the sample. Depression was found in 10% of those without PTSD, in 23% of those with recovered PTSD, and in 61% of those with ongoing PTSD. The Australian POW study also found that anxiety and depression coexisted in a significant portion of the sample.

Etiological Factors

The severity of the stressful experiences of POWs differed in different theaters. Both Korean and Vietnam POWs were interrogated, brain-washed, and tortured, whereas Vietnam POWs were also kept for long periods in solitary confinement. POWs of the Japanese suffered gross physical deprivations, malnutrition, random brutality, hard physical labor, major physical illness, and ultimately a high mortality rate. European theater POWs did not suffer as severely, although resistance fighters (such as the Dutch) had a similar experience to Holocaust survivors. The POW experience in more recent conflicts has generally been briefer and less brutal. Studies addressing etiology have been somewhat piecemeal; World War II POWs provide the most consistent findings.

World War II Prisoners of War

In U.S. POWs, studied by Engdahl and colleagues, generalized anxiety and depressive disorders were related to captivity severity; preexisting psychopathology and family history of psychiatric disorder were not predictive. Chronic depression was predicted by a younger age at time of capture, lower education, more medical symptoms during captivity, and poorer social supports after release. Specific

captivity stressors (weight loss, torture, and disease during captivity) correlated strongly with psychological morbidity and PTSD, whereas age and education were protective. These findings have been largely replicated in the other U.S. series of studies by Sutker and in Canadian, French, and Australian POWs.

Vietnam Prisoners of War

Captivity trauma was greater in POWs captured before 1969 and was associated with more subsequent psychiatric morbidity. Aviators were more resilient to captivity than nonaviators; this was attributed to their higher educational status and military rank and better personal resources. Furthermore, there was no particular coping style indicative of psychiatric illness in this group. Paradoxically, still-serving POWs reported that the harshness of the war experience was associated with a belief of having actually benefited from imprisonment.

The severity of the captivity experience, especially as manifest in gross physical hardship and deprivation, predicted the subsequent short- and long-term psychological morbidity.

Summary

The POW experience varied according to the theater of war; the Japanese were particularly brutal captors. In the early postwar years, excess physical morbidity and mortality occurred particularly in Pacific POWs. Although this declined subsequently, deaths from cirrhosis showed a delayed increase.

Significant psychiatric morbidity has been the most consistent finding in all the studies, including conflicts since World War II, and morbidity has persisted for many years postwar. Studies in the United States have focused specifically on PTSD and show very high rates of prevalence, but sampling bias may, to some degree, contribute to their particularly high rates. Other anxiety disorders and depression are also increased, but psychoses are not. The comorbidity of disorders is common. Furthermore, the high rates of psychological disorder correlate with a range of indices of captivity severity, whereas some sociodemographic factors seem protective. Former POWs who served their countries well and suffered terribly

continue to live with the legacy of the inhumanity of their captors during war.

See Also the Following Articles

Combat Stress Reaction; Concentration Camp Survivors; Korean Conflict, Stress Effects of; Persian Gulf War, Stress Effects of; Prison; Vietnam Veterans, Postwar Experiences and Health Outcomes; War-Related Post-traumatic Stress Disorder, Treatment of.

Further Reading

- Beebe, G. W. (1975). *American Journal of Epidemiology* 101, 400–422.
- Bisson, J. I., et al. (1998). *British Journal of Medicine and Psychology* 71, 247–252.
- Eberly, R. E. and Engdahl, B. E. (1991). *Hospital and Community Psychiatry* 42, 807–813.
- Engdahl, B. and Dikel, T. N. (1997). *American Journal of Psychiatry* 154, 1576–1581.
- Engdahl, B. E., et al. (1991). *Journal of Nervous and Mental Disease* 179, 181–187.
- Engdahl, B. E., et al. (1991). *Social, Psychiatry and Psychiatric Epidemiology* 26, 63–67.
- Goulston, K. J. (1985). *Medical Journal of Australia* 143, 6–10.
- Jorm, A. F., et al. (1997). *Personality and Individual Differences* 23, 371–377.
- Kluznic, J. C., et al. (1986). *American Journal of Psychiatry* 143, 1443–1446.
- Nefzer (1970). *American Journal of Epidemiology* 91, 123–138.
- Neria, Y. and Solomon, Z. (1998). *Journal of Nervous and Mental Disease* 186, 174–182.
- Solomon, Z., et al. (1998). *European Journal of Personality* 12, 271–285.
- Sulway, M. R., et al. (1996). *Neurology* 46, 650–655.
- Sutker, P. B. and Allain, A. N. (1991). *Psychological Reports* 68, 279–284.
- Sutker, P. B. and Galina, Z. H. (1990). *Journal of Consulting and Clinical Psychology* 58, 323–328.
- Tennant, C. and Fairley, M. J. (1997). *Journal of Nervous and Mental Disease* 185, 686–689.
- Tennant, C. and Goulston, K. (1986). *American Journal of Psychiatry* 143, 618–621.
- Tennant, C., et al. (1986). *Psychology and Medicine* 16, 833–839.
- Wilson, J. P. and Raphael, B. (eds.) (1993). *International handbook of traumatic stress syndromes*. New York: Plenum Press.

Problem-Solving Skills Training

A M Nezu and C M Nezu

Drexel University, Philadelphia, PA, USA

T J D'Zurilla

State University of New York, Stony Brook,
NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
A M Nezu, C M Nezu and T J D'Zurilla, volume 3, pp 252–256,
© 2000, Elsevier Inc.

Social Problem Solving
The Problem-Solving Process
Problem-Solving Deficits
Problem Solving as a Moderator of Stress
Efficacy of PST
PST Training Guidelines

Glossary

<i>Effective solutions</i>	Coping responses that achieve one's problem-solving goals, but that also simultaneously maximize positive consequences (i.e., benefits) and minimize negative consequences (i.e., costs).
<i>Problems</i>	Specific existing or anticipated situations that demand responses for adaptive functioning, but that are not met by effective coping responses from the person confronted by them due to the presence of certain obstacles (e.g., ambiguity, uncertainty, conflicting demands, lack of resources, or uniqueness).
<i>Problem orientation</i>	The set of relatively stable cognitive-affective schemas that represent a person's generalized beliefs, attitudes, and emotional reactions about problems in living and one's ability to successfully cope with such problems.
<i>Problem-solving style</i>	The core cognitive-behavioral activities that people engage in when attempting to cope with problems in living.
<i>Social problem solving</i>	The cognitive-behavioral process by which individuals view the nature of problems in living and direct their attempts at coping with such stressors.
<i>Solution</i>	A coping response designed to alter the nature of the problematic situation, one's negative emotional reactions to it, or both.

Social Problem Solving

Throughout history, psychologists and philosophers have argued that an essential part of being human is the capacity to solve problems. Perhaps the most important idea emanating from this belief is the notion that problem-solving ability contributes significantly to social competence and overall psychological well-being, as the ability to cope with and resolve everyday stressful problems has been found to be strongly related to overall personal and social functioning.

Training individuals to become better problem solvers in order to facilitate their ability to cope with stressful situations has been referred to in the psychotherapy and counseling literature as social problem-solving therapy in order to highlight the social and interpersonal context in which real-life problem solving occurs. Social problem solving is defined as the cognitive-behavioral process by which individuals view the nature of problems in living and direct their attempts at coping with such stressors. Relevant problem-solving therapy (PST) goals broadly include altering the stressful nature of the situation itself (e.g., overcoming obstacles to a personal goal), changing one's reactions of distress to such problems (e.g., acceptance that a goal cannot be reached), or both.

Problems can be single events (e.g., obtaining a loan from a bank), a series of related problems (e.g., continuous arguments with a spouse), or chronic situations (e.g., a major chronic illness, such as cancer). Situations become problems when effective responses are required in order for the person to cope adaptively, but when such responses are not immediately available or identifiable due to the presence of various obstacles, including ambiguity, unpredictability, conflicting demands, deficient skills, or lack of resources.

Solutions are coping responses designed to alter the nature of the problematic situation, one's negative emotional reactions to it, or both, whereas effective solutions are those coping responses that not only achieve such goals, but also simultaneously maximize positive consequences (i.e., benefits) and minimize negative effects (i.e., costs).

The Problem-Solving Process

Problem-solving outcomes are largely determined by two general but partially independent dimensions – problem orientation and problem-solving style. Problem orientation is the set of relatively stable

cognitive-affective schemas that represent a person's generalized beliefs, attitudes, and emotional reactions about problems in living and one's ability to successfully cope with such problems. Problem orientation can be either positive or negative. A positive orientation is one that involves a tendency to appraise problems as challenges, be optimistic in believing that problems are solvable, perceive one's own ability to solve problems as strong, and believe that successful problem solving involves time and effort. Conversely, a negative problem orientation is one that involves the tendency to view problems as threats, expect problems to be unsolvable, doubt one's own ability to solve problems successfully, and become frustrated and upset when actually faced with problems.

Problem-solving style refers to those core cognitive-behavioral activities that people engage in when attempting to cope with problems in living. Three differing styles have been identified, one that is adaptive, while the other two reflect maladaptive ways of coping. Rational problem solving is the constructive problem-solving style that involves the systematic and planful application of certain specific skills, each of which makes a distinct contribution toward the discovery of an adaptive solution or coping response in a problem-solving situation. Rational problem solving involves the following four skills: problem definition and formulation, generation of alternatives, decision making, and solution implementation and verification. The goal of problem definition and formulation is to delineate the reasons why a given situation is a problem (e.g., the presence of obstacles), as well as to specify a set of realistic goals and objectives to help guide further problem-solving efforts. The purpose of the generation of alternatives task is to create, using various brainstorming principles, a large pool of possible solutions in order to increase the likelihood that the most effective ideas will be ultimately identified. The goal of decision making is to conduct a systematic cost-benefit analysis of each alternative by identifying and then weighing their potential positive and negative consequences if carried out, and then, based on this evaluation, to develop an overall solution plan. Finally, the purpose of solution implementation and verification is to carry out the solution plan, monitor and evaluate its effectiveness, and troubleshoot if the outcome is unsatisfactory.

Two additional problem-solving styles have been identified, both of which are dysfunctional or maladaptive in nature. An impulsive/careless style involves a generalized response pattern characterized by impulsive, hurried, and careless attempts at problem resolution. Although the individual characterized by this style actively attempts to apply various strategies to address problems, such attempts are narrow, hurried,

or incomplete. For example, a person with this style is likely to consider only a few solution alternatives, often impulsively implementing the first idea that comes to mind. In addition, the narrow range of options and their consequences are scanned quickly, carelessly, and unsystematically.

Avoidance style is a second maladaptive problem-solving pattern, this one characterized by procrastination, passivity, and overdependence on others to provide solutions. This type of individual generally avoids problems rather than confronting them head on, puts off addressing problems for as long as possible, waits for problems to resolve themselves, and attempts to shift the responsibility for solving his or her problems to other people. In general, both styles lead to ineffective or unsuccessful problem resolution.

Problem-Solving Deficits

Important differences have been identified in individuals characterized as effective versus ineffective problem solvers. In general, when compared to their effective counterparts, ineffective problem solvers report a greater number of life problems, more health and physical symptoms, more anxiety, more depression, and more psychological maladjustment. In addition, a negative problem orientation has been found to be associated with negative moods under routine and stressful conditions in general, as well as pessimism, negative emotional experiences, and clinical depression. Persons with a negative orientation also tend to worry and complain more about their health.

In addition, problem-solving deficits have been found to be significantly related to poor self-esteem, hopelessness, suicidal risk, self-injury, anger proneness, increased alcohol intake and substance risk taking, personalities difficulties, criminal behavior, alcoholism, secondary physical complications among persons with spinal cord injuries, premenstrual and menstrual pain, physical health problems, diminished life satisfaction, physical problems among adult cancer patients, and pain severity among adult cardiac patients.

Problem Solving as a Moderator of Stress

How people cope with stressful experiences, including major events (e.g., undergoing a divorce, dealing with the death of a spouse) and daily problems (e.g., continued arguments with a co-worker, limited financial resources) can, in part, determine the degree to which they will experience long-lasting psychological distress, particularly depression. Continued successful attempts at problem resolution will lead to a reduction or minimization of immediate emotional

distress and a reduced likelihood of long-term negative affect (i.e., clinical depression). Alternatively, if one's problem-solving coping skills are ineffective, or if extreme emotional distress impacts negatively on one's coping efforts, resulting in reduced motivation, inhibition of problem-solving performance, or both, then the likelihood of long-term emotional distress will be increased. Further, such negative outcomes can lead to the exacerbation of existing problems and the creation of new ones, which in turn can lead to another major stressful life event, and so forth. As such, how one copes with problems can lead to either an escalation or attenuation of the stress process. For example, research has demonstrated that under similar levels of high stress, individuals with poor problem-solving skills experience significantly higher levels of depression and anxiety than persons characterized by more effective problem-solving skills, supporting the notion that problem solving serves to attenuate the negative effects of stress.

Efficacy of PST

If effective problem-solving skills serve as an important buffering factor regarding the stress process, training individuals in such skills should lead to a decrease in emotional distress and improvement in overall psychological functioning. In fact, PST has been shown to be effective in a wide range of clinical populations, psychological problems, and the distress associated with chronic medical disorders. These include unipolar depression, geriatric depression, distressed primary care patients, social phobia, agoraphobia, obesity, coronary heart disease, adult cancer patients, schizophrenia, mentally retarded adults with concomitant psychiatric problems, HIV risk behaviors, drug abuse, suicide, childhood aggression, and conduct disorder.

In addition to its applicability to a variety of patient populations, PST also appears to be flexible with regard to treatment goals and methods of implementation. For example, it can be conducted in a group format, on an individual and couples basis, and as part of a larger psychosocial intervention package, and can be delivered by telephone. It can also be applied as a means of helping patients to overcome barriers associated with successful adherence to other medical or psychosocial treatment protocols.

PST Training Guidelines

Specific PST therapy objectives include (1) enhancing individuals' positive orientation and application of the four rational problem-solving tasks, and (2) minimizing their negative orientation and tendency to

engage in dysfunctional problem-solving style activities (i.e., impulsive or careless attempts to cope with problems; avoidance of problems). PST interventions include didactic explanations, training exercises, practice opportunities, and homework assignments geared to foster practice between training sessions.

Problem Orientation

Training in this problem-solving component is geared to facilitate the following: positive self-efficacy beliefs (the perception that people can improve their quality of life through effective coping and problem solving), beliefs that problems are inevitable (accepting the notion that it is common and normal to experience a wide range of problems), the ability to identify problems accurately when they occur, and the ability to inhibit emotional reactions that can lead to impulsive reactions or avoidance.

A variety of training approaches can be used to foster a positive problem orientation. One technique is the reverse advocacy role-play strategy. According to this strategy, the therapist pretends to adopt a particular belief about problems and asks the patient to provide reasons why that belief is irrational, illogical, incorrect, or maladaptive. Such beliefs might include the following statements: "Problems are not common to everyone; if I have a problem, that means I'm crazy," "There must be a perfect solution to this problem," "I'll never be the same again." At times when the patient has difficulty generating arguments against the therapist's position, the counselor then adopts a more extreme form of the belief, such as "No matter how long it takes, I will continue to try and find the perfect solution to my problem." This procedure is intended to help patients identify alternative ways of thinking and then to dispute or contradict previously held negative beliefs with more adaptive perspectives.

Patients are also taught to use feelings or emotions as cues that a problem exists by using visual images of a red traffic stop sign as a signal to stop and think. In essence, patients are taught to recognize various situations as problems and to label them as such. Accurately labeling a problem as a problem serves to inhibit the tendency to act impulsively or automatically in reaction to such situations. It also facilitates the tendency to approach or confront problems, rather than to avoid them.

Problem Definition and Formulation

Problem definition can be likened to mapping a guide for the remainder of the problem-solving process. The major focus of this task is to better understand the nature of the problem and to set clearly defined and

reasonable goals. In other words, locating a specific destination on a map makes it easier to find the best route to get there. Training in problem definition and formulation focuses on the following five specific tasks: gathering all available information about the problem, using clear and unambiguous language, separating facts from assumptions, identifying the factors that make the situation a problem, and setting realistic problem-solving goals.

Generation of Alternatives

When generating alternative solutions to a problem, PST encourages broad-based, creative, and flexible thinking. In essence, patients are taught various brainstorming strategies (e.g., the more the better, defer judgment of ideas until a comprehensive list is created). This helps to increase the likelihood that the best or most effective solution ideas will be discovered.

Decision Making

Once a list of alternative options has been generated, the problem solver begins to systematically and thoroughly evaluate the potential for each solution to meet the defined goal(s). Training in this component helps the patient to use the following criteria to conduct a cost-benefit analysis based on the utility of each alternative solution: the likelihood that the solution will meet the defined goal, the likelihood that the person responsible for solving the problem can actually carry out the solution plan optimally, personal (i.e., effects on oneself) and social (i.e., effects on others) consequences, and short- and long-term effects.

Solution Implementation and Verification

This last rational problem-solving task involves first carrying out the solution plan and then monitoring and evaluating the consequences of the actual outcome. PST encourages the patient to practice the performance aspect of solution implementation as a means of enhancing the probability that it will be carried out in its optimal form. Once the plan is under way, the patient is encouraged to monitor the actual results. Using this information allows the individual to evaluate the results by comparing the actual outcome with his or her expectations or predictions about the outcome.

Supervised Practice

After the majority of training has occurred, the remainder of PST should be devoted to practicing the newly acquired skills and applying them to a variety of stressful problems. Beyond actually solving stressful problems, ongoing in-session practice serves

three additional purposes: the patient can receive professional feedback from the therapist, increased facility with the overall PST model can decrease the amount of time and effort necessary to apply the various problem-solving tasks with each new problem, and practice fosters maintenance and generalization of the skills.

The number of practice sessions required after formal PST training often is dependent upon the competency level a patient achieves, as well as on the actual improvement in his or her overall quality of life. In the research protocols that have found PST to be an effective cognitive-behavior therapy intervention, the number of included sessions has ranged from 8 to 12.

See Also the Following Articles

Anxiety; Cognitive Behavioral Therapy; Depression and Manic-Depressive Illness.

Further Reading

- Allen, S. M., Shah, A. C., Nezu, A. M., et al. (2002). A problem-solving approach to stress reduction among younger women with breast carcinoma: a randomized controlled trial. *Cancer* 94, 3089–3100.
- Chang, E. C., D’Zurilla, T. J. and Sanna, L. J. (eds.) (2004). *Social problem solving: theory, research, and training*. Washington, D.C.: American Psychological Association.
- D’Zurilla, T. J. and Nezu, A. M. (1999). *Problem-solving therapy: a social competence approach to clinical intervention* (2nd edn.). New York: Springer.
- D’Zurilla, T. J., Nezu, A. M. and Maydeu-Olivares (2002). *Social Problem-Solving Inventory-Revised (SPSI-R): technical manual*. North Tonawanda, NY: Multi-Health Systems.
- D’Zurilla, T. J., Nezu, A. M. and Maydeu-Olivares, A. (2004). Social problem solving: theory and assessment. In: Chang, E. C., D’Zurilla, T. J. & Sanna, L. J. (eds.) *Social problem solving: theory, research, and training*, pp. 11–27. Washington, D.C.: American Psychological Association.
- Nezu, A. M. (1987). A problem-solving formulation of depression: a literature review and proposal of a pluralistic model. *Clinical Psychology Review* 7, 122–144.
- Nezu, A. M. (2004). Problem solving and behavior therapy revisited. *Behavior Therapy* 35, 1–33.
- Nezu, A. M. and D’Zurilla, T. J. (1989). Social problem solving and negative affective states. In: Kendall, P. C. & Watson, D. (eds.) *Anxiety and depression: distinctive and overlapping features*, pp. 285–315. New York: Academic Press.
- Nezu, A. M. and Perri, M. G. (1989). Problem-solving therapy for unipolar depression: an initial dismantling investigation. *Journal of Consulting and Clinical Psychology* 57, 408–413.

- Nezu, A. M., Nezu, C. M. and Perri, M. G. (1989). *Problem-solving therapy for depression: theory, research, and clinical guidelines*. New York: Wiley.
- Nezu, A. M., Nezu, C. M., Friedman, S. H., Faddis, S. and Houts, P. S. (1998). *Helping cancer patients cope: a problem-solving approach*. Washington, D.C.: American Psychological Association.
- Nezu, A. M., Nezu, C. M., Felgoise, S. H., McClure, K. S. and Houts, P. S. (2003). Project genesis: assessing the efficacy of problem-solving therapy for distressed adult cancer patients. *Journal of Consulting and Clinical Psychology* 71, 1036–1048.
- Nezu, C. M., D’Zurilla, T. J. and Nezu, A. M. (2005). Problem-solving therapy: theory, practice, and application to sex offenders. In: McMurren, M. & McGuire, J. (eds.) *Social problem solving and offenders: evidence, evaluation and evolution*, pp. 103–123. Chichester, UK: Wiley.
- Perri, M. G., Nezu, A. M., McKelvey, W. F., et al. (2001). Relapse prevention training and problem-solving therapy in the long-term management of obesity. *Journal of Consulting and Clinical Psychology* 69, 722–726.

Prolactin and Stress

G Tolis

Hippokrateion Hospital, Athens, Greece, and McGill University, Montreal, Canada

G Rombopoulos, D Kaltsas, E Katounda, V Kaltzidou and N Angelopoulos

Hippokrateion Hospital, Athens, Greece

© 2007 Elsevier Inc. All rights reserved.

Glossary

<i>Catabolism</i>	A wasting process.
<i>Dopamine</i>	A catecholamine neurotransmitter that mediates nerve signaling in the central nervous system.
<i>Opiergic compounds</i>	Opiates or compounds that have an opiate-like effect, for example, morphine, methadone, and heroin.
<i>Prolactin (PRL)</i>	Polypeptide produced in the pituitary with somatolactotropic actions, essential for the preparation of the breast during pregnancy and to assure postpartum lactation.
<i>Somatomammotropic effect</i>	The growth of breast function regulatory effect.

Prolactin (PRL), a somatolactotropic pituitary hormone of paramount importance in lactation, is under tonic inhibition primarily by the tubero-infundibular dopaminergic neurons (TIDA). Psychotropic medications (the class of butyrophenones, i.e., haloperidol) increase PRL levels, whereas dopamine agonists (e.g., apomorphine and bromocriptine) suppress PRL secretion. Recent advances in neurochemistry have made drugs available that have antipsychotic actions without altering PRL secretion (e.g., clozapine). Thyrotropin-releasing hormone (TRH) increases serum PRL levels by a direct effect on the anterior pituitary

gland, as do many of the antipsychotic drugs. The latter act as dopamine receptor antagonists at the pituitary prolactotropes, thereby blocking the normal inhibitory effect of endogenous dopamine on PRL secretion. Dopamine, a key PRL inhibitory factor, is released from the hypothalamus and transmitted to the anterior pituitary gland by way of the hypophyseal portal vessels. Opiates increase PRL levels by a central action that may involve dopamine neurons. Hyperprolactinemia linked to a microprolactinoma is clinically manifested by disturbed gonadal function, decreased libido, and various psychological disturbances (i.e., anxiety and depressive states) that can be ameliorated in part by the administration of dopamine agonists and/or serotonin antagonists. Such observations, as well as the phenomenon of pseudopregnancy, raise the possibility of PRL being involved in central nervous system (CNS) circuits beyond those involved in pregnancy and lactation.

Animal studies have indicated that PRL receptor-deficient mice have hyperprolactinemia, presumably due to the lack of feedback stimulation of TIDA neurons, and that the chronic elevation of PRL secretion occurs in rodents lacking a functional D₂ dopamine receptor. These findings suggest that PRL gains access to the CNS, presumably, via the arcuate-median eminence region, which has an incomplete blood–brain barrier. In experiments transplanting the pituitary to the kidney, there is excess PRL production because the pituitary grafts are remote from the site of dopamine release (in the hypothalamus).

The demonstration that there is PRL mRNA in the hypothalamus and other brain areas suggests that the behavioral effects assigned to PRL may relate to its presence in monoaminergic and peptidergic neurons. Immunohistochemical and *in situ* hybridization

studies revealed that PRL receptors are expressed in the arcuate, periventricular and medial preoptic nuclei both in male and nonpregnant female rats. In addition, during pregnancy and lactation PRL receptor immunoreactivity becomes demonstrable in the paraventricular nucleus (PVN) and ventromedial nucleus. The recruitment of a pleiad of neural structures during pregnancy and lactation assists in the provision of psychosomatic balance, needed for adequate nourishment of the mother and the fetus. The suppression of anxiety, reduced stress response, and expression of maternal behavior are linked to the activities of the PVN and medial preoptic nucleus, whereas the activation of the supraoptic ventromedial and arcuate nuclei are involved in regulating fuel and electrolyte balance.

During late pregnancy and lactation, the acute stress response is attenuated; it is hypothesized that this involves the PVN and the hypothalamic-pituitary-adrenal axis. The observation that PRL has dose-dependent anxiolytic effects in males and nonpregnant female rats (effects neutralized by the administration of antisense nucleotides against the PRL receptor), strongly suggests that PRL has a role as an endogenous anxiolytic. Stress-activated circuits enhance the release of cortical and medullary adrenal biological response modifiers – the glucocorticoids and catecholamines, respectively. When insulin-induced hypoglycemia (IIH) is used as a stressor, there may be an elevation of serum growth hormone (GH) and PRL. The blockade of PRL and enhancement of GH can be effected by dopamine agonists; an earlier clinical observation suggested that cyproheptadine blunts PRL release secondary to IIH, implicating indirectly the participation of serotonergic pathways. Whether the latter neurotransmitter is involved in the PRL release associated with the stress of anesthesia, sleep, nipple stimulation, gynecological exam, or exercise is unknown. Certainly, in contrast to dopamine and dopamine agonists, which block TRH-induced PRL release, serotonin (5-HT) antagonists are ineffective.

Beyond these observations about acute short-lived stress, there are some clinical observations that have stimulated research on psychological-psychiatric distress and PRL dynamics, independently of the activation of the hypothalamic-pituitary-adrenal axis. A woman on a birth-control pill in her teens developed breast engorgement and spontaneous bilateral galactorrhea, which reappeared 5 months later following general anesthesia for an orthopedic procedure. She had postpill amenorrhea with normoprolactinemia and no pituitary tumor; bromocriptine resulted in the cessation of galactorrhea and the reappearance of menses. This and other reports of dopamine agonist use in normoprolactinemic amenorrhea raise the

question: what is normoprolactinemia in aberrant psychological situations? In earlier observations, during acrobatic flight, we observed a significant rise in serum PRL only in novices and not in trained pilots. Transient PRL elevations have been reported during medical interviews, before medical and/or competitive examinations, during argumentative conflicts in neurotic women, and during panic attacks (in the last, it was mentioned that the PRL increase was more consistent than that observed for cortisol, GH, or catecholamines).

In a study carried out in epileptics with a non-lateralized focus following galvanic stimulation of implanted electrodes (in the amygdala and hippocampus), we observed different GH and PRL responses in the corticomedial and basolateral amygdaloid regions. These and other reports (i.e., the different profiles of serum PRL among patients with unipolar vs. bipolar disorders) point to the involvement not only of the infundibular but also of mesocortical and mesolimbic dopamine systems controlling surges, circadian, and sleep-related PRL regulatory circuits. It was recently reported that hyperprolactinemic amenorrheic women had significantly higher symptom questionnaire scores on hostility, anxiety and depression than normoprolactinemic amenorrheic women or women with normal menstrual cycles. Thirty of these women met the *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn.; DSM-III) criteria for major depressive disorder, opening a PRL window for biological psychiatric research. An example along these lines is pseudopregnancy, which is considered a psychiatric condition in DSM-III (hysterical conversion neurosis); that is, the woman, firmly believing that she is pregnant, develops many of the symptoms and signs of pregnancy. In the absence of a prolactinoma, pseudopregnancy had been identified by Hippocrates: "If a woman is not pregnant and has not given birth, produces milk and her menstruation has stopped."

It is beyond any doubt that stress-induced hyper- or normoprolactinemic anovulation operates protectively for both the mother and the potential conceptus, for which the *milieu interieur* could be hostile. In the lactating, breast-feeding mother, the psychosomatic and immune well-being of the family (father, mother, and newborn) complex is preserved because breast-feeding inhibits ovulation and associated side effects. The immunoprotective actions of the somatolactotropic hormones and the maintenance of euthyroidism have been proposed as immunoprotective mechanisms ensuring reduced susceptibility to stress-induced disease. It is hoped that, in women with postpartum depression, hypnosis, psychotherapy, and the proper

choice of psychobiological response modifiers can be selected so that the hormonal-metabolic homeostasis is preserved.

Further Reading

- Bridges, R. S., Numan, M., Ronsheim, P. M., et al. (1990). Central prolactin infusions stimulate maternal behavior in steroid-treated, nulliparous female rats. *Proceedings of the National Academy of Sciences USA* **87**, 8003–8007.
- Dorshkind, K. and Horseman, N. (2000). The roles of prolactin, growth hormone, insulin-like growth factor-I, and thyroid hormone in immune system development and function: insights from genetic models of hormone and hormone receptor deficiency. *Endocrine Review* **21**, 292–312.
- Dutt, A., Kaplitt, M. G., Kow, L. M., et al. (1994). Prolactin, central nervous system and behavior: a critical review. *Neuroendocrinology* **59**, 413–419.
- Fava, G. A., Fava, M., Kellner, R., et al. (1981). Depression hostility and anxiety in hyperprolactinemic amenorrhea. *Psychotherapy and Psychosomatics* **36**(2), 122–128.
- Grattan, D. R. (2002). Behavioural significance of prolactin signalling in the central nervous system during pregnancy and lactation. *Reproduction*. **123**(4), 497–506.
- Hippocrates (1526) *Opera omnia*. Venice: Aldine Press.
- Kelley, K. W. and Dantzer, R. (1991). Growth hormone and prolactin as natural antagonists of glucocorticoids in immunoregulation. In: Plotnikoff, N., Murgo, A., Faith, R. & Wybran, J. (eds.) *Stress and immunity*, pp. 433–452. Boca Raton, FL: CRC Press.
- Pinter, E. J., Tolis, G., Guyda, H., et al. (1979). Hormonal and free fatty acid changes during strenuous flight in novices and trained personnel. *Psychoneuroendocrinology* **4**(1), 79–82.
- Sobrinho, L. G., Simoes, M., Barbosa, L., et al. (2003). Cortisol, prolactin, growth hormone and neurovegetative responses to emotions elicited during an hypnoidal state. *Psychoneuroendocrinology* **28**(1), 1–17.
- Tolis, G., Dent, R. and Guyda, H. (1978). Opiates, prolactin, and the dopamine receptor. *Journal of Clinical Endocrinology and Metabolism* **47**(1), 200–203.
- Tolis, G., Goldstein, M. and Friesen, H. G. (1973). Functional evaluation of prolactin secretion in patients with hypothalamic-pituitary disorders. *Journal of Clinical Investigation* **52**(4), 783–788.
- Tolis, G., Labrie, F., Martin, J. B. and Naftolin, F. (1979). *Clinical neuroendocrinology: a pathophysiological approach*. New York: Raven Press.
- Tolis, G., Stefanis, C., Mountokalakis, T. and Labrie, F. (1984). *Prolactin and prolactinomas: current developments*. New York: Raven Press.
- Tolis, G. and Woods, J. F. (1980). Epilepsy and prolactin release in man: studies during spontaneous seizures or those induced via stereotactically placed electrodes in the human brain. (Abstract, p.465) Program of the 62nd Annual Meeting of the Endocrine Society. Washington.
- Torner, L., Toschi, N., Pohlinger, A., et al. (2001). Anxiolytic and anti-stress effects of brain prolactin: improved efficacy of antisense targeting of the prolactin receptor by molecular modeling. *Journal of Neuroscience* **21**, 3207–3214.

Pro-opiomelanocortin (POMC)

A B Bicknell

University of Reading, Reading, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A B Bicknell and P J Lowry, volume 3, pp 257–265, © 2000, Elsevier Inc.

Introduction

Discovery of POMC

Sites and Tissue-Specific Processing of POMC

Molecular Basis of Tissue-Specific Processing

Regulation of the Expression of the POMC Peptides

Posttranslational Modifications of POMC

Biological Activity of the POMC Peptides

Conclusions

Glossary

Adrenocorticotropin (ACTH)

A 39-amino-acid peptide released from the pituitary in response to stress that acts on the adrenal cortex promoting the synthesis and release of corticosteroids.

α-Melanocyte-stimulating hormone (α-MSH)

A peptide consisting of the first 13 residues of ACTH, produced mainly in the intermediate lobe of the pituitary. This peptide plays a role in the control of skin color in lower animals and in the control of feeding in mammals.

Immuno-precipitation

A purification process in which a specific antibody is used to capture a protein or peptide from a complex mixture (e.g., cell extract).

Prohormone convertases

A family of serine proteases that cleave proteins in the Golgi apparatus/secretory

	granule before they are secreted into the circulation.
<i>Pro-opiomelanocortin (POMC)</i>	A 30 kDa protein synthesized predominantly in the pituitary that is the common precursor to ACTH and several other biologically active peptides.
<i>SDS-PAGE</i>	(Sodium dodecyl sulfate polyacrylamide electrophoresis) A method to separate proteins by size using electrophoresis through a gel matrix.

Introduction

Adrenocorticotropin (ACTH), like the vast majority of peptide hormones, is synthesized as part of a large, biologically inactive precursor protein that is then proteolytically processed to release the active peptides. ACTH is derived from pro-opiomelanocortin (POMC), a 30-kDa protein that is the common precursor to a number of biologically distinct peptides, an N-terminal region (N-POMC 1–76), ACTH, and β -lipotropin (β -LPH). These peptides can be further cleaved to give β -endorphin, corticotropin-like intermediate peptide (CLIP), and N-POMC 1–49. POMC also contains three copies of the melanocyte-stimulating hormone (α -, β -, and γ -MSH) sequence. The name of the protein thus encompasses the names of the peptides it encodes.

Discovery of POMC

The concept of hormone precursors was developed in the 1960s when it was shown that insulin was synthesized as a larger molecule, pro-insulin, and was released by limited proteolysis. During the 1970s evidence accumulated that implied a very close link between ACTH and β -LPH. The two peptides were released in approximately equimolar amounts from the pituitary, while immunocytochemistry showed the two peptides colocalized in the same cells. This evidence suggested that the peptides were derived from the same precursor, and in 1977 this hypothesis was proved to be correct when two groups showed, in two entirely different ways, that indeed ACTH and β -LPH were derived from the same molecule. Mains and Eipper used a pulse chase approach with the mouse pituitary tumor cell line AtT-20, which secretes large amounts of ACTH. They incubated the cells with radiolabeled amino acids (the pulse) and then washed the radioactivity away after a short period (the chase), thus labeling all proteins synthesized during the period of the pulse. By taking the cells at different times after the chase and immunoprecipitating the cellular protein using either an ACTH or β -endorphin antisera followed by SDS-PAGE analysis,

they identified a common 31-kDa protein that became progressively smaller with time after the pulse until it became fully processed.

The second group, headed by Ed Herbert, also used the AtT-20 cell line together with a cell free translation system. Messenger RNA was purified from the cells and expressed in the cell free system. In this way a 28-kDa protein was identified that was recognized by both ACTH and β -endorphin antisera. The apparent difference in size was accounted for by the absence of glycosylation.

Although these experiments gave conclusive evidence that indeed ACTH and β -LPH were derived from the same protein, it was not until 1979, when the cDNA encoding POMC was cloned from both the bovine intermediate lobe and the mouse pituitary cell line AtT-20, that the true structural relationship between ACTH and β -LPH was revealed. Since then the POMC gene has been cloned and sequenced from a variety of organisms. In every species examined, it has been found that there is a single functional copy of the gene, and the overall gene structure, with the exception of the loss of γ -MSH in the salmon, has remained conserved throughout evolution. In the human it consists of three exons and two introns. Exons 2 and 3 are translated, with exon 2 coding for 44 amino acids; the remaining 223 residues are coded for by exon 3. The precursor molecule (pre-POMC) has a 26-amino-acid signal sequence followed by a 76-residue peptide (N-POMC or pro- γ -MSH) that contains one copy of the MSH sequence. This is followed by a joining peptide, ACTH, whose sequence encompasses the second MSH sequence (α -MSH), and finally β -LPH at the C-terminus, the sequence of which includes that for β -endorphin and the third MSH sequence. All the bioactive peptides within POMC are flanked by dibasic amino acid residues, e.g., Arg-Lys, Arg-Arg, Lys-Arg, and Lys-Lys. The structure of POMC and the major peptides released as a result of proteolytic cleavage are shown in [Figure 1](#).

Sites and Tissue-Specific Processing of POMC

The main site of POMC expression is the anterior and intermediate lobes of the pituitary. In the anterior lobe, POMC represents approximately 1% of total protein, while it represents up to 10% in the intermediate lobe. Although these two tissues make up the principle sources of circulating POMC, it is also expressed and processed in several other tissues throughout the body, including the adrenal, gut, reproductive tract, placenta, leukocytes, spleen, lung, liver, thyroid, heart, skin, and brain.

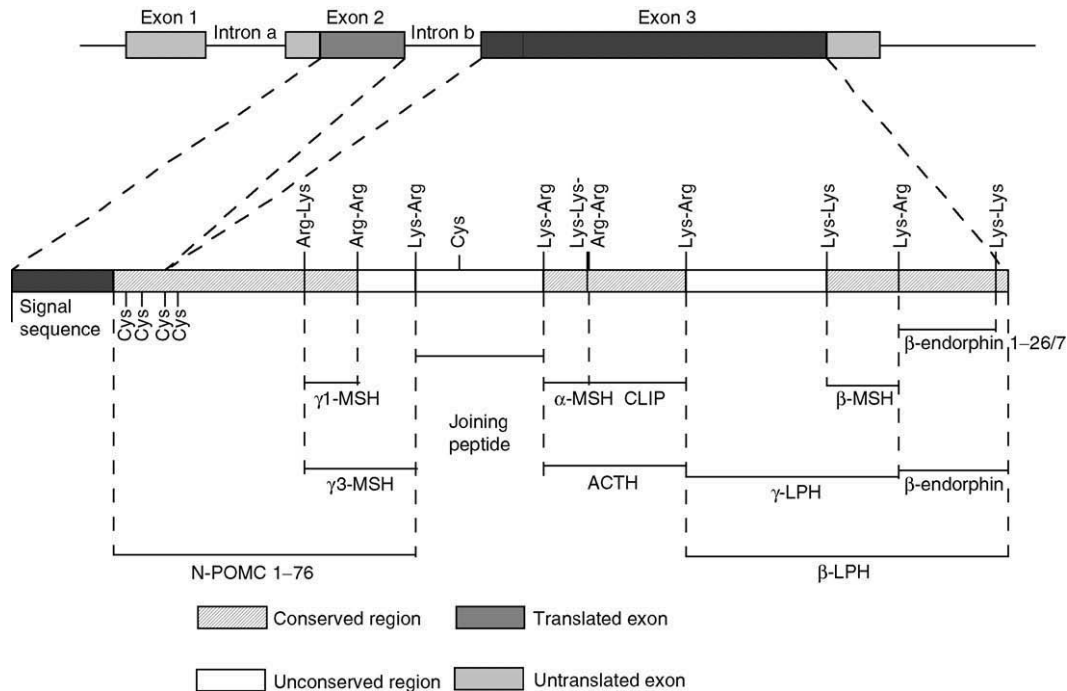


Figure 1 Structure of the POMC gene and the protein it encodes. Positions of the dibasic residues and the Cys residues are shown. Regions of the precursor conserved between species are also shown.

One advantage of producing peptides from a larger precursor is that the precursor can be cleaved in a tissue-specific manner to give a different set of products depending on tissue that the precursor is synthesized in. POMC is the classical example of this so-called tissue-specific processing, with the major sites of synthesis being the anterior and intermediate lobes of the pituitary. Classically, the pituitary can be divided into three distinct tissues:

1. The anterior pituitary or pars anterior contains a mixed population of cells, some of which, named corticotrophs, secrete POMC peptides that are involved mainly in the control of adrenal function.
2. The intermediate lobe or pars intermedia consists solely of melanotrophs that secrete POMC-derived peptides that are involved mainly with skin and hair pigmentation.
3. The posterior pituitary or pars nervosa stores and releases vasopressin and oxytocin produced in the hypothalamus.

In the rat these three tissues can be clearly defined, but in adult humans there is no distinct pars intermedia. It is present in fetal life but disappears soon after birth, leaving just a few scattered cells throughout the anterior lobe, giving the adult anterior lobe the collective name of pars distalis. The region between the anterior and posterior lobes is sometimes called the zona intermedia.

Processing of POMC in the Anterior Pituitary

Processing of the POMC in the anterior pituitary is summarized in [Figure 2](#). The mouse anterior pituitary cell line AtT-20 has been used extensively as an *in vitro* model for processing in the anterior lobe. Using pulse chase experiments it was determined that the major products were a 16-kDa N-terminal fragment, ACTH, and β-LPH. β-endorphin can also be formed from sequential cleavage from β-LPH (not from direct cleavage of POMC) but is only a minor product.

Processing of POMC in the Pars Intermedia

Within the pars intermedia, processing is more extensive, with the sequential cleavage of peptides produced in the anterior lobe. ACTH is cleaved to α-MSH and CLIP, while β-LPH is virtually completely processed to β-endorphin, and γ-LPH. β-endorphin (1-31) is cleaved at the C-terminus to give β-endorphin(1-27), β-endorphin(1-26), and a dipeptide glycylglutamine. The majority of the β-endorphin is N-acetylated. The N-terminal fragment N-POMC 1-76 is partially cleaved at an Arg-Lys site to give N-POMC 1-49 and Lys-γ3-MSH. This latter peptide contains an additional dibasic site (although not in all species) but does not appear to be cleaved to any significant extent.

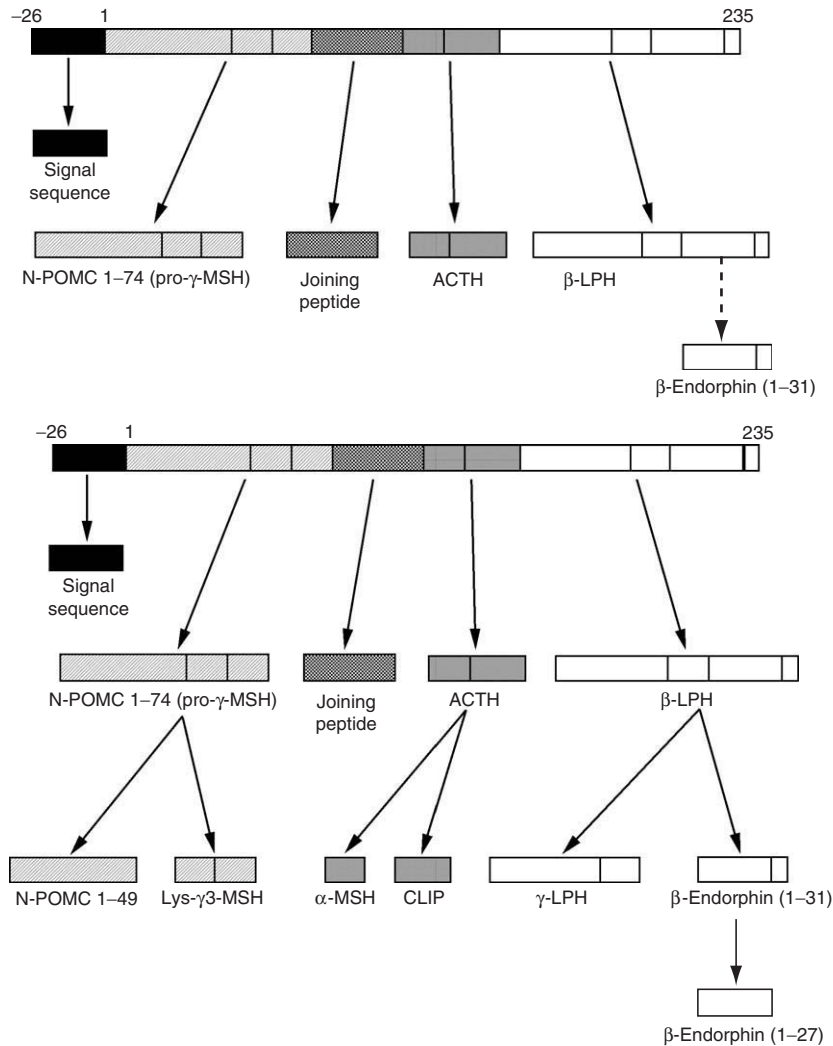


Figure 2 Processing of POMC in anterior and intermediate lobes of the pituitary. Dibasic cleavage sites are shown as vertical lines. Dashed arrow indicates incomplete processing.

Molecular Basis of Tissue-Specific Processing

POMC is itself biologically inactive and must be processed correctly to its biologically active peptides. During and after sorting and packaging of the POMC precursor into the regulated secretory vesicles, and prior to the release of the nascent secretory vesicles from the trans Golgi network, it is endoproteolytically cleaved. Cleavage is normally elicited at dibasic amino acid residues and follows a common route involving three different enzymes:

1. An endoprotease that cleaves on the C-terminal side of the basic amino acid pair.
2. A carboxypeptidase-B-like activity that removes the basic amino acid left after cleavage. Carboxypeptidase E (CPE) (also called CPH) specifically removes the basic residues from the C-terminus of a peptide.

3. A C-terminal amidating enzyme that catalyses the amidation of peptides terminating in glycine. The enzyme peptidylglycine alpha-amidating monooxygenase (PAM) catalyses this reaction and utilizes a glycine situated on the N-terminal side of the dibasic cleavage site as the nitrogen donor for the amide group.

These processing events are carried out in an acidic environment within the trans Golgi network or within the immature secretory vesicles and are necessary for the optimum activity of the processing enzymes as well as to induce conformational changes in the protein to expose the basic amino acid residues for cleavage.

In most cases, specific processing of prohormones occurs on the carboxyl side of pairs of basic amino acids, usually Lys-Arg or Arg-Arg sites. Processing

can also occur at the tetrabasic sequence Arg-X-Lys/Arg-Arg and less frequently at monobasic sites. The enzymes that have implicated in the specific processing of POMC are known as prohormone convertases 1 and 2 (PC1 and PC2). These two enzymes are members of the subtilisin family of Ca^{2+} -dependent serine proteinases.

Identification of the Enzymes Involved in POMC Processing

During the early 1990s the identification of the prohormone converting enzymes was a major aim of researchers in the field. The breakthrough came with the cloning of a yeast gene called KEX2. The product of the KEX2 gene, known as kexin, is a protease that cleaves the yeast alpha factor mating pheromone. Cellular expression of the gene demonstrated that the enzyme belongs to the subtilisin family of Ca^{2+} -dependent serine proteinases. KEX2 has the ability to selectively cleave a number of mammalian prohormones, and this led to the hypothesis that similar mammalian counterparts existed.

The first mammalian gene to be identified as a potential processing enzyme came as a consequence of computer alignment of the amino acid sequences surrounding the serine and catalytically important asparagine of kexin against known protein sequences. By this method, a human gene of unknown function was identified, named *fur*, which lies upstream from the

tyrosine kinase *fps/fes* proto-oncogene. The product of this gene, to become known as furin, was to become the first mammalian processing enzyme identified.

Using the sequence of the two eukaryotic proteases and several other subtilisin sequences, Smeekens and Steiner designed degenerate oligonucleotides corresponding to the conserved active site. By utilizing these in the polymerase chain reaction (PCR) with cDNA isolated from a human insulinoma library as template, they generated a fragment that was then used to probe that library. A full-length cDNA was identified that encoded for a protease that became known as PC2. Concurrently, Seidah and colleagues, using a similar procedure with the sequence of furin and a mouse pituitary library, identified two cDNA clones that were to become known as PC2 and PC1.

Elegant studies performed *in vitro* using cell-based assays and recombinant enzymes, together with studies looking at PC1 and PC2 knockout animals, have revealed the molecular basis of the tissue-specific processing observed in the anterior and intermediate lobes. In the anterior lobe, PC1 is the predominant enzyme present, and due to its more limited proteolytic activity, results in the generation of ACTH. However, in the intermediate lobe, both PC1 and PC2 are present, and their coordinate actions result in the generation of the smaller POMC peptides.

Cleavage of POMC occurs in a specific order and is summarized in [Figure 3](#).

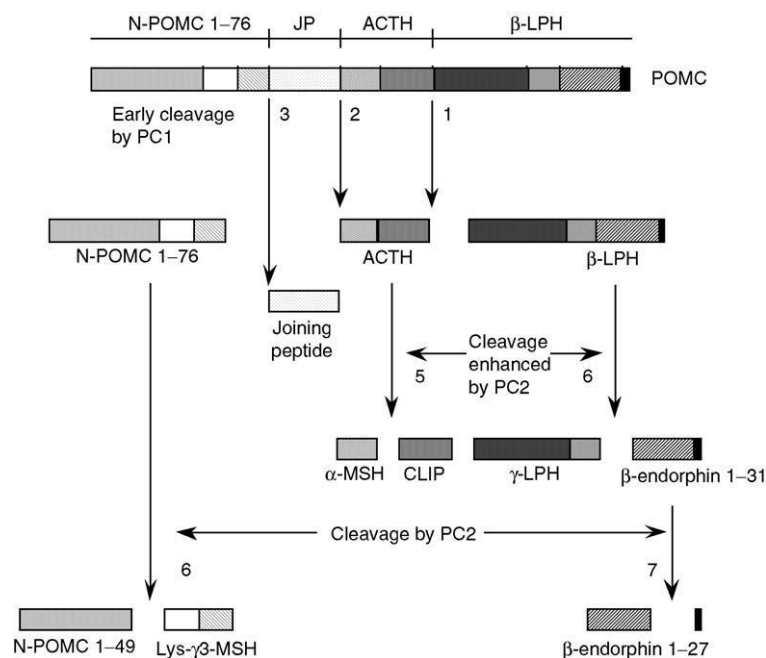


Figure 3 Processing of POMC by PC1 and PC2. The order of cleavage is dictated by the numbers showing that PC2 acts on the products of PC1 cleavage.

Regulation of the Expression of the POMC Peptides

As the processing of POMC occurs in a tissue-specific manner, it follows that the control of POMC expression and secretion of the derived products are also tissue specific.

Anterior Pituitary

With POMC produced in the pituitary, it followed logically that its secretion was under the control of factors released from the hypothalamus. In 1981 Vale identified the first of these as a 41-amino-acid peptide called corticotropin-releasing hormone (CRH). However, CRH does not stimulate ACTH release to the same extent as hypothalamic extracts, and it became clear that at least one other substance stimulated the release of ACTH. This second substance was identified by Lowry, McCann, Antoni and other investigators as arginine vasopressin, which potentiates the ACTH-releasing activity of CRH. Stimulation by CRH results in an increase in POMC mRNA levels as well as an increase in secretion of peptides. The two main factors are also potentiated by the action of several other weak secretagogues, namely, epinephrine, norepinephrine, oxytocin, and angiotensin II.

Observations after treatment with exogenous glucocorticoids or adrenalectomy have long implicated a role for glucocorticoids in the control of POMC peptide secretion. They exert their effects both at the hypothalamus, by decreasing CRH secretion, and at the corticotroph, by decreasing POMC synthesis.

Pars Intermedia

As would be expected, the regulation of POMC expression and secretion in the pars intermedia is quite distinct from that in the pars anterior. Glucocorticoids have no influence on POMC expression; although receptors are present, they do not seem to elicit any control over POMC gene expression.

POMC gene expression is enhanced by stress CRH and β -adrenergic agonists. All exert their actions via cAMP.

Dopamine decreases POMC expression by binding to D2 receptors, which are thought to block the activation of adenylate cyclase, although there may be more direct effects.

Posttranslational Modifications of POMC

POMC and its peptides undergo a number of different modifications following POMC's synthesis and removal of the signal peptide.

Disulfide Bridge Formation

Human POMC contains five Cys residues, one more than every other species. Four occur close to the

N-terminal at positions 2, 8, 20, and 24, while the fifth occurs in the joining peptide. The four N-terminal Cys residues form intramolecular disulfide bridges between 4 and 24 and 8 and 20. The fifth Cys, contained within the joining peptide and unique to humans, leads to its secretion as a homodimer, which may indicate that human POMC is produced as a dimer.

Glycosylation

In all species Asn 65 is N-linked glycosylated in the consensus sequence Asn-X-Ser. O-linked glycosylation also occurs at Thr 45 in the human and porcine pituitary.

Phosphorylation

Serine 31 of ACTH is phosphorylated in the rat, human, and mouse, although its presence has no effect on biological activity.

Acetylation

α -MSH and β -endorphin are found as des-, mono- and diacetylated forms, although the precise form varies between tissue and species. The N-terminal Ser of α -MSH may be acetylated through either the hydroxyl or amide group, acetylation increasing both its biological activity and half life, in contrast to β -endorphin, which reduces its biological opiate activity.

Biological Activity of the POMC Peptides

The peptides contained within the POMC precursor have a wide spectrum of biological activity that is briefly described in the following sections.

Adrenocorticotropin (ACTH)

The main role of ACTH is to promote steroidogenesis within the adrenal cortex. It exerts its actions by binding to the melanocortin 2 receptor (MC2R), leading to a rise in the levels of cAMP and activation of protein kinase A, resulting in secretion of glucocorticoids, androgenic steroids, and, to a lesser extent, mineralocorticoids. The acute effect of ACTH is to increase the rate of conversion of cholesterol to pregnenolone, the first and rate-limiting step in the steroidogenic pathway. The chronic effects of ACTH result in an upregulation of the expression of most of the enzymes involved in the steroidogenic pathway.

ACTH has also been implicated in the maintenance of adrenal weight. The evidence for this is based heavily on the observation that long-term glucocorticoid therapy results in adrenal atrophy. However, since all POMC peptides are downregulated during such treatment and administration of ACTH alone

does not totally reverse the atrophy, it is clear that other POMC peptides play a role in the control of adrenal weight.

N-Terminal Peptides and Adrenal Growth

The role of the N-terminus of POMC in adrenal growth was originally proposed in the early 1980s, when a series of elegant studies demonstrated that N-terminal peptides could stimulate mitogenesis in adrenal cell cultures. It was also shown that immunoneutralization of N-POMC peptides resulted in the inhibition of the compensatory growth response observed following unilateral adrenalectomy.

However, N-POMC 1–76 is itself inactive, but peptides derived from its N-terminus, not containing γ -MSH, are potent adrenal mitogens. Since these shorter peptides are not generated during normal pituitary processing, it was proposed that N-POMC 1–76 is cleaved after secretion. Considerable weight has recently been added to this argument with the identification of an adrenal protease that appears to cleave N-POMC to the shorter forms.

Melanocyte-Stimulating Hormone

There are three copies of the MSH sequence in the POMC precursor: α -MSH in ACTH, β -MSH in β -LPH, and γ -MSH in N-POMC 1–76. The biological role of the MSH peptides in lower vertebrates such as frogs was known before the discovery of POMC. α -MSH produced by the pars intermedia causes the darkening of skin to match a darker background. Both α -MSH and β -MSH share the common core sequence Met-Glu-His-Phe-Arg-Trp-Gly, which is essential for biological activity. γ -MSH contains the sequence Met-Gly-His-Phe-Arg-Trp-Asp and was named due to its similarity to the other MSH sequences. However, it has no melanotropic activity.

The role of the MSH peptides in skin pigmentation in humans is not clear. Injection of the peptides into humans does result in skin darkening. Also, in Nelson's syndrome, a condition in which there are high levels of POMC peptides, skin darkening is observed. However, in the normal human this role is debatable. Humans do not have any pars intermedia tissue, and certainly there are no detectable amounts of α -MSH in the periphery, although mRNA encoding POMC can be detected in human skin.

Role of MSH Peptides in the Regulation of Feeding

Mammals regulate their energy stores relatively constantly despite a large variation in both food availability and physical activity. The control of energy

expenditure has been shown to be the result of a complicated endocrine loop involving a number of neuropeptides, including the melanocortins.

POMC is synthesized in the hypothalamic arcuate nucleus, where it is processed in a manner similar to that observed in the intermediate lobe. It has been generally assumed that α -MSH is the most important peptide since it is the ligand with highest affinity for the two melanocortin receptors – the MC3R and MC4R expressed in the hypothalamus. Genetic and pharmacological disruption of this loop causes obesity in both humans and rodents. The POMC knockout mouse is hyperphagic and obese, whereas infusion of α -MSH (or synthetic agonists) results in anorexia and weight loss.

The complete subtle architecture of this system is still far from clear, but the system's potential as a therapeutic target for the treatment of obesity has made it the focus of intense research over the past few years.

γ -MSH and Steroidogenesis

γ -MSH is generated from N-POMC in the intermediate lobe of the pituitary. It contains a slightly modified core sequence, Met-Gly-His-Phe-Arg-Trp-Asp, and has a biological role distinct from that of α -MSH.

Studies performed by Alastair Brownie in the early 1980s identified γ -MSH as a regulator of hormone sensitive lipase (HSL). This enzyme is found in steroidogenic tissues and breaks down cholesterol esters to free cholesterol, which then feeds into the steroidogenic pathway. Several groups demonstrated that although not directly steroidogenic, this peptide can synergize with ACTH and result in an increase in steroid output. The receptor through which this peptide acts is unclear, since although the MC3R has been shown to be the only melanocortin receptor with significant affinity for γ -MSH, its expression profile and pharmacology suggest that it does not play a role in this system.

Joining Peptide

The joining peptide has no known biological function. Its role, like the C-peptide of insulin, is probably only to provide a spacer between the different peptides. One interesting feature of human joining peptide is that it uniquely contains a cysteine that has been implicated in dimer formation.

β -Endorphin

β -endorphin is one of several endogenous opioids that exert their actions by binding to receptors within the central and peripheral nervous systems. It is part of a larger family whose members include dynorphin and

the enkephalins. The opioids bind to three receptors known as the μ , κ , and δ opioid receptors. The receptors derive their names from the opiate agonists that bind to them: morphine and ketocyclazocine for the μ and κ receptors, and the vas deferens for the δ receptor, where it is located. The receptors differ in their affinity for the different opioids, although not one is specific for a single ligand. β -endorphin and Met-enkephalin bind to the μ and δ receptors, while dynorphin binds to the κ receptor. All the receptors are G-coupled and recognize a common motif consisting of a protonated amine adjacent to an aromatic amine, for example, the tyrosine residue at the N-terminus of the enkephalins. The opioid actions of β -endorphin are dependent on its C-terminal residues. Removal leads to a decrease in both its analgesic properties and binding affinity for the receptors. Thus, the shorter endorphins lack opioid activity, as does N-acetylation of the peptide. The N-terminal portion of β -endorphin is required for binding to the opiate receptors, whereas the C-terminal residues are thought to enhance the binding affinity.

Conclusions

Most peptide hormones are produced as inactive precursor molecules that are subsequently cleaved at dibasic residues to release the active peptides. As the first such multihormone to be discovered, over a

quarter of a century ago, POMC has been extensively studied and characterized. Its role in the control of adrenal function by the actions of ACTH on steroidogenesis, as well as N-POMC on the control of adrenal tonic state, are central to the stress axis as a whole.

See Also the Following Articles

Adrenocorticotrophic Hormone (ACTH); Obesity, Stress and; Peptides.

Further Reading

- Bertagna, X. (1994). Proopiomelanocortin-derived peptides. *Endocrinology and Metabolism Clinics of North America* **23**, 467–485.
- MacNeil, D. J., Howard, A. D., Guan, X., et al. (2002). The role of melanocortins in body weight regulation: opportunities for the treatment of obesity. *European Journal of Pharmacology* **450**, 93–109.
- Raffin-Sanson, M. L., de Keyser, Y. and Bertagna, X. (2003). Pro-opiomelanocortin, a polypeptide precursor with multiple functions: from physiology to pathological conditions. *European Journal of Endocrinology* **149**, 79–90.
- Smith, A. I. and Funder, J. W. (1988). Proopiomelanocortin processing in the pituitary, central nervous system and peripheral tissues. *Endocrine Reviews* **9**, 159–179.
- Steiner, D. F. (1998). The proprotein convertases. *Current Opinion in Chemical Biology* **2**, 31–39.

Prostaglandins

S Moshonov and U Zor

Weizmann Institute of Science, Rehovot, Israel

Z Naor

Tel Aviv University, Tel Aviv, Israel

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S Moshonov, U Zor and Z Naor, volume 3, pp 266–272, © 2000, Elsevier Inc.

Introduction

Characteristics of Prostaglandins

Prostaglandins in the Hypothalamic-Pituitary-Adrenal Axis

Stress and Antistress: Prostaglandins versus Glucocorticosteroids

Glossary

Adrenals

The organs that cap the kidneys. The outer layer, the cortex, synthesizes the corticosteroids, whereas the inner layer, the medulla, synthesizes epinephrine.

Cytokines

Regulatory proteins, such as the interleukins and lymphokines, that are released by cells of the immune system and act as intercellular mediators in the generation of an immune response.

Glucocorticosteroids

Anti-inflammatory steroids; also called glucocorticoids. Cortisol and corticosterone are natural corticosteroids secreted by the adrenal cortex and have strong glucocorticoid activity.

Hypothalamus

A specialized area in the brain that regulates automatic body functions such as temperature, sleep, and water balance. It coordinates the endocrine system.

<i>Noradrenergic cells</i>	Nerve cells that release norepinephrine.
<i>Nonsteroidal anti-inflammatory drugs (NSAIDs)</i>	Drugs (e.g. aspirin, indomethacin) that act by inhibition of the cyclo-oxygenases (COX-1 and COX-2) and hence prevent prostaglandin synthesis.
<i>Phospholipids</i>	Lipid (fat) molecules with a nonlipid phosphate head. Phospholipid bilayers are the basic components of cell membranes.
<i>Phosphorylation</i>	The process of adding phosphorus to a dormant enzyme, thereby activating it. Occasionally, phosphorylation deactivates an active enzyme.
<i>Pituitary</i>	The master endocrine gland that lies just beneath and is attached by the pituitary stalk to hypothalamus. The pituitary controls major endocrine systems, growth, body water and electrolytes, lactation and parturition. The pituitary gland is under central nervous control, which in the case of the anterior pituitary gland is mediated by hypothalamic neurohormones transported from the hypothalamus by the hypophyseal portal vessel system.
<i>Prostaglandins (PGs)</i>	A family of hormones whose chemical structure is identical, except for small changes in the pentane ring; these differences provide the distinguishing nomenclature (prostaglandin D, E, F, and I and thromboxane A). The subscript in the name (e.g., PGD ₂) refers to the number of unsaturated carbon bonds.
<i>Stressors</i>	Any agent that induces stress in the intact animal or in cell or tissues in culture.

Introduction

Stress, generally classified in terms of emotional or physical, presents a very clear pattern of pathophysiological changes. These are the consequences of stimulation by the many biologically active agents – hormones, neurotransmitters, and cytokines – generated by the nervous, endocrine, and immune systems aroused by stress. To a large extent, the systems are interdependent, and their complex networking is still far from being completely understood. The body maintains its healthy state by continuously readjusting to the changing environment, a process known as homeostasis. Stressors push the homeostatic mechanisms beyond normal limits, and the body enters a new, temporary state called stress.

The participation of the main agents of stress, adrenocorticotrophic hormone (ACTH), corticosteroids, and catecholamines (CAs), has been long established. PGs, which are local rather than transported hormones, are not among these main stress hormones

Table 1 Stages in the response to a stress signal that may be affected by prostaglandins

<i>Stages of stress response</i>	<i>Prostaglandin link</i>
Recognition of stress signal	Unclear
Initiation of stress response	Positive
Transmission and modulation of stress response	Possible
Expression of stress symptoms	Positive

that are poured out immediately in response to stress. However, cells from stressed animals produce more PGs than controls do, and PG levels in serum, spinal fluid, and the hypothalamus are elevated during stress. The role of PGs in stress is still largely undefined, but they appear to be involved in at least two stages (Table 1): (1) the initiation of the stress response in the brain and (2) the expression of some symptoms of stress.

Characteristics of Prostaglandins

Background

As first reported in the 1930s, PGs were the most powerful, ubiquitous, diversified, and mysterious biological substances yet discovered. Some 30 years later, they were finally unmasked by the Swedish medicinal chemists Sune Bergström and Bengt Samuelsson, who in 1982 were awarded the Nobel Prize in Physiology with the British pharmacologist John Vane. They isolated, identified, and characterized the PG family, establishing unsaturated lipids derived from the cell membrane as an unexpected class of highly sensitive bioactive regulators. Thanks to Vane and colleagues, the mechanism of action of aspirin, the first of the manufactured NSAIDs and one of the most successful drugs ever to be marketed, was now, some 70 years later, finally elucidated, proving to be a specific inhibitor of PG synthesis.

Synthesis of Prostaglandins

Cell membranes, comprising largely phospholipids, are themselves the major source material for PGs and other biologically reactive eicosanoids (20-carbon unsaturated carboxylic acids) such as leukotrienes (LTs).

Phospholipase A₂ Arachidonic acid (AA), an unsaturated fatty acid, is detached from membrane phospholipids by the enzyme phospholipase A₂ (PLA₂), transformed by cyclo-oxygenases (COXs) into the unstable endoperoxide PGH₂, and then metabolized by the various synthases to PGE₂, PGF_{2α}, PGD₂, prostacyclin (PGI₂), or TXA₂ (Figure 1). The superfamily of the PLA₂ enzymes consists of 14 groups (I–XIV)

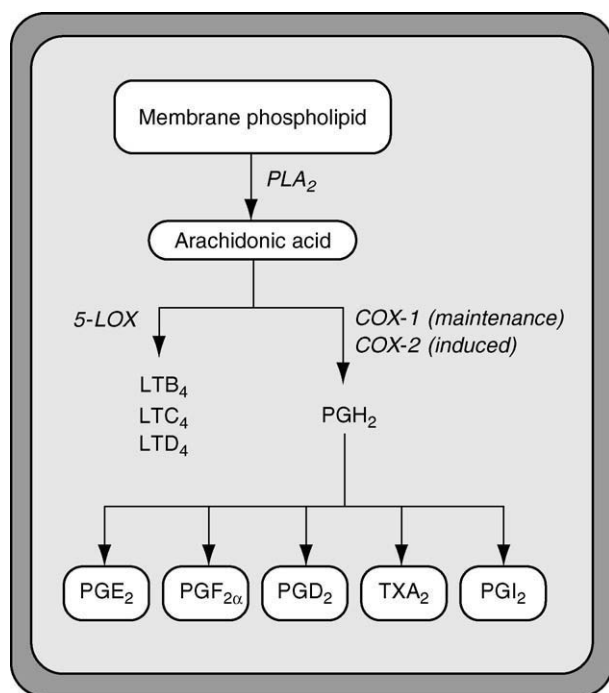


Figure 1 Prostaglandins and their synthesis. COX, cyclo-oxygenase; LOX, lipoxygenase; LT, leukotriene; PG, prostaglandin; PGI₂, prostacyclin; PLA₂, phospholipase A₂; TXA₂, thromboxane A₂.

and at least 19 members. The low-molecular-mass (~12- to 19-kDa) secretory PLA₂ (sPLA₂) proteins are all related evolutionarily, share a common catalytic histidine, and fall into the groups I–III, V, and IX–XIV. A high-molecular-mass (60- to 110-kDa) family of calcium-dependent (requiring micromoles of Ca²⁺ for activation) cytosolic PLA₂ (cPLA₂) proteins use a catalytic serine and have a high specificity to AA. Some are activated by Ser505 phosphorylation, use a novel catalytic dyad of Ser228 and Asp549, and are designated group IV (A, B, C or α , β , γ cPLA₂). The calcium-independent PLA₂ (iPLA₂) (group VI) contains eight ankyrin repeats, has a lipase consensus sequence with an active serine that serves as a nucleophile, plays a role in phospholipids' remodeling and homeostasis via the production of lysophospholipids, and falls into the subgroups (A-1, A-2, and B). Groups VII and VIII PLA₂ are platelet-activating factor (PAF)-acetylhydrolases that cleave PAF to form lyso-PAF and a free acetyl group. mRNA expression of cPLA₂ is enhanced by pro-inflammatory cytokines and inhibited by glucocorticoids (GCs).

An alternative and equally important route for AA is metabolism by lipoxygenases (LOXs) to the exceedingly potent LT derivatives (Figure 1). Like PGE₂ and PGF_{2 α} , LTs mediate inflammatory and allergic reactions, but their contribution to stress is largely uncharted. Other branches of the PG–LT

family are also not discussed in this article; they appear to be less bioreactive and less is known in general about their contribution to pathophysiological mechanisms.

Cyclo-oxygenases The COX enzyme was first known as PGH₂ synthase (PGHS) because of its ability to catalyze the first two reactions in PG synthesis from arachidonate to PGH₂. COX was first isolated in 1976, and its two isoforms, COX-1 and COX-2, were discovered around 1992. The two isoforms (71 kDa), are encoded by different genes, are structurally similar, are composed of some 600 amino acids, and catalyze the same substrate. The enzymes contain a heme group that binds molecular oxygen. COX-1 is mainly found in the endoplasmic membrane, whereas COX-2 is found also in the nuclear envelope. The active sites of the two enzymes differ, which may explain why the enzymes are inhibited by different drugs. COX-1 is constitutive (i.e., always present in almost all cell types) and is associated with normal physiological PG activities, such as protection of the kidney and the stomach lining from damage and maintenance of vascular homeostasis. However, COX-2 is induced by stimuli, both stressful ones such as inflammatory agents and natural ones such as growth factors; it is also found constitutively in a few organs, one of which is the brain. It is not, however, expressed in nerve cells that synthesize the PGs responsible for stress-induced fever, hyperalgesia, and sleepiness. The origin of this stress-induced COX-2 is the endothelial cells lining the blood vessels in the brain, which are accessed easily by the PG-stimulating blood-borne cytokines. Anti-inflammatory GCs, complexed with their specific receptors, penetrate the nucleus, bind to DNA at the hormone response element sequence, regulate transcription, and preferentially suppress the induction of COX-2.

Nature and Functions of Prostaglandins

Prostaglandins are very potent local hormones that are synthesized rapidly by a wide variety of cells in the body. PGs are released into the extracellular space by specific PG transporters. Once released, PGs bind to their cognate receptors, which belong to the G-protein-coupled receptor (GPCR) superfamily. There are four major receptor subtypes for PGE₂ (EP1, EP2, EP3, and EP4), which are encoded by separate genes. There is a major receptor for PGF_{2 α} (FP), for PGD₂ (DP), and for PGI₂ (IP). The various PG receptors differ in their signaling pathways. Sequence homology is found between receptors that share a signaling cascade rather than a ligand. Therefore, the group IP, DP, EP2, and EP4, which act via Gs to elevate cAMP levels, are known as the relaxant receptors because of

their common effect on smooth muscle relaxation. The contractile receptors TP, FP, and EP1 mobilize calcium via Gq and induce smooth muscle contraction. PGs, therefore, act in an autocrine/paracrine manner via cognate GPCRs. Within a short time, usually less than 30 min and in the case of thromboxane A₂ (TXA₂) a mere 30 s, PGs are inactivated (90%) to oxidized metabolites when passing through the lungs; very little PG is normally found in the circulation.

Pathological functions of prostaglandins and connections with stress PGs almost always act in the vicinity of their synthesis. The blood-borne, pro-inflammatory cytokines interleukin (IL)-1, IL-6, and tumor necrosis factor (TNF)- α are released from leukocytes in the blood and macrophages in tissues and are among the agents that locally trigger PG synthesis.

- **Inflammation.** Traumatic, inflammatory, immune, or other foreign stimuli trigger the activation and recruitment of PGs and LTs. They are part of the inflammatory reaction, promoting leukocyte infiltration, edema, and local vasodilation. Stress generates the same arachidonic AA metabolites and cytokines as inflammation.

- **Fever.** The elevation of body temperature, but not the maintenance of core temperature, is controlled by PGE₂ in the preoptic area of the hypothalamus. Stress-induced fever is reduced by the prior administration of aspirin-like drugs, such as indomethacin. The stimulation of PGE₂ synthesis in the endothelial cells of hypothalamic blood vessels by stressors such as lipopolysaccharide (LPS), which mimic the symptoms and the stress of bacterial infection in experimental animals, elevates body temperature. Fever also accompanies the stress of surgery (trauma), pain, and cage switching in rats, restraint in pigs, and open field (but not the stress of restraint) in rats or pigs being transported or frustrated by a lack of reward after exhibiting reward-training behavior.

- **Hyperalgesia.** Heightened sensitivity to pain, or hyperalgesia, is another stress symptom related to PGs. Pain is processed along the spinothalamic tract, reaching the thalamus, limbic system, and frontal cortex before descending through moderating pain-inhibitory nerve pathways. Stress-induced COX-2 is most evident in the spinal cord. Animal models of inflammatory stress and pain, such as carrageenin injected into rat foot pads, show that PGs are generated locally at the sensory nerve endings and in the spinal cord and cause hyperalgesia. COX-2 inhibitors (such as NS398 and DuP 697) and mixed COX-1 and

COX-2 inhibitors (such as aspirin and indomethacin), applied directly to the spinal fluid, are analgesic.

- **Dysmenorrhea.** Dysmenorrhea is a condition with painful menstrual periods with uterine hypercontractility. Increased levels of PGE₂ and PGF_{2 α} were found in the menstrual fluid of women with dysmenorrhea, and COX enzyme inhibitors alleviate the pain.

- **Endometriosis.** Endometriosis occurs when endometrial tissue is found in extrauterine sites; it is associated with pelvic pain and infertility. The increased production of PGE₂ and PGF_{2 α} and overexpression of COX-2 are thought to be associated with endometriosis. Estrogen and angiogenic factors, such as vascular endothelial growth factor (VEGF), may be involved in the proliferative responses induced by the PGs in endometriosis.

- **Cancer.** Various studies have indicated a role for the COX enzymes, PGs, and PG receptors in the pathology of cancers. The overexpression of the COX enzymes and PG receptors and the increased production of PG were found in various cancers. The studies suggested the use of COX enzyme inhibitors in treating some cancers. Because COX-1 is regarded as a constitutive enzyme, whereas COX-2 is the regulated enzyme, attention has been paid to the development of selective COX-2 inhibitors.

Some physiological functions of prostaglandins with connections to stress The various PGs have diverse and frequently opposing actions (Table 2). PGs used in nonpathological functions are synthesized mainly by the constitutive (maintenance) enzyme COX-1.

- **Blood pressure.** PGE₂ and prostacyclin (PGI₂) are vasodilators and have a role in normal blood pressure control. However, TXA₂ is a strong vasoconstrictor. Blood pressure rises in stress and is not affected by PG inhibitors.

- **Gastric ulcers.** In the gastrointestinal system, PGE₂ and PGI₂ protect the stomach lining by inhibiting gastric acid secretion and promoting the healing of ulcers. Prolonged use of aspirin or other NSAIDs to inhibit PG synthesis. Gastric ulcers are enhanced.

- **Preterm birth.** The presence of PGE₂ and PGF_{2 α} , both of which contract the pregnant uterus and are produced by uterine and fetal tissue, is essential for the survival of the fetus. During maternal or fetal stress, these PGs are induced and can lead to preterm birth.

- **Sleep.** A central function of PGD₂ is thought to be sleep induction. Its metabolic enzyme PGD₂ synthase is located in the cerebrospinal fluid and the meninges and choroid plexus, nonneuronal cells lining the whole central nervous system (CNS).

Table 2 Diversity of physiological and pathological effects of the prostaglandins^a

	<i>PGD₂</i>	<i>PGE₂</i>	<i>PGF_{2α}</i>	<i>PGI₂</i>	<i>TXA₂</i>
Blood pressure	Down	Down	No change	Down	Up
Airway	Constrict	Dilate	Constrict	Dilate	Constrict
Uterus (nonpregnant)		Relax	Contract		
Uterus (pregnant)		Contract	Contract		
Kidney (GFR)	Increase	No change		No change	Reduce
Stomach secretions		Inhibit	Stimulate	Inhibit	
Gut motility		Increase	Increase	Increase	Increase
Platelet aggregation			Inhibit		Stimulate
Sleep	Induce				
Fever		Induce			
Pain		Lowers threshold		Lowers threshold	

^aChart is not comprehensive.

Sleepiness is a stress response, and *PGD₂* could be involved in this process.

Stimulation of Prostaglandin Synthesis in Stress

Most stressors stimulate PG synthesis. Increased plasma (or medium) *PGE₂* was induced by heat, pain, trauma (bone fracture), exercise, LPS, restraint, water immersion, and cage switching. A fall in *PGE₂*, however, was associated with cold stress in rats and captivity in frogs. In these captive frogs, *PGF_{2α}* increased, as it also did in individuals subjected to the stress of surgery, students taking examinations, and cells that were heat-stressed. *PGI₂* was also increased by examination and surgery stresses, and it was removed from the aorta of immobilized rats.

In the brain of stressed animals, IL-1-stimulated *PGE₂* releases the corticotropin releasing hormone (CRH) in the hypothalamus, thus activating the hypothalamic-pituitary-adrenal (HPA) axis. Most studies report that *PGE₂* is synthesized by the cytokines in the brain; *PGF_{2α}* and *PGD₂* are also mentioned in such studies. PGs induced by IL-1, IL-6, and TNF- α in the preoptic area also bring about fever and cause hyperalgesia via the spinal cord (and at sensory nerve endings), whereas *PGD₂* in the cerebrospinal fluid and meninges induces sleep. The sources of these PGs are not necessarily neuronal but, rather, originate from the blood vessels, microglia, or other cells lining the brain.

Assay for Prostaglandins in Stress

Types of stress and measurement of prostaglandins

For pharmacological purposes, there is no gross difference between emotional and physical stress. Cell systems galvanized by the body to express and to cope with the stress of, for example, bereavement (emotional or psychological) are ultimately the same ones used in the stress caused by an illness (internal physical) or a supreme athletic effort (external

physical). An element of psychological stress generally accompanies a physical stress.

Experimental models of stress have been devised to mimic human stress-provoking situations, in which animals, or cells or tissues in culture, are subjected to one of a number of unpleasant procedures. In these models, stress behavior, in the presence and in the absence of drugs that modify PG synthesis, is observed and the biochemical changes are monitored. Alternatively, tissue samples are assayed for PGs, *PLA₂*, or COX-2 before and after provoking stress. PGs are commonly assayed using radioimmunoassay techniques, and enzymes are assayed by measuring their activity on labeled substrates or measuring their mRNA levels. Increases in enzyme activity (*PLA₂* and COX) and a rise in concentration of a specific PG have been consistently reported in many different stress states.

Inhibitors of prostaglandins Prostaglandin synthesis is inhibited by two groups of anti-inflammatory agents: NSAIDs and GCs (the very same hormones elicited by stress). NSAIDs interfere only with COX activity; the prime example is aspirin, although indomethacin is routinely used in experimental systems. These drugs inhibit COX-1 and COX-2 to varying degrees. COX-2-selective inhibitors (COXibs) such as Celecoxib (Celebrex) and Rofecoxib (Vioxx), developed in the 1990s, are relatively large inhibitors that penetrate the active site of COX-2 but not that of COX-1. They offered much promise when they first appeared on the market, but were found to be associated with a potential risk of myocardial infarction and sudden cardiac death. More recently, attention has been given to the pharmacology of the PG receptors. Efforts are being made to develop receptor-specific antagonists that will inhibit the biological action of the PG rather than their synthesis. GCs, such as dexamethasone, depress the whole AA metabolite network by partial inhibition of *PLA₂*

synthesis and control stress-induced PGs, in particular by preventing cells from inducing COX-2. The very strong anti-inflammatory action of GCs is due to their wide-ranging anti-inflammatory base. The GC receptor complex interacts directly with and regulates the gene transcription of many inflammatory enzymes and proteins, including COX-2. Inhibitors of PG have been widely employed in the stress models to determine the contribution of PGs to the stress symptoms. This use of PG inhibitors as indirect proof has its limitations because the NSAIDs that can distinguish between the various PGs have not yet been developed and GCs suppress a whole spectrum of agents provoked by stress.

Prostaglandins in the Hypothalamic-Pituitary-Adrenal Axis

Stress Hormone Flow

The perception of distress is instantly transmitted via one or more neural circuits, each with its own neurotransmitter system, to the median eminence (ME) in the hypothalamus. The production of CRH, the major stress hormone released by the hypothalamus, and arginine vasopressin (AVP) is intensified immediately and stimulates the anterior pituitary to release ACTH into the circulation. The adrenal cortex is stimulated by ACTH to produce cortisol and corticosterone, powerful GCs that, by suppressing the production of pro-stress agents, including PGs, keep the sharp stress response from running fatally out of control (Figure 2).

The sympathetic nervous system, also activated in stress, releases CAs into the circulation – norepinephrine (NE) and epinephrine (EPI) – setting the body in the stress-alert mode for flight or fight. CAs stimulate PG synthesis in many tissues, including the brain.

Prostaglandins in the Initiation of Stress Response

Prostaglandin activity has been located in the HPA axis, prior to CRH release. The stress-related PG presence has not been confirmed at other points along the axis. One of the CRH-stimulating neural circuits is noradrenergic. These neurons originate in and/or project to the paraventricular nucleus (PVN) of the hypothalamus and are stimulated by IL-1- and IL-6-induced PGE₂. The rat hypothalamus can synthesize PGE₂, PGF_{2α}, PGI₂, and TXA₂, but, at least *in vitro*, only PGE₂ is produced in response to IL-1 and IL-6, although PGD₂, PGF_{2α}, and TXA₂ also induce CRH release. The noradrenergic circuit that stimulates the HPA axis can stimulate PG synthesis and is inhibited by indomethacin. PGs also seem to be able to directly induce CRH or AVP release. Thus,

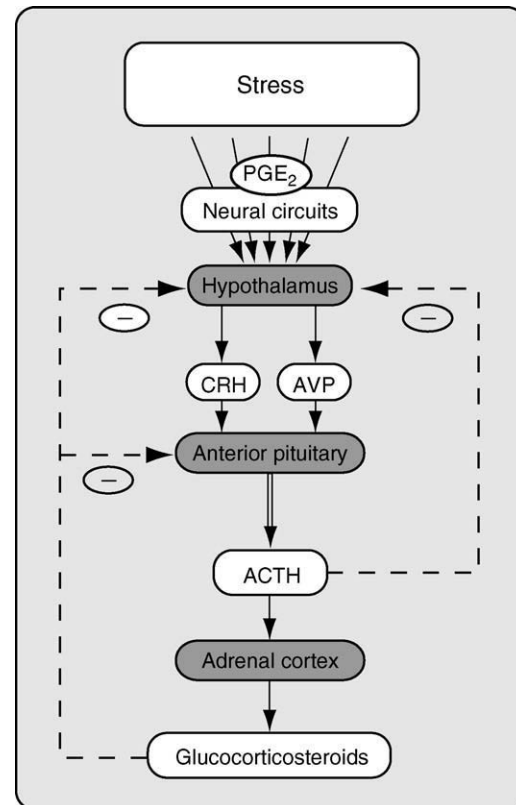


Figure 2 Activation of the hypothalamic-pituitary-adrenal (HPA) axis by stress. Release of adrenocorticotrophic hormone (ACTH) is under the negative control (–) of the corticosteroids it induces. Prostaglandin E₂ (PGE₂) stimulates the release of corticotropin releasing hormone (CRH) or arginine vasopressin (AVP) from the hypothalamus and mediates the noradrenergic activation of the HPA axis.

PGs may activate the HPA axis through the direct stimulation of CRH or AVP neurons or via noradrenergic neurons in the PVN (Figure 3).

Stress and Antistress: Prostaglandins versus Glucocorticosteroids

The stress-induced production of pathological PGs, together with anti-stress GCs, raises the question of their mutual operation during stress. The GC surge elicited by stress effects an inbuilt antistress facet to limit some aspects of the stress response itself. In the absence of GCs, stress responses – high blood pressure, increased heart rate, and fever – are unbridled and could well be fatal. Essentially, GCs return activated blood and other cells to their quiescent state so that they stop producing the cytokines, PGs, and other inflammatory agents that constitute the host defense mechanism.

Under normal physiological conditions, COX-2 is regulated, among others, by endogenous GCs.

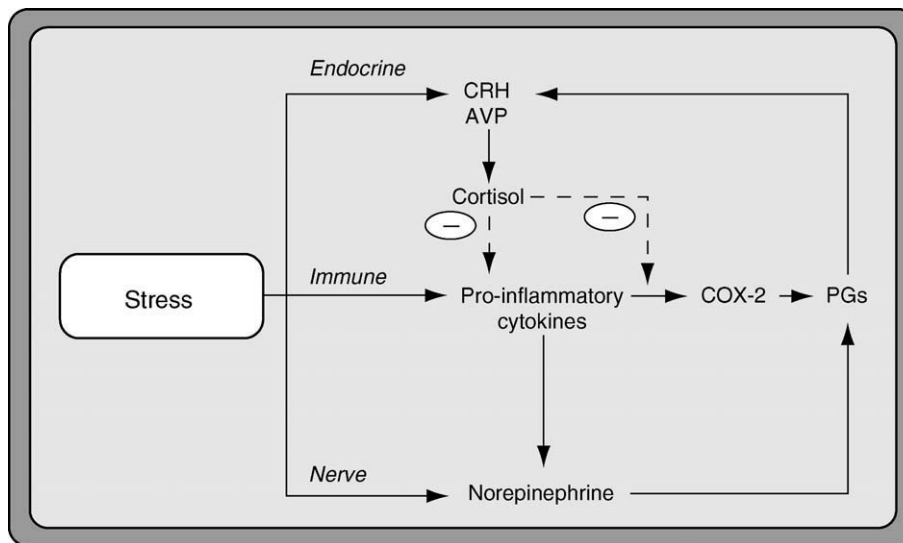


Figure 3 Participation and interdependence of endocrine, immune, and nervous systems in stress. AVP, arginine vasopressin; COX-2, cyclo-oxygenase-2; CRH, corticotropin releasing hormone; PGs, prostaglandins.

Whether the GC surge during stress – a pathophysiological state – affects concomitant PG synthesis has been examined using the rather crude technique of adrenalectomy (ADX) and, more recently, with GC-receptor antagonists such as RU486 (mifepristone). GCs do indeed moderate stress-induced PG production because PG synthesis is elevated in the absence of viable GCs and is reduced again by exogenous dexamethasone, but not by EPI, which is also lacking after ADX.

In the brain, suppression by GCs of PGE₂ along the HPA axis presents an interesting negative feedback response because at least one of the noradrenergic neural circuits leading to ACTH release is mediated via PGE₂. This PGE₂ is stimulated by circulating IL-1, and pro-inflammatory cytokines released into the plasma are also suppressed by GCs. Thus, GCs prevent their own stimulation. The inhibition of IL-1- and PGE₂-induced GCs by GCs could be part of a mechanism for the negative feedback between ACTH and GCs.

In general, and left to themselves, COX-2-stimulated PGs are injurious (they support inflammation allergy). The repression of COX-2 puts GCs into their familiar role of damage control by holding the body's stress response in check and trying to restore homeostasis.

See Also the Following Articles

Congenital Adrenal Hyperplasia (CAH); Caregivers, Stress and; Corticosteroid-Binding Globulin (Transcortin); Corticosteroids and Stress; Cushing's Syndrome, Neuropsychiatric Aspects; Glucocorticoids – Adverse Effects on the Nervous System; Holocaust, Stress Effects of; Hypotension, Hypovolemia, and Septic Shock.

Further Reading

- Bernadini, R., Chiarenza, A., Calogero, A. E., et al. (1989). Arachidonic acid metabolites modulate rat hypothalamic corticotropin-releasing hormone secretion. *Neuroendocrinology* 50, 708–715.
- Brody, T. M., Larner, J. and Minneman, K. P. (1998). *Human pharmacology: molecular to clinical* (3rd edn.). St. Louis, MO: Mosby-Year Book.
- Bugajski, J., Olowska, A., Gadek-Michalska, A., et al. (1996). Effect of indomethacin on the CRH- and VP-induced pituitary-adrenocortical response during social stress. *Life Sciences* 58, 67–72.
- Cowell, A. M. and Buckingham, J. C. (1989). Eicosanoids and the hypothalamic-pituitary axis. *Prostaglandins, Leukotrienes and Essential Fatty Acids* 36, 235–250.
- Elmquist, J. K., Scammell, T. E. and Saper, C. B. (1997). Mechanisms of CNS response to systemic immune challenge: the febrile response. *Trends in Neuroscience* 20, 565–570.
- Gether, U. (2000). Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. *Endocrine Review* 21, 90–113.
- Gilman, A. G., Rall, T. W., Nies, A. S. and Taylor, P. (1990). *The pharmacological basis of therapeutics* (8th edn.). New York: Pergamon Press.
- Goppelt-Strube, M. (1997). Molecular mechanisms involved in the regulation of prostaglandin biosynthesis by glucocorticoids. *Biochemical Pharmacology* 53, 1389–1395.
- Graham, D. J., Campen, D., Hui, R., et al. (2005). Risk of acute myocardial infarction and sudden cardiac death in patients treated with cyclooxygenase 2 selective and non-selective non-steroidal anti-inflammatory drugs: nested case-control study. *Lancet* 365, 475–481.
- Hata, A. N. and Breyer, R. M. (2004). Pharmacology and signaling of prostaglandin receptors: multiple roles in

- inflammation and immune modulation. *Pharmacology and Therapeutics* 103, 147–166.
- Jabbour, H. N. and Sales, K. J. (2004). Prostaglandin receptor signalling and function in human endometrial pathology. *Trends in Endocrinology and Metabolism* 15, 398–404.
- Morimoto, A., Watanabe, T., Morimoto, K., et al. (1991). Possible involvement of prostaglandins in psychological stress-induced responses in rats. *Journal of Physiology* (London) 443, 421–429.
- Narumiya, S., Sugimoto, Y. and Ushikubi, F. (1999). Prostanoid receptors: structures, properties, and functions. *Physiological Reviews* 79, 1193–1226.
- Nasushita, R., Watanobe, H. and Takebe, K. (1997). A comparative study of adrenocorticotropin-releasing activity of prostaglandins E1, E2, F2 alpha and D2 in the rat. *Prostaglandins, Leukotrienes and Essential Fatty Acids* 56, 165–168.
- Oates, J. (1982). The 1982 Nobel Prize in physiology or medicine. *Science* 218, 765–768.
- Rivier, C. (1995). Influence of immune signals on the hypothalamic-pituitary axis of the rodent. *Frontiers in Neuroendocrinology* 16, 151–182.
- Vane, J. R., Bakhle, Y. S. and Botting, R. M. (1998). Cyclooxygenases 1 and 2. *Annual Review of Pharmacology and Toxicology* 38, 97–120.
- Watanabe, T., Morimoto, A., Morimoto, K., et al. (1991). ACTH release induced in rats by norepinephrine is mediated by prostaglandin E2. *Journal of Physiology* (London) 443, 431–439.
- Watanobe, H., Sasaki, S. and Takebe, K. (1995). Role of prostaglandins E1, E2 and F2 alpha in the brain in interleukin 1 beta-induced adrenocorticotropin secretion in the rat. *Cytokine* 7, 710–720.
- Zor, U. (1994). *Lipid mediators in health and disease*. London: Freund.

Proteases in Prokaryotes and Eukaryotic Cell Organelles

E Conway de Macario and A J L Macario

New York State Department of Health and The University at Albany (SUNY), Albany, NY, USA

Published by Elsevier Inc.

Protein Homeostasis

The Archaeal Proteasome

Other Proteases in Archaea

Compartmentalized Proteases in Bacteria

Membrane-Associated Proteolytic Machines in Bacteria

Proteases in Cyanobacteria, Mitochondria, and Chloroplasts

ATP-Independent Proteases

Glossary

- AAA+ proteins** A large family of proteins with ATPase capacity that participate in many cellular processes: AAA stands for ATPases associated with various cellular activities.
- Archaea** One of the three main evolutionary lines of descent or phylogenetic domains. The archaeal organisms are prokaryotes.
- Bacteria** One of the three phylogenetic domains. This domain includes all prokaryotes except archaea.
- Chloroplast** An organelle present in plant cells and photosynthetic protists capable of using

light as source of energy by means of green pigments (chlorophyll) to synthesize carbon compounds from carbon dioxide and water. It is evolutionarily related to cyanobacteria.

Clp

From caseinolytic protease, designates a large family of related enzymes that are also related with the Lon proteases and are present in bacteria, archaea, and eukaryotic organelles.

Eucarya
(*eukarya*)

One of the three main phylogenetic domains. This domain includes all the eukaryotes.

Mitochondrion

An organelle present in animal and plant cells capable of using oxygen to generate energy (aerobes); it synthesizes most of the cellular ATP via oxidative phosphorylation. It is evolutionarily related to the alpha-purple bacteria.

Organelle

A component of eukaryotic cells that resides within the confines of the cell membrane, in the cytosol, but is separated from the cytosol (subcompartment). Organelles contain many components with specialized functions that define their roles, e.g., respiration for the mitochondrion and photosynthesis for the chloroplast. Some organelles are considered to be derived from bacteria that entered into the eukaryotic cell, or its precursor, early in evolution.

Protein Homeostasis

A normal cell has many groups of proteins with distinctive functions. The members of each group must be in their native, functional conformation and within their physiological range of concentration if the cell is to remain fully functional and healthy. Maintenance of all proteins within their physiological conformations and concentration ranges, i.e., protein homeostasis, is achieved by a variety of means. Among these means are those that aim at assisting the nascent polypeptides to fold correctly, namely to achieve a native conformation, or to refold when the mature protein is damaged (unfolded) by stressors such as a sudden increase in temperature (heat shock), or a change in pH or in salinity. Molecular chaperones, discussed in other parts of this Encyclopedia, play a central role in protein folding and refolding.

Another aspect of protein homeostasis is the elimination of damaged molecules that are not repairable by the chaperoning systems, or the degradation of normal molecules that have become unnecessary to the cell after they have carried out their specific functions. Enzymes called proteases, usually forming large multimolecular assemblies such as the proteasome, play a central role in protein degradation and elimination, as discussed in this chapter. Thus the chaperoning and protein-degrading systems are two complementary arms of the whole mechanism responsible for protein homeostasis.

The Archaeal Proteasome

In archaea, the proteasome seems to be the central, if not the exclusive, energy-dependent protease. In contrast, the other prokaryotes, the bacteria, have several of these proteases, such as Lon and Clp (caseinolytic protein; discussed later), and some bacterial species have a proteasome in addition. In fact, the proteasome of one of these bacterial species, *Mycobacterium tuberculosis*, has been shown to be crucial for its survival in the face of stress.

Archaea have a 20S proteasomal particle, with the same overall structure as the eukaryotic counterpart. This overall structure also resembles that of the multimeric chaperonin complex GroEL/S typical of bacteria but also found in some archaea and in eukaryotic cell organelles. All of these complexes are barrel-shaped and have sevenfold rotational symmetry of the subunits arranged around a central cavity, which is accessible only from the ends. Despite the morphological similarities between the archaeal 20S proteasome and the folding chaperonin machine GroEL/S, these complexes differ in several important aspects,

reflecting their very different roles and mechanisms of action. One of the key differences is the size of the mouth of the chamber in its opened state: the mouth in GroEL/S is about four times bigger than that in the proteasome. This difference is very important, because the gate's diameter plays a crucial role in the selection of what goes into the chamber and what is kept outside. Simply on the basis of maximum gate size, we see that entry into the destruction chamber of the proteasome is a highly selective process, allowing the access of only linearized polypeptides brought near by a specific mechanism that includes the attachment of a defined tag to the polypeptide. All of these selection mechanisms operating together determine exactly which polypeptide from among the very many in the intracellular molecular soup will make its way toward destruction. Obviously, a number of mechanisms seem to have evolved in the direction of preserving intracellular structures, a fact that is manifested most evidently in the positioning of the enzymatic sites. These occur inside a chamber without any possibility of digesting surrounding molecules, except those that are allowed inside after careful screening, selection, and preparation.

The archaeal 20S proteasome is composed of 28 subunits, which are related in primary structure and can be classified into 14 families encompassing two superfamilies, alpha and beta. The alpha subunits oligomerize to form the two outer rings, one on each end of the central barrel, i.e., the proteasome core that contains the proteolytic chamber and is formed by two rings of seven beta subunits each. Thus, the similarities with the 20S particle from eukaryotes are evident.

Archaeal and bacterial 20S proteasomes have two to four different subunits (isoforms) in each ring, depending on the species. In contrast, all subunits differ from one another in the eukaryotic proteasome. It should be mentioned, however, that studies in progress are demonstrating more variants of archaeal proteasome subunits, and it is becoming apparent that some species have more isoforms than others. Furthermore, studies in archaea, bacteria, and eukaryotes suggest that certain species, or perhaps many species, have more than one type of 20S particle, each built differently, using different isoforms, and probably having distinct roles. This may be inferred from the fact that in eukaryotes each subunit uniquely occupies a defined position in the ring, suggesting strict control of function and great specialization of the overall structure. Thus, the discovery of distinct types of 20S particles should not be surprising, considering the array of roles that the proteasome plays in a variety of cellular processes.

The archaeal proteasome does not have a 19S particle, in contrast to the eukaryotic 26S proteasome, but it does have a regulatory cap. The archaeal proteasome regulatory factors are called proteasome-activating nucleotidases (Pan), and they are gathered at the entrance of the 20S particle in a manner reminiscent of the 19S particle structure in eukaryotes. The eukaryotic 19S particle has two sections, a lid and a base, the former composed of eight subunits lacking ATPase capacity. These subunits are termed regulatory particle non-ATPases (Rpn). The base has at least six ATPase subunits, which belong to the AAA+, type I proteins, and are called regulatory particle triple A proteins (Rtp). In addition, the base has other subunits that are of the Rpn type. The archaeal regulatory structure, which sits on the alpha rings, is built of Rpt homologs, termed Pan.

AAA+ proteins are proteins involved in many cellular processes, which justifies their name: ATPases associated with various cellular activities, or AAA+.

The final steps of proteolysis, after the proteasome has concluded its work, involve the giant protease TRI (tricorn), which digests the oligopeptides emerging from the proteasome to tri- and dipeptides. These are finally cut into single amino acids by the TRI-interacting factors (peptidases) F1, F2, and F3.

The proteasome in archaea may cooperate with membrane-associated proteases, such as the Lon protease, to translocate and degrade membrane-associated proteins.

Other Proteases in Archaea

Archaea, like bacteria and eukaryotic organelles, possess Clp- and Lon-related proteases. The proteolytic systems of archaea display features that combine bacterial and eukaryotic elements. This duality is just another example of the many structures and functions found in archaea, which seem to be a mixture of bacterial and eukaryotic components, such as those observed in the chaperoning systems and in the gene regulatory machinery.

Compartmentalized Proteases in Bacteria

There are at least five ATP-dependent compartmentalized proteases in bacteria. These bacterial proteases have counterparts in eukaryotic organelles, e.g., mitochondria and chloroplasts. The role of these proteases is to maintain the pool of cellular proteins, such that each species of proteins is present within its physiological range of concentration, and all molecules are in their native, functional conformations. Proteases contribute to this protein homeostasis

by eliminating protein molecules that are abnormal or those that are normal but have become unnecessary, dispensable, after they have done their work and that may interfere with the normal functions of the cell if allowed to stay around. Hence, these proteases are capable of demolishing tertiary and quaternary structures (e.g., they can disassemble multimeric complexes) and degrading their components.

Structure

The active protease structure (i.e., the proteolytic machine that recognizes the client polypeptides marked for degradation, prepares them and delivers them to the proteolytic sites, digests them, and releases the degradation products into the cytoplasm) is a large, barrel-like multimolecular complex. Like the proteasome, the basic bacterial protease structure is composed of multimeric rings stacked coaxially to enclose a central, cylindrical chamber, which is open at both ends. Typically, there are four rings; the central two rings contain the proteolytic sites on their inner surface. The peripheral rings, one on each end of the central proteolytic chamber, serve as substrate selectors and also capture the selected client polypeptide, prepare it (e.g., unfold it completely), thread the linearized substrate through the opening of the central cavity, and translocate it into the cavity. All of these activities, performed by the peripheral rings, require energy from ATP; thus, the components of these rings are ATPases. They belong to the large family of AAA+ ATPases.

Subunits

Several subunits are known to constitute the bacterial protease machines: ClpA, ClpB, ClpC, ClpE, ClpP, ClpX, and ClpY(HslU). Three of the main compartmentalized proteases in bacteria are ClpAP, ClpXP, and ClpXQ (or HslUY); these are composed of a protease (ClpP or ClpQ(HslV)) flanked by an ATPase ring formed by ClpA, ClpX, or ClpY(HslU), respectively. These flanking structures are formed by subunits, which are AAA + ATPases and which function in conjunction with the central structure (ClpP or ClpQ). However, they can also disaggregate protein aggregates by themselves.

Some of the genes encoding Clp proteins are stress inducible, for instance, *clpE*, *clpC*, and *clpP* in *Bacillus subtilis*; hence these proteins are members of the large group of stress proteins.

Rings

The subunits, six or seven of them together, associate to form rings with axial symmetry. Four of these rings

stack on top of each other to build the final structure. Since the rings associate coaxially, the final structure resembles a barrel with a chamber, open at both ends.

All of the subunits in each given ring are identical. The central two rings, which delimit the proteolytic chamber, are each formed by seven ClpP subunits. The two peripheral rings are formed by six subunits, either ClpA or ClpX (for ClpAP and ClpXP, respectively), and they sit on the central two-ring core, one on each end of this core. The peripheral rings in each final structure are identical, either ClpA or ClpX rings.

ClpB forms six-membered rings and interacts with the molecular chaperone machine (Hsp70(DnaK)-Hsp40(DnaJ)-GrpE) to form the Clp/KJE complex. This complex is a disaggregating machine capable of dissolving protein aggregates, particularly when these aggregates contain sHsp in addition to denatured polypeptides, which are the initial cause and key components of the aggregates.

ClpC is a stress protein (i.e., its parent gene is induced by stressors) that is apparently necessary for the survival of *B. subtilis* when this bacterium is faced with stressful conditions. ClpC forms complexes with ClpP, and the adaptor protein MecA (discussed later).

The ATPase peripheral rings play important roles when assembled together with the two central protease rings. These roles are substrate selection, capture, unfolding, and threading, and then translocation and delivery to the proteolytic chamber.

Substrate Selection

Proper selection of which protein will be disassembled and degraded is crucial. In the absence of selection, the entire pool of intracellular proteins (millions of them) would be destroyed rapidly and indiscriminately by the many protease machines extant inside the cell at all times. To avoid this chaotic protein destruction, there are essentially two mechanisms: (1) the enzymatically active sites are separated from the surrounding protein soup by the wall of the cylinder and (2) the peptides to be eliminated must have features that distinguish them from all the others and that make them recognizable by the peripheral rings as client polypeptides.

The proteasome of eukaryotes interacts with the ubiquitin system, which is the system for tagging the polypeptides that will then be recognized by the regulatory particle of the 26S proteasome. Bacteria do not have the ubiquitin system, but they do have tagging mechanisms that work under specific circumstances. In addition, the bacterial proteases can directly recognize partially denatured, damaged polypeptides through

signals consisting of stretches of unstructured sequence, usually near one of the polypeptide's ends. Hydrophobic patches, exposed in these unstructured-sequence segments, are the key as signaling elements.

Also, some client polypeptides are tagged by bacterial tagging mechanisms, such as the SsrA tag system. This mechanism operates when defective polypeptides are produced by stalling ribosomes. In this situation, a short signal of 11 amino acids is added to the C-terminal end of the incomplete polypeptide. SsrA targets the client polypeptide for degradation by ClpXP, ClpAP, and other proteases; it is not specific to any one of them.

Adaptors

Adaptors are proteins that activate proteolytic machines, or that facilitate proteolysis by these machines, by bringing together client polypeptide and machine. The SsbB adaptor joins SsrA-tagged polypeptides together with ClpX (an ATPase) and thus enhances translocation of the client polypeptide into the ClpP proteolytic chamber. Other adaptors are Mub and MecA; the latter interacts with ClpC. MuB inhibits degradation or disassembly of certain potential client polypeptides by occluding the degradation signals present in their sequences.

Another adaptor is ClpS, which interacts with ClpAP to form the ClpAPS complex, a very competent and active protease.

Membrane-Associated Proteolytic Machines in Bacteria

In addition to the protease complexes discussed previously, all of which are in the cytoplasm, bacteria have other protease machines that are associated with their membranes. These machines carry out the degradation of membrane-associated client polypeptides. One of these machines is FtsH(HflB), which forms a very large complex (holoenzyme) with HflKC. FtsH also cooperates with the membrane-integrated metalloprotease HtpX for processing of metal-containing proteins.

Proteases in Cyanobacteria, Mitochondria, and Chloroplasts

In the cyanobacterium *Synechococcus elongatus*, several proteases, Clp-1, Clp-2, Clp-3, and ClpR, and the adaptor ClpS, have been identified. ClpS interacts with ClpA to form ClpAPS, as in other bacteria. Related proteases are found in the chloroplast, presumably evolutionarily related to those found in cyanobacteria.

Mitochondria, which are considered to be derived from ancient alpha purple bacteria – i.e., the ancestors of today's alpha purple bacteria – possess several proteases, including members of the two main AAA+ATPase families. These mitochondrial ATPases are associated with the inner membrane. In the mitochondrial matrix, there are homologs of the bacterial ClpB, ClpX, and Lon, named mtHsp78, mtMcp1, and mLon, respectively. Like their bacterial counterparts, these mitochondrial proteases form a proteolytic chamber with two stacked rings each made of seven identical subunits, and two regulatory rings, one at each end of the core proteolytic chamber, composed of six identical rings, which are distinct from those in the central core rings.

ATP-Independent Proteases

In addition to the ATP-dependent proteases discussed previously, there are others that do not require ATP. For example, one such group of proteases is the DegP (HtrA) family of serine proteases, which encompasses nearly 2000 members. They are widespread and are highly conserved in sequence, being present in the periplasmic space of Gram-negative bacteria and in the ER and chloroplast of eukaryotes.

Interestingly, DegP in the bacterium *Escherichia coli* has either chaperoning or proteolytic ability, depending on the temperature. At physiological temperatures, DegP functions as a chaperone, but at high temperatures it becomes a protease. Like the ATP-dependent proteases, DegP forms oligomeric rings (three subunits per ring) that stack on each other, coaxially, and thus form pairs of hexameric complexes enclosing a central cavity.

See Also the Following Articles

Chaperone Proteins and Chaperonopathies; Heat Shock Genes, Human; Heat Resistance; Heat Shock Response, Overview; Viral Virulence and Stress; Chaperonopathies; Proteases in the Eukaryotic Cell Cytosol.

Further Reading

- Borissenko, L. and Groll, M. (2005). Crystal structure of TET protease reveals complementary protein degradation pathways in prokaryotes. *Journal of Molecular Biology* **346**, 1207–1219.
- Dura, M. A., Receveur-Brechot, V., Andrieu, J.-P., et al. (2005). Characterization of a TET-like aminopeptidase complex from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Biochemistry* **44**, 3477–3486.
- Franzetti, B., Schoehn, G., Hernandez, J.-F., et al. (2002). Tetrahedral aminopeptidase: A novel large protease complex from Archaea. *EMBO Journal* **21**, 2132–2138.
- Guagliardi, A., Mancusi, L. and Rossi, M. (2004). Reversion of protein aggregation mediated by Sso7d in cell extracts of *Sulfolobus solfataricus*. *Biochemistry Journal* **381**, 249–255.
- Mogk, A., Deuerling, E., Vorderwuelbecke, S., Vierling, E. and Bukau, B. (2003). Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. *Molecular Microbiology* **50**, 585–595.
- Maupin-Furlow, J. A., Gil, M. A., Karadzic, I. M., Kirkland, P. A. and Reuter, C. J. (2004). Proteasomes: perspectives from the Archaea. *Frontiers in Bioscience* **9**, 1743–1758 [online journal].
- Maurizi, M. R. and Xia, D. (2004). Protein binding and disruption by Clp/Hsp100 chaperones. *Structure* **12**, 175–183.
- Pallen, M. J., Lam, A. C. and Loman, M. (2001). Tricorn-like proteases in bacteria. *Trends in Microbiology* **9**, 518–521.
- Piszczyk, G., Rozycki, J., Singh, S. K., Ginsburg, A. and Maurizi, M. R. (2005). The molecular chaperone, ClpA, has a single high affinity peptide binding site per hexamer. *Journal of Biological Chemistry* **280**, 12221–12230.
- Schlee, S., Beinker, P., Akhrymuk, A. and Reinstein, J. (2004). A chaperone network for the resolution of protein aggregates: direct interaction of ClpB and DnaK. *Journal of Biological Chemistry* **336**, 275–285.
- Tamura, N., Lottspeich, F., Baumeister, W. and Tamura, T. (1998). The role of Tricorn protease and its aminopeptidase-interacting factors in cellular protein degradation. *Cell* **95**, 637–648.
- Zhang, X., Beuron, F. and Freemont, P. S. (2002). Machinery of protein folding and unfolding. *Current Opinion in Structural Biology* **12**, 231–238.

Proteases in the Eukaryotic Cell Cytosol

E Conway de Macario and A J L Macario

New York State Department of Health and
The University at Albany (SUNY), Albany,
NY, USA

Published by Elsevier Inc.

molecules with ubiquitin-like properties) that tag and prepare the client polypeptide for digestion in the proteasome.

Introduction

Protein-Folding and Protein-Degrading Multimolecular Machines

The Protein Degradation Pathway Involving the Ubiquitin-Proteasome System (UPS)

The Proteasome

Steps Toward Degradation by the Proteasome

The Whole Picture

Glossary

<i>ATP-dependent protease</i>	A protease that requires energy from ATP hydrolysis to perform its functions.
<i>Client polypeptide</i>	A protein destined for degradation, proteolysis (the protease substrate).
<i>Cofactors</i>	Molecules (usually proteins) that cooperate with chaperones in directing the client polypeptide to the protease, and in preparing it (e.g., unfolding it) so that the protease can receive it for digestion.
<i>Protease</i>	An enzyme that digests proteins, breaking them down into components of various lengths and ultimately producing, in many cases, single amino acids. Different proteases differ in complexity: some are formed by a single molecule, whereas others are assemblages of several molecules.
<i>Proteasome</i>	An example of a proteolytic machine, typically present in the cytosol of eukaryotic and archaeal cells, and also in a small number of bacterial species.
<i>Proteolytic machine</i>	A protease that is a multimolecular assembly, in which each component plays a role in the series of events leading to protein digestion and release of cleavage products.
<i>Self-compartmentalized protease</i>	A proteolytic machine that has its active proteolytic sites inside a cavity, inaccessible to surrounding molecules, and thus requiring an active mechanism for entry of the substrate into the cavity for digestion.
<i>Ubiquitin-proteasome system (UPS)</i>	A protein-degrading system involving the proteasome and a series of enzymes. These enzymes are the ubiquitins (or other

Introduction

Small vs. Large Proteases

Protein homeostasis requires production of polypeptides as well as their regulated destruction, when they are no longer needed or are defective. Molecular chaperones play key roles in both the production of functional proteins and the destruction of useless ones. However, the central players in the destruction phase are the proteases. These enzymes occur, like the chaperones, in every cell and cell compartment in all organisms.

Proteases encompass a wide range of structures, from monomeric hydrolases to multimolecular, giant machines with molecular masses on the order of 1 million Daltons. The large molecular assemblies are self-compartmentalized and have their proteolytic sites inside a central chamber that can be reached only by selected substrates.

This article deals with the large, self-compartmentalized, multimolecular proteolytic machines in the three domains of life: archaea, bacteria, and eucarya (or eukarya).

A General Outline of the Process of Proteolysis

Proteins are, in general, degraded via a series of steps: (1) tagging, i.e., the polypeptide destined for degradation must have a tag, which makes it recognizable by the proteases or by the molecules that will direct the polypeptide to the protease; (2) unfolding, which is usually done via an ATP-dependent mechanism by specific components of proteases, such as Lon/Clp in bacteria and the 26S proteasome in eukaryotes (since at this stage the main event is the unfolding of the polypeptide, the molecules that do the unfolding are called unfoldases); (3) ATP-dependent proteases perform the initial enzymatic attack, generating oligopeptides with fewer than 30 amino acids; (4) these oligopeptides are then further digested by endopeptidases, tripeptidyl peptidases, and dipeptidyl peptidases, yielding still shorter oligopeptides (fewer than 10 amino acids long); and (5) these very short oligopeptides are finally digested into still shorter oligopeptides and single amino acids by aminopeptidases, carboxypeptidases, and di- and tripeptidases. Some of these proteases, e.g., TRI (tricorn) and TET (tetrahedral aminopeptidase), are large multimolecular

complexes like the proteasome. For example, TRI, which is present in some archaea and some bacteria (eukaryotes have in the cytosol a protease, tripeptidyl peptidase II (PPDII), which may be the functional equivalent of TRI), takes the oligopeptides produced by the proteasome and digests them to tri- and dipeptides, which are subsequently cleaved into single amino acids by peptidases called TRI-interacting factors F1, F2, and F3. In summary, unfoldases, the proteasome, TRI, and TRI-interacting factors constitute a multimolecular disassembly line. This general schema applies also to other protein-degrading complexes, some of which are discussed in this article.

Protein-Folding and Protein-Degrading Multimolecular Machines

Polypeptides misfolded because of a structural defect (mutation or posttranslational modification) or unfolded because of stress, as well as incompletely folded (nascent) normal polypeptides, in the cytosol confront a dangerous, highly competitive environment, mainly due to overcrowding. All of these kinds of polypeptides tend to entangle, stick to one another, and form aggregates and precipitates, which can be irreversible and toxic to the cell.

As the misfolded or unfolded polypeptides appear in the cytosol, they encounter the molecular chaperone machine formed by three major teammates, the proteins Hsp70(DnaK), Hsp40(DnaJ), and nucleotide exchange factor. The latter is GrpE (glucose-regulated protein E) in prokaryotes, and BAG-1 (BCL2-associated athanogene, where BCL2 stands for B cell lymphoma 2) or HspBP1 (Hsp70-binding protein 1) in eukaryotes. After this encounter, the polypeptide follows one of two alternative pathways, either toward folding or toward degradation, depending to some extent on its own characteristics and on which cofactors join the chaperone machine. In the eukaryotic cytosol, Hip (Hsc70-interacting protein) and Hop (Hsp70/Hsp90-organizing protein) promote folding, whereas BAG-1 and CHIP (C-terminus of Hsc70-interacting protein) direct the polypeptide toward degradation. In both cases, the polypeptide ends up inside a barrel-shaped chamber.

The folding and proteolytic machines are multimolecular assemblies with a central cavity or chamber. In the prokaryotes bacteria, the folding machine is the complex GroEL/S, and in the eukaryotic cytosol it is the CCT (chaperonin-containing TCP-1, also called TriC for TCP-1 ring complex, where TCP stands for tailless complex polypeptide). An equivalent of CCT, named thermosome, occurs in the prokaryotes archaea. In addition, some archaeal species also have GroEL/S-like bacteria.

The proteolytic machine in bacteria is the Clp complex, whereas in the eukaryotes and archaea it is the proteasome: 26S in the former and 20S in the latter. The organelles of the eukaryotic cell also have folding and proteolytic machines; for example, the mitochondrion has equivalents of the bacterial Clp complex for proteolysis, and it harbors equivalents of GroEL and GroES for protein folding.

The Protein Degradation Pathway Involving the Ubiquitin-Proteasome System (UPS)

In the eukaryotic cytosol, the pathway toward degradation includes various steps and involves diverse components. Initially, the molecular chaperone machine goes into action, and then the cofactors BAG-1 and CHIP interact with the machine and with the next group of participants in the pathway leading to the proteasome. This group is the ubiquitin system.

Ubiquitins

In order for a polypeptide that must be degraded to reach the proteasome, it must first be tagged so that the machine can recognize it. The tag is the ubiquitin molecule, and the tagging is done by the ubiquitin system, which consists of diverse enzymes sorted into three groups, E1, E2, and E3. The E1 are ubiquitin-activating enzymes that prepare ubiquitin for interaction with a carrier, the E2 ubiquitin-conjugating enzymes. E3 ubiquitin ligases bring together the polypeptide destined for degradation (substrate or client polypeptide) and E2, at which point the ubiquitin molecule is attached (ubiquitylation) to the client polypeptide. The proper tag that indicates proteolysis is composed of a chain of at least four ubiquitin molecules; peptides tagged with one to three ubiquitins are directed toward other cellular processes and not to degradation by the proteasome (although they may be directed toward degradation by the lysosome). It must be emphasized that ubiquitylation is widespread in the cell and participates in many processes distinct from protein degradation: cell cycle, DNA repair, regulation of gene transcription, apoptosis, and others.

Ubiquitylation

Ubiquitin is a 76-amino-acid polypeptide, and ubiquitylation occurs via formation of an isopeptide bond between an internal lysine of the substrate and the C-terminal glycine (glycine 76) of ubiquitin. Successive additions of more ubiquitins to form a polyubiquitin chain of at least four units, i.e., a good signal for the proteasome, use the same mechanism

and bond type, but the ubiquitin residues involved are glycine 76 and lysine 48: new ubiquitins are attached to an internal lysine of the last ubiquitin molecule in the chain. A polyubiquitylated substrate is then recognized by the proteasome.

Chaperones and Cofactors for Protein Degradation

A pivotal component of the chaperone machine, and thus of the folding and degradation pathway, is Hsp70. This molecule has segments, motifs or domains, that interact with the other components of the machine and, most importantly, with the client polypeptide in need of assistance for folding or refolding, and also with the polypeptide beyond recovery that must be degraded.

The client polypeptide usually displays hydrophobic residues that, in a folded, native protein, lie in the interior of the molecule, inaccessible to other molecules in the immediate surroundings. Incompletely folded, partially unfolded (denatured), and misfolded polypeptides expose hydrophobic patches on their surface, and these attract Hsp70, which has a peptide-binding domain. This domain is flanked, in the linearized Hsp70 sequence, by an ATPase domain on the N-terminal side and by a lid on the C-terminal side, including the sequence glutamate-glutamate-valine-aspartate (EEVD) at the C-terminus.

The N-terminal section of Hsp70 functions as an ATP-binding and ATPase domain; in addition, it is bound by a number of cochaperones or cofactors of the BAG family (e.g., BAG-1 through 6 and HspBP1) and by Hip. With its C-terminal domain, Hsp70 can bind CHIP and Hop.

BAG-1 functions as a nucleotide exchange factor. The various BAG proteins each have on their C-terminal halves the BAG domain that enables them to bind Hsp70 as well as several other proteins (since they participate in a number of cellular processes). In addition to the BAG domain situated near the C-terminus, BAG family members have a ubiquitin-like domain in the middle of the molecule. This domain can bind to the proteasome and can thus direct a client polypeptide already bound to Hsp70 toward degradation.

CHIP has a tetratricopeptide repeat (TPR) domain with several copies of the TPR in its terminal half, a U box near its C-terminus, and a coiled-coil motif in between. This latter motif participates in protein-protein interactions, whereas the TPR domain is involved in the interaction of CHIP with the C-terminus of Hsp70, and the U box enables CHIP to act as an E3 ubiquitin ligase. The activity of CHIP mimics that of an E3 ligase: it binds to the Hsp70/substrate

complex and then the E2 molecule Ubc4/5, and it thus facilitates the enzymatic polyubiquitylation of the substrate with the participation of an E2 enzyme.

A TPR motif or domain consists of a series of TPRs in tandem; the TPRs are imperfect repeats of 34 amino acids each. This motif is present in many proteins and participates in a great variety of molecular interactions.

In summary, both BAG-1 and CHIP bind the C-terminus of Hsp70: BAG-1 via its C-terminal BAG domain and CHIP via its N-terminal TPR domain. CHIP directs the client polypeptide to the proteasome in association with Ubc4/5 or in association with Ubc4/5 and BAG-1. Another chaperone, Hsp90, comes into play for other similar processes involving specific substrates, e.g., steroid hormone receptors. Here, the general mechanism is essentially the same as the one just described for those cases in which Hsp90 does not participate.

The Proteasome

The proteasome is one of the large protease complexes that have a central chamber, roughly cylindrical, in which proteolysis takes place. Other examples of similar proteolytic chambers are the Clp complexes that are found in the cytoplasm of bacteria and inside organelles in eukaryotes.

The proteasome has been found in the cytosol and nucleus of eukaryotic cells, in the cytoplasm of archaea, and in the cytoplasm of some bacteria, such as gram-positive actinomycetes (e.g., *Mycobacterium tuberculosis*). The eukaryotic proteasome is identified as the 26S particle, whereas those of archaea and bacteria are identified as 20S particles, according to their approximate sedimentation coefficients.

Structure

The 26S proteasome is composed of two main particles (or subparticles), the 20S and the 19S particles, both of which are polymers typically composed of 28 and 17–19 subunits, respectively. The 20S is the proteolytic particle, whereas the 19S particle is regulatory. In current terminology, 26S particle denotes a 20S particle with one 19S particle on one end and a 20S particle with two 19S particles, one on each end.

The 20S particle is formed by four stacked rings, two alpha and two beta rings, the latter in the middle, and one alpha ring on each end of the central barrel formed by the central beta rings. Each ring in turn is constituted of seven subunits, alpha and beta for the alpha and beta rings, respectively. All of the alpha subunits and all of the beta subunits are related in primary structure, but not identical. These subunits are arranged in rings with a central cavity, and since

the four rings are stacked coaxially in the overall structure, the 20S proteasome can be described as a cylinder, resembling a barrel.

The 19S particle is composed of at least 17 distinct subunits arranged in two complexes (or subcomplexes), the base and the lid. The base subcomplex sits on the alpha ring of the 20S particle and is formed by nine subunits, of which six are ATPases. The lid is distal to the base subcomplex, in contact with the intracellular milieu, and is formed by eight subunits devoid of ATPase capacity.

Role of the Proteasome Subunits and Particles

While the 20S particle is involved in the actual degradation (proteolysis) of the client polypeptide, the 19S particle has other functions, such as (1) recognition and sorting of the client polypeptides, i.e., those correctly ubiquitinated with chains of four or more ubiquitins; (2) opening of the mouth (gate) of the proteasome to allow entry of the substrate; (3) unfolding of the client polypeptide as much as necessary for it to negotiate the entrance to the proteasome, which is very narrow and does not allow passage of folded proteins; and (4) threading of the extended, unfolded substrate through the gate and transfer of it to the alpha ring of the 20S particle and beyond, so that the polypeptide reaches the proteolytic chamber delimited by the central beta rings. All of these steps require energy from ATP hydrolysis, which is why the base subcomplex has six ATPase subunits.

Steps Toward Degradation by the Proteasome

Calling for Destruction

The polypeptide destined for degradation has short stretches (5–25 amino acids) that constitute hydrophobic patches and mark it for destruction. However, the proteasome does not recognize such signals; for this to happen, a tag, the polyubiquitin chain described earlier, must be added to the client polypeptides. There may be exceptions to this rule. It is possible that some E3 enzymes can interact directly with the proteasome or that some client polypeptides can be recognized directly by the proteasome and engulfed for degradation without participation of E1 and E2 enzymes. These would be examples of what is called ubiquitin-independent protein degradation by the proteasome.

Unfolding, Release of Chaperones, Gate Opening, and Threading

The client polypeptide cannot enter the mouth of the proteasome if it is not opened first and if the

polypeptide is not linearized and freed from the chaperones and cochaperones accompanying it before it is offered to the mouth. Thus, a mechanism for unfolding polypeptides operates at the mouth of the proteasome. Chaperone–cochaperone release, polypeptide unfolding, and gate opening are probably functions of the 19S particle, but the mechanism has not yet been elucidated.

Translocation from the Cytosol into the Chamber

The unfolding, threading, and translocation of the client polypeptide are most likely simultaneous events, although unfolding may partially precede threading, which in turn may begin before the start of translocation and proceed during translocation, and the final moments of the translocation may occur after the preceding events have concluded. By this time, the leading end of the polypeptide will have reached the proteolytic chamber, with proteolysis starting immediately. It is possible that even proteolysis begins almost simultaneously with translocation.

Proteolysis

Unstructured, less stable segments of the incoming, extended polypeptide are probably recognized by the proteolytic sites lining the inside of the proteolytic chamber, and digestion starts immediately. This part of the process is not specific. Other steps, preceding the actual digestion inside the proteasome, are selective: they depend on specific interactions between molecules that possess certain markers that make them interact only in certain combinations and, thus, recognize and process substrates with signals that are specific for only one of these combinations. In these steps, the role of the E enzymes is crucial, as is explained later.

Indeed, the proteolytic systems are finetuned so as to avoid indiscriminate digestion of cellular proteins and to steer toward destruction only those that must be eliminated for the good of the cell. The entire protein degradation process is strictly regulated by a number of mechanisms involving the E enzymes and the 19S proteasome components, as well as other mechanisms that insert specific signals on those polypeptides that are not wanted in the cell.

The proteolytic efficiency inside the proteasome appears to be very high, possibly a consequence of a concentration of enzymatic sites and the ease of encounter and contact between these sites and the extended substrate sliding naked, as it were, in close proximity of the sites. The resulting products are oligopeptides of 3–30 amino acids.

Exit of Digestion Products

Oligopeptide exit from the proteasome seems to be regulated, too, perhaps as strictly as the selection and entrance of the client polypeptide. However, little is known of the exit process. We can assume that rapid escape must be avoided, so that there will be no generation of incompletely digested substrate fragments, i.e., incomplete degradation. Likewise, excessive retention of completely digested fragments would have to be avoided lest the proteolytic chamber and the exit gate be clogged; such clogging would slow the flux of the whole process, including the pace of the proteolytic activity.

Processing and Reuse of Digestion Products

The oligopeptides emerging from the proteasome are cleaved into shorter oligopeptides and finally into single amino acids by aminopeptidases, carboxypeptidases, and di- and tripeptidases, e.g., in archaea TRI and TET, which are large multimolecular complexes like the proteasome, and TRI-related proteases in eukaryotes. The amino acids generated are used again in the synthesis of new polypeptides on the ribosome.

The Immunoproteasome

The proteasome also plays a role in immunity: it processes antigen and produces antigenic oligopeptides for display on the cell surface associated with other molecules as required for triggering a certain group of T lymphocytes. This type of antigen processing is the function of a specialized proteasome, the immunoproteasome, which differs in some details from the other proteasome discussed earlier, which is involved in general protein degradation.

Proteasomes at inflammatory sites undergo induction of regulatory subunits in the 19S particle and changes in the proteolytic subunits in the 20S particle: subunits beta1, beta2, and beta5 are replaced by their immuno counterparts, beta1i (LPM2), beta2i (MECL-1), and beta5i (LMP7), respectively. The 20S particle is assembled *de novo* with the immuno and the regular subunits in, for example, cells stimulated by γ -interferon or tumor necrosis factor. The other, so-called constitutive proteasome is present in normal cells and in immunoprivileged locales, e.g., brain tissue.

Proteins that appear as antigens to the organism are degraded in the immunoproteasome, and the resulting oligopeptides (8–10 amino acids long) are transported from the cytosol to the endoplasmic reticulum (ER) with the participation of TAP (transporter associated with antigen processing) proteins. In the ER, the peptides bind MHC (major histocompatibility complex) class I molecules, and the complex is

inserted on the cell surface. This type of antigen presentation is critical for the action of certain T lymphocytes, e.g., cytotoxic T cells recognize and interact with the peptide–MHC complex, which may lead to the lysis of the cells carrying the complex on the surface as a means of eliminating potentially damaging cells.

It is believed that the immunoproteasome generates correct epitopes from the C-terminus of the protein antigen, but other enzymes are needed for generation of N-terminus epitopes. Some of the latter are LAP (leucine aminopeptidase) and ERAP 1 (ER aminopeptidase 1), which, as its name suggests, is associated with the ER.

Proteasome Pathology

The proteasome, and the immunoproteasome, can be defective and can thus compromise cell physiology and health. Since the proteasome is the major protagonist in the elimination of proteins that are not beneficial to the cell, when the proteasome function is impaired or stops, the effects are immediate and considerable. Likewise, failure of the immunoproteasome will have an impact on several functions of the immune system, with repercussions throughout the entire body.

Proteasome inhibition by chemical compounds, genetic defects in the component subunits, and acquired alterations of these subunits have all been shown to result in proteasome malfunction and in pathological accumulation of protein aggregates. This effect can be even more deleterious in proteinopathies (e.g., Alzheimer's, Parkinson's, and Huntington's disease). The pathogenesis of these diseases includes misfolding and aggregation of genetically defective proteins, which form intra- and extracellular precipitates in various tissues, particularly the brain, resulting in cell death. Proteasome failure can only aggravate this pathogenetic sequence. Also, it has been observed that the proteasome suffers alterations as the process of senescence progresses.

The Whole Picture

The diversity of the proteasome is evidenced by the variety of subunits encoded in the genomes already sequenced. Likewise, the conservation of the proteasome is apparent in the similarity among all of the subunits, particularly within a single type, alpha or beta, and in the overall architecture, which essentially consists of stacked rings of subunits assembled together to enclose a central cavity, with axial gates at both ends, in all species known to have proteasome.

In eukaryotes, proteasomes have been found in the cytosol and nucleus, and in association with the ER

and the cytoskeleton. Also, proteasomes have been found in archaea and in some bacteria, as mentioned earlier.

It has been demonstrated, or it is believed based on some experimental data, that the proteasome plays a key role not only in protein quality control as discussed in this article, but also in several other cellular processes, such as the immune response, cell division, metabolism, DNA repair, and control of gene transcription. In these instances, the ubiquitin system also intervenes.

See Also the Following Articles

Chaperone Proteins and Chaperonopathies; Chaperonopathies; Heat Shock Genes, Human; Heat Resistance; Heat Shock Response, Overview; Proteases in Prokaryotes and Eukaryotic Cell Organelles; Viral Virulence and Stress.

Further Reading

- Ciechanover, A. (2005). Proteolysis: from the lysosome to ubiquitin and the proteasome. *Nature Reviews Molecular Cell Biology* 6, 79–86.
- Ciechanover, A. and Iwai, K. (2004). The ubiquitin system: from basic mechanisms to the patient bed. *IUBMB Life* 56, 193–201.
- Esser, C., Alberti, S. and Hoehfeld, J. (2004). Cooperation of molecular chaperones with the ubiquitin/proteasome system. *Biochimica et Biophysica Acta* 1695, 171–188.
- Geier, E., Pfeifer, G., Wilm, M., et al. (1999). A giant protease with potential to substitute for some functions of the proteasome. *Science* 283, 978–981.
- Glickman, M. H. and Ciechanover, A. (2002). The ubiquitin-proteasome proteolytic pathway: Destruction for the sake of construction. *Physiology Reviews* 82, 373–428.
- Goldberg, A. L. (2003). Protein degradation and protection against misfolded or damaged proteins. *Nature* 426, 895–899.
- Kisselev, A. F., Akopian, T. N., Woo, K. M. and Goldberg, A. L. (1999). The sizes of peptides generated from protein by mammalian 26 and 20S proteasomes: implications for understanding degradative mechanisms and antigen presentation. *Journal of Biological Chemistry* 274, 3363–3371.
- Kloetzel, P. M. and Ossendorp, F. (2004). Proteasome and peptidase function in MHC-class-I-mediated antigen presentation. *Current Opinion in Immunology* 16, 76–81.
- Lum, R., Tkach, J. M., Vierling, E. and Glover, J. R. (2004). Evidence for an unfolding/threading mechanism for protein disaggregation by *Saccharomyces cerevisiae* Hsp104. *Journal of Biological Chemistry* 279, 29139–29146.
- Maupin-Furlow, J. A., Gil, M. A., Karadzic, I. M., Kirkland, P. A. and Reuter, C. J. (2004). Proteasomes: Perspectives from the Archaea. *Frontiers in Bioscience* 9, 1743–1758.
- Muratami, M., Kung, C., Shokat, K. M. and Tansey, W. P. (2005). The F box protein Dsg1/Mdm30 is a transcriptional coactivator that stimulates Gal4 turnover and cotranscriptional mRNA processing. *Cell* 120, 887–899.
- Ortega, J., Heymann, J. B., Kajava, A. V., Ustrell, V., Rechsteiner, M. and Steven, A. C. (2005). The axial channel of the 20 S proteasome opens upon binding of the PA200 activator. *Journal of Molecular Biology* 346, 1221–1227.
- Vaux, D. L. and Silke, J. (2005). IAPs, RINGs and ubiquitylation. *Nature Reviews Molecular Cell Biology* 6, 287–297.
- Zhang, X., Beuron, F. and Freemont, P. S. (2002). Machinery of protein folding and unfolding. *Current Opinion in Structural Biology* 12, 231–238.

Protein Assembly See: Protein Synthesis; Chaperone Proteins and Chaperonopathies; Heat Resistance; Heat Shock Response, Overview; Chaperonopathies; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles.

Protein Folding See: Protein Synthesis; Chaperone Proteins and Chaperonopathies; Heat Resistance; Heat Shock Response, Overview; Chaperonopathies; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles.

Protein Synthesis

M A Brostrom and C O Brostrom

Robert Wood Johnson Medical School, Piscataway,
NJ, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3,
pp 273–282, © 2000, Elsevier Inc., with revisions made by the
Editor.

Phases of Protein Synthesis

Regulation of Translational Initiation

Cellular Stress and the Inhibition of mRNA Translation

Comparable Roles of the Endoplasmic Reticulum and
Cytoplasmic Stress Responses in Recovery of
Translational Activity

eIF2 Kinase as a Focal Point for Regulating Protein
Synthesis during Stress

Glossary

<i>Chaperone</i>	Protein that functions to support other proteins or growing polypeptides in a particular conformation.
<i>Cytoplasmic stress</i>	Condition associated with the accumulation of misfolded or damaged proteins in the cytoplasm.
<i>Endoplasmic reticulum (ER)</i>	Subcellular organelle in which newly synthesized secretory, lysosomal, and membrane proteins are processed and folded.
<i>ER stress</i>	Condition associated with the accumulation of unfolded or underprocessed proteins in the ER.
<i>GRP78/BiP</i>	A glucose-regulated 78-kDa ER-resident chaperone induced in response to ER stress.
<i>HRI</i>	Heme-regulated eIF2 kinase that functions to inhibit translational initiation in erythroid cells.
<i>HSP/HSC70</i>	A 70-kDa cytoplasmic chaperone(s) induced in response to cytoplasmic stress.
<i>PKR</i>	An interferon-inducible, double-stranded RNA-activated protein kinase that phosphorylates eukaryotic initiation factor (eIF) 2 to inhibit translational initiation.
<i>PKR-like ER kinase</i>	An ER membrane-spanning eIF2 kinase that phosphorylates eIF2 in response to ER stress.
<i>Proteotoxic</i>	Condition resulting in the misfolding or damage of cellular proteins.
<i>Translational elongation</i>	Process occurring on 80S ribosomes through which amino acids are transferred from aminoacyl-tRNAs to the growing

polypeptide chain sequentially in accord with the internal codons of mRNA.

Translational initiation

Process by which the 80S ribosome complex carrying the initiator tRNA, met-tRNA, and mRNA properly phased at the initiation codon is assembled.

Protein synthesis refers to the biological process whereby amino acids are assembled by peptide bonding into specific polypeptide sequences in accord with genetic blueprints encoded by deoxyribonucleic acid (DNA).

Phases of Protein Synthesis

The overall process of protein synthesis occurs in several phases, each of which is highly complex and generally viewed as a separate field of study.

DNA-Directed Nuclear Transcription of Temporary Ribonucleic Acid Constructs (Messenger RNA) and the Processing of mRNA into Ribonuclear Particles for Movement to the Cytoplasm

Information transfer from DNA to mRNA is reflected by the synthesis of analogous purine/pyrimidine ribonucleotide sequences arranged in triplet groupings (codons) that specify individual amino acids. General features of the mRNA structure include a 5' to 3' phosphodiester linkage of nucleotides, a 5'-methylated guanylate cap, a 5'-untranslated region, a coding region, and a 3'-polyadenylate tail. Nucleotide base pairing produces variable degrees of internal looping and winding. A number of stress-related hormones, including adrenal steroids and thyroxine, alter the synthesis of mRNAs encoding specific proteins. The relationship of these transcriptional alterations to the general slowing of protein synthesis produced by these hormones in many tissues is unclear.

Ribosomal Synthesis of Protein through Translation of mRNA Codons

This process is divisible into two phases: translational initiation, which relates to the loading of ribosomes onto mRNA, and translational elongation, which refers to the linkage of amino acids on the ribosomal surface by peptide bonding to produce a growing polypeptide. Eukaryotic translational initiation involves the binding of the small (40S) ribosomal subunit near the mRNA cap and its subsequent movement

(scanning) in a 5′–3′ direction until it detects the first AUG (initiator) base sequence. This triplet ordinarily provides the starting point for reading additional codons specifying the amino acid sequences for assembling proteins. Initiation is catalyzed by at least nine distinct proteins termed initiation factors (eIF) composed of some 25 assorted protein subunits. A number of these factors are subject to alterations of catalytic activity through protein phosphorylation by various protein kinase activities. Scanning involves an intricate interplay of these initiation factors with the 40S ribosomal subunit and mRNA. The 40S ribosomal subunit is converted to a 43S preinitiation complex by association with eIF3 and with a ternary complex consisting of eIF2, methionyl initiator transfer RNA (met-tRNA_i), and GTP. Binding of eIF4F (a complex of 4E, 4A, and 4G) and eIF4B produces helicase mediated unwinding (melting) of the cap region of mRNA that permits cap-proximal association of the 43S preinitiation complex. Stabilization of this complex with subsequent scanning of the ribosomal subunit to the initiator AUG appears to require at least two additional factors, eIF1A and eIF1. Upon binding of the anticodon of met-tRNA_i to the initiator AUG by codon–anticodon base pairing, eIF5 stimulates hydrolysis of the GTP associated with eIF2. This hydrolysis is associated with the binding of the initiator met-tRNA_i at the 40S P-(peptidyl binding) site, the release of all initiation factors from the 40S ribosomal subunit, and the binding of the large (60S) ribosomal subunit to form an 80S ribosome competent of performing peptide chain elongation.

Translational elongation involves the addition of amino acids to the growing peptide chain as specified by the sequence of mRNA codons occurring after the initiator AUG. The process requires two elongation factors (eEF) and a series of transfer RNAs, each of which possesses a specific anticodon and is subject to enzymatic aminoacylation with a specific amino acid. Peptide chain assembly proceeds through eEF1-mediated binding of aminoacyl-tRNA molecules to a binding site (A site) on the ribosome such that the anticodon of the tRNA associates with the codon of the mRNA. The carboxyl end of the peptide chain (or the initiator methionine) of the tRNA in the P site is uncoupled from the tRNA and is covalently linked by peptide bonding to the amino group of the aminoacylated tRNA in the A site. Peptide bonding is catalyzed by a peptidyl transferase activity associated with the ribosomal RNA of the 60S subunit. In conjunction with the hydrolysis of GTP, eEF2 supports the translocation of the resultant peptidyl-tRNA from the A site to the P site and the movement of the ribosome three nucleotides (1 codon)

along the mRNA molecule. The A site is then available for binding another aminoacyl-tRNA in accord with the specificity of the next codon on mRNA. The process of peptide bond formation and ribosomal movement continues until a stop codon (UAA, UAG, or UGA) is reached on the mRNA. In conjunction with a release factor, the peptidyl transferase catalyzes the hydrolysis of the peptidyl-tRNA linkage, permitting the release of the completed peptide.

Posttranslational Protein Processing

Newly synthesized polypeptides may be released directly into the cytosol or targeted for translocation into the endoplasmic reticulum (ER) by a short series of amino acids at the amino terminal end of the protein termed the signal peptide sequence. In either event the polypeptides are biologically inactive until folded to proteins with specific three-dimensional structures. The development of biological activity frequently requires the binding or covalent linkage of various prosthetic groups and the association of polypeptides into proteins consisting of multiple (multimeric) subunits. Cotranslational protein folding in the cytoplasm (cytosol plus mitochondria) is assisted by various protein chaperones, including a series of proteins termed heat shock proteins (HSPs) that are induced by conditions that damage cytosolic proteins. Protein folding associated with the ER occurs simultaneously with contranslational translocation and is assisted by various ER resident chaperones, including GRP78/BiP, GRP94, protein disulfide isomerase, calnexin, and others. The signal peptide is normally removed by an endopeptidase. Some proteins are retained in association with the ER membrane, frequently looped back and forth between the cytosolic and luminal faces of the ER in up to seven-fold membrane-spanning units. These proteins are integral to the synthesis of new membrane. A large fraction of the protein processed within the ER lumen is destined for extracellular secretion. These proteins are frequently stabilized by various types of glycosylation, disulfide bond formation, the oxidation of specific proline residues, and assembly into multimeric proteins. Disulfide bond formation is supported by an ER redox potential that is relatively more oxidizing than that of the cytoplasm. Protein processing is also supported by the high Ca²⁺ content of the organelle. Proteins that meet certain folding criteria move by vesicular transport to the Golgi for additional modification, whereas misformed proteins are targeted for degradation. Inhibitors of ER protein processing that slow protein trafficking to the Golgi result in the induction of some chaperones, most notably GRP78 and GRP94.

Regulation of Translational Initiation

The regulation of translational initiation involves the selection frequency of specific mRNA transcripts for translation and alterations in the rates of ribosomal loading on the transcripts.

Frequency of Transcript Selection

Selection frequency governs the frequency at which individual proteins are synthesized in preference to other proteins. This parameter is affected by the relative abundance of each species of mRNA transcript, which in turn is a function of rates of mRNA transcript turnover, by the recruitment of transcripts from ribonuclear particles into active translation, and by the susceptibility of the transcript to cap-proximal melting by the eIF4 complex, permitting loading of 43S preinitiation complexes. Interaction of the cap-binding protein, eIF4E, is influenced by the 5' leader structure of mRNA such that differential rates of utilization occur with different species of mRNA. The activities of various components of the eIF4 complex, including 4E and 4G, are increased by protein phosphorylation occurring in response to mitogens and hormonal growth factors. Two factors, 4E-BP1 and 4E-BP2, have been reported to exist that modify the activity of eIF4E. The dephosphorylated form of 4E-BP1 is an inhibitor of eIF4E. Phosphorylation and dissociation of 4E-BP1 occur in response to growth promoters such as insulin and epidermal growth factor, whereas certain viral infections, including polio and encephalomyocarditis virus, result in dephosphorylation of this protein. A number of protein kinase activities participate in these phosphorylations in cell-free preparations, including protein kinases A and C, S6 ribosomal protein kinase, and certain kinases for which casein serves as a substrate. Continuing efforts are in progress to identify the endogenous protein kinase(s) that catalyzes these phosphorylations. Phosphorylation of either 4E or 4G is associated with increased rates of initiation. While ribosomal loading ordinarily proceeds through mRNA cap-dependent association, it can also occur through internal loading of ribosomes. Internal ribosomal loading is thought to occur on certain mRNA with unusually long 5'-untranslated regions, such as that for GRP78, various HSPs, and some virally derived messages. eIF4E is proposed to promote internal ribosomal loading.

Alterations in Rates of Ribosomal Loading

Formation of the 43S preinitiation complex from the association of the 40S ribosomal subunit with the ternary complex of eIF2, GTP, and met-tRNA, provides a second major regulatory point in translational

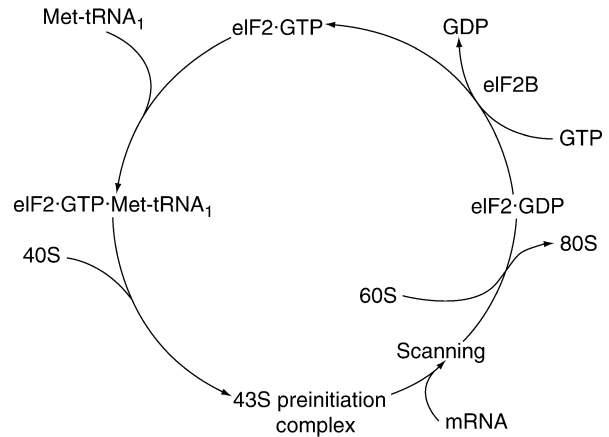


Figure 1 The eIF2 cycle.

initiation (Figure 1). The 43S complex subsequently associates with mRNA, scans to the initiator AUG, and binds the 60S ribosomal subunit to form a monosome capable of translational elongation. At this point the eIF2-associated GTP is converted to GDP, and the factor dissociates from the ribosome. Because the binary eIF2-GDP complex does not support the formation of the ternary complex with met-tRNA_i, a subsequent round of initiation requires the regeneration of eIF2-GTP via the catalytic exchange of GDP for GTP. This exchange is accomplished by eIF2B, a factor typically present at low stoichiometric ratios with respect to eIF2. The affinity of eIF2B for eIF2 is increased greatly by the phosphorylation of Ser-51 of the α subunit of eIF-2 by specific eIF2 kinase activities described later. Increased phosphorylation of eIF2 α is recognized to mediate the translational repression occurring in mammalian cells subjected to a wide variety of physical, chemical, and nutritional stresses. A 20–30% increase in the phosphorylation of eIF2 α is frequently adequate for the sequestration of eIF2B into an inactive complex such that recycling of eIF2 is largely inhibited. Cells, however, tend to vary somewhat in their relative contents of eIF2B and eIF2 and in the degree of eIF2 α phosphorylation required for translational suppression. There is reason to suspect that eIF2B may be subject to additional regulatory inputs. For example, the catalytic activity of the factor is thought to be increased by an insulin-mediated phosphorylation.

Phosphorylation of eIF2 at Ser-51 is catalyzed by four well-characterized protein kinases termed GCN2, PKR, PERK, and HRI. These enzymes comprise a distinct class of protein kinases that share significant sequence homology in their catalytic domains. GCN2 appears to be expressed selectively in yeast and is activated by amino acid starvation. PKR is distributed broadly among mammalian cell

types, inducible by interferon, and activated by low concentrations of double-stranded RNA and by certain viral RNAs. PKR is thought, therefore, to mediate the antiviral effects of interferon. As described later, PKR appears to phosphorylate eIF2 during cytoplasmic stress. PERK functions as an ER membrane-spanning eIF2 kinase that mediates eIF2 phosphorylation during ER stress. The cytoplasmic domain of PERK possesses catalytic activity whereas its luminal domain is thought to sense ER stress. HRI, the heme-regulated inhibitor kinase, is activated by heme deficiency and is expressed predominantly by erythroid cells. Both PKR and HRI undergo an autophosphorylation and dimerization during activation. While less well characterized, PERK probably undergoes similar modifications during activation. PKR, which resides on 60S ribosomes bound to the cytosolic face of the ER, may be capable of cross-dimerizing with PERK.

Most mRNA actively engaged in translation load multiple ribosomes and therefore sediment as polyribosomal assemblies during density gradient centrifugation. The rate of ribosomal loading in conjunction with the rate of peptide bond formation determines the overall rate of translation and the polyribosomal content of cells. Polyribosomal content and size tend to increase whenever the rate of peptide chain elongation is rate limiting with respect to initiation. While some additional recruitment of mRNA from ribonuclear particles into polyribosomes may occur through the phosphorylation and activation of the eIF4 complex in response to mitogens or growth factors, it is unclear that sufficient initiation factors and ribosomal subunits are generally available on a short-term basis to support large increases in initiation through this mechanism. Growth factors also have the potential to alter peptide chain elongation rates through the phosphorylation and activation of eEF1 and, perhaps, through the phosphorylation of ribosomal protein S6 or of various aminoacyl-tRNA synthases. The overall promotion of translation by these modifications as measured by amino acid incorporation ranges on the order of 30–100% depending on cell type over 1–2 h. In contrast, the phosphorylation of eIF2 α in response to various conditions producing cellular stress inhibits initiation by approximately 85–95% within a few minutes in conjunction with reductions in polyribosomal contents. Under these conditions, provision of the 43S preinitiation complex, rather than recruitment of mRNA transcripts through cap melting considerations, becomes rate-limiting on initiation. Suppression of 43S complex formation appears to favor internal ribosomal loading on mRNA transcripts with long 5' leaders typical of various HSPs and of GRP78. Overall, the eIF2 input into the

regulation of translational initiation provides an adjustable braking system that overrides more slowly imposed stimulatory controls. Potential regulation emanating through protein phosphatase activities for the various initiation factors remains largely unexplored.

Cellular Stress and the Inhibition of mRNA Translation

Conditions that either damage intracellular proteins or produce an accumulation of protein-folding intermediates induce cellular stress responses, as defined by the increased expression of one or more HSP or GRP protein chaperones. Protein synthesis is suppressed at translational initiation in a manner that coincides with increased phosphorylation of eIF2 α . Alterations in the content or the degree of phosphorylation of the eIF4 complex have also been observed following varying degrees of cellular damage. Current evidence favors a model in which the rapid sequestration of chaperones by misfolded proteins accumulating during stress signals the activation of an eIF2 kinase (see [Figure 2](#)). Although the subcellular location and degree of protein misfolding determine whether ER resident (GRP) or cytoplasmic (HSP) chaperones are eventually induced, significant misfolding in either subcellular site usually signals increased eIF2 phosphorylation and translational suppression within minutes.

Interruption of ER Protein Processing

ER stress results from an accumulation of under-processed proteins within the organelle that are incompetent of ER to Golgi transport. Accumulation is associated with a variety of treatments that inhibit various processing steps and with genetic variations that result in the expression of unprocessable proteins. Depletion of ER-sequestered Ca²⁺ in response to extracellular Ca²⁺ chelators, Ca²⁺ ionophores, or inhibitors of ER Ca²⁺ uptake interrupts the processing of sugar residues on newly synthesized glycoproteins. Ca²⁺ depletion may also impair the folding of various proteins and the assembly of multimeric proteins in the ER. Mild reducing agents impede the formation of critical disulfide bonds within proteins by shifting the luminal redox potential toward a less oxidizing state such that folding intermediates accumulate. Concentrations of Ca²⁺ depleting agents or reducing agents that are sufficient to inhibit processing do not lower cellular ATP; removal of the agent by washing results in almost immediate restoration of processing and amino acid incorporation. ER stress is also precipitated by the expression of genetically altered unprocessable proteins and by the unbalanced (nonstoichiometric) production of the subunits of

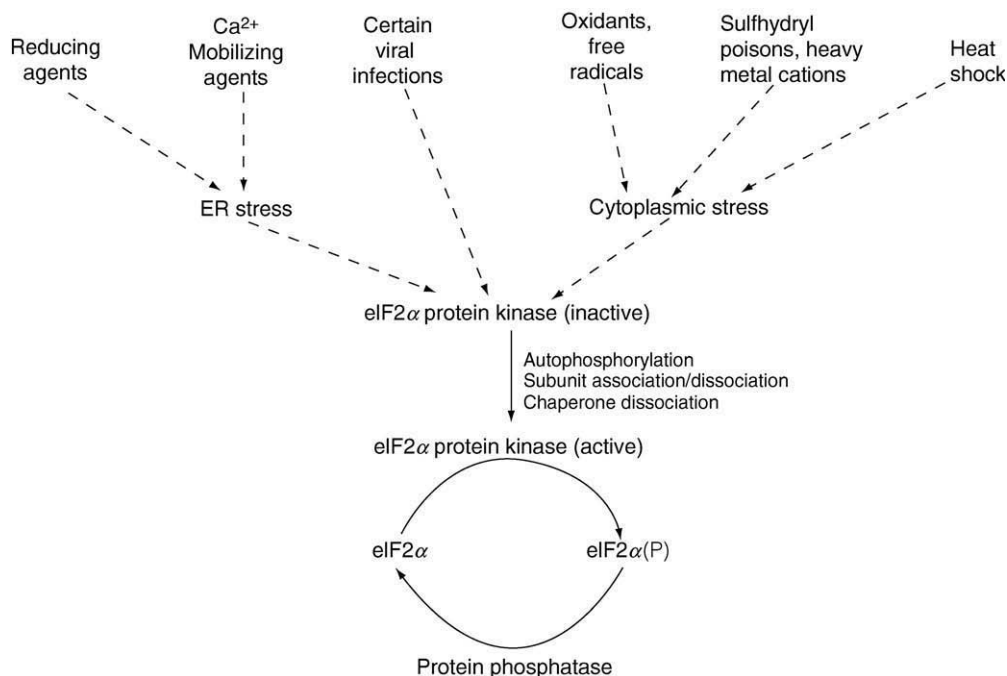


Figure 2 Activation of eIF2 α protein kinase by stress.

multimeric proteins. Tunicamycin, which inhibits asparagine-linked protein glycosylation, and brefeldin A, which produces retrograde coalescence of the Golgi with the ER, result in slowly developing ER stress. These various disturbances of ER function inhibit translational initiation coincident with the phosphorylation of eIF2. Phosphorylation has been ascribed to the activation of PKR and PERK.

A stress response characterized by the increased synthesis of new GRP78 and GRP94 molecules is routinely observed on continued (hours) ER stress. The sequestration of ER chaperones by the under-processed proteins retained within the organelle is thought to initiate the signaling for both translational suppression through eIF2 kinase activation and the induction of new chaperones, albeit by different pathways. The induction of *grp78* mRNA is dependent on the activation of an ER membrane-spanning protein kinase termed Ire1. The ER luminal domains of PERK and Ire1 possess high homology. It is likely that both Ire1 and PERK are inhibited by interactions with GRP78.

Protein Misfolding in the Cytoplasm

Mammalian cells also undergo an acute suppression of mRNA translation during moderate elevations of their optimal ambient temperature or challenge with certain chemical agents. These treatments, which have been termed proteotoxic, promote mis-folding or damage of cytoplasmic proteins. Proteotoxic

agents include sulfhydryl poisons such as sodium arsenite and iodoacetamide, heavy metal cations, and chemical oxidants and free radical generators. The inhibition of mRNA translation in response to most proteotoxic conditions is accompanied by the binding of HSP chaperones to misfolded or damaged proteins, eIF2 phosphorylation, and slowing of translational initiation. Following the initial insult, cells typically express abundant quantities of HSPs of 28–30, 60, 70–72, 90, and 110 kDa mass over the next several hours. These inductions collectively are frequently termed the heat shock response and are accompanied by dephosphorylation of eIF2 and resumption of translation.

The phosphorylation of eIF2 in response to various types of cytoplasmic stress in increased to variable degrees ranging from pronounced phosphorylations with sodium arsenite, low to high phosphorylations with heat stress, and marginal phosphorylations with iodoacetamide. These variations relate to the difficulties in achieving good experimental reproducibility, utilizing conditions that damage cells in a relatively indiscriminate fashion. More stringent conditions that generate relatively complete translational inhibition tend to reduce cell viability, whereas a less rigorous challenge produces only partial suppression. HSP-inducing chemicals observed to promote a rapid phosphorylation of eIF2 include sodium arsenite, Cd^{2+} , Hg^{2+} , *t*-butylhydroperoxide, menadione, and diamide. Arsenite has proven particularly valuable for investigating the regulation of protein synthesis

during cytoplasmic stress as it does not produce ER stress or lower ATP. Reproducible, strong stress responses are produced by the drug as are good recoveries of translational activity during periods of increased HSP expression. While arsenite inactivates many proteins with sulfhydryl groups, inactivation is repaired more readily than that produced by high temperatures, heavy metals, or strong oxidants. Arsenite suppresses translational initiation to degrees that correlate closely with the activation of PKR and the phosphorylation of eIF2. Peptide chain elongation is not affected, energy stores are not depleted, and viability is maintained.

The cytoplasmic chaperone HSC70 interacts with nascent polypeptides in nonstressed cells, most probably to prevent premature folding. In stressed preparations, nascent polypeptides unable to fold properly and mature proteins that experience damaged folding become stably bound to HSC70. Sequestration of this chaperone during cytoplasmic stress, therefore, has been proposed to signal the heat shock response. Inactivation of PKR has been reported to involve interaction of HSP70 with P58, a cellular inhibitor of the kinase. Various HSPs also function prominently in steroid hormone delivery to the nucleus, the regulation of chaperone-dependent protein kinase activities, clathrin uncoating of coated vesicles, and the transport of protein into the mitochondria.

Nutritional Stress and Stress-Related Apoptosis

Amino acid starvation is associated with slowed rates of translational initiation and, at least in yeast, increased eIF2 phosphorylation. The mechanisms signaling phosphorylation under these conditions are unknown, although the depletion of active chaperones may be involved. For example, GRP78 of cultured cells is known to be inactivated by ADP-ribosylation after nutrients in growth media are expended and protein synthetic rates are slowed, eIF2 phosphorylation and translational suppression, attributed to the activation of PKR, have been observed in cells undergoing programmed cell death in response to inflammatory cytokines, growth factor deprivation, and ultraviolet light. The extent to which proteins misfold and stress proteins are induced under any of the just-described conditions is unclear. However, glucose deprivation promotes expression of ER chaperones in cultured cells, presumably as a consequence of decreased protein glycosylation.

Stress and the Activation of HRI

HRI of reticulocyte lysates is activated in response to heat shock, oxidative stress, and addition of

denatured proteins. In lysates the kinase is known to associate with multiple stress chaperones, including HSP90 and HSC70. HSP90 protects HRI during its synthesis and stabilizes the mature kinase such that it can be autophosphorylated and converted to an activatable form without undergoing stress-dependent denaturation. Because HSP90 dissociates on conversion of the kinase to the activatable form, the chaperone does not function in the process by which stress promotes activation. HSC70 also interacts with nascent and mature competent HRI. Unlike HSP90, however, HSC70 remains associated with the fully activatable kinase. Dissociation of this chaperone fosters activation whereas association suppresses activation. In the intact stressed erythroid cell, therefore, sequestration of HSC70 by misfolded proteins may serve to signal activation of HRI, resulting in eIF2 phosphorylation and slowing of protein synthesis. Because development of erythroid stem cells into reticulocytes involves the ejection of the nucleus in conjunction with loss of functional ER, increased expression of ER or cytoplasmic chaperones is not observed in the more differentiated red cell during stress.

The central role of eIF2 kinase in the translational suppression that transpires during viral infection, ER stress, or cytoplasmic stress is modeled in [Figure 2](#). Using this model, misfolded proteins accumulating during stress compete with the kinase for chaperone binding. Dissociation of chaperones, in combination with subunit rearrangements, promotes autophosphorylation and activation of the kinase. eIF2 α is phosphorylated and initiation is slowed. Reversibility of translational inhibition requires reassociation of chaperones to repress kinase activity and a protein phosphatase to dephosphorylate eIF2 α .

Comparable Roles of the Endoplasmic Reticulum and Cytoplasmic Stress Responses in Recovery of Translational Activity

Separate, but closely related, stress response systems exist for the ER, regarding the induction of the GRPs, and for the cytoplasm, pertaining to the induction of the HSPs. As discussed earlier, mRNAs for GRPs and HSPs possess unusual structural features at their 5' ends thought to favor their selective translation during stress. GRP78 and GRP94, the most prominently induced ER stress chaperones, share significant sequence homology with HSP70 and HSP90, respectively. GRP78 is hypothesized to function in early ER protein folding and assembly, in the translocation of proteins from the cytosol to the ER for processing,

and in the retention and degradation of improperly folded proteins. Comparable chaperone activity is exhibited by HSP70 and a closely related form of the protein expressed in nonstressed cells, HSC70. Both prevent incorrect folding of polypeptides during synthesis, permitting delivery to organelles in an unfolded state for translocation. These chaperones may also solubilize or refold denatured or aberrant proteins and/or deliver them to a degradative system. Induction of GRPs or HSPs appears necessary for survival during persistent ER or cytoplasmic stress, respectively.

The induction of either GRPs or HSPs in stressed cells is accompanied by a reduced phosphorylation of eIF2 and a partial resumption of mRNA translation. Induction of either class of stress proteins is sufficient for the development of translational tolerance to subsequent rechallenge of the cells with either ER or cytoplasmic stressors. Cross-tolerance is observed in terms of continued amino acid incorporation into proteins, maintenance of polyribosomal contents, and the lack of increased eIF2 phosphorylation. Translational recovery from inhibition by Ca^{2+} -releasing drugs is partly overturned by antisense oligonucleotides directed against *grp78* mRNA, whereas

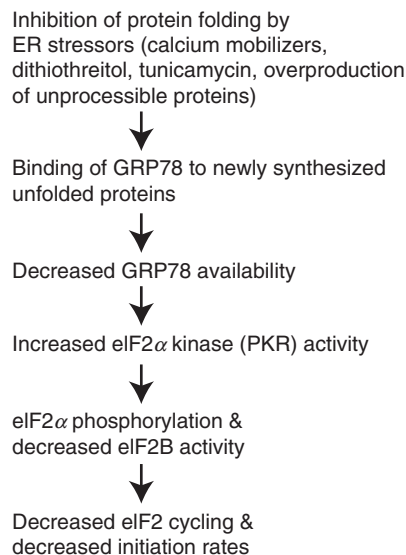
overexpression of GRP78 is sufficient for tolerance to translational inhibition and eIF2 phosphorylation in response to ER stress. Presumably various GRP and HSP chaperones inhibit eIF2 kinase(s) through complexing with a critical component(s) of the enzymes or with another protein(s), which affects activity. If so, the sequence homology of GRP78 and HSP70/HSC70 may be important to the putative dual input.

By regulating eIF2 kinase activity, the two stress chaperone systems permit rates of protein folding or processing to be coordinated with rates of protein synthesis through changes in the rate of recycling of eIF2 and of ribosomal loading onto mRNA. The interactions of GRP78 in regulating translational initiation are modeled in Figure 3.

With the exception of their mutual abilities to influence eIF2 phosphorylation, the two stress systems appear to function independently. For example, sodium arsenite does not induce the GRPs nor does it affect Ca^{2+} retention by the ER. Following induction of GRP78 with Ca^{2+} -mobilizing agents, cells remain responsive to HSP induction by subsequent arsenite treatment. Similarly, cells expressing the cytoplasmic stress response remain susceptible to induction of the GRPs. It should be recognized that

Back Regulation
acute (min)
decreased available GRP78
decreased ribosomal loading on mRNA

Events:



Forward Regulation
long term (h)
induction of GRP78
increased ribosomal loading on mRNA

Events:

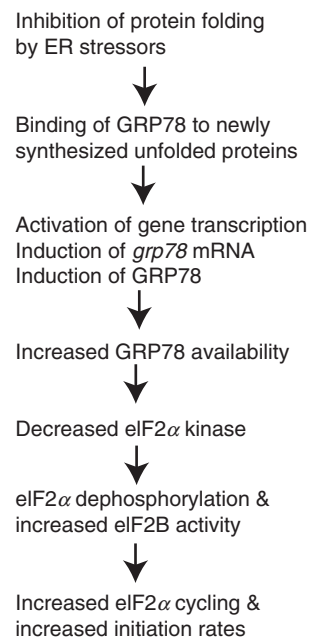


Figure 3 Coordination of rates of protein processing and protein synthesis.

the phosphorylation of eIF2 and the inhibition of translation are not mandatory for the induction of stress proteins.

Low concentrations of ER stressors clearly induce GRP78 in the absence of eIF2 phosphorylation or inhibition of translation. Greater degrees of cytoplasmic stress also appear to be required for the signaling of translational suppression than for induction of the HSPs. These findings are consistent with the existence of pathways to slow the synthesis of protein precursors only when processing or folding is seriously impaired.

eIF2 Kinase as a Focal Point for Regulating Protein Synthesis during Stress

The body of available information with nonerythroid mammalian cells suggests that PERK and PKR function together in governing rates of translational initiation in response to various stimuli including, but not limited to, ER and cytoplasmic stresses. HRI, a heme-regulated protein kinase, is unlikely to mediate translational suppression in stressed nonerythroid cells. Because the induction of either HSPs or GRP78 provides cross-tolerance to either ER or cytoplasmic stressors, it seems clear that PERK and PKR interact in some manner, perhaps through cross-dimerization of their respective kinase domains during activation. It is, of course, quite possible that additional homologous forms of eIF2 kinase remain to be discovered. Putative additional stimuli for eIF2 kinase activation could include hormonal or nutritional alterations or treatments that damage cell membranes.

See Also the Following Articles

Chaperone Proteins and Chaperonopathies; Chaperonopathies; Viral Virulence and Stress.

Further Reading

- Brostrom, C. O. and Brostrom, M. A. (1998). Regulation of translational initiation during cellular responses to stress. *Progress in Nucleic Acid Research and Molecular Biology* 58, 79–125.
- Brostrom, M. A. and Brostrom, C. O. (2003). Calcium dynamics and endoplasmic reticular function in the regulation of protein synthesis: implications for cell growth and adaptability. *Cell Calcium* 34, 345–363.
- Clemens, M. J. (1996). Protein kinases that phosphorylate eIF2 and eIF2B, and their role in eukaryotic cell

- translational control. In: Hershey, J. W. B., Mathews, M. B. & Sonenberg, N. (eds.) *Translational control*, pp. 139–172. Plainview, NY: Cold Spring Harbor Laboratory Press.
- Clemens, M. J. and Elia, A. (1997). The double-stranded RNA-dependent protein kinase PKR: Structure and function. *Journal of Interferon and Cytokine Research* 17, 503–524.
- De Gracia, D. J. and Montie, H. L. (2004). Cerebral ischemia and the unfolded protein response. *Journal of Neurochemistry* 91, 1–8.
- de Haro, C., Mendez, R. and Santoyo, J. (1996). The eIF-2 α kinases and the control of protein synthesis. *FASEB Journal* 10, 1378–1387.
- Duncan, R. F. (1996). Translational control during heat shock. In: Hershey, J. W. B., Mathews, M. B. & Sonenberg, N. (eds.) *Translational control*, pp. 271–293. Plainview, NY: Cold Spring Harbor Laboratory Press.
- Gimenez-Barcons, M., Wang, C., Chen, M., et al. (2005). The oncogenic potential of hepatitis C virus NSSA sequence variants is associated with PKR regulation. *Journal of Interferon and Cytokine Research* 25(3), 152–164.
- Harding, H. P., Zhang, Y. and Ron, D. (1999). Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. *Nature* 397, 271–274.
- Hershey, J. W. B. (1991). Translational control in mammalian cells. *Annual Review of Biochemistry* 60, 717–755.
- Melville, M. W., Tan, S., Wambach, M., et al. (1999). The cellular inhibitor of the PKR protein kinase, P58^{IPK}, is an influenza virus-activated co-chaperone that modulates heat shock protein 70 activity. *Journal of Biological Chemistry* 274, 3797–3803.
- Pain, V. M. (1994). Translational control during amino acid starvation. *Biochimie* 76, 718–728.
- Pain, V. M. (1996). Initiation of protein synthesis in eukaryotic cells. *European Journal of Biochemistry* 236, 747–771.
- Panniers, R. (1994). Translational control during heat shock. *Biochimie* 76, 737–747.
- Pestova, T. V., Borukhov, S. I. and Hellen, U. T. (1998). Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. *Nature* 394, 854–859.
- Proud, C. G. (1992). Protein phosphorylation in translational control. *Current Topics in Cell Regulation* 32, 243–369.
- Welch, W. J. (1993). Heat shock proteins functioning as molecular chaperones: Their roles in normal and stressed cells. *Philosophical Transactions of the Royal Society of London B* 339, 327–333.
- Williams, B. R. (1997). Role of the double-stranded RNA-activated protein kinase (PKR) in cell regulation. *Biochemistry Society Transactions* 25, 509–513.
- Wiseman, R. L. and Balch, W. E. (2005). A new pharmacology—drugging stressed folding pathways. *Trends in Molecular Medicine* 11, 347–350.

Protein Translocation See: Protein Synthesis; Chaperone Proteins and Chaperonopathies; Heat Resistance; Heat Shock Response, Overview; Chaperonopathies; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles.

Proteosome

M Maldonado

University of Pennsylvania, Philadelphia, PA, USA

J Wang

Wyeth Research, Collegeville, PA, USA

© 2007 Elsevier Inc. All rights reserved.

The Ubiquitin-Proteosome System

The Ubiquitin-Proteosome System in the Pathogenesis of Disease

Conclusion

Glossary

<i>26S Proteosome</i>	Large protease complex composed of a 20S proteolytic-containing complex capped by two 19S regulatory particles responsible for degradation of polyubiquitinated proteins.
<i>E3 ubiquitin ligase</i>	Diverse group of enzymes with distinct motifs for recognizing specific substrates that carry out the final transfer of ubiquitin to the bound substrate.
<i>N-end rule</i>	Relates the N-terminal residue to the half-life of an intracellular protein.
<i>Ubiquitin-proteosome system</i>	Major proteolytic pathway for intracellular protein degradation involved in many key cellular processes.
<i>Ubiquitination</i>	Enzymatic process that sequentially covalently links ubiquitin to targeted protein substrates to form a polyubiquitin chain recognized by the 26S proteosome.
<i>Unfolded protein response</i>	Highly specific signaling pathway in the endoplasmic reticulum (ER) that recognizes the accumulation of un- or misfolded proteins caused by disruption of ER homeostasis.

Understanding cellular homeostasis requires knowledge of both synthetic and degradation pathways. Protein degradation, in particular, plays a key regulatory process for many cellular processes. There are two systems involved in intracellular protein

degradation. The lysosomal pathway was first identified, but in the late 1950s, Goldberg and colleagues demonstrated a new type of proteolytic system in lysosome-free reticulocyte cell extracts. This led to the discovery of ubiquitin and the proteosome and the elucidation of the ubiquitin-proteosome system (UPS). The UPS degrades the vast majority (up to 80–90%) of long- and short-lived normal and abnormal intracellular proteins, hence is considered to be the major pathway for intracellular protein degradation. A growing body of evidence shows that this is a highly regulated yet complex system that is central to normal homeostasis, including cell cycle regulation, DNA repair, regulation of immune and inflammatory responses, and cellular response to stress. Proteins targeted by the UPS include mutated and misfolded proteins, cyclins, and tumor suppressors. Derangements of the UPS can lead to many disorders, including malignancies, neurodegenerative diseases, and possible systemic autoimmunity. A better understanding of the UPS process and identification of the components involved in the degradation of key regulatory proteins have led to the development of mechanism-based therapeutics in various diseases.

The Ubiquitin-Proteosome System

Ubiquitin and Ubiquitination

The control of protein turnover must logically, but not necessarily intuitively, be a counterpart to protein synthesis. This dynamic balance, revealed by variability in the degradation rates and half-life of different proteins, relies on the UPS. The proteolysis is energy dependent and specific. Proteins are selectively targeted for degradation by covalent linkage to ubiquitin, a 76-residue protein that is highly evolutionarily conserved in all eukaryotes. The name ubiquitin is a misnomer, having arisen from the original belief that it was found in all cellular organisms. The linkage of ubiquitin to target protein is followed by sequential addition of more ubiquitin moieties to each other to form a polyubiquitin chain that functions as the

recognition signal for a downstream proteasome in the UPS. Poly- or monoubiquitination can be involved in a variety of nonproteolytic functions, such as activating enzymes, modification of histones that regulate transcription, modulation of membrane dynamics, or routing of the tagged proteins to their subcellular compartments. Ubiquitin-binding proteins that have ubiquitin-binding domains recognize some ubiquitinated proteins. Modification of target protein by ubiquitin or a ubiquitin-like protein remodels the surface of target proteins, affecting their stability, interactions with other proteins, activity, and subcellular localization. At least 10 different modifiers have been described in mammalian cells, and conjugation of each modifier to its target has different biological effect.

Ubiquitination Reaction

Ubiquitination of cellular proteins is a highly complex, temporally controlled, and tightly regulated energy-dependent process. The conjugation of ubiquitin is carried out in a cascade reaction involving three enzymes designated E1, E2, and E3 (Figure 1). Ubiquitination begins with the attachment of ubiquitin by E1, also known as the ubiquitin-activating enzyme. Ubiquitin is then transferred to E2, or the ubiquitin-conjugating enzyme. The final step is dependent on one of many E3s (several hundred forms have been identified in the human genome) that are

responsible for catalyzing the addition of ubiquitin to the protein substrate. Chain elongation, or polyubiquitination, can be carried out by the same E3 or by a different ligase, which can be referred to as E4.

Substrate Targeting

The UPS is highly specific and selective. Much of this lies in the diversity of different ubiquitin-protein ligase E3s that can recognize a specific substrate. Many substrates are recognized only after modifications such as misfolding, denaturing, hydroxylation, or phosphorylation. Molecular chaperones such as heat shock protein 70 or chaperon-like molecules such as valosin-containing protein can also act as recognition elements. E3 ligases recognize specific targeting motifs, which can vary from a single amino acid residue to a domain not usually exposed. One of the most important recognition patterns is a destabilizing N-terminal amino acid. A sequence such as arginine and lysine can be a unique feature of a protein, which relates the half-life of an intracellular protein to its N-terminal residue, which has become known as the N-end rule. Phosphorylation of either the substrate or the ligase can also control targeting. Ubiquitin-like proteins can also play a regulatory role by protecting proteins from ubiquitination.

In addition to its substrates, the activity of the ubiquitin system itself can be modulated by a number

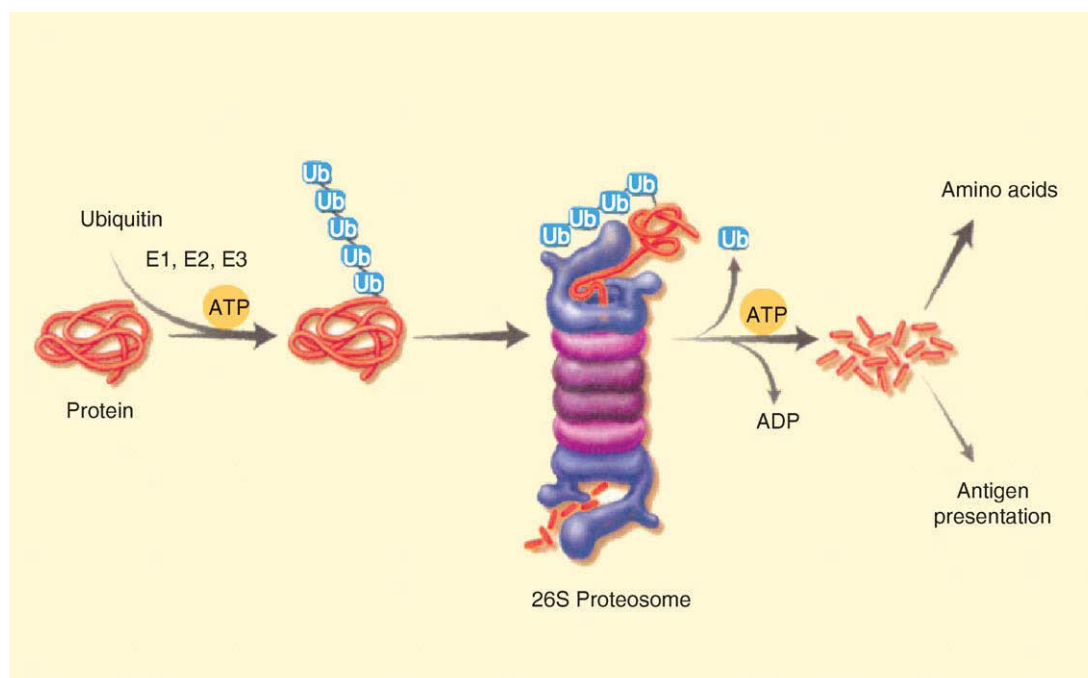


Figure 1 Schematic of the ubiquitin-proteasome system. Ubiquitin is added sequentially to targeted substrates in an energy-dependent enzymatic cascade requiring three enzymes, E1, E2 and E3. Polyubiquitinated proteins are recognized by the 26S proteasome. Cleavage produces small peptides and free ubiquitin. Adapted from Kisselev and Goldberg (2001) with permission from Elsevier.

of factors, including thyroid hormones, glucocorticoid steroids, cytokines, and proteins expressed in malignant cells such as proteolysis-inducing factor (PIF). Interestingly, some factors such as interferon gamma (IFN- γ) not only regulate the modification of substrate, including the I κ B family of proteins, but also modify the components of the enzymatic machinery of the ubiquitin system and the proteasome complex. Similarly, protein kinase C and the tyrosine kinase pathway are involved in both the modification of substrates of UPS and phosphorylation of E1 and E2, thereby exerting their activities with high efficiency.

The Proteasome

The ubiquitinated substrates targeted for proteolysis are recognized by a 26S complex cellular structure, the proteasome (Figure 2). This high-molecular-mass protease is also ATP dependent. Degradation of proteins by proteasomes removes denatured, damaged, or improperly translated proteins and regulates the level of proteins in the cells. The 26S proteasome is composed of two subcomplexes. One or two 19S regulatory particles (RPs) cap either or both ends of the barrel-shaped 20S complex, which carries the proteolytic activity, forming 26S and 30S proteasomes, respectively. The 19S RP is responsible for recognizing the ubiquitinated proteins as proteasome substrates, the opening of the orifice in the 20S

complex that allows entry onto the proteolytic chamber and likely the unfolding of the substrate necessary for entry. Within the chamber, proteins are broken down into short peptides. Without the 19S RP, the 20S complex is extremely inefficient in degrading folded protein substrates. The 20S complex is composed of four stacked rings, two identical outer α rings and two inner β rings. Each ring has seven distinct subunits. The β rings contains multiple catalytic sites including chymotrypsin-like sites that are the target of almost all the synthetic and natural inhibitors of the proteasome. The outer two α rings have no known function. Other subcomplexes can cap the 20S complex, creating a polar arrangement that can affect the product generated.

Degradation Products

After the protein has been degraded, the 19S RP removes the polyubiquitin tag from the substrate protein. Degradation in the proteasome generates short peptides that can have different fates. Some are further degraded by cytosolic peptidases. Some can be transported into the endoplasmic reticulum (ER) for binding to MHC class I molecules and cell surface expression. This latter subset of antigenic peptides may arise from proteasome activator PA28 subcomplex capped polar proteasomes that generate peptides trimmed to properly fit the MHC class I binding grooves. In this regard, hybrid proteasomes (19S

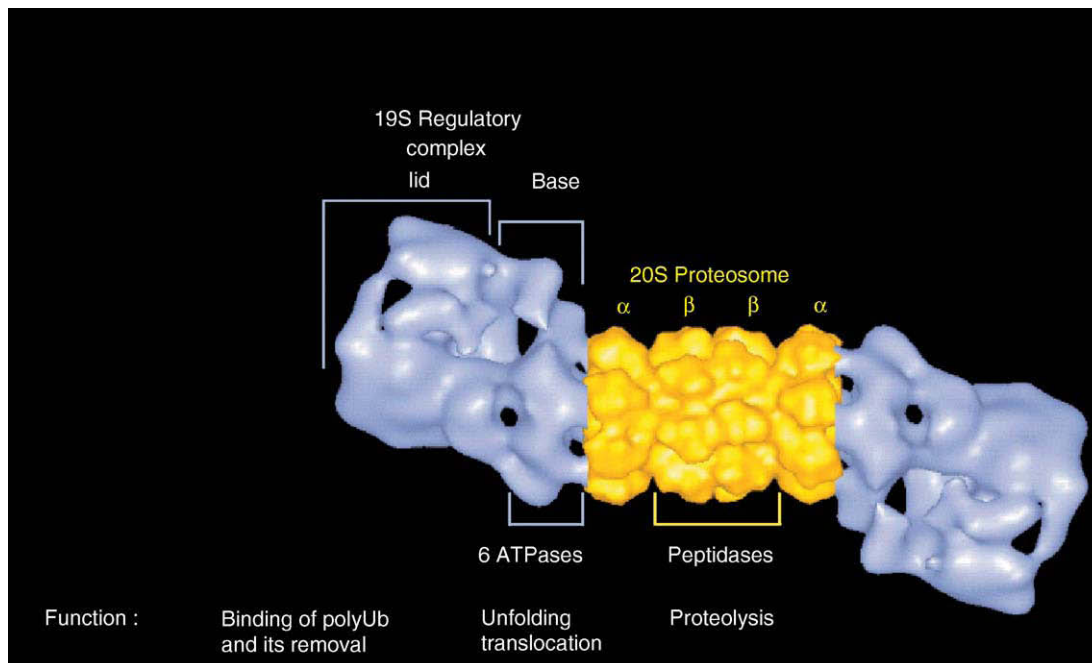


Figure 2 Electron tomographic image and structure of the 26S proteasome microscope. The 26S proteasome is composed of the barrel-shaped 20S complex capped by two 19S subcomplexes. The 19S regulatory particle recognizes the polyubiquitin tag on targeted substrates and unfolds the substrate to enter the proteolytic chamber. The 20S core particle contains the catalytic sites responsible for the proteolysis. Adapted from Kisselev and Goldberg (2001) with permission from Elsevier.

RP-20S proteasome-PA28) may have unique value during immune responses. Peptides are not always the final product of the UPS, as longer, biologically active proteins can be generated from precursors. Free, reusable ubiquitin is also generated. Deubiquitylating enzymes catalyze this reaction.

The Ubiquitin-Proteasome System in the Pathogenesis of Disease

The Role of the UPS in Stress and Disease

Regulation of protein turnover has an important role in various cellular functions, and alterations can lead to many diseases. Viruses have evolved proteins that can exploit this system to their advantage by either enhancing or avoiding ubiquitination. The unfolded protein response (UPR) is a highly specific signaling pathway in the ER that recognizes the accumulation of un- or misfolded proteins caused by disruption of ER homeostasis. Activating the UPR has multiple effects, including enhancing degradation of the luminal content of the ER via the UPS, and has been implicated in numerous disease processes.

Malignancies

Protein degradation via the ubiquitin-proteasome system is an important regulator of cell cycle progression and survival through modulating the function of nuclear factor- κ B (NF- κ B) and pro-apoptotic transcription factor p53. Thus, inhibition of this pathway can be directly cytotoxic and can sensitize several tumor cell types to cytotoxic chemotherapy and radiation. Most of the experience with proteasome inhibition has been gained through work done in this disease area. Preclinical studies have demonstrated that the proteasome inhibitor bortezomib decreases proliferation, induces apoptosis, enhances the activity of chemotherapy and radiation, and reverses chemoresistance in a variety of hematologic and solid malignancy models *in vitro* and *in vivo*. Bortezomib was the first proteasome inhibitor to enter clinical trials. In two phase II trials, bortezomib alone or in combination with dexamethasone elicited durable responses with meaningful survival benefits in patients with recurrent and/or refractory multiple myeloma. In a phase III trial, bortezomib produced significant survival benefits and improved response rates over high-dose dexamethasone at first recurrence and beyond in patients with multiple myeloma. Current studies indicate that bortezomib may serve as induction therapy before stem cell transplantation. In addition, proteasome inhibition with bortezomib also has shown activity and manageable toxicity in mantle cell and other lymphoma.

Neurodegenerative Diseases

Dysregulation of the UPS has been implicated in the pathogenesis of both inherited and acquired neurodegenerative diseases. Accumulation of ubiquitin conjugates and/or inclusion bodies associated with ubiquitin, proteasomes, and certain disease-specific proteins has been demonstrated in a wide variety of neurodegenerative diseases. The list includes neurofibrillary tangles of Alzheimer's disease, brain stem Lewy bodies, the pathological hallmark in Parkinson's disease, Bunina bodies in amyotrophic lateral sclerosis, and nuclear inclusions in CAG repeat expansion disorders such as Huntington's disease, spinocerebellar ataxias, and spinal and bulbar muscular atrophy (Kennedy's disease). This phenomenon is thought to be due to failed attempts of the UPS to process damaged and abnormal proteins, or movement of the proteasome to specific subcellular sites where degradation of abnormal proteins may occur.

There are at least three key players that link aberration in the UPS to the pathogenesis of neurodegenerative diseases. The first is Parkin, a 52-kDa RING finger-containing E3 ligase that ubiquitinates and promotes degradation of several proteins related to the pathogenesis of Parkinson's disease. Mutations in the gene have been found in ~50% of patients with autosomal recessive juvenile Parkinsonism, one of the most common familial form of Parkinson's disease. Second is α -synuclein, a 140-amino-acid residue protein shown to be a major component of Lewy bodies and Lewy neuritis in sporadic Parkinson's disease, dementia with Lewy bodies, and the Lewy body variant of Alzheimer's disease. Third is ubiquitin C-terminal hydrolase (UCH-L1). Mutations of UCH-L1 were discovered in a German family with Parkinson's disease.

The most compelling evidence for the involvement of the UPS in the pathogenesis of Alzheimer's disease derived from the discovery that a frameshift mutation in the ubiquitin transcript resulted in a mutant form of ubiquitin called UBB⁺¹. This mutant has been observed in brains of Alzheimer's disease patients. Expression of UBB⁺¹ in the brain could inhibit the degradation of a polyubiquitinated substrate by the 26S proteasome, leading to accumulation of toxic proteins with neuropathological consequences.

Inflammation and Autoimmune Diseases

Many autoimmune diseases are characterized by immunologically mediated organ damage. Some of these autoimmune diseases may have recognizable environmental triggers, such as infections, but it is not known how or why inflammation is initiated in most and sustained in all. Key to these diseases is the same cells and molecules, such as cytokines, that

are also central to normal immunity. The UPS is now known to regulate many of these, including MHC-mediated antigen presentation, cytokine regulation, and cell cycle, and must therefore also have a role in their pathogenesis and be a logical therapeutic target.

One of the prominent examples for the involvement of the UPS in inflammation is MHC class I antigen processing by antigen-presenting cells (APCs). In addition, the UPS plays a significant role in the regulation of both T cell receptor (TCRs) and costimulatory CD28 signaling through the action of ubiquitin ligases of the Cbl family. CD28 costimulation results in the ubiquitination and degradation of Cbl-b, which eliminates the negative regulators and allows the expression of proinflammatory cytokines and their receptors. However, the most important link between the UPS and inflammation is related to NF- κ B. NF- κ B is a master regulator of many inflammatory cytokine genes, and its activation is mediated through the UPS. NF- κ B is actively inhibited when bound to I κ B. NF- κ B activation follows the degradation of I κ B, which is dependent on ubiquitination of I κ B followed by proteasomal degradation. Hence, alterations in the UPS have profound effects on immune responses, including the regulation of an array of inflammatory cytokines.

INF- γ is a key pleiotropic regulatory cytokine. It controls an inducible proteolytic cascade that provides peptides for MHC class I presentation. Sero-negative spondyloarthropathies (SpA) are a group of diseases characterized by, but not limited to, axial joint inflammatory. Ankylosing spondylitis (AS) is the prototypical SpA. Most patients with AS carry the MHC class I HLA-B27 gene, so its role in disease pathogenesis is an area of intense research. Much interest has been focused on determining the origin and nature of the peptides being presented by HLA-B27 and the cell surface expression of misfolded HLA-B27, key areas in which the UPS is known to play a role.

The UPS is involved in the regulation or induction of apoptosis. Apoptosis has been implicated in both experimental models and patients with systemic lupus erythematosus. Altered clearance of autoantigens is thought to allow for targeting by the immune system and the development of autoimmunity.

The UPS may be involved in the pathogenesis of multiple autoimmune diseases through many different mechanisms. It may therefore be a therapeutic target worth considering. Conversely, targeting the UPS may produce unwanted immunological side effects.

Cardiovascular Diseases

Inflammation, cell proliferation, and apoptosis constitute major feature of atherosclerosis. The UPS

is involved in the development of atherosclerosis through its regulation of these biological pathways. In addition, the UPS is associated with the pathogenesis of acute coronary syndromes through diverse cellular and intracellular targets and modes of actions. Recent experimental evidence indicates that the UPS is functionally active in early atherogenesis despite increased oxidative stress mimicking the initiation stage of atherosclerosis. The pathophysiological role of the UPS in the proliferative aspects of cardiovascular disease was demonstrated by the intense ubiquitin nuclear staining in intimal thickening lesions. Treatment of cultured vascular smooth muscle cells with proteasome inhibitor has led to a dose-dependent inhibition of proliferation and induction of apoptosis. In rodent models, single-dose local administration of the proteasome inhibitor MG-132 for 5 min after endothelial denudation of carotid arteries reduces neointima formation by 75%. Further, proteasome inhibition led to significant reduction in vascular atrophy in DOCA-salt hypertensive rats.

T cells have been identified as a characteristic inflammatory element of atherosclerosis, which accumulate at the rupture-prone sites of the plaque shoulder regions. The UPS is involved in the pathogenesis of atherosclerosis by regulating the activation of T cells through T cell receptors and costimulatory molecule CD28. The UPS may be involved in foam cell formation as well. In an autopsy-based study, increased expression of ubiquitin/ubiquitin conjugates was colocalized with macrophages and apoptotic cells in the area of the lipid core. It has been shown that the expression of p53 in human atherosclerotic plaques was enhanced, and overexpression of p53 resulted in a significant increase in apoptosis and decrease in extracellular matrix in the cap region. Accumulating evidence indicates a strong association of pro-apoptotic protein p53, ubiquitin, and apoptotic neointimal cells in the cap region of unstable plaques from patients with fatal myocardial infarction.

Conclusion

The UPS is a complex system that controls many important aspects of cell function. The covalent linkage of ubiquitin to protein substrates can selectively and specifically alter their fate through proteolytic and nonproteolytic pathways. Loss of normal homeostasis through many mechanisms, including stress, can lead to aberrant cellular function and disease.

See Also the Following Articles

Autoimmunity; Chaperone Proteins and Chaperonopathies; Chaperonopathies; Heat Shock Response, Overview;

Homeostasis; Inflammation; Neurodegenerative Disorders; Parkinson's Disease; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles; Protein Synthesis.

Further Reading

- Adams, J. (2003). The proteasome: structure, function, and role in the cell. *Cancer Treatment Reviews* 29 (Supplement I), 3–9.
- Ciechanover, A. (2003). The ubiquitin proteolytic system and pathogenesis of human diseases: a novel platform for mechanism-based drug targeting. *Biochemical Society Transactions* 31, 474–481.
- Ciechanover, A. (2005). Intracellular protein degradation: from a vague idea thru the lysosome and the ubiquitin-proteasome system and onto human diseases and drug targeting. *Cell Death and Differentiation* 12, 1178–1190.
- Ciechanover, A. and Brundin, P. (2003). The ubiquitin proteasome system in neurodegenerative diseases: sometimes the chicken, sometimes the egg. *Neuron* 40, 427–446.

- Colmegna, I., Sainz, B. Jr., Garry, R. F. and Espinoza, L. R. (2005). The proteasome and its implications in rheumatology. *Journal of Rheumatology* 32, 1192–1198.
- Herrmann, J., Ciechanover, A., Lerman, L. O. and Lerman, A. (2004). The ubiquitin-proteasome system in cardiovascular diseases – a hypothesis extended. *Cardiovascular Research* 61, 11–21.
- Kisselev, A. L. and Goldberg, A. L. (2001). Proteasome inhibitors: from research tools to drug candidates. *Chemistry & Biology* 8, 739–758.
- Richardson, P. G., Mitsiades, C., Hideshima, T. and Anderson, K. C. (2006). Bortezomib: proteasome inhibition as an effective anticancer therapy. *Annual Review of Medicine* 57, 33–47.
- Welchman, R. L., Gordon, C. and Mayer, R. J. (2005). Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Nature Reviews Molecular Cellular Biology* 6, 599–609.
- Zhang, K. and Kaufman, R. J. (2006). The unfolded protein response: a stress signaling pathway critical for health and disease. *Neurology* 66(Supplement 1), S102–109.

Psoriasis

A B- Kirschbaum

Technical University of Dresden, Dresden, Germany

© 2007 Elsevier Inc. All rights reserved.

Immunopathology of Psoriasis
Stress – a Triggering Factor in Psoriasis
Psychobiology of Psoriasis
Summary

Glossary

<i>Keratinocytes</i>	Major cell type of the epidermis making up to 90% of epidermal cells. Keratinocytes are shed and replaced continuously from the stratum corneum. The time of transit from basal layer to shedding is approximately one month although this can be accelerated in conditions of keratinocyte hyperproliferation, i.e., psoriasis.
<i>Phototherapy</i>	Light therapy using specific wavelength of light for a distinct period of time. Generally, lasers, fluorescent light, or full-spectrum light are used.
<i>Superantigens</i>	Molecules that stimulate indiscriminately T-cells. The best characterized superantigens are the microbial toxins from <i>Staphylococcus aureus</i> and <i>Streptococcus pyogenes</i> .

Immunopathology of Psoriasis

Psoriasis (PSO) is a chronic, genetically determined inflammatory skin disease with an estimated prevalence of 2% of the northern population. Apart from allergic contact dermatitis and atopic dermatitis, PSO is the most common chronic skin disease. PSO can start at any age; however, disease onset is most frequently observed in the second decade of life. PSO symptomatology is characterized by erythematous and solid skin plaques with sharply defined borders, covered with fine silvery scales. Although not life threatening, PSO can be a disabling disease with a considerable impact on patients' psychological and social well-being. Psoriatics often feel stigmatized and psychologically stressed by the disfiguring skin disease, which can lead to increased anxiety and depression.

The pathophysiology of PSO is not fully understood. However, accumulating evidence suggests that T-cell-mediated autoimmune processes and the action of pro-inflammatory cytokines cause the hyperproliferation of keratinocytes and assume the psoriatic phenotype. The antigenic peptide responsible for the initial T-cell activation in PSO has not been established yet. However, one current model is that bacterial proteins (streptococcal M proteins) resulting from a preceding infection may act as superantigens. Another important process in PSO pathogenesis

is the trafficking of the activated T cell into the skin. Activated T cells generated from psoriatic skin lesions secrete high concentrations of interleukin (IL)-2, tumor necrosis factor (TNF)- α , and interferon (IFN)- γ , indicating a significant role of T helper 1 (Th₁)-mediated inflammatory processes in the psoriatic skin. The significant role of circulating Th₁-derived cytokines is further supported by the finding that TNF- α and IFN- γ increase the infiltration of T cells and other inflammatory cells such as monocytes into the skin. Locally, IFN- γ stimulates epidermal cell proliferation and keratinocyte hyperplasia, leading to the psoriatic phenotype.

Stress – a Triggering Factor in Psoriasis

Although PSO has been conceptualized as an autoimmune disease primarily mediated by an inappropriate function of T cells, it is broadly accepted that psychosocial stress may affect PSO and may trigger the exacerbation of the disease. For example, 31% of PSO patients reported the onset of PSO in times of increased everyday life stress, and for 71%, PSO symptomatology worsened during stressful life episodes. To further evaluate the role of stress in the onset and exacerbation of PSO, psychocutaneous characteristics in PSO sufferers reporting a strong link between psychosocial stress and skin condition (high stress reactors, HSRs) and in PSO patients without any association between stress and PSO symptoms (low stress reactors, LSRs) have been investigated. Clinically, HSRs suffered from a more severe disease, with a distinct distribution of PSO plaques in emotionally charged body regions (i.e., the face, neck, forearms, hands, and genital regions), affecting the patient's physical appearance and sexuality. It has been proposed that exacerbation of the disease may lead to stronger psychological reactions and, thus, may explain the strong psychocutaneous relationship in this patient group. Psychologically, the HSRs tended to rely more on the approval of others and scored higher on depressed anger and hostility. Others have reported increased levels of anxiety, aggression, obsessionality, and depression in PSO sufferers, with the last being linked to pruritus and sleep difficulties. It is of note, however, that the personality pattern described in psoriatics can also be identified in patients with atopic dermatitis (AD), a chronic (allergic) inflammatory skin disease characterized by clinical features similar to PSO such as severe pruritus or disfiguring skin symptomatology. These observations suggest that the personality type described in psoriatics may not be a PSO-specific personality type but, rather, represents a personality pattern linked to a chronic inflammatory and disfiguring skin disease.

The importance of stress in PSO has been further highlighted by reports suggesting that psychological distress affects treatment outcome in PSO patients. In these studies, it was demonstrated that the level of stress may predict the time it took for photochemotherapy (PUVA: psoralen, P, and long-wave ultraviolet radiation, UVA) to clear PSO symptomatology. Accordingly, stress reduction by stress management, relaxation, or cognitive techniques shortened the time it took to clear PSO symptoms by PUVA and, moreover, improved the clinical severity of PSO.

Psychobiology of Psoriasis

So far we have summarized the immunological and psychological factors crucial in the onset and maintenance of PSO. Based on these findings, it remains to be determined how PSO-relevant psychological and immunological processes are linked and, moreover, by which mechanisms stress may affect skin condition in PSO.

Considerable evidence has emerged showing that the central nervous system (CNS) and the immune system are intimately connected, each regulating the other. One major pathway through which the CNS exerts control over the immune system is via the neuroendocrine system. Neuropeptides released by the brain may act directly on immune cells, or they may affect the release of other hormones from endocrine tissues that can modify immunity. Investigating neuroendocrine-immunity communication pathways in PSO patients and, moreover, studying potential aberrations in the interplay between these systems under stressful conditions may give a hint about how stress can trigger inflammatory processes in PSO and exacerbate PSO symptomatology.

In the last decade, an extensive literature has emerged demonstrating the important immunosuppressive and anti-inflammatory role of the hypothalamus-pituitary-adrenal (HPA) axis. Recently, the role of stress-induced HPA axis activation has been specified – an appropriate reactivity of the HPA axis under stress may be necessary to regulate immune function and to prevent an immune response, for example an inflammatory response, from reaching a level that may be damaging to the host. There are some data suggesting a hyporesponsive or hypoactive HPA axis in PSO patients. For example, attenuated cortisol responses to psychosocial stress or diminished urinary cortisol concentrations in psoriatics have been reported. However, other researchers failed to demonstrate reduced cortisol or adrenocorticotrophic hormone (ACTH) responses to stress in PSO sufferers. One factor that may explain the heterogeneous study results may be the acuity or chronicity of the

ongoing inflammatory process in the patient group investigated. There is evidence that normal cortisol levels can be found in chronic nonprogressive PSO characterized by clinical inactivity of chronic plaques, whereas clinically active phases with fresh eruptions were characterized by increased HPA axis activity. These data suggest that it may be useful to consider inflammation in PSO as a dynamic process with different phases of acuity and remission that may be linked to distinct endocrine response patterns.

In general, the increased activity of the sympatho-adrenomedullary (SAM) system in PSO subjects has been described. Studies using standardized psychosocial stressors, including, for example, free speech or mental arithmetic tasks in front of an audience, found significantly increased epinephrine and norepinephrine levels compared to healthy controls. Because elevated responsiveness of the SAM system with significantly increased levels of catecholamines was also found in other chronic inflammatory skin diseases (i.e., atopic dermatitis), it has been proposed that an overreactive SAM system may be a common feature of chronic inflammatory Th₁-mediated (psoriasis) and Th₂-mediated (AD) inflammatory skin disease. It is well accepted that the SAM system represents a major immunoregulatory system that controls various aspects of immunity. It is suggested that a dysfunctional SAM system may increase the risk of an aberrant immune response, especially under stressful conditions in which the system is activated. In fact, a distinct redistribution of leukocyte subsets or altered cytokine production after stress has been identified in psoriatics. In these studies, it was demonstrated that psychosocial stress yielded an increased number of monocytes, CD4⁺ (Th) cells, CD8⁺ (cytotoxic T) cells, or natural killer (NK) cells, suggesting increased mobilization of these cell types in PSO sufferers. Interestingly, the immunological changes were significantly correlated with stress-induced increases of epinephrine and norepinephrine levels in these patients. The PSO-specific immunological alterations to stress may be of clinical importance. The first leukocytic infiltrate that precedes the developing psoriatic plaque consists primarily of monocytes and CD4⁺ cells. After invasion into the skin, monocytes differentiate into macrophages and can act as antigen-presenting cells. Antigen presentation, however, is necessary for the initial activation of T cells. Monocytes are further able to produce chemokines that increase T-cell trafficking and infiltration into the skin. CD4⁺ cells play a key role in the pathogenesis of PSO, and they are considered to be mainly responsible for the initiation and maintenance of the psoriatic phenotype. For example, the injection of CD4⁺ cells into prepsoriatic skin causes the phenotypic

conversion of previous symptomless skin into psoriatic skin. Activated CD4⁺ cells (and CD8⁺ cells) further produce pro-inflammatory cytokines such as IL-2, IFN- γ , and TNF- α , which are known to play a pivotal role in PSO pathogenesis. It should be noted, however, that, although the stress-induced accumulation of PSO-relevant immunocytes in the peripheral blood of psoriatics has been successfully documented, evidence of stress-induced infiltration of these cells into the (inflamed) skin has not been assessed and, thus, remains speculative. Animal data, however, indicate that acute stress may lead to a pronounced recruitment of immunocytes to the side of inflammation and, further, to a 200–300% higher infiltration of these cells into the (inflamed) skin. In this model, stress hormones seem to be involved because adrenalectomy abolished, and the injection of corticosterone and/or epinephrine enhanced, this effect.

Summary

In this article, important immunological and psychological factors related to the pathogenesis of PSO have been described. It has become clear that immunological abnormalities, such as a dysfunctional Th₁/Th₂ immune response pattern, inappropriate cytokine secretion, and keratinocyte hyperproliferation, play a key role in the onset and maintenance of the disease. On the other hand, stress has been found to aggravate the skin condition in PSO. Regarding the stress-PSO relationship, it is important to keep in mind that stress can trigger the exacerbation of the disease but, in turn, may be a consequence of a worsening of skin condition and the discomfort related to an exacerbation of the disease. This may lead to a *circulus vitiosus* that is often difficult to interrupt.

In the present overview, potential psychobiological pathways, which may be involved in the close relationship of stress and the exacerbation of psoriatic skin inflammation, have been briefly discussed. In PSO patients, distinct aberrations in important immunoregulatory neuroendocrine systems such as the HPA axis and the SAM system may lead to an immunological milieu that may render (genetically disposed) individuals more vulnerable to the development and/or exacerbation of chronic inflammation. The dysfunctional responsiveness of (neuro)endocrine stress systems such as the HPA axis and the SAM system may be of specific pathological relevance when the organism is challenged by a stressor, and an adaptive regulatory control of the immune system is required.

It is of note, however, that, although this model is built on experimental evidence, several parts of it are still speculative, awaiting validation by future

research. Trying to integrate existing facts and promising working hypotheses into a model of psycho-neuroimmunocutaneous communication and, further, to validate such a working model in future research is a challenge for the field of psychodermatology that may lead to new disease concepts and therapeutic regimens of PSO.

See Also the Following Article

Dermatological Conditions.

Further Reading

Darsow, U. and Ring, J. (2001). Neuroimmune interactions in the skin. *Current Opinion in Allergy and Clinical Immunology* 1, 435–439.

Elenkov, I. J. and Chrousos, G. P. (1999). Stress hormones, Th1/Th2 patterns, pro/anti-inflammatory cytokines and susceptibility to disease. *Trends in Endocrinology and Metabolism* 19, 359–368.

Fortune, D. G., Richards, H. L., Kirby, B., et al. (2003). Psychological distress impairs clearance of psoriasis in patients treated with photochemotherapy. *Archives of Dermatology* 139, 458–465.

Griffith, C. E. (2003). The immunological basis of psoriasis. *Journal of the European Academy of Dermatology and Venereology* 2, 1–5.

Griffith, C. E. and Richards, H. L. (2001). Psychological influences in psoriasis. *Clinical and Experimental Dermatology* 26, 338–342.

Krueger, J. G. and Bowcock, A. (2005). Psoriasis pathophysiology: current concepts of pathogenesis. *Annals of the Rheumatic Diseases* 64, 30–36.

Picardi, A. and Abeni, D. (2001). Stressful life events and skin disease: disentangling evidence from myth. *Psychotherapy and Psychosomatics* 70, 118–136.

Psychoanalysis

P Roazen

York University, Cambridge, MA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 283–286, © 2000, Elsevier Inc.

Starting Point

A Science of the Mind

Social Philosophy

Conflictedness of Life

Challenges

Conclusion

In the late 1890s, Sigmund Freud (1856–1939), who had been trained as a Viennese neurologist, created a new field, psychoanalysis, which was designed to understand and treat neurotic afflictions. An essential key to Freud's thinking lies in the character of the last days of the Hapsburg Empire. An unusual gulf between reality and official ideology stimulated a general intellectual revolt, a search for the actualities beneath the pious formulas of proclaimed truth. This revolt was led by those ideally placed to see the discrepancy because they had nothing to gain from accepting the official view, the educated Jews. Mor-

dant irony was an essential weapon for piercing the veil of the structure of formal beliefs. The cultural conflict between East and West that had its vortex in Vienna's cosmopolitan intellectual life and the sense that liberal culture was on the verge of being undermined, were reflected throughout Freud's mature thought.

Starting Point

Freud's starting point as a therapist was the existence of inner conflicts, which interfered with the lives of suffering patients. He proposed that symptoms be looked on as substitute satisfactions, the result of a failure to deal adequately with early childhood patterns. Freud was highlighting persistent infantilism as the ultimate source of adult neurotic problems.

Freud held that neuroses are psychologically meaningful, and he interpreted them as compromise formations between repressed impulses and censoring parts of the mind. One portion of every symptom was understood as the expression of wish fulfillment, and another side represented the mental structure that reacted against the primal wish. Initially, Freud thought that neurotic anxiety arose from sexual sources; specifically, Freud indicted dammed-up sexuality as the physical basis for neurosis.

Freud conceived sexuality so broadly as to include virtually all aspects of childish pleasure-seeking. Fantasies of sexual gratification stemming from early childhood were allegedly the source of adult neurotic dilemmas. Freud proposed that a person's emotional attitude toward his or her parents encapsulated the core problems of neurosis, and he coined the term Oedipus complex to describe a boy's first childish desires for his mother and a girl's earliest affection for her father. Freud understood that someone's emotional attitude toward a family consisted of conflicting emotions involving rivalry and guilt, not just desire. Freud believed that the most troublesome feelings stemmed from emotional problems about which the individual remains unconscious.

Freud was proposing that people have motives that can be operative without their knowing anything about them. His special viewpoint was that of a psychologist, and he sought to pierce the mysteries of memory and false recollections. Freud thought that the compromise formations in constructing our image of the past were just like those in dreaming, as well as the ones underlying neurotic symptomatology and everyday slips of the tongue or pen. He thought that the past lives on in the present and that psychoanalytic treatment consisted in the exploration of each patient's early history.

Freud was ambitious as a theorist, and his notion of neurosis became part of a full-fledged system of thought. A central implication of his approach amounted to an assault on our confidence in our ability to think rationally. Freud was insisting that people are fundamentally self-deceptive. Neurosis was a form of ignorance, and Freud saw it as his task to use the power that came from enlightenment.

A Science of the Mind

Because much as Freud's work can be understood as a critique of the capacity for self-understanding, he was superlatively rationalistic about psychoanalysis itself. He thought he had discovered a science of the mind and that he had uncovered a realm of meaning that could be verified objectively. The technique of free associations, which he relied on during treatment, was one that others could be trained to use. Once patients submitted to the analytic situation, such a commitment could be used by the analyst to promote personal autonomy. Freud was relying on a structured situation in which patients saw the analyst 6 days a week for sessions lasting 50 min each for the sake of forcing people to be free, to use the terminology of Jean-Jacques Rousseau.

One of the chief defects in Freud's approach was his unwillingness to concede the philosophical

underpinnings to his approach. He was convinced that psychoanalysis was capable of transforming thought and undermining previous moral positions, yet he fancied that he had been able to do so without importing any ethical baggage of his own inside his teachings. Yet Freud was clearly expressing a morality of his own; he once explained to a patient that the moral self was the conscious, the evil self being the unconscious. Freud qualified this distinction by maintaining that his approach emphasized not only evil wishes but the moral censorship that makes them unrecognizable. He was insistent that morality was self-evident at the same time that he himself supposedly held to a higher standard of ethics than humanity as a whole.

Because Freud wrote so much about abnormality, it might seem that he would have been obliged to discuss his picture of mental health, but whatever Freud had in mind has to be teased out of his system of thought because he remained loath to deal frontally with a concept such as normality. It is clear that he did not envisage a utopian version of personal happiness; anxiety and despair were to him inevitable parts of the human condition. Freud sought not to eradicate human conflicts but to teach how we can come to terms with them.

Social Philosophy

Although Freud claimed to be intent on steering away from speculative theorizing, he repeatedly allowed himself to become engaged in social philosophy. In each of the last three decades of his life he wrote a book centering on different aspects of the psychology of religion. He made the analogy between religion and obsessional neurosis and pointed out how often outer forms have obliterated the inner religious intention, as with any other self-defeating neurotic structure. The one social coercion Freud felt to be humanly unnecessary was religion. In "The Future of an Illusion" (1927), he stressed the inner instinctual core that strains beyond culture's reach. Freud's theories emphasized the divisions within the human mind; he was differing from the classical liberal tradition to the extent that he saw the individual not as a unit but as an opposed self. However, it is also the case that Freud proposed that there was a central portion of the self deep within humankind that had to remain in opposition to society.

Freud's enlightenment heritage led him to denounce religious belief in so bold a manner. As of 1927, Freud insisted that human helplessness was at the origin of religious conviction; people need religion because of the failure to outgrow the dependency of childhood. Religion is an illusion in the sense that

it is the product of wish fulfillment. Freud skeptically saw religion as a pack of lies, fairy tales that were a product of emotional insecurities. Because religion was based on irrational fears, its unreality may undermine the civilization it currently supports. Illusions are dangerous, no matter how comfortable. Freud ignored his earlier comments on religion, relating it to fears of death and guilt, and he concluded that superstition is intolerable. He was so intolerant of the infantile and the regressive that he had difficulty understanding their constructive functions.

Freud saw the family as the prototype for authority relationships. As he had argued that God the father was needed to allay the deepest fears, so he thought that the Oedipus complex also illuminated the social cohesion of political groups. Freud was suspicious of the masses and disdainful of the lower classes. His elitism lay behind a good deal of his social thinking. Religion always seemed to Freud a more intolerable irrationality than political authority. Politically, he was impressed by the extent of human inner instability and the craving for authority. Although Freud's whole form of therapy was aimed at liberation and independence, politically he was a pessimist.

Conflictedness of Life

In "Civilization and Its Discontents" (1930), Freud eloquently expressed his full sense of the conflictedness of life. He stressed the inevitable pervasiveness of suffering in civilized society. Although he could, as he had in "The Future of an Illusion," write like an eighteenth-century libertarian, here his sense of the inevitable cruelties of life was uppermost in the argument.

For civilization to be powerful enough to protect people from one another and from nature, it must, according to Freud, have at its disposal an equally intense energy. Throughout Freud's thought there is a sense of the limits of life, the truth behind the maxim that one cannot have one's cake and eat it too. Social unity can only be achieved on the ruins of human desires. People need the security of civilized life so deeply that they renounce the gratification of instincts in exchange for society. Freud concluded that the frustration of sexual and aggressive drives are entailed by their very character. Only if society can successfully internalize human instinctuality can civilization be maintained.

Challenges

Alfred Adler

As early as the years immediately before World War I, members of Freud's movement were challenging

some of his central points of view. Adler (1870–1937), for example, was a socialist who went on to found a school of individual psychology apart from Freud's own. Adler had a special concern with the social and environmental factors in disease and highlighted the role of compensations for early defects in his study of organ inferiority; he was proposing that, under the best circumstances, defects in a child could create a disposition toward better performance. Adler was not as exclusively concerned with infantile sexuality as Freud was but was, instead, preoccupied with ego mechanisms and aggressive drives. In contrast to Freud's own lack of interest in politics, Adler sought to improve the world through education and psychotherapy.

Carl Gustav Jung

Jung (1875–1961) led the most painful of the secessions from psychoanalysis. Even during his period of cooperation with Freud from 1906 until 1913, Jung had hesitated to extend the concept of sexuality as broadly as Freud wished, and Jung came to interpret much so-called infantile clinical phenomena as being of secondary rather than causal importance; current difficulties, he held, could reactivate past conflicts. Jung insisted that the past can be used defensively to evade the present, a clinical point that later commanded widespread agreement. Less rationalistic and suspicious of the unconscious than Freud, Jung began to formulate his own views on the compensatory functions of symptoms. No better critique of Freud's excessive rationalism can be found than in Jung's collected works. He proposed that symptoms are always justified and serve a purpose. He was also interested in other stages of the life cycle than the oedipal one. It is still not widely known how early Jung emphasized the central importance of establishing a personal rapport between the patient and analyst if therapy is to be successful.

Further Development

The full-scale development of ego psychology was one of the main developments of psychoanalytic theory since the late 1930s, and it silently incorporated most of the central objections that people such as Adler and Jung had had to Freud's original formulations.

Following Freud's death in 1939, more attempts were made to correct the negativism that had been built into his earlier work. Freud's whole system was designed to explain motivation when a person is in conflict and the ego has relatively failed at its integrative task. As a therapist, Freud was preoccupied with pulling problems apart and tearing fixations asunder, on the assumption that the patient's

ego would be able to put the pieces back together again. For Freud, analysis was automatically synthesis; constructive processes had originally been taken for granted by him.

Freud was a master at understanding the main means of self-deception, but he ignored many processes of self-healing. Therefore, a main trend after his death was to correct this imbalance and to focus on the ego as an agency integrating inner needs and outer realities. The ego has a unifying function, ensuring coherent behavior and conduct. The job of the ego is not just the negative one of avoiding anxiety but also the positive one of maintaining effective performance. The ego's defenses may be adaptive as well as maladaptive. Adaptation is itself bedeviled by anxieties and guilts, but the ego's strength is not measured by the earlier psychoanalytic standard of what in a personality is denied or cut off but, rather, by all the extremes that an individual's ego is able to unify.

At the same time, psychology shifted from the more traditional concern with the defensive ego to the problems of growth and adaptation; it looked for the collective sources of ego development. For instance, there can be a need for a sense of identity to be confirmed by social institutions, as Erik H. Erikson (1902–1994) pointed out; and here organized religion and ritual can play a positive role. Cultural institutions not only can play a constructive role, but they help account for the variety of psychological patterns that can be found in different societies.

In his respect for the dignity of his patients, which made his innovations possible, and in his conviction that despite appearances all people are psychologically one, as well as in his individualistic goals, Freud was a great heir of the enlightenment. He was among those who are ever demanding more freedom. At the same time, however, in the development of psychoanalysis the open-ended quality of liberalism led to a revision of some of its most cherished premises; Freud represents an aspect of liberalism's self-examination. In Freud's quest for an understanding of human feelings, he transcended liberalism and joined hands with thinkers usually associated with traditions alien to it. He demonstrated the degree to which the child lives on within the adult and the way psychological uncertainties can prevent people from ruling themselves.

The humanistically oriented revisionists of Freud's views, for example, Erich Fromm (1900–1980), were trying to inject genuine humanitarianism into a psychoanalytic worldview that appeared with Freud to have ended in therapeutic despair and ethical nihilism. In correspondence and conversation, Freud acknowledged that health was only one value among others and that it could not exhaust morality

as a whole. If he was wary about this whole subject of normality, it was because he realized what kind of quagmire he was in danger of entering. He touched on the subject of normalcy only on the rarest occasions. Freud typically took for granted that the people he liked best to work with were creative and self-disciplined.

Freud feared that the most original and disturbing aspects of his ideas would be destroyed by the widespread popular acceptance of his ideas in North America, yet it is questionable whether he did enough to prevent precisely this outcome. By not providing more hints about normality and not owning up publicly to the wide variety of psychological solutions he found both therapeutically tolerable and humanly desirable, Freud contributed to what he most sought to prevent. He had set out to transform Western values; he was eager to go beyond accepted good and evil. When he assaulted the maxim "love thy neighbor as thyself" as both unrealistic and undesirable, he was explicitly trying to overturn Christian ethics.

Conclusion

It is logically impossible to talk about neurosis without at the same time implying a standard of maturity as well, and yet despite how powerful psychology can be in outlining human defects and weaknesses, it has not been nearly as successful in coming to terms with the positive side of human strength and coherence. In the end, the issue of the significance of normality and its relationship to nihilism has to be left an open question. Freud's psychology did contribute to our understanding of what it can mean to be human. However, it is impossible to attempt to spell out in a definitive way the ideological implications of psychoanalysis. The writers who have been influenced by Freud constitute a wide range of people, from the most radical to some of the most conservative.

It was an old analyst and loyal disciple of Freud, Helen Deutsch (1884–1982), who had the most appropriately philosophical attitude toward the perplexing issue of normality. In her earlier years, when she was one of the most prominent teachers in the history of psychoanalysis, she made it a practice in the course of interviewing prospective analysts for acceptance into training to ask what they thought a normal person would be like. It is, of course, an ultimately unanswerable conundrum and yet one that civilized people are obliged to raise repeatedly. Like all genuine questions in social philosophy, the problem of normality can never be solved; it has to remain a real issue to the extent that we find it intolerable to contemplate a universe lacking in moral values.

See Also the Following Articles

Freud, Sigmund; Psychotherapy.

Further Reading

- Ellenberger, H. F. (1970). *The discovery of the unconscious: the history and evolution of dynamic psychiatry*. New York: Basic Books.
- Freud, S. (1953–1974). Civilization, its discontents (1930). In: Strachey, J. (ed.) *Standard edition of the complete*

psychological works of Sigmund Freud (vol. 21), pp. 59–105. London: Hogarth Press and the Institute of Psychoanalysis.

Freud, S. (1953–1974). The future of an illusion (1927). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vol. 21), pp. 3–56. London: Hogarth Press and the Institute of Psychoanalysis.

Roazen, P. (1975). *Freud and his followers*. New York: Knopf (Reprinted 1992, New York: Da Capo.).

Psychoanalytic Theory See: Psychoanalysis.

Psychological Stressors, Overview

S M Monroe and G M Slavich

University of Oregon, Eugene, OR, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S M Monroe, volume 3, pp 287–293, © 2000, Elsevier Inc.

*Stress
sensitization*

secretions that help to control bodily metabolic activity.

The enhanced and progressive sensitivity of an organism to stress given repeated exposure to stressors.

Historical and General Considerations

Conceptual Progress

Methodological Considerations and Recent Developments

Glossary

Catecholamines Neurotransmitters, including epinephrine, norepinephrine, and dopamine, that promote sympathetic nervous system activity. They may be released in substantial quantities during stressful times.

Corticosteroids Complex chemical compounds produced in the outer layer of the adrenal gland, which include both mineralocorticoids and glucocorticoids. Mineralocorticoids maintain salt and fluid balance in the body, while glucocorticoids have metabolic and anti-inflammatory effects and are important mediators of the stress response.

Neuroendocrine Relating to the nervous and endocrine system, which produces endocrine

Historical and General Considerations**Historical Matters**

It is frequently assumed that psychological stressors are a concern of especially modern origins, or at least that they have become more prominent with recent advances in society and technology. It is also commonly assumed that with the accelerating progress of civilization, more and more people are afflicted with mental and physical disorders. Historical accounts, however, suggest that such ideas about stressors, civilization, and disease have been common for quite some time. Sir Clifford Albutt (1895: 217) expressed such sentiments quite clearly well over 100 years ago:

To turn now . . . to nervous disability, to hysteria . . . to the frightfulness, the melancholy, the unrest due to living at a high pressure, the world of the railway, the pelting of telegrams, the strife of business . . . surely, at any rate, these maladies or the causes of these maladies are more rife than they were in the days of our fathers?

The tendency to view life in stressful terms may be even more basic to human cognition than is readily

apparent. The Greek myth of Sisyphus is enlightening in this regard. The perpetual work of pushing a boulder up a mountain – only to have gravity bring it back down after each and every effort – captures some of the qualities and characteristics linked to modern views of psychological stressors. Perhaps there is something fundamental about the human condition and psyche that fosters a perception of the world as a place rife with unrelenting demands that can never be fully met, resulting in subjective states of fatigue and distress and eventually leading to ill health. Each era may bring its unique colorations to such perceptions and its own attributions regarding their origins.

It is against this psychological backdrop of belief and possible bias in thinking that modern work on psychological stressors must be examined. Psychological stressors and related concepts have been popular explanatory constructs throughout recent, and perhaps not so recent, history. As a result of their subjective allure and apparent explanatory power, these ideas have often been loosely formulated. Owing to conceptual fuzziness and ambiguity, not only has progress in science been slowed, but nonscientific issues, ideas, and biases have been permitted to masquerade as scientific truths.

The concept of psychological stressors is rich with possibilities for shedding light on important matters in adaptation, dysfunction, and disease. The concept is paralleled, though, by the potential pitfalls that accompany its intuitive appeal. The challenge is to translate the fertile ideas about psychological stressors into more precise concepts, definitions, and operational procedures. With more sound definitional and methodological procedures in place, the utility of stress concepts for understanding adaptation and maladaptation, mental and physical disorder and disease, will be better understood.

Early Ideas and Research

A broad foundation for understanding the organism's reactions to challenging environmental circumstances was laid down by Claude Bernard and Charles Darwin during the nineteenth century. Each of these two influential individuals in his own way touched on issues deriving from the tension resulting from ongoing adaptation to changing and challenging environmental circumstances. Yet it was not until the early to mid-twentieth century that such generality and complexity began to be translated into more specific terminology and technology. These efforts can be traced to at least three different lines of thought and research.

The early work of Walter Cannon dealt with ideas about common emotions and their physiological consequences, particularly with respect to the body's maintenance of homeostasis. This line of study was

complemented shortly thereafter by the animal laboratory studies of Hans Selye, wherein acute and severe stressors were systematically investigated. It was in Selye's work that the concept of stress most forcefully emerged. Stress was defined as "the nonspecific response of the body to any demand" (Selye, 1976: 74). Stressors, in turn, were defined as "that which produces stress" (Selye, 1976: 78). Finally, from another vantage point, Adolph Meyer popularized the life chart methodology. This approach emphasized the importance of the dynamic interplay between biological, psychological, and social factors, such that important events within the person's biography became foci of attention for studying health and disease. Collectively, these activities, and the multiple lines of research they generated, served to initiate specific awareness of, and interest in, psychological stressors.

Arising outside of the more purposeful activities of science was another influential development that contributed to the emerging idea that psychological stressors cause both mental and physical disorders. Prior to World War II, psychopathology was commonly attributed to genetic factors or acquired biological propensities; so-called normal people devoid of such taints were thought to be largely invulnerable to mental illness. The experiences during and after World War II dramatically shifted thinking in medical and psychiatric circles to incorporate the idea that severe stress could precipitate breakdown in a previously healthy individual. Once this conceptual shift began, it underscored the multiplicity of health consequences that could be caused by severe stressors. It also opened the door for enlarging conceptual perspectives on psychological stressors by considering how less severe, yet still noxious, aspects of the social and physical environment could contribute to, or precipitate, pathology.

Conceptual Progress

Upon the foundations of stress research and theory laid down by Selye, Cannon, and Meyer, along with the influences of experiences of World War II, modern inquiry into the effects of psychological stressors became a topic of increasing interest and, eventually, of extensive empirical inquiry. Two general themes may be discerned that have underpinned advances in theory: (1) characteristics of psychological stressors and (2) individual differences in response to psychological stressors.

Stressor Characteristics

Despite general agreement about the importance of psychological stressors for health and well-being,

determining exactly what it is about stressful circumstances that is deleterious has proven challenging. An initial question of considerable theoretical importance involved the basic nature of psychological stressors: are they best viewed in a unitary manner as nonspecific demands on the organism (as postulated by Selye), or are psychological stressors more effectively viewed as a class of conditions harboring specific component characteristics of importance? Investigators from two traditions – animal and human research – have addressed this issue, with parallel and sometimes intersecting developments. Although considerable progress has been made, the general topic of elaborating stressor characteristics remains one of central importance in current thinking on psychological stressors.

Animal Laboratory Research

A great deal of work in the 1960s and 1970s was performed to determine whether specific psychological characteristics of stressors possess qualitatively distinct implications for the organism. Initially this work revealed how particular features associated with the environmental stressors might be important for predicting adverse outcomes (as opposed to the more psychologically neutral notion proposed by Selye of general or nonspecific adaptive demands). Such research went on to probe the types of psychological stressors and their effects. It became of central interest to understand in a more differentiated way the effects of diverse psychological stressors.

Animal laboratory studies adopted ingenious designs to differentiate psychological components associated with environmental stressors, with the findings from these studies demonstrating that distinctive psychological characteristics were responsible for many immediate behavioral or physiological responses. For example, specific psychological characteristics of stressors, such as undesirability or controllability, were particularly pertinent for the development of various disorders. It became clear, too, that other characteristics of stressors were important. For example, different parameters of shock administration (acute, intermittent, or chronic) produced different physiological effects in animals. Further, such differences might increase, decrease, or not influence the development of particular diseases. Finally, psychological stressors not only could influence immediate psychobiological functioning, but also could have long-term ramifications through permanent alteration of the organism.

As the importance of specificity of stressor characteristics became more apparent and accepted, questions about the specificity of stress responses also arose. What were the implications of specific stressor

characteristics for different facets of psychological and physiological functioning? Such theoretical developments greatly extended the framework for inquiry, requiring attention to multiple characteristics of stressors in relation to multiple psychological and biological processes and outcomes. Relatively simple, singular response indices (e.g., corticosteroids, catecholamines) were replaced by patterns of behavioral and biological effects or profiles of neuroendocrine responses. More recently, other levels of conceptualization have been proposed. For example, psychological stressors may promote fundamental disruptions in oscillatory regulation of basic biological functions or reversions to earlier modes of functioning.

Overall, research on psychological stressors from animal research has moved beyond unidimensional and linear concepts of stressors and their effects. More recent thinking has adopted a larger framework for understanding the diverse characteristics of stressors that influence particular and varied response systems of the organism. The response systems of interest have expanded from single systems to patterns or profiles of response across multiple indices.

Human Experimental and Field Studies

Investigators of psychological stressors in humans also conducted innovative and insightful studies, both in the laboratory and in the field. Early work tended to focus on the aversive subjective attributes, particularly perception or appraisal, of psychological stressors as evaluated in an experimental setting. Yet at about this same time research on stressful life events began. It is in this area of stress research that activity on psychological stressors perhaps reached its pinnacle, in terms of both productivity and popular interest.

Extrapolating from animal laboratory studies on the one hand, and integrating with Meyer's life chart procedures on the other, Thomas Holmes and Richard Rahe first formulated the idea that distinctive changes in one's life circumstances – specific and documentable life events – could be defined and assessed in an objective manner. The work was initially based on case histories of some 5000 tuberculosis patients, from which they derived a list of 43 life events "empirically observed to occur just prior to the time of onset of disease, including, for example, marriage, trouble with the boss, jail term, death of spouse, change in sleeping habits, retirement, death in the family, and vacation" (Holmes, 1978: 46). The Schedule of Recent Experiences (SRE) was developed and published, and by 1978 more than 1000 publications had utilized this convenient method for

probing a vast range of questions pertaining to stress and illness.

The common feature associated with these disparate life changes – the stressor characteristic of primary concern – was thought to be the degree of social readjustment entailed by the event: “The relative importance of each item is determined not by the item’s desirability, by the emotions associated with the item, nor by the meaning of the item for the individual; it is the amount of change that we are studying and the relationship of the amount of change to the onset of illness” (Holmes, 1978: 47). This viewpoint is consonant with Selye’s ideas about stressors and stress. Hence, the psychologically neutral notion of the readjustment required of life changes was conceptualized as the characteristic responsible vulnerability to a wide variety of psychological and physical maladies.

Much as the emphasis in animal laboratory studies shifted from psychological neutral concepts of any demand, viewpoints within the stressful life events literature began to shift away from the concept of readjustment and toward emphasizing the undesirable characteristics of events. Human studies of life events consequently began to focus on the particular characteristics of psychological stressors and their potentially unique effects. The principle of specificity also was extended from the characteristics of stressors to the specific consequences of such experiences, elaborating theory about the importance of specific psychological stressors for specific responses and eventually for specific types of disorder or disease. A vast literature on this topic has appeared over the past two decades, with diverse conceptualizations of psychological stressors and myriad methods designed to measure them.

The issue of desirability of events, however, along with the more general issue involving stressor characteristics, brought into focus another important subject in the study of psychological stressors: individual differences. What might be viewed or experienced as undesirable by one person could be viewed or experienced as desirable by another. As discussed next, a variety of considerations are invoked to explain variability in effects and outcomes in relation to psychological stressors.

Individual Differences

Despite progress in conceptualizing the component characteristics of psychological stressors, and despite progress in prediction afforded by such work, considerable variability in response to psychological stressors occurs. Even under the most dire of stressful conditions, all animals or individuals do not

necessarily break down. Although a refined understanding of stressor characteristics still may account for some variability in outcomes, other factors may be useful to effectively model effects of psychological stressors. Progress in understanding this issue has again come from both the basic laboratory and human studies of psychological stressors.

Animal Laboratory Research

Although there were characteristic features of physiological responses to the stressors employed in the early paradigm adopted by Selye, not all animals responded to stress in an identical manner. Further, individual differences in response were even more pronounced when the less severe types of stressors were used.

Such variables as prior experience, availability of coping responses, and other aspects of the social and experimental context were found to moderate the influence of psychological stressors. For example, when rats are exposed to electric shock, animals that cannot predict shock occurrence (via warning tones) develop a sixfold increase in gastric ulceration compared to their yoked counterparts (who receive the warning tones). Work along these lines demonstrated the delicate and often subtle interplay among stressor, social context, and resources available to the organism in determining response outcomes. These lines of study, too, suggested that individual differences in susceptibility to psychological stress could be viewed within a developmental perspective in terms of stress sensitization. Laboratory animals repeatedly exposed to severe psychological stressors can become neurobiologically sensitized to the stressors, such that relatively minor degrees of stress eventually acquire the capability of triggering pathogenic responses.

Human Life Stress Research

The importance of individual differences was perhaps more apparent in studies of human life stress and its consequences. A consistent criticism of life events research was the relatively weak association between psychological stressors and disorder. It was assumed that other considerations moderated stress effects, and the elucidation of such factors would increase the predictive capability of disorder following stressful events. Again, there were a number of factors that were believed to moderate the impact of psychological stressors, ranging from environmental factors such as social support to more individual factors such as prior experience and coping. Developmental considerations have also been important in recent theorizing about individual differences in responsivity to

psychological stressors, with the idea that prior exposure to severe psychological stressors renders the individual more susceptible to increasingly lower levels of psychological stress.

A major arena for understanding individual differences in stress susceptibility has been perception. The early and elegant laboratory studies of human stress had indicated the importance of such individual differences in perception, or appraisal, of stressors, and such thinking was readily incorporated into theory and method. Studies of life events, for example, used subjective weights of events experienced by the study participants. Once this avenue of inquiry was opened, it also brought to the forefront a variety of influences on perception, along with other factors that might influence stress responsivity. Thus, research began to focus not only on appraisal of stressors, but also on coping, social support, personality, and other considerations that in theory could moderate the effects of psychological stressors.

As research progressed along these lines, it became clear that making some of these distinctions was easier to do in theory than in method. For example, while it made good sense theoretically to consider an individual's subjective perception of psychological stressors, it was more difficult to employ such information in a scientifically sound manner. When it came to operationalization, serious problems became apparent. For example, owing to depression-based perceptual biases, a depressed person might have a skewed perception of events and rate them as particularly negative (irrespective of the objectively stressful qualities *per se*). Generally, such concerns raised an important paradox for investigations of psychological stressors. Namely, although a large part of the desired knowledge pertained to the individual's idiosyncratic appraisal of psychological stressors, methodological concerns cautioned against direct assessment of such information. Instead, alternative approaches were developed to avoid the pitfalls of using subjective reports in research on psychological stressors, as well as to avoid other problems with these methods that eventually came to light.

Methodological Considerations and Recent Developments

While concepts and methods often intertwine and, united, nurture progress, at times one or the other component may unduly influence development (for good or for bad). This situation appears applicable to research on psychological stressors, in which the methods adopted in animal laboratory research have constrained theory and methods adopted in human

life stress research have misled theory on psychological stressors.

Animal Laboratory Research

The original work of Selye typically employed situations that were overpowering or unavoidable for the animal. Such conditions did not permit an evaluation of behavioral responses or other moderating influences that could influence an animal's adaptation to stressors. Further, it was realized that this methodological paradigm was not informative about psychological stressors that might be of greater ecological and evolutionary relevance (i.e., more typical with respect to the animal's natural environment and evolutionary history). Thus, such an approach readily masked the implications of less severe psychological stressors on physiology and behavior, which in turn might represent a more fertile area of inquiry into stressor effects. Finally, the nature of the stressor employed in the early stress laboratory studies also contributed to the aforementioned difficulty in differentiating physical from psychological effects, which inhibited progress in the arena of conceptual development.

Overall, the range of psychological stressors was constrained by the methods adopted. Theory, in turn, was constrained to account for the consequences of stressors under such restricted and relatively unnatural stressor conditions. More recent research has benefited from methods that encourage assessment of diverse characteristics of psychological stressors and their severities and incorporate the assessment of a wide variety of behavioral and biological response possibilities. Current perspectives based on these broader methodological approaches suggest that the organism's responses are often exquisitely specific nuances of stressors encountered.

Human Studies of Stressful Life Events

The bulk of empirical work on human life stress has been based on self-report checklist methods. The prototype of this approach is the SRE, the instrument that catalyzed research in this area. The popularity of the SRE was likely due to the combination of the intuitive appeal of the stress concept, the apparent objectivity of the method, and the overall impression of scientific legitimacy.

The methodological paradigm launched by the SRE, however, embodied several problems. It became clear that subjects did not report life events in a reliable manner over time and that investigators did not adequately control for the directionality of effects in research designs (e.g., being depressed initially could bring about life events such as trouble at work,

difficulties with spouse, and so on). Indeed, many of the initial items on the SRE were direct indicators of disorder or illness. For example, some of the key criteria for defining clinical depression were represented in the original SRE (e.g., major change in eating habits, major change in sleeping habits). If measures of life events were directly confounded with the presence of disorder or were contaminated by the effects of pre- or coexisting disorder, then clearly general theory about psychological stressors, as well as theory about the characteristics of psychological stressors, rested on flawed information.

In response to these methodological concerns, other investigators designed semistructured interview protocols and developed explicit guidelines, decision rules, and operational criteria for defining and rating life events. These developments further highlighted serious problems with self-report checklist methods. For example, there was too much subjective leeway permitted in defining what constitutes an event with self-report procedures, resulting in considerable variability of content within ostensibly uniform categories of events. In order to have a more firm methodological foundation, the more elaborate and extensive interview and rater-based procedures were employed, which helped to standardize measurement across individuals.

In general, such interview and rater-based approaches have been found to enhance the reliability of life event assessments and to provide stronger predictions of particular kinds of disorders following the occurrence of psychological stressors. Procedures such as these also provide a solid foundation upon which to build in terms of developing taxonomies of psychological stressors and their effects. Although such approaches are more time- and labor-intensive to implement, they represent the current-day gold standard for assessing psychological stressors.

Human Studies Employing Other Measures of Psychological Stressors

There have been other methods in which psychological stressors have been defined and studied. None of these approaches has received the degree of attention devoted to the work on life events, yet each may have useful properties for the study of psychological stressors. Two lines of investigation are noteworthy.

Many investigations have targeted people who experience a specific life event and compared these individuals with controls who do not experience the event. For example, individuals who become unemployed are compared to individuals who do not experience this event in relation to a variety of psychological and physical processes and outcomes. Such work is useful for examining a potentially

more homogenous process with more readily identifiable outcomes. On the other hand, such studies may oversimplify the psychological stressors associated with an event and not specifically articulate the different components within the general event that are most pernicious for health and well-being. For example, the effects can be partitioned into a variety of stressful themes that, although often intercorrelated, may not have uniform effects. Thus, although people who become unemployed in general may experience a loss of self-esteem, loss of income, loss of daily schedule, and so on, each particular situation may pull more or less for heightened responses along these different dimensions. Work sensitive to such variability in the component characteristics will be most useful for research on psychological stressors.

Finally, there also have been efforts to measure psychological stressors through questionnaire or diary methods, inquiring about minor but common daily events, chronic conditions, appraisal processes, and other indicators or correlates of psychological stressors. A promising recent avenue of research involves ecological momentary assessment, where subjects can be prompted throughout the day to respond to queries about their circumstances and psychological states. Such procedures help minimize problems with standard retrospective methods.

In closing, it is appropriate to return to the concerns and caveat with which the article began. The specter of possible biases in the measurement of psychological stressors must be consistently borne in mind, and methods employed must be rigorously attentive to such concerns, to provide a solid empirical foundation upon which theory and research can build for this important area of investigation.

See Also the Following Articles

Affective Disorders; Animal Models (Nonprimate) for Human Stress; Cardiovascular System and Stress; Cognition and Stress; Control and Stress; Coping Skills; Corticosteroids and Stress; Depression Models; Environmental Factors; Immune Suppression; Life Events Scale; Psychosocial Factors and Stress; Selye, Hans; Social Support; Stress, Definitions and Concepts of; Stress Generation.

Further Reading

- Allbutt, C. (1895). Nervous diseases and modern life. *Contemporary Review* 67, 217.
- Brown, G. W. and Harris, T. O. (1989). *Life events and illness*. London: Guilford Press.
- Cohen, S., Kessler, R. C. and Gordon, L. U. (eds.) (1995). *Measuring stress: a guide for health and social scientists*. New York: Oxford University Press.

- Dohrenwend, B. P. (ed.) (1998). *Adversity, stress, and psychopathology*. New York: Oxford University Press.
- Holmes, T. H. and Rahe, R. H. (1967). The social readjustment rating scale. *Journal of Psychosomatic Research* 11, 213–218.
- Lazarus, R. S. (1966). *Psychological stress and the coping process*. New York: McGraw-Hill.
- Monroe, S. M. and Harkness, K. L. (2005). Life stress, the 'kindling' hypothesis, and the recurrence of depression: considerations from a life stress perspective. *Psychological Review* 112, 417–445.
- Post, R. M. (1992). Transduction of psychosocial stress into the neurobiology of recurrent affective disorder. *American Journal of Psychiatry* 149, 999–1010.
- Selye, H. (1976). *The stress of life* (2nd edn.). New York: McGraw-Hill.
- Stone, A. A., Shiffman, S. S. and DeVries, M. W. (1999). Ecological momentary assessment. In: Kahneman, D., Diener, D. & Schwarz, N. (eds.) *Well-being: the foundations of hedonic psychology*. New York: Russell Sage Foundation.
- Weiner, H. (1992). *Perturbing the organism: the biology of stressful experience*. Chicago, IL: University of Chicago Press.
- Weiss, J. M. (1972). Psychological factors in stress and disease. *Scientific American* 226, 104–113.

Psychoneuroimmunology

R Dantzer

Integrative Neurobiology, Bordeaux, France

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R Dantzer, volume 3, pp 294–298, © 2000, Elsevier Inc.

Immune System

Relationships between the Nervous System and the Immune System

Functional Consequences of Relationships between the Nervous System and the Immune System

Immune Influences on Brain Functions

Implications for Psychosomatics

Glossary

<i>Cytokines</i>	Secreted regulatory proteins in response to external insults that control the survival, growth, differentiation, and effector function of tissue cells.
<i>Immune system</i>	A complex network of specialized cells and organs (lymphoid organs) that distinguishes between self and nonself and defends the body against infections from agents such as bacteria, viruses, fungi, and parasites.

Immune System

Immunology deals with understanding how the body distinguishes between what is self and what is non-self. The classic hallmark of the immune system is its ability to recognize and kill foreign substances and at

the same time recognize but not destroy normal host tissues. To protect against potentially pathogenic microorganisms that are foreign to the body (antigens), the host uses white blood cells, or leukocytes, that originate from precursors in the bone marrow. These cells circulate through the blood and lymph system or, for some of them, are fixed in specific tissues. Innate immunity represents the first line of defense against antigens and involves phagocytic cells that are mononuclear (e.g., macrophages) and polymorphonuclear (e.g., neutrophils). These cells capture microbes and digest them, and the mononuclear macrophages also present antigen fragments to lymphocytes. Lymphocytes have highly diverse specific receptors, which allow them to recognize antigens. Adaptive immunity is based on the proliferation and differentiation of antigen-specific lymphocytes into effector cells. The T lymphocytes, so designated because they differentiate and mature in the thymus gland before traveling to the spleen and lymph nodes, are the first of the lymphocytes to be activated. Following contact with antigen-presenting cells such as macrophages, lymphocytes proliferate and destroy the infected cells, in which case they are called cytotoxic T lymphocytes, or they help other lymphocytes to intervene.

There are many different subclasses of T cells that can be distinguished from each other by specialized cell surface molecules. For instance, cytotoxic T lymphocytes are called CD8 positive T cells because they express a glycoprotein known as CD8. CD4 positive T cells, which are markedly depressed in patients suffering from acquired immunodeficiency syndrome, are essential for the activation of macrophages and

B lymphocytes. Other lymphoid cells, called natural killer (NK) cells, are specialized in the recognition of tumor cells. B lymphocytes, so designated because they develop in the bone marrow of adults, differentiate into plasma cells that produce antibodies when they bind antigen and are co-stimulated by a subset of T cells. Antibodies, also called immunoglobulins (Ig), are soluble molecules that attach specifically to antigens, making their destruction possible through a series of complex reactions involving complement molecules. An immune response results in the generation of an increased number of differentiated memory lymphocytes that, upon re-exposure to the same antigen, generates more rapid and effective response. Some antigens induce the production of specialized antibodies, known as immunoglobulin E (IgE), that bind to the surface of mast cells. Recognition of an antigen or allergen by IgE attached in this manner induces the release of various mediators (e.g., histamine and serotonin) that are contained in the secretory granules of mast cells and are responsible for an allergic reaction.

The various cells that are involved in immune reactions communicate with each other by cell-to-cell contact via specialized adhesion molecules and by soluble mediators known as cytokines or interleukins. These two forms of cell-to-cell communication are essential for the initiation and regulation of the host response to infection. Each category of immune cells produces characteristic cytokines. Activated monocytes and macrophages produce the proinflammatory cytokines that are necessary for the development and regulation of the inflammatory response and include interleukin-1 (IL-1), IL-6, and tumor necrosis factor- α (TNF- α). Two subsets of CD4 positive T cells, or T helper (Th) cells, can be distinguished on the basis of their profile of cytokine secretion and the type of immune response they promote. Th 1 cells secrete IL-2, interferon- γ (IFN- γ), and TNF- α , whereas Th2 cells release IL-4, IL-6, and IL-10. Th1 cells are major players in delayed-type hypersensitivity and proinflammatory responses, whereas Th2 cells induce B cell growth and differentiation.

Relationships between the Nervous System and the Immune System

Although the immune system has long been viewed as a self-regulated system, there is now a wealth of evidence in favor of the existence of intricate anatomical and functional relationships within the nervous system. In particular, branches of the autonomic nervous system innervate the spleen and lymph nodes. Immune cells have functional membrane receptors for both classical neurotransmitters and neuropeptides.

This is important because it implies that immune cells proliferate and differentiate in a microenvironment, the composition of which is dependent on the activity of the sympathetic nervous system. In addition, leukocytes themselves are able to produce and release many peptides that have been identified initially in the neuroendocrine and nervous system, such as pro-opiomelanocortin, corticotropin-releasing hormone, growth hormone, and insulin-like growth factors. Conversely, microglial cells, which are the equivalent of macrophages within the central nervous system, are able to synthesize and release proinflammatory cytokines such as IL-1 and TNF- α that are involved in the regulation of the inflammatory response at the periphery, and brain cells also express receptors for these and other cytokines.

Functional Consequences of Relationships between the Nervous System and the Immune System

The commonality of messengers and receptors between the immune system and the nervous system provides new vistas into the old notion that psychological factors can influence the development and evolution of somatic diseases. Objective studies on the influence of psychosocial factors and specific emotional states on immune responses are now possible thanks to the development of quantitative methods for assessing changes in immunity in animals and human subjects. In animal studies, exposure to a variety of stressors, including painful electric shocks, physical restraint, social separation, and the odor of another animal that is stressed, can depress NK cell activity and T and B lymphocyte proliferating response to nonspecific mitogens. Psychological aspects of stressors, particularly the subject's ability to predict and control their occurrence, are more important than their severity, frequency of occurrence, and duration. The effects of stress on immunity depend on the immune response under consideration, and a given stressor can either suppress or enhance different immune responses in the same individual. For example, mice that were exposed for 24 h to the odor of another stressed mouse display a decrease in IL-2 production by splenocytes stimulated with a T cell mitogen, a reduction in NK cell activity, and an increase in IgM and IgG antibody titers against a T cell-dependent antigen. These apparently disparate results can be interpreted as an effect of stress on the Th1/Th2 balance, in the form of activation of Th2 cells and downregulation of Th1 cells. Human studies show that immune functions can be altered by various experimental procedures, such as emotion induction under controlled laboratory conditions

and naturally occurring life events, such as academic examination, difficulties in marital life, and intractable diseases (e.g., cancer, early dementia) affecting the spouse. Major depression is accompanied by alterations in the distribution of T cell subsets and by declines in nonspecific measures of immunity such as NK cell activity and mitogen-induced lymphocyte proliferation. However, there are also signs of immune activation, in the form of increased numbers of leukocytes, neutrophils, and activated T lymphocytes, and an increased production of Th1-like cytokines.

Pathways involved in the mediation of the effects of stress on immunity can be identified by interfering with the activity of the autonomic nervous system and endocrine glands that are known to be involved in neuro-immune interactions. Experimental studies have made it clear that glucocorticoids are not the sole mediator of the effects of stress on immunity. Activation of the autonomic nervous system plays an important role, probably by altering the molecular composition of the microenvironment in which lymphocytes proliferate and differentiate.

The possible consequences of stress-induced immune alterations on the host resistance to infectious disease or the progression of tumor cells are not easy to assess. The relationships between stress and disease are already extraordinarily complex, and in most cases, the importance of the association between stress and illness behavior (symptom awareness and use of health services) is blurring the picture. It is now obvious that there is no global answer to this issue. In each case, a slow and cautious approach, in which the whole set of relationships among a stressor, a specific pathway, immunity, and disease is examined systemically, is more likely to yield valid answers.

Immune Influences on Brain Functions

Infection with foreign agents, in the form of infectious microorganisms or nonreplicating antigens, has long been known to cause a classical stress response, involving activation of the hypothalamic-pituitary-adrenocortical (HPA) axis and increases in catecholamine turnover in the brain and periphery. It is now apparent that these neuroendocrine and neurochemical changes form part of a larger, physiological regulatory system that modulates functional activities of lymphocytes and phagocytes. Molecules that signal the brain that an immune event has been initiated in the periphery have been identified. They are represented by proinflammatory cytokines, such as IL-1 and TNF- α , that are synthesized and released by

activated macrophages and monocytes during the course of an infection. These peripherally released cytokines act on the brain by two pathways: a neural pathway that involves sensory fibers innervating the bodily site of infection and a humoral pathway that involves blood-borne cytokines and activated circulating immune cells. In the first case, the transmission of the immune message to the brain is direct, whereas in the second case, it needs to be relayed by specialized parts of the brain that have a deficient blood-brain barrier and are known as circumventricular organs because of their anatomical location around the brain ventricles. In response to these peripheral signals, glial cells produce proinflammatory cytokines that, by acting on neural target cells, activate the HPA axis, resulting in an increase in circulating glucocorticoids. This response is important for the regulation of the immune response, as glucocorticoids feed back on activated immune cells and attenuate the synthesis of proinflammatory cytokines. Glucocorticoids also act on target cells of proinflammatory cytokines to block the transcription effects of these cytokines. The lack of a glucocorticoid response to an inflammatory stimulus is responsible for development of the disease in experimental models of autoimmune diseases in animals, whether the disease is induced by the administration of specific antigens (e.g., adjuvant-induced arthritis and experimental allergic encephalomyelitis in the Lewis rat) or develops spontaneously (e.g., autoimmune thyroiditis in the obese chicken). In addition to their potent effects on the HPA axis, proinflammatory cytokines also induce fever and nonspecific behavioral responses to infection, such as malaise, pain, fatigue, sleepiness, anorexia, apathy, social withdrawal, irritability, and cognitive deficits. The development and duration of these physiological and behavioral alterations are regulated by molecules called cryogens that oppose the effects of proinflammatory cytokines. Cryogens include neuropeptides, such as vasopressin and α -melanotropin, and anti-inflammatory cytokines, such as the IL-1 receptor antagonist (IL-1RA) and IL-10.

The profound influences of cytokines on brain functions offer insights into understanding why patients afflicted with cancer or hepatitis C experience flu-like symptoms and develop acute psychotic or depressive episodes when injected with recombinant cytokines. Furthermore, many chronic inflammatory conditions have been demonstrated to be also associated with priming of the brain cytokine compartment, so that a local or a systemic immune challenge can precipitate the occurrence of cognitive alterations and mood disorders. This is particularly the case in

aged persons, obese subjects, and patients suffering from atherosclerosis. It could explain the higher prevalence of mood and mild cognitive disorders that is associated with such conditions. In addition to peripheral influences on brain cytokines, it is important to note that proinflammatory cytokines are also induced in the brain by infiltrating blood cells after a brain trauma, a stroke, or a local immune mediated inflammatory response that compromises the blood-brain barrier. The proinflammatory cytokines that are produced in these conditions play a pivotal role in the pathogenesis of the corresponding neurological disease. They induce proliferation of astrocytes and neurodegeneration via necrotic and apoptotic routes, and they are also responsible for the psychopathological and neurological alterations accompanying the disease (e.g., AIDS-related dementia).

Implications for Psychosomatics

Demonstration of the potent behavioral and psychological effects of cytokines has important implications for understanding the relationships between stress and disease. It is no longer tenable to go only one way and claim that life events interacting with personality traits, mood, and psychological variables play a causal role in the pathogenesis of immune-related diseases because of their autonomic and hormonal concomitants. Emotional dispositions and psychological characteristics can be the result of alterations in brain functions caused by still undetectable alterations in immune functions. In the case of cancer, for example, feelings of helplessness and hopelessness that have been related repeatedly to the onset and progression of a tumor may be secondary to central effects of soluble products released by immune cells of the host or even tumor cells during the early stage of the disease, when cancer is still not clinically detectable. The inflammatory response to cancer treatments, especially radiotherapy; can have similar consequences on brain functions.

Research on the effects of psychosocial factors on the onset and progression of diseases ranging from AIDS to allergies and musculoskeletal disorders is now taking a new turn. In an organism in which physiological systems are interconnected by reciprocal regulatory systems, a dysfunction in one system is likely to affect the functioning of other systems. In addition, the dysfunction can be at the level of the regulatory process itself rather than within one of the regulated organs. The immune system has long been represented as a self-regulated system that is relatively insensitive to the host's physiological turmoil. It is now clear that this view is inappropriate, and that reciprocal connections between the brain and the immune system play an important part in the coordinated response to infection and inflammation. The challenge now is to understand how such regulatory, homeostatic mechanisms can be compromised and lead to disease.

See Also the Following Articles

Cytokines; Cytotoxic Lymphocytes; Disease, Stress Induced; Health Behavior and Stress; Hypothalamic-Pituitary-Adrenal; Immunity; Psychosocial Factors and Stress.

Further Reading

- Ader, R., et al. (eds.) (2006). *Psychoneuroimmunology* (4th edn.) New York: Elsevier.
- Konsman, J. P., Parnet, P. and Dantzer, R. (2002). Cytokines and sickness behavior: mechanisms and implications. *Trends in Neurosciences* **25**, 154–159.
- Perry, V. H. (2004). The influence of systemic inflammation on inflammation in the brain: implications for chronic neurodegenerative disease. *Brain, Behavior, and Immunity* **18**, 407–413.
- Segerstrom, S. C. (2005). Optimism and immunity: do positive thoughts always lead to positive effects? *Brain, Behavior, and Immunity* **19**, 195–200.

Psychosocial Factors and Stress

J Siegrist

University of Duesseldorf, Duesseldorf, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J Siegrist, volume 3, pp 299–303, © 2000, Elsevier Inc.

General Background

Overview of Psychosocial Factors Influencing
Stress and Health

Implications for Intervention

Glossary

<i>Self-efficacy</i>	Personal beliefs about one's ability to master challenges.
<i>Self-esteem</i>	Firm feelings and beliefs about one's own worth as a person (influenced strongly by experience on how significant others react toward one's self early in life).
<i>Social anomie</i>	Lack of social rules and orientations in a social group, often associated with feelings of anxiety or helplessness.
<i>Social capital</i>	Patterns of cooperative social exchange in neighborhood and community life characterized by mutual commitment and trust.
<i>Social role</i>	Set of expectations or norms (duties, options) directed toward people who hold important social positions (e.g., family role, work role).
<i>Type A behavior</i>	Complex of behaviors and emotions defined by excessive striving, competitiveness, hostility, and impatience.

General Background

In evolutionary terms, the development of the human brain, particularly its neocortical structures, has critically advanced the species' capacity to cope with environmental challenges and threats. As a consequence, inherited biobehavioral patterns of response to threats to the integrity, survival, and reproduction of the organism have become more variable and complex compared to those of the nonhuman primates. Increasingly, these patterns of response depend on the appraisal of stimulus properties and available opportunities for successful agency and are modulated by prior experience, learned skills, and interpersonal support. This may hold true for responses to natural and man-made disasters and for chronically demanding

socioenvironmental conditions. However, in all these instances, a person's sense of the control over the challenge and over his or her actions in dealing with it must be considered a crucial variable that triggers the quality and intensity of stressful experience. Stressful experience is defined as the experience of a mismatch between the demands put on an individual in a challenging situation and his or her ability to meet the situation with adequate responses. At the level of emotional reactions, this mismatch is likely to elicit feelings of anxiety, irritation, anger, disappointment, or even helplessness and despair. These emotions are often paralleled by sustained neuronal and neuroendocrine activation due to the existence of powerful neuronal circuits that link neocortical information processing with mesolimbic and brain stem activity.

Humans are particularly vulnerable to the distress-enhancing effects of socioenvironmental challenges and threats because the preservation of a person's self relies so heavily on his or her social surrounding. In human infancy, attachment behavior, which develops through reciprocal exchange with a caregiver, is essential for survival and normal psychological functioning. Later in childhood and adolescence, a person's social identity emerges from recurrent exchange with significant others. Moreover, in adult life, core social roles, such as the work role and the family and marital role, serve to link the individual person with a structured, goal-oriented social environment. Thus, important life domains that matter most for the individual person are inherently social in a cooperative or competitive sense.

Placed in this context, human health and well-being are closely related to successful self-regulation in a conducive social environment, and a stressful experience is likely to result from failed opportunities for successful agency, goal achievement, and reward experience in interpersonal exchange. The question then is what analytical concepts have been developed by the social and behavioral sciences to identify relevant socioenvironmental and personal triggers of recurrent stressful experience. The term psychosocial factor is used as a summary label to characterize the still-heterogeneous state of the art in this field. In other words, scientific progress has not yet resulted in a universally accepted taxonomy of stress-eliciting or stress-buffering social and personal (psychological) conditions.

This article presents a brief overview of some of the major concepts of identifying psychosocial influences

on stress and health. The two main procedures mediating this association are (1) enhanced neural/neuroendocrine/neuroimmune activation and (2) initiation of health-damaging behaviors.

There are multiple ways in which social and personal factors interact in influencing health and disease. It is useful to think of a stream of events where the most distal parts, the upstream factors, relate to the society at large or its main institutions (the macrosocial influences) and where the most proximal parts, the downstream factors, concern an individual's close or immediate social surroundings (the microsocial influences) and his or her personal ways of coping with the environment. In the following section, the discussion of psychosocial factors starts with the more distal parts of this stream and then proceeds to more proximal phenomena, including person characteristics. Finally, special attention is given to concepts that focus explicitly on the interaction between person and environment.

Overview of Psychosocial Factors Influencing Stress and Health

Macrosocial and Microsocial Factors

Every human society is characterized by a set of social values, norms, and institutions that are instrumental for the survival and growth of its members. If these values and norms lose their validity and meaning, or if society's patterns of social exchange become unpredictable and unstable, individuals tend to suffer from states of social anomie (lack of rules and orientations). A number of epidemiological investigations have explored the adverse effects on health produced by sociocultural instability, rapid social change, or a high level of social anomie. A majority of these studies found evidence of elevated risks of subsequent physical and mental illness.

One type of study was conducted in native societies that, after a long period of social stability, experienced rapid sociocultural transformation. As a consequence, less solidarity was witnessed among group members, and more deviant behavior, including violence, crime, and disruption of social ties, occurred. Equally distressing, the internalized cultural canon, which had been developed over many generations, broke down. Essential beliefs and ways of interpreting life were lost or fragmented. Under these conditions, cardiovascular risk factors, particularly hypertension, and coronary heart disease manifestation became more prevalent.

Similar observations were made by investigators who followed migrant populations from their original places to new sociocultural environments. Interestingly,

migrants who were able to maintain strong social ties with their traditional groups in the new environment were more likely to be protected against subsequent health risks compared to migrants who deprived themselves from these traditional bonds. This observation led to the conclusion that a stable social network and a particular quality of social exchange, termed social support, may act as a protective resource in coping with stressful circumstances. Social support is defined as the experience of, or access to, social relationships that offer mutual understanding and trust and that recurrently elicit positive emotions.

There is now convincing evidence that negative health effects result from social separation and social isolation, either due to loss of a loved one or due to poor opportunities for developing or maintaining relationships. Again, under these conditions, social support is less likely to be experienced and thus a strong psychosocial resource against stress is lacking.

While close social ties and social support denote downstream phenomena close to the intimate experiences made by individual persons, the more distal, macrostructural conditions of participating in the social fabric were shown to affect health as well. In particular, disinvestment in social capital was found to exert strong effects on morbidity and mortality at the level of whole populations. The term disinvestment in social capital describes a decline in reciprocal social exchange in neighborhood and community life, including a lack of solidarity and trust. For instance, when an area has a higher level of poverty, less investments are made into shared resources of community life, and hostile encounters or competitive struggles are more frequent.

Disinvestment in social capital is one of several health-adverse developments resulting from widening social inequalities. The demonstration of an inverse social gradient of morbidity and mortality within and between populations in a large number of advanced societies must be considered one of the most consistent and important findings of modern social epidemiology: the lower a person's socioeconomic status (in terms of educational attainment, income, and occupational standing), the higher his or her risk of poor health. These differentials are substantial. For instance, mean life expectancy in countries such as Finland or the United Kingdom is shortened by 5 to 6 years among members of the lowest as compared to the highest socioeconomic status group. It is important to note that these social differentials in health are spread across the whole of the society rather than being concentrated on a deprived group at the bottom of the society. So what are the causes of relative social deprivation in health?

Research repeatedly came to the conclusion that access to or quality of health care is not the main cause, as might be expected. Rather, unhealthy behaviors and exposure to stressful material and psychosocial environments, including lack of appropriate resources, must be considered the main determinants of social inequalities in health in advanced societies. These latter conditions include poor living and working conditions, threats to, or experience of, unemployment and income loss, and unstable social networks. Under these circumstances of relative social deprivation, feelings of anger, disapproval, and hopelessness are more frequent, intense, and long-lasting, thus aggravating the burden of stressful experience (see later).

Person Characteristics

Clearly, not every individual reacts to stressful environments in the same way. On the one hand, specific genetic or psychosocial factors may protect a person from adverse effects on health even under highly demanding circumstances. On the other hand, a person's threshold of susceptibility may be very low, thus predisposing him or her to illness onset under conditions of moderate stress. With the introduction of methods of molecular medicine into epidemiological studies, important advances in identifying people at elevated risk of disease were achieved. For instance, combining sociological information on stressful life events with genetic information on specific polymorphisms of the serotonin transporter receptor gene was shown to improve the prediction of early onset of depression in young adults.

One direction of research has focused on psychological traits that predispose an individual to spend high levels of energy. This pattern is functional if demands or threats are faced. However, in less demanding contexts, this may harm the organism by recurrent excessive physiological activation. Originally, researchers were interested in so-called type A behavior, a complex of behaviors and emotions defined by excessive striving, competitiveness, hostility, and impatience. Type A individuals were identified on the basis of specific psychomotor and speech characteristics, such as unrest, hectic, inability to relax, loud speech, and intrusive behaviors. As can be expected from a stress physiological perspective, excessive sympathoadrenergic arousal predisposes these individuals to cardiovascular dysfunction and disease.

Today, however, refined concepts following this direction of research are more successful in predicting future cardiovascular risk. One such concept concentrates on cognitive and affective components of hostility. The former include negative beliefs about

people and negative attitudes toward others, such as cynicism and mistrust, whereas the latter focuses on a frequent and intense experience of anger. A high level of hostility was found to promote health-adverse behaviors such as smoking and alcohol use in adolescence and young adulthood. Moreover, increased sympathetic function in combination with decreased parasympathetic function was observed. This observation could explain at least part of the direct statistical associations between hostility and coronary heart disease incidence that have been reported from several studies.

A similar approach emphasizes cognitive-motivational mechanisms of excessive energy spending. People characterized by overcommitment when facing demands tend to misjudge (e.g., underestimate) challenging stimuli and to expose themselves to multiple obligations. This may be due to their strong desire to control or dominate their environment, to exceed others, and to experience approval. For instance, workaholic behavior in occupational contexts can be interpreted in this context (see later). After a period of continued overcommitment, these individuals are susceptible to a state of vital exhaustion and psychophysiological breakdown, particularly when confronted with loss of control and failure. Overcommitment has been identified as a cardiovascular risk factor in epidemiological and experimental investigations.

A second line of research demonstrates adverse health effects of passive behaviors and associated low levels of energy spending. Importantly, inhibition or withdrawal of activity due to depressive mood and cognitions of helplessness and lack of self-esteem was found not only in reaction to but also preceding any experience of severe threat or loss. Some investigations conclude that individuals characterized by these psychological traits are at elevated risk of suffering from reduced immunocompetence and its pathophysiological consequences. Thus, over- and underactivation of a person's potential of vital energy in relation to self-regulating processes affect health at more fundamental, cellular and subcellular, levels, e.g., by increasing endogeneous oxidation and decreasing DNA repair capacity.

Although these behavioral patterns are influenced, to some extent, by genetic predisposition, evidence indicates that socioenvironmental factors have a strong impact on gene expression via methylation of DNA. Again, combining research on epigenetic processes with sociobehavioral research has promise for advancing the frontiers of knowledge.

In conclusion, distinct personal traits have been identified that must be considered psychological risk factors or protective factors. Protective factors

include a sense of mastery and self-efficacy (i.e., positive beliefs about one's agency) and favorable self-esteem, often paralleled by optimism and trust. Risk factors, as evidenced earlier, include overactive behaviors such as hostility, overcommitment, and excessive striving. These traits carry the risk of precipitating feelings of frustration and helplessness and triggering states of vital exhaustion. Moreover, risk factors include those cognitive, motivational, and affective traits that predispose individuals to withdraw from activity, to experience low levels of personal control and self-esteem, and to suffer from depressive mood and negative attributional style.

Person–Environment Interaction

Traditionally, the disciplines of medical sociology and social epidemiology have emphasized the role of socioenvironmental factors in health and disease, whereas medical psychology and psychosomatic medicine have stressed the importance of personality factors. Therefore, despite the obvious need for a more comprehensive analysis, few attempts were made to combine the two disciplinary approaches. Of these attempts, two developments deserve special attention.

First, in a period of declining impact of psychoanalytic thinking on etiological research, new interest emerged in understanding how socioenvironmental and personal factors interact in promoting health and triggering disease in a life course perspective. For instance, birth cohort studies demonstrate that adverse social circumstances of parents at the time of their children's birth increase the risk of premature mortality in children's later life. Explanations for this refer to elevated vulnerability during pregnancy and experience of cumulative cognitive, behavioral, and social disadvantage during childhood.

Among early psychological risk factors, a disturbed pattern of attachment between mother and child in infancy was shown to be of particular significance for later well-being. However, favorable social circumstances in childhood can compensate for these deficits to some extent by promoting children's resilience.

A second line of extensive research on how person and environment interact in producing stress-related disease has been concerned with the impact of working life on health. In contemporary societies, the work role still defines one of the crucial social roles in adulthood with powerful effects on successful or unsuccessful self-regulation. The adverse effects of job loss and related social isolation on health were mentioned earlier. Additional findings demonstrate that among low-status members of the workforce, stressful, health-damaging jobs in terms of a combination

of high demands and low control are more prevalent. In this context, the model of effort–reward imbalance at work is of interest, as it explicitly combines situational and person characteristics.

This model claims that lack of reciprocity between costs spent and gains obtained at work (high-cost/low-gain conditions) defines a state of emotional distress with special propensity to autonomic arousal and neuroendocrine stress response. Effort at work is spent as part of a socially organized exchange process to which society contributes in terms of rewards (money, esteem, and career opportunities). An imbalance between effort and reward is maintained if no alternative choice is available in the labor market. Moreover, people characterized by a high level of overcommitment and related cognitive distortion of demand appraisal (see earlier discussion) are equally prone to experience an effort–reward imbalance. Prospective epidemiological studies documented that stressful work in terms of high effort, low reward, and low control doubles the risk of incident fatal or nonfatal cardiovascular disease. In addition, onset of depression is more likely to occur.

Implications for Intervention

This article documents a solid body of research on the role of distinct psychosocial factors in triggering stress-related disease (risk factors) and in promoting health (protective factors). It highlights that both upstream (macrosocial) and downstream (microsocial and personal) factors are important in explaining elevated risk of morbidity and mortality. Finally, a combined approach focusing on person–environment interaction, including analyses of epigenetic processes, holds special promise for advancing scientific understanding and its consequences for prevention.

At all levels discussed so far, implications for intervention are obvious. At the macrostructural level, however, it is difficult to imagine how a reduction in the amount of relative social deprivation can be achieved. Policy measures such as limitations in income inequality, labor market regulations, and allocation of means to improve social capital can hardly be realized on the basis of anticipated health gain.

A more realistic approach toward intervention concerns the mesosocial level. For instance, several theory-based concepts of workplace health promotion have been implemented and evaluated. Reductions in blood pressure, heart rate variability, and levels of stress hormones and atherogenic lipids are examples of favorable health outcomes following task reorganization with enlarged decision latitude, improved social support, and improved leadership skills of superiors.

Similarly, more comprehensive intervention programs at the level of community, neighborhood, or school are being developed. Some of these programs pay special attention to sustained health gain in socially deprived groups or in high-risk groups. Examples are early-childhood family education programs, preschool programs, and adolescent health promotion activities. Improving successful parenting in socially deprived families was found to strengthen children's cognitive and emotional development.

Finally, an obvious level of intervention concerns the microsocial and personal level. Several behavioral intervention techniques have been developed and tested in both primary and secondary prevention. These include stress management, promotion of self-control and self-efficacy, personal empowerment, and an increase in social skills and coping effectiveness. These techniques proved to be particularly successful in reducing health-adverse behaviors such as unhealthy diet, smoking, alcohol, and drug consumption. In addition, the provision of social support and the formation of self-help groups are instrumental in reinforcing newly acquired attitudes, motivations, and behaviors. In summary, a rich potential of interpersonal and behavioral intervention approaches is now available that supplements traditional biomedical treatment techniques and that provides new opportunities for primary and secondary prevention. It is important to emphasize that this potential is based on scientific evidence obtained from pioneering research on psychosocial factors, stress, and health.

See Also the Following Articles

Chronic Social Stress: GR Sensitivity in Leukocytes; Hostility; Life Events and Health; Self-Esteem, Stress and Emotion; Social Capital; Social Networks and Social Isolation; Social Support; Type A Personality, Type B Personality.

Further Reading

- Antoniou, A. S. G. and Cooper, C. L. (eds.) (2005). *Research companion to organizational health psychology*. Cheltenham, UK: Edward Elgar.
- Bandura, A. (1997). *Self efficacy: the exercise of control*. New York: Freeman.
- Berkman, L. and Kawachi, I. (2000). *Social epidemiology*. Oxford, UK: Oxford University Press.
- Caspi, A., Sugden, K., Moffitt, T. E., et al. (2003). Influence of life stress on depression. Moderation by a polymorphism in the 5-HTT-gene. *Science* **301**, 386–389.
- Marmot, M. and Wilkinson, R. (eds.) (2005). *Social determinants of health*. Oxford, UK: Oxford University Press.
- Orth-Gomer, K. and Schneiderman, N. (eds.) (1996). *Behavioral medicine approaches to cardiovascular disease prevention*. Mahwah, NJ: Erlbaum.
- Siegrist, J. and Marmot, M. (eds.) (2006). *Social inequalities in health: new evidence and policy implications*. Oxford, UK: Oxford University Press.
- Turner, J. R., Cardon, L. R. and Hewitt, J. K. (eds.) (1995). *Behavior genetic approaches in behavioral medicine*. New York: Plenum Press.
- Weiner, H. (1992). *Perturbing the organism: the biology of stressful experience*. Chicago, IL: University of Chicago Press.

Psychosomatic Heart Disease: Role of Sympathetic and Sympathoadrenal Processes

G W Lambert

Baker Heart Research Institute, Melbourne, Victoria, Australia

T Dawood

Baker Heart Research Institute and Monash University, Alfred Hospital, Melbourne, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

Glossary

Neurogenic essential hypertension
Major depressive disorder

Panic disorder

Elevated blood pressure initiated and sustained by a characteristic elevation in sympathetic nervous activity.

A common and severe mental illness characterized by feelings of hopelessness, helplessness, and worthlessness; often accompanied by suicidal thoughts.

A distressing and disabling condition with patients experiencing recurrent episodes of unexpected intense anxiety of sudden onset.

Mechanisms Underlying Coronary Heart Disease and Myocardial Infarction and Possible Relationship to Stress and Behavior

Sympathoadrenal Function in Panic Disorder

Sympathoadrenal Function in Major Depressive Disorder

I was walking along the road with two of my friends. Then the sun set. Then the sky became a bloody red and I felt a tinge of melancholy, a sucking pain beneath my heart. I stopped, leaned against the railing, dead tired. Over the blue-black fiord and city hung blood and tongues of fire. My friends walked on and I stood again trembling with fright. And I felt as if a loud unending scream were piercing nature. – Edvard Munch

Mechanisms Underlying Coronary Heart Disease and Myocardial Infarction and Possible Relationship to Stress and Behavior

Over recent years, substantial advances have been made in delineating the causes of coronary heart disease. At a community level, the importance of high blood pressure, tobacco smoking, and abnormal blood lipids as causal factors is well established. Recent epidemiological studies also point to an association between psychosocial stressors and increased risk of development of acute myocardial infarction. Mechanistically, the relevant issues linking psychosocial stress and acute cardiac events are whether stress, in one of its numerous forms, can (1) lead to the development of hypertension, (2) cause or aggravate atherosclerosis, and/or (3) trigger a cardiac event in the presence of existing atherosclerosis.

Relation of Stress to the Development of Hypertension

The importance of hypertension in the development of left ventricular hypertrophy, myocardial infarction, heart failure, and sudden death is well established. Indeed, over 90% of patients with heart failure in the Framingham study had a history of hypertension. The longitudinal studies of Timio and colleagues, examining hypertension development and cardiovascular morbidity and mortality in cloistered nuns, provide persuasive evidence linking the stress of daily life to the development of high blood pressure. Essential hypertension is commonly neurogenic, initiated and sustained by the sympathetic nervous system. While activation of the sympathetic nervous system outflows to the kidneys, heart, and skeletal muscle in hypertensive patients has been well documented, a second neural mechanism, reduced reuptake of the sympathetic neurotransmitter, norepinephrine, is also operative. Reuptake of norepinephrine into sympathetic nerves after its release terminates the neural signal. A fault in transmitter inactivation augments the effects of sympathetic nerve traffic. For the sympathetic nerves of the heart, approximately 95% of released norepinephrine is recaptured, so that the heart

is more sensitive than all other organs to impediments in transmitter reuptake. Other cardinal signs of stress common to patients with neurogenic essential hypertension include activation of brain noradrenergic pathways and epinephrine release from the heart.

Relation of Stress to the Development of Atherosclerosis

There exist numerous population studies indicating that psychological disturbances, particularly anxiety and chronic stress, can contribute to atherosclerosis development. In the workplace, insufficient self-regulation of workload, deadlines, and planning and direction of the work have been linked to risk of developing atherosclerosis. Low socioeconomic status or employment level is associated with increased plasma fibrinogen and von Willebrand factor concentrations. Endothelial function is considerably impaired following acute experimental stress.

Relation of Stress to the Triggering of a Cardiac Event

Cardiac sympathetic nerves are preferentially activated by mental stress It is important to realize that sympathetic nervous activity at rest and in response to stress is regionalized. For instance, with even relatively mild experimental mental stress, such as that induced in the laboratory by performing difficult mental arithmetic, the sympathetic nerves of the heart are markedly and preferentially activated, whereas activation of muscle vasoconstrictor sympathetic fibers predominates during the cold pressor test. Mental stress in patients with existing coronary artery disease has been demonstrated to cause anginal chest pain and inadequate blood supply to the heart. The effect of mental stress on cardiac sympathetic nervous activation is greater in the elderly. The importance of cardiac sympathetic activation in these contexts is highlighted by the demonstration, in experimental animals with narrowing of the coronary arteries, that direct electrical stimulation of the cardiac sympathetics is capable of causing electrical instability of the myocardium, triggering disturbances of heart rhythm and cardiac arrest.

General relation between activity of the sympathetic nerves of the heart and sudden death In a range of clinical contexts, stimulation of the cardiac sympathetic outflow has been demonstrated to contribute to myocardial infarction, ventricular arrhythmias, and sudden death. Activation of the sympathetic outflow to the heart is common in patients unexpectedly developing ventricular tachycardia and ventricular fibrillation. Furthermore, myocardial stunning due

to sudden emotional stress is associated with exaggerated sympathetic stimulation. Similarly, in patients with heart failure, there is a high level of stimulation of the cardiac sympathetic nerves, which is a proven determinant of sudden death. β -adrenergic blockers have been recently shown to increase survival in heart failure patients.

Direct relation between mental stress and sudden death in certain circumstances In a rare inherited heart condition, long QT interval syndrome, there is electrical instability of the heart muscle. Mental stress is one proven immediate cause of cardiac arrest in sufferers. Some research linking mental stress to sudden death is disputed because of disagreement over what constitutes a stress, and whether stress can be accurately measured. Recent research showing that rates of sudden, nontraumatic death in people with underlying coronary disease were markedly increased during catastrophic events is free of this criticism, as no finessing is needed in the psychological measurement of stress.

Sympathoadrenal Function in Panic Disorder

Debilitating episodes of recurring, often inexplicable anxiety afflict as many as 12% of the population. These attacks typically are unpleasant and may be accompanied by physical symptoms such as sweating, palpitations, tremor, and a sensation of suffocation. Given the prominent signs of sympathetic activation that occur during a panic attack, panic disorder patients find themselves at the nexus between cardiovascular and psychiatric medicine. It has until recently been felt that although panic disorder is distressing and disabling, it does not constitute a risk to life. Sufferers often fear that they have heart disease because of the nature of their symptoms, but have been reassured that this is not the case. Epidemiological studies, however, indicate that there is an increased risk of myocardial infarction and sudden death in patients with anxiety disorders. This increased risk extends even to premenopausal women, who in general have low coronary risk.

Somatic Sensitization as a Panicogenic Factor

Patients with panic disorder often exhibit hypervigilance to somatic sensations such as an irregular heart-beat or lightheadedness. Given the convergence of neural networks involved in processing viscerosensory and contextual information and their links with brain regions pivotal in the fear network, it is possible

that somatic sensation may be capable of triggering panic attacks. Such a concept is not without precedent. It has been known for some time that the presence of abnormalities of heart rhythm, supraventricular tachycardia in particular, can pre-date and in fact cause panic disorder. More recently, a very high rate of development of panic disorder was observed in patients with implanted cardiac defibrillators. Indeed, factors at play here seem to include the presence of cardiac symptoms and the enhanced vigilance that accompanies them.

Sympathetic and Adrenomedullary Function in Panic Disorder

Resting whole body, cardiac, and muscle vasoconstrictor sympathetic activity is normal in patients with panic disorder. While adrenal medullary secretion of epinephrine, measured by isotope dilution, is typically normal in these patients, there occurs a demonstrable release of epinephrine from the heart. It is probable that during the surges of epinephrine secretion that accompany panic attacks the cardiac sympathetic neuronal vesicles become loaded with epinephrine that is continuously co-released with norepinephrine in the interim periods between attacks. *In situ* synthesis of epinephrine, through activation within the heart of phenylethanolamine-n-methyltransferase (PNMT), perhaps induced by repeated cortisol responses during panic attacks, may also occur. In experimental models of stress, PNMT activation in the heart and other extra-adrenal organs has been demonstrated.

Norepinephrine Reuptake in Panic Disorder

As mentioned previously, reuptake of norepinephrine into sympathetic nerves after its release terminates the sympathetic neural signal. A phenotypic defect in norepinephrine transporter function has been described in patients with panic disorder. Such an abnormality in neuronal norepinephrine reuptake could sensitize the heart to sympathetic activation. The mechanism underpinning the defect in norepinephrine transporter function, and whether treatment modifies the activity of the norepinephrine transporter, remains unknown.

Sympathetic and Adrenomedullary Function during a Panic Attack

During panic attacks triggered cardiac arrhythmias, recurrent emergency room attendances with anginal chest pain and ECG changes of ischemia, coronary artery spasm in attacks occurring during coronary

angiography, and myocardial infarction associated with coronary spasm and thrombosis have been documented. The recurrent episodes of unexpected intense anxiety that characterize a panic attack are accompanied by a pronounced activation of the sympathetic nervous system and adrenomedullary secretion of epinephrine. A concomitant elevation in blood pressure and heart rate also occurs. The characteristic chest pain that occurs during an attack is additionally accompanied by the release of neuropeptide Y into the coronary sinus. Neuropeptide Y is a potent vasoconstrictor with the capacity to cause coronary artery spasm.

Sympathoadrenal Function in Major Depressive Disorder

Central Nervous System Neuronal Activity

The etiology of major depressive disorder (MDD) has been linked, variously, to brain monoaminergic neuronal dysfunction, alterations in monoamine receptor sensitivity, and stress-induced activation of the hypothalamic-pituitary-adrenal axis. The lack of demonstrated relationships between different clinical presentations and biochemical abnormalities has hampered the development of sensitive biochemical markers for MDD. A major difficulty in commonly used peripheral indices of a neurotransmitter (or neurotrophic factor) as an indicator of brain dysfunction is that these compounds are also produced outside the brain. For this reason, venous blood is directly sampled from the brain to study central nervous system (CNS) neurotransmitter turnover, using central venous catheters placed high in an internal jugular vein. Under these circumstances, arteriovenous differences in neurotransmitter concentrations can be extrapolated to estimates of brain production. The observation of global (i.e., noradrenergic, dopaminergic and serotonergic) reduction in both subcortical monoamine turnover and cerebral glucose utilization in patients with MDD is consistent with a reduction in CNS neurotrophic support in patients with MDD.

Depression as a Risk Factor for the Development of Cardiac Disease

There is strong evidence that patients with MDD are at increased risk of developing coronary heart disease. This elevated risk is independent of classical risk factors such as smoking, obesity, hypercholesterolemia, diabetes, and hypertension. The association is present in males and females, across different age groups and in subjects living in different countries. The risk of coronary heart disease is increased 1.5- to

2-fold in those with minor/subsyndromic depression and 3- to 4.5-fold in subjects with major depression. To put this into perspective, the strength of the association between coronary heart disease and MDD is similar to that with smoking and hypercholesterolemia. Also conclusively demonstrated is the adverse effect of depression in patients with heart failure and in patients following myocardial infarction, which materially increases mortality.

Sympathoadrenal Activity in MDD

While the mechanism of increased cardiac risk attributable to MDD is at present uncertain, activation of the sympathetic nervous system may be of prime importance. Chronic stress has been linked with MDD. Examination of indices of sympathetic nervous function in MDD has yielded conflicting results. Some studies indicate a tendency for both the urinary excretion and cerebrospinal fluid level of norepinephrine and its metabolites to be diminished, whereas other reports document elevated plasma levels of norepinephrine and increased rates of norepinephrine spillover to plasma in patients with MDD. It is likely that sympathoadrenal activation is present only in a subset of patients with MDD. Indeed, Gold and colleagues recently demonstrated elevated norepinephrine levels in both the CSF and plasma of patients with melancholic depression.

In addition to possible effects on blood pressure and myocardial stability, sympathetically mediated neural vasoconstriction may exert metabolic effects in skeletal muscle, impairing glucose delivery to muscle and causing insulin resistance and hyperinsulinemia, and in the liver, retarding postprandial clearing of lipids and contributing to hyperlipidemia. A trophic effect of sympathetic activation on cardiovascular growth, contributing to the development of left ventricular hypertrophy, may also occur. Left ventricular hypertrophy is an independent risk factor for cardiovascular morbidity and mortality.

Further knowledge of the mechanisms responsible for generating cardiac risk may pave the way for novel and perhaps relatively simple therapeutic strategies (e.g., low-dose aspirin, β -blockers) to be administered to those with MDD in order to modify cardiac risk. Indeed, defining therapeutic interventions that reduce cardiac risk will be an important step forward in alleviating the burden of depressive illness on the community.

See Also the Following Articles

Antisocial Disorders; Depression and Coronary Heart Disease; Hypertension.

Further Reading

- Bunker, S. J., Colquhoun, D. M., Esler, M. D., et al. (2003). "Stress" and coronary heart disease: psychosocial risk factors. *Medical Journal of Australia* 178, 272–276.
- Esler, M., Jennings, G., Lambert, G., et al. (1990). Overflow of catecholamine neurotransmitters to the circulation: source, fate, and functions. *Physiological Reviews* 70, 963–985.
- Gold, P. W., Wong, M. L., Goldstein, D. S., et al. (2005). Cardiac implications of increased arterial entry and reversible 24-h central and peripheral norepinephrine levels in melancholia. *Proceedings of the National Academy of Sciences USA* 102, 8303–8308.
- Lambert, G., Johansson, M., Agren, H., et al. (2000). Reduced brain norepinephrine and dopamine release in treatment-refractory depressive illness: evidence in support of the catecholamine hypothesis of mood disorders. *Archives of General Psychiatry* 57, 787–793.
- Leor, J., Poole, W. K. and Kloner, R. A. (1996). Sudden cardiac death triggered by an earthquake. *New England Journal of Medicine* 334, 413–419.
- Matsuo, T., Suzuki, S., Kodama, K., et al. (1998). Hemostatic activation and cardiac events after the 1995 Hanshin-Awaji earthquake. *International Journal of Hematology* 67, 123–129.
- Musselman, D. L., Evans, D. L. and Nemeroff, C. B. (1998). The relationship of depression to cardiovascular disease: epidemiology, biology, and treatment. *Archives of General Psychiatry* 55, 580–592.
- Rosengren, A., Hawken, S., Ounpuu, S., et al. (2004). Association of psychosocial risk factors with risk of acute myocardial infarction in 11119 cases and 13648 controls from 52 countries (the INTERHEART study): case-control study. *Lancet* 364, 953–962.
- Schlaich, M. P., Lambert, E., Kaye, D. M., et al. (2004). Sympathetic augmentation in hypertension: role of nerve firing, norepinephrine reuptake, and angiotensin neuro-modulation. *Hypertension* 43, 169–175.
- Spieker, L. E., Hurlimann, D., Ruschitzka, F., et al. (2002). Mental stress induces prolonged endothelial dysfunction via endothelin-A receptors. *Circulation* 105, 2817–2820.
- Timio, M., Saronio, P., Venanzi, S., et al. (1999). Blood pressure in nuns in a secluded order: a 30-year follow-up. *Mineral and Electrolyte Metabolism* 25, 73–79.
- Wilkinson, D. J., Thompson, J. M., Lambert, G. W., et al. (1998). Sympathetic activity in patients with panic disorder at rest, under laboratory mental stress, and during panic attacks. *Archives of General Psychiatry* 55, 511–520.

Relevant Website

www.beyondblue.org.au *beyondblue* is a national, independent, not-for-profit organisation working to address issues associated with depression, anxiety and related substance misuse disorders in Australia.

Psychosomatic Medicine

T Theorell

Karolinska Institutet, Stockholm, Sweden

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by T Theorell, volume 3, pp 304–309, © 2000, Elsevier Inc.

Definitions

Link with Stress Physiology and Epidemiology

Social Environment

Gene Activation

Psychosomatic Attitude

Difference between Behavioral Medicine and Psychosomatic Medicine

Brain Anatomy and Physiology

Studies of Life Changes

Coping and Behavior Patterns

Psychosomatic Treatment

Liaison Psychiatry

Final Comments

Glossary

<i>Alexithymia</i>	Not reading emotions; an inability to identify and discriminate between specific emotions.
<i>Amygdala</i>	Paired almond-shaped nuclei that are part of the limbic system of the brain and are concerned with fear processing.
<i>Anabolism</i>	Rebuilding and repairing of tissues.
<i>Cushing's syndrome</i>	Hyperfunction of the adrenal cortex.
<i>Gastric mucosa</i>	The lining of the stomach wall.
<i>Hippocampus</i>	A paired seahorse-shaped structure that is part of the limbic system of the brain concerned with memory.
<i>Hypothalamic-pituitary-adrenocortical (HPA) system</i>	One of the two main stress response systems of the brain, which subserves the neural (neuroendocrine) control of glucocorticoid secretion by the adrenal glands. Among other actions, the glucocorticoids facilitate energy mobilization.

Definitions

When the concept of psychosomatic medicine was introduced in science it was founded in a medical discipline: psychiatry. Medical conditions (i.e., diagnoses) that were characterized by objective changes and that could be induced by psychological states were defined as psychosomatic. This early definition of psychosomatic included duodenal ulcer, ulcerative colitis, bronchial asthma, essential hypertension, rheumatoid arthritis, thyrotoxicosis, and neurodermatitis. Later definitions have been more focused on the individual's total state. According to such a definition, a psychosomatic condition is a state with somatic and psychological symptoms. The direction of causation is often difficult to establish. Thus, there may be psychological processes influencing somatic mechanisms as well as somatic phenomena that influence psychological states. The latter phenomenon has also been labeled somatopsychic.

Two early authors who were using the old way of defining psychosomatic were Franz Alexander (who was a psychoanalyst and a pupil of Sigmund Freud) and Flanders Dunbar. An important element in the efforts to diagnose conditions in the early days was the specificity concept. Specificity was used to label specific psychological processes that would be associated with specific psychosomatic diagnoses, e.g., a repressed cry in relation to an acute attack of bronchial asthma or an oral frustration during early childhood in relation to duodenal ulcer development in adult years.

A famous example of organic conditions that could be induced by psychological states was observed in the patient Tom, who had a gastric fistula. The fistula made the gastric mucosa available to the eyes of the researchers, which enabled the internists Harold Wolff and Stewart Wolf to conduct psychological experiments: they could study Tom's gastric mucosa while different states were induced in him. When long-lasting states of anger were induced, small ulcers arose, which were associated with elevated acid production, increased motility, and reddening of the mucosa. The opposite condition, a pale and passive mucosa with very little production of acid, was observed when Tom was in a state of long-lasting withdrawn and passive sadness.

Recently, discussion about peptic ulcer etiology has focused on *Helicobacter pylori*, a bacterium that is present in the gastric mucosa of most, but not all, patients with peptic ulceration. *Helicobacter pylori* seems to be necessary for the development of most peptic ulcers. Nevertheless, stress-induced variations in production of acid could increase the likelihood of ulcer development, especially in critical care

situations. That is, the effects of emotional or stress-induced changes in gastric acid production could interact with *Helicobacter pylori* to cause gastric ulceration in some individuals. Robin Warren and Barry Marshall were awarded the 2005 Nobel Prize for the discovery of *Helicobacter pylori*.

In the more modern and wider definition of psychosomatic, the attitude toward the patient's illness is central. If any condition could be more or less psychosomatic, all caregivers should ask themselves whether there are aspects of their patient's condition that could be psychosomatic.

Link with Stress Physiology and Epidemiology

The influence of stress physiology on theoretical modeling in psychosomatic medicine became evident when internists became interested in the field. When psychiatrists collaborated with physiologists and internists, some of the basic mechanisms were revealed, explaining why allergic reactions, wide swings in diabetes mellitus, worsening of hypertension, and many other conditions could be partly psychosomatic.

The next step in the scientific development of the field was taken by epidemiologists, who made it possible to describe psychosomatic conditions in persons outside the doctor's office: the total illness panorama. In this way it became possible to identify those who suffer from psychosomatic conditions but do not consult a caregiver.

Studies of extreme stress were initially an important part of the psychosomatic field, and this is also where the link between stress and psychosomatic medicine was established. In Style's general adaptation syndrome (GAS), stress was seen as the general reaction to a nonspecific challenge or adverse condition that was labeled a stressor (a factor that induces stress). According to this definition, stress is a reaction, and it is a general one. It is evident that this reaction induces arousal, which means that energy is mobilized. There are several parallel phenomena, all of which aid the body in the physical fight-or-flight response. Examples include lowered excretion of water and salt, decreased sensitivity to pain, and decreased inflammatory response in infections. Because energy mobilization (resulting in an elevated blood concentration of glucose and free fatty acid) has the highest priority, anabolism is downregulated. Anabolism is central to the body's central defense of all the organ systems that need constant rebuilding and restoration. If this goes on for a long time (e.g., several months), increased sensitivity to physical and psychological stress in bodily organs is the ultimate result.

This represents a more general theory in contrast to the specificity theory mentioned earlier. Another consequence of long-lasting demands for energy mobilization is that endocrine systems may change their regulatory patterns. One consequence could be that the production of thyroid hormones is stimulated. Thyroid hormones stimulate a long-lasting upregulation of the metabolic rate.

Disturbed regulation of the hypothalamic-pituitary-adrenocortical system (the HPA axis) has also been described in relation to long-lasting arousal. There seem to be at least three kinds of regulatory disturbances in the HPA axis:

1. Lack of inhibition of stimulation. Normally, cortisol inhibits higher centers in the HPA axis when it has been elevated for a long time. When this inhibition does not take place in the normal way, the levels tend to be elevated constantly.

2. Lack of response to stimulation of the HPA axis. This means that plasma cortisol levels are low and that the normal activation does not take place when the HPA axis is stimulated artificially.

3. Unstable levels. In this kind of disturbance, the plasma cortisol levels may be low at rest, but certain situations (that are specific to the individual) trigger a strong HPA axis reaction that may seem to be out of proportion to the importance of the trigger.

The first kind of regulatory disorder is typical of Cushing's syndrome and could be one alternative regulatory disturbance in psychiatric depression. Severe long-lasting psychiatric depression may be associated with constantly high cortisol levels; in most cases, however, the circadian rhythm is preserved.

The second kind of disorder has been shown to exist in a subgroup of patients diagnosed with chronic fatigue syndrome (CFS). Because at least one other group of patients, those with fibromyalgia, share some symptoms that are typical of this regulatory disorder with CFS patients, it is possible that part of the myalgia problem is that they also suffer from this disturbance. It should be pointed out, however, that CFS and fibromyalgia also have differences in symptomatology and that it is not known to what extent psychosomatic processes contribute to them.

The third kind of regulatory disorder is found in posttraumatic stress disorder (PTSD). The key element in the etiology of PTSD seems to be that the long-lasting extreme arousal level increases the number of cortisol receptors of the cell surfaces and also increases their sensitivity. Because the sensitivity is so high, only low concentrations of cortisol are needed to keep the system working at rest. This may be the reason why low serum cortisol levels are found at rest in patients with PTSD.

Disturbances in endocrine functions represent a more general theory that has been of central importance to psychosomatic medicine. Physiology has also contributed to the specificity theory, however. A recent contribution to cardiovascular psychosomatic medicine has been the hostility hypothesis, which states that subjects with a high degree of hostility and cynicism are more likely to develop a myocardial infarction at a relatively young age compared to other subjects. Hostility has been linked to disturbances in the production and blood concentration of serotonin.

Psychosocial processes are mostly not capable of inducing organic illnesses by themselves. There are other pathogenic processes involved that may have a genetic origin. One example that has been discussed recently is the role of the immune system in the etiology of thyrotoxicosis. Excessive production of some of the immune system's interleukines causes abnormal regulation of function of the thyroid gland, resulting in either thyrotoxicosis or hypothyroidism. Still, long-lasting life changes do have a role in facilitating the onset of thyrotoxicosis.

Social Environment

James P. Henry was one of the most active physiologists taking the position that the social environment has to be taken into account in the study of psychosomatic medicine. He represented a group of researchers who increased the understanding of psychosomatic medicine by taking into account both studies of colonies of animals whose social environments were manipulated and natural international variations across the human culture. One important finding was that blood pressure does not rise with increasing age in all cultures.

Gene Activation

Modern psychosomatic research has been influenced by the rapidly growing amount of knowledge about genes and gene activation. Genes that are of particular relevance to psychosomatic medicine are those that regulate monoamines such as thyroxine, noradrenalin, serotonin, and dopamine. This research explores mechanisms and importance of interactions between the environment and genes in the psychogenesis, for instance, of depression. An important landmark in these studies has been the prospective birth cohort study in New Zealand published by Caspi et al. A total of 1037 children were assessed every second year from age 3 up to age 15, every third year to age 21, and finally at the age of 26. Stressful life events occurring after the 21st and before the

26th birthday were assessed by means of a life history calendar. The risk of onset of depression after critical life events was shown to be related to a variant ("polymorphism") in the promoter (control) region of the serotonin transporter gene. One type of allele combination in this region is protective while another increases risk. These studies point at new directions in psychosomatic research.

Psychosomatic Attitude

The modern definition of psychosomatic is related to a treatment attitude toward the patients. The psychosomatic therapist is expected to consider social and psychological factors that may be of importance to the course of the patient's illness and to do this in all cases encountered. This of course means that any condition could be partly or totally psychosomatic. Such awareness should affect all medical treatment.

Difference between Behavioral Medicine and Psychosomatic Medicine

Behavioral medicine and psychosomatic medicine are similar to one another but are not identical. Behavioral medicine deals primarily with phenomena that are possible to record objectively, whereas psychosomatic medicine deals with intrapsychic phenomena as well. Psychosomatic medicine has been introduced mainly by psychiatrists and physiologists (biologists and medical doctors), whereas behavioral medicine has been introduced mainly by behavioral psychologists.

The therapeutic orientation has also been different in the two disciplines: whereas psychodynamically oriented treatments have dominated psychosomatic medicine, behavioral and cognitive psychotherapy has dominated behavioral medicine.

Brain Anatomy and Physiology

A major discovery in brain neurophysiology has changed the basic background of medicine. This is the fact that structural changes in the nerve cells of the hippocampus of the brain may occur as a consequence of traumatic experiences. Such discoveries have been made possible by new techniques for imaging the brain, building on principles such as magnetic resonance and positron emission. According to several studies, the volume of the hippocampus, a structure that is of great importance for the storage of basic memories, may diminish as a consequence of severe traumatic events that cause long-lasting changes in the concentration of several substances in the brain, with resulting PTSD.

Other neuroanatomical studies have shown that the interaction between emotions and cognition

could be localized as traffic between specific structures such as the amygdala (representing nonspecific arousal and anxiety) and hippocampus (memory). Studies have shown that subliminal visual stimulation (pictures presented so rapidly that the person does not become aware of what they represent) may induce activity in the limbic system, e.g., the amygdala. It is known that there are direct connections between the visual cortex and the limbic system, and the cortex may therefore be activated at a later stage than the limbic emotional system. Similar connections exist for auditory stimuli. Such knowledge is of great importance in explaining the fact that sub-conscious processes may induce strong psychosomatic reactions.

Studies of Life Changes

The basic theory behind studies of life changes in psychosomatic research was Selye's nonspecific stress theory. It was later formulated by Holmes and Rahe as a theory of demands for adaptation to any change. A system of life change units was constructed on the basis of average scores from individuals in the normal population on demands for adaptation to specific life changes such as death of spouse, loss of job, and residential move. This line of research has shown that a group of subjects with an accumulation of important life changes during a short period of time will have an increased number of illness episodes during the subsequent period. This association holds on a group level but it is quantitatively relatively weak. It is when coping patterns, and also the perception of the life change by the subjects themselves, are taken into account that predictions get stronger. During later years, studies have shown that for certain types of illnesses and physiological reactions, a distinction could be made between positive and negative life changes – with positive changes in some cases inducing improvement and negative changes inducing deterioration in health. Studies of life changes have been important to the practice of psychosomatic medicine. Particularly in chronic illness rehabilitation, it is very useful to make systemic explorations of ongoing life changes. The patients immediately understand the value of this method, and the exploration may lead to increased insight, which could be very useful. Discussions regarding life changes that the patient has experienced with a concomitant recording of physiological functions may strengthen the insight further.

A sociological school of life change research, represented by Brown and others, has concentrated on the objective aspect of life changes. Special schemes for recording objective aspects of life changes are being

used. By analyzing the structural meaning (not the perception) of a life change, these authors have shown that there is an independent contribution of the change itself to the risk of illness.

Coping and Behavior Patterns

The way in which the individual perceives his or her life situation is of profound importance to his or her health. A basic concept in psychosomatic medicine is alexithymia, an inability to identify and discriminate between specific emotions. According to the theory introduced by Sifneos, primary alexithymia arises when emotions have not had the possibility to develop in the normal way during childhood, with resulting partial or total inability to differentiate specific feelings such as joy, anxiety, sadness, and rage. Secondary alexithymia, however, may arise during major life crises such as traumatic experiences or severe illness episodes. Patients with pronounced alexithymia may be very difficult to treat with verbal psychotherapy, because they do not have access to effective interpretations of what happens in their bodies in relation to emotionally arousing situations. Educational programs are presently being tested. A related concept is *pensée opératoire*, operative thinking, and a propensity to describe practical acts rather than emotions, e.g., in crisis situations.

Another aspect of coping that has been of particular importance to psychosomatic medicine is type A behavior. It was originally described as a behavior pattern that was elicited in certain situations. The cardiologists Friedman and Rosenman had seen it in younger male patients with coronary heart disease. In early stages of the scientific development it was discovered that there were several distinct components in this behavior, such as excessive drive and ambition (to manage difficulties better than others against the worst possible obstacles), hostility/feeling of rush, and excessive work orientation. Later researchers have criticized the definition of general type A behavior and stated that cynical hostility may be the key component. Evaluations of type A behavior intervention programs in secondary coronary prevention have shown that successful long-term programs (mostly lasting for at least a year) for reducing type A behavior may reduce the risk of developing new episodes of coronary heart disease.

Psychosomatic Treatment

Psychosomatic treatment follows different traditions in different countries. In several countries, e.g., Germany, there are several academically based clinics

for psychosomatic treatment, mostly founded in psychoanalytical theory. In other countries, e.g., Sweden, it is believed that psychosomatic treatment should be part of good patient treatment on the whole, and therefore very few psychosomatic departments have been founded.

One of the key elements in psychosomatic medicine, from the time of Freud and his followers to today, has been alexithymia (see earlier discussion). Patients with long-lasting severe psychosomatic disturbances frequently suffer from early (childhood) disturbance of the emotional development. Chronic pain and chronic gastrointestinal functional disorders are examples of conditions that have been shown to be associated with traumatic childhood circumstances. It has been found that traditional verbal psychotherapy may be very difficult in such cases. Alternatives have been tried that could surpass verbal barriers. Such therapies could build either on direct bodily sensations or on indirect ways of helping the patients gain useful insight into emotion-body relationships. One such alternative is the use of arts that could be used either impressively or expressively. An example of the impressive use of music is guided imagery and music (GIM). GIM has been shown to have not only temporary but also long-term effects on endocrine systems. In an uncontrolled study of the use of dance, music, painting, or drama of patients with long-lasting (2 years in most cases) psychosomatic disorders, it was shown that the use of expressive arts can be helpful in gaining useful insight into the interplay between emotions and bodily symptoms. Experiences of art are often less guarded than verbal communications. In this study, the patients went through a catharsis phase of the treatment, mostly during the period 6 to 12 months after the start, in which strong emotions had been evoked and there was evidence of general physiological arousal. The art experiences were often helpful in creating turning points in the therapy.

Liaison Psychiatry

For a long time there has been a strong link between psychosomatic medicine and liaison psychiatry. In liaison psychiatry, specially trained psychiatrists work continuously in somatic medicine and have a special responsibility for patients with psychosomatic conditions. Liaison psychiatry includes areas in which psychological processes are of particular importance to healing of wounds and acceptance of transplants. The latter two areas are of particular importance to surgical wards. Studies during later years have shown that excessive long-lasting mobilization of energy prolongs the healing of wounds.

Final Comments

Psychosomatic medicine has developed rapidly since the introduction of the discipline. Specific theories along with general theories have been tested. In modern psychosomatic medicine there is growing awareness that the social environment is of great importance and also that the etiological processes are very complex. Interactions between physical, chemical, and biological agents (e.g., infections, such as *Helicobacter pylori*) and psychosocial processes cause long-lasting psychosomatic conditions.

See Also the Following Articles

Arthritis; Chronic Fatigue Syndrome; Hippocampus, Overview; Life Events and Health; Psychological Stressors, Overview; Type A Personality, Type B Personality; Ulceration, Gastric.

Further Reading

- Brown, G. W. and Harris, T. (1978). *Social origins of depression: a study of psychiatric disorders in women*. London: Tavistock.
- Caspi, A., Sugden, K., Moffitt, T. E., et al. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 301, 386–389.
- Charney, D. S., Deutch, A. Y., Krystal, J. H., Southwick, S. M. and David, M. (1995). Psychobiological mechanism of posttraumatic stress disorder. *Archives of General Psychiatry* 50, 394–430.
- Dunbar, F. (1954). *Emotions and bodily changes: a survey of literature on psychosomatic interrelationships 1910–1953*. New York: Columbia University Press.
- Friedman, M. and Rosenman, R. H. (1959). Association of specific overt behavior pattern with blood and cardiovascular findings. *Journal of the American Medical Association* 169, 1286–1293.
- Friedman, M., Thoresen, C. E., Gill, J. J., et al. (1986). Alteration of type A behavior and its effect on cardiac recurrences in post myocardial infarction patients: summary results of the recurrent coronary prevention project. *American Heart Journal* 112, 653–655.
- Henry, J. P. and Stephens, P. M. (1977). *Stress, health and the social environment: a sociobiological approach to medicine*. New York: Springer.
- Holmes, T. H. and Rahe, R. H. (1967). The social readjustment rating scale. *Journal of Psychosomatic Research* 11, 213–218.
- Jones, M. P. (2006). The role of psychosocial factors in peptic ulcer disease: beyond *Helicobacter pylori* and NSAIDs. *Journal of Psychosomatic Research* 60, 407–412.
- Kiecolt-Glaser, J. K., Page, G. G., Marucha, P. T., MacCallum, R. C. and Glaser, R. (1998). Psychological influences on surgical recovery: perspectives from psychoneuroimmunology. *American Psychology* 53(11), 1209–1218.
- Le Doux, J. (1999). *The emotional brain*. London: Phoenix.
- Marty, P. and de M'Uzan, M. (1973). La pansée opératoire. *Revue Française de Psychoanalyse* 27, 345–346.
- Marucha, P. T., Kiecolt-Glaser, J. K. and Favagehi, M. (1998). Mucosal wound healing is impaired by examination stress. *Psychosomatic Medicine* 60(3), 362–365.
- McKinney, C. H., Antoni, M. H., Kumar, F. C. and McCabe, P. M. (1997). Effects of guided imagery and music (GIM) therapy on mood and cortisol in healthy adults. *Health and Psychology* 16, 390–400.
- Nemiah, J. C. (1996). Alexithymia: present, past and future. *Psychosomatic Medicine* 58, 217–218.
- Rubin, R. T., Poland, R. E., Lesser, I. M., Winstone, R. A. and Blodgett, A. L. (1987). Neuroendocrine aspects of primary endogenous depression. I. Cortisol secretory dynamics in patients and matched controls. *Archives of General Psychiatry* 44, 328–336.
- Selye, H. (1950). *The physiology and pathology of exposure to stress*. Montreal, Canada: Acta Inc.
- Sifneos, H. (1967). Clinical observations on some patients suffering from a variety of psychosomatic diseases. In: *Proceedings of the Seventh European Conference on Psychosomatic Research*. Basel: Karger.
- Stollman, N. and Metz, D. C. (2005). Pathophysiology and prophylaxis of stress ulcer in intensive care unit patients. *Journal of Critical Care* 20, 35–45.
- Taylor, G. J. and Bagby, R. M. (2004). New trends in alexithymia research. *Psychotherapy and Psychosomatics* 73, 68–77.
- Theorell, T., Konarski, K., Westerlund, H., et al. (1998). Treatment of patients with chronic somatic symptoms by means of art psychotherapy: a process description. *Psychotherapy and Psychosomatics* 67, 50–56.
- Williams, R. B. (1996). Coronary-prone behaviors, hostility and cardiovascular health: implications for behavioral and pharmaceutical interventions. In: Orth-Gomér, K. & Schneiderman, N. (eds.) *Behavioral medicine approaches to cardiovascular disease prevention*. Mahwah, NJ: Erlbaum.

Psychotherapy

G C Smith

Monash University, Clayton, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G C Smith, volume 3, pp 310–315, © 2000, Elsevier Inc.

Introduction

Psychotherapy as an Interpersonal Process

Modification of Feelings, Cognitions, Attitudes, and Behavior

Troublesomeness of Phenomena

The Person Seeks Help

Help Is Sought from a Trained Professional

What Is the Difference between Psychotherapy and Education?

Stress Management

The Schools of Psychotherapy

Methods of Delivery of Psychotherapy

The Efficacy of Psychotherapy

Glossary

<i>Counseling</i>	A professional interpersonal process that overlaps with psychotherapy but that has as its focus specific problems or adjustments.
<i>Psychoanalysis</i>	A theory of mental development embracing the concepts of the unconscious as a determinant of behavior, the use of defenses against the emergence of unconscious tendencies, and the phenomenon of unconscious transference of relationships with past objects to present relationships.
<i>Psychoanalytic practice</i>	The application of a psychotherapeutic method characterized by emphasis on free association, transference, and interpretation, aimed primarily at producing greater self-awareness.
<i>Psychotherapy</i>	An interpersonal process designed to bring about modifications of feelings, cognitions, attitudes, and behavior that have proven to be troublesome to the person seeking help from a trained professional.
<i>Stress disorder</i>	A mental disorder as defined by the International Classification of Diseases and the Diagnostic and Statistical Manual of the American Psychiatric Association.

Introduction

Psychotherapy is an interpersonal process designed to bring about modifications of feelings, cognitions, attitudes, and behavior that have proven to be troublesome to the person seeking help from a trained professional. This time-honored definition, attributed to Strupp, is a condensation of the key issues of psychotherapy. A wider view of psychotherapy includes the concept of helping a person to become his or her best possible self. Psychotherapy is a major component of most stress management programs and treatments for stress-related disorders.

Psychotherapy as an Interpersonal Process

That one person can influence the way another feels, thinks, and behaves is a fundamental assumption in most cultures. Such influence can be either good or bad. The notion that language is involved is implicit. Psychotherapy is a professional refinement of this language-based interpersonal process. It usually consists of a verbal exchange, but may involve writing or dramatic performance as well.

Modification of Feelings, Cognitions, Attitudes, and Behavior

Psychotherapy aims at the modification of existing feelings, cognitions (thoughts), attitudes, and behavior. Psychotherapy avoids the issue of whether these phenomena are inherited or acquired, and it takes as given that they exist and are potentially modifiable. Nevertheless, there is an implied recognition that there has already been a process at work creating these phenomena, with or without the conscious awareness of the individual. Psychoanalysis and its derivative psychotherapies take account of the origin of these phenomena, but focus on their current manifestation in the transference, that is, on what occurs between the patient and therapist.

Troublesomeness of Phenomena

Disturbances of feelings, thoughts, and behavior are universal experiences. Psychotherapy deals with people who feel troubled by their experience of these phenomena, that is, people who are stressed and distressed. They may feel it directly, for example, in the case of depressed mood, or indirectly, for example,

when an upsurge of aggression leads to relationship problems or trouble with authorities. Stressed people are usually troubled by disturbance of their daily function. For the physically ill, stress may be troublesome because of its impact on their illness. The relationship between mental and physical states is clearest in the case of clinical depression. Having an episode of clinical depression at some stage in life at least doubles the risk of developing a major medical disorder such as coronary heart disease, stroke, or diabetes. Once these disorders develop, the presence of depression doubles the mortality rate and greatly increases morbidity and health-care costs, as do poor social support and poor social integration.

The Person Seeks Help

Psychotherapy as a form of interpersonal influence depends on and makes use of the fact that the troubled person seeks out help. In psychotherapies based on psychoanalytic theory, this action is central to the concept of transference. The person seeking help has expectations of the therapist, based on both the realistic knowledge of what the therapist has to offer and on unrealistic expectations transferred to the person of the therapist on the basis of relationship with other key figures in the person's life. This illustrates the fact that there is much going on in a psychotherapeutic relationship that is unsaid yet represented in language at some level and experienced as feelings and attitudes.

Help Is Sought from a Trained Professional

What differentiates psychotherapy from other forms of positive interpersonal influence is that it is offered within a formal set of parameters with an explicit basis for its conduct. That does not necessarily mean that it is valid, but it does mean that it is open to scrutiny. The trained professionals come from a variety of backgrounds, including medicine (particularly psychiatry), psychology, social work, nursing, and education. Specific training in psychotherapy is usually taken as a postgraduate course. There are institutes within the field that are devoted to training in and the development of psychotherapy. It is this postgraduate training that produces the skills and attitudes that differentiates professional psychotherapy and professional counseling from the informal counseling that many professionals provide in their everyday work.

Another implication of the involvement of a trained professional is that a code of ethics applies. Such codes are likely to include statements about respect

for patients or clients, their values, their beliefs, their uniqueness, and their right to self-determination. In particular, psychotherapists are obliged to respect and take account of cultural factors and to avoid discrimination on the basis of such factors and beliefs. There is a consensus that the practice of professional psychotherapy requires ongoing supervision of some sort, in which there can be scrutiny of the application of these ethical principles. Supervision also addresses adherence to protocols. A special feature of supervision of psychoanalytic psychotherapy is that there is also scrutiny of the therapist's reaction to the patient – called countertransference.

What Is the Difference between Psychotherapy and Education?

All forms of professional psychotherapy contain an educational component, as does counseling, even though education is not their aim. Even psychoanalytic psychotherapy, which deemphasizes the didactic component, provides a model of thinking that the patient or client may find that they have adopted. In the sense that psychotherapy is explanatory, it is educational. The difference between psychotherapy and education lies in the fact that psychotherapy deals specifically with those matters that trouble the person requesting it and uses techniques specific to that issue. However, it must be acknowledged that much of modern teaching, with its individualized programs and greater sense of responsibility for the wider concerns of the student, is psychotherapeutic in nature. Many educational institutions are recognizing this by introducing mental health studies into their training. This does not mean that educators will become professional psychotherapists; that would require postgraduate training.

Stress Management

Stress management is the name given to an intervention offered increasingly and to a wide variety of target populations, including those at risk of stress and those who have experienced it. Often given in a group setting, it most frequently includes specific psychotherapy techniques but also imagery and relaxation techniques and an educational component.

The Schools of Psychotherapy

Psychoanalytic Psychotherapy

Professional psychotherapy has its historical roots in the development of psychoanalysis at the end of the nineteenth century. Psychoanalysis as a practice is a form of psychotherapy. Psychoanalytic

psychotherapy is a term used to describe all forms of psychotherapy that are based on the principles of psychoanalysis. The expressions dynamic and psychodynamic can be taken as a synonym for psychoanalytic. The distinction between psychoanalysis and psychoanalytic psychotherapy is indistinct, but it centers on issues of frequency, technique, and goal. Psychoanalysis is usually more intense in frequency and length. Psychoanalytic psychotherapy uses modified analytic techniques and may mix them with nonanalytic techniques. Psychoanalytic psychotherapy is often more problem-focused and may be explicitly supportive.

The issue of whether a form of psychotherapy is described as supportive is a complex one. All psychotherapies are supportive in the sense that they offer a containing framework in which a patient or client can deal with his or her stresses and lessen his or her distress. The concept of nondirectiveness is more pertinent. In psychoanalytic psychotherapy, interventions that contain suggestions or directions are usually avoided because they may interfere with the achievement of greater self-awareness, a major goal of psychoanalytic practice.

As the field of psychoanalysis developed various theoretical streams, so there developed different schools of practice of psychoanalysis and its derivative psychotherapies. Brief dynamic psychotherapy is associated with the names Malan, Sifneos, Strupp, Davanloo, Mann, and Horowitz. Klein, Winnicott, Balint, and Guntrip pioneered object relationships theory and practice. Kohut developed self-theory. Lacan emphasized the linguistic basis of the structure of the unconscious; schools of psychoanalysis based on his teaching have wide support in Europe and South America. Jung, with his emphasis on universal symbols, spawned persisting schools of psychoanalysis.

Psychotherapies Related to Psychoanalysis

Other theorists such as Carl Rogers, Aaron Beck, and Harry Stack Sullivan focused on aspects of the subjective and cognitive experiences with which psychoanalysis dealt. They and others developed forms of psychotherapy based on their theories.

Experiential therapy Client- or person-centered therapy, existential therapy, Gestalt therapy, and emotional-expressive therapies are the most well-known examples of experiential therapy. They have in common the assumption that human nature is growth-oriented, and they focus on subjective experience. For instance, client-centered therapy is described by Rogers as having parallels in existentialism. It

emphasizes the autonomy of clients (and thus avoids use of the term patient). It also emphasizes the attitudinal qualities of the therapist; warmth, empathy, and genuineness.

Cognitive therapy Cognitive therapy assumes that people can develop automatic erroneous thought patterns and beliefs of which they are unaware and that they consequently think and behave in ways that are irrational. Albert Ellis developed the concept of cognitive restructuring in his rational emotive therapy. Beck developed a similar technique, which is the basis for that used by therapists today, in which the patient or client is helped to identify these automatic thoughts and to challenge them, using what has been described as a Socratic dialog. There are both psychoanalytic and phenomenological underpinnings to the therapy. Ryle has developed cognitive analytic therapy, which emphasizes the psychoanalytic component. Analyses of transcripts of therapists using cognitive therapies show that they incorporate elements of other therapies as well. The constructs of behavioral therapy are often so interwoven with those of cognitive therapy that the term cognitive-behavioral therapy has come to be used. Cognitive-behavioral therapy is a major component of stress management programs and of interventions for acute and chronic traumatic stress disorder. It is widely used in the treatment of psychiatric disorders in general.

Interpersonal psychotherapy Interpersonal psychotherapy is an eclectic blend of traditional clinical approaches that has its roots in the work of Sullivan and was operationalized by Gerald Klerman and colleagues. It focuses on current interpersonal issues, particularly those related to grief, interpersonal role disputes, role transitions, and interpersonal deficits. It is, thus, particularly suited to those experiencing stress. It is a brief form of therapy, originally used for the treatment of depression but now applied to many other forms of distress. It, together with cognitive-behavioral therapy, was found to be as effective as antidepressant medication for the treatment of moderate depression in a large study conducted by the National Institute of Mental Health in the United States. Psychodynamic interpersonal therapy emphasizes the use of psychodynamic techniques. It has been shown to be both efficacious and effective in the physically ill; that is, it works in practice as well as in experiments.

Behavioral Therapy

Behavioral therapy has its roots in classical learning theory, with which the names of Pavlov and Skinner are associated. It is also influenced by Bandura's theory of socially determined learning or modeling.

Watson and Wolpe's early clinical work has now been refined and operationalized into a therapy that focuses on current determinants of behavior and draws on the principles of learning to develop individual treatment strategies. Behavioral therapists are aware of the fact that they cannot avoid completely the issue of the meaning of the behavior targeted or the meaning of the relationship between the therapist and the patient.

Methods of Delivery of Psychotherapy

Psychotherapy can be delivered at all ages, and most types of psychotherapy are applicable to all age groups. Melanie Klein and Anna Freud were pioneers of child psychotherapy. There is an emerging discipline of old age psychotherapy.

Individual Therapy

Individual or one-to-one psychotherapy is the commonest application of psychotherapy, but psychotherapy in small groups is an important form.

Couples Psychotherapy

Couples psychotherapy, for those in a relationship, is used by therapists of most schools of psychotherapy. Many therapists use an eclectic approach, combining elements of various techniques. Sexual issues are often an important focus.

Family Psychotherapy

Family psychotherapy, or family therapy as it is now known, has its roots in psychoanalytic theory. It began with the work of Bateson and colleagues in the 1950s on the role of the family in cases of schizophrenia. Salvador Minuchin and colleagues developed techniques based on a psychoanalytic construct of the family, which focused on ways in which one member could be carrying the anxiety for the whole family. More recently, family therapy has been strongly influenced by systems theory. The construct of family is changing rapidly, and given that each member of a family may come with different needs, family therapists have to adapt their technique to the situation.

Group Psychotherapy

Group psychotherapy can be said to have originated during World War II. Psychoanalytically orientated psychiatrists such as Foulkes and Bion used it in England to manage stress-related cases of neurosis more efficiently. The concepts and techniques that they developed are influential today. The Tavistock Institute in the United Kingdom has pioneered models

of applied psychoanalysis for work groups trying to understand their social task. Such work forms part of stress management programs. Other schools of psychotherapy have developed group techniques. In the United States, Rogers extended his client-centered therapy to its application in encounter groups. Lewin developed methods for use in nonclinical situations. Behavioral therapy (e.g., for phobias) can be delivered in groups.

The Efficacy of Psychotherapy

Studies on the efficacy of psychotherapy for stress disorders as defined by international classifications are limited in number. However, the vast literature on the efficacy of psychotherapy in general is highly pertinent because it addresses the issue of stress by implication. Those diagnosed as having depression or anxiety, the most frequently addressed disorders in the research literature, are likely to have declared stress to be their major symptom, and this will have been acknowledged by their clinician when translating symptoms into a named problem or disorder. Furthermore, even when a primary diagnosis of a stress disorder is made, it is highly likely that a simultaneous diagnosis of depression, anxiety, or substance abuse will also be made. Personality disorder is another common comorbidity. The diagnoses of depression and anxiety imply distress, and the seeking of therapy certainly does so. These conditions, rather than the stress, may be the targets of psychotherapy. However, the exploration of perceived stressors and their effects forms an important part of all psychotherapies.

Overview

Health-care providers and consumers in general seek evidence of the efficacy of psychotherapy and justification of its cost. The empirical study of a therapy as complex as psychotherapy, applied as it is to such a wide range of problems and people, is fraught with methodological difficulties. Nevertheless, these difficulties have been addressed sufficiently well for Lambert and Ogles, in their review of efficacy and effectiveness studies, to conclude that many psychotherapies have been shown to have demonstrable effects on clients in clinical trials that are statistically significant and clinically meaningful and that speed up the natural healing process as well as providing coping strategies for future use. Although some clients may achieve meaningful results after brief psychotherapy (21 sessions), many need at least 50 sessions. Other conclusions were that the effects of therapy tend to be lasting, that psychotherapy patients show gains that surpass those resulting from placebo

controls or no treatment, that psychotherapy can be cheaper than medication, and that clients treated with psychotherapy show 25% less medical service utilization. However, they also conclude that average positive effects mask considerable variability in outcomes, that not all are helped, and that some are harmed by inept or inappropriate application.

Systematic and Meta-Analytic Studies

These conclusions were based on studies conducted over the past 80 years. Many of the earlier ones were naturalistic studies of therapy and follow-up. Some used qualitative research of individual cases, a technique now much refined and becoming increasingly important in the study of human behavior and experience in general. These techniques are particularly useful for examining the process of psychotherapy and the experience of patients or clients. Many of the recent studies are experimental; that is, a research plan is developed and clients are recruited for participation and follow-up, with monitoring of therapists as well. Systematic overviews of experimental studies group those that meet strict criteria for inclusion, such as the requirement of random allocation of clients to a treatment or control arm (randomized controlled studies, RCTs). Both qualitative and quantitative reanalyses may be applied. The statistical technique of meta-analysis is often used. It has methodological problems, but it is generally accepted that it provides a useful way of summarizing studies, albeit one that must be supplemented by more qualitative analyses. Early meta-analytical studies found that the average treated person is better off than 80% of the untreated sample (effect size of 0.85 standard deviation units). Lambert and Ogles conclude that studies since then, which include those on other large samples and a reworking of the original study and use of more sophisticated statistical techniques, have supported that finding in general but find lower but still clinically significant effect sizes (0.4–0.6) for both specific and general measures of outcome. The analyses on which these conclusions are based include studies on a wide range of therapies, including all of the mainstream ones, for specific disorders such as anxiety and depression and for general problems such as those of stress and personality. A wide range of severity of psychopathology is included. By way of perspective, effect sizes are similar to those found for the use of antidepressants in depression and pharmacological treatment of anxiety.

With respect to stress and its disorders, Ehlers and Clark conclude that systematic reviews and meta-analyses show that, contrary to the commonly held view, single-session individual debriefing does not

reduce distress and is not effective in preventing posttrauma symptoms. However, a course of up to 16 psychotherapy sessions may be more effective than supportive counseling in ameliorating the severity of acute symptoms and reducing subsequent posttraumatic stress disorder (PTSD), a chronic condition. Bradley et al., in their systematic review of psychotherapy for PTSD, conclude that substantial effects are found for psychotherapeutic intervention, with the majority of individuals showing improvement that was clinically meaningful; this was sustained for at least 6 months.

Which Therapy for Which Patient?

The question of which therapy is best for which client is only slowly being resolved. Relationship factors (trust, warmth, acceptance, and human wisdom) are probably crucial, even in the more technical therapies that generally ignore them. In fact, it is becoming clear that therapists in clinical trials as well as in practice are eclectic in their adaptation of therapy to suit client profiles. Lambert and Ogles conclude that differences in outcome between various forms of therapy are not as pronounced as might have been expected and that it is premature to favor the so-called empirically supported therapies over others. (This term refers to a movement that advocates that only those therapies with proven experimental support should be taught and funded.) Courses of psychotherapy may be shown to be effective in clinical trials, but they are not as efficient in clinical practice, partly because they are too short. Emerging naturalistic studies show a dose–response relationship such that long-term therapy is required, at least to the point of complete resolution of symptoms.

As previously discussed, stress and related disorders constitute risk factors for the development of physical illness and its outcome. Some work has been done on specific psychotherapy techniques for patients with physical and psychological comorbidity. Studies mainly focus on psychoeducational interventions. Health education and stress management programs produce significant positive effects on blood pressure, cholesterol, body weight, smoking behavior, physical exercise, and eating habits. Psychotherapy for depression in diabetes is effective in relieving symptoms and improving control. Early trials of psychotherapy in patients with cancer showed an increase in both well-being and survival time. Improved physical treatments for cancer and the greater availability of psychosocial treatments may explain why later studies are finding no differences in survival time, despite improvement in well-being and quality of life.

What Are the Active Ingredients of Psychotherapy?

The work of Pennebaker and his colleagues has shown that the writing down of accounts of stressful experiences is accompanied by better outcome than is achieved by writing about nonemotional matters. This supports other evidence that suggests that affective experiencing – catharsis – may to be an important ingredient of psychotherapy. Cognitive mastery, making sense of what is happening, also seems to be important. A focus on strategies for behavioral control is the third major factor identified so far. But there is an enormous amount of research that needs to be done if the active ingredients of psychotherapy are to be dissected out and the question “Which therapy for which client by which therapist and how much of it?” is to be answered. Consistent with this realization, researchers are returning to a focus on the ingredients of psychotherapy: what the therapist and patient actually do, and what goes on between them. Systematic reviews and meta-analyses will help answer the questions, but the medical model on which they are based is insufficient to address this most human of experiences, that of someone’s seeking help from a trained professional and entering an interpersonal process designed to bring about the modifications of feelings, cognitions, attitudes, and behavior that have proved troublesome to the person or, indeed, an individual’s seeking greater self-awareness as a way of fulfilling him- or herself better in his or her personal and social commitments.

See Also the Following Articles

Cognitive Behavioral Therapy; Coping Skills; Crisis Intervention; Defensive Behaviors; Family Therapy; Group Therapy; Reenactment Techniques; Trauma Group Therapy.

Further Reading

Ablon, J. S. and Jones, E. E. (2002). Validity of controlled clinical trials of psychotherapy: findings from the NIMH Treatment of Depression Collaborative Research Program. *American Journal of Psychiatry* 159, 775–783.

- Bloch, S. (1996). *An introduction to the psychotherapies*. Oxford: Oxford University Press.
- Bradley, R., Greene, J., Russ, E., et al. (2005). A multi-dimensional meta-analysis of psychotherapy for PTSD. *American Journal of Psychiatry* 162, 214–227.
- Duffy, M. (ed.) (1999). *Handbook of counseling and psychotherapy with older adults* New York: John Wiley.
- Ehlers, A. and Clark, D. M. (2003). Early psychological interventions for adult survivors of trauma: a review. *Biological Psychiatry* 53, 817–826.
- Elkin, I., Shea, M. T., Watkins, J. T., et al. (1989). National Institute of Mental Health Treatment of Depression Collaborative Program: general effectiveness of treatments. *Archives of General Psychiatry* 46, 971–982.
- Hollon, S. D., Thase, M. E. and Markowitz, J. C. (2002). Treatment and prevention of depression. *Psychological Science in the Public Interest* 3, 39–77.
- Kazdin, A. E. (2000). *Psychotherapy for children and adolescents: directions for research and practice*. New York: Oxford University Press.
- Lambert, M. J. and Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In: Lambert, M. J. (ed.) *Bergin and Garfield’s handbook of psychotherapy and behavior change* (5th edn.), pp. 139–193. New York: John Wiley.
- Markowitz, J. C. and Weissman, M. M. (2004). Interpersonal psychotherapy: principles and application. *World Psychiatry* 3, 136–139.
- Ong, L., Linden, W. and Young, S. (2004). Stress management. What is it? *Journal of Psychosomatic Research* 56, 133–137.
- Reinecke, M. A. and Clarke, D. A. (eds.) (2004). *Cognitive therapy across the lifespan*. Cambridge, UK: Cambridge University Press.
- Roth, A. and Fonagy, P. (2005). *What works for whom? A critical review of psychotherapy research* (2nd edn.). New York: The Guilford Press.
- Sandell, R., Blomberg, J. and Lazar, A. (2002). Time matters: on temporal interactions in long-term psychotherapies. *Psychotherapy Research* 12, 39–58.
- Smith, G. C. (2003). The future of consultation-liaison psychiatry. *Australian and New Zealand Journal of Psychiatry* 37, 150–159.
- Spiegel, D. (2002). Effects of psychotherapy on cancer survival. *Nature Reviews Cancer* 2002(May), 383–389.
- Westen, D., Novotny, C. M. and Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin* 130, 631–663.

Psychotic Disorders

J Ventura

University of California, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J Ventura and R P Liberman, volume 3, pp 316–326, © 2000, Elsevier Inc.

Stress in Psychotic Disorders: Conceptual Framework
Stressful Life Events and Psychosis
Natural Disasters and Psychotic Disorders
Expressed Emotion as a Stressor
Treatment as Protection against Stress and Psychosis

Glossary

<i>Brief psychiatric rating scale (BPRS)</i>	One of the most widely used rating scales, nationally and internationally, for measuring psychiatric symptoms at baseline and for assessment of change over time. The BPRS allows for the rapid assessment of symptoms such as depression, anxiety, delusions, and hallucinations.
<i>Cognitive appraisal</i>	An individual's evaluative thought process that instills a life event or situation with a particular meaning.
<i>Coping response</i>	Cognitive and behavioral efforts used to master, tolerate, and reduce the demands of stressful events that tax or exceed an individual's resources.
<i>Expressed emotion (EE)</i>	Attitudes and feelings expressed by a relative during family interactions; an individual with a psychotic disorder would describe this as critical comments, overt or subtle hostility, or emotional overinvolvement in the individual's affairs.
<i>Independent life event</i>	A life event or situation that is not related to an individual's psychotic symptoms and is beyond his or her personal influence, for example, the death of a relative, the loss of a job because the entire company went out of business, or becoming the victim of an unprovoked attack.
<i>Mediator</i>	A variable that explains how, why, or when to expect a relationship between a predictor (e.g., life event) and the criterion (e.g., psychosis).
<i>Prospective studies</i>	Research in which the data are collected on an ongoing regular basis about all study participants, whether they are considered at risk for an illness or not and before the event being studied occurs (e.g., psychotic relapse).

Protective factor

Individual characteristics, which may be biological, intrapersonal, interpersonal, or learned, that reduce or eliminate the risk of developing a psychosis.

Stress

An adverse biological, physiological, or psychological state that develops when the demands on an individual exceed his or her capacity for adaptation.

Vulnerability to psychosis

The enduring characteristics of individuals, which might be genetic, social, behavioral, biochemical, cognitive, or related to brain structure, that increase their susceptibility to developing psychosis.

Stress is a concept on whose definition few scholars or researchers can agree. Yet there is little dispute that stressful events can influence the onset and course of major psychotic disorders. A 1977 vulnerability–stress model by Zubin and Spring proposes that a relationship exists among the level of inherited vulnerability, stress that exceeds a certain threshold, and the onset or relapse of psychotic illness. According to this vulnerability model, individuals with high biological vulnerability are susceptible to the exacerbation of psychotic symptoms even when they experience low levels of stress; however, if the biological vulnerability is low, then high levels of stress are required before symptoms emerge. Differences in the rates of psychotic disorders found in males versus females, urban versus rural areas, and high-stress versus low-stress households support the importance of environmental factors in understanding gene–environment interactions.

Several empirical tests of the vulnerability–stress model have shown that an increase in the frequency of life events occurs in the month before a psychotic exacerbation. However, not all studies have found a reliable association between stressful life events and psychosis. Some researchers believe that the relationship in schizophrenia is weak when compared to more striking findings in patients with mood disorders. Although there is a convergence of evidence that stressful life events are triggers of psychosis, stressful events alone are neither necessary nor sufficient to account for illness onset, nor are they necessary as precursors of relapse. Antipsychotic medication may provide protection against stress and raise the threshold of vulnerability. An interactive vulnerability–stress–protective factors model (Figure 1) hypothesizes the existence of mediators of stress, some of which are considered protective such as

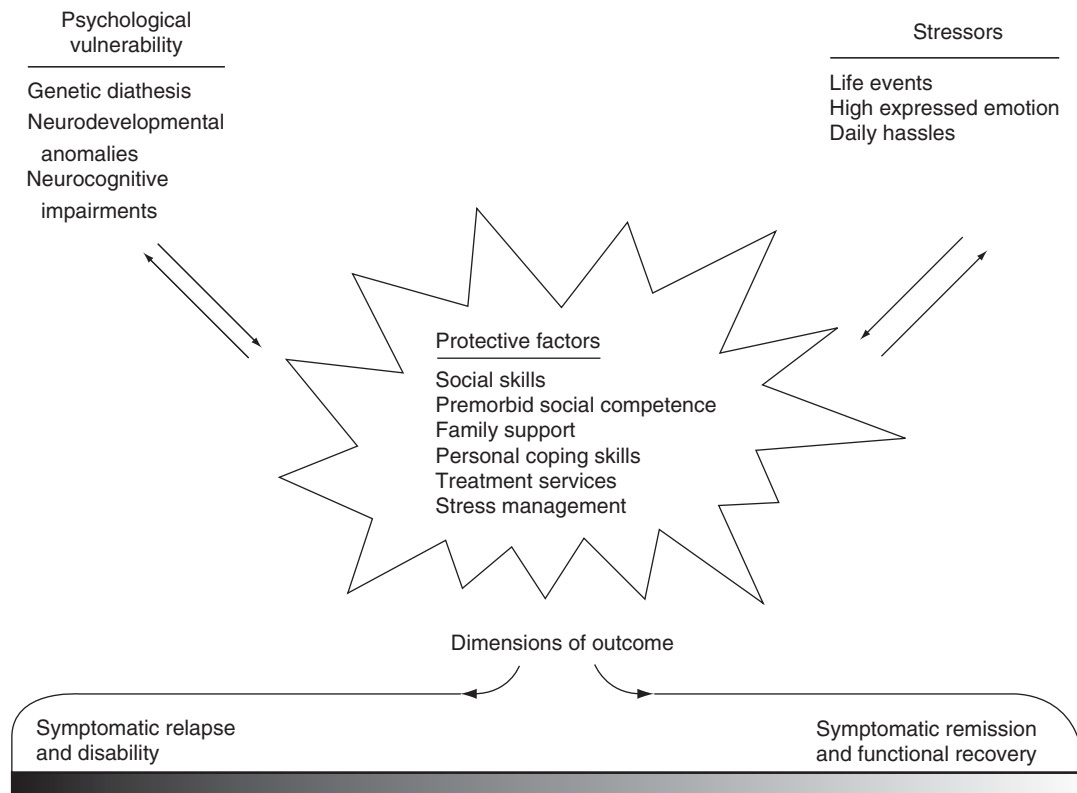


Figure 1 An interactive vulnerability–stress–protective factors model of risk factors for the onset and course of psychotic disorders.

cognitive appraisal of life events, social competence, social and familial support, and premorbid intelligence.

Stress in Psychotic Disorders: Conceptual Framework

All biomedical disorders, including those that manifest themselves by psychotic symptoms, are stress-related biological illnesses. It doesn't matter whether it is the heart or the brain that is involved as the organ from which the illness arises, stressors impinge on the organ's genetic vulnerability to trigger episodes of symptoms, exacerbations, relapses, and disability. Moreover, protective factors in the individual or in the individual's environment can mitigate or buffer the noxious effects of stressors on vulnerable people. When protective factors – such as social competence, social and family support, intelligence, availability, and accessibility to preventive interventions and good quality treatment – outweigh stressors and vulnerability, health may be sustained and disability forestalled.

In this article, we identify some of the stressors that have been implicated in the induction of episodes of illness, relapse, and poor outcome in psychotic disorders. However, it is important to place the role of

stressors within the framework of the vulnerability–stress–protective factors model so as to understand why an individual may experience an episode of illness when exposed to a stressor at one time but not when exposed to the same stressor at another time. In addition, the vulnerability–stress protective–factors framework permits clinical scientists to explain the enormous interindividual variability in susceptibility to stressors – in short, why some people, some of the time, do not experience stress and exacerbations of psychosis in response to certain stressors, but others do.

Stressful Life Events and Psychosis

The life charts of Adolf Meyer, published in 1951, were among the first records of clinical observations that emphasized the importance of psychosocial stressors and the development of pathological conditions. Over the past 3 decades, many researchers have attempted to empirically document the nature of the relationship between stressful life events and psychotic illness. Such studies are based on the notion that higher frequencies of events translate into higher levels of stress for patients. These studies can be grouped into those that compared schizophrenia patients and other psychiatric groups, those that

compared schizophrenia patients with community controls, and those that examined the impact of stressful events on the onset and relapse of psychotic symptoms.

Most research has shown that, compared to schizophrenia patients, other psychiatric groups (usually depressed patients) reported either similar or significantly higher rates of stressful events prior to an episode. These differences were unexpected because there was an assumption that the rate of stressful life events in schizophrenia patients was higher than for patients with other psychiatric disorders. Schizophrenia patients are considered sensitive to even subtle changes in the environment. However, given that schizophrenia patients sometimes withdraw when stressed, their rate of life events may actually be lower than the rate for patients with other psychiatric disorders. Available evidence supports the conclusion that schizophrenia patients report fewer life events than other diagnostic groups. Furthermore, most studies have found that the strength of the relationship between stress and symptoms is more robust in patients with major depression than in patients with schizophrenia.

Most studies that used community controls as a comparison found that schizophrenia patients experienced a higher frequency of stressful events. However, some studies showed no difference. Interestingly, one prospective study found that normal controls actually had higher levels of stressful events than did schizophrenia patients. Several studies have shown an increased frequency of triggering life events in the month prior to the onset or relapse of schizophrenic symptoms. Even studies that did not find a sharp increase in the number of life events in the month prior to onset or relapse found that stressful life events had increased the risk of psychosis. There has been some investigation into whether life events play a more prominent role in the onset of psychosis than in relapse. Although there is no consistent evidence favoring a relationship between stress and onset (compared with stress and relapse), there is some indication that life events play a larger role in psychotic episodes that have a discrete (vs. insidious) onset.

Perhaps more relevant than the absolute frequency of events in patients or controls is the nature of the relationship between life events and psychotic illness. However, most of the studies of stress and psychosis focused on the frequency of events rather than the amount of threat posed. In fact, methods of measuring life event frequency (through open-ended interviews or life event inventories) are readily available. But, in some cases, events that were hoped for or expected (e.g., finding a suitable mate) but did not

occur may also be perceived as stressful. Most researchers believe that the impact of a stressor is determined by an individual's cognitive appraisal and that the appraisal is altered in schizophrenia patients in a way that increases vulnerability to stress and psychosis; yet this important topic has received limited study. Perhaps the popular belief that most schizophrenia patients cannot accurately report their appraisal of life events dampens researchers' enthusiasm for comprehensive study.

Retrospective Studies of Life Events and Psychosis

In a classic and frequently cited 1968 British study, Brown and Birley interviewed schizophrenia patients and their families soon after a hospitalization regarding the 12-week period prior to the onset or relapse of psychosis. Schizophrenic patients reported an increased frequency of independent life events (independent of the insidious onset or occurrence of psychotic symptoms) in the 3-week period prior to the datable onset of their episode. Specifically, 46% of patients reported an independent life event in that 3-week period, compared with 14% for a sample of community controls during a comparable period. Except for the 3-week period prior to an episode, the life event rates between the two groups did not differ significantly. From these results, Brown and Birley concluded that life events play a role in triggering the onset or relapse of psychotic symptoms in schizophrenia.

The most ambitious attempt at a replication of the Brown and Birley findings was a 1987 multinational World Health Organization (WHO) study published by Richard Day and colleagues. The design of this WHO study was virtually identical to Brown and Birley's original work, including the use of retrospective reporting, although normal controls were not included. The life events data were collected in nine centers, five located in developed countries and four in developing countries. Day and colleagues replicated the finding that independent life events occur with increased frequency in the 3 weeks prior to psychotic episodes in schizophrenic patients. Although the lack of a comparison group is a serious concern, the consistency of these findings from different countries lends considerable support to the hypothesis that stressful life events may trigger episodes of schizophrenia.

Questions Concerning Retrospective Studies

Most of the classic research supporting the idea that stress triggers psychosis has been subject to criticism because the data was collected using retrospective methods, in which data regarding life events were

collected after onset or relapse of psychosis. Because retrospective data gathering relies on the patient's memory of life events and the date of onset of psychotic symptoms, it is subject to considerable report distortion, particularly for events of smaller magnitude. Such distortion is likely to bias results when patients (or their families) are searching for an explanation for a known outcome, such as a psychotic episode. In addition, there is the issue of bias in the dating of the onset or relapse of psychosis in these studies. Such a bias favors finding a relationship between stress and psychosis even if there was none.

The prospective studies of the mid-1980s and 1990s, which used a follow-through design in the regular collection of life events data, were aimed at correcting the methodological flaws of the prior retrospective studies. In some studies, life events data were collected as often as biweekly, thereby minimizing errors associated with memory for stressors that preceded relapse. In addition, each patient's symptoms were assessed frequently, minimizing the distortion of an association between the presence of symptoms and certain events.

Prospective Studies of Life Events and Relapse

The first well-designed prospective study of schizophrenia outpatients addressed the question of whether a significant increase in symptoms occurs after independent life events rather than evaluating the frequency of life events in the month preceding symptom relapse. This study, which collected life events data on an ongoing regular basis, focused on the average effects of life events on illness. An increase in negative symptoms was found in response to stressful events but not an increase in positive psychotic symptoms. Based on this stringent test of the relationship between stress and psychosis, Jean Hardesty and colleagues concluded that independent life events have minimal bearing on florid psychotic symptoms in schizophrenia.

A 1989 prospective study by Ventura and colleagues introduced additional methodological improvements in data collection, including (1) careful monitoring of the amount of antipsychotic medication received, (2) the use of both an open-ended style of interviewing and a structured life events inventory, and (3) the classification of life events as independent only when they were both independent of the patient's illness and not within the patient's ability to influence them. Ventura et al. found, as had Brown and Birley, and Day and colleagues, an increase in the number of independent life events in the 4-week period before a relapse or significant exacerbation of positive symptoms; 5 of 11 (45%) patients had at least one independent life

event in the month before relapsing, compared with 1 of 11 (9%) during a comparable period without relapse. Thus, a well-designed prospective study supported the conclusion that stressful life events play a role in triggering psychotic symptoms.

Stress Sensitivity

Researchers and clinicians alike have theorized that schizophrenia patients are sensitive to even very small negative changes in the environment. A series of studies using the Experience Sampling Method (ESM) has provided some evidence for the validity of this hypothesis. The ESM requires that patients wear a beeping watch that alerts them 10 times a day to record real-life activity and their emotional response over a 6-day period. Using the ESM, researchers in the Netherlands found that schizophrenia patients, compared to high-risk relatives and controls, are more likely to experience negative affect in response to minor life events. This negative affectivity is considered a vulnerability marker for future psychosis. In addition, using the ESM the researchers recorded life events in sets of twins, in which one twin scored high on a measure of psychosis proneness and the other scored low. The advantage of this method of research is that the well twin acts as a genetic control for the psychosis-prone twin. Twins who were prone to experiencing psychotic-like experiences were more likely to show sensitivity to minor stressful life events. These findings further support the hypothesis that stress sensitivity acts as a risk factor for psychosis.

General Conclusions on Life Events and Psychosis

Interestingly, similar conclusions about the relationship between stress and psychosis can be drawn from sophisticated prospective studies and the retrospective studies. There is strong evidence from both retrospective and prospective studies to support the triggering hypothesis, even though some studies were able to confirm only an increased risk of psychosis from stressful events. Although life events influence psychotic episodes, the occurrence of stressful life events is usually not sufficient to precipitate an increase in psychotic symptoms. Furthermore, life events are not always necessary for relapse to occur. Some studies found that more than one-half of relapsing patients did not experience a prior life event. Other factors, such as medication reduction/discontinuation or biological factors, may also contribute to the risk of relapse. These conclusions attest to the potential importance of mediators such as, cognitive appraisal and coping efforts, which produce

differences in individual response to stress within groups of psychotic patients. The finding of increased negative symptoms in response to stressful events, instead of psychosis, indicates that not all patients respond to stress in the same way. Prospective research has also shown that life events do not have to be large in magnitude; both major and minor life events are predictors of symptom exacerbation and relapse in schizophrenia. Furthermore, stress sensitivity to daily hassles and minor life events has been associated with increased subjective distress and potential risk for future psychosis in schizophrenia.

Cognitive Appraisal of Stressful Events

Understanding the link between stressors and subsequent psychotic symptoms in schizophrenia may entail looking more closely at potential mediators such as patients' cognitive appraisal of life events. According to stress researchers, an individual's appraisal of an event always should be measured because appraisal determines the stressful nature of the event. In one stress-illness model, the perception of the stressor is schematically represented as a polarizing filter that can magnify or diminish the stressfulness of an event. Richard Lazarus and Susan Folkman hypothesized that appraisal is the first element of a two-phase cognitive-phenomenological process of coping and thereby is a factor in determining how much stress an individual experiences. After decades of research, Lazarus continued to assert that an individual's appraisal is necessary and sufficient in the generation of emotions in response to stressful events. Appraisal thus is relevant to understanding the relationship between stressful life events and subsequent illness.

Research suggests that schizophrenic patients, in particular, might respond in idiosyncratic ways to life events. Appraisal may play a larger mediating role in the relationship between life events and psychotic symptoms in schizophrenia patients than it does in other populations. Both psychoanalytic and cognitive information-processing theories suggest that schizophrenia patients rely on narrow and idiosyncratic meanings of events. Heightened sensitivity to environmental stimuli, coupled with cognitive distortions and psychotic thinking (e.g., persecutory ideas or referential thoughts to either perceived or existing threats) could result in increased stress. Research has shown that schizophrenia patients may overestimate the threatening nature of everyday events or situations and underestimate the effectiveness of their problem-solving ability. This tendency for altered appraisal may increase in patients as their condition worsens.

Can Stress Cause Psychotic Illness?

Although some investigators have argued that stressful life events can cause psychotic episodes in schizophrenia patients, others have argued that life events can only act as triggers of episodes. Central to this issue is whether psychosis can occur in individuals without inherited vulnerability and from stressful events that are definitely independent of the illness. Exposure to wartime combat, natural disasters, and other severe traumatic events that are clearly independent of the individual's influence have been shown to increase the risk of developing psychosis. According to one theory, psychosis can occur even in individuals without inherited vulnerability characteristics, especially if the stressor is severe. However, many individuals have been exposed to all types of stressful situations and catastrophic events and only a handful become psychotic. Some theorists believe that individuals who became psychotic must have some type of underlying vulnerability to illness and that therefore the stress merely acted as a trigger. The possibility also exists that the role of life events may be indirect and occur in conjunction with other risk factors.

Stressful Life Events and Depression in Psychosis

The presence of clinical depression in schizophrenia patients has generated a great deal of interest because of its significance as an early indicator of an impending psychotic relapse, association with an increased risk for suicide attempts, and relationship to psychiatric rehospitalization. Several theories have been put forward to explain the occurrence of depression in schizophrenia. Depression has been considered part of the postpsychotic phase when it occurs after a psychosis remits and as a reaction to the realization that the individual is in trouble or that his or her life has fallen far short of prior expectations. Other researchers view depressive symptoms as part of the acute psychotic process because it has been found to occur simultaneously with the exacerbations of positive symptoms. Depression may be a part of the patient's biological diathesis to schizophrenia and therefore an integral part of the disorder or an aspect of familial genetic liability to disturbances in mood. Some researchers view depression as an unwanted side effect of antipsychotic medication. However, observations by Eugen Bleuler and Emil Kraepelin of depression in schizophrenia during the preneuroleptic era suggest that depression is not simply secondary to antipsychotic medications. In any event, according to the vulnerability-stress-protective factors model, psychosocial stressors such as life events may be expected to increase the chances

of an onset or exacerbation of depressive symptoms in psychotic patients.

Schizophrenia patients have reported a higher frequency of stressful life events during the 6 months before a depressive exacerbation and during the 12-month period prior to a postpsychotic depression. Early parental loss (before age 17) was also found to be more common in depressed than nondepressed chronic paranoid schizophrenia patients. In chronic schizophrenia patients, increased levels of depression have been found to be associated with psychosocial stressors, such as unemployment. Although much of this previous research had some methodological weaknesses, recent work with sound methods confirmed the earlier findings. A well-designed 2000 prospective study using survival analysis showed that stressful life events act as triggers of depressive exacerbation in recent-onset schizophrenia patients. These studies suggest that individuals with schizophrenia who experience undesirable life events may have a vulnerability to depression similar to mood disorder patients and individuals without psychiatric disorders. Thus, stressful events may play an even broader role in psychosis than is typically theorized by the vulnerability–stress–protective factors model.

Natural Disasters and Psychotic Disorders

There are various types of stressful life events that can have a negative impact on schizophrenia patients, ranging from minor daily hassles to major life events. These include catastrophic events such as natural disasters. One reliable predictor of psychiatric morbidity is the degree of exposure to the devastating physical consequences that follow natural disasters. Research on natural disasters showed consistently that the existence of a prior psychiatric disorder (or symptoms) is a risk factor for postdisaster psychopathology. Preexisting disorders such as major depression, posttraumatic stress disorder (PTSD), anxiety disorders, and alcohol dependence have been associated with postdisaster morbidity. One study found 98% of individuals with a prior psychiatric disorder had a diagnosable postdisaster psychiatric disorder, compared with 25% of postdisaster diagnosis in the group without a previous diagnosis. Most studies have been retrospective and therefore subject to recall bias, but these results have been supported in prospective research.

Contrary to what is predicted by the vulnerability–stress–protective factors model, most anecdotal accounts and a few studies indicated that psychiatric patients remain clinically stable or even improve during or after natural or human-made disasters

(e.g., hurricanes, ice storms, nuclear power plant accidents, and missile attacks). During Hurricane Iniki, psychiatric inpatients in Hawaii responded with increased independence, defined as decreased requests for therapy. Similarly, during SCUD missile attacks on Israel throughout the Persian Gulf War, many psychiatric inpatients were described as having coped well, quietly entering sealed concrete bunkers and easily donning gas masks. During a state hospital fire, the patients assisted the hospital staff in the efforts to control the blaze.

In contrast to these prior reports, patients with psychotic disorders in some studies of disasters demonstrated evidence of heightened stress reactivity. On January 17, 1994, a 6.8-magnitude earthquake struck Northridge, California, and provided our research group with an unprecedented opportunity to study the impact of a natural disaster on psychotic patients. Our sample consisted of patients who were previously diagnosed as having a psychotic disorder and were active participants in research projects. As hypothesized, the degree of exposure to the earthquake was a statistically significant predictor of the amount of distress experienced. Schizophrenia patients reported higher levels of avoidance symptoms than did controls when assessed postearthquake. However, the groups did not significantly differ in their mean levels of intrusion symptoms. The specific elevation of avoidance symptoms is noteworthy because the early appearance of avoidance and emotional numbing has been found to be the strongest predictors of disaster-induced psychiatric morbidity, including PTSD. BPRS assessments were compared for patients for whom 1-week preearthquake and 1-week postearthquake ratings were available. We found a significant increase in postearthquake dysphoria but not in psychotic symptoms. In response to the September 11, 2001, terrorist attacks, there was no difference between inpatients in New York who did and inpatients who did not have the opportunity to directly view the disaster through downtown-facing windows. However, patients with a schizophrenia spectrum diagnosis showed evidence of worsening symptoms compared to those with affective disorder or other diagnoses. Although there is some evidence of confirmation of the stress reactivity hypothesis in response to disasters, the stress reactions of psychotic patients are not always consistent and might depend on the amount of exposure, type of disaster, and the setting.

Expressed Emotion as a Stressor

High expressed emotion (EE), the attitudes and feelings expressed by a family member to a patient with a

psychotic disorder, is the most potent stressor precipitating relapses. EE is a misnomer because it does not refer to emotional expressiveness but, rather, to very specific and operationalized types of attitudes and feelings that are associated with stressful interactions within a family living with a severely mentally ill person. These attitudes and feelings take the form of critical comments, hostility, and emotional overinvolvement. Critical comments are complaints about specific behaviors or symptoms of the mentally ill family member, whereas hostility is generally directed at broader and more global personality characteristics. Emotional overinvolvement is overprotectiveness, intrusiveness, self-sacrificing behavior, and hovering over the patient by the family member. Conversely, low EE, which is associated with lower than average relapse rates, may subserve a protective function by insulating the individual from stressors with positive comments, positive reinforcement, and warmth.

Family members' levels of EE are reliably associated with patients' risks of relapse in the subsequent 9–12 months after a psychotic episode. Patients living with or having frequent contact with one or more high EE relatives have relapse rates 2–3 times greater than patients living with or having contact only with low EE relatives. This increased risk is present despite adequate treatment with antipsychotic medication. In a recent meta-analysis of 27 studies, the average rate of psychotic relapse in patients living with one or more high EE relatives was 52%, compared to 22% in patients coming from low EE families. The probability that this result was due to chance is virtually zero. This association of EE and relapse has been found for relapses in bipolar disorder, schizophrenia, and major depression.

Although there is no evidence whatsoever that EE or families have an etiological role in the first episode or development of psychosis, there is solid evidence that there is a causal link between EE and relapse. More than 15 treatment studies in several countries have shown that behavioral, social learning, and psychoeducational interventions that are well-organized, structured, and focused on teaching family members practical methods for understanding and coping with psychosis lead to reductions in EE and in relapses. Thus, the treatment research provides experimental evidence for the role of EE in stress-related relapse.

The association between EE and relapse in psychosis has been consistently found in developing countries (e.g., India) and developed countries (e.g., United States, United Kingdom, and Italy); however, cultural factors do play a role in the proportion of families from various cultures who are high versus low EE. For example, the proportion of high EE

families is significantly lower in rural India than in urban India and is lower among unacculturated Mexican American families than among Anglo-American families. It is likely that cultural factors play a role in the way that families react to and cope with major mental illness in a relative, and this moves the understanding of EE into an interactional framework.

During direct observation of high EE family members interacting to solve a problem, the emergence of subclinical psychopathology in the patient triggered responses by the patient's relatives that might be viewed as incipient signs of high EE. Negative interactional cycles escalated, with efforts by the family to squelch the patient's deviant behavior, leading to more intense symptomatic behavior and greater negativity in the cycles. There is evidence also that enduring personality, attributional, and information processing factors in high EE relatives may contribute to the association between EE and relapses in psychotic disorders. For example, families who perceive the causes of the patients' problems or symptoms as being controllable by the patients (i.e., families with high internal locus of control) are more likely to be high EE. On the other hand, families that are rated as having higher tolerance, flexibility, empathy, and communication are more likely to be low EE.

The role of EE as a stressor in schizophrenia and other psychotic conditions is elucidated by the stress-vulnerability-protective factors model. Certain characteristics of relatives (e.g., attributional style) may predispose them to experience stress when they are faced with abnormal, unpredictable, and even dangerous behavior in their ill family member. They attempt to cope with this stress by using any means at hand – even criticism, hostility, or emotional overinvolvement – and this has counterproductive effects on the ill relative, causing that person stress and leading eventually to relapse after many cycles of stress superimposed on vulnerability in both the patient and relatives. The interactional nature of EE that develops in this synchronous fashion does indeed link stress, vulnerability, and protective factors in both the family members and patient.

Treatment as Protection against Stress and Psychosis

The Role of Antipsychotic Medication in Reducing Psychosis

In most, if not all studies, antipsychotic medications have repeatedly been shown to be effective for the treatment of psychosis and the prevention of relapse in schizophrenia. Brown and Birley reported that a

larger percentage of patients who relapsed while taking an antipsychotic medication had experienced a life event compared to patients who relapsed while not on medication. Additional work by Brown and Birley in 1968 also suggested that “life events and reducing or stopping phenothiazines [antipsychotic medication] contribute as precipitants of acute schizophrenia.” Similar results were published showing that a significantly greater proportion of the patients who relapsed on medication had experienced an independent life event, compared to nonrelapsing patients not on active medication. Leff proposed that regular antipsychotic maintenance medication produces a prophylactic effect against stress by raising a patient’s threshold to relapse; this increased threshold makes the patient less likely to relapse unless exposed to major life stressors.

In a 1992 prospective study, Ventura et al. found that patients who had a psychotic exacerbation while on medication had more independent events in the prior month than did patients who had an exacerbation while off medication. As previously proposed by Leff, Ventura et al. concluded that taking an antipsychotic medication is protective and raises the threshold for relapse. However, another well-designed study did not completely support the hypothesis that life events are more relevant to relapse in patients on maintenance medication.

Behavioral Family Management for Reducing Stress Inherent in High Expressed Emotion

As shown in [Figure 2](#), the daily hassles and challenging or stressful events of everyday life in a family with a relative who has a disabling psychosis can be dealt with in either of two ways. If the family is equipped with good coping and communication skills, these daily problems can be decompressed and solved without the emotional temperature in the family rising above the ill relative’s vulnerability threshold, decreasing the risk of relapse. However, families that lack the know-how and skills to cope with the stressors of everyday life can, all too readily, succumb to a cycle of stress, poor communication, and mounting problems, leading to relapse. Given the large proportion of families that show high EE in the United States and most developed countries, an intervention program is essential to cool the family emotional temperature.

Behavioral family management, initially designed by Liberman, Falloon, and their colleagues at the UCLA Center for Research on Treatment and Rehabilitation of Psychosis, comprises education of the family and patient about the nature of the mental illness affecting the patient, information and advice about using community-based resources for treatment and rehabilitation, planned and structured

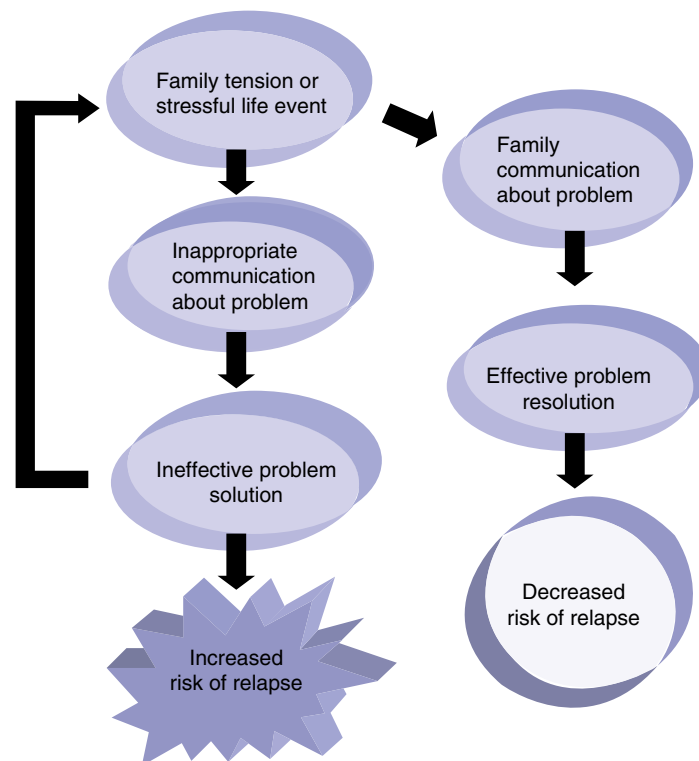


Figure 2 Role of the family in preventing relapse.

training in communication skills, and training in problem solving. Research conducted and replicated on behavioral family management and related treatments has revealed that family-based stress management and skills training is a highly effective addition to the long-term clinical management of psychotic disorders. This modality of treatment has shown superiority in reducing relapse and rehospitalization rates, improving rates of remission of symptoms, enhancing social outcomes, and increasing family and economic benefits. Moreover, the technique can be adapted to be useful in a wide variety of settings, employed by many different disciplines of mental health workers, and effective in a spectrum of patient populations.

Coping with Stressful Life Events

Stressful life events occur often enough in the lives of schizophrenia patients to significantly increase the risk of psychotic relapse. Some research suggested that individuals who experience stress, but who do not relapse, possess psychological characteristics and cognitive functioning that may be associated with effective coping. In particular, the use of problem-focused coping strategies has been associated with good neurocognitive functioning and has been shown to lessen the impact of stressful life events by lowering the risk of relapse. High levels of self-esteem, hope, and insight and the perception of available social support were found in schizophrenia patients who reported using active problem-focused coping strategies. Protective mechanisms, such as effective coping behaviors, in patients and their families may have changed stressful situations into more minor events and reduced the risk of relapse. Positive symptoms have been reduced in patients who consistently applied coping strategies to stressful life events and used problem-solving techniques. These findings support the importance of coping behavior as a protective factor and mediator in the vulnerability–stress–protective factors model.

Despite good evidence for the value of using specific problem-focused coping strategies, schizophrenia patients may fail to use them. This differential use of coping strategies may explain the variability of schizophrenia patients' responses to stressful events. Even after training, they may continue to use ineffective strategies and report that the strategies they used most often were the least effective. Despite the fact that successful coping behaviors may be protective and can be learned, relatively few patients seem to avail themselves of the benefits of using coping strategies. In addition, because many schizophrenia patients have aversive reactions to stressful events, they may respond with avoidance and withdrawal to reduce their acute stress. Thus, further longitudinal research is needed to determine whether short-term or brief avoidance could in some way be adaptive or always leads to problems later.

Further Reading

- Brown, G. W. and Birley, J. L. T. (1968). Crisis and life change and the onset of schizophrenia. *Journal of Health and Social Behavior* 9, 203–214.
- Bustillo, J., Lauriello, J., Horan, W., et al. (2001). The psychosocial treatment of schizophrenia: an update. *American Journal of Psychiatry* 158(2), 163–175.
- DeLisi, L., Cohen, T. and Maurizio, A. (2004). Hospitalized psychiatric patients view the World Trade Center disaster. *Psychiatry Research* 129(2), 201–207.
- Myin-Germeys, I., Krabbendam, L., Jolles, J., et al. (2002). Are cognitive impairments associated with sensitivity to stress in schizophrenia?: an experience sampling study. *American Journal of Psychiatry* 159(3), 443–449.
- Mylin-Germeys, I., van Os, J., Schwartz, J., et al. (2001). Emotional reactivity to daily life stress in psychosis. *Archives of General Psychiatry* 58(12), 1137–1144.
- Ventura, J., Nuechterlein, K. H., Subotnik, K. L., et al. (2000). Life events can trigger depressive symptoms in early schizophrenia. *Journal of Abnormal Psychology* 109, 139–144.



Quality of Life

S M Skevington

University of Bath, Bath, UK

© 2007 Elsevier Inc. All rights reserved.

The Concept of Quality of Life

Assessment of Quality of Life

Dimensions of Quality of Life

Glossary

<i>Subjective</i>	Personal.
<i>Objective</i>	Detached from personal views.
<i>Observable</i>	Visible.
<i>Perceptions</i>	Views.
<i>Functioning</i>	What a person is capable of doing.
<i>Encephalograph (EEG)</i>	Measured electrical changes in brain waves.
<i>Eustress</i>	Positive or “good” stress.

The Concept of Quality of Life

Quality of life (QoL) has become a key goal of contemporary health care. It is often confused with standard-of-living. However, standard-of-living refers to the possession of wealth or material goods. Although a certain number of square feet of living space in United States, or oxen owned in Ethiopia, makes a tangible difference in people's lives, it does not necessarily bring greater happiness or well-being; we can observe this in lottery winners. International figures on annual income per capita collected by the World Bank show that once income exceeds a critical level – \$13,000 in 1995 – the close relationship between subjective well-being and income becomes increasingly loose and dissipated. Furthermore, in the world's richest countries standard-of-living increases do not appear to make a significant difference to people's QoL, a finding that is highly perplexing for policy makers.

A consensus about how to define and measure QoL is still widely debated. In the 1970s, its definitions contained terminological similarities to definitions of

stress. At a time when stress was seen as a phenomenon that exceeded people's resources, in order to provide a good QoL these resources needed to be adequate in terms of satisfying people's wants, needs, and capacities. Since then, definitions have placed greater emphasis on people's subjective perceptions of the important features of their life and, in particular, explored the varied meanings ascribed to these experiences. The ways in which people interpret life's events (e.g., as stressful or pleasant) affects how they see their QoL. Wenger et al. (1984) defined QoL as “an individual's perceptions of his or her functioning and well-being in different domains of life.” Judgments about QoL are now seen as a rich interplay and balance between how people see their internal state, such as the tension in muscles or happiness, and the external events that impinge on them from their environment, such as changing jobs or being bereaved.

Assessment of Quality of Life

The history of QoL assessment measures shows that some of the earliest versions were designed by clinicians for their own use in distributing health-care resources, for example the assessment of Quality Adjusted Life Years (QALYS). This approach excluded a systematic assessment of patients' views, with the assumption that clinicians held a reasonably accurate view of their patients' QoL. However, research has consistently shown that QoL judgments by clinicians and other observers (e.g., spouses or caregivers) often correlate weakly with the views held by the patients about their own QoL. Such proxy information is therefore unreliable, and this is particularly apparent in the case of dementia or stroke, in which the patients' own perceptions can be difficult or impossible to obtain. This discrepancy arises from the observation that many aspects of QoL are simply not open to inspection but are hidden from view like the submerged parts of an iceberg. Furthermore, when measures have limited their contents to subjective reports of observable behaviors such as mobility and

activities of daily living, it has resulted in the assessment of a constricted rather than holistic concept of QoL, due to the exclusion of important but invisible QoL issues such as spirituality and social relationships.

The need for QoL assessment in health has been driven by the problem-solving demands of clinical practice. The aim was to better understand the patient's priorities in order to deliver appropriate care. However, this approach may have provided an unduly negative view of QoL. Patients expect to report their problems during a consultation and to do this succinctly. In these time-limited conditions, they are unlikely to also discuss the positive qualities of life such as the pleasure they get from playing with children, listening to Mozart, or seeing a dewdrop on a leaf. Consequently, many measures have tended to focus on assessing symptoms and dysfunctional aspects of life to the exclusion of the good life. This not only gives health professionals the view that the QoL of their patients is worse than it is, but it also causes the patients who complete the questionnaire to report feeling miserable after focusing unduly on the bad things instead of a more balanced life. Even highly disabled, chronic-pain patients entering a pain management program after many surgical and pharmacological treatments report that chronic pain affects only half the important qualities of life when a range of positive and negative issues are presented for evaluation. The good life was acknowledged by Aristotle and several eighteenth-century philosophers, but has only recently been endorsed again as pertinent to outcomes measurements in health care. Hence, more recent definitions of QoL have sought to emphasize a positive orientation: "The essential characteristics of life, which in the general public is often interpreted as the positive values of life, or the good parts of life, or the total existence of an individual, group or society" (Lindstrom, 1992).

The Example of Sleep as a Quality of Life

Another debate concerns the type of information that constitutes the best measure of QoL, and here sleep provides a suitable example. Researchers have assessed sleep-related QoL by tracing electroencephalograph (EEG) waves, considering this to be an objective indicator. Others assess the timing of sleep and wakening, insomnia patterns, medication use, and so forth, in looking at people's objective perceptions of their sleep. However, we can argue that information derived from either of these channels of inquiry tells us only indirectly about people's QoL. These types of information require further interpretation by the researchers, and their own assumptions may lead

them to draw an erroneous conclusion. Would the people you know see your QoL as good or bad if you told them you slept 3 h a night? What about 12 h a night? Without additional information about whether that sleep was refreshing, it is not possible to draw an accurate conclusion. For this reason the most reliable way to avoid this pitfall is to ask subjective QoL questions such as whether people find their sleep to be refreshing or not. Conclusions about these qualities cannot, therefore, be deduced; direct questioning about the meaning of events is required. Like stressful life events (e.g., divorce), it is the interpretation that matters. Although all these three types of information are themselves valuable and can be used together to form a fuller picture, we contend that QoL is best measured as a subjective phenomenon and that the use of the term should be restricted to this type of information.

The World Health Organisation Definition

Questions have been raised about whether QoL is a global phenomenon or just a Western concept. Although culture has sometimes been included in QoL models, it is usually seen as extraneous and additional, rather than taking a central role in explaining the diverse ways in which different cultural groups interpret QoL. Concerned to understand QoL relating to global health, the World Health Organisation views culture as a "lens" through which people interpret the experiences that affect their QoL. It defines QoL as "an individual's perception of their position in life, in the context of the culture and value systems in which they live, and in relation to their goals, expectations, standards, and concerns" (World Health Organisation Quality of Life Group, 1994: 43). This unusual definition also implies that comparisons are used in helping to make decisions about QoL; that in evaluating their own QoL, people consider how far they have achieved their goals, expectations, and standards, in addition to considering problematic concerns.

Dimensions of Quality of Life

So what does a good QoL consist of, and is there consensus about whether these dimensions are universal? Many hundreds of QoL questionnaires have included the two broad components of physical and mental health identified by Descartes as essential. However, physical health is operationalized in some assessments in terms of symptoms; adding symptoms does not give a clear view of people's QoL because, paradoxically, patients with many symptoms often report a relatively good QoL and, conversely, those

with very few symptoms report it to be poor. Measures also often include estimates of independence or functional status, such as mobility, activities of daily living, and so on. Physicians have traditionally favored those QoL dimensions for which they can independently inspect corroborating evidence alongside the subjective reports. Some issues such as spirituality have largely been ignored as being too difficult to measure and/or irrelevant to health and health care.

The universality of QoL and its dimensions have been addressed by the World Health Organisation Quality of Life (WHOQOL) Group, which has confirmed that physical, psychological, independence, social relations, environmental and spiritual dimensions are important in a wide range of cultures and settings and to sick and well people. Within these broad domains, the group confirms 25 detailed aspects (facets) of QoL that are reliable and valid internationally. A series of novel procedures was used by the WHOQOL Group to create several multilingual instruments and a short version – the WHOQOL-Bref – is now available in over 50 different languages. In addition to the main measures, add-on modules of extra questions have recently become available to support work on spirituality, human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS), older adults, and pain. The WHOQOL results provide support for the view that there is broad recognition and universal agreement about a concept called QoL, even though it may not be labeled as such, and that there are cultural variations in the relative importance of the dimensions included, as well as in how these issues themselves are expressed.

Uses of Quality of Life Assessments

QoL assessments are used for monitoring outcomes (e.g., after policy change). In particular, they are valued widely as a suitable outcome measure in randomized controlled clinical trials. Although outcomes are the main focus here, it is worth noting that QoL can also be investigated as a process. Because psychometric measures are designed to tap into outcomes, the results from these instruments provide merely a snapshot of QoL at one time in an ever-adjusting process. Different theories and techniques may be needed to investigate how and when QoL changes over time and precisely which factors (e.g., stress and eustress) affect positive and negative changes.

Further Reading

Coen, R. F. (1999). Individual quality of life and assessment by carers or 'proxy' respondents. In: Joyce, C. R. B.,

- O'Boyle, C. A. & McGee, H. (eds.) *Individual quality of life: approaches to conceptualization and assessment*, pp. 185–196. Amsterdam: Harwood Academic.
- Inglehardt, R. and Klingemann, H. D. (2000). Genes, culture, democracy and happiness. In: Diener, E. & Suh, E. M. (eds.) *Culture and subjective well-being*, pp. 185–218. Cambridge, MA: MIT Press.
- Layard, R. (2005). *Happiness: lessons from a new science*. London: Penguin Books.
- Lindstrom, B. (1992). Quality of life – a model for evaluating health for all: conceptual considerations and policy implications. *Sozial und Preventivmedizin* 37, 301–306.
- Shin, D. C. and Johnson, D. M. (1978). Avowed happiness as an overall assessment of quality of life. *Social Indicators Research* 5, 475–492.
- Skevington, S. M. (2002). Advancing cross-cultural research on quality of life: observations drawn from the WHOQOL development. *Quality of Life Research* 11, 135–144.
- Skevington, S. M., Lotfy, M. and O'Connell, K. A. (2004). The World Health Organisation's WHOQOL-BREF Quality of Life assessment: psychometric properties and results of the international field trial – a report from the WHOQOL Group. *Quality of Life Research* 13, 299–310.
- Skevington, S. M., Sartorius, N., Amir, M., et al. (2004). Developing methods for assessing quality of life in different cultural settings: the history of the WHOQOL instruments. *Social Psychiatry and Psychiatric Epidemiology* 39, 1–8.
- Skevington, S. M. and Wright, A. (2001). Changes in the quality of life of patients receiving anti-depressant medication in primary care: validating the WHOQOL-100. *British Journal of Psychiatry* 178, 261–267.
- Staquet, M. J., Hayes, R. D. and Fayers, P. M. (1998). *Quality of life assessment in clinical trials: methods and practice*. Oxford: Oxford University Press.
- Wenger, N. K., Mattson, M. E., Furberg, C. D., et al. (1984). Assessment of quality of life in clinical trials of cardiovascular therapies. *American Journal of Cardiology* 54, 908–913.
- World Health Organisation Quality of Life Group (1994). The development of the World Health Organisation Quality of Life assessment instrument (The WHOQOL). In: Orley, J. & Kuyken, W. (eds.). *Quality of life assessment: international perspectives*, pp. 41–60. Berlin: Springer-Verlag.
- World Health Organisation Quality of Life Group (1995). The World Health Organisation Quality of Life assessment (WHOQOL): position paper from the World Health Organisation. *Social Science & Medicine* 41, 1403–1409.
- World Health Organisation Quality of Life Group (1998). The World Health Organization Quality of Life assessment (WHOQOL): development and general psychometric properties. *Social Science & Medicine* 46, 1569–1585.

R

Racial Harassment/Discrimination

D R Williams and S A Mohammed

University of Michigan Ann Arbor, MI, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction: Discrimination Persists
Studies of Discrimination and Health
Understanding the Context of Discrimination and
Responses to It
Measurement Issues
Pathways from Discrimination to Health
Conclusion

Introduction: Discrimination Persists

Discrimination and Racial Harassment

Several lines of evidence converge to suggest that discrimination is an important feature of life in race-conscious societies in which race is an important determinant of access to a broad range of societal benefits and rewards. First, social psychological theory identifies discrimination as a key aspect of intergroup relationships that serve to reinforce the symbolic boundaries that separate one social group from another. Second, there is broad recognition that discrimination persists for socially disadvantaged populations. The American public acknowledges that a high level of discrimination and prejudice exists. National public opinion data reveal that most Americans concede that racial minorities suffer from a fairly high level of racial bias. Blacks are perceived as experiencing the most discrimination, followed by Hispanics and American Indians, and then Asians. They also perceive that some White ethnic/religious groups such as Jews, Italians, and Catholics experience more discrimination than Whites in general. In one national study, for example, two-thirds of the American population reported that blacks suffer a lot or a tremendous amount of discrimination.

Some of the best evidence of the persistence of discrimination comes from audit studies in which

carefully matched Black and White applicants with identical qualifications apply for jobs or housing. Audit studies in employment, for example, find that discrimination favors the White over a Black applicant in one in five audits. Other audit studies suggest that discrimination persists for Blacks and other minorities in the purchase of a broad range of goods and services. For example, one carefully executed study found racial differences in the best price offered for the purchase of a new car. In this study, Black and White testers were sent to automobile dealerships to negotiate the price of a new car. Although all the testers were carefully trained to follow the same script, there were large racial and gender differences in the final price offered for the new car, with White males offered the lowest prices, followed by White females, African American males, and African American females. Similarly, a recent report from the Institute of Medicine documented systematic and pervasive discrimination in the quality and intensity of medical treatment received by Blacks and other minorities in the United States. Thus, there is abundant evidence that discrimination persists in contemporary society. However, not all experiences of discrimination are visible to their intended target. Frequently, victims of discrimination lack all of the details about an interpersonal transaction that they need to know in order to establish that discrimination has occurred. Although self-reports of bias understate the full extent of exposure to discrimination, they capture an important part of discriminatory experiences that are likely to be consequential for health.

Discrimination as a Stressor

This article focuses on aspects of discrimination that are perceived by individuals in society. Subjective experiences of discrimination or unfair treatment are viewed as an important type of stressful life experience that many traditional measures of stress fail to assess. Equity theory suggests that perceptions of unfair treatment can lead to negative emotional reactions and psychosomatic symptoms. Discrimination

can also be stressful because it can reinforce negative social stigma and highlight low social regard. Qualitative descriptions of experiences of discrimination indicate that they can induce considerable levels of psychological distress. Laboratory studies have examined the experimental manipulation of unfair treatment and have found that it leads to elevated levels of psychological distress for a broad range of social groups. Other studies have directly measured physiological responses to racially stressful material in the laboratory setting. These studies have used mental imagery and videotaped vignettes of discriminatory behavior. Most, but not all, of these studies have found that exposure to discrimination leads to cardiovascular and psychological reactivity. An important methodological strength of laboratory studies is that they do not rely on respondents' self-reports of discrimination. However, observed effects in the laboratory are short term, making the extent to which acute physiological arousal under laboratory conditions generalizes to chronic elevation of stress processes and mechanisms in real life unclear. Some recent research documents that high reactivity to acute stressors in laboratory settings is predictive of elevated rates of hypertension several years later. Most of the early laboratory studies focused on Blacks. It is not clear from these studies, however, whether the physiological reactivity to racism-linked experiences is unique or enhanced in comparison to reactivity to other types of stressors.

Studies of Discrimination and Health

Mental health status is the most frequently used health measure in studies of discrimination. Multiple mental health indicators have been used. These include measures of well-being, self-esteem, perceptions of control, psychological distress, anger, and specific psychiatric disorders such as major depression, generalized anxiety, and substance use. These studies have found an inverse relationship between discrimination and mental health – higher levels of discrimination were associated with poorer mental health status. Most of the early research studies in this area were based in the United States and tended to focus on African Americans. Recent studies have documented similar patterns of associations for other minority groups in the United States. For example, studies of Mexican Americans, Chinese Americans, and American Indians have found that perceived discrimination is associated with higher levels of psychological distress. A growing number of international studies documented similar patterns. Studies of Asian immigrants in Canada; Iranians, Turks, and Moroccans in the Netherlands; immigrants in

Finland; Blacks, Asians, and Arabs in Ireland; and multiple ethnic minority populations in the United Kingdom found that discrimination is positively associated with psychological distress. Ongoing research is also exploring these relationships in Australia, Brazil, New Zealand, and South Africa.

Measures of perceived discrimination have also been studied in relationship to multiple measures of physical health. Studies using a single-item global self-rated health measure and other self-ratings of health or checklists of chronic illnesses have also fairly consistently documented that discrimination is related to poorer health status. Two studies found that discrimination makes an incremental contribution above socioeconomic status (SES) in explaining Black-White differences in self-reported health status. Several studies have also examined the association between discrimination and blood pressure levels or hypertension. Some studies found a clear positive association between discrimination and elevated blood pressure, whereas others found that this effect varies with the coping style, sex, SES, or race. Still other studies found no association between discrimination and blood pressure.

Recent research has also paid attention to a broader range of health outcomes. Whereas perceptions of discrimination were unrelated to self-reported heart disease in some early studies, recent research documented a positive association between chronic discrimination and the onset of subclinical heart disease (intima media thickness and coronary calcification). Recent studies also found that maternal exposure to discrimination was associated with preterm deliver and low birth weight. Several studies have also examined the association between perceived racial bias and health behaviors. They have found a positive association between discrimination and cigarette use and alcohol consumption.

Understanding the Context of Discrimination and Responses to It

Some of the inconsistency in findings in studies of discrimination and health could be due to inadequate attention to characterizing the social and psychological context in which discrimination occurs as well as the specific characteristics of the discriminatory incident. First, personality characteristics could predict variations in the association between discrimination and health. Beliefs about the self, such as self-esteem, mastery, racial consciousness, and identity might affect an individual's appraisal and thus the potential threat of a discriminatory experience. Second, contextual variables such as social support could also enhance an individual's capacity to cope and respond

to discriminatory experiences. At the present time, we do not have a clear understanding of the contextual factors that can foster either vulnerability or resilience to perceived bias.

Third, specific characteristics of the discriminatory experience and its particular context can be important determinants of its impact. Key aspects of discriminatory experiences include the type or domain in which it occurs, the magnitude of the incident, its temporal characteristics (acute versus chronic), and the nature of the relationship between this stressor and other race-related and non-race-related stressors. Inadequate attention has also been given to the extent to which the effects of discrimination may vary by the social characteristics of the perpetrator. For example, some limited evidence suggests that the negative health impact of perceived discrimination on African Americans is greater when the perpetrator is also African American than when the perpetrator is White.

Fourth, it is also likely that there are complex interactions between exposure to discrimination and particular coping strategies or styles. In addition to examining traditional problem-focused and emotion-focused measures of coping from the stress literature, research on discrimination needs to attend to discrimination-related measures. Heightened vigilance is one such measure. Because racism is often deeply embedded in the structure and culture of society, it can pose an ever-present threat for socially stigmatized groups. Some individuals respond to this by engaging in psychological and behavioral anticipatory coping strategies to mitigate the perceived dangers in the environment. This can lead to heightened physiological arousal that can adversely affect health. Using denial as a coping strategy is a theoretically important but difficult concept to measure. The literature suggests that, because experiences of discrimination can pose threats to the self, some individuals respond to its occurrence by minimizing or even denying its occurrence. There is also the suggestion in the literature that denying the occurrence of bias is damaging to health.

Measurement Issues

A prerequisite for understanding the association between discrimination and health is the development of valid and reliable measures that fully characterize exposure to discrimination. The literature identifies several issues that affect the appropriate assessment of discrimination.

Perceived Discrimination Is Subjective

An important issue in the assessment of discrimination is the challenge of obtaining valid and reliable

indicators of subjective phenomena such as experiences of discrimination. It is possible that the salience of racial identity or perceived racism could lead some minority group members to perceive incidents as racist that may not be or, worse, could lead to the development of a mind-set in which they perceive experiences to be products of racism when in fact they were not. In fact, some limited evidence suggests that the use of discrimination terminology, in which race or gender is made salient in the assessment of discrimination, can induce respondent biases. This work emphasizes that questions with embedded terminology require both a description of an experience as well as an interpretation of that experience. Moreover, in the cooperative context of a research interview a respondent may make inferences about the questioner's intent and interests that could lead the respondent to try to provide information that is consistent with the questioner's perceived intent. All of this can lead to interpretive response bias that could produce higher estimates of discrimination, driven by the specific terminology used. Thus, questionnaires that ask repeated questions about "racial discrimination" or experiences "because of your race" could produce demand characteristics in which the respondent believes that it is desirable to the interviewer to report such experiences. This could lead to overreports of discrimination.

One promising approach in measuring discrimination seeks to reduce the salience of race by first asking respondents whether they have been treated "unfairly" in multiple domains of life. After respondents have endorsed an experience of unfair treatment, they are asked to attribute a reason. Potential reasons include race and ethnicity, but also allow for other factors such as gender, age, religion, weight, and sexual orientation. This approach enables respondents to report on all instances of unfair treatment, but allows the researcher to separate instances attributed to race from those linked to other reasons. Relatedly, this measurement approach allows for the evaluation of the extent that discrimination based on race is more deleterious to health than discrimination based on other social factors. Inadequate attention has been given to the relative impact of racial discrimination compared to other types of discrimination in the extant research to draw firm conclusions. However, some studies suggest that it is the generic perception of unfair treatment that is deleterious to health, irrespective of the specific attribution regarding the cause of the incident. Research is needed to identify the optimal approaches to the measurement of racial discrimination.

The available evidence suggests that concerns about the overreporting of discrimination may be

overblown. First, making race salient in the assessment of discrimination can lead to underreporting of discrimination. Respondents vary in their thresholds of what constitutes discrimination and may fail to report as discriminatory incidents that were not perceived as very serious. Second, respondents appear to interpret the concept of discrimination as intended by researchers and self-reports of discrimination are consistent with objective experiences. Third, two studies have documented that baseline mental health status (psychological distress and major depression) is unrelated to subsequent reports of discrimination. Fourth, because reporting discrimination can negatively affect self-esteem and perceptions of control, at least some stigmatized group members are likely to minimize and deny experiences of discrimination. Thus, underreporting is as important a threat to the validity of self-reports of discrimination as overreporting. Nonetheless, it is important for researchers to adjust reports of discrimination for potentially confounding psychological characteristics such as social desirability, neuroticism, and other indicators of negative affect. Such adjustment strengthens the methodological rigor of a study and increases the likelihood that observed associations between perceived discrimination and health are not distorted by underlying psychological factors. Including such controls is especially important when the measure of health status is also based on self-report.

Several strategies that have been used to improve the accuracy of the reporting of stressors should also be applied to the study of discrimination. These include the use of cues to memory (such as visual representations and reminders of personally salient events), wording the questions in ways that clearly indicate the domain of the experience being captured, and using a life-events calendar to facilitate dating the onset and resolution of experiences of discrimination.

Perceived Discrimination Is Multidimensional

Capturing the exposure to discrimination in its full multidimensional complexity and assessing the cumulative burden of such exposure over the life course is also another major challenge. Some early studies of discrimination used single-item measures. Such assessment understates the actual level of discrimination. However, many current studies assess exposure to discrimination using retrospective recall to capture exposure within a 30-day, 1-year, 3-year or lifetime time frame.

Comprehensively capturing discrimination requires its assessment in multiple areas of social life. Like other stressors, discrimination is multidimensional and should be assessed in all relevant domains. The stress literature indicates that the most commonly

measured types of stressors include life events, chronic stress, and daily hassles. Life events are major discrete stressors that are readily observable. Daily hassles are minor but often chronic irritations. Chronic stressors are ongoing problems that are typically role-related. There are analogues to all of these types of stressors in current measures of discrimination. Major acute experiences of racial bias are the most commonly assessed type of discriminatory experience. The literature on stress indicates that chronic stressors may be stronger predictors of the onset and course of illness than acute life events. Scales such as the Everyday Discrimination Scale attempt to capture persistent and recurring everyday chronic minor experiences. Future research needs to give more attention to capturing the frequency and duration of chronic discrimination in multiple social contexts, such as employment, educational, and public settings.

Other major types of stressors include traumas, nonevents, and macro-stressors. Developing analogs of these in the assessment of discrimination is a promising direction for future research. Traumas are acute stressors that are severe and overwhelming in impact, such as being kidnapped or raped. Macro-stressors are large-scale systems-related stressors such as economic recessions. Nonevents are desired and expected experiences that fail to occur. The stress literature indicates that the various types of stressors can have independent effects on health. Therefore, assessing all relevant types of stressors is a requisite to an assessment of the full impact of stress on health.

A small but growing body of research on historical trauma and its intergenerational effects on the health of American Indians illustrates the benefit of expanding our definitions of race-related stressors. Historical trauma is viewed as a cumulative and collective psychological wounding over the life span and across generations as the result of the history of genocide and assimilation that American Indians experienced at the hands of Europeans. Studying historical trauma among American Indians is not unlike existing studies of other generational group traumas, such as studies of the effects of the Jewish Holocaust on the physical and mental health, as well as the social, economic, and political contexts, of descendants of Holocaust survivors. It has been argued that historical trauma contributes to a host of issues in the American Indian population, including unresolved grief, alcoholism, depression, anxiety, high rates of suicide, homicide, problematic gambling behaviors, domestic violence, child abuse, poverty, low education levels, and various physical diseases. The literature recognizes that although the concept of historical trauma is theoretically intuitive, there are challenges

in distinguishing historical versus contemporary experiences, understanding how historical trauma is imparted across generations, and defining and measuring the levels of historical trauma. However, scales with good psychometric properties have been developed to assess historical trauma and find that as many as one-half of American Indians think regularly about these historical losses. Moreover, recent empirical studies have found support for an inverse association between historical trauma and health. Clinical interventions to address historical trauma have also been developed. A promising area of research would be to examine the generalizability of this pattern across health outcomes and American Indian cultures, as well as, seeing whether other highly visible traumatic historical events have similar effects for other groups.

Capturing Life Course Exposure

There are important measurement challenges in terms of identifying the appropriate questions that could capture the duration and frequency of discrimination over the life course. An isolated experience of discrimination, just like a random stressor, is unlikely to have long-term negative effects on health. The literature suggests that the key issue is being able to identify discriminatory experiences that are enduring and chronic. A life course perspective is critical in assessing the cumulative burden of discrimination over an individual's lifetime. Measures that do an excellent job of capturing exposure to all of the relevant types of discrimination do not currently exist. Overcoming the challenges linked to declines in the accuracy of recall with the passage of time will not be easy, but this is crucial to obtaining better measures of lifetime exposure to discrimination. Such data are necessary to begin to identify the lag times between exposure to discriminatory stressors and adverse changes in health. Life course assessment of discrimination must also be coupled with an equally comprehensive measurement of other types of stressors. The present-day literature does not provide a clear understanding of how experiences of discrimination relate to other stressors and combine, in additive and interactive ways, to affect specific health outcomes and trajectories of health status over time.

Attributional Ambiguity

Perceived discrimination is based on the perception of the individual. The subjective nature of discrimination and the ambiguity that occurs in much interpersonal interaction can lead to uncertainty regarding the attribution of specific incidents of bias or unfair treatment. It is possible that having to make sense of ambiguous social interactions can be physiologically stressful. The worry and rumination regarding the

causes of experiences of discrimination may be an added burden that can be consequential for the health of nondominant group members. It is thus possible that the degree of ambiguity attending the perception of a discriminatory experience could negatively affect health. This issue deserves careful examination in future research.

Pathways from Discrimination to Health

Researchers have generally found a weaker association between discrimination and health for measures of physical health than for measures of mental health. Physical health outcomes typically assessed are chronic conditions involving complex etiological factors and varying patterns of development and progression over long periods of time. Greater attention needs to be given to the hypothesized pathways by which discrimination might affect health and to identifying the appropriate points along that pathway at which the individual is more vulnerable to exposure to discrimination and other stressors.

The stress literature reveals that one way that stressors can affect health is by giving rise to negative emotional states, such as anxiety and depression, which in turn can directly impact biological processes or patterns of behavior that affect disease risk. Accordingly, one of the pathways by which discrimination can affect physical health outcomes may be indirectly, through psychological distress. Thus, discrimination may lead to elevated psychological symptoms, which, in turn, may lead to chronic physiological arousal. In addition, the negative emotional states created by experiences of discrimination might lead to health behaviors that may ultimately affect disease risk, such as impaired sleep patterns, decreased physical activity, increased substance use, and consumption of more food than usual. It was noted earlier that prior research indicates that discrimination is positively associated with alcohol consumption and cigarette smoking. These and other health-related behaviors may be a pathway by which perceived bias has health consequences. The stress of discrimination and the negative emotional states created by it could also lead to lower levels of adherence to medical recommendations. This latter mechanism has not yet been explored in the literature. The potential effects of discrimination and other stressors on medical compliance emphasize the need to assess the contribution of discrimination not only to the onset of disease but also to its severity and course.

Our understanding is currently limited regarding how exposure to discrimination leads to changes in particular biological responses and health behaviors. Further research is needed to identify the

physiological mediators of the effects of discrimination so that the specific biological pathways by which discrimination can adversely affect health may be studied. Such research should assess the conditions under which specific physiological systems, such as the cardiovascular, neuroendocrine, and immune systems, are affected by particular types of discrimination. We are presently unaware of the genetic and psychological factors that can lead some organ systems to be more vulnerable than others to the effects of discrimination on health.

Conclusion

Scientific research indicates that perceived discrimination is associated with poor health status, with the association being strongest for mental health. However, existing evidence does not clearly indicate the extent to which exposure to perceived discrimination leads to increased risk of disease, the conditions under which this is more or less likely to occur, or the underlying mechanisms and processes linking this stressor to health status. Advances have been made in the conceptualization and measurement of discrimination, but there is still more work to be done. The most urgent need is for the theoretical identification and the empirical verification of the plausible pathways by which experiences of bias can affect various health outcomes. At the same time, this field of research is clearly moving from infancy to childhood. The literature has now identified major gaps in this area, as well as promising conceptual, methodological, and analytic tools that are needed for the rigorous examination of the association between perceived discrimination and multiple indicators of health.

See Also the Following Articles

Cultural Factors in Stress; Refugees, Stress in; Ethnicity, Mental Health.

Further Reading

Blank, R. M., Dabady, M. and Citro, C. F. (eds.) (2004). *Measuring racial discrimination: panel on methods for*

assessing discrimination. National Research Council Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Clark, R., Anderson, N. B., Clark, V. R., et al. (1999). Racism as a stressor for African Americans: a biopsychosocial model. *American Psychologist* 54(10), 805–816.

Dion, K. J. (2001). The social psychology of perceived prejudice and discrimination. *Canadian Psychology* 43(1), 1–10.

Fix, M. and Struyk, R. J. (1993). *Clear and convincing evidence: measurement of discrimination in America*. Washington, DC: Urban Institute Press.

Gomez, J. P. and Trierweiler, S. (2001). Does discrimination terminology create response bias in questionnaire studies of discrimination? *Personality and Social Psychology Bulletin* 27(5), 630–638.

Harrell, J., Hall, S. and Taliaferro, J. (2003). Physiological responses to racism and discrimination: an assessment of the evidence. *American Journal of Public Health* 93(2), 243–248.

Harrell, S. (2000). A multidimensional conceptualization of racism-related stress: implications for the well-being of people of color. *American Journal of Orthopsychiatry* 70(1), 42–57.

Krieger, N. (1999). Embodying inequality: a review of concepts, measures, and methods for studying health consequences of discrimination. *International Journal of Health Services* 29(2), 295–352.

Krieger, N., Smith, K., Naishadham, D., et al. (2005). Experiences of discrimination: validity and reliability of a self-report measure for population health research on racism and health. *Social Science and Medicine* 61(7), 1576–1596.

Paradies, Y. (2006) A systematic review of empirical research on self-reported racism and health. *International Journal of Epidemiology* 35, 888–901.

Whitbeck, L. B., Adams, G. W., Hoyt, D. R., et al. (2004). Conceptualizing and measuring historical trauma among American Indian people. *American Journal of Community Psychology* 33(3–4), 119–130.

Williams, D. R. and Neighbors, H. W. (2002). Racism, discrimination, and hypertension: evidence and needed research. *Ethnicity and Disease* 11, 800–816.

Williams, D. R., Neighbors, H. W. and Jackson, J. S. (2003). Racial/ethnic discrimination and health: findings from community studies. *American Journal of Public Health* 93(2), 200–208.

Rape See: Sexual Assault.

Receptors, Classical See: Membrane Glucocorticoid Receptors; Glucocorticoid Receptor Mutants and Polymorphisms; Corticotropin-Releasing Factor Receptors; Corticosteroid Receptors; Corticotropin Releasing Factor Receptor Deficiency in Mice.

Recovery from Stress

S Sonnentag

University of Konstanz, Konstanz, Germany

C Fritz

Bowling Green State University, Bowling Green, OH, USA

© 2007 Elsevier Inc. All rights reserved.

The Recovery Concept
Relevance of Recovery
Factors Influencing Recovery
Recovery from Job Stress
Conclusion

The Recovery Concept

Research differentiates between two core phenomena related to the stress process: reactivity and recovery. Reactivity refers to the ensuing physiological and psychological changes caused by a stressor and is reflected in the elevation of strain indicators compared to a baseline. Recovery refers to the processes during the poststressor period, when strain indicators return to their baseline levels. Slow recovery from the stressor can be seen as one instantiation of prolonged activity. The duration of the recovery process differs widely among the various indicators. For example, under normal conditions, blood pressure and heart rate recover within a few minutes after the termination of the stressor. However, it can take much longer until neuroendocrine parameters reach the prestressor levels. For example, a literature review based on a total of 77 studies concluded that norepinephrine recovery was complete within 1 h after the termination of the stressor, whereas adrenaline levels often remained elevated, particularly when urinary measures are used. Meta-analytic evidence suggests that salivary cortisol recovery is completed within 40–60 min after the termination of the stressor. Subjective variables such as mood may take hours to return to prestressor levels. Overall, studies show that the time needed

for complete recovery largely depends on person-specific and situation-specific factors.

In laboratory research, recovery most often is initiated by a poststressor rest period, during which the stressor is absent. The degree of recovery can be analyzed with several methods, with curve-fitting and related approaches being the most promising and reliable ones, particularly for analyzing cardiovascular recovery. In field research, it is more difficult to identify recovery periods exactly. For example, when examining recovery from job stress we might regard the weekend as the recovery period. However, this perspective assumes that during the weekend no other stressors (e.g., from the family domain) are present.

Relevance of Recovery

Quick recovery after the termination of a stressor is assumed to be an adaptive response pattern, whereas slow recovery is assumed to be maladaptive. Slow recovery implies that stress responses such as elevated blood pressure or increased levels of catecholamines and cortisol persist over longer periods of time and therefore might harm the organism. For example, McEwen argued that a prolonged allostatic response is one of the core processes that contributes to allostatic overload and subsequently might lead to illnesses such as cardiovascular diseases.

Empirical evidence that the recovery process is crucial for health comes for example from a longitudinal study on nurses. Elevated cortisol levels after work (but not during work) predicted health-care costs 5 years later. Pieper and Brosschot summarized empirical studies demonstrating that prolonged cardiovascular recovery after a stressor predicted general heart rate and blood pressure levels 3–6 years later. A recent study with a community sample points in a similar direction. Individuals who showed slow systolic blood pressure recovery from psychological tasks had elevated clinical systolic and diastolic blood pressure levels 3 years later, even after controlling for the initial clinical blood pressure levels and other variables such

as body-mass index, education, and use of tobacco products. However, before we draw any causal conclusions about slow recovery being the cause of elevated clinical blood pressure levels, we must take into account that cardiovascular risk factors such as race or lack of fitness tend to be related to slow cardiovascular recovery and might cause both slow recovery and high chronic blood pressure.

Factors Influencing Recovery

Research suggests that slow recovery is associated with personal as well as situational factors. For example, individuals with negative emotional dispositions show elevated cardiovascular parameters during poststressor periods. The findings are particularly impressive with respect to hostility. Specifically, hostile persons show elevated levels of systolic blood pressure during sleep. Similarly, in experimental settings that involve the confrontation with a stressor, high hostility is associated with slow systolic blood pressure recovery. In addition to hostility, low self-enhancement (i.e., the tendency of not experiencing highly favorable cognitions about oneself) is related to slow cardiovascular recovery during the poststressor period.

Not only cardiovascular recovery, also cortisol recovery might be hampered by personal factors. Meta-analytical findings show that clinically depressed individuals show higher cortisol levels during the poststressor period than do nondepressed individuals (but not higher baseline or reactivity levels). Moreover, people with low self-esteem show elevated cortisol recovery levels after the experience of failure. In addition, gender and race might also play a role with respect to recovery processes.

With respect to situational factors associated with slow recovery, we can differentiate between acute situational factors present in the specific setting and more chronic situational variables. Acute specific situational factors related to slow recovery refer to the type of stressor and to processes during the poststressor period. Compared with other acute stressors, emotionally involving and anger-provoking situations are associated with a greater delay in cardiovascular and cortisol recovery. It is important to note that slow recovery after the provocation of anger seems to be particularly prominent in men, as opposed to women.

Adrenaline recovery is prolonged after highly demanding situations characterized by a low degree of controllability. With respect to field situations, research suggests that, particularly after periods of high workload and high work intensity, the recovery of adrenaline is delayed.

When it comes to activities that affect recovery during the poststressor period, distraction from the preceding stressor seems to be highly important. Situations that offer distractions from the stressor support faster recovery than situations that allow individuals to ruminate about the stressful event. In addition, music – particularly classical music – helps to reduce blood pressure levels during the poststressor period.

In addition to factors associated with the situation in which the stressor occurs, chronic stressors and other chronic factors may impact recovery processes. People who are exposed to chronic stressors show delayed blood pressure and immunological recovery from acute stressors. Not only chronic stressors but particular stressors in combination with lack of resources are related to slow recovery. For example, in a study with firefighters, it was found that individuals who had experienced major life events during the past 12 months and who received only little social support in various life domains displayed particularly slow cardiovascular recovery. Results from several studies in the work context point in a similar direction. Employees in high-strain jobs (i.e., jobs associated with a high level of demands and low control at the same time) show prolonged cardiovascular and cortisol recovery. However, note that this picture is not consistent across all studies.

Prolonged recovery in people facing stressful life situations is also reflected in research using subjective measures. Individuals in stressful jobs experience a higher need for recovery and feel that it takes them longer to unwind from work. The subjective need for recovery is related to both neuroendocrine parameters and indicators of impaired well-being such as psychosomatic complaints and burnout.

To sum up, personal variables such as hostility or depression are related to slow recovery. Among the most important situational variables that are associated with prolonged recovery are situations that are anger-provoking and emotionally involving and situations characterized by high workload. In addition, chronic background stressors and lack of resources (e.g., low social support) make it more likely that recovery is delayed. With respect to factors that promote recovery, distraction from the stressful situation seems to be highly important.

Recovery from Job Stress

Compared to the vast research on stress at work, research on recovery from work demands is still rare. Often, the research just implicitly assumes that being away from the workplace (such as during vacation, weekends, or during the evening) leads to

recovery from work demands. When an individual is no longer confronted with work demands or work-related demands, recovery can occur. Thus, during leisure time individuals should avoid any work-related demands in order to recover from work demands. We review next empirical evidence on recovery from work demands focusing on recovery during vacations and weekends as well as on recovery during the evening of a workday.

Vacations and Weekends: Effects on Well-Being and Performance-Related Aspects

Vacations as a time away from work give the opportunity for people to recover from work demands and to build new resources. Empirical research on different employee groups and from different countries indicates that vacations are associated with an increase in employee health and well-being. For example, measuring levels of burnout before, during, and after vacation, a study in Israel found a decrease in self-reported burnout after vacation. Furthermore, vacations seem to have a positive impact on psychophysiological indicators such as the immune system or even mortality.

So far, there is only little research on effects of vacation on performance-related outcomes. One study examined the impact of vacations on individual perceptions of effort and task performance at work. Comparison between pre- and postvacation measures indicated no change in task performance. However, study participants reported needing much less effort to get their work done after vacation compared to before vacation. A study measuring changes in employee absenteeism rates found a decrease in absenteeism after vacation compared to prevacation. In summary, although there are the first indications that vacations have a positive effect on performance-related variables, more research seems necessary.

Research further suggests that the positive effects of a vacation subside within a short while after the employee's return to work. This means that well-being and performance-related outcomes return to their prevacation or chronic level. Such fade-out effects have been found within the first few weeks after vacation. Future research should include additional measurement occasions after vacation to better understand the fade-out of recovery effects over time and to possibly reveal factors that may impair or accelerate the fade-out process.

In addition to research on the general effects of vacation, there is some research examining specific variables in the context of a vacation that may promote recovery from work demands. For example, a study found that thinking positively about one's job, pursuing activities that include the experience

of mastery, and relaxation during vacation were associated with higher levels of individual well-being after vacation. In contrast, reflecting about one's job in a negative way and facing high levels of nonwork hassles during vacation, as well as high levels of workload directly after vacation, were detrimental for employee well-being and for performance-related variables directly after vacation. In addition, research revealed a positive relationship between people's satisfaction with the vacation and job and their satisfaction with life after vacation. However, more research is necessary to more clearly understand the processes during vacation that underlie changes in well-being and performance-related outcomes.

So far, research on recovery during the weekend is rare. However, one study found effects of specific experiences during the weekend on well-being and performance-related variables after the weekend. Specifically, reflecting positively about one's job and social activities seemed to promote recovery during the weekend, whereas nonwork hassles during the weekend seemed to be detrimental for recovery from work demands. There clearly is the need for more research on recovery from work demands during weekend to better understand the processes that enhance or hinder recovery processes and subsequent well-being and performance.

Recovery during the Evening of a Workday: Effects on Well-Being and Performance-Related Aspects

Recently, research has started examining how employees unwind from work during the evening of a workday. Research indicates positive effects of recovery during the evening on feelings of well-being at bedtime. Specifically, low-effort activities and social and physical activities seem to promote recovery and increase subsequent well-being. In contrast, household and work-related activities seem to impair recovery from work demands, as indicated by lower levels of well-being at bedtime. In addition, generally experiencing off-work activities as positive or as recovering seems to positively impact well-being at bedtime.

With regard to performance-related aspects, a study demonstrated that higher levels of individual regeneration in the morning due to recovery activities the evening before were positively associated with work engagement and proactive behaviors during the workday. Future research should examine the processes underlying those positive effects in more detail.

There is some empirical evidence that the choice of recovery activities during the evening and their effectiveness may be affected by work conditions. For example, job stressors (e.g., time pressure and low

control) seem to be associated with higher levels of job-related rumination after work, more difficulties detaching from work, and less time spent on sport activities. In addition, long work hours seem to be related to a lower capability to relax after work and again to more difficulties detaching from work and less time spent on sport activities.

Finally, the choice of recovery activities and their effectiveness may be affected by interindividual differences. For example, employees who are high in job involvement seem to detach less from work in the evening than employees low in job involvement. Research further indicates a positive relationship between the individual expectation to benefit from the recovery activities and detachment from work during the evening.

Conclusion

Research on stress recovery is a highly important topic that addresses both fundamental and applied questions. With respect to more fundamental research, theoretical models and empirical studies on recovery will add to our full understanding of how stressors affect the organism. Increasing empirical evidence suggests that not just the immediate reaction to a stressor, but processes related to stress recovery play a crucial role in the development of diseases. With respect to more applied questions, it is important to study in more detail the factors that promote quick recovery. Here, a combination of laboratory studies and studies in real-world field setting seem to be particularly promising.

Further Reading

- Burke, H. M., Davis, M. C., Otte, C., et al. (2005). Depression and cortisol responses to psychological stress: a meta-analysis. *Psychoneuroendocrinology* 30, 846–856.
- Dickerson, S. S. and Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin* 130, 355–391.
- Eden, D. (2001). Vacations and other respites: studying stress on and off the job. In: Cooper, C. L. & Robertson, I. T. (eds.) *International review of industrial and organizational psychology*, pp. 121–146. Chichester, UK: Wiley.
- Fritz, C. and Sonnentag, S. (in press). Recovery, well-being, and performance-related outcomes: the role of workload and vacation experiences. *Journal of Applied Psychology*.
- Ganster, D. C., Fox, M. L. and Dwyer, D. J. (2001). Explaining employees' health care costs: a prospective examination of stressful job demands, personal control, and physiological reactivity. *Journal of Applied Psychology* 86, 954–964.
- Glynn, L. M., Christenfeld, M. and Gerin, W. (2002). The role of rumination in recovery from reactivity: cardiovascular consequences of emotional states. *Psychosomatic Medicine* 64, 714–726.
- Gump, B. B. and Matthews, K. A. (1999). Do background stressors influence reactivity to and recovery from acute stressors? *Journal of Applied Social Psychology* 29, 469–494.
- Hocking Schuler, J. L. and O'Brian, W. H. (1997). Cardiovascular recovery from stress and hypertension risk factors: a meta-analytic review. *Psychophysiology* 34, 649–659.
- Linden, W., Earle, T. L., Gerin, W., et al. (1997). Physiological stress reactivity and recovery: conceptual siblings separated at birth? *Journal of Psychosomatic Research* 42, 117–135.
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* 338, 171–179.
- Pieper, S. and Brosschot, J. F. (2005). Prolonged stress-related cardiovascular activation: is there any? *Annals of Behavioral Medicine* 30, 91–103.
- Sluiter, J. K., Frings-Dresen, M. H. W., Meijman, T. F., et al. (2000). Reactivity and recovery from different types of work measured by catecholamines and cortisol: a systematic literature overview. *Occupational and Environmental Medicine* 57, 298–315.
- Sonnentag, S. (2003). Recovery, work engagement, and proactive behavior: a new look at the interface between non-work and work. *Journal of Applied Psychology* 88, 518–528.
- Stewart, J. C., Janicki, D. L. and Kamarck, T. W. (2006). Cardiovascular reactivity to and recovery from psychological challenge as predictors of 3-year change in blood pressure. *Health Psychology* 25, 111–118.
- Westman, M. and Eden, D. (1997). Effects of a respite from work on burnout: vacation relief and fade-out. *Journal of Applied Psychology* 82, 516–527.

Reductive Stress

J P Kehrer

Washington State University, Pullman, WA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J P Kehrer, volume 3, pp 327–332, © 2000, Elsevier Inc.

Introduction

Reduction–Oxidation Chemistry

Definitions

Assessing Reductive Stress

Causes of Reductive Stress

Effects of Reductive Stress

Energy and Cell Damage

Pyridine Nucleotides and Cell Damage

Regulation of Gene Expression

Conclusion

Glossary

<i>Free radicals</i>	Molecules with one or more unpaired electrons that are not controlled by their immediate environment.
<i>Oxidative stress</i>	A disturbance in the prooxidant-antioxidant balance in favor of the former that enhances the potential for damage.
<i>Reactive oxygen species</i>	Higher-energy-state forms of oxygen or partially reduced forms of molecular oxygen that have increased reactivity.
<i>Redox balance</i>	The balance between oxidation and reduction states of cellular molecules. Because some molecules must be oxidized when another is reduced, the overall redox balance is inviolable. However, this balance can be altered between subcellular locations or cells.
<i>Reductive stress</i>	A disturbance in the prooxidant-antioxidant balance in favor of the latter that enhances the potential for damage; the result of a cell (or subcellular location) becoming relatively more reduced than normal.

Introduction

The ultimate goal of understanding the biochemical and molecular changes associated with stress is to find pharmacological or gene-based methods of protecting at risk tissues from injury or reversing any injury once it has occurred. The induction of stress at the cellular level can occur as the consequence of exposure to various xenobiotics, physical agents such as heat or radiation, or certain diseases. The specific

focus of the current chapter is reductive stress. This type of stress involves alterations in how and where reducing equivalents are generated and used by tissues as cofactors in normal metabolic processes, in various defense systems (particularly oxidant defenses), and in controlling various signal transduction and gene regulation pathways.

Because reducing equivalents are intimately involved in mitochondrial respiration and the generation of energy, reductive stress is most commonly generated through alterations in these pathways, including simple oxygen deprivation. This later process is often followed by reoxygenation, which can produce a sudden burst of oxidants. Thus, reductive stress and oxidative stress are inextricably linked.

Reduction–Oxidation Chemistry

The chemistry of oxidation–reduction reactions is straightforward and easily defined. Specifically, reduction is defined as the addition of one or more electrons to a molecule, whereas oxidation is defined as the removal of one or more electrons. Basic chemical principles dictate that a balance between oxidation and reduction is required in all chemical reactions. Thus, when one substance is oxidized, some other must be reduced.

At the cellular level, life can be defined as the ability to maintain gradients across membranes. Such gradients require an input of energy and may involve ions or molecules. Although at a global level the maintenance of an overall redox balance is inviolable, such a balance is a steady state and can be shifted among different classes of molecules that have different impacts on cell function or between cellular compartments. As a result, individual cells, or areas within a cell, can experience reductant (and obviously concomitant oxidant) imbalances. Significantly, these imbalances do not necessarily deviate from neutrality because, under normal conditions, the presence of substrates used by cells to generate energy means that the intracellular environment of a cell is reduced relative to that found extracellularly.

Because of the reducing nature of the intracellular environment, imbalances toward oxidation tend to be both more common and more studied. A search of the Medline database through March 2006 turned up 38 902 papers using the keyword term oxidative stress but only 43 using the term reductive stress. Despite this huge imbalance in emphasis, the reducing side of the equation is important when it comes to stress and the molecular regulation mechanisms of gene expression.

Definitions

The term oxidative stress was initially defined, by Helmut Sies in 1985, in biological systems as “a disturbance in the prooxidant-antioxidant balance in favor of the former.” This definition was later extended to include the adverse functional consequences of such stress by adding “leading to potential damage.”

The first use of the term reductive stress was by Albrecht Wendel in 1987. Interestingly, the term has never been precisely defined, perhaps because of the obvious inverse relationship to oxidative stress or perhaps because reductive stress has not been widely considered experimentally or theoretically. In general, reductive stress has been related to conditions of respiratory inhibition and refers to situations in which the status of a cell (or subcellular compartment) becomes relatively more reduced. Such situations actually end up enhancing the formation of reactive oxygen species (ROS), making the links between reductive and oxidative stress even stronger.

There is a strong tendency to imply that stress and damage are synonymous. Overcoming this tendency is difficult because damage clearly is accompanied, and usually preceded, by some sort of stress. A definition of damage may help, but there is no consensus as to what constitutes damage in a biological system. The definition of oxidative (and by analogy, reductive) stress given here indicates that stress only enhances the potential for damage, not that damage has occurred. Because the tissue changes that are routinely measured to assess reductive stress have been also observed in situations that are not necessarily detrimental to tissue or cell functions, it is obvious that stress and damage are not synonymous.

Injury, damage, toxicity, and related terms are qualitative descriptors of changes in various indices that are measured to assess cell or tissue function. Just how significant any particular change is in terms of function is rarely assessed, and thus their meaning is a value judgment made by an individual investigator. It is also important to differentiate adaptive responses from acute responses to a stress. Whether or not any observed changes are directly deleterious to function is also important, as is the time frame being considered because the persistence of some initially nondamaging adaptation can become detrimental. One useful working definition of injury is any alteration that diminishes essential functions, whether or not cell death ensues. Of course, it must be remembered that even cell death is not always undesirable, as evidenced by apoptosis under normal conditions.

Assessing Reductive Stress

A measurable shift to a more reduced steady state of some biomolecular pool is a reasonable working definition of reductive stress. The biomolecule selected for such a determination will probably differ from that used to identify the presence of oxidative stress. For example, glutathione disulfide (GSSG) is often measured as an index of oxidative stress because it only forms as the result of oxidative processes. Measuring increases in reduced glutathione (GSH) as an index of reductive stress is not, however, feasible because of cellular controls on GSH content and the large basal levels of this thiol (in contrast to GSSG). This means that small changes would need to be detected on a large background. In general, assessing the relative NAD(P)(H) pools has proven to be the most widely used indicator of reductive stress. Other indicators include an increase in the lactate/pyruvate ratio or the β -hydroxybutyrate/acetoacetate ratio.

Causes of Reductive Stress

Oxygen Deprivation

The proximate cause of reductive stress during oxygen deprivation (e.g., hypoxia or ischemia) is the loss of the terminal electron acceptor (oxygen) and the resultant accumulation of reduced cofactors (such as NADH and NADPH) along with oxidized substrates. Because this process occurs in the mitochondria, it is this organelle that first exhibits measurable changes, mainly the accumulation of the pyridine nucleotide, NADH. In the whole cell, there is also a loss of high-energy phosphates (mainly ATP). Glycolysis can modify the loss of ATP (although not enough to maintain long-term tissue viability) but at the expense of generating a further quantity of reducing equivalents. Because different tissues, as well as tumor cells versus normal cells, have differing glycolytic capacities, the extent of reductive stress induced by oxygen deprivation varies. There is no evidence to suggest that the oxidized substrates or reduced cofactors are directly toxic, but it is possible that excess reducing equivalents in the form of NADH might drive reactions in directions not normally seen, thereby affecting cell function and ultimately viability. Mechanisms of such effects could involve the generation of free radicals or the disruption of redox-regulated signal transduction pathways (Figure 1).

The ability of a tissue to maintain a normal redox balance is believed to be energy-dependent. This concept is supported by studies showing oxidants are more toxic to energy-depleted hypoxic cells than to

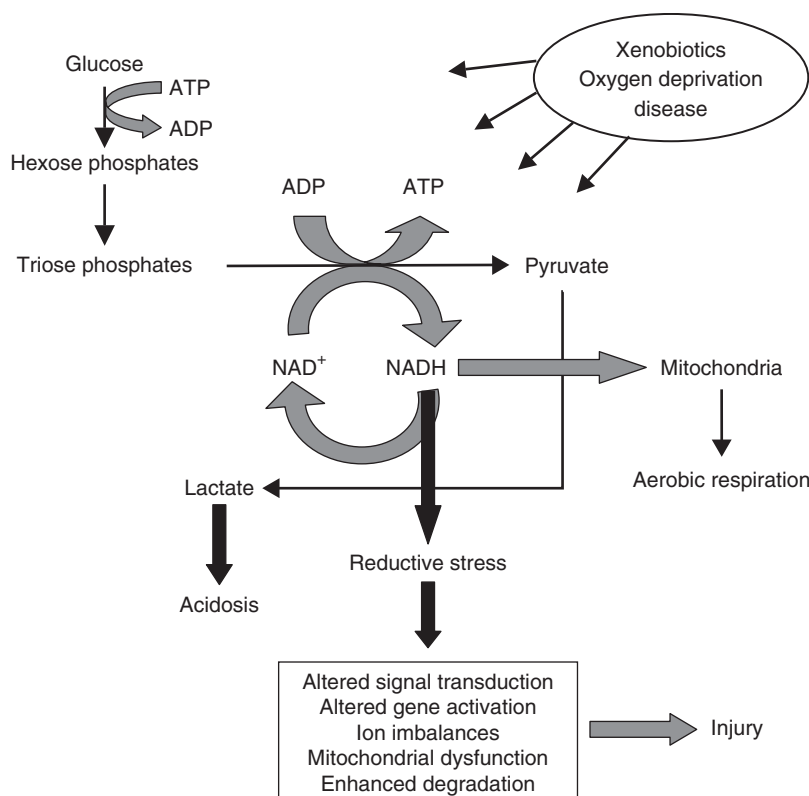


Figure 1 Reductive stress is a situation where the status of a cell (or subcellular compartment) becomes relatively more reduced. Causes of reductive stress can include hypoxia or ischemia and xenobiotics that affect pathways creating and/or using reducing equivalents. As the intracellular milieu becomes more reduced, there can be changes in various signaling pathways that required a defined redox state to be active. Although such effects may be detrimental, the presence of extra reducing equivalents may enhance a cell's ability to respond to an oxidative stress.

normal cells. However, although the production of energy and reducing equivalents are closely linked, they are separate biochemical events and alterations in the supply of reducing equivalents alone has the potential to induce stress.

Xenobiotics

Reductive stress in response to xenobiotics closely resembles that associated with oxygen deprivation. In fact, a widely studied model of reductive stress is termed chemical hypoxia and involves the blockade of mitochondrial respiration at several sites using cyanide and iodoacetate. Although differing from hypoxic hypoxia in some respects (e.g., chemical hypoxia appears to have more effects on NADPH than does oxygen deprivation), the overall biochemistry appears to be similar. Blocking specific mitochondrial electron transport sites with agents such as rotenone or antimycin A may also generate a reductive stress, although it is likely to have some differing effects based on the unique features of these compounds. In fact, such blockage in an aerobic environment may actually increase the generation of ROS.

Effects of Reductive Stress

The gross adverse consequence associated with prolonged oxygen deprivation is obvious – cell or tissue death. However, the biochemical and molecular changes that mediate this gross effect are much less clear. Changes consistent with reductive as well as oxidative stress have been found during hypoxia in whole animals, heart tissue, lung tissue, and chemically hypoxic hepatocytes. It also appears that disrupting cellular energy supplies, and thus reducing equivalent levels, has effects on other redox systems within the cell. For example, mitochondrial thiols are more susceptible to oxidation than are those in the rest of the cell.

Oxidative changes, or an increased susceptibility to oxidation, are consistent findings in various models of reductive stress. The oxidative changes observed in systems in which oxygen metabolism is inhibited could arise from the increased production of ROS or by modifying the cells' ability to respond to other stresses, leading to alterations in redox balance systems. The rate of mitochondrial ROS generation increases with the extent of reduction of mitochondrial

components. Although the low K_m of cytochrome oxidase for oxygen should ensure the removal of virtually all of this molecule during oxygen deprivation, it is possible that the increasingly reduced respiratory chain reacts with residual oxygen to generate ROS. It is also possible that the observed oxidative changes are a consequence of ROS generated on the resumption of respiration. Endogenous antioxidant systems may be impaired, but may continue to function during ischemia. Changes in the content of pyridine nucleotides alone appears not to lead to irreversible injury, but an altered ability to supply reducing equivalents in response to various demands may be critical.

In general, researchers have concluded that oxygen deprivation injury results not only from ATP depletion but also from reductive stress and the consequent generation of ROS by mitochondria. The site of this generation appears to be mainly at complex III, although other cellular sources of ROS appear to exist. It should be noted that not all studies support an increased mitochondrial production of ROS. For example, ROS production may not be increased following hypoxia or ischemia in the heart, brain, or isolated cardiac mitochondria.

Although clearly abnormal and likely to be detrimental in the long term, it is possible that the increased state of reduction that occurs subsequent to oxygen deprivation is actually protective against an acute oxidative stress. Similarly, under conditions of reductive stress, oxidizing agents could, in part, be beneficial because they would provide a sink for excess reducing equivalents.

The detrimental effects of reductive stress may partially be mediated by secondary effects that enhance the generation of ROS. Specifically, reductive stress can result in the intracellular release of iron which, in turn, enhances the production of ROS. Although the mechanisms are not fully understood, excessive NADH generated during a reductive stress can release ferrous iron from ferritin and can also reduce chelated ferric iron. The presence of increased amounts of free iron can, in turn, enhance the transfer of electrons to inappropriate sites, particularly in the presence of oxygen, leading to damage.

Energy and Cell Damage

Damaging events that affect cellular energy levels include mitochondrial dysfunction and the inhibition of glycolysis. Any such event will clearly affect how cells handle their supply of reducing equivalents, which is critical for cellular energy production. As cells adjust to compensate for a damaging event, reducing equivalents are shunted to functions other

than producing energy. Such shifts compound the loss of energy and must be considered potentially damaging events. Cell injury processes targeted to energy production may, therefore, represent a unique challenge for cells to increase their supply of reducing equivalents while maintaining critical redox balance to avoid further disruption of function.

The role of energy in cell injury has been studied in a variety of systems. The importance of secondary messengers in mediating such injury is well recognized. Of significance is calcium, whose intracellular levels appear to be altered in response to redox reactions as well as in the presence of energy deficiency. Changes in the levels of this cation are important for various signal transduction pathways. Thus, the regulation of signal transduction pathways by reductants is not necessarily direct, but may be mediated through secondary messengers such as calcium signaling and protein phosphorylation. Cellular pH is also important because an acidotic environment is protective against hypoxic or ischemic injury.

Pyridine Nucleotides and Cell Damage

The maintenance of reduced intracellular pyridine nucleotides appears to be a key factor in preserving cell viability because chemicals such as menadione, paraquat, and hydrogen peroxide deplete them before the loss of ATP and cytotoxicity. The key pathways maintained by NAD(P)(H) are not known, although modification of NAD(P) reduction status may lead to changes in the redox state of mitochondrial thiol groups. This, in turn, may affect cellular free calcium levels, the mitochondrial permeability transition, and trigger apoptosis.

Regulation of Gene Expression

Genes

The activation of gene expression in response to stress appears to be a universally conserved response in cells and substantial data suggest that oxidants can mediate gene activation. Whether a similar group of reductive stress-responsive genes exists is not known. It has been proposed that *grp* (glucose-regulated protein) genes are more responsive to reductive stress, whereas *hsp* (heat shock protein) genes respond better to oxidative stress.

The role of thiols in the control of cell growth and death is well accepted, although the specific pathways involved are not yet clear. Glutathione, thioredoxin reductase/thioredoxin, protein disulfide isomerase, glutaredoxin, and redox factor-1 have all been shown to affect protein thiols, thereby potentially

regulating protein folding, assembly, activity, and binding as transcription factors to specific DNA sequences. Although most work has focused on changes in terms of oxidants, it is again obvious that the reducing side of the equation could also be important. This is particularly true with thioredoxin, whose roles in regulating cell signaling are becoming increasingly appreciated.

Transcription Factors

At least 64 redox-regulated genes and transcription factors have been identified. More so than most other transcription factors, there is an extensive base of information regarding the activation of NF- κ B. This activation requires phosphorylation of the inhibitory I κ B subunit, which results in its dissociation from the inactive complex and its degradation by proteasome. A role for oxidants in NF- κ B activation is based on the observations that oxidizing conditions activate NF- κ B in several cell types, antioxidants can block this activation, and oxidant production is enhanced by various NF- κ B inducers such as tumor necrosis factor α . However, at the nuclear level, NF- κ B must be reduced in order to bind to DNA. Thus, the overall redox status of specific subcellular sites is crucial for determining the activation state of NF- κ B. Compartmentation may also be critical for the redox responses of NF- κ B-related factor 2 (Nrf-2). It has been suggested that Nrf-2 is retained in the cytosol by Keap-1 but that, when critical thiols on Keap-1 are oxidized, Nrf-2 is released and translocated to the nucleus, where Nrf-2 participates in the transcriptional activation of numerous genes. For Nrf-2 to be fully functional in transcriptional activation, a key cysteine residue must be in the thiol (reduced) form. Although phosphorylation may also be involved, this model further highlights the importance of reductive reactions in controlling signal transduction.

Activator protein 1 (AP-1) is a ubiquitous collection of protein complexes known to regulate transcription in response to environmental stimuli. It is composed of various gene products from the *fos* and *jun* proto-oncogene families. The products of these genes form homodimeric (Jun-Jun) and heterodimeric (Fos-Jun) complexes that bind to DNA, suggesting a crucial role of AP-1 in the control of cell proliferation. In addition to being stimulated by a wide range of xenobiotics or factors that promote cell proliferation, the hypothesis that cellular redox state plays an important role in the activation of AP-1 is well supported. AP-1 appears to be part of a general mechanism of regulation of gene expression following an alteration in redox balance. However, the pathways that control this expression are not clear. Antioxidants strongly activate AP-1, yet growth factor-induced AP-1 activation appears to be

ROS-dependent. Thus, although AP-1 and NF- κ B seem to respond oppositely to antioxidants, reductants and oxidants may have a role in the activation process of both factors.

GSH is clearly important for many of the reduction reactions required for normal cell functioning. In addition, recent data have demonstrated that the thioredoxin system may be the more proximal regulatory factor for several transcription factors including NF- κ B. The overall reduction status of a cell and, in particular, the relative state of both the cytosol and nucleus in terms of thiol redox balance affect the activity of transcription factors. Reductive stress is, therefore, expected to have an effect on their activities.

Conclusion

Reductive and oxidative stress are closely intertwined processes that can affect cells at various levels. Although the oxidative side of the equation has been the most studied, the reductive side has the significant potential to regulate signal transduction pathways and gene expression, as well as more obvious cellular stress responses.

See Also the Following Articles

Apoptosis; Oxidative Stress.

Further Reading

- Allen, R. G. and Tresini, M. (2000). Oxidative stress and gene regulation. *Free Radical Biology and Medicine* **28**, 463–499.
- Arrigo, A-P. (1998). Small stress proteins: chaperones that act as regulators of intracellular redox state and programmed cell death. *Biological Chemistry* **379**, 19–26.
- Cross, J. V. and Templeton, D. J. (2004). Thiol oxidation of cell signaling proteins: controlling apoptotic equilibrium. *Journal of Cellular Biochemistry* **93**, 104–111.
- Dalton, T. P., Shertzer, H. G. and Puga, A. (1999). Regulation of gene expression by reactive oxygen. *Annual Review of Pharmacology and Toxicology* **39**, 67–101.
- Dawson, T. L., Gores, G. J., Nieminen, A-L., et al. (1993). Mitochondria as a source of reactive oxygen species during reductive stress in rat hepatocytes. *American Journal of Physiology* **264**, C961–C967.
- Finkel, T. (2003). Oxidant signals and oxidative stress. *Current Opinion in Cell Biology* **15**, 247–254.
- Forman, H. J., Torres, M. and Fukuto, J. (2002). Redox signaling. *Molecular and Cellular Biochemistry* **234–235**, 49–62.
- Haddad, J. J. (2002). Antioxidant and prooxidant mechanisms in the regulation of redox(y)-sensitive transcription factors. *Cell Signaling* **14**, 879–897.
- Halleck, M. M., Holbrook, N. J., Skinner, J., et al. (1997). The molecular response to reductive stress in LLC-PK1 renal epithelial cells: coordinate transcriptional

- regulation of gadd153 and grp78 genes by thiols. *Cell Stress & Chaperones* 2, 31–40.
- Jones, D. P. (1997). New concepts of the molecular pathogenesis arising from hypoxia. In: King, T. E., Mason, H. S. & Morrison, M. (eds.) *Oxidases and related redox systems*, pp. 127–144. New York: A. R. Liss.
- Kehrer, J. P. and Lund, L. G. (1994). Cellular reducing equivalents and oxidative stress. *Free Radical Biology and Medicine* 17, 65–75.
- Sies, H. (ed.) (1985). *Oxidative Stress*. p. 1. New York: Academic Press.
- Turrens, J. F. (2003). Mitochondrial formation of reactive oxygen species. *Journal of Physiology* 522(2), 335–344.
- Watson, W. H., Yang, X., Choi, Y. E., et al. (2004). Thio-redoxin and its role in toxicology. *Toxicological Sciences* 78, 3–14.

Reenactment Techniques

N Wong

Uniformed Services University of the Health Sciences,
Bethesda, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by N Wong, volume 3, pp 333–334, © 2000, Elsevier Inc.

Rationale for Reenactment

Common Techniques

Glossary

<i>Abreaction</i>	A process by which repressed material, particularly a painful experience or conflict, is brought back into consciousness. In the process the person not only recalls, but also relives the situation and experiences the appropriate emotional response.
<i>Catharsis</i>	The verbalization of ideas, thoughts, and expressed material that is accompanied by an emotional response and that produces a state of relief to the individual.
<i>Psychodrama</i>	A form of group therapy devised by Moreno in the 1920s drawing upon theatrical techniques to explore a patient's personality makeup, interpersonal relationships, conflicts, and emotional patterns.

Rationale for Reenactment

When exposed acutely or chronically to stressful or traumatic situations, some individuals may develop physical and emotional symptoms, which may range from bothersome to totally incapacitating. These symptoms may happen occasionally, or they can be present throughout the day and night. The aim of

having patients reenact their trauma is to prevent, diminish, or totally alleviate any potential or actual symptoms and emotional suffering. In addition to the use of psychodynamic therapy, several categories of reenactment techniques exist and are applied with respect to the clinical indication and ego strength of the individual.

Common Techniques

Debriefing

One of the most commonly utilized techniques to help individuals reenact or recall an acute or intermediate traumatic situation, debriefing can be applied in an individual or group setting. Participants are encouraged to retell their personal experiences and thereby gain an emotional release, or catharsis, that hopefully prevents or reduces the risk for more serious stress reactions. In the process, there is a sharing and education of those experiencing a common trauma that promote ventilation of feelings through the expression of emotions and the response of others.

Psychodrama

Psychodrama is a form of group therapy, but some of its techniques, such as warming up, scene setting, and therapeutic soliloquy, can be used in the individual setting as well for patients suffering from stress. Other psychodrama techniques are role playing, role reversal, therapeutic soliloquy, and mirror. The therapist or leader is an active participant. In role playing, the patient is asked to perform an impromptu dramatization of the traumatic situation. In role reversal, another individual in the group attempts to reenact what the subject may be thinking, or feeling, in the presence of the subject. In therapeutic soliloquy, the individual is encouraged to verbalize side dialogues and actions during a psychodrama scene. The mirror

technique consists of having another person reenact what the subject has done, presenting identical behaviors and expressing similar feelings to show how others experience the patient while the subject sits behind a mirror.

Hypnosis

This modality is especially useful in the reenactment of stressful or traumatic situations causing dissociative amnesia and dissociative fugue with memory loss. Hypnosis is best utilized to explore the events that precipitated the fugue or amnesia. When employing this technique, clinicians should suggest to the hypnotized patients that they may choose to forget some or even all of what was remembered or reenacted in the trance state to prevent them from being overwhelmed by the emergence of unconscious material that the patients are not yet ready to handle. The recaptured or selectively shared data from patients are then dealt with in the fully conscious state to help them cope with the underlying conflicts. It is generally advisable to employ other conscious methods before resorting to the use of hypnosis or chemical means to bring about a reenactment of a traumatic situation. While hypnosis is usually safe to use with most healthy individuals, the therapist or treater should exercise the same degree of caution in probing unconscious material in psychotherapy.

Medication-Facilitated Interviews

Clinicians may use intravenous drug-induced interviews to help patients recall or reenact traumatic situations or conflicts when symptoms may be disabling, or when great resistance is encountered in examining or recalling memories. Commonly utilized drugs include amobarbital, pentobarbital, intravenous lorazepam, and hexobarbital. A disadvantage of these interviews lies in the chemically induced amnesia that often occurs. Also, because of the risk of adverse reactions to the drugs, medication-facilitated interviews should only be performed by trained

individuals with immediate availability of life support systems and the presence of an anesthetist.

Use of Media

Young children may not have the verbal capacity to fully express their anxieties and emotional suffering when exposed to traumatic or stressful situations. Therapy may be conducted via play, painting and drawing, participation in story telling, creating illustrated storybooks, and role playing through which the trauma may be reenacted, explored and clarified, and resolved. For older children, adolescents, and adults, the creation of a documentary video or a play reenacting the trauma can prove extremely therapeutic, providing opportunities for ventilation, abreaction, clarification, reality testing, and mastery.

See Also the Following Articles

Hypnosis; Trauma and Memory.

Further Reading

- Freehill, K. M. (1992). Critical incident stress debriefing in health care. *Critical Care Clinic* 8, 491–500.
- Hanney, L. and Kozlowska, K. (2002). Healing traumatized children: creating illustrated storybooks in family therapy. *Family Process* 41, 37–65.
- Krietemeyer, B. C. and Heiney, S. P. (1992). Storytelling as a therapeutic technique in a group of school-aged oncology patients. *Child Health Care* 21, 14–20.
- Moreno, J. L. (1947). *Psychodrama*. New York: Beacon Press.
- Quittner, A. L., Tolbert, V. E., Regoli, M. J., et al. (1996). Development of the role-play inventory of situations and coping strategies for parents of children with cystic fibrosis. *Journal of Pediatric Psychology* 21, 209–235.
- Rhue, J., Lynn, S. and Kirsch, I. (1988). *Handbook of clinical hypnosis*. Washington, D.C.: American Psychological Association.
- Tosone, C., Gelman, C. R. and McVeigh, L. (2005). Through their own eyes: a media-base group approach to adolescent trauma. *International Journal of Group Psychotherapy* 55, 415–432.

Refugees, Stress in

J D Kinzie

Oregon Health and Sciences University, Portland,
OR, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
J D Kinzie, volume 3, pp 335–337, © 2000, Elsevier Inc.

Scope of the Problem

Stress in Refugees

Stress-Related Disorders in Refugees

Conclusion

Glossary

<i>Depression</i>	A specific psychiatric disorder consisting of lowered mood and symptoms, such as decreased interest in most activity, loss of appetite and weight, poor sleep, fatigue, feelings of worthlessness, poor concentration, and recurrent thoughts of death.
<i>Posttraumatic stress disorder</i>	A specific psychiatric disorder following psychological trauma with symptoms of re-experiencing, such as nightmares and flashback; avoidance behavior, such as attempts to avoid thoughts and memories of the events; and hyperarousal symptoms, such as irritability, startle response, and difficulty sleeping.
<i>Refugee</i>	A person who flees to another country as a result of, or to avoid, political prosecution, war, starvation, or other disasters.
<i>Schizophrenia</i>	A severe mental disorder characterized by loss of contact with reality, hallucination, delusions, and often impairment of social and vocational functioning.

Technically, a refugee is a person who flees to another country usually to escape political prosecution, war, starvation, and other disaster. This definition does not include those displaced internally in their own country. Refugees are distinguished from migrants who voluntarily migrate to another country for economic education or personal reasons. A special group of refugees are the asylum seekers who flee to a foreign country without official recognition or on a temporary visa status and request refugee status usually because of perceived dangers if they return to their homeland. While their claims are reviewed, asylum seekers live in an ambiguous status with the threat of forced repatriation.

Scope of the Problem

At the end of 2004, the United Nations High Commission of Refugees reported 19.2 million persons of concern (including refugees, asylum seekers, stateless, and other). This was a 13% increase from 2003. Those with formal refugee status comprise 48% of the 19.2 million but many stateless persons have not been systemically identified and therefore the total persons of concern is probably much greater than 19.2 million. The areas of southern Africa and CASWANAME (Central Asia, South West Asia, North Africa and Middle East) each hosted 30% of the global refugee population. The Americas only hosted 7%. Afghanistan has been the largest country of origin of refugees despite 3.5 million Afghans in the past 3 years voluntarily repatriated home. The continual civil wars and ethnic strife in Africa and the Middle East indicates the worldwide problems of refugees and asylum seekers will continue to exist.

Stress in Refugees

Types of Stress

The stresses on refugees are diverse, multiple, and frequently catastrophic in nature. The phases of stress are often divided into preflight; flight and separation; asylum; and resettlement. Preflight stress includes poor economic conditions, food shortage, famine, political persecution, social upheaval, and violence. The results can include hunger, malnutrition, physical disability, and physical and psychological trauma. The flight and separation phase can include separation from family and society, sexual assault, violence, and collapse of social support. The result can be multiple physical and social traumas and the sense of being isolated and alone. The asylum phase often includes temporary resettlement in a refugee camp and can involve the threats of being forced to return home, inhospitable living conditions, unemployment, food shortages, and lack of health services. The results can include malnutrition, illness, and despair. Resettlement stress, often in a third country, can include unemployment, social isolation, acculturation problems, limited social ties, prejudice, language barriers, intergenerational conflicts, and minority status. The results can be severe economic hardships, social isolation, delinquency among adolescents, generational, and conflicts.

The amount of direct physical and psychological trauma varies among the refugee groups and the

amount of violence prior to and during their escape. Some, such as the first group of Vietnamese refugees, left because of fear of harm rather than direct harm or threats. Others, for example Cambodian refugees, suffered severe stress over 4 years, which involved a total collapse of their society with starvation, torture, witnessing executions, wanton deaths, forced labor for 4 years, and even violence during the escape and refugee camp experience.

There have been mass murders in the name of ethnic cleansing in Bosnia, and murder, starvation, and random lawlessness with gang violence in Somalia. Refugees fleeing severe ethnic wars, and sadistic acts of cruelty in Africa are often displayed on television. Although technically not refugees (they still stay in their own country), survivors of powerful tsunamis, hurricanes, and earthquakes experience many of the same stress as formal refugees. In general victims of acts of God have less psychiatric disturbance than those victims of man-made disasters, i.e., war.

Vulnerable Groups

At different times, different groups can be more vulnerable to stress or persecution. During armed conflict, young men are often singled out as potential enemies and severely persecuted or executed. Women are particularly vulnerable to sexual abuse (rape) when fleeing from their own country or even in settlement camps. At the country of resettlement, women and the elderly are often isolated at home and experience more difficulty in adjusting. Children acculturate faster as a result of school experience, which can greatly increase the intergenerational conflicts in the family over time.

Protective Factors

Since studies during the bombings of London in World War II, it has been shown that children that stay in the care of their mother, even in dangerous bombing situations, do better than those sent to the safe countryside away from families. An intact family or extended family that can maintain daily routines can provide some protection against the most serious psychological effects. Many studies have shown that having effective social linkages and social support provide a sense of belonging and improved identity, which protects against the stress of social adjustment.

Religious affiliations have been found to be protective in some refugee status, and strong ideological and political commitments have even reduced the effects of torture. The more massive and prolonged the psychological trauma, the less effective are any

protective mechanisms and the more consistently symptoms are found.

Asylum seekers arrive in a country on a temporary visa (or even illegally) and request refugee status. Usually, during a prolonged waiting period, their claims are reviewed, basic services are denied, and there is even the threat of forced repatriation. This group, which has been studied in several countries, has high postmigration stress and often severe pre-migration trauma. With an unclear legal status, this group is very vulnerable and may develop an increased rate of psychiatric symptoms.

Stress-Related Disorders in Refugees

Physical Disorders

The first studies of refugees occurred with Nazi concentration camp survivors who fled to another country after World War II. Their multiple symptoms were thought to be related to brain disease due to head trauma or malnutrition. Later, the symptoms were characterized as concentration camp syndrome (which is similar to posttraumatic stress disorder (PTSD)) and were found primarily to be due to the massive psychological trauma of the holocaust. Nevertheless, many refugees do suffer from head trauma, which may explain some of the complaints of memory and concentration problems.

Although long-term health effects are inconsistently found, many refugees report marked somatic distress, and 15–20% report health impairment. Among various groups of refugees (Cambodians, Vietnamese, Russians, Bosnian, and Somalis) hypertension was found in 40–50% and diabetes (type II) was found in 13–17%. Although dietary changes, refugee status, and effects of aging play a role, severe trauma contributes to high prevalence of these two diseases.

Behavioral Disorders

Refugee populations may be particularly sensitive to alcoholism, drug abuse, delinquency, or antisocial behavior. All the losses, including status and employment, may cause adult males to use alcohol and drugs as a temporary escape in both the camps and in the resettlement country. Asian refugees historically have tended to have lower rates of alcoholism, whereas Central American males tend to have high rates of substance abuse as refugees. Young people in resettlement countries have been found to engage increasingly in antisocial behavior and drug abuse. Refugee children can experience a severe disruption of family relationships and much intergenerational conflict in

their new country. A lack of effective role models, a sense of social isolation, and poor academic performance may turn adolescents to use of drugs and drug dealing in an attempt to acculturate to the host society. Such behavior by adult and juvenile males can increase family stress greatly and lead to family breakup and domestic violence.

Indeed in refugee children from Central America domestic violence (often secondary to alcoholism) is more likely to be the major trauma than war-related violence.

Psychiatric Disorders

The most persistent effects on refugees have been the consistent finding of a high rate of psychiatric disorders. Earlier reports from European refugees after World War II indicate an increased rate of schizophrenia. Later studies have shown that those refugees from the Nazi concentration camps were found to have suffered from a specific disorder characterized by fatigue, irritability, restlessness, anxiety, depression, and frequent nightmares. This concentration camp syndrome was found to be related to massive psychological trauma. The syndrome was largely forgotten; but with a large influx of refugees in the United States in the late 1970s and 1980s, it was found that many refugees suffered from PTSD. PTSD is found to be similar, if not identical, to the syndrome called concentration camp syndrome. By definition, PTSD is the result of a traumatic event in which an individual experiences or confronts death or serious injury and responds with fear, helplessness, or horror. The symptoms can cause re-experiencing such as nightmares, avoidance behavior, or a numbing experience such as efforts to avoid thoughts of the events and hyperarousal symptoms such as difficulty falling asleep, irritability, poor concentration, and startle response. In a clinic treating refugees from Southeast Asia, a 54% rate of PTSD was found among the Vietnamese refugees. The traumatic experience was related both to the Vietnam War and to the escape processes. The 92% rate of PTSD among Cambodian patients all related to the Pol Pot concentration camps. Most clinics have found PTSD to be similarly high, although the rates for Cambodians in the community (non-patients) tend to be lower. A great deal of research and clinical experience indicates that traumatized refugees have high rates of PTSD. Information exists on a number of refugees: Chilean, Salvadorian, Ethiopian Jews, Afghan, Somalis and Bosnians, as well as Tamil and Burmese asylum seekers. A recent community survey of Cambodian refugees two decades after resettlement in the United States found that 62% still suffered from PTSD. This is almost twice as high as among Cambodians living in Cambodia indicating

that refugee status and/or immigration stress contribute to ongoing symptoms. The highly televised bombing of Twin Towers on 11 September 2001 was very disturbing to refugees and many (particularly Somalis and Bosnians) had markedly increased symptoms after repeated viewing. The symptoms are often quite persistent and can recur after periods of quiescence.

Depression characterized by lowered mood, low energy, low interest, poor sleep, and poor appetite is also very high among refugees. PTSD has generally been found to be related to specific traumatic events, whereas depression relates more to postmigration stress. Depression tends to improve over time more than PTSD symptoms.

Schizophrenia, a severe mental disorder characterized by delusions and hallucinations, which is found to be high in refugees from World War II, continues to be increased in current refugees. However, the number involved is not nearly as large as those having PTSD and/or depression.

Conclusion

It is now clear that refugees suffer from much stress, sometimes of catastrophic proportions. This can result in symptoms of PTSD and depression, which are similar regardless of the culture of the refugees or the environment where the refugees resettle. As the number of refugees worldwide increases, the number of refugees impaired and suffering will continue to increase

See Also the Following Articles

Schizophrenia; Social Support; Survivor Guilt; War Stress in the Former Yugoslavia; War-Related Posttraumatic Stress Disorder, Treatment of.

Further Reading

- Boehnlein, J. K. and Kinzie, J. D. (1995). Refugee trauma. *Transcultural Psychiatric Research Review* 32, 223–252.
- Desjarlais, R., Eisenburg, L., Good, B., et al. (eds.) (1995). *World mental health: problems and priorities in low-income countries*. New York: Oxford University Press.
- Jaranson, J. J. and Popkin, M. K. (1998). *Caring for victims of torture*. Washington, DC: American Psychiatric Press.
- Marsella, A., Bornemann, T. and Ekblad, S. (eds.) (1994). *Amidst peril and pain*. Washington, DC: American Psychiatric Association.
- Wilson, J. P. and Drozdek, B. (eds.) (2004). *Broken spirits: the treatment of traumatized asylum seekers, refugees, war and torture victims*. New York: Brunner-Routledge Press.

Regional Blood Flow, Stress Effects

W W Blessing

Centre for Neuroscience, Flinders Medical Centre,
Adelaide, SA, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
W W Blessing, volume 3, pp 338–340, © 2000, Elsevier Inc.

Introduction

Exposure to Cold Causes a Sudden Decrease in Cutaneous
Blood Flow

When the Individual is Frightened, Cutaneous Blood
Flow Falls

Psychotropic Drugs Alter Cutaneous Blood Flow

Cutaneous Blood Flow Decreases During the Febrile
Response to an Infectious Agent

Neuronal Pathways in Brain and Spinal Cord

Co-ordinating Sympathetic Outflow to the
Cutaneous Vascular Bed

Conclusion

Glossary

<i>Clozapine</i>	A drug used to treat schizophrenia, called an atypical antipsychotic agent because it has a reduced tendency to cause muscle rigidity in comparison with classical drugs such as haloperidol.
<i>Raphe/parapyramidal region</i>	Neurons in the midline and ventromedial region of the medulla oblongata with direct axonal projections to the spinal cord.
<i>Thermoregulatory cutaneous vascular beds</i>	Regions of the skin (e.g., fingers in humans and tail in rats) where heat is transferred from the blood to the environment.

Introduction

During the course of evolution, natural selection has shaped behavior as well as psychological and physiological processes so that they operate in an integrated fashion, enabling the individual to successfully undertake the daily life activities necessary for survival and reproduction. During everyday life, the regional distribution of the blood flow is patterned by a combination of central command (brain-initiated) and reflex (organ-initiated) control systems, so that flow to a particular bodily organ is appropriate for the particular activity in which the individual is engaged at a particular time.

Broadly defined, an acute stress may be envisaged as the occurrence of an intrusive event (physical, psychological, or physiological) that could threaten survival of the individual or, in a broader sense, survival of the species. Such events initiate adaptive responses. W. B. Cannon appreciated that the behavioral response to stress (the flight-or-fight response) is accompanied by an acute redistribution of the cardiac output, so that blood is directed to the bodily organs involved in coping with the stressful event. This was an extremely important contribution to physiological thinking, a reminder of the concepts so graphically proposed by Darwin in his book *Expression of the emotions in man and animals*.

Cannon, however, believed the patterns of activity in sympathetic nerves supplying the blood vessels to be diffusely directed rather than regionally focused. The central command process, so Cannon thought, results in a unitary activation of all the sympathetic nerves innervating the adrenal gland and the blood vessels. At the time, this view was based on a general theoretical stance, rather than on experimentally demonstrated changes. Cannon was thinking of a highly mobile animal, actively engaged in fight or flight. In Cannon's scheme, the principal emphasis was on increases in blood flow to vascular beds in skeletal muscle. Measurement of regional blood flow or sympathetic nerve activity in a conscious behaving animal is still technically difficult, so that there is still much to be learnt. In a cat confronted with a second aggressive cat, blood flow to skeletal muscle beds appears to depend on how much the animal actually moves and on whether particular muscle groups are involved in the movement. There is no general increase in blood flow to all skeletal muscles.

Cannon was correct in his view that the response to stressful events includes diversion of blood away from the skin. The short-term (over minutes) nutrient requirements of the skin are not crucial, so that the supplying blood can be diverted elsewhere without deleterious effects to the skin itself. The skin is vulnerable to injury precisely because, as the border between the individual and the external environment, it bears the initial brunt of any actual physical attack upon the individual. If blood is diverted from the skin, the danger of major injury-induced hemorrhage is reduced. This article, drawing on experiments conducted in experimental animals (rabbits and rats), outlines some of the environmental stresses which, via the central command process, activate

cutaneous sympathetic neural outflow so that blood flow to the skin is suddenly reduced.

Exposure to Cold Causes a Sudden Decrease in Cutaneous Blood Flow

The volume of blood flowing through thermoregulatory (heat-exchanging) cutaneous blood vessels (ear pinna in rabbits, tail in rats, fingers and toes in humans) is rapidly reduced when the individual is suddenly exposed to cold (a physical stress), so that heat loss from the body is reduced. The brain, via the various sensory inputs, detects the cold environment and particular brain centers activate the sympathetic nerves innervating the blood vessels, thereby constricting the vessels and reducing the amount of blood that enters the heat-exchanging areas of the body. Since the final control neurons within the central nervous system (the sympathetic preganglionic neurons) are located in the thoracic and upper lumbar segments of the spinal cord, there must be neural connections linking forebrain, midbrain and hindbrain with these spinal neurons. The descending pathways are briefly summarized later in this article.

When the Individual is Frightened, Cutaneous Blood Flow Falls

When the individual detects potentially dangerous events in the external environment (physical or

psychological) special alerting rhythms occur in the hippocampal electroencephalograph followed, a few seconds later by a sympathetically mediated fall in cutaneous blood flow (**Figure 1a**), with a return to baseline levels after a minute or so if the danger is transient. In conscious rabbits (especially bred for use in laboratory experiments and thus familiar with humans) apparently minor stimuli (e.g., the laboratory telephone ringing or an unexpected touch of the fur) often cause cutaneous blood flow to fall suddenly to near zero levels, without change in blood flow to the kidneys, with a small decrease in blood flow to the intestines and abdominal viscera, and with no change in blood flow to vascular beds in skeletal muscle (**Figure 1b**). With relatively minor alerting stimuli there is little or no change in cardiac output. There may be no obvious behavioral change. If the animal is more strongly alerted, either because the stimulus is more threatening or because the animal is already anxious, the cutaneous vasoconstriction is more prolonged. Heat loss via the cutaneous circulation is thereby interrupted for a longer period. Body temperature may increase, a situation referred to as stress-induced hyperthermia.

One way of introducing a more threatening stimulus is to utilize the conditioned fear response. Rats subjected to an electric shock exhibit behavioral change (freezing) and cardiovascular changes. If the electrical shock is delivered soon after the rat has been moved to an unfamiliar cage, similar behavioral

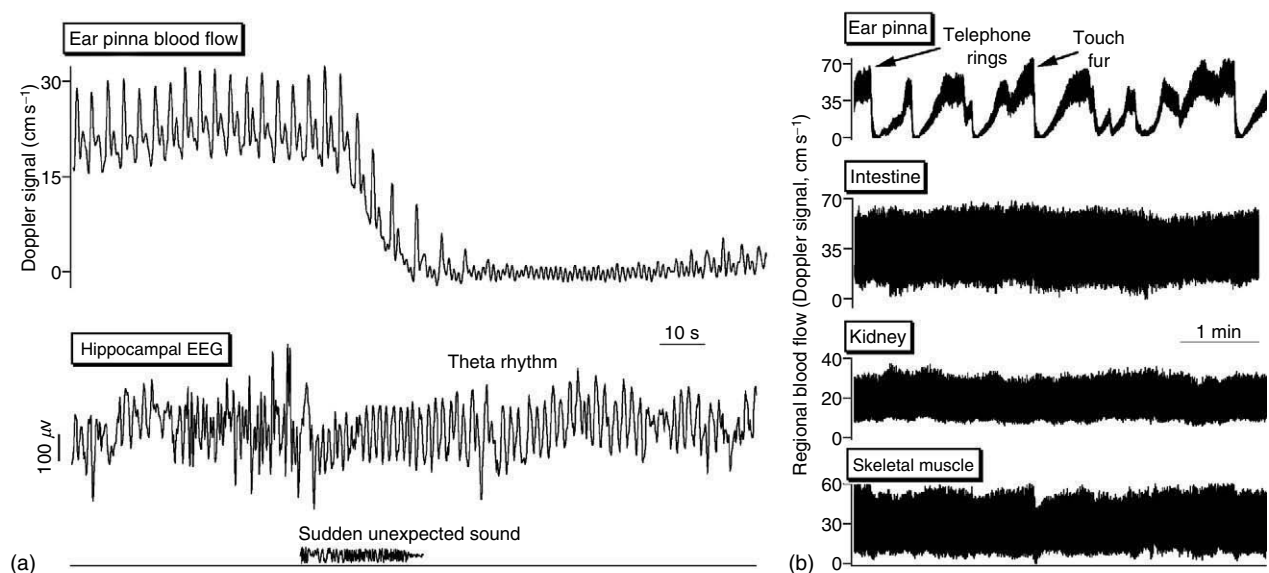


Figure 1 a, Ear pinna arterial blood flow (top panel) and hippocampal electroencephalograph (bottom panel) in a conscious rabbit, recorded before and after onset of a sudden sound. After the sound onset, the electroencephalograph develops a slow (6–7 Hz) regular rhythm indicating that the sound is significant for the rabbit. Approximately 1 s after the sound, ear pinna blood flow falls rapidly to near zero levels. b, Blood flow recorded simultaneously from arteries supplying the ear pinna, intestine, kidney, and skeletal muscle in the hind limb. Ear pinna blood flow falls to near zero levels when the animal detects potentially dangerous environmental stimuli (see Yu and Blessing, 1997: R208).

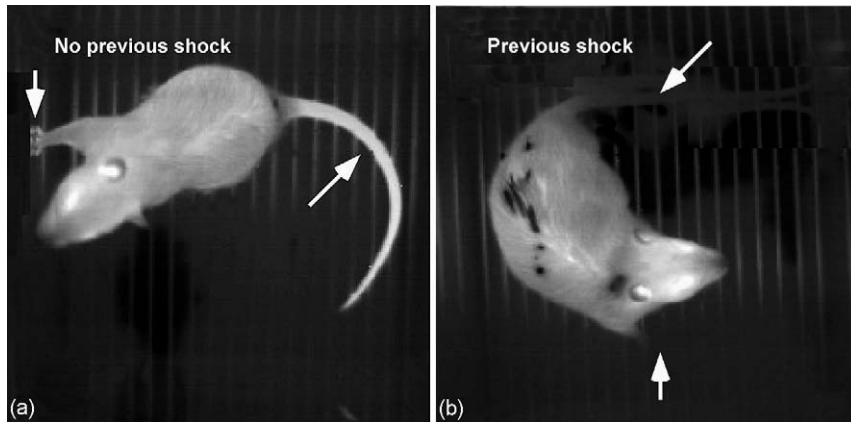


Figure 2 Infrared digital images of one fear-conditioned a, and one sham-conditioned b, rat obtained during re-exposure to the experimental box. Brighter regions indicate increased temperature (i.e., higher blood flow). Note (white arrows) the fall in tail and paw temperature in the fear-conditioned, but not in the sham-conditioned rat during re-exposure to the experimental box (see Vianna and Carrive, 2005: 2505).

and cardiovascular changes occur when the rat is subsequently reintroduced to the same cage, i.e., the rat displays a conditioned fear response. Cutaneous vasoconstriction is a prominent component of the cardiovascular response (Figure 2).

Psychotropic Drugs Alter Cutaneous Blood Flow

There is an intricate interrelationship between neurotransmitters controlling emotional/psychological functioning and neurotransmitters involved in CNS neural pathways regulating cutaneous blood flow. Serotonin (5-hydroxytryptamine, 5-HT) and dopamine are important in both systems. Lysergic acid diethylamide (LSD) and related agents cause hallucinations by an agonist action at subclasses of serotonin receptors (5-HT_{2A} receptors). Ecstasy (methylenedioxymetamphetamine, MDMA) causes release of serotonin in the brain, with subsequent activation of 5-HT_{2A} receptors. These drugs also markedly activate the thermoregulatory sympathetic outflow to the cutaneous vascular bed, causing vasoconstriction and reducing heat loss from the body. Clozapine, a very valuable antipsychotic drug used in the treatment of schizophrenia and schizoaffective disorders to reduce mental stress, reverses the effects of MDMA and LSD-like agents on emotional function. Clozapine also reverses MDMA-induced activation of the cutaneous sympathetic outflow. In drug-free animals, clozapine inhibits resting cutaneous sympathetic outflow, so that cutaneous blood flow increases (Figure 3a), as well as reducing the increase in cutaneous sympathetic discharge normally initiated when the individual detects a frightening stimulus (Figure 3b).

Cutaneous Blood Flow Decreases During the Febrile Response to an Infectious Agent

Infectious agents such as viruses or bacteria cause body temperature to increase (the febrile response). Underlying physiological mechanisms include activation of the cutaneous sympathetic outflow, with constriction of the cutaneous vessels so that body temperature increases partially because heat is no longer transferred from the body to the environment.

Neuronal Pathways in Brain and Spinal Cord Co-ordinating Sympathetic Outflow to the Cutaneous Vascular Bed

Stressful environmental stimuli are detected by the different senses (vision, hearing, touch, temperature, and pain sensitivity, etc.). These afferent stimuli are processed in forebrain centers, including the amygdala and the dorsomedial nucleus of the hypothalamus. Efferent neuronal signals reach the sympathetic preganglionic neurons in the spinal cord by way of a major integrative center in the lower brainstem, in the raphe/parapyramidal region of the rostral medulla oblongata. Neurons in the raphe have direct axonal projections to the sympathetic preganglionic neurons in the spinal cord (Figure 4).

Conclusion

As the individual prepares to cope with the different forms of environmental stress, blood flow to the cutaneous vascular bed is reduced via activation of the relevant sympathetic nerves, a response initiated in the brain and co-ordinated by spinally projecting

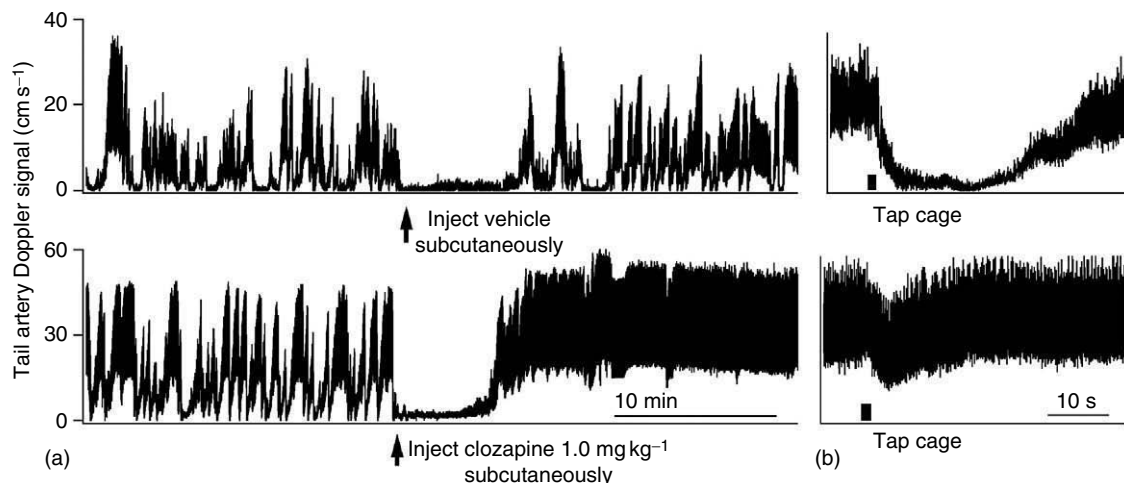


Figure 3 Blood flow to the tail of the conscious freely moving rat measured with a timeframe of minutes (a) and seconds (b). In (a) either vehicle control solution (top panel) or clozapine solution (bottom panel) was administered. Before the injection both traces show frequent sudden falls to near zero levels in response to perception of alarming events by the animal. Tail blood flow falls during the injection procedure in both panels (the rat is alarmed by the procedure). After clozapine, resting blood flow increases and the sudden falls in flow are no longer observed. In (b) the upper panel was obtained from a control, vehicle-treated rat and the lower panel was from a rat treated with clozapine 30 min before. A sudden tap to the animal's cage elicits a rapid fall in tail blood flow, to near zero levels, in the vehicle-treated animal but not in the clozapine-treated animal (see Blessing, 2005: 518).

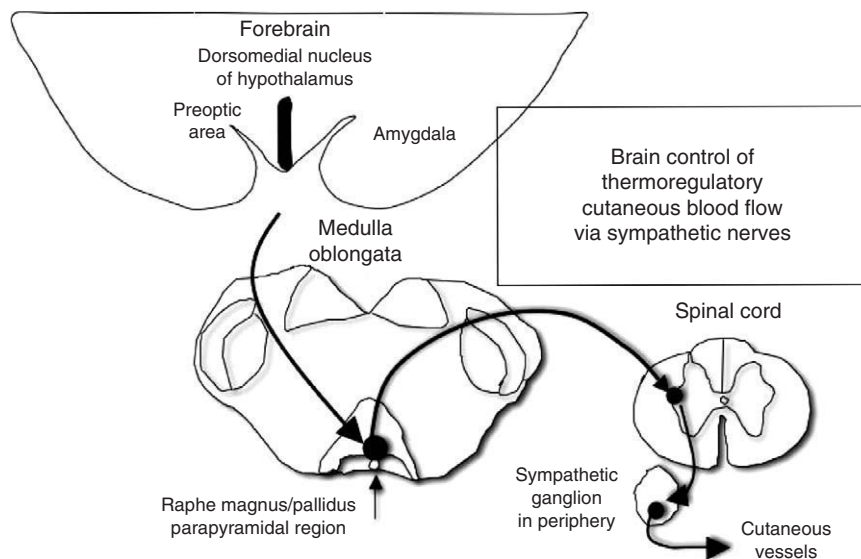


Figure 4 Summary diagram illustrating neural pathways linking brain and sympathetic preganglionic neurons in the spinal cord. The raphe/parapyramidal region of the rostral medulla oblongata is an important relay station in the descending brain pathway mediating stress-induced cutaneous vasoconstriction.

neurons present in the raphe/parapyramidal region of the medulla oblongata. Some drugs used to treat mental illness also interact with brain neurotransmitter systems (serotonin and dopamine) regulating the thermoregulatory cutaneous blood flow to reduce stress-induced sympathetic activation, thereby increasing resting cutaneous blood flow and preventing stress-induced falls in blood flow. Measurement of cutaneous blood flow provides a new biological index of stress.

Acknowledgments

Supported by the National Health and Medical Research Council (NH & MRC) of Australia.

Further Reading

- Blessing, W. W. (1997). *The lower brainstem and bodily homeostasis*. New York: Oxford University Press.
- Blessing, W. W. (2005). Clozapine increases cutaneous blood flow and reduces sympathetic cutaneous vasomotor

alerting responses (SCVARs) in rats: comparison with effects of haloperidol. *Psychopharmacology* 181, 518–521.

Cannon, W. B. (1929). *Bodily changes in pain, hunger, fear and rage*. New York: Harper and Row.

Mancia, G., Baccelli, G. and Zanchetti, A. (1972). Hemodynamic responses to different emotional stimuli in the cat: patterns and mechanisms. *American Journal of Physiology* 223, 925–933.

Vianna, D. M. and Carrive, P. (2005). Changes in cutaneous and body temperature during and after conditioned fear to context in the rat. *European Journal of Neuroscience* 21, 2505–2512.

Yu, Y. H. and Blessing, W. W. (1997). Cutaneous vasoconstriction in conscious rabbits during alerting responses detected by hippocampal theta-rhythm. *American Journal of Physiology* 272, R208–R216.

Relaxation Techniques

W G Whitehouse, E C Orne and M T Orne

University of Pennsylvania School of Medicine,
Philadelphia, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by M T Orne and W G Whitehouse, volume 3, pp 341–347,
© 2000, Elsevier Inc.

Progressive Relaxation

Self-Hypnosis

Autogenic Training

Meditation

Biofeedback-Assisted Relaxation

Conclusion

Glossary

*Behavior
therapy*

A system of psychotherapy that views mental disorders as products of maladaptive learning that are best treated by techniques derived from classical and instrumental conditioning principles.

*Systematic
desensitization*

A technique used in behavior therapy to eliminate specific fears by requiring the patient to imagine different situations involving the feared stimulus while remaining relaxed. To avoid overwhelming the patient, the imagined situations are often mastered successively in an anxiety hierarchy, from least to most fear-provoking.

Progressive Relaxation

For centuries, the prescription rest and relaxation has been a popular treatment for a wide range of maladies. The apparent curative properties of rest were being endorsed by physicians and healers long before

a scientific basis for the association between stress and disease had been established. A seminal contribution to understanding the pathophysiology of stress was the 1929 publication of Edmund Jacobson's *Progressive Relaxation*, an empirically based treatise on the role of neuromuscular tension in various somatic, neurological, and psychiatric disorders. In this work, Jacobson noted that many instances of physical and mental disease include among their symptoms the failure to relax – an ability that is often regained following recovery. Using electrophysiological measures, Jacobson confirmed the presence of neuromuscular tension in various disorders and was able to document the power of relaxation to dissipate stress-related symptoms. He also reported that mental operations, such as imagination, emotion, and recollection, were expressed in muscle contractions, a finding that proved to be an important milestone in the development of the field of psychosomatic medicine. Thus, for the first time, there was compelling scientific evidence to support the widespread belief among medical practitioners that stress and anxiety contributed to disease states and that relaxation-based therapies provided effective treatment options for many patients.

The procedures involved in progressive relaxation (PR) developed directly from Jacobson's scientific studies of muscular tension. A fundamental tenet of this approach is that tension and deep muscle relaxation are physiological opposites that cannot coexist. Thus, to the extent that neuromuscular tension underlies a particular disorder, relaxation should be an effective remedy. Moreover, PR allows patients to become actively involved in their own treatment, a feature that Jacobson regarded as an important psychological benefit of the procedure. The broader objective, however, was to teach patients to incorporate relaxation skills into their lifestyles so that they could become better able to respond appropriately

to future stressors by controlling their tendency to become tense.

Technique

A preliminary condition of training is the use of responses that allow the patient to recognize muscle contraction (i.e., tension) wherever it is present. The procedure requires the patient to alternately tense (i.e., contract) and relax (i.e., extend) each of the major muscle groups of the body. The goal of this exercise is to help the patient to discriminate degrees of muscular tension, which can be controlled as warranted or, in extreme cases, supplanted with deep relaxation.

PR attempts to cultivate whole-body relaxation. Patients are encouraged to assume a comfortable position and relax. Then they are instructed to clench the left or right fist as tightly as possible and to notice the tension as it creeps up from the fist to the hand to the forearm. This is followed by instructions to relax the fist and to pay particular attention to the difference they feel between the tensed and relaxed states. The therapist's instructions merely guide the patient's attention to actual experiences that accompany variations in muscle tension; there is no reliance on suggestive techniques, such as hypnosis. The procedure is repeated using the opposite fist, then the upper arms, the facial muscles, the chest, the stomach, and the lower back; it culminates in the muscles of the lower extremities. Each time the patient is encouraged to become aware of the difference in sensation between tension and relaxation and to allow relaxation to come to predominate.

As originally outlined by Jacobson, the procedure involved upward of 100 training sessions, many focusing on single muscle groups. Pragmatic concerns soon prevailed and abbreviated versions were developed, beginning with the efforts of psychiatrist Joseph Wolpe, who adopted relaxation as a means to countercondition anxiety responses elicited in the context of his behavior therapeutic method of systematic desensitization. Contemporary approaches to PR attempt to promote profound whole-body relaxation in a single session. They routinely involve the alternation between tensing and relaxing groups of muscles, but the procedures take place in a matter of minutes rather than days and they frequently include hypnotic-like suggestions to facilitate relaxation.

Clinical Effectiveness

PR has been applied clinically to various medical disorders in which indicators of stress, such as sustained autonomic arousal or chronic muscle tension, are present. Both controlled and uncontrolled studies

support its effectiveness in the treatment of anxiety disorders, depressive symptoms, anticipatory nausea associated with chemotherapy, tinnitus, insomnia, low back pain, hypertension, and tension headache. The evaluation of the specific contribution of PR in treating these conditions is often complicated by the inclusion of other potentially therapeutic components in the treatment protocols. At the same time, the case for some disorders (e.g., hypertension) generally is inconclusive, owing to the fact that, on an individual basis, stress may not be a primary etiological factor. As Herbert Benson and his colleagues have noted, relaxation can lower an individual's blood pressure, but the magnitude of the effect depends on the extent to which stress plays a role in the first place. A number of studies reported in the literature on the clinical effectiveness of relaxation techniques have failed to make this important initial determination.

Finally, classical PR may not be appropriate for certain disorders, due to its potential to exacerbate symptoms. This has been reported, for example, among individuals suffering from tension headache or myofascial pain disorders, as well as with the elderly, in whom the practice of alternately tensing and relaxing relevant muscle groups actually increased pain and anxiety. Fortunately, when such complications arise, PR can be modified to omit the muscle contraction component of the exercises, which Jacobson recommended only as a means to sensitize patients to the presence of tension in their bodies. In some studies, imaginal relaxation, in which participants are not required to enact muscle tension-release exercises but merely to imagine performing them, was associated with reductions in stress symptomatology. In addition, recent efforts to customize relaxation therapy to the functional abilities of individuals with spinal cord injuries, as determined by the level or site of injury, have produced positive outcomes in several studies.

Self-Hypnosis

The use of hypnosis in stress-related conditions is another technique that seeks to engender deep relaxation, typically with the use of direct suggestions that the subject experience calm, sleepiness, and contentment (see **Hypnosis**). A preponderance of data shows that virtually all hypnotic phenomena that can be experienced in a dyadic relationship with a therapist can also be induced by the patient using self-hypnosis. This is because heterohypnosis is fundamentally self-hypnosis, and vice versa. That is, the experience of hypnosis requires an appropriately hypnotizable individual who is motivated to accomplish the goals targeted by select therapeutic suggestions administered

in a distinctively hypnotic context. By the same token, self-hypnosis can be conceptualized as an extension of heterohypnosis in which the patient carries out a procedure taught in a therapeutic relationship and is able to reinstate feelings that were originally established in heterohypnosis by imagining the words of the therapist during the self-hypnotic exercise.

Self-hypnosis is best learned with the help of an experienced hypnotherapist who outlines the clinical objectives, dispels any misunderstanding about the procedure and its outcomes, teaches techniques to induce the condition of hypnosis, and works with the patient to establish appropriate suggestions and metaphors to facilitate therapy. Routinely, skill in self-hypnosis develops as the patient practices the techniques outside the therapist's office. To encourage regular practice and to optimize the therapeutic value of self-hypnosis, however, it is helpful for the patient to maintain his or her relationship with the therapist. We have found, for example, that in the context of stress and pain management, periodic therapy sessions, or even occasional telephone contact with the therapist, can sustain the patient's motivation to practice self-hypnosis and may even compensate for deficiencies in hypnotic ability that might otherwise produce discouraging outcomes. Attention to motivational issues is, therefore, an important aspect of preparing patients for positive experiences with self-hypnosis.

In addition to the capacity of self-hypnosis to promote deep relaxation, many individuals who practice the technique as a complement to an ongoing therapeutic relationship are also able to favorably modify their perceptions regarding the importance of the stressor, as well as their emotional reactions to it. Accordingly, for patients with sufficient hypnotic ability, self-hypnosis might well be the treatment of choice for conditions (e.g., chronic pain) that fail to respond adequately to relaxation-oriented therapies alone.

Autogenic Training

Closely related to self-hypnosis is the formalized system of auto-suggestive therapy known as autogenic training (AT). It was developed by Johannes Schultz as a form of psychophysiological therapy in which patients can be taught a series of exercises to induce in themselves a profound state of physical and mental relaxation similar to deep hypnosis. The method grew out of Schultz's own clinical work with hypnosis, during which he noted that hypnotized patients commonly reported the sensation of heaviness in their limbs, accompanied by feelings of warmth. These observations were taken and interlaced with his belief

that hypnotic suggestions were successful to the extent that the patient allowed them to happen. Specifically, in Schultz's analysis, hypnotized patients adopted a type of passive concentration, permitting the suggested changes to occur without trying to directly influence their occurrence. Hence, the blueprint for AT evolved. The purpose was (1) to achieve the psychophysiological effects typical of hypnosis without the need to formally induce the condition of hypnosis and (2) to train the patient to self-administer the relevant suggestions, thereby conveying a substantial element of responsibility for treatment from the therapist to the patient.

Technique

AT can be carried out on an individual or group basis, although individual training can more readily adjust to idiosyncratic differences in rate of skill acquisition or other needs. Environmental considerations include a slightly darkened room free of distractions, with a comfortable temperature and a chair, couch, or mattress. The verbal formulas developed by Schultz and used clinically for nearly three-quarters of a century fall into three categories of autogenic exercise: standard, meditative, and special. For illustrative purposes, we outline here only the standard autogenic exercises, whose mastery is a prerequisite for advancement to meditative (involving imagery) and special (designed to treat specific organic or mental disorders) exercises. Moreover, the standard exercises are the most widely used for stress management.

- **Heaviness.** Training in heaviness usually begins with the dominant arm, with the expectation that, as the exercise proceeds, heaviness will begin to generalize to other extremities. The trainee passively concentrates as the therapist repeats calmly, several times, "I am at peace. . . . My right arm (if that is the dominant arm, otherwise, left arm) is heavy. . . . My right arm is heavy. . . . My right arm is heavy. . . . My right arm is heavy." As the right arm becomes heavier, the trainee begins to notice a sense of heaviness in the left arm as well. This is reinforced as the therapist slowly repeats, "My left arm is heavy," a number of times, followed by "Both of my arms are heavy." The exercise is then extended to induce the spreading heaviness in each of the legs, concluding with the formula "My arms and legs are heavy."
- **Warmth.** The second standard exercise is intended to induce the experience of warmth in the arms and legs by increasing blood flow in the extremities. Once heaviness has been achieved, the second formula is combined with the first: "I am at peace. . . . My arms and legs are heavy. . . . I am at peace. . . . My arms and legs are heavy. . . . My right arm is

warm. . . My right arm is warm. . . My right arm is warm.” The procedure continues with repetition of the appropriate verbal formulas and passive concentration as warmth develops in the other arm and in each of the legs, culminating with repetition of the phrase “My arms and legs are warm.”

- **Cardiac regulation.** The third standard exercise focuses on the trainees’ awareness of their heartbeat. Many trainees do not readily perceive their heartbeats unless they are hyperaroused. Such individuals may require postural alterations to sensitize cardiac awareness, such as practicing autogenic exercises with the right hand placed over their heart. The relevant verbal formula that is the focus of passive concentration is “Heartbeat calm and regular.” The objective is not to slow the heart rate but to bring it into a consistent rhythm that supports the already established feelings of heaviness and warmth.
- **Respiration.** The fourth exercise is intended to teach the self-regulation of respiratory functions. Like the cardiac regulatory exercise that precedes it, the trainees’ respiration rates should have been substantially decreased during the heaviness and warmth exercises. However, unlike heartbeat activity, the rate of breathing is relatively easy to modify voluntarily, an inclination that must be resisted if passive concentration is to be maintained. Accordingly, the fourth formula is phrased “It breathes me” to emphasize the importance of passive focus.
- **Abdominal warmth.** The fifth standard exercise attempts to induce a generalized feeling of warmth deep in the region of the solar plexus. Many trainees require an anatomy lesson to localize this target of passive concentration (i.e., halfway between the lower segment of the sternum and the navel) before the therapeutic phrase “My solar plexus is warm” can be employed effectively. The exercise is contraindicated for people with medical conditions involving the viscera of the peritoneal cavity (e.g., diabetes, gastric tumors, and ulcers).
- **Coolness of the forehead.** The final standard autogenic exercise stems from the common observation that a cool cloth applied to the forehead reduces tension – an effect that can also be achieved by a redistribution of the blood flow away from the head (i.e., localized vasoconstriction) and toward the extremities (i.e., peripheral vasodilatation as accomplished in exercise 2). The relevant supplemental autogenic formula to be repeated at this stage of training is simply “My forehead is cool.”

Training is considered complete when the trainee is skilled at the self-administration of the autogenic

program. Each of the standard exercises works toward reinforcing the others to inhibit indicators of sympathetic nervous system activation, thereby promoting a generalized state of physiological and mental relaxation.

Clinical Effectiveness

AT is widely used in conjunction with biofeedback techniques to promote stress reduction, although many clinicians employ autogenic therapy as a solitary method to achieve self-regulation of the autonomic nervous system. Attempts to assess the specific effectiveness of AT by reviewing the relevant scientific and clinical literatures are, however, impeded by several factors (e.g., small nonrandom samples, poor controls, and wide variation in training protocols and extent of training). Quantitative reviews that ignore the idiosyncratic aspects of individual studies and focus, instead, on the magnitude of the treatment effects across studies, compared to no treatment or to other nonpharmacological techniques (e.g., biofeedback, meditation, progressive relaxation, or self-hypnosis) tend to conclude that AT is as efficacious, on the whole, as other biobehavioral interventions for most psychosomatic disorders. There is, however, some evidence that AT may specifically promote decreases in heart rate that are independent of the general lowered respiratory frequencies associated with other relaxation procedures such as Jacobson’s PR technique. Stress-related conditions that are claimed to respond positively to AT include respiratory problems (e.g., bronchial asthma) and circulatory disorders (hypertension, Raynaud’s disease, tachycardia, and cardiac arrhythmia), insomnia, and anxiety.

Meditation

Meditation refers to a collection of practices, generally of Asian origin, that induce a change from the ordinary in a person’s mental focus, with the result that metabolic activity is slowed and the practitioner often feels relaxed and refreshed. Two general forms have been identified, concentrative and nonconcentrative methods.

In the concentrative approach, typified by Transcendental Meditation, the individual is seated in a quiet environment and directs his or her focus on a single repetitive stimulus, such as a word or sound recited mentally as the person takes in or exhales breaths of air. The objective is to remain single-minded. If distracting thoughts intrude, the meditator is directed to simply dismiss them and to bring his or her attention back to the focal stimulus. The technique, along with its religious and philosophical trappings,

was popularized in the western hemisphere in the late 1960s by the Maharishi Mahesh Yogi, whose disciples included many celebrities of the period.

A prominent form of nonconcentrative meditation in use medically is mindfulness meditation, which is derived from Buddhist tradition. The technique does not necessarily lead to a profoundly relaxed state, as is typical of concentrative forms of meditation. Rather, the primary goal is to obtain insight into the self by learning to catalog moment-to-moment changes in experience.

Mindfulness meditation begins with a single object of focus (e.g., breathing) to establish calmness, but the mental focus is gradually expanded to include any ambient stimuli, thoughts, feelings, and physical sensations that enter awareness. Unlike concentrative meditation, practitioners of mindfulness do not regard these experiences as distractions to be ignored but instead dispassionately focus on each until their attention wanders, whereupon they simply note where their thoughts have taken them before redirecting attention back to their in-the-moment experiences. Various formal and informal exercises have been developed to facilitate the adoption of mindfulness as a lifestyle approach to foster an appreciation for the importance of living in the present moment.

Clinical Effectiveness

A number of scientific investigations of Transcendental Meditation were undertaken in the early 1970s by Robert Wallace and Herbert Benson, which documented the ability of meditation to induce deep relaxation and to oppose sympathetic arousal. Among the physiological changes produced during meditation were substantial decreases in the rate and volume of respiration, with correspondingly reduced levels of oxygen consumption and carbon dioxide production. In addition, a slowing of the heart rate, accompanied by increases in skin resistance and decreases in blood lactate levels signified an overall reduction in anxiety or arousal. Furthermore, studies using electroencephalograms (EEGs) confirmed the presence of a low arousal state – a pattern featuring a predominance of alpha activity, with occasional brief transitions toward theta and deep-sleep-like delta frequencies. What was not found in these early studies was evidence that meditation significantly affected blood pressure, which Benson and his colleagues soon realized was due to the fact that their sample of experienced meditators had quite low blood pressure to begin with. Subsequent research determined that meditation was effective in reducing blood pressure in individuals with stress-related hypertension. Recent

studies have also found that meditation is associated with improvements in respiratory and cardiovascular functioning, sleep disturbance, and depression, as well as in the treatment of addictive disorders.

As mentioned previously, nonconcentrative forms of meditation, such as mindfulness, do not uniformly induce relaxation. In fact, although relaxation states often occur when practicing the technique, because the intended goal is greater self-awareness no particular physiological outcome should be anticipated. Nevertheless, mindfulness meditation appears to be notably useful for the management of chronic pain, in which learning to identify the everyday pain sensations apart from their emotional and evaluative components helps to lessen the degree of suffering that is otherwise endured. Preliminary evidence suggests that training in mindfulness meditation may also be of benefit in women with breast cancer by reducing stress and anxiety and improving their outlook for recovery.

Biofeedback-Assisted Relaxation

Many of the bodily processes affected by relaxation techniques are normally unavailable to conscious awareness; hence, they elude any casual attempts at voluntary control. Biofeedback rectifies this situation through the use of sensors and instrumentation capable of detecting, amplifying, and displaying the relevant biological signals. Armed with information regarding the moment-to-moment status of their blood pressure, heart rate, muscle tension, or electrical activity of the brain, patients are in a position to observe any correlations between thoughts and behavior and corresponding changes in these physiological modalities. Once such relationships are discovered, it becomes possible to achieve self-regulation of these systems with the help of biofeedback. In many cases, this self-regulation can be maintained outside the therapist's office (i.e., without reliance on instrumentation) by carrying out the cognitive or motor strategies that were previously demonstrated to affect relevant aspects of the person's physiology.

Biofeedback is used in the treatment of many disorders, some with central nervous system involvement (e.g., cognitive and attentional problems, and depression), others involving the circulatory system (e.g., cardiac arrhythmia, hypertension, and orthostatic hypotension), and others involving localized neuromuscular dysfunction (e.g., urinary or fecal incontinence and tension headache). As an aid to relaxation, however, two modalities are generally preferred: electromyographic (EMG) feedback and temperature feedback.

Electromyographic Biofeedback

EMG biofeedback is employed to monitor and display muscle tension. Although almost any muscle that can be monitored with skin surface electrodes is a viable target for EMG biofeedback, three muscle groups tend to be favored: the frontalis muscle (forehead), the masseter (jaws), and the trapezius (shoulders). The reason these muscles receive special attention is that they are particularly prone to contract during stressful situations. The goal of EMG training is to use the biofeedback information to learn particular cognitive or behavioral strategies that reliably induce relaxation in the targeted muscle site. EMG biofeedback is designed to accomplish the same objectives as Jacobson's progressive relaxation technique while permitting the patient to observe his or her progress.

Temperature Biofeedback

Thermal biofeedback is used to relay information about skin temperature to the patient. To accomplish relaxation, the patient is trained to increase the skin temperature of the hands and/or feet. The instrumentation can be quite simple (e.g., an outdoor thermometer taped to a finger or toe is usually adequate), making it convenient for patients to practice the technique at home. The physiological mechanism that supports the intended alterations in skin temperature is an increase in blood flow in the extremities, which opposes the usual sympathetic response to stress, involving peripheral vasoconstriction.

Clinical Effectiveness

Biofeedback-assisted relaxation is the most commonly employed application of biofeedback techniques. EMG biofeedback procedures have been found to be effective in the treatment of anxiety, insomnia, and various psychosomatic disorders, such as colitis, hypertension, gastric ulcers, and tension headaches. Thermal training has been particularly helpful in the treatment of migraine headache and Raynaud's disease (a vascular disorder involving the sensation of extreme cold in the hands and feet). It is frequently

combined with autogenic phrases, imagery, and breathing exercises.

Conclusion

Whether provoked by palpable adverse life events, microscopic pathogenic agents, or simply our own misperceptions and anxious thoughts, stress is an inescapable part of human experience that demands biological and behavioral adjustments. Unfortunately, when exposed to persistent stress, constitutionally vulnerable individuals or people with inadequate coping resources may suffer harmful health and psychological consequences. The development of relaxation and stress-reduction skills provides an often pleasurable and cost-effective approach to moderating the experience of stress before it takes a serious toll. An important caveat, however, is that each of these techniques has the potential to modify physiological and psychological processes in beneficial and sometimes harmful ways. It is, therefore, strongly advised that individuals motivated to develop relaxation skills consult with a qualified physician, psychologist, or dentist who can recommend an appropriate program that is, at the same time, suited to the person's lifestyle. The most significant benefits of relaxation techniques usually accrue when practice becomes part of a daily routine.

See Also the Following Articles

Anxiety; Cytokines, Stress, and Depression; Depression and Manic-Depressive Illness; Depression Models; Hypnosis.

Further Reading

- Benson, H. and Klipper, M. Z. (2000). *The relaxation response*. New York: Harpertorch.
- Davis, M., Eshelman, E. R. and McKay, M. (2000). *The relaxation & stress reduction workbook* (5th edn.). Oakland, CA: New Harbinger.
- Kabat-Zinn, J. (1991). *Full catastrophe living: using the wisdom of your body and mind to face stress, pain and illness*. New York: Delacorte.
- Lehrer, P. M. and Woolfolk, R. L. (1993). *Principles and practice of stress management* (2nd edn.). New York: Guilford.

Religion and Stress

S Packer

New School for Social Research, New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S Packer, volume 3, pp 348–355, © 2000, Elsevier Inc.

Religion as a Concept

Religion as a Cure for Stress

Religion as a Cause of Stress

Religion as a Correlate of Stress

Glossary

Apocalyptic religion

A religion based on writings prophesying a cataclysmic time when evil forces are destroyed; includes millenarian religions that focus on the end of 1000-year period described in the Book of Revelation.

Cult

A pejorative term used to describe a new religious movement, especially one that is secretive, is socially isolated, is faddish, has unfamiliar rules and rituals, and uses coercive or manipulative or deceptive techniques used to win or keep new converts.

Prophetic religion

A religion that maintains that truth is revealed by prophecy; sometimes referred to as religions of revelation. Prophetic religions include Western monotheistic religions such as Judaism, Christianity, and Islam, as well as Zoroastrianism and Bahai.

Religion

A system of beliefs centered around the concept of a supernatural being or force.

Transcendent religion

A religion that teaches techniques for transcending or transforming the reality of the senses or that denies the existence or importance of mundane reality. Transcendent religions include many Eastern religions and some Western New Age beliefs.

Religion can be both a cause of stress and a cure for stress. Some go so far as to say that it is religion's ability to relieve the very stress that it induces that is partly responsible for its tenacious hold on its truest or newest believers. It has also been noted that religious sentiment and sectarianism rises during times of increased personal or societal stress. Because change and uncertainty are potent causes of stress, on a psychological, sociological, and physiological level, it is

not surprising that some of history's most unusual religious movements have sprung up during the times of rapid social change and uncertainty. Individuals undergoing personal stress and lifestyle shifts are statistically more likely to get involved with unusual or innovative religious movements. Thus, adolescents who have recently left home, the newly divorced or bereaved or relocated, and prison inmates are especially susceptible to the appeal of new religions or radical religions. But charismatic religious groups or cults are not the only religious groups that gain momentum during times of stress; even the ranks of more garden-variety religions grow when the stresses of poverty, old age, war, illness, isolation, or impending death increase.

Religion as a Concept

Religion is anything but a single, all-embracing phenomenon. There are over 100,000 registered religions in the United States alone. The number of registered religions is increasing, not decreasing, and has been on the rise in the United States since the 1960s. The number of people who accept a personal spirituality, rather than a recognized religion, is rising even faster, to the point that spirituality is now a more politically correct and more all-embracing term that bypasses the conservative connotations of the word religion.

Even the meaning of the word religion is a source of contention. For the sake of simplicity, we define religion here as a system of beliefs centered around a supernatural being, power, or force. But, for the sake of accuracy, we must add that many well-respected definitions of religion that include systems of belief do not include a belief in the supernatural and that some well-known religions, such as Confucianism, are not concerned with a supernatural entity.

Given the large number of recognized religions in existence, it stands to reason that religions can differ dramatically from one another. Furthermore, some organized religions include far-flung sects and subsects that bear little superficial resemblance to the parent religion, and some (but certainly not all) religions host a wide variety of views on different issues. It may be tempting to think that religious beliefs, rites, customs, and concerns can be plotted out on a spreadsheet, so that different religions can be directly compared to one another. However, this straightforward approach of scientists leads to an inaccurate appreciation of the vagaries of religion and overlooks the fact that one religion stresses rites of worship,

another concerns itself with right thinking, a third revolves around prayer or penitence, a fourth focuses on gender differences and reproduction issues, a fifth emphasizes the afterlife, and so forth. It is unlikely that each religion has something to say about every aspect of life (or death), even though each religion represents itself as the ultimate authority on human life.

Religion as a Cure for Stress

Both its fiercest critics and its most fervent advocates agree that religion (or at least some religions) can dampen psychological stress. Freud himself, that arch-enemy of religion, admitted as much in his seminal book on *The Future of an Illusion*. Personal testimonials and clergymen's sermons about religion's potential to relieve stress are too numerous to cite here. Religion is so commonly associated with stress that it is virtually a reflex to invoke it at moments of stress. Many images of religion and extreme stress spring to mind immediately, from foot-stomping spirituals of the segregation-era South to the footsteps of the death-row dweller listening to the last words of the seventy-first psalm to the prayers mumbled by chaplains for charging troops who face death on the front. We can even consider the fearlessness of the 9/11 suicide bombers who embraced their own death as (perverse) testimony to the power of faith.

Although not everyone finds relief in religion, there are some compelling sociological statistics supporting religion's overall ability to combat stress. It is well-established that people with strong personal religious beliefs and affiliations are less likely to commit suicide, which can be seen as the ultimate and most extreme response to stress. The fact that some Islamic extremists encouraged suicide missions, in spite of Islam's strong stance against suicide, is an aberration.

Philosophical Aspects

By positing the existence of a supernatural being who actively intervenes in the natural world and who has a plan for each individual human, religion provides an automatic antidote to the stress of being alone and awash in an unresponsive and uncaring universe. Philosophers and theologians alike note that existential angst and dread have plagued twentieth-century society, at least in the West. Ever since Nietzsche declared that "God is dead" as the nineteenth century ended, and as his own symptoms of neurosyphilis progressed, unwavering faith in the supernatural melted away and slowly but surely became the foremost source of stress. Only the lurking specter of a sudden nuclear holocaust overshadowed this stress. Pastoral pop psychologists tried to reverse the tide

that Nietzsche set into motion. Best-selling books, such as Paul Tillich's *Man Is Not Alone*, played on Nietzsche's angst-inducing assertion in its title and appealed to this angst. Similarly, film titles and themes recapitulate the reflexive tendency to call on religion in times of crisis. For instance, the film *God is My Co-Pilot* appeared during World War II. More veiled religious themes bubble to the surface during other times of stress. When the Great Depression cast its shadow, *Lost Horizon* evoked legends of Shangri-La, where Tibetan monks and Christian missionaries fused good works and good karma and reinfused human souls. Then, as the twentieth century came to a close, and as second millennium neared – at the very same time that information technology zoomed ahead and rocked standard channels of communication – religious imagery once again resurfaced in secular guise, this time in the form of angels. New Age angels, which were retrieved from Renaissance-era paintings, appeared on postage stamps.

Psychological Aspects

Antidote to aloneness, abandonment, and dependency

When psychoanalysts try to explain why a belief in the supernatural relieves stress, they look for more personal, rather than societal, explanations. The consistently controversial Freud claimed that a belief in God substitutes for the presence of an all-protective, albeit sometimes punitive, parent. According to such theories, religion recreates the reassuring world of dependency that an infant enjoys. The oceanic feeling of connectedness with the universe that mystics describe could be a way to return to the peacefulness and protectiveness of the womb, where an umbilical cord delivered never-ending nourishment and where amniotic fluid buffered the baby from unkind blows from the world around. Religious redemption concepts that revolve around the return to Eden can be compared to psychoanalytic fantasies about the return to the womb.

More contemporary neo-Freudian psychoanalytic thinkers, such as Karl Winnicott, emphasized the importance of separation-individuation as the critical step in psychological growth. Unlike more orthodox Freudians, who viewed religious reassurance as a sign of psychological immaturity and as an obstacle to self-actualization, theorists such as Winnicott felt that religious belief functions as a useful transitional object, much like a baby blanket or a teddy bear. Both remind the child of the parent's perpetual presence and assuage a child's subjective stress of abandonment during the parents' absence. There are many, many more psychoanalytic theories about the ways that religion relieves stress, each unique and interesting in

its own right and each as untestable by scientific techniques as the tenets of religion itself.

A more pragmatic and less theoretical take on religion as a surrogate dependency is found in the principles behind the Twelve-Step and Recovery Programs and Alcoholics Anonymous. These programs provide a strong social support system and also call on a Higher Power to relieve the psychological and physiological stress of alcohol withdrawal and as an aid in dealing with stresses that were previously ameliorated by addictions. Rather than admonishing their members for depending on this spiritual Higher Power, as more Freudian-influenced therapists did in the past, the Twelve-Step groups see spiritual dependency as far less harmful than chemical dependency. Such pro-spiritual approaches have gained increasing public and professional acceptance and have been adapted to the treatment of many other disorders.

Decreased sense of randomness and uncertainty

There are many other ways that religion relieves personal stress. Religion decreases the sense of randomness, uncertainty, and chaos in the world because it details the logic (or illogic) behind a world plan. It predicts or sometimes prophesies or, at the very least, attempts to explain those events that seem most inexplicable. In doing so, it creates a concept of stability that has been described as an ordered universe. Even religions that foretell the coming of adverse events in the future, such as an impending apocalypse, reincarnation, predestination, or bad karma, have the potential to offer partial relief from stress, simply because they help their members prepare themselves psychologically for adverse events. Providing explanations for events, even without providing a means for altering those events, is referred to as heuristics, and this is one of the most powerful psychological tools that religion possesses. Some critics of psychoanalysis say that its appeal also rests on its heuristic value and argue that unprovable psychoanalytic explanations provide reasons for behavior rather than remedies, in the same way that religion provides reasons for cosmic events and human nature.

Religions further relieve the stress of uncertainty by providing blueprints for behavior that can theoretically change the future. By prescribing prayer, penance, codes of charity, dietary laws, sacrificial rituals, or right thought, those religions reestablish a sense of personal control. More optimistic prophesies, and promises of paradise, confer hope and escape.

Cognitive coping techniques Above and beyond any theoretical parallels with psychoanalysis, many religions also provide specific cognitive techniques that are useful in coping with stress. For instance, the mere

act of acknowledging and articulating the experience of stress can quell some distress. Long before Freud discovered his talking cure, religious prayers and petitions provided collective voices for stress and distress. Hebrew psalms that begin with words such as "From the depths I called out unto Thee" poeticize an acute sense of anguish, whereas Christian recreations of the crucifixion scene dramatize the ordeal of Christ on the cross. These words and images are especially appealing to people who do not have access to other avenues of expression, or who prefer to deflect their own subjective sense of distress by focusing on more universal, cosmic, or collective stresses.

Role models for stress endurance Religious lore offers as role models people who withstood extreme stress and either surmounted that stress, were valued because of their ability to endure stress, or went on to fulfill higher purposes because of those stresses. The Christ-figure is the consummate example of this sort, but it is hardly the only one. Biblical stories about Jonah being swallowed by a whale, Noah's ability to withstand a flood, Daniel's survival in the fiery lion's den, and Job's endurance of the loss of his family and his health all reassure believers that there is a relief for stress or, if there is no immediate relief, that there is a reason or perhaps even a reward in another life. The martyrdom of Christian saints before their beatification, the Buddhist *jataka* stories about Prince Shakyamuni's wandering as a mendicant monk before he achieved enlightenment as he sat under the Bodhi tree, and the image of the Israelite tribes wandering through the desert for 40 years before entering the Land of Israel and acquiring their status as the chosen people are other examples of the pivotal role that stress endurance plays in religious themes.

Consolation, devaluation, and dissociation Some religions provide such a sense of consolation that they have been dubbed the religions of consolation. Christianity's contention that the meek will inherit the Earth reassures people that their this-worldly stress and suffering will be relieved by other-worldly rewards. The promises of future paradise made by prophetic religions dull the pain of the present and appeal to the impoverished, the downtrodden, and the psychologically stressed. Some religions offer so much consolation for contemporary stress and distress that they devalue the material world completely by undermining its ultimate importance or by teaching that the real world is but a delusion and no different from a dream. Buddhism's belief that the material world is nothing more than a delusion, or *maya*, is an extreme example of such transcendent thinking. The Hindu yogin's aspiration to a waking

state of dreamless sleep is another example of transcendence. Some psychotherapeutic techniques train patients to use similar techniques of detachment when confronted with stress-producing stimuli, although it cannot be overlooked that people who enter such dissociative states spontaneously and nonvolitionally develop serious difficulties in life.

Temporal and physical escapes from stress Even world-affirming religions, which affirm the importance of the material world, often offer temporary escape routes from daily stress. Religious holidays and religious services and other sacred times carve out stress-free time during the ordinary workweek and create an opportunity for rejuvenation. For those times when real-world stresses require even more relief than routine religious beliefs or rituals can confer, some religions offer physical as well as temporal retreats where food, clothing, and shelter are available, along with social support, structure, and spiritual exercises. Ashrams, yeshivas, monasteries, convents, and any number of other religious communities provide parallels to the retreats popular among certain Christian denominations. Such religious retreats not only legitimize the need for relief from stress, but also consecrate the choice to retreat from stress. Such religious retreats thereby provide participants with a renewed sense of self-worth, along with a sense of connectedness with others who share their belief system, and may even offer retraining in new vocations or avocations. Some retreats encourage members to contribute to society by doing good works or charity. The modern hospital movement evolved out of the monastic retreats in the Middle Ages, where people who originally sought relief for spiritual, physical, psychological ailments eventually provided care for others after their own recovery. In contrast, psychiatric rest cures and funny farms, which also offered retreats from the real world, stigmatize the participant, pathologize the process, and end social productivity. It is no wonder that religious retreats are often preferred over psychiatric treatment and that religion is often the first defense against stress.

Social Aspects

For some people, and for some religions, religion is a solitary matter and nothing but. For them, it is personal spirituality that matters most, with respect to stress relief and everything else. Although the title of William James's often-republished book is *The Varieties of Religious Experience*, in it James focused exclusively on religion that is experienced in solitude. Nevertheless, religion exists on a social as well as a personal level and acts to relieve social stress (or

produce stress) on both levels as well. By providing a social support network through their communal services and activities, coupled with a sense of collective identity and purpose, religious organizations can directly combat the stress of loneliness, displacement, and *anomie* that Emil Durkheim implicated in his early sociological studies of suicide.

Some religions provide material benefits, such as charity, lodging, employment, social services, and subsidized medical care, which can counteract the intense stresses of economic hardship and ill health and can complement the psychological consolation inherent in religious theory. Religious bureaucracies also offer alternative, and much appreciated, channels for personal and political expression, particularly when access to legitimate political clout is closed. Moreover, organized religion can become powerful enough to challenge the existing political and economic powers and bring about lasting and legitimate social and legal change.

Studies of the conversion process are especially illustrative of the importance of social forces and religion. It often comes as a surprise to people who are ideologically committed to religion to learn that it is the social sway of a religious group that influences individuals to adopt new religious ideas rather than the other way around. In other words, rather than experiencing a life-changing epiphany before converting, similar to the epiphany that the Gospels attributed to Saul of Tarsus on the road to Damascus, converts to new religions are more likely to follow a socially paved path to new religious insights. They accept more and more of the ideas of the group as they gain greater and greater acceptance into that group and as they become more reliant on group members for social support. The process of conversion is more likely to be gradual than sudden; in most people, behavior changes incrementally rather than dramatically. Nor does everyone who gets involved with a new religious group remain committed to that group. Of the many factors that correlate with an individual's long-term involvement with a new religion, one of the most important factors revolves around the role of stress. The more stress relief that an individual experiences at the time of joining a new religious group, the more likely that person is to remain a member of that group. Furthermore, it is the stress that individuals experience when distancing themselves from such groups that typically compels them to return to those groups.

Because it is well known that people who are already in a state of stress stand to achieve the greatest degree of stress relief from a new religion, some such religions make it a point to stress their newest members. Techniques such as sleep deprivation, diet

restriction, social isolation, overwork, enforced silence, sexual abstinence, or even physical or psychological threats make new recruits more receptive to the stress-relieving effects of the religion. Religions that rely on such insidious, quasi-coercive stress-inducing techniques are often denounced as cults and are disdained by nonmembers or former members. Such cultish religions predictably recruit in institutional settings, such as colleges, prisons, and retirement communities, where people are already in high stress states. Because such residents are pre-primed for proselytization, new religions have made major inroads in such places, at times to tragic ends.

It is the intense stress of prison life, rather than an earnest desire to reform or to atone their past acts, that propels some inmates to seek relief through religious conversion. Many ministries maintain an active presence in prisons because they recognize this potential. Sometimes, religious recruitment goes hand in hand with political indoctrination and can lead to involvement in militant religious factions. Many Black Muslims and Seven Percenters, including Malcolm X, traced their conversion to their incarceration. Yet, until 9/11 made the U.S. public more aware of this potential, there was little public opposition to religious recruitment in prisons. Most people assumed that religions strive to instill socially acceptable values that will substitute for criminal behavior, aid in rehabilitation and reentry into society, and thwart future incarceration. The recruitment of prisoners to a new religion was preferred over recruitment by antisocial prison gangs. However, in our post-9/11 world, we must consider the possibility that some extremist religions can instill antisocial attitudes and push some individuals into committing extreme acts.

In contrast, there has been a great public outcry about cult recruitment on college campuses, partly because some cultish religions have the opposite effect of prison-based cult religions. Rather than aiding the recruit's reentry into society, these cults isolate students from society and abort their attempts to achieve the status of other adults in society. Some cults cut off members' contact with families and friends, and cause severe stress in their relatives as a result. Although many legitimate religions were considered to be cults in their early stages – with Methodism being the most notable U.S. example – the mere fact that a small percentage of cults have been associated with mass suicide or homicide is enough to cause public concern and to stimulate psychiatric task forces to investigate these issues. The memories of 800 deaths in Jonestown, the loss of 19 lives in California's Heaven's Gate suicides, the subway poisonings by Japan's Aum Shinrikyo sect,

and the murder of a Hare Krishna defector in North Carolina tend to overshadow studies that show the psychological benefits of conversion in some individuals.

Somatic Aspects

Stress is as much a physiological response as a psychological one. Some religious rites use physical means to achieve higher spiritual states and alleviate both psychological and physiological stress in the process. The muscle stretching systems of yoga, the breathing exercises of Zen, and the controlled movements of Tai Chi are just a few examples of Eastern-influenced mind-body methods that gained popularity in the West. Some of these methods are now taught in continuing education courses sponsored by the American Psychiatric Association.

Although it is impossible to verify the existence of higher spiritual states, it is possible to measure changes in heart rate, respiration, galvanic skin response (sweating), pupil size, secretions of stress hormones such as cortisol or gastric acid, brain waves (through an electroencephalogram, EEG), and muscle tension in people performing those spiritual exercises. Internist Harold Benson's studies of cardiovascular effects of Transcendental Meditation and Tibetan Buddhist chanting appear in juried medical journals and in his popular paperback, *The Relaxation Response*. Japanese psychiatrist Tomio Hirai correlated the EEG and electromyogram (EMG) effects of both Zen meditation and Hindu yoga with reports of spiritual and psychological states and published his results in the curious book *Zen Meditation and Psychotherapy*. These works represent serious research, but there have also been many hyperbolic claims about the benefits of Eastern (and Western) religious systems, leading some professionals to dismiss any and all claims reflexively and prematurely. Recent scientific studies confirming the correlation between stress relief and specific religious rites suggest that both an open mind – and an open eye – can help us to appreciate confirmed the mind-body benefits of religion and also to avoid misleading claims and cults.

Quite the opposite of the calming meditative techniques of Eastern religions are the ecstatic dances, shaking, quaking, rocking, and rolling movement of some sects, which inspired the names Holy Rollers, Ranters, Ravers, Quakers, Shaking Quakers, and Shakers. Such intense activity presumably relieves stress through the same mechanisms that jogging and exercise help more secular devotees. The Dionysian dances of Classical Greece, described in Euripides' play about the Bacchae, were but one of many recurring manifestations of frenzied religious dancing that

serves a related function. Repetitive religious rituals in general are said to relieve stress as well.

Side Effects of Stress Relief

Religion's efficacy at relieving stress does not come without side effects. There are times when religion relieves stress so well that its practitioners are unaware of the demands of the real world and are unable to mount a defense against impending danger. Many social scientists have said that the religious devotion of African Americans shielded them from the pain and poverty of pre-Civil Rights America and delayed the adoption of more appropriate political tactics. Similarly, the Tibetan practice of sending one-third of its youth to Buddhist monasteries left the country defenseless against the Chinese invaders who destroyed temples, massacred monks and nuns, and sent their religious leaders into exile. Some secular Zionists claimed that the insulated and self-satisfied religious infrastructure of some eastern European Jewish communities obscured their awareness of the deadly fate that awaited Jews in Hitler's death camps. An analogous U.S. tragedy occurred when Native American warriors went into battle unarmed, believing that their Ghost Dance ritual would protect them from the guns and arrows of approaching armies. Marx and Engels blamed the mystical and occasionally bizarre religious belief of Russian orthodoxy and schismatics for numbing Russian reactions to the material exploitation by the capitalists and the Czars. This observation prompted Marx to coin his oft-quoted characterization of religion as "the opiate of the masses."

Religion as a Cause of Stress

As effective as religion can be at relieving stress, it can also produce stress. Threats of an afterlife full of hellfire and brimstone or of an impending apocalypse that will destroy the world are obvious stress-producers. On a more subtle level are the many moralistic demands and behavioral codes made by religion, which are often difficult to live up to and, thus, tend to leave some practitioners in a near-constant state of imperfection and incompleteness. Some practitioners adopt even more zealous beliefs and behavior in order to avoid that stress, in an ever spiraling pattern. Skeptics such as psychiatrist John Sargeant observed that some charismatic religions exploit this tandem stress relief–stress production effect when they proselytize and compared this push–pull effect of religion to behavioral conditioning, more nefarious methods of mind control and brainwashing, and even drug addiction.

For sure, religion-induced stress is not limited to perceived threats, nor is it confined to the personal psyche. In spite of its promises of eternal peace, organized religion has indeed injected some of the most realistic threats the world has ever witnessed. The Inquisition, the Crusades, and the wars of religion are but a few testimonies to reality-based stresses posed by religion in the past. Such stresses persist to the present day through suicide cults such as the California-based Heaven's Gate, Islamic terrorist attacks on the World Trade Center, the biological-weapon-wielding Aum Shinrikyo in Japan, the Hindu–Muslim–Sikh conflicts in an atomic-bomb-armed India, the religious-ethnic conflicts of the Balkans that set the stage for World War I, and right-wing religious assassinations in Jerusalem, to name just a few. Although some religions aspire to the day that "the lion will lie down with the lamb," that day has not yet arrived. Religion has been, and probably always will be, as intimately associated with the stress of war and violence as with the proverbial love and peace that it promises. Rising religious fundamentalism promises to produce more political and personal stresses for individuals and for the world at large. At the same time, the religions of reassurance will provide personal stress relief for individuals, creating a never-ending seesaw.

Religion as a Correlate of Stress

The correlation between the rapid rise of new and sometimes radical religious movements and the degree of social stress is nothing less than remarkable; it is cataloged in Norman Cohn's classic *Pursuit of the Millennium* and in his more recent *Chaos, Cosmos, and the World to Come*. A more recent book, *Radical Religion in America*, zeros in on extreme religious schisms that have arisen since 1970, during the times of high stress that followed the social turbulence of the 1960s. It is noteworthy that psychoanalysis, a once-esteemed and all-embracing psychiatric treatment that survives as a mere shadow of its former self, has also been accused of being a cult and of inducing a cultlike devotion that resembles religion more than science. Note that psychoanalysis's popularity peaked in the 1950s, during the times of high stress that followed World War II.

Further Reading

- Cohn, N. (1993). *Cosmos, chaos, and the world to come*. New Haven, CT: Yale University Press.
- Freud, S. (1961). The future of an illusion (1927). In: Strachey, J. (ed.) *Standard edition of the complete*

- psychological works of Sigmund Freud* (vol. 21), pp. 3–56. New York: W. W. Norton.
- Fuller, A. (1994). *Psychology & religion* (3rd edn.). Lanham, MD: Rowman and Littlefield.
- Gallant, M. (ed.) (1989). *Cults and new religious movements*. Arlington, VA: American Psychiatric Association.
- Girard, R. (1993). *Violence and the sacred*. Baltimore, MD: Johns Hopkins University Press.
- Hirai, T. (1989). *Zen meditation & psychotherapy*. New York: Tokyo Publications.
- James, W. (1982). *The varieties of religious experience*. New York: Penguin Books.
- Kakar, S. (1982). *Shamans, mystics, and doctors*. Chicago: University of Chicago Press.
- Kaplan, J. (1997). *Radical religion in America*. Syracuse, NY: Syracuse University Press.
- Kinsley, D. (1996). *Health, healing, and religion*. Upper Saddle River: Prentice-Hall.
- Klass, M. (1995). *Ordered universes*. Boulder, CO: Westview Press.
- Meissner, W. W. (1984). *Psychoanalysis and religious experience*. New Haven, CT: Yale University.
- Ostow, M. (ed.) (1982). *Judaism & psychoanalysis*. London: Karnac Books.
- Packer, S. (1998). Jewish mystical movements and the European ergot epidemics. *Israel Journal of Psychiatry* 35, 227–241.
- Paloutzian, R. (1996). *Invitation to the psychology of religion* (2nd edn.). Boston, MA: Allyn & Bacon.
- Sargant, W. (1957). *Battle for the mind*. New York: Penguin Books.
- Schumaker, J. F. (1992). *Religion and mental health*. New York: Oxford University Press.

9/11, Religion and Stress

S Packer

New School for Social Research, New York, NY USA

© 2007 Elsevier Inc. All rights reserved.

The Who, What, Where, How, and Why of 9/11
Paradoxes about Post-9/11 Stress, Suicide, and Religion
Muslims and Mental Health after 9/11
Jewish Issues Associated with 9/11 Stress
Concerns of Christians and Others after 9/11

Glossary

<i>9/11</i>	September 11, 2001, the day that the World Trade Center was destroyed and the Pentagon was attacked by organized conspirators led in absentia by Osama Bin Laden.
<i>Islam</i>	A prophetic religion that is based on the teachings of the prophet Mohammad, as recorded in the Koran. (Also called Mohammadanism.) Followers of Islam are known as Muslims and may be of any race or ethnicity and may live on any continent. Many, but not all, Muslims are also Arabs.
<i>Posttraumatic stress disorder (PTSD)</i>	A constellation of symptoms that occur after experiencing or witnessing a life-threatening episode. People with PTSD reexperience the original trauma through nightmares or flashbacks and feel detached or numb in daily life.

World Trade Center (WTC)

A complex of office buildings (also known as the Twin Towers after the two tallest buildings) located in lower Manhattan, New York City, that housed 40,000–60,000 workers. It was destroyed on September 11, 2001 (9/11), and the site where the WTC previously stood came to be known as Ground Zero. A religious and political belief that the political state of Israel is the Jewish homeland. Most (but not all) religious Jews are Zionists. Some secular Jews are Zionists, and many secular Jews are not Zionists. There are many different degrees of Zionism.

Zionism

An association between stress and religion has always existed, long before September 11, 2001, came to be known as 9/11 and long before the WTC was destroyed. Before 9/11, religion was most often seen as a cure for stress, but it could also be conceived of as a cause of stress and was certainly recognized as a correlate of stress. But on 9/11, the relative importance of those associations shifted. For most people in the United States, and especially for people who lived and worked in and near New York, religion suddenly shifted to being a serious source of stress. In fact, some very preliminary studies showed that people who were most religious were more likely to experience more stress after 9/11, although the reverse is usually true. In general, religious belief and affiliation offer protection against stress, but the reasons why

religion failed to protect against 9/11-related stress will become obvious during our discussion.

The Who, What, Where, How, and Why of 9/11

What was 9/11, and what made that day so significant to Americans in general and to New Yorkers in particular? On that day, the WTC was destroyed by young Islamic suicide bombers who flew two hijacked airplanes into each of the Twin Towers. At nearly the same time, another hijacked plane hit the Pentagon in Washington, D.C. A fourth passenger plane, bound for San Francisco, was also hijacked, but was commandeered by passengers and diverted from its target destination. That plane crashed outside Pittsburgh and claimed far fewer lives than it was intended to. Each of these events was part of a well-orchestrated and partly successful conspiracy to fell the United States because these Muslim extremists led by Osama bin Laden viewed the United States as an aider and abettor of Zionists who controlled the state of Israel. It should be noted, however, that this Tuesday morning attack on the WTC was not the first. Several years earlier, a van containing a bomb exploded in the WTC parking lot, killing six people and causing commercial chaos in subway-level shops; this culminated in the conviction of a blind sheikh who led this militant Islamic sect.

Some 2000 lives were lost during the second WTC attack on 9/11. Several hundred others died in the Washington area. None of the passengers aboard the hijacked planes survived. Although citizens of many cities, indeed, many countries, were affected by the tragic events of that day, 9/11 was most closely associated with New York City after the silhouettes of the WTC's Twin Towers vanished from the city skyline.

Situated in downtown Manhattan, near New York City's financial district and the world-famous Wall Street, the WTC also impacted ordinary New Yorkers who lived nearby. Those people were not necessarily involved in the financial or insurance or communications businesses that were centered in the WTC. Some New Jersey bedroom communities, short commutes from the WTC, were also impacted after dozens of their residents who worked in the Towers died that day.

The WTC was close to the bustling streets of Chinatown, a mecca for tourists and the culturally curious and home to long-time residents, recent arrivals, and legal and illegal immigrants from mainland China, Taiwan, Hong Kong, Vietnam, and Malaysia. Buddhists, Christians, Catholics, Confucians, Taoists, and Chinese ancestor worshipers coexist on the streets of Chinatown and often intermingle within a single family unit united by its shared Chinese

ancestry but not split by religious affiliation, as Western families often are. The once-lively Little Italy was also within walking distance of the WTC, as was the historically Jewish Lower East Side. In Little Italy, festivals named for regional Roman Catholic saints (such as San Gennaro) rose beyond their religious origins and attracted carnival-loving crowds of all faiths. It was to the Lower East Side, nostalgically known as Delancey Street, that poverty-stricken Jews fled the pogroms at the turn of the twentieth century. They turned their pushcarts into a sprawling Sunday shopping district that gradually gave way to trendy clubs with outrageous names like Lansky Lounge. The once-grand synagogues built by struggling immigrants were often in need of restoration because younger generations abandoned their grandparents' turn-of-the-century terrain and fled to suburbia. The Yiddish theaters that once lined Second Avenue below Fourteenth Street were now reduced to engraved copper stepping-stones that were embedded in sidewalks, engraved with names of yesteryears' Yiddish-speaking stars.

The largely Hispanic neighborhood of Loisada, or Alphabet City, was also a healthy walk from WTC, even if it was not near enough for lunch. Dotted with small Catholic churches, Pentecostal centers, *espirismo* storefronts, and even some *santeria* supply shops, this section was more religiously diverse than might be expected. The few blocks that were left of what were once, respectively, Little Poland and Little Ukraine lay slightly to the west of Loisada and due north of the Twin Towers, and were home to ornate Orthodox churches and even some shops catering to neo-pagan practitioners. Little India, which should rightfully have been called Little Bangladesh because its restaurateurs actually came from Muslim country of Bangladesh, was caddy-corner from Little Ukraine. And so on and so forth.

South Street Seaport, an old-time fish market that had been transformed into a mall-like tourist attraction, was an after-work refuge, replete with busy bars, shops, and eateries. Tribeca was a newly minted, mostly loft-living community on the lower west side of Manhattan; pricey, arty Soho lay a little to the east of Tribeca and the Towers. Battery Park City, with its combination of luxury condos and subsidized housing, was also a stone's throw away and was populated by young families in need of affordable housing and by financial firms that needed crash apartments for out-of-town deal makers and overnights. The West Village was not too far away either, with its authors, artists, and actors and its ever-active gay presence, which survived and thrived in spite of the devastation of the AIDS epidemic. The East Village, which was not quite as close

to the WTC as the West Village, was home to students, musicians, aspirants of all sorts, and also addicts, to judge by the reports of drug-related arrests. Several colleges and universities dotted the lower Manhattan landscape as well, including New York University.

Paradoxes about Post-9/11 Stress, Suicide, and Religion

There was a reason why we have delved into the details of these different downtown neighborhoods and business districts. We need to understand who lived downtown, why they were affected by 9/11, and why religion impacted some more than others. Studies have shown that the closer a person lived to Ground Zero, where the WTC once stood, the greater his or her chance of developing PTSD. That finding is not so surprising, considering how devastated the landscape was and how disturbing it was to view the carnage on a day-to-day basis and to smell smoldering flesh for weeks or maybe months later. What was surprising was that some studies showed that people with strong religious beliefs were more likely, not less likely, to experience stress post-9/11, even though religion ordinarily provides protection against stress. There were reasons why religion produced stress in this situation.

Other surprises awaited researchers after 9/11. Significantly, it turned out that people who received emergency debriefing were more likely, not less likely, to suffer long-lasting psychological symptoms of PTSD. It is now believed that rehearsing the traumatic event consolidated bad memories and made them more permanent than they might have been had the memories simply been allowed to fade away over time. An even bigger surprise concerned the prevalence of PTSD overall; post-9/11 PTSD was not nearly as widespread as first predicted.

Perhaps the biggest surprise overall concerned the role that religion played as a promoter of, rather than a protector against, suicide. Muslims, who ordinarily oppose suicide, might have even more reason for such surprise. Before we examine any specific associations between religion and suicide, let us look at some general associations between stress and religion.

Religion is often sought out in times of stress. It typically acts as a psychological salve. Repeatedly, religion has been shown to diminish the chances of completed suicide, which is the ultimate measurement of stress's effects. New religions are also known to thrive during times of social stress; recruitment to cults is especially successful during times of rapid social change. Moreover, religious participation generally decreases both criminal activity and criminal

recidivism (although high-profile exceptions to this trend have admittedly come to light).

Yet the events during and after 9/11 contradicted some of these basic assumptions (and I deliberately used the word assumptions because some of these givens about religion are not necessarily historically factual). For one thing, the mere fact that the suicide bombers committed both suicide – and homicide! – in the name of religion seems to stand established wisdom on its head. After all, study after study shows that religious faith and affiliation protect people from suicide and help people through severe depression, financial devastation, even grave and disabling illness. Religion as a generic concept is supposed to instill a reverence for life, not a disregard for life (although some Eastern religions devalue daily life to the point that they dismiss everyday existence as *maya*, or illusion, and some fundamentalist religious sects place more value on the heavenly afterlife rather than on the earthly this-life).

Yet here was a situation in which the strong religious beliefs of the suicide bombers not only did not protect them from suicide but actually pushed them to suicide. Moreover, their religious belief promoted not only suicide but also homicide – and promoted not just homicide but mass homicide. On the surface, the bombers' behavior appears paradoxical. No wonder some studies showed that people with strong religious beliefs felt more stressed after 9/11 than those with lesser religious commitment. Their confidence in the protective effect of religion was challenged. Any preexisting assumptions about the automatic ability of religion to counteract impulses to hurt oneself or others could no longer be accepted as absolute fact.

If we hypothesize that at least some of the people with strong religious adherence had gravitated to religion because they needed an extra buffer against stress in the first place, then it becomes even easier to appreciate why many felt more stressed after the 9/11 suicide attacks. If their psychological foundations were already shaky, but were buttressed by religious belief that now wobbled, then they would feel shakier still. They might need to distance themselves from the perpetrator religion (Islam), to prove that their own religion still provided them protection. It would be to their psychological advantage to deride Islam at face value, to draw a distinction between their own religion and Islam, and even to wage war against Muslims in order to protect their own psychological turf and to fortify their own mental foundation.

To understand the motives of these Muslim suicide bombers, we must realize that they acted on the earnest belief that they would earn themselves and their

families eternal paradise for their valiant acts of self-sacrifice. That reward system stands in direct contradiction to the hellfire and brimstone that presumably awaits devout Christians who choose suicide. The social prestige provided by those Islamic suicides is also the polar opposite of the social ostracism faced by Jewish families whose relatives elect to suicide. Because Jews who commit suicide cannot be buried in a Jewish cemetery, their survivors are deprived of the comfort of a collective mourning ritual, and the individual is deprived of an identity and a commemoration after death.

However, it should be said that 9/11 was by no means the first time that suicide was committed in the name of religion. Nor was Islam the first religion to encourage suicide in the name of faith. Christ's willing self-sacrifice on the cross could arguably be said to be a sort of self-accepted suicide. It is less controversial to claim that the desert saints who starved to death in the third and fourth centuries were also committing suiciding in the name of faith. Many Christian martyrs have been beatified because of their willingness to endure torture unto death in the name of religion. The Hindu *sati*, a cultural *cum* religious rite that requires the widow to climb on to the funeral pyre of her deceased husband and to join him in cremation and reincarnation, was a once-respected (but now suspect) religious rite of self-sacrifice. Jews who chose suicide over surrender at mountaintop fortress of Masada or who refused to convert to Christianity when the Spaniards offered them a choice between the cross or the Inquisition could be said to have chosen suicide, albeit under coercion. This was quite different from the Muslim suicide volunteers who went to their deaths willingly and equally willingly agreed to take the lives of many others.

Suicide through self-starvation is an acceptable, even respectable, end for some ascetic sects of Jains in India. In the Middle Ages, the Cathars practiced self-starvation as a religious rite. And let us not forget the soldiers who volunteered to serve in wars based on religion knowing that they might not survive the battle but believing in the utility of their self-sacrifice.

9/11 challenged the standard associations between stress and religion in other important ways as well. As previously mentioned, recruitment to new religions or cults typically peaks at times of social stress, as does religious involvement in general. But there is no evidence that such a trend took place in the wake of 9/11 because of very specific interventions. If anything, recruitment to psychiatric treatment peaked instead, as prevention efforts to avert the predicted outbreak of PTSD diverted interest away from religion and into more mainstream therapeutic approaches. Public

health and social service organizations made massive outreach efforts in the downtown area. Emergency debriefing stations were set up in lobbies in lower Manhattan, where residents were informed about both public and private psychiatric services. The Red Cross subsidized psychiatric treatment, even for those who could otherwise afford it. Posters recommending mental health care appeared in subway cars and at bus stops and reminded commuters that mental health care was not only easily available but was also readily affordable. As it turned out, these seat-of-the-pants psychiatric treatments did not accomplish what they purported and now seem to have increased the stress symptoms and PTSD. But they may have prevented religious-cult recruitment in an otherwise susceptible population.

For anyone who came of age at a time when public service announcements reminded Americans that the "family that prays together stays together" and when religion was the first referral made to someone suffering from stress or distress, seeing these subway posters was startling. These posters signaled a significant shift in social values as well as in clinical care. We might say that making a first-line recommendation for psychiatric care, rather than for religious rituals, was a sure sign that psychiatric treatment had finally come of age and was de-stigmatized enough to appear in public places outside of student counseling services.

Another explanation for the strong emphasis on mental health and PTSD in the immediate aftermath of 9/11 may have reflected the professional make-up of New York City. New York City is home to an unusually high number of mental health and social service professionals whose services are available at a moment's notice. Plus, psychotropic medications had become increasingly available and acceptable during the decade that preceded 9/11. The public had heard about the benefits of Prozac, and about its hazards as well, and had voted for the former. Unfortunately, it took approximately 2 years for researchers to realize that the mental health-mediated debriefings meant to prevent PTSD had actually increased its occurrence. In contrast, those who were treated with medications that decreased physiological reactions fared better.

We might hypothesize that astute politicians and public health officials had an even greater goal when they choose to promote mental health treatment rather than religion to assuage stress. City bureaucrats, who are acutely aware of the intense role that religion plays in politics at any given time, may have realized that an excessive emphasis on religion could polarize people even more at this crucial time. Such

polarization could create serious social unrest, catalyze riots, trigger interreligious rifts, and provoke other unpredictable reactions. Perhaps someone realized that it was socially safer, if not clinically wiser, to shift the focus to universal psychiatric approaches and to deflect attention from particularistic religious relief.

Even though religion was not recommended as a first-line antidote for post-9/11 stress, we would be hard-pressed to deny the daunting impact of religion-related stress. Admittedly, we would be equally hard-pressed to find complete and accurate data on this topic at this early date, and so we must make do with speculation and hypotheses, seeded by anecdotal reports from the *New York Times* and garnered from personal clinical experiences.

Muslims and Mental Health after 9/11

Most obviously, we expect that Muslims would be affected most dramatically by the events of 9/11, not just because 9/11 was perpetrated by Muslims who acted against other Americans in the name of their faith but also because of the repercussions against Muslims (and against those who were mistaken for Muslims). At this point, it is important to point out that many different people from many different countries and from many different ethnic groups embrace Islam around the world and that the choices made by the suicide bombers were the choices made by but a few members of a multimillion-member religion. Yet all American-based Muslims were forced to face scrutiny by the American public. Many were subjected to police searches, involuntary incarceration, or automatic deportations, all of which caused stress.

Once the United States declared war on Iraq, American Muslims faced further stress, regardless of their personal political persuasions or religion affiliations. There were some similarities to the indignities endured by Japanese Americans during World War II, when their allegiance to the United States was suspect after Japan bombed Pearl Harbor, to the point that even U.S. citizens were moved out of their homes and resettled in closely guarded camps. Because Muslim Americans are sometimes (but certainly not always) identifiable through physical appearance or religious garb or Mideastern surnames, almost any Muslim can become an easy target of suspicion, discrimination, or outright abuse. Guilt through association can cause an assumption of guilt.

People who are in the United States on a student visa or a J-1 visa (as many physicians in training are)

also carry legal documents that identify their country of origin. It would not be surprising if stress-related disorders increased among Muslims after 9/11, but it is not known with certainty if this actually occurred. The U.S. medical literature is strikingly devoid of studies on this subject at the time of this writing. In fact, it is doubtful that the full extent of the medical or mental health effects on Muslims will ever be known with certainty. It is likely that many Muslims, especially those who are working or studying in the United States on revokable visas, will avoid treatment altogether for fear of calling attention to themselves and risking being reported to the authorities or of appearing to be less qualified to compete in the work or educational arenas that influence their visas. It is probable that Muslims whose immigration papers are not complete or perfect will follow the paths taken by other illegal immigrants in the United States and will also be left uncounted and untreated by the official U.S. health-care system.

Furthermore, it is not clear whether psychiatric symptoms alone, such as stress, anxiety, or insomnia, could accurately measure the distress experienced by this very diverse religious group, whose members come from a wide variety of cultures, educational backgrounds, and countries of origin, and who speak many different languages. Some cultural groups do not express distress in psychological terms but prefer physical descriptions instead, sometimes complaining of dizziness, sleep problems, or vague aches and pains. Even people who are aware of psychiatric symptoms per se are still not immune to medical illnesses that increase in the presence of psychological or social stress. Cardiovascular and cerebrovascular events are notorious for their responsiveness to stress, as are some gastrointestinal diseases, autoimmune disorders, and even neurological symptoms.

In many instances, increases in drug and alcohol use are more accurate and immediate indicators of stress than are purely psychological symptoms and must be studied with as much diligence as medical or psychiatric symptoms. Alcohol consumption in New York City skyrocketed after 9/11, especially among women, as both medical statistics and bar revenues attested. But it is important to be aware of the baseline alcohol and drug consumption of any particular culture before determining if there has been a significant increase in use. For instance, Islam (apart from renegade Sufi mystics) generally prohibits alcohol. Therefore, even a modest increase in alcohol consumption in Muslims after 9/11 carries very serious clinical implications and cannot be meaningfully compared to increases that occurred in religious or ethnic groups that do not outlaw alcohol. On the

other hand, some Islamic cultures permit public marihuana use or even the social consumption of opium-based tea, although these activities are outlawed in the United States.

Although I cannot make claims to any population-based data on the use of psychiatric services by Muslims at this time, I can comment on recent trends in my own private practice in downtown Manhattan. Two distinctive gender-based trends emerged among Muslims.

First, there was a sudden surge in foreign-born Muslim males living in the United States on student visas who requested outpatient treatment for anxiety, insomnia, concentration problems, and vague perceptions of persecution by peers and co-workers. In some instances, patients needed an official return-to-work clearance, which they thought could be completed in a single visit. At other times, people who had longstanding behavior issues that could easily have been identified by college advisors much earlier were suddenly sent for psychiatric evaluations.

In each instance, these men were asked specifically if they thought that their new onset of psychiatric problems resulted from religious or cultural issues. If relevant, each was asked if he felt that the school or employer's request for psychiatric clearance was prompted by increased security concerns that were specific to Muslims. In each instance, the prospective patient denied any such association, but, in one case, the patient refused to fill out a form that inquired about native language, religion, and culture and made a point of stating that he did not consider religion to be relevant. When told that many patients, particularly college students, do not complete that section, he continued to insist that doctors do not need to know his religion.

Denial is not unusual in psychiatric practice, and a patient's repeated denial of the importance of an otherwise important issue can lead to a fertile investigation of psychological conflicts. Such denial usually takes place within the psychiatric interview itself and is addressed, often quite productively, during the therapy session. What was unusual in this situation was the prospective patient's repeated refusal to fill out standard medical history forms prior to the appointment.

Occasionally, the private outpatient psychiatrist encounters a paranoid patient who refuses to complete paperwork at the time of the first interview, but, more commonly, someone who is that paranoid avoids an appointment altogether and is brought to an institutional setting involuntarily. Moreover, pure paranoid patients typically show other observable signs of paranoia; their eyes scan the room and they sit near the door, if they sit down at all. They may refuse to enter

enclosed spaces in the first place (which comes as welcome relief to the psychiatrist, who is equally ill-at-ease in being with a potentially volatile paranoid person).

Yet, in the situations described here, the Muslim men did not exhibit other observable signs of paranoia. Their previous social functioning prior to the onset of their psychiatric symptoms was definitely not suggestive of serious preexisting psychiatric illness. In each instance, these men were well-educated and performed well in the past, making it unlikely that their circumscribed suspicions about filling out forms was part and parcel of more global paranoia and appeared to be based on legitimate fears about incarceration or deportation or the loss of his job or student status or visas required for school or work, even though each denied this possibility. This paranoic behavior can be compared to the guardedness and evasiveness seen in some African American men who have had unpleasant experiences with legal authorities in the past.

Second, I noted an even greater increase in the number of Muslim American women seeking psychiatric treatment, albeit for very different reasons. These women typically sought treatment because of stress brought on by family conflicts that arose as a result of their defiance of the gender-role restrictions traditional in Muslim families. In some cases, romantic involvements with non-Muslim men had already caused serious family rifts and threatened the patient with the loss of both family and financial support. It is difficult, and also probably inaccurate, to attribute these situations solely to the aftershocks of 9/11. It is much more likely that Americanization in general was responsible for rattling the religious traditions of these new immigrants, just as Americanization and the increased social choices afforded by U.S. society have shaken the foundations of nearly every ethnic group that arrived on U.S. shores. Still, I wonder if the Muslim-mediated suicide bombings of 9/11 shook the faith of at least some Muslim women and gave them tacit permission to question the overall correctness of preexisting, previously unquestionable cultural and religious traditions. Alternatively, perhaps the widespread publicity about psychiatric treatment that came in the wake of 9/11 also made these women aware of the availability of secular psychiatric treatment for personal stress.

Jewish Issues Associated with 9/11 Stress

Last but not least, let me turn to specific stresses experienced by some Jews after 9/11. Although there are no studies to date that show that New York City Jews were

at increased risk for post-9/11 PTSD, there are reasons why Jews might experience different stresses than other religious groups. For one thing, some Muslim extremists said that they had bombed the United States specifically because of its support for Zionists in Israel. Moreover, some semiofficial spokespeople stated that they had targeted New York City in particular not just because of its centrality to the financial and communications industries but because of its disproportionately high Jewish population. Such extremists link all Jews to extremist Zionists, regardless of their personal political persuasions.

Jews might feel themselves targeted by the Muslim extremists, but they might also have residual PTSD because of personal or cultural connections to the Holocaust, when 6 million Jews perished at the hands of the Nazis. Those who smelled the burning bodies in the WTC – as did anyone who lived downtown, often for months afterward – typically compared the scene to the crematoria where Jewish bodies were burned en masse after being gassed in places such as Auschwitz. Such scents were described in detail in accounts by Holocaust survivors. Ironically, 2000 people perished in a single day at WTC on 9/11. That very same number – 2000 – was gassed every single day at Auschwitz when the ovens were at their peak performance. This strange synchronicity did not go unnoticed by those who knew the details of the Holocaust, and it served only to highlight the horrors of the Holocaust and to reopen old wounds that refused to heal.

Some Jews were pretraumatized because friends or relatives had been hurt or killed during recurrent terrorist attacks in Israel. Many felt personally threatened when they saw that the same terror could take place on supposedly safe U.S. soil. Among Jews who were familiar with Jewish history, 9/11 was compared to Kristallnacht, the day that marked the beginning of the Holocaust, when Jewish shops were shattered by Nazi storm troopers in a foreshadowing of the horror that was to come a few years later. The fact that this was the second, not the first, attack on the WTC was used to fortify the comparisons with Kristallnacht.

It is curious that a disproportionate number of Jews lived in the expansive Grand Street Co-ops in lower Manhattan. It is also curious that New York University, a university often identified as having a high proportion of Jewish students, dots downtown Manhattan. With its sprawling campuses and dormitories, New York University spans the area from the Seaport to Fifteenth Street, with some dorm rooms overlooking Ground Zero. It is possible that the presence of these ethnic and religious populations skewed statistics and influenced the high rate of PTSD among

those who lived near Ground Zero. A more detailed demographic study of post-9/11 stress syndromes could clarify these possibilities.

Concerns of Christians and Others after 9/11

Is worth noting that members of some religions and ethnic groups became Muslims by proxy and were subjected to tauntings and even beatings because they were mistaken for Muslims. Turban-wearing Sikhs and Jains from the Indian subcontinent were targeted by aggressive youths in the days after the bombings and some were seriously injured. It is conceivable that people who practice many other religions but who appear to be of Mideastern or Muslim origin realized that they were at risk and so have suffered increased stress since 9/11.

It is likely that even all-American Christians also suffered from religion-related stress after 9/11, simply because most Americans like to think of themselves as religiously tolerant and become uneasy when religious prejudices are provoked, as they were by the events of 9/11. Let us not forget that the United States was founded for the expressed purpose of freedom of religion. When religion-inspired acts such as the terrorist bombings and hijackings of 9/11 threaten to make folly of this basic American freedom, everyone, regardless of his or her religion, feels stress.

In summary, Although there is much that remains uncertain about the overall impact of 9/11, both with respect to religion and stress, and to its many other ramifications, some things are certain. For one thing, 9/11 brought renewed attention to religion and its role in both promoting and protecting against stress. Unlike the Oklahoma bombing, which was strictly secular, this mass murder was intimately linked to religion and always will be, regardless of what is revealed about the individual bombers' personal psychologies in the future. This sad event serves as a reminder that the age-old association between stress and religion cannot be put on the shelf and relegated to the past; it is very much a part of the present. The subject of stress and religion is as relevant to the twenty-first century as it was to the first century CE.

Further Reading

- Dervic, K., Oquendo, M. A. and Grunebaum, M. F. (2004). Religious affiliation and suicide attempt. *American Journal of Psychiatry* 161, 2303–2308.
- Durkheim, E. (1951). *Suicide*. Spaulding, J. A. & Simpson, G. (trans.). New York: Free Press.

- Falsetti, S. A., Resnick, P. A. and Davis, J. L. (2003). Changes in religious belief following trauma. *Journal of Trauma and Stress* 16, 391–398.
- Galea, S., Vlahov, D. and Resnick, H. (2003). Trends of probably post-traumatic stress disorder in New York City after the September 11 terrorist attacks. *American Journal of Epidemiology* 158, 514–524.
- Koenig, H. G., McCullough, M. E. and Larson, D. B. (2001). *Handbook of religion and health*. New York: Oxford University Press.
- Schuster, M. A., Stein, B. D. and Jaycox, L. (2001). A national survey of stress reactions after the September 11, 2001 terrorist attacks. *New England Journal of Medicine* 345, 1507–1512.
- Shalev, A. Y. (2004). Further lessons from 9/11: does stress equal trauma? *Psychiatry* 67, 174–177.
- Wolinsky, F. D., Wyrich, K. W. and Kroenke, K. (2003). 9–11, personal stress, mental health, and sense of control among older adults. *Journal of Gerontology, Psychology, and Social Sciences* 58, S146–150.

Remodelling of Neuronal Networks by Stress

E Fuchs

German Primate Center and Medical School, University of Göttingen, Göttingen, Germany

G Flügge

German Primate Center, Göttingen, Germany

© 2007 Elsevier Inc. All rights reserved.

Early Ideas about the Stress Response

Stress Circuits

Activation of Brain Systems

Structural Plasticity of the Adult Brain

Remodeling of Brain Cells by Stress

Stress Suppresses Neurogenesis in the Adult Dentate Gyrus

Plasticity of Astrocytes

Plastic Changes of Brain Cells Are Reflected by Alterations in Gene Expression

Structural Changes Are Reversible

Conclusions

Glossary

BrdU

(5-Bromo-2'-deoxyuridine)

A thymidine analog that is incorporated into the DNA of a dividing progenitor cell during the S phase. Fully differentiated neurons do not divide and cannot integrate the label. Depending on the survival time after application, BrdU is a marker of proliferating cells (short survival time) and their progeny (longer survival time). BrdU labeling is a nonisotopic immunocytochemical method.

Early Ideas about the Stress Response

When Selye formulated his stress theory approximately 70 years ago, stress was thought to have a

merely endocrine character, and noxious stimuli of physical or chemical nature were the primary stressors discussed. However, subsequent research demonstrated that stress-induced activation of the endocrine system leads to increased release of glucocorticoids (cortisol in primates including humans, corticosterone in rodents) from the adrenal cortex and that these peripheral stress hormones also profoundly influence brain function. Moreover, stress also increases the activity of the sympathetic nervous system to enhance release of norepinephrine from peripheral nerve endings. In addition, stress modulates the activities of, for example, central nervous monoaminergic circuits. These have diverse effects on many brain cells. Great progress has been made in the understanding of the biology of stress reactions and the neurobiological alterations that occur as a consequence of exposure to challenging events in the environment.

Stress Circuits

One of the most intensively investigated stress circuits is the limbic-hypothalamic-pituitary-adrenocortical (LHPA) system, which determines the magnitude and duration of hormonal, neural, and behavioral reactions to stressful experiences. The LHPA system combines both brain and endocrine components and can be regarded as a classic neuroendocrine circuit that integrates cognitive, emotional, neuroendocrine, and autonomic inputs. It consists of the primary brain regions that process emotional information, the limbic system and the hypothalamus, as well as pituitary and adrenal glands as endocrine organs. In concert with other neurohormonal actions such as stimulation of the adrenomedullary system (release of adrenaline from the adrenal medulla), LHPA activation leads to an increased release of corticosteroid hormones from the adrenal cortex into the bloodstream. These

steroid hormones are potent modulators of cell physiology, mood, and behavior and serve both to alert the organism to an environmental or physiological challenge and to defend its homeostasis.

Within the LHPA system, the hippocampus plays a central role. This brain structure has a high density of corticosteroid receptors and modulates glucocorticoid release from the adrenal cortex via a negative feedback loop exerting inhibitory effects on the LHPA axis. In addition to its intimate involvement in neuroendocrine regulation, the hippocampus plays an important role in emotional processing and in key aspects of learning and memory. While short-term activation of the LHPA system is essential for vital functions, constant hyperactivity of the circuit results in chronically elevated glucocorticoids, a maladaptive and potentially neuropathological condition. As demonstrated in animal studies, chronic stress leads to an impairment of the negative feedback control mechanism of corticosteroid release from the adrenal cortex, partly because of a reduced expression of hippocampal corticoid receptors.

Activation of Brain Systems

Stress does not have a global effect on all brain areas. Rather, depending on the type of stressor, different neuronal circuits are activated. Limbic brain regions comprising, for example, the hippocampus and amygdala as well as the prefrontal cortex are sensitive to stressors such as restraint or anxiogenic external stimuli, for example, a novel environment. Common to these stressors is the fact that they require, prior to initiation or inhibition of the stress response, a central nervous processing of information coming from different sensory organs. This information processing involves limbic and cortical regions, and its outcome depends on previous experience. In contrast, physiological threats such as exposure to ether result in a direct activation of efferent visceral pathways, in part mediated by the paraventricular nucleus of the hypothalamus. In this case, the rapid activation of brain stem and hypothalamus circumvents cognitive-emotional processing via the higher brain regions.

The intraneuronal communication that finally leads to the stress reaction is brought about by neurotransmitter systems. Rapid neuronal responses to stress are elicited by, for example, glutamate (an excitatory neurotransmitter) and GABA (γ -aminobutyric acid, an inhibitory neurotransmitter). These modulators act on specific receptors and ion channels in neuronal membranes to directly control neuronal firing. This fast neurotransmission is modulated by slow-acting neurotransmitters such as the monoamines (for example, the catecholamines dopamine, norepinephrine,

and epinephrine and the indolamine serotonin). Activation of brain monoamine systems is a major component of the stress response, and turnover of monoamines is increased during stressful experiences. Stress-induced alterations in the monoamine system are regarded as the basis for stress-related behavior and are of special interest because the hyperactivity of catecholaminergic and serotonergic neurons that occurs during stress may induce an imbalance between the different neurotransmitter systems. This imbalance may contribute to psychopathologies such as major depression.

Another example of substances that modulate neuronal activity and are involved in mediation of stress responses are neuropeptides, for example corticotropin-releasing hormone (CRH). This is thought to play a major role in the stress response. Neuropeptides as well as monoamines exert their actions via specific receptors located in the membranes of neurons. Some of these receptors are also found in glia cells. Many of these receptors belong to the family of G-protein-coupled receptors (GPCRs) that are linked to intracellular second messenger systems via guanosine 5'-triphosphate (GTP)-binding proteins. The intracellular signal transduction pathways that are regulated by GPCRs bring about many of the short- and long-term effects of stress on neuronal activity and regulate most aspects of neuronal functioning, including metabolism and gene expression. Thus, intracellular signal transduction systems play a key role with respect to stress-mediated regulation of neuronal functioning and structure, and, on the cellular level, they represent the first step of a cascade of events in the brain that leads to changes in neural cells, including alterations in gene transcription and morphology of neurons. Adverse effects of stress are supposed to result partly from (1) sustained activation or inhibition of certain neurotransmitter systems and intracellular pathways and/or (2) chronic exposure to neuroactive substances such as the glucocorticoids. These steroid hormones interact with intracellular receptors, forming complexes that bind to specific sites on DNA to regulate transcription.

Structural Plasticity of the Adult Brain

Approximately a century ago, it was finally established that the brain consists of neurons and glia (neural cells) rather than being a continuous syncytium. Thereafter, a rather static view of the brain prevailed. It was thought that electrical and chemical information was processed by fixed neuronal circuits. Ramon y Cajal stated that "in adult centers the nerve paths are something fixed, ended, immutable. Everything may die, nothing may be regenerated" (Cajal

1928, p. 750). However, in recent years, this view has been gradually revised on the basis of studies demonstrating that neuronal circuits and connections between cells in the brain are subject to lifelong modifications and reorganizations. An impressive example of the consequences of alterations in sensory inputs and changes in circulating hormones comes from research using ground squirrels. In these hibernating animals, synaptic contacts of mossy fibers with CA3 hippocampal pyramidal neurons are altered in many aspects of their structure during different stages of the torpor–activity cycle. This clearly demonstrates the capacity of the adult brain to adjust its neuronal circuits to altered hormonal and/or sensory input. Based on such findings, it was hypothesized that similar structural modifications may also occur during stressful life events when hormonal levels are altered (for example, elevated glucocorticoids, reduced gonadal steroids) and when sensory input is changed. It has transpired that this view is accurate.

This ability of the brain to perform functionally relevant adaptations following various challenges is called plasticity. Neuronal plasticity is absolutely mandatory for adequate functioning of an individual in a continuously changing environment. The dynamic processes of neuroplasticity are based on the capability of neurons to adapt to alterations in the internal and/or external environment by modifying specific structures and function; they do this by changing their entire cell surfaces, including the synapses, so that the locations of receptors and other molecules imbedded in the plasma membrane are reorganized. Adaptive or experience-dependent plasticity reveals itself in many forms, ranging from changes in gene expression to changes in neurotransmitter release and in behavior. Even under normal or undisturbed conditions, contacts between neurons are continuously replaced and renewed in the adult and differentiated brain. Enhanced axonal outgrowth and collateral sprouting on the presynaptic site may lead to the formation of new synapses, and existing contacts between neurons may be eliminated by terminal retrograde degeneration. The number of postsynaptic sites on a neuron can be increased or decreased by alterations in the size of its dendritic tree or the spine density on the dendrites.

Remodeling of Brain Cells by Stress

Probably the most thoroughly investigated stress-induced change in neuronal morphology is the regression of apical dendrites of pyramidal neurons, which was first demonstrated in the hippocampus. After this finding, dendritic remodeling of CA3 pyramidal neurons has been repeatedly documented after chronic stress as well as after corticosterone administration.

Similar shortening of dendritic branches, although to a smaller extent, has been observed in granule cells of the dentate gyrus and in pyramidal cells of the hippocampal region CA1, both in chronically stressed and in corticosterone-treated rats. It is assumed that, as a result of the reduced surface area of the neurons, synaptic contacts are also diminished. Indeed, a significant loss of synapses on CA3 pyramidal cells, as well as profound changes in the morphology of the afferent mossy fibers that terminate on these neurons, was detected in chronically stressed or corticosterone-treated animals. Interestingly, even a brief social defeat stress with a long time delay thereafter can significantly reduce the length of apical dendrites but increase the length and complexity of the primary dendrites at the basis of the pyramidal neurons. These data suggest that a brief social conflict is sufficient to drive a dynamic reorganization of neuronal networks with site-selective elimination as well as *de novo* growth of dendritic branches, changes that persist for several weeks after the acute stress exposure.

Contrary to conclusions from initial studies proposing an impact of chronic stress on the number of pyramidal neurons in the hippocampus, recent stereological investigations failed to detect a statistically significant loss of these neurons following prolonged hypercortisolism resulting from stress exposure, corticosteroid administration, or aging. However, it should be noted that the number of hippocampal interneurons or of neurons in the hilus were not specifically examined in these studies. Recently, evidence was provided that long-term psychosocial stress reduces the number of parvalbumin-immunoreactive neurons that are regarded as GABAergic interneurons. This stress effect was observed in the dentate gyrus and in hippocampal region CA3, whereas subfield CA1 was not affected. Thus, the possibility that long-term stress may induce loss of interneurons without affecting the number of principal cells in the hippocampus cannot be ruled out.

Chronic stress alters dendritic morphology not only in the hippocampus but also in the medial prefrontal cortex (mPFC), which modulates various higher cognitive and emotional functions and contributes to regulation of LHPA system activity. Recent reports demonstrated that 3 weeks of either daily restraint stress or daily corticosterone injections led to changes in pyramidal neurons of the mPFC. Up to 20–35% retraction of the distal dendritic branches was seen, together with a significant (16%) decrease in apical dendritic spine density. Interestingly, even repeated vehicle injections resulted in similar, although less pronounced, changes, indicating that with respect to morphology of neurons, the mPFC is even more sensitive to external influences than the hippocampus.

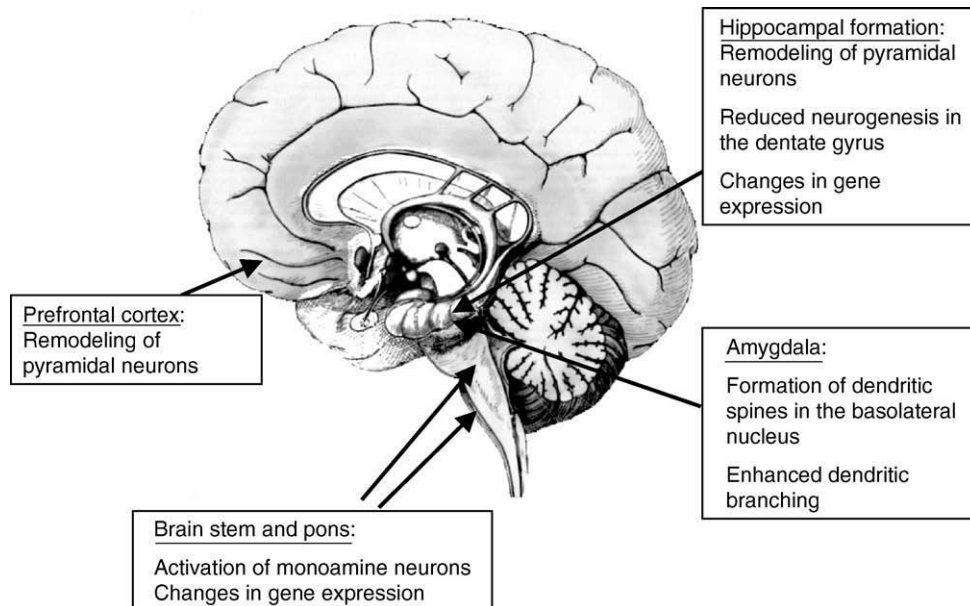


Figure 1 Examples of stress-induced changes in the brain. Hippocampal formation: remodeling of pyramidal neurons characterized by shrinkage of dendrites of pyramidal neurons in region CA3; changes in mRNA and protein reflect that stress regulates expression of genes involved in neuronal differentiation. Prefrontal cortex: remodeling of pyramidal neurons. Amygdala: in the basolateral nucleus, formation of new spines on the dendrites of spiny neurons. Brain stem and pons: activation of monoamine neurons, e.g., those that produce norepinephrine; changes in gene expression.

The amygdala plays a crucial role in regulating fear and anxiety and is essential for the formation of affective memories. Increasing evidence supports a critical role of this brain region in activating the LHPA system. A recent study investigated the impact of chronic stress on the dendritic pattern of amygdaloid neurons. In contrast to findings in the hippocampus, spiny neurons in the basolateral complex of the amygdala exhibited enhanced dendritic branching in response to chronic stress. Examples of stress-induced changes in the brain are summarized in [Figure 1](#).

Stress Suppresses Neurogenesis in the Adult Dentate Gyrus

For a long time, a central hypothesis in neuroscience was that in the mammalian brain, the production of neurons occurs only during development and stops before puberty, and new neurons cannot be formed in the adult brain. This widely held belief has been challenged in recent years by extensive evidence from many mammalian species, including nonhuman primates as well as humans, showing that even in the adult brain, certain areas retain the capability to generate new neurons. Adult spontaneous neurogenesis takes place only in certain regions such as the subgranular zone of the dentate gyrus in the hippocampal formation and the subventricular zone at the lateral ventricle. In the dentate gyrus, newly generated cells

become incorporated into the granule cell layer and attain the morphological and biochemical characteristics of granule neurons. The neuronal nature of these cells is demonstrated by the formation of synapses on the cell bodies and dendrites, extension of axons into the CA3 region, and generation of action potentials.

These observations on neurogenesis were facilitated by the advent of a novel method for detecting cell proliferation and migration. Using the thymidine analog 5-bromo-2'-deoxyuridine (BrdU), which is incorporated into the DNA of dividing cells during the S phase of mitosis, proliferating cells and their progeny can be labeled. BrdU can be visualized by immunocytochemical techniques, and when used in combination with other specific cell markers, it allows determination of the phenotype of the newly generated cells. Using such methods, it has been demonstrated that the majority of newborn cells in the dentate gyrus differentiate into neurons. In other words, it has been shown that the BrdU-labeled cells express the neuronal marker neuron-specific nuclear protein. A smaller proportion of the newborn cells express markers for glial cells such as glial fibrillary acidic protein. Although the new neurons comprise a minuscule proportion of the total neuronal population, their continuous addition over the entire life span implies considerable structural changes within the neural network. The magnitude and ubiquity of adult neurogenesis across vertebrates suggest that it is

functionally significant and not merely a vestige of development.

Stress is a very potent suppressive factor of adult neurogenesis, although a large number of environmental and endogenous parameters also have an impact on the formation of new neurons in the adult brain. The suppression of cell turnover by stress predicts that the age of the cell population, the connectivity of the neurons, and the resulting properties of neuronal circuits in a stressed individual might be substantially different from control situations, with potentially important functional consequences. Several different types of stressful experiences, such as exposure to the odor of a predator, social subordination, or resident–intruder stress, have been shown to inhibit neurogenesis in the dentate gyrus in several species, including nonhuman primates. Furthermore, it appears that the susceptibility of neurogenesis to stress is age dependent in that effects of stress on cell proliferation in the dentate gyrus are more pronounced in old animals than in young ones.

It has become increasingly evident that the antecedents of many illnesses begin in fetal life and, furthermore, that prenatal conditions can have a significant impact on either health or disease in the postpartum period. In a recent study that investigated whether prenatal stress can alter neural, hormonal, and behavioral status in nonhuman primates, pregnant rhesus monkeys were acutely stressed on a daily basis for a quarter of their 24-week gestation using an acoustic startle protocol. At 2 to 3 years of age, hippocampal volume, cytogenesis in the dentate gyrus, and cortisol levels were evaluated in the offspring from stress and control pregnancies. Prenatal stress reduced hippocampal volume and inhibited neurogenesis in the dentate gyrus. These changes were associated with higher cortisol levels in the blood of the monkey offspring, lower levels of exploration, and higher levels of motor behavior. These findings indicate that also in primates, a prenatal environment can alter behavior, may deregulate neuroendocrine systems, and affects the hippocampal structure in a persistent manner.

Plasticity of Astrocytes

The most abundant type of cell within the central nervous system is the glia. In the adult mammalian brain, there are 10 to 50 times more glia cells than neurons. Astrocytes are the most prominent type of glia cells, accounting for about one-third of brain mass. Recent studies have revealed that, in addition to their housekeeping functions, astrocytes are dynamic regulators of synaptogenesis and synaptic strength. These cells control neuronal production, for example, in

the adult dentate gyrus. Thus, abnormalities in glial functioning are likely to contribute to the impairment of structural plasticity in the brain. In line with these ideas, it was recently demonstrated that long-term stress significantly decreased both the number and the volume of astroglial somata. In addition, whether antidepressant treatment with the selective serotonin reuptake inhibitor fluoxetine could afford protection from these stress effects was examined. Indeed, fluoxetine prevented a stress-induced decrease in the number of astrocytes but did not counteract shrinkage of the astroglial somata.

Morphological changes in astrocytes most probably have functional significance with respect to neuron–glia interaction, and because this interaction supports neuronal functioning, interneuronal communication will also be affected by the stress-induced changes. The reduced number or weakened activity of astrocytes could change levels of extracellular glutamate, thus leading to high concentrations of this excitatory neurotransmitter that may have excitotoxic effects. An upregulation of the glial glutamate transporter (GLT-1) in the hippocampus has been reported after chronic stress, and it has been suggested that this might be a compensatory mechanism to control for increased extracellular glutamate concentrations caused by stress. Interestingly, a treatment with the antidepressant tianeptine can block the stress-induced upregulation of GLT-1, and tianeptine also reverses stress-induced dendritic remodeling in CA3. Therefore, the modulation of glial GLT-1 expression in hippocampal region CA3 can be considered a regional neurochemical correlate of dendritic remodeling. These changes in the structure and molecular plasticity of astroglia in response to stress and antidepressant treatment support the notion that glial changes may contribute to the pathophysiology of stress-related disorders such as major depression as well as to the cellular actions of antidepressants.

Plastic Changes of Brain Cells Are Reflected by Alterations in Gene Expression

The stress-induced changes in neuronal structure imply that brain cells adjust their biosynthetic pathways to the new requirements. For example, concomitant with retraction of dendrites, there is probably a reduced need for membrane proteins. Related adaptational processes take place partly on the level of gene transcription. Because various methods, namely, cDNA microarrays, serial analysis of gene expression, and subtractive hybridization, are currently used to

identify genes that are differentially regulated by stress, such genes still appear rather heterogeneous. However, it is already clear that some observations from gene transcription studies are consistent with what has been found with other methods.

In hippocampal formation, expression of several genes known to be involved in neuronal differentiation was found to be downregulated by chronic social stress. Such genes included those encoding the membrane glycoprotein M6a, the CDC-like kinase 1, and a gene encoding a distinct subunit of certain G-proteins, GNAQ. All these genes are known to be involved in neurite outgrowth and neuronal differentiation, thus supporting the view that alterations in neuronal morphology and/or formation of neurons are primary effects of stress, at least in hippocampal formation. Furthermore, expression of neural cell adhesion molecule (NCAM) was found to be downregulated after chronic restraint stress in the hippocampus. NCAM is known to regulate neurite outgrowth and target recognition in the developing nervous system by mediating cell adhesion and regulating signal transduction.

Neuronal plasticity is accompanied by dynamic changes in elements of the cytoskeleton. Alpha-tubulin, which is the major component of microtubules, can be posttranslationally modified, and tyrosinated and acetylated alpha-tubulin are considered markers of dynamically changing and stable forms of microtubules, respectively. Restraint stress for 4 days decreased the expression of tyrosinated alpha-tubulin and increased expression of the acetylated form in the hippocampus. This is in line with the view that stress alters the morphology of pyramidal neurons.

Plastic changes are not restricted to forebrain areas. In a brain stem region, the rat dorsal raphe nucleus, which contains a large number of serotonergic neurons innervating the forebrain, long-term social stress increased the expression of genes involved in regulation of neurotransmitter release (synaptosomal-associated protein 25 and synaptic vesicle glycoprotein 2b). These data support the view that stress increases activity of neurons in the dorsal raphe nucleus.

Data from mRNA or protein expression studies sometimes appear contradictory. However, when interpreting such data, in which brain region, on which cellular level (neurons versus glia), and in which cellular compartment (neuronal cell body versus area of the dendrites versus axon terminal areas) the stress-induced changes were detected should be considered. Finally, besides the sex and the age of the investigated species, both the type and the duration of the stressor play a crucial role with respect to effects of stress on gene expression.

Structural Changes Are Reversible

As shown in animal studies, apical dendrites of hippocampal pyramidal neurons that were shortened by restraint stress or glucocorticoid exposure regained their normal length within 3 weeks after treatment. Several *in vivo* magnetic resonance imaging studies in humans suggest a correlation between hippocampal volume reduction and cumulative glucocorticoid exposure, although exceptions have also been reported. In Cushing's patients who suffer from hypercortisolism, smaller hippocampal volumes returned to normal after successful surgery that led to normal glucocorticoid levels. Thus, up to a certain point, structural changes may be at least partially reversible and may permit restoration of normal functioning. However, it appears that the appropriate intervention should take place within a certain time before the changes become irreversible.

The mechanisms that are responsible for hippocampal volume loss have not yet been identified. Massive neuronal loss following exposure to repeated episodes of hypercortisolemia can be excluded, as no major cell loss was apparent in experimental animals as well as in human postmortem brain tissue. Neurogenesis has been regarded as another mechanism that might affect the volume of the hippocampus. However, hippocampal neurogenesis adds relatively few neurons per day, and those only to the granule cell layer of the dentate gyrus, where an equivalent number of neurons also die. Alternately, hippocampal volume loss might be due to alterations in amounts of the dendritic, axonal, and synaptic components or to changes in levels of glia. Furthermore, stress may induce a shift in fluid balance between the ventricles and brain tissue. This assumption is supported by numerous clinical studies reporting lower volumes of different brain structures in conjunction with enlarged ventricles in patients with affective disorders. However, it remains to be determined whether changes in fluid content cause the aforementioned neuroplastic changes in the neuronal network or just accompany them.

Conclusions

The diverse forms of stress-induced changes in neural cells are of particular interest. They show that even the adult and differentiated brain is a plastic organ. This lifelong plasticity is a prerequisite for the brain to adapt to environmental changes. Plasticity of the neuronal networks is the basis for learning and memory. On the other hand, extensive neuroplastic changes induced by stress may also lead to imbalances in central nervous neural networks. A promising aspect with regard to therapy is that many of these

processes appear to be reversible. Reversibility of structural as well as functional plasticity has already been demonstrated, for example, in response to pharmacological treatments. Thus, progress in our understanding of neural plasticity has profound implications for the treatment of a number of neurodegenerative and psychiatric disorders.

See Also the Following Articles

Acute Trauma Response; Corticosteroid Receptors; Glucocorticoid Effects on Memory: the Positive and Negative; Glucocorticoids – Adverse Effects on the Nervous System; Memory and Stress; Maternal Deprivation; Peptides.

Further Reading

- Abumaria, N., Rygula, R., Havemann-Reinecke, U., et al. (2006). Identification of genes regulated by chronic social stress in the rat dorsal raphe nucleus. *Cellular and Molecular Neurobiology* 26, 145–162.
- Alfonso, J., Frasch, A. C. and Flügge, G. (2005). Chronic stress, depression and antidepressants: effects on gene transcription in the hippocampus. *Reviews in Neuroscience* 16, 43–56.
- Cajal, S. R. (1928). *Degeneration and regeneration of the nervous system*. London: Oxford University Press.
- Czeh, B., Simon, M., van der Hart, M. G., et al. (2005). Chronic stress decreases the number of parvalbumin-immunoreactive interneurons in the hippocampus: prevention by treatment with a substance P receptor (NK1) antagonist. *Neuropsychopharmacology* 30, 67–79.
- Czeh, B., Simon, M., Schmelting, B., Hiemke, C. and Fuchs, E. (2005). Astroglial plasticity in the hippocampus after chronic psychosocial stress and concomitant fluoxetine treatment. *Neuropsychopharmacology*, Dec. 14 [Epub ahead of print].
- de Kloet, E. R., Vreugdenhil, E., Oitzl, M. S. and Joëls, M. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Reviews* 19, 269–301.
- Flügge, G., van Kampen, M. and Mijster, M. J. (2004). Perturbations in brain monoamine systems during stress. *Cell and Tissue Research* 315, 1–14.
- Fuchs, E. and Czeh, B. (2005). Adult neurogenesis in rodents and primates: functional implications. In: Steckler, T., Kalin, N. H. & Reul, H. (eds.) *Handbook of stress and the*

- brain Part 1: The neurobiology of stress. Techniques in the behavioral and neural sciences* (Vol. 15), pp. 711–727. Amsterdam: Elsevier.
- Fuchs, E., Czeh, B., Kole, M. H. P., Michaelis, T. and Lucassen, P. J. (2004). Alterations of neuroplasticity in depression: the hippocampus and beyond. *European Neuropsychopharmacology* 14, 481–490.
- Herman, J. P. and Cullinan, W. E. (1997). Neurocircuitry of stress: central control of the hypothalamo-pituitary-adrenocortical axis. *Trends in Neuroscience* 20, 78–84.
- Kempermann, G. (2002). Regulation of adult hippocampal neurogenesis – implications for novel theories of major depression. *Bipolar Disorders* 4, 17–33.
- Mason, J. W. (1968). A review of psychoneuroendocrine research on the pituitary-adrenal cortical system. *Psychosomatic Medicine* 30, 576–609.
- McEwen, B. S. (1999). Stress and hippocampal plasticity. *Annual Review of Neuroscience* 20, 49–70.
- Mitra, R., Jadhav, S., Bruce, S., et al. (2005). Stress duration modulates the spatiotemporal patterns of spine formation in the basolateral amygdala. *Proceedings of the National Academy of Sciences USA* 102, 9371–9376.
- Popov, V. I. and Bocharova, L. S. (1992). Hibernation-induced structural changes in synaptic contacts between mossy fibres and hippocampal pyramidal neurons. *Neuroscience* 48, 53–62.
- Reul, J. M., Sutanto, W., van Eekelen, J. A., Rothuizen, J. and de Kloet, E. R. (1990). Central action of adrenal steroids during stress and adaptation. *Advances in Experimental Medicine and Biology* 274, 243–256.
- Sapolsky, R. M. (1999). Glucocorticoids, stress, and their adverse neurological effects: relevance to aging. *Experimental Gerontology* 34, 721–732.
- Selye, H. (1936). A syndrome produced by diverse noxious agents. *Nature London* 138, 32.
- Stanford, S. C. (1993). Monoamines in response and adaptation to stress. In: Stanford, S. C. & Salmon, P. (eds.) *Stress. From synapse to syndrome*, pp. 281–331. London: Academic Press.
- Watanabe, Y., Gould, E. and McEwen, B. S. (1992). Stress induces atrophy of apical dendrites of hippocampal CA3 pyramidal neurons. *Brain Research* 588, 341–345.
- Wellman, C. L. (2001). Dendritic reorganization in pyramidal neurons in medial prefrontal cortex after chronic corticosterone administration. *Journal of Neurobiology* 49, 245–253.

Renal and Adrenocortical Actions of Dopamine

B C Williams, Y-C Lo and S W Walker

University of Edinburgh, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 356–359, © 2000, Elsevier Inc.

Renal Actions of Dopamine

Adrenocortical Actions of Dopamine

Conclusions

Glossary

*Aromatic
L-amino acid
decarboxylase
(L-AAAD)
(EC 4.1.1.28)*

Previously known as hydroxytryptophan decarboxylase or dopa decarboxylase, L-AAAD catalyzes the decarboxylation of L-5-hydroxytryptophan and L-dopa (L-3,4-dihydroxyphenylalanine) to 5-hydroxytryptamine and dopamine, respectively. It requires pyridoxal phosphate as coenzyme and occurs most abundantly in the kidney, liver, gastrointestinal tract, adrenal gland, and brain.

*Dopamine
(3,4-
dihydroxy-
phenylethyl-
amine)*

A biogenic amine, first identified as an important neurotransmitter in the central nervous system and later discovered to exert actions in peripheral tissues, particularly the cardiovascular system, the renal system, and the adrenal gland.

*Dopamine
receptors*

Plasma membrane proteins that mediate the biological effects of dopamine through the activation of an intracellular second messenger system. Many different classes of dopamine receptors have been defined according to their pharmacological properties using selective agonists and antagonists and/or their molecular structure determined by molecular cloning techniques. Dopamine receptors in peripheral tissues are still distinguished from those in the central nervous system. Although they appear to be similar, they may not be identical, thus D₁ and D₂ denote dopamine type 1 and type 2 receptors, respectively, in the central nervous system, whereas DA₁ and DA₂ denote what are considered to be the equivalent (and perhaps related) receptors in peripheral tissues.

Natriuresis

The loss of sodium ions from circulating plasma into the urine through the action of endogenous chemicals or drugs on the

sodium transport process occurring in the proximal and distal segments of the renal tubules.

*Zona
glomerulosa*

A narrow layer of cells in the outer region of the adrenal gland, situated directly beneath the adrenal capsule, whose specific function is to secrete the steroid hormone aldosterone, which acts on mineralocorticoid receptors in the distal convoluted tubule of the kidney to cause sodium retention and potassium loss.

Renal Actions of Dopamine

Much of our current understanding of the renal actions of dopamine derives from the pioneering work of Leon Goldberg from 1964 onward, which stimulated subsequent research throughout the world.

Effects of Dopamine on the Renal Vasculature

Infusion of dopamine *in vivo* in several animal models leads to vasodilatation and to an increase in renal blood flow with a less marked rise in the glomerular filtration rate (GFR). The limited studies conducted in humans have confirmed these effects, and it appears that the major effect of dopamine is to increase selectively blood flow through the renal cortex rather than through the renal medulla, indicating that dopamine causes a redistribution of renal blood flow, which may reflect the relative distribution of dopamine receptors in the renal cortex versus the renal medulla vasculature. Dopamine receptors that mediate dopamine-induced renal arterial dilatation have been partially characterized using pharmacological agents. In the *ex vivo*-perfused rat kidney, the DA₂ receptor agonist bromocriptine causes vasodilatation, and this response is inhibited by metoclopramide, a DA receptor antagonist. In contrast, in renal microvessels from rabbits, the dopamine-induced relaxation of afferent and efferent glomerular arterioles could be mimicked by fenoldopam and SKF 87516, two specific DA₁ receptor antagonists. In humans, the infusion of fenoldopam causes renal vasodilatation, and in human renal artery segments, dopamine-induced relaxation can be blocked by (+)-sulpiride and SCH 23390, two DA₁ receptor antagonists. It therefore appears that both DA₁ and DA₂ receptors play a role in dopamine-induced vasodilatation and that their respective contribution may be species specific.

Effects of Dopamine on the Renal Tubule

In pharmacological doses, dopamine is a potent natriuretic agent in the kidney. The mechanisms for this natriuretic action appear to be multiple, involving increases in renal plasma flow and GFR, a redistribution of blood flow, and inhibition of the tubular transport of sodium. Dopamine infusion also increases urine cyclic AMP excretion in humans and this effect parallels the increased excretion of sodium into the renal tubules. There is now clear evidence for the involvement of DA₁ receptors in the natriuretic response to dopamine. Infusions of the DA₁ receptor antagonist SCH 23390 via the renal artery in the conscious uninephrectomized dog decrease urine flow rate and sodium excretion, and this effect is inhibited by the DA₁ receptor agonist fenoldopam. These *in vivo* studies have been supported by *in vitro* experiments, which have demonstrated specific binding sites for the DA₁ antagonist SCH 23390 in the proximal tubule of several species, and dose-dependent inhibition of the enzyme Na⁺K⁺-ATPase in the thick ascending limb of the loop of Henle by the DA₁ receptor agonist fenoldopam, which can be blocked by the DA₁ receptor antagonist SCH 23390. The role of DA₂ receptors in the natriuretic action of dopamine still requires further clarification, although these receptors have been identified clearly in the rat proximal tubule.

Effects of Dopamine on Renin Release

The effects of dopamine on renin release are variable, depending on the species and the experimental model. Intrarenal infusion of dopamine in the conscious dog causes a dose-dependent increase in renin release, which is attenuated by the DA₂ receptor antagonist sulpiride. Similarly, in both the isolated perfused rat kidney and dispersed rat renal cortical cell preparations, dopamine stimulates renin release via the activation of DA₂ receptors. In humans, the intravenous infusion of dopamine has been shown either to have no effect on plasma renin activity (PRA) or to increase PRA. In contrast, oral administration of the DA₂ receptor agonist bromocriptine, in humans, causes a decrease in PRA. It has been hypothesized that the stimulatory effects of dopamine on renin release may relate to a direct postsynaptic effect on the renal juxtaglomerular cells and that the inhibitory actions of dopamine on renin release may be presynaptic. Further studies are essential to clarify the actions of dopamine on renin release, especially in humans.

Intrarenal Sources of Dopamine

Renal plasma concentrations of dopamine are low compared to the concentrations of dopamine

required to cause the effects described earlier. It is also apparent that the concentration of dopamine found in urine is vastly greater than what could be accounted for by filtered plasma dopamine. This led to the hypothesis that there were renal sources of dopamine. One hypothesis suggested that dopamine in the kidney was largely derived from the dopamine conjugates sulfate and glucuronide, which are formed in the adrenal medulla and deconjugated in the kidney by the renal enzymes arylsulfatase and β-D-glucuronidase, respectively, releasing free dopamine. Although dopamine conjugates may be one source of renal dopamine, they do not appear to be the major source, as sodium chloride loading in adrenalectomized rats still leads to a marked increase in urine dopamine. A second hypothesis suggested that dopaminergic nerves, which were shown by immunofluorescence studies to be located primarily at the glomerular vascular pole, could release dopamine for local actions on renal hemodynamics in the kidney. In support of this hypothesis, it was demonstrated that renal nerve stimulation in the rat kidney produced vasodilatation that could be inhibited by the DA₂ receptor antagonist sulpiride. It seems likely that dopaminergic nerves could play a role in the regulation of glomerular tuft blood flow and in the regulation of renin release from the juxtaglomerular cells in the afferent arteriole. It is less likely that this source of dopamine plays a role in the natriuretic action of dopamine in the renal tubules. The major source of renal dopamine, which is of specific relevance to the natriuretic effects of dopamine, is undoubtedly that produced by the action of the enzyme aromatic L-amino acid decarboxylase (L-AAAD) on circulating L-dopa. This enzyme is located in the renal tubules, with higher levels in the proximal as compared to the distal segments. An increase in sodium chloride intake causes an increase in renal dopamine and urine dopamine, which may reflect an increase in the uptake of L-dopa from plasma into the proximal tubule (which is a sodium-dependent process), an increase in L-AAAD activity, or both. When the specific inhibitor of L-AAAD, carbidopa, is administered to humans, urine sodium excretion falls as a consequence of the decrease in renal dopamine synthesis.

Pathophysiology of Renal Dopamine

Because the renal actions of dopamine (increased GFR, vasodilatation, and natriuresis) are beneficial for the maintenance of normal renal function, its possible role in several different pathophysiological states, where renal function may be compromised, has been investigated. In some patients with essential hypertension, urinary dopamine excretion is reduced, and an increased sodium intake produces a

paradoxical fall in urinary dopamine in comparison to normal subjects, where sodium loading leads to an increase in urinary dopamine. This led to the hypothesis that the increased retention of sodium in some forms of human hypertension may relate to a defect in the kidney to produce dopamine in order to excrete the excess sodium from the body. Similarly, patients with chronic renal failure show reduced dopamine excretion compared to normal subjects and an impaired renal vascular response to dopamine infusions, which may relate to a loss of dopamine receptors within the kidney.

Adrenocortical Actions of Dopamine

Dopamine acts as a tonic inhibitor of adrenocortical function, and its effects appear to be restricted to the cells in the zona glomerulosa, where it acts specifically to inhibit the secretion of the principal mineralocorticoid, aldosterone.

In Vivo Studies

In 1975, it was observed that the DA₂ receptor agonist bromocriptine inhibited the plasma aldosterone response to furosemide (a diuretic) in humans. This was the first evidence to indicate a role for dopamine as an inhibitor of aldosterone secretion *in vivo*. Further studies in humans then demonstrated that administration of metoclopramide, a nonselective DA receptor antagonist, led to an increase in plasma aldosterone concentration, which was independent of changes in PRA and circulating sodium, potassium, and ACTH concentrations. The stimulatory effect of metoclopramide on aldosterone secretion in humans can be inhibited by the orally active DA agonist ibopamine, which by itself acts as an inhibitor of aldosterone secretion. In addition, experimental evidence suggested that alterations in angiotensin-induced aldosterone responsiveness to changes in dietary sodium chloride may be related to dopaminergic modulation. Although these studies indicate that dopamine can act as a tonic inhibitor of aldosterone secretion, it is not clear whether these effects are mediated primarily by the central nervous system (because the DA agonists and antagonists used can cross the blood-brain barrier) or whether they may relate to more direct actions on the cells in the adrenal zona glomerulosa.

In Vitro Studies

A direct inhibitory action of dopamine on angiotensin II-induced aldosterone secretion was first reported in 1979 by McKenna, who used isolated bovine zona glomerulosa cells. Pharmacological experiments using rat adrenal zona glomerulosa homogenates

later gave evidence for the presence of DA₁ receptors, which were coupled to activation of the G_s-protein, leading to the stimulation of cyclic AMP, and DA₂ receptors, which appeared to couple to the G_i-protein, causing a decrease in cyclic AMP. Compelling evidence suggests that the inhibitory effect of dopamine on aldosterone secretion is mediated by DA₂ receptors, which, when activated, inhibit intracellular Ca²⁺ concentrations or reduce transmembrane Ca²⁺ fluxes within zona glomerulosa cells. In particular, dopamine reduces angiotensin II-stimulated Ca²⁺ influx and angiotensin II-induced inositol phosphate production in rat zona glomerulosa cells. In addition, dopamine inhibits voltage-dependent T-type channels in cultured rat zona glomerulosa cells. No molecular studies have been reported yet on DA₂ receptors in the adrenal cortex, although DA₁ receptors have been shown to be present using *in situ* hybridization techniques to localize DA₁ receptor mRNA and immunohistochemistry to localize DA₁ receptor protein. The function of DA₁ receptors in the adrenal cortex requires further clarification, although tissue culture experiments in the rat zona glomerulosa suggest that activation of this receptor could lead to a stimulatory effect on aldosterone secretion.

Adrenocortical Sources of Dopamine

The concentrations of dopamine that are required to inhibit aldosterone secretion are far in excess of the concentrations of dopamine that circulate in plasma. Although dopamine is synthesized in the adrenal medulla, there is still no clear evidence that dopamine secreted by the adrenal medulla can affect steroid secretion in the zona glomerulosa directly. However, the enzyme L-AAAD, which catalyzes the conversion of circulating L-dopa to dopamine, has been located in the adrenal zona glomerulosa, and there is good evidence to suggest that circulating L-dopa may be the source of adrenocortical dopamine. The effects of dopamine within the adrenal cortex may therefore relate to a paracrine control mechanism, which is mediated by the activity of L-AAAD and the activation of specific DA receptors within the adrenal zona glomerulosa.

Conclusions

The principal actions of dopamine in the kidney are to increase renal blood flow and to promote natriuresis. The natriuretic effects of dopamine may be reinforced by its inhibitory action on aldosterone secretion in the zona glomerulosa cells of the adrenal cortex. These effects of dopamine are, to a great extent, mediated by the enzyme L-AAAD, which is located in the renal cortex and in the adrenal cortex.

The factors that control the expression and function of L-AAAD within the kidney and adrenal cortex, and the further development of selective DA₁ and DA₂ receptor agonists and antagonists, will be of paramount importance to our future understanding of the dopaminergic control of renal and adrenocortical function.

See Also the Following Articles

Adrenal Cortex; Aldosterone and Mineralocorticoid Receptors; Dopamine, Central.

Further Reading

- Aherne, A. M., Vaughan, C. J., Carey, R. M. and O'Connell, D. P. (1997). Localization of dopamine D_{1A} receptor protein and messenger ribonucleic acid in rat adrenal cortex. *Endocrinology* 138, 1282–1288.
- Buu, N. T. and Lussier, C. (1990). Origin of dopamine in the rat adrenal cortex. *American Journal of Physiology* 258, F287.

- Goldberg, L. I. (1972). Cardiovascular and renal actions of dopamine: potential clinical applications. *Pharmacology Reviews* 24, 1–29.
- Goldberg, L. I., Volkman, P. H. and Kohli, J. D. (1978). A comparison of the vascular dopamine receptor with other dopamine receptors. *Annual Review of Pharmacology and Toxicology* 18, 57–79.
- Kebabian, J. W. and Calne, D. B. (1979). Multiple receptors for dopamine. *Nature (London)* 277, 93–96.
- Lee, M. R. (1982). Dopamine and the kidney. *Clinical Science* 62, 439–448.
- Lee, M. R. (1993). Dopamine and the kidney: ten years on. *Clinical Science* 84, 357–375.
- Missale, C., Lombardi, C., DeCotiis, R., et al. (1989). Dopaminergic receptor mechanisms modulating the renin-angiotensin system and aldosterone secretion: an overview. *Journal of Cardiovascular Pharmacology* 14, S29.
- Muller, J. (1987). *Regulation of aldosterone biosynthesis, physiological and clinical aspects*. New York: Springer Verlag.
- Williams, B. C. (1986). Dopamine in the kidney. In: Winlow, W. & Markstein, R. (eds.) *The neurobiology of dopamine systems*, pp. 385–401. Manchester, UK: Manchester University Press.

Reproduction, Effects of Social Stress On

C A Shively

Wake Forest University School of Medicine,
Winston-Salem, NC, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
C A Shively, volume 3, pp 360–365, © 2000, Elsevier Inc.

Luteal phase deficiency

Abnormalities of the luteal phase of the menstrual cycle that may include reduced sex hormone concentrations and changes in phase length.

Stereotypy

A locomotor (e.g., pacing) or self-directed (e.g., biting) behavior that is performed precisely the same way, repeatedly in quick succession.

Social Stress and the Likelihood of Conception
Social Stress and Pregnancy Outcome
Long-Term Health of the Offspring

Glossary

<i>Functional hypothalamic chronic anovulation</i>	Inhibition of ovulation due to stress.
<i>Hypothalamic-pituitary-adrenocortical axis</i>	A major stress reactive system in the body involving the hypothalamus, pituitary, and adrenal glands.
<i>Hypothalamic-pituitary-gonadal axis</i>	The pathway of the reproductive system that controls gonadal function.

Reproductive failure rates are uniformly high across the mammalian order. From the perspective of human medicine, reproductive failure is often viewed as evidence of a disorder that requires treatment. However, mammologists and evolutionary biologists understand reproductive failure as an evolutionary safeguard against wasting energy on reproduction when the likelihood of a successful outcome is relatively low. From this perspective, it makes sense that the reproductive system would be especially sensitive to environmental variables that might influence the outcome of reproductive effort. Female mammals have a greater investment in reproduction than males, and the female reproductive system is particularly sensitive to environmental factors signaling adverse conditions for successful reproduction. Thus, there is less data available

concerning stress effects on reproductive function of males than on that of females. For species that live in groups, social factors are key aspects of the environment. One type of social environmental influence to which the reproductive system of mammals appears quite sensitive is social stress. Social stress is defined here as a social environmental factor that perturbs homeostasis and elicits behavioral or physiological change. The purpose of this article is to examine the evidence of parental social stress on reproductive health, including fertility, pregnancy outcome, neonatal health, and the lifelong health of the offspring.

Social Stress and the Likelihood of Conception

Physiological Mechanisms

The mechanisms by which social stress may influence the likelihood of conception are not completely understood. However, stress-induced activation of the hypothalamic-pituitary-adrenocortical (HPA) axis inhibits the hypothalamic-pituitary-gonadal (HPG) axis at multiple levels. Corticotropin-releasing hormone (CRH) inhibits hypothalamic gonadotropin-releasing hormone (GnRH). CRH stimulates pituitary adrenocorticotropin (ACTH) secretion, which in turn increases adrenal glucocorticoid secretion. Glucocorticoids inhibit GnRH secretion, as well as pituitary luteinizing hormone and ovarian estrogen and progesterone secretion. CRH also stimulates β -endorphin, which in turn inhibits GnRH secretion. Gonadal dysfunction results from alterations in GnRH pulsatility from the hypothalamic pulse generator impairing ovulation. Other neurotransmitters that mediate stress effects on the GnRH pulse generator include dopamine, neuropeptide Y, norepinephrine, γ -aminobutyric acid (GABA), and endogenous opioid peptides.

Dopamine receptors appear to mediate stress-induced transcription of hypothalamic CRH and proopiomelanocortin, from which a number of peptides are derived, including ACTH and endogenous opioids. The results of a series of nonhuman primate studies suggest that dopamine is involved in the effect of one type of social stress on the reproductive system. A major social organizing mechanism of primate society is the social status hierarchy. Among group-living primates, low social status, or social subordination, appears stressful and is associated with lower reproductive output in females. Socially subordinate female cynomolgus monkeys (*Macaca fascicularis*) receive more aggression, maintain a high level of vigilance, are groomed less by conspecifics, and spend more time alone than their dominant counterparts. Their adrenal glands hypersecrete cortisol, and

they are insensitive to glucocorticoid negative feedback. The central dopaminergic function of these females is also deranged, as evidenced by insensitivity to haloperidol, lower cerebrospinal fluid levels of homovanillic acid (HVA), a dopamine metabolite, and lower D₂ receptor binding potential in striatum observed by positron emission tomography. In males, the relationship between social subordination stress and reproductive output is less consistent than in females. Associations between social status and offspring sired are observed in some years but not others. Likewise, the relationship between low social status and activated HPA function is less apparent in male than in female Old World primates. However, central dopaminergic function of subordinate males is different than that of dominants, as evidenced by lower cerebral spinal fluid (CSF) HVA concentrations and lower striatal D₂ receptor binding potential. Thus, social subordination stress may alter central dopaminergic function, increase HPA activity, and inhibit HPG function.

Another mechanism that may mediate stress effects on conception involves reactive oxygen species (ROS). ROS can modify cell function and endanger cell survival. Because these are normal consequences of oxygen metabolism, ROS must be continuously deactivated by antioxidant defense mechanisms to avoid oxidative stress. There is some evidence that psychosocial stress may increase ROS, decrease DNA repair, and alter rates of apoptosis. Thus, oxidative damage may occur during stress. High levels of ROS are found in the semen of infertile men and the peritoneal and follicular fluid of infertile women.

Inhibition/Delay of Sexual Maturation

Two hypotheses address the relationship between social stress and sexual maturation. The social suppression hypothesis posits that stress suppresses sexual development. In support of this hypothesis, it has been widely observed in monkeys that low social status is associated with delayed sexual maturation and that dominant females reproduce at an earlier age than subordinate females. Likewise, girls from higher social classes enter puberty earlier than girls from lower social classes. Delayed sexual maturation following severe stress (e.g., war) has been documented in boys and girls, though is not well studied. In contrast, the psychosocial acceleration theory posits that under certain types of stressful conditions in which low-quality parental investment signals that parental investment does not promote reproductive success, sexual maturation may accelerate so that reproduction may occur earlier and more often. There is some evidence that family warmth and positivity may

delay, and family conflict and coercion may accelerate, pubertal development. Thus, the nature of the psychosocial stressor may be critical in predicting effects on sexual development.

Inhibition/Delay of Ovulation

Low social status is associated with inhibition of ovulation in baboons and macaques (*Macaca* spp.). Female cynomolgus monkeys (*M. fascicularis*) have menstrual cycles that are very similar to those of women. Socially subordinate female cynomolgus monkeys have low luteal phase progesterone concentrations, low follicular phase estradiol concentrations, a greater proportion of menstrual cycles with impaired luteal function, and a greater proportion of anovulatory menstrual cycles. Furthermore, when female monkeys change social status, their menstrual cycle characteristics change also (Table 1), suggesting that the reproductive system of these females is exquisitely sensitive to their social environment.

Socially suppressed ovarian function in these monkeys is associated with deleterious effects on the skeletal and cardiovascular systems, suggesting that socially induced hormonal impairments unaccompanied by amenorrhea have far-reaching effects on health.

Social stress may also inhibit ovulation in women, a condition referred to as functional hypothalamic amenorrhea or functional hypothalamic chronic anovulation (FHCA). Diagnosis is by exclusion of all other potential organic causes. The best indicator that FHCA is stress related is that it is generally accompanied by hypercortisolemia. Note that socially subordinate female monkeys that are stressed do not necessarily have more anovulatory cycles than their nonstressed dominant counterparts. Thus, in extrapolating to women, it is important to consider that stress-induced suppression of reproductive function may not manifest as frank amenorrhea, complicating diagnosis. In women, perturbations in GnRH pulsatility may result in ovarian dysfunction manifested

as amenorrhea, polymenorrhea, oligomenorrhea, or luteal phase deficiency, the latter of which may go undetected unless infertility issues arise. FHCA also appears reversible in women. The prevalence of FHCA varies depending on the subpopulation considered. Estimates in the general population range from 2 to 7%, whereas in high-risk populations, such as women going to college or joining the armed forces and women with depression, the prevalence may be as high as 70%. Under these circumstances ovulatory impairment appears to coincide with psychosocial stress and precede nutritional stress. Women that develop FHCA are more likely to be anxious, be perfectionists, and have difficulty coping.

Impairment of Hypothalamic-Pituitary-Testicular Function

During periods of social instability, low social status and aggression received are both associated with suppressed plasma testosterone concentrations in male macaques and baboons. In men, stressful events (e.g., work stress, exams, or competitions) may decrease testosterone levels, which can lead to decreased libido, diminished muscle mass, change in hair growth, and decreased spermatogenesis. Decreased testosterone may be due to perturbations in GnRH pulsatility or to direct effects of cortisol on testicular Leydig cell testosterone production. Social stress (e.g., death of a close family member) can result in oligoasthenozoospermia, which may remain undetected unless infertility issues arise. Decreased sperm number and quality are associated with decreased fertility.

Social Stress and Pregnancy Outcome

Physiological Mechanisms

There are several ways in which neuroendocrine stress responsivity may affect pregnancy. CRH regulates the HPA responses to stress. During pregnancy

Table 1 Effect of changing social status on menstrual cycles^a

Initial status	Subordinate		Dominant		Final status effect $p \leq$
Final status	Subordinate	Dominant	Subordinate	Dominant	
Number of cycles	15.2 (0.94)	15.4 (1.03)	14.4 (1.52)	15.1 (1.1)	0.60
Percent ovulatory	68.1 (9.19)	91.4 (3.36)	81.6 (6.34)	88.4 (4.42)	0.03
Percent impaired	19.4 (5.21)	4.2 (1.65)	8.4 (2.8)	8.1 (2.60)	0.08
Percent anovulatory	12.5 (6.00)	4.5 (2.21)	9.9 (5.44)	3.5 (2.76)	0.17
Progesterone (ng/ml)	6.7 (0.96)	10.4 (1.72)	8.1 (0.98)	9.6 (0.97)	0.03

^aMeans and standard errors. Social status of 42 female cynomolgus monkeys living in small social groups was documented for 2 months. The constituency of the social groups was changed such that half of the subordinates became dominant and half of the dominants became subordinate, and social status and menstrual cycle characteristics were documented for 24 months.

it also coordinates and controls the physiology of parturition. CRH is produced in the placenta, and placental CRH is stress sensitive. Maternal pituitary-adrenal hormones stimulate placental CRH secretion. Maternal stress is positively associated with circulating ACTH concentrations, and maternal social support is negatively associated with circulating concentrations of ACTH, β -endorphin, and cortisol. Maternal β -endorphin secretion during the third trimester is associated with increased fetal heart rate reactivity and decreased uroplacental blood flow, the latter being consistent with fetal hypoxia. Elevated CRH levels are associated with preterm labor and preterm delivery. Onset of the birthing process also involves oxytocin, vasopressin, and prostaglandins, as well as HPA activation. Secretion of these hormones may be altered by stress, thus affecting the timing of birth. Indeed, uteri from preterm deliveries appear to be hypersensitive to oxytocin. Stress-induced catecholamine and oxytocin may reduce the threshold of uterine contractility, thus resulting in preterm labor.

There is substantial evidence that HPA axis peptides (including ACTH and β -endorphin) affect the development of fetal brain and behavior and that perinatal exposure to ACTH and its analogues permanently alters growth and behavior. Thus, maternal stress may alter the trajectory of development beginning *in utero*.

Neuroendocrine stress responses are known to suppress the immune system; however, these processes have been little studied during pregnancy. Pregnancy involves immunosuppression in order to avoid fetal rejection. Thus, stress-related immunosuppression may be superimposed on the normal physiological immunosuppression of pregnancy, potentially affecting pregnancy outcome. Stress-induced immunosuppression increases vulnerability to infection. Infection, in turn, may induce preterm delivery. Preeclampsia is also associated with altered immune function and thus may be affected by stress-induced immunosuppression. Preterm birth is associated with enhanced expression of pro-inflammatory cytokines. Pro-inflammatory cytokines stimulate prostaglandin and metalloprotease synthesis in gestational tissues, which promote spontaneous labor and rupture of membranes. They also stimulate fetal-placental production of steroids and CRH, which promote parturition.

Norepinephrine and epinephrine are stress-responsive catecholamines that decrease blood flow to the uterus. Decreased uterine blood flow may cause hypoxia, hypotension, bradycardia, and malformations in the fetus. Hypoxia may result in increased ROS production. Oxidative stress due to increased ROS and/or decreased antioxidant defense mechanisms has been implicated in miscarriage and preeclampsia.

Thus, maternal, placental, and fetal neuroendocrine, immune, and cardiovascular processes link maternal stress during pregnancy with adverse pregnancy outcomes.

Failure of Implantation

It is difficult to determine failure of implantation, especially to discriminate it from ovulatory difficulties. Both social status and current ecological conditions affect the likelihood of successful implantation in yellow baboons. Under poor environmental conditions, socially dominant females (with low social stress) are more likely to achieve successful implantation than their subordinate, stressed counterparts. Likewise, in an *in vitro* fertilization study, implantation rates were reportedly higher in women with relatively low cardiovascular reactivity to a standardized laboratory stressor.

Spontaneous Abortion

There is an association between spontaneous abortion and psychological stress in pigtail macaques (*M. nemestrina*). Low social status, psychological stress, low control, and low social support have been associated with spontaneous abortion in humans. In 192 women with recent spontaneous abortions, it was found that recent negative life events significantly increased the likelihood of spontaneous abortion of chromosomally normal conceptions. In perhaps one of the largest studies available ($n = 214,108$), adverse pregnancy outcomes, including spontaneous abortion, were most frequent in women working in jobs characterized by high demand and low control. In another study of several thousand women, stressful work increased spontaneous abortion by a factor of 2.45 (95% confidence interval [CI] 1.03–5.81) in older women, 2.96 (95% CI 1.16–7.52) in smokers, and 2.27 (95% CI 0.97–5.27) in primigravid women. Thus, stressful life events, work characterized by high demand and low control, and stressful work may increase rates of spontaneous abortion.

Maternal Stress and Neonatal Mortality

Associations of neonatal mortality and maternal social stress are well documented in nonhuman primates. Rhesus macaque infants born to socially dominant females have a higher rate of survival than those born to socially subordinate mothers. Higher mortality also has been observed among offspring of psychologically stressed pigtail macaques. Chimpanzee and gorilla offspring born to subordinate mothers are subject to greater neonatal mortality than those born to dominants, in part due to infanticide. Maternal psychological stress and low social status have

been associated with neonatal mortality in humans. In a large study ($n = 214,108$), adverse pregnancy outcomes, including stillbirth, were most frequent in women working in jobs characterized by high demand and low control. Maternal risk factors for sudden infant death syndrome (SIDS) include low socioeconomic status (SES), unmarried status, teen age, less education, depression, drug use, and ethnicity. These characteristics are also risk factors for low birth weight, and low birth weight is a powerful predictor of SIDS. While it is difficult at this time to tease apart the relative contributions of these risk factors, all may involve increased maternal stress during gestation.

Maternal Stress and Neonatal Morbidity

The effects of maternal social stress during pregnancy on neonatal morbidity are more subtle and difficult to determine than neonatal mortality. Measures of neonatal morbidity include labor and delivery complications, gestational age, Apgar scores, birth weight, and neuromotor and neurobehavioral development during the first 30 days.

A series of studies in squirrel (*Saimiri boliviensis*) and rhesus monkeys demonstrated that mild maternal stress (e.g., removal from a home cage daily followed by 10 min of white noise during mid-late gestation) results in lower birth weight, impaired neuromotor development (as measured by a modified Brazelton Newborn Behavioral Assessment Scale), and impaired neurobehavioral development of attention and orienting. Early gestation stress was associated with more profound impairments to growth and neuromotor development during the first postnatal month than late gestation stress. These impairments are mimicked by administering ACTH to pregnant females for 2 weeks at mid-pregnancy, suggesting that activation of the HPA axis is instrumental in the deleterious effects of maternal prenatal stress on infant morbidity. Similar decrements in offspring development continue into adolescence (see later).

A large number of studies have been conducted to test the hypothesis that prenatal maternal stress causes infant morbidity in humans, and the balance of the literature supports this hypothesis. Birth weight and gestational age are the most commonly measured outcome variables. Low birth weight can be due to preterm delivery (less than 37 weeks gestation with an appropriate infant weight for gestational age) or to retarded fetal growth (inappropriate weight for gestational age). Maternal anxiety and depression result in low birth weight and perhaps smaller head size.

The magnitude of the effect is comparable to that of smoking. Composite measures of maternal prenatal stress (including some combination of trait or state anxiety, depression, life events, social support, and self esteem) best predict preterm delivery/low gestational age. Likewise, the association between psychosocial stressors and birth weight is strongest when composite measures of maternal stress (e.g., exposure to stressors such as life event scales, social support, anxiety, health risk behaviors) were used. The timing of the stress is important. Low birth weight and shorter gestational age have been observed when the stressors occur in the first trimester, whereas preterm delivery is associated with third trimester stress. In one of the largest studies available ($n = 214,108$), adverse pregnancy outcomes, including light-for-date birth weight and full-term low birth weight, were most frequent in women working in jobs characterized by high demand and low control.

Long-Term Health of the Offspring

Maternal Stress and Offspring Central Nervous System (CNS) Function

The effects of maternal stress during gestation have been shown to affect CNS function of rhesus monkeys (*M. mulatta*) through adolescence (4 years of age). Six-month-old offspring from control mothers and mothers that experienced a daily mild stressor (three random white noise bursts over 10 min) were evaluated during the latter half of pregnancy. In a response to a novel environment, offspring of prenatally stressed mothers exhibited more disturbance behavior (e.g., clinging to attachment object, self-directed stereotypic behaviors) and lower levels of gross motor and exploratory behavior than those of nonstressed mothers. At 8 and 18 months, prenatal stress-derived monkeys had higher CSF concentrations of metabolites of dopamine and norepinephrine, suggesting altered CNS function. The behavioral responses of these monkeys to social separation stress were evaluated. Disturbances in exploratory, locomotor, and clinging behavior observed at 6 months of age continued, and disturbances in social play were also observed. The effects of prenatal stress on later behavioral responses to stress were more profound for social than for nonsocial behavior, and suggest that offspring from prenatally stressed mothers are behaviorally hyperresponsive to stress. Between 6 months and 1 year of age, male offspring of prenatally stressed mothers had smaller, and females had larger, corpus callosum. At 14–19

months of age, prenatal stress-derived monkeys had higher ACTH and cortisol levels at baseline and under anesthesia. Furthermore, in response to stress, prenatal stress-derived offspring had higher ACTH and similar cortisol responses as control offspring. These observations suggest that the HPA axis of offspring of primate mothers stressed during pregnancy hyperresponds to stress later in life, and that the components of this stress-responsive system are dysregulated. The behavior of these prenatal stress-derived offspring was characterized again at 4 years of age, a time corresponding to human adolescence. At this age, the behavioral pathology of the prenatal stress-derived offspring was profound and included greater locomotion, stereotypies (pathological repetitive behaviors such as rocking), clinging, and self-clasping. These monkeys explored less in a novel environment and engaged in less social interaction in a playroom environment. At 5–7 years of age, *in vivo* imaging studies revealed a perturbed striatal dopaminergic system, which is critically involved in drug abuse vulnerability and sensitive to adult social stresses. Taken together, these observations suggest that prenatal stress can have profound and long-term effects on the behavior and neurobiology of primate offspring that is particularly apparent in response to stress.

Similar data from human studies are difficult to acquire. However, the results of a number of studies suggest that prenatal stress may perturb CNS function of newborns. There are reports of increased fetal activity associated with maternal anxiety and catastrophic stress due to earthquake. Anxious, type A, or depressed mothers have newborns who cry more and are more difficult to soothe. First trimester maternal stress and anxiety have been associated with low scores on cognitive and psychomotor tests, poor adaptation to novel environments, and difficult behavior in infants 8 months of age. At 9 years of age, attention deficits and aggressiveness were more common in the offspring of mothers who were anxious during pregnancy. Maternal stress is also linked to offspring psychopathology. Schizophrenia, in particular, has been linked to prenatal stress. Maternal exposure to death of a loved one, influenza infection, or catastrophic stressors increase the risk of the offspring developing schizophrenia, the onset of which is most common between 16 and 30 years of age.

The Significance of Low Birth Weight to Health in Later Life

Cardiovascular disease (CVD) is the leading killer of men and women. Low birth weight has been

associated with all-cause mortality, coronary heart disease, stroke, type II (non-insulin-dependent) diabetes mellitus, hypertension, and the metabolic syndrome. The metabolic syndrome is a cluster of characteristics (e.g., abdominal fat deposition, hyperglycemia, hyperinsulinemia, hypertension, and an atherogenic lipid profile) that greatly increase risk of CVD. Given the patterning of associations, it appears that low birth weight may be associated with lifelong derangement of the metabolic function of several major systems that culminates in increased risk of cardiovascular disease. Hypotheses currently being tested to explain fetal programming of adult health include the possibility that suboptimal environmental influences on fetal growth lead to permanent changes in organogenesis (i.e., particularly those involved with energy metabolism) that may serve to preserve the growth of certain organs, like the brain, to the expense of the function of other systems. Prenatal exposure to excess cortisol is considered a likely mechanism by which maternal stress during gestation may lead to fetal programming of cardiovascular, metabolic, and neuroendocrine disorders in adult life.

Exaggerated maternal stress responses can bathe the fetus in excessive glucocorticoids, resulting in an adult programmed for exaggerated HPA responses to stress. If this adult gestates young, she may also produce offspring with exaggerated stress responses, and thus intergenerational transmission of stress reactivity and subsequent ill health may be perpetuated. Indeed, low-birth-weight women are more likely to have low-birth-weight babies.

See Also the Following Articles

Childbirth and Stress; Chronic Social Stress: GR Sensitivity in Leukocytes; Social Stress, Animal Models of; Stress Induced Anovulation.

Further Reading

- Amiel-Tison, C., Cabrol, D., Denver, R., et al. (2004). Fetal adaptation to stress: Part II. Evolutionary aspects; stress-induced hippocampal damage; long-term effects on behavior; consequences on adult health. *Early Human Development* 78(2), 81–94.
- Berga, S. L. (1997). Behaviorally induced reproductive compromise in women and men. *Seminars in Reproductive Endocrinology* 15, 47–53.
- Buitelaar, J. K., Huizink, A. C., Mulder, E. J., et al. (2003). Prenatal stress and cognitive development and temperament in infants. *Neurobiology of Aging* 24(supplement 1), S53–S60.

- Burton, G. J. and Jauniaux, E. (2004). Placental oxidative stress: from miscarriage to preeclampsia. *Journal of the Society for Gynecologic Investigation* 11(6), 342–352.
- Drake, A. J. and Walker, B. R. (2004). The intergenerational effects of fetal programming: non-genomic mechanisms for the inheritance of low birth weight and cardiovascular risk. *The Journal of Endocrinology* 180(1), 1–16.
- Federenko, I. S. and Wadhwa, P. D. (2004). Women's mental health during pregnancy influences fetal and infant developmental and health outcomes. *CNS Spectrum* 9, 198–206.
- Huizink, A. C., Mulder, E. J. and Buitelaar, J. K. (2004). Prenatal stress and risk for psychopathology: specific effects or induction of general susceptibility? *Psychological Bulletin* 130, 115–142.
- Kaplan, J. R. and Manuck, S. B. (2004). Ovarian dysfunction, stress, and disease: a primate continuum. *Institute of Laboratory Animal Resources Journal* 45, 89–115.
- Koenig, J. I., Kirkpatrick, B. and Lee, P. (2002). Glucocorticoid hormones and early brain development in schizophrenia. *Neuropsychopharmacology* 27(2), 309–318.
- Mulder, E. J., Robles de Medina, P. G. and Huizink, A. C. (2002). Prenatal maternal stress: effects on pregnancy and the (unborn) child. *Early Human Development* 70(1–2), 3–14.
- Roberts, A. D., Moore, C. F., DeJesus, O. T., et al. (2004). Prenatal stress, moderate fetal alcohol, and dopamine system function in rhesus monkeys. *Neurotoxicology and Teratology* 26(2), 169–178.
- Schneider, M. L., Roughton, E. C., Koehler, A. J., et al. (1999). Growth and development following prenatal stress exposure in primates: an examination of ontogenetic vulnerability. *Child Development* 70(2), 263–274.
- Seckl, J. R. (2004). Prenatal glucocorticoids and long-term programming. *European Journal of Endocrinology* 151(supplement 3), U49–U62.
- Sullivan, F. M. and Barlow, S. M. (2001). Review of risk factors for sudden infant death syndrome. *Paediatric and Perinatal Epidemiology* 15(2), 144–200.
- Wasser, S. K. and Barash, D. P. (1983). Reproductive suppression among female mammals: implications for biomedicine and sexual selection theory. *The Quarterly Review of Biology* 58, 513–538.
- Zitzmann, M. and Nieschlag, E. (2001). Testosterone levels in healthy men and the relation to behavioural and physical characteristics: facts and constructs. *European Journal of Endocrinology* 144(3), 183–197.

Reproductive Dysfunction in Primates, Behaviorally Induced

J L Cameron

University of Pittsburgh and the Oregon Regional Primate Research Center, Beaverton, OR, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 366–372, © 2000, Elsevier Inc.

Overview

Stress-Induced Impairment of the Hypothalamic-Pituitary-Testicular Axis in Male Primates

Stress-Induced Impairment of the Hypothalamic-Pituitary-Ovarian Axis in Female Primates

Social Status as a Determinant of Reproductive Capacity Mechanisms Underlying Stress-Induced Reproductive Dysfunction in Primate Species

Individual Susceptibility to Stress-Induced Suppression of Reproductive Function

Summary

Glossary

Hypothalamic-pituitary-adrenal axis

The complex of (1) specialized cells in the hypothalamus that secrete the neuro-endocrine peptide, corticotropin-releasing hormone, (2) cells in the anterior pituitary that secrete adrenocorticotropin (ACTH), and (3) the adrenal gland that secretes glucocorticoid and mineralocorticoid hormones in response to ACTH.

Endogenous opioid peptides

Peptides released both within the brain and from pituitary cells that inhibit binding of morphinelike compounds to receptors.

Individual variability in stress responsiveness

Difference in physiological responses among individuals in a particular species to a given stress.

<i>Reproductive hormones</i>	Pituitary gonadotropins, including luteinizing hormone and follicle-stimulating hormone, and gonadal steroid hormones, including estradiol, progesterone, and testosterone.
<i>Social stress</i>	Stress arising from interactions with others or the perception of other's attitudes.
<i>Social status</i>	Placement of a particular animal within a social hierarchy, with subordinate animals deferring to the wishes of more dominant animals.

Both acute and chronic psychological and social stresses can impair reproductive hormone secretion in a variety of nonhuman primate species. This impairment can be subtle, consisting of a mild suppression in reproductive hormone secretion, or can be dramatic, causing a complete suppression of fertility and reproductive behavior. Among individuals there are marked differences in the responsiveness of the reproductive axis to social stresses, with factors contributing to this variability including the type of stress experienced, the magnitude and duration of stress, the perception of the stress by the individual, the social status of the individual, the concurrent level of aggressive behavior displayed by the individual, seasonal cues, and the prior level of activity within the reproductive axis (Figure 1). Mechanisms contributing to social stress induction of reproductive dysfunction include activation of the adrenal axis, increased secretion of endogenous opioids, increased prolactin release, and changes in sensitivity to gonadal steroid hormone feedback; however, in most forms of social stress the pathways leading to reproductive dysfunction remain to be elucidated.

Overview

While many forms of physical stress, such as energy restriction, temperature stress, infection, pain, and injury, have been clearly associated with the impairment of reproductive function, the effects of behaviorally induced stresses, i.e., psychological and social stresses, on the activity of the reproductive axis have been less well studied. The topic of behaviorally mediated stresses on reproductive function is of particular interest in primate species, which live in complex social groups and have higher cortical brain areas similar to humans, making study of these species particularly useful in understanding how such stresses may impact on reproduction and fertility in humans. This article provides a review of what is known currently about the role of psychological and social stresses in modulating the activity of the reproductive axis in nonhuman primate species. The review will show that the impact of psychological and social stresses on reproductive function is associated with significant variability among individuals and will detail the factors known to influence this variability in stress responsiveness.

Stress-Induced Impairment of the Hypothalamic-Pituitary-Testicular Axis in Male Primates

A variety of acute psychological stresses have been shown to cause the acute suppression of circulating luteinizing hormone (LH) and testosterone levels in male nonhuman primates. Examples include studies showing that the restraint of male rhesus monkeys suppresses pulsatile LH and testosterone release, and that mean testosterone concentrations decline rapidly in male rhesus monkeys in response to aggressive attacks by other monkeys. However, in other studies

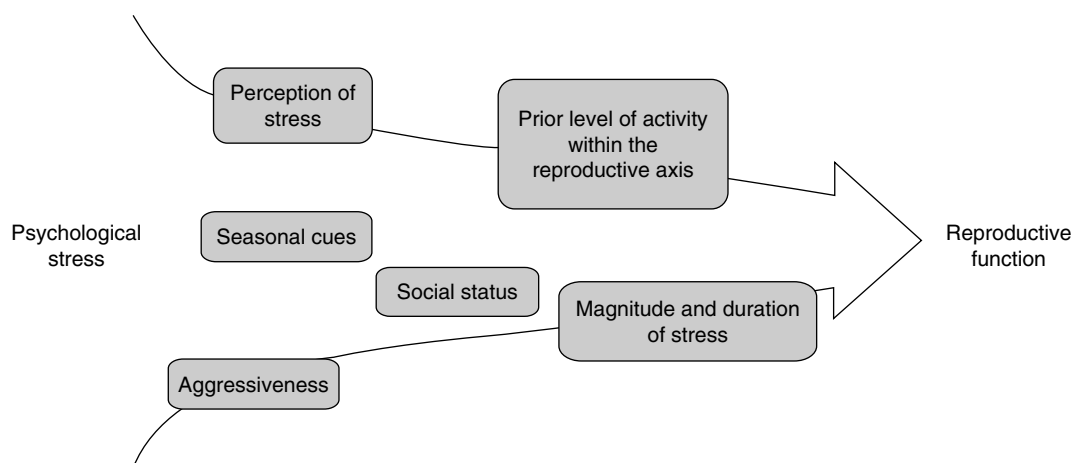


Figure 1 Schematic diagram of variables important in determining whether psychological stress will stimulate or inhibit reproductive function within an individual.

the acute response to stress in male primates has been shown to be more variable. In baboons, darting and capture have been shown to lead to an immediate suppression of testosterone in some animals, but to an initial increase in plasma testosterone levels followed by a much more subtle decline in other animals. Similarly, in response to pairing with an unfamiliar male animal, some male squirrel monkeys have an acute release of testosterone, whereas other males do not show this initial rise in testosterone. In this case, monkeys showing acute testosterone release are the behaviorally dominant monkeys in the pairs, whereas subordinate monkeys fail to show the acute testosterone release. Although many acute psychological stressors can suppress reproductive hormone secretion in male primates, not all such stresses have this effect. For example, our laboratory has used the acute psychological stress of exposing male rhesus monkeys to the sight of leather capture gloves to examine the physiological sequelae of acute psychological stress and has shown that although heart rate and cortisol release are elevated acutely in this paradigm and that immune cell function is suppressed for a number of hours, LH and testosterone secretion are not influenced. Although this stress evokes physiological stress responses in several systems, it seems likely that monkeys perceive this stress as being less threatening than a stress such as an aggressive attack by other monkeys. Thus, the perception of severity of stress is likely to be important in determining whether the stress will lead to a suppression of reproductive hormone secretion.

A number of chronic social stresses have also been associated with a marked and sustained suppression of reproductive hormone secretion in males, again with several investigations showing a marked dichotomy in response between dominant versus subordinate males, and other investigations showing that variability in the response of the reproductive axis to stress is dependent on environmental factors. Examples of prolonged stresses shown to lead to decreased testosterone secretion include social stress associated with loss or removal of the dominant male in the social group, housing males in all male social groups, and housing males in isolation. Interestingly, several social factors appear to modulate the response of the reproductive axis to stress in such conditions of chronic social stress. In a number of studies, a high degree of correlation between aggressiveness and testosterone titers, with the more aggressive males showing higher circulating levels of testosterone, has been reported. There also appears to be an interaction between the social status and the level of environmental stress, such that in periods of social stability there can be no difference in plasma testosterone levels

between dominant and subordinate animals, but in periods of social instability the dominant animals can show higher plasma testosterone levels than the subordinate males.

Based on the ability of both acute and chronic social stress to suppress reproductive hormone secretion, it has been posited that social status (i.e., dominance rank) plays an important role in determining the lifetime reproductive success in primates, with subordinate animals experiencing a greater degree of social stress and having a lesser degree of reproductive success. However, evidence for this hypothesis is weak, and it would appear that again multiple factors, including current dominance rank, time of year, magnitude of stress, aggressiveness of the animal, and level of activity of the reproductive axis prior to stress exposure, can all play roles in modulating reproductive success, in the same way these factors influence reproductive hormone secretion. Thus there are a number of instances in which subordinate males have similar long-term reproductive success when compared to dominant males.

Stress-Induced Impairment of the Hypothalamic-Pituitary-Ovarian Axis in Female Primates

Chronic social stress can also impair reproductive function in female nonhuman primates, which often show an increase in the length of the menstrual cycle and an increased incidence of anovulation in response to being maintained in stressful situations. In general, stress-induced lengthening of the cycle is characterized predominantly by a lengthened follicular phase and a normal length luteal phase. Social stresses associated with the inhibition of reproductive function include moving females to new social groups, housing in isolation, and even moving singly housed females to a new room where they can only interact with other animals through visual, auditory, and olfactory cues.

A number of studies have shown that dominant females have greater reproductive success than subordinate females in various nonhuman primate species. It has been widely assumed that subordinate females experience more stress than dominant females and that this stress plays a role in the relative suppression of reproductive function. This assumption has been born out by studies showing that subordinate female macaques housed in small group settings have larger adrenals and an enhanced adrenal responsiveness to adrenocorticotropin (ACTH) stimulation accompanying an increased incidence of anovulatory cycles. However, as in males, evidence shows that dominance rank may only determine reproductive success

when environmental resources are limiting, at least in some species. This supports the notion that a number of environmental and social factors play a role in determining the impact of stress exposure on reproductive function.

Social Status as a Determinant of Reproductive Capacity

In the preceding paragraphs, evidence has been presented showing that exposure to stress can suppress reproductive hormone secretion in nonhuman primate species that show normal adult reproductive function in nonstressed conditions. It seems likely that the ability of stress to impair reproductive function in these animals may well be analogous to the ability of psychological stress to impair reproductive function in humans. In addition, however, there are several species of nonhuman primates in which social interactions play a more absolute role in determining the activity of the reproductive axis, and this appears to be particularly true in the case of females. Specifically, in talpoin monkeys (*Miopithecus talapoin*) and marmosets, only the dominant female in a social group is reproductively active and all subordinate females are reproductively inactive until removal of the dominant female from the social group. As discussed later, evidence shows that social stress cues play an important role in this form of regulation of the reproductive axis.

In social groups of common marmosets (*Callithrix jacchus*), only the dominant female is reproductively active and bears offspring. As the female offspring mature in their natal group they remain reproductively quiescent, essentially held in a prepubertal state. These subordinate females display no signs of ovulation, they have low circulating concentrations of LH and ovarian steroid hormones, and the ovaries remain dramatically smaller than those of the dominant female, containing only small follicles and stromal tissue. The infertility can be reversed by several weeks of treatment with a pulsatile gonadotropin releasing hormone (GnRH) regimen, indicating that the suppression of reproductive function is at a central level to block the GnRH drive to the reproductive axis. Similar findings have been reported in female talpoin monkeys.

In these species, subordinate males also appear to have a suppression of reproductive axis activity in comparison to the dominant males, but the impairment of the reproductive axis is not as severe in males as in females. In male marmosets, the subordinate animals living in social group settings have lower circulating concentrations of LH and testosterone. In male talpoin monkeys, there is a general trend for the most dominant males to display the greatest

amount of sexual behavior and the most subordinate males to show little or no sexual behavior when housed in social groups. When housed in social groups of multiple males and females there is also a trend for the dominant males to have higher circulating concentrations of testosterone than the most subordinate males, but on a day-to-day basis this relationship between social rank and plasma testosterone concentrations is not always apparent. Interestingly, when male talpoin monkeys are moved to individual cages the correlation between social rank and plasma testosterone concentrations completely disappears. It thus appears that social rank modulates the activity of the reproductive axis, but that this effect is only apparent when monkeys are in a social setting where dominance hierarchies are in place.

The two most common routes by which social cues can suppress reproductive function in mammals are by behavioral intimidation and by pheromonal cues. Although there is some evidence that pheromonal cues play a role in the reproductive quiescence in the subordinate animals of these species, ablation of the olfactory epithelium and vomeronasal organ does not prevent suppression of reproductive function in subordinate animals, clearly indicating that other cues play a role in maintaining reproductive suppression. Behavioral intimidation of subordinate animals is apparent in both of these species. When marmosets are first placed in social groups there is a brief period where they receive an increased amount of harassment and intimidation from dominant animals, and in talpoin monkeys the place within the social hierarchy of a particular animal is established by the direction and frequency of aggressive interactions between various animals. The most dominant animals receive the least amount of aggression from others in the social group, whereas the most subordinate animals receive the most aggression. In addition, subordinate animals visually monitor more dominant animals at very frequent intervals. Subordinate animals also show an increase in circulating concentrations of the stress hormones, prolactin and cortisol. Cumulatively, data discussed earlier indicate that the stress associated with subordinate status within a social group at least partially underlies the suppression of reproductive function in these species of primates.

Mechanisms Underlying Stress-Induced Reproductive Dysfunction in Primate Species

Activation of the Adrenal Axis

In many cases, social stress-induced reproductive dysfunction is accompanied by an activation of the

hypothalamic-pituitary-adrenal axis, and it is possible that in these cases hormones of the adrenal axis may play a role in suppressing reproductive hormone secretion, and ultimately sexual behavior. Several secretory products of the adrenal axis have been shown to be capable of suppressing activity of the reproductive axis. Corticosteroids have been shown to both directly inhibit GnRH secretion and suppress the pituitary responsiveness to GnRH, although these effects of corticosteroids require high levels sustained over a several-week time course. Corticotropin-releasing hormone (CRH), the central hypothalamic-releasing factor of the adrenal axis, has also been shown to be a potent inhibitor of LH secretion when it is administered in pharmacological experiments. In addition, for some forms of physical stress, the administration of CRH antagonists can prevent the stress-induced suppression of reproductive hormone secretion.

Despite the evidence that elevated activity of the adrenal axis often accompanies suppressed reproductive hormone secretion in various forms of psychological and social stress in primates, there is no evidence to date that in these psychological stress conditions that activation of the adrenal axis is indeed causing the suppression of the reproductive axis. A causal relationship would only be established if it could be shown that blocking activation of the adrenal axis reversed the psychological stress-induced inhibition of the reproductive axis, and such experiments have not been performed to date. In addition, it is important to recognize that in at least several situations where other physical stresses are associated with both an increase in adrenal hormone secretion and a decrease in reproductive hormone secretion that such studies have shown that activation of the adrenal axis is not causing the suppression of the reproductive axis. Thus, the role that activation of the adrenal axis actually plays in causing suppression of the reproductive axis in any conditions of social or psychological stress remains to be elucidated.

Other Mechanisms

Evidence shows that the increased release of endogenous opioid peptides plays at least a partial role in suppressing the activity of the reproductive axis in some forms of social stress. There are several reports that administration of the opioid receptor antagonist, naloxone, can restore or partially restore LH secretion in conditions of psychological stress, including the suppression of LH and testosterone secretion in chair-restrained rhesus monkeys and after anesthetic darting in olive baboons. Naloxone can also restore LH secretion when administered to ovariectomized, subordinate marmosets. However, because opioid receptor antagonists are not effective in restoring activity of the reproductive axis in all forms of social

stress-induced reproductive dysfunction, clearly other mechanisms are also involved in transmitting information about social stress to the reproductive axis.

Subordinate talpoin monkeys show a marked elevation in circulating prolactin levels compared to dominant animals, and it is possible that this increase in prolactin may play a role in suppressing the activity of the reproductive axis in this species, in that bromocryptine, a dopamine agonist, which reduces prolactin secretion, is able to at least partially restore reproductive hormone secretion. However, prolactin does not appear to mediate stress-induced suppression of the reproductive axis in all species. Even in another primate species in which subordinate animals show marked suppression of the reproductive axis, the marmoset, prolactin levels are not elevated and do not appear to play a role in mediating the inhibition of reproductive function.

In marmosets, evidence shows that the social suppression of reproductive function is associated with an increased sensitivity of the reproductive axis to estradiol negative feedback, as shown by the ability of low levels of estradiol to inhibit LH secretion in subordinate, but not dominant, female marmosets. In talpoin monkeys, removal of the subordinate female from the social group leads to a restoration of reproductive hormone secretion and increased sensitivity of the reproductive axis to estradiol-positive feedback. The neural mechanisms that underlie changes in sensitivity to estradiol feedback in these conditions remain unknown.

A review of all the studies examining social stress-induced suppression of the reproductive axis suggests that there are likely to be multiple neural systems that contribute to the inhibition of reproductive function in times of social and psychological stress. Clearly, there appear to be species differences in the neural pathways leading to reproductive inhibition. However, perhaps more interesting to consider is the possibility that within many species a similar set of pathways is involved in the suppression of reproductive axis activity, but that differential activation or inhibition of specific pathways within this complex is dependent on the specific stress experienced, as well as within a specific species. A great deal more work is needed in this field to dissect the neural mechanisms underlying behaviorally induced reproductive dysfunction and how this varies with the type of stress and with species.

Individual Susceptibility to Stress-Induced Suppression of Reproductive Function

A concept that arises again and again in this area of investigation is that in most all species some animals appear to be more susceptible to the stress-induced

suppression of reproductive function than others. As discussed in this article, subordinate social status is associated with a greater inhibition of the reproductive axis, either in response to an imposed stress or in the normal living situation, in a diversity of primate species, including rhesus monkeys, cynomolgus monkeys, baboons, talpoin monkeys, and marmosets. A difficult question to answer that needs to be addressed in future studies is whether these animals are intrinsically more susceptible to stress-induced reproductive dysfunction or whether they experience more stress and thus have greater inhibitory drive to their reproductive axis.

Summary

As detailed in this review, both short- and long-term exposure to a variety of psychological and social stresses can suppress reproductive hormone secretion and lead to a decrease in reproductive behaviors and fertility in a number of nonhuman primate species. Although there appear to be some species differences in the degree or route of inhibition of the reproductive axis caused by exposure to such stresses, there also appear to be several general principles that are common to most species. Perhaps most important of these is the concept that the effect of a particular stress at a particular time on the reproductive axis of an individual animal can be modulated by a great number of variables. These modulating variables include social status, the magnitude and duration of stress, perception of the stress, aggressive behavior, seasonal cues, and the prior level of activity within the reproductive axis (see [Figure 1](#)). Thus, although it is possible to measure group mean responses to specific stresses and make conclusions about the effect of a stress on the reproductive axis in that species, such a mean assessment may be of little use in determining whether an individual animal will experience a suppression of reproductive function in response to that stress. An increased understanding of the mechanisms by which psychological and social stresses suppress the activity of the reproductive axis may well be achieved by focusing future studies on these individual differences in response to stress.

See Also the Following Articles

Opioids; Primate Hierarchies and Personality; Social Status and Stress.

Further Reading

- Abbott, D. H., O'Byrne, K. T., Sheffield, J. W., et al. (1989). Neuroendocrine suppression of LH secretion in subordinate female marmoset monkeys (*Callithrix jacchus*). In: Eley, R. M. (ed.) *Comparative reproduction in mammals and man*, pp. 63–67. Nairobi: National Museums of Kenya.
- Adams, M. R., Kaplan, J. R. and Koritnik, D. R. (1985). Psychosocial influences on ovarian endocrine and ovulatory function in *Macaca fascicularis*. *Physiology and Behavior* 35, 935–940.
- Bowman, L. A., Dilley, S. R. and Keverne, E. B. (1978). Suppression of oestrogen-induced LH surges by social subordination in talpoin monkeys. *Nature (London)* 275, 56–58.
- French, J. A., Abbott, D. H. and Snowdon, C. T. (1984). The effect of social environment on estrogen excretion, scent marking, and sociosexual behavior in tamarins (*Saguinus oedipus*). *American Journal of Primatology* 6, 155–167.
- Keverne, E. B., Meller, R. E. and Eberhart, A. (1982). Dominance and subordination: concepts or physiological states? In: Chiarelli, O. (ed.) *Advanced views in primate biology*, pp. 81–94. Main Lectures of the VIIIth Congress of the International Primatological Society, Florence, 7–12 July 1980. New York: Springer-Verlag.
- Nishida, T. (1983). Alpha status and agonistic alliance in wild chimpanzees (*Pan troglodytes schweinfuthii*). *Primates* 24, 318–336.
- Rowell, T. E. (1970). Baboon menstrual cycles affected by social environment. *Reproduction Fertility And Development* 21, 133–141.
- Sade, D. S., Cushing, K., Cushing, P., et al. (1976). Population dynamics in relation to social structure on Cayo Santiago. *Yearbook of Physical Anthropology* 20, 253–262.
- Sapolsky, R. M. (1982). The endocrine stress-response and social status in the wild baboon. *Hormones and Behavior* 16, 279–292.
- Wilson, M. E., Gordon, T. P. and Bernstein, I. S. (1978). Timing of births and reproductive success in rhesus monkey social groups. *Journal of Medical Primatology* 7, 202–212.

Resistance

L M Zabarenko

Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
L M Zabarenko, volume 3, pp 373–375, © 2000, Elsevier Inc.

Introduction
Definitions
Implications for Technique
Implications for Theory
Current Views
Recent Research
Conclusion

Glossary

<i>Intellectualization</i>	The use of intellectual activities to exert control over anxiety and reduce tension. This proclivity is often encountered as a defense against instinctual drives in adolescence and is enlisted as resistance in insight-oriented therapy.
<i>Psychoanalysis</i>	A discipline discovered by Sigmund Freud which is (1) a psychotherapeutic technique, (2) a systematized theory of human behavior including development and (3) a method of investigating the mind.
<i>Resistance</i>	A familiar but paradoxical phenomenon in psychotherapy and psychoanalysis. The patient, having entered treatment, opposes the process in ways which can be conscious, pre-conscious or unconscious.
<i>Transference</i>	The displacement of patterns of feelings originally experienced in childhood onto a person involved in a current relationship.
<i>Transference resistance</i>	The usually unconscious use of transference affects as a resistance to therapeutic progress.

Introduction

By the time people consult a mental health professional for relief from the consequences of stress, they are usually highly motivated. Psychological suffering; anxiety and depression; and loss of functions such as the inability to work, engage fully in interpersonal relations, or care for themselves have often reached profound levels. It is remarkable, then, to discover that many patients also resist treatment and that such

opposition is a reliable phenomenon that can be expressed in a whole universe of behaviors. The patient may be silent, withhold essential information from the therapist, flood the sessions with intellectualizations or dreams so that reporting consumes all the time, arrive late for appointments or fail to come, delay paying the fee or refuse to pay at all, and find it impossible to follow directions about medications. Many additional forms of resistance have been reported, including eagerness to make the best use of the time in the sessions, “talking over the treatment every day with some intimate friend,” what Freud called a “therapeutic leak,” using evasive speech (or lapsing into technical jargon) or clichés, or mumbling and accusing the therapist of suggesting ideas too strongly. There is also documentation of resistance as it occurs in the termination phases of psychoanalysis in children and in adults.

Patients can be aware of resistance, but often they are not; that is, resistance may be conscious, preconscious, unconscious, or a mix of all three. Regardless of the etiology, resistance tends to be most pronounced in the insight-oriented psychotherapies, those aimed not only at symptomatic relief but also at self-knowledge.

Definitions

In *The Interpretation of Dreams*, Freud was crystal clear: “Whatever disturbs the progress of the work is resistance”; this concept is still current. Stressing that resistance is ubiquitous and pervasive, Moore and Fine (1990: 168–169) also point out that “analysis of resistances . . . has become a hallmark of psychoanalytic treatment.”

Resistance as a Defense

In psychoanalysis, the most ambitious form of psychotherapy, patients are often unconsciously anxious about revealing themselves, dreading possible rejection by the analyst. They may also fear health itself as alien, dangerous, and possibly more menacing than being sick.

Under such circumstances, almost anything can be drafted to slow the process of change, as the lists described earlier attest. In addition, patients may treat the therapist as an enemy and alternately comply with and rebel against his or her presumed wishes; for example, some patients report memories or dreams but cannot work with them via associations. Among the feelings defended most keenly are those toward the therapist (i.e., transference resistances), but unconscious resistance may also be evoked in the analyst or therapist

and these constitute a manifestation of countertransference problems. These can interfere with the therapist's ongoing empathic resonance with the patient, resulting in technical misjudgments such as inappropriate content or timing of interpretations.

Resistance Resulting from Shifts in the Therapeutic Alliance and as Distinct from Negative Therapeutic Reaction and Negative Transference

Because it is pervasive, resistance should be clearly distinguished from other therapeutic phenomena, such as negative therapeutic reactions or negative transference. According to Moore and Fine (1990: 168–169), a negative therapeutic reaction is a “clinical response. . . . Following a period of what seems to be constructive and effective therapeutic technique and understanding, the patient's condition paradoxically worsens.” Freud's position was that “Every partial solution that ought to result . . . in an improvement or a temporary suspension of symptoms produces . . . for the time being an exacerbation of (the) illness.” However, when the patient sees the therapist as malevolent, exploitive, or manipulative, these are negative transference feelings.

History

Sigmund Freud is generally credited with the identification and explication of resistance in contemporary psychological thought. His first mention of the entity came in a discussion of Elizabeth von R., a patient who was sent to him in the fall of 1892 and became his “first full-length analysis of a hysteria.” Instead of using hypnosis, as he had previously in his work with Josef Breuer, Freud encouraged this patient to associate freely, to say whatever came to mind. When she was silent and he inquired what was going on in her head, she replied, “Nothing.” Freud refused to accept this answer, believing that it was the same species of willful forgetting that had caused her hysterical symptoms in the first place.

Another patient, the obsessional neurotic known as “the rat man,” began treatment in late 1907. This man described the conflict that results in resistance by quoting Nietzsche to Freud: “‘I did this,’ says my Memory. ‘I cannot have done this,’ says my Pride and remains inexorable. In the end – Memory yields.” Describing a kind of punishment practiced in Asia, the patient rose from the couch and begged Freud not to insist that he go on. Freud explained that “The overcoming of resistances is a law of the treatment” and helped him to finish the story; the perpetrator was tied down, a pot with rats in it was turned upside down on his buttocks, and the rats bored their way into his anus.

By 1912 Freud was confident enough to emphasize the pervasiveness of the phenomenon: “(It) accompanies the treatment step by step. Every single association, every act of the person . . . must reckon with the resistance . . . (and it) represents a compromise between the forces that are striving towards recovery and the opposing ones.” Everything that distracts the patient from following the fundamental rule of free association is an obstruction.

Implications for Technique

Because resistance is a dramatic and undeniable example of the operation of the unconscious, its appearance can signal when therapeutic work is approaching painful and important areas of the illness. Such nodal points are often thick with cues about the etiology of the neurosis, and tact and timing are essential when attempting to bring the resistance to the patient's awareness. Careful and accurate interpretations at this point can inform both participants greatly and free up the therapeutic log jam.

Implications for Theory

At first thinking that resistance was simply an automatic defense, Freud saw it chiefly as an obstacle to the work. However, as the implications and extent of unconscious resistances became clearer, these led Freud to reconsider his topographic model of the psyche as the relationship of conscious, unconscious, and preconscious mentation and to incorporate these concepts into a new stance that stressed a structural view of the mind as composed of the id, ego, and superego.

One good way to grasp how profoundly the discoveries about resistance changed psychoanalytic theory is to contrast the taxonomies Freud used to categorize them. In 1895, he thought resistances appeared under three circumstances: (1) if “there is a personal estrangement . . . if the patient feels that she has been neglected, insulted, too little appreciated,” (2) if the patient dreads becoming too dependent on the physician or influenced by him, or (3) if the patient is frightened at finding “she is transferring on to the figure of the physician the distressing ideas which arise from the content of the analysis.” Coincidentally, this discussion marks the first use of the term transference in the psychoanalytic sense.

Twenty-eight years later in a famous passage, Freud found he required a five-tier classification system. Three species of resistance came from the ego – repression, resistance, transference resistance – and the unwillingness to renounce the gains of being ill was “an assimilation of the symptom into the ego.”

Id resistance, approached by “working through” the analytic discoveries, and superego resistance completed the list. The latter, “the last to be discovered, seemed to spring from guilt or the need for punishment and opposes every move toward success.”

Current Views

The understanding of resistance has remained a salient feature of psychotherapy and psychoanalysis and a prominent focus of psychoanalytic technique. However, with the epigenesis of post-Freudian theories, each school has proposed new ways of thinking about the etiology of resistance and offered corresponding technical recommendations for dealing with it.

Self-psychologists believe that resistance is evoked by failure of the therapist’s empathic attunement and that such failure will mirror the narcissistic patient’s early deprivation. Object relations theorists stress the extent to which resistance occurs when there has been insufficient recognition of the real as well as the unique transference relationship between the therapist and the patient. Brenner’s revised structural proposals see resistance mainly as a defense that is a compromise formation in the patient’s conflict between wishing to recover and wishing to remain ill. Schafer feels that resistance can be traced to insufficient understanding of transference–countertransference currents as these occur in analytic work.

Recent Research

The explosive proliferation of research in the neurosciences and the deepening sophistication of findings in cognitive psychology have profound implications for clinicians facing resistance in their daily work. There are venerable journals in each area, but they are not notably accessible to the others and few acknowledge or attempt to integrate the contributions from all three areas. *Neuro-psychoanalysis* is one of the most scholarly and useful of a group of new journals whose aim is to do this. First published in 1999, it is staffed by large and distinguished editorial boards in neuroscience and psychoanalysis. The lively ongoing disagreements in this journal and the roster of established experts as authors testify to the relevance of neuroscience to analytic and dynamic psychotherapeutic work and simultaneously to the triangulation of cognitive psychology, neuroscience, and psychotherapy.

For example, recent research in cognitive psychology has underscored the importance of continuity in forming psychological constructs and offers an appealingly quantitative explanation of the fact that resistance that disrupts continuity (i.e., missing

sessions or coming late) can delay or even disrupt therapeutic progress. The experimental demonstration that remembering itself can cause inhibition suggests that when patients’ recovery of important memories falter and transference constructs might be invoked, the therapist could also benefit from information from relevant cognitive principles that would enrich understanding the difficulty and inform technique.

Clinician–researcher–scholars capable of interdisciplinary synthesis are rare but not unknown. In a 2003 summary of his work, for example, Levin illustrates the importance of learning in psychotherapy and the links among learning, psychoanalysis (especially but not exclusively transference), and the neurosciences, an emphasis he has espoused for decades. Citing recent research findings, he suggests using these to augment existing ways of thinking about resistance – that resistance occurs because “patients defend themselves against more novelty than they can process at a given time” (Levin, 2003: 82), a cognitive and neural overload.

Conclusion

The importance of the patient’s resistance to treatment has escalated as the disciplines of psychotherapy and psychoanalysis have matured and become theoretically and technically more diverse. However, the observation that there are obstacles to cure has remained consistent, as has the indefatigable work of finding more effective ways to understand and clear away these barriers to progress.

See Also the Following Article

Psychoanalysis.

Further Reading

- Brenner, C. (1982). *The mind in conflict*. New York: International Universities Press.
- Dewald, P. A. (1980). The handling of resistances in adult psychoanalysis. *International Journal of Psychoanalysis* 61, 61–70.
- Eder, M. (1930). Dreams as resistance. *International Journal of Psychoanalysis* 11, 40–47.
- Fraley, R. C. and Roberts, B. W. (2005). Continuity: a dynamic model for conceptualizing the stability of individual differences in psychological constructs across the life course. *Psychological Review* 112, 60–74.
- Freud, S. (2000). The interpretation of dreams (1900). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vols. 4–5), pp. 1–630. New York: W. W. Norton.

- Freud, S. (2000). Notes upon a case of obsessional neurosis (1909). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vol. 10), pp. 153–257. New York: W. W. Norton.
- Freud, S. (2000). The dynamics of transference (1912). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vol. 12), pp. 98–108. New York: W. W. Norton.
- Freud, S. (2000). The ego and the id (1923). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vol. 19), pp. 1–66. New York: W. W. Norton.
- Freud, S. (2000). Inhibitions, symptoms and anxiety (1926). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vol. 20), pp. 77–175. New York: W. W. Norton.
- Gay, P. (1988). *Freud: a life for our time*. New York: W. W. Norton.
- Kohut, H. (1971). *The analysis of the self*. Madison CT: International Universities Press.
- Levin, F. (1985). The need for a psychoanalytic learning theory. Paper presented at the meeting of the American Society of Adolescence, Dallas, TX, May 17.
- Levin, F. (2003). *Psyche and brain. The biology of talking cures*. Madison, NJ: International Universities Press.
- Moore, B. and Fine, B. (eds.) (1990). *Psychoanalytic terms and concepts*. New Haven, CT: American Psychoanalytic Association and Yale University Press.
- Schafer, R. (1997). Vicissitudes of remembering in the countertransference: fervent failure, colonisation and remembering otherwise. *International Journal of Psychoanalysis* 78, 1151–1163.
- Stone, L. (1973). On resistance to the psychoanalytic process. In: Rubinstein, B. B. (ed.) *Psychoanalysis and contemporary science*, pp. 42–73, New York: Macmillan.
- Velting, H. and van Knippenberg, A. (2004). Remembering can cause inhibition: retrieval-induced inhibition as a cue independent process. *Journal of Experimental Psychology* 30, 315–318.

Restraint Stress

R J Servatius

New Jersey Medical School and New Jersey Health Center, D. V. A. Medical Center, East Orange, NJ, USA

G Salameh and K M Coyle

New Jersey Medical School, Newark, NJ, USA

W P Paré

D. V. A. Medical Center, Perry Point, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 376–377, © 2000, Elsevier Inc.

Methods of Restraint

Stressor Intensity

Stressor Parameters

Glossary

Chronic stress state A persistent presence of physiological and psychological symptoms of the stress response following exposure to an extreme stressor.

Inescapable stress A stressor that is delivered to an animal whereby it cannot escape or avoid by performing any particular behavior.

Methods of Restraint

Loose Restraint

A method of restraining an animal so that free movement is restricted, yet some movement is still possible. Loose restraint is often used in conjunction with other procedures, not necessarily related to stress research, where it is necessary that the animal's movements be minimized, e.g., while taking a blood sample or administering an injection. In rodents, loose restraint is usually accomplished by placing the rodent in a small chamber or in a cloth. In nonhuman primates, restraint is accomplished by making the size of an individual cage smaller.

Tight Restraint

This method completely eliminates an animal's ability to move. Techniques for rodents include jacket or harness restraint or securing the rodent's limbs as it is lying prone. For nonhuman primates, tight restraint is accomplished by placing the subject in a chair with straps attached.

Supine Restraint

A variation of tight restraint for rodents in which the animal is tied down while lying on its back.

Combination Restraint Stress with Other Stressors

Restraint stress is combined with another type of stressor such as cold, heat, water immersion, and electrical shock to the tail.

Stressor Intensity

Hans Selye performed the pioneering work with restraint stress. The use of restraint, as an experimental procedure, was initially developed by Serge Bonfilio in France and by David Brodie in the United States. Subsequently, restraint stress has been utilized extensively, alone or in combination with other stressors, in order to measure its physiological, psychological, and pharmacological effects. Similar to pharmacological studies, the capacity of a particular treatment to produce a credible effect is reinforced by what is known as the “dose-response” relationship. The basic tenet of this relationship is that increasing doses of a particular treatment elicit increasing degrees of that treatment’s effect. Therefore, based on integrated stress research spanning the fields of psychology, neuroscience, endocrinology, immunology, gastroenterology, cardiology, and pharmacology, one may grade a particular stress protocol’s “dose” or intensity. In their two comprehensive reviews of the restraint stress literature, Paré and Glavin have provided the means to gauge the intensity of various experimental manipulations, including the parameters of restraint type, duration of restraint, and number of stress episodes, among others. Intensity is verified through associated effects on behavioral and physiological stress respondents. Thus, the methods of restraints can be graded as combination > supine > tight > loose.

Stressor Parameters

Duration of Restraint

While the type of stress most accurately predicts the intensity of a stress episode, the duration of the episode can greatly affect its intensity as well. In general, increasing the duration of a particular restraint session

increases the degree of stress-associated symptoms. However, with the milder restraint types (loose restraint), longer duration bouts may be associated with decreased responsiveness—evidence of habituation.

Number of Restraint Episodes

As with the duration of a restraint session, increasing the number of sessions has the general effect of inducing more pathological consequences. However, this generalization is qualified by studies demonstrating diminished physiological responsiveness with multiple stressful episodes.

Other Factors

Several factors will influence the degree or duration of stress responses to restraint greatly.

1. Gender: pathophysiological consequences from restraint stress are more apparent when restraint stress episodes occur in the activity phase of estrus as opposed to the inactivity phase.
2. Cycles: the effects of restraint stress appear to differ as a function of time of day and season.
3. Age: although the amplitude of glucocorticoid responses in aged rats does not differ from young rats exposed to restraint stress, the duration of these elevations is longer in aged rats.
4. Strain: some rat strains (e.g., Wistar-Kyoto) are more vulnerable to the effects of restraint, whereas others are more resistant.

See Also the Following Articles

Animal Models (Nonprimate) for Human Stress; Immobilization Stress; Selye, Hans.

Further Reading

- Glavin, G. B., Paré, W. P., Sandbak, T., Bakke, H. and Murison, R. (1994). Restraint stress in biomedical research: an update. *Neuroscience and Biobehavioral Reviews* 18, 223–249.
- Paré, W. P. and Glavin, G. B. (1986). Restraint stress in biomedical research: a review. *Neuroscience and Biobehavioral Reviews* 10, 339–370.

Revenge Fantasies

M Horowitz and S Meffert

University of California, San Francisco, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Biological Issues

Social Issues

Psychological Issues

Treatment Issues

Conclusion

Revenge fantasies have been an important human activity since the beginning of recorded history. The Bible, Shakespeare, novels, and modern television and movies motivate their characters with fantasies and plans for seeking vengeance. The more traumatic an experience, the more intense are the revenge fantasies. Everyone is so familiar with revenge fantasies that there is no need to review their many forms. What is of concern is how to control revenge fantasies, how to reduce preoccupation with them, and how to prevent maladaptive action.

Biological Issues

The theme of revenge fantasies is so pervasive in human history, so common throughout populations that it seems to function almost like an instinct or drive. There is an adage that revenge is a dish best served cold. The indication of a biological function is not that revenge is cold; rather, it is a hot thirst that seems to be slaked when the revenge is taken. Animals other than human beings may take a revenge for a recent attack. Whenever a human being is caught up in a primal revenge action plan it is possible to use higher psychological processes to increase self-governance and to provide social organizational structures to prevent damage.

Social Issues

Genocidal campaigns are usually based on propaganda that incites bloodlust on the basis of revenge fantasies. General secular values are needed to counteract some fundamentalist religious values that indicate the need for blood revenge, as in killing whole populations. The evolution of society itself shows this history over several centuries. Social schemas of understanding and compassion have been used for this purpose. Our focus is mainly on psychological issues, so we comment here that, after a traumatic experience shared by a population group, it is important to

emphasize to that population what is adaptive and maladaptive in terms of themes of retaliation versus those of adequate compensation or preparation so that the catastrophe will not happen again.

As recently noted by the World Health Organization, there is a divide in humanitarian aid services between mental and social health, particularly around the issue of trauma. Many debates have centered on the idea that posttraumatic stress disorder (PTSD) requires an understanding of individual pathology and that treatment that should be indexed to a particular culture, population, and era; critics argue that PTSD diagnosis and treatment are therefore not generalizable, particularly not across cultures. Such debates have become so heated that humanitarian aid efforts often have to choose between using individual, trauma-based mental health treatment and broader psychosocial public health-oriented treatment to address community-level mental health problems. The problem with splitting treatment into individual and social components is that it may obscure their heavily intertwined nature by compartmentalizing them into separate categories.

One of the most potentially destructive aspects of trauma is the impulse for retaliation that follows. The tragic aspect of this impulse is that it is much more often enacted against the close social contacts of the traumatized individual than the original aggressor; although the focus of revenge fantasies may be the hated enemy, the target is often loved ones, such as spouses and children. Research with Vietnam veterans shows that veterans with symptoms of psychological trauma are at high risk of committing violent acts toward their spouses and children.

One horrifying tactic of warfare capitalizes on this feature of human nature by raping enemy women or women who belong to the same social group as that of the enemy. This appears to be calculated not only to infect bloodlines; humiliate, torture, and sometimes kill the victim; but also to destroy social bonds in the enemy group. The rapists understand that the desire for revenge and anger against the group they represent will cause the community to reject the raped women from their social roles of wife, mother, and relative. The rapists count on the idea that feelings of anger and revenge will triumph over the social bonds of love and affection, leading to the internal destruction of the community. Unfortunately, they are often right.

Psychological Issues

Emily Dickinson may have led her cloistered, unmarried, and childless life because she felt betrayed. She became preoccupied with love and death in her poetry; in which she could only fantasize revenge. Here, for example, is her poem #339F.

I like a look of Agony,
 [perhaps the look she would like to imagine on the
 face of this man]
 Because I know it's true
 [not false seductive expression, which she may have
 felt lured by]
 Men do not sham Convulsion
 [meaning the signs of suffering, although they may
 feign love]
 Nor simulate, a Throe

 The eyes glaze once – and that is Death –
 [she is imagining as she writes that this man's death
 was painful, as he deserved]
 Impossible to feign
 The Beads upon the Forehead
 By Homely Anguish Strung

One of the worst betrayals is when a trusted person exploits us. In modern days, some men insist on going bareback into intercourse (i.e., they refuse to use a condom). Some do so knowing that they risk transmitting a disease they know they have. The sexual partner finds out later that he or she has been infected with human immunodeficiency virus (HIV), hepatitis C, syphilis, gonorrhea, chlamydia, or genital herpes. Revenge fantasies then become a frequent compensation and sometimes a violent enactment.

The recipient of the disease may have felt love but that turns to hate. Revenge fantasies accumulate; forgiveness is beyond imagination. Still, the goal is to avoid violence or identifying with the selfish lover. Struggling to master a betrayal does not mean either taking revenge or giving up any course of compensatory action.

An effective response to a trauma is a course of action that may create more good than evil. Telling a community circle about the bareback perpetrator does so because it may prevent similar crimes. But figuring out the differences between a maladaptive hunger for revenge and an effective course of action is difficult. Expressions from the Bible such as “an eye for an eye, a tooth for a tooth” are uttered frequently. In our contemporary understanding of this language, they have a bloody, excessive connotation, which was not meant when the text was written. Then it meant do not annihilate or destroy the betrayer excessively but rather take a measured and reasoned course by retaliating by taking only an eye for an eye or a tooth for a tooth. In other words, do not massacre.

The perpetrator of a trauma that leads to revenge fantasies is seldom a total monster; we can try to understand his motives. Having some understanding does not mean blindly trusting the same abusing person or group again. Revising beliefs and values can lead to increased skill at seeing other humans as blends of good and evil, caring and selfishness. We get better at reading their intentions and learn when and how to be self-protective.

This work of reviewing memories of a trauma or betrayal can be painful. Intense negative emotions are rekindled. The strength of revenge fantasies may lead to repeated intrusive episodes, making it hard to concentrate our attention elsewhere. Our anger, fear that no one can be trusted, and despair over the coldness of the world can cause us to feel bitter or to impulsively engage in acts of hostility. Sometimes our aggression is displaced on to some weaker target. The crippling interpersonal effect of revenge fantasies is that, whether fulfilled or not, they often result in more destruction of those we love than of our despised enemies. Such venting leads to guilt and shame.

How can we avoid recycling trauma and revenge that causes further trauma in this manner? Thinking of a betrayal over and over again, without progress in healing; only increases bitterness and diminishes our chances at happiness. Our goal must be a point of completion in our processing of the memory. This may involve efforts to achieve a new more correct and realistic judgment. A plan for making a measured response at the right target or at no target can result.

A memory of an injury or an insult is fertile ground for growing a seed of rage into a towering weed that can strangle love and compassion. Some people or some groups always seem to carry a chip on their shoulder; they expect frustration from others and even nice gambits from others can be twisted and become misinterpreted as not enough. They feel as if the whole world has already betrayed them and refuses to make up for it. Each new betrayal reinforces this irrational attitude.

Seething with anger does have an energizing function because possessing targeted anger often makes us feel strong, whereas fear makes us feel weak. That is why states of self-righteous indignation are so popular on talk radio and why the revenge ideas are easy to sow and hard to give up.

After a betrayal, a person with such character traits of bitterness and resentment often enters a state of self-righteous indignation and then stays in this state because it feels like energy or fuel for the self. The burning anger helps such people feel more solid instead of feeling frail, empty, or apathetic. Breaking free of such snags is difficult. We can, however, learn to recognize the trap and step away from it. With the

recognition that we can be strong without being in a self-righteous rage, we can begin to analyze the situation rather than replaying an endless loop of revenge that goes nowhere.

A character trait of bitterness can result from unfortunate experiences in childhood or adolescence. Rightly or wrongly, many children regard their parents as omnipotent and so blame their parents for any suffering they felt as a consequence of unfortunate events. Any new frustration that occurs in their adult lives just adds to the trait of bitterness.

Treatment Issues

There are many evidence-based treatments for depression and anxiety, yet there remain very few treatments for the anger and interpersonal violence that often stem from trauma and revenge fantasies – with the exception of dialectical behavioral therapy, which generally aims to treat anger that results in self-harm. Violent behavior and the enactment of revenge fantasies often find their treatment, or lack thereof, in the criminal justice system, sometimes with court-mandated anger-management programs, for whose success there is little empirical support. Many factors may contribute to this situation, one of which could involve a conflicted attitude that society has toward anger, violence, and revenge. Particularly with cultural constructions of masculinity, the qualities of assertiveness, forcefulness, and even a measure of anger or the implication of violence are sometimes valued as reflecting ambition, competence, or efficiency in achieving life goals. Hollywood, of course, is a frequent purveyor of these stereotypes, with the hot-tempered violent hero with a chip on his shoulder from some previous wrong. These concepts stand in contrast to other cultural norms, such as the many aphorisms and maxims that underscore the destructive effects of anger and revenge.

As an adult, a person can realize the truth of parental limitations. These limitations, seen through adult mindfulness, are usually different from adolescent visions of how bad the parents had been. As a child, the person regards a parental lapse as being done on purpose. As an adolescent, he or she may show contempt for grownups. As an adult, the person can revise his or her entire internalized story of what happened and the explanations for how and why it happened. For example, a mother's depression and substance abuse can be seen as a terrible difficulty and struggle rather than as a deliberate abandonment of a child.

The story of a person's own life can be revised in the direction of having more sympathy for, and less blame toward, parental figures. This does not mean

memories and attitudes toward parents are changed to happy ones. Many adults had unerasable childhood tragedies and or deficiencies in parenting. On adult reflection, there may be a realistic appraisal of the past events, but that realistic version may still fall far short of some ideal or even good-enough parenting.

Realistic appraisal may lead to sadness, even a delayed mourning for childhood losses. The person must mourn a childhood that cannot ever be repeated with a healthier-minded parent. Passing through that grief makes the person ready to seize what the here and now might provide and to look to the near future for what is possible. New, warm, loving relationships in various areas of one's life can help one to heal the wounds of childhood. The person learns to forgive him- or herself because the person was not to blame, as he or she might have supposed as a child, and to forgive others for the neglect they could not prevent.

Suppose the person had an abusive and neglectful father, became bitter, and now unconsciously hoped to find an ideal father as an adult. The person hopes that new candidates will make up for the damage done in the past, but they all turn out to fail the mission. At first, the new candidate for caring seems amiable and supportive, but the good qualities fall far short of the person's immense expectations. Sometimes others are provoked to fail; an otherwise inexplicable hatred is then felt toward the contemporary nice person and the inappropriate anger ruins the relationship. Efforts to change can stop this vicious cycle and reduce even the character trait of being constantly bitter.

The general wisdom of pursuing forgiveness does not condone neglect or abuse; it simply understands the humanness in all parties. The person avoids demonizing the targets of his or her moral indignation. Compassion is sought, but this does not mean passive resignation to enduring wrongs. Turning the other cheek may only invite more slaps in the face.

Giving up self-righteousness is not the same thing as giving up moral indignation. Moral indignation is a response to seeing people break rules when they could follow them. A good person can use moral indignation and do something to prevent evil and promote compassion or virtue. He or she may use moral indignation to gather a like-minded group together to encounter and counteract the exposure of children to abuse or neglect. This kind of indignation is effective and adaptive, whereas chronic bitterness with revenge fantasies is inefficient and maladaptive.

After experiencing an intense trauma or betrayal, the person has to ask him- or herself whether the perpetrator is a friend, an enemy, or perhaps an unreliable narcissist. The person modifies his or her expectations according to a realistic appraisal. The

person must learn to recognize enemies only when they exist. Even in these situations, the best aim is to temper wrath with reason because the sword of destructiveness often cuts backward.

It is helpful to ask questions. Why did the enemy attack or deceive? How was I harmed? If I counter-attack, what are the likely consequences? How does a cycle of betrayal, attack, and counterattack end? In answering these questions, the person finds that the extreme of an ideal self and a demonic enemy is almost always wrong. The person must clarify more realistic central appraisals.

Suppose there is a persistent enemy. One ideal scenario is to trust in a truce based on a pact of non-belligerence. Can a person realistically find a way so that the enemy will respect the rules? In the dreaded train of thought, the enemy is divested of human qualities and is viewed as a constantly malignant force. In an ideal train of thought, the enemy keeps his or her word of honor. These opposing extremes set the stage for a moderated train of thought. The rational analysis does not automatically arrive at truce or warfare. The enemy may remain the enemy. The rational appraisal leads to preparation for defense and continued alertness to signs of increasing danger. It rejects an attitude of unending battle and vendettas. It sets up bad consequences for breaking a truce and good consequences for keeping it. The person may fight the enemy but plan to accept the enemy's child.

Conclusion

The best adage ever coined against brutality is the golden rule: do unto others as you would want them to do unto you; do not treat others as you would not want to be treated. An excellent quality of such adages is that they can become a socially shared morality or ethic that can be taught by any religion or culture.

A new adage might be: when you feel bitter, seek the good. This adds something to existing popular adages such as: perform random acts of kindness and senseless acts of beauty. Usually people want to do good when they feel good or are in love. The goal is to do good when we are less elated. The idea of bringing goodness into the world when we feel demeaned, insulted, angry, and frustrated is a useful one to highlight because this stance can counteract self-righteous rage and revenge impulses.

This attitude – when you feel bitter, do good – counteracts two normal but malignant responses to

being insulted. One response is to cave in and feel weak, deflated, and degraded; the other is the opposite, a result of role reversal, to adopt a strong menacing role from which to lash out at others who are weaker. For example, a child who sees a civil war squad murder his parents may well grow up to become a civil war soldier ready to commit similar atrocities or worse.

Adages help keep us alert, so here are some other versions. When you feel bitter, do good. Do not wait until you feel good to act with compassion for others. Protect others as you want to be protected. The best cure for an enraged heart is comforting others. Do not pass hatred on; fight it with compassion.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Disasters and Mass Violence, Public, Effects of; Domestic Violence; Male Partner Violence; Posttraumatic Stress Disorder in Children; Posttraumatic Stress Disorder, Delayed; Posttraumatic Stress Disorder, Neurobiology of; Violence; War Stress in the Former Yugoslavia; War-Related Posttraumatic Stress Disorder, Treatment of; War, Suicide and Sacrifice; Posttraumatic Stress Disorder – Clinical; School Violence and Bullying.

Further Reading

- Cardozo, B. L., Kaiser, R., Gotway, C., et al. (2003). Mental health, social functioning, and feelings of hatred and revenge of Kosovar Albanians one year after the war in Kosovo. *Journal of Traumatic Stress* 16, 351–360.
- Dickenson, E. (1924). *The Complete Poems*. Boston: Little, Brown.
- Horowitz, M. J. (1998). *Cognitive psychodynamics: from conflict to character*. New York: John Wiley.
- Horowitz, M. (2001). *Stress response syndromes* (4th edn.). Northvale, NJ: Aronson.
- Horowitz, M. (2004). *Treatment of stress response syndromes*. Washington, DC: American Psychiatric Press.
- Horowitz, M. J. (2005). *Understanding psychotherapy change*. Washington, DC: American Psychological Association.
- Orth, V., Montad, L. and Maercker, A. (2006). Feelings of revenge, retaliation motive, and posttraumatic stress reactions in crime victims. *Journal of Interpersonal Violence* 21, 229–243.
- van Ommeren, M., Saxena, S. and Saraceno, B. (2005). Mental and social health during and after acute emergencies: emerging consensus? *Bulletin of the World Health Organization* 83(1), 71–77.
- Young, A. (1997). *The harmony of illusions: inventing post-traumatic stress disorder*. Princeton, NJ: Princeton University Press.

Rheumatic Disorders

A T Masi and J C Aldag

University of Illinois College of Medicine at Peoria
(UICOMP), Peoria, IL, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction

Prototype Rheumatic Disorders and Stressors

Possible Stressor Specificities to Particular Pathways

Profoundly Complex Patterns of Stress Responses

Stressors Related to the Physiopathogenesis of Rheumatoid Arthritis

An Evolving Research Challenge in Rheumatoid Arthritis

Summary

Glossary

Ankylosing spondylitis

A form of arthritis mainly affecting the spine and sacroiliac joints and, to a lesser degree, those of the girdles and lower extremities. A striking association occurs with the genetic factor HLA-B27 and familial aggregation is strong. Mechanisms remain obscure, but typical enthesopathic lesions suggest that biomechanical stresses could activate inflammatory pathways in its physiopathogenesis.

Biomechanics

The study of the application of mechanical laws and the actions of internal or external forces upon the living body or its structures. Biomechanical forces are believed to be important contributors to a number of rheumatic disorders, especially ankylosing spondylitis and osteoarthritis.

Central sensitization

A biopsychological process whereby sensory signals or stimuli are amplified in intensity with expansion of the receptive fields through central nervous system (CNS) mechanisms.

Central sensitivity syndromes

Chronic, noninflammatory rheumatic disorders believed to constitute an interrelated or overlapping family of symptomatic and dysfunctional conditions. The causes are unknown, but may result from central sensitization mechanisms and may be amplified by environmental, personal, or socioeconomic stresses.

Enthesopathy

A pathological process occurring at sites of attachment or insertion of joint capsules, ligaments, or muscle tendons into bony tissues, which is typically found in ankylosing spondylitis and related spondyloarthropathy conditions.

Fibromyalgia syndrome

A common and prototypic example of the family of central sensitivity syndromes. It is defined by widespread pain and the physical examination finding of multiple foci of localized tenderness on firm palpation at characteristic points (tender points), particularly of myofascial tissues. Pain, stiffness, and systemic symptoms of poor sleep and fatigue, anxiety or depression are often aggravated by insufficient or excessive physical activity and by environmental, personal, or socioeconomic stresses.

Microvascular endothelium

The single cell lining layer of the luminal surface of the microvascular system, which regulates transport of macromolecules and blood components into the interstitium. Endothelial cells are regulated by the neuroendocrine and immune (NEI) systems and contribute to inflammation by their activation leading to enhancement of inflammatory cell migration from the lumen into the interstitium, and by their secretion of mediators of inflammation.

Osteoarthritis

The most common form of arthritis, particularly in older ages, which results from biomechanical stresses and biochemical changes that damage articular cartilage and also cause lesions in the surrounding osteoligamentous tissues of the joint. The process may be either primary resulting from chronic excessive repetitive joint impacting or secondary to gross trauma or inflammatory arthritis conditions. Weight-bearing joints are mainly affected and it is often referred to as degenerative joint disease.

Physiopathogenesis

A process which progresses from chronic disordered physiology or dysregulated function of organs or systems (presymptomatic phase) to later stages of pathological, symptomatic disease states. The implication of this term is a transition from initial chronic physiological dysregulations (preclinical phase) to later clinical disease states. Alternatively, the term, pathophysiology, implies that initial pathological processes alter physiological functions secondarily. Bidirectional derangements often occur in many diseases, but to differing degrees, and depending upon the phases of the processes.

Rheumatoid arthritis

A chronic systemic disease primarily affecting small joints of the extremities in a symmetrical fashion and characterized

Rheumatic disorders

by inflammatory changes in the synovial membranes. Without effective therapy, the pathological process extends beyond the synovial tissue and inflammation erodes adjacent joint cartilage and bone, leading to deformities and disability. The presymptomatic causes of this disease leading to inflammatory pathways are unknown, although such pathology predominates in the clinical phase. Women are affected about three times as often as men, which is opposite to ankylosing spondylitis, and is an important clue to respective predisposing mechanisms.

The general nosological term for conditions affecting the musculoskeletal system that include symptoms of pain, stiffness, and other related abnormalities pertaining to joints or nonarticular (myofascial) structures. These disorders include a wide variety of conditions characterized by degeneration, inflammation, or metabolic derangements of connective tissues, joints, or other musculoskeletal structures

Spondyloarthropathy disorders

A group of overlapping arthritic diseases affecting extremity or sacroiliac joints and the vertebral spine to varying extents. Ankylosing spondylitis is the prototypic disorder in this group that mainly affects the sacroiliac joints and the spine, but can also involve lower extremity and girdle joints. They share clinical and radiological features, particularly enthesopathy lesions and increased presence of the cell surface HLA-B27 antigen marker.

Introduction

Historical Contributions to Stress Concepts

Modern concepts of regulatory biology may be traced to Claude Bernard (1813–78), who pioneered the doctrine of a constancy of the *milieu interieur*. Walter Cannon (1871–1945) further promoted and broadened the concept of physiologic organization and regulations, especially in his 1929 report on homeostasis. That original concept referred to an ideal setting of effectors controlling the *milieu interieur*. However, modern concepts of homeostasis infer the occurrence of different settings of the effectors as well as the accumulation of stress impacts or allostatic load on the physiological control mechanisms. The paradigm of allostatic load proposes that stress mediators have protective as well as damaging effects on the organism. The accumulation and overexposure of the body to stresses can promote adverse effects on various systems by the mediators.

In the next decade, Hans Selye (1907–82) developed experimental (rodent) models of generalized stress reactions to noxious agents. He demonstrated opposite changes in the thymus (atrophy) and adrenals (hypertrophy) as part of the organism's responses to injuries and chemical intoxications. Those experiments revealed counterregulatory relations between the hypertrophy of adrenal glands and involution of lymphatic tissues, which he originally considered to be nonspecific.

In 1946, Selye published his synthetic theory of the general adaptation syndrome and diseases of adaptation. Relative adrenal insufficiency was proposed to be a contributory factor in stress-related or immunologically mediated diseases. Selye's research and that of others indicated that the adrenal glands are closely connected with the organism's resistance to noxious stimuli or stresses.

In humans, adrenal gland enlargement was recognized to result from a wide variety of conditions, including those of increased physiological demands (pregnancy, cold exposure, exercise, and low oxygen tension), hypermetabolic disturbances (hyperthyroidism, and excessive insulin administration), infections, burns, and other shock-producing states as well as multiple toxic drugs. The concept of stress reactions became recognized largely under Selye's influence. His theory offered a unifying mechanism to interpret complex sets of metabolic, psychoneurophysiological, immunological, and other adaptive changes that characterized responses to a wide variety of noxious or perturbing stimuli.

The above-mentioned and other concepts of stress biology will be related to selected examples of rheumatic disorders. Human stress responses are profoundly complex, but their diverse impacts may be inferred to contribute to symptoms, physiopathogenesis, or outcomes of particular rheumatic disorders.

A Broad Spectrum of Rheumatic Disorders

Rheumatic disorders encompass a broad spectrum of conditions which affect the musculoligamentous (myofascial) and skeletal structures of the body as well as other soft connective tissues and internal organ control systems. These conditions include over 100 medical diagnoses. The diseases usually have defining clinical features, which vary in patterns of pain, psychological behaviors, limitations of movements, localizations of inflammation, or structural involvements in the body. The patterns of symptoms or pathology result from varied, but often overlapping primary or secondary pathways in the causations of these diseases. Classification of these different disorders is often based upon defining symptoms or clinical patterns, rather than upon specific etiology.

Nosology of Rheumatic Disorders and Potential Stressors

The term, rheumatism, is derived from the Greek *rheuma*, flux, inferring that symptoms or manifestations can *flow* from one part of the body to another. The expanded classification of *Arthritis and Rheumatic Diseases* is the overriding medical label for this broad spectrum of conditions and diseases. The role of stress in such a large number and diverse patterns of rheumatic disorders cannot be fully addressed in this limited review. Particular examples and principles of stress reactions affecting this family of conditions are offered. Research indicates that the general population is often exposed to physical or psychological stressors, which vary considerably among individuals and subgroups. The patterns of stressors tend to relate to particular rheumatic disorders, presumably via their respective mechanistic pathways. Examples of types of stressors affecting rheumatic diseases are provided.

Prototype Rheumatic Disorders and Stressors

Features of Fibromyalgia Syndrome

A type of noninflammatory rheumatic condition associated with stress mechanisms is fibromyalgia syndrome (FMS). This disorder is manifested by musculoligamentous (myofascial) pain, fatigue, nonrestorative sleep, affective features, and decreased functional ability. This condition is an identifiable clinical entity, despite the absence of objective physical deformity, inflammation, or structural lesions. Its clinical symptoms follow rather characteristic patterns. Its course tends to be associated with the degrees of subjectively perceived stresses or objectively experienced physical aggravating factors. The causal pathways of FMS are unknown, but are believed to involve central nervous system (CNS) and autonomic nervous system (ANS) activation and sensitization. The more prominently implicated stressors are believed to operate via these physiological pathways.

The particular stress factors which aggravate FMS differ in individuals, perhaps due to their predispositions and personalized circumstances. They include variously described stresses in personal relationships, demands of daily living as well as particular environmental or occupational hardships. Additional individualized stressors include cold or unusually warm temperatures, trauma, physical impacting or other biopsychological stresses, and biomechanical strains on the body. Data suggest that sufficient intensity of stress can cause central sensitization, as reflected in posttraumatic stress disorder (PTSD).

FMS is believed to be but one recognized clinical entity in a family of central sensitivity syndromes. These disorders compose a spectrum of clinically overlapping and co-occurring noninflammatory conditions. They frequently manifest with myofascial pain, fatigue, nonrestorative sleep, hypersensitivity to various stimuli, and affective features. The conditions are somewhat diverse, involving different systems, but tend to have overlapping physiological stressors and aggravating environmental or occupational factors. Criteria have been developed to classify the particular syndromes, based upon relatively defining clinical features.

Features of Rheumatoid Arthritis

The main inflammatory rheumatic condition believed to be affected by stress mechanisms is rheumatoid arthritis (RA). This disease is characterized clinically by pain, objective swelling, and inflammatory damage in multiple extremity joints (i.e., polyarticular) in a symmetrical pattern. Blood samples of RA patients clearly show evidence of systemic inflammation. Increased levels of biomarkers (e.g., acute phase proteins) and mediators of inflammation (e.g., inflammatory cytokines) are found in the blood and joints during active phases of this disease.

In genetically susceptible people, the underlying systemic mechanisms believed to be operating in RA are dysregulations of the neuroendocrine and immune (NEI) as well as the microvascular endothelial (MVE) systems and their interactions. The more specific stressors are believed to further perturb the already dysregulated NEI and MVE systems or place increased demands upon joints. However, CNS-mediated psychosocial stresses also affect pain behaviors and disease activity in RA, as may occur generally in other disorders.

Stressors of the RA clinical status include particular physical (e.g., overuse of joints) or subjective (e.g., poor sleep and fatigue) factors as well as systemic circumstances, which overtax the physiological balance or energy reserves of the body. Aggravating or stressor factors can increase both symptoms and objective manifestations of RA. Such stressors are believed to operate via enhanced perturbations of the already dysregulated physiopathogenetic pathways in RA as well as by greater direct physical demands upon compromised joints.

Features of Osteoarthritis

Various biomechanical stressors operate in rheumatic disorders, to cause either aggravated symptomatology or accelerated tissue damage. Osteoarthritis (OA) and ankylosing spondylitis (AS) are diseases, which result in damage and deformity to peripheral or spinal joints. The mechanisms are believed to be primarily

caused by biomechanical stressors, at least in their initial stages. Later phases of these diseases may include some degrees of inflammation, either localized articular (OA) or combined joint and systemic (AS), but to lesser degrees than in RA.

Structurally, the basic elements of body design are balanced systems of compression and tension in its functioning components. Biomechanical stresses which alter such balance, e.g., excessive compression (chronic overloading from obesity) or excessive tension (suspected axial myofascial hypertonicity in AS) can contribute to accelerated joint damage or increased symptomatology.

In diarthrodial (freely moving) joints, movements occur across the smooth specialized cartilage tissue surfaces covering the ends of connecting bones (e.g., knuckles of the hands, hips, or knees). This remarkable cartilage tissue provides combined compressional elasticity (i.e., water trapping) and tensile strength (i.e., resistance to tearing). Although a popular lay concept is that joints are primarily designed to absorb impacts, their main function is to permit smooth movements in transmission of dynamic forces between connecting bones, ligaments, and muscles. Friction and impacting of joints are normally minimal, since surface movements (velocity vectors) are parallel to the joint contours, which have covering fluid films. The coefficient of friction across normally articulating joint surfaces is miniscule (Figure 1).

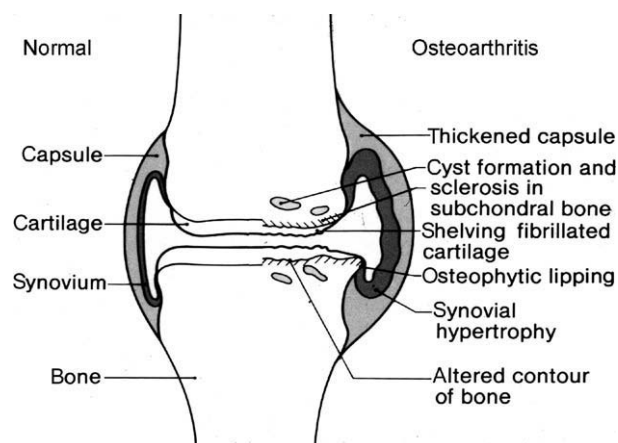


Figure 1 Schema of normal versus osteoarthritis joint. Anatomical features of a normal diarthrodial joint are schematically shown. The cartilage is a specialized tissue which normally provides compressible elasticity and resistance to tearing in smooth joint movements. Muscles and ligaments are attached to the joint capsule and bones, providing essential dynamic support and transfers of forces. In osteoarthritis, cartilage undergoes loss of tissue and function and the articulating bones undergo gross alterations in contours as well as defects in microarchitecture. The synovial lining of the joint becomes secondarily inflamed from reactions to the cartilage breakdown products and increased strains of a biomechanically stressed joint. Synovial changes in osteoarthritis are less manifest than in rheumatoid arthritis, where inflammation is a primary pathological process in the joint.

When joints are chronically overimpacted or overloaded, as in OA, the cartilage tissues begin to fail. Cartilage loses its normal elasticity, proper opposing surface conformations, and tensile integrity. The later stages of OA are characterized by substantial losses of cartilage substance, deformities of bones and joints (Figure 1). The consequences are loss of mechanical joint efficiency, pain on movements, secondary muscle atrophy, and decreased dynamic support. Thus, chronically decompensated or unbalanced biomechanical stresses can contribute to joint damage and later failure. Gross trauma can cause acute injury and loss of joint integrity, leading to subsequent chronic joint failure, as described above.

Features of Ankylosing Spondylitis

In AS, lesions usually begin in the sacroiliac joints (SIJs) or the lumbosacral vertebrae, at the base of the spine, where the truncal compressional forces are greatest (Figure 2). These joints have minimal, but essential, normal degrees of motion. They are primarily designed to provide stability as well as transfer of forces between the trunk and lower extremities and vice versa. The vertebral lesions in AS are mainly localized at the attachment tissues (enthesopathy), which result in secondary calcifications and immobilization by bony fusion of the connecting vertebrae. Typically, the spinal lesions progress from the base to higher levels, as limitations in movements and fixation occur below. In the SIJs, the pathological lesions progress from initial cartilage narrowing, to localized erosions of this tissue, and to final total cartilage loss, with obliteration of the joints by an intra-articular bony fusion process.

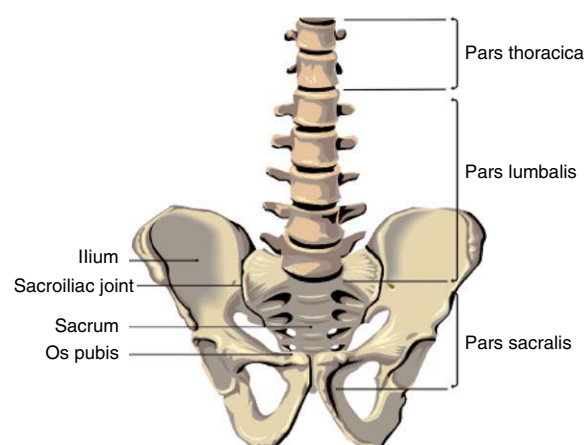


Figure 2 Skeletal structures relevant to ankylosing spondylitis. Skeletal structures of the lower vertebral spine, sacrum, and pelvis are shown. In ankylosing spondylitis, pathological changes typically occur first in the sacroiliac joints and the lumbosacral spine, which bear the greatest truncal loads. Biomechanical stresses are believed to play a major role in development and localization of tissue changes in this disease (see text).

Altered biomechanics is suspected to be important, if not primary, in the physiopathological processes in AS. One concept favored by these authors is that the SIJ lesions develop in a setting of excessive compressional forces across these joints, whereas the vertebral lesions result from excessive tensional stresses on the attachment tissues. In contrast, the popular concept is that inflammation is the primary mechanism at these sites. However, neither the initiating mechanisms nor explanation for the specific tissue localizations are elucidated in this inflammatory theory. If the above-mentioned biomechanical alterations are the main predisposing factors in AS, secondary inflammatory mechanisms could logically occur following chronic microinjury.

All organisms on this planet have evolved and adapted appropriately to resist forces of gravity. Accordingly, all normal tissues and body parts must possess sufficient cellular and structural integrity to resist gravity as well as to maintain an appropriate balance of stability and flexibility mechanisms to perform the vast spectrum of life's daily activities. Structural integrity of the musculoskeletal system is required at micro- and macroanatomical levels. Also, efficient biomechanical connectivity of the myofascial and osteoligamentous tissues is essential. Efficient transfers of physical energies and forces requires dynamically normal organization in the performance of the multitude of body movements and resistance activities. Compromised integrity of either the musculoskeletal structures or their dynamic organization predisposes to particular diseases like OA and AS as well as to accelerated joint damage in RA and other rheumatic conditions.

Possible Stressor Specificities to Particular Pathways

In order to better interpret how stressors can affect particular rheumatic disorders, it is important to understand the main mechanisms or systems alterations, which determine these respective conditions. For example, CNS pathways are believed to be the main mechanisms operating in FMS. They include enhanced emotional and sensory perceptual activation via autonomic and cognitive systems, rather than by alternative inflammatory or biomechanical stress mechanisms. Instead, dysregulations or activations of the NEI and MVE systems are believed to be the primary pathways in RA, leading to enhanced immunological and inflammatory reactivity. Whereas, biomechanical stressors are believed to cause the primary structural and tissue lesions in OA and AS.

Types of stressors that typically aggravate a particular rheumatic disorder may be related to the host's susceptibility to the respective condition. Stressor specificity

relationships to rheumatic disorders might be metaphorically described as a particular glove fitting the owner's hand, at least in part. If stress-disorder specificity does occur, the phenomenon may reflect a potential biological band-width resonance between the stressors and reactor systems (as suggested by Dr. Richard Imrich). Future investigations may permit a partitioning of stresses and their reactive effects according to degrees of stressor intensity and frequency and categories, in relation to the host's adaptabilities versus vulnerabilities, under particular environmental and psychosocial circumstances. For example, experimental and human data suggest that repeated stress exposures under controlled and supportive conditions tend to increase resistance to adverse reactions (favorable conditioning) and vice versa (enhanced vulnerability).

Although certain stressors tend to primarily aggravate one or another rheumatic disease pathway, they can often indirectly or secondarily affect other disease systems. Thus, particular types of stressors are common in related rheumatic disorders, but they are not strictly compartmentalized in their effects. Stressors can have varied degrees of overlapping primary or secondary influences due to complex interactions of the host and environmental circumstances in a particular disorder.

An extensive spectrum of psychological, systemic, physical, or biomechanical loading factors can influence rheumatic disorders. Each condition tends to have its respective pattern of host susceptibilities, somatic developmental influences, psychosocial predispositions, or correlated behaviors. The different disorders show relative reactivity to the varied environmental stimuli, stresses, or challenges. Individuals differ in multiple and complex ways, and are affected differently by the respective heterogeneous stressors. Therefore, only general interpretations can be made regarding the main types of stressors affecting particular rheumatic disorders, and considerably more research is needed in this area.

Profoundly Complex Patterns of Stress Responses

Central Stress Systems

Different stressors of sufficient intensity can elicit relatively patterned responses via multiple effectors. The intensity of stressors and their consequent responses may be either relatively limited or more greatly enhanced. Limited psychophysiological stresses may primarily stimulate CNS and ANS pathways. Enhanced systemic stressors may also significantly activate the hypothalamic-pituitary-adrenal (HPA) axis and related neuroendocrine effectors. These include the two adrenal systems, i.e., the adrenocortical

(corticosteroids) and the adrenomedullary (adrenaline). Other HPA axis components contribute to stress-related responses, which are intended to diminish acute tissue injuries and modulate overstimulation of the immunological and inflammatory responses.

Neurodynamic models of the brain propose that sensory circuits can be tuned, for purposes of selectively sensitizing the cortices to desired or expected inputs. Such dynamics enhance selective attention and intentional learning. However, selective sensitivity and central sensitization in people susceptible to FMS and related conditions could possibly overamplify signals perceived to be aversive. In overly distressed people, such mechanisms would need to be modulated or unlearned, in order to relieve sensory and psychic pain induced by stressful psychosocial stimuli.

Although ANS dysfunction may explain many of the multisystem features of FMS, a fundamental dilemma is whether such activated status is constitutionally predisposed or developmentally acquired. Some data indicate that HPA or ANS dysfunctions may be acquired responses in particular people (possibly innately susceptible) who are exposed to various stressful psychosocial or other life circumstances, particularly during childhood or developmental stages. Carefully controlled prospective studies in family settings will be needed to better assess the dynamic impacts of stress mechanisms on familial versus environmental backgrounds, and their interactions. Such data can clarify the respective roles of these components in FMS, related disorders, and symptomatic patterns of illnesses generally.

Responses to major emotional or physical stresses are physiologically integrated by the HPA axis, the adrenal medulla, and the sympathetic nervous system (SNS). People differ in their susceptibility to physical stresses or psychoneuroendocrine perturbations. Some people have limited versus robust stress reactions, depending upon the integrity of their physiological adaptation. Healthy, well-conditioned, younger people generally react more favorably to systemic stresses (i.e., less allostatic load), than elderly or frail individuals, who may more likely show adverse reactions and consequences.

Variability in Central Nervous System Stress Responses

Individuals vary in cognitive or perceptual sensory sensitization and in their SNS activation, due to genetic, developmental, and psychosocial or behavioral conditioning. Stresses that perturb these CNS pathways can amplify many subjective complaints by central sensitizing mechanisms. Such symptoms include greater intensity and distribution of pain, anxiety, or dysphoria, increased arousal and nonrestorative

sleep, easy fatigue, and generally diminished functional ability. More limited intensities of psychosocial, physical, and environmental stresses may not be sufficient to perturb the majority of individuals in the population, nor be significant factors to predictably alter NEI-MVE interactions in the majority of people. However, people more susceptible to the preceding limited stressors (i.e., those having low stress tolerance) may nevertheless experience significantly enhanced symptomatology or dysfunctions.

In FMS patients, stress susceptibility pathways seem to be related mainly to increased central sensory sensitization mechanisms. Such pathways can enhance perceptual pain and autonomic emotive reactions, but not necessarily affect systemic neuroendocrine or immune mechanisms that predispose to inflammation. Susceptibilities to particular rheumatic disorders seem to be amplified by complementary stressors. In FMS, the more evident stressors seem to primarily impact central sensitization pathways. However, various stressors can secondarily influence sleep and physical conditioning, and can affect other body systems. In addition, emotional or biomechanical factors can also amplify symptomatology in FMS and related disorders.

Stress Reactivity and Systemic Pathways

In subsets of RA patients, stressors are believed to impact the HPA axis and MVE pathways. These systems may possibly show greater stress responsiveness in RA patients due to: (1) innate deficiencies in counterregulatory competence; (2) behavioral exposures or constitutional susceptibility to endothelial microvascular activation, or (3) constitutionally increased immunological or inflammatory activation, which are described further below.

Biomechanical Stresses and Anatomical Structures

Although, OA and AS are strongly affected by biomechanical stresses and gravitational forces, other rheumatic conditions can also be influenced by such factors. In the spinal column, vertebral bodies act mainly as compressional and supportive members, whereas the intervertebral anterior longitudinal ligament acts as a binding or tensional element. In the spine, forces are transmitted at specialized junctional sites, where spinal ligaments attach to the vertebral bodies (i.e., entheses). Such anatomical design allows both stability and mobility via the shared functions of the compressional and tensional osteoligamentous elements. However, such passive osteoligamentous structures alone do not permit sufficient stability to maintain anatomical equilibrium against gravitational forces, let alone provide for stabilized movements

or resistance activities. It is the skeletal and particularly postural muscles that provide the requisite supportive and movement functions in the body via their active contractile mechanisms and mechanical tone.

Complex Integrated Homeostasis and Stress Responses

Complex net interactions occur among the different homeostatic systems and subsystems. In physiological balance, prompt compensation occurs among the systems in response to varied changes. A primary versus secondary perturbation is difficult to distinguish, unless the pathways are known to be predetermined. In chronic disease, e.g., RA, dysregulations in the NEI and MVE systems occur concurrently. Their sequences of perturbations are not known and cannot be determined in cross-sectional analyses. The respective sequential changes must be analyzed over the course of time in prospective studies. Such longitudinal studies can better differentiate initial or primary alterations from secondary pathways, and are a priority objective for future research. Integrative longitudinal analyses offer greater promise to uncover complex mechanisms in chronic diseases with multifactorial determinants, than reductionist analyses which mainly address single-factor pathogenic mechanisms.

Stressors Related to the Physiopathogenesis of Rheumatoid Arthritis

The major example of a rheumatic disorder in this article is RA. It illustrates important principles pertaining to the group of systemic inflammatory diseases and how they may be influenced by various stressors. Considering only the adult onsets of RA in this discussion, the clinical spectrum is decidedly heterogeneous in patterns of manifestations and in degrees of severity. Such disease heterogeneity (phenotypes) may be attributable to a combination of innate genetic predispositions, somatic developmental influences as well as interactions with environmental factors and their respective stressors.

Primary versus Secondary Stress Pathways in RA

A fundamental objective in clinical research is to differentiate primary determinant factors from secondary reactions in the development and course of disease. Both the presymptomatic and clinically active phases of disease need to be critically investigated. Additionally, a priority challenge is to determine normal relations and interactions among the relevant physiological systems in health that become altered in the disease process. Generally, little research

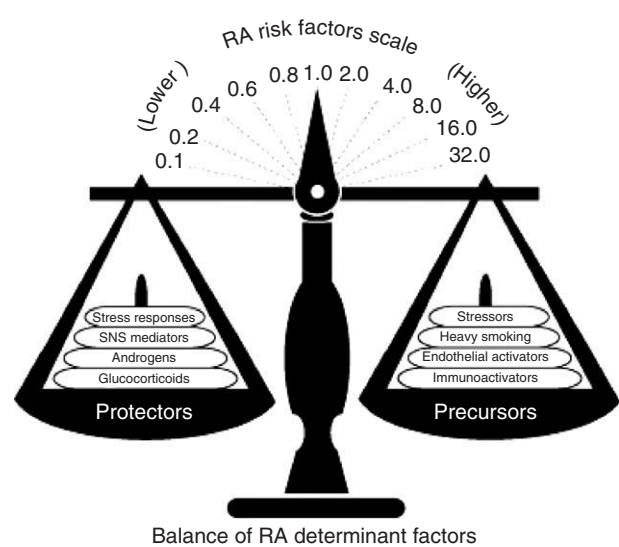


Figure 3 Balance of rheumatoid arthritis determinant factors. The recognized precursor risk factors and counteropposing protective factors for onset of rheumatoid arthritis are schematically shown. Normally, counterregulation occurs between activation of the immunological and microvascular endothelial systems and the anti-inflammatory systems, including glucocorticoids, androgens, and sympathetic nervous system (SNS) mediators. In RA, heavy cigarette smoking and other stressors are believed to dysregulate normal physiological balance and enhance activation of inflammatory pathways (see text).

data are available on such early physiological interactions in clinically healthy subjects or in susceptible people before clinical onset. Most available data pertain to the pathological alterations that occur during active phases of clinical disease, without knowing the physiopathogenetic pathways (Figure 3).

Stressors may Amplify Dysregulations in RA

Under normal circumstances, counterregulatory balance occurs between the NEI systems. Once disease manifests, definite immunological activation and inflammation become evident, which, in turn, stimulate the HPA axis. In active disease, blood levels of the main inflammatory cytokines are significantly elevated (i.e., interleukin (IL)-1 β , IL-6, and tumor necrosis factor (TNF)- α), and they centrally stimulate the HPA axis. The HPA axis products normally act to counter-regulate overactivation of inflammatory pathways (Figure 3). However, such HPA axis actions are suspected to be relatively deficient in some people susceptible to RA, particularly younger women with onsets in their premenopausal ages.

In humans, cortisol is secreted from the adrenal cortex as the main anti-inflammatory product of the HPA axis, whereas in rodents, corticosterone is the main glucocorticoid. Blood levels of these corticosteroids are tightly controlled by negative feedback

mechanisms at the hypothalamic and pituitary (HP) levels. The concept has evolved that relative deficiency or insufficiency of glucocorticoid responsiveness permits inappropriately exaggerated activation of the immunological and microvascular endothelial systems. Increased predisposition to inflammatory disorders may also result from natural susceptibilities via other host immunological mechanisms or added predispositions from varied stressors.

Firm data are not available on mechanisms or pathways whereby particular stressors can alter the normal psychoneuroendocrine physiology and influence susceptibility to RA. However, one study of 2490 Vietnam Theater Veterans diagnosed as having PTSD showed a fivefold increased likelihood to have postwar RA. That study and other data may indicate that people with relatively lower HPA axis and SNS stress responsiveness may be more susceptible to both PTSD and autoimmune disorders, like RA. Alternatively, stressors that predispose to PTSD may suppress HPA axis responsiveness via inhibition of corticotropin-releasing hormone (CRH). In turn, counterregulation of immunological activation could be diminished. Research on presymptomatic risk factors for onset of RA indicate that increased predisposition is related to evidences of: (1) increased immunological activation, (2) physiologically lower adrenal cortical reserve in younger women, and (3) history of heavy cigarette smoking in females and males, all of which may impact stress responsiveness of the NEI and MVE systems.

An Evolving Research Challenge in Rheumatoid Arthritis

In growth, maturation, and later aging into the senescent period, marked changes normally occur in the NEI and MVE systems. Attempts have been made to correlate such important physiological changes with risks of developing RA in the population. In women, incident risks of RA onset increase rather regularly with aging, from adolescence to the people in their seventies. However, in males aged 15–50 years, the risks are level and low, about one-fifth as high as in females. Such differential susceptibility patterns indicate that males are relatively protected during their younger adult and more masculine years. From such data, adrenal and gonadal androgens have been implicated as being protective of RA, perhaps by exerting greater control of immunological activation.

Estrogens control endothelial cell activation more effectively than androgens. Premenopausal women have significantly lower natural risk of coronary artery disease (CAD) and myocardial infarction (MI) than male counterparts. Sufficient intensity and duration of cigarette smoking exposures are stressful

vasculotoxins that increase risks of CAD and RA in both sexes. The smoking effect on CAD is magnified in women by use of oral contraceptives (OCs), which provide supraphysiological or pharmacological doses of estrogens. However, OCs alone have not been proven to amplify risks of RA. Also, pregnancy typically improves RA activity, which argues against female sex hormones being a primary risk factor.

Pharmacological doses of estrogens have been shown to increase risks of CAD and MI in both females and males. However, the mechanism is likely to be due to increased thrombogenic risks, rather than by microvascular endothelial or inflammatory pathways. Interpretation of such clinical-epidemiological risk patterns are difficult. Nevertheless, tentative hypotheses may be drawn.

First, heavy cigarette smoking is a stressor which increases risks for a large spectrum of microvascular-related disorders, including RA, and equally so in males and females.

Second, males have significantly lower natural risk of RA during their younger adult ages, suggesting greater control of immunoactivation during those years of highest androgenicity. A significant association of PTSD and RA is reported in Vietnam Theater Veterans, suggesting that stress or diminished tolerance thereto may be a susceptibility factor. Typically, stress reactions result in decreased androgen levels, which could affect susceptibility to RA in males.

Third, estrogen levels *per se* are not correlated with increased RA risks, because younger and pregnant women have considerably lower incidence than postmenopausal women. Also, OCs have not been proven to increase RA risk. Thrombogenic pathways, enhanced by pharmacological doses of estrogens, do not seem to be a primary mechanism in RA, in contrast to immunoactivation and MVE perturbations.

Summary

Interfaces of the immunological and microvascular endothelial systems seem to be primarily incriminated in the inflammatory pathways in RA. Their activation may be amplified permissively by relatively insufficient HPA axis counterregulatory controls. Sufficient cortisol responsiveness is needed to control capillary permeability and transmigration of inflammatory cells into the interstitium. This potent hormone is also required to control overactivation of the immunological system and inflammatory cytokine production, which are key mediators in the RA process.

Stressors and environmental activators may be suspected as being contributory co-factors in the complex web of causation of RA (Figure 3). Heavy cigarette smoking has known microvascular toxic effects. Also, stressful life circumstances likely compromise

normal HPA axis counterregulatory control functions and diminish androgen levels, particularly in males. The glucocorticoids, cortisol in humans and corticosterone in rodents, are the most potent anti-inflammatory products of the HPA system and in the stress response. However, many additional steroid hormones are secreted by the adrenal cortex, in addition to those products of the gonads. Also, catecholamines are secreted by the adrenal medulla and ANS, which are normally immunosuppressive. Besides glucocorticoids, these additional products of the HPA axis, gonads, and ANS can also modulate activation of the immunological–endothelial systems and control their interactions. Further research is needed on the complex relations of stress reactions to risks and course of RA and other rheumatic disorders.

Acknowledgments

Support for this project was provided by the Department of Medicine, University of Illinois College of Medicine at Peoria, and a grant from the MTM Foundation.

Further Reading

- Boscarino, J. A. (2004). Posttraumatic stress disorder and physical illness. *Annals of the New York Academy of Sciences* 1032, 141–153.
- Cannon, W. B. (1929). Organization for physiological homeostasis. *Physiological Reviews* 9, 399–431.
- Felson, D. T. and Zhang, Y. (1998). An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. *Arthritis and Rheumatism* 41, 1343–1355.
- Freeman, W. J. (2003). Neurodynamic models of brain in psychiatry. *Neuropsychopharmacology* 28, S54–S63.
- Masi, A. T. (2000). Hormonal and immunologic risk factors for the development of rheumatoid arthritis: an integrative physiopathogenetic perspective. *Rheumatic Disease Clinics of North America* 2, 775–803.
- Masi, A. T. and Aldag, J. C. (2005). Integrated neuroendocrine immune risk factors in relation to rheumatoid arthritis: should rheumatologists now adopt a model of a multiyear, presymptomatic phase? *Scandinavian Journal of Rheumatology* 34, 342–352.
- Masi, A. T. and Chang, H. J. (1999). Cigarette smoking and other acquired risk factors for rheumatoid arthritis. In: Kaufman, L. D. & Varga, J. (eds.) *Rheumatic diseases and the environment*, pp. 111–127. New York: Chapman & Hall.
- Masi, A. T. and Walsh, E. G. (2003). Ankylosing spondylitis: integrated clinical and physiological perspectives. *Clinical and Experimental Rheumatology* 21, 1–8.
- Masi, A. T., Aldag, J. C. and Jacobs, J. W. G. (2005). Rheumatoid arthritis: neuroendocrine immune integrated physiopathogenetic perspectives and therapy. *Rheumatic Disease Clinics of North America* 31, 131–160.
- Masi, A. T., DaSilva, J. A. P. and Cutolo, M. (1996). Perturbations of hypothalamic-pituitary-gonadal (HPG) axis and adrenal androgen (AA) functions in rheumatoid arthritis. *Baillière's Clinical Rheumatology* 10, 295–332.
- Masi, A. T., White, K. P. and Pilcher, J. J. (2002). Person-centered approach to care, teaching, and research in fibromyalgia syndrome: justification from biopsychosocial perspectives in populations. *Seminars in Arthritis and Rheumatism* 32, 71–93.
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* 338, 171–179.
- Robin, E. D. (1979). Claude Bernard. Pioneer of regulatory biology. *JAMA* 242, 1283–1284.
- Selye, H. (1936). Thymus and adrenals in the response of the organism to injuries and intoxications. *British Journal of Experimental Pathology* 17, 234–240.
- Selye, H. (1946). The general adaptation syndrome and the diseases of adaption. *Journal of Clinical Endocrinology and Metabolism* 6, 117–230.
- Straub, R. H. and Besedovsky, H. O. (2003). Integrated evolutionary, immunological, and neuroendocrine framework for the pathogenesis of chronic disabling inflammatory diseases. *FASEB Journal* 17, 2176–2183.
- Yunus, M. B. (2005). The concept of central sensitivity syndromes. In: Wallace, D. J. & Clauw, D. J. (eds.) *Fibromyalgia and other central pain syndromes*, pp. 29–44. Philadelphia: Lippincott Williams & Wilkins.
- Walker, J. G., Littlejohn, G. O. and McMurray, N. E., et al (1999). Stress system response and rheumatoid arthritis: a multilevel approach. *Rheumatology* 38, 1050–1057.

Ribosomes See: Protein Synthesis; Chaperone Proteins and Chaperonopathies; Heat Resistance; Heat Shock Response; Overview; Chaperonopathies; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles.

Risk Factors for Stress See: Life Events Scale; Life Events and Health.

S

Salivary Cortisol

C Kirschbaum and D H Hellhammer

Technical University of Dresden, Dresden, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 379–383, © 2000, Elsevier Inc.

Methodological Issues

Salivary Cortisol Stress Responses

Field Studies and Ambulatory Assessments

Psychoendocrine Stress Research in Infants and Children

The hypothalamic-pituitary-adrenal (HPA) axis responds rapidly and rather specifically to a wide range of environmental and internal demands often referred to as stress. It is thought that the HPA response to stress plays a pivotal role in the organism's attempt to maintain function through change, as expressed in the allostasis model. Although the HPA hormones corticotropin releasing hormone (CRH) and adrenocorticotrophic hormone (ACTH) may exert additional adaptive effects, accumulating evidence suggests that cortisol is the main hormone responsible for allostatic stress responses.

Following stimulation by ACTH, cortisol is synthesized and secreted from the adrenal glands and released into the circulating blood, where it is bound rapidly to carriers such as corticosteroid-binding globulin (CBG), albumin, and erythrocytes. Only a small fraction (2–15%) of cortisol released remains unbound, or free. According to the free hormone concept, it is only this hormone fraction that brings about the multitude of genomic cortisol effects in peripheral tissues and in the brain. Although in blood both bound and free cortisol can be measured, only free cortisol appears in saliva. Hence, the measurement of salivary cortisol provides an index of the biologically active fraction of this steroid hormone.

Methodological Issues

For decades, research on the acute and chronic effects of stress has employed cortisol levels as an index of the individual response to stress. Due to methodological reasons, this research was slowed by the need for multiple blood samples to measure cortisol levels repeatedly. Blood sampling is not only labor-intensive (and expensive) but sometimes also counterproductive in stress research because the procedure of venipuncture itself can lead to significant HPA activation. The advent of advanced biochemical assays capable of measuring cortisol reliably in the lower nanomolar range, however, helped to overcome this problem.

Since the mid-1980s, the assessment of free cortisol in saliva has become an increasingly important tool in stress research because it has a number of advantages over the measurement of cortisol in blood or urine. Of prime interest to the stress researcher is the noninvasiveness of sampling and the ability to obtain samples at short intervals from subjects over the entire life span without raising ethical problems associated with venipuncture. The measurement of cortisol levels in saliva is the method of choice for field studies or ambulatory assessments in the natural environment of the subjects because medical personnel or laboratory facilities are not required for collecting and storing the samples. Salivary cortisol is not measured only in humans, it also serves as an endocrine correlate of stress in animals, including pigs, monkeys, elephants, sheep, tree shrews, guinea pigs, and goats.

Sampling, Storage, and Biochemical Analysis

Samples are usually obtained by the subjects or patients; infants and young children (<2 years old) require additional help from parents or hospital staff. Subjects either salivate directly into small containers or use devices such as the Salivette, which is currently the most frequently employed sampling method.

Stimulation of saliva flow by gently chewing on the device or by adding a few crystals of sugar-sweetened Kool-Aid (for use with infants and small children) ensures adequate sample volume. Because most immunoassays available today provide false high values with low pH samples, saliva should not be sampled after stimulation with citric acid. The storage and mailing of samples are uncomplicated. Because cortisol is rather stable in saliva, samples can be stored at room temperature for at least 4 weeks without a significant reduction in cortisol concentration or mailed to a laboratory at ambient temperature, even overseas. However, in order to keep samples over prolonged periods and to avoid molding, it is recommended that samples be stored at -20°C or lower. For the biochemical analysis of salivary cortisol, several methods are now available that allow reliable measurement of the steroid hormone. Immunoassays with radioactive and other tracers are the method of choice for most laboratories. An increasing number of commercial assay kits designed for use with serum or plasma now also provide instructions for use with saliva. In a few instances, the use of alternative analytical techniques may be required. For example, the simultaneous detection of several steroid hormones and their metabolites in saliva can be achieved with high-pressure liquid chromatography (HPLC). Also, for processing samples that contain high levels of compounds (e.g., prednisone) that cross-react significantly with most antibodies raised against cortisol for use with immunoassays, HPLC may be the adequate analytical technology. Gas chromatography-mass spectrometry is used only for validating cortisol results obtained with other assay methods.

Comparison with Cortisol Levels in Blood

Because cortisol is a small (molecular weight, MW, 362) and highly lipid-soluble molecule, the unbound hormone can pass easily through the lipid-bilayer membranes of nucleated cells. This allows free cortisol to appear in all bodily fluids, including blood, cerebral spinal fluid, urine, sweat, semen, and saliva. Cortisol bound to carriers is usually excluded from these body compartments.

Most important for stress research, cortisol levels measured in saliva agree very well with the amount of free cortisol in blood. Correlations between salivary and unbound blood cortisol levels usually explain more than 80% of the total variance observed ($r \geq 0.90$). This high agreement is due to the fact that cortisol enters the cell and the oral cavity mainly by passive diffusion. It is therefore independent of transport mechanisms and saliva flow rate, in contrast to other components also found in saliva

(e.g., immunoglobulin A and catecholamines). Absolute levels of free cortisol are lower in saliva due to a relative abundance of the cortisol-metabolizing enzyme 11- β hydroxysteroid dehydrogenase.

Circadian Rhythm and Normal Values

As in blood, salivary cortisol levels follow a robust circadian rhythm that has been studied only in awake subjects so far. Peak levels are usually found in the early morning hours, with decreasing values thereafter. Several secretory episodes can be detected throughout the day, even in the absence of psychological or physical stress.

Stimulated by an individual's waking up in the morning, there is a 50–100% increase in cortisol levels, peaking about 30 min after awakening. This response shows remarkable stability over weeks and months in the same individual and has been found to relate to prolonged stress exposure and the feeling of burnout. After the morning peak, cortisol levels decrease steadily until a second major secretory episode occurs in response to larger meals at lunch time. As with the wakening response, meal-related peaks appear to be more pronounced (approximately 1.5-fold) in free cortisol compared to total blood cortisol. During afternoon and evening hours, saliva cortisol continues to decrease with no further major secretory episode. [Figure 1](#) depicts the circadian rhythm of cortisol in saliva representing normal values typically observed in a healthy population.

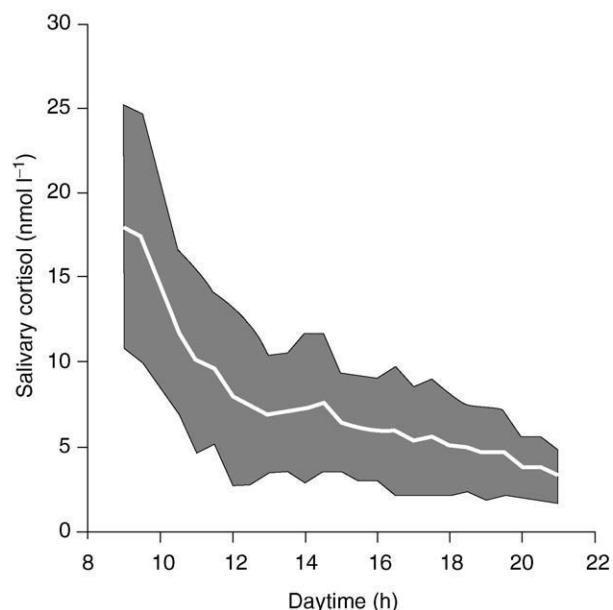


Figure 1 Normal values of salivary cortisol during the day (white line: mean levels; shaded area: means ± 1 SD).

Salivary Cortisol Stress Responses

Response Kinetic

In contrast to other neuronal and endocrine response circuits, it takes several minutes before an HPA stress response becomes visible in the periphery. After the onset of a psychosocial stressor or strenuous exercise, ACTH levels start to rise within less than 5 min and show peak values shortly after the cessation of stress (<5 min). The cortisol response lags by 5–20 min, with maximal cortisol levels 10–30 min after stress cessation. Although cortisol exerts a strong feedback action at the level of the pituitary, hypothalamus, and hippocampus, prolonged periods of stress can lead to cortisol secretion over several hours. For example, in marathon runners, salivary cortisol was observed to increase monophasically for more than 4 h, peaking approximately 30 min after the run was finished.

The transfer of cortisol from plasma to saliva occurs quickly. Within less than a minute, cortisol injected intravenously appears in saliva, and the peak concentration in saliva lags by less than 2–3 min compared to levels measured in blood. Thus, it is not surprising to find similar stress response kinetics in saliva and in blood. For example, total plasma cortisol and salivary cortisol levels rise in parallel in response to a laboratory stress task that consists mainly of an anticipation period (3 min), public speaking (5 min), and mental arithmetic (5 min) in front of an audience. Following the cessation of stress, however, free cortisol tends to increase further for another 15–20 min, whereas total cortisol levels

are unchanged or start to decline (Figure 2). A similar response pattern was observed in plasma after meals, bicycle ergometry, and injection of ACTH(1–24) and during sleep.

Discrepancy between Salivary Cortisol and Total Plasma Stress Responses

Although salivary cortisol levels agree well with plasma-free cortisol, they show only moderate correlations with total cortisol levels under certain circumstances. Although the adrenal cortex was stimulated by ACTH and secreted significant amounts of cortisol, subjects may show blunted or completely absent cortisol responses as measured in saliva. This phenomenon can be observed in women using oral contraceptives (OC) (Figure 3). The use of oral contraceptives leads to an increase in CBG levels and hence to a larger number of binding sites for cortisol in the circulating blood. Therefore, more cortisol released from the adrenals is bound by CBG in these women than in non-OC users, and lower levels of the biologically active free fraction appear in saliva (as well as in target tissues). Similar discrepancies between total blood and salivary cortisol levels should be expected in all cases of increased or decreased CBG production.

Field Studies and Ambulatory Assessments

One of the major advantages of salivary cortisol is that samples can be obtained in the field, that is, in the natural (social) environment of the subjects/patients

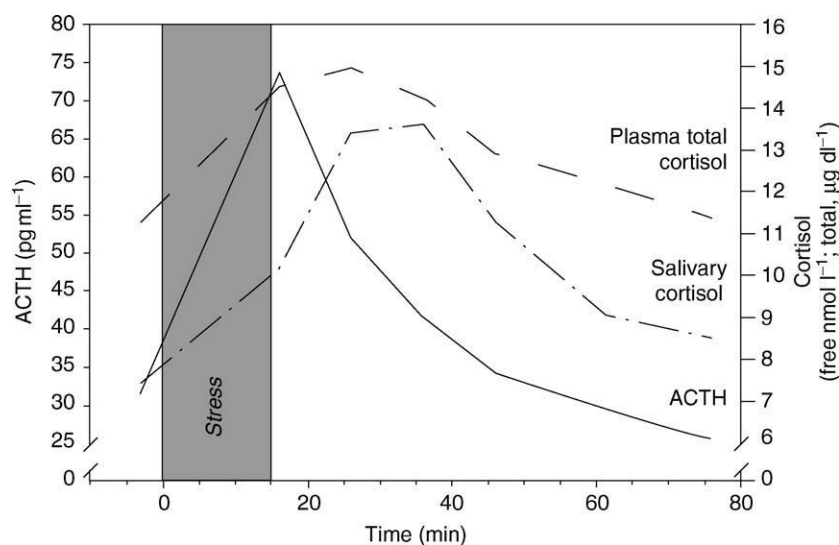


Figure 2 Mean ACTH, total plasma cortisol, and salivary cortisol responses to a brief psychosocial stress test (Trier social stress test). Note that there are different units of hormone concentrations for total plasma and salivary cortisol.

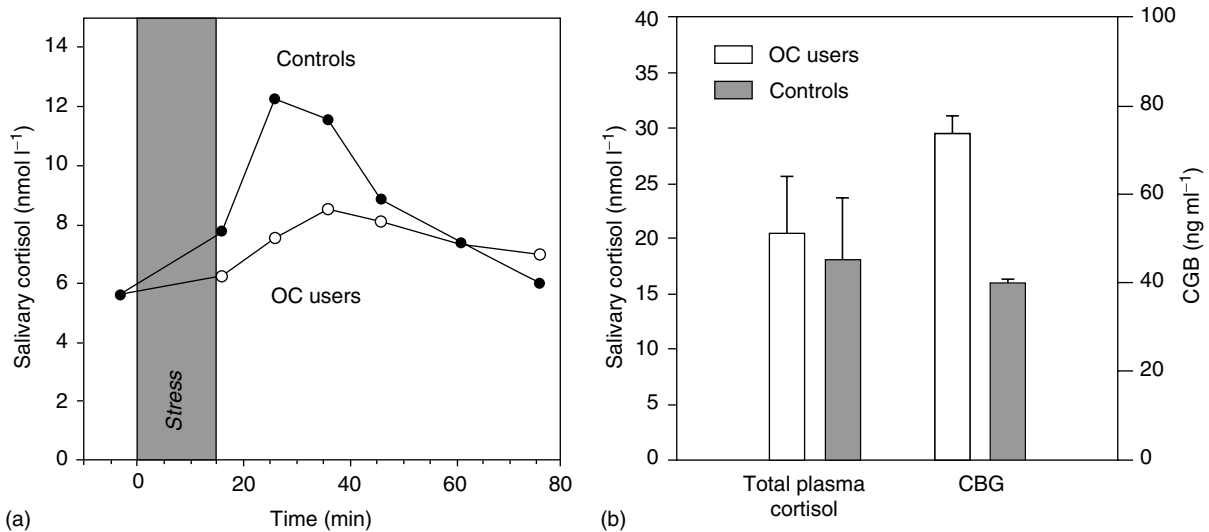


Figure 3 Cortisol responses to stress in women using oral contraceptives (OC) and nonusers. a, Salivary cortisol levels; b, Total plasma cortisol and corticosteroid-binding globulin (CGB) levels. The blunted free cortisol response pattern in OC users is not due to lower cortisol secretion but rather a consequence of enhanced CGB levels. AUC, area under the curve.

as well as in special settings outside a psychological or medical laboratory. This advantage has allowed researchers to perform a number of studies that would have been very difficult or even impossible with blood sampling. With saliva, large cross-sectional studies can be performed by mailing sampling devices and instructions to the subjects, who return the material by regular mail to the laboratory. Populations of 500–1000 subjects or more can be investigated easily by this method.

Of special interest to the stress researcher is the ease of obtaining cortisol readings from subjects repeatedly before and after exposure to real-life stress situations. For example, the impact of the first three parachute jumps on acute HPA responses and subsequent pharmacological tests was studied with saliva samples obtained at the airfield at 20-min intervals over 10 h. In addition, the same subjects collected saliva samples at identical times on a control day to relate the stress-induced changes to the individuals' normal circadian rhythms. Furthermore, ambulatory (or momentary) assessments of cortisol responses to daily stress can be performed independently of the researcher. Using electronic devices such as preprogrammed wristwatches and pocket-size (computer) booklets, researchers can beep study subjects several times a day and request a report on current, past, or anticipated stressors. Twenty minutes later, they beep the subjects again and ask them to take a saliva sample. Several studies with momentary assessments of stress and salivary cortisol have been published supporting the view that even minor

changes in negative affect can result in increased cortisol levels, whereas a positive affect has the opposite effect.

Psychoendocrine Stress Research in Infants and Children

For ethical reasons, invasive measures are usually prohibited for use with healthy infants or children. Circumventing these difficulties, the availability of salivary cortisol has stimulated research into the psychoendocrinology of stress in infants and children significantly. A rapidly growing literature shows that the human HPA axis begins to respond to psychosocial or physical stimulation with an activation of the axis very early in life. In combination with appropriately modified protocols, the acute response to experimental stress can thus be monitored in children and compared directly with adult response patterns, which will eventually lead to a more detailed understanding of the ontogenesis of the human HPA stress response.

Acknowledgments

This work was supported by a grant from the Deutsche Forschungsgemeinschaft (Ki 537/6-1).

See Also the Following Articles

Allostasis and Allostatic Load; Hypothalamic-Pituitary-Adrenal; Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress.

Further Reading

- Brooks, F. S. and Brooks, R. V. (1984). Cortisol and cortisone in saliva. In: Read, G. F., Riad-Fahmy, D., Walker, R. F. & Griffiths, K. (eds.) *Immunoassays of steroids in saliva*, pp. 322–326. Cardiff, UK: Alpha Omega.
- Buske-Kirschbaum, A., Jobst, S., Wustmans, A., et al. (1997). Attenuated cortisol response to psychosocial stress in children with atopic dermatitis. *Psychosomatic Medicine* 59, 419–426.
- Cook, N. J., Read, G. F., Walker, R. F., Harris, B. and Riad-Fahmy, D. (1992). Salivary cortisol and testosterone as markers of stress in normal subjects in abnormal situations. In: Kirschbaum, C., Read, G. F. & Hellhammer, D. H. (eds.) *Assessment of hormones and drugs in saliva in bio-behavioral research*, pp. 147–162. Seattle, WA: Hogrefe & Huber.
- Deinzer, R., Kirschbaum, C., Gresele, C., et al. (1997). Adrenocortical responses to repeated parachute jumping and subsequent h-CRH challenge in inexperienced healthy subjects. *Physiology & Behavior* 61, 507–511.
- Ekins, R. (1990). Measurement of free hormones in blood. *Endocrine Review* 11, 5–46.
- Follenius, M. and Brandenberger, G. (1986). Plasma free cortisol during secretory episodes. *Journal of Clinical Endocrinology and Metabolism* 62, 609–612.
- Gunnar, M. R. (1989). Studies of the human infant's adrenocortical response to potentially stressful events. *New Directions in Child Development*, 3–18.
- Kirschbaum, C. and Hellhammer, D. H. (1989). Salivary cortisol in psychobiological research: an overview. *Neuropsychobiology* 22, 150–169.
- Kirschbaum, C., Kudielka, B. M., Gaab, J., et al. (1999). The impact of gender, menstrual cycle phase and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis. *Psychosomatic Medicine* 67, 154–162.
- Kirschbaum, C., Pirke, K.-M. and Hellhammer, D. H. (1993). The “Trier Social Stress Test” – a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology* 28, 76–81.
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* 338, 171–179.
- Pruessner, J., Wolf, O. T., Hellhammer, D. H., et al. (1997). Free cortisol levels after awakening: a reliable biological marker for the assessment of adrenocortical activity. *Life Sciences* 61, 2539–2549.
- Spangler, G. and Scheubeck, R. (1993). Behavioral organization in newborns and its relation to adrenocortical and cardiac activity. *Child Development* 64, 622–633.
- van Eck, M., Berkhof, H., Nicolson, N., et al. (1996). The effects of perceived stress, traits, mood states, and stressful daily events on salivary cortisol. *Psychosomatic Medicine* 58, 447–458.
- Walker, R. F., Joyce, B. G., Dyas, J. and Riad-Fahmy, D. (1984). Salivary cortisol. I: Monitoring changes in normal adrenal activity. In: Read, G. F., Riad-Fahmy, D., Walker, R. F. & Griffiths, K. (eds.) *Immunoassays of steroids in saliva*, pp. 308–316. Cardiff, UK: Alpha Omega.

Salt Appetite

R S Weisinger

La Trobe University, Victoria, Australia

N Chen

University of Melbourne, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R S Weisinger and D A Denton, volume 3, pp 384–392,

© 2000, Elsevier Inc.

Introduction

Environmental Stress and Salt (Sodium Chloride) Appetite

Hormones of the Hypothalamic-Pituitary-Adrenal Axis

Associated with Stress and Salt Appetite

Brain Mechanisms Involved in Stress-Induced Salt Appetite

Conclusion

Glossary

<i>Homeostasis</i>	The maintenance of a relatively constant internal environment.
<i>Hypertension</i>	Abnormally or excessively high blood (arterial) pressure.
<i>Hypothalamic-pituitary-adrenal (HPA) axis</i>	The integrated association between the hypothalamus (an area of the brain that integrates the various automatic body reactions), the pituitary gland (a small gland at the base of the brain that controls hormone release from the adrenal gland), and the adrenal gland.
<i>Ingestive behavior</i>	Taking food or fluid into the body as nourishment.
<i>Lesion</i>	The removal of or irrevocable damage to a localized region of body tissue.

Sympathetic nervous system

One of the two divisions of the autonomic nervous system. It innervates involuntary processes, which are largely beneath consciousness (e.g., the actions of the visceral organs, heart, blood vessels, glands, and smooth muscle). The activation of the sympathetic nervous system prepares the body for fight or flight and works antagonistically with the second autonomic division, the parasympathetic nervous system.

Introduction

Stress is often seen as a major contributor to disease, with specific effects on several of the body's physiological and behavioral systems. Its effect on the cardiovascular system may result in increased mortality. In addition, stress or the hormonal consequences of stress cause several changes in ingestive behavior. Evidence suggests that stress or the hormones of stress can directly influence the intake of salt (sodium chloride). An increased intake of salt is widely accepted to be a factor in the causation of hypertension and, in some instances, it magnifies the cardiovascular effects of stress. Evidence regarding the brain mechanisms involved in the changes in salt appetite caused by stress and the hormones of stress are described.

Environmental Stress and Salt (Sodium Chloride) Appetite

Environmental stress can result in increased salt appetite, as has been shown in several ways. One of the first demonstrations was made while attempting to fit jackets, designed for carrying small infusion pumps, on wild rabbits (see [Figure 1](#)). This minor restraint on free movement caused the rabbits to become agitated, and, over the course of several days, a spectacular increase in salt appetite was observed – their intake of a concentrated salt solution increased by over 15-fold. This intake of salt approximated a daily turnover of 30% of extracellular fluid sodium, sodium in the tissue fluid outside the cells. In similar experiments in which restraints were placed on mice (see [Figure 2](#)), an increase in salt intake was also observed. This salt intake was eliminated by the subcutaneous administration of the drug captopril; this drug prevents the formation of the peptide hormone angiotensin II. Angiotensin II, produced peripherally (i.e., in the circulation) and in the brain, is involved in both sodium and cardiovascular homeostasis, and it is one of the hormones known to be elevated in stressful situations. Increased levels of angiotensin II are

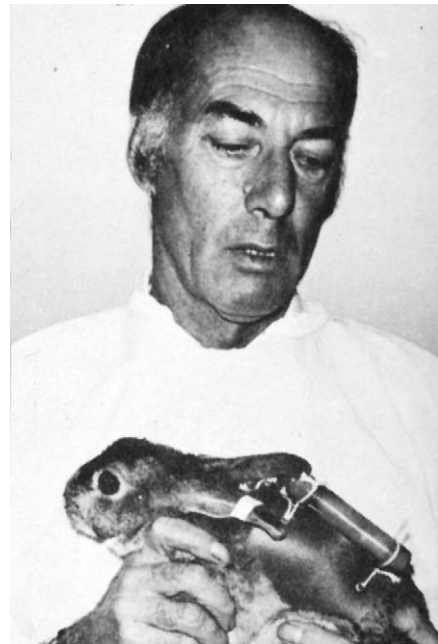


Figure 1 Jack Nelson holding a wild rabbit in a jacket designed to carry infusion pumps.



Figure 2 Restrained mouse on running wheel.

also thought to explain the increased salt intake of mice kept in a cold environment for several hours a day. It is well known that rats from one colony become over-reactive when introduced into another colony. An increased salt appetite has been documented in these stressed animals.

The sympathetic nervous system is responsible for the fight or flight response in animals confronted with stressful situations. Drugs such as clonidine or reserpine, which interfere with the actions of the sympathetic nervous system, block the characteristic stress-induced increase in salt intake. This suggests that the increase in salt intake involves the increased activity of the sympathetic nervous system.



Figure 3 Two adult male Sprague-Dawley rats in conflict during crowding.

Crowding, an imposed social stress, was also found to increase salt intake in rats. Following a control period during which rats were housed alone in a small individual cage, two rats were placed together into a different cage (see [Figure 3](#)). The intake of concentrated salt solution increased over several weeks.

In summary, animals maintained in stressful situations have been shown to manifest an increase in salt intake. This increased salt intake is mediated by increased levels of the hormones associated with stress, namely, the hormones of the hypothalamic-pituitary-adrenal (HPA) axis, primarily adrenocorticotrophic hormone (ACTH) and the mineralocorticoid/glucocorticoid hormones secreted by the adrenal gland. Angiotensin II, another hormone whose levels are increased in stressful situations, can stimulate the secretion of ACTH and the adrenal hormones. Interfering with the actions of angiotensin II can block salt appetite induced by stress. At present, the contribution of the sympathetic nervous system to salt appetite induced by stress seems paradoxical given that the activation of this system causes increased blood pressure.

Hormones of the Hypothalamic-Pituitary-Adrenal Axis Associated with Stress and Salt Appetite

It is well documented that hormones of the HPA axis (see [Figure 4](#)) can cause salt appetite. These include corticotropin releasing hormone (CRH) from the paraventricular nucleus of the hypothalamus, ACTH from the pituitary gland at the base of the brain, and the adrenocortical hormones (often referred to as steroid hormones), including the mineralocorticoid hormones, aldosterone and deoxycorticosterone and the glucocorticoid hormones, corticosterone and cortisol. The name of the former suggests their

important role in the retention of minerals (i.e., salt) in the body and the name of the latter reflects their role in glucose metabolism. CRH and the more recently discovered urocortin (UCN) are thought to have a major physiological role in the response to stress, including the stimulation of ACTH secretion. ACTH and angiotensin II are important in stimulating the release of the mineralocorticoid hormones from the adrenal cortex.

Adrenocortical Hormones

There are at least 20 known steroid hormones with numerous functions as well as catecholamines produced and secreted by the adrenal gland. Work, started over 50 years ago, has demonstrated that the administration of mineralocorticoid hormones (aldosterone or deoxycorticosterone) causes salt appetite. Mineralocorticoid-induced salt appetite has been shown in rats, mice, rabbits, hamsters, and sheep. Sheep treated with deoxycorticosterone shift their preference from a low-salt diet to a high-salt diet, and the increased intake of salty food is associated with an increase in both water intake and blood pressure. In addition, infusion of aldosterone into the brain causes increased blood pressure in rats.

Mineralocorticoid hormones may also influence salt appetite in the rat by its interaction with angiotensin II. The synergy between the mineralocorticoid hormones and angiotensin II may be due to the ability of the mineralocorticoid hormones to increase the number or sensitivity of angiotensin II receptors. Evidence contrary to the proposal that salt appetite is caused by a synergy between mineralocorticoids and brain angiotensin II has been reported in sheep. Data regarding the synergy hypothesis are inconclusive in the pigeon and rabbit.

The administration of either cortisol or corticosterone increased salt appetite in rabbits and mice. The concurrent administration of both cortisol and corticosterone increased appetite by more than the sum of the individual intakes. This result suggests a synergy between the two glucocorticoid hormones. In humans, the administration of cortisol increased blood pressure but did not alter salt preference.

The glucocorticoid hormones may contribute to salt appetite by potentiating the appetite caused by mineralocorticoid treatment. The administration of corticosterone or dexamethasone, a potent synthetic glucocorticoid hormone, for example, enhances salt appetite induced by aldosterone or deoxycorticosterone in rats. A similar influence has been reported in mice.

In experiments in sheep, the intravenous infusion of a combination of steroids, a treatment that mimics

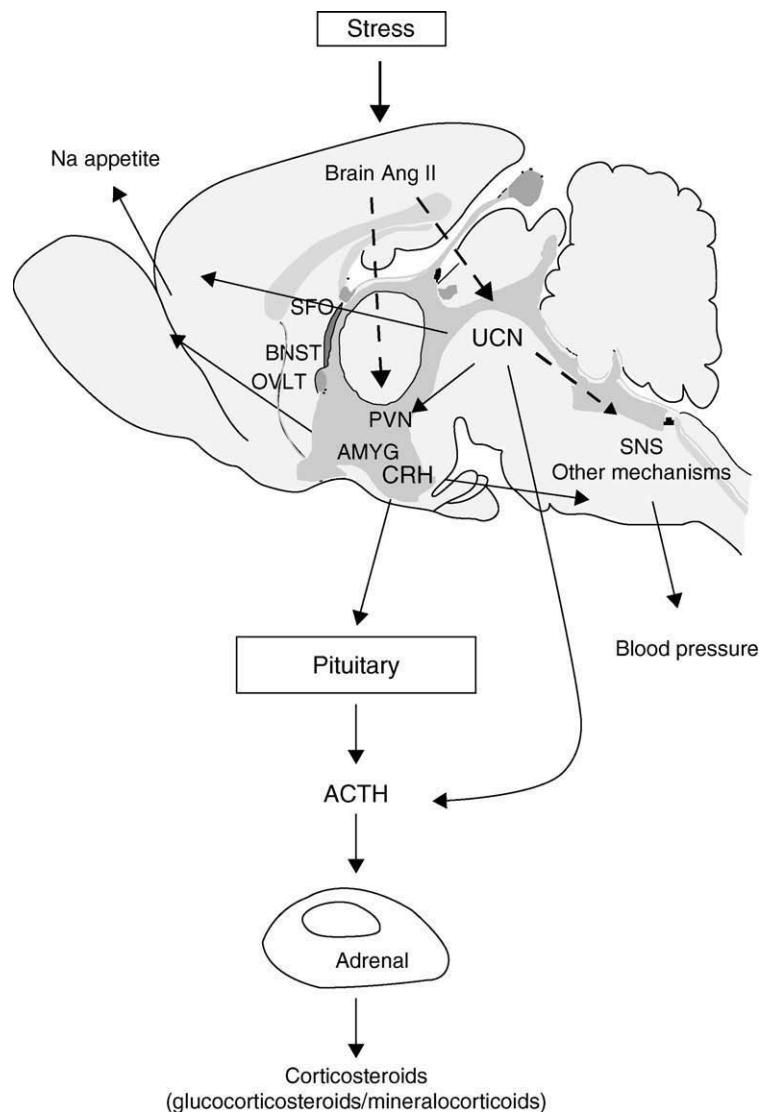


Figure 4 Some of the mechanisms affecting salt appetite and blood pressure. ACTH, adrenocorticotrophic hormone; AMYG, amygdala (a laterally situated structure shown in central perspective); Ang II, angiotensin II; BNST, bed nucleus of the stria terminalis; CRH, corticotropin releasing hormone; OVLT, organum vasculosum of the lamina terminalis; PVN, paraventricular nucleus of the hypothalamus; SFO, subfornical organ; SNS, sympathetic nervous system; UCN, urocortin.

the effects of ACTH administration, increased salt appetite by an amount similar to that observed in sheep treated with ACTH. Blood pressure was also increased. The steroids infused included both mineralocorticoid (aldosterone and deoxycorticosterone) and glucocorticoid (corticosterone, cortisol, and 11-deoxycortisol) hormones and produced blood levels of these hormones similar to those observed during ACTH administration. Blood pressure was also elevated in this instance. However, the infusion of the steroids into the arterial blood directly supplying the brain or directly into the brain did not increase blood pressure. This suggests that, in contrast to the influence of aldosterone on blood pressure in the rat,

the steroid-induced increase in blood pressure in sheep is not due to a central effect of the steroids.

In summary, an increased intake of salt is produced by the administration of mineralocorticoid and/or glucocorticoid hormones. Several synergies between adrenal hormones in regard to the stimulation of salt appetite have been reported. Also, mineralocorticoid hormones may work in conjunction with angiotensin II to stimulate salt intake.

Adrenocorticotrophic Hormone

ACTH was isolated from mammalian hypothalamus in the 1940s–1960s; it has the following properties:

- It is a 39-amino-acid peptide, derived from pro-opiomelanocortin and found primarily in the pituitary gland but also found in the brain (e.g., the arcuate nucleus).
- Its levels are elevated during stress.
- It controls the adrenal corticosteroid responses to stress.

The administration of ACTH causes a large increase in salt intake in several species such as sheep, rabbits, rats, and mice. When rats were offered a choice among concentrated solutions of sodium chloride, potassium chloride, calcium chloride, and magnesium chloride, their appetite was primarily for the sodium chloride solution. The intake of water was also increased. By the fifth day of treatment, the salt intake (and output) was equivalent to the daily turnover of total body salt. In mice, the increase in salt intake caused by ACTH was not altered by treatment with captopril. The observation that the stress-induced (but not ACTH-induced) salt intake was blocked by captopril suggests that the stress-induced production of angiotensin II precedes the increase in ACTH; that is, in stressful situations, angiotensin II has an important role in the stimulation of ACTH secretion.

Angiotensin II, administered into the circulation or directly into the brain, can stimulate the secretion of ACTH, possibly by stimulating the secretion of CRH from the paraventricular nucleus of the hypothalamus. The increase in water intake caused by ACTH has been shown to occur in rats and mice that were offered only water to drink. Results showed that ACTH caused an increase in the extracellular fluid sodium concentration, which caused the enhanced water intake. The administration of ACTH into the circulation causes a large and sustained increase in blood pressure in sheep, rats, and humans. In humans, blood pressure was elevated even when the dietary salt intake was low; however, much larger increases in blood pressure occurred when salt intake was high. In both sheep and rat studies, ACTH-induced increases in salt appetite and blood pressure were eliminated by adrenalectomy, suggesting that this mechanism is dependent on the adrenal hormones. In sheep, ACTH treatment also increased the heart rate.

Corticotropin Releasing Hormone and Urocortin

CRH was isolated from sheep hypothalamus and characterized in the early 1980s; it has the following properties:

- It is a 41-amino-acid peptide, primarily from the paraventricular nucleus of the hypothalamus but also found in other brain nuclei and in peripheral tissues.

- Its levels are elevated during stress.
- It coordinates responses to stress: increases the synthesis and release of ACTH from the anterior pituitary and influences the autonomic nervous system (e.g., increases sympathetic activity).
- Its actions are mediated by two receptor subtypes: CRH-R1, CRH-R2 (G-coupled, seven-transmembrane, bound receptors).

UCN was isolated from rat midbrain and characterized in 1990s; it has the following properties:

- It is a 40-amino-acid peptide, found in the brain (e.g., in the Edinger–Westphal nucleus, lateral superior olive, lateral hypothalamus, and supraoptic nucleus) and peripheral tissues (e.g., the pituitary, heart, and gastrointestinal tract).
- Its levels are elevated during stress.
- It coordinates responses to stress.
- Its actions are mediated primarily by CRH-R2 receptors.

CRH and UCN are found in brain areas known to be important for body sodium homeostasis (e.g., the hypothalamus; septal area; organum vasculosum of the lamina terminalis, OVLT; and amygdala) as well as in systems associated with autonomic function (e.g., the parabrachial nucleus and locus coeruleus). CRH and UCN are increased in the brain by stress and are involved in the initiation of the behavioral and physiological responses to stress. The major actions of CRH and UCN include the release of ACTH and the activation of the sympathetic nervous system. CRH also causes increased locomotor activity and modulation of the immune response. UCN shares a 45% homology with CRH, has CRH receptor-binding characteristics, and has a potency suggesting that it may mediate some of the actions, particularly of a behavioral nature, attributed to CRH.

The actions of CRH and UCN appear to be mediated by at least three CRH receptors ($-R1$, $-R2\alpha$, and $-R2\beta$) located in brain. CRH-R1 receptors are located in various regions of the cerebellum, several cortical areas, and the anterior and intermediate lobes of the pituitary. CRH-R2 receptors are located in the lateral septal nucleus, medial and cortical nuclei of the amygdala, bed nucleus of the stria terminalis, hippocampus, hypothalamic regions, and the serotonergic neurons of the midbrain raphe nuclei. CRH-R2 receptors are important in integrating hypothalamic neuroendocrine functions, as well as the autonomic and behavioral actions of central CRH. CRH-R2 α receptors are almost exclusively in the brain; CRH-R2 β receptors are found in brain and periphery. CRH-R2 receptors have a 6- to 40-fold greater affinity for UCN than for CRH.

The administration of ovine CRH in the brain of rabbits increased sodium intake, caused a sustained increase in both cortisol and corticosterone, and transiently increased blood pressure by 10 mmHg. The administration of CRH into the circulation did not influence salt intake. In recent experiments in sheep, however, prolonged central (see Figure 5) or peripheral (see Figure 6) administration of CRH or UCN failed to stimulate, and in fact inhibited, salt appetite despite stimulating the secretion of ACTH, as evidenced by the increased plasma levels of cortisol. Furthermore, the prolonged administration in the brain of CRH and UCN decreased food intake and need-free salt intake in sheep, rats, mice,

and baboons. The decrease in salt intake in baboons was transient, but the decrease in food intake was sustained throughout the infusion period. The mechanisms responsible for the inhibition of salt appetite are unclear. In sheep, the prolonged central or peripheral administration of either CRH or UCN increased blood pressure, but only UCN increased heart rate.

The administration of CRH and UCN could be expected to increase salt intake due to their stimulatory influence on the secretion of ACTH or activity of the sympathetic nervous system. It appears that the hormones associated with stress may have both excitatory (ACTH and adrenal hormones) and

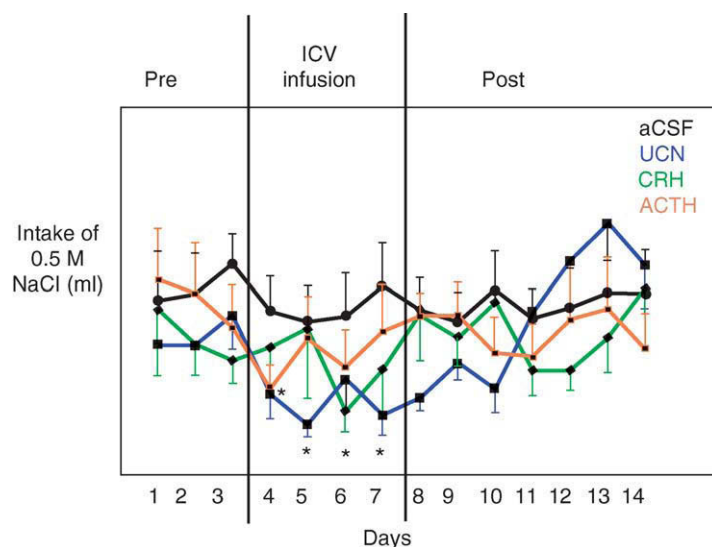


Figure 5 The effect of intracerebroventricular (ICV) infusions of aCSF, UCN, CRH, and ACTH on 0.5 M NaCl intake in sheep. * indicates $p < 0.05$. aCSF, artificial cerebrospinal fluid; ACTH, adrenocorticotropic hormone; CRH, corticotropin releasing hormone; UCN, urocortin.

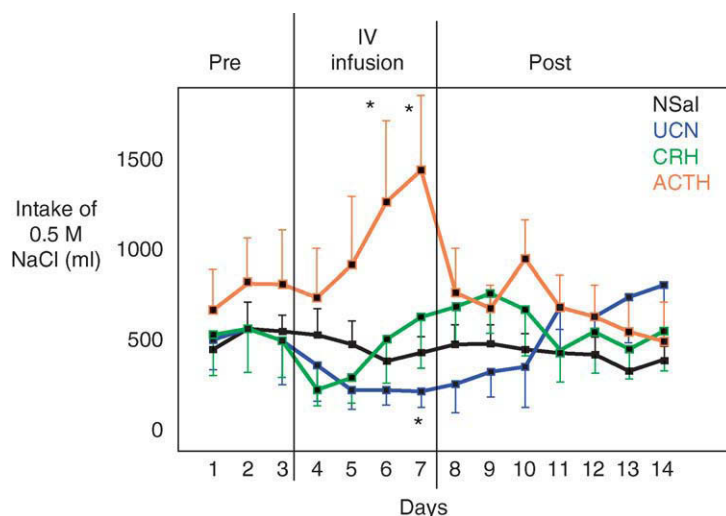


Figure 6 The effect of intravenous (IV) infusions of NSal, UCN, CRH, and ACTH on 0.5 M NaCl intake in sheep. ACTH, adrenocorticotropic hormone; CRH, corticotropin releasing hormone; NSal, normal saline; UCN, urocortin.

inhibitory (CRH and UCN) influences on salt intake. At present, the explanation for the inhibitory influence of CRH and UCN is not known. Presumably, the inhibition of salt appetite is caused by the influence of CRH on systems other than those influencing ACTH secretion or sympathetic nervous activity.

Combined infusion of Urocortin and Adrenocorticotrophic Hormone

The observation that CRH and UCN increased the secretion of ACTH, as evidenced by the increase in plasma cortisol but decreased salt appetite, suggested that ACTH-induced salt appetite was inhibited. Administering the two peptides together tested the possibility that UCN inhibits ACTH-induced salt appetite. Indeed, the ACTH-stimulated salt intake was completely blocked by UCN, administered into either the brain or the circulation (see [Figure 7](#)). Additional observations indicated that UCN, given systemically, reduced the increase in blood pressure and enhanced the increase in heart rate caused by ACTH. In contrast, UCN, administered in the brain, did not alter the increase in blood pressure or the increase in heart rate caused by ACTH. The mechanisms responsible for the inhibitory role of UCN on salt appetite are still unknown. However, it is thought that UCN may play a role in steroid-induced salt appetite at the amygdala.

It is notable that prolonged peripheral treatment with a combined infusion of ACTH and UCN produced a marked increase in heart rate (see [Figure 8](#)), which was at least additive in magnitude when compared with heart-rate increases due to the infusion of either hormone alone. Plasma cortisol was also

elevated with both the central and peripheral infusions of UCN in conjunction with the peripheral administration of ACTH. The major increase in heart rate in response to the combined administration of peripheral ACTH and UCN is probably via the direct action of CRH-R2 receptors on the heart. The inhibitory effect of UCN on blood pressure may be explained by the UCN-induced secretion of natriuretic peptides, molecules that increase excretion of sodium. These results indicate that sympathetic nervous system control of the cardiovascular system cannot be the only mechanism responsible for these stress hormone-induced changes. See [Table 1](#).

Brain Mechanisms Involved in Stress-Induced Salt Appetite

Although the brain mechanisms responsible for salt appetite induced by stress have not yet been conclusively identified, there are several major candidates. The first involves midline brain tissue located on the front wall of the third ventricle (a donut-shaped, fluid-filled cavity). The circumventricular organs (i.e., the subfornical organ and the OVLT), notable for their lack of blood-brain barrier, are thought to be crucial to body fluid and mineral homeostasis in rats. A lesion to either of these areas causes a marked reduction in salt intake by the salt-depleted animal. However, sheep with lesions in this brain area responded normally to salt depletion. Evidence also indicates that these brain areas may contribute to salt appetite induced by treatment with the mineralocorticoid hormones. For example, there appear to be mineralocorticoid receptors within these brain regions and deoxycorticosterone treatment

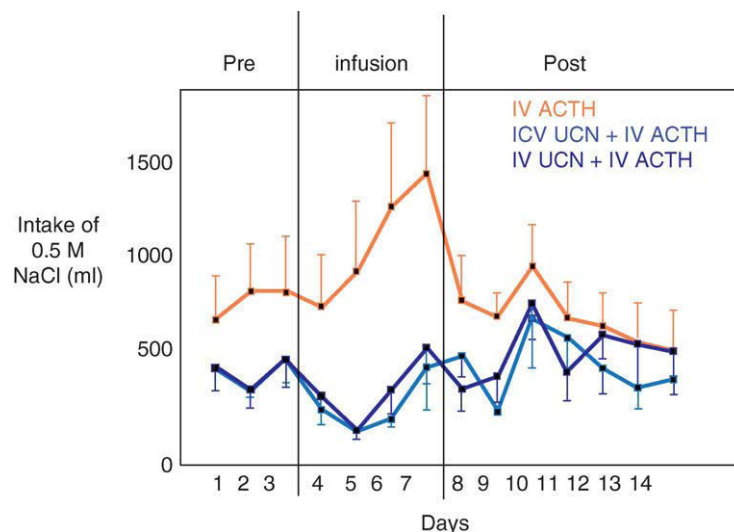


Figure 7 The effect of intracerebroventricular (ICV) and intravenous (IV) infusions of UCN on ACTH-induced 0.5 M NaCl intake in sheep. ACTH, adrenocorticotrophic hormone; UCN, urocortin.

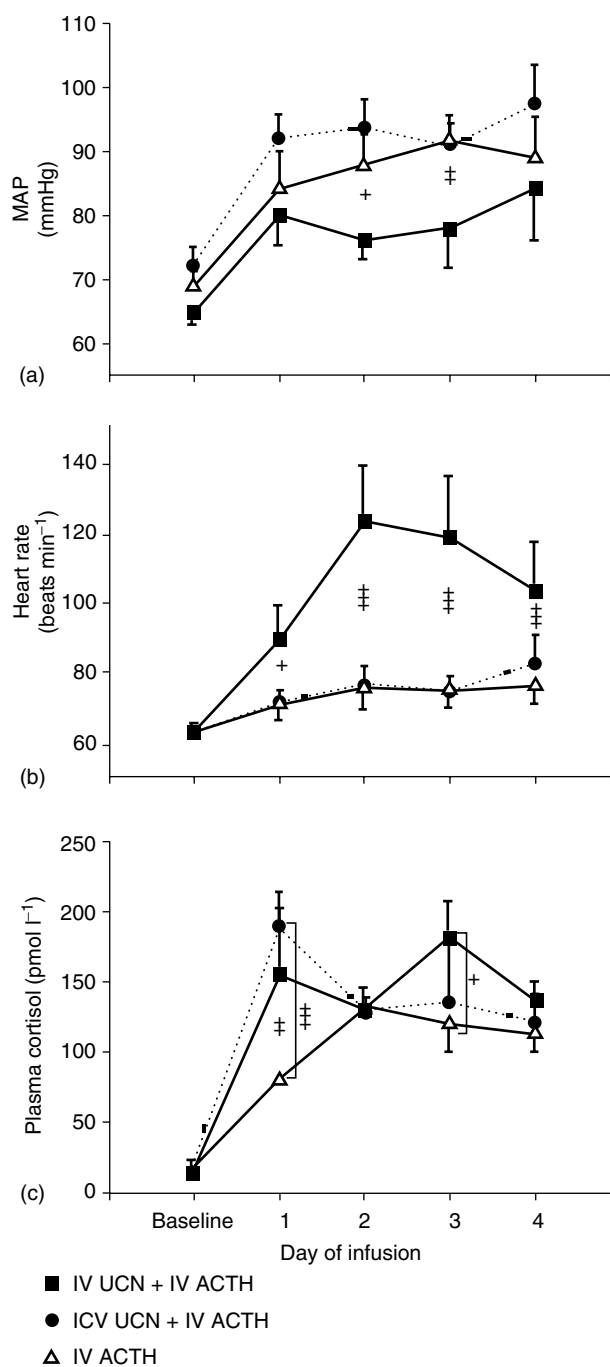


Figure 8 The effects of peripheral ACTH and central UCN infusions in sheep on a, blood pressure; b, heart rate; c, plasma cortisol concentration. ACTH, adrenocorticotrophic hormone; ICV, intracerebroventricular; IV, intravenous; MAP, mean arterial pressure; UCN, urocortin.

increases neuronal activity at these brain areas. However, salt appetite induced by treatment with deoxycorticosterone is unchanged or enhanced in rats with lesions of this brain region.

The second candidate involves brain regions located some distance from the midline and behind the

blood-brain barrier – the amygdala and bed nucleus of the stria terminalis. These brain areas are involved in the neural and hormonal responses to stress, especially those related to fear and anxiety. Increased neuronal activity has been reported in the medial amygdala of animals subjected to stress. Interference with mineralocorticoid receptor formation, with antisense to its mRNA, in the amygdala has been shown to interfere with the expression of deoxycorticosterone-induced salt appetite. Rats with lesion of the medial amygdala have impaired aldosterone and deoxycorticosterone-induced salt appetite, and the corticosterone potentiation of aldosterone-induced salt appetite is eliminated. Despite the evidence that salt appetite caused by salt depletion is mediated by the synergy between aldosterone and angiotensin II, rats with lesion of the medial amygdala do not have impaired salt appetite caused by salt depletion. Rats with lesion of the central nucleus of the amygdala or the bed nucleus of the stria terminalis (an area of brain intimately related to the central nucleus of the amygdala and often called the extended amygdala) show deficits in both deoxycorticosterone and salt-depletion-induced salt appetite. It seems clear at present that the several different brain areas mentioned earlier participate in the control of salt appetite. In rats, the subfornical organ and OVLT appear to be important to salt appetite caused by salt depletion, the medial amygdala seems to be important to salt appetite stimulated by the adrenocortical hormones, and the central nucleus of the amygdala and the bed nucleus of the stria terminalis seem to be involved in both. Further work is required to determine how these brain areas interact and what their roles are in modulating salt appetite during stress. Also, the brain mechanisms involved in salt appetite in other species still needs to be determined.

Conclusion

Overall, the data reviewed here show that stress can cause increased salt intake. Also, some, but not all, of the hormones associated with the stress response (e.g., ACTH and the adrenocortical hormones) can induce salt appetite. CRH and UCN inhibit salt appetite, even that caused by ACTH. The increased intake caused by ACTH appears to be mediated by the action of the adrenocortical hormones acting on receptors located in the amygdala and bed nucleus of the stria terminalis. At present, however, a role for adrenocortical hormones acting on receptors located in the circumventricular organs cannot be ruled out. An increase in angiotensin II or activity of the sympathetic nervous system caused by stress may also stimulate salt intake, possibly by the stimulation of the HPA

Table 1 Physiological effects of urocortin in sheep^a

	CRH		UCN		ACTH		ACTH + UCN	
	ICV	IV	ICV	IV	ICV	IV	ICV	IV
Salt appetite	↓	↓	↓	↓	↓	↑↑↑	↓	↓
Corticosteroids ^b	↑	↑↑	↑	↑↑	↔	↑↑	↑↑	↑
MAP	↑	↑	↑	↑	↔	↑↑↑	↑↑↑	↑↑↑
Heart rate	↔	↔	↑↑↑	↑↑↑	↔	↑↑	↑↑	↑↑↑↑

^aArrows indicate direction and relative magnitude of change. ACTH, adrenocorticotrophic hormone; CRH, corticotropin releasing hormone; ICV, intracerebroventricular; IV, intravenous; MAP, mean arterial pressure; UCN, urocortin.

^bPlasma cortisol level.

axis. The mechanisms responsible for the inhibitory actions of CRH and UCN are as yet unknown.

Hormones associated with the stress response also increase blood pressure, and evidence suggests that the enhanced intake of salt may augment any increases in blood pressure caused by the stress. For example, in rats, increased blood pressure caused by stress is augmented by high salt intake. Similar observations have been made in dogs and baboons. In humans, salt sensitivity and family history may also contribute to the effect of high salt intake on stress-induced hypertension.

Figure 4 depicts some of the mechanisms by which stress may influence both dietary salt intake and blood pressure. Other factors will no doubt emerge with further research, including species differences.

Acknowledgments

This work was supported by the National Health and Medical Research Council of Australia (NHMRC grant number 217011) and the Deutsche Bank Postgraduate Scholarship.

Further Reading

Axelrod, J. and Reisine, T. D. (1984). Stress hormones: their interaction and regulation. *Science* **224**, 452–459.
 Denton, D. A. (1982). *The hunger for salt: an anthropological, physiological and medical analysis*. Berlin: Springer-Verlag.

Denton, D. A., Coghlan, J. P., Fei, D. T., et al. (1984). Stress, ACTH, salt intake and high blood pressure. *Clinical and Experimental Hypertension: Theory and Practice* **A6**, 403–415.
 Denton, D. A. and Nelson, J. F. (1980). Kare, M. R., Fregly, M. J. & Bernard, R. A. (eds.) *Biological and behavioral aspects of salt intake*, pp. 229–246. New York: Academic Press.
 Gaillard, R. C. and Al-Damluji, S. (1987). Stress and the pituitary-adrenal axis. *Baillière's Clinical Endocrinology and Metabolism* **1**, 319–354.
 Glavin, G. B. (1985). Stress and brain norepinephrine: a review. *Neuroscience and Biobehavioral Reviews* **9**, 233–243.
 Makara, G. B. (1985). Mechanisms by which stressful stimuli activate the pituitary-adrenal system. *Federation Proceedings* **44**, 149–153.
 Saavedra, J. M. (1992). Brain and pituitary angiotensin. *Endocrine Reviews* **13**, 329–380.
 Schulkin, J. (1991). Sodium hunger: the search for a salty taste. Cambridge, UK: Cambridge University Press.
 Weisinger, R. S., Blair-West, J. R., Burns, P., et al. (1996). The role of angiotensin II in ingestive behaviour: a brief review of angiotensin II, thirst and Na appetite. *Regulatory Peptides* **66**, 73–81.
 Weisinger, R. S., Blair-West, J. R., Burns, P., et al. (2000). The inhibitory effect of hormones associated with stress on Na appetite of sheep. *Proceedings of the National Academy of Sciences USA* **97**(6), 2922–2927.
 Weisinger, R. S., Blair-West, J. R., Burns, P., et al. (2004). Cardiovascular effects of long-term central and peripheral administration of urocortin, corticotropin-releasing factor, and adrenocorticotropin in sheep. *Endocrinology* **145**(12), 5598–5604.

Schizophrenia

M van den Buuse and D Copolov

The Mental Health Research Institute of Victoria,
Parkville, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by J M Crook and D L Copolov, volume 3, pp 393–397,
© 2000, Elsevier Inc.

Introduction

Schizophrenia: Definition and Etiology

The Effects of External Stressors: Expressed Emotion and
Life Events

The Hypothalamic-Pituitary-Adrenocortical (HPA) Axis
and Schizophrenia

Stress, Early Development, and Schizophrenia:
Neurobiological Findings

Treatment Implications of the Vulnerability-Stress
Model of Schizophrenia: Psychosocial
Treatment Strategies

Conclusion

Glossary

<i>Adrenocorticotropin (ACTH)</i>	The pituitary hormone that stimulates the synthesis and secretion of adrenal corticosteroids and especially the glucocorticoids.
<i>Corticotropin-releasing hormone (CRH)</i>	The 41-amino-acid residue peptide released from hypothalamic neurons that mediates neural control of ACTH secretion. CRH has also been implicated in the central control of emotion and behavior.
<i>Dexamethasone suppression test (DST)</i>	A test designed to determine the responsiveness of the HPA axis to the negative feedback effects of the potent synthetic glucocorticoid dexamethasone. Major mental disorder is often associated with a blunted response to the DST.
<i>Expressed emotion (EE)</i>	Emotion expressed by close relatives toward family members with schizophrenia.
<i>Hypothalamic-pituitary-adrenocortical (HPA) axis</i>	The system that involves CRH-induced ACTH release, which in turn leads to adrenal glucocorticoid secretion. The latter moderates ACTH release by way of negative feedback regulation.
<i>Magnetic resonance imaging (MRI)</i>	An imaging technique commonly used primarily to produce high-quality images of brain.

Ultra-high risk (UHR) A cohort of individuals identified as being at ultra-high risk to develop psychosis.

Introduction

Stress may have an impact on the course and severity of schizophrenia during a number of stages along the neurodevelopmental trajectory, including presymptomatic and early symptomatic phases as well as the phases associated with chronicity and relapse. Extensive neurobiological research has now identified molecular pathways and brain regions that are likely to be involved in the interplay between stress and schizophrenia. The interaction between these pathways has not, however, been the subject of pharmacological intervention studies. Most of the therapy-oriented research in this field has involved psychosocial approaches to reducing stress and minimizing the impact of life events in order to reduce symptom severity and the risk of relapse.

Schizophrenia: Definition and Etiology

Schizophrenia is an enigmatic illness of unknown cause, heterogeneous symptomatology, and unpredictable course, with a variable treatment response. While the precise definition of schizophrenia poses numerous difficulties, there is general agreement that the disorder involves a breakdown of integration among emotions, thoughts, and actions due to underlying neurobiological abnormalities. Common symptoms include disorganized thought processes, hallucinations, delusions, inappropriate or blunted affect, social withdrawal, and poverty of speech. Cognitive impairments, including deficits in working memory, attention, and executive functioning are also often found in the disorder. Because many of these symptoms occur in other disorders, they cannot be regarded as diagnostic when taken individually. According to most diagnostic systems, the diagnosis of schizophrenia requires specific psychopathologies to be present for at least 6 months.

The onset of schizophrenia is usually during late adolescence or early adulthood. Gender differences exist, with women on average having a later onset and a more benign course of the illness. The lifetime prevalence of schizophrenia is around 1% of the general population across all ethnic backgrounds, nationalities, and socioeconomic classes. Most people with the disorder are unemployed. Many live in substandard

accommodation and experience drug or alcohol dependencies. Approximately 10% of people with schizophrenia commit suicide. Thus, schizophrenia represents a significant public health problem, from both an economic and a social point of view.

The Effects of External Stressors: Expressed Emotion and Life Events

Genetic factors play a major role in the development of schizophrenia; however, they alone do not explain the illness. First, multiple genes appear to be involved in the etiology of schizophrenia, collectively increasing susceptibility to the illness but each contributing only a small percentage to the risk. Furthermore, while family, twin, and adoption studies strongly support this genetic contribution, the finding of only 35–50% concordance for schizophrenia among monozygotic twins clearly supports the involvement of environmental factors. Epidemiological studies have identified several such adverse factors that may increase incidence of schizophrenia, such as pre- or perinatal obstetric complications, maternal influenza or rubella infection, or malnutrition and psychological trauma during pregnancy. For example, among both men and women conceived at the height of the Dutch famine in 1945, the risk of later developing schizophrenia was double that of individuals born before or after this period. Prenatal exposure to famine was also associated with decreased intracranial volume and white matter hyperintensities. These data have now been confirmed in a separate Chinese study, in which it was found that prenatal exposure to the 1959–61 famine more than doubled the risk of later developing schizophrenia. Both studies thus show that early stress caused by prenatal nutritional deficiency may play a role in the pathogenesis of schizophrenia. It should be noted that other nonstress factors could also play a role in the apparent link between famine and schizophrenia.

In addition to early environmental disturbances causing an increased risk, environmental factors later in life, such as drug abuse or psychosocial stress, may precipitate the disorder in predisposed individuals. For example, there is a significant relationship between the use of drugs of abuse, such as cannabis, and the exacerbation of psychosis and schizophrenia. The common denominator of all these early and late environmental factors is that they all induce stress, i.e., increased central and systemic release of stress hormones and disturbance of physiological and neurochemical homeostasis. The development of schizophrenia therefore may involve a disruption of

early development, by either genetic or stress factors, which induces subtle changes along the neurodevelopmental trajectory. These changes could take place in synaptogenesis, synaptic pruning, axonal myelination, and neuronal proliferation in various brain regions, including cortex and hippocampus, but if they precede the onset of the disorder they either remain silent until the onset of the disorder or lead to very minor neuropsychiatric or behavioral changes in childhood and adolescence. Such changes have been reported and include a tendency for delayed neurological development, reduced social functioning, lower IQ, and neurological soft signs. However, while these associations are statistically significant, many individuals in the normal population also display such abnormalities. As a result, the presence of such mild abnormalities is not of clinical utility in the prediction of the development of schizophrenia. The emergence of psychotic symptoms may be triggered by additional stress, which may also be responsible for a later aggravation or reemergence of symptoms. One of the most influential versions of the neurodevelopmental hypothesis of schizophrenia thus comprises at least two principal components, highlighting postulated early changes in the brain, followed by later disturbances in psychological and physiological homeostasis – the so-called two-hit hypothesis of schizophrenia.

Early attempts to characterize the existence of stress-reactive schizophrenia were driven by the finding that attempts to overcome negative symptoms by social stimulation sometimes induced the reemergence of positive, psychotic symptoms of schizophrenia. This has since been tested through the study of expressed emotion (EE) – emotion expressed by close relatives toward family members with schizophrenia. Many, but not all, studies have shown that patients who experience a family environment with high EE, especially of a type that is chronic and critical in nature, have a greater propensity for psychotic relapse compared to patients from families with lower levels of EE.

Although EE studies have provided some of the strongest evidence for the stress reactivity of schizophrenia, the EE measure has been frequently misunderstood and misapplied. One of the most common mistakes is to conclude that critical and overinvolved caregivers are the major cause of poor clinical outcomes, with the result that family members are blamed for causing or exacerbating the illness. Such an approach fails to take into account the iterative relationship between high EE and patient relapse, which includes the impact of psychiatric illness on

the patient's carers. In support of the iterative nature of this relationship is the established correlation between EE and family burden.

In comparison to EE, which is relatively ambient and consistent, thereby providing a probe for the impact of chronic stress on schizophrenia, life events are more discrete and less predictable, thereby providing a greater insight into the effect of acute stress on the condition. Although the majority of studies support an association between the propensity to experience episodes of schizophrenia and either the frequency of both major life events and daily hassles or their negative subjective appraisal, most of these studies have been retrospective and could therefore be affected by a search after meaning, which leads individuals to bias their recall in favor of giving undue importance to life events in order to explain their psychotic symptoms.

The anxiety associated with the psychotic experience itself may also contribute to relapse and symptom severity. Clearly, responses to delusions of persecution and danger or to hallucinations of a derogatory and intrusive nature may involve severe distress, which in turn may further exacerbate psychotic symptoms.

The Hypothalamic-Pituitary-Adrenocortical (HPA) Axis and Schizophrenia

There is good evidence that HPA axis dysfunction, as evidenced by nonsuppression in the dexamethasone suppression test (DST), occurs in approximately one in five subjects with schizophrenia, which is more than in normal subjects but less than in patients with depression. Post mortem studies have shown reduced glucocorticoid receptor expression in the hippocampus of subjects with schizophrenia, which might contribute to reduced negative feedback to the HPA axis and thus play a role in DST nonsuppression.

DST nonsuppression appears to be more common in patients who are in the acute or initial phases of their illness than in stable or chronic phases. This may explain some inconsistency in the literature, which could be attributable to differences between prodromal, first episode, and chronic patients. In addition, antipsychotic treatment has been shown to normalize at least some of the HPA axis disturbances, such as nonsuppression in the DST. The initial phase of psychosis has been shown to be associated with HPA axis hyperactivity, shown by elevated circulating cortisol cerebrospinal fluid CRH levels and impaired DST. First episode patients show enlarged pituitary volumes as measured by magnetic resonance imaging (MRI). This might reflect HPA axis hyperactivity, since

enhanced CRH and adrenocorticotropin (ACTH) secretion leads to increased size and number of pituitary ACTH-producing cells. These effects could be the consequence of the previously mentioned distressing psychotic experience or, alternatively, could be involved in the induction of psychosis. To distinguish between these possibilities, pituitary volumes have been measured in a cohort of individuals identified as being at ultra-high risk (UHR) to develop psychosis. Importantly, these subjects were symptom-free and did not receive antipsychotic medication at the time of pituitary volume measurement. Within this UHR cohort, pituitary volume positively predicted future transition to psychosis, and subjects with the greatest increase in pituitary volume also showed the shortest time between being scanned and psychosis onset. Theoretically, UHR individuals with enlarged pituitary volumes might benefit from preventative treatments that target HPA axis overactivity, such as glucocorticoid receptor antagonists or CRH receptor antagonists.

While there is clear evidence for disturbances in HPA axis function in a subset of individuals with schizophrenia, the mechanism by which these disturbances mediate illness development and symptomatology is less clear. Chronically enhanced levels of glucocorticoid hormones are neurotoxic, for example, leading to macroscopic volume reduction and cell loss in the hippocampus as well as reduced neuronal volume and dendritic arborization. These effects could be associated with some of the cognitive impairments seen in patients with schizophrenia and could also explain some of the structural changes seen in the brains of patients with schizophrenia, such as reduced cortical and hippocampal volumes.

There appears to be a synergistic relationship between HPA axis activity and dopaminergic activity in the brain. Glucocorticoids increase dopamine metabolism in the striatum and dopamine release in the nucleus accumbens. If these effects apply to all areas of the brain, HPA axis disinhibition is likely to contribute to dopaminergic hyperactivity, which appears to be central to psychosis symptomatology. Positron emission tomography studies have shown that patients with schizophrenia display greater increases in central dopamine release in response to certain challenges, including low-dose amphetamine. All antipsychotic drugs share dopamine receptor antagonism as part of their pharmacological mode of action. Conversely, abnormalities in dopaminergic activity influence HPA axis activity and render an individual hyperresponsive to stress. The vulnerability-stress model of schizophrenia suggests that the illness is potentiated by stress due to this interplay with brain dopaminergic activity.

Stress, Early Development, and Schizophrenia: Neurobiological Findings

In an attempt to elucidate neurobiological mechanisms involved in the interaction of environmental factors and schizophrenia, animal model studies have simulated major life events during early stages of development. These animals, mostly rats, are then tested as adults in behavioral paradigms with relevance to schizophrenia. Developmental disruptions include neonatal stress induced by maternal deprivation, prenatal maternal stress, developmental stress induced by placental insufficiency, antimitotic treatment or vitamin D depletion, and prenatal or neonatal infection. Several studies have also used neonatal brain lesions, such as in the ventral hippocampus, frontal cortex, or amygdala. Such early interventions induce behavioral effects in the adult such as hyperlocomotion, enhanced sensitivity to psychotomimetic drugs (amphetamine, phencyclidine), disruptions of pre-pulse inhibition, and learning and memory deficits.

Another approach has been to study the influence of variations in maternal care. In these rat studies, dams were classified according to the level of maternal care, and behavioral and molecular changes in the offspring were examined. Offspring from mothers that displayed high levels of pup licking, grooming, and nursing differed in the level of DNA methylation from offspring from mothers that displayed low levels of these behaviors. These effects persisted into adulthood and resulted in specific changes in the regulation of glucocorticoid receptor expression. Offspring from mothers that displayed lower levels of maternal care showed lower levels of glucocorticoid receptor expression and greater plasma corticosterone increases to stress than offspring from mothers that displayed higher levels of maternal care. These studies show that early environmental factors, in this case the level of maternal care, induce long-term molecular and biochemical alterations leading to differences in glucocorticoid receptor expression and alterations in HPA axis activity. Such changes in stress responsiveness could be involved in the role of stress in schizophrenia.

Treatment Implications of the Vulnerability-Stress Model of Schizophrenia: Psychosocial Treatment Strategies

Psychosocial treatments are critical elements in the management of schizophrenia, provided that they are – as they need to be – in tandem with the key-stone of treatment, antipsychotic medication. Stress management has received relatively little research

attention in the treatment of the disorder, but there are suggestions that this form of treatment reduces the prevalence of hospitalization. Psychosocial treatment modalities that have received greater validation include family therapy, assertive community treatment, psychoeducation, social skills training, and cognitive-behavioral therapy. To the extent that these modalities improve coping skills, the positive outcomes they have been associated with in terms of relapse prevention and symptom amelioration might, in part, be a consequence of reduction in stress levels.

Conclusion

A variety of epidemiological, clinical, and neurobiological studies have suggested a nontrivial link between stress and schizophrenia, both during development and during the initial and established phases of the disease. These studies provide clues about brain mechanisms that could be the target of future new pharmacological treatments and about psychosocial strategies aimed at enhancing coping skills and reducing the impact of stress on the individual. However, as the available evidence is still controversial, further studies are clearly needed to better our understanding of the role of stress in schizophrenia.

See Also the Following Articles

Dexamethasone Suppression Test (DST); Hypothalamic-Pituitary-Adrenal; Pre-pulse Inhibition; Psychotic Disorders.

Further Reading

- Bustillo, J., Lauriello, J., Horan, W. and Keith, S. (2001). The psychosocial treatment of schizophrenia: an update. *American Journal of Psychiatry* 158, 163–175.
- Corcoran, C., Walker, E., Huot, R., et al. (2003). The stress cascade and schizophrenia: etiology and onset. *Schizophrenia Bulletin* 29, 671–692.
- Cotter, D. and Pariante, C. M. (2002). Stress and the progression of the developmental hypothesis of schizophrenia. *British Journal of Psychiatry* 181, 363–365.
- Garner, B., Pariante, C. M., Wood, S. J., et al. (2005). Pituitary volume predicts future transition to psychosis in individuals at ultra-high risk of developing psychosis. *Biological Psychiatry* 58, 417–423.
- Gispén-de Wied, C. C. (2000). Stress in schizophrenia: an integrative view. *European Journal of Pharmacology* 405, 375–384.
- Hulshoff Pol, H. E., Hoek, H. W., Susser, E., et al. (2000). Prenatal exposure to famine and brain morphology in schizophrenia. *American Journal of Psychiatry* 157, 1170–1172.

- McGrath, J. J., Feron, F. P., Burne, T. H., Mackay-Sim, A. and Eyles, D. W. (2003). The neurodevelopmental hypothesis of schizophrenia: a review of recent developments. *Annals of Medicine* 35, 86–93.
- Miller, P., Lawrie, S. M., Hodges, A., et al. (2001). Genetic liability, illicit drug use, life stress and psychotic symptoms: preliminary findings from the Edinburgh study of people at high risk for schizophrenia. *Social Psychiatry and Psychiatric Epidemiology* 36, 338–342.
- Mueser, K. T. and McGurk, S. R. (2004). Schizophrenia. *The Lancet* 363, 2063–2072.
- Pantelis, C., Yucel, M., Wood, S. J., McGorry, P. D. and Velakoulis, D. (2003). Early and late neurodevelopmental disturbances in schizophrenia and their functional consequences. *Australia and New Zealand Journal of Psychiatry* 37, 399–406.
- Van den Buuse, M., Garner, B. and Koch, M. (2003). Neurodevelopmental models of schizophrenia: effects on prepulse inhibition. *Current Molecular Medicine* 3, 459–471.
- Van den Buuse, M., Garner, B., Gogos, A. and Kusljic, S. (2005). Importance of animal models in schizophrenia research. *Australia and New Zealand Journal of Psychiatry* 39, 550–557.
- Weaver, I. C., Cervoni, N., Champagne, F. A., et al. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience* Aug 7(8), 847–854.

School Stress and School Refusal Behavior

C A Kearney, L C Cook and G Chapman
University of Nevada, Las Vegas, NV, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by C A Kearney and C Pursell, volume 3, pp 398–401, © 2000, Elsevier Inc.

School Stress and Linkage to School Refusal Behavior
Major Characteristics of School Refusal Behavior
Etiology and Major Functions of School Refusal Behavior
Assessment of School Refusal Behavior
Treatment of School Refusal Behavior

Glossary

- School refusal behavior* The substantial child-motivated refusal to attend school and/or difficulties remaining in classes for an entire day.
- School stress* Unpleasant physical and cognitive symptoms in response to global and specific school-related stressors.

School Stress and Linkage to School Refusal Behavior

Children and adolescents become anxious about many things, but youths become particularly stressed by difficult family and school situations. School-related stressors may involve global and specific issues. Global school-related stressors include large class size, threats of victimization and violence, changes in academic and social status, excessive

homework, and tedious curricula. Specific school-related stressors include teachers and other school officials, peer interactions, examinations, performances before others, social gatherings, and use of cafeterias and public restrooms.

School-related stressors are quite prevalent among youths, and many youths experience concomitant physiological and cognitive symptoms as a result. Common physiological symptoms include headaches, stomachaches, shortness of breath, trembling, dizziness, and heart palpitations. Common cognitive symptoms include catastrophic thoughts about losing control, going insane, suffering humiliation or harm, experiencing negative evaluation from others, appearing foolish, having few friends, having trouble concentrating, and lacking competence in various areas. If school-related stress becomes severe, then some youths may refuse to attend school to avoid these unpleasant symptoms.

Major Characteristics of School Refusal Behavior

School refusal behavior is the substantial child-motivated refusal to attend school and/or difficulties remaining in classes for an entire day. Specifically, the behavior involves youths ages 5–17 years who (1) are completely absent from school, (2) attend school but then leave school during the school day, (3) go to school following intense behavior problems such as tantrums in the morning, and/or (4) display unusual distress during the school day that leads to pleas for future nonattendance. School refusal behavior may also be defined by its timeline. For example,

self-corrective school refusal behavior refers to absenteeism that remits spontaneously within 2 weeks. Acute school refusal behavior lasts from 2 weeks to 1 year, and chronic school refusal behavior lasts longer than 1 year. Greater chronicity is related to greater difficulty reintroducing a child to school.

Approximately 5–28% of school-age youths refuse school at some time in their lives. School refusal behavior is seen equally in boys and girls and among families of various income levels. The short-term consequences of school refusal behavior include declining academic status, social alienation, increased risk of legal trouble, family conflict, and severe disruption in a family's daily routine. The long-term consequences of school refusal behavior include school dropout and subsequent economic deprivation, occupational and marital problems, alcohol abuse and criminal behavior, and mental disorders.

School refusal behavior is marked by considerable symptom heterogeneity. Common internalizing problems include general and social anxiety/shyness, depression and social withdrawal, fear, fatigue, and somatic complaints such as stomachaches, headaches, nausea, and tremors. Common externalizing problems include tantrums, verbal and physical aggression, clinging, reassurance seeking, refusal to move, noncompliance, and running away from school or home.

School refusal behavior is also marked by many different diagnoses that could cause or be the result of problematic absenteeism. In this population, common comorbid diagnoses include separation anxiety disorder, generalized anxiety disorder, depression, and oppositional defiant disorder, among others. In addition, school refusal behavior has been linked to limited developmental disabilities such as learning disorders and pervasive developmental disabilities such as Asperger's disorder, autism, and mental retardation.

Etiology and Major Functions of School Refusal Behavior

The specific causes of school refusal behavior are often unclear for a particular case. However, common triggers include the onset of a new school year, entry into a new school building, school-related trauma, teacher or class changes, family illness or transition, parental psychopathology, and marital conflict or divorce. Child factors such as behavioral inhibition, excessive shyness, and association with deviant peers can also be influential. Unfortunately, identifying the etiology in this population is often a fruitless task. As such, clinicians may focus instead on the primary functions or reasons why youths refuse school:

1. To avoid school-related objects and situations that provoke a general sense of negative affectivity (dread, anxiety, depression, and physiological symptoms).
2. To escape aversive social and/or evaluative situations at school.
3. To receive attention from significant others outside of school.
4. To obtain tangible reinforcement outside of school.

The initial two functions refer to youths who refuse school for negative reinforcement, or to avoid something unpleasant at school. These functions are closely related to the concept of school stress. The latter two functions refer to youths who refuse school for positive reinforcement, or to pursue someone or something more alluring outside school. Some youths also refuse school for multiple reasons, and those who do often require complex and extended treatment.

Assessment of School Refusal Behavior

Assessing youths with school refusal behavior often involves defining a child's exact behavior problems, discovering primary and secondary functions for absenteeism, and identifying effective treatment. To do so, assessment methods may include interviews, child self-report and parent/teacher measures, direct observation, daily monitoring, and discussions with school officials. Interviews may be structured or unstructured in nature. A common structured interview for this population is the Anxiety Disorders Interview Schedule for Children, which has child and parent versions and helps clinicians identify various diagnoses. In addition, the interview contains a special section on school refusal behavior to help clinicians pinpoint the frequency and intensity of the behavior and related stress-based symptoms. Unstructured interviews may also be used to assess youths with school refusal behavior. Questions may relate to school-related stressors that affect a particular child, degree of avoidance, physiological and cognitive symptoms, attention-seeking, and pursuit of tangible reinforcers outside of school.

Several child self-report and parent and teacher measures also pertain to youths with school refusal behavior. The most common address negative affectivity, depression, general and social anxiety, somatic complaints, overactivity, aggression, noncompliance, demands for attention, social problems, fear, self-concept, family dynamics, and the function of school refusal behavior. Gathering information from multiple sources is recommended because school refusal behavior can include many subtle symptoms.

Direct observation of a child's and family's morning activities at home may also help define the primary symptoms and maintaining variables of school refusal behavior. Pertinent behaviors include:

1. Avoidance in the form of clinging, refusal to move, running away, and/or noncompliance to requests.
2. Physiological reactivity such as stomachaches, headaches, abdominal pain, tremors, and nausea/vomiting.
3. Cognitive distortions or verbalizations about discomfort related to school.
4. Sudden changes in behavior.
5. Pleas to stay home from school.
6. Increased family conflict, especially following a curtailment of the child's social activities.
7. Significant changes in parental behavior.
8. Teacher reports of differences in a child's behavior at school.

The daily assessment of a child's school refusal behaviors and attendance is also important. This helps gauge familial compliance to homework assignments, increase insight into a child's problem, and identify positive or negative changes in behavior across time. For example, family members may provide daily ratings with respect to a child's anxiety, depression, and overall distress (general feelings of dread or being upset). In addition, parents may rate a child's noncompliance (not listening to parent commands) and disruption to their daily routine. Parents can also list other child behavior problems and time missed from school.

Finally, school officials may be contacted for additional information, including course schedules, grades, written work, and required make-up work; goals and attitudes of school officials and peers regarding a child; procedures and timelines for reintegrating a child into school; potential obstacles to reintegrating a child into school; confirmation of past school refusal behavior; general social or other behaviors of a child in school; a map or outline of the school (e.g., lockers, cafeteria, library); feedback about treatment effectiveness; disciplinary procedures and procedures for contacting parents; and rules about absenteeism, conduct, or leaving school areas.

Treatment of School Refusal Behavior

Treatment for youths with school refusal behavior may be linked to the function of the behavior. For example, treatment for youths who refuse school to avoid stimuli provoking negative affectivity (i.e., school stress) often involves reducing unpleasant physiological symptoms and exposing a child to

various school-related items and situations. Specifically, this involves psychoeducation about the nature and process of anxiety (i.e., linking feelings, thoughts and behaviors), developing an anxiety and avoidance hierarchy of situations and stimuli, teaching somatic management skills to decrease negative emotional arousal (i.e., relaxation training and breathing retraining), systematic imaginal and *in vivo* exposure to anxiety cues identified on a graded hierarchy (progressive reintroduction to the school setting), and having a child praise him- or herself for coping with transient negative emotions.

For youths who refuse school to escape aversive social and/or evaluative situations, treatment strategies involve building or refining social skills and reducing social anxiety in key interactive situations at school. Specifically, this involves psychoeducation about the nature and process of anxiety (i.e., linking feelings, thoughts, and behaviors), teaching a child to identify cognitive distortions (e.g., all-or-none thinking, catastrophization, and overgeneralization) and negative self-statements in anxiety-provoking situations, learning methods to change distorted thoughts to helpful coping thoughts, graduated exposure to anxiety-provoking social situations through in-session practice with a therapist (including modeling and role-play of different situations), and between-session practice of skills in real-life social situations.

For youths who refuse school for attention, treatment strategies involve parent-based techniques to shift parental attention away from school refusal behaviors and toward appropriate school attendance behaviors. Specifically, this involves restructuring parent commands, establishing fixed routines regarding a child throughout the day, implementing negative consequences for school refusal behaviors, implementing positive consequences for school attendance behaviors, and, in some cases, forced school attendance.

For youths who refuse school for tangible reinforcement outside school, treatment strategies involve improving a family's problem-solving abilities, reducing conflict, increasing rewards for school attendance, and decreasing rewards for school absenteeism. Specifically, this involves establishing times and places when family members can negotiate problem solutions, defining behavior problems specifically, designing written parent-child contracts to address problems, implementing contracts, training in communication skills, training in peer-refusal skills (i.e., helping a child refuse offers to miss school), and escorting a child to school or to each class. For all treatments, relapse-prevention training is also necessary.

See Also the Following Articles

Attention-Deficit/Hyperactivity Disorder, Stress and; Neurodevelopmental Disorders in Children.

Further Reading

- Berg, I. and Nursten, J. (1996). *Unwillingly to school* (4th edn.). Washington, DC: American Psychiatric Press.
- Kearney, C. A. (2001). *School refusal behavior in youth: a functional approach to assessment and treatment*. Washington, DC: American Psychological Association.
- Kearney, C. A. and Albano, A. M. (2000). *When children refuse school: a cognitive-behavioral therapy approach/therapist's guide*. New York: Oxford.

- Kearney, C. A. and Silverman, W. K. (1996). The evolution and reconciliation of taxonomic strategies for school refusal behavior. *Clinical Psychology: Science and Practice* 3, 339–354.
- Kearney, C. A. and Silverman, W. K. (1999). Functionally-based prescriptive and nonprescriptive treatment for children and adolescents with school refusal behavior. *Behavior Therapy* 30, 673–695.
- King, N. J., Ollendick, T. H. and Tonge, B. J. (1995). *School refusal: assessment and treatment*. Boston, MA: Allyn and Bacon.
- Silverman, W. K. and Albano, A. M. (1996). *Anxiety disorders interview schedule for DSM-IV: child version*. New York: Oxford.

School Violence and Bullying

J Juvonen

University of California, Los Angeles,
Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Hostile School Behavior: Prevalence and Concerns
Connection between Bullying and Violent Offending
Antecedents and Consequences of Hostile Experiences
The Effects of Witnessing Hostile Behaviors

Bullying and school violence are significant sources of stress in American schools. In a recent national survey, children ages 8–15 rated bullying as one of the major concerns affecting their lives. Another survey indicated that almost two-thirds of public secondary school students thought that a shooting could take place in their school.

In contrast to violent behaviors that involve use of physical force, bullying entails the abuse of psychological strength, which can be achieved by various means. Bullies intimidate their victims by relying on name-calling, exclusion, threats, and/or spreading of rumors. Because the perpetrators of violent acts as well as bullying abuse their power, both forms of hostile behavior cause psychological distress in that the actual and potential victims of peer maltreatment or abuse experience anxiety. Repeated experiences of hostile behavior by schoolmates are, in turn, associated with increased feelings of depression and either withdrawn or aggressive behavior.

Hostile School Behavior: Prevalence and Concerns

According to the most recent national data, 14% of students ages 12–18 reported having been bullied in school during the previous 6 months. The estimates for bullying vary depending on how broadly the term is defined and the time frame provided. Using daily (as opposed to monthly or yearly) reports of bullying experiences, almost 50% of young teens reported at least one incident of bullying during a 2-week period in urban middle schools. Up to 70% of secondary students reported having experienced bullying at some point during their school careers. Although the experiences of bullying are fleeting for most students, an estimated 5–10% of students confront bullying on a weekly, if not daily, basis for 1 or more school years.

The prevalence estimates of bullying also vary by grade level and school type. Compared to other grade levels, middle school students reported the highest rates of bullying, and sixth-grade students were most fearful of their safety. Although girls reported being more concerned about their safety at school or on their way to and from school, boys reported more incidents of violence and bullying. Boys were also the perpetrators of the most violent incidents.

Approximately 70% of public schools reported one or more violent incidents during a school year. Although the most serious forms of violence are less common, one-fifth of public schools reported at least one incident involving aggravated assault, sexual assault, rape, or robbery per year. Violent incidents

are most likely to take place during the high school years and in large urban schools.

Students in urban public schools are most likely to fear for their safety. Black and Latino students reported feeling more fearful of being attacked than did White students. Although some hostile incidents appeared to be racially motivated, this does not necessarily mean that students were worst off in ethnically mixed schools. Recent analyses of middle schools suggest that greater ethnic diversity (i.e., a larger number of ethnic groups and their relative balance) is associated with feelings of safety.

Connection between Bullying and Violent Offending

Children who bully in childhood are at risk of becoming violent offenders. Research conducted in Norway shows that boys identified as bullies in their teenage years were four times as likely to commit violent acts in their twenties than other boys. Long-term studies conducted in the United States reveal that when childhood bullying is associated with more serious subsequent fighting, the risk of committing assaults and rape significantly increases. Hence, at least the physical forms of bullying and violent offending among males reflect a spiral of increasingly serious hostile behaviors.

Both childhood bullying and school violence are predicted by personality dispositions (e.g., impulsivity), environmental factors beyond school (e.g., problematic relationships with parents), and school-related situations (e.g., perceptions of unfair treatment by teachers or peers). These risk factors often interact in that students with certain dispositions are more likely to encounter situations that prompt them to act aggressively. For example, impulsive students are likely to act in a way that increases their chances of someone's bumping into them in a hallway, and, after experiencing an unintended push, they are more likely to react with hostility. The hostile response is often promoted by a perception that the other person intended to cause harm.

Many aggressive students develop a generalized sense of mistrust. Hence, students commit hostile acts to protect themselves. For example, compared to non-bullied students, bullied students are more likely to carry a weapon to school. Fear for their safety is the most common reason provided by students who carry weapons to school.

Other students' hostile acts are motivated by a desire to maintain a sense of control and to show their power. These bullies or violent offenders enjoy at least temporary admiration among their peers because of their dominant status. Fellow students rarely

challenge their mean and hurtful behaviors because they are fearful of the reactions of these dominant ringleaders. Passivity and acceptance of toughness, in turn, promote social norms that encourage and maintain bullying or even violent behavior. Thus, the school environment, including the reactions of fellow peers (victims or bystanders), plays a critical role in motivating hostile behavior in school.

Antecedents and Consequences of Hostile Experiences

Because of the low prevalence rates of serious violent incidents in school, little is known about the factors that increase the risk of being targeted. However, research on bullying shows that socially withdrawn and passive children are at risk of getting bullied. After repeated experiences of bullying, these submissive victims become even more withdrawn. These students most often suffer in silence, and teachers and parents are unaware of their plight or distress. Instructors may interpret the quiet signs of distress as a lack of interest in schoolwork. Indeed, teachers rate victims as being just as disengaged from schoolwork as bullies.

Students who respond to bullying with unregulated aggression (i.e., aggressive victims) are also at a high risk for continued peer maltreatment. In these cases, bullying experiences are related to extreme peer rejection and severe school difficulties. Aggressive victims are presumed to have underlying emotional difficulties, which are also used to help explain why their difficulties continue. Thus, in the cases of both submissive and aggressive victims, the causes and consequences of bullying form a vicious circle.

Although aggressive victims resemble violent offenders most closely, the victims of bullying are unlikely to resort to violence to get back at their tormentors. Nevertheless, the tragic school shooting incidents since the late 1990s highlight that chronic victims of bullying, like violent offenders, have trouble effectively dealing with social failures (e.g., public ridicule or rejection by a romantic partner). Some violence prevention and antibullying interventions are therefore specifically designed to help students better cope with bullying incidents.

The Effects of Witnessing Hostile Behaviors

The hostile behaviors of schoolmates affect not only those who are victimized but also those who witness peer abuse or maltreatment in school. For example, analyses of daily incidents of bullying in middle school show that social anxiety is increased when students see someone else being bullied. Students

who witness violent incidents are likely to feel especially vulnerable and unsafe. Thus, although there may be relatively few hostile students in any one school, one or two can terrorize entire grade levels and a couple of violent events may cause widespread fear. In view of such exposure effects, the level of concern about bullying and violence in American schools is not unrealistically high.

Part of the distress caused by observed events is likely to be related to the school's response system (or lack thereof). Although the parents' concerns about the safety of their children at school may be alleviated by the presence of metal detectors and campus officers, students continue to worry about their safety because they know that someone can smuggle a weapon to campus or they are fearful of someone bullying them in sites lacking monitoring (e.g., restrooms and locker rooms). In addition to feeling vulnerable, an unsafe or noncaring school climate is likely to breed mistrust and alienation that, in turn, increase the risk of students carrying weapons and setting off a hostile response.

In summary, school violence and bullying cause, and are also partly caused by, feelings of vulnerability that reach beyond personal experiences of hostile acts. This can be depicted as a following process:

Perceived vulnerability of self → Hostile
response to perceived threat → Unsafe school
climate

To promote feelings of safety and trust, schools are relying more and more on prevention models that focus not only on at-risk youth but also on

helping improve the school climate. A caring climate in schools, in which students feel they belong, are respected, and are listened to and in which teachers help mediate hostile incidents, can buffer the safety concerns and distress of the entire student body.

See Also the Following Articles

Adjustment Disorders; Depression Models; School Stress and School Refusal Behavior.

Further Reading

- Jimerson, S. R. and Furlong, M. (eds.) (2006). *Handbook of school violence and school safety: from research to practice*. Mahwah, NJ: Lawrence Erlbaum.
- Juvonen, J. and Graham, S. (eds.) (2001). *Peer harassment in school: the plight of the vulnerable and victimized*. New York: Guilford Press.
- Juvonen, J., Nishina, A. and Graham, S. (2006). Ethnic diversity and perceptions of safety in urban middle schools. *Psychological Science* 17, 393–400.
- Leary, M. R., Kowalski, R. M., Smith, L., et al. (2003). Teasing, rejection, and violence: case studies of the school shootings. *Aggressive Behavior* 29, 202–214.
- Smith, P. K., Pepler, D. and Rigby, K. (eds.) (2004). *Bullying in schools: how successful can interventions be?* New York: Cambridge University Press.

Relevant Websites

- U.S. Department of Education and U.S. Department of Justice. (2005). Indicators of school crime and safety: 2005. National Center of Education Statistics. <http://nces.ed.gov>.

Seasonal Changes in Stress Responses

R J Nelson and L B Martin II

Ohio State University, Columbus, OH, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R J Nelson and D L Drazen, volume 3, pp 402–408, © 2000, Elsevier Inc.

Introduction

Seasonal Energetic Adaptations

Seasonal Adjustments in Glucocorticoid Secretion

Why Are Stress Responses Seasonally Variable?

Glossary

Glucocorticoids

A class of steroid hormones secreted from the adrenal cortices that serves to promote energy use. Corticosterone is the glucocorticoid usually secreted by rodents and birds, whereas cortisol is typically secreted by primates.

Melatonin

An indole-amine hormone secreted from the pineal gland at night that is important in the regulation of daily and seasonal biological rhythms.

Photoperiod

Day length or the amount of light per day.

<i>Photo-periodism</i>	The response of plants to the length of day and night, including the ability to determine day length (usually by measuring night length) in both plants and animals; first proposed by botanists Garner and Allard in 1922.
<i>Pineal gland</i>	An endocrine gland that secretes melatonin, usually located between the telencephalon and diencephalon in vertebrates; also called the epiphysis.

Introduction

Stressors are not constant but, rather, vary over time. In addition to daily changes in stress, seasonal changes in the quality and quantity of stress are common. Many energetic adaptations have evolved that attenuate the stress response during the winter when energy shortages may limit the ability of organisms to cope successfully with specific types of stressors. The temporal organization of such stress responses therefore must be considered to obtain a holistic understanding of the mechanisms underlying the stress response.

The main endocrine features of the stress response involve hormones that act to suppress energy storage and promote energy utilization from adipose and liver stores. Epinephrine and the glucocorticoids are the major hormonal components of the stress response. Animals require a relatively steady supply of energy to sustain biological functions. Although these energy demands are somewhat continual, both daily and seasonal fluctuations in energy requirements occur as physiological and behavioral activities wax and wane. Animals typically do not feed continuously to supply energy to cells; rather, adaptations have evolved to allow animals to store and conserve energy when they are engaged in nonfeeding activities. Eventually, however, the chronic need for energy depletes these stores and requires animals to obtain food. In most habitats, significant daily and seasonal fluctuations in food availability occur. For example, outside of the tropics, food availability is generally low during the winter when thermogenic energy demands are typically high. Consequently, energy intake and energy expenditures are never perfectly balanced because of fluctuations in both energy requirements and availability. Whenever a significant imbalance of energy is perceived, a stress response in the form of secreted glucocorticoids ensues.

Reproduction is sufficiently demanding that a stress response is often associated with breeding in terms of elevated glucocorticoid concentrations. That is, glucocorticoids are released to support the energetic needs of reproduction including territorial defense or courtship behaviors. However, when

energy is insufficient to support reproduction and thermogenesis or other energetic demands, the stress response becomes prolonged. A chronic stress response results in pathology. Most animals restrict breeding to specific times of the year when food is abundant and survival and reproductive success are most likely. The inhibition of winter breeding is one of the central components of a suite of energy-saving seasonal adaptations evolved among nontropical animals. In addition to the seasonal onset and cessation of reproductive function among nontropical vertebrates, it is also well-established that endocrine, metabolic, growth, neural, immune, and thermogenic processes undergo seasonal changes to promote winter energy conservation. Mechanisms in virtually every physiological and behavioral system have evolved to cope with the winter energetic bottleneck resulting from increased thermogenic energy requirements during periods of low energy availability. Presumably, animals possessing these energy-conserving adaptations enhance their survival and ultimately increase their reproductive success. Generally, glucocorticoid concentrations are higher during the winter than the summer and individuals display more elevated stress responses during the winter than the summer. However, because coping strategies are engaged in anticipation of the winter stressors, many individuals do not show exaggerated stress responses during the winter.

How do animals determine time of year? In most cases, the photoperiod (day length) acts as the principal external cue to determine time of year. Photoperiod measurement is important for organizing the seasonal collapse and regrowth of the reproductive system in mammals and birds from temperate and boreal latitudes. Indeed, changes in day length provide the most error-free signal for predicting favorable conditions. Day length is encoded by the hormone melatonin, which is secreted primarily at night from the pineal gland.

Seasonal Energetic Adaptations

Many aspects of life are energetically challenging, and individuals have evolved timing mechanisms to parse out energetically demanding activities to different times of year. For example, birds engage in predictable cycles of migration, breeding, parenting, molting, and overwintering. These energetically expensive activities generally do not overlap, and, again, all these activities elicit classical stress responses (i.e., elevated activation of the hypothalamic-pituitary-adrenal (HPA) axis).

One 1936 study is particularly informative in understanding the basis of Selye's thinking about stress

responses. Rats were housed in low-temperature conditions for 2 days, 2 weeks, or 2 months. After 2 days in these conditions, Selye noted the hypertrophy of the adrenal cortex; reduced eosinophil (a type of immune cell circulating in the blood) numbers; and atrophy of the thymus, spleen, and lymph nodes. In other words, the rats exhibited a classic stress response. After 2 weeks in the low temperatures, it was discovered that rats no longer displayed these symptoms; essentially, they had become cold-adapted and could tolerate low temperatures without exhibiting any further stress responses. After approximately 2 months of exposure to low temperatures, however, the rats lost this acquired resistance to low temperatures and died.

This process of coping with stressors was termed the general adaptation syndrome (GAS), and consists of three stages: (1) the alarm reaction, (2) resistance, and finally, (3) exhaustion. During the alarm reaction, stressors are detected. Coping with the stressor occurs during the resistance stage. The exhaustion stage is characterized by the termination of the stress response and the onset of stress pathology, which in extreme cases can lead to death. Although Selye believed that the exhaustion phase was due to the termination of the stress response, current ideas about stress indicate that the pathological effects of stress are due to the prolonged activation of the stress response. These three stages of the GAS correspond to the response of the rats to low temperatures after 2 days (alarm reaction), 2 weeks (resistance), and 2 months (exhaustion), respectively. If individuals have the opportunity to gradually move to low temperatures or are allowed to prepare for low-temperature conditions by first being housed for 8–10 weeks in short day lengths, then no stress responses occur.

Seasonal Adjustments in Glucocorticoid Secretion

Glucocorticoids are released in response to stressful stimuli and can affect reproduction, digestion, growth, multiple types of immune function, and virtually any other physiological process that requires significant energy. Predictably, seasonal variation in glucocorticosteroid secretion is more often the rule than the exception, probably because of the pivotal role these hormones play in orchestrating the distribution of resources among competing physiological activities.

Most animals studied to date display some seasonal variation in stress responses. For instance, Indian barnacles (*Balanus balanoides*) exhibit elevated antioxidant enzyme activity (a component of this species' stress response) prior to the monsoon season

compared with during it. Further, almost all wild amphibians, reptiles, and birds studied to date show marked changes in glucocorticoid concentrations during the year, with maximal values observed during the breeding season. Individuals of most mammalian species exhibit similar basal glucocorticosteroid concentrations year-round in the wild. However, this apparent pattern may reflect the paucity of studies on wild rodents rather than a lack of seasonality of this parameter. Indeed, rodents, such as deer mice (*Peromyscus* sp.), bred and maintained in the lab exhibit high corticosteroid concentrations when kept in long-day (breeding) conditions. Other species, however, such as Siberian hamsters (*Phodopus sungorus*), display elevated glucocorticosteroid concentrations when housed in short photoperiods.

Although it is clear from these studies that stress responses are temporally labile, two questions remain: (1) Is seasonal variation omnipresent within and among species (e.g., over a species' range) and (2) are other aspects of the stress response as seasonally labile as corticosteroid hormone concentrations? In terms of spatial variability, the results are mixed. Presumably, if photoperiod has significant effects on stress responses, then animals living at higher latitudes would exhibit higher circulating corticosteroid concentrations during breeding than animals at low latitudes simply because of the large variation in annual photoperiod. Redpolls (*Carduelis flammea*) are birds that live in the challenging arctic conditions of Alaska. Both males and females showed the typical marked elevation in blood corticosterone concentrations when captured for approximately 1 h, indicating that they responded similarly to many other vertebrates. However, the amount of the elevation in postcapture corticosterone varied across the year; corticosterone concentrations were generally blunted during January and maximal during June when birds were breeding. Generally, birds that had the highest fat stores secreted the lowest amounts of glucocorticoids postcapture. Similar patterns have been reported for other birds living at high latitudes, including the white-crowned sparrows (*Zonotrichia leucophrys gambelii*), snow buntings (*Plectrophenax nivalis*), and Lapland longspurs (*Calcarius lapponicus*). Across latitudes, however, results are mixed. Some species show little variability in circulating glucocorticoid concentrations (e.g., pied flycatchers, *Ficedula hypoleuca*, and mountain chickadees, *Poecile atricapilla*), whereas others (e.g., house sparrows, *Passer domesticus*, and white-crowned sparrows, *Zonotrichia leucophrys*) exhibit spatial variation. In common with spatial variability, the social system of animals can also affect seasonal variation in glucocorticoid responses. A study of greylag geese (*Anser anser*)

revealed that social status and physiological stress are not related in the same way in all seasons of the year.

In addition to hormones, other aspects of the glucocorticoid stress response show significant spatial and temporal flexibility. In particular, the HPA axis appears to be substantially malleable among wild songbirds. Individuals of different species can apparently modulate the levels of glucocorticoids they produce at different times of the year by changing the sensitivity of adrenergic tissues to upstream regulatory hormones such as corticotropin releasing hormone (CRH) or adrenocorticotrophic hormone (ACTH). Alternatively, some species are capable of modulating steroid-binding globulin levels in their blood across the year instead, although it remains unclear whether these globulins promote or inhibit the actions of the glucocorticosteroids to which they bind. Finally, corticosteroid receptors also appear to be seasonally variable. In house sparrows, membrane corticosteroid receptors are lowest during the nesting season, whereas cytosolic receptors are lowest during winter.

Why Are Stress Responses Seasonally Variable?

Presumably, stress responses help animals prepare for or recover from stressors. Why, then, are stress responses variable during the year when all individuals could benefit from optimal stress responses? Several hypotheses have been proposed. First, glucocorticoid stress responses are expensive and hence may not be affordable at all times of the year. Indeed, pyruvate, a readily usable energy source, blocked the stress-evoked suppression of cell-mediated and humoral immune activity in mice (*Mus musculus*). Second, high corticosterone concentrations, particularly when maintained for extended periods (i.e., chronic), can be immunosuppressive; over the short term, however, glucocorticosteroids can be immunoenhancing. In general, winter survival in small animals is expected to require a positive balance between short-day enhancement of immune defenses and glucocorticoid-induced immunosuppression. Immunosuppression caused by high glucocorticosteroids may be induced by many winter-related factors, including overcrowding, increased competition for scarce resources, low temperatures, reduced food availability, increased predator pressure, and lack of shelter. Moreover, winter breeding, with the potentially concomitant elevation in sex steroid hormones, may also compromise the immune system. Presumably, this explains why winter breeding predominantly occurs when other environmental stressors such

as temperature and food availability are not severe. Overall, the balance between enhanced immune function (i.e., to the point where autoimmune disease becomes a danger) and stress-induced immunosuppression (i.e., to the point where opportunistic pathogens and parasites overwhelm the host) must be met for animals to survive and become reproductively successful.

Several studies support this hypothetical construct. For instance, hamsters maintained in short photoperiods (and made to experience short-term stress) show more rapid wound healing and trafficking of leukocytes to the skin compared to unstressed short-day animals (or stressed and unstressed long-day hamsters). In one passerine bird, cutaneous immune responses to a T-cell mitogen (phytohemagglutinin, PHA) were less dramatic in animals induced into breeding than in those in a nonbreeding condition. Moreover, corticosterone appears to be immunosuppressive in the same species only in winter (not summer) conditions. Such immune–endocrine interrelationships are not always uniform between the sexes within a species. For example, acute stress enhances cell-mediated immune activity (delayed-type hypersensitivity to dinitrofluorobenzene) in female, but not male, Siberian hamsters. Indeed, ample evidence indicates that immune activity can vary dramatically between sexes; thus, it is plausible that the glucocorticoid mediators of immune defenses might be labile as well.

One other hormone that is integrally involved in vertebrate photoperiodism, melatonin, can affect the seasonal interplay between corticosterone and immune function. Generally, melatonin enhances immune activity, whereas glucocorticosteroids (at chronically high levels) compromise immune activity. Under certain circumstances, melatonin treatment may ameliorate the immunosuppressive effects of glucocorticosteroids. For instance, Siberian hamsters attenuated the duration of fever during simulated short winter day lengths, presumably to conserve energy. Hamsters housed in long days were injected with saline or melatonin before lights off each day for 1 or 6 weeks; then fever was assessed following injections of bacterial lipopolysaccharide (LPS). The fever duration was attenuated (32%) only in hamsters that decreased body mass, increased glucocorticoid concentrations, and displayed gonadal regression in response to 6 weeks of melatonin. These results suggest that long-term exposure to long-duration melatonin signals induce the physiological changes required for short-day immune responses, perhaps via interactions among hormones downstream from melatonin, such as cortisol and leptin.

In addition to their effects on immune function, corticosteroids can dramatically affect behavior, particularly behaviors associated with reproduction. Generally, stress responses are up- and down-regulated across seasons, but the way in which this happens depends on the social and physical environments in which species live. For instance, songbirds living at high latitudes are thought to mount weak corticosterone responses to stress because the abandonment of offspring that it might lead to would be maladaptive given that these species typically only have one opportunity per year to breed. In the tropics, however, corticosterone–behavior relationships may be different because reproductive efforts are generally less constrained by time.

Considered together, both the incidence and responses to stressors vary on a seasonal basis. The environmental regulation of seasonal changes in many traits, including reproduction, metabolism, and immunity, is primarily mediated by photoperiod. From a physiological prospective, the pineal hormone melatonin appears to coordinate seasonal adjustments in stress responses including changes in reproductive, metabolic, and immune function. To what extent the immune-enhancing effects of melatonin are unique to seasonally breeding rodents and whether it is a generalized feature of seasonal adjustments in stress responses among humans remain largely unspecified. Although physiological responses are important mediators of seasonal and photoperiodic changes in stress responses, behavioral alterations might also have an important role in reducing their energetic costs. The careful integration of the behavioral and physiological mechanisms underlying the seasonality of stress responsiveness at both ultimate and proximate levels should provide important and novel insights into the interaction among seasonal environmental factors, stressors, immune function, and the pattern of diseases and mortality. To date, most efforts have not advanced much beyond describing the seasonal changes that occur in responses to stressors. The next step is to identify the cellular and molecular mechanisms that mediate the effects of season and photoperiod on stress responsiveness and link variability in these mechanisms to fitness.

See Also the Following Articles

Alarm Phase and General Adaptation Syndrome; Allostasis and Allostatic Load; Glucocorticoids, Role in Stress;

Immune Response; Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress.

Further Reading

- Bilbo, S. D., Dhabhar, F. S., Viswanathan, S. A., et al. (2002). Short day lengths augment stress-induced enhancement of skin immune function. *Proceedings of the National Academy of Sciences USA* **99**, 4067–4072.
- Breuner, C. W. and Orchinik, M. (2002). Downstream from corticosterone: seasonality of binding globulins, receptors, and behavior in the avian stress response. In: Dawson, A. (ed.) *Avian endocrinology*, pp. 385–399. New Delhi: Narosa Publishing.
- Breuner, C. W., Orchinik, M., Hahn, T. P., et al. (2003). Differential mechanisms for regulation of the stress response across latitudinal gradients. *American Journal of Physiology* **285**, R594–R600.
- Martin, L. B., Gilliam, J., Han, P., et al. (2005). Corticosterone suppresses immune function in temperate but not tropical house sparrows (*Passer domesticus*). *General and Comparative Endocrinology* **140**, 126–135.
- Neigh, G. N., Bowers, S. L., Pyter, L. M., et al. (2004). Pyruvate prevents restraint-induced immunosuppression via alterations in glucocorticoid responses. *Endocrinology* **145**, 4309–4319.
- Nelson, R. J., Demas, G. E., Klein, S. L., et al. (2002). *Seasonal patterns of stress, immune function, and disease*. New York: Cambridge University Press.
- Nelson, R. J. (2004). Seasonal immune function and disease responses. *Trends in Immunology* **24**, 187–192.
- Niyogi, S., Biswas, S., Sarker, S., et al. (2001). Seasonal variation of antioxidant and biotransformation enzymes in barnacle, *Balanus balanoides*, and their relation with polycyclic aromatic hydrocarbons. *Marine Environmental Research* **52**, 13–26.
- Prendergast, B. J., Yellon, S. M., Tran, L. T., et al. (2001). Inhibition of cellular immune function by *in vitro* melatonin is modulated by photoperiodic history in Siberian hamsters. *Journal of Biological Rhythms* **16**, 224–233.
- Romero, L. M. (2002). Seasonal changes in plasma glucocorticoid concentrations in free-living vertebrates. *General and Comparative Endocrinology* **128**, 1–24.
- Romero, L. M. (2004). Physiological stress in ecology: lessons from biomedical research. *Trends in Ecology and Evolution* **19**, 249–255.
- Sapolsky, R. M., Romero, L. M. and Munck, A. (2000). How do glucocorticoids influence stress responses?: integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrine Reviews* **21**, 55–89.
- Yellon, S. M., Kim, K., Hadley, A. R., et al. (2005). Time course and role of the pineal gland in photoperiod control of innate immune cell functions in male Siberian hamsters. *Journal of Neuroimmunology* **161**, 137–144.

Seasonal Rhythms

L M Romero

Tufts University, Medford, MA, USA

© 2007 Elsevier Inc. All rights reserved.

Evidence for Seasonal Glucocorticoid Rhythms
Mechanisms Regulating Annual Rhythms of
Glucocorticoids
Why Do Seasonal Glucocorticoid Rhythms Exist?
Adaptive Significance

Glossary

<i>Basal glucocorticoids</i>	Plasma concentrations of glucocorticoids in the absence of a stressor.
<i>Corticotropin-releasing hormone (CRH)</i>	Hormone released by the hypothalamus.
<i>Glucocorticoid-binding globulin (CBG)</i>	Proteins in the plasma that bind glucocorticoids.
<i>Hypothalamic-pituitary-adrenal (HPA) axis</i>	The neuroendocrine pathway leading from the brain to the adrenals that culminates in glucocorticoid release.
<i>Permissive effect</i>	An indirect effect whereby a hormone indirectly allows another system to work better.
<i>Stressor</i>	A stimulus that elicits a release of glucocorticoids.

Evidence for Seasonal Glucocorticoid Rhythms

Glucocorticoid secretion is often considered an obligatory response to stressors. In a laboratory setting, stressors applied in similar circadian, developmental, and psychological contexts within the same species produce remarkably consistent magnitudes of glucocorticoid secretion. Recent evidence, however, indicates that glucocorticoid release is modulated seasonally in the majority of free-living species that have been studied to date. In other words, the magnitudes of both basal and stress-induced glucocorticoid concentrations vary throughout the year. An example, data for white-crowned sparrows (*Zonotrichia leucophrys*), is presented in [Figure 1](#).

Many recent studies have measured glucocorticoids in free-living wild animals by directly sampling from freshly caught individuals. Plasma samples

collected within a few minutes of capture are assumed to reflect prestress glucocorticoid concentrations because glucocorticoids start to increase 2–3 min after initiation of a stressor. Stress-induced glucocorticoid concentrations then reflect the animal's response to capture, handling, captivity, etc. and are assumed to be a measure of the hypothalamic-pituitary-adrenal (HPA) axis's ability to respond to a stressor.

These studies provide clear evidence that a majority of reptile, amphibian, avian, and mammal species seasonally modulate glucocorticoid concentrations. Males and females sometimes differ in whether they show seasonal rhythms, and occasionally a species will show a seasonal rhythm in basal but not stress-induced glucocorticoid concentrations (or vice versa). However, approximately 90% of all species that have been studied have at least one sex with a seasonal rhythm of glucocorticoid release.

The evidence for an annual glucocorticoid rhythm is generally much stronger for basal concentrations, primarily because more studies have focused on this time point. Furthermore, avian species often show a seasonal rhythm in basal but not stress-induced glucocorticoid concentrations, suggesting that basal and stress-induced glucocorticoid concentrations are regulated differently. Laboratory studies show that basal and stress-induced glucocorticoids have different regulatory mechanisms and serve different physiological functions that are mediated by separate receptors. It is not surprising, therefore, that the HPA axis might be regulated differently under basal and stress conditions, but these regulatory pathways can be further dissociated into different annual cycles.

Of those species that do show a seasonal rhythm, the seasonal peak is usually during the breeding period. There are three notable exceptions. First, the most robust rhythm in birds is a nadir during the prebasic molt when they replace their feathers. This is thought to protect the bird from the protein catabolism effects of glucocorticoids during the protein mobilization of feather replacement. Second, three desert-breeding birds show a seasonal glucocorticoid peak during the rainy season rather than the breeding period, suggesting that seasonal rhythms can be timed to environmental cues as well as seasonal physiological cues. Third, even though most mammal species show a seasonal rhythm, there is no consensus season when mammals tend to have elevated glucocorticoid concentrations.

The capacity of glucocorticoid-binding globulins (CBGs) has also been found to vary seasonally and

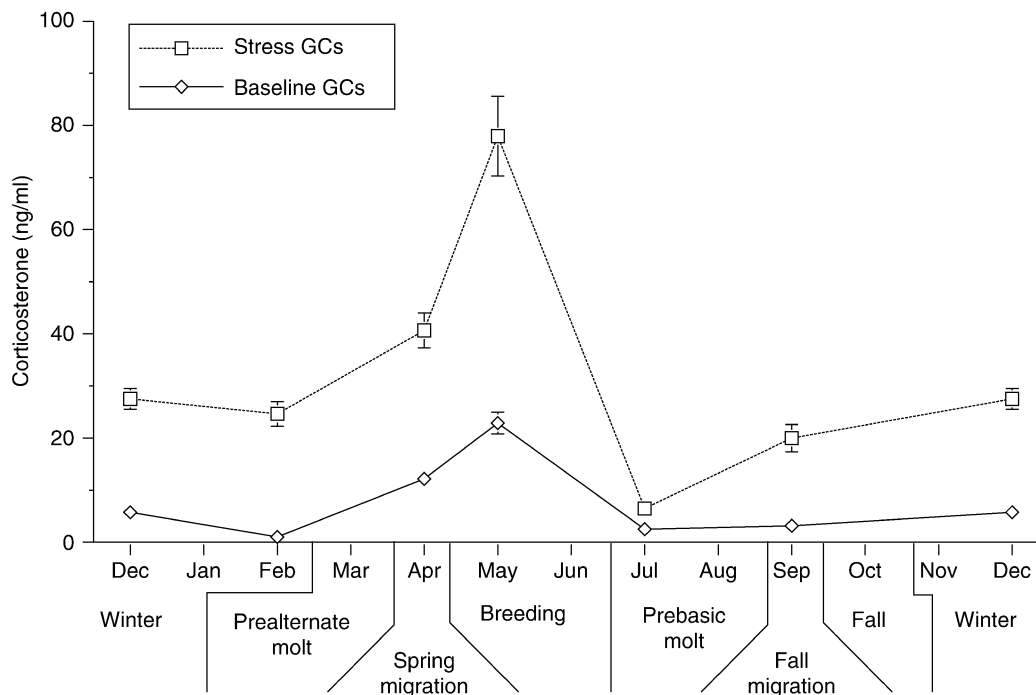


Figure 1 Seasonal rhythm of baseline and stress-induced corticosterone from white-crowned sparrows. Reprinted from Romero, L. M. (2002). Seasonal changes in plasma glucocorticoid concentrations in free-living vertebrates. *General and Comparative Endocrinology* 128, 1–24 with permission from Elsevier.

to be regulated by gonadal testosterone. Because only free steroid is generally considered to be available for diffusion into tissues, free steroid concentrations may be more biologically relevant than total steroid concentrations. When total and free steroid concentrations are compared, the seasonal rhythm in total glucocorticoids often changes or disappears entirely.

Many wild species held in long-term captivity also show seasonal rhythms, but these rhythms do not always match what is found in free-living animals. This likely reflects that both domestication and captivity alter HPA axis function in ways that are different for each species, making it difficult, if not impossible, to ascertain natural annual variation from these populations.

Mechanisms Regulating Annual Rhythms of Glucocorticoids

Progress is just beginning in understanding how HPA axis function is regulated seasonally. What induces these changes, be it photoperiod, temperature, food availability, etc., is currently unknown. Plasma glucocorticoid concentrations often are correlated with adrenal mass, and early studies showed increases in adrenal mass during breeding. Consequently, the ability of adrenal tissue to respond to adrenocorticotrophic hormone (ACTH) changes seasonally. A few studies in birds indicate that there are multiple regulatory

points in the HPA axis. Some species seasonally regulate glucocorticoid release from the adrenal, others seasonally regulate ACTH release from the pituitary, and still others regulate CRH and arginine vasotocin (the avian congener of arginine vasopressin) release from the hypothalamus. These regulatory control points may be related to how sensitive each species is to adverse environmental conditions.

Glucocorticoids also interact with the gonadal system. Although glucocorticoids inhibit gonadal hormone release, exogenous testosterone elevates glucocorticoids in free-living birds. This suggests a complex interaction between the gonadal and adrenal systems, with glucocorticoids downregulating gonadal androgen release concurrent with gonadal androgens elevating glucocorticoid release. Gonadal androgens, therefore, could potentially be an important physiological regulator of seasonal glucocorticoid rhythms.

Why Do Seasonal Glucocorticoid Rhythms Exist?

Modulation of glucocorticoid release presents an interesting paradox: if glucocorticoid release is so important for survival, then how do animals survive stressors at certain times of the year when they essentially fail to release glucocorticoids? There have been three major hypotheses proposed to explain seasonal glucocorticoid rhythms.

Energy Mobilization Hypothesis

The energy mobilization hypothesis focuses on the metabolic effects of glucocorticoids. Glucocorticoids play an important role in energy mobilization, especially during stress, so glucocorticoid concentrations should be highest during energetically costly times of the year. According to this hypothesis, high glucocorticoid concentrations during breeding result from breeding exerting the highest energetic demands during the year. Females of many species, for example, expend substantial energy during breeding, and breeding males of many species have increased energy costs associated with testosterone and territorial defense. Moreover, increased energy mobilization could explain seasonal peaks during nonbreeding periods if these periods are more energetically costly than breeding.

Two examples, however, mammalian hibernation and avian molt, highlight the insufficiency of the energy mobilization hypothesis in providing a universal explanation for seasonal glucocorticoid rhythms. Hibernating mammals require dramatically increased fat stores to survive the winter, but it is not clear that these energy requirements are greater or less than lactation. Second, feather replacement during molt in birds requires substantial energy expenditure, yet this period is the nadir of the annual rhythm. The energy mobilization hypothesis can be a powerful explanatory mechanism for certain species during some times of the year, but it clearly is insufficient to fully explain seasonal glucocorticoid rhythms.

Behavior Hypothesis

The behavior hypothesis focuses on the behavioral effects of glucocorticoids. It posits that annual glucocorticoid rhythms result from different requirements for expressing (or not expressing) glucocorticoid-mediated behaviors at different times of the year. For example, an important glucocorticoid-mediated behavior in wild animals might be fleeing an area and relocating during storms. This may be an excellent strategy for most of the year, but could be catastrophic for an individual's overall fitness when relocation requires abandoning young. Consequently, the behavior hypothesis suggests that the need to seasonally regulate the expression of those behaviors drives glucocorticoid rhythms.

The behavior hypothesis, however, does not fit a more general annual rhythm in glucocorticoid concentrations. Glucocorticoid concentrations are highest in most species during breeding, exactly when glucocorticoid concentrations should be lowest if seasonal glucocorticoid rhythms were primarily to regulate the behavioral consequences of glucocorticoid release.

Consequently, this hypothesis may be useful in explaining modulation of glucocorticoid concentrations within a single season, but not across seasons.

Preparative Hypothesis

The preparative hypothesis focuses on the coordinating aspects of glucocorticoid physiology and posits that seasonal rhythms in glucocorticoid concentrations serve to modulate the priming of other stress pathways during periods with different potential exposure to stressors. Glucocorticoids are only one part of the stress response. Other hormones (especially epinephrine and norepinephrine), neurotransmitters (e.g., 5-HT, CRH), opioid peptides, and cytokines (e.g., IL-6), as well as other brain functions, exert their effects rapidly with the onset of stress without involving the pituitary or adrenals. However, glucocorticoids have a permissive effect on many of these systems and thereby facilitate better performance under stress. Higher basal glucocorticoid concentrations may augment the priming effects on these systems in preparation for further stressors. This is the traditional explanation for timing the circadian peak for the beginning of the active period. These permissive effects may help prepare the organism for subsequent stress responses.

If the preparative hypothesis is correct, higher glucocorticoid concentrations should coincide with seasons in which predictable stressors are more likely. The available data fit this prediction moderately well. The risk of exposure to many stressors in wild animals likely varies seasonally. During the breeding season, for example, predation risk may increase when caring for young; competition for mates, with the risk of frequent fights, increases; and disease may become more prevalent when breeding individuals congregate. Consequently, glucocorticoid concentrations may vary seasonally in order to mediate changes in preparedness in the nonglucocorticoid pathways of the stress response. However, the existence of predictable seasonal increases in stressor risk has not yet been demonstrated.

Adaptive Significance

Proper concentrations of glucocorticoids seem crucial for survival. Too little can result in even mild stressors killing animals, but too much can lead to suppressed gonadal and immune function and ultimately disease. Consequently, a balance must be struck between needing glucocorticoids to survive stressors and modulating glucocorticoid secretion to prevent deleterious exposures. What is becoming clear is that there is a seasonal rhythm in what is the

right amount of glucocorticoids. Understanding why the right amount changes seasonally will help us understand why glucocorticoids are released during stress and how they help animals survive. The three explanatory hypotheses presented earlier are derived from different aspects of glucocorticoid physiology. The energy mobilization hypothesis emphasizes glucocorticoid's metabolic effects; the behavior hypothesis emphasizes the desired acute behavioral effects of glucocorticoids; and the preparative hypothesis emphasizes the regulatory role that glucocorticoids have in other stress systems. Integrating these hypotheses into a cohesive physiology will provide the foundation for understanding how elevated glucocorticoid concentrations aid in survival.

See Also the Following Articles

Circadian Rhythms, Effects of Prenatal Stress in Rodents; Circadian Rhythms, Genetics of; Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress; Corticosteroids and Stress; Cortisol Awakening Response; Corticosteroid Receptors; Glucocorticoid Negative Feedback; Hypothalamic-Pituitary-Adrenal.

Further Reading

Guillette, L. J., Cree, A. and Rooney, A. A. (1995). Biology of stress: interactions with reproduction, immunology and intermediary metabolism. In: Warwick, C., Frye, F. &

- Murphy, J. B. (eds.) *Health and welfare of captive reptiles*, pp. 32–81. London: Chapman & Hall.
- Romero, L. M. (2002). Seasonal changes in plasma glucocorticoid concentrations in free-living vertebrates. *General and Comparative Endocrinology* **128**, 1–24.
- Sapolsky, R. M., Romero, L. M. and Munck, A. U. (2000). How do glucocorticoids influence stress-responses? Integrating permissive, suppressive, stimulatory, and adaptive actions. *Endocrine Reviews* **21**, 55–89.
- Tyrrell, C. L. and Cree, A. (1998). Relationships between corticosterone concentration and season, time of day and confinement in a wild reptile (tuatara, *Sphenodon punctatus*). *General and Comparative Endocrinology* **110**, 97–108.
- Wingfield, J. C. (1994). Modulation of the adrenocortical response to stress in birds. In: Davey, K. G., Peter, R. E. & Tobe, S. S. (eds.) *Perspectives in comparative endocrinology*, pp. 520–528. Ottawa: National Research Council of Canada.
- Wingfield, J. C. and Romero, L. M. (2001). Adrenocortical responses to stress and their modulation in free-living vertebrates. In: McEwen, B. S. & Goodman, H. M. (eds.) *Handbook of physiology; section 7: The endocrine system; vol. IV: Coping with the environment: neural and endocrine mechanisms*, pp. 211–234. New York: Oxford University Press.
- Wingfield, J. C., Hunt, K., Breuner, C., et al. (1997). Environmental stress, field endocrinology, and conservation biology. In: Clemmons, J. R. & Buchholz, R. (eds.) *Behavioral approaches to conservation in the wild*, pp. 95–131. Cambridge, UK: Cambridge University Press.

Secretagogue

G Fink

Mental Health Research Institute, Melbourne, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G Fink, volume 3, pp 409–411, © 2000, Elsevier Inc.

Principles

Secretagogues Involved in the Neuroendocrine Response to Stress

Multiple Secretagogues in Other Systems: Exemplified by Growth Hormone Secretion

Sympathetic System: Role in Fight or Flight and Model for Exocytosis

Receptors and Intracellular Signaling Cascades

Conclusion

Glossary

Catecholamines

Adrenaline and noradrenaline are catecholamines. Neurotransmitters of the sympathetic nervous system (noradrenaline) and the adrenal medulla (adrenaline), they rapidly prepare the body for flight or fight in response to a stress. This is achieved mainly by adrenaline or noradrenaline stimulation of glucose release from the liver, increased heart rate, and redistribution of the blood circulation (by tissue-selective vasomotor control) from nonessential (e.g., the gut) to essential organs such as the brain, heart, and skeletal muscles.

Chemical neurotransmitter

A compound synthesized and released by a nerve cell (neuron) that activates or inhibits another neuron or other type of excitable target cell such as muscle.

<i>Chromagranins</i>	A class of acidic proteins present in large secretory granules (vesicles) in neuroendocrine tissues and neuroendocrine-derived tumors. Chromagranins were first identified in the chromaffin cells of the adrenal medulla, sympathetic nerves, and the pituitary gland. One of the chromagranins, secretogranin II, is the precursor of secretoneurin, a neuropeptide with a wide distribution in brain through all vertebrate phyla.
<i>Endocrine glands</i>	Glands that secrete signaling molecules, termed hormones, into the bloodstream, which transports the hormones to their target organs and tissues, where they exert their effects.
<i>Exocrine glands</i>	Glands that secrete to the outside of the body (i.e., into the gut, nasopharynx, bronchial tree, genitourinary system, and skin) as opposed to endocrine glands, which secrete hormones into the bloodstream. The acinar cells of the pancreas, which secrete gut enzymes, are key examples of exocrine cells in the context of this topic.
<i>Exocytosis</i>	The release of a secretory product, such as neurotransmitters, in packets or quanta. The secretory products are packaged by the Golgi apparatus of the cell in membrane-bound vesicles; exocytosis involves fusion of the vesicular membrane with the cell membrane, resulting in the release of the quantum of product inside the vesicle. This is thought to be the main mechanism of secretion in most secretory cells, including neurons.
<i>Ligand</i>	A molecule that binds or bonds specifically to another substance/compound – in biology, usually a receptor.
<i>Neurohormone</i>	A compound synthesized and released by a nerve cell that reaches its target by transport in special vessels, such as the hypothalamo-hypophyseal vessels in the case of the hypothalamic-pituitary system, or the systemic circulation for peptides such as oxytocin and vasopressin whose targets are in the periphery. Most, but not all, neurohormones are peptides.
<i>SNAP-25</i>	Synaptosome-associated protein of 25 kDa; one of the SNARE proteins involved in calcium-regulated exocytosis.
<i>SNARE proteins</i>	Soluble N-ethylmaleimide-sensitive factor attachment protein receptors; involved in calcium regulated exocytosis.
<i>Stimulus secretion coupling</i>	The link between a stimulus, usually membrane depolarization (axon potentials in the case of nerve cells) under physiological conditions, and the release of a neurotransmitter or neurohormone.

Principles

A secretagogue is an agent that promotes the secretion of hormones, neurohormones, chemical neurotransmitters, enzymes, or other compounds synthesized and secreted by cells. The selectivity or specificity of a secretagogue is determined by the structure of the secretagogue and that of its receptors on or within the cell. Secretagogues may be endogenous biological agents natural to the animal, exogenous compounds obtained from plants, or artificially synthesized pharmacological agents. The activity of exogenous secretagogues is thought to be due to the fact that the agent, by virtue of its chemistry and conformation, is capable of activating receptors specific to an endogenous secretagogue. There are numerous examples of this type of pharmacological mimicry: for example, acetylcholine (endogenous) and nicotine or muscarine (plant derived), opiates (plant derived), and the endogenous opioids – enkephalins and endorphins. The different classes of receptors that mediate secretagogue action are outlined later in the article.

Some secretagogues are nonspecific and will stimulate the release of most secretory substances. The most widely known and commonly used (experimentally) nonspecific secretagogue is a high concentration of extracellular potassium ions. High extracellular potassium concentrations trigger secretion by depolarizing the cell membrane, which induces a massive increase in intracellular calcium ions. Studies with calcium ionophores (compounds that induce calcium influx into cells) and calcium channel modulators in adrenaline-secreting chromaffin cells suggest that it is the rate of calcium ion influx rather than the absolute level of intracellular calcium ion reached that determines the amount of catecholamine released. Secretagogue-induced depolarization of cell membranes is associated with the release of neurotransmitters, neurohormones, some pancreatic and salivary gland enzymes, and the pancreatic hormone insulin. The coupling of stimulus-induced membrane depolarization with secretion is termed stimulus–secretion coupling, a term coined by W. W. Douglas based on his work on the neurohypophysis and the adrenal medulla.

Secretagogues Involved in the Neuroendocrine Response to Stress

Several secretagogues are involved in organizing the biological response to stress. Thus, stress, most frequently generated by an event recognized by the senses as dangerous or potentially dangerous, activates neuronal systems in higher brain centers that activate the main components of the limbic system, which include the hippocampus, amygdala, and

hypothalamus. This activation and the transmission of signals between nerve cells (neurons) involve the release of chemical neurotransmitters at nerve junctions (synapses). Nerve impulses generated as a consequence of stress activation of higher brain centers and the limbic system are transmitted to the paraventricular nuclei (PVN) of the hypothalamus, the major stress nuclei of the brain. These nuclei comprise neurons that synthesize and secrete corticotropin-releasing hormone-41 (CRH-41) and arginine vasopressin (AVP) and project to the median eminence of the hypothalamus, where they terminate on the primary plexus of the hypophyseal portal vessels. Released from PVN terminals in the median eminence and transported to the anterior pituitary gland by the hypophyseal portal vessels, CRH-41 and AVP act synergistically to stimulate the release of adrenocorticotropin (ACTH) into the systemic circulation. The PVN neurons, therefore, constitute the final common pathway neurons of the central neuroendocrine stress control system. Final common pathway neurons was a term first used by Sherrington to describe the alpha motor neurons of the spinal cord that project to and control skeletal muscle activity.

In humans and rodents, CRH-41 is the main secretagogue for ACTH release; its action is potentiated by AVP. In sheep, the situation is reversed, with AVP being the main ACTH secretagogue and CRH-41 potentiating the action of AVP. In addition to these ACTH-releasing factors, there is experimental evidence that under certain conditions the brain may inhibit ACTH release. A likely candidate for an ACTH-inhibiting factor is the cardiac hormone atrial natriuretic peptide (ANP), which is also synthesized in PVN neurons and released into hypophyseal portal blood.

Transported by the circulation to the adrenal glands, ACTH is the main secretagogue for stimulating the release of adrenal corticosteroids, and especially the glucocorticoids, cortisol in humans and corticosterone in rodents. Under certain situations, steroids can be secretagogues. Thus, estrogen, in its positive feedback mode for gonadotropin release, induces the release of gonadotropin-releasing hormone and prolactin, albeit possibly by indirect actions. Estrogen can also potentiate compound 48/80- or substance P-induced release of serotonin and histamine from mast cells. Glucocorticoids are potent inhibitors of ACTH release; this action is exerted directly on the anterior pituitary gland, where glucocorticoids inhibit the action of CRH-41 and AVP, as well as on the hypothalamus, where they inhibit CRH-41 and AVP release. This inhibitory effect of glucocorticoids is the basis for their negative feedback control of ACTH release. Whether corticosteroids are secretagogues is

not known; however, by exerting pronounced effects on protein synthesis and enzyme activity in nerve and other cells, corticosteroids modulate the action of secretagogues.

In addition to what might be called the classical secretagogues of the brain-pituitary-adrenal system, other secretagogues implicated in the control of stress-induced glucocorticoid secretion include urocortin (a homolog of CRH-41), angiotensin II, and cytokines such as tumor necrosis factor- α (TNF- α) and interleukin-1 β (IL-1 β). TNF- α and IL-1 β both stimulate ACTH release and are two of several secretagogues that link the immune and neuroendocrine defense systems. Extreme stress, such as that produced by multiple trauma or septic shock, results in a paradoxical decrease of the secretion of ACTH in the presence of increased cortisol secretion. It is postulated that this paradox may be brought about by the inhibition of ACTH secretion by ANP and by the direct adrenocortical stimulation of glucocorticoid secretion by the peptide endothelin-1.

Multiple Secretagogues in Other Systems: Exemplified by Growth Hormone Secretion

The existence of more than one secretagogue for pituitary hormones is not confined to ACTH. Thus, growth hormone (GH) release is regulated by two hypothalamic peptides or neurohormones: GH-releasing hormone, which stimulates GH release, and somatostatin, which inhibits GH release. Ghrelin, a peptide secreted by the stomach, is also an endogenous ligand for the GH secretagogue receptor. Ghrelin appears to be present also in the hypothalamus, but its precise mechanism of action and interactions with GH-releasing hormone and somatostatin needs to be elucidated. Finally, GH release is also induced by increased plasma concentrations of certain amino acids, such as arginine and lysine (which can therefore be regarded as GH secretagogues). However, while arginine (administered intravenously) is used as a provocative test of GH secretion in human subjects, there appears to be no robust evidence that amino acids administered orally stimulate detectable sustained GH release.

Sympathetic System: Role in Fight or Flight and Model for Exocytosis

The sympathetic component of the autonomic nervous system and the adrenal medulla play an equally important role in the stress response, especially in the acute phase when autonomic activation prepares the animal for fight or flight. That is, in response to stress, the brain through the hypothalamus and hindbrain

centers activates the sympathetic nervous system, which, through the release of noradrenaline at sympathetic nerve terminals and adrenaline from the adrenal medulla, stimulates several physiological changes that include an increase in heart rate and blood pressure and the rapid mobilization of glucose production by the liver.

Stimulation of adrenaline secretion from the adrenal medulla by the preganglionic secretagogue acetylcholine has been used extensively as a model for the investigation of secretagogue action and stimulus-secretion coupling. Based on the seminal electrophysiological studies of Bernard Katz and associates and confirmed by ultrastructural (electron microscopical) findings, it has long been accepted that stimulus-secretion coupling results in the release of neurotransmitters, hormones, and enzymes by way of exocytosis. Work on the adrenal medulla and sympathetic nerves proved important for elucidating the mechanism of exocytosis and the recycling and reuse of vesicular membranes, the characterization of the chromagrins, and the identification of peptides involved in exocytosis such as syntaxin, SNAP-25 and L- and N-type calcium channels. In brief, secretory vesicles formed at the trans-Golgi network of neuroendocrine and endocrine cells undergo translocation, docking, and priming before they are ready to fuse with the plasma membrane and release their secretory product into the extracellular space. This process, termed regulated exocytosis, is controlled by Ca^{2+} (using synaptotagmin) and is mediated by SNARE proteins. For details about this and vesicle priming, depriming and cell-membrane fusion see Stojilkovic (2005).

Receptors and Intracellular Signaling Cascades

There are two broad classes of receptors – cell surface and nuclear. Located intracellularly, the nuclear receptor superfamily binds steroid (e.g., cortisol, corticosterone, aldosterone, estradiol, testosterone) and thyroid hormones, derivatives of vitamins A and D, and xenobiotics. Cell surface receptors are composed of the following classes (examples of ligands):

1. Ligand-gated ion channels (e.g., nicotinic acetylcholine receptor)
2. Receptor tyrosine kinases (e.g., insulin, insulin-like growth factor, epidermal growth factor)
3. Receptor serine/threonine kinases (e.g., activin, inhibins)
4. Receptor guanylate cyclase (e.g., atrial natriuretic peptide)

5. G-protein-coupled receptor superfamily (e.g., adrenergic agents, muscarinic acetylcholine agents, peptide hormones and neurohormones, glycoproteins, glucagon, parathyroid hormone)
6. Cytokine receptors (e.g., growth hormone, prolactin, leptin)

With the exception of glucocorticoids, cell surface receptors are the targets of all the secretagogues involved in the neuroendocrine response to stress. Thus, CRH-41, vasopressin, ACTH, adrenaline, and noradrenaline all bind to specific G protein-coupled receptors, while ANP binds to guanylate cyclase receptors. Acetylcholine, the preganglionic secretagogue that triggers adrenaline release from the adrenal medulla, acts by way of nicotinic acetylcholine receptors on adrenal medullary cells.

Secretion induced by ligand activation of non-voltage-dependent cell surface receptors is triggered by one of a series of intracellular signaling cascades, the most well known of which are the adenylyl cyclase, phosphatidyl inositol, and kinase systems. As indicated previously, secretagogue-induced secretion involves a rapid increase in the intracellular concentrations of calcium ions, triggered either by effects on the cell surface calcium channels in the case of voltage-dependent receptors, or by the intracellular signaling cascades in the case of non-voltage-dependent receptors. In addition to stimulating secretion, activation of cell surface receptors frequently results in gene transcription that can lead to synthesis of secretory product or a key enzyme involved in the synthesis of secretory product. The latter is exemplified by the fact that acetylcholine, which induces the secretion of adrenaline from the adrenal medulla and noradrenaline from postganglionic neurons of the sympathetic system, also induces the synthesis of tyrosine hydroxylase, the rate-limiting enzyme for the synthesis of noradrenaline and adrenaline. This action of acetylcholine appears to be mediated by both nicotinic voltage-dependent receptors and muscarinic G protein-coupled receptors.

Conclusion

The mechanism of action of secretagogues is the subject of intensive research and is one of the major themes of modern molecular biology and molecular pharmacology. Calcium ions play a pivotal role in the action of secretagogue-induced secretion by exocytosis. Although a good deal is known about secretagogue action, a more rigorously defined understanding is likely to emerge from modern molecular genetic and microphysiological approaches.

See Also the Following Articles

Adrenal Medulla; Feedback Systems; Glucocorticoid Negative Feedback; Glucocorticoids, Role in Stress; Neuroendocrine Systems.

Further Reading

- Craske, M., Takeo, T., Gerasimenko, O., et al. (1999). Hormone-induced secretory and nuclear translocation of calmodulin: oscillations of calmodulin concentration with the nucleus as an integrator. *Proceedings of the National Academy of Sciences USA* **96**, 4426–4431.
- Douglas, W. W. (1968). Stimulus secretion coupling: the concept and clues from chromaffin and other cells. *British Journal of Pharmacology* **34**, 453–474.
- Douglas, W. W. (1973). How do neurones secrete peptides? Exocytosis and its consequences, including “synaptic vesicle” formation, in the hypothalamo-neurohypophyseal system. *Progress in Brain Research* **39**, 21–39.
- Douglas, W. W. (1974). Involvement of calcium in exocytosis and the exocytosis-vesiculation sequence. *Biochemical Society Symposia* **39**, 1–28.
- Eder, U., Fischer-Colbrie, R., Kogner, P., et al. (1998). Levels and molecular forms of chromagranins in human childhood neuroblastomas and ganglioneuromas. *Neuroscience Letters* **253**, 17–20.
- Fink, G., Robinson, I. C. A. F. and Tannahill, L. A. (1988). Effects of adrenalectomy and glucocorticoids on the peptides CRF-41, AVP and oxytocin in rat hypophyseal portal blood. *Journal of Physiology* **401**, 329–345.
- Fink, G., Dow, R. C., Casley, D., et al. (1992). Atrial natriuretic peptide is involved in the ACTH response to stress and glucocorticoid negative feedback in the rat. *Journal of Endocrinology* **135**, 37–43.
- Gerasimenko, O. V., Gerasimenko, J. V., Petersen, O. H., et al. (1996). Short pulses of acetylcholine stimulation induce cytosolic Ca^{2+} signals that are excluded from the nuclear region in pancreatic acinar cells. *Pflügers Archive* **432**, 1055–1061.
- Larsen, P. R., Kronenberg, H. M., Melmed, S. and Polonsky, K. S. (eds.) (2003). *Williams textbook of endocrinology* (10th edn.) Philadelphia, PA: Saunders.
- Leitner, B., Schneitler, C., Klocker, H., et al. (1998). Formation and sequence analysis of secretoneurin, a neuropeptide derived from secretogranin 11, in mammalian, bird, reptile, amphibian and fish brains. *Neuroscience Letters* **248**, 105–108.
- Leitner, B., Lovisetti-Scarniorn, P., Heilmann, J., et al. (1999). Subcellular localization of chromagranins, calcium channels, amine carriers, and proteins of the exocytotic machinery in bovine splenic nerve. *Journal of Neurochemistry* **72**, 1110–1116.
- Mogami, H., Nakano, K., Tepikin, A. V., et al. (1997). Ca^{2+} flow via tunnels in polarized cells: recharging of apical Ca^{2+} entry through basal membrane patch. *Cell* **88**, 49–55.
- Petersen, O. H. (1996). Can Ca^{2+} be released from secretory granules on synaptic vesicles? *Trends in Neuroscience* **19**, 411–413.
- Petersen, O. H. and Iwatsuki, N. (1978). The role of calcium in pancreatic acinar cell stimulus-secretion coupling: an electrophysiological approach. *Annals of the New York Academy of Sciences* **307**, 599–617.
- Petersen, O. H., Petersen, C. C. and Kasai, H. (1994). Calcium and hormone action. *Annual Review of Physiology* **56**, 297–319.
- Stojilkovic, S. S. (2005). Ca^{2+} -regulated exocytosis and SNARE function. *Trends in Endocrinology and Metabolism* **16**, 81–83.
- Turnbull, A. V., Dow, R. C., Hopkins, S. J., et al. (1994). Mechanisms of activation of the pituitary-adrenal axis by tissue injury in the rat. *Psychoneuroendocrinology* **19**, 165–178.
- Vliagotis, H., Dimitriadou, V., Boucher, W., et al. (1992). Estradiol augments while tamoxifen inhibits rat mast cell secretion. *International Archives of Allergy and Immunology* **98**, 398–409.

Seizures See: Epilepsy.

Selective Serotonin Reuptake Inhibitors (SSRIs)

M Gitlin

Geffen School of Medicine at the University of California,
Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Serotonin and the Selective Serotonin Reuptake Inhibitors
Stress and the Brain

Stress and Serotonin

Antidepressants, SSRIs, and Stress Responses

Glossary

<i>Brain-derived neurotrophic factor (BDNF)</i>	A brain peptide that enhances growth of neurons.
<i>Hippocampus</i>	A brain structure vital for learning and memory.
<i>Selective serotonin reuptake inhibitors (SSRIs)</i>	A class of antidepressants that have the dominant initial biological effect of increasing serotonin in the synaptic cleft.
<i>Serotonin</i>	A neurotransmitter that is involved in the regulation of mood, anxiety, and stress responsiveness, among other effects.

Serotonin and the Selective Serotonin Reuptake Inhibitors

Selective serotonin reuptake inhibitors (SSRIs) are the dominant class of antidepressants prescribed in the Western world. Although they are not more effective than other classes of antidepressants, their ease of use (most patients show optimal benefit from one daily dose of any of the SSRIs), their relatively benign side effect profile, and the extraordinary marketing campaigns by the pharmaceutical manufacturers of the individual agents have made them the most commonly prescribed antidepressants and among the most commonly prescribed agents in all of medicine.

All SSRIs share the common trait of blocking the serotonin transporter, the receptor on the presynaptic neuron that regulates the reuptake of serotonin back into the presynaptic neuron. The functional consequence of blocking serotonin reuptake is to allow serotonin to remain in the synaptic cleft longer, thereby enhancing its likelihood of hitting the postsynaptic serotonin receptor. The end result of this is to enhance serotonin neurotransmission. The strength of serotonin activity at serotonin receptors is

inversely proportional to the number of serotonin transporter molecules present at the presynaptic membrane. Although the mechanism by which SSRIs ameliorate depression is far more complex than this, the initial enhancement of serotonergic function by reuptake blockade is generally thought to be the first step in a cascade of biological effects.

Although originally released and marketed as treatments for depression, SSRIs have now been demonstrated to be effective in a wide variety of psychiatric disorders, including panic disorder, social anxiety disorder, obsessive-compulsive disorder, generalized anxiety disorder, posttraumatic stress disorder (PTSD), and some of the symptoms of certain personality disorders. Thus, these powerful serotonin enhancers seem to be effective treatments in a wide variety of mood and anxiety disorders. Many, but not all, of these disorders can be triggered or exacerbated by acute or chronic stress. For some individuals – specifically those whose depressive episodes are triggered by acute or chronic stressful life events – depression can be described as the end product of failed adaptation to chronic emotional stress. Because of the overlap among anxiety disorders, depression, and stress, the remainder of this article reviews all three, even though, of course, these are similar but not identical constructs or clinical conditions.

Stress and the Brain

Stress alters the brain in a variety of ways. Two of these stress-related effects – decreases in hippocampal volume and the diminishing of the activity of a brain peptide, brain-derived neurotrophic factor (BDNF) – are highly relevant in understanding the potential effects of antidepressants and SSRIs, in particular in their altering responses to stress.

A decrease in the volume of the hippocampus, the brain structure vital for learning and memory, is among the most consistent consequences of stress. In rodents, experimental chronic stress causes the loss of hippocampal neurons. Similar findings have been demonstrated in humans suffering from depression and in other stress-related psychiatric disorders such as PTSD. Lower hippocampal volumes in depressed patients compared with normal people are seen bilaterally, although this finding is more consistent on the left versus right hippocampus. Similar findings are seen with PTSD, even in those patients not suffering from comorbid depression. In addition, in two studies, the reduction in hippocampal volume

correlated with the length of the depressive illness. It is as yet unclear whether these changes in actual brain structure seen with stress versus psychiatric disorders such as depression or PTSDs are similar or identical.

The second important stress-related effect on brain function is in the decreasing activity of the brain peptide BDNF. BDNF plays a major role in the survival, growth, and protection of neurons, both in the central and the peripheral nervous systems. The role of BDNF in neuronal function continues into adult life; thus, BDNF should be considered as playing an ongoing regulatory role for optimal neuronal function in youngsters and adults. Normally, neurons are rewired and shaped (with the sprouting of dendrites, called dendritic arborization) throughout adult life. BDNF seems integrally involved with this process. In addition, the production of new neurons in adults, called neurogenesis, seems also to be regulated in part by BDNF. Of note, BDNF shows the greatest expression in the hippocampus, the brain region most consistently demonstrated to show morphological changes in stress and depression.

In animals, a variety of stressors (such as forced immobilization in rats or learned helplessness following exposure to inescapable shock) consistently cause a dramatic reduction in BDNF expression in the hippocampus. There is some evidence that the prolonged glucocorticoid secretion (see **Glucocorticoids, Effects of Stress on, Glucocorticoids, Role in Stress**) that is consistently observed in association with severe stress may mediate some of the effects of stress and depression on decreased BDNF function.

Human studies of BDNF are more difficult to obtain because direct measurements of brain are not possible to obtain and the relationship between blood and brain levels of BDNF is still uncertain. Nonetheless, in one study, serum BDNF levels in depressed patients were reported to be lower than those seen in normal controls. In addition, in another study, serum BDNF levels correlated inversely with depression rating scale scores.

Stress and Serotonin

Central nervous system responses to stress are mediated by multiple neurotransmitter systems (see **Primates: Rearing and Effects of Stress on Primate CNS Function, Stress and CNS Arousal; Genomic Contributions**). Consistent evidence suggests that the serotonin system plays an important role in the brain's response to stress, thereby further suggesting that modifying serotonergic function may alter responses to stress. In animals, low serotonin is associated with increased sensitivity to environmental cues and greater responsiveness to threat. In humans,

this implies that individuals with low serotonin may be more easily traumatized by stressful life events (or have a lower threshold for showing the effects of traumatic events) than those with higher baseline serotonin levels.

The likelihood of stress triggering depressive symptoms and disorders seems to be moderated in part by different profiles of the gene that helps regulate the serotonin transporter (i.e., the receptor that regulates the reuptake of serotonin into the presynaptic neuron). In a seminal study, young adults with either one or two copies of the short allele of the gene that helps regulate the serotonin transporter were more likely to become depressed (as measured by either symptoms or disorders) but only when stressed by life events such as employment problems, health issues, housing or relationship issues, and so forth. In addition, childhood maltreatment similarly predicted adult depression only among those with the short allele of the serotonin transporter gene. The implication of this study is that responses to stress and the vulnerability to becoming depressed due to the stress can be predicted by a genetic measure of serotonin function.

Similarly findings can be seen in stressed non-human primates. Infant macaque monkeys raised in peer groups were compared to those reared with their mothers. (Lack of normal mothering can be considered an early stress or trauma in both nonhuman primates and in humans.) The biological marker for stress was adrenocorticotrophic hormone (ACTH) and cortisol levels, both of which are altered in stressful conditions (see **Adrenocorticotrophic Hormone (ACTH), Hypocortisolism and Stress**). Monkeys with the short allele of the serotonin transporter gene showed greater alterations in ACTH and cortisol levels during separation from mothers than those with the longer allele. Thus, these experimental results in monkeys are entirely consistent with the results seen in young adult humans.

Another clue that serotonin may be intimately linked to the effects of stress was demonstrated in a study of infant rhesus monkeys who were raised in mother-reared versus peer-reared conditions. Early maternal deprivation was associated with decreased serotonin metabolites in the cerebrospinal fluid (CSF), a marker for serotonin function in the central nervous system. Furthermore, this decrease in serotonergic function stayed stably lower for the 5 months of the study, implying a potentially long-term effect of stress and trauma on brain serotonin activity.

The regulation of BDNF by specific neurotransmitters is assuredly complex. However, there is some evidence that serotonin plays a role in neurogenesis and may, therefore, be linked with BDNF function. As an example, experimental lesions of the serotonin

system have been shown to decrease neurogenesis in animals. Serotonin agonists increase neurogenesis, an effect blocked by serotonin antagonists. Thus, medications that alter serotonin function, such as SSRIs, can theoretically be expected to increase neurogenesis. Not only do antidepressants enhance BDNF function, but BDNF has been reported to be a potent neurotrophic factor for norepinephrine and serotonin neurons. Thus, if these preliminary data hold true, the relationship between antidepressants and BDNF is best described as reciprocal or bidirectional.

Antidepressants, SSRIs, and Stress Responses

Given that markers of stress and stress-related effects on brain function include both decreasing hippocampal size and diminishing BDNF activity, antidepressants might exert therapeutic effects by preventing or reversing both these effects. Evidence supports both these possibilities.

Antidepressants have been demonstrated to increase neurogenesis in rats. Of note, the antidepressant classes that have shown these positive effects include SSRIs, tricyclic antidepressants, and monoamine oxidase (MAO) inhibitors. The positive effects on neuronal remodeling and neurogenesis seem to require 2–4 weeks of treatment, similar to the time frame of clinical response to these agents in depressed individuals. Electroconvulsive therapy (ECT), with its well-documented antidepressant effects, also increases BDNF expression, as does transcranial magnetic stimulation (TMS), a more recently described antidepressant treatment. The mood stabilizers lithium and valproate, used primarily in the treatment of bipolar disorder, show similar effects. Of note, psychotropic agents from other classes such as antipsychotics and opiates are not associated with increased BDNF activity, implying that the effects may be relatively specific to agents with antidepressant or mood-stabilizing effects.

In addition, as hypothesized, antidepressants also reverse the decreased neurogenesis seen in association with chronic stress and can block this effect of chronic stress when administered preventively. This effect, seen in animal experiments, is entirely consistent with the human evidence that antidepressants administered on a long-term maintenance basis can prevent depressive episodes in individuals vulnerable to them.

It must be acknowledged, however, that even in SSRI studies not all the data demonstrate the salutary effects of these agents on increasing BDNF, neurogenesis, or reversing hippocampal atrophy. In some studies, the expected positive effect is simply not

observed. In other rodent studies, the behavioral effects of antidepressants were seen but the enhancement of BDNF function was not, implying that BDNF enhancement was not necessary for the antidepressant behavioral effects.

Interestingly, one factor other than antidepressants that can increase neurogenesis/cellular remodeling in the hippocampus is exercise. In animals, exercise plus antidepressants enhanced BDNF function more than antidepressants. This may partially explain the mild antidepressant effect of aerobic exercise in humans.

Similar observations supporting the role of SSRIs in reversing stress-induced changes in the central nervous system have been made with anxiety models in animals and anxiety disorders in humans. Maternally deprived animals with increased acoustic startle responses and enhanced activity of the hypothalamic-pituitary-adrenal (HPA) axis showed a reversal of these effects with the SSRI paroxetine over a 3- to 4-week time frame (similar to the expected time to clinical effect in anxiety disorders in humans.) In humans, in one study, paroxetine both improved symptoms of PTSD but also enhanced cognitive performance and increased hippocampal volume.

Although all antidepressants may ultimately demonstrate the type of positive effect on reversing and/or preventing stress responses, there may be some differences across antidepressant classes. As one example, although some studies are positive, there are a number of other studies suggesting that tricyclic antidepressants do not consistently increase BDNF, as do other antidepressant classes such as SSRIs.

Consistent with the notion that there may be differences across antidepressant classes in modulating stress responses, a recent study compared the effect of an SSRI to reboxetine, an antidepressant that enhances norepinephrine (but not serotonin) in normal individuals responding to positive and negative images. Although both antidepressants diminished responses to negative emotional material, only the SSRI (citalopram) abolished the increased startle response in the context of negative affective images. The implication of this observation is that antidepressants with strong effects on serotonin, such as SSRIs, may be able to selectively block reactions to negative affective situations that may therefore enhance individuals' capacity to resist the negative effects of stress by blunting the affective and biological response to it.

Because virtually all the research discussed in this chapter is recent, only further studies, many of which are ongoing, will establish how consistently antidepressants reverse stress-related brain changes and whether these effects are common to all antidepressant classes and not just to SSRIs.

See Also the Following Articles

Adrenocorticotrophic Hormone (ACTH); Antidepressant Actions on Glucocorticoid Receptors; Cytokines, Stress, and Depression; Depression Models; Glucocorticoids, Effects of Stress on; Glucocorticoids, Role in Stress; Hippocampus, Overview; Depression, Immunological Aspects; Primates: Rearing and Effects of Stress on Primate CNS Function; Hypocortisolism and Stress; Stress and CNS Arousal; Genomic Contributions.

Further Reading

- Campbell, S., Marriott, M., Nahmias, C., et al. (2004). Lower hippocampal volume in patients suffering from depression: a meta-analysis. *American Journal of Psychiatry* 161, 598–607.
- Caspi, A., Sugden, K., Moffitt, T. E., et al. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 301, 386–389.
- Duman, R. S. (1998). Novel therapeutic approaches beyond the serotonin receptor. *Biological Psychiatry* 44, 324–335.

- Duman, R. S., Heninger, G. R. and Nestler, E. J. (1997). A molecular and cellular theory of depression. *Archives of General Psychiatry* 54, 597–606.
- Duman, R. S., Nakagawa, S. and Malberg, J. (2001). Regulation of adult neurogenesis by antidepressant treatment. *Neuropsychopharmacology* 25, 836–844.
- Harmer, C. J., Shelley, N. C., Cowen, P. J., et al. (2004). Increased positive versus negative affective perception and memory in healthy volunteers following selective serotonin and norepinephrine reuptake inhibition. *American Journal of Psychiatry* 161, 1256–1263.
- Hashimoto, K., Shimizu, E. and Iyo, M. (2004). Critical role of brain-derived neurotrophic factor in mood disorders. *Brain Research Reviews* 45, 104–114.
- Kramer, P. D. (2005). *Against depression*. New York: Viking Penguin.
- Nemeroff, C. B., Bremner, J. D. and Foa, E. B. (2006). Post-traumatic stress disorder: a state-of-the-science review. *Journal of Psychiatric Research* 40, 1–21.
- Russo-Neustadt, A. A. and Chen, M. J. (2005). Brain-derived neurotrophic factor and antidepressant activity. *Current Pharmaceutical Design* 11, 1495–1510.

Self-Esteem, Stress, and Emotion

D Roger

University of Canterbury, Christchurch, New Zealand

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D Roger, volume 3, pp 412–416,

© 2000, Elsevier Inc.

The Concept of Self-Esteem

Stress and Personality

Stress and Self-Esteem

Conclusion

Glossary

Emotional inhibition

The inhibition of the expression of experienced emotion.

Factor analysis

A statistical technique for identifying clusters of related items (factors) in questionnaires and for identifying redundant items in preliminary questionnaires. Deciding how many factors to extract is usually based on either an eigenvalue-1 criterion or a scree test,

Hypothalamic-pituitary-adrenal (HPA) axis

with the latter providing the more conservative estimate.

The link between the hypothalamus, the pituitary, and the adrenal cortex and medulla. The activation of the adrenal medulla in response to demands, including perceived stress, results in elevated secretion of catecholamines such as epinephrine and norepinephrine. Elevated secretion of adrenal corticosteroids such as cortisol also occurs in these circumstances via the release of corticotrophic hormones from the pituitary gland.

Neuroticism

The tendency toward emotional lability or emotional sensitivity.

Personality traits

Enduring, stable predispositions to behave in a particular way.

Rumination

The tendency to persist in thinking about emotionally upsetting events that have happened in the past or might occur in the future; also called emotional rehearsal.

Self-esteem

An evaluative measure of how individuals feel about themselves; includes perceived attractiveness, competence, and confidence.

<i>States</i>	Unstable aspects of personality, such as moods, that vary according to circumstances.
<i>Type</i> <i>A behavior pattern</i>	(TABP) Behavior, such as hostile competitiveness and time urgency, that is thought to increase individuals' susceptibility to coronary heart disease.

Self-concept is a central theme in personality psychology. For the purpose of controlled studies, it is usually operationalized as self-esteem, and a number of self-report psychometric scales have been developed to assess the degree to which individuals value themselves. These self-evaluative perceptions are a reflection of the way individuals feel about their attractiveness and their competencies, and the concept of self-esteem has an intuitive logic in the sense that almost everything people do is evaluated in terms of how successful or valued they feel they are.

Self-esteem clearly has important psychological implications, particularly for stress and personality, but the research has been hampered by conceptual and methodological problems. Many of the early definitions were overinclusive, and the assessment of self-esteem has been marred by psychometric shortcomings. However, work over the past 2 decades has begun to address these issues, leading to a clearer and more precise understanding of self-esteem.

The Concept of Self-Esteem

As early as 1890, William James described self-esteem as being dependent on the match between aspirations (which he called pretensions), on the one hand, and achievements or successes, on the other; the greater the discrepancy between aspiration and actual achievement, the lower self-esteem is likely to be. This suggests a constantly changing process, but self-esteem can also be defined as a relatively permanent predisposition or trait. Indeed, James acknowledged that people tend to have a more or less consistent view of their own self-esteem, developed over time and relatively unaffected by circumstance. Other early theories on the self-concept suggested that the opinions of significant others constituted a kind of mirror in which we view and evaluate ourselves. Although this social self-esteem might vary according to the criticism or the praise we receive, these perceptions became generalized to form an enduring composite of self-attitudes.

Recent trends in research on the self have revived the distinction between relatively stable trait self-esteem and unstable state self-esteem and have provided evidence to show that the distinction is empirically meaningful. For example, assessing

self-esteem repeatedly over successive days yields groups with relatively stable or unstable self-esteem. When these two groups are given either positive or negative feedback on their social skills, they differ in their responses, with the unstable high self-esteem subjects generally having more favorable responses to positive feedback but more defensive reactions to negative feedback. The results of these studies suggest that people with unstable high self-esteem are in fact characterized by a poorly developed self-concept and are more reliant on external evaluations and opinions than those with stable high self-esteem.

If the distinction between stable and unstable tendencies were not made, all of these subjects would have had apparently high self-esteem, and the questionnaire responses of the unstable subjects masked an underlying uncertainty and lack of confidence. This approach has been echoed in research on perceived control, which at the extremes of a normal distribution distinguishes between subjects with either an internal or an external locus of control. The former consider their behavior to be a consequence of their own actions, whereas the latter typically attribute events in their lives to luck, chance, or powerful others. However, a further distinction can be drawn between congruent and defensive externals – those who genuinely believe in fate or luck (congruents) and those who say they do in order to protect themselves in the event of failure (defensives). The defensive subjects are identified by their scores on measures of social desirability, which are thought to reflect a desire for approval.

In the context of self-esteem, a similar distinction can be drawn between subjects with high self-esteem and a high need for approval and those with high self-esteem but a low need for approval, based on their scores on social desirability. When these groups are asked to perform ambiguous tasks that allow believable success and failure feedback to be given and are then invited to indicate what they would like their ideal performance to be on a subsequent task, the subjects whose ideal is most diminished by the failure feedback on the analogous task are those with defensive high self-esteem. By contrast, high self-esteem subjects with a low need for approval are unaffected by such experimental manipulations.

More recent research has shown that the degree to which self-esteem is affected by feedback is moderated by other psychological variables, including emotional intelligence (EQ). Self-esteem is positively correlated with measures of EQ, but when individuals with either high or low EQ are exposed to laboratory manipulations that induce negative affect, those with high EQ showed less reduction in self-esteem than those with low EQ.

These effects have been borne out in more naturalistic settings that have incorporated other correlates of self-esteem, such as the degree of perceived support from significant others. A study of distress among cancer patients showed that a perceived lack of social support led to lower self-esteem, which in turn was linked to greater distress. Not surprisingly, self-esteem is also strongly negatively correlated with depression, and a related study of the link between negative marital interactions and depression in arthritis sufferers showed that the relationship was significantly moderated by self-esteem. Conversely, prospective research designs have shown that having social support is associated with high self-esteem, which in turn increases optimism while decreasing depression.

The overall findings confirm that the construct does indeed have two facets: a relatively stable, resilient self-regard, and a more changeable aspect that depends on external circumstances. Although people with genuinely high underlying self-esteem are less affected by a perceived lack of competence in particular situations than those with either low self-esteem or unstable high self-esteem, exposure to self-critical conditions results in diminished self-esteem. An important question raised by these findings is whether self-esteem can systematically be enhanced by targeted training interventions. One study that suggested that it can be involved the random allocation of physically disabled women to two training groups, one of which included a self-esteem enhancement program. The follow-up showed that the women in the experimental group that received the self-esteem training had significantly improved self-esteem as well as lower levels of depression.

However, one of the problems with experiments on self-esteem is the question of reliability. Self-esteem is typically assessed by self-report questionnaires, and it is difficult to ensure that people will respond honestly, particularly where the questions have a personal, evaluative connotation. The problem of response bias is inherent in questionnaire assessment techniques and can only adequately be overcome by using some form of control scale such as social desirability.

Another psychometric problem inherent in scale construction is the distinction between general and multidimensional structure. For example, one of the more widely used self-esteem scales is the 10-item Rosenberg Self-Esteem Inventory. This has the advantage of being extremely brief and easy to administer, but it is usually considered to be unidimensional and it does not break down into different domains of self-esteem. Other popular measures that offer systems for scoring responses into subscales covering areas such as academic, family, and social self-esteem are also thought to represent a single general factor,

despite the breakdown into different domains. By contrast, many of the more recent self-esteem scales are predicated specifically on multidimensional models of self-concept. The subscales within these questionnaires typically cover broad domains of physical, social, and academic self-esteem, and there is evidence to show that the subscales are empirically discriminable. However, the most satisfactory of these models are also hierarchical and include a higher-order general factor. Indeed, when conservative techniques such as scree tests are used for extracting factors rather than an eigenvalue-1 criterion, a single general factor is usually indicated for self-esteem questionnaires.

In principle, higher-order general factors in personality are more likely to be independent of one another than lower-order ones. Neuroticism and extraversion, for example, are postulated as statistically independent dimensions. However, self-esteem correlates significantly negatively with neuroticism and positively with extraversion, and it correlates significantly with a wide range of other personality variables as well. This is perhaps not altogether surprising in view of the degree of item-overlap among them. Neuroticism scales are a case in point. They include items that are explicitly concerned with self-esteem, such as "Are you troubled with feelings of inferiority?" and "Do you worry a lot about your looks?"

Setting aside the effects of item-overlap, it may be that the patterning of these other personality dimensions is, in fact, determined by individual differences in self-esteem – in other words, that self-esteem might represent a highest-order construct. What is clear from the research findings is that self-esteem exerts a powerful influence on behavior, and this can be captured by a compensatory model that formalizes the relationship between self-esteem and other aspects of personality, with particular reference to stress and health.

Stress and Personality

Stress has been defined in a variety of ways, with each definition having important implications for the way in which it is measured. Stimulus definitions, for example, led to the development of life-event scales for assessing stress. These instruments continue to be used, despite widespread evidence that life-event reporting is fundamentally flawed. By contrast, response definitions have often focused on physiological changes, and this has proved to be a particularly useful avenue of research for linking stress and illness. Recent advances in psychoneuroimmunology have shown that sustained activation of the HPA axis following exposure to stress may have deleterious consequences; the elevation in adrenal catecholamines

has a significant impact on cardiovascular function, and there is good evidence for the role of adrenal corticosteroids in compromising immune competence.

However, wide individual differences in stress reactions suggest that physiological response mechanisms alone provide only a partial answer. A number of researchers proposed that aspects of personality might play a significant role, and early work on the concept of hardiness suggested that having an internal locus of control may serve as a buffer against stress. Hardiness is a composite of three dimensions: control, commitment, and challenge. Early research had suggested that hardiness, and particularly the control component, might be important in buffering individuals against stress. Unfortunately, subsequent findings using locus of control to investigate the role of personality in stress and illness have been equivocal, as have findings with other variables such as neuroticism and extraversion.

Locus of control is a dimension of personality, but the stimulus properties of control can also be investigated by manipulating the degree of control available over environmental factors. Pioneering experiments in the context of stress used pairs of monkeys, one having control over shock avoidance (the executive monkey) and a yoked companion monkey that did not. The findings suggested that the executive monkey was more likely than the yoked monkey to develop gastrointestinal lesions through anticipatory stress arousal. However, other researchers obtained the opposite result, and it has been pointed out that the relationship between control and stress is more complex; control tends to be harmful primarily when it involves engaging in active and effortful coping responses. The picture becomes even more complicated when individual differences in perceived control are included; having an internal locus of control has been shown to be associated with a greater susceptibility to stress among subjects who experienced high levels of uncontrollable life events.

In view of the inconclusive findings from research on the role of personality in stress and illness, it has been argued that the personality constructs used in earlier studies were inappropriate because they had not been developed specifically in the context of stress research. A number of scales for assessing aspects of emotional response styles have emerged from stress research programs, and subsequent validation studies of these scales have shown that behavior such as rumination and emotional inhibition is strongly related to physiological indices of stress. Rumination in particular has been shown to prolong heart-rate recovery and corticosteroid secretion, and recent findings have shown a significant link between rumination and suppressed immune function. Other work

has implicated the tendency to inhibit the expression of emotion in a range of physiological response mechanisms, including delayed muscle tension recovery following exposure to stress. Research on emotion control is thus set within the context of adrenomedullary and adrenocortical response mechanisms mediated by personality, and the findings suggest a new definition of stress as a preoccupation with emotional upset.

Apart from emotional inhibition and rumination, other important factors with implications for stress include coping styles and the type A behavior pattern (TABP). It has been pointed out that there is confusion in the literature over the precise definition of coping, but coping inventories have typically isolated three coping domains involving rational (or problem-focused), emotional, and avoidance strategies. More recent research has uncovered a fourth domain entitled detachment within a four-factor questionnaire, and factor analyses of the four-factor structure has subsequently combined the detached and emotional components into a single bipolar scale that significantly moderates the impact of adaptational stress.

In the same way, a more detailed analysis of TABP has resulted in a reformulation of the construct as a normally distributed dimension of achievement orientation. This work has followed from earlier inconclusive attempts to inadequately distinguish between toxic and nontoxic components of TABP. The more recent research has yielded questionnaires comprising discrete, inversely correlated toxic and nontoxic subscales, in which the former is characterized by hostile competitiveness and the latter by more cooperative achievement styles.

Research with the new scales has shown a consistent pattern of intercorrelations among them and suggested a broader category of toxic behavior that might draw the different strands together. This composite risk index of toxic behavioral tendencies (rumination and emotional inhibition, a tendency towards toxic achievement styles, and a reliance on emotional rather than detached coping) is a potentially powerful tool for predicting deterioration in health status during adaptation, but it does not offer a clear understanding of why some individuals behave in a toxic fashion. Placing the findings on personality and stress in the context of self-esteem offers a way of explaining the findings.

Stress and Self-Esteem

Researchers have speculated that rumination may be the result of factors such as gender or a lack of social support. More fundamentally, however, ruminating about emotional upset may represent an attempt to

preserve self-esteem. TABP has, in fact, been interpreted in this way, by describing type As as having rigid and unrealistic ways of evaluating their own self-worth based on unrealistically high expectations of themselves. A failure to meet these expectations leads to low self-esteem, which these individuals then attempt to compensate for by engaging in the hard-driving, competitive, and aggressive behavior that characterizes TABP. If compensatory behavior is associated with toxic achievement motivation, a positive association might be expected between self-esteem and nontoxic achieving, and this has indeed been found to be the case.

These findings have recently been extended in research examining the costs of attempting to salvage self-esteem when unrealistic expectations are not met. These studies have focused on the goals of efforts to maintain self-esteem and have shown that individuals with purely self-validating goals respond to threats to self-worth in ways that affect both mental and physical health. Clearly, high self-esteem has positive effects on health only when it is pursued in ways that are not contingent on self-validation, a finding that echoes the distinction made earlier between trait and state self-esteem. Noncontingent self-esteem corresponds to a relatively enduring sense of self-worth that is not dependent on others' opinions. Similarly, problem-focused coping strategies are ineffective for highly self-critical perfectionists, but are effective for those with low self-criticism; in the former group, high achievement motivation is undermined by self-criticism, which in turn compromises coping efforts.

Self-esteem has also been shown to act as a buffer against both anxiety and depression. High self-esteem plays a pivotal role in providing protection against anxiety provoked by perceived threats, and in a study that used measures of both optimism and self-esteem to predict postpartum depression, optimism was associated with less depression up to 2 weeks postpartum, but the association with self-esteem was still evident 4 weeks later. Even when the effects of optimism and initial levels of depression were parcelled out, the association between self-esteem and postpartum depression remained.

The stress-buffering effects of self-esteem have been directly tested in experimental studies that used both psychological and physiological measures of stress. These experiments compared groups of subjects who differed in their self-esteem, and the results showed that subjects with high self-esteem suffered less distress, performed better on a cognitive task, and showed less physiological arousal in response to laboratory stressors. The physiological measure in

these studies was heart-rate elevation and recovery, but other research has used direct measures of immune competence to investigate the effects of self-esteem. For example, assays of natural killer (NK) cell activity showed that NK cytotoxicity was significantly suppressed as a consequence of negative self-evaluation.

Conclusion

In conclusion, the evidence implicating self-esteem in human stress responses is unequivocal. Low self-esteem is significantly associated with deleterious psychological, physiological, and immunological consequences, whereas high self-esteem serves a protective or buffering function, and these findings hold true even when the effects of other personality factors are parcelled out. Recent research suggests that self-esteem may act as a superordinate construct that determines the patterning of a wide range of lower-order factors. This model suggests that low self-esteem may govern a range of stress-related personality factors, including emotional inhibition and rumination, the tendency to engage in toxic patterns of achievement behavior such as hostile competitiveness, and the attribution of control externally rather than internally. It is argued that these patterns of behavior are used to compensate for the negative feelings resulting from low self-esteem and that engaging in them serves to perpetuate a maladaptive cycle that becomes self-reinforcing.

Further Reading

- Bracken, B. A. (ed.) (1996). *Handbook of self-concept: developmental, social, and clinical considerations*. New York: John Wiley.
- Crocker, J. and Park, L. E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin* 130, 392–414.
- Hughes, R. B., Robinson-Whelan, S., Taylor, H. B., et al. (2004). Enhancing self-esteem in women with physical disabilities. *Rehabilitation Psychology* 49, 295–302.
- Rector, N. and Roger, D. (1997). The stress buffering effects of self-esteem. *Personality and Individual Differences* 23, 799–808.
- Roger, D. (1999). The role of cognitive rumination, coping styles and self-esteem in moderating adaptational responses to stress. In: Mervielde, I., Deary, I., De Fruyt, F. & Ostendorf, F. (eds.) *Personality psychology in Europe* (vol. 7), pp. 351–364. Tilburg: Tilburg University Press.
- Roger, D. and Najarian, B. (1998). The relationship between emotional rumination and cortisol secretion under stress. *Personality and Individual Differences* 24, 531–538.

- Schroeder, D. H. and Costa, P. T. (1984). Influence of life event stress on physical illness. *Journal of Personality & Social Behaviour* 46, 853–863.
- Step toe, A. (1983). Stress, helplessness and control: the implications of laboratory studies. *Journal of Psychosomatic Research* 27, 361–367.

- Thomsen, D. K., Mehlsen, M. Y., Hokland, M., et al. (2004). Negative thoughts and health: associations among rumination, immunity, and health care utilization in a young and elderly sample. *Psychosomatic Medicine* 66, 363–371.

Self-Harm See: Suicide, Biology of; Eating Disorders and Stress.

Selye, Hans

B Tuchweber and P Bois

University of Montreal, Montreal, Quebec, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G Jasmin and P Bois, volume 3, pp 417–418, © 2000 Elsevier Inc.

Overview

Education and Awards

Selye's Message

Overview

Hans Selye (Figure 1), born in Vienna in 1907, is internationally acknowledged as the father of stress and, arguably, as the greatest physician of all time. Selye's work on stress has had an enormous impact on biomedical sciences and the practice of medicine. His extraordinary energy resulted in the publication during his lifetime of more than 1450 scientific papers and 22 books, of which 15 are monographs that have been translated into many languages. It is noteworthy that in the past 3 years, more than 20 years after his death in 1982, his publications have received more than 695 citations. His concern about gaining public understanding of this work led him to write several books for the layman, some of which gained great popularity, such as *The stress of life* and *Stress without distress*. Selye was an accomplished linguist able to present his work in many parts of the world, where he was eagerly sought as a speaker. As a lecturer and teacher he inspired countless listeners and students worldwide.

Selye was driven by an insatiable inquisitiveness for the pursuit of truth. Well served by a legendary memory, a great mind, and intuitive skills, he understood the importance of reflection following experimentation and gathered in his laboratories the brain power that was necessary to challenge established truths and generate new paradigms.

Education and Awards

Selye's scientific education began in 1924 at the German University of Prague, where he obtained a



Figure 1 Hans Selye.

PhD in chemistry and a doctorate in medicine. As was customary for medical students at that university, he did part of his training abroad, at the University of Paris and the University of Rome. Selye undertook postdoctoral studies in Prague and then was nominated assistant in experimental pathology in charge of the histology laboratory. In 1931, he was awarded a Rockefeller Research Fellowship to work at Johns Hopkins University in Baltimore. A year later he moved to McGill University in Canada, where he was nominated lecturer and assistant professor in the Department of Biochemistry. In 1937 he joined the Department of Histology, where he was nominated associate professor in 1941. It was in 1945 that Selye moved to the Université de Montreal, where he founded and directed the Institut de Médecine et de Chirurgie expérimentales. He held the position of professor and director until his retirement in 1976. Afterwards, Selye founded and became president of the International Institute of Stress in Montreal as well as professor emeritus until his death in 1982.

Selye was the recipient of many honorary degrees and awards from many countries, including 17 honorary degrees from universities in the United States and abroad, 72 medals and awards (e.g., the American Academy Achievement Golden Award, International Kittay, Killam Award), and 12 diplomas. He was elected Fellow of the Royal Society of Canada in 1941 and in 1968 was appointed Companion of the Order of Canada.

Selye's Message

What was Selye's message? Although the starting points were different, the central theme of all his research from 1936 onward was stress and adaptation to stress. His landmark article published in 1936 in *Nature* titled 'A syndrome produced by diverse noxious agents' described the general adaptation syndrome in response to stress with its three stages (alarm, resistance, exhaustion). Very early during his investigations, Selye recognized the significance of the adrenal gland in adaptation and also demonstrated the importance of the pituitary-adrenocortical axis in the stress reaction. He classified the steroids and named the products of the adrenal cortex corticoids, which he subdivided into glucocorticoids and mineralocorticoids. The key role of glucocorticoids in the treatment of inflammatory and immune disorders was also demonstrated.

It is of interest that in Selye's institute, two postdoctoral fellows, Claude Fortier and Roger Guillemin, initiated work on the mechanisms of hormonal regulation of the pituitary-adrenocortical axis. Guillemin

later identified the presence of the corticotropin-releasing factor (CRF) in the hypothalamus that stimulated the secretion of adrenocorticotropin (ACTH) by the pituitary gland. He was awarded the Nobel Prize in 1977 for this significant contribution to the understanding of the mechanisms of the stress response.

Further studies by Selye lent support to the view that diseases of adaptation were maladies in which derangements of the adaptation syndrome play a key role. This included the demonstration of hormonal conditioning, the influence of hormones upon reactivity, factors influencing the development of cardiac necroses and related lesions, hormonal production of hypertension, nephrosclerosis, and generalized collagen diseases.

Selye's basic studies were conducted using traditional physiological techniques in animals but were followed by many investigations by some of his graduate students and others with sophisticated techniques including neurochemistry and molecular biology. Selye did not pay much attention to these methods, which he called *in vitro* approaches. Instead, he put his trust in nature, "alone capable," as he used to say, of replying to his questions. He wrote a book on this subject, entitled *In vivo: The case for supramolecular biology*, in which he made a point of his generalist approach.

Selye was one of the great men who were not limited by the bounds of a single discipline, in his case physiology, but who covered several fields (endocrinology, immunology, metabolism) and created new disciplines (neuropsychimmunology, neuroendocrinology). His work had an impact on human development, cardiovascular disease, central nervous system depression, and the physiology of mind-body interactions. His article on the stress response published in *Nature* in 1936 has been recognized as a psychiatry classic because it led to the study of the effects of stress and hormones, mainly corticosteroids on brain function. Selye acquired worldwide fame through his description of the physiological and biological effects of stress. Although he worked only with animals and never conducted clinical research on or treated humans, the body of knowledge he generated has had a profound impact on our understanding of the consequences of stress on human development, modulation of immune and endocrine response, and psychobiology. There is no doubt that millions of people have benefited and are benefiting today from the concepts he developed.

See Also the Following Article

Alarm Phase and General Adaptation Syndrome.

Further Reading

- Selye, H. (1936). A syndrome produced by diverse nocuous agents. *Nature* 138, 32.
- Selye, H. (1947). *Encyclopedia of endocrinology*. Montreal: Acta Inc.
- Selye, H. (1956). *The stress of life*. New York: McGraw-Hill.
- Selye, H. (1964). *From dream to discovery*. New York: McGraw-Hill.

- Selye, H. (1967). *In vivo: The case of supramolecular biology*. New York: Liveright.
- Selye, H. (1971). *Hormones and resistance*. New York: Springer-Verlag.
- Selye, H. (1976). *Stress in health and diseases*. Boston, MA: Butterworth.
- Selye, H. (1979). *The stress of my life, a scientist's memoirs*. New York: Van Nostrand Reinhold.

Separation, Marital See: Divorce; Children of; Domestic Violence; Life Events Scale; Life Events and Health; Male Partner Violence; Marital Conflict; Marital Status and Health Problems; Marriage.

Sepsis, Acute Respiratory Distress Syndrome, and Glucocorticoid Resistance

G U Meduri

University of Tennessee Health Science Center
Department of Medicine, Memphis, TN USA

© 2007 Elsevier Inc. All rights reserved.

*Systemic
inflammation*

The release of inflammatory cytokines in the systemic circulation manifesting clinically with fever and increased heart rate, respiratory rate, and white cell blood count.

Introduction

Acute Respiratory Distress Syndrome

Systemic Inflammation-Associated Glucocorticoid Resistance

Nuclear Factor κ B and Glucocorticoid Receptor α Interaction

Systemic Inflammation-Induced Glucocorticoid Resistance Can Be Improved with Prolonged Glucocorticoid Supplementation

Findings of Randomized Trials

Conclusion

Glossary

*Hypoxemic
respiratory
failure*

*Multiple organ
dysfunction*

Sepsis

A life-threatening reduction in blood oxygenation requiring assisted mechanical ventilation.

Systemic inflammation-associated impairments of vital organs: brain, lung, cardiovascular system, renal, liver, and hematopoietic system.

Systemic inflammation induced by infection.

Introduction

Hundred of thousands of American are affected every year by severe sepsis (700,000) and acute respiratory distress syndrome (ARDS; 190,000) creating a significant short-term and long-term burden on public health and health-care economics. Dysregulated systemic inflammation with persistent elevation of circulating inflammatory cytokines over time is the pathogenetic mechanism for the dysfunction and failure of the vital organs, the leading cause of death in patients with sepsis and ARDS. In sepsis and ARDS, the downregulation of inflammation is *conditio sine qua non* to their resolution. An improved understanding of the cellular mechanisms that initiate, propagate, and restrict inflammation in sepsis and ARDS is essential to make advances on present treatment modalities and to develop more reliable methods of monitoring disease progression and responses to therapy.

It is now appreciated that at cellular level transcription factors nuclear factor (NF)- κ B, which is

activated by inflammatory signals, and the glucocorticoid receptor (GR) α , which is activated by endogenous or exogenous glucocorticoids (GCs), have diametrically opposed functions (stimulatory vs. inhibitory) in regulating inflammation. NF- κ B is recognized as the principal driver of the inflammatory response, being responsible for the transcription of more than 100 genes, including tumor necrosis factor (TNF)- α and interleukin(IL)-1 β (Figure 1). Once activated, NF- κ B and GR α can mutually repress one

another through a protein–protein interaction that prevents their DNA binding and subsequent transcriptional activity. Known interactions between NF- κ B and the activated glucocorticoid receptor (GC-GR α) are depicted in Figure 1. The activation of one transcription factor in excess of the binding (inhibitory) capacity of the other shifts cellular responses toward the increased (dysregulated) or decreased (regulated) transcription of inflammatory mediators over time (Figure 2). The role of GC

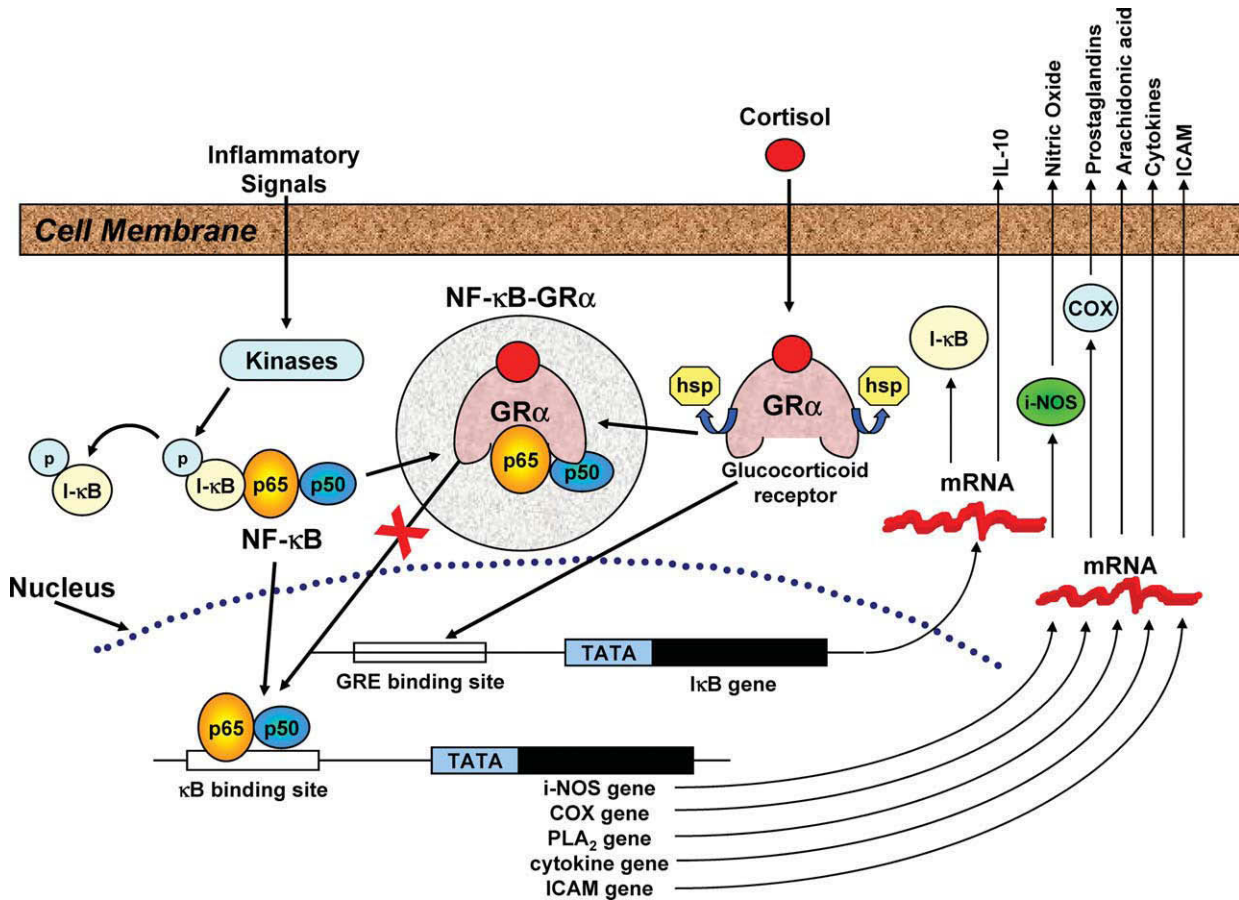


Figure 1 Positive regulatory loop interaction between NF- κ B and the activated glucocorticoid receptor (GC-GR α). When cells are stimulated by inflammatory signals, specific kinases phosphorylate the inhibitory protein I- κ B and cause its rapid degradation. The activated form of NF- κ B then moves to the nucleus, initiating the transcription of mRNA of inflammatory cytokines, chemokines, cell adhesion molecules, and inflammation-associated enzymes (cyclooxygenase, phospholipase A2, and inducible nitric oxide). Cortisol or exogenous glucocorticoids freely cross into the cytoplasm and bind to their specific glucocorticoid receptors (GR α) to form GC-GR α . GC-GR α complexes may influence NF- κ B activity in five major ways: (1) physically interacting with the p65 subunit to form an inactive (GC-GR α -NF- κ B) complex, (2) inducing the synthesis of the inhibitory protein I- κ B α via interaction with glucocorticoid-responsive element DNA in the promoter of the I- κ B gene, (3) blocking the degradation of I- κ B α via enhanced synthesis of IL-10, (4) impairing the TNF- α -induced degradation of I- κ B α , and (5) competing for limited amounts of GR co-activators such as cAMP response element binding protein (CREB)-binding protein (CBP) and steroid receptor coactivator (SRC-1). I- κ B α , in addition to holding NF- κ B in an inactive cytoplasmic state, also translocates into the nucleus, where it binds activated NF- κ B complexes to induce their export to the cytoplasm. GC-GR α may also decrease the stability of mRNA of several inflammatory cytokines and other molecules. The products of the genes that are stimulated by NF- κ B activate this transcription factor. Thus, TNF- α and IL-1 β both activate and are activated by NF- κ B by forming a positive regulatory loop that amplifies and perpetuates inflammation. Reproduced with permission from Meduri, G. U., Muthiah, M. P., Carratu, P., et al. (2005), Nuclear factor-kappaB- and glucocorticoid receptor alpha-mediated mechanisms in the regulation of systemic and pulmonary inflammation during sepsis and acute respiratory distress syndrome: evidence for inflammation-induced target tissue resistance to glucocorticoids, *Neuroimmunomodulation* 12(6), 321–338.

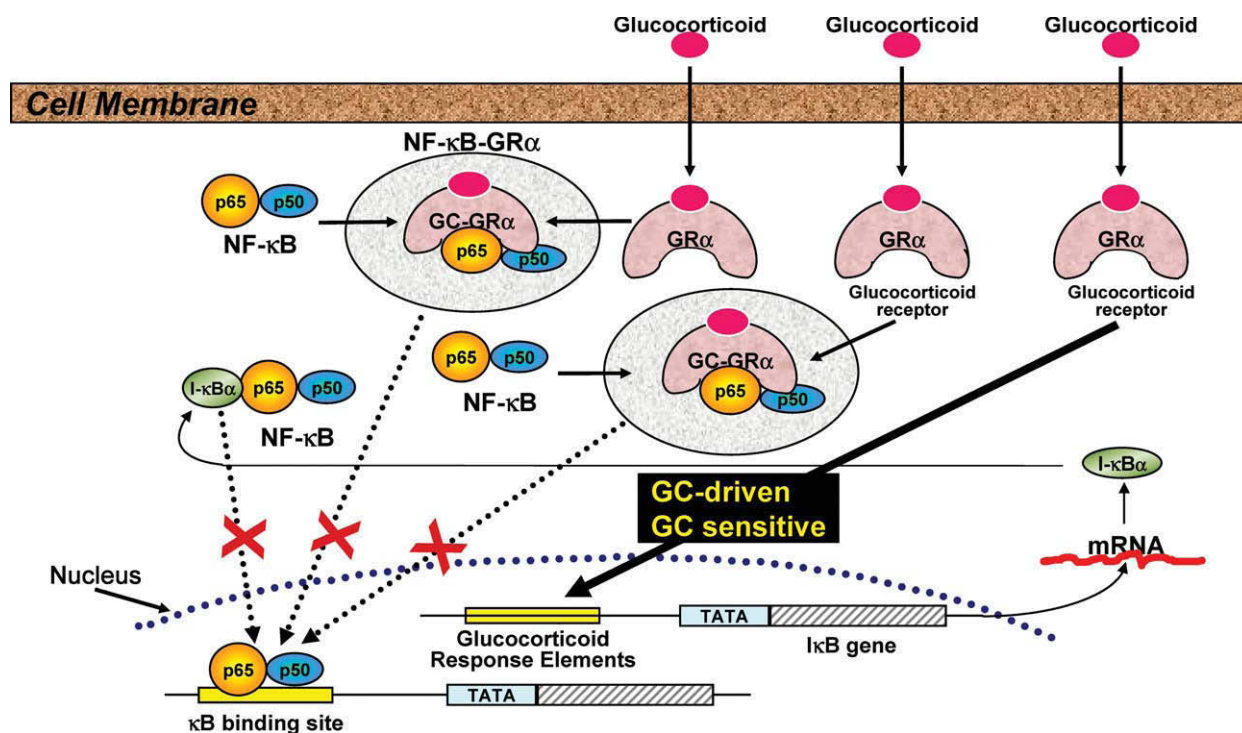


Figure 2 Regulated response interaction between NF-κB and the activated glucocorticoid receptor (GC-GRα). GC-GRα activation sufficient to maintain NF-κB levels in homeostasis and achieve a reduction in the transcription of inflammatory mediators over time is shown. In improvers, both GC-GRα binding to NF-κB and nuclear GC-GRα binding increased significantly over time, indicating an excess activation of GC-GRα compared to NF-κB (GC-GRα-driven response). Reproduced with permission from Meduri, G. U., Muthiah, M. P., Carratu, P., et al. (2005), Nuclear factor-kappaB- and glucocorticoid receptor alpha-mediated mechanisms in the regulation of systemic and pulmonary inflammation during sepsis and acute respiratory distress syndrome: evidence for inflammation-induced target tissue resistance to glucocorticoids, *Neuroimmunomodulation* 12(6), 321–338.

supplementation in patients with life-threatening systemic inflammation is undergoing a cyclical reassessment, stirred by a new understanding of the role of GCs in regulating inflammation and the positive results of recent randomized controlled studies.

Acute Respiratory Distress Syndrome

ARDS is a life-threatening form of hypoxemic respiratory failure (mortality of 30–60%) that develops within 12–48 h of a patient's acquiring an acute illness complicated by severe systemic inflammation. The most common condition precipitating ARDS is severe sepsis, and the mortality rate for sepsis-induced ARDS has not significantly decreased in the last 20 years. In ARDS, the evolution (regulated vs. dysregulated) of systemic inflammation in the first week of mechanical ventilation determines the physiological progression of the disease and the final outcome. Patients failing to improve in the first week of mechanical ventilation (nonimprovers) have, compared to improvers, a persistent elevation in circulating and bronchoalveolar lavage (BAL) levels of inflammatory cytokines and chemokines, markers of

alveolocapillary membrane (ACM) permeability and fibrogenesis. In sepsis and ARDS, the downregulation of inflammation is essential for the restoration of homeostasis and survival.

Systemic Inflammation-Associated Glucocorticoid Resistance

GCs as end effectors of the hypothalamic-pituitary-adrenal (HPA) axis are the most important physiological inhibitors of inflammation, affecting hundreds of genes involved in stress-related homeostasis. However, endogenous GCs are not always effective in suppressing life-threatening systemic inflammation, and the degree of cortisolemia frequently correlates with the severity of illness and the mortality rate. Unquestionably, the elevation of GC secretion in non-survivors of sepsis and ARDS is inadequate to meet the needs of the concurrent inflammatory response and its adverse systemic effects. Its failure to suppress inflammation could be due to tissue resistance to GCs, the inadequacy of the level and duration of endogenous GC elevation to suppress a systemic inflammatory response gone awry, or both.

The concept of acquired GC resistance was first introduced by Kass and Finland in 1957. These investigators suggested that increased blood cortisol levels in nonsurvivors of sepsis might reflect a blocking of steroidal activity or transport as a consequence of the infection. In this situation, a small increase in blood levels with a low (equal to or less than physiological) dose of exogenous GCs was believed to be sufficient to facilitate the passage of steroids into host cells. GC resistance was originally described as a primary inherited familial syndrome and was recently recognized as an acquired condition. Among others, acquired immune tissue-specific GR resistance has been described in patients with asthma, acquired immunodeficiency syndrome (AIDS), severe sepsis, and sepsis-induced ARDS.

In vitro studies showed that cytokines may induce resistance to GCs by reducing GR binding affinity to cortisol and/or GC response elements (GREs). Such abnormalities of GR function were demonstrated in T cells incubated with a combination of interleukin (IL)-2 and IL-4, IL-1, IL-6, and interferon (IFN)- γ , or IL-13. GC resistance was induced in a cytokine concentration-dependent fashion and was reversed by the removal of cytokines. GR-mediated resistance in the presence of systemic inflammation was also studied in experimental models of sepsis and sepsis-induced ARDS. In a sheep model of sepsis-induced ARDS, the maximal binding capacity of GR decreased continuously after endotoxin infusion, whereas there was a marked elevation of cortisol levels. The reduced GR binding correlated negatively ($r = -0.87$, $p < 0.01$) with phospholipase A₂ (PLA₂) activity, a gene that is stimulated by NF- κ B. In a rat model of septic shock, GR blockade by mifepristone (RU 486) exacerbated the physiological and pathological changes induced by endotoxemia. PLA₂ activity in rats with 80% GR blockade was more marked than in those with 50% GR blockade. The monocytes of patients with sepsis developed near-total GC resistance *in vitro* when cytokines, especially IL-2, were added.

Nuclear Factor κ B and Glucocorticoid Receptor α Interaction

New evidence indicates that a primary pathogenetic mechanism explaining GC resistance or insensitivity in patients with life-threatening systemic inflammation is the imbalance between excessive NF- κ B activation by inflammatory signals and deficient GR α activation by endogenous or exogenous glucocorticoids. Using an *ex vivo* model of systemic inflammation, our group investigated intracellular upstream and downstream events associated with DNA binding

of NF- κ B and GR α in naïve peripheral blood leukocytes (PBLs) stimulated with longitudinal plasma specimens obtained from 28 patients with sepsis-induced ARDS and correlated the physiological progression with laboratory findings. Exposure of naïve cells to longitudinal plasma samples led to divergent directions in NF- κ B and GR α activation that reflected the severity of systemic inflammation (defined by plasma TNF- α and IL-1 β levels). The activation of one transcription factor in excess of the other shifted cellular responses toward decreased (GR α -driven) or increased (NF- κ B-driven) transcription of inflammatory mediators over time (Figures 2–3).

Plasma samples from patients with declining inflammatory cytokine levels over time elicited a progressive increase in all measured aspects of GC-GR α -mediated activity ($p = 0.0001$) and a corresponding reduction in NF- κ B nuclear binding ($p = 0.0001$) and transcription of TNF- α and IL-1 β (regulated, GR α -driven response; Figure 2). In contrast, plasma samples from patients with sustained elevations in inflammatory cytokine levels over time elicited only modest longitudinal increases in GC-GR α -mediated activity ($p = 0.04$) and a progressive increase in NF- κ B nuclear binding over time ($p = 0.0001$) that was most striking in nonsurvivors (dysregulated, NF- κ B-driven response; Figure 3). These findings demonstrate that insufficient GC-GR α -mediated activity is an important mechanism for early loss of autoregulation (i.e., downregulation of NF- κ B activation) in nonimprovers. Deficient GR α activity in naïve cells exposed to plasma from patients with dysregulated inflammation was observed despite elevated circulating cortisol and adrenocorticotrophic hormone (ACTH) levels, implicating inflammatory cytokine-driven excess NF- κ B activation as an important mechanism for target-organ insensitivity and/or resistance to cortisol.

Our findings are in agreement with two longitudinal studies that investigated NF- κ B binding activity directly in the peripheral blood mononuclear cells of patients with sepsis or trauma. In both studies, nonsurvivors (compared to survivors) had a progressive increase in NF- κ B activity over time. In one longitudinal study, nonsurvivors of septic shock had by days 2–6 a 200% increase in NF- κ B activity compared to day 1. Similarly, using our *ex vivo* model we found that NF- κ B binding activity on day 3 of ARDS clearly separated patients by outcome. In an immunohistochemistry analysis of lung tissue, areas with severe (vs. mild) ARDS had a higher nuclear uptake of NF- κ B (13 ± 1.3 vs. 7 ± 2.9 ; $p = 0.01$) and a lower ratio of GR α : NF- κ B nuclear uptake (0.5 ± 0.2 vs. 1.5 ± 0.2 ; $p = 0.007$). Thus, an uninhibited increase in NF- κ B activation in circulating cells over time

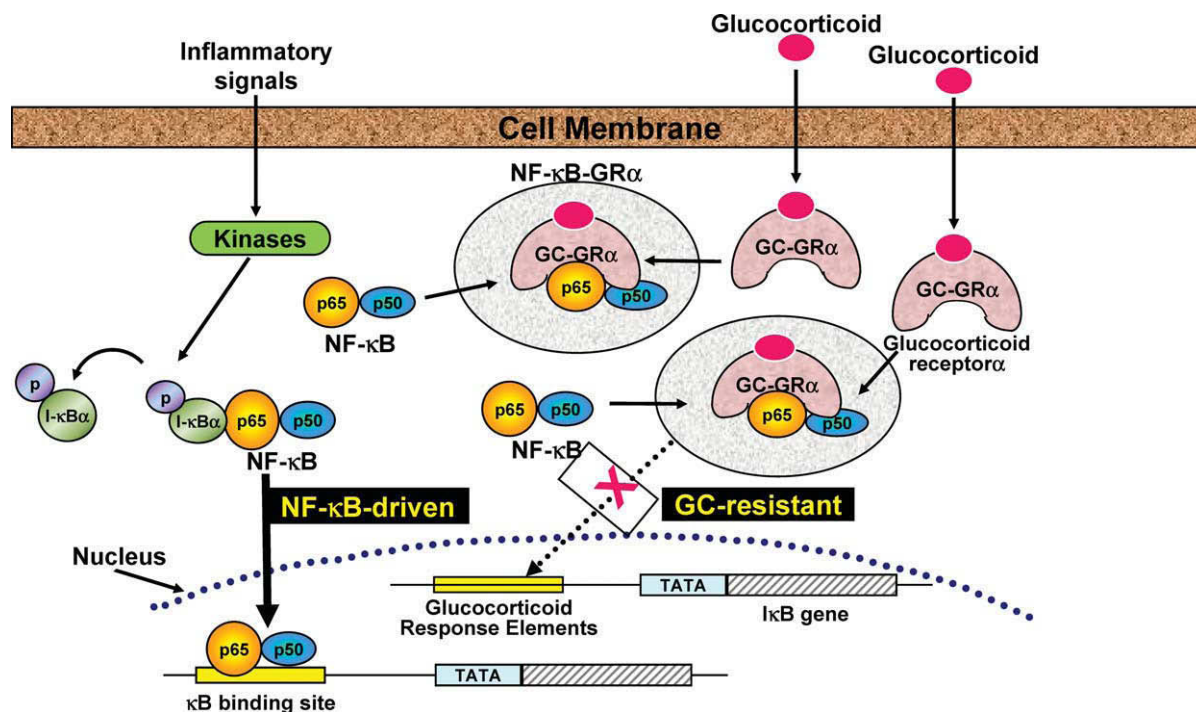


Figure 3 Dysregulated response interaction between NF-κB and the activated glucocorticoid receptor (GC-GRα). An excess of NF-κB activation is shown, leading to the protracted transcription of inflammatory mediators over time. In nonimprovers, GC-GRα binding to NF-κB was modestly increased, whereas nuclear NF-κB binding increased substantially over time and nuclear GC-GRα and cytoplasmic IκBα levels declined (NF-κB-driven response). Reproduced with permission from Meduri, G. U., Muthiah, M. P., Carratu, P., et al. (2005), Nuclear factor-kappaB- and glucocorticoid receptor alpha-mediated mechanisms in the regulation of systemic and pulmonary inflammation during sepsis and acute respiratory distress syndrome: evidence for inflammation-induced target tissue resistance to glucocorticoids, *Neuroimmunomodulation* 12(6), 321–338.

appears to be a significant premortem pathogenetic component of lethal sepsis and ARDS.

Systemic Inflammation-Induced Glucocorticoid Resistance Can Be Improved with Prolonged Glucocorticoid Supplementation

Our findings place GC-GRα-mediated downregulation of NF-κB activity as a critical factor for the reestablishment of homeostasis during the acute, life-threatening systemic inflammation that accompanies sepsis and ARDS. A subgroup of 11 patients with dysregulated inflammation and deficient GC-GRα activity were randomized to prolonged GC supplementation initiated on ARDS day 9 ± 3 . Quantitatively adequate and prolonged hormonal supplementation restored GR anti-inflammatory function, decreasing the production of inflammatory cytokines, inflammatory cytokine-driven HPA axis activity, and inflammatory cytokine-driven organ dysfunction. Normal PBL exposed to plasma samples collected during GC versus placebo treatment exhibited rapid, progressive, and significant increases in

GRα-mediated activities (GRα binding to NF-κB, GRα binding to GRE DNA, stimulation of inhibitory protein IκBα, and stimulation of IL-10 transcription) and significant reductions in NF-κB κB binding and transcription of TNF-α and IL-1β. During GC treatment, the relationship between NF-κB and GC-GRα signaling pathways changed from an initial NF-κB-driven and GC-GRα-resistant state to a GC-GRα-driven and GC-GRα-sensitive one. In aggregate, the findings indicate that systemic inflammation-induced GC resistance is an acquired generalized process mediated by excess NF-κB activation and potentially reversible by increasing GC-GRα activation with quantitatively adequate and prolonged GC supplementation.

Findings of Randomized Trials

At present, the administration of low-dose GCs is the only treatment shown in randomized trials to have a significant positive biological and physiological effect across the longitudinal spectrum of the severity of sepsis-associated systemic inflammation: severe sepsis, septic shock, early ARDS, and unresolving ARDS.

Table 1 Prolonged glucocorticoid treatment in sepsis and ARDS: effect on mortality^a

Severity of sepsis	Number of RCTs	Number of patients (death of total)			Source ^b
		Treated	Control	p value	
Severe sepsis	1	0 of 23 (0%)	7 of 23 (30%)	0.009	5
Septic shock	5	7 of 22	12 of 19		2
		3 of 20	4 of 20		3
		6 of 23	10 of 21		4
		82 of 151	91 of 149		1
		8 of 20	12 of 20		10
Total from meta-analysis		106 of 236 (45%)	129 of 229 (56%)	0.01	8
Early ARDS	1	50 of 63 (21%)	16 of 28 (43%)	0.03	6
Unresolving ARDS	2	2 of 16 (12%)	5 of 8 (62%)	0.03	7
		32 of 90 (36%)	31 of 90 (34%)	NS	9

^aARDS, acute respiratory distress; RCTs, randomized controlled trials.^bData from:

1. Annane, D., Sebille, V., Charpentier, C., et al. (2002), Effect of treatment with low doses of hydrocortisone and fludrocortisone on mortality in patients with septic shock, *Journal of the American Medical Association* **288**(7), 62–871.
2. Bollaert, P. E., Charpentier, C., Levy, B., et al. (1998), Reversal of late septic shock with supraphysiologic doses of hydrocortisone, *Critical Care Medicine* **26**(4), 645–650.
3. Briegel, J., Forst, H., Haller, M., et al. (1999), Stress doses of hydrocortisone reverse hyperdynamic septic shock: a prospective, randomized, double-blind, single-center study, *Critical Care Medicine* **27**(4), 723–732.
4. Chawla, K., Kupfer, Y., Goldman, I., et al. (1999), Hydrocortisone reverses refractory septic shock, *Critical Care Medicine* **27**, A33.
5. Confalonieri, M., Urbino, R., Potena, A., et al. (2005), Hydrocortisone infusion for severe community-acquired pneumonia: a preliminary randomized study, *American Journal of Respiratory Critical Care Medicine* **171**(3), 242–248.
6. Meduri, G. U., Golden, E., Freire, A. X., et al. (2005), Methylprednisolone infusion significantly improves lung function in patients with early acute respiratory distress syndrome: results of a randomized controlled trial, *Chest* **128**, S129.
7. Meduri, G. U., Headley, S., Golden, E., et al. (1998), Effect of prolonged methylprednisolone therapy in unresolving acute respiratory distress syndrome: a randomized controlled trial, *Journal of the American Medical Association* **280**, 159–165.
8. Minneci, P. C., Deans, K. J., Banks, S. M., et al. (2004), Meta-analysis: the effect of steroids on survival and shock during sepsis depends on the dose, *Annals of Internal Medicine* **141**(1), 47–56.
9. Steinberg, K. P., Hudson, L. D., Goodman, R. B., et al. (2006), Efficacy and safety of corticosteroids for persistent acute respiratory distress syndrome, *New England Journal of Medicine* **354**(16), 1671–1684.
10. Yildiz, O., Doganay, M., Aygen, B., et al. (2002), Physiological-dose steroid therapy in sepsis, *Critical Care* **6**(3), 251–259.

All randomized studies showed significant physiological improvement and a reduction in the duration of stays in the intensive care unit. The overall effects on mortality are shown in [Table 1](#).

In prolonged low-dose GC treatment, maximal benefits are obtained by (1) avoiding the premature removal of an effective drug, leading to rebound inflammation and physiological deterioration, and (2) preventing potential complications associated with GC treatment (infection surveillance and avoidance of paralysis). Uncontrolled studies have also shown a likely benefit in patients with viral pneumonia-induced ARDS, including SARS. The mortality rate in 849 cases of SARS treated with antiviral therapy and GC supplementation was only 7.7%. In hantavirus pulmonary syndrome, the mortality rate decreased from 47% in patients treated with antiviral therapy to 13% when GC supplementation was also included. The findings for severe viral pneumonia and the results of recent randomized controlled trials (RCTs) for ARDS indicate that prolonged methylprednisolone use might be useful in the prevention of morbidity and mortality associated with avian influenza.

Conclusion

Dysregulated systemic inflammation has enormous health-care consequences. A new understanding of the cellular mechanisms responsible for inflammation-associated GC resistance provides a new rationale for investigating prolonged low-dose GC supplementation. The consistent positive results of small- to moderate-size randomized studies are very encouraging. Federal financial support is urgently needed to conduct additional RCTs investigating this inexpensive off-patent treatment in patients with life-threatening systemic inflammation.

Acknowledgments

This work was supported by Grants from the Assisi Foundation of Memphis.

Further Reading

Annane, D., Sebille, V., Charpentier, C., et al. (2002). Effect of treatment with low doses of hydrocortisone and fludrocortisone on mortality in patients with septic

- shock. *Journal of the American Medical Association* 288(7), 62–871.
- Annane, D., Sebille, V., Troche, G., et al. (2002). A 3-level prognostic classification in septic shock based on cortisol levels and cortisol response to corticotropin. *Journal of the American Medical Association* 283(8), 1038–1045.
- Bollaert, P. E., Charpentier, C., Levy, B., et al. (1998). Reversal of late septic shock with supraphysiologic doses of hydrocortisone. *Critical Care Medicine* 26(4), 645–650.
- Briegel, J., Forst, H., Haller, M., et al. (1999). Stress doses of hydrocortisone reverse hyperdynamic septic shock: a prospective, randomized, double-blind, single-center study. *Critical Care Medicine* 27(4), 723–732.
- Chawla, K., Kupfer, Y., Goldman, I., et al. (1999). Hydrocortisone reverses refractory septic shock. *Critical Care Medicine* 27, A33.
- Cheng, V. C., Tang, B. S., Wu, A. K., et al. (2004). Medical treatment of viral pneumonia including SARS in immunocompetent adult. *Journal of Infection* 49(4), 262–273.
- Confalonieri, M., Urbino, R., Potena, A., et al. (2005). Hydrocortisone infusion for severe community-acquired pneumonia: a preliminary randomized study. *American Journal of Respiratory Critical Care Medicine* 171(3), 242–248.
- Kass, E. H. and Finland, M. (1957). Adrenocortical hormones and the management of infection. *Annual Review of Medicine* 8, 1–18.
- Meduri, G. U., Golden, E., Freire, A. X., et al. (2005). Methylprednisolone infusion significantly improves lung function in patients with early acute respiratory distress syndrome: results of a randomized controlled trial. *Chest* 128, S129.
- Meduri, G. U., Headley, S., Golden, E., et al. (1998). Effect of prolonged methylprednisolone therapy in unresolving acute respiratory distress syndrome: a randomized controlled trial. *Journal of the American Medical Association* 280, 159–165.
- Meduri, G. U., Muthiah, M. P., Carratu, P., et al. (2005). Nuclear factor-kappaB- and glucocorticoid receptor alpha-mediated mechanisms in the regulation of systemic and pulmonary inflammation during sepsis and acute respiratory distress syndrome: evidence for inflammation-induced target tissue resistance to glucocorticoids. *Neuroimmunomodulation* 12(6), 321–338.
- Meduri, G. U., Tolley, E. A., Chrousos, G. P., et al. (2002). Prolonged methylprednisolone treatment suppresses systemic inflammation in patients with unresolving acute respiratory distress syndrome: evidence for inadequate endogenous glucocorticoid secretion and inflammation-induced immune cell resistance to glucocorticoids. *American Journal of Respiratory Critical Care Medicine* 165(7), 983–991.
- Meduri, G. U. and Yates, C. R. (2004). Systemic inflammation-associated glucocorticoid resistance and outcome of ARDS. *Annals of the New York Academy of Sciences* 1024, 24–53.
- Minneci, P. C., Deans, K. J., Banks, S. M., et al. (2004). Meta-analysis: the effect of steroids on survival and shock during sepsis depends on the dose. *Annals of Internal Medicine* 141(1), 47–56.
- Rhen, T. and Cidlowski, J. A. (2005). Antiinflammatory action of glucocorticoids – new mechanisms for old drugs. *New England Journal of Medicine* 353(16), 1711–1723.
- Steinberg, K. P., Hudson, L. D., Goodman, R. B., et al. (2006). Efficacy and safety of corticosteroids for persistent acute respiratory distress syndrome. *New England Journal of Medicine* 354(16), 1671–1684.
- Yildiz, O., Doganay, M., Aygen, B., et al. (2002). Physiological-dose steroid therapy in sepsis. *Critical Care* 6(3), 251–259.

Septic Shock See: Hypotension, Hypovolemia, and Septic Shock; Sepsis, Acute Respiratory Distress Syndrome and Glucocorticoid Resistance.

Serotonin

B C Williams and Y-C Lo

University of Edinburgh and Western General Hospital,
Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3,
pp 419–423, © 2000, Elsevier Inc.

Serotonin and Adrenocortical Function

Serotonin and Renal Function

Conclusion

Glossary

5-Hydroxytryptamine (5-HT) receptors

Plasma membrane proteins that mediate the biological effects of serotonin through the activation of an intracellular second-messenger system. There are at least seven different classes of 5-HT receptors, which are defined according to their pharmacological properties using selective agonists and antagonists and/or their molecular structure as determined by molecular cloning techniques; some of these classes of 5-HT receptors have been divided further into subclasses. Not all of the 5-HT receptors identified have yet been assigned a clear biological function.

Aromatic-L-amino acid decarboxylase (L-AAAD)

A chemical (EC 4.1.1.28) that catalyzes the decarboxylation of L-5-hydroxytryptophan and L-dopa (L-3,4-dihydroxyphenylalanine) to serotonin and dopamine (3,4-dihydroxyphenylethylamine), respectively. It requires pyridoxal phosphate as a coenzyme and occurs most abundantly in the kidney, liver, gastrointestinal tract, adrenal gland, and brain; previously known as hydroxytryptophan decarboxylase or dopa decarboxylase.

Serotonin (5-hydroxytryptamine, 5-HT)

A biogenic amine (first isolated from serum by Irvine Page in 1948) shown to affect vascular tone. Its other name (adopted by Erspamer in the 1940s) was enteramine because it was found to be associated with enterochromaffin cells of the gastric mucosa. By the 1950s, the chemical structure was established and serotonin/enteramine was renamed 5-hydroxytryptamine.

Zona fasciculata

A broad layer of cells in contact with the zona glomerulosa and forming the main mass of the adrenal cortex. These cells

secrete the steroid hormones cortisol (in humans) and corticosterone (in rodents). These steroids act on glucocorticoid receptors, which are distributed very widely in mammalian tissues, and produce many different biological actions, including the breakdown of carbohydrates, proteins, and fats.

Zona glomerulosa

A narrow layer of cells in the outer region of the adrenal gland, situated directly beneath the adrenal capsule, whose specific function is to secrete the steroid hormone aldosterone, which acts on mineralocorticoid receptors in the distal convoluted tubule of the kidney to cause sodium retention and potassium loss.

Serotonin and Adrenocortical Function

In Vitro Effects of Serotonin in the Adrenal Cortex

The stimulatory effect of 5-HT on adrenocortical steroidogenesis in several species, including the rabbit, guinea pig, rat, and human, was discovered by Rosenkrantz in 1959. Based on the sodium-retaining effects and chromatographic properties, the steroid stimulated by 5-HT was identified as aldosterone, which is synthesized in the zona glomerulosa cells of the adrenal cortex. Concurrent experiments demonstrated that rat serum (but not rat plasma) stimulated steroid production in the rat zona glomerulosa, indicating that a component of the blood-clotting system (later identified as 5-HT) was involved. These landmark experiments led to subsequent studies by numerous workers to determine the role of 5-HT in adrenocortical function. Most of these studies were carried out in rat models for the reason that the zona glomerulosa in this species can be cleanly dissected from the zona fasciculata and zona reticularis and, following enzymatic disaggregation of the tissue, the zona glomerulosa cells can be obtained with less than 5% contamination by other cell types. When rats are placed on a low sodium chloride diet, the sensitivity of the aldosterone response to 5-HT is increased in the zona glomerulosa. Conversely, when either a high sodium chloride diet or a low potassium chloride diet is given, the zona glomerulosa cells become less sensitive to 5-HT. This clearly indicates a dependency of the aldosterone response to 5-HT on the sodium/potassium status, which is also observed for other zona glomerulosa-specific stimuli, such as angiotensin II and plasma potassium. In contrast, no

direct steroidogenic response to 5-HT is found in cells of the zona fasciculata or zona reticularis. The stimulation of aldosterone by 5-HT is associated with an increase in cAMP in rat zona glomerulosa cells, and these responses are inhibited by 5-HT antagonists such as methysergide, cyproheptadine, and ketanserin, indicating the presence of 5-HT cell surface receptors in the zona glomerulosa. Although this receptor was identified tentatively as a 5-HT₂ receptor (now renamed the 5-HT_{2A} receptor) based on the effects of ketanserin, subsequent cloning of this receptor did not support this notion. No mRNA for this receptor (which is coupled to phospholipase C) has been found in zona glomerulosa cells, and 5-HT does not stimulate phospholipase C in these cells. The 5-HT receptor in rat zona glomerulosa cells still awaits clear identification. In human zona glomerulosa cells, the 5-HT receptor has been partially characterized, pharmacologically, as a 5-HT₄ receptor (which couples to adenylate cyclase), and the 5-HT₄-selective agonist zacopride is a potent stimulus to aldosterone in human zona glomerulosa tissue. In both rat and human zona glomerulosa cells, 5-HT-stimulated aldosterone secretion is also coupled to Ca²⁺ influx and is inhibited by verapamil, suggesting the involvement of voltage-dependent Ca²⁺ channels.

Ex Vivo Effects of Serotonin in the Adrenal Cortex

In the isolated perfused rat adrenal gland, 5-HT increases the perfusion flow rate, which correlates with the increased secretion of both aldosterone and corticosterone. No 5-HT receptors have been characterized in the adrenal vasculature to date, but because the blood flow rate is an important determinant for steroid secretion *in vivo*, these potential vascular effects of 5-HT in the adrenal cortex could be important physiologically.

In Vivo Effects of Serotonin on Adrenocortical Steroidogenesis

Infusions of high doses of 5-HT in humans cause an increase in aldosterone but not cortisol secretion. Because plasma levels of 5-HT are normally extremely low, due to the active 5-HT reuptake mechanism in the platelets, the potential physiological relevance of this observation is unclear. The administration of the 5-HT precursors L-tryptophan and 5-hydroxytryptophan (5-HTP) in humans produces an increase in both aldosterone and cortisol secretion, and it is thought that these effects are mediated by the activation of the hypothalamic–pituitary–adrenal (HPA) axis because it has been demonstrated clearly in rat tissues *in vitro* that 5-HT can stimulate both corticotropin releasing hormone (CRH) release and adrenocorticotrophic

hormone (ACTH) release. Intraperitoneal injections of 5-HTP in rats lead to an increase in both aldosterone and corticosterone secretion and plasma renin activity *in vivo*. The corticosterone response to 5-HT can be completely inhibited by the co-administration of dexamethasone, whereas the aldosterone response to 5-HTP is only partially inhibited. The co-administration of the angiotensin-converting enzyme inhibitor captopril partially inhibits the aldosterone response to 5-HTP treatment and does not affect the corticosterone response. These studies therefore indicate that corticosterone (a zona fasciculata steroid) can be regulated by 5-HT *in vivo* by the stimulation of the HPA axis, whereas the aldosterone response is more complex, involving the activation of the HPA axis, the renin–angiotensin system, and, in addition, direct effects in the zona glomerulosa. The effects of 5-HT on renin release are thought to be centrally mediated because 5-HT does not stimulate renin release in the kidney. The notion that the direct effects of 5-HT on zona glomerulosa cells could be an important physiological control mechanism for aldosterone secretion is supported by studies that demonstrate that infusions of the 5-HT₄-selective agonist zacopride in humans increase plasma aldosterone but not plasma cortisol.

Sources of Serotonin in the Adrenal Gland

The physiological source of 5-HT that regulates aldosterone secretion in the zona glomerulosa has not been clearly identified to date. Although 5-HT has been located in the adrenal medulla of some species, including the rat and pig by immunohistochemical techniques, the centripetal blood flow in the adrenal gland from the outer cortex to the inner medulla argues against any significant influence of medullary secretory products on adrenocortical function. In the rat, mast cells (which store 5-HT) have been found in the subcapsular portions of the adrenal artery, and drugs that cause mast-cell degranulation can stimulate aldosterone secretion in the perfused rat adrenal gland. Physiological mechanisms leading to the specific release of 5-HT from these mast cells within the adrenal cortex, however, have not been identified. Other studies have focused on platelets as a possible physiological source of adrenocortical 5-HT. The use of ¹⁴C-radiolabeled 5-HT to track the platelet deposition of 5-HT *in vivo* in rats has demonstrated that 5-HT can potentially be released in some endocrine tissues, including the thyroid gland and the adrenal gland, but the release mechanisms are unclear. Perhaps the most likely source of adrenocortical 5-HT is the enzyme L-AAAD, which has been found, by immunohistochemistry, to be specifically associated with the zona glomerulosa in the rat adrenal cortex.

This enzyme converts 5-HTP (which circulates in plasma and is not stored in platelets) to 5-HT and can be inhibited by carbidopa. The incubation of rat adrenal zona glomerulosa tissue *in vitro* with 5-HTP, at physiological concentrations, leads to an increase in aldosterone and 5-HT, and both are completely inhibited in the presence of carbidopa. Furthermore, this enzyme in the zona glomerulosa appears to be regulated by changes in the dietary sodium chloride intake in parallel with observed changes in aldosterone secretion. Thus, the enzyme activity is increased by a low sodium chloride intake and is decreased by a high sodium chloride intake.

Serotonin and Renal Function

Effects of Serotonin on Renal Vasculature

The effects of 5-HT on the renal circulation are variable, depending on the species studied. Most investigations have been carried out in rats or dogs, with few comprehensive studies in humans. The administration of 5-HT to rats *in vivo* (by infusion or intraperitoneal injection) increases renal vascular resistance and reduces the renal blood flow rate. These effects are also observed in the *ex vivo*-isolated perfused rat kidney, are independent of intact renal nerves, and are inhibited by ketanserin, a 5-HT_{2A} antagonist. In the dog, similar effects are observed, although they are clearly more dependent on renal innervation, in that the stimulation of the sympathetic nerves results in a vasodilator response to 5-HT rather than a vasoconstriction. It has been suggested that 5-HT_{1D} receptors may be more important in mediating 5-HT-induced renal vasoconstriction than 5-HT_{2A} receptors in the dog because the 5-HT agonist sumatriptan can reduce renal blood in the presence of a ketanserin blockade. The 5-HT receptor that mediates renal vasodilatation in the dog has not been characterized, although it is thought to be associated with the vascular endothelium and to be coupled to nitric oxide production. Most studies in humans have indicated either no change or only a modest reduction in renal plasma flow and glomerular filtration rate (GFR) in response to intravenous 5-HT. The physiological role of 5-HT in modulating renal hemodynamics in humans is therefore still unclear.

Effects of Serotonin on Renal Tubule Function

In rats, intravenous or intraperitoneal administration of 5-HT causes a decrease in the urine flow rate and a reduction in sodium excretion. Both the antidiuretic and the antinatriuretic effects of 5-HT are thought to involve a direct effect on tubular function, although the antinatriuretic could be partially mediated by

5-HT-stimulated aldosterone secretion. Similarly, in dogs, 5-HT can produce both antidiuresis and antinatriuresis, and these effects are inhibited by the non-selective 5-HT₁ antagonist methysergide. In humans, most studies show that 5-HT infusions lead predominantly to an antinatriuretic effect, with no significant or only modest antidiuresis. Similar responses are observed in response to 5-HTP.

Renal Sources of Serotonin

Although blood platelets store 5-HT, which can be released in certain disease states characterized by intravascular coagulation, it is unlikely that platelet 5-HT contributes to the effects of 5-HT on renal function in normal physiological conditions. Renal tissue, however, can synthesize 5-HT from the precursor 5-HTP via the action of L-AAAD, which is found throughout the renal cortex and also in the renal medulla. Immunostaining for L-AAAD is particularly intense in the renal tubules of the rat kidney, the site of action for 5-HT in relation to its antinatriuretic effect.

Actions of Serotonin Renal Prodrugs

Renal prodrugs are essentially precursors of biologically active endogenous compounds that are preferentially metabolized to the active compound in the kidney. An intravenous infusion of either 5-HTP or the renal prodrug γ -L-glutamyl-5-hydroxy-L-tryptophan (γ -glutamyl-5-HTP) in humans reduces sodium excretion and increases urinary 5-HT excretion due to their metabolism to 5-HT within the kidney. In the presence of carbidopa, which inhibits the activity of L-AAAD, the sodium retention caused by these drugs and their metabolism to 5-HT are completely abolished.

Conclusion

5-HT is a stimulus for both mineralocorticoid (aldosterone) and glucocorticoid (cortisol and corticosterone) secretion. Within the adrenal cortex, 5-HT is a specific stimulus for aldosterone secretion in the zona glomerulosa. In the kidney, the renal hemodynamic effects of 5-HT still require clarification, especially in humans, whereas the antinatriuretic effects of 5-HT are reproducible in several species, including humans. The enzyme L-AAAD is likely to play a key role in the local production of 5-HT in the adrenal cortex and in the kidney. It seems logical that the actions of 5-HT within the adrenal cortex and the kidney are coordinated in order to maintain sodium homeostasis, and it will therefore be important to understand the mechanisms that regulate the activity of L-AAAD in these tissues.

See Also the Following Articles

Adrenal Cortex; Aldosterone and Mineralocorticoid Receptors; Glucocorticoids, Overview.

Further Reading

- Adler, S. (1977). Serotonin and the kidney. In: Essman, W. B. (ed.) *Serotonin in health and disease* (vol. 4), pp. 99–137. New York: Spectrum.
- Burns, N., Brett, L., Olverman, H. J., et al. (1996). The role of L-aromatic amino acid decarboxylase in serotonin-stimulated aldosterone secretion in response to salt intake. *Endocrine Research* **22**, 577–578.
- Collis, M. G. and Vanhoutte, P. M. (1977). Vascular reactivity of isolated perfused kidneys from male and female spontaneously hypertensive rats. *Circulation Research* **41**, 759–767.
- Contesse, V., Hamel, C., Lefebvre, H., et al. (1996). Activation of 5-hydroxytryptamine(4) receptors causes calcium influx in adrenocortical cells: involvement of calcium in 5-hydroxytryptamine-induced steroid secretion. *Molecular Pharmacology* **49**, 481–493.
- Davies, E., Edwards, C. R. W. and Williams, B. C. (1991). Serotonin stimulates calcium influx in isolated rat adrenal zona glomerulosa cells. *Biochemical and Biophysical Research Communications* **179**, 979–984.
- Davies, E., Rossiter, S., Edwards, C. R. W., et al. (1991). Serotonergic stimulation of aldosterone secretion in the rat *in vivo*: role of the renin-angiotensin system. *Journal of Endocrinology* **130**, 347–355.
- Davies, E., Rossiter, S., Edwards, C. R. W., et al. (1992). Serotonergic stimulation of aldosterone secretion *in vivo*: role of the hypothalamo-pituitary adrenal axis. *Journal of Steroid Biochemistry and Molecular Biology* **42**, 29–36.
- Erspamer, V. (1966). Peripheral physiological and pharmacological actions of indolealkylamines. *Handbook of Experimental Pharmacology* **19**, 245–359.
- Hinson, J. P., Vinson, G. P., Pudney, J., et al. (1989). Adrenal mast cells modulate vascular and secretory responses in the intact adrenal gland of the rat. *Journal of Endocrinology* **121**, 252–260.
- Holmes, M. C., Di Renzo, G., Beckford, U., et al. (1982). The role of serotonin in the control of secretion of corticotropin releasing factor. *Journal of Endocrinology* **93**, 151–160.
- Lefebvre, H., Contesse, V., Delarue, C., et al. (1993). Effect of the serotonin-4 receptor agonist zacopride on aldosterone secretion from human adrenal cortex: *in vivo* and *in vitro* studies. *Journal of Clinical Endocrinology and Metabolism* **77**, 1662–1666.
- Li Kam Wa, T. C., Freestone, S., Samson, R. R., et al. (1993). A comparison of the renal and neuroendocrine effects of two 5-hydroxytryptamine renal prodrugs in normal man. *Clinical Science* **85**, 607–614.
- Li Kam Wa, T. C., Freestone, S., Samson, R. R., et al. (1994). The antinatriuretic action of γ -L-glutamyl-5-hydroxy-L-tryptophan is dependent on its decarboxylation to 5-hydroxytryptamine in normal man. *British Journal of Clinical Pharmacology* **38**, 265–269.
- McCubbin, J. W., Kaneko, Y. and Page, I. H. (1962). Inhibition of neurogenic vasoconstriction by serotonin. *Circulation Research* **11**, 74–83.
- Muller, J. (1987). *Regulation of aldosterone biosynthesis: physiological and clinical aspects*. New York: Springer-Verlag.
- Osim, E. E. and Wyllie, J. H. (1983). Loss of 5-hydroxytryptamine from mammalian circulating platelets. *Journal of Physiology* **340**, 77–90.
- Shoji, T., Tamaki, T., Fukui, K., et al. (1989). Renal haemodynamic responses to 5-hydroxytryptamine(5-HT): involvement of the 5-HT receptor subtypes in the canine kidney. *European Journal of Pharmacology* **171**, 219–228.
- Stier, C. T., McKendall, G. and Itskovitz, H. D. (1984). Serotonin formation in non-blood perfused rat kidneys. *Journal of Pharmacology and Experimental Therapeutics* **228**, 53–56.
- Vinson, G. P., Pudney, J. A. and Whitehouse, B. J. (1985). The mammalian adrenal circulation and the relationship between adrenal blood flow and steroidogenesis. *Journal of Endocrinology* **105**, 285–294.
- Whiting, M. V. and Cambridge, D. (1995). Canine renovascular responses to sumatriptan and 5-carboxyamido-tryptamine: modulation through endothelial 5HT₁-like receptors by endogenous nitric oxide. *British Journal of Pharmacology* **114**, 969–974.

Serotonin in Stress

S Kusljic and M van den Buuse

Mental Health Research Institute, Parkville, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

Stress and Psychopathology: Possible Role of Serotonin
Stress, Serotonergic Drugs, and Human Psychopathology
Effect of Stress on Serotonin Parameters
Effect of Serotonergic Drugs on Stress Responses
Conclusion

Glossary

<i>Hypothalamic-pituitary-adrenocortical (HPA) axis</i>	The pathway that regulates glucocorticoid hormones secretion in the periphery.
<i>Raphe nuclei</i>	Groups of brain cells that are the main source of serotonin in the forebrain.
<i>Serotonin (5-hydroxytryptamine, 5-HT)</i>	A widely distributed monoamine that has been involved in almost every function of the brain.

Stress and Psychopathology: Possible Role of Serotonin

Serotonin (5-hydroxytryptamine, 5-HT) is a widely distributed monoamine that has been involved in almost every function of the brain. It is known that serotonin acts through a complex receptor system (5-HT₁–5-HT₇), with the number of receptors showing regional variation in terms of their distribution. The brain serotonergic system originates predominantly from cell groups located in either the dorsal or the median raphe nuclei, but some cell bodies are located in the ventrolateral region of the midbrain. Serotonergic neurons have a major influence on the regulation of neuroendocrine function, cardiovascular and respiratory activity, sleep, pain sensitivity, and mood. Serotonin furthermore acts as a growth factor during embryogenesis, and serotonin receptor activity forms a crucial loop in the pathway leading to changes in brain structure. There is extensive evidence for a role of serotonin in human illnesses associated with disrupted brain functioning.

Similarly, stress is involved in many aspects of human disease. At least part of this involvement could be caused by or associated with effects of stress on the serotonergic system and its widespread functions. For example, some of the physiological changes

associated with the stress response include enhanced cardiovascular output and respiration, modulation of the immune response, increased cerebral perfusion, and mobilization of energy to maintain brain function. Also more subtle changes in behavioral functioning caused by stress, e.g., arousal, anxiety, and cognition, may involve brain serotonin. Psychological stress, furthermore, can have an important role in the onset and course of mental illness, including depression, anxiety, and schizophrenia.

The influence of stressful life events in humans has received widespread attention. Exposure to early life stress is associated with neurobiological changes in children and adults, which may underlie the increased risk of psychopathology. There is substantial evidence that early life events influence serotonergic function in the brain as well as brain development and subsequent adult behavior. For example, the effects of early traumatic events on the developing brain and the biological abnormalities found in subjects with schizophrenia include hyperactivity of the hypothalamic-pituitary-adrenocortical (HPA) axis and serotonin abnormalities, and structural changes to the brain such as damage to the hippocampus, a brain region enriched in serotonin receptors. Stress accelerates both the release and turnover of brain serotonin. This can lead to restoration of homeostasis, whereas chronic stress can induce a long-lasting imbalance in central nervous neurotransmitter systems, including serotonin. Abuse and neglect during early childhood is linked to low levels of serotonin and risk for hostility, aggression, substance abuse, and suicide.

Several other lines of evidence have linked brain serotonin activity with human psychopathology. A polymorphism in the promoter region of the serotonin transporter gene has been suggested to contribute to anxiety in humans, and the presence of an allele encoding a short sequence in this promoter region increases the risk for the development of a depressive disorder. Studies in a population of rhesus monkeys have shown that the short sequence allele corresponds to low concentrations of the serotonin metabolite 5-hydroxyindolacetic acid in the cerebrospinal fluid. Mutations in the gene for the enzyme tryptophan hydroxylase may also predispose to mood disturbances. Alterations in the several synaptic proteins, transcription factors, and proteins involved in neuronal growth/differentiation that are modified in their expression in experimental models of chronic stress can be reversed by treatment with antidepressants, thus suggesting an important role for the serotonin system.

Stress, Serotonergic Drugs, and Human Psychopathology

Evidence for the importance of the serotonin system in depression and anxiety disorders has increased substantially over the recent years, and stress plays a major role in the development of these disorders. The first effective antidepressants (tricyclic antidepressants and monoamine oxidase inhibitors) such as imipramine and iproniazid relied at least in part on their ability to increase serotonin and norepinephrine levels in the synapse. Research on the biological effects of these drugs led to the hypothesis of a deficiency in both these cerebral monoamines in depression. Antidepressants act not only on depressive symptoms but also on anxiety. Most of the evidence suggests that depletion of serotonin in the brain increases anxiety in humans. Thus, drugs that increase the levels of serotonin have anxiolytic effects. A new generation of antidepressants that selectively inhibit the neuronal reuptake of serotonin (selective serotonin reuptake inhibitors [SSRIs]) specifically focused on serotonin neurotransmission with little or no activity on noradrenergic mechanisms. Several drugs, including fluoxetine, citalopram, and paroxetine, had the same clinical efficacy, yet appeared to be better tolerated and lack most of the adverse side effects of tricyclic antidepressants and monoamine oxidase (MAO) inhibitors. Originally introduced for the treatment of depression, the SSRIs have been used as the first-line treatment for a broad range of psychiatric disorders, including anxiety, obsessive-compulsive disorder, social phobia, and posttraumatic stress disorder. The target of SSRIs is the presynaptic serotonin transporter protein, which is responsible for clearance of synaptic and plasma serotonin.

Effect of Stress on Serotonin Parameters

Stress increases concentrations of serotonin and its metabolites in several brain regions indicating increased turnover rates of the neurotransmitter. Moreover, stress increases the activity of tryptophan hydroxylase, the biosynthetic enzyme confined to serotonin-producing cells. Therefore, studies of the effect of stress on serotonergic function in laboratory animals may lead to a better understanding of the biological bases of psychopathology.

Repeated exposure of rats to forced swimming increased serotonin concentrations in the striatum but reduced serotonin levels in the lateral septum. Chronic restraint stress in rats revealed increased serotonin turnover in the hippocampus, whereas chronic social defeat in mice upregulated expression of the MAO-A gene in the raphe nuclei. MAO-A is an isoenzyme that

preferentially oxidizes serotonin. An upregulated expression of the MAO-A gene accelerates serotonin degradation and is associated with enhanced expression of the serotonin transporter.

Furthermore, early life stress such as maternal deprivation can affect several aspects of serotonin transmission later in life. For example, maternal deprivation at postnatal day 3 has long-term effects on the serotonergic responsiveness of CA1 pyramidal neurons in the hippocampus of the rat. Adult CA1 hippocampal cells from maternally deprived rats compared to nondeprived rats show a significantly attenuated response to serotonin. This attenuation is also seen in adult rats subjected to high doses of exogenously administered corticosterone or to chronic stress.

Corticosterone regulates hippocampal physiology, particularly in the CA1 region, and the 5-HT_{1A} receptor effector pathway. However, prolonged elevation of the corticosteroid concentration does not alter 5-HT_{1A} receptor mRNA levels in pyramidal neurons. In addition, in rats that were subjected to unpredictable stressors for 3 weeks and that displayed increased adrenal weight and attenuated body weight compared with controls, the 5-HT_{1A} receptor mRNA expression was not changed after chronic stress exposure. Therefore, stress may alter serotonergic function possibly by posttranslational modification of the specific serotonin receptor subtype or changes in the serotonin receptor signal transduction pathway.

Effect of Serotonergic Drugs on Stress Responses

Serotonin and HPA Axis Activity

The primary neuroendocrine response to acute stress includes increased secretion of cortisol in humans or corticosterone in rodents. Glucocorticoid hormones are the final effectors of the HPA axis that participate in the control of homeostasis and the response of the body to stressors. Corticosterone acts on both the high-affinity mineralocorticoid receptors and the lower affinity glucocorticoid receptors. Mineralocorticoid receptors are found in some limbic brain structures that are also rich in serotonin innervation, such as the hippocampus, whereas glucocorticoid receptors are found in several brain regions, including frontal cortex and hypothalamic paraventricular nucleus. Glucocorticoids play a key regulatory role in the termination of the stress response by exerting negative feedback at the levels of hypothalamus and pituitary, and there is increasing evidence to link serotonin with HPA regulation. Serotonin-containing neurons directly innervate corticotropin-releasing hormone (CRH)-containing cells located in the paraventricular

nucleus of the hypothalamus. Serotonergic innervation of the paraventricular nucleus modulates the release of CRH, leading to the release of adrenocorticotrophic hormone (ACTH), which triggers glucocorticoid secretion from the adrenal cortex. Furthermore, administration of the serotonin precursors L-tryptophan and 5-hydroxytryptophan stimulates ACTH and cortisol release, as does the serotonin releasing agent fenfluramine. 5-HT_{1A} and 5-HT_{2A} receptors are the main serotonin receptors mediating serotonergic stimulation of the HPA axis. Activation of 5-HT_{1A} and 5-HT_{2A} receptors increases plasma corticosterone levels. Conversely, the stimulatory effect of serotonin on CRH secretion was antagonized by the 5-HT₁ and 5-HT₂ receptor antagonist metergoline and by the selective 5-HT₂ receptor antagonists ketanserin and ritanserin. Furthermore, administration to rats of the 5-HT_{1A} agonist, 8-hydroxy-2-(di-n-propylamino) tetralin (8-OH-DPAT), resulted in activation of the HPA axis as evidenced by increased plasma ACTH and corticosterone concentrations. Chronic administration of 8-OH-DPAT resulted in increased CRH concentrations in the piriform cortex, entorhinal cortex, amygdala, and hippocampus. However, administration of 8-OH-DPAT had no effect on the 5-HIAA/5-HT ratio. In contrast, 8-OH-DPAT treatment was able to decrease plasma corticosterone responses following acoustic stress, conditioned fear, and cocaine administration. This indicates that the predominant effect of 8-OH-DPAT on adrenocortical responses is mediated at postsynaptic sites not involved in the regulation of cerebral biogenic amine metabolism.

Acute administration of SSRIs stimulates the HPA axis, whereas continuous treatment desensitizes it. The immediate effect of SSRIs on the HPA axis function occurs in the hypothalamic paraventricular nucleus via postsynaptic serotonin receptors on CRH neurons. It is likely that serotonin 5-HT_{1A} and 5-HT_{2A} postsynaptic receptors are indirectly stimulated by SSRI administration.

Physiological Responses: Blood Pressure/Body Temperature

A typical neuroendocrine response also involves the increased secretion of catecholamines (epinephrine and norepinephrine) from the sympathetic nervous system and adrenal medulla. Central catecholamine neurons play a critical role in cardiovascular responses to stress-related stimuli. Hypertension and increased heart rate are key components of the cardiovascular response to stress, leading to the essential increase in tissue perfusion that has to occur in stressful situations. Among the complex interactive systems that contribute to the central regulation of cardiovascular

function, many areas are under the direct influence of serotonergic pathways. Both blood pressure and heart rate may be increased or decreased by the activation of distinct brain serotonin receptors. For example, pharmacological activation of central 5-HT₃ receptors evokes a significant hypotensive effect in rats. By contrast, activation of 5-HT₂ receptors in the lateral hypothalamus and in the paraventricular nucleus increases blood pressure and heart rate. Intraventricular injection of the 5-HT_{2C} agonist (1-(3-chlorophenyl)piperazine) (mCPP) elicited a significant increase in blood pressure in nonstressed animals. However, injection of mCPP did not change the hypertensive or heart rate responses induced by restraint stress. Furthermore, cardiovascular effects caused by the central administration of mCPP were blocked by pretreatment with the 5-HT_{2C} antagonist SDZ SER 082. Conversely, administration of SDZ SER 082 blunted stress-induced hypertension without modifying stress-induced increase in heart rate. Several serotonergic drugs with full or partial 5-HT_{1A} receptor agonist activity, such as 8-OH-DPAT, buspirone, and MDL 73,005EF, were shown to reduce both blood pressure and heart rate responses to open-field novelty stress (Figure 1). This effect was seen without changes in the exploratory locomotor activity response in this test.

Administration of serotonergic drugs may affect other physiological responses associated with stress, such as body temperature changes. Acute administration of 3,4-methylenedioxymethamphetamine (MDMA, ecstasy) is associated with rapid hyperthermia that can last several hours and is linked to monoamine release from nerve endings in the brain. In addition to hyperthermia, other, often fatal, physiological problems occur, including tachycardia, renal failure, edema, and thrombocytopenia. MDMA acts by releasing monoamines, including serotonin, but the receptor mechanisms underlying MDMA-induced hyperthermia and cutaneous vasoconstriction remain unclear. MDMA-induced hyperthermia is associated with reduced cutaneous blood flow, which may contribute to the hyperthermia by diminishing normal heat transfer from the body to the environment. The specific 5-HT_{2A/2C} agonist DOI also causes marked hyperthermia, suggesting the possibility that stimulation of 5-HT_{2A/2C} receptors might cause sympathetically mediated cutaneous vasoconstriction. The 5-HT_{2A} receptor antagonist ketanserin reversed the cutaneous vasoconstriction induced by DOI in rabbits and tail vasoconstriction and hyperthermia in rats. Furthermore, administration of the antipsychotic clozapine reverses hyperthermia and sympathetically mediated cutaneous vasoconstriction induced by MDMA in rabbits and rats. These mechanisms

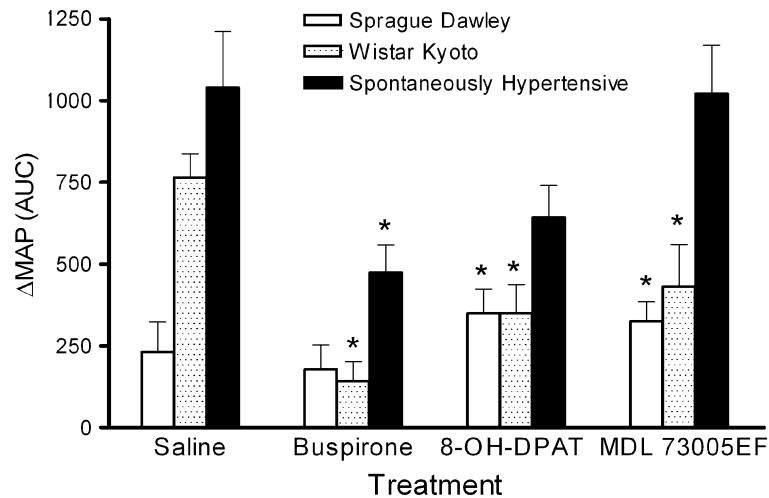


Figure 1 Blood pressure increases caused by open-field novelty stress are modulated by serotonergic drugs in Sprague-Dawley rats, Wistar-Kyoto rats (WKY) and spontaneously hypertensive rats (SHR). Pretreatment with 0.5 mg/kg of the 5-HT_{1A} receptor partial agonist and dopamine D₂ receptor antagonist, buspirone, reduced the blood pressure increases in WKY and SHR. Pretreatment with 0.25 mg/kg of the 5-HT_{1A} receptor full agonist, 8-OH-DPAT, enhanced the blood pressure increase in Sprague Dawley rats, but reduced it in WKY, with no significant change in SHR. This same pattern of strain differences in responding was seen with pretreatment with the 5-HT_{1A} receptor partial agonist, MDL 73,005EF. The rats were placed in 90 cm open field while blood pressure was measured with radiotelemetry. For more data and details, see Van den Buuse and Wegener (2005).

support a role of serotonin mechanisms in stress-induced changes in body temperature. In addition, clozapine and ketanserin, by interactions with 5-HT receptors or by other serotonergic mechanisms, may be of therapeutic benefit in humans suffering from life-threatening hyperthermia such as that induced by ingestion of MDMA.

Conclusion

Serotonin is a widely distributed monoamine that has been involved in almost every function of the brain. Stress influences 5-HT activity, and this interaction could be involved in the development of stress-related illnesses, including depression and schizophrenia. Conversely, serotonergic drugs modulate activity of the HPA axis and physiological responses to stress. Thus, there is a direct reciprocal relationship between stress and activity of the serotonin system in the brain.

See Also the Following Articles

Adrenocorticotrophic Hormone (ACTH); Blood Pressure; Corticosteroid Receptors; Depression and Manic-Depressive Illness; Disease, Stress Induced; Glucocorticoids, Effects of Stress on; Glucocorticoids, Overview; Glucocorticoids, Role in Stress; Hypothalamic-Pituitary-Adrenal;

Psychological Stressors, Overview; Schizophrenia; Serotonin; Selective Serotonin Reuptake Inhibitors (SSRIs).

Further Reading

- Buwalda, B., Kole, M. H., Veenema, A. H., et al. (2005). Long-term effects of social stress on brain and behavior: a focus on hippocampal functioning. *Neuroscience and Biobehavioral Reviews* **29**, 83–97.
- Carrasco, G. A. and Van de Kar, L. D. (2003). Neuroendocrine pharmacology of stress. *European Journal of Pharmacology* **463**, 235–272.
- Dinan, T. G. (1996). Serotonin and the regulation of hypothalamic-pituitary-adrenal axis function. *Life Sciences* **58**, 1683–1694.
- Flügge, G., Van Kampen, M. and Mijster, M. J. (2004). Perturbations in brain monoamine systems during stress. *Cell and Tissue Research* **315**, 1–14.
- McEwen, B. S. (2003). Early life influences on life-long patterns of behavior and health. *Mental Retardation and Developmental Disabilities Research Reviews* **9**, 149–154.
- Saphier, D. and Welch, J. E. (1995). Effects of the serotonin_{1A} agonist, 8-hydroxy-2-(di-n-propylamino)tetralin on neurochemical responses to stress. *Journal of Neurochemistry* **64**, 767–776.
- Van den Buuse, M. and Wegener, N. (2005). Involvement of serotonin_{1A} receptors in cardiovascular responses to stress: a radio-telemetry study in four rat strains. *European Journal of Pharmacology* **507**, 187–198.

Serotonin Transporter Genetic Modifications

K Sugden

Social, Genetic and Developmental Psychiatry Centre,
London, UK

© 2007 Elsevier Inc. All rights reserved.

The Serotonin Transporter

The *5-HTT* Gene

Functional Polymorphisms of the *5-HTT* Gene

The Serotonin Transporter-Linked Polymorphic Region
and Stress

*Transcription
factor binding
site*

A consensus sequence within a gene, promoter, or regulatory region that is recognized and bound by a transcription factor that then regulates the transcription of that gene.

*Variable
number of
tandem repeats
(VNTR)*

A polymorphism within a DNA sequence that comprises varying numbers of repeating units. VNTRs can be microsatellites (usually repeat units of 2–8 bp) or minisatellites (more than 8 bp). VNTRs are hypermutable and often give rise to many alleles of different lengths.

Glossary

Enhancer

A short region of DNA that can be bound with proteins (such as transcription factors) to enhance transcription levels of genes. An enhancer does not necessarily have to be located close to the genes it affects, although some enhancers are found within intronic sequences.

*Messenger
RNA (mRNA)*

The primary sequence transcribed from the DNA template. mRNA has intronic sequences (i.e., sequences not represented in the transcript exiting the nucleus) removed by splicing and encodes the amino acid sequence of a protein.

*Poly-
adenylation*

The cleavage of mRNA and addition of 50–200 adenosine residues, a poly-A tail. Polyadenylation is an important post-transcriptional modification that ensures the appropriate termination of the molecule at the 3' end and helps in both the export of the mRNA from the nucleus and protection from exonucleases.

*Polyadenyla-
tion site*

The point of polyadenylation of newly transcribed mRNA.

*Serotonin
(5-hydroxy-
tryptamine,
5-HT)*

A protein present in many tissues, especially blood and nervous tissue; it stimulates a variety of smooth muscles and nerves and functions as a neurotransmitter.

*Single
nucleotide
polymorphisms
(SNPs)*

Single-base changes in a DNA sequence, for example a C to T substitution. SNPs can be found in the coding (exonic) and noncoding region (e.g., untranslated region, promoter, or intron) of a gene, and coding SNPs can be synonymous (no change in resultant amino acid) or non-synonymous (change in resultant amino acid).

Transcription

The synthesis of an RNA copy from a sequence of DNA; the first step in gene expression.

The Serotonin Transporter

The serotonin transporter (*5-HTT*) is important in the termination and modulation of serotonergic neurotransmission by sodium-dependent reuptake of serotonin into the presynaptic neuron, which thus terminates the synaptic actions of serotonin and recycles it into the neurotransmitter pool. The transporter is the site of action of a number of commonly used antidepressants, such as selective serotonin reuptake inhibitors (SSRIs), fluoxetine (ProzacTM), sertraline (ZoloftTM), and paroxetine (PaxilTM), and tricyclic antidepressants. A number of psychostimulants such as cocaine, methylphenidate, and MDMA (ecstasy) may also interact with *5-HTT*. For this reason, *5-HTT* has become one of the most widely studied genes in the field of behavior, especially in the context of psychiatric disorders such as depression, neuroticism, and suicidal behavior.

The *5-HTT* Gene

The gene encoding the *5-HTT* protein in humans has been mapped to the chromosomal location 17q11.2 and consists of 14 exons spanning approximately 38 kb. Several studies have focused on screening the genomic region for variation, and a number of polymorphisms have been characterized. Currently, at least 62 SNPs with frequency data within the gene, untranslated, or promoter region of *5-HTT* are listed in the Entrez SNP database (dbSNP). Most of these SNPs have not yet been investigated in the context of complex behavior, and any role they might have in regulation of the gene is still unknown.

Functional Polymorphisms of the *5-HTT* Gene

There are some regions within *5-HTT* that have been demonstrated to affect the expression of the gene.

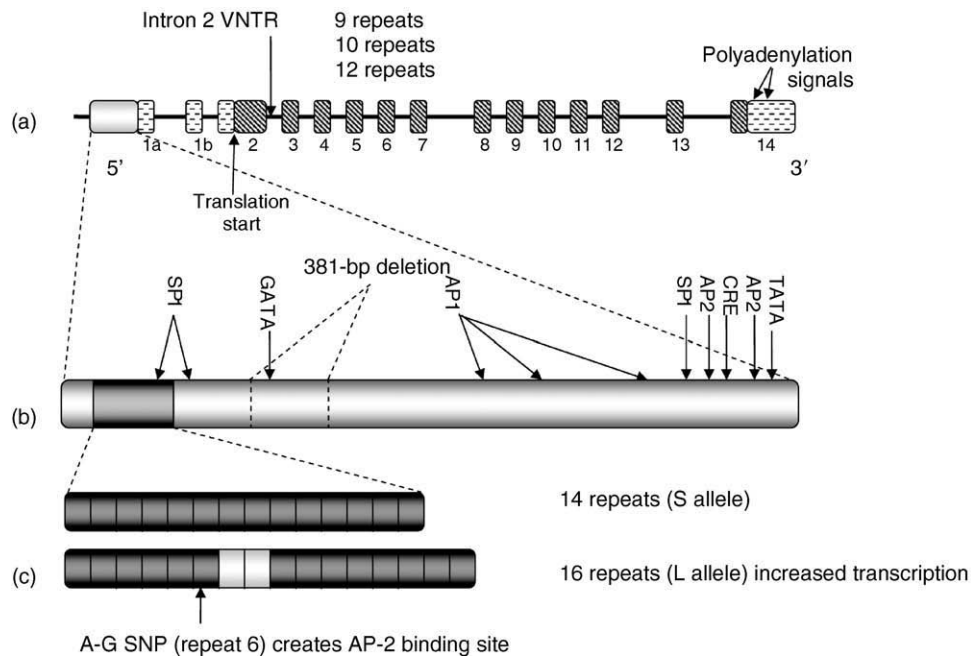


Figure 1 Serotonin transporter gene (not to scale). a, Organization of the gene; b, promoter region; c, *5-HTTLPR* as either the S or L allele. In a, the coding exons are represented by diagonally marked boxes; noncoding exons are represented by dashed boxes. The position of the promoter (light grey) and intron 2 VNTR are shown. In b, the position of some postulated transcription factor binding sites are shown by arrows; the region of the 381-bp deletion is demonstrated by dashed lines, as is the position of the *5-HTTLPR* (dark grey box). In c, the light boxes represent the position of the two repeat units that are inserted in the L allele; the position of the A-G SNP is also shown.

The most common of these are described next (see [Figure 1](#)).

The Serotonin Transporter-Linked Polymorphic Region

One of the most documented variants within *5-HTT* is the serotonin transporter-linked polymorphic region (*5-HTTLPR*; [Figure 1c](#)), a polymorphic GC-rich repetitive element within the proximal 5' regulatory region. In humans (and indeed in primates), the *5-HTTLPR* is polymorphic, consisting of a variable number of 20- to 23-bp repeats. The *5-HTTLPR* exists as two common alleles: the short allele (S, 14 repeats) and the long allele (L, 16 repeats). There is evidence that this polymorphic region has functional significance; in JAR (human carcinoma) cells, the L allele confers higher activity than the S allele, whereas lymphoblastoid cells with the LL genotype have 1.4–1.7 times the amount of mRNA as SS or SL cells. Furthermore, the uptake of labeled serotonin in LL genotype cells is 1.9–2.2 times that in cells carrying one or two copies of the S variant. More recently, it has been shown that the *5-HTTLPR* gene actually exhibits a number of alleles of varying repeat numbers. Furthermore, the S and L alleles are themselves variable, so that,

although their length is similar, SNPs occur within their repeating units. In total, at least 14 alleles of varying length and sequence have been identified, four different S alleles (14-A to 14-D), six different L alleles (16-A to 16-F), and alleles consisting of 13, 15, 18, 19, 20, and 22 repeats. The longer alleles appear to exhibit a silencer effect on the gene, having lower activity than 14- and 16-repeat variants; however, these alleles are extremely rare and appear to be found in specific populations only, such as the 20-repeat allele in the Japanese population. Recently, it has been demonstrated that an A-G SNP (dbSNP reference number rs25531) within the sixth repeat of the *5-HTTLPR*, represented by the 14-D and 16-D alleles, creates a binding site for the transcription factor AP-2 that might affect the rate of transcription of the gene.

Recently however, the functional effect of the *5-HTTLPR* polymorphism has been called into question because *in vivo* experiments failed to find the same genotypic effect as that observed in cell lines. For example, the *5-HTTLPR* genotype has no effect on the 5-HTT binding of a serotonergic ligand, a measure of 5-HTT availability, in human brain. In other words, if the *5-HTTLPR* is driving differential

transcription, it has no eventual effect on *5-HTT* availability. Similarly, some experiments suggest that *5-HTT* binding in the brain does not differ by *5-HTTLPR* genotype. If these results are robust, then it may be that the expression patterns in cell-based assays do not accurately reflect the pattern of expression *in vivo*. Indeed, it has been shown that the regulation of the *5-HTT* depends on the presence of the substrate (serotonin, 5-HT) rather than on any genetically driven effect. This might be an indication of the many levels of complex regulation that genes are subject to *in vivo* that may not be accurately modeled by cell-line *in vitro* approaches; the functional effect of the *5-HTTLPR* observed in cells might be adapted by the complex intra- and extra-cellular factors that brain neurons are exposed to.

Intron 2 Variable Number of Tandem Repeats

Along with the *5-HTTLPR*, there are other proposed transcriptional regulatory regions within the *5-HTT*. One example is a VNTR within intron 2 that has been shown to have differential enhancer activities in the mouse embryo and in human embryonic stem (ES) cells (Figure 1a). The VNTR consists of 9–12 copies of a 16- or 17-bp repeat, the 12-repeat allele (STin12) having greater enhancer activity than the 10-repeat allele (STin10) in both model systems. This ability to regulate expression is cell-specific in humans, in that it is observed in ES cells but not HeLa cells. However, as in the case of the *5-HTTLPR*, the regulation is much more complex than the effect of repeat length; the VNTR does not comprise a succession of perfect repeats, rather there are seven varying repeat sequences and the differences between alleles is described by the deletion of specific repeat elements. The functional effect of the VNTR is regulated by at least two transcription factors – CTCF-binding factor (CTCF) and Y box binding protein (YB-1) – that bind to a sequence that is deleted in the nine-repeat allele (STin9). YB-1 is inhibitory, so binding to the VNTR results in decreased expression. However, the addition of YB-1 results in a release from inhibition of the STin9 VNTR in COS-7 cells because the inhibitory site is deleted. This effect is antagonized by CTCF and is not observed with either the STin10 or STin12.

This evidence suggests that *5-HTT* expression is differentially regulated by these two regions. However, little is known about their combined effect on transcription. The only evidence to date indicates that the expression of the gene is highest in lymphoblastoid cells having both *5-HTTLPR* L/L and STin12/STin12 genotypes. The expression is lowest in cells with at least one S and one STin10 allele. Based on previous evidence of higher expression of *5-HTT* with

LL- and STin12/STin12-compared to S- and STin10-containing genotypes, this finding could be said to fit an *a priori* hypothesis; but some replication is necessary before any firm conclusions on the synergistic action of the *5-HTTLPR* and intron 2 VNTR can be made.

Other Regulatory Elements

The transcriptional effect of the *5-HTTLPR* and intron 2 VNTR is further compounded by the observation of splicing variation within the 5' untranslated region, so that two alternative forms of *5-HTT* mRNA exist. Within the human brain and placenta, exon 1B is alternatively spliced and appears in transcripts to a varying degree, whereas exon 1A seems to be the principal transcription initiation point. This type of alternative splicing could help to specify cell-specific activity of the *5-HTT* and ensure appropriate levels of neurotransmission in different tissues. This tissue-specific splicing of 5' untranslated exons is also observed in the brain and gut in rats, indicating that it is a common mechanism of regulating protein expression. Furthermore, tissue-specific mRNA splicing might also be precipitated through multiple polyadenylation sites found in exon 14. This mechanism dictates alternate transcription stop sites that give rise to mRNA transcripts of varying length that, like alternative splicing of exons, might regulate tissue-specific expression. It has also been proposed that the structure of the *5-HTTLPR* is likely to result in a 381-bp deletion in the promoter region (Figure 1b) and that this deletion is present in 20–60% of genomic DNA derived from human brain and blood. Specific breakpoints suggest that the event is similar to V(D)J rearrangement often observed in antibody genes. This rearrangement results in looping out the 381-bp DNA sequence so that the two breakpoints are adjacent and are subsequently joined together, resulting in the deletion of the looped-out sequence.

The consequences of this deletion are not known, but it is suggested that the event is likely to be regulated by tissue- and *5-HTTLPR*-dependent mechanisms. However, this effect of mosaicism has yet to be consistently observed in multiple populations.

The Serotonin Transporter-Linked Polymorphic Region and Stress

The *5-HTT* has been implicated in the stress response because serotonin concentrations are increased within the synapse when under stress. Given that the *5-HTTLPR*, through gene-regulatory mechanisms, might affect the efficiency by which serotonin is

transported back in to presynaptic neuron, the *5-HTTLPR* is an ideal polymorphism candidate for studies of stress and its outcomes.

Animal studies have proven invaluable in this context because it is possible to accurately control both psychosocial and physiological effectors, unlike in humans. In Rhesus monkeys, there are variations in response to early rearing experience that are dependent on the primate *5-HTTLPR* genotype. Monkeys that possess an S allele have significantly lower cerebrospinal fluid (CSF) 5-hydroxyindoleacetic acid (5-HIAA, a serotonin metabolite) concentrations than their homozygous L-allele counterparts. This suggests that serotonin reuptake is lowered in monkeys with an S allele, but the effect is specific to those who were peer-reared (early stressful environment) rather than parent-reared. No differential effect of genotype is seen between parent-reared monkeys. This suggests that early experience has some effect on the functioning of the serotonin system within primates and that the *5-HTTLPR* genotype becomes important only when stressful events are experienced. There also exists a knockout mouse for the serotonin transporter, a genetically manipulated mouse that is deficient in the *5-Htt* gene. Stressed mice that are heterozygous or homozygous for the knockout mutation (i.e., mice that have lower or negligible levels of 5-HTT) have a greater response to stress (measured as increased levels of stress hormone adrenocorticotrophic hormone, ACTH) than homozygous wild-type mice. The levels of ACTH are the same in knockout and wild-type mice when they are not stressed, again suggesting that it is the variability in 5-HTT density that participates in interindividual differences in stress response.

In humans, neuroimaging research suggests that individual differences in stress perception are mediated by genetic variations in the *5-HTTLPR*. Individuals with one or two copies of the S allele of the *5-HTTLPR* exhibit greater amygdala neuronal activity to fearful stimuli than individuals homozygous for the L allele. This suggests that the SS- and SL-genotype individuals respond to the same stressful event to a greater magnitude than the LL individuals. In addition, it has been demonstrated that there is a differential response to brain-derived neurotrophic factor (BDNF) by *5-HTTLPR* genotype. BDNF is involved in activating the stress response, so this suggests that variability in *5-HTT* may influence variability in other downstream molecular processes that 5-HTT is associated with. Some of the strongest evidence to date that the *5-HTTLPR* can modulate the human stress response comes from a longitudinal study of almost 1000 individuals from Dunedin, New Zealand, in which it was shown that variation in the *5-HTTLPR* can modify the effect of stress on

the outcome of depression. Stressful life events, such as the breakdown of a marriage or unemployment can lead to depression; however, this is not true for all individuals who experience stress. It was found that individuals who had the LL genotype of the *5-HTTLPR* were protected against the effects of stress in that their rates of depression were much lower than SS- and SL-genotype individuals with similar levels of stressful life events. This effect was seen whether the stress was experienced proximally (within the previous 5 years) or earlier (in childhood). The rates of depression in individuals who had not experienced stress were similar regardless of genotype. Consequently, evidence from these diverse sources suggests that variation in the *5-HTTLPR* can influence psychopathological reactions to stress and that it can account for some of the variability in individual response.

See Also the Following Articles

Gender and Stress; Genetic Factors and Stress; Genetic Polymorphisms in Stress Response; Social Status and Stress.

Further Reading

- Benjamin, A. B., Ebstein, R. P. and Belmaker, R. H. (eds.) (2002). *Molecular genetics and the human personality*. Washington, DC: American Psychiatric Publishers.
- Bennett, A. J., Lesch, K. P., Heils, A., et al. (2002). Early experience and serotonin transporter gene variation interact to influence primate CNS function. *Molecular Psychiatry* 7, 118–122.
- Caspi, A., Sugden, K., Moffitt, T. E., et al. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 301(5631), 386–389.
- Hariri, A. R., Mattay, V. S., Tessitore, A., et al. (2002). Serotonin transporter genetic variation and the response of the human amygdala. *Science* 297(5580), 400–403.
- Heils, A., Teufel, A., Petri, S., et al. (1996). Allelic variation of human serotonin transporter gene expression. *Journal of Neurochemistry* 66(6), 2621–2624.
- Klenova, E., Scott, A. C., Roberts, J., et al. (2004). YB-1 and CTCF differentially regulate the 5-HTT polymorphic intron 2 enhancer which predisposes to a variety of neurological disorders. *Journal of Neuroscience* 24(26), 5966–5973.
- Kraft, J. B., Slager, S. L. and McGrath, P. J. (2005). Sequence analysis of the serotonin transporter and associations with antidepressant response. *Biological Psychiatry* 58(5), 374–381.
- Lesch, K. P. (1998). Serotonin transporter and psychiatric disorders: listening to the gene. *Neuroscientist* 4(1), 25–34.
- Lesch, K. P. (2001). Serotonin transporter: from genomics and knockouts to behavioral traits and psychiatric disorders. In: Briley, M. & Sulser, F. (eds.) *Molecular*

- genetics of mental disorders*, pp. 221–267. London: Martin Dunitz.
- Murphy, D. L., Li, Q., Engel, S., et al. (2001). Genetic perspectives on the serotonin transporter. *Brain Research Bulletin* 56(5), 487–494.
- Nakamura, M., Ueno, S., Sano, A., et al. (2000). The human serotonin transporter gene linked polymorphism

(5-HTTLPR) shows ten novel allelic variants. *Molecular Psychiatry* 5(1), 32–38.

Relevant Website

Entrez SNP database. <http://www.ncbi.nlm.nih.gov/SNP/>

Sex Differences in Human Stress Response

B M Kudielka

University of Trier, Trier, Germany

D H Hellhammer

University of Trier, Trier, Germany

C Kirschbaum

Technical University of Dresden, Dresden, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by B M Kudielka, D H Hellhammer, C Kirschbaum, volume 3, pp 424–428, © 2000, Elsevier Inc.

Sex Differences in Health and Disease
Sex Differences in Subjective Aspects of the Human Stress Response
Sex Differences in Physiological Aspects of the Human Stress Response

Glossary

Cortisol

A glucocorticoid hormone produced by the adrenal cortex. Cortisol exhibits a pronounced circadian rhythm and responds to a wide range of stressors. It has many physiological effects and virtually every single nucleated cell in the body has cortisol receptors. It is predominantly (90–95%) bound to binding proteins in blood; only 5–10% of the total plasma cortisol circulates as biologically active, unbound, free cortisol.

Hypothalamic-pituitary-adrenal (HPA) axis

A central regulatory and control system of the organism that connects the central nervous system with the endocrine system. Under stress, the hypothalamus secretes corticotropin releasing hormone (CRH), which provokes the release of adrenocorticotrophic hormone (ACTH) from the pituitary. ACTH triggers cortisol secretion from the adrenal cortex.

Lymphocytes

The functioning of the axis is controlled by several negative feedback loops.

White blood cells responsible for the specific responses of the immune system to antigens. Based on the tissue where the stem cells mature to the specific lymphocyte, they are called B cells or T cells. B cells mature in the bone marrow and form the cellular basis for the humoral immunity (e.g., antibody production), whereas thymus-derived T lymphocytes (e.g., T helper cells and T suppressor cells) represent the effector cells of the specific cellular immunity.

Sex steroids

Hormones produced in the gonads and, to a lesser degree, in the adrenal cortex. Estradiol and testosterone play major roles in the female and male reproductive system. In women, sex steroid concentrations show typical variations during the menstrual cycle. Sex steroids also influence the peripheral and central nervous system.

Sympathetic-adrenal-medullary (SAM) axis

Sympathetic nerves, originating in the central nervous system, stimulate the adrenal medulla, which in turn secretes epinephrine and norepinephrine. These catecholamines mobilize energy-producing mechanisms quickly and downregulate less important organ functions.

Sex Differences in Health and Disease

Men and women are at differential risks for a number of illnesses. This observation is reflected in the sex-specific prevalence rates for several diseases. Women suffer more often from autoimmune illnesses such as lupus erythematosus, multiple sclerosis, and neurodermatitis, whereas men are more prone to develop coronary heart diseases or infectious diseases. Concerning psychiatric disorders, women more often develop anxiety, depression, phobias, panic disorders,

or obsessive-compulsive disorder, whereas men more often are prone to antisocial behavior, substance abuse, or suicide. Although women consistently appear to report more physical and somatoform symptoms than men, the life expectancy for women is approximately 7 years longer than it is for men.

In addition to the clear sexual dimorphism with respect to certain diseases, sex differences exist in psychological and biological aspects of the human stress response. Investigation into sex-specific stress responses may help to reveal mechanisms that underlie somatic, psychosomatic, and psychiatric illnesses. Data showing that several diseases are associated with altered stress responses support this assumption.

On a behavioral level, psychiatric disorders such as depression and anxiety are interrelated with increased symptom reporting. On a physiological level, several of these disorders are accompanied by dysfunctional endocrine regulations. Major depression, susceptibility to infectious diseases, and cardiovascular problems seem to be associated with a hyperactive HPA axis, whereas autoimmune processes are associated more often with a hyporeactive HPA axis. The observation that (premenopausal) women (in contrast to men) seem to be protected from coronary heart disease parallels the finding that men show higher blood pressure changes and catecholamine increases under psychological stress. In addition to the described sex dimorphism with respect to disease and sex-specific stress responses in humans, it is generally accepted that exposure to stress can cause or intensify numerous diseases.

Thus, differences in health status may, at least in part, be related to sex differences in stress responses. Sex-specific stress responses do not seem to be epiphenomena of sex-specific prevalence rates of several disorders but, instead, appear to reflect causal links.

Sex Differences in Subjective Aspects of the Human Stress Response

Following Lazarus and Folkman, the extent of experienced stress is not determined by a situation or an event itself but is, instead, mediated by a three-phasic appraisal process. The importance or threat of the characteristics of a situation or an event is assessed in a primary appraisal loop. Simultaneously, in a secondary appraisal the amount of available coping capacity or resources is estimated. A final reappraisal determines the extent of experienced stress.

Several studies, ranging from everyday hassles to posttraumatic stress reactions, conclude that women subjectively experience more stress than men. A socio-cultural approach focusing on external environmental

factors emphasizes that women's lives are characterized more often by multiple social roles and lower socioeconomic status, which are both positively correlated with stressful life events. In addition to the finding that women experience more stress in everyday life, stress vulnerability and worry disposition also appear to be higher in women compared to men. This leads to the assumption that sex differences exist in people's readiness to perceive and appraise situations as stressful. This suggestion is supported by the finding that self-descriptions concerning self-esteem or self-confidence, self-efficacy, and sense of competence are slightly but consistently lower in women than in men. Although men repeatedly showed a more favorable locus of control and attribution style in questionnaire-derived studies, Miller and Kirsch concluded that data about sex differences in attributions, locus of control, and perceptions of control during stressful episodes remain inconsistent. Recent findings by Schulz and co-workers revealed that women's higher average scores in perceived stress, chance control orientation, stress-susceptibility, depressivity, and bodily complaints, as well as their lower scores in self-esteem and self-concept of ability, may be attributed to sex differences in worry disposition.

A higher sensitivity and attention to nonspecific stress symptoms appear to bring about more episodes of perceived stress in women. In contrast to women, men seem to be less sensitive in their perception of subtle bodily symptoms, and they are less prone to appraise bodily or emotional changes as significant. Concerning physical and somatoform symptoms, higher symptom reporting in women is described as a general phenomenon rather than one restricted to certain types of symptoms. This effect persists even after adjusting for psychiatric comorbidity such as depression and anxiety, disorders that are more prevalent in women and associated with increased reports of physical symptoms. The suggestion that women are more perceptive about emotions is supported by the finding that women outperform men at all ages in the encoding and decoding of emotions.

The differences in coping styles between men and women are small. Whereas women appear to prefer emotion-focused solutions including the greater use of social support and supportive networks, men focus more often on instrumental or problem-oriented coping. Reports on sex-specific effectiveness of coping strategies contradict the former notion that women's coping strategies could be inadequate or less effective for alleviating distress.

Investigations of perceived stressfulness after confrontation with various standardized laboratory

stressors repeatedly led to similar reports of feelings of distress in men and women. Whereas performance in laboratory stress tests was not associated with sex, women tend to judge their own efforts and performances as poorer than men do.

The assessment of subjective aspects of the human stress response is always based on self-reports measured by standardized questionnaires, interviews (structured or half-structured), and open (unstructured) reports. Therefore, it should be considered that sex-related response differences to stress could also reflect sex-specific biases in the willingness (or capability) to report distress.

Sex Differences in Physiological Aspects of the Human Stress Response

Different regulatory systems of the human body are activated by stress. Endocrine stress-responsive systems such as the HPA and SAM axes help the organism adapt to increased demands and maintain homeostasis after challenge. In addition to hormonal responses, stress also provokes changes in cardiovascular and immunological parameters.

Hypothalamic-Pituitary-Adrenal Axis

Investigation of HPA axis responses to different types of stressors revealed that physical exercise ($>70\%$ $\text{VO}_{2\text{max}}$) and psychological challenges are potent activators of the HPA axis. Whereas the majority of studies did not reveal any sex differences in response to physical exercise, confrontation with psychosocial stimuli leads to higher adrenocorticotrophic hormone (ACTH) and free cortisol levels in adult men than in women. This sex difference is conserved after menopause and in elderly individuals. Total cortisol secretion, however, does not differ between sexes in young adults. Although men and women show similar total cortisol responses after stress, a sexual dimorphism exists in the amount of bioavailable free cortisol, an observation calling for simultaneous measurements of total and free cortisol levels. These findings can, at least in part, be explained by greater ACTH pulses in men as well as by the higher sensitivity to ACTH of the adrenal cortex in women.

Observed sex differences in HPA axis stress responses may be due to sexual dimorphisms in brain functions and circulating sex steroid and corticosteroid-binding globulin (CGB) levels. Prime candidates for explaining sex differences in HPA axis stress responsiveness are the gonadal steroids. Among them, estradiol seems to exert modulating effects on HPA axis functioning, including HPA axis responsiveness and sensitivity to glucocorticoid negative feedback.

Animal studies consistently revealed that estradiol has a strong stimulatory influence on HPA axis functioning with modulatory effects on mineralocorticoid and glucocorticoid receptors.

Although human data are more contradictory than evidence from animal literature, gonadal steroids also seem to exert important modulatory effects on HPA axis functioning in humans. In a series of studies from our laboratory, free cortisol responses systematically varied as a function of menstrual-cycle phase. Women in the luteal phase show free cortisol stress responses comparable to men, whereas women in the follicular phase and women taking oral contraceptives have significantly lower free cortisol responses. Although the total plasma cortisol net stress responses differ neither between men and women nor between different menstrual-cycle phases, the absolute levels are elevated in women with oral contraceptive use. A short-term estradiol treatment potentiated HPA axis responses to acute psychosocial stress in young men. Furthermore, a 2-week estradiol substitution seemed to diminish the endocrine response of postmenopausal women to exogenous corticotropin releasing hormone (CRH) after dexamethasone premedication, as observed in women with and without short-term estradiol treatment. In addition, a comparison with a third group of premenopausal women indicates that the feedback sensitivity of elderly estradiol-treated women approximates the endocrine response of premenopausal (untreated) women. However, this treatment did not change the acute HPA axis stress responses to psychosocial stress (Trier Social Stress Test). In sum, it should be acknowledged that there is still a paucity of data on the effects of gonadal steroids on the regulation of the human HPA (and SAM) axis. There is a definite need for future in-depth studies on these issues.

Sympathetic-Adrenal-Medullary Axis

Studies investigating SAM axis responses to stress showed that stress-induced epinephrine increases are higher in men than in women, suggesting that men are more sensitive to sympathetic stimulation via emotional arousal. In contrast, no sex effects are reported for norepinephrine concentrations. However, contradictory results also exist.

Reproductive hormones appear to have a significant impact on catecholamine stress responses. This assumption is confirmed first by findings of altered catecholamine levels throughout the menstrual cycle with enhanced epinephrine levels during the luteal phase. Second, comparisons between pre- and postmenopausal women revealed that stress-related epinephrine responses are increased in postmenopausal women

and that estradiol treatment reduced catecholamine reactivity to stress. Third, short-time estradiol administration modulates catecholamine reactivity in healthy young men, although the direction of the effect remains contradictory. Although the majority of studies point to an estradiol-mediated reduction in SAM axis responses to laboratory stressors, a few carefully conducted studies show sex steroids to have no or a stimulating effect.

Cardiovascular Responses

In a meta-analysis, Stoney and coworkers summarized sex differences in stress-related heart-rate and blood pressure changes. They concluded that women have higher heart rates at rest as well as slightly larger increases during challenge, whereas men exhibit higher systolic blood pressure changes. Allen and colleagues also showed higher heart-rate increases in women and higher blood pressure changes in men after confrontation with various laboratory stress tasks, supporting the conclusion that women are more likely to be cardiac reactors and men are more prone to be vascular reactors. We aggregated heart-rate responses across five different studies in healthy children and in younger and elderly adults. Accordingly, enhanced heart-rate stress reactivity can be observed in girls and young adult females compared to boys and young adult men. However, in older men and women peak heart-rate responses are comparable, with only men returning to prestressor baselines during recovery.

Although the data are inconsistent, reports of blood pressure changes as a function of menstrual-cycle phase point to sex steroids having an impact on cardiovascular functioning. The fact that coronary heart disease morbidity and mortality rates increase in women after menopause, a time characterized by the marked decrease of reproductive hormones, confirms the protective role of sex steroids. In line with this idea, reports show that postmenopausal women have a higher cardiovascular responsiveness than premenopausal women.

Whereas postmenopausal estrogen use predominantly reduces the risk of coronary heart disease in women, exogenous estrogen use in clinical trials or hormonal replacement therapy studies do not consistently confirm the protective role of estrogens. The apparently conflicting results of exogenous estrogen use on cardiovascular functioning could be explained by a dual action of estrogens. Estrogens may increase the risk of thrombo-embolic events due to an adverse influence on coagulation, and simultaneously they may help to prevent atherogenic events through favorably changed lipid and lipoprotein levels.

Immunological Responses

In general, immunological activity appears to be higher in women than in men. Whereas levels of circulating immunoglobulins are higher and antibody responses to a variety of pathogens are larger in females, sex differences in cell-mediated immunity can vary depending on the immunological parameter studied. Sex differences in basal immunological functioning may result from the stimulation of lymphoid tissues by sex steroids. In general, estrogens appear to suppress T-cell-dependent immune function, but can increase antibody reactivity to antigens. Testosterone (and other androgens) has an immunosuppressive effect and increases susceptibility to infection, although it may help to prevent the development of autoimmune processes. In sum, sex hormones have been shown to modulate a large variety of mechanisms involved in the immune response. Evidence is accumulating that shows that estrogenic and androgenic actions on cells can be mediated not only through classical intracellular receptors but also through membrane receptors on cell surfaces of compounds of the immune system. Wunderlich and coworkers investigated possible cross-talk of nongenomic testosterone signaling with other genotropic signaling pathways in the rat. Furthermore, in addition to direct receptor-mediated influences, indirect effects on immunological functioning and disease processes may be exerted by sex-steroid-induced changes in other biological systems that are interrelated with the regulation of the immune system. For example, the androgenic steroid hormone dehydroepiandrosterone sulfate (DHEAS) was recently shown to inhibit apoptosis in human peripheral blood lymphocytes through a mechanism independent of either androgen or estrogen receptors.

Although there is a voluminous literature showing the impact of acute laboratory stress on rapid immune cell changes, sexual dimorphisms in immune responses to stress in humans have not been extensively examined and the reported results remain contradictory. Whereas Kiecolt-Glaser and colleagues found that women are more likely to elicit immunological changes after controversial discussions with their husbands, as reflected by a summary immune change score (based on killer cell lysis, blastogenic responses to two mitogens, the proliferative response to a monoclonal antibody to the T₃ receptor, percentage of macrophages, and number of T lymphocytes), Scanlan and coworkers concluded that immunological parameters appear to be more affected by psychosocial stressors (hassles and caregiving) in men than in women. Recently, Pehlivanoglu and colleagues observed a shift from cellular to humoral immunity (only) in men and showed evidence that acute stress

may cause different immunological changes across the female menstrual cycle. In addition to these sexual dimorphisms, the use of oral contraceptives (OC) can also alter stress-induced immune responses in women. Furthermore, sex and OC intake may exert differential effects on the immune system by modulating glucocorticoid sensitivity of pro-inflammatory cytokine production.

In sum, immune functions at rest and under stress differ significantly between men and women, apparently due to sex-related differences in sex steroid levels.

See Also the Following Articles

Adrenal Medulla; Autoimmunity; Coping Skills; Depression Models; Economic Factors and Stress; Gender and Stress; Heart Disease/Attack; Immune Response; Menopause and Stress; Social Status and Stress; Social Support.

Further Reading

- Allen, M. T., Stoney, C. M., Owens, J. F., et al. (1993). Hemodynamic adjustments to laboratory stress: the influence of gender and personality. *Psychosomatic Medicine* 55, 505–517.
- Cleary, P. D. (1987). Gender differences in stress-related disorders. In: Barnett, R. S., Biener, L. & Baruch, G. K. (eds.) *Gender and stress*, pp. 39–72. New York: Free Press.
- Kiecolt-Glaser, J. K., Malarkey, W. B., Chee, M., et al. (1993). Negative behavior during marital conflict is associated with immunological down-regulation. *Psychosomatic Medicine* 55, 395–409.
- Kirschbaum, C., Schommer, N., Federenko, I., et al. (1996). Short term estradiol treatment enhances pituitary-adrenal axis and sympathetic responses to psychosocial stress in healthy young men. *Journal of Clinical Endocrinology and Metabolism* 81, 3639–3643.
- Kroenke, K. and Spitzer, R. L. (1998). Gender differences in the reporting of physical and somatoform symptoms. *Psychosomatic Medicine* 60, 150–155.
- Kudielka, B. M., Buske-Kirschbaum, A., Hellhammer, D. H., et al. (2004). Differential heart rate reactivity and recovery after psychosocial stress (TSST) in healthy children, younger adults, and elderly adults: the impact of age and gender. *International Journal of Behavioral Medicine* 11, 116–121.
- Kudielka, B. M., Buske-Kirschbaum, A., Hellhammer, D. H., et al. (2004). HPA axis responses to laboratory psychosocial stress in healthy elderly adults, younger adults, and children: impact of age and gender. *Psychoneuroendocrinology* 29, 83–98.
- Kudielka, B. M. and Kirschbaum, C. (2005). Sex differences in HPA axis responses to stress: a review. *Biological Psychology* 69, 113–132.
- Kudielka, B. M., Schmidt-Reinwald, A. K., Hellhammer, D. H., et al. (1999). Psychological and endocrine responses to psychosocial stress and dexamethasone/corticotropin-releasing hormone in healthy postmenopausal women and young controls: the impact of age and a two-week estradiol treatment. *Neuroendocrinology* 70, 422–430.
- Lazarus, R. S. and Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.
- Lutzky, S. M. and Knight, B. G. (1994). Explaining gender differences in caregiver distress: the roles of emotional attentiveness and coping styles. *Psychology and Aging* 9, 513–519.
- McCrudden, A. B. and Stimson, W. H. (1991). Sex hormones and immune function. In: Ader, R., Felten, D. L. & Cohen, N. (eds.) *Psychoneuroimmunology*, pp. 475–793. San Diego, CA: Academic Press.
- Miller, S. M. and Kirsch, N. (1987). Sex differences in cognitive coping with stress. In: Barnett, R. S., Biener, L. & Baruch, G. K. (eds.) *Gender and stress*, pp. 278–307. New York: Free Press.
- Pehlivanoglu, B., Balkanci, Z. D., Ridvanagaoglu, A. Y., et al. (2001). Impact of stress, gender and menstrual cycle on immune system: possible role of nitric oxide. *Archives of Physiology and Biochemistry* 109, 383–387.
- Polefrone, J. M. and Manuck, S. B. (1987). Gender differences in cardiovascular and neuroendocrine response to stress. In: Barnett, R. S., Biener, L. & Baruch, G. K. (eds.) *Gender and stress*, pp. 13–37. New York: Free Press.
- Raison, C. L. and Miller, A. H. (2003). When not enough is too much: the role of insufficient glucocorticoid signaling in the pathophysiology of stress-related disorders. *American Journal of Psychiatry* 160, 1554–1565.
- Rohleder, N., Wolf, J. M. and Kirschbaum, C. (2003). Glucocorticoid sensitivity in humans – interindividual differences and acute stress effects. *Stress* 6, 207–222.
- Scanlan, J. M., Vitaliano, P. P., Ochs, H., et al. (1998). CD4 and CD8 counts are associated with interactions of gender and psychosocial stress. *Psychosomatic Medicine* 60, 644–653.
- Schulz, P., Schlotz, W., Wolf, J., et al. (2002). [Gender differences in stress-related variables: the influence of worry-disposition]. *Zeitschrift für Differentielle und Diagnostische Psychologie* 23, 305–326.
- Stoney, C. M., Davis, M. C. and Matthews, K. A. (1987). Sex differences in physiological responses to stress and in coronary heart disease: a causal link? *Psychophysiology* 24, 127–131.
- Takahashi, H., Nakajima, A. and Sekihara, H. (2004). Dehydroepiandrosterone (DHEA) and its sulfate (DHEAS) inhibit the apoptosis in human peripheral blood lymphocytes. *Journal of Steroid Biochemistry and Molecular Biology* 88, 261–264.
- Whitacre, C. C. (2001). Sex differences in autoimmune disease. *Nature Immunology* 2, 777–780.
- Wunderlich, F., Benten, W. P., Lieberherr, M., et al. (2002). Testosterone signaling in T cells and macrophages. *Steroids* 67, 535–538.
- Young, E. A. (1995). The role of gonadal steroids in hypothalamic-pituitary-adrenal axis regulation. *Critical Reviews in Neurobiology* 9, 371–381.

Sex Steroids, Response to Stress and Susceptibility to Depression

M V Seeman and L E Ross

University of Toronto, Toronto, Ontario, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M V Seeman, volume 3, pp 429–434, © 2000, Elsevier Inc.

Prolactin

A hormone secreted by the anterior pituitary gland whose function, among others, is to stimulate the growth of milk-secreting glands in the breast in preparation for, or maintenance of, lactation.

Psychophysiological disorders

Disorders such as chronic fatigue syndrome, irritable bowel syndrome, and fibromyalgia, which are stress-related body symptoms

Introduction

Gender Differences in Depression and Stress Syndromes
Sex Steroids
Stress
Conclusion

Glossary

<i>Circadian rhythms</i>	Twenty-four-hour sleep–wake cycles.
<i>Estrous cycles</i>	Cyclic sexual receptivity in female animals that is accompanied by specific hormonal periodicity and that lays the groundwork for impregnation.
<i>Genomic effects</i>	Direct effects on DNA in cell nuclei.
<i>Growth hormone</i>	A polypeptide hormone secreted by the anterior lobe of the pituitary gland that promotes tissue and skeletal growth and influences the metabolism of proteins, carbohydrates, and lipids.
<i>Hypothalamic-pituitary-adrenal axis</i>	The hormonal circuit from the hypothalamus to the anterior pituitary gland to the adrenal gland and back to the hypothalamus.
<i>Hypothalamic-pituitary-gonadal axis</i>	The hormonal circuit from the hypothalamus to the anterior pituitary gland to the ovaries or testes and back to the hypothalamus.
<i>Hypothalamic-pituitary-thyroid axis</i>	The hormonal circuit from the hypothalamus to the anterior pituitary gland to the thyroid gland and back to the hypothalamus.
<i>Melatonin</i>	A hormone that is secreted by the pineal gland and that regulates the onset and timing of sleep.
<i>Neurotransmitters</i>	Chemical messengers that cross the gap between neurons to either excite or inhibit the neighboring neuron.
<i>Posttraumatic stress syndrome</i>	A set of psychiatric symptoms that develop as a reaction to a traumatic event.

Introduction

From a physiological point of view, both depression and anxiety are the product of complex alterations of the hypothalamic-pituitary-thyroid axis, of prolactin, growth hormone, and melatonin irregularity, of sleep, activity, appetite, and temperature disturbance, of hypothalamic-pituitary-gonadal axis and hypothalamic-pituitary-adrenal axis deregulation, and of multiple neurotransmitter disequilibrium. Given the gender ratio of the prevalence of depression and anxiety and the generally accepted role that stressful life events and powerful emotions play in the precipitation of mood disorders, this article focuses on the effects of ovarian hormones on the body's response to stress.

Gender Differences in Depression and Stress Syndromes

Women are more susceptible than men to all forms of depression. Even in bipolar affective (manic-depressive) illness, in which the gender prevalence is equal, the frequency of depressive episodes is higher in women. Chronic, subthreshold depression (dysthymia) and episodic major depressive disorder are at least twice as common in women as they are in men. These sex differences have been replicated in all major door-to-door epidemiological surveys and in treatment-seeking population samples from around the world, with the exception of reports from countries where women are known to have reduced access to medical services. Strikingly, this sex ratio does not emerge until adolescence, when gonadal hormones are activated. Before puberty, depressive episodes are as frequent in boys as they are in girls.

Anxiety may be a symptom of a depressive episode or it may constitute the major component of a variety

of anxiety syndromes such as posttraumatic stress disorder (PTSD), phobia, panic disorder, agoraphobia, obsessive-compulsive disorder, and generalized anxiety. With the exception of obsessive-compulsive disorder, in which the gender ratio is equal, adult women are more susceptible than men to all the preceding *Diagnostic and Statistical Manual of Mental Disorders* (DSM)-defined syndromes. As in depression, the sex differences in most forms of anxiety emerge only after adolescence.

The aftermath of stressful life events may be experienced in the psychological realm, as in PTSD, or primarily in the somatic realm, as in the various psychophysiological disorders. Both forms are more prevalent in women. When exposed to a major trauma, approximately twice as many women as men develop PTSD, and women make up about 85% of PTSD sufferers whose symptoms last for over 1 year after exposure. After the World Trade Center attack, gender differences in PTSD could be explained by differences in marital status (more common among women living alone without a permanent relationship) and education (more common among women with relatively little education and low income), but significantly more women experienced major depression and significantly more men were diagnosed with a substance abuse disorder. These differences persisted after adjustment for demographic variables.

Both depression and anxiety are more severe and more frequent during the premenstrual, postpartum, and menopausal periods of women's lives. An example of experimentally elicited menstrual phase vulnerability is the Van Goozen et al. report of a study involving 58 healthy women. Deliberately arranged anger provocation evoked most emotionality in those women who were premenstrual at the time of the experiment. Bloch et al. used an analogous design to investigate susceptibility to depression in response to exogenous estrogen and progesterone among women with a history of postpartum depression. Five of eight women with a history of postpartum depression, and none of the control women, with no psychiatric history, developed significant mood symptoms during hypogonadism induced via administration of a gonadotropin releasing-hormone agonist.

In summary, depressive and anxious states, as well as stress-provoked somatic conditions, are two to three times as common in women as in men, but there is no consistent sex difference in their prevalence prior to puberty. In women, symptom severity and frequency are associated with times in the reproductive cycle when the circulating level of ovarian hormones (estrogens and progestins) are falling.

Sex Steroids

Male-Female Differences

Gender differences in illness prevalence can result from a combination of inborn, learned, and socially imposed factors, but what distinguishes the sexes during fetal development are sex-specific hormones that exert powerful effects on the fetal brain and contribute to its sexual differentiation. While the hormonal milieu does not differ appreciably for boys and girls between birth and puberty, their brains are sexually differentiated from early gestation. The dramatic hormonal changes at puberty, superimposed on a dozen or more years of differential sex role experience, exert a profound effect on the adolescent brain. In addition to median levels of sex steroids, what also differs after puberty (and this may turn out to be the crucial difference) is that only women undergo menstrual cycles. Both sexes are subject to pulsatile changes in hormone level and circadian fluctuations, but the monthly repetitive hormonal cycles interrupted by pregnancy are unique to women. It is not until very old age, when the testes stop producing testosterone, that the hormonal environment of the male brain and that of the female brain once again approach the par that existed prior to puberty. Even then, receptor densities for the various hormone agonists and their site specificity remain gender distinct.

These biological differences may play out in the context of differential exposure to social and/or psychological stress. It has been proposed that women may experience a greater frequency of stressful life events than men. Alternatively, women may be more likely than men to interpret life events as stressful. However, a recent twin study by Kendler et al. revealed that men and women were equally likely to experience an episode of major depression associated with a major life event, although the categories of significant events differed between the sexes (i.e., interpersonal events were more relevant to women; work problems were more relevant to men).

In summary, although exposure to the impact of different experiences modulates the internal cascade of neurochemical events in which sex hormones play such a crucial part, neuroendocrine differences and their sex-specific rhythms are arguably central to male-female differences in the vulnerability to depression and anxiety. Hormones determine how the male and female brain processes the experience to which it is exposed, in particular, how it differentially responds to stressful experience and powerful emotions.

Sex Steroids and Neurotransmitter Systems

Sex steroids operate through genomic effects on neurotransmitter receptor protein or through direct membrane effects on the electrical firing rate of neurotransmitter-secreting cells. The genomic effect is relatively slow (minutes to hours); the membrane effect is rapid (microseconds to minutes).

It is well recognized that the serotonergic system is important in the maintenance of mood. The effectiveness of selective serotonin reuptake inhibitors (SSRIs) in the treatment of both depression and anxiety has highlighted the serotonin system. Estrogens exert multiple effects on serotonin function, many of which are site specific. In the amygdala, a key brain structure involved in the mediation of environmental stress, estrogens increase serotonin turnover. Estrogens increase the density of serotonin 5-hydroxytryptamine 2A (5HT_{2A}) receptors in rat cerebral cortex and nucleus accumbens. They decrease 5HT_{1A} mRNA in the medial amygdala, piriform cortex, and perirhinal cortex, suggesting a possible pathway through which mood and memory are modulated. Throughout the brain, estrogens increase the rate of degradation of monoamine oxidase, which, in turn, leads to increased brain availability of both serotonin and catecholamines. The interaction between estrogen and the serotonin system may have clinical implications: there is recent evidence that SSRIs are effective in reducing vasomotor symptoms associated with estrogen decline during menopause, and estrogen therapies may hold promise in management of hormone-dependent psychiatric conditions, including postpartum depression. Finally, there is some evidence that concurrent estrogen therapy may augment or accelerate response to SSRIs in women.

Estrogen-responsive and progesterone-responsive neurons use the inhibitory γ -aminobutyric acid (GABA) as a transmitter, producing an anxiolytic affect. Estrogen also induces choline acetyltransferase and potentiates excitatory amino acids such as glutamate and aspartate. The estrogen effect on the dopaminergic system is complex and phasic, causing alterations in pre- and postsynaptic dopamine receptors and differential distribution of the isoforms of the dopamine D₂ receptor messenger RNA. More specific to stress and depression, estrogens block both the synthesis and the depletion of catecholamines, including norepinephrine. Norepinephrine plays a major role in the autonomic stress response. Estrogens work via endogenous opiate peptides to exert an activating influence on the autonomic sympathetic response to stress, which results in vasoconstriction, quickened heartbeat, and sweating, common symptoms of anxiety.

In summary, ovarian hormones exert multiple, complex effects on neurotransmitter systems understood to be involved in depression and anxiety syndromes.

Effects on Membrane Stability

Because of their mainly activational influences, estrogens are proconvulsants in susceptible individuals, decreasing seizure threshold and increasing the magnitude of seizure activity. In experimental animals, this appears to be a direct consequence of depolarization of cellular membranes. Progesterone, on the other hand, raises the seizure threshold, an effect opposite to that of estrogen and probably responsible for catamenial or premenstrual epilepsy, when progesterone levels drop. Seizure activity (e.g., electric shock treatment) has been known for a very long time to ameliorate depressive mood.

In summary, the net effect of progesterone is to reduce anxiety but to deepen depression. The opposite is true for estrogen: levels of circulating estrogens correlate with increased alertness in animals, reduced distractibility, and shorter reaction times. The two hormones maintain a balance or homeostasis that, when disturbed, can tip a vulnerable person in the direction of mental states that are diagnosed as depression or anxiety.

Other Variables

Estrogens increase blood flow to the brain and within the brain, which speeds metabolism and increases the general efficiency of neural networks and increases the effectiveness of drug delivery. Estrogens stimulate glucose transport and metabolism, promoting the ability of neuronal networks to perform their work.

Neuronal responsiveness to steroid stimulation results from both genetic and membrane effects at multiply distributed sites involving large neural circuitry. Network effects may differ depending on the particular configuration of a specific circuit. Hormonal effects on neural network properties appear to be state dependent. In other words, effects may differ widely depending on the light-dark cycle, the sleep cycle, the activity cycle, the glucose level, the noise level, the presence or absence of an animal of the opposite sex, the general stress level, and the specific type of stressor.

Genomic Effects

Among the many genes suspected to come under the influence of estrogens are those responsible for neuronal development, migration, growth, and survival, for dendritic spine formation and density, and for antioxidant activity, synaptic plasticity,

neurodegeneration, and apoptosis secondary to glucose deprivation or other toxic event.

Steroid hormones are important epigenetic regulators. In a tissue-specific manner, they control the dosage effects of genes. They can turn genes off and on. They bind to their own receptors, a family of DNA-binding proteins located within the cell nucleus. The human estrogen receptor (ER- α) has been characterized both at the cDNA level and at the genomic level. It is transcribed from three different promoter regions. The expression of these three ER- α mRNAs appears to be regulated in a tissue-specific and, perhaps, age-specific way.

ER- α is sparsely distributed in key cognitive regions. This made it difficult to understand why estrogen replacement therapy sometimes improved cognitive functions until the recent discovery of a new estrogen receptor (ER- β), localized in brain regions associated with learning, memory, and emotion.

Once estrogen enters the cell nucleus and binds to the ER expressed in that cell, a conformational change takes place. The receptor's DNA binding domain is exposed and attaches to specific sequences of nucleotides in the nuclear DNA, thus regulating the transcription of a target gene. Adapter proteins or other transcription factors, which are cell specific, modulate the strength of the activation of the gene. Genetic transcription is one way in which estrogens modify the structure and function of brain cells.

Membrane Receptors

Estrogen also produces nongenomic effects by binding to neuronal membrane receptors. This effect is rapid, with a latency of seconds to minutes, and disappears quickly. Estrogen effects on dendritic spines in the CA1 region of the hippocampus are a good example of a rapid response. Estrogen given to an ovariectomized rat produces an increase in the number of spines on apical dendrites of CA1 pyramidal neurons. Cyclic changes in spine density can be seen throughout the estrus cycle of rats, suggesting the formation and destruction of synapses in quick time, presumably, therefore, mediated via membrane receptors. One mechanism through which dendritic spine density is increased is neurotrophin brain-derived neurotrophic factor (BDNF). Estradiol downregulates BDNF in cultured hippocampal neurons within 24 h of exposure. This increases excitatory tone in pyramidal neurons, which may lead to as much as a twofold increase in dendritic spine density.

In summary, estrogens, by a variety of means, help to increase neuronal connections. If nerve impairment occurs, they foster repair and delay or prevent secondary atrophy.

Stress

The Stress Response

The human organism has evolved an elaborate mechanism for dealing with perceived threat. Cortisol, epinephrine, and norepinephrine are automatically released, mobilizing the necessary energy for fight or flight. This primarily involves the limbic-hypothalamic-pituitary-adrenal-gonadal-thyroid axis. The integration of external stimuli converges on the medial parvocellular division of the paraventricular nucleus of the hypothalamus (mpPVN), which secretes corticotropin-releasing hormone (CRH). This hormone binds to its receptors in the anterior pituitary, stimulating adrenocorticotrophic hormone (ACTH)-secreting cells. ACTH is released into the circulation and, in the adrenal cortex, activates the synthesis and release of glucocorticoids whose many actions are mediated via specialized receptors affecting and regulating genes throughout the body. At the same time, stress leads to the modulation of the secretion of thyroid-releasing hormone and gonadal-releasing hormone from the hypothalamus with subsequent downstream effects on the thyroid and the gonads. Appetite and the need for sleep are suppressed. Growth stops. The overall result is that the organism is prepared to meet the perceived threat.

A large part of the brain is devoted to coping with stress. The amygdala is also placed on alert and activates the hypothalamic-pituitary-adrenal (HPA) axis. The prefrontal cortex is involved in assessing the danger and putting a halt to the stress response when the danger is past.

Men and women show differential HPA axis responses to psychosocial stress, with women of reproductive age showing less pronounced HPA axis responses to stress than men of the same age. However, this response varies with menstrual cycle phase and menopausal status, suggesting interactions between the HPA axis and gonadal hormones.

Termination of the Stress Response

The level and duration of glucocorticoid activity must be adequate to the organism's need to deal effectively with the perceived stressor. Glucocorticoids (helped along by excitatory amino acids) are, however, potentially harmful and may cause cell death through the promotion of apoptosis. A mechanism needs to be in place for their rapid elimination. Otherwise, what is adaptive becomes, under conditions of severe or ongoing stress, an agent of pathological change. Hippocampal neurons, particularly rich in steroid receptors, are exquisitely vulnerable to glucocorticoid-induced atrophy.

Many negative feedback loops exist within the stress activation system, which effectively puts a stop to the glucocorticoid stress response once it is no longer needed. Besides inhibition of the amygdala by the thinking brain, the prefrontal cortex, mounting concentrations of cortisol in the brain automatically shut off production of CRH. Sex steroids do the same. The rapidity of both stress activation and stress termination is affected by the circadian rhythm in plasma glucocorticoid levels, being most responsive when steroid levels are at their lowest, in the late evening in diurnal animals.

Stress Response Termination in Depression

Efficient termination of the stress response may be impaired in depression. For instance, the left prefrontal cortex has been shown in functional imaging studies to be impaired in major depression. As mentioned previously, depression and stress have long been associated. Adverse life events trigger depressive episodes, and depression itself constitutes an ongoing stressor to the organism. The experience of depression is one that precludes pleasure and hope, prevents sleep, and diminishes appetite, metabolic activity, autonomic function, neuroendocrine processes, growth, and reproduction.

There has been a long tradition of cortisol measurement in depressive states, culminating in the mid-1970s with the introduction of the Dexamethasone Suppression Test (DST) as a diagnostic test. The assumption was that suppression would be deficient in depression, and this has turned out to be true in at least some forms of depression.

This fact may explain some of the cognitive impairment that accompanies depression when glucocorticoid levels remain high and result in atrophy of hippocampal cells. Gonadal steroid receptors, with variations from one time period to another, are also expressed in the hippocampus. In animal experiments, androgens have been shown to downregulate glucocorticoid receptors in the hippocampus. Estrogens have also been reported to protect against glucocorticoid-induced hippocampal atrophy. Although both androgens and estrogens are potentially protective, both female rats and female monkeys undergoing various forms of experimentally induced stress fare better than their male counterparts with respect to secondary hippocampal damage.

In summary, a sustained stress response produces hippocampal degeneration and atrophy of apical dendrites of hippocampal pyramidal neurons. The interaction among glucocorticoids, neuronal growth factors, neurotransmitters, and sex steroids is complex and still in the process of clarification. Acute responses to stress, when prolonged, cause damage.

Conclusion

It would appear from animal work that women might have an advantage over men with respect to neuronal protection in the face of stress. Why, then, are they more prone than men to depression and anxiety?

The integrity of circadian rhythms and sleep-wake cycles seems to play a part in depression. People with depression feel worse in the morning than in the evening. They cannot sleep. They wake early. Resynchronizing the sleep cycle is an effective treatment for depression. The stress response cascade is modulated by circadian rhythm, and estrogens appear to be important circadian pacemakers. Is it possible that the monthly periodicity of ovarian hormones destabilizes these regular daily rhythms and thus contributes to women's special vulnerability to depression? The monthly cycling may sensitize target receptors and may in fact kindle these mechanisms in such a way that high circulating levels of ovarian hormones are needed to maintain equilibrium. This means that vulnerability to stress-related episodic depression and anxiety is heightened during periods of relative hormone decline.

That being said, it is also important to keep in mind that the physiological stress response cascade and its sequelae may be triggered by emotions other than fear of perceived threat. It may be precipitated by strong feelings of many sorts, such as frustration, empathy, or guilt, for instance, and these feelings may be relatively ubiquitous in women, prone as women are to relatively close interpersonal attachments and larger social networks when compared to men. Nurturant proclivities may be mediated by prolactin, produced in the anterior pituitary. The activities of this hormone (more active in women) may help to explain the sex ratio of adult depressive disorders.

In summary, women are more prone than men to suffer from anxiety and depression in their many forms. The mechanisms involved may reflect a tendency to hyperarousal or prolonged arousal when responding to a stressor. The sex differences may be due to relative inefficiency in turning off the stress signal, perhaps resulting from monthly kindling by fluctuations in ovarian hormones or, alternatively, from a relatively wider range of stimuli eliciting the stress response in women.

See Also the Following Articles

Anxiety; Depression and Manic-Depressive Illness; Estrogen; Gender and Stress; Sex Differences in Human Stress Response.

Further Reading

- Bloch, M., Schmidt, P. J., Danaceau, M., et al. (2000). Effects of gonadal steroids in women with a history of postpartum depression. *American Journal of Psychiatry* 157, 924–930.
- Bora, S. H., Liu, Z., Kecojevic, A., Merchenthaler, I. and Koliatsos, V. E. (2005). Direct, complex effects of estrogens on basal forebrain cholinergic neurons. *Experimental Neurology* 194, 506–522.
- Dalla, C., Antoniou, K., Drossopoulou, G., et al. (2005). Chronic mild stress impact: are females more vulnerable? *Neuroscience* 135, 703–714.
- Davydov, D. M., Shapiro, D., Goldstein, I. B. and Chic-DeMet, A. (2005). Moods in everyday situations: effects of menstrual cycle, work, and stress hormones. *Journal of Psychosomatic Research* 58, 343–349.
- Fitch, R. H. and Denenberg, V. H. (1998). A role for ovarian hormones in sexual differentiation of the brain. *Behavioral and Brain Sciences* 21, 311–352.
- Holbrook, T. L., Hoyt, D. B., Stein, M. B. and Sieber, W. J. (2002). Gender differences in long-term posttraumatic stress disorder outcomes after major trauma: women are at higher risk of adverse outcomes than men. *Journal of Trauma* 53, 882–888.
- Huttner, R. P. and Shepherd, J. E. (2003). Gonadal steroid, selective serotonin reuptake inhibitors, and mood disorders in women. *Medical Clinics of North America* 87, 1065–1076.
- Kajantie, E. and Phillips, D. I. W. (2005). The effects of sex and hormonal status on the physiological response to acute psychosocial stress. *Psychoneuroendocrinology* 31, 151–178.
- Kendler, K. S., Thornton, L. M. and Prescott, C. A. (2001). Gender differences in the rates of exposure to stressful life events and sensitivity to their depressogenic effects. *American Journal of Psychiatry* 158, 587–593.
- McEwen, B. S. and Magarinos, A. M. (2001). Stress and hippocampal plasticity: implications for the pathophysiology of affective disorders. *Human Psychopharmacology* 16(S1), S7–S19.
- Merchenthaler, I., Lane, M. V., Numan, S. and Dellovade, T. L. (2004). Distribution of estrogen receptor alpha and beta in the mouse central nervous system: *in vivo* autoradiographic and immunocytochemical analyses. *Journal of Comparative Neurology* 473, 270–291.
- Razandi, M., Pedram, A., Merchenthaler, I., Greene, G. L. and Levin, E. R. (2004). Plasma membrane estrogen receptors exist and function as dimers. *Molecular Endocrinology* 18, 2854–2865.
- Steiner, M., Dunn, E. and Born, L. (2003). Hormones and mood: from menarche to menopause and beyond. *Journal of Affective Disorders* 74, 67–83.
- VanGoozen, S. H., Frijda, N. H., Wiegant, V. M., Endert, E. and Van de Poll, N. E. (1996). The premenstrual phase and reactions to aversive events: a study of hormonal influences on emotionality. *Psychoneuroendocrinology* 21, 479–497.
- Weissman, M. M., Neria, Y., Das, A., et al. (2005). Gender differences in posttraumatic stress disorder among primary care patients after the World Trade Center attack of September 11, 2001. *Gender Medicine* 2, 76–87.

Sex-Specific Effects of Early Social Stress in Mammals: A Study in Guinea Pigs

S Kaiser and N Sachser

University of Muenster, Münster, Germany

© 2007 Elsevier Inc. All rights reserved.

Prenatal Stress
Nonsocial versus Social Stressors
A Study in Guinea Pigs
Comparison with Other Nonhuman Mammals
Neuroendocrine Mechanisms of Prenatal
Social Stress
Adaptive Function of Prenatal Social Stress

Glossary

<i>Demasculinization</i>	Loss or attenuation of male-typical traits.
<i>Feminization</i>	Enhancement of female-typical traits.
<i>Maternal effects</i>	Mothers' control over offspring development.
<i>Prenatal stress</i>	When stressors acting on pregnant females result in significant consequences for the offspring during fetal development and/or later in life.
<i>Social environment</i>	Part of an individual's environment that is constituted by other members of the same species.
<i>Social stress</i>	Stress due to members of the same species.

Prenatal Stress

Internal and external stressors constantly challenge organisms. Animals and humans respond to environmental challenge by activating a variety of coping systems (e.g., the hypothalamic-pituitary-adrenocortical [HPA] and the sympathetico-adrenomedullary [SAM] system). These systems are activated to maintain the internal milieu within a normal physiological range (i.e., homeostasis). This is crucial for normal pre- and postnatal development. The activation of the HPA axis results in an elevated secretion of glucocorticoids (e.g., cortisol, corticosterone), while the activation of the SAM axis causes an increase in serum catecholamine concentrations (e.g., epinephrine, norepinephrine). If stressors act on a pregnant female, concentrations of maternal hormones will change. Because the internal environment of the pregnant animal constitutes much of the embryo's external environment, it is reasonable to suggest that the former may affect fetal development. It is well-known that stressors acting on the organism during the pre- and/or perinatal phase can have distinct and long-term effects on behavior, reproductive functions, and the immune as well as the neuroendocrine and autonomic systems. For example, it has been reported numerous times that stress applied to pregnant rats results in demasculinization (i.e., lower levels of copulatory behavior) and feminization (i.e., higher levels of lordosis behavior, a female-typical sexual behavior) of male offspring compared to sons of nonstressed mothers.

Nonsocial versus Social Stressors

In most experimental studies on the effects of early stress, pregnant nonhuman mammals were subjected to nonsocial stressors (e.g., regimens of heat, bright light, or restraint). However, most of these stressors do not exist in the natural world of the animals and so it is questionable whether they represent ecologically relevant stimuli.

In their natural habitat, animals have to cope with a variety of social stressors that are indicative of their ecological niche. They have to adjust to stressors in the physical aspect of their environment (e.g., inclement weather) as well as in the biotic world that surrounds them (e.g., predators, food shortage). A major part of an individual's biotic environment consists of other members of the same species, which can be defined as his or her social world. In fact, a majority of human and nonhuman animals' daily expectations, motivations, and behaviors are directed to encounters with conspecifics. On the one hand, this social world can support good welfare and health (e.g., through the effects of social support). On the other hand, it can result in severe degrees of stress that can eventually

lead to disease and even death (e.g., in the case of social defeat, social instability, or crowding). Thus, the social environment can represent a most effective stressor, and during pregnancy it can be crucial for the development of the offspring as well.

A Study in Guinea Pigs

In nonhuman mammals, studies on prenatal social influences on offspring development have been conducted in several species, including mice, rats, bank voles, squirrel monkeys, and pigs. The most comprehensive data so far have come from guinea pigs. In this species social stimuli have widespread neuroendocrine effects, and it is argued that this species is well suited for the study of socioendocrine effects throughout the life span and can be a valuable complement to nonhuman primate research in this area. Guinea pigs in prenatal stress research were kept in two different social situations: in a stable and an unstable social environment. In the stable social environment, the group composition (one male, five females) was kept constant; in the unstable social environment, two females from different groups were exchanged every third day. The instability of the social environment during pregnancy and lactation alters the development of the offspring in a sex-specific fashion (for a summary see [Table 1](#)).

Table 1 Summary of the effects of early social stress in guinea pigs^a

	Females	Males
Behavior	SF masculinized	SM infantilized
Body weight	SF = CF	SM = CM
Serum cortisol	SF = CF	SM delayed decrease with age
Serum testosterone	SF >>>> CF	SM = CM
Adrenal TH	SF >>>> CF	SM <<<< CM
Adrenal weight	SF >>>> CF	SM = CM
Leucocytes	SF << CF	SM = CM
Granulocytes	SF << CF	SM = CM
ER- α ARC	SF >>>> CF	SM = CM
ER- α MPOA	SF > CF	SM = CM
ER- α CA1	SF >> CF	SM << CM
AR ARC	SF >>>> CF	SM <<<< CM
AR MPOA	SF >>>> CF	SM << CM
AR CA1	SF >> CF	SM = CM

^aCF, adult control females; SF, early stressed females; CM, control males; SM, early stressed males; DHEA, dehydroepiandrosterone; DHEAS, dehydroepiandrosteronesulfate; adrenal TH, activities of the adrenal tyrosine hydroxylase; ER- α , estrogen receptor- α , neurons with nuclear staining; ARC, arcuate nucleus; MPOA, medial preoptic area of the hypothalamus; CA1, pyramidal cell layer of Ammon's horn of the hippocampus; AR, androgen receptor, neurons with nuclear staining. Statistics: </>, $p < 0.1$; <</>>, $p < 0.05$; <<</>>>, $p < 0.01$; <<<</>>>>, $p < 0.001$; =, $p > 0.1$ (for details see text).

Effects on the Development of the Daughters

Daughters whose mothers lived in an unstable social environment show a most conspicuous behavioral masculinization: they display significantly higher amounts of male-typical courtship and play behavior later in life than daughters whose mothers lived in a stable social environment. The behavioral masculinization of prenatally stressed adult female guinea pigs is accompanied by increased testosterone concentrations. (Even though testosterone is mainly found in male mammals, produced predominantly by the testis, it can also be found in females. In female mammals, testosterone is secreted by the ovaries as well as the adrenal cortex; however, in much lower concentrations compared to males.) Furthermore, daughters whose mothers lived in an unstable social environment show elevated activities of the SAM system compared to daughters whose mothers have lived in a stable social environment. This activation of the SAM system is indicated by greater adrenal tyrosine hydroxylase activity. This enzyme catalyzes the rate-limiting step of the catecholamine (e.g., epinephrine, norepinephrine) biosynthesis. Finally, daughters whose mothers lived in an unstable social environment show a male-typical expression pattern of androgen receptors (ARs) in parts of the limbic system, that is, significantly more neurons with nuclear staining are present in the medial preoptic area (MPOA) and the arcuate nucleus (ARC) of the hypothalamus. Furthermore, an upregulation of ARs in the CA1 region of the hippocampus and of estrogen receptor- α (ER- α) in the MPOA, ARC, and CA1 is found in daughters whose mothers lived in an unstable social environment. The MPOA plays an essential role in the control of many aspects of sexual behavior, such as masculine behavior and sexual preference; the ARC is important for the regulation of gonadotropin-releasing factors; and the hippocampus is involved in the regulation of vegetative functions as well as in learning and memory.

Effects on the Development of the Sons

Different effects are found in early stressed males. Sons whose mothers lived in an unstable social environment display behavioral patterns, that are usually shown only by very young male guinea pigs (e.g., resting with bodily contact) more frequently and until an older age than sons whose mothers lived in a stable social environment. In addition, the courtship behavior of these sons is frequently integrated into play behavior. These effects can be regarded as behavioral infantilization, which corresponds with delayed development of the adrenocortical system (i.e., the typical decrease in concentrations of cortisol,

the main glucocorticoid in guinea pigs, with age is delayed) and – in contrast to daughters – decreased activity of the SAM system indicated by lower tyrosine hydroxylase activity. Finally, males whose mothers lived in an unstable social environment show a downregulation of AR expression in the MPOA and the ARC of the hypothalamus and of ER- α in the CA1 region of the hippocampus compared to males whose mothers lived in a stable social environment.

Comparison with Other Nonhuman Mammals

As these studies in guinea pigs show, prenatal factors can influence male and female offspring in different ways: the behavior of the female offspring is masculinized, whereas the behavior of the male offspring is infantilized. Interestingly, similar effects have been found in other studies investigating effects of prenatal social stress on offspring development, even though different stressors and different species were used. One common trait of prenatally stressed females can be described as a masculinization. In guinea pigs, this masculinization affects behavior, hormonal state, and sex steroid receptor distribution in specific parts of the limbic system. Aspects of masculinization also exist in mice and bank voles. A second characteristic feature of females prenatally exposed to social stress is a more-or-less severe impairment of reproductive function later in life. For example, in mice, less receptivity, delayed puberty, and altered estrus cycles are described. In male offspring, prenatal social stress frequently leads to a less pronounced expression of male-typical traits and/or a delay in development. In guinea pigs, this affects behavior, endocrine systems, and sex steroid receptor distribution in specific parts of the limbic system and is designated as infantilization. In other species (e.g., mice, rats and bank voles), similar effects are described in terms of demasculinization (i.e., the loss or attenuation of male-typical traits) and feminization (i.e., the expression of female-typical traits). Furthermore, some studies report an impairment in the function of reproductive organs, poorer motor abilities, altered emotional behavior, and effects on learning and memory.

In those few studies that consider offspring of both sexes, sex-specific or even sex-reversed effects are described. For example, in squirrel monkeys, prenatally stressed daughters show higher levels of immunoglobulin G (IgG), whereas in sons IgG concentrations decline. However, in most studies either the effects of prenatal social stress are investigated in only one sex or the sex is not identified.

Neuroendocrine Mechanisms of Prenatal Social Stress

A single neuroendocrine pathway clearly cannot explain the consequences of prenatal social stress. Rather, a variety of mechanisms is probably involved. For many aspects, prenatal social stress affects the offspring in a sex-specific and sex-reversed way. Furthermore, at least in guinea pigs, female offspring seem to be more susceptible to this stress than male offspring (see Table 1). However, a higher responsiveness to prenatal social stress in males has also been shown.

Pituitary Adrenocortical System

There is general agreement that maternal steroids play a decisive role in shaping fetal brain development. Most authors favor a role of the maternal hypothalamic pituitary adrenocortical (HPA) system in mediating the effects of the social environment during pregnancy on the offspring's behavior. Prenatal social stress often results in an activation of the maternal HPA system, and as a consequence the secretion of glucocorticoids is elevated. Thus, in male offspring mainly two different mechanisms are discussed: the high maternal glucocorticoid concentrations can influence the development of the fetus both directly and indirectly. In the male fetus a testosterone surge during a critical period of sexual differentiation is necessary for the masculinization of brain and behavior. It is known that high glucocorticoid concentrations can suppress testosterone secretion. Thus, the elevated glucocorticoid concentrations resulting from prenatal stress can eliminate or attenuate the testosterone surge in the male fetus. This can cause the behavioral demasculinization of the prenatally stressed males.

In female offspring a different mechanism is discussed. An activation of the HPA system triggers the secretion not only of glucocorticoids but also of androgens from the adrenal cortex. (The adrenal cortex consists of three different zones: the outermost zona glomerulosa [secreting mineralocorticoids like aldosterone], the middle zona fasciculata [secreting glucocorticoids], and the innermost zona reticularis [secreting androgens].) Since androgens are able to cross the placenta, they might cause a masculinization of the female embryonic hypothalamus during a critical period of sexual differentiation.

Sympathetic Adrenomedullary System

The preceding hypotheses assume that the pituitary adrenocortical system is involved during pregnancy in mediating the influence of the social environment on the development of the offspring, and numerous experiments support this view. However, there are

also indications that the HPA axis does not necessarily need to be involved. At least in some cases, an alternative pathway, which may include the maternal sympathetic adrenomedullary system, seems more likely, as can be seen from studies in guinea pigs. Surprisingly, pregnant females of this species living in stable and unstable social conditions do not differ in plasma concentrations of cortisol. Even the introduction into a group of unfamiliar conspecifics does not result in increased cortisol levels 2 h later. Thus, ongoing activity of the pituitary adrenocortical system in pregnant females living under stable and unstable social conditions is probably mostly comparable. In contrast to cortisol, serum dehydroepiandrosterone (DHEA) and dehydroepiandrosterone sulfate (DHEAS) concentrations are lower in pregnant female guinea pigs living in an unstable social environment compared to controls. DHEA and its sulfate are, in addition to cortisol, the major hormones produced by the adrenal cortex. DHEAS *per se* has no effect, but it serves as a hormone pool for DHEA. Although DHEA is a weak androgen, it can exert both estrogenic and androgenic effects. This may be due to the fact that both DHEAS and DHEA are converted to testosterone and 17 β -estradiol in nonendocrine cells such as macrophages.

It has been reported that pregnant female guinea pigs living in an unstable social environment show a behavioral activation. This could be the reason for the reduced androgen concentrations of such females, because a behavioral activation usually is correlated with an increased activity of the SAM system. If so, the decreased DHEAS and DHEA concentrations probably result from the elevated level of catecholamine production in the adrenal medulla, since it is well known that catecholamines suppress DHEA(S) secretion. Due to low serum levels of maternal DHEAS, lower serum levels of estrogens or androgens might be expected in the fetus. If so, a significant impact on fetal brain development might occur.

A recent experiment supports the hypothesis that there is a relationship between the decreased androgen concentrations of pregnant females living in an unstable social environment and the behavioral infantilization of their male offspring. The decreased androgen concentration is experimentally mimicked by application of the antiandrogen flutamide during the sensitive phase of fetal CNS sexual differentiation. Sons of mothers who have received the antiandrogen during pregnancy show a similar behavioral infantilization later in life, as do sons whose mothers have lived in an unstable social environment during pregnancy.

The results regarding the masculinized daughters of females living in an unstable social environment are confusing. In mammals, masculinization occurs if sufficient androgens are available during a critical period of sexual differentiation. In the case of insufficient concentrations of androgens, a female pattern develops. In guinea pigs, however, pregnant females who have lived in an unstable social environment show decreased androgen concentrations, but the daughters are behaviorally masculinized. Thus, for guinea pigs it seems unlikely that elevated maternal androgens resulting from social stress during pregnancy bring about a behavioral masculinization of their daughters. Nevertheless, androgens do seem to be involved: masculinized daughters show male-typical patterns of androgen receptor expression in the MPOA and the ARC, two brain areas that are well known to play an essential role in the control of sex-typical behaviors and the regulation of gonadotropin-releasing factors. However, at present the neuroendocrine mechanism is not clear.

Adaptive Function of Prenatal Social Stress

The characteristic traits of individuals who were exposed to environmental stimuli (stressors) during pregnancy (e.g., masculinization of daughters, feminization of sons, heightened levels of anxiety, increased cortisol responses to stressors) are often regarded as deviation from the norm or even as pathological. Alternatively, and in accordance with current evolutionary theory, these traits might represent adaptive maternal effects as well. That is, mothers try to maximize their own Darwinian fitness by efficiently adjusting their offspring to the current environmental conditions.

Effects of Prenatal Social Stress in Guinea Pigs: Pathology or Adaptation?

If prenatal influences on behavior indeed are adaptive, the following questions arise: What is the benefit of being a masculinized daughter? What is the benefit for sons who show a less pronounced expression of male-typical traits and/or a delay in development? So far, there are no empirical data to answer these questions, but reasonable hypotheses in accordance with evolutionary theories on maternal effects can be derived.

Fluctuations of natural populations of small mammals with alternating states of high and low population density are a well-known phenomenon. They also have been documented in natural populations of the feral ancestor of the domestic guinea pig. In this species, population densities can change considerably, even from year to year. Under high-density conditions,

social instability is a common trait, whereas low population densities are characterized by stable social situations. Hence, pregnant female guinea pigs can experience completely different degrees of social stability not only under experimental laboratory conditions but also in their natural habitats.

If a pregnant female that lives in a high-density population has the possibility to preprogram her daughter so that she will fit optimally into a high-density situation, what should she do? A reasonable answer is to masculinize her female offspring in order to make them more robust and/or more competitive. These traits could help her daughter more efficiently defend important resources, such as food and shelter, which are scarce in high-density situations. However, a characteristic feature of females that are masculinized behaviorally due to prenatal stress is a more-or-less severe impairment of reproductive function later in life. Thus, under high-density conditions, there is a trade-off between the benefits of a behaviorally and endocrinologically masculinized phenotype and the costs of decreased reproductive success. Nevertheless, under such conditions, the chances of reproductive success would be expected to be greater in masculinized than in nonmasculinized females. Under low-density conditions, sufficient resources are available and fewer competitive abilities are needed, so it is more beneficial to invest time and energy into reproductive effort rather than to build and maintain a male-typical phenotype to defend resources. Thus, under low-density conditions, chances of reproductive success would be greater in nonmasculinized females compared to masculinized ones.

Similarly, if mothers program their male offspring prenatally such that a less pronounced expression of male-typical traits and/or a delay in development occur, they will provide them with a promising reproductive strategy under conditions of high density.

See Also the Following Articles

Adrenal Medulla; Behavior, Overview; Corticosteroids and Stress; Environmental Factors; Gender and Stress; Hypothalamic-Pituitary-Adrenal; Social Stress, Animal Models of; Steroid Hormone Receptors.

Further Reading

- Birkhead, T., Schwabl, H. and Burke, T. (2000). Testosterone and maternal effects-integrating mechanisms and function. *Trends in Ecology & Evolution* 15, 86–87.
- Cushing, B. S. and Kramer, K. M. (2005). Mechanisms underlying epigenetic effects of early social experience: the role of neuropeptides and steroids. *Neuroscience and Biobehavioral Reviews* 29, 1089–1105.

- de Kloet, E. R., Sibug, R. M., Helmerhorst, F. M. and Schmidt, M. (2005). Stress, genes and the mechanism of programming the brain for later life. *Neuroscience and Biobehavioral Reviews* 29, 271–281.
- Dufty, A. M. Jr., Clobert, J. and Müller, A. P. (2002). Hormones, developmental plasticity and adaptation. *Trends in Ecology & Evolution* 17, 190–196.
- Gandelman, R. (1992). *Psychobiology of behavioral development*. New York: Oxford University Press.
- Kaiser, S. and Sachser, N. (2005). The effects of prenatal social stress on behaviour: mechanisms and function. *Neuroscience and Biobehavioral Reviews* 29, 283–294.
- Von Holst, D. (1998). The concept of stress and its relevance for animal behavior. In: Lehman, D. S., Hinde, R. & Shaw, E. (eds.) *Advances of the study of behavior*, pp. 1–131. London: Academic Press.
- Ward, I. L. (1972). Prenatal stress feminizes and demasculinizes the behavior of males. *Science* 175, 82–84.
- Weinstock, M. (1997). Does prenatal stress impair coping and regulation of hypothalamic-pituitary-adrenal axis? *Neuroscience and Biobehavioral Reviews* 21, 1–10.

Sexual Assault

N C Feeny and T J Linares

Case Western Reserve University, Cleveland, OH, USA

E B Foa

University of Pennsylvania, Philadelphia, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by N C Feeny and E B Foa, volume 3, pp 435–439,

© 2000, Elsevier Inc.

Prevalence of Traumatic Events and Sexual Assault

Psychological Sequelae of Sexual Assault

Treatment of Sexual Assault-Related PTSD

Predictors of PTSD

Conclusion

Glossary

Arousal

A cluster of PTSD symptoms that includes concentration problems, anger, exaggerated startle response, sleep disturbance, and being overly alert.

Avoidance

A cluster of PTSD symptoms including behavioral and psychic avoidance, such as avoidance of thoughts, feelings, and reminders of the trauma; emotional numbing; loss of interest in activities; feeling disconnected from others; psychogenic amnesia; and a sense of fore-shortened future.

Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV)

A reference manual used by mental health professionals that includes descriptions and diagnostic criteria for psychiatric disorders.

Posttraumatic stress disorder (PTSD)

An anxiety disorder that develops in response to a traumatic event and is characterized by three clusters of symptoms re-experiencing, avoidance, and arousal. PTSD is diagnosed after a 1-month period of sustained symptoms following the traumatic event.

Prevalence

How widespread a disorder or an event (e.g., sexual assault) is at a given time.

Re-experiencing

A cluster of PTSD symptoms that includes experiencing the trauma in the form of nightmares, flashbacks, intrusive distressing thoughts, or becoming intensely emotionally upset or having physiological arousal upon exposure to reminders of the trauma.

Trauma

An event experienced or witnessed, in which an individual is subjected to the threat of death, injury, or physical integrity, and during which an individual feels terrified, horrified, or helpless.

Prevalence of Traumatic Events and Sexual Assault

Estimates of exposure to traumatic events vary, but on the whole, they are quite high. Several studies have examined rates of traumatic experiences in the general population. In a nationally representative sample of approximately 6000 people, ages 15–54, 60% of men and 51% of women reported experiencing at least one traumatic event in their lifetime. Similar findings emerged from a second study using a large, racially diverse sample; 74% of men and 65% of women reported experiencing at least one traumatic event. Somewhat lower prevalence was found in a sample of young adult health maintenance organization (HMO) members ($n = 1007$): 43% of men and

39% of women reported a traumatic experience in their lifetime. This discrepancy may be accounted for by differences in the samples; members of an HMO may be of higher socioeconomic status than the general population, which may have shielded them from a high-risk level of exposure to trauma. Studies examining rates of trauma exposure internationally have found similarly high levels of exposure. In a large sample of four cities in Mexico, 76% of the participants reported at least one traumatic event in their lifetime, and the majority of the participants experienced multiple versus single traumas.

Another large epidemiological study examined rate of exposure to traumatic events among a representative sample of 4000 women. Consistent with other results, almost 70% of women reported having experienced a traumatic event during their lifetime. In addition to these high rates of single trauma exposure, prevalence data have shown women to be at high risk for experiencing multiple episodes of violence throughout their lives.

Women are more likely than men to experience the traumas of rape and sexual assault. In the literature, rape is typically considered oral or vaginal penetration, while sexual assault is a more broad term that encompasses the use of pressure and coercion to force sexual contact. Lifetime prevalence estimates of rape and sexual assault among women vary widely across studies, ranging from 2 to almost 30%. Among adult women, several studies estimated prevalence of completed rape to be 20%. Others, however, have documented much lower rates, ranging from 2 to 14%.

In an epidemiological study of 4008 women, the prevalence of completed rape was 13%; other types of sexual assault were reported by 14% of the sample. Prevalence of childhood sexual abuse is similarly high. In a nationally representative sample of 2626 adults, almost one-third (27%) of the women reported a history of sexual abuse (e.g., being touched, grabbed, kissed, shown genitals) during childhood. Of these women who reported childhood sexual abuse, 49% (13% of the entire sample) reported actual or attempted intercourse. Among adolescents, one study found that 43% of the sample had experienced a traumatic event, of which rape accounted for 2% of the traumas. Further, female adolescents were significantly more likely to report a completed rape than males, which highlights young women's risk for sexual assault. International samples also report similar levels of sexual assault in women. Of the types of trauma reported by women in the Mexico study, nearly 15% reported some type of sexual assault (10.5% sexual molestation, 3.9% sexual assault). Since sexual

assault is less likely to be reported to the police than other crimes, it is safe to assume that the rates mentioned above are an underestimation of the actual prevalence of sexual assault.

Psychological Sequelae of Sexual Assault

Posttraumatic Stress Disorder (PTSD)

Sexual assault is associated with a variety of psychological difficulties. Among the most common are symptoms of posttraumatic stress, a set of psychological and physical symptoms that follow a traumatic experience. In some cases, these symptoms persist beyond the immediate aftermath of the trauma and develop into posttraumatic stress disorder (PTSD), an anxiety disorder that includes symptoms of arousal, avoidance, and re-experiencing, which lasts for more than 1 month and causes significant impairment in social or occupational functioning. This disorder affects from 1 to 12% of the general population and close to 25% of those exposed to a traumatic event. Women appear to be twice as likely to develop PTSD as men (10% of women vs. 5% of men).

Retrospective and prospective studies show female victims of sexual assault to be at particular risk for developing PTSD. For example, Norris examined the impact of 10 common traumatic events (e.g., robbery, assault, disaster) in a large, racially diverse sample. Of all the traumatic events surveyed, sexual assault yielded the highest rate of PTSD (14% lifetime). Similarly, Resnick et al. found that 12% of rape victims had current PTSD compared to 3% of victims of nonviolent traumas. Using a national sample of 1500 women, rape and sexual assault led to the highest lifetime (39%) and current (13%) rates of PTSD. In a community sample of 400 adult females, of women who had experienced a completed rape, 57% met lifetime diagnostic criteria and 17% met criteria for current PTSD. Lower rates (37% lifetime and 11% current) were found in females who had experienced a nonsexual assault.

Two prospective assessment studies have examined the prevalence of PTSD following sexual or nonsexual assault. In a study of female rape victims, 94% of the women met symptom criteria for PTSD within 2 weeks following the assault. Three months after the sexual assault, rates of PTSD had declined considerably; however, 47% continued to meet PTSD criteria. Women who did not meet criteria for PTSD 3 months after the assault showed a steady reduction in symptoms over the 12 weeks of assessment. Minimal symptom improvement was seen for those women

who continued to meet PTSD criteria 3 months after the trauma. Follow-up assessments conducted 6 and 9 months after the assault revealed that 42% of women continued to meet criteria for PTSD at each time point. In a second study, Riggs et al. found that 90% of rape victims and 62% of nonsexual assault victims met symptom criteria for PTSD within 2 weeks of an assault. At the 1-month assessment, the rate was 60% for rape and 44% for nonsexual assault survivors. At 3 months, the rate had dropped to 51 and 21% for rape and nonsexual assault survivors, respectively. Similar findings have been shown in adolescents, with those who experienced a sexual assault being seven times more likely to meet PTSD criteria compared to all other reported trauma types. These results, taken together, suggest that sexual assault is more likely to produce PTSD than other simple assaults.

Prevalence of other Psychological Reactions

Reactions to sexual assault other than PTSD are also common. Anxiety, anger, depression, substance abuse, and dissociation occur with high frequency and are discussed in the following sections.

Anxiety Anxiety and fear are among the most common reactions to sexual assault, but are often discussed under the rubric of PTSD. After an assault, women are often quite afraid of rape- and assault-related situations and also experience general anxiety at high levels. Such fears have been documented to last up to 16 years postassault. Indeed, in one study, only 23% of sexual assault victims did not show elevated fear 1 year after the assault. Although victims' fearfulness declined over time, they remained more fearful than nonvictimized controls 1 year after the assault. Thus, fear and anxiety levels remain elevated in many victims of sexual violence long after the assault.

Anger

Elevated anger levels are often seen among sexually assaulted women compared to nonvictimized controls. In a prospective study of 116 rape and other crime victims and 50 matched nonvictimized controls, Riggs et al. found that level of anger was related to aspects of the assault such as the use of a weapon and the victim's response to the attack. In addition, elevated anger levels were related to PTSD symptoms 1 month after the assault. Similarly, Feeny et al. found anger levels 1 month after an assault to be predictive of PTSD symptom severity at 3 months postassault. The authors hypothesized that anger interferes with

recovery by preventing emotional processing of the traumatic event.

Depression Depression is another common reaction to sexual assault, although some evidence suggests that it is less persistent than fear and anxiety. In a study of psychological reactions to rape, Frank and Stewart found that almost half of their sample was diagnosed with major depression; however, over 3 months' time, these symptoms declined considerably. Retrospective examinations of depression among trauma victims, however, suggest more enduring symptoms. For example, one study found that an average of almost 22 years postassault, rape victims were more depressed than nonvictims.

In the North Carolina catchment area study, the lifetime prevalence rate of depression was significantly related to sexual assault: women who had been sexually assaulted were more than twice as likely to manifest a major depressive episode as women who had not been sexually assaulted. In addition, comorbidity of PTSD and depression is high: 48% of those who met lifetime criteria for PTSD also met lifetime criteria for depression, and 43% evidenced current comorbidity of the two disorders. In a sample seeking treatment for PTSD, 63% met criteria for current major depression. These results suggest that depression should be routinely assessed in victims of trauma who are seeking treatment.

Substance abuse Female assault victims, particularly those who develop PTSD, are also at risk for substance abuse. In a national probability sample of women, crime victims (sexual and nonsexual assault) were at increased risk of developing an alcohol use disorder, independent of their PTSD status. Moreover, those diagnosed with PTSD were 3 times more likely to have serious alcohol problems than those without the disorder and 14 times more likely than those without a history of criminal victimization. In the North Carolina catchment area study, lifetime prevalence of substance (alcohol and drugs) abuse and dependence was significantly related to the presence of sexual assault. In a recent longitudinal study, recent and past assault were associated with two to three times the risk of alcohol abuse in women, even after controlling for past alcohol use, age, race, and education.

Additionally, the prevalence of PTSD in patients enrolled in a substance abuse treatment program ranges from 12 to 34%, and prevalence rates are significantly higher among women, ranging from 33 to 59%. Among psychiatric outpatients, the rate of victimization among alcoholic women was higher

than among nonalcoholic women, suggesting a specific link between early victimization and later alcohol abuse among adult women. In summary, being a victim of a trauma appears to substantially increase the risk of developing a substance use disorder and vice versa: substance abuse seems to increase the likelihood of victimization.

Dissociation The construct of dissociation typically focuses on three clinical entities: alterations in memory (e.g., aspects of the trauma not consciously accessible), in identity (e.g., disengagement between self and environment), and in consciousness. Dissociation has been conceptualized to reflect emotional disengagement from trauma memories, which has been thought to hinder recovery. Indeed, several studies suggest that dissociation does hinder the processing of the traumatic event and subsequent natural recovery process. Trauma victims with PTSD exhibit more dissociative symptoms than those without PTSD. Moreover, if dissociation is present during or immediately after a trauma it is predictive of later psychopathology. Similarly, acute stress disorder, which emphasizes dissociative symptoms, has been found to be predictive of later posttrauma symptoms. The role of dissociation in promoting emotional disengagement needs further exploration in studies that evaluate risk factors in trauma exposed individuals.

Treatment of Sexual Assault-Related PTSD

Several cognitive behavioral treatment (CBT) programs have been developed to treat PTSD related to sexual assault, including exposure therapy, stress inoculation training (SIT), cognitive therapy (CT), eye movement desensitization and reprocessing (EMDR), and cognitive processing therapy (CPT). In this section, outcome studies, which evaluate the efficacy of these treatments with assault victims, are reviewed.

Recent studies examining the use of CBT with female sexual and physical assault victims have utilized well-controlled designs, and all but one utilized many of the methodological gold standards proposed by Foa and Meadows, including presence of PTSD diagnosis, assessment of PTSD severity, reliable and valid measures of related psychopathology, randomization to treatment conditions, and presence of control or comparison groups.

Prolonged exposure (PE) is an exposure therapy program that involves confrontation aimed at ameliorating PTSD, which typically includes mentally reliving the traumatic event repeatedly and *in vivo*

confrontation with trauma-related situations, which evoke fear but are not objectively dangerous. In the first controlled study of exposure for rape-related PTSD, the efficacy of PE, SIT, supportive counseling (SC), and a wait-list control was compared for female victims of sexual assault. PTSD symptoms were assessed at pretreatment, posttreatment, and follow-up evaluations with psychometrically sound interviews and self-report measures administered by trained clinicians who were blind to treatment assignment. PE and SIT showed significant pre-post reductions on re-experiencing and avoidance clusters of PTSD, while SC and wait list did not. Also, at the end of treatment 50% of patients in SIT and 40% of those in PE no longer met criteria for PTSD; in contrast, only 10% of SC patients and none in the wait list no longer met diagnostic criteria. At follow-up, patients in the PE group tended to show further reduction in PTSD symptoms, whereas patients in SIT and SC did not show further improvement.

In a second study, Foa et al. obtained further support for the efficacy of a modified PE in treating rape-related PTSD. This study compared the efficacy of PE to SIT, a combination of both treatments (PE/SIT), and a wait-list condition. All active treatments resulted in substantial symptom reduction and were superior to the wait-list condition in ameliorating PTSD and related symptoms. PE was superior to the other treatments: it effected a greater reduction of anxiety and depression and resulted in fewer drop-outs than SIT and PE/SIT. In addition, the number of patients who achieved good end-state functioning tended to be larger for the PE group than for the SIT and PE/SIT groups. In a third study, Foa et al. examined the efficacy of PE relative to a program that includes PE and cognitive restructuring (PE/CR). Results indicated that PE and PE/CR were superior to wait list at posttreatment on measures of PTSD and depression, with these gains maintained at follow-up. The addition of the cognitive restructuring component did not appear to enhance treatment outcomes; as with most posttreatment outcomes, PE and PE/CR did not differ from one another. Additionally, no differences in outcome were found among patients who were treated by experienced or recently trained counselors, suggesting that PE could be easily disseminated into community settings.

EMDR is a form of exposure accompanied by saccadic eye movements that has been the focus of considerable controversy due to initial claims of remarkable success in only a single session. Of the number of studies evaluating the efficacy of EMDR, only a few utilized well-controlled designs and thus yielded interpretable results. One well-controlled

study evaluated the efficacy of EMDR relative to a wait-list control condition for PTSD in female rape victims. Three sessions of EMDR resulted in greater improvement for PTSD symptoms (57% reduction in independently evaluated PTSD at posttreatment and 71% in self-reported PTSD at follow-up), relative to the wait-list condition (10% reduction at posttreatment). However, in contrast to other controlled studies, one therapist conducted all treatments, and therefore the contribution of therapist and treatment effects were confounded. A more recent study comparing PE and EMDR found that both treatment groups had significant reductions in PTSD symptoms at posttreatment. At 6-month follow-up, PE resulted in more adaptive scores on a number of anxiety and depression measures than those seen for participants who received EMDR. This study provides further support for the use of exposure techniques in the treatment of sexual assault.

Another recent study compared EMDR, PE, and relaxation training. This study found that PE effected the largest reductions in symptoms of re-experiencing and avoidance compared to EMDR and relaxation training and also effected the quickest reduction in avoidance symptoms. This study reported that EMDR and relaxation training did not differ from each other in efficacy or speed of change. Additionally, recent meta-analyses have found no evidence that the use of eye movements, integral to this treatment, have any effect on treatment outcomes. Thus, the use of the exposure techniques, which are a part of this treatment protocol, appears to account for the treatment outcomes using EMDR.

CPT is a program that was developed specifically for use with rape victims and includes cognitive restructuring and an exposure variant (i.e., writing the traumatic memories and reading the account aloud to the therapist). In a quasi-experimental design, CPT was compared to a naturally occurring wait-list control group. Overall, women who received CPT improved significantly from pre- to posttreatment (about 50% reduction in PTSD symptoms), whereas the wait-list group did not show significant improvement. In a second study, Resnick et al. compared the efficacy of a 12-session CPT, a 9-session PE, and a wait-list control in rape victims with chronic PTSD. After 9 weeks of PE and 12 of CPT, both treatments were highly effective in reducing PTSD symptoms. At follow-up assessments 3 and 9 months posttreatment, treatment gains were maintained.

A treatment specifically designed for women who experienced childhood sexual abuse has also been investigated; it is similar to CPT in that it is a combination treatment utilizing exposure and cognitive

components. This treatment consists of skills training in affect and interpersonal regulation followed by a modified PE protocol (STAIR-modified PE). In a controlled study, relative to a wait-list condition, participants receiving STAIR-modified PE showed significant improvement in symptoms of PTSD, as well as improvements in affect regulation. This study provided evidence for the use of CBT procedures with childhood abuse survivors; however, future studies should compare the usefulness of the addition of the skills training as compared to PE alone, as previous studies utilizing PE have not found beneficial effects of adding components beyond exposure techniques.

In summary, the CBT program that has received the most empirical support in controlled studies is PE, and its efficacy in reducing PTSD symptoms is at least as high as that of comparison CBT programs. Additionally, recent evidence suggests that it can be successfully used in community settings. Thus, at present, PE is the recommended psychosocial treatment choice.

Predictors of PTSD

As is apparent from the results reviewed previously, not all women who have been sexually assaulted develop chronic PTSD; thus the identification of factors that are involved in hindering recovery has been the focus of much investigation. Foa and Riggs proposed to classify predictors into three categories: pretrauma, trauma, and posttrauma variables. In terms of pretrauma variables, it appears that poor psychological functioning prior to the trauma renders the individual vulnerable to developing chronic disturbances posttrauma. In addition, a prior history of traumatic events in child- or adulthood intensifies the response to subsequent trauma. Factors related to the trauma itself, including trauma severity, also appear to predict the development of PTSD, as well as posttrauma variables such as social support.

In a prospective study of 277 female assault victims, Amir et al. tested a predictive model of PTSD that included all three domains thought to be relevant to the prediction of PTSD. Several variables emerged as significant predictors at 1 and 3 months post-assault: trauma history, trauma severity, cognitions and complex emotions, basic emotions, and early PTSD symptoms. Trauma history had an impact on cognitions and emotions, which, in turn, influenced PTSD symptom severity 4 and 12 weeks after the assault. Trauma severity influenced PTSD symptoms 2 weeks after the assault.

Recent meta-analyses provide additional information regarding the predictors of PTSD. One

meta-analysis found that the effects of predictors varied based on the sample under investigation; however, three variables were consistently significant predictors of PTSD: psychiatric history, childhood abuse, and family psychiatric history. Also, the largest effect sizes were found with predictors that were at or around the time of the trauma (e.g., social support following trauma). A second meta-analysis reviewed 476 studies that investigated PTSD predictors. Of the reviewed studies, 68 evaluated several of the hypothesized major predictors of PTSD, which include prior trauma, prior psychological adjustment, family history of psychopathology, perceived life threat, posttrauma social support, emotional response to trauma, and dissociation to trauma. Of these predictors, dissociation at the time of the trauma was the largest predictor of later developing PTSD. Indeed, the predictors with the largest effect sizes were those that were related more proximally to the trauma itself (e.g., dissociation and emotional response), similar to the finding in the Brewin et al. analyses. These results underscore the complexity of the mechanisms underlying the development of PTSD.

Conclusion

The majority of people in the United States are exposed to at least one traumatic experience in their lifetime; however, only a minority develop long-lasting psychological difficulties, including chronic PTSD. Importantly, not all traumas are equal in causing the development of PTSD; traumas inflicted by another person, especially sexual assaults, are particularly prone to produce lasting disturbances.

Once established, PTSD is unlikely to remit unless treated. The most empirically validated type of therapy for this disorder is CBT, and among the various CBT programs, PE has gained the most empirical support.

See Also the Following Articles

Anger; Anxiety; Depression and Manic-Depressive Illness; Dissociation; Drug Use and Abuse; Trauma and Memory.

Further Reading

- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental health disorders* (4th edn.). Washington, D.C.: American Psychiatric Association.
- Davidson, J. and Foa, E. B. (eds.) (1993). *Posttraumatic stress disorder: DSM-IV and beyond*. Washington D.C.: American Psychiatric Press.
- Foa, E. B. and Rothbaum, B. O. (1998). *Treating the trauma of rape*. New York: Guilford Press.
- Foa, E. B., Dancu, C. V., Hembree, E. A., Jaycox, L. H., Meadows, E. A. and Street, G. P. (1999). A comparison of exposure therapy, stress inoculation training, and their combination for reducing posttraumatic stress disorder in female assault victims. *Journal of Consulting and Clinical Psychology* 67(2), 194–200.
- Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M. and Nelson, C. B. (1995). Posttraumatic stress disorder in the national comorbidity survey. *Archives of General Psychiatry* 52, 1048–1060.
- Resnick, H. S., Kilpatrick, D. G., Dansky, B. S., Saunders, B. E. and Best, C. L. (1993). Prevalence of civilian trauma and posttraumatic stress disorder in a representative national sample of women. *Journal of Consulting and Clinical Psychology* 61(6), 984–991.

Sexual Dysfunction

W T O'Donohue and L Woodward Tolle
University of Nevada, Reno, Nevada, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by W T O'Donohue, L G Regev and D A Henderson, volume 3, pp 440–444, © 2000, Elsevier Inc.

Etiology
Interventions
Summary

Glossary

<i>Anxiety</i>	A symptom of stress.
<i>Spectatoring</i>	Monitoring one's sexual arousal, looking for signs of failure.
<i>Stress</i>	A subjective, phenomenological experience that is thought of by the individual experiencing it to be aversive; stress is generally thought to be a response to some external event or situation.
<i>Stressor</i>	External condition or situation that can cause a stress response in a given individual.

The *Diagnostic and Statistical Manual of Mental Disorders* (4th edn.; DSM-IV-TR) characterizes sexual dysfunctions as either disturbances that occur in one of the four phases of the sexual response cycle or by pain that is associated with sexual intercourse. The four phases of the sexual response cycle are desire, excitement, orgasm, and resolution. The desire phase consists of fantasies about and the motivation to engage in sexual activity. The excitement stage includes a subjective sense of sexual pleasure and corresponding physiological changes, with major changes in males being tumescence and erection, and in females vasocongestion in the pelvis, vaginal lubrication and expansion, and the swelling of the external genitalia. The orgasm phase consists of a peaking of sexual pleasure accompanied by the release of sexual tension and contraction of the perineal muscles and reproductive organs. In males, ejaculation of semen occurs in this phase, and in females, there are rhythmic contractions of the wall of the vagina. Finally, the resolution phase involves a sense of muscular relaxation and well-being. Males are physiologically refractory to another erection and orgasm for a period of time; however females can respond to additional stimulation almost immediately.

Sexual dysfunction can occur in any one of these phases or more than one, and factors such as age, experience, frequency and chronicity of the symptoms, subjective distress, and the symptoms' effects on other areas of functioning must be considered. The sexual dysfunction disorders as outlined in the DSM-IV-TR include sexual desire disorders (hypoactive sexual desire disorder and sexual aversion disorder), sexual arousal disorders (female sexual arousal disorder and male erectile disorder), orgasmic disorders (female orgasmic disorder, male orgasmic disorder, and premature ejaculation), sexual pain disorders (dyspareunia and vaginismus), sexual dysfunction due to a general medical condition, substance-induced sexual dysfunction, and sexual dysfunction not otherwise specified.

The prevalence of sexual dysfunctions is difficult to accurately report because, given the sensitivity of the issue, it is estimated that most cases go untreated or unreported. Of the epidemiological studies that have been conducted, there has been a large variability in terms of the prevalence of various sexual dysfunctions found in different populations. Other factors that might be contributing to these different findings could be the use of different assessment methods, the definitions used, or simply the varying characteristics of the samples used. Studies have estimated that the lifetime incidence of sexual problems is anywhere between 10 and 52% in males and 25 and 63% in females, whereas other studies have found ranges of only 5–13% in males and females.

Etiology

Research has demonstrated that the onset of sexual dysfunction varies greatly in that it could be present from the initial onset of sexual functioning or it could occur after a period of normal sexual functioning. Sometimes sexual dysfunction is not limited to certain types of stimulation, situations, or partners, and sometimes it is. Several considerations must be taken into account when making clinical judgments about the presence of a sexual dysfunction. Among these considerations are the individual's ethnic, cultural, religious, and social backgrounds. Another factor that adds a considerable degree of subjectivity to the diagnosis is the adequacy of sexual stimulation. There are no norms for this or for any of the factors clinicians must render judgments about. These factors may all influence individuals' sexual desire, expectations around sexual activity, and attitudes about their and their partner's performance.

There are also several risk factors associated with sexual dysfunctions that might be biological or psychological in nature. Biological factors can either directly lead to sexual problems or indirectly contribute. Some of the most common direct biological risk factors for sexual dysfunction include vascular disease, diabetes, hormone levels, the use of alcohol, and the use of certain medications (for example antihypertensives). Common indirect biological risk factors include age, cigarette smoking, chronic pain, heart disease, cancer, and other medical procedures that might alter an individual's appearance, leading to body-image insecurities and possible subsequent problems with normal sexual functioning. Psychological risk factors for sexual dysfunction include the presence of mood disorders, anxiety disorders, substance use disorders, eating disorders, and, largely, stress. Various negative life stressors adversely affect sexual functioning, including job stress, financial stress, bereavement or the loss of a loved one, divorce or relationship problems, problems conceiving, and marriage. Much of the literature examining the risk factors associated with sexual dysfunction in humans is correlational; however, studies using animals have provided more substantial evidence about the effects of stress on sexual functioning. Specifically, researchers have found that both physical stress (e.g., shock or immobilization) and social stress (e.g., overcrowding, other aggressive animals attacking, or isolation) result in copulatory disorder (equivalent to male erectile disorder in humans) in male mice and that, when faced with more mild stressors, mice copulatory behavior decreased. These findings are consistent in research conducted with human sexual behavior when faced with life stressors. Sexual dysfunction has been found to be more prominent among couples who report more severe and greater stressors in their lives compared to couples who report having fewer life stressors. In addition, sexual desire has been found to be lower or absent in individuals who report higher levels of occupational stress. Finally, research has discovered that sexual dysfunction is more likely in couples in which one or both partners have a diagnosed anxiety disorder.

Anxiety and stress, although appearing to frequently have different operational definitions, are frequently used interchangeably in the literature. Some studies examining the effects of anxiety on sexual arousal have produced the interesting findings that anxiety was found to negatively affect sexual arousal in individuals with sexual dysfunction but may actually enhance arousal in individuals with normal sexual functioning. It is possible that when sexually functional men and women become anxious during sexual activity, their autonomic arousal increases and they focus more on erotic cues, which enhances their

arousal. Individuals with sexual dysfunction, on the other hand, become anxious, and when their autonomic arousal increases, they focus on not being able to perform (which Masters and Johnson called "performance anxiety" or spectating) and subsequently fail to become aroused.

When a sexual dysfunction is present, it often acts as an additional source of stress in the individual's or couple's lives, serving only to further exacerbate the dysfunction. This appears to be especially true in couples who are trying to conceive because there is not only pressure to perform sexually for pleasure and intimacy but also to get pregnant. Arguments, blame, embarrassment, and a lack of communication about sexual dysfunction can also increase stress and cause the sexual problem to remain unimproved.

Sexual Desire Disorders

The sexual desire disorders include hypoactive sexual desire disorder and sexual aversion disorder. Hypoactive sexual desire disorder is characterized by a deficiency or absence of sexual fantasies and desire for sexual activity, taking into consideration age and the individual's personal history. Sexual aversion disorder is characterized by a persistent or recurrent extreme aversion to and avoidance of all, or almost all, genital sexual contact with a sexual partner. Possible stress-related causes of both of the described disorders include anxiety over difficulty getting aroused or climaxing, marital stress, fear of intimacy, fear of losing control, guilt or religious prohibitions, fear of sexually transmitted diseases, fear of blood or pain during intercourse, posttraumatic stress disorder following rape or sexual assault, or a secondary trigger for panic disorder. Women are twice as likely to experience the desire disorders and are also more likely to report higher levels of psychological distress than men, more marital dissatisfaction than men, and greater self-reported stress levels than men, specifically in the areas of related to home and finances. Most frequently, hypoactive sexual desire disorder develops in adulthood, after a period of normal sexual functioning, in association with stressful life events or interpersonal difficulties.

Sexual Arousal Disorders

The sexual arousal disorders consist of female sexual arousal disorder and male erectile disorder. Female sexual arousal disorder is characterized by a persistent or recurrent inability to attain, or maintain until completion of the sexual activity, an adequate lubrication and swelling response of sexual excitement. Male erectile disorder is characterized by a persistent or recurrent inability to attain, or to maintain until

the completion of the sexual activity, an adequate erection. Women who experience female sexual arousal disorder often report little or no subjective sense of arousal and are unlikely to seek professional treatment because this problem can frequently be remedied with the aid of a lubricant. Without intervention, female sexual arousal disorder can result in painful intercourse, avoidance of sexual activity, and relationship problems. With male erectile disorder, different patterns can be seen with this sexual dysfunction. Some individuals may have difficulty obtaining an erection, some are able to obtain an erection but lose it when attempting penetration, some are able to sustain an erection on penetration but cannot on thrusting, and other males are able to sustain an erection only during self-masturbation or on awakening. Erectile difficulties are frequently associated with sexual anxiety, fear of failure, anxiety over sexual performance, and a decreased subjective sense of sexual excitement and pleasure. Erectile dysfunction can cause marital and relationship problems and may be the cause of infertility. The introduction of Viagra has caused male erectile disorder to become well known and in some cases obsolete.

Orgasmic Disorders

The orgasmic disorders consist of female orgasmic disorder, male orgasmic disorder, and premature ejaculation. Female orgasmic disorder is characterized by a persistent or recurrent delay in, or absence of, orgasm following a phase of normal female arousal taking into consideration the woman's age, sexual experience, and adequacy of sexual stimulation that she receives. Male orgasmic disorder is characterized by a persistent or recurrent delay in, or absence of, orgasm following a phase of normal male sexual arousal taking into consideration the man's age and his adequacy in focus, intensity, and duration. Premature ejaculation is characterized by a persistent or recurrent ejaculation with minimal stimulation before, on, or after penetration and before the individual wishes it. Stress-related etiological factors for women with female orgasmic disorder can include marital dissatisfaction, high levels of distress, and interpersonal stress. In addition, female orgasmic disorder has been found to adversely affect women's body image and self-esteem, which can then exacerbate the sexual dysfunction. Male orgasmic disorder is associated with performance anxiety, spectatoring, previous sexual traumas, and sexual and marital dissatisfaction. Premature ejaculation is associated with embarrassment and reluctance to engage in a sexual relationship due to fear of rejection and ridicule. Males with premature ejaculation often socially isolate themselves to prevent these events from taking place.

Sexual Pain Disorders

Sexual pain disorders include dyspareunia and vaginismus. Dyspareunia is characterized by recurrent or persistent genital pain associated with sexual intercourse in either a male or a female. Vaginismus is characterized by recurrent or persistent involuntary contraction of the perineal muscles surrounding the outer third of the vagina when penetration with a penis, finger, tampon or speculum into the vagina is attempted. Stress-related etiological factors associated with dyspareunia include anxiety, guilt from religious inhibitions, poor body image, and marital or relationship dissatisfaction. Vaginismus is associated with anxiety, interpersonal stress, and prior sexual abuse or traumatization.

Interventions

If stress is found to be an underlying cause of the sexual dysfunction, treatment will include strategies designed to decrease stress.

Psychoeducation

Psychoeducation can be a powerful tool in alleviating sexual dysfunction, not only in its ability to inform individuals but also to address false beliefs, unrealistic expectations, and lack of knowledge about physiology, anatomy, and normal sexual functioning, and it can, in and of itself, relieve anxiety associated with doing it wrong or feeling abnormal. As a result, psychoeducation can lead to rapid improvements in only a couple of sessions. Excerpts from human sexuality textbooks, complete with diagrams and pictures, can help provide a comprehensive overview of sexual functioning. Addressing common myths associated with human sexual functioning can also be helpful in educating individuals with sexual dysfunction.

Systematic Desensitization

Systematic desensitization is an effective therapy strategy designed to reduce anxiety. The individual is first equipped with relaxation skills and then is asked to come up with a hierarchy of increasingly anxiety-provoking stimuli or situations. When using this technique, an individual is placed in a deeply relaxed state and is presented with a series of gradually increasing anxiety-provoking situations using imaginal exposure. The therapist starts with the least anxiety-provoking stimulus on the hierarchy and then slowly works up to more anxiety-inducing stimuli. The therapist asks the individual to imagine the stimulus, and when the individual experiences anxiety during the exposure exercise, the therapist asks the individual to allow him- or herself to get rid of the

image and then induces the original relaxed state. This procedure is used for longer and longer periods of time until the individual is able to retain the image without experiencing any anxiety; this is continued until the hierarchy is completed successfully.

Cognitive Restructuring

Intrusive negative thoughts can largely contribute to sexual dysfunctions. Ultimately, these thoughts interfere with sexual arousal and enjoyment because individuals are no longer able to concentrate on the sexual activity. Frequently, these negative thoughts may revolve around anxiety over their performance, one partner's dissatisfaction, or negative expectations about the sexual activity itself. In addition, individuals with multiple life stressors may experience thoughts concerning these stressors that may be a part of their sexual dysfunction and that can similarly disrupt sexual performance. In using cognitive restructuring, the therapist helps the individual identify these thoughts and then come up with other, more sensual or erotic thoughts to replace with these during intercourse. In this way, the individual can learn to turn the previously anxiety-provoking situation into a pleasurable one. In some cases, individuals have a difficult time coming up with the negative thoughts they typically have during intercourse. In these situations, it is useful to have them keep a diary, monitoring these thoughts as they occur and recording them following sexual activity. Another way to challenge individuals' negative thoughts is to help them determine whether the thoughts have any factual basis. This helps individuals let go of some negative and false assumptions that they might be carrying that are preventing them from enjoying intercourse.

Sensate Focus

Sensate focus is an effective tool used to help individuals with sexual dysfunction, specifically in cases in which performance anxiety appears to play a key role. Simply focusing on maintaining an erection or achieving an orgasm as the main goal of sexual intercourse can often cause anxiety to increase, particularly if there is a past history of the individual being unable to do so. Feelings of failure and disappointment can lead to increased anxiety about the next sexual encounter, and more preoccupation with achieving success often leads to subsequent failures. Sensate focus seeks to reduce anxiety by allowing individuals to take this pressure to perform out of the way by simply not allowing sexual intercourse for a time. A series of exercises are completed by the couple that include nonsexual touching and caressing while clothed, progressively moving through a series of increasingly demanding sexual activities over the

course of treatment. The individuals are taught to focus on the sensations they experience rather than engaging in spectating. Because the pressure to aspire to the ultimate goal of maintaining an erection or achieving an orgasm is no longer the focus, the couple is able to restructure its goals in sexual encounters and thus reduce the anxiety and stress associated with performance anxiety.

Couples Therapy

Interpersonal, marital, and couples issues appear to play a central role in the development and/or maintenance of several sexual dysfunctions. Many of these issues are significantly improved by the implementation of communication skills training. Most sexual partners have a difficult time discussing their sexual relationship with one another. There are often fears about hurting one another's feelings, saying the wrong thing, or not being able to feel that they can ask for more of what they want and less of what they do not want. An improvement in overall communication can increase relationship satisfaction while creating a more effective line of communication to discuss more delicate matters, such as the couple's sexual relationship. Conflict resolution skills are important in maintaining a fair and even balance in times when there might otherwise be power struggles, and these can also help to build trust and respect. Enhanced communication not only alleviates stress from the relationship but also can assist in bringing the couple closer, both of which can contribute to a more satisfying sexual relationship.

Pharmacological Interventions

As mentioned previously, Viagra has become a household name, and it and its competitor medications are making a huge impact on sexual dysfunction. Not only does it significantly alleviate erectile dysfunction in a large proportion of men (the female equivalent to Viagra is beginning to be developed and has seen promising results in alleviating female arousal disorder), but it is also helping to decrease the social stigma surrounding sexual dysfunction and is allowing individuals to feel more comfortable about disclosing this problem to their physicians. The negative side to this is that, because it has become so convenient to get a prescription for this type of sexual enhancement drug, it is beginning to be abused by individuals who do not have sexual dysfunction and are simply using it recreationally. Naturally, physicians do not want to make it difficult for individuals who honestly may be suffering from erectile dysfunction to receive available support and help, so they do not withhold prescriptions from individuals reporting that they need the

drug. Thus, for now, no effective solution has been proposed to prevent this abuse from taking place.

Summary

Sexual dysfunctions are very common, affecting a large percentage of the population. Although epidemiological data are inconclusive, it is possible that only a small proportion of individuals suffering from sexual dysfunctions are being effectively treated. Stress appears to play a large role in the development, persistence, exacerbation, and maintenance of sexual dysfunctions. Several evidence-based interventions have been found to be effective with this population, including systematic desensitization, couples therapy, and sensate focus. In addition, pharmacological interventions have been helpful in alleviating the symptoms of sexual dysfunction. Further research is needed to continue to examine the role of stress in sexual dysfunctions and to development effective treatments to help prevent them.

Further Reading

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental health disorders* (4th edn.). Washington, D.C.: American Psychiatric Association.

Barlow, D. (ed.) (2001). *Clinical handbook of psychological disorders: a step by step treatment manual*, (3rd edn.). New York: Guilford Press.

Fasolo, C., Mirone, V., Gentile, V., et al. (2005). Premature ejaculation: prevalence and associated conditions in a sample of 12,558 men attending the andrology prevention week 2001 – a study of the Italian Society of Andrology. *Journal of Sexual Medicine* 2(3), 376–382.

Heiman, J. R. (2002). Sexual dysfunction: overview of prevalence, etiological factors and treatments. *Journal of Sex Research* 39(1), 73–78.

Lagana, L., Classen, C., Caldwell, R., et al. (2005). Sexual difficulties of patients with gynecological cancer. *Professional Psychology: Research and Practice* 36(4), 391–399.

Leiblum, S., Brown, C., Wan, J., et al. (2005). Persistent sexual arousal syndrome: a descriptive study. *Journal of Sexual Medicine* 2(3), 331–337.

Mohan, R. and Bhugra, D. (2005). Literature update: a critical review. *Sexual and Relationship Therapy* 20(1), 115–122.

O'Donohue, W., Fisher, J. and Hayes, S. (eds.) (2003). *Cognitive behavior therapy: applying empirically supported techniques in your practice*. New York: John Wiley.

O'Donohue, W. and Geer, J. H. (1993). *Handbook of sexual dysfunctions: assessment and treatment*. Needham, MA: Allyn & Bacon.

Sexual Offenders

F M Saleh

University of Massachusetts Medical School,
Worcester, MA USA

H M Malin

Institute for Advanced Study of Human Sexuality,
San Francisco, CA, USA

© 2007 Elsevier Inc. All rights reserved.

General Considerations

Definitional Considerations

Treatment of Sex Offenders

Concluding Remarks

Glossary

Ephebophilia

A paraphilia in which the object of sexual attraction is a pubescent youth (child or adolescent), but not yet a fully sexually mature person, of either sex; also called hebephilia.

Paraphilia

Recurrent, intense sexual urges and fantasies, sometimes accompanied by sexually driven behaviors, occurring over a period of at least 6 months, involving one or more of the following: (1) a nonconsenting partner (e.g., children or a rape victim), (2) an object or body part (e.g., fur or feet), or (3) a sensory/emotional state (e.g., pain or humiliation) not typically associated with the archetypal sexual-response cycle. Paraphilias may be obligatory, in which case they are necessary components for completing the sexual-response cycle, or nonobligatory, in which case they may sometimes be incorporated in the sexual-response cycle but are not always essential for its completion. Paraphilias may be either pathological or non-pathological depending on the degree to which they impair the paraphile's social functioning and emotional well-being. They may be dangerous (such as lust

	murder or autoasphyxiophilia) or benign (such as transvestic fetishism or foot partialism) depending on their potential for lethality or adverse physical or social impact on self or others.
<i>Pedophilia</i>	A paraphilia involving sexual arousal to a prepubescent child or children of either sex.
<i>Perversion</i>	Once a psychiatric term roughly equivalent in meaning to paraphilia; now generally a pejorative term used by non-professionals to describe “kinky” sex or other expressions of sexuality (e.g., homosexual behavior) of which they disapprove.
<i>Phenomenology</i>	In the study of paraphilias or other psychiatric phenomena, the scientific method in which the phenomena experienced by the individual, as measured by observation and/or self-report, are described and classified rather than explained or interpreted.
<i>Sex offender</i>	A legal term denoting an individual who has broken a law governing the conduct of sexual behavior; the term is usually reserved for individuals who have been convicted of a crime involving sexual behavior.
<i>Sexual-response cycle</i>	The sequence of stages and events typically occurring in the progression of sexual activity from the beginning of sexual arousal to orgasm and resolution. Based on work by Masters and Johnson and amplified by other sexologists, the sexual-response cycle is sometimes arbitrarily divided into five stages or phases: (1) desire phase (sometimes called libido or drive); (2) excitement phase (sometimes called arousal), in which sexual arousal normally builds; (3) plateau (the point of highest arousal sustainable short of triggering orgasm); (4) orgasm (or climax); and (5) resolution (sometimes called the refractory period), in which arousal returns to prebaseline desire and during which an additional stimulus does not immediately lead to a repetition of the cycle.

General Considerations

It is universally assumed to be axiomatic that victims of sexual assault in its various manifestations suffer from the stress of their experiences to one degree or another. However, little or no scientific investigation has been conducted that sheds light on the relationship of stress to the experience of having been sexually victimized. Among important unstudied and,

therefore, unanswerable questions are: What experiences are most likely to cause toxic stress, who is at greatest risk for developing stress-induced sequelae, and what factors might protect against or mitigate the adverse consequences of stress resulting from sexual victimization? It may be concluded, however, that those who work with and care for the sexually victimized are highly sensitized to the importance of stress-related components of sexual assault, whether the assault takes the form of uninvited genital exposure or rape with the threat of death.

The situation with respect to sex offenders is roughly equivalent, although there is the added variable that cultures typically do not concern themselves much with the impact of stress on sex offenders, who are generally regarded as willful evildoers and not as psychiatrically ill. In this article, it is not our intent to explore to any great degree the ongoing controversy of whether sex offenders are criminals or patients and the degree to which the psychiatric impairment in some sex offenders may influence culpability for their offending behavior.

The question of whether individuals who engage in illegal paraphilia-driven behaviors should be considered as having diminished criminal culpability given their psychiatric diagnosis is controversial and complex. Over a century ago, the eminent British psychiatrist Henry Maudsley expressed concern that society often fails to even consider such a possibility. He wrote, “If the law cannot adjust the measure of punishment to the actual degree of responsibility . . . that is no reason why we should shut our eyes to the facts; it is still our duty to place them on the record, in the confident assurance that a time will come when men will be more able to deal more wisely with them.”

By definition, however, sex offenders are criminals with offenses ranging from misdemeanors to felonies. The term sex offender is strictly a legal term. There is no psychiatric diagnosis of sex offender, although there is some overlap between psychiatry and the law in that some individuals who have diagnosable psychiatric illness do commit sex offenses. Some sex offenders, including but not limited to those with paraphilias, may be more likely than others to commit sex offenses as a result of their psychiatric syndromes. Although severe comorbid psychiatric conditions may mitigate social response (e.g., the punishment may be less severe for an exhibitionist responding to command hallucinations), sex offenders are generally assumed to have adequate volitional control over their offending behaviors and are, thus, expected to conform their behavior to the expectations of society. Those who do not are likely to feel the full measure of retribution social systems are capable of bringing to

bear in dealing with them. That response, inevitably, induces toxic stress.

Indeed, with the diminishing emphasis on rehabilitation and the ascendancy of retribution in American penology, punishment is meted out with the specific intention of subjecting sex offenders to pain, suffering, and the attendant stress the retributive penal philosophy guarantees. The response of society to sex offenses is deliberately and systematically to induce intensive stress in the offender by prescribing a continuum of ever harsher punishment ranging, at minimum, from humiliation and social opprobrium to often-lengthy prison sentences and, occasionally, death. These stressors are intentionally inflicted with full knowledge of the adverse consequences of the stress they engender.

The underlying philosophy of retributive penology appears to be threefold. First, it is axiomatic that in serious offenses, retribution by the state is both appropriate and desirable because the state itself and the social values it defends are directly aggrieved by the crime. Second, it is axiomatic that the state has the duty to punish on behalf of individuals more specifically aggrieved by the crime, to bring miscreants to justice as their proxy, which prevents the degeneration of retribution into vengeance, which, in turn, risks sowing the seeds of vendetta and the further destabilization of society. Finally, retributive penology is motivated by the premise that institutionalized punishment promotes deterrence by example. That is, not only is the offender punished but those who would follow in his criminal footsteps, witnessing his punishment, will be deterred from committing similar crimes. In some cases, there is a bonus – a sex offender incarcerated for life or executed for his crimes will not himself recidivate. We mention, parenthetically, an observation first made by John Money that certain paraphilias are the only mental illnesses deemed to merit the death penalty rather than psychiatric treatment in some jurisdictions.

With respect to perpetrators of sex offenses, the authors are not aware of any peer-reviewed research or literature directly addressing the issue of stress in the offender, although it seems equally axiomatic that both victims and offenders may experience toxic stress when the sex offending takes place. There have been limited studies of stress in prison populations, which may provide leads for research in this area. It is thought, for example, that inmates or corrections officers with a certain psychological makeup or privileged position are immunized against the stress of the prison setting to a greater degree than other inmates or corrections officers in the same setting.

Part of the difficulty in assessing the impact of stress on both victims and sex offenders is that neither

is a homogeneous group. In this article, we raise some issues relevant to stress in sex offenders and leave to others the parallel task of defining stress-related issues in the victims of sex offenses.

Definitional Considerations

Sex Offender

As we have pointed out, the term sex offense is a legal term. What constitutes a sex offense differs from jurisdiction to jurisdiction. As we pointed out in the preceding section, sex offenders are not a homogeneous group, with the sole exception of the fact that they have broken some law governing sexual behavior in a given culture or jurisdiction. It is sometimes not well understood that cross-cultural as well as intra-cultural considerations must be taken into account in defining sex offenses as well stress with respect to atypical sexual urges, fantasies, and behaviors.

For example, in some jurisdictions in the United States and in similarly less sexually tolerant foreign cultures, certain sexual behaviors such as premarital or extra-marital sexual contact, homosexual acts, cross-dressing, and heterosexual oral-genital or anal sexual contact are sex crimes and their practitioners are thus sex offenders. In virtually all cultures and jurisdictions, some sexual behaviors such as sexual contact with children (although the age of majority varies widely) are crimes and the individuals who engage in them are sex offenders as well.

Whether or not such criminal behaviors always, or even usually, engender stress depends on myriad considerations, including the degree of enforcement of the laws prohibiting the behavior, potential punishments, and whether individuals who engage in them find their own behavior personally acceptable (ego-syntonic) or unacceptable (ego-dystonic). Just as sex offenders are not a homogeneous group, the resultant stress from sex offending is not homogeneous.

Perhaps the only constant in the term sex offense and its equivalents in other cultures is that it is a legal construct. As such, it is not necessarily congruent with psychiatric or psychological views of these acts. Indeed, the differences between legal and psychiatric conceptualizations of sexual behaviors often leads to considerable dissonance and clouds the social debate about what should be done about sex offenders; they can lead to violent culture clashes, complicating considerably our understanding of stress as it relates to sex offending.

Phenomenology

Phenomenology, as we use the term here, is the personal, internal experience of the sexual fantasy or

behavior without respect to explanations for the behavior itself. Phenomenology, which has its roots in existential philosophy and, by extension, existential psychology, concerns itself with a scientific approach based on the phenomena themselves (here sexual drives, feelings, and motivated behaviors), without attempting to further explain them, as, for example, a psychodynamic approach might do. It also fails to embrace *in toto* the tenets and conventions of behaviorism. This theoretical approach has profound implications for treatment of “sex offenders.”

Phenomenology is premised on the concept that behavior is determined by the way a person perceives reality rather than on some external objective reality. Thus, phenomenology does not rely on attempts to explain or interpret in order to classify sex offending. Nor does it rely overly much on changing behaviors themselves with the expectation that changing behaviors will change behavior. As a theory of psychology, phenomenology draws heavily on the tenets of existential psychology, which posits that psychology should concern itself with the organism’s own experience of existence in the world. It draws heavily on the philosophical work of Karl Jaspers and the psychological application of that work by existential theorists such as Medart Boss, Ludwig Binswanger, and Viktor Frankl.

Paraphilias

Sex offenses and sex offenders are often classified as paraphilic or nonparaphilic.

The term, paraphilia, (from Greek *para-* beside + *philos* love) is an English translation of a term first proposed by Wilhelm Stekel in his book *Sexual Aberrations*, which first appeared in English about 1925. Reportedly, Stekel believed that a new term with less pejorative connotations than perversion (from Latin for turning around) would be helpful in examining and ameliorating these mental illnesses. Freud used the term, but it was not in widespread use in the psychiatric literature until the 1950s.

The *Diagnostic and Statistical Manual* (4th edn., text revision, DSM-IV-TR) of the American Psychiatric Association proposes that the “essential features of a paraphilia are recurrent, sexually arousing fantasies, sexual urges or behaviors generally involving 1) non-human objects, 2) the suffering or humiliation of oneself or one’s partner, or 3) children or other non-consenting persons, that occur over a period of at least 6 months” (Criterion A; 2000: 566). DSM-IV-TR suggests that for some individuals “paraphilic fantasies or stimuli are obligatory for erotic arousal and are always included in sexual activity.” For other individuals, paraphilic “preferences appear only

episodically (e.g., during periods of stress)” (2000: 566). Criterion B for paraphilias, according to the DSM-IV-TR, is that they “must cause clinically significant impairment in social, occupational or other important area of functioning” (2000: 566). Thus, the DSM-IV-TR definition of paraphilia is at variance with the definition that we proposed previously, in that behavior alone is sufficient to make a diagnosis. We insist on a definition of paraphilia that presupposes motivated behavior that is sexually driven by recurrent, intense sexual urges and fantasies.

As an example, consider the variety of motivations driving the behavior of genital exposure:

- Preparing to bathe, dress, urinate, or otherwise engage in the gamut of biologically or socially based needs ordinarily considered to be nonsexual.
- Engaging in nude sunbathing, swimming, or other nonsexual recreational pursuits alone or with consenting others in lawful locations.
- Permitting a genital examination or other procedure for medical reasons.
- Engaging in sexual behavior with oneself or with consenting others.
- Attempting to attract the attention of nonconsenting people with or without the intention to initiate sexual contact.
- Attempting to shock, embarrass, humiliate, punish, frighten, or harass a nonconsenting person.
- Engaging in a social or cultural protest or making a political statement.
- Responding to a dare, exposing oneself as a penalty in a game or as a requirement for consensual initiation into social group.
- Engaging in a marginally socially accepted behavior, such as streaking or flashing, or participating exposed in a parade, festival, or other activity in which others are likely to do the same thing.
- Modeling for sculpture, painting, photography, or other artistic endeavors such as making films or participating in theatrical performances, either for remuneration or for altruistic reasons.
- Stripping for remuneration at a club or prenuptial party.
- Being sexually assaulted or intimidated into exposing oneself.
- Submitting to a lawful strip search.
- Preparing to commit a sexual assault on a nonconsenting person.
- Exposing oneself secondary to disinhibition caused by exogenous substances.
- Exposing oneself in response to command hallucinations.
- Exposing oneself secondary to dementia or traumatic brain insult.

Surely, most of these instances of genital exposure cannot be considered paraphilic. Some are prosocial, whereas some are antisocial. Some have negative or neutral erotosexual valence, whereas some are highly erotic. Paraphilic behavior, influencing as it does the sexual-response cycle, is by definition erotosexual. Yet only a subset of the exposing behaviors listed here that have positive erotosexual valence might be motivated by recurrent, intense sexual urges and fantasies outside some presumptive or defined norm. Only those that satisfy this criterion can be considered paraphilic.

Etiology There is no known etiology of paraphilias, but some investigators believe there are associations between paraphilias and other life experiences. The most common association reported is between paraphilias and early childhood experiences. Theories change and track dominant schools of psychological thought. For example, when psychodynamic theory was the predominant psychiatric theory, fur fetishism was explained as a sequella of the birth experience, in which an infant became fixated on the pubic hair of the mother during childhood. In the light of present-day understanding, such an explanation seems unlikely. However, more satisfying explanations have not yet been forthcoming.

More recently, various researchers have postulated a relationship between the development of paraphilias and childhood physical or sexual abuse. Such theories have led to concepts such as the cycle of abuse, in which it is postulated that the abused becomes the abuser by some mechanism as yet not clearly delineated. There is currently considerable debate about the proposition that such abuse may have occurred in preverbal children and the memories recovered many years later. Empirical studies have not been kind to the cycle of abuse construct, although it is at the foundation of a number of still widely used treatment regimens.

Although a good statistical correlation between having been abused as a child and perpetrating abuse as an adult has not been demonstrated, in some sexually abusive individuals there is a considerable history of childhood sexual, physical, and/or emotional abuse or neglect. The recalled or sometimes corroborated sexual biographies of a great many paraphilic individuals contain experiences in which the patient's paraphilia seems closely linked to these experiences. For example, individuals who received frequent spankings, enemas, or even highly embarrassing physical examinations as children have sometimes reported that as adults these unpleasant experiences have become highly pleasurable with positive erotic valence and may be incorporated as obligatory or nonobligatory paraphilic components of their sexual-response

cycle. One possible explanation for this transformation may be the mechanism of opponent process theory of acquired motivation proposed by Richard Solomon and his colleagues, in which, like skydiving, an intensely fear-laden experience gradually becomes highly pleasurable, sought-after exhilarating activity. Other investigators have suggested that paraphilias are reinforced by masturbation to deviant fantasies. Unfortunately, all these ideas, as well as others, await solid empirical evidence. At this point, we can say with certainty only that the genesis of paraphilias remains largely unknown.

Prevalence There are in excess of 50 named paraphilias in the sexological literature. Some appear to be quite rare, whereas others, especially in their nonobligatory form, seem relatively common. The paraphilias most frequently coming to the attention of clinicians are pedophilia (which typically includes ephebophilia, although they are different paraphilias), voyeurism, and exhibitionism, according to DSM-IV-TR. This is hardly surprising, given that acting on these paraphilias is illegal in most jurisdictions, and individuals acting on these paraphilias are the most likely to seek clinical treatment, typically mandated by the judiciary. Given the distribution of pornographic materials with legal paraphilic content, for example, films and magazines devoted to transvestic fetishism, leather fetishism, or the milder forms of sadism and masochism, we might speculate that these paraphilias are among the more common. Yet sadism and masochism are relatively infrequently seen in clinical practice.

Erotic materials depicting other kinds of paraphilias, for example zoophilia, are typically suppressed. And although zoophilia is illegal in most jurisdictions, it seems to be less aggressively prosecuted, even when it comes to the attention of the authorities, than, say, sexual behavior involving children. It is difficult to extrapolate from any available data the prevalence of a given paraphilia or an all-inclusive prevalence of paraphilias in the American culture. DSM-IV-TR reports that, except for sexual masochism, with a sex ratio of 20 males to each female, paraphilias are almost never diagnosed in females. DSM-IV-TR further reports that approximately half of all individuals with paraphilia seen in clinics are married.

Treatment of Sex Offenders

In practice, the treatment of sex offenders may depend less on careful differential diagnosis than on the convictions that the treatment provider holds about sex offenders. Treatment providers with a behavioral view rely more heavily on psychological treatments

designed to change behaviors than do treatment providers who subscribe to a more medical model approach. These providers tend to use more pharmacologically based interventions than their behaviorist colleagues. The authors rely on an approach that takes advantage of an appropriate mix of interventions, following a careful differential diagnosis.

Psychological Treatments

The various psychological therapies, whether individual or group therapies, are designed to change behaviors through a variety of strategies including challenging patients' distorted thought and belief patterns (cognitive behavioral therapy), developing strategies to anticipate and resist deviant sexual cravings (relapse prevention), and replacing maladaptive behaviors with more functional behaviors (behavioral therapy). More traditional psychodynamic and psychoanalytic psychotherapy treatment modalities have been largely abandoned in the treatment of sex offenders because of their lack of demonstrated efficacy despite decades of attempts.

Behavior therapy Behavioral therapy is guided by learning theory, specifically social learning theory. In contrast to psychodynamic therapies, behavior therapies are primarily concerned with the aberrant behavior itself and not its underlying cause. That is, behavior therapy aims to decrease and extinguish deviant sexual arousal through a set of techniques, such as systematic desensitization, aversion therapy, biofeedback, minimal arousal conditioning with aversive conditioning, and covert sensitization. (Masturbatory satiation, a form of arousal reconditioning, is no longer widely used.)

Covert sensitization, a form of aversion therapy, has been widely used in the treatment of paraphilic patients. It is postulated that the pairing of deviant sexual fantasies with the mental image of its humiliating consequences (i.e., the aversive experience) will dissuade a patient from acting out sexually. In covert sensitization, deviant arousal or excitement is paired with a visualized distressing thought or fantasy. To give an example, a voyeur will be asked to imagine being taken into custody as soon as he begins to contemplate peeping into other people's windows.

In contrast to the aversive behavioral therapies, operant conditioning methods use positive reinforcers. For example, positive olfactory conditioning employs pleasant aromas with nondeviant sexual stimuli.

Cognitive-behavioral therapy Cognitive-behavioral therapy hypothesizes that paraphilias are maintained by distorted cognitions and reinforced by masturbation to inappropriate fantasy. The aim of cognitive

therapy is to change maladaptive thoughts and beliefs through techniques such as cognitive restructuring and thought stopping and to eliminate reinforcers that maintain distorted cognitions. The technique of cognitive restructuring teaches patients to identify, challenge, and ultimately replace erroneous (i.e., distorted) beliefs with adaptive and pro-social cognitions. Thought stopping is a technique that aims to decrease the frequency and duration of deviant sexual thoughts.

More often than not, a combination of one or more of these therapies is used. Victim empathy training has traditionally been an important component of all behavioral treatment and involves helping a patient take on the perspective of his victim(s). In addition, assertiveness and a variety of social skills training programs have been deemed to be important components in treatment programs, although recent research seems to indicate that empathy or social skills training have little or no impact on sexual recidivism in treated populations.

Relapse prevention A relapse is a recurrence of undesirable behavior after a period that has been free of the undesirable behavior. The term "relapse" is not synonymous with "recidivism" which indicates re-arrest for an illegal behavior. The objective of relapse prevention is to help a patient maintain behavioral changes. In order to reduce the risk of inappropriate sexual behavior, patients are taught to identify their deviant cycle and to develop strategies to implement when the cycle is triggered.

Group therapy Group psychotherapy is a widely accepted treatment modality used in the treatment of paraphilic and nonparaphilic sex offenders. It incorporates relapse prevention and cognitive-behavioral techniques, as well as empathy and social skills training. Group therapy creates a supportive milieu that is conducive to frank discussion and that provides guidance regarding effective relapse prevention strategies. It also encourages the development of a social support network and of healthy adult relationships. One of the primary advantages of group therapy over individual therapy is that it allows patients to be challenged by members of the group rather than by the therapist. This is important given that patients may respond and be more willing to change if challenged by one of their peers.

Effectiveness of psychological treatments The effectiveness of the psychological treatments in treating paraphilias is difficult to assess. One measure of treatment success has traditionally been recidivism. Some studies have concluded that these treatments

have positive effects; other studies have found minimal or no effect with behavioral therapies. The interpretation of these studies may be confounded by a number of variables such as non-standardized treatment delivery or heterogeneous populations. Also, studies of the treatment of sex offenders generally do not differentiate paraphilic from nonparaphilic offenders.

Nor is it clear that treatment delivered in one setting, say a prison, generalizes to another setting, such as the community. State-dependent learning, for example, may impact psychological treatments that depend on a psycho-educational approach. Relapse prevention, a widely used psychological treatment modality, may be less effective than previously believed. In one randomized clinical trial that compared the reoffense rates of patients in an inpatient relapse prevention treatment program with the reoffense rates in two untreated prison control groups, no significant differences among the three groups in rates of sexual or violent reoffending over an 8-year follow-up period were found.

Biological Treatments

The authors believe that with certain patients, for example those with dangerous paraphilias or those at acute risk of victimizing someone, pharmacological intervention should be the treatment option of first resort. Biologically based treatments should also be made available to paraphilic patients if, for example, cravings for deviant sexual activities become overpowering or when specific symptoms are not responsive to less invasive treatment modalities (e.g., behavioral therapy).

Surgery Among the biologically based treatments, we must differentiate between surgical and pharmacological interventions. Stereotaxic hypothalamotomy is only of historical interest. On the other hand, the orchiectomy (surgical removal of the gonads) data are particularly relevant in that they provide the basis for our understanding of the benefits of the currently used testosterone-lowering agents. Orchiectomy consistently lowers the rate of recidivism. Orchiectomy has been rendered largely unnecessary with the advent of the pharmacological treatment.

Pharmacotherapy For ethical and medical reasons, randomized, double-blind, placebo-controlled pharmacological treatment studies cannot be conducted with symptomatic paraphilic patients who reside in the community. Such studies could only be justified on patients with relatively innocuous paraphilias. Nevertheless, the number of medications used to treat paraphilias has been steadily increasing. These

medications can be divided into two general classes: those that act by lowering available androgens (e.g., progesterone derivatives and the gonadotropin releasing hormone analogs) and those that do not directly reduce androgen activity but reduce libido and/or sexual function as a side effect (e.g., some serotonin reuptake inhibitors). As with all pharmacological treatments, the choice of which medication to use is primarily based on the patient's presenting symptoms, his or her concomitant psychiatric/neurological disorders, and the results of the psychosexual and medical workup. A regrettable shortcoming of pharmacotherapy, however, is that it does not selectively reduce the risk for sex offending without generally impairing overall sexual functioning.

Concluding Remarks

There are a number of studies in the literature suggesting that treatment can be helpful, particularly in paraphilic samples. One such study reported rates of recidivism following treatment in a primarily paraphilic sample that were well below the rates reported in the world literature for comparable untreated cases. In another, a meta-analysis that reviewed data from 43 studies examining the sex offense recidivism rate for mixed sex offenders without differentiating between paraphilic and nonparaphilic patients found that recidivism, on average, was approximately 12% lower for treatment groups than for nontreatment groups.

Measuring treatment effectiveness presents significant problems. Recidivism, defined as some subset of involvement with the criminal justice system during or after the completion of treatment, is the most commonly used end point for determining treatment effectiveness in sex offenders. But not all paraphiles are sex offenders and not all sex offenders are paraphiles. Clearly, it is not possible to use recidivism as a measure of treatment efficacy with patients whose paraphilias are not illegal. Defining recidivism and dealing with important but largely unknown variables such as the base rate for reoffending and other methodological impediments greatly influence the interpretation of the available data and are central to answering questions about which treatment modalities work and which do not.

Sex offenses are prevalent and those motivated by paraphilia-driven phenomenology often constitute significant psychiatric disorders. Other, nonparaphilic sex crimes may have their roots in serious psychopathology as well. All too frequently, patients manifesting these conditions are simply ostracized, stigmatized, and punished when what is called for is enlightened psychiatric care.

Ignoring phenomenology and defining even observable manifestations of human sexuality in purely behavioral terms can only lead to egregious errors. Nowhere is this more apparent than in the arena of sex offender treatment. The phrase itself seems to us to be misleading, if not an oxymoron. With or without civil commitment statutes, clinicians are often called on to treat sex offenders. For example, a phenomenological view of rape implies that rape is not a monolithic sex offense but a highly variable one. Rape may be an opportunistic act, carried out by an individual without a sense of conscience. It may be the result of command hallucinations. It may also be paraphilic. It seems reasonable to assume, as in all psychiatric treatment, that how we approach the treatment of a rapist, or for that matter any sex offender, must depend to some degree on differential diagnosis, which, in turn, depends on the phenomenology underlying the behavior.

See Also the Following Articles

Antisocial Disorders; Sexual Assault.

Further Reading

- Abel, G. G., Mittleman, M., Becker, J. V., et al. (1988). Predicting childmolesters' response to treatment. *Annals of the New York Academy of Sciences* 528, 223–224.
- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th edn., text rev.). Washington, DC: American Psychiatric Association.
- Berlin, F. S. (1989). The paraphilias and depo-provera: some medical, ethical and legal considerations. *Bulletin of the American Academy of Psychiatry and the Law* 17(3), 233–239.
- Berlin, F. and Malin, H. (1991). Media distortion of the public's perception of recidivism and psychiatric rehabilitation. *American Journal of Psychiatry* 148(11), 1572–1576.
- Bradford, J. M. W. (2000). The treatment of sexual deviation using a pharmacological approach. *Journal of Sex Research* 37(3), 248–257.
- Freund, K. (1980). Therapeutic sex drive reduction. *Acta Psychiatrica Scandinavica* 62(supplement 287), 1–39.
- Frost, J. J. and Mayberg, H. S. (1986). Alteration in brain opiate receptor binding in man following arousal using C-11 carfentanil and positive emission tomography. *Journal of Nuclear Medicine* 27, 1027 (abstract of presentation at the 33rd annual meeting of the Society of Nuclear Medicine).
- Greenberg, D. M., Bradford, J. M. W., Curry, S., et al. (1996). A comparison of treatment of paraphilias with three serotonin reuptake inhibitors: a retrospective study. *Bulletin of the American Academy of Psychiatry and the Law* 24(4), 525–532.
- Hanson, R. K., Gordon, A., Harris, A. J., et al. (2002). First report of the collaborative outcome data project on the effectiveness of psychological treatment for sex offenders. *Sex Abuse* 14(2), 169–194.
- Krafft-Ebing, R. von. (1965). *Psychopathia sexualis* (1886). New York: G. P. Putnam's Sons.
- Langelüddeke, A. (1963). *Die Entmannung von Sittlichkeitsverbrechern in Deutschland*. Berlin: DeGruyter.
- Laws, D. R., Hudson, S. M. and Ward, T. (2000). *Remaking relapse prevention with sex offenders: a sourcebook*. Thousand Oaks, CA: Sage.
- Marshall, D. R. and Laws, W. L. (2003). A brief history of behavioral and cognitive behavioral approaches to sexual offenders. Part 1: Early developments. *Sexual Abuse: A Journal of Research and Treatment* 15(2), 93–120.
- Money, J. (1985). *The destroying angel*. Buffalo, NY: Prometheus Books.
- Money, J. (1986). *Venuses penuses: sexology, sexosophy and exigency theory*. Buffalo, NY: Prometheus Books.
- Moser, C. (2002). Are any of the paraphilias in the DSM mental disorders? *Sexual Behavior* 31(6), 490–491.
- Murphy, W. D. and Carich, M. S. (2001). Cognitive distortions and restructuring in sexual abuser treatment. In: Carich, M. S. & Stephens, E. M. (eds.). *Handbook for sexual abuser assessment and treatment*, pp. 65–75. Brandon, VT: Safer Society Press.
- Ortman, J. (1980). The treatment of sexual offenders: castration and antihormone therapy. *International Journal of Law and Psychiatry* 3, 443–451.
- Quinsey, V. L., Harris, G. T., Rice, M. E., et al. (1993). Assessing treatment efficacy in outcome studies of sex offenders. *Journal of Interpersonal Violence* 8, 512–523.
- Rosler, A. and Witztum, E. (1998). Treatment of men with paraphilia with a long acting analogue of gonadotropin-releasing hormone. *New England Journal of Medicine* 338, 416–465.
- Saleh, F. M. and Berlin, F. S. (2003). Sex hormones, neurotransmitters and psychopharmacological treatments in men with paraphilic disorders. *Journal Child Sexual Abuse* 12(3–4), 53–74.
- Saleh, F. M. and Guidry, L. L. (2003). Psychosocial and biological treatment considerations for the paraphilic and nonparaphilic sex offender. *Journal of the American Academy of Psychiatry and the Law* 31, 486–493.
- Saleh, F. M., Niel, T. and Fishman, M. (2004). Treatment of paraphilia in young adults with leuprolide acetate: a preliminary case report series. *Journal of Forensic Sciences* 49(6), 2–6.
- Stürup, G. K. (1968). Treatment of sexual offenders in Herstedvester, Denmark: the rapist. *Acta Psychiatrica Scandinavica* 44, 5–63.

Shift Work See: Sleep Loss, Jet Lag, and Shift Work.

Sickle Cell Disease and Stress

K Midence

Institute of Medical and Social Care Research,
Bangor, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3,
pp 445–446, © 2000, Elsevier Inc.

Medical Complications

Stress and Chronic Illness

Sickle Cell Disease-Related Stressors

Coping Styles and Sickle Cell Disease

Glossary

Sickle cell disease (SCD) An inherited blood disorder including sickle cell anemia, sickle cell hemoglobin disease, and sickle β -thalassaemia. SCD affects around 1 in 400 African Americans, and approximately 50,000 individuals have sickle cell anemia in the United States. SCD is also found in Europe, India, Africa, Central and South America, and the Caribbean.

Vaso-occlusive episodes Recurrent and painful crises are caused by the sickle hemoglobin molecules, which may stick together under low oxygen tension, distorting the red blood cells into a pathognomonic and rigid sickle shape and reducing their normal elasticity.

SCD, like all chronic illnesses, can be considered as a major stressor to which children and their families have to adjust by reducing the stresses associated with the condition and developing appropriate coping skills to ensure that the affected children and their families can lead a normal life.

Medical Complications

Vaso-occlusive crises caused by sickled red cells are responsible for chronic and acute damage to tissue and organs and cause recurrent and unpredictable

episodes of acute pain, often very violent, in the affected parts of the body, most usually the arms, legs, joints, and back, causing 91% of hospital admissions. Other complications include abdominal pain and enlargement of the liver or spleen, infections, delayed growth and onset of puberty, problems with pregnancy, priapism, avascular necrosis, and strokes, which may also occur in childhood. Effective clinical management requires medical and psychological skills. Medical treatment is mainly supportive and involves rapid and adequate analgesia, hydration, and treatment of any precipitating events such as infection and inflammatory conditions.

Stress and Chronic Illness

Wallander and Varni suggested that families with a chronically ill child are at greater risk of maladjustment because they are exposed to a greater number of stressful situations. These may happen to the child directly (e.g., missing school because of complications) or indirectly, affecting the family (e.g., marital disagreements and financial difficulties). They also suggested a set of factors associated with adjustment to chronic illness, including intrapersonal (e.g., functional independence), interpersonal (e.g., coping style of the family), and social-ecological (e.g., family functioning and socioeconomic status). This approach emphasizes the need to identify the specific stresses related to a disease and also generates the hypothesis that increased frequency and severity of stressors will result in poor adjustment. Adjustment could then be defined in terms of how well an individual deals with the typical stressful situations associated with a specific chronic illness.

Sickle Cell Disease-Related Stressors

For children and adolescents with SCD, the most frequent and severe stress is probably the experience of pain, both mild and severe; the uncertainty of painful crises; its adverse effect on their daily lives; and the risk of a sudden death. Pain crises and hospitalizations have been reported by children with SCD as

the most stressful aspects of their lives. SCD can disrupt school attendance and activities (e.g., sports), which may have a negative effect on their social life and academic performance. In adults with SCD, adjustment has been related more closely to life stress (e.g., financial and work) and social support. SCD can limit or disrupt a wide range of activities because exercise, cold exposure, and other factors can precipitate sickling.

However, family life and employment are the most significant areas of social life in adults with SCD. An increased personal responsibility for coping and financial status play important roles in coping with SCD, and most of those affected belong to ethnic groups that are already at a significant socioeconomic disadvantage. It has been suggested that the more psychosocial stressors within the family, the less effectively the SCD patient and the family cope with the condition and that families with more social support available appear to cope more favorably with the stressors associated with SCD.

Psychological research suggests that life stress and psychosocial adversity are the most likely precursors of maladjustment and psychopathology in adults with SCD. Thompson and colleagues found that over half of adults with SCD in their study met the criteria for poor adjustment, with 40 and 32% showing symptoms of depression and anxiety, respectively. Better adjustment was associated with lower levels of perceived daily stress, higher efficacy expectations about illness tasks, and less use of palliative coping and negative or passive coping strategies, but it was not associated with type of SCD or measures of severity. Family functioning was related to psychological status, with higher family support and lower family conflict and control among those with better adjustment.

Coping Styles and Sickle Cell Disease

Research has shown that active coping strategies (i.e., addressing stressors) are associated with good adjustment, whereas avoidance coping strategies (i.e., ignoring or avoiding stressors) are associated with poor adjustment. Passive strategies and negative thinking have been associated with lower activity levels and greater treatment requirements. Thompson and colleagues provided a transactional stress and coping model that views chronic illness as a potential stressor to which the individual and the family attempt to adapt. Thus, the transactional stress and coping model includes parameters that are illness-specific as well as common demographic and adaptational parameters that are assessed across illnesses.

Coping is hypothesized to be a function of the transactions of illness parameters, demographic

parameters, and adaptational processes. Illness parameters include type, severity, and age of onset. Demographic parameters include patient age, gender, and socioeconomic status. Adaptational processes include cognitive processes of stress appraisal; expectations of efficacy; and health locus of control, self-esteem, and causal attributions.

Thompson and coworkers found that the pain-coping strategies of children with SCD were characterized by negative thinking, which accounted for a 20% increment. Variables of the transactional stress and coping model accounted for 30% of the variance in child-reported symptoms and 45–49% of the variance in mother-reported behavior problems. Thompson and colleagues also found that higher levels of stress and lower levels of self-worth accounted for a significant portion of the variance in self-reported symptoms.

Future research needs to investigate illness-specific patterns of adaptation and to develop interventions aimed at enhancing methods of managing cognitive appraisals of illness-related stress.

See Also the Following Articles

Avoidance; Coping Skills.

Further Reading

- Midence, K. and Elander, J. (1994). *Sickle cell disease: a psychosocial approach*. Oxford: Radcliffe Medical Press.
- Midence, K. and Elander, J. (1996). Adjustment and coping in adults with sickle cell disease: an assessment of research evidence. *British Journal of Health Psychology* 1, 95–111.
- Thompson, R. J., Gil, K. M., Abrams, M. R., et al. (1992). Stress, coping, and psychological adjustment of adults with SCD. *Journal of Consulting and Clinical Psychology* 60, 433–440.
- Thompson, R. J., Jr., Gil, K. M., Burbach, D. J., et al. (1993). Role of child and maternal processes in the psychological adjustment of children with sickle cell disease. *Journal of Consulting and Clinical Psychology* 61, 468–474.
- Thompson, R. J., Jr. and Gustafson, K. E. (1996). *Adaptation to chronic childhood illness*. Washington, DC: American Psychological Association Press.
- Thompson, R. J., Jr., Gustafson, K. E., Gil, K. M., et al. (1998). Illness specific patterns of psychological adjustment and cognitive adaptational processes in children with cystic fibrosis and sickle cell disease. *Journal of Clinical Psychology* 54, 121–128.
- Wallander, J. L. and Varni, J. W. (1992). Adjustment in children with chronic physical disorders: programmatic research on a disability-stress-coping model. In: La Greca, A. M., Siegel, L. J., Wallander, J. L. & Walker, C. E. (eds.) *Stress and coping in child health*. New York: Guilford Press.

Sleep Loss, Jet Lag, and Shift Work

E Van Cauter

University of Chicago, Chicago, IL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by E V Cauter, volume 3, pp 447–448, © 2000, Elsevier Inc.

Stress and Sleep Loss

Stress and Jet Lag

Stress and Shift Work

Glossary

Endogenous circadian rhythm

An oscillation with a period of approximately 24 h generated by a pacemaker mechanism located in the suprachiasmatic nuclei of the hypothalamus. This endogenous rhythm persists in the absence of any periodic variations in the environment (e.g., light–dark cycle and rest–activity cycle).

Jet lag syndrome

The malaise experienced by the majority of travelers during the days following a flight across time zones. The syndrome usually includes symptoms of daytime fatigue, difficulties initiating and maintaining sleep, deterioration of mood, vigilance, and cognitive function and may also involve gastrointestinal disorders. The syndrome is believed to result from the misalignment between bodily rhythms and the 24-h periodicities in the environment.

Sympathetic and parasympathetic tone

The autonomous nervous system plays a crucial role in the stress response and includes the sympathetic and parasympathetic nervous systems. Nearly all organs receive both sympathetic and parasympathetic inputs. During stress, sympathetic activation is increased and thus the balance of sympathetic versus parasympathetic control (i.e., the sympatho-vagal balance) is shifted toward higher sympathetic as compared to parasympathetic control.

The evolution of industrial society has been associated with major changes in the behavioral control of the sleep–wake cycle and in the environmental light–dark cycle. The availability of electrical lighting has made extended activity hours and around-the-clock operations possible. Misalignments among

endogenous circadian rhythms, sleep–wake cycle, and light–dark cycle are highly prevalent in the working population and also affect millions of travelers who use rapid air transportation across time zones. While it is the common perception that sleep loss, jet lag, and shift work are stressful, there are remarkably few studies of either subjective or objective markers of stress in these conditions.

Stress and Sleep Loss

Voluntary sleep curtailment is a hallmark of modern society. In the United States, normal sleep duration has decreased from approximately 8.5 h in 1960 to an average of less than 7.0 h in 2006. To maximize the time available for work and leisure activities, it has become socially desirable to curtail sleep to the shortest amount tolerable. Around-the-clock operations imply chronic partial sleep loss for the large population of shift workers and the repeated occurrence of total sleep deprivation for key personnel.

Although multiple studies have shown that total sleep deprivation or repeated sleep curtailment by as little as 2–3 h per night results in marked decrements in mood and performance, subjective indices of stress in subjects studied in the laboratory remain generally low. In particular, self-ratings on scales of tense and nervous remain stable across extended periods of total sleep deprivation under comfortable and sedentary laboratory conditions. Similarly, ratings of perceived stress remain low and stable in subjects submitted to up to 1 week of partial sleep deprivation in the laboratory. Sleep deprivation does not appear to be a classical stressor, resulting in an immediate activation of the sympathetic nervous system or in the rapid release of stress hormones. Indeed, the only known immediate effect of acute total sleep deprivation on physiological function is the absence of responses normally elicited by sleep onset or awakening from sleep, such as declines and increases in heart rate. However, sleep loss, even when only partial, appears to be associated with an alteration in the regulation of the hypothalamic–pituitary–adrenal (HPA) axis, the major neuroendocrine transducer of stress. Indeed, in sleep-deprived subjects, the normal morning-to-evening decline of blood levels of the stress hormone cortisol is slowed down and, as a result, evening cortisol levels are elevated. This alteration suggests that the rate of recovery of the HPA axis from exogenous stimulation by stressors may be affected by sleep loss. An elevation of cardiac

sympathetic activity occurs following recurrent sleep restriction. Thus, sleep loss may indeed result in an activation of the major biological mechanisms involved in the stress response. However, even though sleep loss appears to be interpreted biologically as a stressor, subjectively, the subjects do not seem to experience sleep restriction as stressful since they consistently rate their perceived stress as low, irrespective of bedtime duration. This apparent dissociation between the biological and psychological response to sleep loss may only occur in the unchallenging environment of the laboratory. The combination of increased sympathetic nervous activity and elevated evening cortisol levels in a state of sleep debt may play a role in the deterioration of parameters of glucose tolerance that is associated with short sleep.

Stress and Jet Lag

Jet lag syndrome is a condition of desynchrony between the near 24-h (i.e., circadian) endogenous rhythm generated in the suprachiasmatic nuclei of the hypothalamus and the external light-dark and rest-activity cycles that arises following rapid transportation across time zones. The syndrome usually includes symptoms of fatigue, sleep disturbances, reduced cognitive and psychomotor performance, and gastrointestinal tract disorders. The syndrome subsides as adaptation to local time is achieved. The 24-h secretory rhythms of hormones of the HPA axis adapt at a slower rate (averaging 1 day per hour of time difference) than diurnal variations in catecholamine levels or sympatho-vagal tone. Adaptation to jet lag nearly always involves some degree of sleep loss and therefore alterations of HPA regulation typical of sleep deprivation are to be expected.

Stress and Shift Work

To meet the demands of around-the-clock operations, millions of shift workers live in a condition of chronic sleep debt and desynchrony between endogenous rhythms and environmental periodicities. Indeed, studies have shown that shift workers get significantly less sleep per scheduled sleep period than daytime workers, with many sleeping on average less than 5 h per work day. In contrast to occasional jet lag, shift work is a major health hazard, involving an increased risk of cardiovascular illness, gastrointestinal disorders, psychosocial complaints, sleep disturbances, substance abuse, reduced immune function, and infertility.

Shift workers thus live in a situation of conflicting zeitgebers, which almost never allow a complete shift of the circadian system. Indeed, several studies have

shown that workers on permanent or rotating night shifts do not adapt to these schedules, even after several years. In particular, the temporal profile of secretion of the stress hormone cortisol shows little, if any, adaptation to the work schedule. There is evidence for increased blood pressure and increased sympathetic versus parasympathetic tone during night shift as compared to day shift activity in emergency rooms of hospitals. Under chronic conditions, such disturbances could result in cardiovascular illness.

See Also the Following Articles

Industrialized Societies; Sleep, Sleep Disorders, and Stress; Workplace Stress.

Further Reading

- Adams, S. L., Rose, D. M., Weiss, J., et al. (1998). Ambulatory blood pressure and monitoring of emergency physicians before, during, and after a night shift. *Academic Emergency Medicine* 5, 871-877.
- Bliwise, D. L. (1996). Historical change in the report of day-time fatigue. *Sleep* 19, 462-464.
- Bonnet, M. and Arand, D. (1995). We are chronically sleep deprived. *Sleep* 18, 908-911.
- Broman, J. E., Lundh, L. G. and Hetta, J. (1996). Insufficient sleep in the general population. *Neurophysiologie Clinique* 26, 30-39.
- Czeisler, C. A., Johnson, M. P., Duffy, J. F., et al. (1990). Exposure to bright light and darkness to treat physiologic maladaptation to night work. *New England Journal of Medicine* 322, 1253-1259.
- Dinges, D., Pack, F., Williams, K., et al. (1997). Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. *Sleep* 20, 267-277.
- Ekstrand, P., Bostrom, P. A., Arborelius, M., et al. (1996). Cardiovascular risk factors in commercial flight aircrew officers compared with those in the general population. *Angiology* 47, 1089-1094.
- Folkard, S., Minors, D. S. and Waterhouse, J. M. (1985). Chronobiology and shift work: current issues and trends. *Chronobiologia* 12, 31-54.
- Knutsson, A., Akerstedt, T., Orth-Gomer, K., et al. (1986). Increased risk of ischaemic heart disease in shift workers. *Lancet* 2, 89-92.
- Leprout, R., Copinschi, G., Buxton, O., et al. (1997). Sleep loss results in an elevation of cortisol levels the next evening. *Sleep* 20, 865-870.
- Roden, M., Koller, M., Pirich, K., et al. (1993). The circadian melatonin and cortisol secretion pattern in permanent night shift workers. *American Journal of Physiology* 265, R261-R267.
- Spiegel, K., Leprout, R., L'Hermite-Balériaux, M., et al. (2004). Impact of sleep duration on the 24-hour leptin

profile: relationships with sympatho-vagal balance, cortisol and TSH. *Journal of Clinical Endocrinology and Metabolism* 89, 5762–5771.

Spiegel, K., Knutson, K., Leproult, R., et al. (2005). Sleep loss: a novel risk factor for insulin resistance and type 2 diabetes. *Journal of Applied Physiology* 99, 2008–2019.

Tepas, D. I. and Carvalhais, A. B. (1990). Sleep patterns of shiftworkers. *Occupational Medicine* 5, 199–208.

Weibel, L., Spiegel, K., Follenius, M., et al. (1996). Internal dissociation of the circadian markers of the cortisol rhythm in night workers. *American Journal of Physiology* 270, E608–E613.

Sleep, Sleep Disorders, and Stress

A N Vgontzas, S Pejovic and M Karataraki

Penn State University College of Medicine, Hershey, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A N Vgontzas, E O Bixler and A K Kales, volume 3, pp 449–457, © 2000, Elsevier Inc.

Normal Sleep
Sleep Disorders
Sleep, Sleep Disorders, and Physical and Emotional Stress
Sleep and Stress System
Conclusion

Glossary

Addison's disease An endocrine disease that is associated with aldosterone and cortisol (hormones) by the adrenal glands.

Apnea Cessation of breathing lasting at least 10 s.

Cataplexy A brief and sudden loss of muscle control without loss of consciousness usually precipitated by strong emotions.

Continuous positive airway pressure (CPAP) A small apparatus which through a soft plastic mask blows air gently through the nose into the throat to keep the airway open.

Cosinor analysis Least square approximation of time series using a cosine function of known period.

Dihydroxy-phenylacetic acid (DOPAC) A metabolite of dopamine.

Dihydroxy-phenylglycol (DHPG) A metabolite of norepinephrine produced by the action of the enzyme monoamine oxidase (MAO) which can be measured in urine.

Hypersomnia Disorder of excessive diurnal and nocturnal sleep.

Hypnagogic Occurrence of an event while drifting into sleep.

Hypoxia

A pathological condition in which the body as a whole (generalized hypoxia) or region of the body (tissue hypoxia) is deprived of adequate oxygen supply.

Of unknown causation.

Idiopathic Microneurography

A method to record the traffic of impulses in human nerves with percutaneously inserted needle electrodes.

Narcolepsy

Sleep disorder associated with excessive daytime sleepiness, involuntary daytime sleep episodes, disturbed nocturnal sleep, and cataplexy (sudden loss of muscle tone).

Parasomnia

Is any sleep disorder such as sleepwalking, teeth grinding, night terrors, rhythmic movement disorder, rapid eye movement (REM) behavior disorder, restless leg syndrome, and sleep talking, characterized by partial arousals during sleep or during transitions between wakefulness and sleep.

Polysomnography

The continuous and simultaneous recording of physiological variables during sleep, e.g., electroencephalography, electrooculography, electromyography, electrocardiography, respiratory flow, etc.

Pulsatile analysis

A global analysis of the mechanical function of the cardiovascular system, including the determinants of cardiac output.

Slow-wave sleep (SWS)

Sleep characterized by electroencephalogram waves of duration slower than 4 Hz and high amplitude.

Stage rapid eye movement (REM) sleep

The stage of sleep with highest brain activity characterized by enhanced brain metabolism and vivid dreaming.

Sleep appears to be an important element in maintaining equilibrium or homeostasis that is vital for the survival of living organisms. In this article, we review the association between sleep, sleep disorders, and stress, which has been defined as a state of disharmony, or threatened homeostasis.

Normal Sleep

Humans spend at least one-third of their lives asleep, yet there is little understanding of why we need sleep and what mechanisms underlie its capacities for physical and mental restoration. However, there has been a significant increase of empirical knowledge that is useful in the evaluation and management of most sleep complaints and their underlying disorder.

The interaction of circadian effects, i.e., usual time to go to sleep and amount of prior wakefulness (homeostatic response), determines the onset and amount of sleep. Natural sleep–wake rhythms cycle at about 25 h rather than coinciding with the solar 24-h schedule. As a result, many people depend on external cues to keep their diurnal cycle on time. The normal diurnal clock resists natural changes in its pattern by more than about 1 h per day, which explains the sleep difficulties that usually accompany adaptation to new time zones or switches in work shifts.

Individuals differ considerably in their natural sleep patterns. Most adults in nontropical areas are comfortable with 6.5–8.0 h daily, taken in a single period. Children and adolescents sleep more than adults, and young adults sleep more than older ones. Normal sleep consists of four to six behaviorally and electroencephalographically (EEG) defined cycles, including periods during which the brain is active (associated with rapid eye movements, called REM sleep), preceded by four progressively deeper, quieter sleep stages graded 1–4 on the basis of increasingly slow EEG patterns. Slow-wave sleep (SWS; also referred to as deep sleep; stages 3 and 4) gradually lessens with age and usually disappears in the elderly.

Sleep Disorders

Sleep disorders are common in the general population and are associated with significant medical, psychological, and social disturbances. Insomnia, the most common sleep disorder, most often reflects psychological disturbances. Excessive daytime sleepiness is the predominant complaint of most patients evaluated in sleep disorders clinics and often reflects organic dysfunction. Narcolepsy, idiopathic hypersomnia, and sleep apnea are the most common disorders associated with excessive daytime sleepiness. Narcolepsy and idiopathic hypersomnia are chronic brain disorders with an onset at a young age; sleep apnea occurs predominantly in middle-aged men and (to a lesser degree) women and is associated with obesity and cardiovascular complications. The parasomnias, including sleepwalking, night terrors, and nightmares, have benign implications in childhood but often reflect psychopathology or significant stress in adolescents and adults and organicity in the elderly.

Sleep, Sleep Disorders, and Physical and Emotional Stress

“Stress is life and life is stress” in Hans Selye’s words. If it is a fact that there is no life without stress, it is also true that there is no life without sleep. Thus, it is not surprising that the association between sleep and physical and emotional stress has been long recognized. Hippocrates observed early on the relationship between mental and physical health and sleep and stated that sleeplessness is a sign of pain and suffering and may lead to mental illness while sleeping during the day (daytime sleepiness) is an indication of illness. Also, sleep has been an ancient remedy in the hands of physicians to combat emotional or physical stress, e.g., infection.

Acute or short-term sleep disturbance is usually associated with a variety of situational stresses (work-related, interpersonal, or financial difficulties) or with medical problems such as pain, cardiopulmonary or gastrointestinal disorders, or the febrile prodromes to influenza. Various drugs, including those that humans have been using for hundreds of years to decrease stress and dysphoria such as alcohol, nicotine, and caffeine, can affect adversely the quantity and quality of sleep. Stressful life events at the onset of sleep disorders are quite common, particularly in those disorders that appear to be of emotional origin, such as insomnia. It has been found that 75% of the chronic insomniacs had experienced some stressful life event at about the time insomnia began. Stressful life events are present, although to a lesser degree compared to insomniacs, in other sleep disorders, such as narcolepsy or sleep apnea, or adult sleepwalking, night terrors, and nightmares. The presence of stressful life events at the onset of organic disorders, such as narcolepsy or sleep apnea, is considered coincidental. However, it is interesting to note that, for example, it has been proposed that in narcolepsy it is the combination of genetic factors and stress, including emotional stress, which leads to the manifestation of the disorder. In this respect, a recent study showed that a majority of people with narcolepsy had a stressful life event at onset of their illness.

Psychological distress appears to be quite common among patients with sleep disorders. In some of these disorders, such as chronic insomnia or adult parasomnias, it appears to be primary or etiologic. In other sleep disorders, such as sleep apnea, narcolepsy, or idiopathic hypersomnia, it appears to be secondary or reactive to the chronic physical problems associated with the sleep disorder. Furthermore, it has been reported that in insomniacs, it is not only stress, but also how they handle stress that leads to chronic difficulties in initiating and maintaining

sleep. In particular, insomniacs handle external stress and conflict by internalization of their emotions which results in a combination of emotional arousal and physiological activation. This mental and physiological hyperarousal leads to difficulty in initiating sleep, whether at the beginning of the sleep period or when returning to sleep following awakening. Fear of sleeplessness and performance anxiety further intensifies the emotional arousal. All of these factors through behavioral conditioning lead to a vicious circle that perpetuates insomnia.

Because of the high frequency of emotional distress associated with sleep disorders, stress management techniques are routinely recommended to all sleep-disorder patients. Specifically, in insomnia, besides recommending sleep hygiene measures, medication, or psychotherapy, a series of stress management measures are routinely recommended, such as: recognizing the association between stressful events and sleeplessness; ventilating conflict and anger to avoid internalization of emotions; addressing daily worries a few hours before bedtime; becoming tolerant of occasional sleeplessness; avoiding rumination over sleep difficulty; and relaxation techniques. Increasing empirical evidence suggests that cognitive behavioral therapy (CBT) for insomnia is the most effective non-pharmacological approach in the management of chronic insomnia. Besides using educational and behavioral techniques, CBT involves cognitive restructuring that aims at identifying and altering faulty beliefs and attitudes about sleep that in themselves are believed to perpetuate the vicious cycle of insomnia, fear of sleeplessness, emotional arousal, and further sleep disturbance. However, the CBT approach does not address long-term personality patterns such as internalization of psychological conflict and inhibition of emotions associated with insomnia. In the sleep disorders of organic origin, such as narcolepsy or sleep apnea, supportive therapy, including educating and advising patients and their families gently but honestly about the nature of the chronic disorder and that the symptoms are not under voluntary control, is of utmost importance. Furthermore, helping the patient to adjust his or her lifestyle to the changes imposed by the sleep disorder, i.e., sleep or cataplectic attacks in narcolepsy, or enhancing his/her compliance to the chronic use of cumbersome and often, anxiety-provoking therapies, such as continuous positive airway pressure (CPAP) in sleep apnea, is an important dimension of the overall management of the sleep-disorder patient.

Sleep disturbance is a very frequent symptom in most of the psychiatric disorders. Depression, anxiety disorders, and acute onset schizophrenia are frequently associated with sleep disturbance. Depression, in

particular melancholia, is associated with early morning awakening, while sleep loss is the early sign of an acute onset schizophrenia or an early sign of relapse. In contrast, sleep disturbance or sleep loss can lead to mental illness. For example, it has been reported that chronic insomnia is associated with a higher incidence of major depression. Also, acute sleep loss can lead to a manic episode in a patient with bipolar disorder. Thus, it is not surprising that sleep therapy or hypnotherapy was very common in the practice of mental health professionals in the nineteenth and early twentieth centuries. Today, the improvement of the quality and quantity of sleep of a person under emotional distress is one of the therapeutic priorities of the practicing clinician.

Sleep and Stress System

In mammalian organisms, including humans, the stress system consists of components of the central nervous system (CNS), including: the corticotropin-releasing hormone (CRH) neurons of the hypothalamic paraventricular nucleus; and several, mostly norepinephric nuclei of the brain stem, and their peripheral limbs, the hypothalamic-pituitary-adrenal (HPA) axis and the peripheral autonomic system, whose main function is to maintain homeostasis, both in the resting and stress states.

Normal Sleep and Stress System

Although the association between stress and sleep has been noticed for hundreds of years, a more systematic and scientifically based approach in the relation between sleep and stress system has only existed since the 1980s. It was in 1983 that Dr. Weitzman and his colleagues reported that sleep, in particular SWS, appears to have an inhibiting influence on the HPA axis and cortisol secretion ([Figure 1](#)). Since then, several studies have replicated this finding. In a reverse mode, activation of the HPA axis and central administration of CRH can lead to arousal and sleeplessness. In normal sleepers, wakefulness and stage 1 sleep (light sleep) accompany cortisol increases, while SWS is associated with declining plasma cortisol levels. In addition, in healthy people, induced sleep disruption (continuous arousals) is associated with significant increases of plasma cortisol levels. Furthermore, mean plasma cortisol level was significantly higher in a group of subjects with a shorter total sleep time than another group with a longer total sleep time. Finally, it has been suggested that the effect of aging on the levels and diurnal variation of human adrenocorticotrophic activity could be involved in the etiology of poor sleep in the elderly.

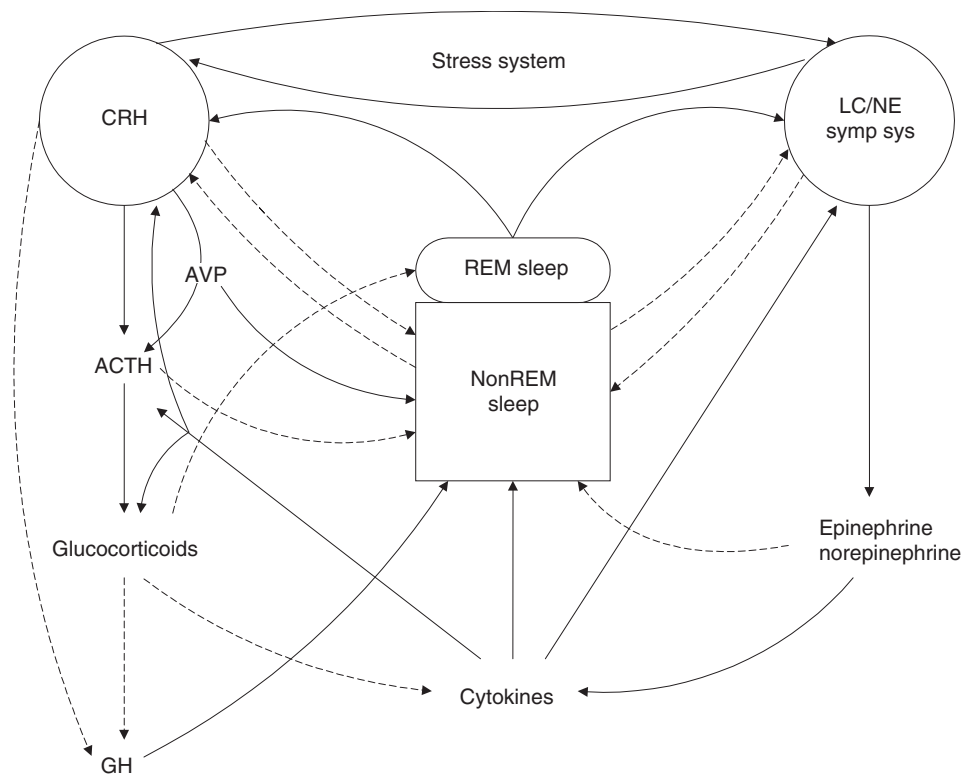


Figure 1 A simplified, heuristic model of the interactions between central and peripheral components of the stress system with sleep and REM sleep. CRH–ACTH–cortisol and LC/NE systems are two main arousal systems that their activation leads to wakefulness. NonREM sleep is associated with an inhibitory effect on these two wakefulness promoting systems, whereas during REM there is an activation of both limbs of the stress system. ACTH, corticotropin; AVP, arginine vasopressin; CRH, corticotropin-releasing hormone; GH, growth hormone; LC/NE symp sys, locus ceruleus-norepinephrine/sympathetic system. A solid line denotes promotion/stimulation; a broken line denotes disturbance/inhibition.

REM sleep, which is a state of CNS activation that resembles unconscious wakefulness (paradoxical sleep), appears to be associated with a higher activity of the HPA axis (Figure 1). An early study showed that urinary 17-hydroxy corticoids were increased during REM epochs in urological patients. One study in healthy, normal sleepers showed that the amount of REM sleep was positively correlated with 24-h urinary free-cortisol (UFC) excretion. These results are consistent with the co-existence of stress system activation, and REM sleep increases in patients with melancholic depression. In rats, chronic mild stress resulted in disruption of REM sleep, including a reduced latency to the onset of first REM period, providing support for the validity of the proposed association between stress and mechanisms underlying endogenous depression, including REM sleep alterations.

Although the effects of sleep on the sympathetic nervous system have been studied less frequently, it appears that in humans during sleep, catecholaminergic activity is decreased (Figure 1). In particular, catecholaminergic activity appears to be suppressed

during SWS, while there is an activation of the sympathetic nervous system during REM sleep. It has been found that in healthy, normal sleepers, the amount of REM sleep correlated to 24-h excretion of urinary catecholamines while sympathetic-nerve activity measured using microneurography, was elevated during REM sleep. Interestingly, in rats, locus ceruleus (LC) neurons fire most frequently during waking, less during nonREM (NREM) sleep, and are silent during paradoxical sleep.

The stress system is closely linked with immune responses. In particular, three pro-inflammatory cytokines (tumor necrosis factor- α (TNF- α), interleukin (IL)-1, and IL-6) have a strong stimulating effect on HPA axis (Figure 1). Glucocorticoids, the endproducts of the HPA axis, inhibit the production of all three cytokines, while catecholamines, the other products of the stress system, stimulate IL-6 secretion. In animals, several studies have demonstrated the potential role of these cytokines, particularly IL-1 and TNF- α in sleep regulation. In humans, these cytokines appear to be increased at nighttime. A study in young, healthy, normal sleepers showed that the amount and

quality (depth) of sleep correlated negatively with the overall daytime secretion of IL-6. These data suggest that a good night's sleep is associated with decreased secretion of IL-6, a good sense of wellbeing, and that good sleep is associated with decreased exposure of tissues to the pro-inflammatory and potentially detrimental actions of IL-6.

The role of corticotropin-releasing hormone in sleep-wake regulation CRH, which is produced and released from parvocellular neurons of the paraventricular nucleus, is the key regulator of the HPA axis. Release of CRH is followed by enhanced secretion of adrenocorticotrophic hormone (ACTH) from the anterior pituitary and cortisol from the adrenal cortex. CRH exerts various influences on behavior, including wakefulness and sleep.

Several studies in animals have shown that intracerebroventricular administration (ICV) of CRH induces increased waking. Also, specific CRH receptor blockade reduces spontaneous awakening and rat strains deficient in the synthesis and secretion of hypothalamic CRH spend less time awake than do control counterparts. Finally, CRH concentration in the brains of rats exhibits circadian fluctuations and is highest when rats are most active. Several reports indicate that CRH is excitatory in the LC, hippocampus, cerebral cortex, and some portions of the hypothalamus. Spontaneous discharge rates in the LC are highest during arousal and lowest during sleep. Also, it has been reported that the brain β_1 -adrenergic system is involved in the restraint stress-induced increase in arousal in rats.

In humans, the majority of studies suggest that the sleep of young individuals is rather resistant to the arousing effects of CRH. In contrast, middle-aged individuals responded to an equivalent dose of CRH with significantly more wakefulness and suppression of SWS compared to baseline. Based on these findings, we concluded that middle-aged men show increased vulnerability of sleep to stress hormones, possibly resulting in impairments in the quality of sleep during periods of stress. These findings suggest that changes in sleep physiology associated with middle-age play a significant role in the marked increase of prevalence of insomnia in middle-age. Also, peripheral administration of CRH is associated with REM suppression, which is stronger in the young than in the middle-aged. The administration of ACTH and its analogs in humans has been associated with general CNS activation consisting of a decreased sleep period time and sleep efficiency, and an increase of sleep latency. Continuous administration of ACTH produced a marked reduction in REM sleep.

Arginine vasopressin (AVP) is a peptide which acts as a co-factor of CRH in the activation of the HPA system. Several experiments in rats have suggested that the continuous administration of AVP causes sleep disturbance.

Steroid effects on sleep The administration of glucocorticoids has been reported to be associated with a robust suppression of REM sleep (Figure 1). Also, in some studies, it has been reported that the continuous or pulsatile nocturnal administration of cortisol paradoxically was associated with a modest increase of SWS in addition to the well-established decrease of REM sleep. The opposite effects of cortisol on SWS compared to CRH are believed to be mediated by a feedback inhibition of CRH from cortisol. In contrast, the glucocorticoid antagonist, mifepristone (RU-486), led to a significant worsening of sleep in normal controls. The administration of mineralocorticoids is not associated with significant changes of sleep structure.

A study in patients with Addison's disease recently demonstrated that an evening replacement dose of hydrocortisone was necessary for proper expression of REM sleep (*vide infra*) suggesting that glucocorticoids have some permissive action for this sleep parameter, possibly reflecting an inverse u-shaped dose-response curve.

Interestingly, in clinical practice, the use of glucocorticoids is associated with sleep disturbance. In fact, in a multicenter study in which steroids were used on a short-term basis, sleep disturbance was one of the most common side effects.

Aging, HPA axis, and sleep Old age is associated with marked sleep changes consisting of increased wakefulness, minimal amounts of SWS, declining amounts of REM sleep, and earlier retiring and rising times. Some studies have shown that older adults have elevated cortisol levels at the time of the circadian nadir and have higher basal cortisol levels than younger adults. It is difficult to discern whether the latter changes are associated with aging or increased medical morbidity common in this group. It has been suggested that the effect of aging on the levels and diurnal variation of human adrenocorticotrophic activity could be involved in the etiology of poor sleep in the elderly. Higher evening cortisol concentrations are associated with lower amounts of REM sleep and increased wakefulness. More recently, it was shown that older women without estrogen replacement therapy (ERT), when subjected to mild stress, showed greater disturbances in sleep parameters than women receiving ERT, suggesting that female hormones protect women's sleep from external stressors.

Sleep Deprivation and Stress System

If sleep is important for our sense of wellbeing, then it is conceivable that sleep deprivation represents a stress to humans and should be associated with activation of the stress system. However, several studies that have assessed the effects of one night's sleep deprivation on the HPA axis have shown that cortisol secretion is either not or minimally affected by sleep following prolonged wakefulness. Two studies have reported somewhat antithetical results, with the one study showing that cortisol secretion is elevated the next evening following sleep deprivation, and the other study showing a significant decrease of plasma cortisol levels the next day and recovery night. The study that showed decrease of plasma cortisol levels following sleep loss indicated that this inhibition of the HPA axis activity was associated with an enhanced activity of the growth hormone axis. The finding that sleep deprivation leads to lower cortisol levels postdeprivation (primarily during the subsequent night of sleep) suggests that lowering the level of HPA activity, which is increased in depression, may be the mechanism through which sleep deprivation improves the mood of depressed individuals. Prolonged sleep deprivation that results in death is associated with increased plasma norepinephrine levels, and higher ACTH and corticosteroid levels at the later phase of sleep deprivation. It is postulated that these increases are due to the stress of dying rather than to sleep loss.

A recent study demonstrated that a week of partial sleep restriction did not affect significantly the 24-h cortisol secretion either in men or women. However, there was a significant effect in terms of the circadian secretory pattern of the hormone as indicated by the significant time effect, and the time and gender interaction effect. Specifically, the peak cortisol secretion in the morning was lower after sleep restriction than baseline and this difference was stronger in men than in women. In both genders, the peak of cortisol secretion was shifted by 2 h earlier than baseline (08.00 versus 06.00). Sleep restriction did not affect significantly the nadir value of cortisol secretion, neither the afternoon, nor evening presleep levels.

The inconsistency in cortisol secretion after total or partial sleep loss may be explained by methodological differences in these studies, i.e., nonstressful deprivation is not associated with increased cortisol secretion. Also, the higher evening cortisol levels reported in studies where sleep was curtailed at the beginning of the night may represent a change of the 24-h-secretory pattern, e.g., a shift of the nadir point later at night.

Although sleep deprivation appears to have beneficial effects on the mood of depressed individuals, in

healthy individuals, sleep deprivation leads to next-day fatigue and somnolence, decreased concentration and attention, and increased vulnerability to accidents. These effects appear to be mediated partially by the increased levels of IL-6, which is related to the stress system. It has been shown that following sleep deprivation, there is an increase of IL-6 levels during the daytime and a decrease during the first night of sleep following sleep deprivation. The increased levels of IL-6 during the daytime can explain the somnolence and fatigue experienced by the individuals after one night's sleep loss, whereas the nighttime undersecretion of IL-6 might be responsible for the better quality (depth) of their sleep. The exact mechanisms of this elevation are not known. However, alterations of the stress system following sleep deprivation may be responsible for this elevation. For example, the decreased levels of cortisol postdeprivation and/or possible increase of catecholaminergic activity can lead to elevation of plasma IL-6 levels. It has been shown that sleep deprivation in humans appears to enhance norepinephrine and dopamine activity.

Sleep Disorders and Stress System

Although sleep disorders/disturbances with their various physical and mental effects on the individual should be expected to affect the stress system, information regarding the effects of sleep disorders/disturbances on this system is lacking.

Insomnia and stress system Insomnia, a symptom of various psychiatric or medical disorders, may also be the result of an environmental disturbance or a stressful situation. When insomnia is chronic and severe, it may itself become a stress that affects the patient's life so greatly that it is perceived by the patient as a distinct disorder itself. Either way, as a manifestation of stress or a stressor itself, insomnia is expected to be related to the stress system. There is a paucity of data regarding the activity of the stress system in insomniacs. Few studies have measured cortisol levels in poor sleepers or insomniacs, and those results are inconsistent. The majority of these studies reported no difference between controls and poor sleepers in 24-h cortisol or 17-hydroxy steroid excretion. In a study in 15 young adults with insomnia, it was demonstrated that 24-h UFC levels were positively correlated with total wake time ([Table 1](#)). In addition, 24-h urinary levels of catecholamines and their metabolites, dihydroxyphenyl glycol (DHPG), and 3,4-dihydroxy-phenylacetic acid (DOPAC) were positively correlated with percent stage 1 sleep and wake time after sleep onset. Norepinephrine tended to correlate positively with percent stage 1 sleep and

Table 1 Sleep disorders and the stress system

Insomnia	Increased 24-h secretion of ACTH-cortisol, more intense in the evening; positive correlations between 24-h UFC and urinary catecholamines and polysomnographic indices of severity of the disorder
Sleep apnea	Increased excretion of urinary catecholamines; increased nighttime plasma catecholamines; increased daytime plasma IL-6 and TNF- α levels; mild hypercortisolemia compared to obese controls
Narcolepsy	Increased daytime plasma IL-6 and TNF- α levels
Idiopathic hypersomnia	Mild hypocortisolism; enhanced ACTH secretion to exogenous CRH; central CRH deficiency

ACTH, adrenocorticotrophic hormone; CRH, corticotropin-releasing hormone; IL-6, interleukin-6; TNF- α , tumor-necrosis factor- α ; UFC, urinary free-cortisol.

wake time after sleep onset and negatively with percent SWS. It was concluded that in chronic insomnia, the activity of both limbs of the stress system, i.e., the HPA axis and the sympathetic system, relates positively to the degree of objective sleep disturbance. However, in this study also, the total amount of the 24-h UFC or catecholamines excretion was not different from normative values. The results on the positive correlation between urinary catecholamines and polysomnographic indices of sleep disturbance in chronic insomnia are consistent with previous studies that showed elevated urinary catecholamines in chronically stressed subjects who experience sleep disturbance or in healthy subjects whose sleep was affected by nocturnal aircraft noise.

In a recent controlled study in which objective sleep testing and frequent blood sampling was employed, 24-h ACTH and cortisol plasma concentrations were significantly higher in people with insomnia than matched healthy controls. Within the 24-h period, the greatest elevations were observed in the evening and first half of the night. Also, people with insomnia with a high degree of objective sleep disturbance (sleep time <70%) compared to those with a low degree of sleep disturbance secreted a higher amount of cortisol. Pulsatile analysis revealed a significantly higher number of peaks per 24 h in insomniacs than in controls ($p < 0.05$), while Cosinor analysis showed no differences in the circadian pattern of ACTH or cortisol secretion between people with insomnia and controls. Thus, insomnia is associated with an overall increase of ACTH and cortisol secretion, which, however, retains a normal circadian pattern. Also, this increase relates positively to the degree of objective sleep disturbance. These findings are consistent with a disorder of CNS hyperarousal rather than one of sleep loss, which is usually associated with no change or a decrease in cortisol secretion, or a circadian disturbance. Two recent studies found that cortisol was elevated in primary insomnia with objectively documented sleep disturbance during the evening and nighttime. However, another study found that people with insomnia without evidence of objective

sleep disturbance did not report differences from controls. These findings suggest that objective sleep disturbance may be a useful index of the biological severity of the disorder.

Chronic activation or dysregulation of the stress system may play a significant role in the poor mental and physical health associated with chronic insomnia. Insomnia is associated with a higher risk of development of psychiatric disorders, such as major depression and anxiety disorders. Also, people with insomnia tend to have more health complaints than normal controls, both as children and adults, and they do appear to have poorer health in other respects. Psychosomatic-type illnesses, such as allergy, asthma, colitis, hypertension, migraine headaches, and ulcers, are reported by about twofold more people with insomnia than controls. Consistent with the above finding, a recent study that used a large general random sample of men and women (1754 people) reported that insomnia is independently associated with medical disorders such as hypertension. Also, nonspecific physical symptoms, such as headache, diarrhea, constipation, stomach discomfort, palpitations, shortness of breath, nonspecific pain, tiredness, and weakness, are also more common in people with insomnia. It can be hypothesized that these physical complaints, including those labeled as psychosomatic, reflect the chronic effects of an activation/dysregulation of the stress system in various body organs and functions. Furthermore, these data suggest that the therapeutic goal in insomnia should not be just to improve the quality or quantity of nighttime sleep. Rather, they suggest that the common practice of prescribing hypnotics only for patients with chronic insomnia at most is of limited efficacy. Sedative antidepressants alone or in combination with the newer antidepressants, e.g., selective serotonin reuptake inhibitors (SSRIs), appear to have a normalizing effect on measures, i.e., cortisol. Furthermore, the focus of psychotherapeutic and behavioral modalities, including sleep hygiene measures, should not be to just improve the emotional and physiological state of the people with insomnia

pre- or during sleep, but rather to decrease the overall emotional and physiologic hyperarousal and its underlying factors, present throughout the 24 h sleep-wake period. Finally, objective measurements of sleep using simpler methods such as actigraphy may prove useful in predicting the severity of the disorders and, therefore, the urgency for medical intervention.

Disorders of excessive daytime sleepiness and stress system Obstructive sleep apnea, the most common sleep disorder associated with excessive daytime sleepiness and fatigue, is associated with nocturnal hypoxia and sleep fragmentation. The latter conditions should be expected to be associated with an activation of the stress system. Indeed, it has been shown that urinary catecholamines, as well as plasma catecholamines measured during the nighttime, are elevated in people with sleep apnea compared to controls (Table 1). Also, using microneurography, it has been shown that obstructive apneas are associated with a surge of sympathetic nerve activity. It has been proposed that sympathetic activation in sleep apnea is one of the mechanisms leading to the development of hypertension, a condition commonly associated with sleep apnea. With similar logic, it was expected that sleep apnea would be associated with an activation of the HPA axis also. However, the few studies that have assessed the plasma cortisol levels in people with sleep apnea have failed to show any differences from controls. Also, no differences were reported in the excretion of cortisol following the abrupt withdrawal of CPAP, a definitive treatment for sleep apnea. However, whether these results represent an adaptation of the HPA axis to a chronic physical stress or are a result of incomplete assessment of the HPA axis, remains open.

Narcolepsy and idiopathic hypersomnia are two primary sleep disorders of bipolar disorders currently treated with medication such as stimulants and modafinil. The association of the two other organic disorders of excessive daytime sleepiness, narcolepsy and idiopathic hypersomnia, with the stress system has not been assessed. Preliminary data from a study that assessed the responsiveness of plasma cortisol and ACTH to the exogenous administration of CRH in individuals with idiopathic hypersomnia indicated reduced 24-h UFC levels compared to controls while ACTH response to CRH tended to be significantly higher in the patients compared to controls (Table 1). These preliminary findings suggest a subtle hypocortisolism and inferred central CRH deficiency in patients with idiopathic hypersomnia. A central CRH deficiency in these patients is consistent with

their clinical profile of increased daytime sleepiness and deep nocturnal sleep (generalized hypoarousal).

The stress related cytokines (TNF- α , IL-1, and IL-6) have been suggested as potential mediators of excessive daytime sleepiness (Table 1). In 1997, a report on the morning plasma levels of TNF- α , IL-1 β , and IL-6 in patients with sleep apnea, narcolepsy and idiopathic hypersomnia showed that TNF- α was significantly elevated in people with sleep apnea and narcolepsy compared to that in normal controls. Plasma IL-1 β concentrations were not different between the sleep disorder patients and controls, whereas IL-6 was markedly and significantly elevated in patients with sleep apnea compared to that in normal controls. The primary factor in the elevation of plasma TNF- α values was the degree of nocturnal sleep disturbance, whereas the primary determinant for IL-6 levels was body mass index. These findings suggest that TNF- α and IL-6 might play a significant role in mediating sleepiness and fatigue in disorders of excessive daytime sleepiness in humans. Also, it was hypothesized that IL-6 plays a role in mediating sleepiness in obese patients, even those without obstructive sleep apnea, who experience, as confirmed also in the sleep laboratory, a significant degree of daytime sleepiness and fatigue. In 1996, another study reported that the circadian rhythm of TNF- α was significantly altered in patients with sleep apnea and that the peak concentration that occurred during the night in normal control subjects was not present in patients. Rather, the patients exhibited increased TNF- α concentrations in the afternoon, the time period during which concentrations in normal control subjects were at a minimum. Furthermore, in this study, it was reported that the alteration in the cytokine profiles were not normalized by the CPAP therapy. Since then, several studies have confirmed that TNF- α and IL-6 are elevated in disorders of excessive daytime sleepiness, i.e., sleep apnea and narcolepsy independently of confounding variables such as obesity. Interestingly, a recent study reported that neutralizing TNF- α is associated with a significant decrease of sleepiness in patients with sleep apnea.

Conclusion

In conclusion, the collective findings on the association between sleep and stress system appear to suggest that sleep is a major antistress mechanism that is vital in the preservation of the physical and mental homeostasis of the human species. Also, sleep loss appears to be a physical and mental stress to humans, whereas sleep disorders/disturbances may lead to a significant dysregulation of the stress system. More knowledge

on the association between sleep, sleep disorders, and stress system would lead to a better understanding of the mechanisms that underlie sleep's capacities for physical and mental restoration and to improve treatment of sleep disorders/disturbances.

Further Reading

- Chrousos, G. P. (1995). The hypothalamic-pituitary-adrenal axis and immune-mediated inflammation. *New England Journal of Medicine* 332, 1351–1362.
- Chrousos, G. P. and Gold, P. W. (1992). The concepts of stress and stress system disorders. *JAMA* 267, 1244–1252.
- Edinger, J. D. and Means, M. K. (2005). Cognitive-behavioral therapy for primary insomnia. *Clinical Psychology Review* 25, 539–558.
- García-Borreguero, D., Wehr, T. A., Larrosa, O., et al. (2000). Glucocorticoid replacement is permissive for rapid eye movement sleep and sleep consolidation in patients with adrenal insufficiency. *Journal of Clinical Endocrinology and Metabolism* 85, 4201–4206.
- Kales, A. and Kales, J. (1984). *Evaluation and treatment of insomnia*. New York: Oxford University Press.
- Krueger, J. M. and Obál, F. (1994). Sleep factors. In: Saunders, N. A. & Sullivan, C. E. (eds.) *Sleep and breathing*, pp. 79–112. New York: Dekker.
- Opp, M. R. (1995). Corticotropin-releasing hormone involvement in stressor-induced alterations in sleep and in the regulation of waking. *Immunology* 5, 127–143.
- Rechtschaffen, A. and Kales, A. (1968). *A manual of standardized terminology, techniques, and scoring system for sleep stages of human subjects*. NIH Report No. 204. Bethesda, MD: National Institutes of Health.
- Rechtschaffen, A., Bergmann, B. M., Everson, C. A., et al. (1989). Sleep deprivation in the rat: X. integration and discussion of the findings. *Sleep* 12, 68–87.
- Rodenbeck, A., Heuther, G., Ruther, E., et al. (2002). Interactions between evening and nocturnal cortisol secretion and sleep parameters in patients with severe chronic primary insomnia. *Neuroscience Letters* 324, 154–163.
- Rodenbeck, A., Cohrs, S., Jordan, W., et al. (2003). The sleep-improving effects of doxepin are paralleled by a normalized plasma cortisol secretion in primary insomnia. A placebo-controlled, double-blind, randomized, cross-over study followed by an open treatment over 3 weeks. *Psychopharmacology* 170, 423–428.
- Somers, V. K., Dyken, M. E., Mark, A. L., et al. (1993). Sympathetic-nerve activity during sleep in normal subjects. *New England Journal of Medicine* 328, 303–307.
- Steiger, A. and Holsboer, F. (1997). Neuropeptides and human sleep. *Sleep* 20, 1038–1052.
- Steiger, A., Antonijevic, I. A., Bohlhalter, S., et al. (1998). Effects of hormones on sleep. *Hormone Research* 49, 125–130.
- Van Cauter, E. (2005). Endocrine physiology. In: Kryger, M. H., Roth, T. & Dement, W. C. (eds.) *Principles and practices of sleep medicine* (4th ed., pp. 266–282). Philadelphia, PA: Elsevier Saunders.
- VanCauter, E., Leproult, R. and Plat, L. (2000). Age-related changes in slow wave sleep and REM sleep and relationship with growth hormone and cortisol levels in healthy men. *JAMA* 284, 861–868.
- Vgontzas, A. N., Papanicolaou, D. A., Bixler, E. O., et al. (1997). Elevation of plasma cytokines in disorders of excessive daytime sleepiness: role of sleep disturbance and obesity. *Journal of Clinical Endocrinology and Metabolism* 82, 1313–1316.
- Vgontzas, A. N. and Kales, A. (1999). Sleep and its disorders. *Annual Review of Medicine* 50, 387–400.
- Vgontzas, A. N., Bixler, E. O., Lin, H.-M., et al. (2001). Chronic insomnia is associated with nyctohemeral activation of the hypothalamic-pituitary-adrenal axis: clinical implications. *Journal of Clinical Endocrinology and Metabolism* 86, 3787–3794.
- Vgontzas, A. N., Zoumakis, M., Bixler, E. O., et al. (2003). Impaired nighttime sleep in healthy old vs. young adults is associated with elevated plasma IL-6 and cortisol levels: physiologic and therapeutic implications. *Journal of Clinical Endocrinology and Metabolism* 88, 2087–2095.
- Vgontzas, A. N., Zoumakis, M., Bixler, E. O., et al. (2004). Adverse effects of modest sleep restriction on sleepiness, performance, and inflammatory cytokines. *Journal of Clinical Endocrinology and Metabolism* 89, 2119–2126.

Smoking and Stress

F J McClernon

Duke University Medical Center, Durham, NC, USA

D G Gilbert

Southern Illinois University-Carbondale, Carbondale, IL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D G Gilbert and F J McClernon, volume 3, pp 458–465, © 2000, Elsevier Inc.

*Situation x
trait adaptive
response
(STAR) model*

An integrative model that emphasizes the interaction of situational factors with genetic and personality factors in determining who is likely to become a smoker, how rewarding smoking will be, and when and how often smoking is likely to be rewarding to the individual as a coping tool.

Tar

The mass of some 5000 components of tobacco smoke, excluding nicotine and gases (e.g., carbon monoxide).

Introduction

Personality Traits, Genetic, and Biological Predictors of Smoking and Stress

Smoking as a Coping Response

Smoking Cessation, Stress, and Coping

The Psychobiology of Smoking, Coping, and Stress
Diathesis

Summary

Glossary

<i>Acetylcholine</i>	A neurotransmitter in the brain that has receptors on nerve cells that nicotine stimulates and promotes the firing of.
<i>Affect</i>	Moods, emotions, and feelings, including negative affect (e.g., anxiety, depressive mood, and anger) and positive affect (e.g., euphoria, vigor, and pleasant feelings).
<i>Cognitive dysfunction</i>	The impairment or biasing of information processes such as perceptual, attentional, and memorial functions that can be caused by temporary physiological and mood states, as well as by longer-lasting forms of psychopathology, neurological abnormality, or temperamental bias.
<i>Dopamine</i>	A neurotransmitter in the brain having a number of functions, including the enhancement of attention and signaling of reward in response to appetitive stimuli.
<i>Neuro-transmitter</i>	A chemical released at the synapse (junction) between nerves that promotes the conduction of nerves impulses from one nerve to another.
<i>Nicotine</i>	The primary psychoactive component of tobacco smoke; it binds to endogenous nicotinic acetylcholine receptors, which results in a variety of neuronal actions.

Introduction

Smoking remains the number one preventable cause of death and disability in the United States and is a burden on people, economies, and health-care systems worldwide. Knowledge about the relation of smoking and stress has expanded exponentially in the last 3 decades as a result of new findings in the fields of genetics, neuroscience, and clinical research. This research has revolutionized thinking about who smokes and why it is so difficult for many to quit. Genes have been found to play a significant role in determining who becomes and remains a smoker. Some of the genes that predispose some people to smoke predispose the same individuals to experience higher than normal levels of stress and negative mood states. These and related findings indicate that there are many clinically and theoretically important relations among smoking, stress, and coping.

Personality Traits, Genetic, and Biological Predictors of Smoking and Stress

Evidence shows clear and substantial relations among smoking, stress-related personality traits, and psychopathology. Further, all these characteristics are largely heritable (passed on from generation to generation via genes) and share common genetic bases.

Stress, Personality, Psychopathology, and Smoking

Individuals with specific psychiatric problems and personality traits leading to frequent or intense personal distress are substantially more likely to smoke. Smoking rates are disproportionately high in individuals with psychiatric conditions, including major depressive disorder, various anxiety disorders, schizophrenia,

antisocial personality disorder, attention deficit hyperactivity disorder (ADHD), and various substance abuse disorders, including alcoholism. The prevalence of smoking in individuals with a lifetime history of these conditions is two to three times higher than in unaffected individuals. The personality traits associated with the increased prevalence of smoking reflect dispositions that are similar in nature to psychiatric disorders. For example, the personality trait of neuroticism (the general tendency to experience negative affect and stress) has been found to be associated with higher prevalence of smoking. Smoking (and drug abuse more broadly) is also much more prevalent among individuals high in impulsiveness or sensation-seeking traits. In addition to finding relations between these traits and smoking among adults, longitudinal studies have shown that children and adolescents with these traits are more likely to smoke as adults (as well as to use other drugs). Taken together, these data suggest that smoking behavior is not randomly distributed in the adult population and also support the notion that nicotine plays some role in the psychological life of the smoker beyond simple addiction or withdrawal alleviation.

Identical Genes Related to Smoking, Psychopathology, and Personality Traits

Results from family, twin, and molecular genetic studies demonstrate that genetic factors play an important role in smoking behavior and stress responses. Genes account for approximately half of the variability in who will smoke and who will not. Further, many of the same genes that predispose individuals to smoke are also responsible, in part, for predisposing individuals to the stress-related personality traits and psychiatric disorders previously mentioned. For instance, the same genes that contribute to a vulnerability to neuroticism also appear to dispose individuals to clinical depression and to smoking. Similarly, the same genes that contribute to impulsivity and sensation seeking also contribute to a variety of compulsive behaviors, including drug abuse, antisocial behavior, and smoking.

The identification of some of the many genes that contribute to smoking and smoking-related traits has been one of the major accomplishments of recent years. Genetically influenced biological differences between people in (1) brain neurotransmitter receptors and (2) liver enzymes related to nicotine metabolism appear to play roles in determining smoking. Some of the most exciting insights in this area are related to the neurotransmitter dopamine. Dopamine is intimately associated with coping, stress, natural reward, and the pleasurable effects of abused drugs, including nicotine. A number of genes combine to

determine a person's brain dopamine system functioning. Those individuals with low dopaminergic system functioning levels are more likely than others to smoke, abuse alcohol and other drugs, and have impulsive and sensation-seeking personality traits. Thus, these dopamine genes appear to contribute to the association of smoking with these personality traits and substance-abuse patterns.

Smoking as a Coping Response

These personality traits and psychiatric conditions are associated with distressing cognitive and emotional states. To the degree smoking is effective at ameliorating these undesirable cognitive or affective states, coping by means of smoking will occur and will be experienced as rewarding by a smoker. It is not surprising, then, that those individuals who frequently experience cognitive and affective states that can be improved by smoking are more likely than others to find smoking to be a useful and reinforcing tool (and thus smoke more often).

Although smokers as a group report smoking to reduce negative affect; cope with stress; reduce weight or appetite; and enhance stimulation, cognitive performance, and pleasure, the relative importance of these smoking motivations vary across (and also within) smokers. Consistent with the coping model, those who are depression prone are more likely to say they smoke to reduce depressive moods and thoughts, whereas those with attentional problems say they smoke more often to enhance attention and, in turn, those who are prone to anxiety are more likely to say they smoke to alleviate anxiety. Similarly, individuals who are motivated to keep from gaining weight value the fact that smoking helps maintain a lower body weight. Thus, the motivations to smoke are to some extent consistent with personality traits of the smokers. This tendency to smoke for reasons related to personality traits is called trait-adaptive responding because smoking reflects an adaptive coping response related to the smoker's specific personality traits (see [Figure 1](#)).

Smoking/Nicotine as a Coping Response

Withdrawal alleviation As previously suggested, one theory proposes that smoking alleviates stress only to the extent that nicotine withdrawal is stressful to smokers. These withdrawal alleviation theories are supported by the simple fact that smokers find abstinence from smoking stressful and smoking alleviates this stress. A related theory suggests that smokers so come to associate withdrawal-induced stress with negative states that any negative state – whether during withdrawal or not – cues smoking behavior.

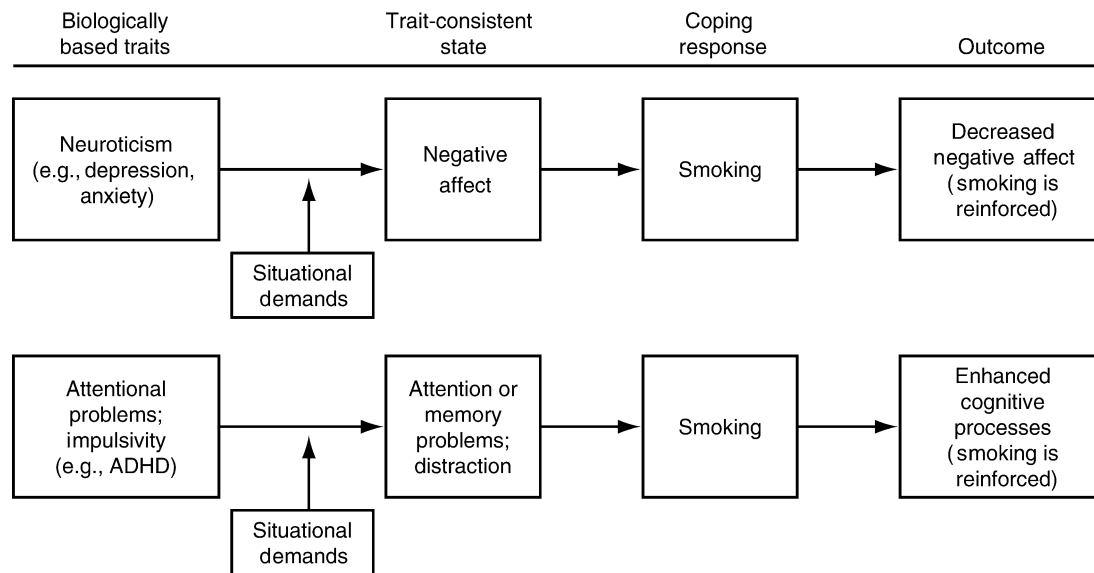


Figure 1 Biologically based traits increase the frequency and severity of trait-consistent states. To the extent that these states are improved by smoking/nicotine, smoking is reinforced.

Although withdrawal alleviation models of smoking and stress are compelling, they cannot provide a complete picture of nicotine dependence. The withdrawal alleviation model suggests the stress of withdrawal is temporary; once withdrawal subsides the now-ex-smokers will return to their presmoking affective and cognitive states. Indeed, many studies have reported temporary increases in negative affect, stress, and difficulty concentrating following quitting that return to or improve beyond prequit levels in a matter of weeks. However, these studies have been limited by a number of factors that might preclude accurate characterization of withdrawal in smokers. First, because these studies do not take into account data of smokers who relapse (i.e., smokers for whom quitting is most difficult), the findings minimize the negative affect and distress the majority of smokers experience on quitting. Second, these studies have often not included a continuing-to-smoke control group, so other factors affecting mood and cognition, such as the passage of time, are also not accounted for.

When these two limitations are eliminated from research designs, a different picture of withdrawal emerges. Recent longitudinal studies that minimize relapse and include adequate control groups have observed transient increases in negative affect, stress, and difficulty concentrating that do not return to baseline levels, or the approximate levels of these symptoms in a continuing-to-smoke control group. These findings suggest that smokers may smoke, in part, to reduce underlying distressful cognitive and affective states that were present prior to becoming a

smoker, which, on quitting smoking, are rerevealed. Thus, smokers may smoke in order to alleviate withdrawal symptoms to a degree, but also smoke in order to self-medicate underlying trait-based problems. It should be noted, however, that well-designed studies to date have accurately assessed the effects of quitting smoking on negative affect for only 45 days of abstinence. Thus, it is possible that after a longer period of time effects of quitting do resolve in most quitters.

Absolute inherent effects nicotine A growing body of research suggests that nicotine can, under specific situations, improve mood and cognitive performance, even in smokers who are only minimally deprived of nicotine or in nonsmokers. Laboratory studies show that smoking and other forms of nicotine delivery usually reduce negative moods and emotions in situations in which stressors are ambiguous and/or not dominant, such as situations in which an individual is anticipating experiencing stress. In contrast, when the context includes clear, immediate, and potent threat stimuli, negative moods are not impacted by smoking or nicotine.

Many experts and some experimental evidence suggest that nicotine can also indirectly promote positive affect and minimize negative affect in situations in which nicotine contributes to goal achievement. For example, when they need to concentrate for a long period of time to accomplish a goal, such as writing a paper, smokers say they smoke to sustain their concentration and thereby achieve their goal. Laboratory studies have also shown nicotine to facilitate cognitive functions, including vigilance and memory,

in both habitual tobacco users and in individuals with no prior exposure to nicotine. In addition, other studies have shown that nicotine may improve affect and reduce stress by helping smokers focus on emotionally positive or neutral distractors. Research suggests that nicotine is especially effective at reducing stress when there is a substantial degree of choice as to whether or not to attend to stressful cues and when some of these nonstressful cues are emotionally positive.

A two-factor model of smoking A combined two-factor model proposes that smokers smoke in order to (1) prevent or alleviate withdrawal and (2) gain beneficial effects inherent in nicotine. Among heavy smokers, the acute effects of abstinence resolve across the first week or two of abstinence; thereafter, underlying psychological and pathological traits are rerevealed and greater stress, negative affect, and difficulty concentrating are experienced. Similarly, the attention and mood-enhancing effects of most habitually used stimulant drugs (caffeine, amphetamine, and cocaine) appear to reflect the combined influence of such inherent and withdrawal-alleviation mechanisms.

Habit, sensory aspects, and other factors Smokers may smoke for numerous reasons other than those already described. Many smokers find the sensory aspects of smoking pleasurable (e.g., the feel of smoke in lungs and the taste), and these sensory cues further reinforce smoking behavior. Further, many smokers, regardless of why they started smoking in the first place, may continue to smoke, largely out of habit. Finally, recent evidence suggests that nonnicotine chemicals in tobacco smoke may have anti-depressant qualities that may further ameliorate negative affect and reinforce smoking behavior.

Smoking Cessation, Stress, and Coping

Based on the empirical evidence showing that smoking cessation is frequently a stressful process resulting in more coping failures than successes, clinicians and researchers have spent much time and energy on minimizing smoking-cessation-related stress and maximizing coping.

Psychological and Behavioral Therapies for Smoking Cessation

Most of the widely used smoking cessation programs include several hours of skills training in cognitive and behavioral techniques designed to help the smoker cope not only with the cravings to smoke but also

with negative mood states frequently associated with quitting and known to promote relapse to smoking. Such techniques include behavioral relaxation training and learning to think about stressful situations in a manner that is less likely to produce negative emotional states. Currently popular pharmacological treatments, such as nicotine gum, nicotine patches, and the antidepressant drug Zyban, also include a complex therapeutic package that includes behavioral relaxation and coping components, along with information about the cessation process and social support.

Antidepressant Medication Therapies for Smoking Cessation

Consistent with the view that smoking motivation and the effects of smoking cessation are mediated by brain mechanisms related to negative affect and depression, one antidepressant medication, Zyban, has been found to be beneficial in helping individuals maintain smoking abstinence. This U.S. Food and Drug Administration (FDA)-approved smoking-cessation drug is also marketed under a different brand name as an antidepressant medication.

The Psychobiology of Smoking, Coping, and Stress Diasthesis

During withdrawal from nicotine, smokers with depressive tendencies have been found to exhibit brain wave (via electroencephalogram, EEG) asymmetry profiles characteristic of depressed individuals. These abstinence-induced EEG asymmetries appear to last for at least a month after quitting. The EEG asymmetry seen in these individuals is a relatively greater EEG slowing of the left (compared to the right) frontal cortex subsequent to quitting. Left-hemisphere processing is associated with active coping, verbal behavior, and positive affect, whereas right-hemisphere processing is more associated with negative mood states and withdrawal. These and related findings suggest that nicotine may act as an antidepressant in part because in individuals prone to depression it tends to increase the activation more in the left than in the right frontal cortex. Overall, nicotine tends to increase brain activation as measured by EEG, whereas smoking abstinence tends to decrease EEG activity in both hemispheres, although to a greater extent in the left relative than in the right frontal cortex. This decrease in EEG activation corresponds to the sustained decrease in ability to concentrate also found to last for at least a month after quitting smoking.

The EEG-activating effects of smoking are thought to reflect the effects of nicotine on brain receptors for the neurotransmitter acetylcholine. Nicotine has the ability to stimulate cholinergic receptors throughout the body. The stimulation of these cholinergic receptors can result in a cascade of events, including increases in heart rate, blood pressure, blood cortisol concentrations and the activation of other neurotransmitter systems in the brain. Another important brain system influenced by nicotine uses dopamine as its neurotransmitter. Dopamine (DA) is associated with brain regions involved in attention, emotion, coping, and reward. DA is involved in coping in both positive and negative emotional states. Nicotine-induced increments in brain mesolimbic DA activity may increase the ability of the organism to avoid aversive stimuli and to obtain positive ones.

The effects of smoking on EEG, mood, and attention discussed so far have been found to be due primarily to the effects of nicotine; however, recent evidence suggests that an unidentified component of tobacco smoke (other than nicotine) inhibits monoamine oxidase (MAO) activity in the brain in the same manner as some antidepressant medications. Thus, it may be that some of smoking's antidepressant, mood-enhancing, and cognitive benefits come from the inhibition of MAO activity by something other than nicotine.

Summary

Environment, genetics, personality, and psychopathologic traits interact to determine smoker status and the type of smoker an individual becomes. Individuals genetically and otherwise predisposed to negative affect, sensation seeking, and/or suboptimal cognitive performance are more likely to be exposed to smoking environments, start smoking in order to seek negative affect relief or as an act of sensation seeking, and find smoking reinforcing. The widely held belief among smokers and nonsmokers that smoking reduces negative affect and enhances cognition leads predisposed individuals to experiment with smoking as a form of self-medication, to minimize negative affect, and to maximize cognitive functioning. The affect-modulating and cognitive-enhancing effects of nicotine are more reinforcing to those who need and value them the most – those in stressful circumstances and those with a vulnerability to cognitive and affective dysfunction. When individuals who are predisposed to smoking try to quit, they lose the beneficial effects of nicotine and are especially vulnerable to experiencing the cognitive or affective state to which they are predisposed. Thus, depression-prone

individuals are more likely to become depressed, anxiety-prone individuals are more prone to anxiety, and attentionally disordered quitters are disposed to suffer more severe attention problems. The enhanced severity of cognitive and affective responses to quitting contributes to the failure of vulnerable individuals to maintain abstinence after efforts to quit smoking. Smoking in individuals who are vulnerable to smoking can thus be seen as an attempt to adapt to their situation by altering their cognitive and/or affective state.

However, affective and cognitive traits also alter the probability of smoking by more indirect mechanisms. Both affective and cognitive traits influence long-term goals by a number of indirect mechanisms. For example, risk-taking and other reward and sensation-seeking behaviors make certain individuals temperamentally insensitive to cues of potential punishment (e.g., health warnings about smoking) and promote cognitive, social, and behavioral sequences that reinforce smoking and other risk-taking behavior. The sensation-seeking life style, which includes smoking, drinking, and a focus on pleasure, may be so reinforcing to individuals with a strong genetic disposition that their cognitive motivational schemas include no plans to quit any risky behaviors. In contrast, negative-affect-prone individuals are likely to focus on the immediate relief of negative affect produced by smoking rather than on the long-term health consequences. Individuals who are more psychobiologically stable will see less benefit from smoking and be less inclined to take up or continue smoking because the perceived risks outweigh the perceived gains.

See Also the Following Articles

Alcohol and Stress: Social and Psychological Aspects; Alcohol, Alcoholism and Stress: A Psychobiological Perspective; Dopamine, Central; Drug Use and Abuse; Health Behavior and Stress; Comorbid Disorders and Stress.

Further Reading

- Comings, D. E., Ferry, L., Bradshaw-Robinson, S., et al. (1996). The dopamine D2 receptor (DRD2) gene: a genetic risk factor in smoking. *Pharmacogenetics* 6, 73–79.
- Gilbert, D. G. (1995). *Smoking: individual differences, psychopathology, and emotion*. Washington, DC: Taylor and Francis.
- Gilbert, D. G. (1997). The situation x trait adaptive response (STAR) model of substance use and craving. *Human Psychopharmacology: Clinical and Experimental* 12, S89–S102.

- Gilbert, D. G., McClernon, F. J., Rabinovich, N. E., et al. (1998). Effects of smoking abstinence on mood and craving in men: influences of negative-affect-related personality traits, habitual nicotine intake, and repeated measurements. *Personality and Individual Differences* 25, 399–423.
- Gilbert, D. G., McClernon, F. J., Rabinovich, N. E., et al. (2002). Mood disturbance fails to resolve across 31 days of cigarette abstinence in women. *Journal of Consulting & Clinical Psychology* 70, 142–152.
- Gilbert, D. G., McClernon, F. J., Rabinovich, N. E., et al. (2004). Effects of quitting smoking on EEG activation and attention last for more than 31 days and are more severe with stress, dependence, DRD2 A1 allele, and depressive traits. *Nicotine & Tobacco Research* 6, 249–267.
- Kandel, D. B. and Davies, M. (1986). Adult sequelae of adolescent depressive symptoms. *Archives of General Psychiatry* 43, 255–262.
- Kassel, J. D. (1997). Smoking and attention: a review and reformulation of the stimulus-filter hypothesis. *Clinical Psychology Review* 17, 451–478.
- Kassel, J. D., Stroud, L. R. and Paronis, C. A. (2003). Smoking, stress, and negative affect: correlation, causation, and context across stages of smoking. *Psychological Bulletin* 129, 270–304.
- Kendler, K. S., Neale, M. C., MacLean, C. J., et al. (1993). Smoking and major depression: a causal analysis. *Archives of General Psychiatry* 50, 36–43.
- Levin, E. D., Conners, C. K., Sparrow, E., et al. (1996). Nicotine effects on adults with attention-deficit/hyperactivity disorder. *Psychopharmacology* 123, 55–63.
- Scarr, S. and McCartney, K. (1983). How people make their own environments: a theory of genotype x environment effects. *Child Development* 54, 424–435.
- Shiffman, S. A. (1986). Cluster analytic classification of smoking relapse episodes. *Addictive Behaviors* 11, 295–307.

Social Capital

I Kawachi

Harvard School of Public Health, Boston, MA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by I Kawachi, volume 3, pp 466–467, © 2000, Elsevier Inc.

Forms and Functions of Social Capital

Measurement of Social Capital

Social Capital and Population Health

Mechanisms Linking Social Capital and Health

Negative Effects of Social Capital

Glossary

<i>Bonding social capital</i>	Social connections that exist within homogeneous groups, for example the strong ties that connect family members, close friends, or neighbors who are alike with respect to race/ethnicity, social class, and religion.
<i>Bridging social capital</i>	Social connections that cross race/ethnic, class, and other boundaries.
<i>Social capital</i>	The resources available to individuals and groups through their social connections and membership in community networks.

Forms and Functions of Social Capital

Social capital is the resources available to individuals through their social connections and membership in community networks. In contrast to financial capital, which resides in people's bank accounts, or human capital, which is embodied in individuals' investments in education and job training, social capital inheres in the structure and quality of social relationships between individuals.

The concept was originally developed in sociology and political science. According to James Coleman, social capital can take on several different forms, such as levels of trust within a social structure, norms and sanctions, information channels, and appropriable social organizations. An example of an appropriable social organization is the case of a residents' association in an urban housing project that formed initially for the purpose of pressuring builders to fix various problems (e.g., leaks and crumbling sidewalks). After the original problems were solved, the organization remained as available social capital within the community to improve the quality of life of residents in other ways.

In fields outside of health, social capital has been applied in a variety of settings ranging from the smooth and effective functioning of government institutions to the ability of communities to solve the

problems of collective action, such as intervening to prevent the occurrence of juvenile delinquency. More recently, empirical studies have begun to link social capital to population health outcomes.

Measurement of Social Capital

Two approaches have been used to measure social capital: (1) surveys and (2) direct social observations. In the survey-based approach, researchers inquired about a range of attitudes (trust of others and expectations of reciprocity and mutual aid) as well as behaviors (e.g., membership in community organizations, levels of voluntarism, and intensity of informal social interactions with neighbors) that are thought to be indicative of social capital. By contrast, direct social observation may involve the use of experimental methods, such as the letter-drop experiment in which stamped addressed envelopes are deliberately dropped on street corners to determine the proportion of letters that are subsequently mailed to the addressee by strangers (an experimental indicator of reciprocity). More commonly, direct observational approaches emphasize the observable characteristics of communities that may serve as proxies for the extent of social capital, such as the density of voluntary associations within a given locality.

Given the well-established tradition in community psychology of studying the characteristics of cohesive communities, there is legitimate debate over the extent to which the concept of social capital represents old wine in new bottles. In community psychology, concepts such as the psychological sense of community, community competence, and neighboring all predate the current vogue for the term social capital. In addition, there is an extensive prior literature in health psychology and social epidemiology on the related constructs of social support and social integration. However, an important distinction between social support and social capital is that, whereas the former variable is measured at the individual level, the latter is most often conceptualized and measured at the community or area level. That is, social capital has most often been considered to be a property of the collective (neighborhoods, cities, or even regions).

If social capital is indeed a collective property, a critical question is whether aggregated responses to social surveys (e.g., in response to questions about trusting neighbors) genuinely represent a group characteristic or whether they are confounded by other characteristics of the individual respondents that correlate with perceptions of social capital. Some evidence suggests that aggregated survey responses actually capture genuine place-based differences. An analysis of the Community Survey of the Project

on Human Development in Chicago Neighborhoods reported that even after controlling for the socio-demographic attributes of individual respondents (age, sex, race/ethnic status, income, and schooling), there were significant between-neighborhood variations in levels of social capital as measured by perceptions of trust.

Finally, research has increasingly pointed out the important distinction between bonding and bridging social capital. Bonding social capital refers to the social connections that exist within homogeneous groups, for example the strong ties that connect family members, close friends, or neighbors who are alike with respect to race/ethnicity, social class, and religion. By contrast, bridging social capital refers to social connections that cross race/ethnic, class, and other boundaries.

Social Capital and Population Health

Empirical studies have linked social capital to a growing number of population health outcomes, including mortality, mental health, health-related behaviors, self-rated health, and disability. In one of the first ecological analyses, indicators of social capital at the U.S. state level (levels of trust, perceptions of reciprocity, and density of civic associations) were strongly correlated with all-cause and cause-specific mortality rates. Subsequent investigations extended these findings down to the neighborhood level. In a study based in the 343 neighborhoods of Chicago, the levels of trust, reciprocity, and group membership were found to be significantly inversely correlated with mortality rates for residents, even after controlling for the degree of economic deprivation.

Increasingly, studies of social capital and health have used multilevel analysis, with interesting results. In one study of mental health among African American residents of a disadvantaged community in Alabama, higher bonding social capital was found to be associated with worse mental health outcomes, whereas bridging social capital was found to be associated with better mental health. In other words, within the context of a highly disadvantaged community, a greater level of social connection and involvement may be associated with worse health outcomes because of the strains associated with the obligation to help others in such settings. In another study analyzing self-rated health within the 2000 Social Capital Community Benchmark Survey (conducted in 40 U.S. communities), a complex interaction was found between community levels of social trust and individual levels of trust. Among individuals who expressed a high level of trust of others, there was a positive influence (on self-rated health) of living in

communities with higher levels of trust. On the other hand, the opposite was found for individuals expressing mistrust of others in their community. For these individuals, the higher the community's level of trust, the worse they felt.

Mechanisms Linking Social Capital and Health

Several mechanisms have been hypothesized through which social capital can influence health. They include:

1. The ability of a community to undertake collective action (e.g., organizing to pass local ordinances that restrict smoking in public places or lobbying to fight the closure of local clinics).
2. The ability of a community to enforce healthy norms by exercising informal social control over deviant behaviors (e.g., underage drinking, smoking, and drug abuse).
3. The faster diffusion of innovation through information channels (e.g., new knowledge about a new cancer screening test).
4. The exchange of social support between members of the community (e.g., the provisions of emotional support and stress-buffering or the exchange of instrumental aid, such as cash loans, during emergencies).

It is important to emphasize, however, that not all kinds of social capital result in beneficial outcomes for health. Like any form of capital (e.g., financial capital), social capital can be used for pro-social as well as antisocial ends and can result in deleterious as well as beneficial health outcomes.

Negative Effects of Social Capital

Social capital is not a panacea for health promotion. Strong social ties within a group or a community (high bonding social capital) can imply a high level of social exchange, which in turn translates into role strain for certain members of the in-group (the support providers, who are often women). Strong bonding social capital can often coexist with conflict between the group and outside social groups (low bridging social capital). In some instances, high social capital within a community has been channeled to

exclude outsiders from entering the community – for example, by enforcing residential segregation. Strong levels of social cohesion can also imply restrictions on the freedom of individual expression as well as a downleveling of achievement norms. In addition, some forms of social capital, such as criminal gangs, may provide benefits to its members, but do very little to foster social cohesion with the rest of society. Finally, some have criticized the concept of social capital for encouraging the blaming of victims (or, in this case, the blaming of communities) for their problems. These criticisms emphasize the caution warranted in applying the concept of social capital in health promotion.

See Also the Following Articles

Community Studies; Social Support.

Further Reading

- Coleman, J. S. (1990). *Foundations of social theory*. Cambridge, MA: Harvard University Press.
- Kawachi, I., Kennedy, B. P., Lochner, K., et al. (1997). Social capital, income inequality, and mortality. *American Journal of Public Health* 87, 1491–1498.
- Kawachi, I., Kim, D. J., Coutts, A., et al. (2004). Reconciling the three accounts of social capital. *International Journal of Epidemiology* 33(4), 682–690.
- Lochner, K., Kawachi, I., Brennan, R. T., et al. (2003). Social capital and neighborhood mortality rates in Chicago. *Social Science & Medicine* 56(8), 1797–1805.
- Lochner, K., Kawachi, I. and Kennedy, B. P. (1999). Social capital: a guide to its measurement. *Health and Place* 5, 259–270.
- Mitchell, C. U. and LaGory, M. (2002). Social capital and mental distress in an impoverished community. *City & Community* 1(2), 195–215.
- Szreter, S. and Woolcock, M. (2004). Health by association?: social capital, social theory, and the political economy of public health. *International Journal of Epidemiology* 33(4), 650–667.
- Subramanian, S. V., Kim, D. J. and Kawachi, I. (2002). Social trust and self-rated health in US communities: a multilevel analysis. *Journal of Urban Health* 79(4, supplement 1), S21–S34.
- Subramanian, S. V., Lochner, K. and Kawachi, I. (2003). Neighborhood differences in social capital in the US: compositional artifact or a contextual construct? *Health and Place* 9, 33–44.

Social Conflict See: Social Networks and Social Isolation; Income Levels and Stress; Social Status and Stress; Environmental Factors.

Social Networks and Social Isolation

L F Berkman

Harvard School of Public Health, Boston, MA, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction

Theoretical Orientations

Health, Social Networks, and Social Integration

A Conceptual Model Linking Social Networks to Health

Conclusion

Glossary

<i>Appraisal support</i>	Help in decision making, giving appropriate feedback, or help deciding which course of action to take.
<i>Emotional support</i>	The amount of loving and caring, sympathy and understanding, and/or esteem and value available from others.
<i>Instrumental support</i>	Help, aid, or assistance with tangible needs such as getting groceries, getting to appointments, and cooking; aid in kind, money, or labor.
<i>Social networks</i>	The web of social relationships that surrounds an individual and the characteristics of those ties.
<i>Social participation and engagement</i>	The enactment or interaction of social ties in real-life activity. Meeting with friends and family, attending social or religious functions, and participating in occupational or social roles or in group recreation are examples.

Introduction

Over the last 25 years, there have been dozens of articles and books on issues related to social networks and social support. It is now widely recognized that social relationships and affiliation have powerful effects on physical and mental health for a number of reasons.

When investigators write about the impact of social relationships on health, many terms are used loosely and interchangeably, including social networks, social support, social ties and social integration. The aim of this article is to (1) discuss theoretical orientations from diverse disciplines that we believe are fundamental to advancing research in this area, (2) briefly review findings related to mortality and biological pathways leading to poor health, (3) provide a set of definitions of networks and aspects of networks and

support, and (4) present an overarching model that integrates multilevel phenomena.

Theoretical Orientations

There are several sets of theories that form the bedrock for the empirical investigation of social relationships and their influence on health. The earliest theories came from sociologists such as Emile Durkheim, as well as from psychoanalysts such as John Bowlby, who first formulated attachment theory. A major wave of conceptual development also came from anthropologists, including Bott, Barnes, and Mitchell, and from quantitative sociologists, such as Fischer, Laumann, Wellman, and Marsden, who, along with others, developed social network analysis. This eclectic mix of theoretical approaches coupled with the contributions of epidemiologists form the foundation of research on social ties and health.

Durkheim's contribution to the study of the relationship between society and health is immeasurable. Perhaps most important is his contribution to the understanding of how social integration and cohesion influence suicide. Durkheim's primary aim was to explain how individual pathology was a function of social dynamics. In light of recent attention to upstream determinants of health, Durkheim's work reemerges with great relevance today.

Social Network Theory: A New Way of Looking at Social Structure and Community

During the mid-1950s, a number of British anthropologists found it increasingly difficult to understand the behavior of either individuals or groups on the basis of traditional categories such as kin groups, tribes, or villages. Barnes and Bott developed the concept social networks to analyze ties that cut across traditional kinship, residential, and class groups and to explain behaviors they observed such as access to jobs, political activity, and marital roles. The development of social network models provided a way to view the structural properties of relationships among people.

Network analysis "focuses on the characteristic patterns of ties between actors in a social system rather than on characteristics of the individual actors themselves and use these descriptions to study how these social structures constrain network member's behavior" (Hall and Wellman, 1985: 26). Network analysis focuses on the structure and composition of

the network and on the contents or specific resources that flow through those networks. The strength of social network theory rests on the testable assumption that the social structure of the network itself is largely responsible for determining individual behavior and attitudes by shaping the flow of resources that determine access to opportunities and constraints on behavior.

Health, Social Networks, and Social Integration

From the mid-1970s to the present, there have been a series of studies consistently showing that the lack of social ties or social networks predicts mortality from almost every cause of death. There are now a number of studies, relating outcomes to morbidity and biomarkers of stress or resilience to inflammation, metabolic and neuroendocrine functioning, and cardiovascular reactivity, that also support the hypothesis that social networks impact health and well-being directly through multiple pathophysiological mechanisms. Social networks and support may both directly affect health outcomes and buffer the effects of other stressful experiences.

Here, we briefly review the mortality studies. In the first of these studies from Alameda County, men and women who lacked ties to others (in this case, based on an index assessing contacts with friends and relatives, marital status, and church and group membership) were 1.9–3.1 times more likely to die in a 9-year follow-up period than those who had many more contacts. Another study, in Tecumseh, Michigan, showed a similar strength of positive association for men, but not for women, between social connectedness/social participation and mortality risk over a 10–12 year period; this study has the additional strength that it controlled for some biomedical predictors assessed from physical examination (e.g., cholesterol, blood pressure, and respiratory function).

Similar results have been reported from several studies in the United States and from three in Scandinavia. Investigators working on a study in Evans County, Georgia, found risks to be significant in older White men and women without social networks, even when biomedical and sociodemographic risk factors were controlled for, although some racial and gender differences were observed. Two studies in Sweden reported significantly increased risks among socially isolated adults, and a study of men and women in eastern Finland found that social connections were related to mortality risk for men, but not for women, independently of standard cardiovascular risk factors.

Studies of older men and women confirm the continued importance of these relationships into late life. Furthermore, two studies of large cohorts of men and women in a large health maintenance organization and 32,000 male health professionals suggested that social networks are, in general, more strongly related to mortality than to the incidence or onset of disease.

Two more recent studies in Danish men and in Japanese men and women further indicated that aspects of social isolation or social support are related to mortality. Virtually all these studies found that people who are socially isolated or disconnected from others are between two and five times more likely to die from all causes, compared to those who maintain strong ties to friends, family, and community.

Social networks and support have been found to predict a very broad array of other health outcomes from survival post-myocardial infarction to disease progression, functioning, and the onset and course of infectious diseases. Several studies suggested that social isolation is related to C-reactive protein (especially in men), measures of allostatic load, fibrinogen and cardiovascular reactivity, and neuroendocrine function. Although observational studies have been reasonably consistent in linking social isolation to outcome, psychosocial interventions have not been as successful at producing changes in social isolation or support that translate into better health outcomes.

A Conceptual Model Linking Social Networks to Health

Although the power of using measures of networks or social integration to predict health outcomes is indisputable, the interpretation of what the measures actually measure has been open to much debate. Hall and Wellman have appropriately commented that much of the work in social epidemiology has used the term social networks metaphorically because rarely have investigators conformed to more standard assessments used in network analysis. This criticism has been duly noted, and several calls have gone out to develop a second generation of network measures.

A second wave of research developed in reaction to this early work and as an outgrowth of work in health psychology, which has turned the orientation of the field in several ways. These social scientists focused on the qualitative aspects of social relations (i.e., their provision of social support or, conversely, the detrimental aspects of relationships) rather than on the elaboration of the structural aspects of social networks. Most of these investigators followed the assumption that what is most important about networks is the support functions they provide. Although

social support is among the primary pathways by which social networks may influence physical and mental health status, it is not the only critical pathway. Moreover, the exclusive study of more proximal pathways detracts from the need to focus on the social context and structural underpinnings that may significantly influence the types and extent of social support provided.

In order to have a comprehensive framework in which to explain these phenomena, it is helpful to move upstream and return to a more Durkheimian orientation to network structure and social context. It is critical to maintain a view of social networks as lodged within those larger social and cultural contexts that shape the structure of networks.

Conceptually, social networks are embedded in a macro-social environment in which large-scale social forces may influence network structure, which in turn influences a cascading causal process beginning with the macro-social to psychobiological processes to impact health. In this framework, social networks are embedded in a larger social and cultural context in which upstream forces are seen to condition network structure.

Networks may operate at the behavioral level through at least five primary pathways: (1) provision of social support, (2) social influence, (3) social engagement and attachment, (4) person-to-person contact, and (5) access to resources and material goods. These psychosocial and behavioral processes may influence even more proximate pathways to health status, including direct physiological responses; psychological states, including self-esteem, self-efficacy, and depression; health-damaging behaviors such as tobacco consumption or high-risk sexual activity, health-promoting behavior such as appropriate health service use and exercise, and exposure to infectious disease agents such as human immunodeficiency virus (HIV), other sexually transmitted diseases (STDs), and tuberculosis.

The Assessment of Social Networks

Social networks might be defined as the web of social relationships that surround an individual and the characteristics of those ties. Network characteristics are:

- Range or size: the number of network members.
- Density: the extent to which the members are connected to each other.
- Boundedness: the degree to which they are defined on the basis of traditional structures such as kin, work, and neighborhood.
- Homogeneity: the extent to which individuals are similar to each other in a network.

Related to network structure, the characteristics of individual ties include:

- Frequency of contact.
- Multiplexity: the number of types of transactions or support flowing through ties.
- Duration: the length of time an individual knows another
- Reciprocity: the extent to which exchanges are reciprocal.

Downstream Social and Behavioral Pathways

We next briefly review the primary pathways by which networks may influence health.

Social support Moving downstream, we now consider the mediating pathways by which networks might influence health status. Most obviously, the structure of network ties influences health via the provision of many kinds of support. This framework immediately acknowledges that not all ties are supportive and that there is variation in the type, frequency, intensity, and extent of support provided. For example, some ties provide several types of support, whereas other ties are specialized and provide only one type. Social support is typically divided into emotional, instrumental, appraisal, and informational support.

Emotional support is related to the amount of “love and caring, sympathy and understanding and/or esteem or value available from others” (Thoits 1995: 64). Emotional support is most often provided by a confidant or intimate other, although less intimate ties can provide such support under circumscribed conditions. Instrumental support refers to help, aid, or assistance with tangible needs such as getting groceries, getting to appointments, phoning, cooking, cleaning, and paying bills. House identifies instrumental support as aid in kind, money, or labor. Appraisal support is help in decision making, help by giving appropriate feedback, or help in deciding which course of action to take. Informational support is the provision of advice or information in the service of particular needs. Emotional, appraisal, and informational support are often difficult to disaggregate and have various other definitions (e.g., self-esteem support).

Perhaps even deeper than support are the ways in which social relationships provide a basis for intimacy and attachment. Intimacy and attachment have meaning not only for relationships that we traditionally think of as intimate (e.g., between partners and between parents and children) but also for more extended ties. For instance, when relationships are

solid at a community level, individuals feel strong bonds and attachment to places (e.g., neighborhood) and organizations (e.g., voluntary and religious).

Social influence Networks may influence health via several other pathways. One pathway increasingly recognized is social influence. Shared norms around health behaviors (e.g., alcohol and cigarette consumption, sexual practices, and health-care use) might be powerful sources of social influence with direct consequences for the behaviors of network members. For instance, the relationships of adolescent girls with both their male and female friends and with their parents are associated with pregnancy risk. The social influence that extends from the network's values and norms constitutes an important and underappreciated pathway through which networks impact health.

Social engagement and attachment A third and more difficult-to-define pathway by which networks may influence health status is by promoting social participation and social engagement. Participation and engagement result from the enactment of potential ties in real-life activity. Getting together with friends, attending social functions, participating in occupational or social roles, joining in group recreation, attending church are all instances of social engagement. Thus, through opportunities for engagement, social networks define and reinforce meaningful social roles, including parental, familial, occupational, and community roles, which in turn provide a sense of value, belonging, and attachment. Several recent studies suggest that social engagement is critical in maintaining cognitive ability and in reducing mortality.

In addition, network participation provides opportunities for companionship and sociability. Some have argued that these behaviors and attitudes are not the result of the provision of support *per se* but are the consequence of participation in a meaningful social context in and of itself. One reason measures of social integration or connectedness may be such powerful predictors of mortality over long periods of follow-up is that these ties give meaning to individuals' lives by virtue of enabling them to participate fully, to be obligated (in fact, often to be the providers of support), and to feel attached to their community.

Person-to-person contact Another behavioral pathway by which networks influence disease is by restricting or promoting exposure to infectious disease agents. In this regard, the methodological links between epidemiology and networks are striking. What is perhaps most remarkable is that the same network characteristics that can be health-promoting

can at the same time be health-damaging if they serve as vectors for the spread of infectious disease. Efforts to link mathematical models applying network approaches to epidemiology are in their infancy and have started to appear over the last 10 years.

The contribution of social network analysis to model of disease transmission is the understanding that in many, if not most cases, disease transmission is not spread randomly throughout a population. Social network analysis is well suited to the development of models in which exposure between individuals is not random but rather is based on geographic location, sociodemographic characteristics (e.g., age, race, and gender), or other important characteristics of the individual such as socioeconomic position, occupation, and sexual orientation. Furthermore, because social network analysis focuses on characteristics of the network rather than on characteristics of the individual, it is ideally suited to the study of the diffusion of transmissible diseases through populations via bridging ties between networks or to uncovering the characteristics of ego-centered networks that promote the spread of disease.

Access to resources and material goods Surprisingly little research has sought to examine differential access to material goods, resources, and services as a mechanism through which social networks might operate. This, in our view, is unfortunate given the work of sociologists showing that social networks operate by regulating an individual's access to life opportunities by virtue of the extent to which networks overlap with other networks. In this way, networks operate to provide access or to restrict opportunities in much the same way as the social status does. Perhaps the most important among studies exploring this tie is Granovetter's classic study of the power of weak ties that, on the one hand, lack intimacy but, on the other hand, facilitate the diffusion of influence and information and provide opportunities for mobility.

Conclusion

This article has reviewed the ways in which social isolation, lack of support, and disengagement may be chronically stressful social experiences and are linked to poor health outcomes. With the development of multilevel framework, we are struck by two issues of profound importance. The first is the upstream question of identifying the conditions that influence the development and structure of social networks. Such questions have been the substantive focus of much of social network research, especially in relationship to urbanization, social stratification,

and culture change. Yet little of this work has been integrated with health issues in a way that might guide us in the development of policies or interventions to improve public health. Recent work relating social cohesion to economic inequality has begun to help us decipher the complex interrelationships between these social experiences. Of particular interest are the ways in which social policies may impact the form and function of social networks. For instance, work/family policies and housing policies may impact network and supportive relationships in important ways, ultimately having a health impact.

The second major issue relates to the downstream question. Many investigators have assumed that networks influence health via social support functions. Our framework makes clear that this is but one pathway linking networks to health outcomes. Furthermore, the work on conflict and stress points out not only that are not all relationships positive in valence but that some of the most powerful impacts on health that social relationships may have are through acts of abuse, violence, and trauma. Fully elucidating these downstream experiences and how they are linked to health and via which biological mechanisms remains a major challenge in the field.

Acknowledgments

This paper is adapted from Berkman, L. F. and Glass, T. (2000), Social integration, social networks, social support and health, In Berkman, L. F. & Kawachi, I. (eds) *Social epidemiology*. New York: Oxford University Press; and Berkman, L. F. (2001), Social integration, social networks, and health. In Smelser, N. J. & Baltes, P. B. (eds), *International encyclopedia of the social and behavioral sciences*. New York: Elsevier.

See Also the Following Article

Social Support.

Further Reading

Bassuk, S., Glass, T. and Berkman, L. F. (1999). Social disengagement and incident cognitive decline in community-dwelling elderly persons. *Annals of Internal Medicine* 131, 165–173.

- Berkman, L. (2001). Social integration, social networks and health. In: Smelser, N. J. & Baltes, P. B. (eds.) *International encyclopedia of the social and behavioral sciences* (vol. 21), pp. 14327–14332. New York: Elsevier.
- Berkman, L. F. and Glass, T. (2000). Social integration, social networks, social support and health. In: Berkman, L. F. & Kawachi, I. (eds.) *Social epidemiology*, pp. 137–173. New York: Oxford University Press.
- Berkman, L. F. and Syme, S. (1979). Social networks, host resistance, and mortality: a nine-year follow-up of Alameda County residents. *American Journal of Epidemiology* 109, 186–204.
- Bearman, P. and Bruckner, H. (1999). Peer effects on adolescent sexual debut and pregnancy: an analysis of a national survey of adolescent girls. The National Campaign for the Prevention of Teen Pregnancy.
- Cohen, S., Underwood, S. and Gottlieb, B. (2000). *Social support measures and intervention*. New York: Oxford University Press.
- Durkheim, E. (1897). *Suicide*. New York: Free Press.
- Fischer, C. S., Jackson, R. M., Steuve, C. A., et al. (1977). *Networks and places*. New York: Free Press.
- Glass, T., Mendes de Leon, C., Marottoli, R., et al. (1999). Population based study of social and productive activities as predictors of survival among elderly Americans. *British Medical Journal* 319, 478–483.
- Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology* 78, 1360–1380.
- Hall, A. and Wellman, B. (1985). Social networks and social support. In: Cohen, S. & Syme, S. L. (eds.) *Social support and health*, pp. 23–41. Orlando, FL: Academic Press.
- Kawachi, I., Colditz, G. A., Ascherio, A., et al. (1996). A prospective study of social networks in relation to total mortality and cardiovascular disease in men in the U.S.A. *Journal of Epidemiology Community Health* 50, 245–251.
- Pennix, B. W., van Tilburg, T., Kriegsman, D. M., et al. (1997). Effects of social support and personal coping resources on mortality in older age: the Longitudinal Aging Study, Amsterdam. *American Journal of Epidemiology* 146, 510–519.
- Seeman, T. (1996). Social ties and health: the benefits of social integration. *Annals of Epidemiology* 6, 442–451.
- Sugisawa, H., Liang, J. and Liu, X. (1994). Social networks, social support and mortality among older people in Japan. *Journal of Gerontology* 49, S3–S13.
- Thoits, P. (1995). Stress, coping, and social support processes: where are we? What next? *Journal of Health and Social Behavior* (extra issue), 53–79.

Social Status and Stress

D de Ridder

Utrecht University, Utrecht, The Netherlands

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D de Ridder, volume 3, pp 468–473, © 2000, Elsevier Inc.

Social Status, Health, and Stress

Indicators of Social Status

Social Status and Exposure to Stressful Conditions

Social Status and Vulnerability to Stressful Conditions

Glossary

<i>Coping resources</i>	The personal (e.g., self-esteem and sense of control) and social (most notably, social support) resources an individual has at his or her disposal to counter the adverse effects of the experience of distress.
<i>Coping strategies</i>	The cognitive and behavioral attempts an individual may employ when faced with the experience of distress.
<i>Exposure</i>	The actual confrontation with stressful situations (events) of chronic stressful conditions.
<i>Social status</i>	Any attribute indicating the social position of an individual in a group or in the society. The most important attribute is socioeconomic status, comprising income, education, and occupation.
<i>Vulnerability</i>	The lack of personal and social coping resources, which increases the risk of a greater adverse impact of stressful conditions.

Social Status, Health, and Stress

The relationship between social status and health is intriguing because it challenges existing models of how a distal variable such as social status might affect health. Research in the past decades focused on explaining the association between low social status, especially low socioeconomic status (SES), and poor health, assuming that the greater distress of living in deprived social and material conditions is an important factor in the strong and consistent association between health and social status. The extensive literature on SES and health provides some indications why the theoretical framework of stress should be considered an interesting candidate in furthering our

understanding of the association between SES and health. First, the linearity of the relationship between SES and health implies better health outcomes as individuals ascend the SES continuum, so that even relatively affluent groups exhibit worse health than their higher SES counterparts. This suggests that factors other than the obvious role of inadequate financial resources or poor living conditions are involved and that SES might also affect psychosocial factors. Identifying these factors may be considered a search for more health-proximal processes so as to bridge the gap between indicators of social status and the associated health outcomes. Second, the analysis of potential measurement artifacts shows that the association between social status and health is robust and is little affected by social selection factors, which would be the case if those who are in poor health were at a higher risk of getting a lower social status. Social selection effects are rare, however, and limited to young children who are unable to attend school as a result of a serious and disabling disease. The absence of social selection effects suggests that the strong association between social status and health is explained by social causation mechanisms – social status affects health instead of the other way around. Particular elements of the stress-coping paradigm offer a framework to study the causal chain between social status at one end of the continuum and health at the other one, particularly those elements that deal with conceptualizing environmental risk for health (exposure to stressors) and individual responses to explain the impact of these stressors (vulnerability to stress).

Indicators of Social Status

Traditionally, most social-epidemiological research has employed the criteria of income, occupation, and education for assessing social status. Less frequently, indicators of material wealth such as car ownership, house ownership, the number of computers in the house, or the number of annual holidays have been used. Interestingly, the sociocultural component of SES expressed in educational attainment appears to have a stronger impact than the more socioeconomic elements of income and occupational status. Recent research indicates that the relationship between income and health varies substantially by level of education. Education improves health, and its effects are larger at low levels of income; education

also reduces the strength of the income–health relationship. The linear gradient relationship between income and health is thus more characteristic of groups with higher levels of education, suggesting that education might improve skills needed to cope effectively with stress. This is shown, for example, in the finding that the mother’s educational status is a better predictor of family stress than both family income or the occupational status of the father.

Recently, it has been argued that it is not the absolute level of SES that is important for understanding its association with health and distress but rather inequality from relative standing. There is suggestive evidence that subjective social status (assessed by a ladder with 10 rungs representing where people stand in society in terms of who are the best off, have the most money, have most education, and have the best jobs) is more strongly linked to health in terms of self-rated health, heart rate, and cortisol habituation to repeated stress than the traditional objective measures of SES (education, income, and occupation). The so-called hierarchy stress perspective argues that the gradient-like association between SES and health reflects social ordering rather than material deprivation *per se*. According to this perspective, higher income leads to less stress, less status anxiety, and more perceived control. These underlying mechanisms are psychological and related to relative rather than absolute material deprivation. Social ordering has effects that are no less substantial than absolute deprivation because relative deprivation leads to worse health both directly through physiological pathways (e.g., cortisol levels) and indirectly through maladaptive coping behaviors (e.g., smoking or alcohol consumption). It is noteworthy that in countries with a high degree of income inequality (e.g., the United States) the SES–health gradient is steeper than in economically more egalitarian European countries. Thus, although SES is a reliable and robust indicator of health, SES as such provides few cues for understanding its relationship with stress. In the words of the social epidemiologists Brown and Harris (1978; p. 10), “It is not that the ‘demographic type’ measures are of no use, it is that they are not enough. What is required is their combination with concepts and measures dealing directly and in detail with the immediate . . . experience of the individual.”

Such an approach is found in social role research, advancing the hypothesis that the lower social status of, for example, the nonemployed affects their sense of meaning and purpose in life and has much less to do with the direct material consequences of their income and employment status. Striking findings in this area have been reported 30 years ago on the

disadvantaged social role of nonemployed married women, who generally reported poorer health than both employed married women and men – although it is not clear to what extent social changes since the 1970s concerning the role of women make those observations still valid. These studies revealed that women who confront the double role of parent and employee reported lower levels of psychological distress than women who are devoted to the role of homemaker. These findings suggest that regardless the actual level of stressful experiences due to overload, the social status of worker is somehow beneficial for the employed women’s health. It should be noted, however, that, although working women may be more protected from distress than nonemployed women, they are still confronted with more distress than employed men.

The results from social role studies are corroborated by findings on particular social groups in our Western community who cannot claim high social status regardless of their income, education, or employment status. Research on the elderly, the unmarried, and nonwhites puts results from social role studies in a somewhat different perspective, suggesting that it is not only the meaning associated with particular roles that accounts for the greater distress of those low in social status but also their ranking order in the social group. Individuals from disadvantaged social groups have been reported to be in poorer health and facing more distress, which may be accounted for by their lower social ranking, although direct tests of this assumption in humans are not available. Animal studies have shown that subordination tends to be associated with a chronically overactive stress response (e.g., higher blood pressure). This pattern is believed to reflect the classic picture of dominance hierarchies as linear pecking orders in which resources are unevenly distributed, inequalities are maintained through aggression and intimidation, and subordinates are subject to the most severe resource limitations, the fewest opportunities for coping, and the greatest physical and social stressors. However, recent studies suggest that the relationship between social rank and patterns of the stress response may be more diverse than has been assumed hitherto. More specifically, it appears that the negative effects of low ranking on stress responses may be attenuated by complex affiliative relationships that may help low-status stressed individuals to cope. It is unknown, however, to what extent these studies on social ranking in animals are directly relevant for humans because humans often belong to multiple hierarchies and tend to value most the one in which they rank highest.

Social Status and Exposure to Stressful Conditions

The common approach to social status and psychological distress suggests that those who are in a disadvantaged social position are confronted with more stressors than their more privileged counterparts, either as a result of exposure to more frequent or more intense negative life events or of living in chronic stressful conditions associated with continuous exposure to daily hassles. Thus, this approach reflects the notion that poorer health associated with lower social status results from social causation, emphasizing the role of structural and material conditions in causing distress. The life event approach to stress dictates that major life events (e.g., divorce or being fired) impose a breach in behavioral routines, cause distress, and therefore require adaptive efforts, eventually exhausting the system. The chronic stress approach takes a somewhat different perspective and states that is the repeated exposure to minor events or daily hassles (e.g., refused services or poor transportation facilities) that ultimately causes wear and tear on the individual.

Life Events

The assumption that people with low social status report more life events than those with higher status is corroborated only weakly by empirical research. Insofar as individuals of lower SES have reported higher numbers of life events, most of these events are directly related to financial problems. There is also a higher report of uncontrollable life events, reflecting to some extent the conditions of poverty and deprivation (e.g., discharge due to becoming obsolete) in individuals with lower social status. Still, even when a higher frequency of life events is reported, this generally fails to account for the higher levels of self-reported distress; no more than 10% of the variance in distress is explained by exposure to life events. Adopting an approach that allows for a contextual evaluation of events shows slightly different results, demonstrating that it is not so much the number of events or even their quality (e.g., uncontrollability) that counts but their potential influence on a number of crucial life areas. Research by social epidemiologists Brown and Harris describing the onset of depression in working-class women clearly demonstrated that these women were confronted with life events that had the potential of generating a chain of other stressful events. The most typical event these women reported involved a teenage pregnancy, which caused these young women to drop out of school and engage in an untimely and unwanted marriage, eventually leaving them divorced, socially isolated,

financially deprived, and without a significant chance of entering the labor market or controlling their future in another way – leaving them in a state of hopelessness that is characteristic of depression. Other studies of lower-class women have demonstrated that the lives of these women are fraught with a large number of network events, meaning that they consider the life events experienced by significant others in their social network as if they concerned them directly, which may further explain the higher levels of distress reported by these women. Nevertheless, it is clear that, although the number of major life events appears to be somewhat higher in those occupying low social status, the higher levels of exposure to major life events are not a sufficient explanation to account for the higher levels of distress.

Daily Hassles and Chronic Stressful Conditions

It appears that it is not so much the incidence of particular major life events that is causing high levels of distress in low-social-status individuals but the continuous confrontation with a large number of repetitive small events that mark living in deprived circumstances. The lower an individual is on the SES continuum, the greater the amount of hassles and the greater the time needed to address the basic tasks of living, including shopping in poorly provided facilities, lack of good transportation, poor access to social and health-care facilities, and refused services. Of course, such difficulties in accomplishing daily routines have their echo in social relations. It is not surprising, therefore, that minor conflicts in relationships in work, marriage, and parenting appear to be more common in low-SES individuals. Still, it is hard to imagine that the experience of minor stress is a simple question of individual exposure. Improving the understanding of how chronic stress affects individuals of lower SES may not so much lie in the study of individual experience of these minor stressors but in identifying the environmental features of low-SES neighborhoods themselves, implying that low SES may have an impact on a community level too. Interestingly, research shows that neighborhood-level indicators predict distress above the effects of individual-level indicators. This suggests that the more proximal environment in which people live may provide important information for identifying the environmental features that cause distress. Those in the lower ends of the social-class distribution disproportionately live in deprived neighborhoods, occupy jobs characterized by high demands and low control, and live in social environments in which they are exposed to violence, conflict, abuse, crowding, and noise. All these environmental features may be considered to increase the experience of distress, although it

remains to be answered how these unhealthy environments get under an individual's skin (Taylor & Repetti, 1997). Adopting a perspective that conceptualizes social status as a proxy to understanding the stressful features of particular environments raises the issue of how situational demands should be distinguished from subjective appraisals of these objective demands. This issue constitutes a classic problem in the cognitive approach to stress formulated by Richard Lazarus and deals with the important question of how demands in the environment are appraised and handled by the individual.

Social Status and Vulnerability to Stressful Conditions

No matter how unhealthy and environment is, not all individuals sharing the same environment are affected by that environment in a similar way, suggesting that factors concerning individual susceptibility to these environmental threats are at work. The approach highlighting a differential vulnerability to stressful conditions may improve our understanding of why individuals of lower social status experience more distress. Vulnerability is reflected in the availability of resources that may be helpful in countering the adverse effects of exposure to stress, including the repertoire of coping strategies, social support, and beliefs about control and competence. Because exposure to environmental threats fails to offer a sufficient explanation for the experience of greater distress in low-SES individuals, it has been hypothesized that the availability of coping resources varies with social status. The question is whether and why lower SES is associated with lower availability of these and other coping resources.

Coping

The employment of adequate cognitive and behavioral coping efforts constitutes one of the most important resources for buffering the adverse effects of psychological distress. Individuals who find constructive ways of coping with stress, such as taking direct action or finding meaning in their experience, are better able to withstand the negative effects of stressful circumstances than individuals who believe their coping resources to be minimal or who believe that the situation is beyond their control and rely on avoidance coping strategies, which in the long run often turn out to compromise their health. Generally, people of low social status are more inclined to adopt avoiding and emotion-focused strategies when they are confronted with stress than are their more advantaged peers. Education level and family income directly relate to the frequency of problem-focused

coping for dealing with financial strain and other hassles, with those of low social status reporting less problem-focused coping. Why low SES promotes generally ineffective coping strategies is unknown, although it has been suggested that the acquisition of coping strategies may to some extent be shaped by the continuous confrontation with few possibilities to control their environment. Although research that directly links social status and coping is rare, it has been shown that individuals of low social status are more sensitive to stress and think of it as unpleasant, disruptive and beyond their control. In contrast, individuals with higher social status appear to be less bothered by distressing events and more often regard it as a challenge to do something about it. It thus seems that the appraisal of potentially upsetting situations as uncontrollable and overwhelming may activate dysfunctional coping patterns. More generally, it seems that low-SES environments reduce individuals' reserve capacity to manage stress, thereby increasing negative emotions and cognitions with subsequent effects for dealing with these stressful conditions themselves.

Social Support

The adverse effect of chronic exposure to stressful environments may be partially offset if an individual has at least one supportive person in his or her neighborhood. Living in low-SES communities may compromise the effective use of social contacts, however. Although the social network of low-SES individuals generally is not smaller than the network of those with better education and higher incomes, the networks of low-SES individuals tend to be more homogeneous, with a large proportion of kin, which may affect the type of support provided and its perceived helpfulness. If an individual is not willing to share particular problems with family members, the possibilities for social support decrease dramatically. The larger proportion of kin in the social networks of low-social-status individuals also has consequences for mutual dependency. Asking for help on one occasion implies some kind of obligation to provide support when a family member asks for it on another occasion. The larger proportion of family in the network also bears costs because it may promote the experience of stress when a family member experiences distress. In addition, the rare prospective research in this area demonstrates that the experience of distress may affect the availability of social support. Studies of economic stressors such as financial hardship or unemployment confirm a pattern in which economic stressors lead to increased marital stress and decreased social support.

Social Distribution of Coping Resources

It is not clear why low-SES individuals are at a disadvantage with regard to the availability of coping resources. Vulnerability is thought to be partly genetically determined and partly acquired during childhood and adolescence by accumulating negative experiences, which may be difficult to compensate for. Lack of financial resources and poor material conditions are, of course, more common in low-SES individuals, but these factors do not appear to be the most important determinants of vulnerability. No evidence is available that limited access to coping resources is a direct result of deficient life conditions. Low self-esteem and low personal control are acquired during repetitive and accumulating negative events causing a negative spiral in which an individual learns to be helpless and hopeless. Low chances of employment or low educational attainment do not by themselves cause vulnerability to distress, but this does not mean that vulnerability does not operate at a social level. Particular resources and coping strategies may be distributed differently among the social classes and also particular beliefs about stress and health may be socially shaped and transmitted. The employment of ineffective coping strategies seems to depend on strong beliefs about what constitutes stress and what can be done about it. Beliefs reflect, at least in part, socially shared assumptions about stress, coping, and health and may be considered cultural models of the individual and his or her social group. A significant example concerns cigarette smoking as a way of coping. Although smoking is widely recognized as a health-threatening habit in both low-SES and high-SES individuals, in low-social-status individuals smoking has often turned into one of their few available coping options and creates a feeling of rest and control when faced with distress. Further research on the social distribution of coping resources is certainly required.

See Also the Following Articles

Aging and Psychological Stress; Cultural Factors in Stress; Economic Factors and Stress; Education Levels and Stress; Marital Status and Health Problems; Minorities and Stress; Smoking and Stress; Social Capital.

Further Reading

- Abbott, D. H., Keverne, E. B., Bercovich, G. B., et al. (2003). Are subordinates always stressed?: a comparative analysis of rank differences in cortisol levels among primates. *Hormones and Behavior* **43**, 67–82.
- Adler, N. E., Boyce, T., Chesney, M. A., et al. (1994). Socioeconomic status and health: the challenge of the gradient. *American Psychologist* **49**, 15–25.
- Adler, N. E., Epel, E. S., Castellazzo, G., et al. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy white women. *Health Psychology* **19**, 586–592.
- Antonovsky, A. (1967). Social class, life expectancy and overall mortality. *Millbank Quarterly* **45**, 31–73.
- Brown, G. W. and Harris, T. (1978). *Social origins of depression: a study of psychiatric disorder in women*. London: Tavistock.
- Chen, E., Matthews, K. A. and Boyce, W. T. (2002). Socioeconomic differences in children's health: how and why do these relationships change with age? *Psychological Bulletin* **128**, 295–329.
- Cohen, S., Kaplan, G. A. and Salonen, J. T. (1999). The role of psychological characteristics in the relation between socioeconomic status and perceived health. *Journal of Applied Social Psychology* **29**, 445–468.
- Feldman, P. J. and Steptoe, A. (2004). How neighborhoods and physical functioning are related: the roles of neighborhood socioeconomic status, perceived neighborhood strain, and individual health risk factors. *Annals of Behavioral Medicine* **27**, 91–99.
- Gallo, L. C. and Matthews, K. A. (2003). Understanding the association between socioeconomic status and health: do negative emotions play a role? *Psychological Bulletin* **129**, 10–51.
- Sapolsky, R. M. (2004). Social status and health in humans and other animals. *Annual Review of Anthropology* **33**, 393–418.
- Taylor, S. E. and Repetti, R. L. (1997). Health psychology: what is an unhealthy environment and how does it get under the skin? *Annual Review of Psychology* **48**, 411–447.
- Wilkinson, R. G. (1999). Health, hierarchy, and social anxiety. In: Adler, N. E., Marmot, M., McEwen, B. & Stewart, J. (eds.) *Special issue on socioeconomic status and health in industrial nations: social, psychological, and biological pathways*, *Annals of the New York Academy of Sciences*, 48–63.

Social Stress, Animal Models of

J Bugajski and A Gadek-Michalska

Polish Academy of Sciences, Krakow, Poland

A J Bugajski

Jagiellonian University, Krakow, Poland

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J Bugajski, volume 3, pp 474–478, © 2000, Elsevier Inc.

Social Stress Models

Limbic Stress Processive Pathways

Adaptation to Stress

Transmitter Systems Adaptation

CRH Systems Adaptation

Vasopressin Systems Adaptation

Glossary

<i>Adaptation</i>	The situation in which repeated exposure to the same stimulus elicits diminishing responses from the body.
<i>Corticotropin-releasing hormone (CRH)</i>	Neurohormone secreted by the hypothalamus whose functions are related to stress, a primary activator of stress hormones.
<i>Limbic-hypothalamic-pituitary-adrenal (LHPA) axis</i>	The pathway regulating the secretion of stress hormones by brain limbic structures, hypothalamus, pituitary, and adrenal gland and involving release of corticotropin-releasing hormone (CRH) from the hypothalamus, adrenocorticotrophic hormone (ACTH) from the pituitary gland, and corticosteroid hormones from the adrenal gland.
<i>Neurotransmitters</i>	Chemical messengers released from different kind of nerve endings involved in the transmission of a nerve impulse.
<i>Social stress</i>	A highly complex set of neuroendocrine changes in response to the immediate social environment of the animal.
<i>Vasopressin (AVP)</i>	Neurohormone secreted by the hypothalamus, a major regulatory peptide, with CRH, for the secretion and synthesis of ACTH; initiates the hypothalamic-pituitary-adrenal (HPA) response to stressors.

Conventional models of physical stressors in laboratory animals show little validity with chronic human psychopathologies. Models of social stress based on the naturalistic events of interactions between conspecifics present a realistic model to include chronic

psychosocial stress in animals. Repeated or chronic psychosocial stress induce adaptation of various neurotransmitter and neuropeptide regulatory systems involved in the hypothalamus-pituitary-adrenal (HPA) axis response. Chronic stress impairs the cholinergic and adrenergic system-induced HPA response. The corticotropin-releasing hormone (CRH)- and vasopressin (AVP)-induced activation of the HPA axis follows distinct time courses. Complex transcription factor interactions and mechanisms accompany the activation of these neuropeptides during psychosocial stress.

Social Stress Models

The stress response is usually presented as adaptive physiological and psychological processes activated in animals and humans by various kinds of stressors. It is regarded as an alarm system that is initiated whenever the dynamic equilibrium of the living organism, called homeostasis, is threatened by physical and psychological events that are known as stressors. Current understanding of physiological mechanisms of the central stress response system is based on both clinical data and animal reactions to various physical stressors, such as footshock, forced swim, or immobility, which may elicit behavioral and psychological responses different from those evoked by psychosocial stressors. In socially organized mammals, however, the predominant stressors are not physical events but arise from the immediate social environment of the animal. The majority of stress stimuli in humans is of a psychosocial nature and contributes to the development and expression of disease. In order to study mechanisms involved in the etiology of human affective disorders, various animal models of social stress are used. Apart from genetic factors that predispose for psychopathologies, environmental stress plays an important role in the etiology of these mental diseases.

Common rodent and nonhuman primate models of social stress used in the laboratory with a focus on social hierarchy models contribute in complementary ways to the understanding of social stress. Animal models of social stress involve single, intermittent, or chronic exposure of a subject animal to a conspecific. The most frequently used models for rodents are the social defeat and the social colony models. In the social defeat model, a male intruder is introduced into the territory of another male, the resident

conspecific, which attacks and defeats the intruder. The desired outcome of the social conflict is ensured since the resident usually has higher body weight, is familiar with fighting, and frequently belongs to a more aggressive strain than the intruder.

In a social colony, male and female rats are living together in seminatural conditions and a social structure develops, with a single male being the dominant rat and the others subordinates. Social defeat stress induces major functional and structural changes in subregions of the hippocampal formation. Defeated or subordinate animals have an increased activation of HPA axis, demonstrated by an increase in blood adrenocorticotrophic hormone (ACTH) and corticosterone levels. Consequently, animal models that involve natural social context seem to be more appropriate because they represent situations that individuals may encounter in their everyday lives.

Crowding typically induces a social stress reaction with prominent psychosocial components, avoiding the physical components present in most typically used stressors and mimicking emotional state alterations. Randomly selected animals are placed together in housing cages such that each has less than a standard amount of space. In crowded rats, no agonistic interactions leading to dominance or subordination are encountered. The prototypical marker of the presence and magnitude of social stress is the activity of the HPA axis. Many responses to social stress appear to be quite consistent across various animal models, with only a few exceptions that may depend on the paradigm or species studied. Psychological stress is a good animal model for investigating the various behavioral and neurohormonal changes related to stressful circumstances.

Psychological stressors are associated with the activation of the sympathetic and adrenomedullary system and the HPA axis. These systems respond to psychological stress with increased output of noradrenaline from the locus coeruleus (LC) and noradrenaline and adrenaline from the adrenal medulla. The hypothalamus produces CRH, which activates the anterior pituitary gland to secrete ACTH, which in turn stimulates the adrenal cortex to secrete corticosteroids, primarily cortisol in humans and corticosterone in rodents. Negative emotional responses disturb the regulation of the HPA system. Relatively pronounced HPA activation is common in depression, with episodes of cortisol secretion being more frequent and persistent among the depressed than among other psychiatric patients and normal subjects. Chronic stress may induce a state of hyporesponsiveness of the HPA axis whereby cortisol secretion is attenuated.

Limbic Stress Processive Pathways

Psychological and neurogenic stressors requiring higher-order sensory processing of signals prior to initiation of a stress response recruit the limbic stress processive pathway. These processive stressors constitute stimuli that become stressful only by comparison with previous experience. Different types of emotional experience may involve different pathways and arise from activity initiated as a result of thoughts and memories in the cortex or as a result of activity initiated in the hypothalamus. The components of the limbic-sensitive pathways include prefrontal cortex, septum, extended amygdala, hippocampus, and hypothalamic nuclei. In contrast, the limbic-insensitive pathway is recruited by stressors that represent an immediate threat to homeostasis and does not require limbic structures. The amygdala and the hippocampus normally inhibit hypothalamic activity, but stress blocks this circuitry and thus leads to stimulation of hypothalamic releasing hormones such as the stress peptide. Both the hippocampus and the medial prefrontal cortex (mPFC) play an important role in the negative feedback regulation of HPA activity during physiological and behavioral stress. Chronic restraint and psychosocial stress can induce atrophy of hippocampal apical dendrites of CA3 pyramidal neurons, depletion of synaptic vesicles in mossy fiber terminals in CA3, and impairment of hippocampal-dependent learning tasks. Chronic elevations in corticosterone also induce a suppression in synaptic plasticity in the CA1 hippocampal field.

The amygdala is also a major extrahypothalamic source of CRH-containing neurons with high expression levels for the CRH-1 and CRH-2 receptors. Psychological stressors can activate the amygdala CRH system without evident activation of the hypothalamic CRH system; the amygdala CRH system is much more sensitive to psychological stressors than the hypothalamic paraventricular nucleus (PVN). The amygdala, in concert with other extrahypothalamic limbic regions, manifests the behavioral responses to stress, while the hypothalamic CRH system is important for modulating neuroendocrine responses to stress. The hippocampus is an important target of circulating adrenal cortical hormones that act on their receptors and are abundant in this structure. Stress affects glucocorticoid receptor and mineralocorticoid receptor gene expression in the hippocampus in a stressor- and gender-specific manner. Chronic exposure to high corticosteroid levels, during prolonged stress, can have a detrimental effect on hippocampal integrity and function and its excitation and plasticity.

Adaptation to Stress

Adaptation of a response to given challenge is one of the key features of the normal ability of an organism to cope with the external world. Social factors, as unavoidable real life events, are powerful modulators of the stress response. Adaptation to stressful challenges involves activation of neural, neuroendocrine, and neuroendocrine-immune mechanisms. The ability of the body to increase or decrease vital functions to a new state is termed allostasis, or stability through change. Allostasis actively maintains homeostasis and promotes adaptation. When the stress response is inadequate or excessive and the cost of restitution of homeostasis is too high, the resulting condition has been termed by McEwen allostatic load.

Physiological stress responses are accepted as beneficial in the short term but can be detrimental if over-activated or prolonged. Stress modulators regulate the extent, duration, and long-term impact of acute stress responses. Stress hormones are crucial mediators of both adaptive and maladaptive responses. Chronic stress induces changes in neuroendocrine and neuronal systems at several levels: in morphology of brain nuclei and neurons and in neurotransmitter and second messenger systems including single molecules. This also takes place in limbic brain structures that are under the influence of stress-activated systems. Within the limbic-hypothalamic-pituitary-adrenal (LHPA) system, chronic stress impairs the feedback-control mechanisms associated with decreased gene expression of corticosteroid receptors and structural alterations in the hippocampal formation. Prolonged stress reduces the inhibitory feedback exerted by the hippocampus over the hypothalamic CRH-producing PVN and consequently results in hypersecretion. Sustained stress leads to adrenal hyperactivity. LHPA axis alterations develop in all stressed animals independently of their position in the social hierarchy, whether dominants, subordinates, residents, or intruders, indicating that the HPA axis is more sensitive to the stressful nature of the situation and less modulated by the individual differences in the appraisal of the situation. Dysfunction of systems regulating CRH activity in the central nervous system (CNS) and LHPA axis dysregulation often contributes to the etiology of several psychiatric disorders, such as anxiety, panic disorder, posttraumatic stress disorder, and depression.

Challenging signals from the environment are communicated to the final cellular level of adaptation to stress to turn cascades of gene expression. Genes turned on by stress transcribe the information in their DNA code into messenger RNA (mRNA), which then

is translated into proteins making up the structure and adaptive functions of the brain and body. Proteins regulate the structural and energy dynamics of the cell, as well as the neurotransmitters, hormones, and signals of the psychoneuroendocrine system in mind-body medicine. A significant number of the genes in the human genome can be spliced together in different ways to generate alternative sets of proteins that modulate physiological adaptation in response to challenge and stress from purely physical to psychosocial levels.

Transmitter Systems Adaptation

Cholinergic System

Acute stress activates not only the HPA axis but also the septohippocampal cholinergic system in male rats. In the brain, acetylcholine (ACh) is one of the neurotransmitters clearly involved in the stimulation of hypothalamic release of CRH. Acute stress of immobilization increases the synthesis, binding, and release of ACh in the hippocampus and limbic structures. Chronic stress induces reduction of choline uptake and an increase in the number of muscarinic binding sites in the septohippocampal cholinergic system, which is involved in the pituitary-adrenal response to stressful stimuli. The social stress of crowding (21 rats placed in a cage designed for 7) for 3, 7, and 14 days considerably reduces the secretion of corticosterone induced by ICV stimulation of cholinergic muscarinic receptors. The reduced responsiveness of the HPA axis may be caused by a homologous desensitization and downregulation of central muscarinic receptors by an increased secretion of ACh during crowding stress. Homologous desensitization of muscarinic receptor may be regulated by β -adrenergic receptor-like kinase.

The function of the cholinergic system depends on the ACh gene, which codes for ACh and the acetylcholinesterase gene, which codes for the enzyme that hydrolyzes ACh and terminates the function of ACh excitatory neurotransmission in the CNS. Psychosocial stress can alter the cholinergic gene expression or splicing to generate alternative proteins that are involved in a variety of normal mind-body functions, as well as pathologies. Such stress-induced alternative gene splicing may constitute a major pathway in the modulation of gene expression by creative psychological and social processes.

Monoaminergic System under Stress

An important component of the stress response in the brain is the activation of monoaminergic systems.

The LHPA axis interacts with a variety of monoaminergic systems, including brain stem noradrenergic and serotonergic systems. Endocrine and central brain stem monoaminergic responses to stress appear to be precisely coordinated to facilitate adaptive physiological responses. Stress-induced stimulation of the noradrenergic/adrenergic, the serotonergic, and the dopaminergic system enables the brain to flexibly respond to challenges by reorganizing its neuronal networks. Acute stress stimulates noradrenaline (NA) release from terminal neurons in the LC, leading to depletion of NA, but repeated or prolonged stress restitutes NA levels in the brain.

Stress raises activity of phenylethanolamine-N-methyltransferase (PNMT), a specific enzyme that converts NA into adrenaline and accelerates turnover of brain NA and adrenaline.

In addition, stress-induced increased expression of the tyrosine hydroxylase gene may elevate NA concentrations. The release of catecholamines in the response to stressors is followed by an increase in the expression of genes that encode catecholamine-synthesizing enzymes by transcriptional mechanisms in the LC and the adrenal medulla. The persistence of transcriptional activation depends on the duration and severity of the stress. In response to an acute stressor, catecholamines stimulate some limbic regions, such as the hippocampus, amygdala, and hypothalamus, whereas the higher cognitive centers in the prefrontal cortex are inactivated. During stress, high concentrations of NA and/or adrenaline cause a decrease in a number of adrenergic receptors, termed downregulation, and low agonist concentrations upregulate adrenergic receptors. Adrenoreceptor (AR) downregulation may result from reduced expression of the receptor gene or proteolytic degradation of receptor molecules. A reduced number of functional receptors results in desensitization of the respective receptor system and impairs its biological response to an agonist. Receptor molecules may be desensitized by phosphorylation through protein kinases, rapid uncoupling from its effector unit, sequestration from the cellular membrane, and internalization into the cytoplasm and intracellular vesicles. Chronic psychosocial stress reduces the number of adrenergic receptors in brain regions that regulate emotional behavior depending on the duration of the stress period and brain regions. In the LC, α_2 -AR binding is already reduced 2 days after the onset of the stress and remains at a low level. Diminished autoreceptor activity impairs feedback inhibition of NA release and, consequently, higher levels of the neurotransmitter in projecting regions of the LC. During chronic psychosocial stress, α_2 -AR expression

alters from initial downregulation to later upregulation, inducing first high and then low levels of NA.

Stress-induced downregulation of brain β -ARs is regarded as an important mechanism of adaptation to stress. Repeated or chronic stress induces high levels of endogenous agonist NA, and adrenaline causes rapid β -AR internalization into the cells and their functional desensitization. This is considered one of the biochemical mechanisms in receptor adaptation performed in order to prevent dangerous effects of the persisting stress-induced high levels of catecholamines. The psychosocial stress of crowding for 3–7 days considerably reduces the secretion of corticosterone elicited by ICV β -adrenergic receptor agonist and evokes progressively lesser diminution of that response after 14 and 21 days of crowding. In brain cortex α_2 -ARs are temporarily downregulated after 2 days of stress but upregulated after 4 weeks. Stress rapidly desensitizes the β -AR system via receptor sequestration and internalization, but some of these changes are only transient since internalized β -ARs can reinsert into the plasma membrane and restore normal receptor numbers. Social crowding stress only moderately diminishes the secretion of corticosterone evoked by ICV α_1 -adrenergic agonist. Adrenergic receptors, like the cholinergic muscarinic receptor, are linked to G-protein and are regulated by phosphorylation, which regulates desensitization of receptors to their agonists. Phosphorylation of the β -adrenergic receptor by cAMP-dependent kinase and protein kinase C (PKC) during stress is intimately involved in the desensitization of the receptor. The brain, and specifically the hippocampus, have large neuronal plasticity for a successful adaptation to a stressful situation that occurred previously. Changes that occur in the hippocampus are accompanied by alterations in the other brain regions functionally related to the hippocampus. Reacting and changing neurotransmitter systems in brain areas interact in a coordinated fashion aimed at maintaining stability through change in physiological coping mechanisms during stress.

CRH Systems Adaptation

The most well-known hypothalamic peptide whose functions are related to stress is the CRH, which may be regarded as a primary activator of the LHPA system. Produced by many neurons throughout the brain, CRH functions as neurotransmitter. CRH and noradrenergic neurons of the central nervous stress system innervate and stimulate each other in a positive feedback mechanism. CRH given ICV in rats activates the HPA axis and the sympathetic nervous

system. Stress activates hypothalamic as well as extra-hypothalamic CRH neuronal pathways. In response to a stressful stimulus, CRH is released into the hypothalamic-portal blood from neurons in the PVN of the hypothalamus, the site of convergence for stress signals in the brain. Hypothalamic CRH acts as a major mediator of the mammalian stress response by stimulating pituitary pro-opiomelanocortin (POMC) gene expression and ACTH secretion that, in turn, stimulates release of glucocorticoid from the adrenal gland. CRH stimulates POMC gene transcription in pituitary corticotrope by several transcription factors. The transcriptional activation of the pituitary POMC gene by CRH is associated with inhibition of nuclear transcription factor NF- κ B DNA-binding activity and with activation of the activator protein-1 (AP-1) factor. CRH activates target pituitary cells and neurons via distinct genes encoding membrane CRH-R1 and CRH-R2 receptors. Stress causes marked increases in the expression and cell content of CRH-R1 receptors, which predominantly mediate the CRH-induced hormonal response to stress. CRH released within the PVN increases cAMP levels, and sequentially facilitates CRH transcription, via cAMP response-element binding protein (CREB) phosphorylation and later limits the transcription via induction of the inducible cAMP early repressor (ICER). Fine modulation of CRH expression depends on a complex balance between activation by phospho-CREB and repression by ICER late during stress. Induction of ICER could serve as a protective mechanism to limit transcription of hypothalamic CRH and avoid pathological consequences of excessive expression of neuropeptide in response to stress.

Precise regulation of circulating glucocorticoids is critical for preserving homeostasis under stress conditions. Excessive CRH production may lead to HPA axis dysfunction and psychiatric disturbances. The inhibitory isoforms of CRE modulator induced during stress contribute to the decline in CRH gene transcription during persistent stimulation.

Vasopressin Systems Adaptation

In the initiation and modulation of the HPA axis response to acute and psychosocial stress both CRH and AVP are involved. These two major regulatory peptides for the secretion and/or synthesis of ACTH are colocalized in the parvocellular neurons in the PVN in both humans and rodents. These peptides, colocalized in the same secretory granules at the nerve endings of the median eminence and coreleased into the portal circulation, exert synergistic effects on ACTH secretion from the anterior pituitary. Apart

from presence in the hypothalamus, CRH and AVP are localized in several other limbic and brain structures. A variety of acute and chronic stressors alters their concentrations in the hypothalamus and several other brain regions.

In pituitary corticotrope, the vasopressin V1b receptor isoform is specifically expressed and plays a crucial role in regulating HPA axis activity during stress and in resting conditions. Under stress conditions, both the hypothalamic AVP/V1bR and the CRH/CRH R1 system are required to maintain normal HPA axis regulation. Acute stress results in a rapid increase of CRH heterogeneous nuclear RNA (hnRNA) at the time of maximal ACTH secretory response. The stress-induced AVP hnRNA response in the parvocellular neurons is slower in onset than that of primary CRH transcript, with a maximum at 2 h after challenge. Differences in timing of CRH and AVP hnRNA responses indicate distinct mechanisms underlying stress-induced activation of these genes. Physical or psychological stressors activate the CRH and AVP in the parvocellular neurons of the PVN, the major regulators of ACTH secretion. The AVP gene is upregulated by cAMP. The promoter of the AVP gene contains cAMP response element (CRE), and the promoter activates of transfected AVP gene are upregulated with cAMP. The dual CRH/AVP regulatory system may benefit the organism in coping with various stressful events. The release of CRH and AVP peptides is regulated differently depending on the nature and the duration of the imposed stress. Prolonged stressors induce a gradual shift from a CRH-predominant stress response to an AVP-predominant response after chronic stress. These adaptational changes could partly depend on glucocorticoid receptor changes induced by repeated stress in the PVN and the hippocampus and the greater responsiveness of AVP than CRH to chronic stimulatory inputs. CRH and AVP regulate MR and GR in the hippocampus and anterior pituitary, structures with the highest levels of MR and GR mRNA and receptor protein in the CNS. These GRs have an important role in the regulation of the LHPA axis and stress adaptation and the pathophysiology of stress-related disorders.

AVP and CRH bind to specific pituitary corticotrop membrane receptors that activate different second messengers. The rat pituitary AVP receptor is coupled to phospholipase C, and the CRH receptor to an adenylate cyclase. Repeated stress induces rapid adaptation of the CRH mRNA response and desensitization of the pituitary ACTH response, which may be caused by negative feedback mechanisms, as well as depletion of the releasable ACTH pool or down-regulation of pituitary CRH receptors. On the other

hand, chronic stress or repeated exposure to identical short-term stressors may induce an increased pituitary responsiveness due to sensitization of pituitary corticotrops to CRH. The social stress of crowding for 3 days facilitates the ACTH response to CRH and corticosterone response remains unaffected, indicating that the HPA system is fully sensitive or even hyperactive. This may reflect predominantly moderate upregulation of pituitary CRH system. Crowding stress considerably reduces the AVP-induced ACTH and corticosterone response. This results from desensitization and/or decrease in the number of AVP receptors in the anterior pituitary corticotrops or from desensitization of the AVP-stimulated intracellular inositol triphosphate and the PKC signaling pathway. AVP plays a primary role in the regulation of the pituitary adrenal axis during adaptation to stress. Chronic stress increases the expression of AVP in parvicellular neurons of the PVN, and secretion of AVP but not CRH into the pituitary portal circulation. Changes in density and sensitivity of pituitary AVP receptors may play a major role in adaptation of the HPA axis to crowding stress, as with adaptation to chronic psychosocial stress.

See Also the Following Articles

Animal Models (Nonprimate) for Human Stress; Crowding Stress; Psychosocial Factors and Stress; Sex-Specific Effects of Early Social Stress in Mammals: a Study in Guinea Pigs.

Further Reading

- Barr, C. S., Newman, T. K., Shannon, C., et al. (2004). Rearing condition and rh5-HTTLPR interact to influence limbic-hypothalamic-pituitary-adrenal axis response to stress in infant macaques. *Biological Psychiatry* 55, 733–738.
- Bartolomucci, A., Palanza, P., Sacerdote, P., et al. (2005). Social factors and individual vulnerability to chronic stress exposure. *Neuroscience and Biobehavioral Reviews* 29, 67–81.

- Blanchard, R. J., McKittrick, C. R. and Blanchard, D. C. (2001). Animal models of social stress: effects on behavior and brain neurochemical systems. *Physiology and Behavior* 73, 261–271.
- Bugajski, J. (2000). Social stress. In: Fink, G. (ed.) *Encyclopedia of Stress* (Vol. 3), pp. 474–478. San Diego, CA: Academic Press.
- Bugajski, J., Borycz, J., Glód, R., et al. (1995). Crowding stress impairs the pituitary-adrenocortical responsiveness to the vasopressin but not corticotropin-releasing hormone stimulation. *Brain Research* 681, 223–228.
- Buwalda, B., Kole, M. H. P., Veenema, A. H., et al. (2005). Long-term effects of social stress on brain and behavior: a focus on hippocampal functioning. *Neuroscience and Biobehavioral Reviews* 29, 83–97.
- Flugge, G., van Kampen, M. and Mijster, M. J. (2003). Perturbations in brain monoamine systems during stress. *Cell Tissue Research* 315, 1–14.
- Fuchs, E. and Flugge, G. (2003). Chronic social stress: effects on limbic brain structures. *Physiology and Behavior* 79, 417–427.
- McEwen, B. S. (1998). Stress, adaptation, and disease. Allostasis and allostatic load. *Annals of the New York Academy of Sciences* 840, 33–44.
- McEwen, B. S. (2000). The neurobiology of stress: from serendipity to clinical relevance. *Brain Research* 886, 172–189.
- Rossi, E. L. (2004). Stress-induced alternative gene splicing in mind-body medicine. *Advances in Mind-Body Medicine* 20, 12–19.
- Sabban, E. L. and Kvetnansky, R. (2001). Stress-triggered activation of gene expression in catecholaminergic systems: dynamics of transcriptional events. *Trends in Neuroscience* 24, 91–98.
- Tamashiro, K. L. K., Nguyen, M. M. N. and Saki, R. R. (2005). Social stress: from rodents to primates. *Frontiers in Neuroendocrinology* 26, 27–40.
- Tanoue, A., Ito, S., Honda, K., et al. (2004). The vasopressin V1b receptor critically regulates hypothalamic-pituitary-adrenal axis activity under both stress and resting conditions. *The Journal of Clinical Investigation* 113, 302–309.
- Watanabe, Y., McKittrick, C. R., Blanchard, D. C., et al. (1995). Effects of chronic social stress on tyrosine hydroxylase mRNA and protein levels. *Molecular Brain Research* 32, 176–180.

Social Support

T C Antonucci

University of Michigan, Ann Arbor, MI, USA

J E Lansford

Duke University, Durham, NC, USA

K J Ajrouch

Eastern Michigan University, Ypsilanti, MI, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by T C Antonucci, J E Lansford and K J Ajrouch, volume 3, pp 479–482, © 2000, Elsevier Inc.

Theoretical Perspectives

Social Support as a Mechanism for Coping with Stress

Social Support as a Source of Stress

Summary and Conclusion

Glossary

<i>Anticipated support</i>	Support perceived to be available if the need for it arises in the future.
<i>Appraisal support</i>	Feedback relevant to self-evaluation.
<i>Emotional support</i>	The provision of empathy, reassurance, liking, and respect.
<i>Informational support</i>	Help with problem solving and advice.
<i>Instrumental support</i>	Aid with services and the tasks of daily living.
<i>Social network</i>	The group of individuals with whom an individual may exchange support.

This article examines the concept of social support and its relationship to stress. A brief review of several illustrative theoretical perspectives of social support is presented first. Next, an outline and overview of the research focusing on social support as a coping mechanism is provided, followed by a discussion of social support as a source of stress and a brief summary and conclusion.

Theoretical Perspectives

Social support has received a great deal of attention in the last several decades. There are some points that are now relatively well accepted: social support tends to be cumulative over the life course; it is often multi-generational; it is best when an individual both provides and receives support, although not necessarily in the context of the same relationship; and it can include exchanges between family and friends and between formal and informal providers.

Research thus far has offered interesting and sometimes quite useful evidence for how social relations affect the individual's socialization, development, coping, and general well-being. The concept of social support has been attractive to researchers in a wide variety of disciplines, including epidemiology, sociology, psychology, social work, and health care. The related theoretical perspectives or frameworks offered by each of these disciplines are too numerous to mention here; instead, illustrative examples are provided.

Epidemiologists offered some of the earliest theorizing about the role of social support and health. Researchers in this field suggested that social relations help protect the individual from disease. This perspective was initially considered quite revolutionary because these researchers claimed that physical as well as mental health could be influenced by social relations. More recently, other researchers have found empirical support for this perspective.

Kahn and Antonucci extended this notion by suggesting that individuals are surrounded by convoys of family and friends. The resulting interaction and socialization processes teach them the roles and responsibilities of childhood, adulthood, and old age. This convoy of relations is assumed to have a positive effect, helping the individual successfully meet the challenges of life. However, it is recognized that convoys may not always be positive and sometimes can expose individuals to stress rather than serve as protection. This theoretical framework suggests that social support is best understood within a life course perspective and that social interactions are cumulative and developmental.

These theoretical frameworks were general and broad; other theories have taken a more practical approach to understanding social support. Cantor's hierarchical theory of social support suggests that a natural hierarchy of people exists from whom an individual prefers to receive support. People prefer to receive support from a spouse and children first and then turn to friends and neighbors if the former are unavailable. Only if all of these informal sources of support are unavailable are people likely to seek support from a formal provider (e.g., government and community institutions). A somewhat complementary theory was offered by Litwak, who suggested that people turn to those who are most able to provide the specific goods or services they need rather than to those to whom they feel closest. Both of these theoretical perspectives have a practical task-oriented focus.

Some theories have been more psychologically focused. For example, Pearlin argued that others step in when needed and help the individual achieve his or

her goal. This provides the individual with a feeling of ability or personal achievement, which not only aids in coping with a specific event but also provides the individual with a more generalizable sense of mastery. This theoretical perspective is similar to that offered by Antonucci and Jackson, who suggested that there is a life-span developmental benefit to the constant repetition of supportive interactions. Receiving support over time leads to viewing oneself as worthy, cared for, and loved. Under optimal conditions, this exchange provides the individual with a sense of personal efficacy that is critical as the individual faces the challenges or crises of life.

The types of theoretical perspectives that have guided research in this area can be summarized as follows: (1) general broad-based theories suggesting that social relations affect an individual's ability to resist stress and disease; (2) more practical, task- or person-specific theoretical frameworks that provide guidelines for determining who is the logical choice for helping a person cope with a specific problem, task, or crisis; and finally (3) more global, developmental perspectives that are designed to explain how an individual develops the ability to cope with stress over time. We turn now to a consideration of the evidence that examines how social support influences the individual's experience of stress.

Social Support as a Mechanism for Coping with Stress

It is well established that social relationships influence individuals' psychological well-being by providing love, intimacy, reassurance of worth, tangible assistance, and guidance. Across the life span, the lack of high-quality relationships is associated with negative physical and psychological consequences such as anxiety, depression, loneliness, and poor health. Studies of social support and well-being generally reflect two lines of research. First, support is thought to have direct effects on health maintenance and recovery, enhancing well-being regardless of the level of stress experienced. Second, support has been conceptualized as having indirect effects on well-being by acting as a buffer against stress. According to this second approach, during stressful events social support mediates the association between stress and adjustment so individuals who have high-quality support are not as negatively affected by stress as those who have less adequate support. Research provides evidence for both models. In this section, however, we focus on conceptualizing support as a buffer against stress.

Empirical studies provide evidence that social support can reduce the negative impact of stress associated with physical illness, unwanted or unexpected

life events, and developmental transitions. For example, supportive relationships with family members and friends can help children cope with stress associated with school transitions, young adults adjust to the birth of a child, or older adults negotiate retirement. In particular, having a confidant appears to be an especially helpful resource in coping with stress. Recent evidence highlights race, ethnicity, and culture as important factors in links between stress and social support. For example, studies suggest that social support can buffer ethnic minorities from the stress associated with discrimination.

Social support is most effective as a buffer against stress when the support matches the needs elicited by the stressor. Researchers discuss four main types of support: emotional, appraisal, informational, and instrumental. Different types of support are more appropriate as buffers against some stressful experiences than others. For example, bereavement may call for emotional support, and illness may call for instrumental support.

Researchers also distinguish between the perceived availability of support and the actual receipt of support. Available or anticipated support is the belief that social network members are willing to provide support if the need for it arises in the future. Anticipated support and received support have different effects on psychological well-being. It appears that the perception that support is available if needed has direct positive effects on well-being but that actually receiving support is not always beneficial. In particular, received support may not act as a buffer against stress because receiving support may be a signal that an individual's own coping resources have been inadequate in resolving the stressful situation. Individuals are likely to try to deal with stress on their own before turning to others for help. Thus, receiving support may indicate that their own efforts have failed and in some sense counteract the beneficial effects of the support. In addition, receiving support may place burdens on the support network and lead to conflicts among members. Anticipated support, on the other hand, promotes individuals' own coping efforts, helps individuals maintain hope that others are available if needed, and reflects a sense of caring and commitment from the social network.

It is important to recognize that the presence of a social relationship does not necessarily indicate that the relationship is supportive. The overall quality of the relationship is more important than its mere existence. For example, Major and colleagues found that supportive relationships could help women cope with the stressful experience of having an abortion. If these supportive relationships were also perceived as being high in conflict, however, the women showed

more distress than if the relationships were low in conflict. Among women who did not perceive their relationships as supportive, there was no association between conflict in these relationships and distress. Thus, in considering whether relationships can act as buffers against stress, it is important to consider their negative as well as supportive characteristics.

To summarize, a great deal of evidence suggests that social support can help individuals cope with stress. To be most beneficial, this support should match the needs elicited by the stressor and be provided or expected in the context of a high quality relationship.

Social Support as a Source of Stress

The fact that social support can have negative consequences for the individual has only recently been addressed, and researchers are now openly calling for more attention to this issue. Social networks can become a source of conflict, stress, and sometimes health-damaging effects. One important dimension to this approach is the finding that larger social networks do not guarantee happiness, nor do they necessarily create a buffer from stress. For example, larger social networks are associated with greater stress among women, especially due to role strain. A case in point is role overload, a situation in which demands are placed on an individual that exceed his or her capacities to cope with them. For example, the responsibilities of informal caregivers to seriously impaired relatives may compete with other demands of a social network, resulting in a significant amount of stress for the individual. Having more social contacts "not only provides more potential resources but also may create additional demands on time and increase the probability of interpersonal conflicts" (Cohen and Syme, 1985: 12). Nevertheless, role theorists, such as Coontz, have noted that multiple roles are also associated with multiples sources of satisfactions, reward and accomplishments, thus suggesting the picture is not a simple one.

The most recent findings about social relations as a source of stress provide some insight into biopsychosocial links. When social relations are negative in nature, physiological reactivity becomes heightened. For instance, hostile spousal interactions increase heart rate, blood pressure, and neuroendocrine reactivity, as well as compromising immune functions. Moreover, negative social relations earlier in the life course directly affect physiological reactions later in life, contributing to allostatic load, which in turn may accelerate disease and poor mental health. This new body of research suggests that social relations influence physiological stress indicators and

that, furthermore, the effects may be lifelong and cumulative.

Assistance from family is often thought to be the most reliable and consistent source of social support, yet family members who provide social support during times of crisis are themselves the sources of stress at other times. Moreover, negative interactions are more likely to involve kin than nonkin. Family also tends to make up the majority of individuals' social networks as they age. Thus, the source of stress is often found through the relationships that unfold within the confines of family support situations. For example, divorced women with social networks dominated by family may be more criticized for their failed marriage and thus experience significantly more, rather than less, stress.

Morgan conducted focus group discussions with recently widowed women and men to explore whether social networks actually ease adjustment to widowhood. The discussions revealed that a significant portion of negative outcomes derived from relationships with family members. Morgan stated that his findings may generate useful hypotheses about the role of family and friends as sources of support by suggesting that a distinction be made between positive and negative aspects of relationships. Support from family often produces stress when discrepancies exist between expectations and actual performance.

High levels of social support have also been found to cause stress. For example, research with the elderly demonstrates that such high levels of support from caregivers may cause stress and thus be harmful to the elder's well-being. Similarly, research has shown that family assistance to adolescent mothers helps them cope with many stressful aspects of their experience, but this support often is accompanied by conflict, frustration, disappointment, and stress. Too much support may foster dependency, causing the receiver to lose autonomy and hence develop low self-esteem. Another example is chronic financial strain. Stack found that survival in a poverty-ridden community depended on financial support from social network members. Support of this type, although generally positive, could under certain conditions be problematic. Having others expect support can also limit the success of the support provider. The network members they turn to for support in the face of chronic financial stress are likely to be facing financial difficulties of their own. The issue is complex. Social support can help minorities cope with the harmful mental and physical health consequences of racial discrimination. Jackson, Williams, and Torres hypothesize that, although negative health outcomes are associated with the stress of discrimination and the related experience of social inequality and racism,

these effects can be successfully buffered by social support. At the same time, it has been noted that having a strong group identity may elevate individuals' perceptions of discrimination and in some ways make them more vulnerable to its effects.

Summary and Conclusion

Many theorists from several different disciplines have developed frameworks regarding the association between social support and stress. Empirical investigations of these theories have revealed that social support can help individuals cope with stress but that social support also can be a source of stress. These theories and research suggest that social support can be a complicated matter with both positive and negative implications for individuals and their ability to cope with the stress and challenges of everyday life.

See Also the Following Articles

Caregivers, Stress and; Familial Patterns of Stress; Health and Socioeconomic Status; Sex Differences in Human Stress Response; Social Capital.

Further Reading

- Antonucci, T. C. and Jackson, J. S. (1987). Social support, interpersonal efficacy, and health: a life course perspective. In: Carstensen, L. L. & Edelstein, B. A. (eds.) *Handbook of clinical gerontology*, pp. 291–311 New York: Pergamon.
- Cantor, M. H. (1979). Neighbors and friends: an overlooked resource in the informal support system. *Research on Aging* 1, 434–463.
- Cohen, S. and Syme, S. L. (1985). Issues in the study and application of social support. In: Cohen, S. & Syme, S. L. (eds.) *Social support and health*, pp. 3–22 New York: Academic Press.

- Coontz, S. (1992). *The way we never were: American families and the nostalgia trap*. New York: BasicBooks.
- Jackson, J. S., Williams, D. R. and Torres, M. (2003). Discrimination, health and mental health: the social stress process. In: Maney, A. (ed.) *Socioeconomic conditions, stress and mental disorders: toward a new synthesis of research and public policy*, pp. 86–146. Bethesda, MD: NIH Office of Behavioral and Social Research.
- Kahn, R. L. and Antonucci, T. C. (1980). Convoys over the life course: attachment, roles and social support. In: Baltes, P. B. & Brim, O. (eds.) *Life-span development and behavior* (vol. 3), pp. 254–283. Boston: Lexington.
- Kiecolt-Glaser, J. K. and Newton, T. L. (2001). Marriage and health: his and hers. *Psychological Bulletin* 127, 472–503.
- Litwak, E. (1985). *Helping the elderly: the complementary roles of informal networks and formal systems*. New York: Guilford.
- Major, B., Zubek, J. M., Cooper, M. L., et al. (1997). Mixed messages: implications of social conflict and social support within close relationships for adjustment to a stressful life event. *Journal of Personality and Social Psychology* 72, 1349–1363.
- McEwen, B. S. (2000). Allostasis and allostatic load: implications for neuropsychopharmacology. *Neuropsychopharmacology* 22, 108–124.
- Morgan, D. L. (1989). Adjusting to widowhood: do social networks really make it easier? *The Gerontologist* 29, 101–107.
- Pearlin, L. I. (1985). Social structure and processes of social support. In: Cohen, S. & Syme, S. L. (eds.) *Social support and health*, pp. 43–60 New York: Academic Press.
- Seeman, T. E. and Crimmins, E. (2001). Social and environmental effects on health and aging: integrating epidemiologic and demographic perspectives. In: Seeman, T. E. (ed.) *Special issue on population health and aging: strengthening the dialogue between epidemiology and demography*. *Annals of the New York Academy of Sciences*, 954, pp. 88–117. 88–117.
- Sellers, R. M. and Shelton, J. N. (2003). The role of racial identity in perceived racial discrimination. *Journal of Personality and Social Psychology* 84, 1079–1092.
- Stack, C. B. (1974). *All our kin*. New York: Harper and Row.

Social Support in Trauma

K O'Donnell and A Steptoe

University College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

Social Support and Traumatic Events
Fluctuating Social Support in Times of Crisis
Social Support and Posttraumatic Stress Disorder

Glossary

<i>Perceived support</i>	Perceptions of helping and support behavior by others.
<i>Received support</i>	The actual experience of helping and support behaviors.
<i>Social embeddedness</i>	The extent to which the individual has active ties with family, relatives, friends, work colleagues, and other members of the community. The size, density, and

	reciprocity of the social network are all important.
<i>Support deterioration</i>	The reduction in perceived support and social embeddedness that may take place following trauma and disaster.
<i>Support mobilization</i>	The activation of helping and support behaviors in times of trauma.

Social Support and Traumatic Events

Everyone has to deal with traumas of varying forms in his or her life. These can be personal, such as the diagnosis of a serious illness, involvement in an accident, or unexpected death of a relative; or traumas can occur on a larger scale in the community when disaster strikes. Disaster can strike unpredictably at any time and in any place. Even when disasters do not lead directly to injury, they can have devastating effects on mental and physical health, on economic well-being, and on the structure of a society. The impact of disasters depends on many factors, including the location, context, and the actual form the disaster takes. The location of a disaster affects people's access to help and the availability of resources. For example, the 2004 Indian Ocean tsunami devastated remote areas of several countries, and it took many days before assistance was received. The context of a disaster is very influential; in some cases, survivors can lose their possessions, homes, and livelihoods, with long-term consequences for health and well-being. The form of the disaster – whether it is a natural disaster, involves humanmade objects (such as trains or nuclear power plants), or is deliberate (in the case of terrorism) – also make a difference. A recent authoritative review of some 160 samples of disaster victims concluded that mental health outcomes were worse for young people, victims from developing countries (compared to those from developed countries), and victims who experienced mass violence such as terrorism or shooting sprees (compared with natural or technological disasters). At the same time, there is great variation in responses to disaster. Worse mental health outcomes and more stress symptoms are commonly observed in women (compared with men), ethnic minority groups, people exposed to more severe stress, people with other severe stresses in their lives, and people with previous psychiatric problems.

Research on how people cope with disaster and trauma and rebuild their lives has focused on the psychological and material resources that are mobilized to help them. Prominent among these is social support. It may seem obvious that social support is a good thing in situations of trauma and disaster. The news media regularly depict neighbors, friends, and even strangers rallying round when a disaster strikes.

But the reality is far more complex. Often, the acute mobilization of support is not sustained in the long term, and the people who need support most are not necessarily the ones who receive it.

Social support is multidimensional. Two broad types of support can be offered in times of trauma: emotional support (comfort, sympathy, and reassurance) and practical support (food, shelter, money, advice, etc.). Social support takes place within a social and community context, and the extent to which the individual is socially embedded (has relationships within the community) is crucial. There are also important differences between received and perceived support that need to be taken into consideration. Social support is related to the individuals' socioeconomic circumstances. For example, contrary to the widely held belief that social support might be stronger in lower-socioeconomic groups due to extended ties and the values placed on social connectedness, research has shown that it is often actually lower in these groups. Mickelson and Kubzansky explain that, whereas middle-class communities can pool their resources to help those in need in times of stress, in more disadvantaged populations these resources are already stretched. Social support levels are very fluid. They depend on whether the trauma affects one individual or affects a whole social network. Studies of how social support operates during disaster and trauma has led to the recognition of the importance of concepts such as support mobilization and support deterioration.

Fluctuating Social Support in Times of Crisis

There is typically a strong mobilization of helping behavior when disaster strikes. However, the research shows that certain stressful events in people's lives can lead to a decline in social support. Divorce, the death of a loved one, the loss of a job, and illness are not only stressful in themselves but can also lead to a significant change in the availability of social contacts. There are indications that in these circumstances the decline in social contacts can itself be an additional stressor. For the individual whose partner was his or her main source of emotional support, comfort, and advice, the partner's death means that this support can no longer be called on to help in the coping process.

Similar factors may operate in disaster situations. Early research by Kaniasty and Norris was carried out with victims of floods in Kentucky in the 1980s. These disasters occurred at the same time as ongoing research that involved measures of social support, making it possible to compare support levels in the

same community both pre- and posttrauma. It was found that the victims of floods reported that less social support was available after the disaster than before. Their perceptions and expectations of support were not fulfilled. Consequently, the psychological impact of the disaster was magnified; people were not only stressed by their losses but also by the deterioration of perceived social support. This disaster took place in an impoverished rural community where there were few surplus resources.

Later research carried out in the context of two major hurricanes (Hurricane Hugo and Hurricane Andrew) broadly supports these findings. In the short-term (up to 12 months after the disasters), poor mental health is positively associated with greater direct exposure to the hurricane and is negatively related to perceived support. But perceived support was lower in people with greater direct exposure because those who were more in need were more likely to feel that support had fallen short. In the longer run, mental health was more closely associated with perceptions of support than with the extent of exposure to the disaster.

Kaniasty and Norris also examined the different ways in which social support is influenced by the context of the crisis, studying two towns in Mexico that were affected by severe floods and mudslides in 1999. The towns had very different experiences of the disaster. In one (Teziutlan), a whole mountain region was destroyed, so survivors were moved to a new community outside the original city. Flooding was worse in the second town (Villahermosa), but in this case victims were not moved. These communities were studied at several time points from 6 months to 2 years after the disaster. This allowed changes in the ratings of social support to be examined. At 6 months after the disaster, the people of Teziutlan reported less perceived support and less social embeddedness than the residents of Villahermosa. At the 2 year follow-up, the reported levels of social support in Villahermosa had improved markedly, whereas in Teziutlan they remained below normal. Social support was particularly low among women and people of lower social status.

Social Support and Posttraumatic Stress Disorder

The consensus in the literature is that social embeddedness and received support are beneficial for the mental health of disaster victims. Lack of social support is also a strong risk factor for posttraumatic stress disorder (PTSD), an effect that has been observed across a range of study populations. Interestingly, it appears that the level of social support

available at the time of the trauma and after the trauma is crucial to the development of PTSD symptoms. The levels of social support measured before the trauma are less predictive, probably because perceptions of support may not match reality. But relationships can be complicated. One study of police officers in New Zealand examined the effect of support from peers, from supervisors and from people outside work on the development of PTSD symptoms following exposure to severe trauma. Although levels of support did help to explain some of the variation in PTSD, it was found that being able to talk to others about the negative aspects of the work was actually associated with more, rather than fewer, PTSD symptoms. Fostering the right atmosphere for social support in organizations in which staff may have to deal with high levels of trauma can be of benefit.

The extent to which the observational research on social support and trauma can be translated into active support services by health professionals is controversial. The fact that the people who report low support from their friends and families experience greater psychological distress than others does not mean that remedial support by health and emergency personnel will necessarily be beneficial. There is growing evidence, for example, that some early psychological interventions such as debriefing can do more harm than good. Similarly, research on patient support groups indicates that these can have deleterious as well as beneficial effects. Victims of disasters can be reluctant to seek or accept psychological help. It may be that social support interventions are best delayed for a number of weeks after a disaster, prioritizing more acute needs. It is only after the disaster situation has settled down to some extent that victims are prepared to focus on emotional needs. Further research is needed to extend knowledge of how communities affected by disaster could be supported to aid the sustained mobilization of social support, whether support interventions are possible, and which approaches are most beneficial for those at risk.

See Also the Following Articles

Combat, Acute Reactions to; Disaster Syndrome; Disasters and Mass Violence, Public, Effects of; Earthquakes, Stress Effects of; Floods, Stress Effects of; Social Support; 9/11, Religion and Stress; Crisis Intervention; Loss Trauma; Posttraumatic Stress Disorder – Clinical.

Further Reading

Brewin, C. R., Andrews, B. and Valentine, J. D. (2000). Meta-analysis of risk factors for posttraumatic stress disorder in trauma-exposed adults. *Journal of Consulting and Clinical Psychology* 68, 748–766.

- Galea, S., Ahern, J., Resnick, H., et al. (2002). Psychological sequelae of the September 11 terrorist attacks in New York City. *New England Journal of Medicine* 346, 982–987.
- Helgeson, V. S. and Cohen, S. (1996). Social support and adjustment to cancer: reconciling descriptive, correlational and interventional research. *Health Psychology* 15, 135–148.
- Kaniasty, K. and Norris, F. H. (1993). A test of the social support deterioration model in the context of natural disaster. *Journal of Personality and Social Psychology* 64, 395–408.
- Mickelson, K. D. and Kubzansky, L. D. (2003). Social distribution of social support: the mediating role of life events. *American Journal of Community Psychology* 32, 265–281.
- Norris, F. H., Friedman, M. J., Watson, P. J., et al. (2002). 60,000 disaster victims speak. Part I: An empirical review of the empirical literature, 1981–2001. *Psychiatry* 65, 207–239.
- Norris, F. H., Baker, C. K., Murphy, A. D., et al. (2005). Social support mobilization and deterioration after Mexico's 1999 flood: effects of context, gender and time. *American Journal of Community Psychology* 36, 15–28.
- Stephens, C., Long, N. and Miller, I. (1997). The impact of trauma and social support on posttraumatic stress disorder. A study of New Zealand police officers. *Journal of Criminal Justice* 25, 303–314.

Somatic Disorders

F Creed

University of Manchester, Manchester, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by F Creed, volume 3, pp 483–485, © 2000, Elsevier Inc.

Psychosomatic Disorders

Stress and the Onset of Organic Disorders

Medically Unexplained Symptoms

Anxiety and Depressive Disorders

Glossary

<i>Anxiety disorder</i>	A disorder of mood characterized by tension and worry, often accompanied by physical symptoms of tension, including palpitations, butterflies in the stomach, and breathing difficulties.
<i>Chronic social difficulties</i>	Ongoing environmental stressors such as unemployment, serious financial or health difficulties, and ongoing serious difficulties in close relationship.
<i>Depressive disorder</i>	A disorder of mood accompanied by sleep, appetite disturbance, lack of energy and concentration, and possibly ideas of suicide.
<i>Life event</i>	A discrete change in the person's immediate environment, such as a bereavement or a marital separation.
<i>Medically unexplained symptoms</i>	Symptoms of pain or function (e.g., bowel or breathing disturbances) that cannot be explained on the basis of underlying organic disease.

Organic disease

A physical disease understood on the basis of recognized tissue damage.

Psychosomatic medicine

The study of biological, psychological, and social aspects of all diseases.

Psychosomatic Disorders

Psychosomatic disorders were previously thought to be a separate group of disorders in which stress and psychological distress led to disease. Examples were peptic ulcer, arthritis, and dermatitis. It is now recognized that all diseases may have social, psychological, and behavioral aspects as well as physical organic features. The World Health Organization defines psychosomatic medicine as “the study of biological, psychological and social variables in health and disease.”

This is best illustrated by the example of heart disease. Heart disease is caused by many factors. These include inherent biological factors, including age, sex, and genetic enhancement, that increase the chances of a heart attack; physiological factors, such as raised blood pressure, smoking, alcohol consumption, and increased blood cholesterol level; and psychological and social factors, which include depression, anxiety, and certain forms of stress, which are associated with an increased chance of developing a heart attack or increased chance of dying thereafter. Stress may lead to anxiety or depressive disorders, which can lead to a fast heart rate and increased chance of irregular heart beats, which may be fatal soon after a heart attack. Lack of social support – in the form a close person with whom all difficulties or problems can be shared – also contributes to increased chance of further heart attacks.

Thus, the modern view of psychosomatic mechanisms includes a role for stress, depression, and lack of social support alongside biological factors in the causation or outcome of disease. The relative importance of environmental stress varies in different conditions.

Stress and the Onset of Organic Disorders

Stress is much less important than the other biological or physical causative factors in the development of a heart attack. There is no clear link between stressful life events (e.g., marital separation and loss of job) and having a heart attack. There is a link with more chronic stress – 40% of heart attack patients have experienced a chronic social difficulty concerning work, money, or housing compared to 17% of healthy controls. The greatest difference seems to be in the excessive hours worked and lack of proper holidays. This seems to be something that individuals bring on themselves by working excessively hard (i.e., the type A personality).

In other diseases, different patterns have been reported. In peptic ulcer, for example, life events are important, especially those that engender goal frustration. These events involve working hard toward a particular goal for many years and then being thwarted by a change of circumstances at the last moment. Such an event occurred in over half (54%) of people with peptic ulcer, compared to only 9% of healthy comparison groups.

Another stress is that of a major earthquake. There was an increase in the number of people with stomach ulcer after the Hanshin-Awaji earthquake; for most people, the ulcer developed in conjunction with an infectious agent in the stomach, *Helicobacter pylori*, but among those physically injured following the earthquake stomach ulcers developed independently of this infectious agent. Thus, there is a clear relationship with stress, which must be considered alongside other risk factors for peptic ulcer – *H. pylori* infection, drugs taken for arthritis smoking, and an inherited predisposition.

There is a discernible increase in environmental stress prior to the onset of stroke, but the effect is weak – the biological factors listed as risk factors for heart disease are much more important. In one study, 24% of stroke patients and 12% of controls had experienced a severe life event in the year prior to the stroke. This is a much smaller difference than that found in peptic ulcer.

It is easy to assess life events prior to the onset prior of a heart attack or a stroke because these conditions develop at a clear point in time. In multiple sclerosis, however, the onset is much less clearly defined. It

does appear that the experience of a severe life event during the 6 months prior to the onset of multiple sclerosis is very considerable – 62% of multiple sclerosis patients had experienced such an event, compared to only 15% of a control group. If this result is confirmed, it suggests that the experience of stress probably interacts with the immune system, which may have been sensitized by exposure to an infection during childhood or early adulthood.

Stress and the Immune System

One of the most likely ways that stress leads to the onset of organic disease is by the modification of the immune system. This has been shown in very well-controlled studies of susceptibility to infection with the cold virus. An impressive feature is the dose-response relationship between an index of psychological stress and chance of infection; that is, the more stress experienced, the greater likelihood that a cold will develop after experimental exposure to the virus. The psychological stress in these studies comprised a combination of stressful life events (environmental stress) and the degree of distress (symptoms of anxiety and depression).

In order to exclude the interference with biological variables, these studies controlled for a number of variables that might explain the stress-infection association: smoking, alcohol, exercise, diet, quality of sleep, number of white cells in the blood, and total immunoglobulin levels. The researchers also excluded the effect of personality variables of low self-esteem, personal control, and degree of extroversion and introversion. When this is done, there is a clear link between the stress levels of the caregivers of people with Alzheimer's disease and immune response, measured by the antibody response to influenza vaccination.

Medically Unexplained Symptoms

By contrast to the pattern observed in organic disease, there seems to be a very clear increase in stress just before the onset of certain disorders in which no organic disease can be found to explain the symptom. Typical disorders are irritable bowel syndrome, chronic fatigue syndrome, and fibromyalgia. The symptoms may be severe (e.g., pain) and may lead to an operation, such as an appendectomy although the appendix appears to be quite normal. The pain may also lead to many investigations that fail to reveal any recognized organic disease to explain it. Approximately two-thirds of such patients have experienced recent environmental stress, compared to less than one-quarter of people with organic disease.

or healthy control subjects. This pattern of recent stress is similar to the stress that occurs prior to depressive disorders.

Anxiety and Depressive Disorders

The most common reaction to environmental stress is depression or anxiety, which, if severe and/or prolonged, amounts to anxiety or depressive disorder. Bereavement is a typical environmental stress that may be followed by the onset of depression; it may also be associated with the increased onset of serious physical illness. The mortality of widowers has been shown to increase 40% during the 6 months following the loss of their wives.

This means that the combination of a physical (organic) disorder and concurrent anxiety or depressive disorder is common; approximately one in five people with a serious and/or disabling physical illness have concurrent anxiety or depression. Depressed people with physical illness experience more severe pain, view their illness more seriously, and feel that the chance of recovery is lower. This view of the illness can lead to a delay in the return to normal activities, such as return to work or resuming a normal sexual relationship after a heart attack.

A woman with rheumatoid arthritis, for example, whose husband also has a disabling illness can be described as being under stress, partly because of her own illness and partly because of her husband's illness. These two stresses may lead to depression, in which case the pain worsens and the illness appears to be worse. Her going to a rheumatology clinic may lead to a change in the treatment for her rheumatoid arthritis, but only if the depression is recognized and treated will the patient's health begin to return. For

this reason, modern medicine is increasingly keen to identify anxiety and depressive disorders in patients with chronic diseases, as well as the accompanying environmental stress.

Further Reading

- Creed, F. H. (1993). Stress and psychosomatic disorders. In: Goldberger, L. & Breznitz, S. (eds.) *Handbook of stress: theoretical and clinical aspects*, pp. 496–510. New York: Macmillan.
- Cohen, S., Tyrrel, D. A. J. and Smith, S. P. (1991). Psychological stress and susceptibility to the common cold. *New England Journal of Medicine* 325, 606–612.
- Dickens, C. M., McGowan, L., Percival, C., et al. (2004). Lack of a close confidant, but not depression, predicts further cardiac events after myocardial infarction. *Heart* 90(5), 518–522.
- Frasure-Smith, N., Lesperance, F. and Talajic, M. (1993). Depression following myocardial infarction: impact on 6-month survival. *Journal of the American Medical Association* 270, 1819–1825.
- Hemingway, H. and Marmot, M. (1993). Evidence based cardiology: psychosocial factors in the aetiology and prognosis of coronary heart disease. Systematic review of prospective cohort studies. *British Medical Journal* 318, 1460–1467.
- Matsushima, Y., Aoyama, N., Fukuda, H., et al. (1999). Gastric ulcer formation after the Hanshin-Awaji earthquake: a case study of *Helicobacter pylori* infection and stress-induced gastric ulcers. *Helicobacter* 4, 94–99.
- Vedhara, K., Cox, N. K., Wilcock, G. K., et al. (1999). Chronic stress in elderly carers of dementia patients and antibody response to influenza vaccination. *Lancet* 353, 627–631.
- World Health Organisation (1964) *Psychosomatic Disorders. Thirteenth Report of the WHO Expert Committee on Mental Health*. Geneva: World Health Organisation.

Space Travel See: Pressures, Effects of Extreme High and Low; Space, Health Risks of.
--

Space, Health Risks of

V A Convertino

U.S. Army Institute of Surgical Research, Fort Sam
Houston, TX, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction

Stress Environment of Space

Impact on Body Fluid and Electrolyte Balance

Impact on Nutrition

Impact on the Cardiovascular System

Impact on Thermoregulation

Impact on the Skeletal Muscle System

Impact on the Skeletal System

Impact on Neurovestibular Function

Impact on the Immune System

Summary

Glossary

<i>Atrophy</i>	A diminution or degeneration of bodily tissues or organs.
<i>Baroreceptors</i>	Neural structures located throughout the cardiovascular system that transmit nerve signals to the central nervous system in response to changes in blood pressure and initiate an involuntary nerve response (baroreflex) that brings arterial blood pressure back toward normal.
<i>Cardiac output</i>	The quantity of blood pumped by the heart each minute; the product of the amount of blood pumped each time the heart contracts (stroke volume) and the number of times the heart beats (heart rate).
<i>Hydrostatic pressure</i>	The pressure created from the weight of a column of fluid due to the presence of gravity. On Earth, hydrostatic pressure increases 1 mmHg for each 13.6 mm distance below the surface of the fluid column.
<i>Ionizing radiation</i>	Energy imparted by radiation to an atom that results in the ejection of an electron and the creation of an ion.
<i>Maximal O₂ uptake</i>	The maximal capacity for consuming and using oxygen (aerobic capacity) by the body tissues.
<i>Muscle fiber</i>	An individual muscle cell.
<i>Orthostatic intolerance</i>	The inability to tolerate the upright posture due to low blood pressure and reduced blood flow to the brain, leading to symptoms such as lightheadedness,

dizziness, blurred vision, and possible fainting.

Stress

A challenge to homeostasis that is accompanied by a change in physiological function.

Introduction

Human physiology and health have evolved and adapted in the presence of Earth's gravity. The physical stress imposed by movement in the presence of gravity manifests itself in a dramatic elevation in energy exchange, reflected by such physiological responses as increased oxygen uptake, respiration, heart rate, and sweating. The magnitude of physical stress associated with the development of the force required by muscles to perform various physical activities such as moving one's body or lifting or throwing objects is dependent on the level of gravity opposing these actions. Thus, gravity has traditionally been considered the primary constant stress factor for influencing adaptation and maintaining the health of the human body on Earth. However, the evolution of travel and habitation of humans in space, and the associated absence of gravity, has imposed new physiological adaptations via a reduction in hydrostatic pressure gradients within the cardiovascular system, a decreased weight load on the muscles and bones, and a lower energy requirement and/or output. The space environment is unique in that the physiological stress normally associated with the maintenance of health on Earth is altered as a result of reduced gravitational accelerations and factors related to confinement in a closed ecological environment. This article describes numerous physiological adaptations and responses associated with travel to space and planets with lower gravitational forces than Earth that can compromise normal function and health, particularly on return to Earth.

Stress Environment of Space

Space exposes astronauts to extreme environmental and psychological conditions imposed by a variety of stressors that result in a pronounced compromise to their safety and health, both in the space environment and on return to Earth. Immediately on entry into orbit, the relative absence of gravity results in unloading of the body tissues; external light-dark cycles occur every 90 min, disrupting normal circadian rhythmicity and ultraviolet light is nearly abolished;

microflora proliferate in the air and water and carbon dioxide levels within the enclosed space craft are chronically elevated. The crew members are required to exercise vigorously during long-duration space missions, often more intensely than they did before flight in an effort to maintain their baseline fitness and reduce the physiological decompensation associated with space flight. Psychological stress levels that can lead to significant neuroendocrine effects on physiological functions may become elevated when diverse crew members are required to exist in relatively confined living quarters, with high noise and vibration levels and with altered sleep and eating patterns. Indeed, there are reports suggesting that interpersonal conflict and depression may have contributed to early mission termination. Perhaps the most threatening factor to the health of individuals during extended travel in space is the significant elevation in ionizing radiation. For instance, the dose of ionizing radiation received during a 6-month stay on the International Space Station can be 15-fold greater than living 1 year in Houston, Texas. Depending on the type and dose of exposure, ionizing radiation can lead to damaged nerve cells in the central nervous system, cataracts, reduced fertility, and potential carcinogenic hazards. Hence, it is clearly a misconception to view space as a condition of reduced stress. The physiological responses to any one or a combination of the stressors associated with space missions can result in a loss in function and a subsequent compromise of health, particularly on return to a higher-gravity environment such as Earth.

Impact on Body Fluid and Electrolyte Balance

In the absence of gravity, altered hydrostatic pressure gradients within the cardiovascular system represent a potent stressor experienced immediately on exposure to space, resulting in a redistribution of body fluids that increases perfusion in the upper body regions while decreasing perfusion in the lower body regions. Increased filling of the heart and central circulation activate volume receptors that initiate an inhibition of vasopressin (a water-retaining hormone) and the renin-angiotensin-aldosterone system (a salt-retaining hormone response) that regulate body fluid and electrolyte balance. These neuroendocrine responses result in the increased excretion of water (diuresis) and sodium (natriuresis) from the kidneys. Although there may not be clinically significant alterations in electrolyte concentrations in the body fluids, plasma volume is reduced within 24–48 h of exposure to space. The resulting elevation in hematocrit (ratio of red blood cells to total blood volume)

inhibits the secretion of a hormone responsible for stimulating red blood cell production (erythropoietin) and subsequently induces a more gradual decrease in red blood cell mass. The result of kidney excretion of sodium chloride and reductions in plasma and red blood cell volumes is a resetting of the total circulating blood volume by 10–15% below the preflight level with little change in the primary electrolyte and protein concentrations in the body fluids.

Impact on Nutrition

In addition to the involuntary loss of body fluids and electrolytes, space alters dietary intake through changes in eating behaviors. The most prominent evidence that minimal nutritional requirements are compromised in space is a significant average reduction of approximately 7% (~4 kg) in body weight. However, the average nutritional composition of the typical diet during space missions is similar to that on Earth with ~16–18% protein, ~17–25% fat, ~60% carbohydrate, ~1% fiber, and ~3% minerals. Average daily fluid (>2000 ml) and sodium (>3500 mg) intakes are at or above recommended minimal requirements, whereas the average dietary calcium intake of approximately 900 mg day⁻¹ is below the recommended dietary allowance (RDA) of 1000–1200 mg day⁻¹. The average daily baseline energy expenditure in space has been estimated at 2500 kcal, although the actual daily energy expenditure is much often higher than this because of the performance of physical exercise and infrequent extravehicular (outside the spacecraft) work activities (EVA). When the data are standardized for body size (weight), it becomes clear that space inhibits hunger because there is a significant and consistent reduction in daily energy intake and a negative nitrogen balance. Reduced caloric intake (i.e., failure to consume available food) in space has been associated with crewmembers' decreased appetite, feeling of fullness in the abdomen, nausea, and preoccupations with critical mission tasks.

Impact on the Cardiovascular System

The loss of circulating blood volume contributes directly to a reduction in cardiac filling, as evidenced by the reduction in echocardiographic measures of end-diastolic (input) and stroke volumes (output) in astronauts following space flight. Despite compensatory elevations in heart rate, a subsequent reduction in cardiac output can result in low blood pressure (hypotension) if compensatory peripheral vasoconstriction is not adequate. In addition to reduced circulating vascular volume, there is evidence of

alterations in cardiovascular structures that occur during adaptation to space. Magnetic resonance imaging of cardiac muscle mass before and after space missions has revealed evidence of a smaller myocardial mass (cardiac atrophy), which could contribute to the reduced myocardial contractility and stroke volume that is observed when astronauts return to Earth from space. In the relative absence of gravity and posture-related hydrostatic pressure gradients, the space environment elicits much smaller fluctuations in pressures within the cardiovascular system than those that naturally occur during change in posture in the terrestrial gravity environment. Without the dynamic stimulation of the baroreceptors (pressure receptors) from such pressure fluctuations, space is associated with a compromise to reflexes that regulate blood pressure and flow (baroreflexes) only on return to Earth. There is also evidence of elevated sympathetic nerve activity and circulating catecholamines during exposure to the space environment. Higher sympathetic nerve activity in the presence of lower blood volume is associated with the elevated constriction of the arteries in the peripheral circulation, a condition that represents a lower capacity to maintain blood pressure during standing. The condition of less circulating blood volume can be exacerbated by an increased venous compliance of the lower extremities that can lead to a greater accumulation of blood away from the brain and heart and has also been reported following exposure to space.

There are at least two documented functional compromises to the health and safety of astronauts that are associated with these cardiovascular alterations during space. On return to Earth, some degree of cardiovascular instability due to the development of low arterial blood pressure during standing (orthostatic hypotension) has been reported in many individuals who have traveled in space. Significant orthostatic hypotension has been described in 40–50% of crew members in the U.S. space program and is associated with symptoms such as dizziness, lightheadedness, and blurred vision. In some astronauts, this condition has led to fainting (syncope) and is defined as orthostatic intolerance. Orthostatic hypotension and intolerance after return from space are associated with reduced cardiac output, which is associated with low circulating blood volume, which is further complicated by the increased pooling of blood in the lower extremities due to increased venous compliance (extension of the veins) and attenuated elevation in heart rate due to cardiac baroreflex impairment. Compromised orthostatic performance in astronauts is of particular operational interest because of its potential to increase the risk of serious injury from falling, particularly during critical times such as an

emergency egress from the spacecraft. This potential risk has raised great concern and, subsequently, led to vigorous research to counteract the problem with various countermeasures.

On return from space, astronauts experience more than a 20% reduction in their maximal oxygen uptake (VO_2max ; aerobic capacity). The measurement of VO_2max provides an integrated index of the functional integrity of the cardiorespiratory system and its mechanisms that regulate the delivery to and use of oxygen by the working muscles. The reduction in VO_2max on return to Earth can be explained entirely by a decrease in cardiac filling and stroke volume associated with lower circulating blood volume. The cardiovascular effect of the reduction in VO_2max on return to Earth is highlighted if physical work is performed in the upright (i.e., orthostatic) posture, as evidenced by higher heart rates and lower stroke volumes measured in astronauts who exercise in the upright compared to the supine posture after their return from space. The operational consequence of reduced aerobic reserve is an increased stress on the cardiovascular system with any physical effort during or after return from a space mission. Although there is no documented case of an adverse cardiac event on return to Earth after a space mission, a potential health consequence of increased stress on the heart, particularly during intense exercise or physical work, could manifest itself in myocardial ischemia or abnormal cardiac rhythms.

Impact on Thermoregulation

On Earth, the body core temperature fluctuates within $\pm 0.5^\circ\text{C}$ as a consequence of circadian rhythm, with higher temperatures during the day than the night. Body temperature circadian rhythm is altered in space such that the minimum night temperature is elevated by 0.2°C and the maximal daily temperature is reduced by 0.2°C . Despite a change in the natural light–dark cycles to 90 minutes, the change in body temperature is in part a direct effect of low gravity because body temperature circadian rhythm can be altered by changes in gravity in the presence of normal light–dark cycles. While in space, the average sweat rate is 10% less and insensible water exchange is lower during the performance of physical work because convective heat flow and sweat dripage are reduced in the low-gravity environment by the formation of an observed sweat film on the skin surface. This results in high levels of sustained skin wetness that acts to suppress sweating. Consequently, the average sweat rate during exercise is 10% less in space and crew members demonstrate a faster rise in body core temperature for a given absolute level of

physical work after return from space. After return to Earth, skin blood flow is reduced at rest (probably associated with elevated sympathetic vasoconstriction) and the sensitivity of active vasodilation and sweating are impaired during exercise. Although there are few measurements of body temperatures recorded on astronauts during their daily activities in space, recorded alterations in the response to body temperature during daily exercise suggest a change in thermoregulation that could compromise the capacity for heat exchange and lead to higher thermal loads during EVA. The functional consequence of alterations in thermoregulation following space flight is an impaired ability to dissipate body heat during physical work, with a higher risk of heat injury.

Impact on the Skeletal Muscle System

A significant stress associated with the relative absence of gravity in space is the unloading and subsequent disuse atrophy of skeletal muscles, especially those of the lower extremities. The magnitude of the resulting decrease in limb circumference has been used as a gross index of muscle atrophy. In addition, muscle biopsies taken from astronauts before and after space missions have revealed significant muscle atrophy as evidenced by reduction in cross-sectional areas (CSAs) of slow-twitch and fast-twitch muscle fibers of the vastus lateralis (a quadriceps muscle). The relative reduction in muscle fiber CSAs is approximately 50–60% greater in larger fast-twitch fibers, which are associated with force development (i.e., strength), than in smaller slow-twitch fibers, which are associated with posture and endurance. Muscle atrophy in space is also accompanied by reduced capillary-to-fiber ratio in both slow- and fast-twitch fibers after flight, a condition that can contribute to lower oxygen delivery and use and the cardiovascular effects of limiting $\text{VO}_{2\text{max}}$.

In addition to changes in muscle morphology, space-flight causes histochemical alterations. Although there are no significant changes in anaerobic (glycolytic) enzyme activities in muscles following space flight, the activities of enzymes associated with aerobic metabolic pathways (succinate dehydrogenase, citrate synthase, β -hydroxyacyl-CoA dehydrogenase) and the use of oxygen are reduced in both slow- and fast-twitch muscle fibers. Thus, prolonged muscle unloading and reduced energy requirements associated with space causes muscle atrophy with subsequent compromises in the capacity to generate force and in the delivery and utilization of oxygen at the level of muscle fibers.

The alterations in cellular morphology and histochemistry of skeletal muscle resulting from space

travel are associated with detriments in function. Muscle strength can be decreased by 20% for knee extensors and 10% for knee flexors. In addition to the loss in muscle strength, an increase in muscle fatigue is evidenced by a 50% increase in motor nerve activity (i.e., amplitude of electromyographic recordings) required to develop an equal force in calf skeletal muscles following a space mission compared with preflight measures.

The reductions in the ability to develop and maintain dynamic muscle forces that are required for normal daily functions on Earth represents a clinical consequence of muscle atrophy associated with space. For instance, a 10% decrease in force generation in the muscle groups of the lower extremities associated with adaptation to space is associated with a 10–15% reduction in distances of long and high jumps. The atrophy and fatigue of postural muscles can limit the ability to function during simple standing. The significant loss of muscle structure and function could predispose individuals to serious injury during physical activities or render them incapable of performing routine tasks required for normal living such as standing, stair climbing, and getting out of bed.

Impact on the Skeletal System

The result of the space environment on the integrity of the skeletal structure is clear; bone mass and bone mineral density (BMD) decrease in space. The most profound cause of bone loss in space is the dramatic unloading of bone with the relative absence of gravitational forces and dynamic mechanical forces exerted by skeletal muscles during movement on Earth. Bone loss begins immediately on unloading and is associated with a rapid and sustained loss of calcium from the body. The state of the skeletal system is dictated by bone mass turnover, which is determined by the net balance between resorption and formation. Bone resorption increases within the first few days of space, with an overall increase in bone resorption of 50% during long-duration flights, as indicated by elevated urinary calcium excretion. Although not all data support the notion of changes in bone formation in space, there exists some biochemical evidence that suggests a possible reduction in bone formation such as slight reductions in osteocalcin and bone-specific alkaline phosphatase. Thus, it appears that bone loss in space is due to increased bone resorption compounded by a slowly developing decrease in bone formation. Bone calcium can be lost at rates of up to 140 mg day^{-1} during flight. During 4–14 months in space, whole body BMD decreases at a rate of approximately 0.6% per

month despite the implementation of physical exercise countermeasures of $1\text{--}2\text{ h day}^{-1}$.

In addition to the unloading of bone with the absence of gravitational forces and dynamic mechanical forces exerted by skeletal muscles, other characteristics of the space environment are not conducive to the maintenance of a healthy skeletal structure. The near abolishment of ultraviolet light within the spacecraft environment can contribute to the reported reduction in vitamin D required for normal bone formation and breakdown (resorption), leading to the inhibition of calcium absorption in the intestine, parathyroid hormone release, and mineralization of bone. Ambient CO_2 levels, which are normally 0.03% on Earth, are chronically elevated to approximately 0.4% inside the spacecraft during nominal crew activity, with peaks of 1–3% during physical exercise workouts or visits from additional crew members. These high ambient CO_2 levels in the spacecraft can result in a more acid state within the body tissues, requiring bone resorption (bone breakdown at the expense of maintaining bone density) to provide bases such as carbonates and phosphates to buffer the respiratory acidosis. Although not consistently reported in all astronauts, the psychological stress of space travel may also contribute to bone loss by the direct effect that increased cortisol (stress hormone) has on stimulating bone resorption. In addition, abnormal nutritional habits reported in some crew members, such as reduced energy and calcium intake and a high sodium diet, may also contribute to the increased bone loss in space.

Health risks associated with bone degradation in space have been proposed by some flight surgeons. The increased blood and urinary calcium concentrations can increase the risk of kidney stones. There are medical reports of the formation of kidney stones in space that have resulted in infection and discomfort that could threaten the success of a space mission. Because the qualitative changes in bone in space are similar to those reported in patients with disuse osteoporosis, an increased risk of bone fracture on return to Earth represents one of the greatest health concerns of long exposures to space.

Impact on Neurovestibular Function

Neurovestibular function is important for maintaining normal posture and balance on Earth. There are several anecdotal reports by astronauts that indicate significant impairment to neurovestibular function after returning from space. While driving a car on a curved ramp, one astronaut voluntarily pulled off the road because he was experiencing a strong feeling

of tilt. In another instance, an astronaut fell to the ground while attempting to walk across a room to his bed in the dark. The gravitational unloading of the vestibular system in space causes numerous adaptations that can compromise the postural stability of an individual on return to Earth. Astronauts with their eyes closed significantly overestimate the degree of tilt when placed in a chair that is rotated sideways or tilted in a roll direction after returning from space. Experiments following space flight clearly demonstrate that astronauts show the greatest postural instability when they must rely on vestibular rather than visual information to maintain posture. Because of its importance in normal daily functions such as walking, running, navigation, and driving a car, impaired neurovestibular function associated with adaptation to space may increase the risk of injury from accidents and falls following return to Earth. Fortunately, most data indicate rapid improvement in balance and posture within hours of returning to Earth.

Impact on the Immune System

There is strong evidence that humans who fly in space have compromised immune function. During the Apollo missions, 15 of 29 astronauts reported bacterial or viral infections either during flight or within the first week after returning to Earth. During the infamous Apollo 13 mission, Fred Haze suffered from a severe urinary tract infection. The examination of cell-mediated immune responses measured from blood samples drawn from various astronauts and cosmonauts after their return from space revealed a lower immune response as indicated by an inhibited mitogen-induced formation of leukocytes, a severe decrease in the ability of leukocytes to produce interferon- α/β , an altered leukocyte subset distribution and interferon production, and reduced natural killer cell activity. Although the cause(s) of compromised immune function in space is unclear, the stress associated with confined environments, intense work schedules, exposure to radiation, poor nutrition, sleep deprivation, and/or the absence of gravity have all been associated with immune suppression. In space, more frequent illnesses associated with bacterial or viral infections can compromise the successful completion of a space mission by reducing the capacity to perform physical work. When back on Earth, a significant risk to the health of a space traveler is that immune dysfunction might make the crew more susceptible to more chronic tissue radiation damage and to the future development of tumors and cancers.

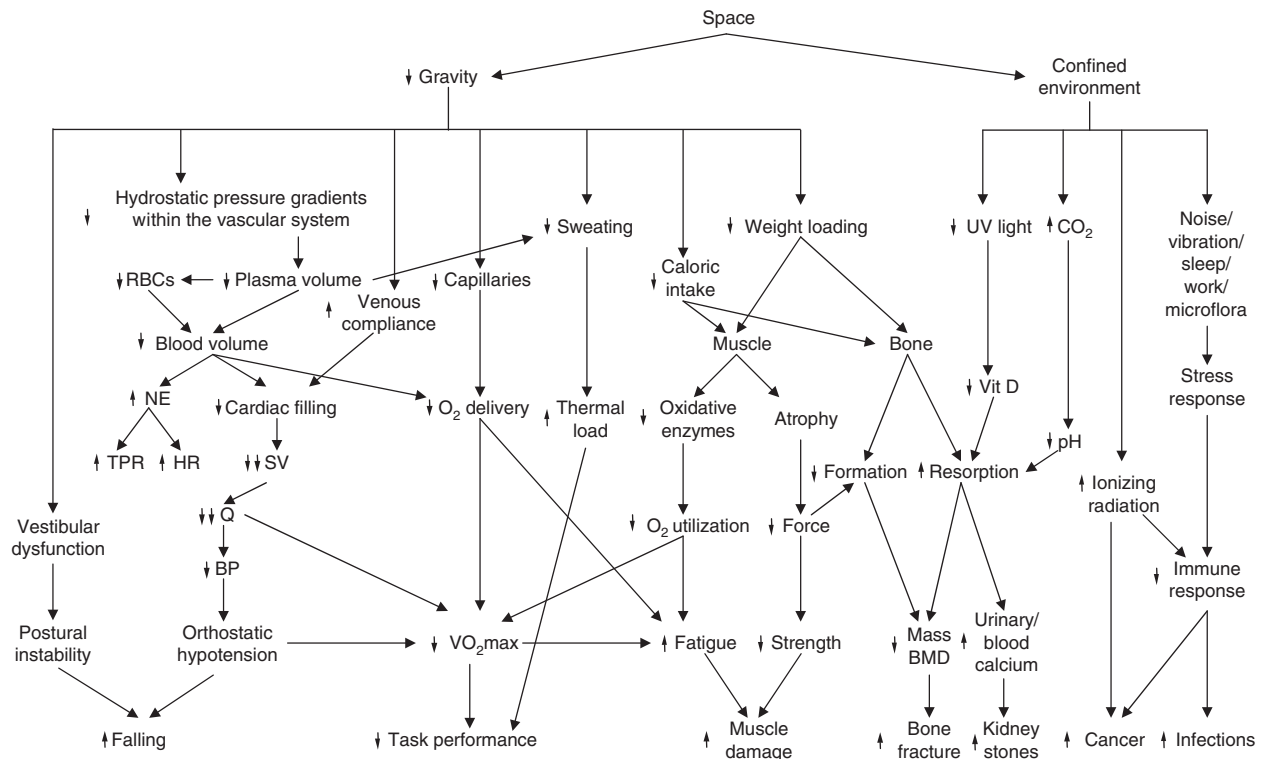


Figure 1 Cascade diagram of the relationships between physical and physiological stresses associated with the space environment and alterations in physiological functions associated with compromised health status. BMD, bone mineral density; BP, blood pressure; HR, heart rate; NE, norepinephrine; Q, cardiac output; RBCs, red blood cells; SV, stroke volume; TPR, total peripheral resistance; UV, ultraviolet; Vit D, vitamin D; VO₂max, maximal oxygen uptake.

Summary

The reduction of gravity in space presents an environmental stress that has profound effects on the structure and function of virtually all human tissues, resulting in physiological responses often seen in conditions of compromised health associated with aging, deconditioning, or inactivity-related diseases. The cascade of physical, physiological, and psychological responses associated with exposure to space is complex and interactive (Figure 1).

Although the absence of gravity is probably the most significant cause of physiological adaptations in space, the confined environment of space travel with changes in light, ambient gases, ionizing radiation levels, bacterial growth, and diverse personalities living together in confined quarters for long periods of time in the presence of high noise and vibration levels, intense work schedules, and altered sleep and eating patterns can contribute to organ dysfunctions. The space environment presents numerous physical and chemical conditions that can have profound effects on physiological health, the most significant effects of which occur on return to Earth's gravity.

Further Reading

- Buckey, J. C. (2006). *Space physiology*. New York: Oxford University Press.
- Buckey, J. C., Lane, L. D., Levine, B. D., et al. (1996). Orthostatic intolerance after spaceflight. *Journal of Applied Physiology* 81, 7–18.
- Convertino, V. A. (1990). Physiological adaptations to weightlessness: effects on exercise and work performance. *Exercise and Sports Science Reviews* 18, 119–165.
- Convertino, V. A. (1991). Neuromuscular aspects in development of exercise countermeasures. *Physiologist* 34 (supplement), S125–S128.
- Convertino, V. A. (1996). Clinical aspects of the control of plasma volume at microgravity and during return to one gravity. *Medicine and Science in Sports and Exercise* 28, S45–S52.
- Convertino, V. A. (1996). Exercise and adaptation to microgravity environments. In: Fregly, M. J. & Blatteis, C. M. (eds.) *Handbook of physiology. Sec. 4: Environmental physiology. III: The gravitational environment*, pp. 815–843. New York: Oxford University Press.
- Convertino, V. A. (1997). Conditions of reduced gravity. In: Low, P. A. (ed.) *Clinical autonomic disorders* (2nd edn., pp. 429–440). Philadelphia: Lippincott-Raven.

- Convertino, V. A. (2000). Effects of microgravity on exercise performance. In: Garrett, W. E. Jr. & Kirkendall, D. (eds.) *Exercise and sport science*, pp. 459–469. Philadelphia: Lippincott Williams & Wilkins.
- Convertino, V. A. (2002). Mechanisms of microgravity-induced orthostatic intolerance and implications of effective countermeasures: overview and future directions. *Journal of Gravitational Physiology* 9, 1–12.
- Convertino, V. A. (2002). Planning strategies for exercise and nutrition countermeasures for long duration space flight. *Nutrition* 18, 880–888.
- Edgerton, V. R. and Roy, R. R. (2000). Gravitational biology of the neuromotor systems: a perspective to the next era (invited rev.). *Journal of Applied Physiology* 89, 1224–1231.
- Edgerton, V. R., Zhou, M.-Y., Ohira, Y., et al. (1995). Human fiber size and enzymatic properties after 5 and 11 days of spaceflight. *Journal of Applied Physiology* 78, 1733–1739.
- Fitts, R. H., Riley, D. R. and Widrick, J. J. (2000). Microgravity and skeletal muscle (invited rev.). *Journal of Applied Physiology* 89, 823–839.
- Lesnyak, A., Sonnenfeld, G., Avery, L., et al. (1996). Effect of SLS-2 spaceflight on immunologic parameters of rats. *Journal of Applied Physiology* 81, 178–182.
- Levine, B. D., Lane, L. D., Watenpaugh, D. E., et al. (1996). Maximal exercise performance after adaptation to microgravity. *Journal of Applied Physiology* 81, 686–694.
- Oman, C. M., Pouliot, C. F. and Natapoff, A. (1996). Horizontal angular VOR changes in orbital and parabolic flight: human neurovestibular studies on SLS-2. *Journal of Applied Physiology* 81, 69–81.
- Prisk, G. K. (2000). Microgravity and the lung (invited rev.). *Journal of Applied Physiology* 89, 385–396.
- Schneider, S. M. and Convertino, V. A. (2005). Physiological systems and their responses to conditions of microgravity and bed rest. In: Tipton, C. M. (ed.) *ACSM's advanced exercise physiology*, pp. 595–619. Baltimore: Lippincott Williams & Wilkins.
- Stein, T. P., Leskiw, M. J. and Schluter, M. D. (1996). Diet and nitrogen metabolism during spaceflight on the shuttle. *Journal of Applied Physiology* 81, 82–97.
- Turner, R. T. (2000). What do we know about the effects of spaceflight on bone? (invited rev.). *Journal of Applied Physiology* 89, 840–847.
- West, J. B. (2000). Historical perspectives: physiology in microgravity. *Journal of Applied Physiology* 89, 379–384.

Spinal Cord Injury, Physical Stress of

S Bhatia and M Quigley

Allegheny General Hospital, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

Cardiovascular Complications
Respiratory Complications
Gastrointestinal Complications
Urologic Complications
Skin Complications
Bone and Joint Complications

Glossary

<i>Autonomic nervous system</i>	Part of the nervous system that regulates several vital functions through two mutually opposing systems, the sympathetic system and the parasympathetic system.
<i>Bradycardia</i>	A slow heart rate, usually less than 60 beats per minute in adults.
<i>Cauda equina</i>	A bundle of spinal nerve roots below the termination of the spinal cord at the first lumbar vertebra. Typically it is more resistant to injury than the cord itself.

Conus medullaris

The tapering terminal part of the spinal cord at the level of the first lumbar vertebra that contains the nerves that serve distal lower extremity function and sphincter control.

Detrusor

The muscle present in the bladder wall that contracts and pushes urine out during micturition.

Ileus

Loss of muscular contractions in the bowel wall causing atony and dilatation of bowel loops.

Osteomyelitis
Priapism

An infection of bone, usually bacterial.
Involuntary, persistent erection of the penis.

Spinal cord injuries are relatively uncommon and predominantly affect young males (80–85%). Overall, about 55% of injuries affect the cervical spine and 15% each affect the thoracic, thoracolumbar, and lumbosacral spine.

Apart from weakness, paralysis, and sensory changes, in the acute stage (depending on the level of injury) certain physiological changes are often evident. These include spinal shock, paradoxical respiration,

low body and high skin temperature, priapism, and a sweating level. Spinal shock is usually seen in injuries above the T6 level and is characterized by hypotension and bradycardia without hypovolemia due to loss of vascular sympathetic tone. The higher the level of injury and the more complete the injury, the greater the intensity and duration of spinal shock. The motor and sensory functions usually recover (if any recovery is possible) within hours, but the autonomic component of spinal shock (loss of sympathetic function) may persist for days to months. Paradoxical respiration is the result of diaphragmatic contraction in the face of intercostal muscle paralysis. The skin is warm because of a loss of sympathetically mediated peripheral vasoconstriction; this also causes the core temperature to fall. Priapism may result from an unopposed parasympathetic function, and sweating is lost because of loss of sympathetic activity.

Cardiovascular Complications

Cardiovascular complications are most common with complete injuries of the spinal cord in the cervical or high thoracic regions.

Hypovolemia and Hypotension

Spinal shock is a diagnosis of exclusion because spinal cord injury is often accompanied by other injuries. Extra emphasis should be placed on detection of any other injuries, as the lack of sensation and loss of abdominal muscle tone may make physical examination unreliable. Injuries to the chest, abdomen, pelvis, and long bones must be actively sought and excluded as a source of blood loss, hypovolemia, and hypotension. Hypovolemia and hypotension should be corrected rapidly. Hypotension in particular may increase the area of secondary injury in the spinal cord and worsen outcome. Crystalloids, colloids, and blood products may be used to replace volume as needed. Clinicians must be aware, however, that relatively low systolic blood pressure is normal in an acute high spinal cord injury, and the presence of otherwise normal cardiovascular function as indicated by adequate urinary output suggests that the patient is hemodynamically stable. Later in the course of the injury, hypotension may occur with upright posture. This is due to venous pooling and the lack of sympathetic output. It can be prevented by elastic stockings, abdominal binders, or small doses of vasopressors.

Bradycardia

Unopposed vagal activity with the loss of sympathetic activity causes bradycardia. It may be severe, especially in young patients with a high vagal tone.

Atropine or glycopyrrolate may be used to reverse bradycardia. If the patient is intubated, suctioning the endotracheal tube may sometimes provoke severe bradycardia.

Autonomic Hyperreflexia

Autonomic hyperreflexia is characterized by massive sympathetic outflow in response to certain noxious stimuli, usually genitourinary manipulations such as a change or irrigation of an indwelling Foley catheter. In severe cases, marked elevations in blood pressure, bradycardia, flushing, headache, nausea, and cardiovascular collapse may occur. There is severe vasoconstriction below the level of the cord injury and vasodilatation above it. This unopposed reflex sympathetic activity is usually seen well after the period of the spinal shock and is treated by withdrawing the initiating stimulus and administering vasodilators. Spinal anesthesia eliminates this reflex sympathetic outflow and should be considered for patients undergoing all major genitourinary procedures despite the patient's anatomic anesthesia below the injured level.

Venous Thrombosis

Immobilization and venous stasis, coagulation abnormalities, decreased fibrinolytic activity, and endothelial damage have been implicated in the high incidence of deep venous thrombosis in patients with complete spinal cord injury. Deep venous thrombosis usually occurs in the first 2 to 3 weeks after a spinal cord injury. A high index of suspicion is necessary, and prophylactic caval filters may be considered for high-risk patients.

Over the long term, cardiovascular function declines at an accelerated rate, and asymptomatic cardiovascular disease is also more prevalent in patients with spinal cord injury. This has been linked to the presence of dyslipidemia, enforced rest with limited exercise options, insulin resistance and hyperinsulinemia, and elevated percentages of body fat.

Respiratory Complications

Complete lesions above the C3 level are accompanied by a loss of diaphragmatic and intercostal muscle activity. These patients are dependent on artificial ventilation, initially through an endotracheal tube and later through a tracheostomy. Patients with mid-cervical lesions may require ventilatory assistance initially but can usually be ventilator-free long term. They remain at high risk for repeated pulmonary infections, requiring mechanical ventilation, however, due to lack of respiratory reserve. Patients with low cervical and high thoracic injuries have normal diaphragm

function but lack intercostals and abdominal muscle tone. Respiratory mechanics and efficiency are altered, as is the ability to breathe deeply or cough. Bronchial secretions are also increased because of the inability to clear airways; atelectasis, hypoxemia, and repeated infections are common.

Gastrointestinal Complications

In the acute stage of cervical or high thoracic spinal cord injury, gastric dilatation and ileus may be seen. These may persist for up to a week in complete cervical injuries. Abdominal distention, absence of bowel sounds, and abdominal x-rays showing dilated jejunal, ileal, and large bowel loops help make the diagnosis. Large volumes of fluids and electrolytes may be sequestered in the bowel and may need replacement. Cushing ulcers may also be seen in the gastrointestinal (GI) tract and may cause GI bleeding and bowel perforation. All spinal cord injury patients should receive prophylaxis with H₂ blockers or proton pump inhibitors. Long-term effects on the GI tract include gastro-esophageal reflux, delayed gastric emptying, altered colonic motility, and loss of control of defecation.

Urologic Complications

Initially after high spinal cord injury, during the period of spinal shock, the bladder is completely areflexic, but the sphincter tone may be high, necessitating a Foley catheter to allow for urinary drainage. After the period of spinal shock, detrusor contractions may return, but often these movements are not coordinated with the sphincter mechanisms, resulting in poor bladder emptying. Different patterns of incoordination may be seen with conus medullaris/cauda equina injuries. Urinary infections, especially with ureteral reflux, are major problems causing renal damage, and management is targeted to the avoidance of these. Clean intermittent catheterization remains the cornerstone of management.

Skin Complications

Pressure ulcerations are an area of breakdown leading to ischemia and necrosis of the skin. Prolonged

exposure to pressures just slightly above capillary filling pressure initiates a series of events leading to tissue necrosis and ulceration and is common in dependent areas such as the back and ischial tuberosities. Patients with spinal cord injury are also subject to venous stasis, atrophy of fat and muscle, moist perineum from fecal and urinary incontinence, and increased shearing stresses, which increase the risk of pressure ulceration. If left untreated there is invasion of deep fascia, muscles, bone, and joints. When this occurs, osteomyelitis is common and is a source of septicemia.

Bone and Joint Complications

Removal of weight bearing as a result of spinal cord injury results in skeletal adaptive responses characterized by loss of bone mineral and a negative calcium balance. This may be prevented by exercise therapy, electrical stimulation, and antiresorptive agents. An exercise regimen will also prevent joint contractures and maintain joint mobility.

See Also the Following Articles

Hypotension, Hypovolemia, and Septic Shock; Spinal Cord Injury, Psychological Stress of.

Further Reading

- Broah, S. L. (ed.) (2005). *Managing spinal cord injury: a guide to living well with spinal cord injury*. Washington, D.C.: NRH Press.
- Fehlings, M. G. and Baptiste, D. C. (2005). Current status of clinical trials for acute spinal cord injury. *Injury* 36 (supplement 2), B113–B122.
- Heimbürger, R. F. (2005). Return of function after spinal cord transection. *Spinal Cord* 43, 438–440.
- Krause, J. S., Winkler, T., Steins, S. A., et al. (2001). *Spinal cord injury desk reference: guidelines for life care planning and case management*. New York: Demos Medical Publishing.
- Liverman, C. T., Altevogt, B. M., Joy, J. E., et al. (eds.) (2005). *Spinal cord injury: progress, promise and priorities*. Washington, D.C.: National Academies Press.
- Norenberg, M. D., Smith, J. and Marcillo, A. (2004). The pathology of human spinal cord injury: defining the problems. *Journal of Neurotrauma* 21, 429–440.

Spinal Cord Injury, Psychological Stress of

S D Schnakenberg-Ott and M R Lovell

University of Pittsburgh Medical Center, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

Prevalence

Psychological Effects of Spinal Cord Injury

Psychological Assessment of Patients with Spinal Cord Injury

Treatment of Emotional Distress in Spinal Cord Injury Patients

Life Satisfaction and Coping with Injury

Glossary

Anti-depressants

The pharmacological classification of drugs typically used to treat mood disorders; include serotonin reuptake inhibitors and tricyclics.

Cognitive-behavioral therapy

An approach to psychotherapy that employs a variety of techniques in an attempt to change behavior or mood by disputing negative or irrational cognitive thought processes and replacing them with more rational or positive thought processes.

Neuropsychological evaluation

An assessment that typically involves the administration of several standardized cognitive tests that are sensitive to the effects of brain dysfunction.

*Paralysis
Paraplegia
paralysis*

A complete or partial loss of function. The paralysis of the lower extremities as the result of an injury below the first thoracic vertebrae.

Posttraumatic stress disorder (PTSD)

An anxiety disorder that is characterized by the reexperiencing of an extremely traumatic event such as a natural disaster, war, or rape. This syndrome is often accompanied by increased arousal, recurrent thoughts or images, and avoidance of stimuli associated with the trauma.

*Premorbid personality characteristics
Quadriplegia*

The level of psychological or personality functioning before the onset of a disorder or condition.

The paralysis of all four extremities as the result of an injury above the thoracic vertebrae.

Spinal cord injury (SCI)

Trauma to the spinal cord from a sudden blow to the spine resulting in vertebral fractures and dislocations; can result in varying degrees of paraplegia and quadriplegia.

Spinal cord injury (SCI) is generally due to a sudden blow to the spine resulting in vertebral fractures and dislocations. This type of trauma can result in varying degrees of paraplegia (the paralysis of the lower extremities as the result of an injury below the first thoracic vertebrae) and quadriplegia (the paralysis of all four extremities caused by an injury above the thoracic vertebrae). Consequently, SCI victims may endure problems with breathing and bowel or bladder control. Vulnerability to secondary medical problems such as chronic pain, respiratory problems, heart failure, and decubitus ulcers are also very common among SCI patients.

Prevalence

There is an estimated 11 000 new cases of SCI reported each year in the United States, and in 2005, there were reportedly 250 000 Americans living with SCI. Motor vehicle accidents account for a majority of the injuries, followed by falls, acts of violence (e.g., gunshot wounds and assaults), and sport activities.

Young adult men are primarily affected by SCIs; however, the rates are increasing among women. A 1998 study conducted by Schackelford, Farley, and Vines comparing men and women with SCI found that approximately 82% of SCIs occur in men compared to a 20% occurrence rate in women. The prevalence of SCIs also appears higher in Whites (62.9%), followed by African Americans (22%) and Hispanics (12.6%). The life expectancy rate for SCI patients has steadily increased during the past 50 years, primarily as a result of advanced medical care and the presence of emergency rehabilitation services. Approximately 85% of SCI patients go on to live 10 years or more if they survive the first 24 h after their injury. Recent statistics indicate that the average age of people living with SCI is 40 years. Because SCI is a lifelong disability, the need for health-care services will continue to grow as patients with SCI live longer.

Psychological Effects of Spinal Cord Injury

In addition to secondary medical complications, many SCI patients also experience emotional problems such as depression, anxiety, and social isolation. Psychological distress may stem from diminished independence, functional impairment, chronic pain, and loss of social and work status. Depression is thought to be the most common psychological disorder among SCI

patients. A review of the literature suggests that depression rates among SCI patients vary, with a majority of the reports indicating that between 20 and 45% of people with SCI suffer from depression. Individuals with a history of psychological disorders and substance use may be at increased risk of experiencing depression following SCI.

People with SCI appear to be at greater risk for developing anxiety disorders compared to the general population. Scivoletto et al. found in 1997 that SCI patients who presented with greater medical needs were often more anxious and depressed than SCI patients with less severe medical problems. High levels of psychological distress were also indicated for SCI patients who lived alone, had lower levels of education, and were unemployed. Individuals with a disability (e.g., polio or SCI) were also found to report lower levels of life satisfaction compared to nondisabled people. In a 1999 study examining the correlates of stress in a sample of 187 SCI patients, Gerhart and colleagues found that depressive symptoms, poor perceived well-being and poor life satisfaction were strongly associated with earlier and future stress. Interestingly, medical conditions such as pressure sores and upper-extremity pain were not found to be associated with stress in this sample.

Because SCI is often the result of a traumatic and life-threatening accident, patients with SCI may also suffer from posttraumatic stress disorder (PTSD). Radnitz et al. reported a PTSD prevalence rate of 10–15% among SCI patients, whereas other researchers reported rates up to 40%. Interestingly, Radnitz and colleagues concluded that individuals with quadriplegia were diagnosed with PTSD less frequently than individuals with paraplegia, therefore suggesting that there may be an increased vulnerability for PTSD among paraplegic SCI patients. According to the researchers, the aforementioned finding may reflect the fact that quadriplegia involves injury to the upper region of the spinal cord, thereby damaging the nerve fibers that affect sympathetic arousal thought to be associated with PTSD. In a 2003 study that attempted to identify the risk factors of PTSD in SCI patients, Nielsen did not find significant relationships between level of paralysis and PTSD or trauma etiology and PTSD. However, when considering demographic variables, the author found that marriage was associated with a decreased likelihood of SCI patients developing PTSD.

There is a high risk of pre-injury substance abuse among SCI patients compared to the general population. Many patients undergo some sort of substance abuse counseling as part of a postinjury rehabilitation program; however, this does not always prevent patients from returning to substance use following hospital discharge. In addition, one 2004 study

found that substance use at the time of injury was related to a negative response to problem solving in SCI patients. SCI patients are often prescribed narcotic and analgesic medication to treat chronic pain, and unfortunately, the potential for abuse of prescription drugs by some SCI patients is high. In a study examining alcohol and marijuana abuse in a sample of 123 SCI patients, it was found that patients who abused alcohol perceived their health status to be poorer, were more clinically depressed, and experienced more stress in their lives than patients with no alcohol abuse. Marijuana users also reported higher levels of stress and depression.

Suicide rates among SCI patients are five times higher than among nondisabled people, and suicide is the leading cause of death in the SCI population. Pre-injury personality characteristics, inadequate coping skills, history of substance abuse, and a poor social network have been identified in the literature as risk factors for suicidal ideation in patients with SCI. Young, adult, White males appear to have the highest rate of suicide within the SCI population. It is not surprising that most SCI patients exhibiting suicidality meet diagnostic criteria for depressive disorders. Kishi et al. examined acute-onset suicidal ideation and delayed-onset suicidal ideation in a group of SCI patients as well as patients suffering from stroke, brain injury, or myocardial infarction. Compared to patients who were identified as having delayed-onset suicidal ideation, patients with acute-onset suicidal ideation exhibited greater risk factors, including a history of psychological problems and substance abuse as well as poor social support. Therefore, it is possible that suicide rates could be significantly reduced if patients at risk are identified early in rehabilitation programs.

Psychological Assessment of Patients with Spinal Cord Injury

A comprehensive evaluation, including a thorough assessment of personality, substance use, and emotional and cognitive functioning, should be conducted for patients with acute SCI. With respect to personality testing, it may be important to identify characteristics that may interfere with rehabilitation and those individuals at risk for developing significant emotional difficulties following SCI. Chronic pain has been associated with depression and negatively affects coping. Therefore, assessing pain in SCI patients should also be an important component to the postinjury psychological evaluation. Pain could also negatively impact the rehabilitation process, especially if uncontrolled or untreated. The presence of chronic pain can also lead some patients to self-medicate with alcohol or illicit drugs. Because people

with SCI are often prescribed narcotics for pain relief, the potential of abusing prescription drugs is often present, thereby requiring physicians and mental health providers to routinely assess for improper use of prescription drugs.

A majority of SCI patients often have an underlying head injury and demonstrate poor performance on neuropsychological tests. A brain injury can be associated with lingering neurocognitive and neurobehavioral sequelae that can negatively impact rehabilitation. Therefore, following SCI, patients may need to undergo at least a brief assessment of cognitive functioning. Patients exhibiting significant cognitive difficulties on brief assessment or patients with any known brain injury should be referred for a more comprehensive cognitive evaluation, preferably by a neuropsychologist. A neuropsychological evaluation may provide additional information regarding the patient's cognitive strengths and weaknesses that can aid in determining prognosis for rehabilitation. Schmitt and Elliott found that SCI patients with higher verbal learning skills reported greater levels of self-adjustment and acceptance of their injury. A neurocognitive evaluation may also aid in determining what academic or work accommodations need to be implemented should the patient be able to resume school instruction or employment.

Treatment of Emotional Distress in Spinal Cord Injury Patients

Psychological services should be available for injured individuals and their families while the patients are in the hospital and following discharge. Psychotherapeutic intervention should focus on adjustment, self management, and skills training. Several studies conducted by Craig and colleagues have shown that employing cognitive-behavioral strategies as a form of treatment for SCI patients with significant levels of depression and anxiety are effective. Cognitive-behavioral therapy attempts to change behavior or mood by disputing negative or irrational cognitive thought processes and replacing cognitive distortions with more rational or positive thought processes. Cognitive-behavioral techniques are not only beneficial in the treatment of mood disorders, but have also been suggested efficacious for managing pain. Overall, cognitive-behavioral therapy may be potentially cost effective by decreasing the number of hospital readmissions and drug use.

In addition to individual psychotherapeutic intervention, group therapy may also be a valuable form of treatment for SCI patients, especially in a hospital setting. Groups that are psychoeducational in nature and focus on specific themes, such as coping with the injury, vocational issues, sexual dysfunction,

substance abuse, or secondary medical complications may be of great benefit to acutely injured patients. Therapists might also consider inviting family members to attend a separate support group to address the emotional, logistical, and financial issues associated with having a family member with SCI.

Issues related to sexual functioning in the SCI population is becoming more recognized and discussed throughout SCI literature. Of course, the level of physical impairment or paralysis following SCI can affect sexual functioning. Psychotherapeutic intervention that addresses both adaptation and coping with the physical and functional changes in sexuality following SCI is necessary. Erectile dysfunctions in men and physiological changes in women are common features of SCI. Obviously, people with SCI are not immune to or incapable of acquiring sexually transmitted diseases (STDs), and women with SCI may become pregnant. Therefore, mental health providers may also find themselves in a position to educate SCI patients on safe sex practices.

Pharmacological treatment may also be advantageous for the treatment of emotional disorders (e.g., depression and anxiety) in patients with SCI. Because they have fewer side effects compared to tricyclic antidepressants, serotonin reuptake inhibitors (SSRIs) usually have a more benign side-effect profile. However, few studies have been conducted that demonstrate the efficacy and possible side effects of antidepressant therapy in SCI patients.

Life Satisfaction and Coping with Injury

Factors such as premorbid personality characteristics and available social support can strongly influence coping with SCI. Frank and Elliott found that patients who believed they had some control over their situation and events experienced less depression and were better adjusted than patients who believed they had little control over their life situation and health status. Other protective factors against depression and anxiety appear to be age and the number of years since the injury occurred.

Life satisfaction has been shown to be higher in SCI patients with adequate social support than in SCI patients with poor social support systems. Length of time living with SCI was also positively correlated with greater life satisfaction. Similarly, a 2002 study found that social support was associated with lower levels of hopelessness and depression. Ongoing health-care services are also necessary because the life span of individuals with SCI continues to rise, thereby increasing the probability of secondary medical conditions. One study found that SCI patients who received regular follow-up medical care reported better health, independence, and fewer

severe secondary medical complications, and demonstrated lower depression scores than SCI patients who did not receive follow-up care. In essence, adequate health care delivered by well-trained physicians, nurses, mental health-care providers, vocational rehabilitation specialists, and other essential staff is pertinent in order to help facilitate a longer life expectancy and increased quality of life among SCI survivors.

See Also the Following Articles

Depression Models; Spinal Cord Injury, Physical Stress of.

Further Reading

- Beedle, A. and Kennedy, P. (2002). Quality of social support predicts hopelessness and depression post spinal cord injury. *Journal of Clinical Psychology in Medical Settings* 9(3), 227–324.
- Craig, A. and Hancock, K. (1998). Living with spinal cord injury: longitudinal factors, interventions, and outcomes. *Clinical Psychology and Psychotherapy* 5, 102–108.
- Craig, A., Hancock, K. and Dickson, H. (1994). Spinal cord injury: a search for determinants of depression 2 years after the event. *British Journal of Clinical Psychology* 33, 221–230.
- DeVivo, M., Black, K., Richards, S., et al. (1991). Suicide following spinal cord injury. *Paraplegia* 29, 620–627.
- Dijkers, M., Abela, M. and Gordon, W. (1995). The aftermath of spinal cord injury. In: Stover, S. L., DeLisa, J. A. & Whiteneck, G. G. (eds.) *Spinal cord injury: clinical outcomes from the model systems*, pp. 185–212. Gaithersburg MD: Aspen.
- Dreer, L., Elliott, T. and Tucker, E. (2004). Social problem-solving abilities and health behaviors among persons with recent onset spinal cord injury. *Journal of Clinical Psychology in Medical Settings* 11(1), 7–13.
- Dunn, M., Love, L. and Ravesloot, C. (2000). Subjective health in spinal cord injury after outpatient healthcare follow-up. *Spinal Cord* 38, 84–91.
- Frank, R. and Elliott, T. (1989). Spinal cord injury and health locus of control beliefs. *Paraplegia* 27, 250–256.
- Gerhart, K., Weitzenkamp, D., Kennedy, P., et al. (1999). Correlates of stress in long term spinal cord injury. *Spinal Cord* 37, 183–190.
- Hess, D., Marwitz, J. and Kreutzer, J. (2003). Neuropsychological impairments after spinal cord injury: a comparative study with mild traumatic brain injury. *Rehabilitation Psychology* 48, 151–156.
- Kemp, B. and Krause, J. (1999). Depression and life satisfaction among people ageing with post-polio and spinal cord injury. *Disability and Rehabilitation* 21, 241–249.
- Kennedy, P. (2001). Spinal cord injuries. In: Johnston, D., Johnston, M., Bellack, A. & Hersen, M. (eds.) *Health psychology. Vol. 8: Comprehensive clinical psychology*, pp. 445–462. Amsterdam: Elsevier Science.
- Kennedy, P., Lowe, R., Grey, N., et al. (1995). Traumatic spinal cord injury and psychological impact: across sectional analysis of coping strategies. *British Journal of Clinical Psychology* 34, 627–639.
- Kishi, Y., Robinson, R. and Kosier, J. (2001). Suicidal ideation among patients during the rehabilitation period after life threatening physical illness. *Journal of Nervous and Mental Disease* 189, 623–628.
- McColl, M. A. and Rosenthal, C. (1994). A model of resource needs of aging spinal cord injured men. *Paraplegia* 32, 261–270.
- National Spinal Cord Injury Statistical Center (1995). *Spinal cord injury: facts and figures at a glance*. Birmingham, AL: National Spinal Cord Injury Statistical Center.
- Nielsen, M. S. (2003). Prevalence of posttraumatic stress disorder in persons with spinal cord injuries: the mediating effect of social support. *Rehabilitation Psychology* 48, 289–295.
- North, N. (1999). The psychological effects of spinal cord injury: a review. *Spinal Cord* 37, 671–679.
- Radnitz, C., Hsu, L., Tirch, D., et al. (1998). A comparison of posttraumatic stress disorder in veterans with and without spinal cord injury. *Journal of Abnormal Psychology* 107, 676–680.
- Richards, J. S. (2005). Spinal cord injury pain: impact, classification, treatment trends, and implications from translational research. *Rehabilitation Psychology* 50, 99–102.
- Richards, J. S., Kewman, D. G. and Pierce, C. A. (2000). Spinal cord injury. In: Frank, R. G. & Elliott, T. R. (eds.) *Handbook of rehabilitation psychology*, pp. 11–27. Washington DC: American Psychological Association.
- Schmitt, M. M. and Elliott, T. R. (2004). Verbal learning ability and adjustment to recent-onset spinal cord injury. *Rehabilitation Psychology* 49, 288–294.
- Scivoletto, G., Petrelli, A., Di Lucente, L., et al. (1997). Psychological investigation of spinal cord injury. *Spinal Cord* 35, 516–520.
- Shackelford, M., Farley, T. and Vines, C. L. (1998). A comparison of women and men with spinal cord injury. *Spinal Cord* 36, 337–339.
- Young, M. E., Rintala, D. H., Rossi, D., et al. (1995). Alcohol and marijuana use in a community-based sample of persons with spinal cord injury. *Archives of Physical Medicine and Rehabilitation* 76, 525–532.

Relevant Website

Neurosurgery Today (2003). Spinal cord injury FAQ. <http://www.neurosurgerytoday.org>.

Startle Response

C O Ladd, P M Plotsky and M Davis

Emory University Medical School, Atlanta, GA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 486–494, © 2000, Elsevier Inc.

Characterization of the Startle Response
Mediation of the Startle Reflex
Modulation of the Startle Reflex By Stress
Other Startle Phenomena

Glossary

Amygdala

A bilateral limbic structure located just anterior to the ventral tip of the hippocampus. The amygdala is a collection of perhaps 13 different nuclei, including the central nucleus of the amygdala, the basolateral amygdala nucleus, and the lateral amygdala nucleus. This structure is critically involved in fear as well as attention, aggression, sexual behavior, and certain aspects of appetitive behavior, such as secondary reinforcement.

Bed nucleus of the stria terminalis (BNST)

A structure surrounding the anterior commissure that is related anatomically and functionally to the amygdala. It is also believed to play a role in fear or anxiety as well as sexual arousal.

Corticotropin releasing hormone (CRH)

A 41-amino-acid peptide that coordinates the behavioral, endocrine, and autonomic indices of the mammalian stress response by acting as a neurohormone and neurotransmitter (neuromodulator) in the brain. CRH and its receptors are expressed in both the amygdaloid nuclei and the BNST.

Nucleus reticularis caudalis pontis

A group of large neurons in the pontine reticular formation that forms the final nucleus in the brain stem involved in the primary acoustic startle circuit.

The mammalian startle response consists of a rapid extension and then flexion of a series of muscles in response to an abrupt stimulus that is usually auditory but may be somatosensory, vestibular, or visual. As a simple, stimulus-controlled behavior, the mammalian startle response has been used to investigate the neuroanatomical and cellular basis for learning and behavior.

Characterization of the Startle Response

Acoustic Startle

The acoustic startle reflex (ASR) typically is elicited by a brief auditory stimulus (90–120 dB, 50 ms duration). The minimum stimulus required to elicit a response in normal rats is approximately 80–90 dB, provided that the stimulus reaches full intensity within approximately 12 ms of the stimulus onset. Even very intense auditory stimuli do not elicit a startle reflex if they have long rise times (e.g., >25 ms). In mammals, the amplitude of the startle response is proportional to the duration of the stimulus (up to 8 ms) with a direct positive relationship between startle amplitude and stimulus intensity. In rodents, high-frequency auditory stimuli are more effective than low-frequency stimuli in eliciting startle. The startle reflex has been observed in every mammal tested and is characterized by an extremely short latency. In rats, electromyographic (EMG) recordings in hindlimb muscles indicate a response latency of 8 ms following onset of an auditory stimulus.

Air Puff Startle

The startle reflex can also be elicited by an air puff that contains both tactile and acoustic components. Across laboratories, the duration of the air puff stimulus ranges from 100 to 120 ms and the stimulus intensity ranges from 9 to 75 psi. Studies in hearing-impaired rodents have shown that the auditory component of a low-intensity air puff stimuli (12.5 psi) is primarily responsible for eliciting the startle response. With more intense air puff stimuli, the auditory and tactile stimuli synergize to produce large-amplitude startle responses. Air puff startle has been used not only to elicit and measure the startle reflex, but also as an unconditioned stressor. For example, in rats, air puff stimuli can induce ultrasonic vocalizations, activate the hypothalamic-pituitary-adrenal (HPA) axis, and increase defensive withdrawal responses.

Human Startle

In humans, startle involves primarily a flexion response that is most pronounced in the upper half of the body, particularly the face, neck, and shoulders, with the minimum response being flexion of the sternocleidomastoid muscle (a latency of 14 ms). The response begins with blinking of the eyes (a latency of 35 ms) and extends caudally, causing forward head movement, facial grimacing with widening of the mouth and occasional baring of the teeth, elevation

of the shoulders, abduction of the arms, bending of the elbows, pronation of the forearms, flexion of the fingers and trunk, abdominal contraction, and knee flexion. Thus, the human startle response primarily involves flexor muscles, but may also include extensors, depending on the subject's posture at the time of stimulus onset. Startle in humans is almost always measured via the amplitude of the eyeblink, which is highly graded in magnitude and shows the same types of plasticity as the whole body response in rodents.

Rodent Startle

In rodents, startle can result in prominent extension of the forelimbs and hindlimbs to the point where all four feet come off the ground. Smaller startle responses involve flexion of the back so that the body shortens and rises into a characteristic hunched posture. In both cases, startle amplitude is measured as a downward movement of a cage that is detected by transducers sensitive to rapid movement that put out a voltage proportional to the magnitude of the startle reflex (e.g., accelerometers, piezoelectric disks, load cells, and phonograph cartridges). Startle amplitude is typically defined as the maximal transducer output during a brief period of time (e.g., 200 ms) after onset of the eliciting stimulus. By sampling cage output over brief periods of time, startle can be readily discriminated from spontaneous or drug-induced changes in locomotor or stereotyped behavior.

Mediation of the Startle Reflex

Anatomical Structures

Because the ASR has such a short latency, a simple neural pathway must mediate it. In 1982, it was proposed that acoustic startle in the rat was mediated by four synapses: three in the brain stem (the ventral cochlear nucleus, an area just medial and ventral to the ventral nucleus of the lateral lemniscus, and the nucleus reticularis pontis caudalis) and one on to motoneurons in the spinal cord. Electrolytic lesions of these nuclei eliminated acoustic startle, whereas single-pulse electrical stimulation of these nuclei elicited startlelike responses with a progressively shorter latency as the electrode was moved farther down the startle pathway. However, it has now been found that the selective destruction of cell bodies without a concomitant loss of fibers of passage in the ventral nucleus of the lateral lemniscus or the area just ventral and medial to it did not eliminate startle, whereas comparable lesions in the nucleus reticularis pontis caudalis still completely eliminated startle. These data strongly suggested, therefore, that an additional synapse in the ventral nucleus of the lateral lemniscus or the area just ventral and medial might not be necessary for

relaying auditory information from the cochlear nucleus to the pontine reticular formation. However, all data still pointed to the critical importance of the nucleus reticularis pontis caudalis in the acoustic startle reflex. Until very recently, it has been unclear how auditory information gets to this traditionally non-auditory part of the brain stem. It is now known that in both rodents and humans there is a small group of very large cells (35 μm in diameter) embedded in the cochlear nerve, called cochlear root neurons. These neurons receive direct input from the spiral ganglion cells in the cochlea, making them the first acoustic neurons in the central nervous system. They send exceedingly thick axons (sometimes as wide as 7 μm) through the trapezoid body, at the very base of the brain, directly to an area just medial and ventral to the lateral lemniscus and continue on up to the deep layers of the superior colliculus. However, they give off thick axon collaterals that terminate directly in the nucleus reticularis pontis caudalis, exactly at the level known to be critical for the acoustic startle reflex. Most important, the chemical destruction of these cochlear root neurons completely eliminates the acoustic startle reflex. Although damage to the auditory root, where the cochlear root neurons reside, has not been fully ruled out, other tests indicated that these animals could clearly orient to auditory stimuli (e.g., suppression of licking) and had normal compound action potentials recorded from the cochlear nucleus.

Although there is still some disagreement, the acoustic startle pathway may consist of only three synapses on to the (1) cochlear root neurons, (2) neurons in the nucleus reticularis pontis caudalis, and (3) motoneurons in the facial motor nucleus (pinna reflex and perhaps the eyeblink) or spinal cord (whole-body startle; [Figure 1](#)).

In humans, the obicularis oculi muscles that are innervated by the facial motor nucleus mediate the eyeblink component of startle. Because the facial motor nucleus receives a direct projection from the nucleus reticularis pontis caudalis, the eyeblink component of startle could involve a pathway consisting of cochlear root neurons, the nucleus reticularis pontis caudalis, and the facial motor nucleus. However, this simple neural pathway may not be able to account for the relatively long latency of the eyeblink reflex (35 ms), so other pathways have been suggested.

Neurotransmitters that Mediate the Startle Reflex

Auditory nerve to cochlear root neuron synapse At the present time, there are no direct data on the identity of the neurotransmitter that may mediate startle at this level. Preliminary experiments in rodents suggest it may be glutamate acting on non-*N*-methyl-D-aspartate (NMDA) receptors because a local infusion

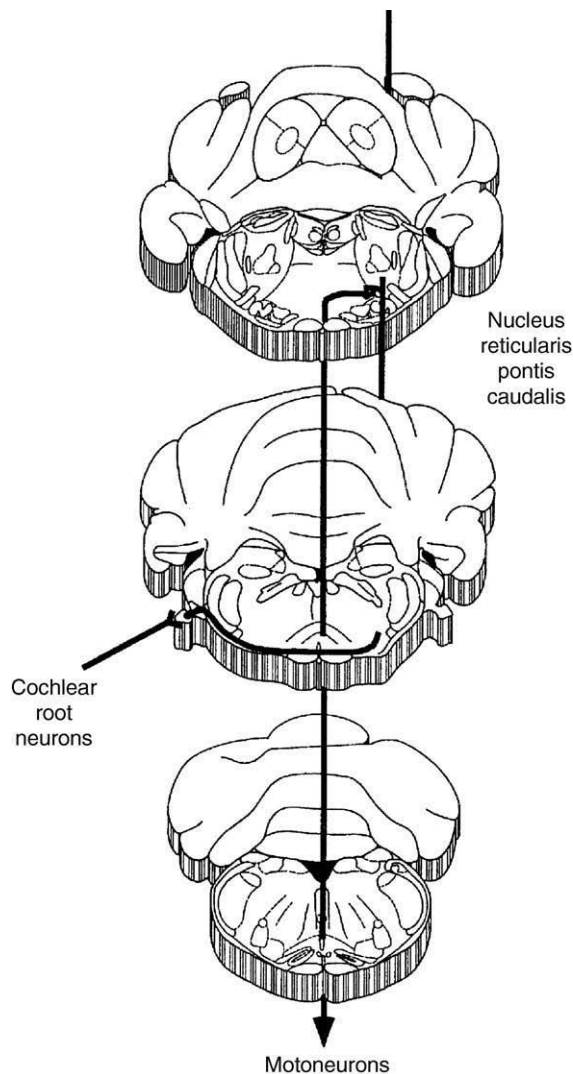


Figure 1 Schematic diagram of a primary acoustic startle pathway. Auditory nerve fibers synapse onto cochlear root neurons embedded in the auditory nerve. Axons from these cells project through the ventral acoustic stria and send projections to the nucleus reticularis pontis caudalis (PnC). Projections of cells in the PnC form the reticulospinal tract, which make monosynaptic and polysynaptic synapses in the spinal cord. These cells also project to the facial motor nucleus in areas critical for the pinna component of startle in rats and perhaps the eyeblink component of startle in humans (not shown). Lesions of the cochlear root neurons, the ventral acoustic stria, the PnC, or reticulospinal tract eliminate the acoustic startle response.

of the non-NMDA antagonist CNQX in the vicinity of the root neurons depressed startle, whereas local infusion of the NMDA antagonist AP5 did not.

Cochlear root neuron to reticulospinal neuron synapse
Glutamate acting on both NMDA and non-NMDA receptors may well be the neurotransmitter that mediates startle at the level of the nucleus reticularis pontis caudalis. Local infusion of either CNQX or

AP5 in rats chronically implanted with bilateral canulas aimed at the nucleus reticularis pontis caudalis significantly reduced startle amplitude by as much as 70–80%. AP5 and CNQX attenuated startle over a dose range, which indicated that the nucleus reticularis pontis caudalis is extremely sensitive to these compounds. Thus, small changes in neurotransmitter activity in this area may have a profound effect on the expression of the acoustic startle reflex.

Reticulospinal neurons to motor neuron synapse

Intrathecal infusion (into the subarachnoid space of the spinal cord to allow drugs to diffuse in the vicinity of motoneurons) of the non-NMDA antagonist CNQX or the NMDA antagonist AP5 each reduced the amplitude of the whole-body startle reflex in a dose-dependent manner. When EMG activity in the hindlimbs was used to define startle, intrathecal administration of a combination of AP5 and CNQX completely eliminated all EMG components of the startle response, consistent with the hypothesis that excitatory amino acids actually mediate startle at the spinal level. Moreover, when the EMG components of the startle reflex were separated into early or late components, CNQX preferentially eliminated the early components (a mean latency of 8.29 ms), whereas AP5 preferentially eliminated the later components (a mean latency of 14.57 ms or longer).

The fact that startle appears to involve both non-NMDA and NMDA receptors at the spinal level can be interpreted in a number of ways. The simplest possibility is that glutamate is released from reticulospinal tract terminals, which form monosynaptic connections onto motoneurons expressing both NMDA and non-NMDA receptors. The short-latency EMG components then reflect the activation of non-NMDA receptors, whereas the longer-latency EMG components reflect the activation of NMDA receptors. A second possibility is that reticulospinal neurons make both monosynaptic and polysynaptic connections via interneurons onto spinal motoneurons. The short-latency EMG components then reflect the monosynaptic connections on to non-NMDA receptors, whereas the longer-latency EMG components reflect polysynaptic connections involving NMDA receptors. A third possibility is that the short-latency EMG components of startle reflect the activation of non-NMDA receptors by the short-latency startle pathway described previously, whereas the later components reflect activity in other brain-stem structures, which then eventually activate NMDA receptors on spinal motoneurons. This idea can be tested by the direct stimulation of reticulospinal neurons in conjunction with intrathecal infusion of excitatory amino acid antagonists.

In summary, the excitatory amino acid glutamate may well mediate startle at each of the three central synapses along the acoustic startle pathway. NMDA and non-NMDA receptors appear to be involved at both the spinal cord and nucleus reticularis pontis caudalis. Processes involved in the release of glutamate, as well as changes in the sensitivity of glutamate receptors, are expected to affect startle and startle plasticity.

Modulation of the Startle Reflex By Stress

The Fear-Potentiated Startle Effect

In rats, the amplitude of the startle reflex is increased when elicited by an auditory stimulus shortly after the onset of a light paired previously with a foot shock. This fear-potentiated startle effect occurs only following prior light–shock pairings and not when lights and shocks have been presented in an unpaired or random relationship, indicating its dependence on prior Pavlovian fear conditioning. It is probable that the light produces a state of fear, which increases reflexive behavior, because drugs such as buspirone or valium, which

reduce fear and anxiety in humans, block the increase in startle in the presence of the light but do not systematically alter startle in the absence of light when appropriate doses are used. When the eyeblink component of the startle reflex is measured in humans, fear-potentiated startle can be produced using very similar conditioning procedures to those used in rodents. Humans also show an increase in startle amplitude simply when they are told that a shock might occur after presentation of a light or when they see threatening pictures such as a gun or an attacking dog. Conversely, pleasant pictures shown to humans or a light paired with food shown to rats reduce startle amplitude.

Figure 2 summarizes our knowledge of the neural pathways involved in the fear-potentiated startle effect using a visual conditioned stimulus. Visual information goes through the visual thalamus to the perirhinal cortex, which then projects to the amygdala, which has a direct projection to the nucleus reticularis pontis caudalis in the startle pathway. The lateral posterior nucleus of the thalamus, which receives heavy input from the retina, projects directly to the perirhinal cortex. Lesions of the lateral posterior nucleus of the thalamus, combined with lesions of the dorsal and ventral nuclei of the lateral geniculate, completely block

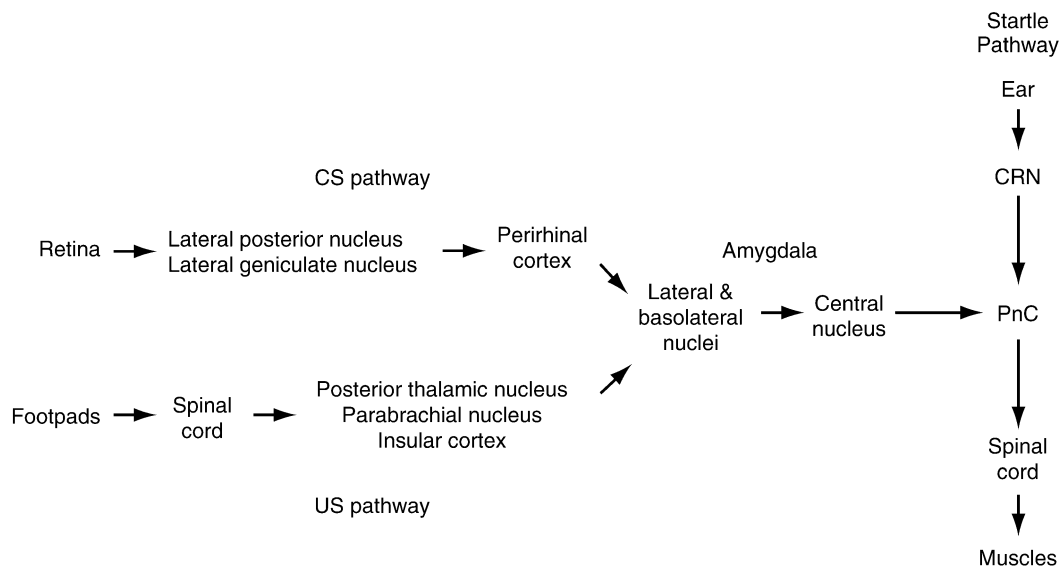


Figure 2 Working hypothesis of the neural pathways involved in fear-potentiated startle using a visual-conditioned stimulus and a foot shock-unconditioned stimulus. Visual information from the retina goes in parallel to the dorsal lateral geniculate nucleus and the lateral posterior nucleus in the thalamus. Lesions of both of these nuclei together after light–shock pairings eliminate fear-potentiated startle to a light. Lesions of either area alone or lesions of the visual cortex alone, which receive inputs from the dorsolateral geniculate nucleus, do not block the expression of fear-potentiated startle to a light. The lateral posterior nucleus projects directly to the perirhinal cortex, whereas the dorsolateral nucleus projects to the perirhinal cortex indirectly via the visual cortex. The perirhinal cortex projects to lateral and basolateral nuclei of the amygdala, which project to the central nucleus. The central nucleus has direct projections to the nucleus reticularis pontis caudalis, an obligatory part of the acoustic startle pathway; however, indirect projections via a synapse in the rostral midbrain probably are involved as well. Shock information is relayed to the lateral and basolateral amygdala via parallel inputs arising from the insular cortex and posterior thalamic nucleus. Lesions of both these areas together prior to light–shock training block the acquisition of fear-potentiated startle. However, lesions of both of these areas following training have no effect. Lesions of either area alone fail to block acquisition or expression of fear-potentiated startle. CRN, cochlear root neurons; CS, conditioned stimulus; PnC, nucleus reticularis pontis caudalis; US, unconditioned stimulus.

fear-potentiated startle using a visual-conditioned (but not an auditory-conditioned) stimulus. Lesions of either structure alone have no effect. Furthermore, posttraining lesions of the visual cortex, to which the dorsal nucleus of the lateral geniculate projects, do not affect fear-potentiated startle. Hence, visual information gets to the perirhinal cortex either directly from the lateral posterior nucleus or indirectly from the dorsal nucleus of the lateral geniculate via the visual cortex. Lesion data suggest that either one of these pathways is sufficient to mediate fear-potentiated startle using a visual-conditioned stimulus. The perirhinal cortex has heavy reciprocal projections with the lateral and/or basolateral nuclei of the amygdala. Relatively small lesions of the perirhinal cortex completely block fear-potentiated startle, provided the lesion includes an area of perirhinal cortex both dorsal and ventral to the rhinal sulcus.

Selective destruction of cell bodies in the lateral and basolateral nuclei of the amygdala completely block fear-potentiated startle using either a visual- or an auditory-conditioned stimulus when the lesions were made either before or after training. Lesions of the central nucleus (CeA), innervated by the basolateral nucleus of the amygdala, block the expression of fear-potentiated startle using either a visual- or an auditory-conditioned stimulus. The CeA has both direct and indirect projections via the central gray to the nucleus reticularis pontis caudalis, and lesions at several points along this pathway block the expression of fear-potentiated startle.

When shock is paired with light during fear-potentiated startle training, shock information is relayed to the basolateral amygdala via two parallel pathways. These include the caudal granular/dysgranular insular cortex and the posterior intralaminar nuclei of the thalamus. Pretraining lesions of both of these areas together, but not either one alone, blocked the acquisition of fear-potentiated startle. However, posttraining combined lesions of these same brain areas did not prevent the expression of conditioned fear.

Sensitization of Startle by Foot Shocks

When rats are presented with a series of startle stimuli before and after a series of 10 foot shocks presented at a 10-s interstimulus interval, startle shows a progressive increase in amplitude that peaks in approximately 10 min. The magnitude of this startle sensitization effect is directly related to the number or intensity of intervening shocks. Moreover, shocks given prior to presenting any startle stimuli also lead to an elevation of startle, indicating that enhanced startle results from sensitization rather than simply from

dishabituation. Like fear-potentiated startle, lesions of the amygdala block the shock sensitization of startle.

Noise-Enhanced Startle

The startle reflex is increased when elicited in the presence of steady background noise as well as following a period of sustained high-level noise. Although this may reflect an effect of noise stress, noise-enhanced startle is not blocked by lesions of the amygdala and shows inconsistent sensitivity to anxiolytic compounds.

Light-Enhanced Startle

When the startle reflex is elicited by a loud sound 5–20 min after a bright light has been turned on, there is an increase in the startle reflex. This effect is also decreased significantly by drugs that reduce anxiety in humans and may indicate an unconditioned anxiogenic effect of bright light that enhances startle. Humans show a significant increase of startle amplitude (i.e., of the eyeblink response) in the dark. The species difference may reflect fear of the light in a nocturnal species (rats) compared to fear of the dark in a diurnal species (humans). When the lights go off, many people feel more anxious, especially if they were afraid of the dark when they were young. In patients with posttraumatic stress disorder, startle is increased in the dark to a greater degree than that seen in combat controls. These individuals report that the darkness makes them think of being back at their guard post in Vietnam and anxious about being hit by an incoming mortar.

Like fear-potentiated startle, local infusion of the glutamate antagonist NBQX into the basolateral nucleus of the amygdala blocks light-enhanced startle. However, unlike fear-potentiated startle, NBQX infused into the BNST, but not the CeA, blocks light-enhanced startle. The BNST also receives input from the basolateral nucleus of the amygdala and projects directly to the startle pathway at the level of the nucleus reticularis pontis caudalis. Thus, a pathway going from the visual thalamus to basolateral amygdala to BNST to nucleus reticularis pontis caudalis may mediate light-enhanced startle.

Corticotropin Releasing Hormone-Enhanced Startle

The intraventricular administration of the peptide CRH produces a variety of behavioral and neuroendocrine effects similar to those seen during fear and anxiety. Intraventricular administration of the CRH antagonist α -helical CRH(9–41) blocks the behavioral and neuroendocrine effects of natural stressors

or conditioned fear. The intraventricular administration of CRH increases the acoustic startle reflex. This effect can be blocked by the benzodiazepine chlordiazepoxide, suggesting that the excitatory effect of CRH on startle reflects an anxiogenic effect of the hormone. This effect still occurs after adrenalectomy, indicating that it is not dependent on the release of corticosterone from the adrenal glands. Thus, CRH-enhanced startle represents yet another example of a long-lasting, unconditioned, anxiogenic effect. Chemically induced lesions of the BNST, but not of the amygdala, block CRH-enhanced startle. Moreover, local infusion of CRH into the BNST, but not into the CeA, increases startle amplitude. Hence, CRH-enhanced startle is more similar to light-enhanced startle than fear-potentiated startle in terms of its dependence on the BNST versus the CeA.

Drug Withdrawal

Several studies in both rats and humans have shown that alcohol withdrawal increases acoustic startle amplitude. In rats, nicotine withdrawal or withdrawal from chronic diazepam also increases startle amplitude. In contrast, naloxone-precipitated opiate withdrawal depresses startle markedly. Withdrawal from stimulants such as cocaine or amphetamine has little effect on startle.

Maternal Separation

Exposing rat pups to moderate periods of maternal separation (180 min/day during postnatal days 2–14) increases the adult HPA response to acute air puff startle compared to nonhandled or normally reared rat pups. These animals also show an elevation of the acoustic startle reflex. In contrast, brief (15-min) periods of maternal separation during that same neonatal period renders these animals hypo-responsive to air puff startle (and other psychological stressors) as adults compared to nonhandled rat pups. This alteration in sensitivity and reaction to an acute air puff stimulus may be due to persistent changes in the central GABA, norepinephrine, and CRH systems, as well as in hippocampal glucocorticoid receptor expression.

Other Stressors

In rats, startle is increased during frustrative nonreward, but is decreased following cold swim stress or food, water, or sleep deprivation. In humans, the eye-blink component of startle is increased when elicited in the presence of noxious odors or heat (47–49°C) applied to the thumb. However, anticipation of giving a speech produces a pronounced decrease in startle amplitude, even though this is a very anxiogenic event

for humans. Although the pattern of data is by no means clear, it appears that stressors that direct attention inward (food, water, sleep deprivation, and hypothermia) depress startle, whereas stressors that direct attention outward (threat of shock, a predator, and a harmful environment) tend to increase startle amplitude.

Other Startle Phenomena

Habituation and Dishabituation

When a startle stimulus is repeated relatively rapidly (e.g., once every 10–15 s), the amplitude of the startle reflex will generally decline over repeated test trials, a phenomenon called habituation. In general, the rate of response decrement within a test session is faster with short intervals than with long intervals between stimuli (interstimulus intervals) and with weak startle stimuli than with strong startle stimuli. In contrast, across-session habituation (e.g., measured several hours later or the next day) is generally greater following long intertrial intervals than short intertrial intervals and loud startle stimuli than weak startle stimuli, provided that a common test interval or test intensity is used to assess the level of habituation.

If, following habituation, a novel intense stimulus is presented, the startle amplitude may increase above its habituated level, a phenomenon called dishabituation. It is generally believed that dishabituation is a special case of sensitization and that habituation and sensitization represent separate processes that summate to form the final response amplitude at any given time. In fact, a loud startle stimulus may produce both habituation and sensitization, which jointly determine startle amplitude.

By eliciting the startle reflex electrically via the implantation of tiny electrodes at different levels of the startle pathway, researchers concluded that short-term habituation may result from a change in transmission at the nucleus reticularis pontis caudalis. Because lesions in a variety of brain areas fail to alter the rate of short-term habituation of the startle reflex, it may involve processes intrinsic to the startle pathway, such as receptor desensitization or the buildup of a very local inhibitory process.

However, brain areas extrinsic to the startle pathway itself may mediate long-term habituation. Leaton and Supple have shown that lesions of the cerebellar vermis, which is not part of the acoustic startle pathway, significantly attenuates the decrease in startle amplitude seen across daily test sessions, without affecting the rate of response decrement within test sessions. This blockade of long-term habituation was observed with two different stimulus intensities and

cannot be explained by ceiling or floor effects caused by cerebellar lesions.

Sensitization of startle may also involve brain areas extrinsic to the startle pathway. Borszcz and colleagues found that loud startle stimuli produce sensitization that can best be described as fear of the experimental context in which startle is measured. Reintroducing animals into this context produces a good deal of freezing, a reliable index of fear. Fear of the context elevates startle on subsequent test sessions, leading to a reduction in the amount of long-term response decrement. Treatments such as lesions of the amygdala or drugs such as diazepam, which reduce fear in many situations, facilitate the degree of long-term response decrement, presumably by blocking sensitization and hence allowing long-term habituation to be revealed.

Prepulse Inhibition

The amplitude of the startle response is reduced when the strong startle-eliciting stimulus is preceded shortly (30–500 ms) by a weaker stimulus, or prepulse. This phenomenon, termed prepulse inhibition (PPI), may be a measure of the gating or filtering of sensory information. PPI may be mediated by prepulse inputs to the inferior and superior colliculi, which activate the pedunculopontine tegmental nucleus. This structure, in turn, sends inhibitory cholinergic projections to the caudal pontine reticular nucleus.

A cortico-pallido-striato-pontine circuit connecting limbic cortical regions and subcortical dopaminergic systems to the pedunculopontine tegmental nucleus modulates PPI. These pathways are primarily glutamatergic, dopaminergic, or GABAergic. Both glutamate and dopamine have been implicated in the modulation of PPI, glutamate as a facilitator via subcortical NMDA receptors and dopamine as an inhibitor via D2 receptors in the nucleus accumbens. In the presence of excess mesolimbic dopamine, the pedunculopontine tegmental nucleus is inhibited both directly and indirectly (through the ventral pallidum) by accumbal efferents. The inhibition of the pedunculopontine tegmental nucleus would then block the prepulse from activating this structure and thus reduce its inhibitory effect.

The disruption of prepulse inhibition in rats by dopamine agonists or stress has received particular consideration as an animal model of the putative attentional impairments underlying the positive symptoms of schizophrenia thought to be associated with excessive dopamine transmission. Reduced PPI is also observed in other disorders characterized by sensorimotor and cognitive gating deficiencies, including Huntington's disease, obsessive-compulsive disorder,

attention deficit disorder, and Tourette's syndrome. Schizophrenic patients are often unable to sustain voluntary attention or filter irrelevant thoughts and stimuli. Several neuroimaging studies of schizophrenics have demonstrated anatomical, metabolic, and cellular abnormalities in the cortico-striato-pallido-pontine circuitry modulating PPI, which may underlie these deficits. This model is currently being used to investigate the genetic and early environmental contributions to PPI deficiency, as well as to screen potential antipsychotics.

See Also the Following Articles

Evolutionary Origins and Functions of the Stress Response; Fear; Fight-or-Flight Response.

Further Reading

- Boulis, N. M., Kehne, J. H., Miserendino, M. J. D., et al. (1990). Differential blockade of early and late components of acoustic startle following intrathecal infusion of 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX) or D, L-2-amino-5-phosphonovaleric acid (AP-5). *Brain Research* 520, 240–246.
- Davis, M. (1984). The mammalian startle response. In: Eaton, R. C. (ed.) *Neural mechanisms of startle behavior*, pp. 287–351. New York: Plenum.
- Davis, M. and Astrachan, D. I. (1978). Conditioned fear and startle magnitude: effects of different footshock or back-shock intensities used in training. *Journal of Experimental Psychology: Animal Behavior Processes* 4, 95–103.
- Davis, M., Falls, W. A., Campeau, S., et al. (1993). Fear-potentiated startle: a neural and pharmacological analysis. *Behavioural Brain Research* 58, 175–198.
- Davis, M., Gendelman, D. S., Tischler, M. D., et al. (1982). A primary acoustic startle circuit: lesion and stimulation studies. *Journal of Neuroscience* 6, 791–805.
- Davis, M., Parisi, T., Gendelman, D. S., et al. (1982). Habituation and sensitization of startle reflexes elicited electrically from the brainstem. *Science* 218, 688–690.
- Dawson, M. E., Schell, A. M. and Bohmelt, A. H. (1999). *Startle modification: implications for neuroscience, cognitive science, and clinical science*. Cambridge, UK: Cambridge University Press.
- Dunn, A. J. and Berridge, C. W. (1990). Physiological and behavioral responses to corticotropin-releasing factor administration: is CRF a mediator of anxiety or stress responses. *Brain Research Reviews* 15, 71–100.
- Ebert, U. and Koch, M. (1992). Glutamate receptors mediate acoustic startle input to the reticular brain stem. *NeuroReport* 3, 429–432.
- Engelmann, M., Thrivikraman, K. V., Su, Y., et al. (1996). Endocrine and behavioral effects of airpuff startle in rats. *Psychoneuroendocrinology* 21, 391–400.
- Geyer, M. A., Swerdlow, N. R., Mansbach, R. S., et al. (1990). Startle response models of sensorimotor gating

- and habituation deficits in schizophrenia. *Brain Research Bulletin* **25**, 485–498.
- Grillon, C. and Davis, M. (1997). Fear-potentiated startle conditioning in humans: effects of explicit and contextual cue conditioning following paired versus unpaired training. *Psychophysiology* **34**, 451–458.
- Grillon, C., Morgan, C. A., Davis, M., et al. (1998). Effect of darkness on acoustic startle in Vietnam veterans with post-traumatic stress disorder. *American Journal of Psychiatry* **155**, 812–817.
- Grillon, C., Pellowski, M., Merikangas, K. R., et al. (1997). Darkness facilitates the acoustic startle reflex in humans. *Biological Psychiatry* **42**, 461–471.
- Groves, P. M. and Thompson, R. F. (1970). Habituation: a dual process theory. *Psychological Review* **77**, 419–450.
- Koch, M. and Schnitzler, H.-U. (1997). The acoustic startle response in rats: circuits mediating evocation, inhibition and potentiation. *Behavioural Brain Research* **89**, 35–49.
- Landis, C. and Hunt, W. (1939). *The startle paradigm*. New York: Farrar and Rinehart.
- Lang, P. J. (1995). The emotion probe. *American Psychologist* **50**, 372–385.
- Lee, Y. and Davis, M. (1997). Role of the hippocampus, bed nucleus of the stria terminalis and amygdala in the excitatory effect of corticotropin releasing (CRH) hormone on the acoustic startle reflex. *Journal of Neuroscience* **17**, 6434–6446.
- Lee, Y., Lopez, D., Meloni, E. G., et al. (1994). A primary acoustic startle reflex circuit: role of auditory root neurons and the nucleus reticularis pontis caudalis. *Society for Neuroscience Abstracts* **20**, 1009.
- Lee, Y., Lopez, D. E., Meloni, E. G., et al. (1996). A primary acoustic startle circuit: obligatory role of cochlear root neurons and the nucleus reticularis pontis caudalis. *Journal of Neuroscience* **16**, 3775–3789.
- Lopez, D. E., Merchan, M. A., Bajo, V. M. and Saldana, E. (1993). The cochlear root neurons in the rat, mouse and gerbil. In: Merchan, M. A. (ed.) *The mammalian cochlear nuclei: organization and function*, pp. 291–301. New York: Plenum Press.
- Miserendino, M. J. D. and Davis, M. (1993). NMDA and non-NMDA antagonists infused into the nucleus reticularis pontis caudalis depress the acoustic startle reflex. *Brain Research* **623**, 215–222.
- Swerdlow, N. R., Caine, S. B., Braff, D. L., et al. (1992). The neural substrates of sensorimotor gating of the startle reflex: a review of recent findings and their implications. *Journal of Psychopharmacology* **6**, 176–190.
- Swerdlow, N. R. and Geyer, M. A. (1998). Using an animal model of deficient sensorimotor gating to study the pathophysiology and treatment of schizophrenia. *Schizophrenia Bulletin* **24**, 285–301.
- Walker, D. L. and Davis, M. (1997). Anxiogenic effects of high illumination levels assessed with the acoustic startle paradigm. *Biological Psychiatry* **42**, 461–471.
- Walker, D. L. and Davis, M. (1997). Double dissociation between the involvement of the bed nucleus of the stria terminalis and the central nucleus of the amygdala in light-enhanced versus fear-potentiated startle. *Journal of Neuroscience* **17**, 9375–9383.
- Yeomans, J. S. and Franklin, P. W. (1996). The acoustic startle reflex: neurons and connections. *Brain Research Reviews* **21**, 301–314.

Steroid Hormone Receptors

M Beato

Centre de Regulació de Genòmica, UPF, Barcelona, Spain

J Klug

University of Giessen, Giessen, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M Beato and J Klug, volume 3, pp 495–504, © 2000, Elsevier Inc.

Historical Development

Steroid Hormone Receptors Form a Subgroup within the Nuclear Receptor Family

Domain Structure of Steroid Hormone Receptors

Individual Family Members

Heterooligomeric Complex and Ligand and DNA Binding

Activation of Transcription

Involvement of the Glucocorticoid Receptor in the Stress Response

Cross-Talk with Other Signal Transduction Pathways

Glossary

Chromatin

A complex of nuclear DNA and basic proteins called histones. Every 146 bp of nuclear DNA is wrapped around a histone octamer into a nucleosome. A 30-bp DNA linker connects one nucleosome to the next, forming a nucleosome necklace. Intracellular transcription factors that are induced by the specific and high-affinity binding of ligands to exert positive or negative effects on the expression of target genes by different mechanisms.

Steroid hormone receptors (SHRs)

<i>Steroid hormones</i>	Lipophilic hormones derived from cholesterol and synthesized in the adrenal cortex (glucocorticoids, mineralocorticoids, and adrenal androgens), the testes (testicular androgens and estrogen), and the ovary and placenta (estrogens and progestagens).
<i>Transcription factors (TFs)</i>	Nuclear proteins interacting with the regulatory regions of genes, promoters, and enhancers. They can be divided into general transcription factors, forming the transcription initiation complex required for the transcription of all genes, and sequence-specific transcription factors, which bind to specific DNA sequences present in only a subset of genes and carry activation or repression domains that serve to activate or repress transcription.

Steroid hormone receptors (SHRs) are ligand-activated transcription factors and the key mediators of steroid hormone action. Lipophilic steroid hormones pass the cell membrane of target cells by simple diffusion and bind to soluble intracellular SHRs that are complexed to chaperones, such as Hsp90, or to a small fraction of membrane-tethered SHRs. Chaperones are proteins that help other proteins fold and prevent aggregation. The binding of agonistic or antagonistic ligands leads to different allosteric changes of receptors, making them competent to exert positive or negative effects on the expression of target genes by three mechanisms. First, ligand-bound SHR complexes can bind to short, specific DNA sequences in the vicinity of target genes, termed hormone response elements (HREs). The HRE-recruited hormone–receptor complexes are then able to relay activating or repressing signals to the target-gene transcription machinery (Figure 1a). Second, the SHR response can be integrated into intracellular signal transduction pathways that transmit an extracellular growth factor signal via mitogen-activated protein kinases (MAPKs) to a nuclear transcription factor that activates a target gene (Figure 1b). Third, through protein–protein interactions, SHRs can repress the activity of many genes that are switched on during the stress response (transrepression) (Figure 1c).

Historical Development

The gonads and adrenal gland produce five major groups of steroid hormones: estrogens, progestagens, androgens, glucocorticoids, and mineralocorticoids. The effects and clinical symptoms associated with a reduced level of these hormones were known long before the hormone concept was developed. Aristotle, for instance, identified the consequences of castration in humans and birds, and in 1855 Addison described

the symptoms of chronic adrenal insufficiency. Moreover, since the nineteenth century it has been known that female sex steroid hormones play a pivotal role in promoting mammary tumor growth, similar to androgens promoting the growth of prostatic neoplasms. Between the tenth and sixteenth centuries, Chinese alchemists developed empirical methods to purify steroids to near homogeneity. In the first half of the twentieth century, it became possible to isolate individual steroids from extracts derived from source tissues by employing biological complementation assays. When tritium and ^{14}C -labeled hormone compounds of high specific activity became available in the late 1950s, it became possible to follow the fate of the steroid hormones up to their target tissues. Thus, it could be shown that the hormone itself, not a metabolite, produced the response via the activation of gene expression mechanisms. The concept that steroid hormones are involved in transcriptional control was also triggered by the observation of ecdysone-induced giant chromosome puffs by Clever and Karlson in 1960. Some 10 years later the groups of Jensen and Gorski had established a two-step model that involved the binding of the hormone to specific high-affinity SHRs within the target cells, followed by the activation of the hormone–receptor complex in order to induce the expression of hormone responsive genes (Figure 1a). After another 10 years, the hormone receptors for glucocorticoids and estrogens were cloned and thus became the first molecularly defined transcription factors for RNA polymerase II. At around the same time, the cloning of some SHR targets, such as the human metallothionein gene and the mouse mammary tumor virus genome, led to the identification of the first HREs. Within the last 15 years, many details of the SHR signal transduction pathway, including new mechanisms, have been discovered (Figures 1b–c).

Steroid Hormone Receptors Form a Subgroup within the Nuclear Receptor Family

In addition to the hormone receptors for glucocorticoids and estrogens, which were cloned first, receptors for the remaining three major steroid hormone classes have been identified and characterized extensively, as well as a second estrogen receptor and three estrogen-related receptors. All SHRs are characterized by a central DNA-binding domain (DBD) that targets the receptor to the HREs and by a ligand-binding domain (LBD) that is required for switching the receptor's functions. When the *v-erbA* oncogene was cloned, it turned out to also contain a DBD and a LBD that were homologous to the cognate regions of

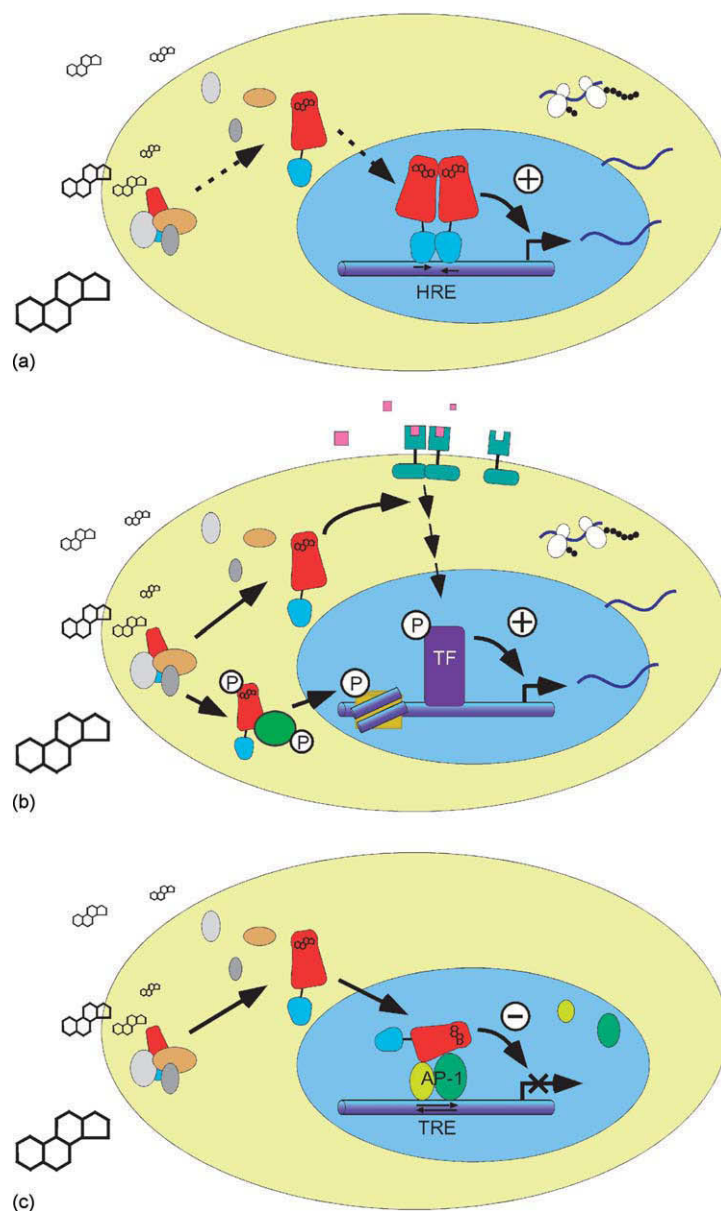


Figure 1 Three implementations of the SHR signal transduction pathway. a, Ligand-bound SHR complexes binding to short, specific DNA sequences in the vicinity of target genes; b, SHR response being integrated into intracellular signal transduction pathways; c, SHRs repressing the activity of genes that are switched on during the stress response. In a, the lipophilic steroid hormones (depicted by the chemical formula showing the central sterane building block of all steroid hormones) pass the cell membrane of target cells by simple diffusion and bind to soluble intracellular SHRs that are complexed to chaperones such as Hsp90 (shown as ellipses with various colors). Binding of the hormone leads to an allosteric change of the receptor, resulting in the dissociation of chaperones (dashed arrow). The liganded SHR-complexes dimerize, bind to hormone response elements (HREs, dashed arrow) in the vicinity of target genes, and transmit activating (or repressing) signals to the target genes' transcription machinery. In b, an extracellular polypeptide growth factor signal (pink squares) is received by the target cell via membrane receptors with kinase activity that dimerize on ligand binding. The dimerized growth factor receptors relay the signal through a cascade of mitogen-activated protein kinases (MAPKs; arrows) to sequence specific transcription factors (TFs) that are induced (or repressed) by phosphorylation (P) to activate the target gene. Hormone-activated SHRs can bypass a growth factor receptor-mediated signal by activating MAPKs and vice versa. In addition, hormone receptors form complexes with cytoplasmic kinases (green ellipsoid) such as ERK and MSK1 that are activated (phosphorylated) on hormone binding and activate (phosphorylate) the receptor. The complex of activated receptor and kinase(s) is recruited to the target promoter where it phosphorylates histones (brown rectangle) as a prerequisite for chromatin remodeling and transcriptional activation. In c, many immunoregulatory genes as well as genes involved in inflammation are positively regulated by the heterodimeric sequence-specific transcription factor activator protein 1 (AP-1) recognizing a TPA response element (TRE) within the promoter of a target gene. Through protein-protein interactions, ligand-activated SHRs are able to repress in trans the AP-1 activity (transrepression).

the known SHRs. Later on, the *c-erbA* locus was identified as the receptor for the thyroid hormone T_3 , which is also a hormone found in the nucleus like steroid hormones. The identification of another structurally related receptor for the vitamin A metabolite retinoic acid solidified the concept of a nuclear receptor superfamily. Today, more than 520 different nuclear receptors have been described from more than 120 metazoan species, among them receptors for the known nuclear hormones and a vast number of so-called orphan receptors with no, or unknown and possibly novel, ligands. In order to avoid the coining of several trivial names for the same gene, an official nomenclature system has been in use since 1999 (NUREBASE). The SHRs and steroid hormone-like receptors form the nuclear receptor subfamily 3 (NR3), consisting of three groups: A, B, and C (Figure 2a). NR3 includes two receptors for estrogens ($ER\alpha$ and $ER\beta$, or NR3A1 and NR3A2), three estrogen-related orphan receptors ($ERR\alpha$, $ERR\beta$, and $ERR\gamma$, or NR3B1, NR3B2, and NR3B3), and one receptor each for the other major steroid hormone classes: glucocorticoids (GR, or NR3C1), mineralocorticoids (MR, or NR3C2), progestagens (PR, or NR3C3), and androgens (AR, or NR3C4).

Domain Structure of Steroid Hormone Receptors

All SHRs are modular proteins composed of distinct regions, as shown in Figure 2b (top line). When the chicken estrogen receptor cDNA became available, comparison with the human estrogen and glucocorticoid receptor sequences led to the region A–F nomenclature, which is used for all members of the whole nuclear receptor superfamily today. From numerous functional and structural analyses, it became clear that these distinct regions correspond to functional and structural units called domains. Region C (the DBD) and region E (the LBD) display a high degree of sequence conservation, whereas no significant conservation is detected between paralogous SHRs for the regions A/B, D, and F.

Region D is considered a flexible hinge region between the DBD and LBD. Its very N-terminus is an integral part of the DBD and is involved in DBD dimerization. The F region is essential for hormone binding in the PR, GR, and AR but not in the $ER\alpha$. This domain is also important for the discrimination between agonistic and antagonistic hormone ligands. The A region is highly conserved between chicken and human estrogen receptors, but this distinction is much less clear in the other steroid receptors. Therefore, regions A and B are combined into an A/B region in most cases.

Regions C and E are not only responsible for DNA and ligand binding, respectively, but they do encode other functions as well (Figure 2b, bottom). The intracellular distribution of steroid receptors is the result of nuclear-cytoplasmic diffusion and cytoplasmic-nuclear shuttling. At equilibrium, the majority of steroid receptors are in the nucleus due to the presence of nuclear localization signals (NLS), which are believed to be required for nuclear pore recognition. A constitutive NLS is located at the border of regions C and D, whereas a second NLS in the LBD is ligand-dependent. Intracellular localization is different for GR and MR because hormone-induced nuclear translocation has been reported in these cases.

Most SHRs bind as homodimers to their cognate response elements within the promoters of target genes, which is a specific characteristic of this subfamily within the nuclear receptor superfamily. It is also well known that the predominant form of the GR is monomeric in solution and that dimerization occurs only after binding to an HRE, accounting for the cooperativity observed during DNA binding. Responsible for dimerization are a weak dimerization domain encoded within the DBD and a strong dimerization interface provided by the LBD.

Ligand binding confers transcriptional competence on to SHRs that is exerted in most receptors by two independent transactivation functions: a constitutively active one in the A/B region located close to the DBD, referred to as activation function 1 (AF-1; also called τ -1 or enh-1), and a ligand-inducible activation function in the LBD, called AF-2. AFs connect the receptor to the transcription apparatus via direct interactions with basal transcription factors, sequence-specific transcription factors, and/or coregulators.

Individual Family Members

The human estrogen receptor (now called $hER\alpha$, Figure 2a) was cloned from a cDNA expression library produced from the breast cancer cell line MCF-7. A second estrogen receptor, $ER\beta$, has been characterized that is very similar to $ER\alpha$ in terms of structure and function but also shows subtle and important functional differences. Both receptors bind the ligands estradiol, diethylstilbestrol, estriol, and estrone with high affinity. cDNA clones for $hERR\alpha$ and $hERR\beta$ were isolated from a human testis cDNA library using the $hER\alpha$ DBD as a probe. Mouse $ERR\gamma$ was identified by yeast two-hybrid screening, using the transcriptional coactivator GR interacting protein 1 as bait. It is still unclear whether the ERRs are true orphan receptors in search of a ligand or are ligand-independent.

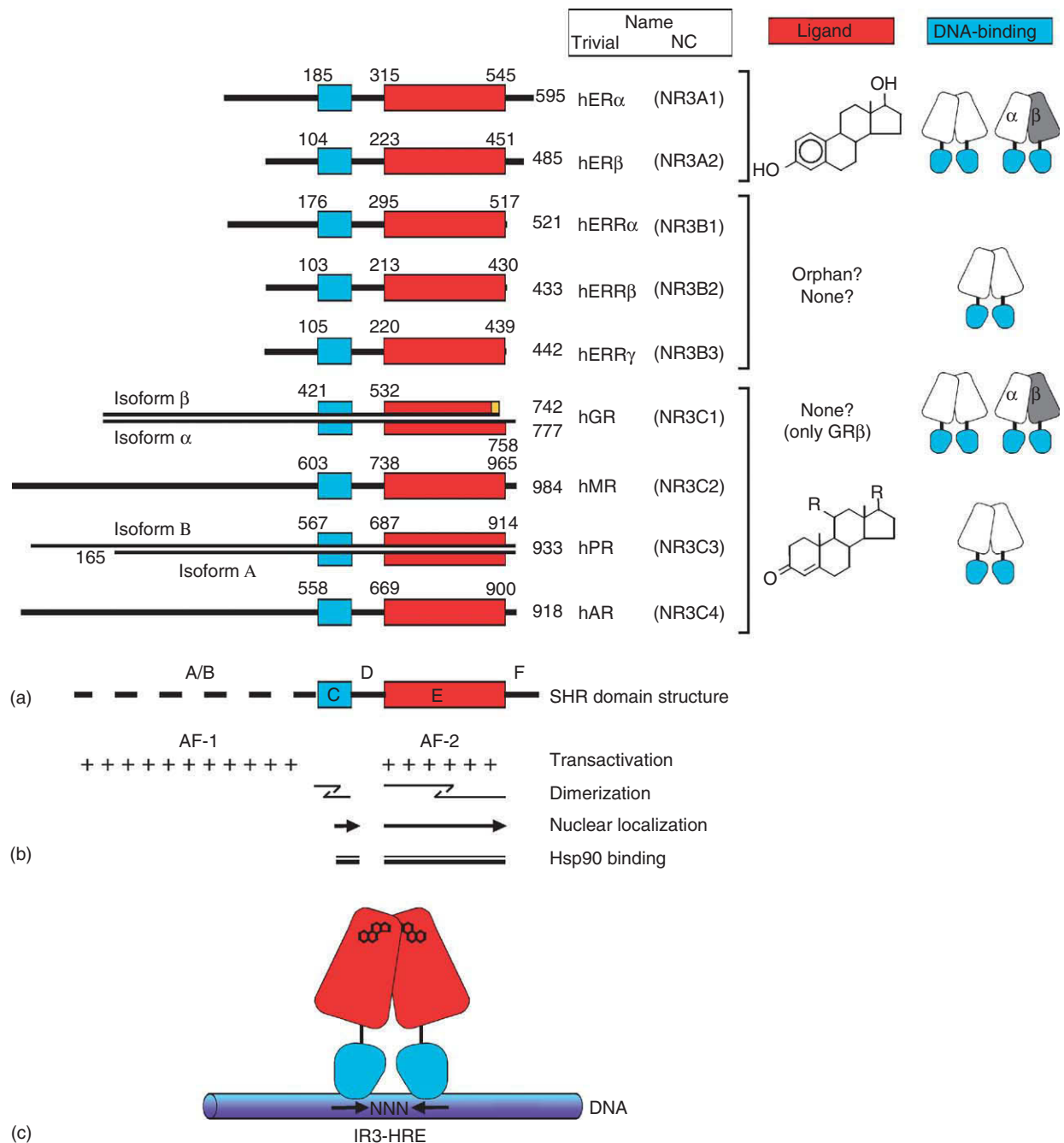


Figure 2 The human steroid hormone receptor family. a, Structure of all family members, cognate ligand, and DNA-binding mode; b, SHR domain structure and structure–function relationships; c, an SHR homodimer bound to its cognate hormone response element (HRE). In a, the numbers refer to amino acid positions. Highlighted are the DNA-binding domain (DBD; shown in blue) and the ligand-binding domain (LBD; shown in red). The A/B and F domains are drawn to scale, and the name of each receptor is indicated (trivial and name according to nomenclature scheme). In c, the HRE is of an inverted repeat (IR) type (palindrome), and half-sites (arrows) are separated by three unspecified (N) nucleotides (IR3-HRE); the hormone ligand is indicated by a sterane formula. AF-1 and -2, transcription activation functions 1 and 2; AR, androgen receptor; ER, estrogen receptor; ERR, estrogen-related receptor; GR, glucocorticoid receptor; h, human; MR, mineralocorticoid receptor; PR, progesterone receptor.

For the GR, two different classes of cDNAs have been described in humans, termed hGR α and hGR β , which are the result of alternative splicing from a single gene transcript. Both isoforms are identical up to amino acid 727 and then diverge, with hGR α being

slightly larger (777 amino acids) than hGR β (742 amino acids). In contrast to hGR α , which is commonly referred to as the bona fide hGR, hGR β was long dismissed as a cloning artifact but has now been shown to be expressed at modest but varying levels

in a range of tissues. Each variant mRNA is translated from at least eight initiation sites into multiple GR α and possibly GR β isoforms, amounting to a multitude of GR monomers and probably a plethora of homo- or heterodimers of unknown function. hGR β does not bind hormones, is transcriptionally inactive, and therefore acts as a ligand-independent negative regulator of glucocorticoid action in transfection experiments. However, the β isoform is not conserved across species, and functional evidence that implicates hGR β in glucocorticoid insensitivity is still missing. hGR α shows a high affinity for the artificial glucocorticoid dexamethasone, moderate affinity for the physiological steroids cortisol and corticosterone, and low affinity for mineralocorticoids and progesterone. In contrast to all other SHRs, GR α dimerizes only weakly.

The progesterone receptor was the first SHR shown to exist in two common isoforms generated by differential promoter use. One promoter initiates transcription at positions +1 and +15 of the gene, which gives rise to the longer isoform B (Figure 2a). The second promoter initiates hPR transcripts between +737 and +842, encoding the 164-amino-acid residues shorter hPR form A. Only PR-B harbors a third activation function (AF-3) at its specific N-terminus, which functions promoter- and cell-specifically. Both isoforms, therefore, display differential target gene specificity. First regarded as an exception in the family, differential promoter use and alternative splicing are now considered the rule for all members. Both PR isoforms show a high affinity for the natural ligand progesterone and the synthetic agonist R5020. Alternative splicing can also generate truncated forms of PR that are found in breast cancer.

The human mineralocorticoid receptor (hMR) and androgen receptor (hAR) were cloned 2–3 years after the GR was described. The hAR exists in two isoforms, hAR-A and hAR-B, which are structurally analogous to the two hPR isoforms. In contrast to hPR-A/B, hAR-A is expressed at substantially lower levels than the B form and its contribution to androgen action is not known. The hAR binds the two naturally occurring ligands dihydrotestosterone and testosterone with high affinity, whereas the hMR shows high and equivalent affinity for aldosterone, physiological corticosteroids, and progesterone.

Heterooligomeric Complex and Ligand and DNA Binding

Heterooligomeric Complex

In 1965, the groups of Jensen and Gorski showed that specific and high-affinity binding sites for estrogen

existed in large heterooligomeric complexes migrating with 9S on low-salt sucrose gradients. This non-DNA-binding complex can be transformed into a more compact and protease-resistant 4S DNA-binding form by high salt and temperature or hormone. The 9S heterooligomeric receptor complexes were later shown to be composed of the two chaperones – heat shock protein (Hsp)90, recently identified as a member of the GHKL family of ATPases, and Hsp70 – plus a number of co-chaperones.

Hsps form an evolutionarily very conserved group of stress-inducible proteins. Contrary to their name, Hsps are strongly induced not only by heat shock, but are also expressed constitutively at physiological temperatures at significant levels. Hsps belong to a larger family of chaperones that are known to stabilize the native structure of other polypeptides by controlled binding to them.

Most investigations have been performed with the heterooligomeric GR complex that contains a dimer of Hsp90 and a substoichiometric amount of Hsp70 and one or two molecules of p23. The last two are required for the assembly of the final heterooligomeric complex. The crystal structure of full-length yeast Hsp90 in complex with an ATP analog and the yeast p23 homolog has been solved recently. For PR and AR, the LBD is sufficient for an interaction with Hsp90 that also requires ATP. In case of the GR, a seven-amino-acid segment overlapping the N-terminus of the LBD, which is located on the surface of the LBD at the opening of the ligand-binding pocket, is sufficient and required for Hsp90 binding. In case of the ER, in addition to the LBD, part of the hinge region D stabilizes the Hsp90–ER complex. From the known crystal structures of a number of LBDs of nuclear receptors and several biochemical studies, it becomes clear that Hsp90 recognizes two hydrophobic surfaces at opposite sides of the LBD located in the vicinity of the ligand-binding pocket. The large distance between the two hydrophobic surfaces probably accounts for the fact that two Hsp90 molecules are bound to SHRs. The LBD of the receptor must be in a complex with Hsp90 in order to have high ligand-binding activity. Hsp90 binding seems to be a characteristic of SHRs because other nuclear receptors, such as the retinoid, thyroid hormone, and vitamin D₃ receptors, do not associate efficiently with these Hsps.

The complexes also contain one of a number of tetratricopeptide (TPR) domain immunophilins that bind to a TPR acceptor site at the C-terminus of Hsp90. Immunophilins are characterized by their peptidylprolyl isomerase (PPIase or rotamase) domains and can function as chaperones like Hsps. SHR complexes contain one of three immunophilins (FKBP51,

FKBP52 or Cyp-40) or PP5, a protein phosphatase that contains both a TPR domain and a PPIase homology domain. The PPIase domain of immunophilins is a binding site for immunosuppressant drugs such as FK506 or cyclosporine A that inhibit PPIase activity on binding.

In hormone-free cells, steroid receptors shuttle between the cytoplasm and nucleus. The distribution may be predominantly cytoplasmic (GR) or nuclear (ER α). Steroid-dependent transformation of the GR triggers its rapid translocation to the nucleus, possibly involving movement along microtubules driven by the motor protein dynein. In wheat germ extracts, it was shown *in vitro* that wheat immunophilins can link the GR–Hsp90 heterocomplex to cytoplasmic dynein via conserved TPR and PPIase domain interactions. Immunoadsorption of GR from cell lysates of L cells coimmunoadsorbs Hsp90, the dynein intermediate chain, FKBP52, PP5, and dynactin (a subunit of the dynein-associated dynactin complex). If the microtubules are stabilized by taxol, α -tubulin is coimmunoadsorbed as well.

All these data are derived from *in vitro* experiments with cell extracts in which the cellular microenvironments are destroyed, so that the observed complexes may be artificial. But a number of *in vivo* observations indicate that the heterooligomeric receptor complex is also of functional importance. Increased levels of FKBP51 have been reported to be the common cause of glucocorticoid resistance in three New World primates, due to delayed translocation of the GR to the nucleus. Likewise the overexpression of FKBP51 in COS cells reduces the hormone binding activity of GR and hence its transcriptional activity. In yeast cells, the overexpression of FKBP51 has no effect, whereas FKBP52 enhances GR-dependent transcription that is dependent on interaction with Hsp90 and PPIase activity. It could be shown that the binding of FKBP52 to Hsp90 potentiates glucocorticoid-dependent reporter gene activity in yeast and mammalian cells that is dependent on PPIase activity and increases the hormone binding affinity of the receptor. Moreover, differential binding of FKBP51 and -52 to dynein corresponds to a differential effect on nuclear translocation. The co-expression of dynactin and the PPIase domain, as well as PPIase domain-swapping experiments, demonstrated the relevance of the dynactin complex for receptor function. Hsp90 and the co-chaperone p23 have been shown to colocalize *in vivo* at genomic HREs in a hormone-dependent manner. p23 seems to be involved in disrupting receptor-mediated transcriptional activation *in vivo* and *in vitro* and, thus, it could be involved in terminating the hormone response.

Therefore receptor-associated chaperones and co-chaperones can no longer be viewed as simple repressors that have to be removed by hormone in order to transform the non-DNA-binding heterooligomeric complex into a DNA-binding entity; they are, instead, viewed as chaperone complexes that accompany all steps in the life of these client proteins: folding, assembly of quaternary structures, intracellular transport, and targeting for proteolytic degradation.

Ligand Binding

Although not proven formally, it is believed that the lipophilic steroid hormones, as well as synthetic compounds with agonistic or antagonistic effects, enter the target cell by simple diffusion and bind to SHRs within the heterooligomeric complex. For hER α and hPR, the three-dimensional structure of their ligand-binding domains complexed with 17 β -estradiol and the antihormone raloxifene or progesterone, respectively, have been determined. Like other known LBDs of nuclear receptors, the LBDs of SHRs are folded into a three-layer antiparallel α -helical sandwich that creates a wedge-shaped molecular scaffold with the ligand-binding cavity at the narrower end of the domain (Figure 3). This cavity is completely partitioned from the external environment and is closed by helix 12 of the LBD, operating as a lid after the hormone has entered the binding pocket. The relocation of this amphipathic helix 12 over the hormone-binding site generates a new surface that allows coactivators to bind to the LBD, thereby mediating the activity of AF-2, which forms the core of helix 12. The LBD dimerization interfaces are formed by α -helices that line up and/or intertwine and differ slightly between ER α and PR.

DNA Binding

The receptors for progestagens, glucocorticoids, mineralocorticoids, and androgens bind to the same HRE that is composed of two half-sites (AGAACA), which are arranged as inverted (palindromic) repeats (IR) with the half-sites being separated by three nucleotides (IR-3). The HRE half-site recognized by both ERs (AGGTCA) deviates in two nucleotides from this DNA sequence (Figure 3c). One HRE half-site is recognized by one receptor monomer, and the homodimer binds in a head-to-head orientation to an IR-3 element. The non-ERs recognize their half-site mainly by a hydrophobic interaction with the methyl group of the thymine in position 4 of the complementary strand (AGAACA) that is not present in the DNA sequence recognized by the ER. The half-site spacing of 3 bp is such that an SHR homodimer binds both half-sites on the same face of the DNA. In addition to IR-3 elements, only the AR is able to selectively

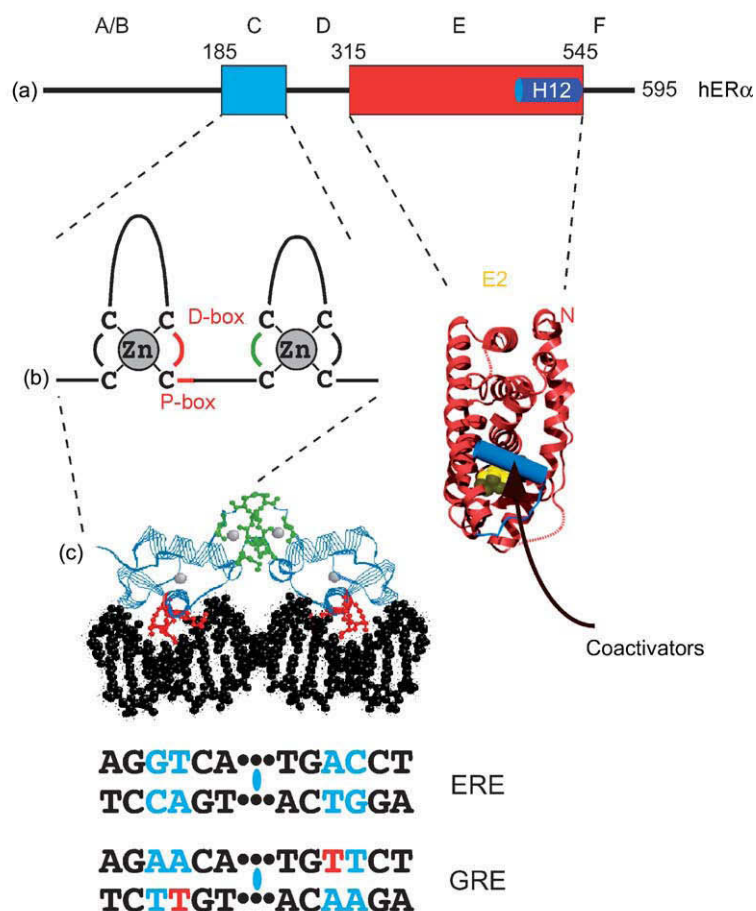


Figure 3 Structure of the human estrogen receptor α . a, Functional domains (A/B to F) of hER α at the primary structure level, with the DNA-binding domain (DBD) and the ligand-binding domain (LBD) highlighted; b, an exploded view of the secondary structure of the hER α DBD bound to DNA, followed by the sequences of the consensus estrogen response element (ERE) and a consensus glucocorticoid response element (GRE); c, crystal structure of the hER α DBD complexed with 17 β -estradiol (E2). In a, The numbers refer to amino acid positions. DBD is highlighted in blue and the LBD in red. The A/B and F domains are drawn to scale. In b, the secondary structure of the DBD is characterized by two steroid hormone receptor-specific zinc fingers. Four cysteines each tetrahedrally coordinate one zinc ion (grey). The Proximal box (P-box) responsible for specific DNA recognition is shown in red; the Distal box (D-box) mediating DBD dimerization is shown in green. In the crystal structure in c, the four amino acid side chains of the P-box (red) are shown interacting with DNA bases (black). These residues are part of an α -helix responsible for specific DNA-sequence recognition that is positioned within the DNA's major groove (perpendicular to the paper plane). A second amphipathic α -helix can be seen to cross the recognition helix with the red P-box residues at a right angle. Between the two helices, the green D-box residues promoting DBD dimerization are located. Underneath the DNA, the sequences of the consensus ERE and a consensus GRE are shown. Both half-sites are organized as an inverted repeat with 3-bp spacing (IR-3). The symmetry axis is indicated by a blue ellipse. The bases that differ between ERE and GRE are highlighted in blue, and the thymine important for the discrimination between both response elements is highlighted in red. In the crystal structure in d, E2, a natural ligand, is shown in yellow. Decisive for a productive interaction with coactivators is the proper positioning of helix 12 (H 12) over the ligand-binding pocket. (We thank R. Hubbard and A. Pike, University of York, UK, for providing this part of the figure.)

recognize direct repeats of the two half-sites separated by three nucleotides (ADR-3 elements). In contrast to other nuclear receptors binding to DR elements, the AR homodimer also binds in a head-to-head orientation to an ADR-3 element.

The DBD of SHRs comprises approximately 80 amino acids encoded by region C plus some 14 N-terminal amino acids of region D. Domain C contains two patterns that are reminiscent of, but clearly distinguishable from, the zinc-finger motifs

first observed in the *Xenopus laevis* transcription factor IIIA. The SHR zinc fingers are also able to tetrahedrally coordinate a zinc atom, but are of the type (Cys2-Cys2). Only very few amino acids, termed the proximal (P-)box, within the first SHR zinc finger are responsible for specific recognition of the cognate HRE. Another set of amino acids, called the distal (D-)box within the second SHR zinc finger, forms the weak dimerization interface of the DNA-binding domain (Figure 3b).

Activation of Transcription

To regulate transcription, agonist-liganded SHRs must talk to the general transcription machinery. This can be achieved either by a direct contact between SHRs and the general transcription factors (GTFs), which are building up the transcription preinitiation complex at the transcription start point, or by means of other cellular factors that do not bind to DNA directly but are recruited to the promoters by the SHRs via protein–protein interactions. Depending on the effect of these coregulators on gene expression, they can be divided into coactivators, which promote transcription mediated by SHRs, and corepressors, which suppress SHR-dependent gene transcription. Coregulators, other sequence-specific transcription factors, and chromatin factors all function as interpreters between SHRs and the general transcription machinery.

Coactivators

Coactivators were originally called transcriptional intermediary factors because they are supposed to bridge between DNA-bound sequence-specific transcription factors and GTFs. The suspicion that SHRs may require coactivators for activation was derived from the squelching phenomenon – the observation that an excess receptor can inhibit its own transactivation as well as transcription by other transactivators. One possible explanation for this behavior is that additional factors, which are present in limited amounts and required for transactivation, are trapped by excess SHRs in unproductive complexes. Conversely, by overexpressing coactivators, it is possible to reactivate a target promoter that has been squelched by excess SHR.

Today, more than 20 of these coactivators are known that were quickly identified within a decade. An important and well-characterized class of coactivators is the steroid receptor coactivator (SRC, or p160) family of coactivators, consisting of its three members SRC-1, SRC-2 (also known as transcription intermediary factor 2 or glucocorticoid receptor-interacting protein 1), and SRC-3. The three related proteins have partial overlapping specificity for hormone receptors and cell type. SRCs enhance hormone-dependent transcriptional activities of SHRs without altering the basal activity of a target promoter. The biological functions of SRCs have been revealed by gene-disruption experiments in mice. Whereas SRC-1-deficient mice display normal growth and fertility but partial resistance to a number of steroid hormones, SRC-2 knockout mice have reduced fertility due to gonadal failure and are leaner because the ratio of white and brown adipose tissue is decreased. The inactivation of SRC-3 has more

dramatic effects and results in growth retardation and smaller body size as well as resistance to growth hormones (GHs) and estrogen. Like SHRs, coactivators are modular proteins with common functional domains such as one that possesses acetyltransferase activity. For the binding to agonist-liganded SHRs, one or several copies of an α -helical LXXLL signature motif, also called a nuclear receptor (NR) box, are required. After the binding of an agonistic ligand to the SHR, helix 12, containing the AF-2 core of the receptor, is relocated, exposing a hydrophobic cleft and an adjacent ionic clamp as the principal determinants for the interaction with an NR box.

In vivo coactivators are organized into complexes that are poised for interaction with SHRs in response to appropriate stimuli. Many of the characterized complexes share subunits, suggesting that a directed or cyclical assembly of coactivator complexes takes place. Current data support a model in which the initial binding of ligand results in the dissociation of corepressors and recruitment of chromatin remodeling machines such as SWI/SNF. The binding of SRCs and other coactivators such as CBP/p300 then results in local acetyltransferase activity and the opening of chromatin in the vicinity of the target promoter. C-terminal SRC-domains can recruit other coactivators, such as the coactivator-associated arginine methyltransferase-1 (CARM1) and cAMP responsive element binding protein (CREB)-binding protein (CBP), and the related protein p300 for complementing and potentiating histone acetyltransferase activity, respectively.

Coactivators are subject to a number of posttranslational modifications for modulating coactivator function. Acetylation and ubiquitination appear to alter the half-life of molecules, whereas phosphorylation is essential for the optimal activity of coactivators within the cell. Phosphorylation of SRCs occurs in response to a plethora of stimuli, including epidermal growth factor, steroid hormones, cytokines, and increased intracellular cAMP. In turn, multiple kinases, such as MAPKs and I κ B kinases, that are involved in a number of cellular signaling pathways are responsible for a distinct phosphorylation pattern on SRC-3, regulating its activity and, more important, its specificity. In contrast to SHRs, SRCs have a broad tissue distribution and coactivate an ever-expanding number of transcription factors. Therefore, coactivators, like SRCs, might serve as integrators of multiple afferent stimuli in order to produce an appropriate cellular response.

Cyclin D1, the regulatory subunit of a holoenzyme that phosphorylates and inactivates the retinoblastoma protein tumor suppressor and promotes progression through the G1-S phase of the cell cycle, has

been identified as a coactivator for ER α recently. The *cyclin D1* gene is frequently overexpressed in human breast cancer and is capable of inducing mammary tumorigenesis when overexpressed in transgenic mice. Cyclin D1 binds to the N-terminal half of the ligand binding domain of ER α and enhances ligand-induced recruitment of ER α to an estrogen responsive element (ERE). More important, cyclin D1 antagonizes the repression of ER α activity mediated by the tumor suppressor gene product BRCA1, which is functioning as a corepressor of ER α in this context. Thus cyclin D1 and BRCA1 constitute an antagonizing pair of coregulators for ER α .

Corepressors

For some members of the nuclear receptor family, such as the thyroid hormone receptor (TR) and the retinoid acid receptors (RARs), that form heterodimers with the retinoid X receptor (RXR), it is well known that they repress basal transcription in the unliganded state. It has been shown that this silencing effect is mediated by corepressors that were called nuclear receptor corepressor (N-CoR) and silencing mediator for retinoid and thyroid hormone receptors (SMRT). N-CoR and SMRT associate with DNA-bound TR–RXR or RAR–RXR heterodimers in the absence of hormone but not in its presence. SMRT and N-CoR are multiprotein complexes that exhibit histone deacetylase activity. Both corepressors are also involved in steroid hormone action because it was demonstrated that the overexpression of N-CoR and SMRT represses the partial agonistic activity of both ER α bound to the antihormone tamoxifen and of PR bound to the antagonist RU486. Recognition by transcriptionally incompetent SHRs of corepressors is mediated by short amphipathic helical signature motifs, called CoRNR boxes, which are similar to the NR boxes of coactivators. Corepressor activity is mediated by histone deacetylation. Because corepressors lack intrinsic deacetylation domains, they recruit histone deacetylases for this task. Apparently, corepressors and coactivators are functionally similar proteins, so it is not surprising that a corepressor can also turn into a coactivator, as shown for N-CoR and RAR.

Recently, the tumor suppressor gene product BRCA1 has been shown to be a corepressor of ER α . BRCA1 is implicated in DNA repair mechanisms, and germ-line mutations in the BRCA1 gene predispose women to ovarian and breast cancer. BRCA1 inhibited estradiol-induced endogenous gene expression at levels of BRCA1 that are within the physiological range. Moreover, BRCA1 was shown to bind to ER α directly *in vitro* and the binding of BRCA1 to an

estrogen responsive gene was reduced by estradiol *in vivo*. BRCA1 appears not to act on specific estrogen responsive genes or gene subsets but blocks the ability of estradiol to alter the expression of the vast majority of all estradiol responsive genes in MCF7 cells. BRCA1 downregulates the coactivator p300, and excess p300 rescues BRCA1 repression of ER α like cyclin D1 can. The physical association of cyclin D1 with ER α may exclude the direct association of BRCA1 with the receptor or alter the conformation of BRCA1 and/or the binding of other BRCA1-associated proteins and thereby contributes to activation at an ERE.

Altogether, SHRs can exist in a multitude of conformations, depending on the nature of bound ligand. The various interaction surfaces are mainly determined by helix 12 and most likely by the C-terminus beyond helix 12, the F region. The transactivation by SHRs appears to be not only a simple net activation but the sum of relief from repression by corepressors and activation by coactivators. The switch from the repressed to the activated state is initiated by the hormone ligand through an allosteric change in SHR structure followed by an intricate interplay of coregulators.

Interactions with Sequence-Specific Transcription Factors

SHRs control the activity of natural promoters through positive and negative interactions with sequence-specific transcription factors. Particularly interesting is an interaction that represses the activity of both partners, as exemplified in the case of hGR and the heterodimeric transcription factor AP-1. Similar interactions have been described between GR and the p65 subunit of the transcription factor nuclear factor (NF)- κ B and the transcription factor GATA-1. From experiments with numerous SHR mutants, it can be concluded that repression is most likely mediated by a protein–protein interaction with an SHR monomer that requires the N-terminal portion of the receptor. Because most immunoregulatory genes as well as genes involved in inflammation are positively regulated by the transcription factors AP-1 and NF- κ B, it is conceivable that the immunosuppressive and anti-inflammatory activities of glucocorticoids are mediated through the inhibition of AP-1- and NF- κ B-mediated transactivation by GR (Figure 1c).

A novel biological action of GR was revealed by constructing mice with the hepatocyte-specific loss of GR that have a severe reduction in body weight due to decreased bone length and size of internal organs, very reminiscent of mice lacking the signal transducer and activator of transcription (STAT)5. In both mice strains, growth hormone (GH) signaling via

the JAK–STAT pathway is impaired. GR directly interacts with both phosphorylated and unphosphorylated forms of STAT5, and both proteins are detectable at promoters of GH-regulated genes *in vivo*, which is essential for GH signaling. As shown by dimerization defective GR mutant (GR^{dim}) mice, GH signaling is independent of GR DNA binding and most likely mediated by protein–protein interaction with STAT5 transcription factors.

Interaction with Chromatin Factors

The interaction of SHRs with DNA, the general transcription machinery, coregulators, and sequence-specific transcription factors takes place in the nucleus with its DNA compacted into chromatin. Genetic analyses have demonstrated a widespread involvement of chromatin structure in gene regulation in general. For the GR, it was shown that components of the so-called SWI–SNF complex, a set of pleiotropic transactivators that counteract the repressive functions of chromatin, are required for transactivation in yeast. In human cells lacking brahma (hBrm), the homolog of the SWI–SNF complex subunit SWI2, transactivation by GR is weak and can be enhanced selectively by the ectopic expression of hBrm. Like SWI2 in yeast, hBrm is part of a large multiprotein complex that mediates the ATP-dependent disruption of a nucleosome. A direct interaction between nuclear receptors and components of hBrg1, a brahma-related ATPase complex, has been demonstrated, and chromatin immunoprecipitation experiments indicate that the hormone receptors recruit these ATP-dependent chromatin remodeling complexes to their target genes. In the case of the MMTV promoter, activation by glucocorticoids and progestins depends on the correct organization of the promoter sequences in nucleosomes so that some HREs of a larger cluster get exposed. The binding of the receptor to the exposed HREs leads to the recruitment of ATP-dependent chromatin remodeling activities that displace H2A–H2B dimers from the promoter and allow the binding of the ubiquitous transcription factor NF1 and the recruitment of more receptor molecules, followed by coactivators and the transcription preinitiation complex. Optimal activation requires the inclusion of linker histones that help to position nucleosomes and generate a more suitable platform for receptor binding.

Apart from the SWI/SNF complex and related ATP-dependent chromatin remodeling machines, increasing evidence has implicated the modification of histones in the regulation of gene transcription. Histones are composed of a histone fold domain, involved in wrapping the DNA in nucleosomes, and an N-terminal tail rich in lysine that protrudes out of

the nucleosome. The acetylation of tail lysines reduces the affinity of histone tails for DNA greatly and is believed to render DNA more accessible to transcription factors while still maintaining a nucleosomal architecture. Moreover, some of the identified coactivators for SHRs, such as CBP/p300 and SRC1, as well as the largest subunit of the pivotal GTF TFIID, TAF_{II}230/250, turned out to be histone acetyl transferases (HATs). All of these observations led to the current two-step model for transcriptional activation by SHRs that is built on the hormone-mediated recruitment of coactivators and other transcription factors with HAT activity, resulting in the local destabilization of repressive histone–DNA interactions, followed by direct or indirect interactions with the basal transcription machinery. Histone arginine methyl transferases, such as CARM1, are also implicated in gene activation by steroid hormones. and recent evidence suggests a participation of kinases that phosphorylate core and linker histones. Altogether, these new findings emphasize the importance of chromatin remodeling in gene regulation by steroid hormones.

Recently, a detailed kinetic analysis by chromatin immunoprecipitation (ChIP) analyses of the assembly of transcription initiation complexes on the estrogen responsive promoter of the pS2 gene has revealed a sequence of cycles of formation of multiprotein complexes involving protein-modifying enzymes and ATP-dependent chromatin remodeling complexes, followed by ubiquitin/proteasome-mediated degradation of the receptor with a cycling amplitude on a time scale of minutes. The dynamic nature of transcriptional activation has also been assessed by single live-cell imaging of transcription factors tagged with fluorescent proteins. Contrary to the ChIP results, fluorescence recovery after photobleaching analyses have shown that SHRs and interacting coregulators rapidly exchange on large artificial HRE arrays on a time scale of seconds. In attempting to reconcile these contradictory results, we have to consider that ChIP is a composite freeze-frame of promoter occupancy in millions of cells, whereas live-imaging focuses on the chronology of events in a single cell.

Involvement of the Glucocorticoid Receptor in the Stress Response

Studies of human patients and adrenalectomized animals have shown that glucocorticoids have a plethora of functions, the most important being the stimulation of gluconeogenesis; decrease of glucose uptake by peripheral tissues; increase of glycogen deposition; increase of lipolysis; stimulation of protein and nucleic acid synthesis, predominantly in the liver; suppression of immune and inflammatory responses;

inhibition of fibroblast proliferation and extracellular matrix deposition; maintenance of normal cardiovascular function and blood pressure; and regulation of calcium absorption. Many of these functions should prevent an overreaction of defense mechanisms to stress and be conducive to coping with the stress response. In line with this hypothesis, the synthesis and release of glucocorticoids are elevated during stress, and people with chronic adrenal insufficiency show “general languor and debility, remarkable feebleness of the heart’s action, irritability of the stomach and a peculiar change in the color of the skin,” as Addison described it in 1855. Although none of these symptoms is life-threatening, patients with Addison disease cannot withstand stress and are prone to suffering an acute adrenal crisis that can lead to shock or coma.

Until recently, it was believed that all these glucocorticoid effects were of a genomic nature, mediated by its receptor binding to glucocorticoid response elements (GREs) in the target genes. Given the multiple important functions of glucocorticoids, it was therefore not surprising to learn that GR-deficient mice, which were generated using gene knockout techniques, are not viable and die shortly after birth.

However, the cause of death – respiratory failure due to lack of lung maturation – does not seem to depend on genomic effects mediated by GREs. Contrary to GR-deficient mice, homozygous mice that express a mutated GR that is not able to dimerize appear almost normal and healthy under standard laboratory conditions. This dimerization-defective GR (GR^{dim}) can no longer bind efficiently to palindromic GREs, and these mutant mice do not respond to the administration of glucocorticoids by the induction of well-characterized hormone responsive genes. Clearly, GRE-mediated gene activation is not required for development and survival. On the other hand, GR^{dim} mice show a selective impairment of spatial memory in the water maze. It is known that the stress that is associated with a learning task – here orientation in a maze – leads to the activation of GRs in the brain, which is required for facilitated storage of the acquired information. Therefore, the facilitating effects of corticosterone on spatial memory seem to depend on DNA binding of the GR.

There is also increasing evidence that altered GR signaling is an important cause of the pathogenesis of major depression. The most consistent biological finding in patients with depression is hyperactivity of the hypothalamic-pituitary-adrenal (HPA) axis. Mutant mice with a 50% GR gene dose reduction have a disinhibited HPA axis like human patients and show increased helplessness after stress exposure, a behavioral correlate of depression in mice. In contrast, mice that overexpress GR demonstrate a reduced

helplessness response after stress challenge. Also, human carriers of two polymorphisms in the GR gene had an increased risk of developing a major depressive mode.

Cross-Talk with Other Signal Transduction Pathways

Steroid hormones can either act via their cognate receptors that alter gene expression after binding to HREs in target genes (termed genomic or nuclear action) or act via membrane-bound or -localized SHRs or non-SHRs (termed nongenomic or cytoplasmic action). Whereas genomic action is slow and takes 30–60 min, nongenomic action is rapid and takes a couple of seconds or minutes. We now briefly summarize the important findings on nongenomic actions that involve the classical steroid receptors.

The existence of such nongenomic actions became apparent in experiments with GR knockout mice that die postnatally and with mice that express a dimerization-defective mutant GR and are viable. The dimerization-defective GR shows impaired GRE-dependent transactivation while functions that require cross-talk with other transcription factors remain intact. Moreover, many studies employing different techniques and assays have shown that a small pool of estrogen (some 2%) and ARs are associated with the plasma membrane. More specifically, both receptors have been identified in the caveolae, together with a number of signaling molecules. One of the actions of steroid receptors thought to take place in the caveolae is the estrogen-mediated rapid increase in endothelial nitric oxide synthase (eNOS) activity. The 110-kDa caveolin-binding protein striatin has been identified as a molecular interaction hub that organizes an ER α -eNOS membrane signaling complex engaging also the G-protein G_{zi}. The disruption of the ER α -striatin interaction abolishes the rapid activation of MAPK, Akt kinase (also called protein kinase B), and eNOS by estrogen, whereas genomic signaling is not affected. But other proteins such as caveolin-1, modulator of nongenomic action of estrogen receptor (MNAR), Shc, or insulin-like growth factor-1 receptor are also implicated in tethering ER α to the inner face of the plasma membrane bilayer. Moreover, the steroid-independent palmitoylation of a cysteine residue within the E domain has been observed. Palmitoylation facilitates caveolin-1 binding as well as membrane localization. Interestingly, A/B and C domain-deleted ER α still localizes to the membrane and signals to extracellular receptor kinase (ERK) like the wild-type receptor. Current evidence suggests that ER α membrane localization is required for kinase activation.

A well-known rapid response of steroid receptors is the activation of the c-Src–p21ras–ERK2 pathway by ligand-bound AR, ER α (but not ER β), GR, MR, and PR-B. For all five SHRs MAPK activation is abrogated in the presence of antihormones, but this is less clear for MR. Although various mechanisms have been proposed, the interaction of SHRs with the kinase c-Src is an important aspect of cross-talk. Whereas AR interacts with the SH3 domain of c-Src, ER α interacts with the SH2 domain. AR activation of the ERK pathway is more effective if, in addition to AR binding to the SH2 domain, ER α simultaneously binds to the SH3 domain. In cells expressing both receptors, stimulation with androgen or estrogen induces a productive ER α –AR–c-Src complex that can be prevented by treating with either antiandrogens or antiestrogens. PR-B does not bind to c-Src directly, but requires the presence of ER α . In the absence of progestins, PR-B (but not PR-A) binds to ER α . Progestins release ER α from this complex and thus trigger the interaction of ER α with c-Src. In case of ER α , cross-talk is exemplified by epidermal growth factor (EGF) treatment of ovariectomized mice. This results in nuclear translocation and an altered phosphorylation state of ER α . Transient transfection studies showed that EGF, thus, can activate ER α .

The activation of the GR elicits apoptosis in T cells, but can paradoxically also promote thymocyte survival. This dual effect is reminiscent of the actions of the T-cell antigen receptor, which is essential for thymocyte survival but can also cause apoptosis. Interestingly, the activation of the T-cell antigen receptor blocks glucocorticoid-induced apoptosis, thus implying functional cross-talk between these two distinct signaling systems. The regulation of programmed cell death by glucocorticoids also involves the regulated expression and alternative splicing of the *bcl-X* gene and crosstalk with the STAT5 family of transcription factors. In lymphocytes, the apoptotic effect is mediated by a selective interaction of GR with activated STAT5a, whereas in mammary cells, equipped with STAT5b, glucocorticoids act as antiapoptotic factors.

The downstream targets of the rapid action of steroids remain to be fully uncovered, but in many cases the SHRs themselves – and their coactivators – are targeted. Therefore, one function of these rapid actions seems to be the potentiation of genomic action. All SHRs have been known to be phosphoproteins for many years. Most of the identified phosphorylation sites are serine and threonine residues, but some of the family members are also phosphorylated on tyrosine. In case of the chicken PR, for instance, four phosphorylated serines have been identified common to both isoforms, PR-A and PR-B. The

two N-terminal sites are only moderately phosphorylated in the absence of hormone, whereas, after hormone treatment, an increase in the phosphorylation of these sites and the appearance of two new phosphorylation sites are observed. The mutation of these serine residues results in a cell- and promoter-specific variation of receptor activity when tested in transfection experiments. In case of ER α ERK, phosphatidylinositol 3-kinase, p90^{RSK}, a kinase downstream of Akt2, and Akt2 itself phosphorylate discrete residues of the receptor, upregulating its transcriptional activity or stability.

See Also the Following Articles

Adrenal Insufficiency; Androgen Action; Corticosteroids and Stress; Estrogen; Glucocorticoids, Overview; Membrane Glucocorticoid Receptors.

Further Reading

- Addison, T. (1855). *On the constitutional and local effects of disease of the suprarenal capsules*. London: D. Highley.
- Baulieu, E.-E. and Kelly, P. A. (1990). *Hormones: from molecules to disease*. New York: Chapman and Hall.
- Beato, M., Herrlich, P. and Schütz, G. (1995). Steroid hormone receptors: many actors in search of a plot. *Cell* **83**, 851–857.
- Cato, A. C. B., Nestl, A. and Mink, S. (2002). Rapid actions of steroid receptors in cellular signaling pathways. *Science's Signal Transduction Knowledge Environment* **138**[Online].
- Gronemeyer, H. and Laudet, V. (1995). Transcription factors 3: nuclear receptors. *Protein Profile* **2**, 1173–1308.
- Nuclear Receptors Nomenclature Committee. (1999). Nuclear receptors nomenclature committee: a unified nomenclature system for the nuclear receptor superfamily. *Cell* **97**, 161.
- Levin, E. R. (2005). Integration of the extranuclear and nuclear actions of estrogen. *Molecular Endocrinology* **19**, 1951–1959.
- Lu, N. Z. and Cidlowski, J. A. (2006). Glucocorticoid receptor isoforms generate transcription specificity. *Trends in Cell Biology* **16**, 301–307.
- Mangelsdorf, D. J., Thummel, C., Beato, M., et al. (1995). The nuclear receptor superfamily: the second decade. *Cell* **83**, 835–839.
- McKenna, N. J. and O'Malley, B. W. (2002). Combinatorial control of gene expression by nuclear receptors and coregulators. *Cell* **108**, 465–474.
- Metivier, R., Reid, G. and Gannon, F. (2006). Transcription in four dimensions: nuclear receptor-directed initiation of gene expression. *EMBO Reports* **7**, 161–167.
- Parker, M. G. (1991). *Nuclear hormone receptors: molecular mechanisms, cellular functions, clinical abnormalities*. London: Academic Press.
- Pratt, W. B., Morishima, Y., Murphy, M., et al. (2005). Chaperoning of glucocorticoid receptors. In: Gaestel, M.

- (ed.) *Handbook of experimental pharmacology*. Vol. 172: *Molecular chaperones in health and disease*, pp. 111–138. Berlin: Springer.
- Wu, R. C., Smith, C. L. and O'Malley, B. W. (2005). Transcriptional regulation by steroid receptor coactivator phosphorylation. *Endocrine Reviews* 26, 393–399.
- Wintermantel, T. M., Berger, S., Greiner, E. F., et al. (2005). Evaluation of steroid receptor function by gene targeting

in mice. *Journal of Steroid Biochemistry and Molecular Biology* 93, 107–112.

Relevant Website

NUREBASE, <http://www.ens-lyon.fr/LBMC/laudet/nurebase>

Steroid Hydroxylases

F H de Jong

Erasmus University Medical Centre, Rotterdam,
The Netherlands

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
F H de Jong, volume 3, pp 505–507, © 2000, Elsevier Inc.

Introduction

Cholesterol Side-Chain Cleavage Enzyme: CYP11A1

17-Hydroxylase: CYP17A1

21-Hydroxylase: CYP21A2

11 β -Hydroxylases: CYP11B1 and CYP11B2

Aromatase: CYP19A1

Steroid Catabolism by CYPs

Glossary

<i>Cytochromes P450</i>	Enzymes involved in the hydroxylation of endogenous or exogenous substrates.
<i>Hydroxylation of steroids</i>	Replacement of hydrogen (H) atoms in steroids by hydroxyl (OH) groups.
<i>Localization of steroidogenic enzymes</i>	Specific organs, cell types, or intracellular organelles in which the enzymes, responsible for the biosynthesis of steroids, can be found.
<i>Regulation of steroidogenesis</i>	Processes that play a role in increasing or decreasing the amounts of steroids formed.
<i>Steroid biosynthesis</i>	The process by which hormonally active steroids are formed from cholesterol.

Introduction

The cytochrome P450 (CYP) genes form a family, encoding a class of heme-thiolate enzymes active in the hydroxylation of endogenous and exogenous substances. They are involved in detoxification and catabolism of pharmacological substances and in

biosynthesis of steroids. Homology between the various members is relatively high: CYPs in the same subfamily usually are about 70% similar. These enzymes are instrumental in conveying molecular oxygen to designated positions in the steroid skeleton, in this way introducing hydroxyl groups and determining which final hormonal steroids will be produced in the process of steroid biosynthesis. The specific roles of these enzymes in the pathway of steroidogenesis are indicated in **Figure 1**. The electrons necessary to drive this process are supplied by NADPH or NADH and are passed to mitochondrial CYPs by way of the flavoprotein adrenodoxin reductase to the iron-sulfur protein adrenodoxin. For the CYPs in the microsomal cell fraction, the role of adrenodoxin is taken over by P450 oxidoreductase; cytochrome b5 may play an additional role to support 17,20-lyase activity.

Cholesterol Side-Chain Cleavage Enzyme: CYP11A1

The enzyme CYP11A1 performs the conversion of cholesterol to pregnenolone, by first hydroxylating cholesterol at the 20 α and 22 carbon atoms and then lysing the bond between these carbon atoms. The enzyme, of which only one form is known, is localized in the adrenal glands, the gonads, the placenta, and the brain. Intracellularly, the location is in the inner mitochondrial membrane. Apparently, this localization is extremely important for the activity of the enzyme: transfection studies with constructs that direct the protein into the endoplasmic reticulum failed to show activity.

Regulation of the expression of CYP11A1 differs in different tissues. In organs where an acute stimulation of CYP11A1 transcription can take place under the influence of adrenocorticotrophic hormone (ACTH) or gonadotropins, the adrenals, and the gonads, the responsible cyclic AMP-responsive elements differ from

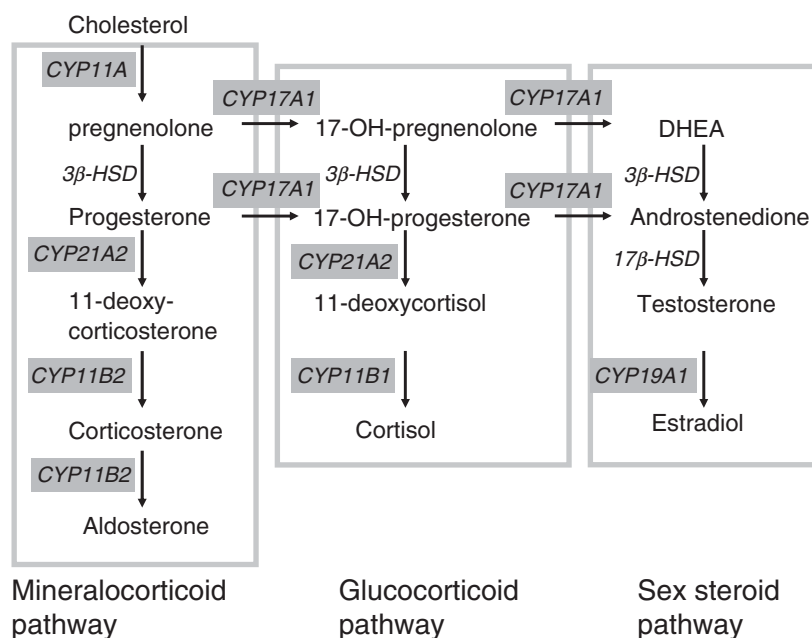


Figure 1 Localization of steroidogenic enzymes along the pathways leading to the biosynthesis of mineralocorticoids, glucocorticoids, and sex steroids. Shaded boxes indicate the various cytochromes P450 (CYPs). DHEA, dehydroepiandrosterone.

those in the placenta. Furthermore, a number of transcription factors, such as SF-1 and Sp-1, play a role in the stimulation of the transcription of CYP11A1. This transcriptional regulation of the amount of enzymes in the glands takes a relatively long time (2–4 h), whereas the stimulation of steroidogenesis may be much faster. These fast effects of increased cAMP levels are mediated through effects on the transport of cholesterol into the mitochondria by the steroidogenic acute regulatory protein (StAR), which enables cholesterol transport to the mitochondrial inner membrane. StAR is not expressed in brain or placenta. It has become apparent that the inability to lyse the C20–C22 bond of cholesterol, as present in congenital lipoid adrenal hyperplasia, may be due to inactivating mutations in CYP11A1, but can also be caused by inactivating mutations of StAR, indicating the crucial importance of this protein in steroidogenesis.

17-Hydroxylase: CYP17A1

After the conversion of cholesterol to pregnenolone in the mitochondria, further conversion of pregnenolone takes place in the endoplasmic reticulum by the conversion to 17-OH pregnenolone by CYP17A1 or to progesterone by 3β-HSD. Progesterone, in turn, can be converted to 17-OH progesterone by the same CYP17A1, which also contains 17,20-lyase activity to convert 17-OH pregnenolone to dehydroepiandrosterone (DHEA) and 17-OH progesterone to

androstenedione. For this reason, CYP17A1 is at the bifurcation of the steroidogenic pathways to glucocorticosteroids and the sex steroids. CYP17 is expressed in the adrenal glands and in the gonads. The production of DHEA, which might be important as a neurosteroid through its interactions with γ-aminobutyric acid (GABA) and N-methyl-D-aspartate (NMDA) receptors, in the brain is an enigma so far because CYP17A1 has not been detected in the brain, which has led to alternative hypotheses for the mechanisms responsible for the production of this steroid involving reactive oxygen species. The expression of the CYP17A1 gene is mainly regulated by cAMP. Addition of cAMP derivatives can stimulate the transcription of the gene; this activity can be inhibited by the addition of phorbol esters, which act via protein kinase C and growth factors such as epidermal growth factor (EGF) and fibroblast growth factor (FGF).

The separate regulation of 17-hydroxylase and 17,20-lyase activity of the enzyme has been resolved on the basis of the finding of mutated genes, causing the production of 17-hydroxylated steroids, without removal of the C20,C21 side chain. These mutations are localized in the region of the protein, which interacts with the additional electron donor cytochrome b5; lyase activity appears to be more vulnerable to the disruption of interaction with this redox partner than the 17-hydroxylase reaction. Similarly, the changing ratios between the production of glucocorticoids and adrenal androgens, mainly DHEA

sulfate, during human development is likely to be due to a changed expression of cytochrome b5.

The pivotal role of CYP17A1 in steroidogenesis is emphasized by the effects of inactivating mutations in this enzyme: a complete block of 17-hydroxylation leads to the inability to synthesize cortisol, whereas aldosterone will be produced in abundance, giving rise to hypertension. Furthermore, sex steroids cannot be produced, causing undervirilization in boys and absence of puberty in both sexes.

21-Hydroxylase: CYP21A2

CYP21A2 catalyzes the conversion of progesterone to 21-hydroxyprogesterone (deoxycorticosterone, DOC) and of 17-OH progesterone to 17,21-dihydroxyprogesterone (deoxycortisol, compound S). These reactions are a prerequisite to synthesize mineralocorticoids and glucocorticoids; the enzyme is expressed in the cortex of the adrenal gland and is not present in the gonads. Using reverse transcription polymerase chain reaction (RT-PCR), expression of CYP21A2 has also been detected in the brain stem, but not in other regions of the brain.

The expression of CYP21A2 is stimulated by cAMP- and protein kinase C-dependent pathways. In this way, the production of deoxycorticosterone in the zona glomerulosa of the adrenal cortex may be stimulated by angiotensin II, whereas the production of deoxycortisol in the zonae fasciculata and reticularis is ACTH dependent.

Due to the fact that an inactive pseudogene, CYP21A1P, exists in the neighborhood of the CYP21A2 gene on chromosome 6, frequent exchanges between these genes have taken place, leading to impairment of the functioning of the resulting protein, the inability to produce aldosterone, corticosterone, and cortisol, and overproduction of adrenal androgens (in humans). These enzyme deficiencies can lead to salt loss in newborn babies and severe masculinization in girls. In female patients with milder forms of the defect, late onset virilization or hirsutism may be found.

11 β -Hydroxylases: CYP11B1 and CYP11B2

The further conversion of 21-hydroxylated steroids to the main mineralocorticoids and glucocorticoids (aldosterone and cortisol [humans] or corticosterone [rat and mouse]) is affected by CYP11B2 and CYP11B1, respectively. The genes encoding these members of the CYP family are found in close vicinity on chromosome 8q and probably arise from ancient gene duplication. The localization of these enzymes in the adrenal cortex is different: CYP11B1

is localized in mitochondria of the zona fasciculata, whereas CYP11B2 is found in mitochondria of the zona glomerulosa. Furthermore, using RT-PCR, mRNAs encoding these enzymes have been detected in various regions of the brain and in the heart.

The regulation of the expression of the two genes is also different. The expression of CYP11B1 is regulated by the ACTH-driven production of cAMP, followed by activation of a cAMP-responsive element, whereas CYP11B2 is stimulated mainly by angiotensin II, which stimulates the production of inositol triphosphate and diacylglycerol by phospholipase C. This leads to increased levels of intracellular calcium and increased activity of phosphokinase C. In this way, the expression of the two 11-hydroxylases can be regulated separately in the zona glomerulosa and zona fasciculata, respectively, of the adrenal cortex.

The clinical results of deficiencies of either of these enzymes can be predicted on the basis of the steroids, the biosynthesis of which will be blocked. Defective CYP11B1 gives rise to low levels of cortisol that can only be increased by supply of exogenous glucocorticoids. An insufficient production of aldosterone, caused by a mutation in CYP11B2, can be partially compensated by overproduction of deoxycorticosterone, which has mineralocorticoid activity.

Aromatase: CYP19A1

The final steroid-hydroxylating enzyme, which plays a role in the biosynthesis of steroids, is CYP19A1. It is responsible for the hydroxylation of carbon atom C19 of testosterone and androstenedione, the subsequent cleavage of the bond between C19 and C10, and subsequent redistribution of the electrons over the A ring to form the aromatic steroids estradiol and estrone, respectively. Only one copy of the gene has been found. It has a complex structure at the 5' end, allowing start of transcription at a number of sites, under the control of various promoter sequences. This means that in various tissues, different mechanisms can be used for the stimulation of the transcription of the gene: in the gonads and adrenals, a cAMP-mediated mechanism is responsible for the stimulation of CYP19A1 by gonadotropins and ACTH, respectively, whereas in fat cells, cortisol can stimulate the production of estrogens. CYP19A1 is also expressed in a large number of other tissues, such as muscle, bone, and brain. In the latter, estrogens formed by this enzyme may play a role in sexual differentiation.

Defects of CYP19A1 are very rare. However, the clinical picture in patients with mutations in this gene shows profound changes of parameters of gonadal function, bone metabolism, and aspects of the

metabolic syndrome, indicating the role of estrogens in the regulation of numerous aspects of homeostasis.

Steroid Catabolism by CYPs

Apart from their role in the biosynthesis of steroid hormones, enzymes of the CYP family also play important roles in the breakdown of steroids. The most important enzymes in this respect are those of the CYP1 and CYP4 subfamilies. Members of the CYP1 subfamily are important in the hydroxylation of mainly estrogens, where both C2 and C4 may be attacked. CYP4 family members hydroxylate corticosteroids and androgens at the 6 position and can also attach a hydroxyl group at C16. It is important to be aware of the fact that expression of all of these enzymes can be induced by xenobiotics, leading to aberrant metabolism of the steroid hormones and, for example, changes in the ratios between the serum levels of steroids produced in the adrenal gland, which might lead to erroneous conclusions of disturbed steroidogenesis.

See Also the Following Articles

11 β -Hydroxysteroid Dehydrogenases; Estrogen; Steroid Hormone Receptors.

Further Reading

de Ronde, W., Pols, H. A. P., van Leeuwen, J. P. T. M., et al. (2003). The importance of oestrogens in men. *Clinical Endocrinology* 58, 529–542.

Fujieda, K., Okuhara, K., Abe, S., et al. (2003). Molecular pathogenesis of lipoid adrenal hyperplasia and adrenal hypoplasia congenita. *Journal of Steroid Biochemistry and Molecular Biology* 85, 483–489.

Grinberg, A. V., Hannemann, F., Schiffler, B., et al. (2000). Adrenodoxin: structure, stability, and electron transfer properties. *Proteins* 40, 590–612.

Lewis, D. F. (2004). 57 varieties: the human cytochromes P450. *Pharmacogenomics* 5, 305–318.

Lisurek, M. and Bernhardt, R. (2004). Modulation of aldosterone and cortisol synthesis on the molecular level. *Molecular and Cellular Endocrinology* 215, 149–159.

Manna, P. R., Wang, X. J. and Stocco, D. M. (2003). Involvement of multiple transcription factors in the regulation of steroidogenic acute regulatory protein gene expression. *Steroids* 68, 1125–1134.

Nebert, D. W. and Russell, D. W. (2002). Clinical importance of the cytochromes P450. *Lancet* 360, 1155–1162.

New, M. I. (2003). Inborn errors of adrenal steroidogenesis. *Molecular and Cellular Endocrinology* 211, 75–83.

Shimada, T. and Fujii-Kuriyama, Y. (2004). Metabolic activation of polycyclic aromatic hydrocarbons to carcinogens by cytochromes P450 1A1 and 1B1. *Cancer Science* 95, 1–6.

Simpson, E. R., Misso, M., Hewitt, K. N., et al. (2005). Estrogen—the good, the bad, and the unexpected. *Endocrine Reviews* 26, 322–330.

Stratakis, C. A. and Bossis, I. (2004). Genetics of the adrenal gland. *Reviews in Endocrine and Metabolic Disorders* 5, 53–68.

You, L. (2004). Steroid hormone biotransformation and xenobiotic induction of hepatic steroid metabolizing enzymes. *Chemico-Biological Interactions* 147, 233–246.

Steroid Synthetic Pathways See: Steroid Hydroxylases; 11 β -Hydroxysteroid Dehydrogenases.

Steroidogenesis Inhibitors See: 11 β -Hydroxysteroid Dehydrogenases.

Strain Differences in Stress Response in Rodents

P Mormède

INRA – Université Victor Segalen, Bordeaux, France

© 2007 Elsevier Inc. All rights reserved.

Strain Differences

Sources of Genetic Variability

Systems Genetics and Genetic Reference Populations

Glossary

<i>Inbred</i>	Obtained by continuous breeding of closely related individuals so as to preserve desirable traits in a stock.
<i>Inbred strain</i>	A strain of organisms that has been maintained by sibling (sister × brother) mating for 20 or more consecutive generations. Except for the sex differences, mice of an inbred strain are as genetically alike as possible, being homozygous at virtually all of their loci.
<i>Isogenic</i>	Genetically identical (except for sex); coming from the same individual or from the same inbred strain.
<i>Intercross</i>	A cross between two animals that have the same heterozygous genotype at designated loci; for example, between sibling F1 hybrids that were derived from an outcross between two inbred strains.
<i>Quantitative trait</i>	A phenotype that is influenced by multiple genes. Its variation is therefore continuous (skin color, weight, etc.).
<i>Quantitative trait locus (QTL)</i>	A polymorphic site on a chromosome containing alleles that differentially influence the expression of a quantitative trait.
<i>Recombinant inbred (RI) strains</i>	Strains constructed by crossing two inbred strains to produce an F1 generation, followed by 20 or more consecutive generations of brother × sister mating. However, the F1 generation may be produced in other ways, such as crossing mice of either existing RI strains, advanced intercross lines, or multiple inbred strains (e.g., an eight-way cross).
<i>Systems genetics</i>	A method for integrating across scales of organization using genetic reference populations. It summarizes results on molecular networks (gene expression data) associated with single-nucleotide polymorphisms SNPs and structural and functional variations.

Strain Differences

Inbred strains of mice and rats are invaluable animal resources for studying the genetic bases of phenotypic variation. They have been obtained through brother-sister mating for more than 20 generations, so all individuals have the same genome (isogenic) and each individual is homozygous at each locus. In a given environment, strain differences result from differences at the genome level.

Large variations in hypothalamic-pituitary-adrenal (HPA) axis activity have been described among inbred strains of rats and mice. Many of these studies are related to the pathophysiology of immune diseases, in which the HPA axis could play a critical role together with immune mechanisms. The Fischer 344 (F344) and Lewis rat strains particularly were studied in this respect, the Lewis female rats bearing an HPA axis that is less reactive to most stimuli and being more sensitive to induced autoimmune diseases such as streptococcal cell wall-induced arthritis or experimental allergic encephalitis. The sensitivity of the Lewis strain can be restored by the removal of the adrenal glands or by treatment with a glucocorticoid antagonist, and F344 rats become resistant when treated with glucocorticoid hormones. They are also studied in relation with the metabolic effects of corticosteroid hormones in the context of nutrition, obesity, and cardiovascular diseases.

Sources of Genetic Variability

It is not easy to detect the primary change(s) responsible for the functional variations of the HPA axis because phenotypic differences may be multiple and the components of the HPA axis interact strongly with numerous facilitation and feedback mechanisms at every level. For instance, when compared to F344 rats, the inbred brown Norway (BN) rats display a number of differences in their HPA axis function, with a different diurnal pattern of plasma corticosterone levels, a faster recovery after restraint stress, an adrenal gland of larger size but less reactive to adrenocorticotrophic hormone (ACTH), and mineralocorticoid receptors (MRs) and glucocorticoid receptors (GRs) hypersensitive to their ligands.

Two main strategies are available for exploring the molecular bases of genetic variation. The detailed exploration of phenotypic differences may suggest candidate mechanisms and/or genes to analyze. For

instance, the comparison of BN and F344 rat strains suggested that the MR is a major source of genetic variability. Therefore, the gene encoding MR (named Nr3c2) was cloned in the two strains and sequenced, revealing in the BN strain a mutation in the coding region with an amino acid change in the protein (Y73C) that increases the transactivating properties of MR agonists.

Another approach, known as quantitative trait loci (QTL) analysis, is based on the association between polymorphic anonymous markers and the phenotypic value of the trait under study in a segregating population (such as F2 or backcross). In this approach, there is no need for any assumption about the gene(s) involved. In the case of the BN/F344 comparison, a QTL analysis confirmed the linkage between the mutation in the Nr3c2 gene and several MR-related phenotypes and disclosed several other loci contributing to phenotypic variability of MR- and GR-related phenotypes. Another study in a BN $\times$ Wistar Kyoto hyperactive (WKHA) intercross revealed the contribution of a locus on chromosome 2 influencing poststress levels of renin and aldosterone. In this locus the angiotensin receptor 1b was mapped, which is responsible for angiotensin II (Ang II) effects on aldosterone secretion, and, indeed, differences were found in the response of adrenocortical cell to Ang II stimulation *in vitro*. This example illustrates the power of the QTL approach to reveal candidate genes possibly involved in phenotypic variation, as well as the multiplicity of mechanisms underlying genetic variation in corticosteroid hormone secretion and action.

Systems Genetics and Genetic Reference Populations

An extension of this QTL approach is the use of recombinant inbred (RI) strains, which are constructed by crossing two inbred strains to produce an F1 generation, followed by 20 or more consecutive generations of brother–sister mating. From a single F1 population, as many strains as possible can be derived. Each strain contains a unique, approximately equal proportion of genetic contributions from the two progenitor inbred strains. Instead of having a single individual for each genetic combination, as in an F2 population, each strain represents an unlimited number of animals with the same genotype, so the phenotype can be measured with more accuracy (in several animals instead of a single F2 animal) and different phenotypes measured across studies are cumulative and can be compared. Furthermore, precise genotyping of the strains with sequence-derived single

nucleotide polymorphisms allows precise QTL mapping. For instance the most used set of recombinant inbred strains with approximately 80 strains has been obtained from the C57BL/6 and DBA/2 mouse strains, and much larger sets (up to 1000 strains) are under construction. All data about these strains are integrated in a unique database named WebQTL that allows a large range of data mining from genetic correlation across multiple traits and QTL analysis. As an example, plasma corticosterone levels after an injection of saline (which can be considered as a mild stressor) was studied in 19 strains. In the WebQTL database of published phenotypes, genetic correlations are found between these levels and a number of phenotypes related to neuroanatomical and neurofunctional measures, response to drugs of abuse, and immune and inflammatory processes. This can be also correlated with gene expression data in different tissues. Therefore, these genetic reference populations allow a systems genetics approach that integrates all these data from diverse sources to explore genetic variation in HPA axis and in its target organs as a complex functional system.

See Also the Following Articles

Corticosteroid Receptors; Hypothalamic–Pituitary–Adrenal; Genetic Variation of HPA Axis Activity and Function in Farm Animals; Genetic Testing and Stress.

Further Reading

- Chesler, E. J., Lu, L., Wang, J., et al. (2004). WebQTL: rapid exploratory analysis of gene expression and genetic networks for brain and behavior. *Nature Neuroscience* 7, 485–486.
- Churchill, G. A., Airey, D. C., Allayee, H., et al. (2004). The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nature Genetics* 36, 1133–1137.
- Marissal-Arvy, N., Lombes, M., Petterson, J., et al. (2004). Gain of function mutation in the mineralocorticoid receptor of the Brown Norway rat. *Journal of Biological Chemistry* 279, 39232–39239.
- Mormède, P., Courvoisier, H., Ramos, A., et al. (2002). Molecular genetic approaches to investigate individual variations in behavioral and neuroendocrine stress responses. *Psychoneuroendocrinology* 27, 563–583.
- Peirce, J. L., Lu, L., Gu, J., et al. (2004). A new set of BXD recombinant inbred lines from advanced intercross populations in mice. *BMC Genetics* 5, 7–8.

Relevant Websites

Jackson Laboratory. www.jax.org.
WebQTL. www.genenetwork.org.

Stress and Anxiety: Treatment with 2nd Generation Antipsychotic Drugs

J S Ballon

Stanford University, Stanford, CA, USA

J A Boyd and D A Wirshing

VA Greater Los Angeles Healthcare System,
Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Second-Generation Antipsychotic Drugs as Anxiolytic Agents in Animal Models of Anxiety
Akathisia – More Prevalent than it Appears?
The 5HT_{1A} Receptor – an Anxiety Receptor for Antipsychotic Drugs?
Histamine – Antipsychotic Drugs as Soporific Agents
Cortisol – How do Stress Hormones Respond to Dopamine Blockade?
Clinical Reports of Second-Generation Antipsychotic Drugs in Anxiety Disorders
Future Directions

Glossary

<i>Akathisia</i>	A state of inner restlessness characterized by pacing, inability to sit still, or other constant urge for movement.
<i>Antipsychotic</i>	A medication used to treat psychosis, i.e. delusions or hallucinations.
<i>Anxiolytic</i>	A medication or treatment used to decrease anxiety.
<i>Anxiogenic</i>	A factor which causes or increases levels of anxiety.
<i>Cholinergic</i>	Related to the neurotransmitter acetylcholine.
<i>Dopaminergic</i>	Related to the neurotransmitter dopamine.
<i>GAD</i>	Generalized anxiety disorder.
<i>Histaminergic</i>	Related to the neurotransmitter histamine.
<i>Hypercortisolemia</i>	An elevated level of the glucocorticoid cortisol in the blood.
<i>OCD</i>	Obsessive-compulsive disorder.
<i>Presynaptic</i>	A receptor located before the connection of nerves (synapse) which influences the nature of the synapse.
<i>SANS</i>	Scale for the assessment of negative symptoms.
<i>Serotonergic</i>	Related to the neurotransmitter serotonin.

Second-generation antipsychotic medications are a heterogeneous group of medications that work at varying levels at many receptors, including dopaminergic, serotonergic, cholinergic, and histaminergic. Varying potencies at these receptors, and specific receptor subtypes, are responsible for not just the

antipsychotic and antimanic properties that these medications all share, but may also account for varying levels of anxiolytic effect. This article will explore some mechanistic answers for this anxiolytic effect including the effect on cortisol, decreased levels of akathisia compared to first-generation antipsychotic drugs, histaminergic effects, and increased potency at the serotonergic (5HT) 1A receptor.

Second-Generation Antipsychotic Drugs as Anxiolytic Agents in Animal Models of Anxiety

Initial efficacy for anxiolytic properties for antipsychotic drugs was controversial, but first observed in rat data. Bruhwyler et al. demonstrated that rats in an open-field paradigm had greater amount of exploration and decreased defecation while on high doses of clozapine. This was in contrast to the effect of haloperidol in a similar paradigm. Additionally, in an elevated plus-maze protocol, rats on a low dose of clozapine demonstrated behavior consistent with anxiogenic properties as they spent less time in the open part of the maze compared with rats who received only the vehicle injection.

Inoue et al. demonstrated that olanzapine and clozapine worked to prevent the acquisition of freezing, but not the expression of freezing, in a conditioned-fear paradigm. This was contrasted however with benzodiazepines, selective serotonin-reuptake inhibitors (SSRIs) and 5HT_{1A} antagonists which both inhibit the acquisition and expression of freezing. However, there was a slight separation from haloperidol and chlorpromazine in the response to the conditioned stimulus.

Further evidence of anxiolytic properties of second-generation antipsychotic drugs was shown in a conflict schedule model in which rats on olanzapine demonstrated similar behavior to those on chlordiazepoxide, but those on haloperidol and risperidone did not. In contrast, risperidone was shown to have short-term anxiolytic properties in a study by Nowakowska et al., however these effects were not seen after 7 days of treatment.

Akathisia – More Prevalent than it Appears?

Akathisia is a troubling side effect that is seen in upwards of 90% of younger patients treated with

first-generation antipsychotic drugs. In a mixed inpatient population, the rate of akathisia in patients on first-generation antipsychotic medications was 18% and pseudoakathisia was 24%. Though the prevalence tends to decrease with age, even in patients aged over 65 years, it is seen in up to 15% of patients.

These data are especially problematic as the core symptoms of akathisia are severe internal restlessness which generally leads to increased motor activity including pacing, marching, and ultimately can manifest as a severe dysphoria. Other common symptoms of akathisia include tension, panic, irritability, and impatience. Suicidality has been linked to akathisia in patients on high potency first-generation antipsychotic drugs. Akathisia is difficult for patients to describe subjectively and can, therefore, often be confused for anxiety. Patients who develop akathisia are also at a much higher risk for nonadherence to treatment and therefore are at a higher risk for relapse.

Second-generation antipsychotic drugs have shown a decreased rate of akathisia. This is considered to be related to the presynaptic blockade of 5HT_{2A} receptors which lead to increased levels of dopamine in the nigrostriatal pathway compared to first-generation antipsychotic drugs. In several trials, olanzapine has been shown to have decreased rates of extrapyramidal side effects including akathisia. In a large meta-analysis, olanzapine, risperidone, and placebo were shown to have decreased rates of akathisia compared with haloperidol.

In addition to the decreased rates of akathisia seen with second-generation antipsychotic drugs, there has been a benefit shown in treatment emergent extrapyramidal symptoms, including akathisia, when patients are switched from haloperidol to olanzapine. Costa e Silva et al. demonstrated an 83% decrease in akathisia symptoms, as measured by the Barnes Akathisia Scale, in patients switched from haloperidol to olanzapine in a large, multicenter trial in Latin America. It may be that some of the clinical observations of anxiety in patients taking first-generation antipsychotic medications is related to under diagnosed akathisia and that anxiolytic effects of second-generation antipsychotic drugs are related in part to the decreased rate of akathisia seen.

The 5HT_{1A} Receptor – an Anxiety Receptor for Antipsychotic Drugs?

The 5HT_{1A} receptor has long been associated with anxiolytic benefit. Buspirone is a partial agonist at 5HT_{1A} receptor and has been considered by many to be an anxiolytic agent. Atypical antipsychotic medications have varying degrees of activity at the

5HT_{1A} receptor, with the strongest affinity seen by aripiprazole and ziprasidone. 5HT_{1A} agonists do not typically occupy the receptor at clinically significant doses, but do alter the overall activity at the receptor. Aripiprazole, ziprasidone, quetiapine, and clozapine have been shown to have partial agonist activity at 5HT_{1A}, with aripiprazole and ziprasidone the most potent. 5HT_{1A} agonism has been linked to increased release of dopamine in the striatum, though it is not clear if this is clinically significant.

Ziprasidone has shown anxiolytic effects in non-psychotic individuals. In a double-blind, randomized trial (30 subjects), ziprasidone was compared to diazepam for anxiety prior to dental procedures. Ziprasidone showed a more rapid, and a slightly more potent, anxiolytic effect compared to diazepam which had a later onset and duration of action. Both were effective, with ziprasidone showing a lesser effect on blood pressure and not showing any greater sedation than placebo, compared to diazepam which was substantially sedating for over 2 hours.

Histamine – Antipsychotic Drugs as Soporific Agents

Histaminergic agents have long been used as treatments for anxiety and insomnia. Histamine antagonists, such as hydroxyzine, have been used as successful treatments in generalized anxiety disorder (GAD), and are often used in posttraumatic stress disorder (PTSD) and acute anxiety. Olanzapine and quetiapine are potent antagonists of the H₁ receptor. H₁ antagonists have also been shown to reduce the level of cortisol, which may be an indirect mechanism for achieving an anxiolytic effect. In this way, the antihistaminergic antipsychotic drugs share characteristics with some of the tricyclic antidepressant drugs, such as clomipramine and imipramine. Clomipramine, a highly histaminergic tricyclic antidepressant, remains one of the most powerful medications for obsessive-compulsive disorder (OCD) currently available. Recent work has identified antipsychotic drugs as useful in augmentation of OCD treatment. Risperidone has the most robust data, but there are reports regarding the use of quetiapine and olanzapine. Antipsychotic drugs have not shown benefit as monotherapy however, but in combination with serotonergic agents provide additional benefit in treatment refractory cases. The data for patients with comorbid tic disorders shows the most benefit from treatment with antipsychotic drugs.

Antipsychotic drugs are also being increasingly prescribed for insomnia. In patients with insomnia secondary to PTSD, olanzapine has shown a benefit as

an augmenting agent, targeting sleep and nightmares. There have been some initial studies regarding alpha-1 antagonists such as prazosin in decreasing nightmares associated with PTSD. Olanzapine and quetiapine are potent alpha-1 antagonists as well, and in that way may provide a benefit in treatment of nightmares and insomnia associated with PTSD. Additionally, the antihistamine properties of these medications also make them sedating and while that may be a treatment limiting side effect in certain circumstances, makes them beneficial agents in the treatment of sleep. Despite these benefits, caution should be exercised in using these medications solely for treatment of insomnia as there are many other treatment options and the side effect profile of these medications restricts their use to only the most refractory cases in which polypharmacy is needed.

Cortisol – How do Stress Hormones Respond to Dopamine Blockade?

Cortisol is a hormone that is typically associated as a biomarker for stress that has also been shown to be elevated in depression and anxiety. It is also been shown to be elevated overall in schizophrenia. A correlation has been seen between increased negative symptoms, measured by the Scale for the Assessment of Negative Symptoms (SANS), and elevations in cortisol, though overall psychopathology scores, medication dosage, and type of medication needed do not necessarily correlate with cortisol levels. Cognitive symptoms, often considered to be most predictive of future prognosis in schizophrenia, are also correlate severity with increased cortisol levels.

The worsened cognitive symptoms with hypercortisolemia may be related to long-term damage to the hippocampus that is seen with chronic elevations in cortisol. In the primate brain, high concentrations of glucocorticoid receptors are seen in the hippocampus and in the primates given chronic dexamethasone treatment, there is an approximately 30% reduction in size and segmental volumes of the hippocampus over time. Similar effects on the hippocampus are seen in primates exposed to animal paradigms of chronic stress. Schizophrenia patients, even at initial presentation, are known to have structural deficits of the hippocampus.

In a double-blind, randomized, cross-over trial, olanzapine and quetiapine showed decreased cortisol levels compared to haloperidol and placebo. The mechanism behind such a direct action on cortisol is not fully understood, but may be most prominently related to the action on serotonergic receptors, with some contribution from histaminergic or adrenergic

blockade. The decreased level of cortisol may relate to the antidepressant and anxiolytic action seen with these agents over the first-generation antipsychotic drugs.

Clinical Reports of Second-Generation Antipsychotic Drugs in Anxiety Disorders

The clinical literature on using antipsychotic medications for anxiety disorders is fairly limited. The mainstay of treatment are primarily the SSRIs and serotonin-norepinephrine reuptake inhibitors (SNRIs) and thus antipsychotic medications are typically used adjunctively when those medications are not fully effective. Other medications that are also utilized include benzodiazepines and antihistaminergic medications and buspirone. Atypical antipsychotic medications are overall used mostly as adjunctive agents in the treatment of refractory anxiety.

A double-blind, placebo-controlled trial of risperidone augmentation in GAD showed decreased core anxiety symptoms on the Hamilton Rating Scale for Anxiety. There have been several small, open label studies in GAD and social anxiety disorder that have had conflicting and ultimately inconclusive results. Risperidone, ziprasidone, olanzapine, quetiapine, and aripiprazole all have open-label studies associated with their use in refractory GAD, but thus far none have been demonstrated as effective monotherapy. While the data is scant, what can be surmised is that there is a subset of patients with severe GAD who are aided by these medications. Particularly, the medications that are most soporific will induce sleep and can be calming to the patients suffering from GAD. Further work and double-blind studies are needed to determine the ultimate clinical benefit from these agents.

In PTSD, atypical agents have most been targeted for their effects on sleep and insomnia. Nightmares and other intrusive thoughts are a common feature of this disorder and have been shown to be generally responsive to atypical antipsychotic drugs. In war survivors in Croatia, monotherapy with fluphenazine, risperidone, quetiapine, or olanzapine led to decreased symptoms of PTSD. The more sedating antipsychotic drugs have shown an increased benefit with promoting more restful sleep with the most data being shown for quetiapine. Other data have been shown in olanzapine and risperidone. Again, these agents have shown the most promise when used adjunctively to serotonergic antidepressants.

There are a number of articles in which antipsychotic drugs are used to control severe OCD symptoms, however, the paradox is that there are also a

number of articles suggesting that the atypical antipsychotic medications may worsen obsessive compulsive symptoms and even bring them on *de novo*. The thought is that the 5HT_{2A} antagonism is the mechanism by which the anxiety/OCD symptoms can worsen.

Future Directions

There is growing support and evidence that antipsychotic drugs may provide anxiolytic benefit for a variety of disorders – psychotic disorders as well as PTSD, OCD, and GAD. The future challenges we face in using this class of agents so broadly is to do so carefully, as there are many untoward side effects that are emerging, most concerning are diabetes, weight gain, and dyslipidemia. Chronic usage of atypical antipsychotic drugs requires close health maintenance and monitoring for subtle changes in metabolic activity. More rare side effects, such as neuroleptic malignant syndrome and dystonia, must also not be forgotten, as they are reported in all the atypical antipsychotic medications, though at frequencies that are less than the first-generation, high-potency agents. As the treatment of anxiety disorders has shifted more towards the primary care setting, education into the less common side effects is important in those prescribing these agents. Additional consideration must be made towards the relatively high cost of these medications compared to other alternatives. Atypical antipsychotic drugs do not characteristically show addictive potential, though there have been reports of diversion of quetiapine, particularly in the prison population. Overall, these agents do show promise, particularly for the refractory patient, and when used judiciously provide a helpful adjunct to therapy for severe anxiety disorders.

Further Reading

- Ahearn, E. P., Winston, E., Mussey, M. and Howell, T. (2003). Atypical antipsychotics, improved intrusive symptoms in patients with posttraumatic stress disorder. *Military Medicine* 168(9), x.
- Baker, R. W., Ames, D., Umbricht, D. S., Chengappa, K. N. and Schooler, N. R. (1996). Obsessive-compulsive symptoms in schizophrenia: a comparison of olanzapine and placebo. *Psychopharmacology Bulletin* 32(1), 89.
- Bartzokis, G., Lu, P. H., Turner, J., Mintz, J. and Saunders, C. S. (2005). Adjunctive risperidone in the treatment of chronic combat-related posttraumatic stress disorder. *Biological Psychiatry* 57(5), 474.
- Brawman-Mintzer, O., Knapp, R. G. and Nietert, P. J. (2005). Adjunctive risperidone in generalized anxiety disorder: a double-blind, placebo-controlled study. *Journal of Clinical Psychiatry* 66(10), 1321.
- Cohrs, S., Röher, C., Jordan, W., et al. (2006). The atypical antipsychotics olanzapine and quetiapine, but not haloperidol, reduce ACTH and cortisol secretion in healthy subjects. *Psychopharmacology* 185(1), 11.
- Inoue, T., Tsuchiya, K. and Koyama, T. (1996). Effects of typical and atypical antipsychotic drugs on freezing behavior induced by conditioned fear. *Pharmacology, Biochemistry and Behavior* 55(2), 195.
- Nemeroff, C. B. (2005). Use of atypical antipsychotics in refractory depression and anxiety. *The Journal of Clinical Psychiatry* 66(Suppl 8), 13.
- Nowakowska, E., Chodera, A., Kus, K. and Rybakowski, J. (1999). Some behavioural effects of risperidone in rats: comparison with haloperidol. *European Neuropsychopharmacology* 9(5), 421.
- Parker, K. J., Schatzberg, A. F. and Lyons, D. M. (2003). Neuroendocrine aspects of hypercortisolism in major depression. *Hormones and Behavior* 43(1), 60.
- Pierre, J. M., Shnyder, I., Wirshing, D. A. and Wirshing, W. C. (2004). Intranasal quetiapine abuse. *American Journal of Psychiatry* 161(9), 1718.
- Pollack, M. H., Simon, N. M., Zalta, A. K., et al. (2006). Olanzapine augmentation of fluoxetine for refractory generalized anxiety disorder: a placebo controlled study. *Biological Psychiatry* 59(3), 211.
- van Lier, S., Vermetten, E., Geuze, E. and Westenberg, H. G. M. (2006). Pharmacotherapy for disordered sleep in post-traumatic stress disorder: a systematic review. *International Clinical Psychopharmacology* 21(4), 193.
- Wirshing, D. A. (2004). Schizophrenia and obesity: impact of antipsychotic medications. *The Journal of Clinical Psychiatry* 65(Suppl 18), 13.
- Wirshing, D. A., Boyd, J. A., Meng, L. R., Ballon, J. S., Marder, S. R. and Wirshing, W. C. (2002). The effects of novel antipsychotics on glucose and lipid levels. *Journal of Clinical Psychiatry* 63(10), 856.
- Wirshing, W. C. (2001). Movement disorders associated with neuroleptic treatment. *Journal of Clinical Psychiatry* 62(Suppl 21), 15.

Stress and CNS Arousal: Genomic Contributions

A C Ribeiro and D W Pfaff

Rockefeller University, New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

Arousal

Neuroanatomical and Neurochemical Basis of Arousal

Amplified States of Arousal

Significance of Studying Arousal

Promising Future Lies Ahead in Arousal Research

Glossary

Arousal The fundamental property of the central nervous system that determines the amplitude of responses to any given stimulus.

Arousal crescent A group of primitive neurons residing in the ventral and medial edges of the brain stem that have ascending projections that alert the brain.

Arousal-promoting mechanisms Monoaminergic cell groups (dorsal raphe nuclei, locus coeruleus, substantia nigra and ventral periaqueductal gray, and tuberomammillary nuclei) that have ascending projections to the thalamus, hypothalamus, and cerebral cortex and whose chemical products favor states of behavioral arousal.

Sleep-promoting mechanisms A small group of neurons residing in the ventrolateral preoptic nucleus that fire predominantly during sleep and send inhibitory projections to the monoaminergic cell groups.

Arousal

Stress responses, whether defined endocrinologically or behaviorally, depend for their expression on the fundamental arousal of the central nervous system (CNS). The relation between stress and arousal is not simple. In terms of a person's psychological feelings, moderate and increasing levels of arousal can be a good thing. However, too much or for too long is both tiring and unpleasant (**Figure 1**).

In the workplace, the relationship between stress and arousal has been defined by occupational psychologists (**Table 1**). Too much demand with too little control can be so stressful as to cause illness. Therefore, it is useful to explore CNS mechanisms underlying the concept of arousal. Here we briefly summarize its operational definition, neuroanatomy, functional genomics, and neurochemistry.

Generalized Arousal

Although intuitively simple to understand, the concept of arousal has not been an easy one to define scientifically. Because all our behaviors and responses to stimuli depend on the underlying state of arousal we are in, it becomes critically important to quantitatively define arousal and to have ways to accurately measure it.

Arousal refers to the degree of stimulation necessary to drive a certain behavior in an animal. The particular stimulus or behavior can vary, but the underlying arousal state determines the magnitude of the response and even if the behavior will occur at all. Another way of envisioning this concept is that arousal is the force or motivation behind a given response, whereas the stimulus dictates what the response modality is.

The overall arousal state of an organism depends on generalized arousal and many specific forms of arousal, such as hunger, thirst, sex, sleep, and fear (**Figure 2**). If we express this concept in the framework of a mathematical equation, then arousal (A) is a function of the general arousal (A_g) state and each individual specific arousal component (A_{s1-n}). Influencing each of these variables is the underlying temperament of the individual, which we express as a constant (K) for each variable ($g, 1-n$). Because these parameters (thirst, hunger, sex, sleep, fear, etc.)

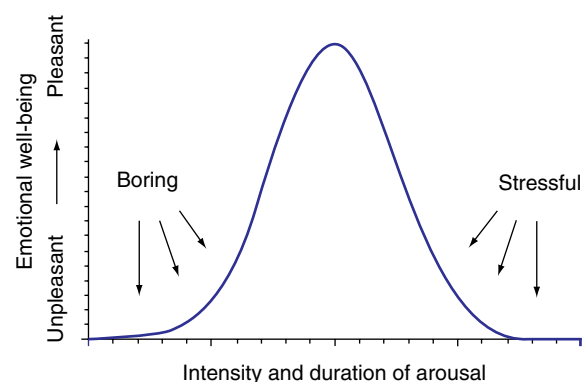


Figure 1 Prolonged arousal can lead to stress.

Table 1 Relationship between occupational demands and employee's control

		<i>Demand for arousal, alertness, and vigilance</i>	
		<i>High</i>	<i>Low</i>
Degree of control by employee	High	Stimulating	Boring
	Low	Stressful	Restful

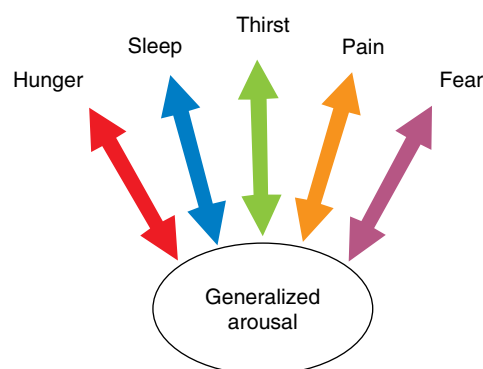


Figure 2 Representation of the arousal equation. Each specific arousal (e.g., pain) depends on the underlying generalized arousal state of the animals, as well as on other specific arousal components (e.g., hunger motivation).

often influence one another, the Arousal Equation can quickly become very complex:

$$A = f(K_g A_g + K_{s1} A_{s1} + K_{s2} A_{s2} + \dots + K_{sn} A_{sn})$$

Traditionally, scientists studying arousal often disagreed on a definition of arousal. Some viewed arousal as one large behavioral phenomenon, where others viewed it as a collection of multiple specific components. The concept outlined in the Arousal Equation unifies, for the first time, these two schools of thought and provides a precise, experimentally feasible manner of determining the arousal state of an animal.

Operational Definition and Quantitative Measures of Arousal

Generalized arousal can be experimentally determined and measured. By definition, an animal is considered to be more aroused if it has increased responsiveness to stimuli (S) from the environment (any sensory modality), increased voluntary motor (M) activity, and also increased ability to react emotionally (E). These three components encompass large classes of varied sensory modalities, which invariably give us a better measure of generalized CNS arousal. This is in sharp contrast to classic neuroscience, which sought to characterize the specific responses of an organism to a specific stimulus. Behavioral arousal is produced by select brain pathways and key populations of neurons being activated, eventually leading to changes in gene-expression patterns.

Using a statistical test called principal components analysis has enabled us to screen large amounts of data (e.g., behavioral responses to the three modalities of stimuli – sensory, motor and emotional reactivity – for each mouse) and uncover trends in the data. The use of this mathematical tool has revealed that generalized arousal accounts for approximately

one-third (from 29.7 to 45%, depending on the experimental conditions) of all the data in arousal-related activities.

Application of Information Theory to Understanding the Neurobiology of Arousal

Introduced by Claude Shannon in the late 1940s, information theory, also known as Shannon's communication theory, establishes that the probability that an event will occur is inversely proportional to the information carried in that event. Consider, for example, the act of flipping a coin. There is a 50% chance of the coin landing heads up versus tails up; therefore, the possible information carried in this system is maximal – the more uncertain an outcome is, the more information transmittal there is. In contrast, if an event has a very high probability of occurring, there is very little room for doubt and very little information carried by the system. In neurobiological terms, the more unexpected the stimulus, the more dramatic the response (e.g., startle) will be. Likewise, the more nonintellectually stimulating environment, the less alert the animal will be and the less likely it is to respond to a stimulus (i.e., the less aroused it is).

Within the CNS, neuronal firing patterns are also governed by the principles of information theory. At any given time, there is a certain probability that a neuron will fire. The closer this probability is to 0.5 (or 50%), the more aroused the animal will be and the larger the repertoire of possible responses it can mount.

Neuroanatomical and Neurochemical Basis of Arousal

Proper arousal responses are critically important for nearly every aspect of our lives. Given this primal role, it is not surprising that nature has evolved ways to ensure that its regulation is safeguarded. This is accomplished by having multiple, redundant, and independent but cooperative mechanisms contributing to the overall arousal state.

Primordial Arousal-Promoting Neurons in the Arousal Crescent

A population of neurons in the reticular formation plays a key role in the regulation of arousal. These neurons originate in the ventral area of the brain stem, at the junction with the spinal cord, and extend bilaterally in a crescent-shaped form through the brain stem and eventually project ventrally into the thalamus and hypothalamus. This neuronal group is among the most evolutionarily conserved networks in animals, adding support to the hypothesis that it plays a crucial role in modulating arousal. These

neurons have a keen ability to sense and respond to a variety of arousing stimuli arising from within the body and the environment, including changes in blood pressure, pain, stress, and hormone levels. These neurons, mostly glutamatergic in nature, have ascending projections to the thalamus, hypothalamus, and basal forebrain, as well as descending projections to the spinal cord. These projections terminate at biologically relevant cell groups that are also involved in arousal control (e.g., the basal forebrain). Projections from this arousal crescent can travel via two routes: a ventral one and a dorsal one. The more ventral one is evolutionarily more primitive and terminates in the cholinergic basal forebrain. The more dorsal route terminates in the thalamus and hypothalamus and is more recent, evolutionarily speaking; it is especially well developed in primates.

Ascending Arousal Activating Pathways

In the 1940s and 1950s, Moruzzi, Magoun, and other neurophysiologists identified a key population of arousal-promoting neurons residing in the brain-stem reticular formation. These cholinergic neurons, found predominantly in the laterodorsal tegmental nucleus and the pedunculopontine tegmental nucleus, have extensive excitatory projections to the thalamus, which in turn project to widespread areas of the cerebral cortex. Acetylcholine release in both the thalamus and the cortex is highest during behaviorally aroused states of wakefulness and rapid eye movement (REM) sleep. In addition to the projections to the thalamus and eventually the cortex, brain-stem reticular neurons also project to the basal forebrain and hypothalamus, which contributes to stabilizing the aroused state.

Monoaminergic nuclei include the serotonergic dorsal raphe nuclei, adrenergic locus coeruleus, and dopaminergic substantia nigra in the brain stem, as well as the histaminergic tuberomammillary nucleus in the posterior hypothalamus. These nuclei all exhibit the highest discharge rates with maximal neurotransmitter release during wakefulness, lower discharge rates in non-REM sleep, and cease firing during REM sleep. The silencing of these neurons, and thus the absence of monoamine neurotransmission, plays a permissive role on REM sleep. These monoaminergic nuclei have ascending projections to the thalamus, hypothalamus, and cerebral cortex that favor states of behavioral arousal (e.g., wakefulness and REM sleep).

Descending Arousal Pathways

In antiphase to these arousal-promoting mechanisms, there are select brain regions that actively promote

cortical synchronization, characteristic of non-REM sleep. One such brain nucleus is the hypothalamic ventrolateral preoptic nucleus (VLPO), whose neurons send inhibitory γ -aminobutyric acid (GABA)ergic and galaninergic projections to the monoaminergic cell groups, thus silencing them.

The lateral hypothalamus also contains neurons that are important for vigilance control. Wake-active orexin-producing neurons have been shown to have excitatory projections to monoaminergic cell groups. Significantly, monoaminergic projections to the VLPO express orexin receptors, and the lack of orexin binding at these receptors causes behavioral state instability associated with narcolepsy. In addition, intermingled with these neurons are melanin-concentrating hormone–GABA-containing neurons that are most active during REM sleep.

Neurochemicals Influencing Arousal

Acetylcholine Originally called vagusstoff (in reference to its release after stimulation of the vagus nerve), acetylcholine was the first neurotransmitter to be discovered. It is produced by the enzyme choline acetyltransferase and can bind to two classes of receptors: nicotinic and muscarinic. The major sources of acetylcholine in the CNS are the basal forebrain, pontine reticular formation, and medial septum. Cholinergic neurons from the reticular formation provide a major excitatory input, stimulating cortical activation, and the discharge patterns of these neurons are highest during behavioral aroused states.

Noradrenaline/norepinephrine In the CNS, norepinephrine is primarily produced by the locus ceruleus neurons, and these neurons promote behavioral arousal, primarily via their projections to the forebrain, brain stem, and spinal cord. Like other monoaminergic cell groups in the brain, locus ceruleus neurons discharge maximally during active waking, decrease their firing rates during non-REM sleep, and are silent during REM sleep. Norepinephrine inhibits VLPO sleep-promoting neurons by binding at α_2 -adrenoreceptors. Also, significantly, norepinephrine increases the female sexual behavior lordosis by modulating the firing rates of ventromedial hypothalamic neurons, and sexual interactions between mice lead to increased norepinephrine release in the hypothalamus.

Serotonin Produced in the dorsal raphe nuclei of the brain stem, serotonin is a monoaminergic neurotransmitter that, like norepinephrine, promotes arousal. Serotonergic neurons fire maximally during waking (quiet waking), decrease their firing rate during non-REM sleep, and cease firing during REM sleep.

Many mood disorders, anxiety, and depression have been linked to deficiencies in serotonergic neurotransmission. Due to the large number of serotonin receptors (14) and their varied functions, studying the pharmacology of serotonin can be quite challenging.

Histamine Produced by tuberomammillary neurons in the posterior hypothalamus, histamine is a biogenic amine involved in vigilance control. With discharge patterns similar to serotonergic and noradrenergic neurons, maximal histamine release occurs during the waking state, and the administration of anti-histamines is often accompanied by drowsiness. Like norepinephrine, histamine also increases sexual behavior in female mice by increasing the firing rates of ventromedial hypothalamic neurons. In addition, sleep-promoting VLPO region neurons send inhibitory GABAergic and galaninergic projections to histaminergic neurons, thus establishing a mutually exclusive, reciprocal interaction between the arousal and sleep-promoting brain regions.

Dopamine Neurons containing dopamine in the ventral periaqueductal gray matter have recently been identified as a major dopaminergic input to modulate arousal. These neurons have the highest discharge rates during wakefulness and have extensive projections to brain regions involved in sleep regulation. Dopaminergic neurons from the ventromedial nucleus and the substantia nigra may also contribute to promote arousal.

Orexin/hypocretin The recently discovered neuropeptides orexin/hypocretin are produced by a small population of neurons in the lateral hypothalamus perifornical area. The two peptides have the same precursor protein and, in addition to playing a role in stimulating appetite, they promote wakefulness. Orexinergic neurons have direct excitatory projections to the tuberomammillary nucleus, and a deficiency in either the orexin peptides or the receptors leads to behavioral state instability and narcolepsy.

Prostaglandin D₂ Prostaglandin D synthase catalyzes the isomerization of the common prostanoid precursor prostaglandin H₂ to the endogenous somnogen prostaglandin D₂. The involvement of prostaglandin D₂ in sleep regulation has been studied for the past 25 years. In fact, it was among the first endogenous sleep-promoting substances to be identified. Prostaglandin D₂ is thought to be involved in homeostatic sleep regulatory mechanisms because animals that lack prostaglandin D synthase or prostaglandin D₂ receptor fail to exhibit a sleep rebound in

response to sleep deprivation. The somnogenic effects of prostaglandin D₂ are hypothesized to be mediated by the nucleoside adenosine, and infusion of prostaglandin D₂ into the subarachnoid space increases extracellular levels of adenosine.

Adenosine The nucleoside end product of energy metabolism, adenosine, has been shown to accumulate in a wake-dependent manner in the basal forebrain. The sleep-promoting effects of prostaglandin D₂ have been shown to be mediated by adenosine 2A receptors in the VLPO. Adenosine has been shown to disinhibit sleep-active VLPO neurons by blocking inhibitory inputs to these neurons and also by directly exciting these neurons. In addition, adenosine can inhibit wake-active neurons in the tuberomammillary nucleus. Interestingly, the most widely consumed arousal-promoting drug by humans, caffeine, works by blocking adenosine 2A receptors.

Steroid hormones Estrogens are produced by the ovaries and are secreted shortly before ovulation. The major role of estrogen is to prepare the female body for courtship, mating, and, eventually, maternal care. As such, estrogens modulate most of the neurochemicals and pathways mediating arousal to ensure that reproduction occurs. Examples are seen in cholinergic neurotransmission, in which estrogens not only increase the rate of acetylcholine synthesis but also increase the number of cholinergic receptors on target neurons. Similarly, estrogens increase the expression of genes involved in norepinephrine production and enhance the responses of dopamine and histamine. Conversely, the levels of the enzyme that produces the endogenous somnogen prostaglandin D₂ is decreased after estrogen treatment. In addition to the well-established roles of estrogens on sexual behaviors, there are many nonreproductive effects of estrogens on behavior, including anxiety, food intake, learning, memory, mood disorders, and sleep.

Integration of Signals

The neuroanatomical pathways discussed so far refer to mechanisms mediating arousal and sleep-wakefulness induction; there are also underlying circadian and homeostatic forces that determine when sleep can occur. The homeostatic drive for sleep, also known as process S, can be envisioned as an increased pressure that progressively gets larger the longer the animal is kept awake and that eventually reaches threshold levels, which trigger sleep. This mechanism is involved in modulating the day-to-day sleep distribution patterns, as well as sleep rebound in response to extended wakefulness. Homeostatic

sleep-regulatory molecules (e.g., prostaglandin D₂, adenosine, and nitric oxide) accumulate in a wake-dependent manner and eventually stimulate sleep. Also communicating with these effector pathways is the circadian component, which determines when, along the 24-h light–dark cycle, sleep and arousal are most appropriate. In fact, the suprachiasmatic nucleus, the circadian pacemaker, has many direct and indirect projections to both sleep-promoting and arousal-promoting brain regions and thus contributes to the fine-tuning of arousal.

Amplified States of Arousal

Sex and Fear

Given the elaborate symphony of mechanisms and chemicals underlying arousal regulation, we quickly come to the conclusion that critical biological functions must rely on it. One such function is reproduction. The propagation of the species, our ultimate goal, is intimately dependent on proper levels of arousal at the appropriate times. The primal function that estrogens have is to prepare the body for reproduction. As such, redundant mechanisms have evolved that ensure that, once the hormonal milieu is optimal and ovulation is near, females exhibit the behaviors that attract mates and result in fruitful copulation. Although conceptually ideal, this scenario does not always occur. Imagine, if you will, that the animals are exposed to a stressful stimulus or are in an unfamiliar place; it may not be feasible to invest time and energy in reproduction when they should be worried, really, about survival. This is, in fact, what the literature tells us occurs in rodents. In the protective shelter of a running-wheel cage or a home cage in our arousal-monitoring equipment, after estrogen treatment the animals exhibited increased arousal (e.g., increased running-wheel activity). However, when the animals were tested in an open field or placed in an unfamiliar cage, the arousal behavior was quickly replaced by fear, resulting in lowered responsiveness by the animal.

A fine line exists between which estrogen-mediated mechanisms are activated, depending on the environmental circumstances the animal finds itself; therefore, caution must be exercised when analyzing and interpreting such behavioral responses.

Significance of Studying Arousal

Arousal-related disorders affect people of all ages and genders. They comprise a wide spectrum of ailments, ranging from problems of vigilance control (e.g., narcolepsy) to hyperactivity disorders (e.g.,

attention-deficit hyperactivity disorder) and, in extreme cases, to devastating vegetative states, all of which can have detrimental and costly consequences.

Disorders of Reduced Arousal

Sleep disorders Forty million Americans are chronically affected by sleep disorders, and an additional 20–30 million report episodic sleep problems. Disorders of sleep range from the most common insomnias to sudden infant death syndrome to hypersomnias. Lack of research and adequate public education has led to a rising epidemic of sleep disorders, especially those associated with obesity (e.g., sleep apnea) and also to a chronically sleep-deprived society. Disorders of arousal mechanisms can be disastrous and costly, causing poor job performance, lack of attention in classrooms, and traffic accidents, all resulting from loss of vigilance.

Extreme hypo-arousal states The most extreme and devastating states in which arousal mechanisms are compromised can be seen in patients in vegetative, comatose, and stuporous states. Although the redundancy of neuroanatomical and neurochemical pathways minimize the possibility that this will occur, situations still exist in which the ability to be aroused is compromised. Only by understanding and characterizing these mechanisms can there be hope that someday it may be possible to stimulate certain key elements of this pathway (perhaps neurons in the arousal crescent) to reverse or ameliorate this condition.

Disorders of hyperarousal Cases of excessive arousal can be seen in patients suffering from attention-deficit hyperactivity disorder and posttraumatic stress disorder. Although they do not appear to be as debilitating as disorders of decreased arousal, being too aroused can interfere with a person's ability to function in society, for example, being able to learn in school or to focus at work. Likewise, the inability to repress unwanted detrimental memories in post-traumatic stress disorder can interfere with the healing process and the ability to resume normal activities.

Promising Future Lies Ahead in Arousal Research

Arousal mechanisms underlie all behavioral aspects of our existence. Although nature has done an astonishing job of ensuring that arousal mechanisms function optimally, as can be seen by the many independent but redundant pathways regulating arousal, situations still exist in which this system

is compromised. Only by knowing the intricacies of this regulation can researchers develop pharmacological and surgical therapies to reversing these problems.

See Also the Following Articles

Anxiety; Circadian Rhythms, Genetics of; Coping and Stress: A Lens and Filter Model; *Drosophila* Genes and Anoxia; Genetic Factors and Stress; Genetic Predispositions to Stressful Conditions; Heat Shock Genes, Human; Heat Shock Proteins: HSP60 Family Genes; Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress; Expression Profiling of Stress Responsive Gene Patterns; Gene Environment Interactions in Early Development; Genetic Variation of HPA Axis Activity and Function in Farm Animals; Genetic Polymorphisms in Stress Response; Genetic Testing and Stress; Neuroticism Genetic Mapping of; Neuroticism Response to

Stress, Genetic Mapping of Mice; Serotonin Transporter Genetic Modifications.

Further Reading

- Borbely, A. A. (2001). From slow waves to sleep homeostasis: new perspectives. *Archives of Italian Biology* 139 (1–2), 53–61.
- Garey, J., Goodwillie, A., Frohlich, J., et al. (2003). Genetic contributions to generalized arousal of brain and behavior. *Proceedings of the National Academy of Sciences USA* 100(19), 11019–11022.
- McEwen, B. S. (2004). *The end of stress as we know it*. Washington, DC: Dana Press/Joseph Henry Press.
- Mong, J. A. and Pfaff, D. W. (2004). Hormonal symphony: steroid orchestration of gene modules for sociosexual behaviors. *Molecular Psychiatry* 9(6), 550–556.
- Pfaff, D. W. (2006). *Brain arousal and information theory: neural and genetic mechanisms*. Cambridge, MA: Harvard University Press.

Stress and Inflammation See: Inflammation.

Stress Effect of Assisted Reproduction

F M Helmerhorst, R Sibug and E R de Kloet
Leiden University Medical Center, Leiden, Netherlands

© 2007 Elsevier Inc. All rights reserved.

The Selye Alarm Reaction
Selye and Reproduction
Stress and Implantation

Glossary

<i>Amenorrhoea</i>	Absence of a menstrual period.
<i>Anorexia nervosa</i>	Relentless pursuit of thinness.
<i>Anovulation</i>	Absence of ovulation.
<i>Blastocyst</i>	Development of embryo with an inner (embryoblast) and outer (later placenta) cell mass.
COS	Controlled ovarian stimulation, by e.g. FSH in order to develop more than one ovarian follicle.
CRH	Corticotropin-releasing hormone.

<i>Endometrium</i>	Inner mucous membrane of the uterus.
<i>flt-1</i>	VEGF receptor.
<i>GnRH</i>	Gonadotropin-releasing hormone.
<i>Gonadotropins</i>	Tropic hormones (e.g. LH and FSH) that stimulate gonadal growth, steroidogenesis and gametogenesis.
<i>Hypercorticism</i>	Increased concentration of cortisol.
<i>KDR/flk-1</i>	VEGF receptor.
<i>LH</i>	Luteinizing hormone.
<i>FSH</i>	Follicle-stimulating hormone.
<i>hCG</i>	Human chorionic gonadatropin.
<i>hFSH</i>	Human follicle-stimulating hormone.
<i>hLH</i>	Human luteinizing hormone.
<i>HPA</i>	Hypothalamic-pituitary-adrenal.
<i>Perinatal</i>	Around birth
<i>VEGF</i>	Vascular endothelial growth factor.

The Selye Alarm Reaction

In Knobil and Neill's *Physiology of Reproduction*, Ferrin defines stress "as the state in which the general adaptation syndrome has been evoked and in which

an increased demand requires the body to perform adaptive functions to re-establish normalcy" (2006: 2627). The three phases differentiated by Selye are the alarm reaction, the first phase in which the stressor affects the balanced system and then attempts to neutralize the stressor by catecholamine release; the second phase, in which the hypothalamic-pituitary-adrenal (HPA) axis acts to stabilize the situation by corticosteroid release; and, when this fails, the third and fatal stage.

Selye and Reproduction

The classic example of the Selye effect in reproduction is still observed when women develop amenorrhea on exposure to extremely violent and life-threatening situations, such as occurred during World War II. Corticotropin releasing hormone (CRH) inhibits gonadotropin-releasing hormone (GnRH) secretion in the hypothalamus, which leads to absence of follicle-stimulating hormone (FSH) and luteinizing hormone (LH) production and then to ovarian quiescence. Simultaneously, hypercortisolism is established. A similar mechanism of an activated HPA axis may lead to amenorrhea among endurance athletes and young women suffering from anorexia nervosa. However, in a consultation room of an infertility clinic, the perception of stress among normally cycling subfertile women is quite different from what we just have described as the Selye effect. The frequently expressed position of subfertile couples is that stress is involved in their infertility. The proof of their conviction as expressed by their friends and described in popular magazines is in the frequently occurring observation that there is a causal relation between taking a long weekend abroad and the desirable pregnancy. "Doctor, how can we avoid stress before and during the *in vitro* fertilization treatment?" asked a couple who was infertile for 5 years.

Three randomized controlled intervention trials illustrate the present situation between stress and assisted reproduction. One study by Kemeter and Feichtinger is low powered and flawed by methodology (e.g. no notification of the allocation concealment, doubts about randomization method), Tarabusi et al. and de Klerk et al. used surrogate endpoints (measuring depression and/or anxiety). From these sparse data we may at least conclude that a relationship between stress as defined in those publications and a hard clinical outcome after assisted procreation has not been properly investigated.

Stress and Implantation

In De Kloet et al., evidence is presented that suggests that stress-related events that occur between conception

and the postnatal period can impact the offspring and permanently change their brains and behavior. The difficulties we encounter studying these processes and extrapolating from animal data to the human situation is illustrated by the fact that challenges that occur as early as ovarian hyperstimulation applied to facilitate blastocyst implantation may affect postnatal life.

Singleton pregnancies from assisted reproduction have a significantly worse perinatal outcome than nonassisted singleton pregnancies, although the prevalence of very low birth weight and very preterm birth is low. Very preterm birth is associated with later neuromotor and cognitive impairment, reduced school performance, and psychiatric morbidity. Young adults who are born very prematurely have different personality types than their term-born peers. This may be associated with an increased risk of psychiatric difficulties.

Subfertility cannot explain the risk of poor perinatal outcome after ovarian stimulation. Assisted reproduction is nearly always preceded by ovarian stimulation, in contrast to cryopreserved and thawed embryo transfer (ET). Because birth weight after cryopreserved and thawed ET is normal or even increased compared to control values, we might hypothesize that ovarian stimulation (with or without an IVF or intracytoplasmic sperm insemination procedure) is associated with low birth weight and might be considered a stress factor in the implantation process.

Blastocyst implantation is intimately associated with vascular permeability and angiogenesis. The vascular endothelial growth factor (VEGF) is a potent stimulator of these two processes. It has four isoforms in rodents, VEGF115, VEGF120, VEGF164, and VEGF188, which bind to the two tyrosine kinase receptors, flt-1 and KDR/flk-1. Our studies showed that VEGF120 and VEGF164 are highly expressed in the uterus. VEGF120, like VEGF164, showed a very distinct transcript signal and also exhibited a spatio-temporal distribution; the mRNA expression of VEGF, flt-1, and flk-1 correlates spatially and temporally with changes in angiogenesis and vascular permeability at implantation sites.

In our study, we tested the hypothesis that treatment with gonadotropins to induce controlled ovarian stimulation (COS) evokes a stress response, whose effects persist after conception. Adult female CD1 mice were treated with urinary human FSH (hFSH) and urinary human chorionic gonadotropin (hCG), recombinant hFSH and recombinant hLH, or with saline alone. The end point tested was the expression of VEGF120 during the mouse peri-implantation period. Urinary gonadotropic treatment caused lower levels of VEGF120, flt-1, and flk-1 mRNA levels, reduced the size of the embryo implantation site,

delayed implantation and prolonged the gestational period; both urinary hFSH and urinary hCG contributed to the adverse effects. Recombinant gonadotropin treatment did not alter any of the parameters studied. In addition, our recent gene-expression profiling study using microarray technology showed that urinary gonadotropins, but not recombinant gonadotropins, exerted differential effects on gene-expression patterns during the murine peri-implantation period.

The CRH system is another factor that is hypothesized to play a role systemically and/or locally in blastocyst implantation. Studies in rodents have shown that CRH expression is 3.5-fold higher in the implantation sites than interimplantation sites and that the injection of a CRH receptor 1 (CRH-R1) antagonist resulted in a 50–70% reduction in the number of implantation sites. Moreover, the administration of antibody against CRH inhibits the attachment of the blastocyst to the implantation site. The activation of CRH receptors during the implantation phase has been demonstrated in the human endometrium. We, therefore, tested in the same system whether the urinary and recombinant gonadotropins influence the CRH system during the mouse peri-implantation period. The levels of CRH and CRH-R1 expression were not influenced, suggesting that the CRH system is not involved in mediating the adverse effects of urinary gonadotropins on the implantation process.

The urinary gonadotropins did not show a significant effect on birth weight, but we observed a positive correlation between birth weight and length of gestation, suggesting that the longer period of gestation allowed time for the fetuses to compensate for the delayed growth and achieve comparable birth weights.

Further Reading

- Brundu, B., Loucks, T. L., Adler, L. J., et al. (2006). Increased cortisol in the cerebrospinal fluid of women with functional hypothalamic amenorrhea. *Journal of Clinical Endocrinology and Metabolism* **91**, 1561–1565.
- De Klerk, C., Hunfeld, J. A. M., Duivenvoorden, H. J., et al. (2005). Effectiveness of a psychosocial counselling intervention for first-time IVF couples: a randomized controlled trial. *Human Reproduction* **20**, 1333–1338.
- De Kloet, E. R., Sibug, R. M., Helmerhorst, F. M., et al. (2005). Stress, genes and the mechanism of programming the brain for later life. *Neuroscience and Biobehavioral Reviews* **29**, 271–281.
- Ferin, M. (2005). Stress and the reproductive system. In: Knobil, E. & Neill, J. D. (eds.) *Physiology of reproduction* (3rd edn., pp. 2627–2694). Boston: Elsevier.
- Helmerhorst, F. M., Perquin, D. A. M., Donker, D., et al. (2004). Perinatal outcome of singletons and twins after assisted conception: a systematic review of controlled studies. *British Medical Journal* **328**, 261–265.
- Kapiteijn, K., de Bruijn, C. S., de Boer, E., et al. (2006). Does subfertility explain the risk of poor perinatal outcome after ivf and ovarian hyperstimulation? Manuscript.
- Kemeter, P. and Feichtinger, W. (1986). Prednisolone supplementation to Clomid and/or gonadotropin stimulation for in-vitro fertilization: a prospective randomized trial. *Human Reproduction* **7**, 441–444.
- Sibug, R. M., Datson, N., Tijssen, A. M. I., et al. (2006). Effects of urinary and recombinant gonadotropins on gene expression profiles during the murine peri-implantation period. *Journal of the Society for Gynecologic Investigation* **13**(2, supplement), 357A.
- Sibug, R. M., de Koning, J., Tijssen, A. M., et al. (2005). Urinary gonadotropins but not recombinant gonadotropins reduce expression of VEGF120 and its receptors flt-1 and flk-1 in the mouse uterus during the peri-implantation period. *Human Reproduction* **20**, 649–656.
- Sibug, R. M., Helmerhorst, F. M., Tijssen, A. M. I., et al. (2002). Estrogen reduces vascular endothelial growth factor 164 expression in the mouse nucleus paraventricularis of the hypothalamus. *Neuroscience Letters* **333**, 199–202.
- Sibug, R. M., Helmerhorst, F. M., Tijssen, A. M. I., et al. (2002). Gonadotropin stimulation reduces VEGF (120) expression in the mouse uterus during the peri-implantation period. *Human Reproduction* **17**, 1643–1648.
- Tarabusi, M., Volpe, A. and Facchinetti, F. (2004). Psychological group support attenuates distress of waiting in couples scheduled for assisted reproduction. *Journal of Psychosomatic Obstetrics and Gynecology* **25**, 273–279.

Stress Effects, Overview

A Steptoe

University College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A Steptoe, volume 3, pp 510–511, © 2000, Elsevier Inc.

Glossary

<i>Coping responses</i>	Cognitive and behavioral responses mobilized in an effort to manage stressful transactions.
<i>Individual-response specificity</i>	The proclivity of an individual to respond to a range of challenges with a consistent profile of physiological activation.

The effects of stress are manifest in four distinct domains: physiology, behavior, subjective experience, and cognitive function. The physiological effects of stress include alterations in neuroendocrine, autonomic nervous system, and immune function and are of great interest to biomedical researchers because of their implications for disease. The behavioral manifestations of stress include aggressive responses, freezing, and displacement activities in animals; health-related behaviors such as smoking and alcohol consumption in humans; and deficits in the performance of complex tasks that are particularly relevant to public safety in areas such as transport and pollution control. Subjective experiences during stress include distress and feelings of dissatisfaction and anger, together with anxiety and depressive responses, which may contribute to psychiatric ill health. The cognitive manifestations of stress include alterations in information processing and attentional and memory functions, which can, in turn, influence performance and decision-making capability.

Stress effects are patterned both within and across these four domains. Attempts to single out a gold-standard measure that provides a consistent indicator of stress responses in all situations have proved unsuccessful. For example, research in the tradition initiated by Hans Selye has emphasized the hypothalamic-pituitary-adrenocortical axis and the output of corticosteroids as central elements of physiological stress responses. Although it is true that this pathway is very sensitive to stressors, it does not necessarily relate closely to other responses. Thus, differences in the stress-induced inhibition of immune responses with varying stressor predictability and controllability

have been found to be independent of corticosteroid responses. Particularly low levels of cortisol have been recorded in victims of posttraumatic stress disorder, suggesting that some manifestations of stress are independent of heightened adrenocortical function. Similarly, it is commonplace in human studies to find weak or negligible correlations between the magnitude of subjective distress and the activation of physiological indices of stress.

Several factors appear to be important in understanding these variations in stress effects. The first is that the pattern of change depends on the nature of the stressful experience and the coping responses mobilized to manage the situation. Thus, naturalistic studies of antagonistic encounters in animals have drawn a distinction between active behavioral responses involving fighting and escape that are associated with sympathetic nervous system activation and more passive conservation-withdrawal behaviors that tend to be linked with parasympathetic responses. At the level of subjective disturbance, depressed mood tends to be elicited by loss events and experiences, whereas anxiety is more characteristic as a response to threatening occurrences.

Second, stress effects vary in their time course and duration. Coping responses that are effective acutely may become maladaptive when stimulation is prolonged. Some important differences in responses emerge only with chronic aversive stimulation, whereas other effects only arise after stimulation is terminated. For instance, studies of human information processing have shown that performance may be maintained during stress exposure due to the exertion of additional effort and that only later manifestations of disturbance emerge. Chronic or repeated stimulation may lead to the failure of regulatory physiological feedback systems, with damaging consequences. Thus, research on socioeconomic status and stress responsivity indicates that differences across the social gradient are more apparent in rates of post-stress recovery than in the magnitude of acute biological responses.

When the implications of stress research for health and well-being are considered, the variations in response between individuals are striking. Two people exposed to the same threatening situation may differ substantially in the magnitude and duration of stress responses, and stress-related health problems may emerge in several contrasting ways both physically and mentally. Some of these variations result

from differences in temperament, social resources, and the effectiveness of the coping responses that the individual brings to bear on the stressful transaction. Others relate to differences in vulnerability, which may arise through both genetic and experiential factors. Genetic variations both in central neurotransmitter metabolism and in peripheral receptor sensitivity contribute to variations in biological response. Animal studies have shown that experiences early in life, such as gentle handling or maternal separation, are associated with decreases or increases in neuroendocrine responses to stressful stimulation in adult life. Clinical investigations of adult psychiatric disorder have uncovered variations in childhood trauma and abuse that sensitize the individual to subsequent stressful events.

Another relevant factor is individual-response specificity. Identified through research on acute psychophysiological responses, this is the tendency for some people to react with a consistent physiological response profile across a range of challenges. One person may be particularly responsive in terms of blood pressure and will show exaggerated blood pressure responses to a range of situations, whereas another is more responsive hormonally or gastrointestinally. Such variations may signal the existence of a special vulnerability or lack of resistance in a particular biological system. Background factors such as nutritional status, physical fitness level, age, gender, and socioeconomic status all serve to moderate the extent to which stress responses may have long-term adverse effects. Hence, the prediction of stress effects in any particular case depends not only on assessing the responses in the four domains with appropriate measures but also on understanding the background factors that predispose to poor adaptation or alternatively bolster the individual's resilience.

Further Reading

- Cannon-Bowers, J. A. and Salas, E. (eds.) (1998). *Making decisions under stress*. Washington, DC: American Psychological Association.
- Charmandari, E., Tsigos, C. and Chrousos, G. (2005). Endocrinology of the stress response. *Annual Review of Physiology* 67, 259–284.
- Charney, D. S. (2004). Psychobiological mechanisms of resilience and vulnerability: implications for successful adaptation to extreme stress. *American Journal of Psychiatry* 161, 195–216.
- Marmot, M. G. and Wadsworth, M. E. J. (eds.) (1997). *Special issue on fetal and early childhood environment: long-term health implications British Medical Bulletin* 53(1).
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* 338, 171–179.
- Pacak, K. and Palkovits, M. (2001). Stressor specificity of central neuroendocrine responses: implications for stress-related disorders. *Endocrine Reviews* 22, 502–548.
- Raison, C. L. and Miller, A. H. (2003). When not enough is too much: the role of insufficient glucocorticoid signaling in the pathophysiology of stress-related disorders. *American Journal of Psychiatry* 160, 1554–1565.
- Steptoe, A. (in press). Psychophysiological contributions to behavioral medicine and psychosomatics. In: Cacioppo, J. T., Tassinari, L. G. & Bernston, G. (eds.) *The handbook of psychophysiology* (3rd edn.). New York: Cambridge University Press.
- Steptoe, A. and Marmot, M. (2005). Socioeconomic position and coronary heart disease: a psychobiological perspective. In: Waite, L. J. (ed.) *Aging, health, and public policy*, pp. 133–150. New York: Population Council.
- Suomi, S. J. (2003). Gene-environment interactions and the neurobiology of social conflict. *Annals of the New York Academy of Sciences* 1008, 132–139.
- Yehuda, R. (2002). Current status of cortisol findings in post-traumatic stress disorder. *Psychiatric Clinics North America* 25, 341–368.

Stress Generation

J E Roberts

University at Buffalo, The State University of New York,
Buffalo, NY, USA

J A Ciesla

Vanderbilt University, Nashville, TN, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by J E Roberts and J A Ciesla, volume 3, pp 512–517,
© 2000, Elsevier Inc.

Life Stress in Psychopathology
Generation of Stressful Life Events
Generation of Stress Perceptions and Threat
Issues in Stress Assessment

Glossary

<i>Appraisal</i>	A cognitive process that involves evaluating the degree of threat associated with a life stressor, as well as the implications of that stressor.
<i>Camberwell life events and difficulties schedule</i>	An interview-based procedure for evaluating the presence, severity, and origins of stressful events and difficulties.
<i>Diatheses</i>	Vulnerability factors that increase the stressfulness and negative consequences of life events.
<i>Difficulties</i>	Ongoing stressful situations that persist for at least 3 months.
<i>Independence</i>	The degree to which a person's behavior or characteristics might have brought about a stressful life event.
<i>Matching events</i>	Acute major negative events that arise from or are thematically related to ongoing chronic stressors or role conflicts.

Life Stress in Psychopathology

Links between life stress and mental disorders such as depression have been drawn throughout history. For example, in the seventeenth century Robert Burton described how melancholia often was precipitated by environmental adversities, including interpersonal losses. Indeed, the very phenomenology of depression involves feelings of loss and ruminations on the hopelessness of one's life situation. When depressed, one experiences the world as impoverished and as imposing impossible demands. In such a frame of mind, ordinary hassles become insurmountable obstacles, deprivations, or failures over which people despair.

Thus, depression itself magnifies the apparent stress and strain of life.

Investigators have typically examined the role that life stress might play in the onset and maintenance of mental disorders. In other words, the focus has been on how life stress can result in psychopathology. However, it is also clear that life stress often results from individuals' behavior or characteristics. Many stressful life events and difficulties, therefore, are not random occurrences that happen to befall the unfortunate, but rather are at least in part created by the same individuals who suffer from them.

Likewise, certain individuals tend to react to negative life events with greater feelings of stress and distress than others. Increasingly, investigators are exploring the manner in which characteristics of the person and the environment can breed stressful life events and difficulties, as well as the more subjective stress reactions to those events. This generated stress, in turn, potentially contributes to other adverse psychological outcomes, such as episodes of depression. Although this is largely studied in the context of depression, an important area of future research involves determining whether other forms of psychopathology also contribute to stress generation.

As suggested earlier, stress generation can involve two separate processes. First, characteristics or behavior of the individual can contribute to the occurrence of actual life events. In other words, for various reasons, some people are more prone to encountering negative life events and difficulties than others. Second, characteristics of the individual can contribute to the level of threat or perceived stress associated with particular life events. In other words, given seemingly identical life events, some people will experience more severe subjective stress reactions than others.

Generation of Stressful Life Events

It is clear that people tend to select or even forge environments that reflect their personalities and genetic predispositions. Individuals with various temperaments create social-environmental niches that are consistent with these characteristics, and in turn these environments are likely to be associated with various types of life events. For example, studies have demonstrated that persons with higher levels of the personality characteristic neuroticism are more prone to experiencing negative life events, whereas those who are higher in extroversion are more prone to experiencing positive life events. Likewise, it is

quite possible that cognitive diatheses associated with depression, such as negative attributional styles or dysfunctional beliefs about interpersonal relationships, can lead to future stressful life events. For example, an individual who is cognitively biased to interpret ambiguous social encounters as rejection is likely to create the very abandonment so feared. The person's negative interpretations are likely to lead to conflict and rejection. In a similar vein, there is strong evidence that individuals genetically predisposed to depression are at heightened risk for negative life events. For example, the nondepressed biological relatives of depressed persons have been shown to be at greater risk of experiencing major life stressors than relatives of nondepressed individuals. Several quantitative behavioral genetic studies have confirmed this genetic contribution to stressful life events. Overall it is apparent that negative life events are not entirely randomly distributed and, instead, tend to cluster in certain persons. Furthermore, these individuals seem to be the same persons who are most vulnerable to psychological disorders, particularly depressive episodes, following life stress.

Brown and colleagues were among the first to explore the role that psychiatric symptomatology might play in generating stressful life events. Initial efforts by this research team were directed at establishing that negative life events could play a causal role in the development of subsequent depressive episodes. It was therefore critical that these investigators focused on events that clearly arose independently of psychiatric symptomatology. Across a number of studies, these and other researchers demonstrated that major acute stressful life events that were independent of depression itself (as opposed to events such as hospitalization for depression or losing one's job because of concentration difficulties) contributed to the onset of episodes of this disorder. However, these investigators also established that individuals with subclinical depressive conditions, as well as other psychiatric conditions, were at greater risk of encountering environmental adversities than those who were without symptomatology. In other words, individuals with subclinical depression or other psychiatric conditions were more likely to be faced with major acute stressful life events than were asymptomatic individuals. These acute stressors, in turn, were associated with the onset of episodes of clinical depression. Interestingly, subclinical depression and other psychiatric conditions failed to predict future depressive episodes in the absence of life stressors. Apparently, stress generation was largely responsible for the exacerbation of depressive symptoms, leading subclinical experiences to develop into full depressive episodes.

Furthermore, subclinical depression and other psychiatric conditions were associated with the generation of life events that are particularly depressogenic. Specifically, these conditions bred acute life events that matched either ongoing life difficulties (e.g., a conflictual marriage) or role conflicts (e.g., a conflict between one's career and family). These types of matching events have proven to be extremely potent contributors to the onset of episodes of clinical depression. In other words, matching events seem to be considerably more likely to result in clinical depression than other types of negative life events. Unfortunately, depression-prone individuals appear to generate the very types of stressors that are most likely to exacerbate their psychiatric conditions.

In a highly influential paper, Hammen also suggested that stress generation might be an important mechanism through which depression can lead to future depression. Across a number of studies based on adult, adolescent, and child samples, Hammen and colleagues have demonstrated that depression-prone individuals experience a greater frequency of subsequent life stressors compared to those with other psychiatric conditions, those with medical illnesses, and healthy control participants. Interestingly, this research suggests that interpersonal stressors, in particular, are generated by depression-prone individuals. Interpersonal stressors largely include conflicts and tension between people. Importantly, Hammen's work has demonstrated that stress generation can take place even during periods of remission, suggesting that factors beyond depressive symptomatology itself are likely to be at work. In other words, stress generation appears to be driven by some correlate or correlates of depression rather than depressive symptomatology itself.

What might these correlates be? Studies have explored aspects of personality pathology and interpersonal functioning, such as poor interpersonal problem solving and excessive reassurance seeking. For example, symptoms of cluster A (odd and eccentric) and cluster B (dramatic and erratic) personality pathology have been found to contribute to the generation of interpersonal stress, which in turn predicted the development of depressive symptomatology. Other research has focused on two personality styles implicated in vulnerability to depression, sometimes referred to as interpersonal relatedness (a style marked by excessive dependence on relationships) and self-definition (a style marked by excessive reliance on achievement and individuality in maintaining self-esteem), and has found that self-definition is particularly related to stress generation. In terms of interpersonal functioning, one study found that the generation of minor social stressors mediated the association between excessive

reassurance seeking and the development of depressive symptoms, whereas another found that the generation of interpersonal stressors mediated the association between poor interpersonal problem solving and subsequent depressive symptoms. Another line of research conducted by Kendler and colleagues has found that neuroticism and genetic liability to depression contribute to stress generation, whereas childhood sexual abuse and variation in the serotonin transporter gene increase individuals' sensitivity to the depressogenic effects of life stressors. Together, these studies suggest that personality pathology, interpersonal functioning, and some forms of genetic liability are associated with future depression by increasing the likelihood that individuals encounter stressful life events (particularly interpersonal stressors), whereas traumatic stress early in life and other types of genetic variability increase stress sensitivity. It will be important for future work to determine whether these characteristics contribute to stress generation and sensitivity on their own, or whether their effects become more potent in combination with pre-existing depressive symptomatology.

In addition to these characteristics of the person, it is likely that characteristics of the person's social environment are also related to stress generation. Brown and colleagues have documented how other ongoing aspects of the social world are associated with risk for future acute life events. These investigators developed the concept of life structure as being important in understanding the origins of life stress. In many cases, individuals who are faced with various forms of deprivation during childhood (e.g., parental loss, abuse, or neglect) become locked into an ongoing conveyor belt of environmental adversity and psychological vulnerability. These childhood environmental experiences put individuals at risk for further negative environmental experiences during later stages of life. For example, parental lack of care, neglect, and abuse during childhood can lead to teen pregnancy and marriage to an undesirable partner. In turn, these experiences set the stage for a variety of additional acute stressful life events and difficulties that are likely to occur in the future, such as marital violence, economic hardship, and infidelity. Simultaneously, this conveyor belt of adversity would continuously foster and reinforce psychological vulnerabilities, such as feelings of helplessness and low self-esteem, making it increasingly difficult to cope with stressors.

In summary, acute negative life events are more likely to be encountered by individuals with certain characteristics. In particular, this type of stress generation is associated with some forms of genetic predisposition, neuroticism, depression, and, quite

possibly, cognitive characteristics that are thought to increase risk for depression. Further, these generated life events appear to be characterized by interpersonal conflict and a resonance with other ongoing life stressors or role conflicts. It is presently unclear why depression-prone individuals are at risk for generating life stress. Important directions for future work include examining theoretically plausible mediators and moderators of this effect. For example, Hammen and colleagues have begun to explore the possible roles of interpersonal functioning, comorbidity, and personality pathology, while Brown and colleagues have focused on the individual's unfolding social-environmental context. Future work also needs to examine stress generation in psychological disorders beyond depression.

Generation of Stress Perceptions and Threat

Individuals differ in the degree to which they experience negative reactions to seemingly identical life events, i.e., stress reactions are determined not only by the nature of the event, but also by the nature of the individual. Lazarus and Folkman proposed that stress itself is best thought of as a process in which the individual actively mediates the relationship between life events and psychological distress. In response to a life event, individuals are thought to first engage in a primary appraisal of the potential negativity or threat. Different individuals are likely to appraise the same event in quite different ways. One person may view a job demotion as an irreversible career blemish, while another may view this event as only a minor setback. Subsequent to this initial threat appraisal, individuals are thought to engage in another assessment. This secondary appraisal involves further evaluation of the situation and the degree to which the person has the resources available to effectively deal with the event. In a number of ways, this involves cognitive, emotional, and social aspects of the individual. Finally, following primary and secondary appraisal, different coping strategies can be initiated. Individuals are known to differ in their coping efficacy and typical coping styles, both of which are likely to effect psychological adjustment to life stress. In describing this stress process, Lazarus has been influential in identifying multiple ways in which properties of the individual can affect vulnerability to life stress.

Lazarus' model highlights the importance of individual differences in appraisal processes as contributing to the generation of stress reactions, but what underlies and influences these appraisals? From a theoretical perspective, other types of vulnerability

(so-called diatheses) may contribute to these negative appraisals. Diathesis stress models posit that certain characteristics increase the stress that individuals experience in response to negative life events. One such vulnerability factor that has received empirical support is attributional style.

According to Abramson and colleagues, individuals are motivated to understand and make sense of the causes of life events and tend to have relatively stable characteristic attributional styles. Persons who attribute negative life events to causes that are global (i.e., have far-reaching implications), stable (i.e., are predictive of future negative events), and internal (i.e., are caused by the self) are more likely to experience psychological distress than individuals who attribute negative life events to causes that are specific (i.e., are limited in life domain), unstable (i.e., are not likely to occur again), and external (i.e., are caused by the environment). While this is true of negative events, the opposite pattern is considered dysfunctional in the context of positive life events.

In addition to negative attributional styles, rigid, inappropriate, and perfectionistic beliefs about the self and the world, known as dysfunctional attitudes, likely contribute to the generation of stress perceptions and threat. Many of these beliefs concern contingencies of self-worth that are unrealistic in nature (e.g., one must be perfect or admired by all in order to be a person of worth or deserving of love). These cognitions are thought to interact with life stressors to result in psychological disorders. In this sense, a negative life event may be experienced as distressing in itself, but may also be considered by the individual to be an indication of a lack of self-worth. Through these beliefs, an event such as being laid off due to a factory closing may become a sign of one's personal incompetence and lack of value, generating more negative appraisals and perceived stress than would be generated by someone without these dysfunctional beliefs. In this way, the manner in which an individual's self-esteem is maintained and structured influences how much perceived stress is generated by particular classes of negative life events.

It is also apparent that environmental factors can contribute to these appraisal processes and, consequently, to the degree of stress generated by the occurrence of negative life events. For example, individuals with inadequate social support are likely to have more severe stress reactions than those with more supportive close relationships. Brown and colleagues have found that individuals who were "let down" during the negative event itself were more likely to develop episodes of depression than those who received expected crisis support. In the former situation, individuals' positive expectations about their close relationships would be shattered, leading to disappointment and

a reevaluation of the nature and quality of these relationships. Consequently, the event and its surrounding context would likely be appraised as considerably more threatening than if support had been forthcoming.

Likewise, individuals who are already faced with significant life stressors, particularly severe ongoing difficulties, are more likely to suffer from more severe stress reactions in response to additional negative life events than those without these additional burdens. Making matters more complicated, Coyne and Wiffen have argued that it is likely that these environmental factors are closely associated with the cognitive vulnerabilities discussed previously. For example, to some extent attributions about the implications of negative interpersonal events or dysfunctional beliefs about relationships likely arise from conflictual interpersonal relationships, such as unstable marriages. It becomes less clear whether it is primarily the cognitive characteristic (e.g., negative attributional style or dysfunctional beliefs about relationships) or the environmental factor (e.g., conflictual marriage) that contributes to the generation of stress reactions following an acute negative life event (e.g., a major argument with one's spouse).

In summary, characteristics of the individual and his or her social environment contribute to the degree of stress and distress that is generated in response to negative life events. These characteristics include appraisal processes that, in turn, are likely to be driven by underlying cognitive diatheses (such as negative "attribution styles" and dysfunctional beliefs), as well as environmental characteristics (such as inadequate social support and chronic stressors). An important area of future investigation involves attempting to disentangle the overlapping roles of these cognitive and environmental characteristics.

Issues in Stress Assessment

The notion of stress generation mandates that careful attention be paid to a number of issues in the assessment of life stress. First, it becomes crucial to examine the dimension of event independence. Presumably, generated events are those that were at least somewhat dependent on the individual's behavior or characteristics. Second, as previously discussed, stress generation involves two different processes: the first reflects the generation of actual negative life events, whereas the second reflects the generation of stress perceptions and reactions. Assessment procedures must be able to distinguish between these two processes. Third, it is becoming clear that in terms of objective events, stress generation is specific to certain classes of events, namely, events that match ongoing problems or that involve interpersonal

conflict. Assessment procedures need to be able to cleanly distinguish these classes of events from others.

In contrast to the stress generation studies conducted by Brown, Hammen, and their colleagues, the vast majority of research on the role of life stress in depression has relied on self-report checklists to assess life events. Unfortunately, self-report measures of life stress likely combine both relevant and irrelevant aspects of the stress construct, making it difficult to advance knowledge in this area. First, self-report measures of life events are unable to partition the occurrence of objective life events from the perception of life stress. As a result, it is unclear whether a participant's score is the result of actual changes in the environment or the perception of changes. A second major limitation of self-report life event measures is that they are unable to assess the degree to which events were influenced by the person or by other independent causes. Clearly, in order to investigate the hypothesis that individuals contribute to their own stress, it is necessary to establish that a given stressor was the result of an individual's behavior and was not completely independent or random. Finally, self-report measures make it difficult, if not impossible, to classify events into particular domains (e.g., interpersonal and matching versus nonmatching) in order to examine specific types of events.

Importantly, Hammen and colleagues' work suggests that depression generates interpersonal conflict events, but not other types of events. As such, these interpersonal conflict events should be examined separately from other types of events. It is notoriously difficult to make these distinctions on the basis of self-report checklists. For example, events that seemingly reflect noninterpersonal experiences (such as loss of one's job) often can involve an interpersonal conflict (such as a fight with one's boss). More in-depth probing is required to determine whether or not an event had significant interpersonal overtones.

Clearly, advancements in this area of inquiry are going to require the use of more sophisticated interview-based procedures that can more cleanly assess particular dimensions and domains of stressful life events. The Camberwell Life Events and Difficulties Schedule (LEDS) is the most well-developed and methodologically sound system. The LEDS first involves careful questioning about potential events in a variety of different life domains, as well as in-depth probing about the context surrounding each event. Events are then rated on a number of dimensions, including threat and independence, by a team of individuals who did not conduct the interview and who are kept blind to the psychological reactions to the events. In order to evaluate the likely individual meaning and importance of the event, ratings of degree of threat (stress) are made on the basis of the context

surrounding the event. So, for example, a pregnancy would be rated quite differently if it occurred in the context of a happy marriage and good financial resources than it would if it occurred in the context of marital violence, financial hardship, and poor living quarters. Of particular importance is the fact that the LEDS includes more than 5000 examples of different events that help anchor ratings of threat and independence. In addition, guidelines are provided to help determine what counts as an event. These guidelines set thresholds for inclusion and exclusion. Consequently, they help prevent individuals who report a number of very minor and trivial incidents from receiving higher scores than those who experienced the same incidents, but who failed to mention them during the interview. The LEDS is currently the only interview that includes these important features.

The notion of stress generation implies that the stressful life event in part resulted from the person's behavior or characteristics. As such, it is crucial that the dimension of life event independence is examined. The LEDS classifies events on a 12-point scale that includes categories that vary in terms of their independence of the person's behavior. Point 1 refers to events that are completely independent, such as accidents occurring to family members; point 2 refers to events for which it is impossible to completely rule out some aspect of the subject's behavior in leading to the event, e.g., a close friend deciding to move to another city; point 3 refers to events that possibly arose through the subject's behavior, but is unlikely, e.g., a child's academic difficulties; point 4 refers to physical illnesses of the subject; point 5 refers to events in which the subject consented or complied with external circumstances, e.g., agreeing to care for a sick relative; point 6 refers to intentional acts of the subject, e.g., deciding to return to school; point 7 refers to events that probably arose as a result of the subject's negligence, e.g., an automobile accident while subject was intoxicated; point 8 refers to events involving arguments or breaking off contact with someone after tension; point 9 refers to breaking off contact with someone without a precipitating argument or period of tension; point 10 refers to the subject's love events; point 11 refers to the subject's partner's love events, e.g., a partner's affair; and point 12 refers to events that are clearly the result of depression, e.g., fatigue, irritability, and concentration problems contributing to reprimands from one's boss and, eventually, to being fired.

In summary, research is beginning to document the ways in which individuals prone to psychological disorders, particularly depression, contribute to the stress that they experience in their lives. In turn, this generated stress might effect the onset, maintenance,

and recurrence of these conditions. Our developing understanding of these complicated transactional pathways suggests that more refined evaluations of life stress are required to advance knowledge in this area. In-depth interview-based approaches, such as the LEDS, are required to evaluate life events in terms of their severity, independence, content, and relation to other ongoing life problems.

See Also the Following Articles

Life Events and Health; Life Events Scale; Genetic Predispositions to Stressful Conditions; Psychosocial Factors and Stress.

Further Reading

- Abramson, L. Y., Alloy, L. B. and Metalsky, G. I. (1989). Hopelessness depression: a theory-based subtype of depression. *Psychological Review* **96**, 358–372.
- Brown, G. W. and Harris, T. O. (1978). *The social origins of depression*. New York: Free Press.
- Brown, G. W. and Harris, T. O. (eds.) (1989). *Life events and illness*. New York: Guilford Press.
- Coyne, J. C. and Whiffen, V. E. (1995). Issues in personality as diathesis for depression: the case of sociotropy-dependency and autonomy-self-criticism. *Psychological Bulletin* **118**, 358–378.
- Daley, S. E., Hammen, C., Davila, J. and Burge, D. (1998). Axis II symptomatology, depression, and life stress during the transition from adolescence to adulthood. *Journal of Consulting and Clinical Psychology* **66**, 595–603.
- Davila, J., Hammen, C., Burge, D., Paley, B. and Daley, S. E. (1995). Poor interpersonal problem solving as a mechanism of stress generation in depression among adolescent women. *Journal of Abnormal Psychology* **104**, 592–600.
- Hammen, C. (1991). Generation of stress in the course of unipolar depression. *Journal of Abnormal Psychology* **100**, 555–561.
- Kendler, K. and Kartowski-Shuman, L. (1997). Stressful life events and genetic liability to major depression: genetic control of exposure to the environment? *Psychological Medicine* **27**, 539–547.
- Lazarus, R. S. and Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.
- Monroe, S. M. and Roberts, J. E. (1990). Conceptualizing and measuring life stress: problems, principles, procedures, progress. *Stress Medicine* **6**, 209–216.
- Monroe, S. M. and Simons, A. D. (1991). Diathesis-stress theories in the context of life stress research: implications for the depressive disorders. *Psychological Bulletin* **110**, 406–425.
- Potthoff, J. G., Holahan, C. J. and Joiner, T. E. (1995). Reassurance seeking, stress generation, and depressive symptoms: an integrative model. *Journal of Personality and Social Psychology* **68**, 664–670.
- Roberts, J. E. and Monroe, S. M. (1994). A multidimensional model of self-esteem in depression. *Clinical Psychology Review* **14**, 161–181.
- Scarr, S. and McCartney, K. (1983). How people make their own environments: a theory of genotype – environment effects. *Child Development* **54**, 424–435.

Stress Hyporesponsive Period

S Levine

University of California, Davis, CA, USA

E R de Kloet

Leiden/Amsterdam Center for Drug Research, Leiden, Netherlands

G Dent

AstraZeneca Pharmaceutical Co., Wilmington, DE, USA

M S Schmidt

Max Planck Institute of Psychiatry, Munich, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S Levine, G Dent, and E R de Kloet, volume 3, pp 518–526, © 2000, Elsevier Inc.

Stress Hyporesponsive Period

Stress Hyporesponsive Period, the Adrenal, and Corticosterone

Stress Hyporesponsive Period and the Pituitary

Stress Hyporesponsive Period and the Brain

Corticosteroid Feedback

Conclusion

Glossary

Maternal deprivation

Separation of mother and infant during the stress hyporesponsive period, which needs to last for at least 8 h for immediate and persistent effects on the neuroendocrine regulation of the hypothalamic-pituitary-adrenal axis.

Stress hyporesponsive period (SHRP)

A period of reduced adrenal corticosterone and pituitary adrenocorticotrophic hormone release in response to stress, lasting in the rat from postnatal day 4 to 14 and in the mouse from postnatal day 1 to 12.

Based primarily on the pioneering work of the late Hans Selye, the stress response has become somewhat synonymous with the release of hormones from the pituitary and the adrenal glands. Thus, in most adult mammals stimuli presumed to be stressful result in a systematic release of adrenocorticotrophic hormone (ACTH) and the subsequent secretion of glucocorticoids from the adrenal. This simplistic view of the pituitary-adrenal axis as first described by Selye has been elaborated on extensively. The regulation of the stress hormones clearly involves specific peptides synthesized and stored in the brain, corticotropin-releasing factor (CRF) and arginine vasopressin (AVP), and brain-derived neurotransmitters, such as norepinephrine. Thus, the brain must be included as a critical stress-responsive system. However, the sequences of responses observed consistently in the adult are in many ways very different in the developing organism. Abundant evidence indicates that the rules that govern the activity of the hypothalamic-pituitary-adrenal (HPA) axis in the adult are very different in the neonate.

Stress Hyporesponsive Period

In 1950, a report appeared that first indicated that the neonatal response to stress deviated markedly from that observed in the adult and thus created a field of inquiry that has persisted for over 4 decades. Using changes in adrenal ascorbic acid as the response to stress, Jailer reported that the neonate did not show any response to stress. By the early 1960s, Shapiro placed a formal label on this phenomena, designating it the stress nonresponsive period (SNRP). It is important to note that for the most part the database for this description was the inability of the rat pup to show significant elevations of corticosterone following stress. There was one study that received little attention at the time but did raise important questions concerning the validity of the notion of an SNRP. In that study, in addition to exposing the pup to stress and its demonstrating a lack of a corticosterone response, another group was injected with ACTH. These pups also failed to elicit a corticosterone response, which indicated that one of the factors contributing to the SNRP could be a decreased sensitivity of the adrenal to ACTH. Therefore, it was conceivable that other components of the HPA axis might be responsive to stress. The resolution of this question was dependent on the availability of relatively easy and inexpensive procedures for examining other components of the HPA axis. The methodological break through, which altered most of endocrinology and had a major impact on our understanding the ontogeny of the stress response, was the development of radioimmunoassay (RIA) procedures.

The initial impact of the RIA was to change the designation of this developmental period for the SNRP to the stress hyporesponsive period (SHRP). This change was a result of studies that showed a small but significant rise in corticosterone when measured by the RIA. Thus, although the response of the adrenal was reduced markedly during the SHRP, the adrenal was capable of releasing low levels of corticosterone.

It has been argued that the biological function of the SHRP is to protect the developing organism, specifically the brain, from an excess of circulating glucocorticoids. There is abundant evidence that high levels of glucocorticoids suppress neural growth; thus, under basal circumstances, it seems beneficial for the neonate to suppress the activity of the stress system. However, in situations that threaten health or even survival, an activation of the HPA system might be crucial. The HPA axis during development is unique in that it maintains a delicate balancing act between the regulation of stable levels of corticosterone and the ability to respond to systemic stimuli that may be life threatening. Thus, the neonate can more precisely discriminate between stressors that require higher-level cognitive processing and life-threatening events.

Indeed, when investigators began to examine other components of the neonate's HPA axis it became apparent that the SHRP was by no means absolute and that the different components of the axis responded differentially and were stimulus-specific. The question addressed in this article is whether the concept of the SHRP is still valid. In order to confront this question, we examine the development of several components of the HPA axis: the adrenal, the pituitary, and the brain.

Stress Hyporesponsive Period, the Adrenal, and Corticosterone

This article focuses on how the development of the different components of the HPA axis bears on the issue of the SHRP. It is generally agreed that in response to most stressors the neonate fails to elicit adrenocortical response or does so minimally. There are several features that characterize the function of the pup's adrenal. The first and most obvious characteristic of the adrenal is that basal levels of corticosterone during the SHRP are considerably lower than observed immediately following parturition and after the pup emerges from the period of adrenal quiescence, which is normally between postnatal day 4 and 14 in the rat (postnatal day 1 and 12 in the mouse). Further, numerous investigators have reported that the neonate can elicit a significant increase in plasma levels of corticosterone following a variety of challenges. However, invariably the

magnitude of the response is small compared to older pups that are outside the SHRP and, of course, to the adult. Thus, whereas the reported changes in corticosterone levels following stress in the adult can at times exceed $50 \mu\text{g dl}^{-1}$, rarely does the infant during the SHRP reach levels that exceed $10 \mu\text{g dl}^{-1}$. These levels are reached only under special circumstances. Thus, the ability of the neonatal adrenal to secrete corticosterone seems to be impaired markedly.

There is an important caveat in making the assumption that the reduced level of corticosterone following stress indicates a reduction in biological activity. Corticosterone exists in the circulation in two forms, bound and unbound. The large majority of corticosterone in the adult is bound to corticosteroid-binding protein (CBG) and other binding proteins. Only a small fraction exists in the free form, which is considered to be the biologically active form. Another aspect of the SHRP in rodents is the relative absence of CBG during the SHRP. Thus, although absolute values of corticosterone, which normally include both bound and unbound hormone, are very low in the absence of CBG, the fraction of corticosterone that is available in the free form for binding to corticosteroid receptors may actually be higher than is observed in the adult. Further, the clearance of corticosterone from the circulation is significantly slower in the pup and, as a consequence, corticosterone is available for a more prolonged period. Thus, the biologically active corticosterone does have a more prolonged time to exert its effects in the periphery and the brain.

Although there appear to be rate-limiting factors that act developmentally to limit the secretion of corticosterone in the neonate, evidence indicates that the adrenal is actively suppressed during the SHRP. It has been extensively documented that certain aspects of the rodent maternal behavior play an important role in regulating the neonate HPA axis. In particular, two specific components of the dam's caregiving activities seem to be critical; licking/stroking and feeding. Numerous studies have demonstrated that feeding is in part responsible for the suppression of the pup's capacity to both secrete and clear corticosterone from the circulation. Thus, removing the mother from the litter for 24 h (and thus fasting the pup) results in a significantly higher basal level and a further increase in the secretion of corticosterone following stress or administration of ACTH. The initial activation of the HPA axis occurs after between 4 and 8 h of maternal absence. This process is likely mediated by metabolic signals because replacement with glucose or blockade of ghrelin receptors during deprivation prevents HPA axis activation. We have postulated that one of the consequences of maternal

deprivation is to increase the sensitivity of the adrenal to ACTH. This has been demonstrated in several ways. (1) Following deprivation, significantly lower doses of ACTH are required to induce the adrenal to secrete corticosterone. (2) Although the levels of ACTH are equivalent between deprived and non-deprived pups under certain experimental conditions, the levels of corticosterone are greater in deprived pups. (3) Studies indicate that following a mild stress (injection of isotonic saline) there is an increase in *c-fos* gene expression in the adrenal cortex of the deprived neonate, whereas the nondeprived pup exhibited almost no detectable levels of *c-fos* mRNA. If maternally deprived pups are provided with food during the period of maternal deprivation, both basal and stress levels of corticosterone no longer differ from mother-reared pups.

Stress Hyporesponsive Period and the Pituitary

The concept of an absolute SHRP regarding the response of the pituitary following stress in the neonate is much more problematic. Whether the pituitary can show an increase in ACTH in response to stress is dependent on numerous factors. Among these are the age of the neonate, the type of stress imposed, and, once again, maternal factors. The early findings concerning the stress response of the pituitary suggested that there was a deficiency in the neonate's capacity to synthesize ACTH. Thus, as a result, the pup should exhibit a reduction in the magnitude of the ACTH stress response. However, sufficient data indicate that the pituitary of the neonate does have the capacity to synthesize and release ACTH that closely resembles the adult response. What seems to discriminate the neonate from the adult is that for the pup the response of the pituitary is much more stimulus-dependent. Further, the ability to terminate the stress response is also not fully developed and does not mature until quite late in development. Perhaps the earliest demonstration that the neonate can indeed mount an ACTH response to at least some types of challenges was a study that challenged neonates with an injection of endotoxin throughout the period from birth to weaning. At all ages, the neonate exhibited a significant elevation of ACTH that, beginning day 5, was equivalent to the adult. Of interest is that, although there was a robust ACTH response, the corticosterone response was reduced markedly from day 5 and did not begin to approach adult value until approximately day 15. The difficulty with this study is that very large doses (90% lethal dose, LD90) were used. It has been reported more recently that

the administration of interleukin (IL)-1 elicited an ACTH response in pups as early as postnatal day 6. The peak of the response followed a similar time course to that of the adult, although the magnitude of the response was significantly lower early in development. The reduced response in day 6 neonates cannot, however, be interpreted as a reduction in the neonate's capacity to produce ACTH. Three hours following adrenalectomy (ADX), a robust increase in ACTH occurs as early as day 5, presumably due to the absence of a corticosterone negative feedback signal. This magnitude of the ACTH response is as great as that seen in older neonates at day 18, which are well out of the SHRP. Under some circumstances, the pup can show an even greater ACTH response early in development than later. In contrast, brief periods of maternal separation, exposure to novelty, injections of isotonic saline, and restraint for 30 min all failed to elicit an ACTH response in normally reared pups until they escaped from the SHRP.

Recent studies in the mouse have revealed that the low and suppressed ACTH response during the SHRP is largely mediated by an enhanced glucocorticoid receptor (GR)-mediated negative feedback. Treatment with the GR antagonist RU486 resulted in a pronounced increase of basal ACTH and corticosterone values in 9-day-old pups to levels greater than in the adult (see Figure 1). As a different study observed, because only a small effect of RU486 is observed when it is applied locally at the paraventricular nucleus (PVN), a direct GR-mediated suppression of ACTH production and release at the pituitary level seems likely, at least toward the end of the SHRP. These data also indicate that the hypoactivity of the neonatal stress system may be restricted to the pituitary-adrenal axis.

However, so far there is little known about why and how neonates discriminate between different classes of stimuli, whereas older pups, which have escaped from the SHRP, and adults appear to respond in a similar manner regardless of the stress-inducing stimulus. Several hypotheses could account for this phenomenon. First, it could simply be a matter of stimulus intensity. Thus, the neonate may be less responsive to stimuli of lower intensities and therefore require a more intense stressor to activate the neuroendocrine cascade that eventually leads to a release of ACTH. Second, it has been well documented that different stimuli activate distinct neural pathways that lead to the release of CRF and thus ACTH. It is conceivable that these neural pathways mature differentially or are activated during development distinctly differently than in adulthood.

A third factor appears to contribute to the reduced capacity of the mother-reared pup to respond to milder stress-inducing procedures. The role of mother's caregiving activities on the developing adrenal was discussed earlier. Evidence shows that maternal factors can also actively inhibit the release of ACTH. Maternally deprived mouse pups show a marked increase in ACTH secretion between 4 and 8 h following removal of the mother. Pups deprived of maternal care for 24 h showed an increase in ACTH following an injection of saline as early as day 6. The effects of maternal deprivation were even more apparent at days 9 and 12. Although in subsequent experiments the response at day 6 was not reliable, significant increases in ACTH were replicated in 9- and 12-day-old neonates. These increases are also observed following 30 min of restraint. In contrast, nondeprived pups failed to show an acute release of ACTH following stress. Whereas feeding was required in order to reduce the sensitivity

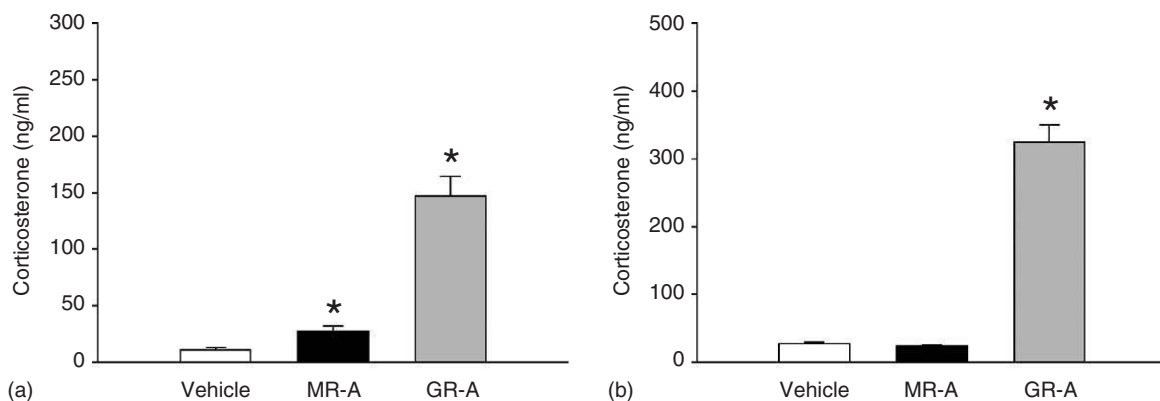


Figure 1 Nine-day-old mouse pups during the SHRP: a. plasma corticosterone; b. ACTH. Animals were injected intraperitoneally (IP) 16 and 8 h before testing with the vehicle (polyethylene glycol, $n = 10$), the MR antagonist (MR-A) RU28318 ($n = 15$), or the GR antagonist (GR-A) RU38486 ($n = 15$). Data represent mean $\pm$ SEM. * Significant from vehicle-treated animals, $p < 0.05$. Adapted from Schmidt et al., *Endocrinology* 146(3):1458–1464, with permission. Copyright 2005, The Endocrine Society.

of the adrenal, anogenital stroking can reverse the increased sensitivity of the HPA axis following deprivation. Thus, different components of the mother's behavior appear to be involved in regulating different components of the endocrine stress response. These data suggest that the dam's behavior was actively inhibiting part of the neuroendocrine cascade that ultimately results in the peripheral endocrine responses to stress. Thus the capacity to respond even to mild stimuli is present early in development but is observable only if the maternal inhibitory factors are not present. It is important to note, however, that although the pup can be induced to show an endocrine response to stress during the SHRP, maternal inhibition is not the only rate-limiting factor. If we examine the body of data on the ACTH responses in pups, it emerges that, even when the infant responds to mild stress during the SHRP, the magnitude of the response is always considerably lower in the SHRP than that of the older pups (day 18) and adults.

Stress Hyporesponsive Period and the Brain

The brain is clearly a stress-sensitive organ. It is not the purpose of this article to review all the changes in the brain that have been shown to occur in response to stress in the neonate. For this review, the focus is on the specific neural markers that have a direct effect on the regulation of the pituitary. Several of the early studies reported that during postnatal development there is a deficiency in the neonate's capacity to increase gene transcription and peptide expression of ACTH secretagogues, even though there is a striking increase in ACTH release. However, recent data indicate that this is not completely correct. A mild stressor in the form of a saline injection results in a fast and significant increase of CRF mRNA expression in the PVN in pups during the SHRP within 30 min rather than the 4 h used to demonstrate CRF mRNA in adult animals. There is also clear evidence that CRF is involved in the stress responses during the SHRP. The administration of a CRF antagonist blocked the ACTH response in neonates following exposure to an endotoxin. In addition, recent studies with CRF receptor 1-deficient mice demonstrated that a functioning CRF system at the pituitary level is essential for deprivation-induced changes in corticosterone secretion.

Corticosteroid Feedback

Negative feedback has two modes of operation. First, the proactive mode involves the maintenance

of basal levels of HPA activity and is, in the adult, predominantly mediated by the high-affinity mineralocorticoid receptors (MRs) for corticosterone in higher brain regions. Second, the reactive mode facilitates the termination of stress-induced HPA activity, which also involves the lower-affinity GRs localized in higher brain areas as well as the PVN and the pituitary corticotrophs. Both feedback modes are operative during the SHRP. Early studies reported a pronounced increase of circulating ACTH in response to ADX in neonates, a process that was attributed to the MR. During the SHRP, the MR are highly expressed in the hippocampus, have adultlike properties and, as observed in the adult, have low amounts of corticosterone mainly occupy MRs. However, in a recent study, the treatment of neonatal mice with the MR antagonist RU28318 resulted in only a small increase in circulating corticosterone, indicating a modulatory role of the MRs. In contrast, the GR antagonist RU486 caused a profound increase of ACTH and corticosterone to adultlike levels, indicating that GRs play a predominant role in suppressing the HPA axis activity in the neonate. Because a local blockade of the GRs in the PVN of a neonatal rat resulted only in minor changes of corticosterone secretion, the most likely site of GR-mediated negative feedback in the neonate under basal conditions is in the pituitary.

The reactive feedback, probably mediated by GRs in the brain, seems to be a late-developing process because during neonatal life the stress-induced ACTH levels are not terminated as efficiently as in adulthood. Low levels of GRs are present in the brain around birth; the levels rise slowly and do not achieve adult levels before 4 weeks of age. The affinity of GRs is higher perinatally than at later ages. The microdistribution of GRs changes dramatically in some regions during the first week of life. In some regions of the hypothalamus, GR expression is very dense, but it disappears close to detection levels in later life. Other regions, such as the dentate gyrus of the hippocampus, do not express GRs until day 12.

Twenty-four hours of maternal deprivation impairs the reactive (rather than the proactive) control of HPA activity in the neonates. This was demonstrated in a study design using ADX animals and corticosterone substitution either at the time of ADX or 3 h after ADX, when the removal of the adrenals has triggered a profound ACTH response. The latter condition tests reactive feedback, which appeared during maternal deprivation to be resistant to corticosterone. Corticosterone replacement at the time of ADX, which tests the efficacy of proactive feedback, was not affected by maternal deprivation. The resistance to corticosterone could be due to two interacting

processes. The first represents the hyperdrive of signals from the stress circuitry innervating the PVN neurons triggered by maternal deprivation. The second process may be related to an effect of maternal deprivation on the development of corticosteroid receptors. Indeed, maternal deprivation downregulates hippocampal MR expression as well as GR expression in discrete regions of the hippocampus and the PVN. These changes in GR could be restored if the deprived neonate received aspects of maternal care, such as stroking and feeding. Thus, the altered characteristics clearly may be involved in the impaired feedback following maternal deprivation.

Conclusion

Is there an SHRP? The concept of the SHRP in its most traditional form describes a period in development when most aspects of the HPA axis appear to be downregulated, resulting in an organism that either fails to respond or at best shows a response that is reduced markedly compared to older pups and adults. This may still be applicable to the response of the adrenal if we take total corticosterone as the index of the stress response. Therefore, the SHRP in fact represents an adrenal hyporesponsive period, primarily caused by a potent corticosterone feedback signal that ensures quiescence in the face of common daily stressors. The existing information on the ontogeny of the HPA axis does not support, therefore, the concept of an SHRP as it was formulated in the early 1970s.

Acknowledgments

This work was supported by a grant from the National Institutes of Mental Health, MH-45006 (to S.L.), a NATO collaborative research grant (CRG 97.0477), the Netherlands Organization for Scientific Research (NWO), 554–545 and the Royal Netherlands Academy of Arts and Sciences.

See Also the Following Articles

Corticosteroid-Binding Globulin (Transcortin); Corticotropin Releasing Factor (CRF); Glucocorticoid Negative Feedback.

Further Reading

- Bohn, M. C. (1984). Glucocorticoid-induced teratologies of the nervous system. In: Yanai, J. (ed.) *Neurobehavioral teratology*, pp. 365–387. New York: Elsevier.
- de Kloet, E. R., Sibug, R. M., Helmerhorst, F. M., et al. (2005). Stress, genes and the mechanism of programming the brain for later life. *Neuroscience and Biobehavioral Reviews* 29(2), 271–281.
- de Kloet, E. R., Rosenfeld, P., Van Eekelen, J. A. M., Sutanto, W. and Levine, S. (1988). Stress, glucocorticoids and development. *Progress in Brain Research* 73, 101–120.
- Levine, S. (1994). The ontogeny of the hypothalamic-pituitary-adrenal axis: the influence of maternal factors. *New York Academy of Sciences* 740, 275–288.
- Levine, S., Glick, D. and Nakane, P. K. (1976). Adrenal and plasma corticosterone and vitamin A in rat adrenal glands during postnatal development. *Endocrinology* 118, 1177–1179.
- Meaney, M. J., Diorio, J., Francis, D., et al. (1994). Environmental regulation of the development of glucocorticoid receptor systems in the rat forebrain: the role of serotonin. *New York Academy of Sciences* 740, 260–274.
- Rosenfeld, P., Suchecki, D. and Levine, S. (1992). Multifactorial regulation of the hypothalamic-pituitary-adrenal axis during development. *Neuroscience and Biobehavioral Reviews* 16, 553–568.
- Rosenfeld, P., Van Eekelen, J. A. M., Levine, S., et al. (1993). Ontogeny of corticosteroid receptors in the brain. *Cellular and Molecular Neurobiology* 13, 295–321.
- Sapolsky, R. M. and Meaney, M. (1986). Maturation of the adrenocortical stress response: neuroendocrine control mechanism and the stress hyporesponsive period. *Brain Research Reviews* 11, 65–76.
- Schmidt, M., Levine, S., Oitzl, M. S., et al. (2005). Glucocorticoid receptor blockade disinhibits pituitary-adrenal activity during the stress hyporesponsive period of the mouse. *Endocrinology* 146(3), 1458–1464.

Stress in University Students

B Andrews and J Hejdenberg

Royal Holloway University of London, Egham, UK

© 2007 Elsevier Inc. All rights reserved.

Conceptual Issues in Student Stress Research
The Impact of Commonly Identified Student Stressors
Implications for Intervention

Glossary

<i>Cross-sectional research</i>	Investigations in which all factors are assessed at one time.
<i>Homesickness</i>	Distress from longing for home during one's absence from it.
<i>Life-stage transition</i>	Moving from one life stage to another (e.g., from childhood to young adulthood) with attendant changes (e.g., from being materially dependent on others to becoming more independent).
<i>Prospective/longitudinal research</i>	Investigations in which factors are assessed at more than one time; usually carried out to identify factors that might predict future outcomes.

Pressures facing university students have been a cause of concern because of their impact both on mental health and academic performance. In recent years, student mental health services in the United Kingdom and United States have reported increases in the severity of the presenting problems reported by students. There is also evidence from studies in the United Kingdom that university students report more mental health symptoms than age-matched controls, although there is a lack of information on diagnosable mental disorders in students internationally. In some countries, apparent increases in student distress have been linked to decreases in state funding for students and widening student participation; widening access to higher education has resulted in more diverse student populations, thereby increasing the probability of students in the system with elevated levels of life problems. Given the paucity of information on specific sources of stress and student physical health, the focus in this article is on the stressors associated with the university experience that relate to poor mental health and academic performance.

Conceptual Issues in Student Stress Research

In research on university students, the term stress has commonly been used to describe subjective responses

to an experience or situation. Homesickness and worries associated with meeting course requirements are good examples, although non academic worries are often included. The term has also been used to refer to life events and difficulties such as bereavement or financial hardship that may be considered universally threatening without the proviso in assessment that the event caused stress, upset, or worry. In the main, the literature has emphasized students' subjective worries rather than reports of actual life difficulties. Some authors have argued for the advantages of assessing stress as the amount of subjective worry associated with individual stressors. The individuality in the experience of stress, such as the availability of coping skills and/or personal vulnerability, is recognized by using such a method. The approach is useful for descriptive purposes, but research into the effects of stress on student mental health and achievement demands a more objective assessment of life problems to avoid confounding cause and effect.

The Impact of Commonly Identified Student Stressors

Before reviewing the literature, it should be noted that much of the relevant research has been carried out with narrow samples of students in vocational studies or other specialized subjects. There is therefore the problem of the generalizability of vocational samples, such as medical and nursing students, to the whole student population. The general use of cross-sectional designs, with only one data collection point, is also problematic because causal direction cannot be inferred. The relevant studies described here have therefore been selected because they represent broader populations of university students, and, whenever possible, prospective studies are described.

The sources of stress most commonly identified by students in surveys worldwide include the academic, financial, and relationship domains, although it should be noted that student reports are constrained by the areas initially defined by the investigators. Researchers have also identified the transition to university and paid term-time employment as additional sources of stress. To evaluate the existing evidence of the effects of stress on student mental health and achievement, four of these stressors (transitions, academic and financial pressures, and term-time employment) are discussed here. These have been selected because they are specifically related to the student experience and have been the main focus of investigations into the impact of stress on students.

Transition to University

Becoming a student involves major life changes for young people. Faced with new responsibilities and social roles inherent to the university experience, this transition has the potential to cause distress. The literature has focused on the negative aspects of this transition, but it can also provide the opportunity to learn new competencies and to develop new and rewarding social relationships. For example, longitudinal studies of life-stage transitions have demonstrated their positive effects on self-esteem in young adulthood. However, few longitudinal student studies have been undertaken, and studies that have followed students through the transition to university with a survey before entry are uncommon. Studies that have surveyed students before university entry have shown (usually) short-term increases in mental health symptoms after university entry, which on the face of it appears to be indicative of the stress associated with this experience. In one important study of homesickness, Fisher and Hood measured psychological disturbance in students 3 months before university entry and again 6 weeks into their first term. The levels of psychological disturbance, which included depressive and obsessional symptoms and absent-mindedness, increased in the period after the transition to university. The increases were most apparent in students reporting homesickness, although these students had higher symptom levels to start with. Using an objective measure of residence change, there was no difference in psychological disturbance between students still living with parents and those living on campus. This research demonstrated that, at least in the short term, entering university has consequences in terms of emotional disturbance and cognitive impairment regardless of residence status. This suggests that factors other than the transition of moving from home contribute to student distress.

A subsequent longitudinal study by Andrews and Wilding assessed anxiety and depressive symptoms in students and showed a significant increase in symptoms from 1 month before college entry to midcourse. However, when they looked categorically at the rates of high depression and anxiety in the students, a rather more complex picture emerged. Of the students who were relatively symptom-free before entry, 29% had developed high levels of depression or anxiety symptoms (subsequently explained by factors such as financial hardship and relationship problems). In contrast, approximately one-third of those with high levels of depressive or anxious symptoms before entry were relatively symptom-free by midcourse. This indicates that transitions at this life stage should not be viewed as exclusively negative. There is also a need to consider the more positive aspects of the university

experience and the ways in which they may act to increase student well-being.

Academic Pressures

The university years are a time of continuous assessments for students in most countries, and as such they can put a considerable amount of pressure to achieve on even the hardest of students. As already discussed, one of the problems in assessing the impact of the stress associated with academic work is disentangling cause from effect. Although academic pressures may have a negative effect on emotional well-being, anxiety and depression can affect academic work through symptoms such as high arousal and lack of motivation and concentration. In the absence of in-depth enquiries in this area, it is usually not possible to ascertain whether academic stress is an antecedent or consequence of mental health problems. A more useful line of enquiry into the impact of academic study *per se* is to consider specific academic events such as taking exams. This experience is recognized by leading life events researchers as a universally threatening experience that can increase the likelihood of physical and mental disorders. The effect of studying for and taking examinations has been demonstrated in a recent longitudinal study of undergraduates carried out by Surtees and colleagues at the University of Cambridge in England. Levels of anxiety and depressive symptoms peaked during examination periods, although this suggests that the negative impact is likely to be short-lived for the majority who pass successfully.

Financial Problems

Most young people have not managed their own finances before starting university, and this can often come as a rude shock. Financial hardship is a common problem facing students, and one that is expected to worsen in countries where state funding is being reduced in favor of students paying their own fees and living costs. It is therefore of particular concern that relevant studies have consistently indicated the adverse effects of financial difficulties on student mental health. There is also emerging evidence of the negative impact financial problems and their attendant consequences can have on academic achievement. Much of this research has been carried out in the United Kingdom, where there have been rapid changes to student funding in the past few years.

To address the problems of causality evident in cross-sectional research, Andrews and Wilding recently carried out a longitudinal study of over 350 undergraduate students in different university faculties to consider the impact of different types of stressful life experiences, including financial difficulties. The main

experiences assessed included major financial difficulties, serious personal illness or injury, serious illness and death of close others and serious relationship problems and separations. Of the various adverse experiences, only financial difficulties (experienced by 21%) and personal illness or injury (experienced by 6%) contributed significantly to increased levels of depression midcourse, after adjusting for depressive and anxiety symptoms present before university entry. This initial evidence suggests that the direction of causality is from financial problems to depression. The odds of becoming depressed were three times higher for students with financial problems. This study also considered the impact of stressful life experiences on academic attainment and found that the only adverse experience that predicted poor exam performance was financial difficulties. Depression appeared to mediate this relationship when prior academic attainment was controlled.

Paid Term-Time Employment

An additional stressor related to financial hardship is that of paid work. Large-scale studies have shown that students reporting difficulties in keeping up with bills or having financial problems are significantly more likely to undertake term-time employment. Being obliged to work in addition to full-time study is a further pressure on students. So far, research in this area has focused on the effects of term-time work on academic performance rather than on mental health. Large-scale surveys of university students on both sides of the Atlantic have found that academic performance decreases as the hours of paid work increase. There is evidence from student research in the United States by Pascarella and colleagues that paid term-time work can impede the basic cognitive processes necessary for academic achievement such as reading comprehension and critical thinking. Subsequent studies have found that students who work longer hours during term-time achieve lower end-of-year exam averages than other students. In a recent study, this effect still held after relevant demographic factors and prior academic performance were taken into account.

Implications for Intervention

Student counseling services are a common feature of university life, along with other educational support services and pastoral care from personal tutors and

advisors. However, to our knowledge there is little available information on the systematic evaluation of the effectiveness of such services. Some real-life difficulties that cause students distress may not be amenable to the kinds of services that counselors provide, particularly when they involve financial crises and the pressures of juggling the need to earn money with academic demands. Money management courses would appear to be relevant, but more needs to be known about their availability and effectiveness. Counseling services may be more relevant in dealing with stressors of a more intrapsychic nature. Most services deal with issues of attachment and homesickness, and many provide support in managing the stress experienced by students around exam times. Given the available evidence of the adverse effects of these common problems on student mental health and academic attainment, there is an obvious need for well-controlled trials to evaluate how well university support services deal with different types of student stressors.

See Also the Following Articles

Economic Factors and Stress; Psychosocial Factors and Stress.

Further Reading

- Andrews, B. and Wilding, J. M. (2004). The relation of depression and anxiety to life-stress and achievement in students. *British Journal of Psychology* 95, 509–521.
- Fisher, S. (1994). *Stress in academic life: the mental assembly line*. Buckingham, UK: Open University Press.
- Grant, A. (2002). Identifying students' concerns taking a whole institutional approach. In: Stanley, N. & Manthorpe, J. (eds.) *Students' mental health needs: problems and responses*. London: Jessica Kingsley.
- Pascarella, E. T., Edison, M. I., Amaury, N., et al. (1998). Does work inhibit cognitive development during college? *Educational Evaluation and Policy Analysis* 20, 75–93.
- Royal College of Psychiatrists (2003). *The mental health of students in higher education*. Royal College of Psychiatrists Council report CR112, London: RCP.
- Stewart-Brown, S., Evans, J., Patterson, J., et al. (2000). The health of students in institutes of higher education: an important and neglected public health problem? *Journal of Public Health Medicine* 22, 492–499.
- Surtees, P. G., Wainwright, N. W. J. and Pharoah, P. D. P. (2002). Psychosocial factors and sex differences in high academic attainment at Cambridge University. *Oxford Review of Education* 28, 21–38.

Stress Induced Anovulation

S L Berga and T L Loucks

Emory University School of Medicine, Atlanta, GA, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction

Neuroendocrine Mechanisms Linking Cognition, Mood,
Behavior, and Gonadotropin-Releasing Hormone Drive
Pathogenesis of Stress-Induced Anovulation
Behavioral, Nutritional, and Metabolic Influences on the
Reproductive Axis
Synergism among Stressors
Treatment Considerations
Summary

Glossary

<i>Allostasis</i>	A sustained adjustment that promotes survival in response to chronic challenge; may increase long-term health burden.
<i>Anovulation</i>	The cessation of ovulation.
<i>Cognitive-behavior therapy</i>	A form of psychoeducation or talk therapy that addresses nonadaptive thoughts, cognitions, attitudes, and conceptualizations.
<i>Eumenorrhea</i>	A normal pattern of menstrual bleeding; can occur in the absence of ovulation.
<i>Gonadotropin-releasing hormone (GnRH)</i>	Decapeptide produced and released in a pulsatile manner from hypothalamic neurons; stimulates the release of pituitary luteinizing hormone (LH) and follicle stimulating hormone (FSH).
<i>Hypothalamic hypogonadism</i>	A condition characterized by suppression of GnRH drive and, in turn, pituitary LH and FSH, resulting in reduced or absent gonadal steroidogenesis and gametogenesis.
<i>Hypothalamic hypothyroidism</i>	An allostatic adjustment in thyroidal function that conserves energy expenditure and is characterized by relatively normal levels of pituitary thyroid stimulating hormone (TSH) in the presence of suppressed levels of thyroid hormones triiodothyronine (T3) and thyroxine (T4).
<i>Limbic-hypothalamic-pituitary-adrenal axis</i>	Neuroendocrine circuit that mediates stress perception and stress responses.
<i>Metabolism</i>	Biochemical processes that regulate energy expenditure and storage.
<i>Secondary amenorrhea</i>	Cessation of menses in a woman of reproductive age who was previously eumenorrheic.
<i>Stress</i>	Physical and psychological stimuli that challenge the status quo and elicit homeostatic or allostatic behavioral responses.

Introduction

A wealth of scientific and clinical evidence supports the notion that stress causes reproductive compromise and increases health burden. Neuroendocrine, metabolic, and behavioral responses to acute stress represent transient homeostatic adaptations that promote survival in the face of perceived challenge; chronic stress elicits allostatic (sustained) adjustments that also promote survival but at greater health cost (Figure 1). Reproductive compromise impedes population replenishment and increases acute and chronic health burden in women, men, and children due to metabolic dysfunction, obesity, diseases of aging, preterm delivery, birth defects, costly and risky infertility therapies, and, through epigenetic mechanisms, it imprints the next generation. Stress is the most common and most commonly underappreciated cause of reproductive dysfunction (Table 1). Stress-induced anovulation (SIA), often termed functional hypothalamic amenorrhea (FHA) or functional hypothalamic chronic anovulation, causes infertility and increases acute and chronic health burden.

Behaviors that chronically activate the limbic-hypothalamic-pituitary-adrenal (LHPA) axis and/or chronically suppress the hypothalamic-pituitary-thyroidal (HPT) axis compromise the hypothalamic-pituitary-gonadal (HPG) axis in both women and men by reducing hypothalamic gonadotropin-releasing hormone (GnRH) drive. Individuals with functional (not due to defined organic conditions) forms of hypothalamic hypogonadism typically engage in a combination of behaviors in response to psychogenic stress that concomitantly induce intermittent or chronic energy imbalance. Further, chronic adrenal activation may increase the energetic cost of common activities such as running. In women, functional hypothalamic hypogonadism exists on a spectrum that includes polymenorrhea (menstrual interval < 24 days), eumenorrhea with reduced luteal progesterone secretion, and anovulatory eumenorrhea, oligomenorrhea (menstrual interval 735 days), or amenorrhea (absence of menstruation for >3–6 months). In men, oligoasthenospermia (very low sperm count with low motility and abnormal morphology) may result, but this is typically not clinically evident unless fertility is being sought or severe testosterone deficiency results in muscle wasting or other phenotypic alterations.

This article focuses primarily on SIA/FHA in women. SIA/FHA typically results from psychogenic stress coupled with a mild energy imbalance and represents an allostatic adaptation, that is, a stable

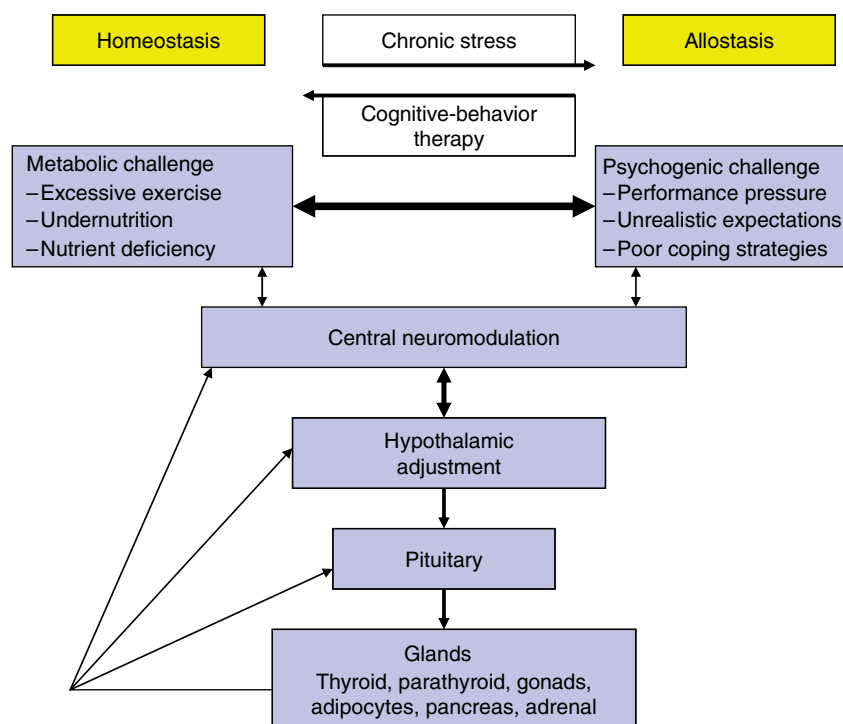


Figure 1 Conceptualization of synergism between metabolic and psychogenic stress in the pathogenesis of stress-induced anovulation.

Table 1 Common causes of anovulation/amenorrhea^a

	Percentage
Stress-induced anovulation/ functional hypothalamic amenorrhea	34
Hyperandrogenism/PCOS	29
Hyperprolactinemia	13
Premature menopause	12
Asherman's syndrome	5
Other	7

^aBased on data presented in Reindollar, R. H., Novak, M., Tho, S. P. T., et al. (1986), Adult-onset amenorrhea: a study of 262 patients, *American Journal of Obstetrics and Gynecology* **155**, 531–543. PCOS, polycystic ovarian syndrome.

change in behaviors and secretory patterns that promotes acute survival but at some health cost. SIA/FHA affects roughly 5% of women of reproductive age; less severe forms of hypothalamic hypogonadism are more common and are less clinically evident (Figure 2).

Stress-induced anovulation is theoretically reversible, but reproductive recovery appears to depend on the restoration of eucortisolemia and at least partial recovery from functional hypothalamic hypothyroidism. Hormone replacement strategies have limited benefit for women with SIA/FHA because they do

not promote recovery from allostatic endocrine adjustments in the adrenal and thyroidal axes. Indeed, the rationale for the use of sex steroid replacement is based on the erroneous assumption that functional forms of hypothalamic hypogonadism represent only or primarily a loss of sex steroid exposure due to reduced GnRH drive. Further, the replacement of sex hormones masks deficits that accrue from chronically altered adrenal and thyroidal functions. Long-term deleterious consequences of SIA/FHA probably include an increased risk of cardiovascular disease, osteoporosis, depression, other psychiatric conditions, dementia, and neurodevelopmental compromise in offspring. Although fertility can be restored with exogenous administration of gonadotropins or pulsatile GnRH, fertility management alone does not engender adrenal and thyroidal recovery. Pregnancy in the face of ongoing psychogenic stress and metabolic imbalance may increase the likelihood of poor obstetrical, fetal, or neonatal outcomes. In contrast, behavioral and psychological interventions that address problematic behaviors and attitudes have the potential to facilitate reproductive recovery along with adrenal and thyroidal recovery. In short, full endocrine recovery offers better individual, maternal, and child health.

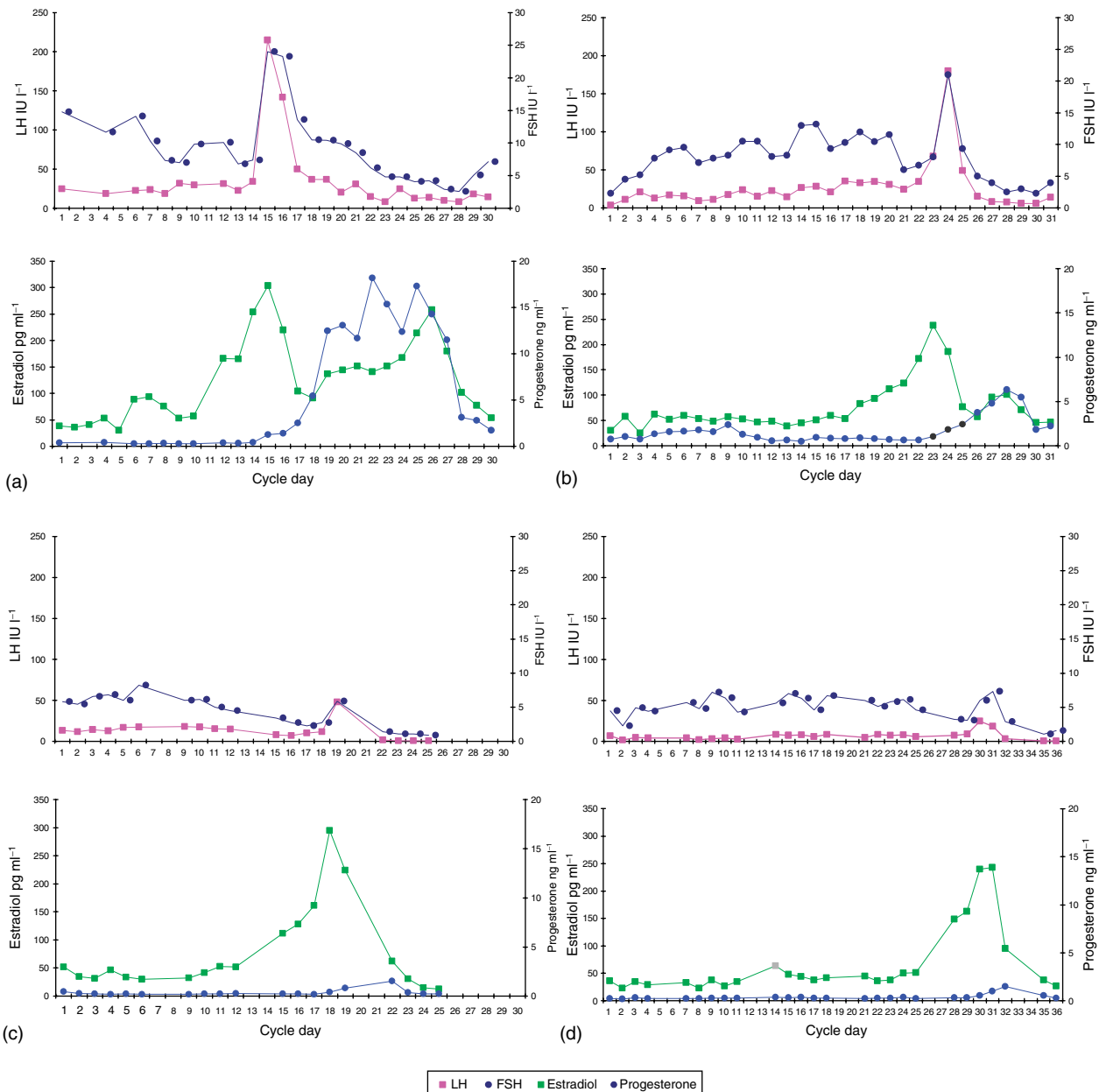


Figure 2 Continuum of ovarian function in women with preserved menstrual cyclicity. a, Ovulatory eumenorrhea; b, luteal insufficiency; c–d, eumenorrheic anovulation. Pituitary gonadotropin (LH and FSH) and ovarian sex steroid (estradiol and progesterone) levels were obtained daily for one cycle. FSH, follicle-stimulating hormone; LH, luteinizing hormone. Note that $\text{FSH} > \text{LH}$, but $\text{FSH} < 20 \text{ IU/L}$, thus the cause of hypogonadism is hypothalamic rather than reduced ovarian reserve.

Neuroendocrine Mechanisms Linking Cognition, Mood, Behavior, and Gonadotropin-Releasing Hormone Drive

Our conceptualization of the bidirectional interaction between behavior and gonadal function has been refined by scientific insights into the mechanisms mediating neuroendocrine responses to thoughts, feelings, and behaviors.

The proximate cause of hypothalamic forms of hypogonadism, including SIA/FHA, is reduced hypothalamic GnRH input. Gonadal function depends directly on secretion from the hypothalamus of GnRH as pulses. Declines in pulsatile GnRH secretion reduce pituitary secretion of luteinizing hormone (LH) and follicle-stimulating hormone (FSH) and wholly or partially compromise folliculogenesis.

Decreased GnRH drive is a common cause of anovulation and amenorrhea. Decrements in central GnRH–LH/FSH drive exist on a continuum, however, and may vary from day to day. Because of this potential variability in GnRH drive, ovarian compromise exists on a spectrum and may present as polymenorrhea, eumenorrhea, oligomenorrhea, or amenorrhea. Clinically, the decreased ovarian function can be occult or obvious. In men, decreased central GnRH drive may cause oligoasthenozoospermia. Typically, gonadal compromise in men is clinically occult unless fertility is sought and the compromise is sufficiently significant to cause infertility. However, severe hypothalamic hypogonadism in men may present as decreased libido, diminished muscle mass, or altered hair (beard) growth.

The most common cause of reduced GnRH drive is functional; that is, it is not due to identifiable organic causes such as hypothalamic tumors or pituitary adenomas. Functional hypothalamic hypogonadism can be defined as a common and theoretically reversible form of gonadal compromise in which psychophysiological responses to life events activate central neural networks and thereby alter glandular secretion and metabolism.

The mechanisms mediating the disruption of central GnRH drive are poorly understood and complex. GnRH neurons are diffusely distributed in the medial basal hypothalamus. Most, but not all, GnRH axons project to the median eminence, allowing pulses of GnRH to be released into the portal vasculature. GnRH neurons communicate with one another via synapses. Although GnRH neurons are endogenously pulsatile, their activity must be synchronized by GnRH-to-GnRH synapses for the GnRH bolus released into the portal vasculature to be of sufficient magnitude to trigger pituitary release of LH and FSH. GnRH neurons receive synapses from neurons that contain GnRH, the endogenous opioid peptide β -endorphin, neuropeptide Y (NPY), and catecholamines. Other factors that modulate the frequency or activity of the GnRH pulse generator are thought to exert their effects indirectly by acting through the neuronal systems that have direct synaptic connections or by interacting with glial cells that interpose between synapses. For instance, progesterone appears to slow the frequency of pulsatile GnRH release by increasing hypothalamic opioidergic tone, but it also may cause glial cells to envelope the GnRH neurons and to reduce thereby the GnRH-to-GnRH connections. Peripheral substances may gain access to the GnRH neuronal network via specialized neurovascular cells that line the fenestrated blood–brain barrier at the level of the hypothalamus and median eminence, so the modulation of GnRH secretion is not

Table 2 Putative modulators of GnRH drive^a

CRH	Metabolic signals
Opioids	Sex steroids
Adrenergic	Gonadal peptides
GABA	Growth factors
Dopamine	Glial cells
Serotonin	GnRH
Immune	Other

^aCRH, corticotropin releasing hormone; GABA, γ -aminobutyric acid; GnRH, gonadotropin-releasing hormone.

confined to central factors. The brain–gut axis appears to communicate the metabolic status of an individual to the hypothalamic GnRH neurons through multiple mechanisms. A partial list of putative neuromodulators of GnRH drive is shown in [Table 2](#).

Progress has been made in quantifying endocrine patterns for research purposes. However, in humans, GnRH pulsatile secretion can be inferred only from the pattern of LH secretion in the circulation. Blood samples must be obtained via an indwelling intravenous catheter at intervals of 10–15 min for durations of 12–24 h. Even so, inherent limitations exist in estimating GnRH secretion from peripheral LH patterns. The tedious nature of quantifying central GnRH drive explains why this is not done routinely during clinical evaluation. Technical limitations also plague the recognition and quantification of stress and metabolic status. The accuracy of psychometric inventories for assessing and quantifying stress, mood, and cognitive patterns is inherently constrained by reporting biases, whereas endocrine or biophysical indices of stress and metabolic status are technically cumbersome and expensive to collect. Neuroimaging techniques have afforded a new window into the neurochemistry and neuroanatomy of behaviors and thoughts, but these techniques have not yet been used to understand the pathogenesis of stress-induced reproductive compromise.

Although innumerable animal studies have demonstrated that activation of the hypothalamic–pituitary–adrenal (HPA) axis by a variety of stressful paradigms induces reproductive compromise, only a few studies have elucidated the mechanisms mediating the disruption of GnRH drive. Direct evidence exists for corticotropin releasing hormone (CRH), β -endorphin, dopamine, and arginine vasopressin. Recently, NPY has been implicated as serving a neuromodulatory link between metabolic deficits induced by diet and exercise and reduced GnRH drive. A key inhibitory neurotransmitter system in the brain uses γ -aminobutyric acid (GABA). GABA opens the same potassium channels in neurons of the mediobasal

hypothalamus as μ -opioid receptor agonists such as β -endorphin. GABA inhibited GnRH gene expression in rats and suppressed pubertal GnRH increase in juvenile female rhesus monkeys, but the role of GABA in human hypothalamic hypogonadism remains unclear. Serotonin neurons may also play a role in modulating GnRH drive.

Pathogenesis of Stress-Induced Anovulation

The best biochemical evidence supporting the concept that stress impairs ovarian function in women is the consistent demonstration that women with hypogonadotropic hypogonadism not due to defined organic conditions have higher cortisol levels than eumenorrheic, ovulatory women. There is direct evidence that this relationship holds in women with athletic amenorrhea as well. In a study by Loucks et al., an inverse relationship was observed between the degree of ovarian compromise and cortisol levels. When compared with eumenorrheic but sedentary women, eumenorrheic athletes had less luteal progesterone secretion, as evidenced by lower urinary levels of pregnanediol-glucuronide, fewer LH pulses in a day, and higher cortisol levels. Furthermore, amenorrheic athletes that were anovulatory had the fewest LH pulses in a day and the highest cortisol levels despite comparable levels of exertion and fitness. An inverse relationship holds between adrenal activation (independent of the life events or behaviors that initiate or sustain this activation) and suppression of the hypothalamic GnRH drive to the ovary, as evidenced by marked reduction in LH pulse frequency.

Other hypothalamic outputs also are altered in women with SIA/FHA. Given the neuroanatomical integration of the hypothalamus, this is predictable. The purpose of the hypothalamus is to generate an endocrine action plan to preserve the organism in the face of challenge. Part of the action plan involves metabolic mobilization. However, metabolic mobilization involves more than an increase in cortisol secretion. In SIA/FHA, the thyroid axis differs from that of eumenorrheic women in that thyrotropin (TSH) is not increased in response to decrements in thyronine and thyroxine, which indicates an altered hypothalamic set point akin to what is seen in hospitalized patients who develop what is referred to as sick euthyroid syndrome. In athletic women, a similar alteration in the thyroid axis was seen only in those who had compromised ovarian function. The secretory patterns of growth hormone, prolactin, and melatonin also differed from those of eumenorrheic women. The constellation of neuroendocrine aberrations that accompany SIA/FHA strongly suggests that central

neurotransmission has been chronically altered to forge an allostatic state. The aim of allostatic adjustments is to allow the individual to cope with chronic as opposed to acute challenge. In this context, then, the hypothalamus links the external environment, the internal milieu, and gonadal function.

The process of recovering from the allostatic state of SIA/FHA is also interesting but less well documented. We showed that women in the process of recovering from functional hypothalamic hypogonadism displayed cortisol levels identical to those of eumenorrheic women before there was complete recovery of GnRH drive. Further, a marked increase in thyrotropin-stimulating hormone (TSH) drive occurred as a prelude to increases in thyroxine and thyronine. We recently reported that women with SIA/FHA treated with cognitive-behavior therapy (CBT) displayed a return of ovarian function and ovulation and reduced cortisol levels but had only partial recovery from hypothalamic hypothyroidism and no weight gain. These data, coupled with our psychometric data showing that women with SIA/FHA differ from eumenorrheic women (EW) in their attitudes toward eating and desire for thinness, raise the possibility that intermittent or mild energy imbalance may persist after CBT-induced reproductive recovery. In short, hypothalamic recovery involves a set of interlinked readjustments, including LHPA restoration, return of GnRH pulsatility, and at least partial remission of hypothalamic hypothyroidism, but the duration and sequence of recovery is just now beginning to be described and the role of metabolic factors as mediators of recovery remains unclear.

The characteristic hypothalamic alterations associated with SIA/FHA only become problematic when ongoing challenges elicit a chronic rather than acute response. The long-term consequences of persistent HPA activation have been studied in animal models and hippocampal neuron loss has been documented. Stress appears to increase the risk of dementia and other neurodegenerative disorders. Other long-term sequelae of persistent metabolic mobilization are largely unknown, but there is no reason to assume that such a process is benign. For instance, recent data obtained in women with weight-restored anorexia nervosa who remained amenorrheic indicated that exogenous sex steroid replacement was unable to stimulate appropriate bone accretion. The investigators speculated that the ineffectiveness of hormone exposure was due to ongoing metabolic derangements such as increased cortisol exposure, altered growth hormone action, or hypothalamic hypothyroidism. Because of the concomitant endocrine and metabolic disturbances, hypothalamic hypogonadism must be regarded as a condition deserving

clinical attention even when fertility is not an immediate goal.

The central neuromodulators responsible for the initiation and maintenance of the disruption of GnRH are difficult to identify in humans. First, the factors that initiate allostasis may differ from those that maintain allostasis. To study initiating factors, we would need to intensely monitor populations at risk for the development of hypothalamic hypogonadism or try to induce hypothalamic hypogonadism in a nonhuman primate model. Once a chronic hypogonadal state had been reached, it theoretically would be possible to identify the agents that maintain this disruption by administering antagonists that cross the brain–blood barrier, by performing lumbar punctures and obtaining cerebrospinal fluid (CSF), or by performing neuroimaging studies with an appropriate ligand. To date, efforts to identify these neuromodulators in humans have yielded inconsistent results. Thus, naloxone, an opioidergic blocker, increased LH pulse frequency or levels in some, but not all, women with SIA/FHA. Also, the infusion of metoclopramide, a dopamine receptor blocker, to women with FHA accelerated LH pulse frequency, whereas that of eumenorrheic women remained constant. These data suggest that there may be dopaminergic as well as opioidergic inhibition of GnRH drive. To explore the hypothesis that the reduction in GnRH drive was maintained by CRH, vasopressin, β -endorphin, or a combination of these factors, we performed lumbar punctures to obtain rostral CSF in women with SIA/FHA and those with eumenorrhea. This approach revealed increased CRH in the CSF of women with depression, but we found that CRH levels were identical in women with SIA/FHA and eumenorrhea. Vasopressin levels were similar, but, surprisingly, β -endorphin levels were lower in women with SIA/FHA. Further, we evaluated free cortisol in the CSF of these same women and found that these levels were approximately 30% higher in those women with stress-induced anovulation. In aggregate, these findings (Figure 3) support the notion that the preservation of CSF CRH levels in the presence of elevated CSF cortisol reflects chronic stress-induced (allostatic) resistance to negative feedback suppression by cortisol.

Other putative central neurotransmitters that may contribute to the initiation and maintenance of SIA/FHA are the serotonergic and GABAergic systems. Several lines of evidence suggest that serotonergic function gates stress reactivity and metabolism. First, we showed in our monkey model of SIA/FHA that stress-sensitive monkeys had larger cortisol elevations and reduced prolactin responses to fenfluramine, a serotonergic agonist. These data were interpreted as indicating that central serotonergic

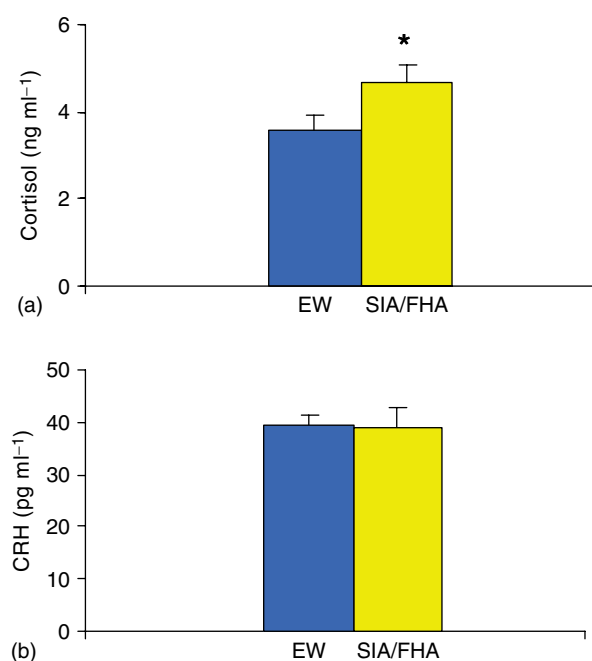


Figure 3 Mean $\pm$ SE cerebrospinal fluid hormone levels in eumenorrheic women (EW) and women with SIA/FHA, demonstrating allostatic resistance to cortisol feedback suppression of CRH. a, Cortisol levels; b, CRH levels. CRH, corticotropin releasing hormone; SIA/FHA, stress-induced anovulation/functional hypothalamic amenorrhea. From Brundu, B., Loucks, T. L., Adler, L. J., et al. (2006), Increased cortisol in the cerebrospinal fluid of women with functional hypothalamic amenorrhea, *Journal of Clinical Endocrinology and Metabolism* **91**, 1561–1565.

tone was reduced in stress-sensitive as contrasted with stress-resilient monkeys. Second, serotonergic neurons, which express leptin receptors, terminate on GnRH neurons. Serotonin has satiety effects similar to those of leptin and may modulate the response to orexigenic and anorectic signals. Moenter's lab exposed fasting female mice to saline, leptin, or fluoxetine, a selective serotonin reuptake inhibitor (SSRI) antidepressant, and followed estrus cycle length. Leptin and fluoxetine prevented fasting-induced cycle lengthening, whereas the serotonin receptor antagonist metergoline blocked cycle rescue by fluoxetine. Treating leptin-deficient *ob/ob* and leptin receptor-deficient *db/db* mice with fluoxetine did not normalize body weight or rescue estrus cycles. Further, there is abundant evidence of a role for the serotonergic system in the regulation of stress reactivity and feedback inhibition of the HPA axis by cortisol. In addition, variation in transporter-facilitated uptake of serotonin has been implicated in anxiety in humans and in animal models of anxiety and stressfulness. Human serotonin reuptake transporter (SERT) gene transcription is modulated by a common

polymorphism in its upstream regulatory region. The short variant reduces the transcriptional efficiency of the SERT gene promoter, resulting in decreased SERT expression and serotonin reuptake. The short (s) variant polymorphism is associated with an increased incidence of affective disorders in humans and maladaptive behavior in monkeys coincident with increased HPA responsivity to stress. Further, the long/short genotype produces a phenotype similar to that of the short/short genotype. In humans, the SERT polymorphisms moderated the influence of stressful life events on depression in a representative birth cohort. Stress-sensitive monkeys had a lower expression of SERT mRNA and monoamine oxidase A (MAO-A) in the caudal region of the dorsal raphe nucleus, whereas no differences were detected between stress-sensitive and stress-resilient monkeys in the mRNA for the serotonin 1A autoreceptor or MAO-B. The determinants of stressfulness in women with SIA/FHA are incompletely understood, and this is certainly one of the variables that warrants further investigation.

Judd et al. reported that a single 2 mg dose of alprazolam, a GABA receptor agonist, decreased cortisol levels and increased LH pulse frequency from 0.8 to 2.0 pulses/8 h in women with stress-related anovulation, whereas its administration decreased LH pulse frequency in eumenorrheic women in the follicular phase. These data also suggest a possible indirect role for GABA neurons in the stress-induced neuromodulation of GnRH pulsatility. Not surprisingly, the neurochemistry of stress and SIA/FHA is far from simple, and firm mechanistic conclusions are not possible at present.

Behavioral, Nutritional, and Metabolic Influences on the Reproductive Axis

Identification of factors that activate the adrenal axis and suppress the thyroidal and ovarian axes can be challenging. It has been suggested that different stressors and behaviors activate somewhat different central mechanisms, such that the signals that alter hypothalamic function are specific to the type of stressor. Indeed, this variability may well explain in part why the neurochemistry of stress is so complex. Psychosocial dilemmas are seen as activating those central pathways subserving perception, whereas exercise and weight loss are generally viewed as disturbing metabolic regulation. Although it seems logical that specificity in the neural or peripheral cascades mediating responses to specific stressors exist, we currently have no method for clearly delineating psychogenic from metabolic stress. Psychogenic stress has a metabolic cost and metabolic stressors, such as

food restriction and excessive exercise, are often initiated to cope with psychogenic stress. However, until proven otherwise, the safest assumption is that stress comes in flavors and that some individuals are more sensitive than others to the same stressor or set of stressors.

Behavioral Influences

A number of behavioral and other psychogenic factors including exercise, low weight and weight loss, affective and eating disorders, various personality characteristics, drug use, and external and intrapsychic stresses have been associated with functional hypothalamic hypogonadism. Given individual variation in metabolism, autonomic tone, habitus, aptitudes, attitudes, and psychological valences, what is stressful to one person may be more or less so to another. Therefore, it is not surprising to find behavioral heterogeneity in the pathogenesis of SIA/FHA. Any given stressor, when the dose is large enough, can probably activate the central neural pathways leading to the disruption of GnRH. In clinical research, the trend has been to study single stressors and to partition as separate populations women with exercise amenorrhea, anorexia nervosa, and idiopathic amenorrhea. Populations studied in clinical research settings may not be entirely representative of all women with SIA/FHA because research subjects must meet relatively strict inclusion and exclusion criteria. In general, women with SIA/FHA do not report or do not have an easily identified solitary stressor. Typically, there are multiple, seemingly minor, stressors, such as a combination of job or school pressures, poor eating habits, and increased energy expenditure through activity or exercise.

To understand the role of psychological variables, such as attitudes and expectations, in the pathogenesis of SIA/FHA, we compared three groups of women: those with eumenorrhea and demonstrable luteal adequacy; those with SIA/FHA unrelated to excessive exercise, weight loss, an eating disorder, drug use, or an affective disorder; and those with anovulation due to an identifiable organic cause. Being amenorrheic, regardless of cause, was associated with a compromised sense of psychological equilibrium as reported on psychometric inventories, but, as a group, only women with SIA/FHA differed from the other groups on scales that measured unrealistic expectations and dysfunctional attitudes (defined as attitudes likely to impair coping responses). For instance, women with SIA/FHA were both highly perfectionistic and sociotrophic (having a high need for social approval). Because perfectionism interferes with social approval or acceptance, one interpretation is that the concomitant high drive for perfectionism and sociotrophy

creates an intrapsychic conflict that women with SIA/FHA may not possess the appropriate coping skills to resolve. Another interpretation is that the expectation of simultaneously being perfect and garnering social approval is an unrealistic expectation of self and others. Our earlier study suggested that women with SIA/FHA had trouble relaxing and having fun, attributes that may further predispose them to value performance at the expense of psychological needs. Although women with SIA/FHA do not typically meet the criteria for an eating disorder, they display many attitudes and behaviors similar to women with eating disorders, such as a drive for thinness and disordered eating. What appears to separate women with undifferentiated SIA/FHA from those with an eating disorder is the degree of disturbance, including the degree of food restriction, weight loss, bingeing, and perfectionism. However, direct comparisons have not been conducted for any of these psychological attributes, so these interpretations are based on impressions from extant literature. The bottom line is that attitudes and expectations engender behaviors such as aberrant food intake and excessive exercise that further challenge the hypothalamus to maintain homeostasis. Whether the identifiable behavior is performance pressure, exercise, or irregular food intake, the end result is the same, that is, the disruption of the GnRH pulse generator to the extent that anovulation occurs. It stands to reason, then, that the key to recovery is to change both the behaviors and attitudes that have initiated and now sustain reduced GnRH drive. Once the hypothalamic GnRH pulse generator has been disrupted, it may take a prolonged duration of energy balance and improved psychological equilibrium for the hypothalamic allostasis to reverse and for ovulatory function to return.

The available data suggest that any behavior or expectation that concomitantly activates to a sufficient degree the HPA and HPT axes has the potential to disrupt the GnRH pulse generator. Thus, the list of behaviors and attitudes associated with the development of SIA/FHA is expected to be diverse and extensive. Generally, a mix of multiple, seemingly minor psychogenic and metabolic stressors appears to be more deleterious to reproductive function than a solitary stressor.

Nutritional and Metabolic Influences

Nutritional and metabolic signals play critical roles in the elaboration of homeostatic and allostatic responses to ongoing challenges and can influence the reproductive system at many levels. Research on overnutrition has focused primarily on specifying its effects on gonadal function, with some attention

having been given to the secondary consequences of obesity on endometrial development and central hypothalamic drive. Few studies have characterized the role of nutritional signals in the modulation of reproductive function in men, but available evidence suggests that undernutrition is as deleterious to reproductive competency in men as it is in women.

Metabolic imbalance occurs when energy expenditure exceeds energy intake (negative energy balance). The brain is the most metabolically active tissue in the body, requiring 16 times more energy per unit mass than muscle tissue. Because humans have much bigger brains relative to body size than do other primates or other species, they use more of their daily energy intake than any other species to supply their brain; humans are estimated to use 25%, monkeys 8%, and rodents 5%. Indeed, as a species, humans may be uniquely sensitive to energy imbalance. Metabolism reflects the interactions of the neuroendocrine system (e.g., HPA and HPT axes), body composition (adipokines), and the enteric system (insulin, ghrelin, etc.). Given the multitude of redundant metabolic and nutritional signals communicating energy status to the brain, it is difficult to specify the independent role of any one factor or signal. Signals reflecting energy stores, recent nutritional information, and specific classes of nutrients are integrated in the central nervous system, particularly the hypothalamus, to coordinate energy intake and expenditure. Chronic energy deficiency alters thyroidal function to slow metabolism and correct negative energy balance. Food intake is influenced by availability, emotional state, social cues, and learned behaviors. Although peripheral signals convey information about energy stores and immediate energy availability to the hypothalamus, particularly the arcuate nucleus, these two categories of signals are not exclusive. For instance, leptin and insulin, which are actively transported across the blood-brain barrier, potentiate satiety signals. Thus, the level of any one signal *per se* may be less important than the action of that hormone, and the action is likely to be gated by the constellation of other metabolic signals. Putative orexigenic signals include ghrelin, NPY, orexins A and B, melanin-concentrating hormone (MCH), and agouti-related peptide. Putative anorectic and satiety signals include (but are not limited to) cortisol, CRH, insulin, glucose, resistin, leptin, proopiomelanocortin (POMC), cocaine- and amphetamine-regulated transcript (CART) peptide, peptide YY (PYY), and glucagon-like peptide (GLP)-1. Adipokines (adipocyte-derived hormones) implicated in energy regulation include leptin, adiponectin, and resistin. Leptin is the dominant long-term energy signal informing the brain of adipose

energy reserves; it also functions as a satiety signal. Adiponectin acts as an insulin-sensitizing agent by reducing hepatic glucose production and is reduced in obesity. Resistin is linked to insulin tolerance and decreases glucose uptake by adipocytes. Ghrelin, a 28-amino-acid acylated hormone that is also a growth hormone secretagogue, is produced by the gastrointestinal tract, especially the fundus of the stomach. Plasma ghrelin levels rise during fasting and immediately before anticipated mealtimes and then fall within an hour of food intake, suggesting that ghrelin is important for meal initiation. Given the plethora of nutritional and metabolic signals, specifying their independent effects is challenging, particularly in humans.

Relevant human studies on the role of these factors in reproduction are few. Depression, like stress, involves LHPA axis dysregulation and is associated with loss of appetite; resistin levels correlated with free cortisol levels and adiponectin correlated with insulin sensitivity in a study of depressed humans. In a rodent model, the central LHPA axis challenge of hypoxia and the LHPA axis feedback challenge of dexamethasone independently decreased body weight and increased adiponectin levels, whereas dexamethasone, but not hypoxia, increased resistin. In women, ghrelin levels were reported to be increased in anorexia nervosa and exercise amenorrhea. However, Troisi et al. observed lower ghrelin in women with anorexia or bulimia that correlated with serum cortisol levels. In these studies, subjects and controls differed markedly with regard to body weight and leptin levels, food intake was not assessed, and the diurnal pattern of ghrelin was not determined; for instance, De Souza et al. and Tanaka et al. obtained only a single morning's blood sample after an overnight fast and Misra et al. determined ghrelin levels only overnight. Of note, Miljec et al. reported ghrelin insensitivity in response to an infusion of ghrelin in women with anorexia nervosa compared to eumenorrheic women, but appetite was measured only by self-report and not by food intake. Women with SIA/FHA are more active than most eumenorrheic women and tend to consume less fat and more carbohydrates, but reproductive recovery from SIA/FHA did not require weight gain and was not associated with a change in standard metabolic variables such as leptin.

A balance between energy intake and expenditure is critical to survival. Every action, even thinking, has an energetic cost. Given the fundamental nature of metabolism, it is not surprising that organisms have evolved a complex and redundant signaling system to gate appetite, food-seeking behavior, and fuel storage. Although stress increases energetic demand, the acute behavioral response is often a blunting of

appetite and a reduction in food intake so as to shift attention away from maintenance/sustenance needs and toward coping responses. Chronic reductions of appetite and energy intake in the face of an increase in energy demand probably represent appetite or metabolic allostasis; because of ensuing hypothalamic hypothyroidism, undernutrition may develop without weight loss. The exact mechanisms subserving the multiple associations between reproduction and nutrition and the applicability of animal studies to human syndromes remain to be determined. Given the considerations outlined here, including the high metabolic demands of the human brain, energetic imbalance may have more serious reproductive consequences for humans than for other animals.

Psychological states generate metabolic demands. Psychological states recognized as stressors activate the LHPA axis. When this activation is acute and limited, the adrenal releases a bolus of cortisol, a hormone with multiple metabolic properties. If the stressor is chronic and mild to moderate, then there is modest activation of the LHPA such that cortisol is mildly elevated but the circadian pattern of cortisol is preserved. With extreme unrelenting stress, cortisol secretion is elevated and there is loss of the circadian pattern. Conditions that chronically activate the LHPA axis have been associated with a spectrum of reproductive impairment.

In summary, there is a plethora of metabolic signals that reflect the acute and chronic nutritional state of a given individual. These signals convey information to important tissues and organs and thereby engender homeostatic and allostatic responses. A key target tissue is the brain. Both undernutrition and overnutrition impact reproductive processes. Given the energetic demands of reproductive processes, however, undernutrition has the potential to induce more significant reproductive compromise.

Synergism among Stressors

The notion that stress comes in neurochemical flavors is supported by animal studies suggesting that there are subtle but distinct differences in the neuroendocrine responses to different stress paradigms. Further, neuroendocrine and metabolic responses to acute exercise were greater in men whose HPA axis did not suppress when they were given dexamethasone before the exercise challenge. These data indicate that the degree of HPA activation potentiates the neuroendocrine and metabolic responses to subsequent challenge. Conversely, Altemus et al. showed that lactating women are hyporesponsive to exercise challenge. Taken together, these data buttress the

notion that responses to a given stressor are gated not only by the stressor type but also by host factors, including hormonal and metabolic states. Thus, some individuals are clearly more reactive to similar stressors than others. Obviously, emotional valence and expectations also determine the extent to which psychosocial variables serve as a psychogenic stressors. Our preliminary investigations in women who developed SIA/FHA unrelated to weight loss, exercise, and definable psychiatric disorders indicated that a primary factor that distinguished women with SIA/FHA from those with definable causes of anovulation and those who were ovulatory was the presence of unrealistic expectations. Fioroni also found that women with SIA/FHA, compared to eumenorrheic women or those with other causes of anovulation, held more negative attributions about recent life events. Kirschbaum found that men who did not habituate when exposed to repeated psychogenic challenge viewed themselves as less attractive, had lower self-esteem, and reported being in a depressed mood more often. Apparently, unachievable ambitions or other cognitive distortions create vulnerability to life's inevitable challenges and probably heighten responsivity to metabolic stressors such as exercise or food restriction. Potentially, the converse is true. Energetic imbalance may augment reactivity to psychogenic stressors.

There can be no doubt that weight loss and exercise serve as metabolic stressors. In monkeys trained to run, it was shown that caloric supplementation reversed the anovulation induced by training. Interestingly, the monkeys did not spontaneously develop a compensatory increase in appetite and had to be bribed with colorful candy to consume more calories. On the other hand, modest dietary restriction accompanied by small amounts of exercise greatly increased the proportion of monkeys who become anovulatory when presented with social stress. A prospective study of unselected women demonstrated that exercise and weight loss caused anovulation. It is likely that sufficient exercise and weight loss, independent of psychogenic stress, can alter metabolism to the point that GnRH pulsatility is disrupted. Loucks and Thurma quantified the amount of energy restriction needed to impact GnRH drive by studying eumenorrheic women in the follicular phase. Energy balance was achieved by providing 45 kcal/kg of lean body mass (LBM)/day. Graded daily energy deficits of 10, 25, and 35 kcal/kg were then experimentally induced for 5 days. An energy deficit of 33% had no impact on LH pulse frequency, whereas an energy deficit of approximately 75% induced a decline in LH pulse frequency of approximately 40%. The induction of an energy deficit resulted in a graded increase in the 24-h cortisol level. At an energy

availability of 10 kcal/kg of LBM (75% deficit), the mean 24-h cortisol was increased by approximately 30%, which is the amount of increase typically seen in women with SIA/FHA. Much like stress-sensitive monkeys, women whose luteal phase progesterone levels were lowest at the initiation of the energy restriction showed the greatest response to the imposed metabolic challenge. In most real-life situations, except for extreme circumstances such as war or famine, metabolic deficits are not imposed but, rather, initiated by individuals in response to self-imposed expectations including drive for thinness. Most women with SIA/FHA, when carefully evaluated, display more than one trait, state, or behavior capable of activating stress response cascades or inducing a mild metabolic deficit, but most do not have a profound metabolic deficit that alone would explain the reduction in central GnRH drive. Many of the behaviors, such as exercise, that independently suppress central reproductive drive but only at more extreme levels, may be initiated as coping responses to psychosocial dilemmas. Because of the synergism between metabolic and psychogenic stressors, a combination of multiple, small magnitude, mixed stressors may be potentially more disruptive of reproductive function than a single large stressor limited to one category.

To better understand how a combination of seemingly minor psychogenic and metabolic stressors might synergistically disrupt GnRH drive as demonstrated in our monkey model, we compared endocrine responses to submaximal exercise in women with SIA/FHA to those in eumenorrhea. Women with SIA/FHA displayed a larger increase in cortisol than ovulatory eumenorrheic women in response to exercise. Further, glucose responses between the two groups were divergent in that women with SIA/FHA showed a 10% decrease in glucose whereas eumenorrheic women had a modest 3% increase in glucose, presumably to cope with the energetic demand of exercise. Interestingly, these two groups did not differ at baseline with regard to cortisol or glucose levels. The decrement in glucose seen in SIA/FHA but not eumenorrheic women suggests latent metabolic compromise and indicates that SIA/FHA women are unable to meet energetic demands of ongoing activities. Further, it is likely that the drop in glucose activates the HPA axis and is at least partly responsible for the sustained hypercortisolemia characteristic of SIA/FHA. Because metabolic signals modulate GnRH pulsatility, exercise-induced metabolic imbalance also probably contributes to ongoing reproductive suppression. These results also reveal why the endocrine effects of a stressor, such as exercise, depend on the preexisting neuroendocrine and metabolic state of the individual.

Treatment Considerations

SIA may not be recognized unless the menstrual interval is markedly short, long, irregularly irregular, or absent. Likewise, luteal phase insufficiency due to decreased hypothalamic drive may not be noted unless infertility results. Even then, luteal insufficiency is notoriously difficult to document unless it is recurrent. Based on the foregoing concepts, we speculate that the more clinically evident the ovarian compromise, the greater is the hypothalamic challenge and the more profound are the associated adrenal and thyroidal derangements and sex steroid deprivation.

If a woman with functional hypothalamic hypogonadism is seeking to become pregnant, ovulation induction can be accomplished technically with the exogenous administration of pulsatile GnRH therapy or gonadotropins. The administration of exogenous GnRH therapy is advantageous insofar as it diminishes the risk of ovarian hyperstimulation and multiple gestation associated with gonadotropins. Clomiphene citrate is an option, but it may be less effective because of its hypothalamic site of action and the hypothalamus of women with SIA/FHA is already not responding to decreased sex steroid secretion. Concerns have been raised that ovulation induction may place women with SIA/FHA at increased risk for premature labor and intrauterine growth restriction. The parenting skills of women with SIA/FHA may be impaired because they are already overwhelmed and stressed prior to pregnancy and delivery and their psychological precariousness may place their children at risk for poor psychosocial development. Further, a recent study showed that children born to mothers with clinically occult hypothyroidism due to autoimmune thyroiditis had a mean full-scale intelligence quotient that was seven points lower than the control population. The women with clinically silent hypothyroidism had a 30% reduction in thyroxine, which is roughly what is observed in women with SIA/FHA. Of critical importance is the fact that maternal thyroxine is the only source of fetal thyroxine in the first trimester and the predominant fetal source in the second and third trimesters. Because the fetal brain requires an appropriate amount of thyroxine for neurogenesis, even small deficits in thyroxine may induce neurodevelopmental deficits. Increased maternal cortisol may also have independent effects on fetal neurodevelopment and organogenesis. Recent evidence showed that severe stress, such as that associated with the unexpected death of a child, increased the risk of congenital anomalies of the cranial neural crest eightfold. Further, stress and its endocrine concomitants and even brief undernutrition have been implicated as a cause

of preterm delivery. It is not known if the endocrine concomitants associated with SIA/FHA pose a similar risk, but this is clearly a potential hazard if ovulation induction is undertaken before amelioration of the allostatic changes in the adrenal and thyroidal axes.

A popular approach to a woman with SIA/FHA who is not seeking immediately to become pregnant is to offer her hormone replacement. This approach is based on the presumption that sex steroid deprivation is the primary therapeutic issue. There are inherent limitations with this approach, however. First, data indicate that exogenous sex steroid exposure does not fully promote bone accretion or cardioprotection in the presence of ongoing metabolic derangements. Second, the ongoing insults to the brain from chronic amplification of stress cascades go unchecked. Further, estrogen therapy does not correct hypothalamic hypothyroidism. In short, hormone therapy may mask potentially deleterious processes that are unlikely to be ameliorated by hormone exposure alone. It is critical to remember that SIA/FHA is more than a disorder of reduced GnRH secretion.

Hormone therapy *per se* is unlikely to be harmful, but more than hormone administration is needed – the stress process needs to be interrupted. Although psychopharmacological approaches have not been well studied, they probably could be used on an interim basis in special circumstances. The study by Judd suggested that a short course of alprazolam might be effective in reducing HPA activation and permitting hypothalamic-pituitary-ovarian (HPO) recovery. However, this approach is not recommended for a woman hoping to conceive. The optimal intervention is to reverse the stress process so that the hypothalamus recovers and gonadal function resumes. An integral goal of the treatment plan for women with functional hypothalamic hypogonadism is to help them identify and ameliorate the sources of psychogenic and metabolic stress and to provide emotional support while they learn coping mechanisms other than dieting or exercising. Nonpharmacological interventions such as stress management, relaxation training, or psychoeducation empower individuals by fostering self-care and competency. In this regard, nonpharmacological therapies have the potential to produce long-term benefits on psychological and, thereby, physical health. Behavioral therapies acknowledge the wisdom of the body and understand that SIA/FHA represents an endocrine adaptation that can be reversed with appropriate psychogenic and behavioral modifications.

Given these considerations, we recently studied whether CBT aimed at ameliorating problematic attitudes and behaviors facilitated ovarian recovery in

normal weight women with SIA/FHA. Women with SIA/FHA were randomly assigned to observation versus CBT groups. CBT consisted of 16 visits with a physician, therapist, or nutritionist during 20 weeks. The two groups were followed for return of menses for up to 8 weeks following the intervention. Regardless of menstrual pattern, estradiol and progesterone levels were monitored at weekly intervals for 4 weeks before and after observation or CBT. Approximately 88% of those who underwent CBT had evidence of

ovulation, whereas only 25% of those in the observation group did. **Figure 4** illustrates the recovery of sex steroid secretion in a woman in the CBT group who showed ovarian recovery and in a woman in the observation group who did not have recovery of ovarian function. Interestingly, when it occurred, ovarian recovery was not associated with significant weight gain. This does not mean that subjects did not alter food intake or energy expenditure, however. Improved nutrition generally restores the thyroidal

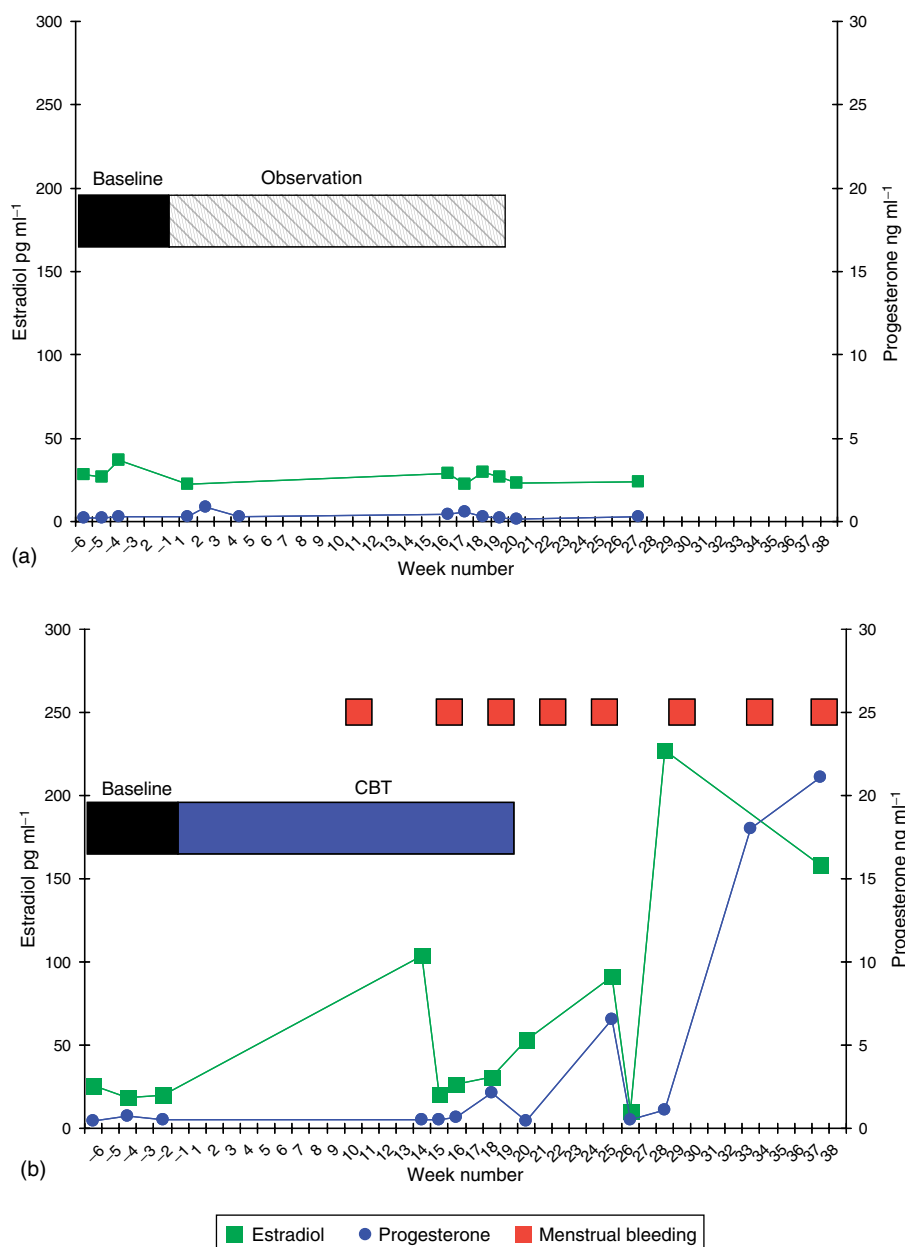


Figure 4 Serum estradiol and progesterone levels and menstrual bleeding in women with SIA/FHA. a, A patient who was observed and did not recover; b, a patient who was treated with cognitive-behavior therapy and recovered. CBT, cognitive-behavior therapy; SIA/FHA, stress-induced anovulation/functional hypothalamic amenorrhea. From Berga, S. L., Marcus, M. D., Loucks, T. L., et al. (2003), Recovery of ovarian activity in women with functional hypothalamic amenorrhea who were treated with cognitive behavior therapy, *Fertility and Sterility* 80, 976–981.

axis, thereby leading to an increased basal metabolic rate and a need for more calories. Our earlier studies showed that women who experienced a spontaneous recovery from SIA/FHA were eucortisolemic, had increased but not fully restored LH pulsatility, and markedly increased TSH levels in the face of persistent reductions in T_3 and T_4 levels. We recently compared metabolic variables in women who did and who did not recover from SIA/FHA after CBT intervention and observed that ovarian recovery was associated with a decline in cortisol and increase in TSH but that thyroidal hormones and leptin levels were comparable between recovered and unrecovered women. Taken together, these findings suggest that the thyroidal axis, which gates metabolic tone, remains restrained by an as yet uncharacterized factor(s) after HPA and HPO recovery or recovers more slowly than the HPA and HPO axes, similar to the recovery timeline related to acute illness. Ultimately, behavioral and psychological interventions that address problematic behaviors and attitudes permit the resumption of ovarian function along with recovery of the LHPA and HPT axes. In short, full endocrine recovery offers better individual, maternal, and child health.

Summary

Stress is one of the most common and most commonly underappreciated causes of infertility and reproductive compromise in men and women. Stress-induced hypothalamic hypogonadism increases the acute and chronic health burden for individuals and their offspring. Our findings regarding the pathogenesis of SIA/FHA reveal a powerful synergism between metabolic imbalance and psychogenic challenge and underscore that seemingly mundane psychological issues evoke not only reproductive compromise but also concomitant neuroendocrine and metabolic allostasis. These mechanistic pathobiological insights have been harnessed to develop therapeutic options aimed at ameliorating the sustaining cognitive and behavioral variables. By showing that CBT reverses SIA/FHA, we can appreciate the inherent neural plasticity that forms the rationale for coupling mindfulness approaches with the judicious use of psychopharmacological and technical interventions.

Acknowledgments

Portions of the work presented herein were funded by grants from the National Institutes of Health (R01-MH50748 to SLB and RR-00046 to the General Clinical Research Center at the University of Pittsburgh).

See Also the Following Article

Allostasis and Allostatic Load.

Further Reading

- Altemus, M., Deuster, P. A., Galliven, E., et al. (1985). Suppression of hypothalamic-pituitary-adrenal axis responses to stress in lactating women. *Journal of Clinical Endocrinology and Metabolism* 80, 2954–2959.
- Alvero, R., Kimzey, L., Sebring, N., et al. (1998). Effects of fasting on neuroendocrine function and follicle development in lean women. *Journal of Clinical Endocrinology and Metabolism* 83, 76–80.
- Ariyasu, H., Takaya, K., Tagami, T., et al. (2001). Stomach is major source of circulating ghrelin and feeding state determines plasma ghrelin-like immunoreactivity levels in humans. *Journal of Clinical Endocrinology and Metabolism* 86, 4753–4758.
- Badman, M. K. and Flier, J. S. (2005). The gut and energy balance: visceral allies in the obesity wars. *Science* 307, 1909–1914.
- Barr, C. S., Newman, T. K., Schwandt, M., et al. (2004). Sexual dichotomy of an interaction between early adversity and the serotonin transporter gene promoter variant in rhesus macaques. *Proceedings of the National Academy of Sciences USA* 101, 12358–12363.
- Barr, C. S., Newman, T. K., Shannon, C., et al. (2004). Rearing condition and rh5-HTTLPR interact to influence limbic-hypothalamic-pituitary-adrenal axis response to stress in infant macaques. *Biological Psychiatry* 55, 733–738.
- Becker, A. E., Grinspoon, S. K., Klibanski, A., et al. (1999). Eating disorders. *New England Journal of Medicine* 340, 1092–1098.
- Bennett, A. J., Lesch, K. P., Heils, A., et al. (2002). Early experience and serotonin transporter gene variation interact to influence primate CNS function. *Molecular Psychiatry* 7, 118–122.
- Berga, S. L. (2002). Disorders of gonadotropin secretion. In: Wass, J. & Shlet, S. M. (eds.) *Oxford textbook of endocrinology and diabetes*, pp. 1119–1129. Oxford University Press.
- Berga, S. L. and Daniels, T. L. (1991). Use of the laboratory in disorders of reproductive neuroendocrinology. *Journal of Clinical Immunoassay* 14, 23–28.
- Berga, S. L., Daniels, T. L. and Giles, D. E. (1997). Women with functional hypothalamic amenorrhea but not other forms of anovulation display amplified cortisol concentrations. *Fertility and Sterility* 67, 1024–1030.
- Berga, S. L., Loucks, A. B., Rossmann, W. G., et al. (1991). Acceleration of LH pulse frequency in functional hypothalamic amenorrhea by dopaminergic blockade. *Journal of Clinical Endocrinology and Metabolism* 72, 151–156.
- Berga, S. L., Loucks, T. L. and Cameron, J. L. (2001). Endocrine and chronobiological effects of fasting in women. *Fertility and Sterility* 75, 926–932.
- Berga, S. L., Loucks-Daniels, T. L., Adler, L. J., et al. (2000). Cerebrospinal fluid levels of corticotropin-releasing

- hormone in women with functional hypothalamic amenorrhea. *American Journal of Obstetrics and Gynecology* 182, 776–784.
- Berga, S. L., Marcus, M. D., Loucks, T. L., et al. (2003). Recovery of ovarian activity in women with functional hypothalamic amenorrhea who were treated with cognitive behavior therapy. *Fertility and Sterility* 80, 976–981.
- Berga, S. L., Mortola, J. F., Gorton, L., et al. (1989). Neuroendocrine aberrations in women with functional hypothalamic amenorrhea. *Journal of Clinical Endocrinology and Metabolism* 68, 301–308.
- Berga, S. L., Mortola, J. F. and Yen, S. S. C. (1988). Amplification of nocturnal melatonin secretion in women with functional hypothalamic amenorrhea. *Journal of Clinical Endocrinology and Metabolism* 66, 242–244.
- Berga, S. L. and Yen, S. S. C. (2004). Reproductive failure due to central nervous system-hypothalamic-pituitary dysfunction. In: Strauss, J. & Barbieri, R. (eds.) *Yen and Jaffe's reproductive endocrinology: physiology, pathophysiology, and clinical management* (5th edn., pp. 564–569). Philadelphia: Elsevier Saunders.
- Bethea, C. L., Pau, F. K. Y., Fox, S., et al. (2005). Sensitivity to stress-induced reproductive dysfunction linked to activity of the serotonin system. *Fertility and Sterility* 83, 148–155.
- Bethea, C. L., Streicher, J. M., Mirkes, S. J., et al. (2005). Serotonin-related gene expression in female monkeys with individual sensitivity to stress. *Neuroscience* 132, 151–166.
- Biller, B. M. K., Federoff, J. H., Koenig, J. I., et al. (1990). Abnormal cortisol secretion and responses to corticotropin-releasing hormone in women with hypothalamic amenorrhea. *Journal of Clinical Endocrinology and Metabolism* 70, 311–317.
- Brundu, B., Loucks, T. L., Adler, L. J., et al. (2006). Increased cortisol in the cerebrospinal fluid of women with functional hypothalamic amenorrhea. *Journal of Clinical Endocrinology and Metabolism* 91, 1561–1565.
- Bullen, B. A., Skrinar, G. S., Betins, I. Z., et al. (1985). Induction of menstrual disorders by strenuous exercise in untrained women. *New England Journal of Medicine* 312, 1349–1353.
- Cameron, J. L., Nosbisch, C., Helmreich, D., et al. (1990). Reversal of exercise-induced amenorrhea in female cynomolgus monkeys (*Macaca fascicularis*) by increasing food intake. Paper presented at the 72nd annual meeting of the Endocrine Society, Atlanta, GA. Abstract # 1042.
- Cameron, J. L., Weltzin, T. E., McConahan, C., et al. (1997). Slowing of pulsatile luteinizing hormone secretion in men after forty-eight hours of fasting. *Journal of Clinical Endocrinology and Metabolism* 73, 35–41.
- Caspi, A., Sudgen, K., Moffitt, T. E., et al. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 301, 386–389.
- Challis, J., Sloboda, D., Matthews, S., et al. (2000). Fetal hypothalamic-pituitary-adrenal (HPA) development and activation as a determinant of the timing of birth, and of postnatal disease. *Endocrine Research* 26, 489–504.
- Champoux, M., Bennett, A., Shannon, C., et al. (2002). Serotonin transporter gene polymorphism, differential early rearing, and behavior in rhesus monkey neonates. *Molecular Psychiatry* 7, 1058–1063.
- Cummings, D. E., Purnell, J. Q., Frayo, R. S., et al. (2001). A preprandial rise in plasma ghrelin suggests a role in meal initiation in humans. *Diabetes* 50, 1714–1719.
- Daniels, T. L., Cameron, J. L., Marcus, M. D., et al. (1997). Recovery from functional hypothalamic amenorrhea (FHA) involves resolution of hypothalamic hypothyroidism. Paper presented at the 79th annual meeting of the Endocrine Society, Minneapolis, MN. Abstract #P1–378.
- DeSouza, M. J., Miller, B. E., Loucks, A. B., et al. (1998). High frequency of luteal phase deficiency and anovulation in recreational women runners: blunted elevation in follicle-stimulating hormone observed during luteal-follicular transition. *Journal of Clinical Endocrinology and Metabolism* 83, 4220–4232.
- de Boo, H. A. and Harding, J. E. (2006). The developmental origins of adult disease (Barker) hypothesis. *Australian and New Zealand Journal of Obstetrics and Gynaecology* 46, 4–14.
- Drew, F. L. (1961). The epidemiology of secondary amenorrhea. *Journal of Chronic Diseases* 14, 396–407.
- Dudas, B. and Merchenthaler, I. (2006). Three-dimensional representation of the neurotransmitter systems of the human hypothalamus: inputs of the gonadotropin hormone-releasing hormone neuronal systems. *Journal of Neuroendocrinology* 18, 79–95.
- Fioroni, L., Fava, M., Genazzani, A. D., et al. (1994). Life events impact in patients with secondary amenorrhea. *Journal of Psychosomatic Research* 38, 617–622.
- Fries, H., Nilius, S. J. and Pettersson, F. (1974). Epidemiology of secondary amenorrhea. II: A retrospective evaluation of etiology with regard to psychogenic factors and weight loss. *American Journal of Obstetrics and Gynecology* 118, 473–479.
- Fuller, R. W. (1996). Serotonin receptors involved in regulation of pituitary-adrenocortical function in rats. *Behavioural Brain Research* 73, 215–219.
- Genazzani, A. D., Gamba, O. and Petraglia, F. (1998). Estrogen replacement therapy modulates spontaneous GH secretion but does not affect GHRH-induced GH response and low T3 syndrome in women with hypothalamic amenorrhea associated to weight loss. *Journal of Endocrinological Investigation* 21, 353–357.
- Giles, D. E. and Berga, S. L. (1993). Cognitive and psychiatric correlates of functional hypothalamic amenorrhea: a controlled comparison. *Fertility and Sterility* 60, 486–492.
- Golden, N. H., Lanzkowsky, L., Schebendach, J., et al. (2002). The effect of estrogen-progestin treatment on bone mineral density in anorexia nervosa. *Journal of Pediatric and Adolescent Gynecology* 15, 135–143.
- Goodman, R. L. (1989). Functional organization of the catecholaminergic neural system inhibiting luteinizing hormone secretion in anestrus ewes. *Neuroendocrinology* 47, 203–250.
- Grinspoon, S., Miller, K., Coyle, C., et al. (1999). Severity of osteopenia in estrogen-deficient women with anorexia nervosa and hypothalamic amenorrhea. *Journal of Clinical Endocrinology and Metabolism* 84, 2049–2055.

- Haddow, J. E., Palomaki, G. E., Allan, W. C., et al. (1999). Maternal thyroid deficiency during pregnancy and subsequent neuropsychological development of the child. *New England Journal of Medicine* **341**, 549–555.
- Hansen, D., Lou, H. C. and Olsen, J. (2000). Serious life events and congenital malformations: a national study with complete follow-up. *Lancet* **356**, 875–880.
- Heisler, L. E., Tumber, A. J., Reid, R. L., et al. (1994). Vasopressin mediates hypoglycemia-induced inhibition of luteinizing hormone secretion in the ovariectomized Rhesus monkey. *Neuroendocrinology* **60**, 297–304.
- Holsboer, F. and Barden, N. (1996). Antidepressants and hypothalamic-pituitary-adrenocortical regulation. *Endocrine Review* **17**, 187–205.
- Hurley, D. M., Brian, R., Outch, K., et al. (1984). Induction of ovulation and fertility in amenorrheic women by pulsatile low-dose gonadotropin-releasing hormone. *New England Journal of Medicine* **301**, 1069–1074.
- Judd, S. J., Wong, J., Saloniklis, S., et al. (1995). The effect of alprazolam on serum cortisol and luteinizing hormone pulsatility in normal women and in women with stress-induced anovulation. *Journal of Clinical Endocrinology and Metabolism* **80**, 818–823.
- Kaplan, J. R. and Manuck, S. B. (2004). Ovarian dysfunction, stress, and disease: a primate continuum. *ILAR Journal* **45**, 89–115.
- Kasuya, E., Nyberg, C. L., Mogi, K., et al. (1999). A role of γ -aminobutyric acid (GABA) and glutamate in control of puberty in female rhesus monkeys: effect of an antisense oligodeoxynucleotide for GAD67 messenger ribonucleic acid and MK801 on luteinizing hormone-releasing hormone release. *Endocrinology* **140**, 705–712.
- Khoury, S. A., Reame, N. E., Kelch, R. P., et al. (1987). Diurnal patterns of pulsatile luteinizing hormone secretion in hypothalamic amenorrhea: reproducibility and responses to opiate blockade and a α_2 -adrenergic agonist. *Journal of Clinical Endocrinology and Metabolism* **64**, 755–762.
- Kirschbaum, C., Prussner, J. C., Stone, A. A., et al. (1995). Persistent high cortisol responses to repeated psychological stress in a subpopulation of healthy men. *Psychosomatic Medicine* **57**, 468–474.
- Kiss, J. and Halasz, B. (1985). Demonstration of serotonergic axons terminating on luteinizing hormone releasing hormone neurons in the preoptic area of the rat using a combination of immunocytochemistry and high resolution autoradiography. *Neuroscience* **14**, 69–78.
- Klibanski, A. N., Biller, B. M. K., Schoenfeld, D. A., et al. (1994). The effects of estrogen administration on trabecular bone loss in young women with anorexia nervosa. *Journal of Clinical Endocrinology and Metabolism* **80**, 898–904.
- Kojima, M., Hosoda, H., Date, Y., et al. (1999). Ghrelin is growth hormone releasing acylated peptide from stomach. *Nature* **402**, 656–660.
- Kuljis, R. O. and Advis, J. P. (1989). Immunocytochemical and physiological evidence of a synapse between dopamine and luteinizing hormone releasing hormone-containing neurons in the ewe median eminence. *Endocrinology* **124**, 1579–1581.
- Leonard, W. R. and Robertson, M. L. (1994). Evolutionary perspectives on human nutrition: the influence of brain and body size on diet and metabolism. *American Journal of Human Biology* **6**, 77–88.
- Leranth, C., MacLusky, N. J., Shanabrough, M., et al. (1988). Catecholaminergic innervation of luteinizing hormone-releasing hormone and glutamic acid decarboxylase immunopositive neurons in the rat medial preoptic area. *Neuroendocrinology* **48**, 581–602.
- Lesch, K. P., Bengel, D., Heils, A., et al. (1996). Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* **274**, 1527–1531.
- Li, S. and Pelletier, G. (1996). Further studies on the mechanism of action of the endogenous benzodiazepine receptor ligand octadecaneuropeptide on gonadotropin-releasing hormone gene expression in the rat brain. *Neuroendocrinology* **64**, 79–84.
- Lopez, J. F., Chalmers, D. T., Little, K. Y., et al. (1998). Regulation of serotonin, glucocorticoid and mineralocorticoid receptor in rat and human hippocampus: implications for the neurobiology of depression. *Biological Psychiatry* **43**, 547–573.
- Loucks, A. B., Laughlin, G. A., Mortola, J. F., et al. (1992). Hypothalamic-pituitary-thyroidal function in eumenorrheic and amenorrheic athletes. *Journal of Clinical Endocrinology and Metabolism* **75**, 514–518.
- Loucks, A. B., Mortola, J. F., Girton, L., et al. (1989). Alternations in the hypothalamic-pituitary-ovarian and hypothalamic-pituitary-adrenal axes in athletic women. *Journal of Clinical Endocrinology and Metabolism* **68**, 402–411.
- Loucks, A. B. and Thurma, J. R. (2003). Luteinizing hormone pulsatility is disrupted at a threshold of energy availability in regularly menstruating women. *Journal of Clinical Endocrinology and Metabolism* **88**, 297–311.
- Loucks, T. L., Dube, J., Laychak, K., et al. (2004). Metabolic and endocrine responses to submaximal exercise challenge in women with functional hypothalamic amenorrhea. Paper presented at the 51st annual meeting of the Society of Gynecologic Investigation, Houston, TX. Abstract #605.
- Lupien, S. J., Nair, N. P., Briere, S., et al. (1999). Increased cortisol levels and impaired cognition in human aging: implication for depression and dementia in later life. *Reviews of Neuroscience* **10**, 117–139.
- MacLusky, N. J., Naftolin, F. and Leranth, C. (1988). Immunocytochemical evidence for direct synaptic connections between corticotropin-releasing factor (CRF) and gonadotropin-releasing hormone (GnRH)-containing neurons in the preoptic area of the rat. *Brain Research* **439**, 381–395.
- Mancini, F., Basanta-Henry, P., Corominal, A. G., et al. (2006). Do metabolic variables gate reproductive recovery in functional hypothalamic amenorrhea? Paper presented at the 88th annual meeting of the Endocrine Society, Boston, MA. Abstract # P2–531.
- Marcus, M. D., Loucks, T. L. and Berga, S. L. (2001). Psychological correlates of functional hypothalamic amenorrhea. *Fertility and Sterility* **76**, 310–316.

- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* **338**, 171–179.
- McEwen, B. S. (2002). Sex, stress, and the hippocampus: allostasis, allostatic load, and the aging process. *Neurobiology of Aging* **23**, 921–939.
- McMillen, I. C., Schwartz, J., Coulter, C. L., et al. (2004). Early embryonic environment, the fetal pituitary-adrenal axis, and the timing of parturition. *Endocrine Review* **30**, 845–850.
- Mikkelsen, J. D., Hay-Schmidt, A. and Kiss, A. (2004). Serotonergic stimulation of the rat hypothalamo-pituitary-adrenal axis: interaction between 5-HT1A and 5-HT2A receptors. *Annals of the New York Academy of Sciences* **1018**, 65–70.
- Miljic, D., Pekic, S., Djurovic, M., et al. (2006). Ghrelin has partial or no effect on appetite, growth hormone, prolactin, and cortisol release in patients with anorexia nervosa. *Journal of Clinical Endocrinology and Metabolism* **91**, 1491–1495.
- Miller, D. S., Reid, R. R., Cetel, N. S., et al. (1983). Pulsatile administration of low-dose gonadotropin-releasing hormone: ovulation and pregnancy in women with hypothalamic amenorrhea. *Journal of the American Medical Association* **250**, 2937–2941.
- Misra, M., Miller, K. K., Kuo, K., et al. (2005). Secretory dynamics of ghrelin in adolescent girls with anorexia nervosa and healthy adolescents. *American Journal of Physiology, Endocrinology and Metabolism* **289**, E347–E356.
- Morreale de Escobar, G., Obregon, M. J. and Escobar del Rey, F. (2000). Is neuropsychological development related to maternal hypothyroidism or to maternal hypothyroxinemia? *Journal of Clinical Endocrinology and Metabolism* **85**, 3975–3987.
- Munoz, M. T., Morande, G., Garcia-Centenera, J. A., et al. (2002). The effects of estrogen administration on bone mineral density in adolescents with anorexia nervosa. *European Journal of Endocrinology* **146**, 45–50.
- Murphy, D. L., Li, Q., Engel, S., et al. (2001). Genetic perspectives on the serotonin transporter. *Brain Research Bulletin* **56**, 487–494.
- Nachtigall, L. B., Boepple, P. A., Pralong, F. P., et al. (1997). Adult-onset idiopathic hypogonadotropic hypogonadism – a treatable form of male infertility. *New England Journal of Medicine* **336**, 410–415.
- Petraglia, F., Vale, W. and Rivier, C. (1986). Opioids act centrally to modulate stress-induced decrease in luteinizing hormone in the rat. *Endocrinology* **119**, 2445–2450.
- Petrides, J. S., Mueller, G. P., Kalogeras, K. T., et al. (1994). Exercise-induced activation of the hypothalamic-pituitary-adrenal axis: marked differences in the sensitivity to glucocorticoid suppression. *Journal of Clinical Endocrinology and Metabolism* **79**, 377–383.
- Pike, K. M. and Rodin, J. (1991). Mothers, daughters, and disordered eating. *Journal of Abnormal Psychology* **100**, 198–204.
- Prevot, V., Croix, D., Bouret, S., et al. (1999). Definitive evidence for the existence of morphological plasticity in the external zone of the median eminence during the rat estrous cycle: implication of neuro-glio-endothelial interactions in gonadotropin-releasing hormone release. *Neuroscience* **94**, 809–819.
- Quigley, M. E., Sheehan, K. L., Casper, R. F., et al. (1980). Evidence for increased dopaminergic and opioid activity in patients with hypothalamic hypogonadotropic amenorrhea. *Journal of Clinical Endocrinology and Metabolism* **50**, 949–954.
- Raff, H. and Bruder, E. D. (2006). Adiponectin and resistin in the neonatal rat: effects of dexamethasone and hypoxia. *Endocrine* **29**, 341–344.
- Reame, N. E., Sauder, S. E., Case, G. D., et al. (1985). Pulsatile gonadotropin secretion in women with hypothalamic amenorrhea: evidence that reduced frequency of gonadotropin-releasing hormone secretion is the mechanism of persistent anovulation. *Journal of Clinical Endocrinology and Metabolism* **61**, 851–858.
- Reindollar, R. H., Novak, M., Tho, S. P. T., et al. (1986). Adult-onset amenorrhea: a study of 262 patients. *American Journal of Obstetrics and Gynecology* **155**, 531–543.
- Romero, L. M., Plotsky, P. M. and Sapolsky, R. M. (1993). Patterns of adrenocorticotropin secretory release with hypoglycemia, novelty, and restraint after colchicine blockade of axonal transport. *Endocrinology* **132**, 199–204.
- Sapolsky, R. M. (2000). Glucocorticoids and hippocampal atrophy in neuropsychiatric disorders. *Archives of General Psychiatry* **57**, 925–935.
- Sapolsky, R. M. (2005). The influence of social hierarchy on primate health. *Science* **308**, 648–652.
- Sapolsky, R. M. and Krey, L. C. (1988). Stress-induced suppression of luteinizing hormone concentrations in wild baboons: role of opiates. *Journal of Clinical Endocrinology and Metabolism* **66**, 772–776.
- Schwartz, M. W. and Morton, G. J. (2002). Keeping hunger at bay. *Nature* **418**, 595–597.
- Shanan, J., Brezinski, A., Sulman, F., et al. (1965). Active coping behavior, anxiety, and cortisol steroid excretion in the prediction of transient amenorrhea. *Behavioral Sciences* **10**, 461–465.
- Spratt, D. I., Longcope, C., Cox, P. M., et al. (1993). Differential changes in serum concentrations of androgens and estrogens (in relation with cortisol) in postmenopausal women with acute illness. *Journal of Clinical Endocrinology and Metabolism* **76**, 1542–1547.
- Stein-Behrens, B., Mattson, M. P., Chang, I., et al. (1994). Stress exacerbates neuron loss and cytoskeletal pathology in the hippocampus. *Journal of Neuroscience* **14**, 5373–5380.
- Suh, B. Y., Liu, J. H., Berga, S. L., et al. (1988). Hypercortisolism in patients with functional hypothalamic amenorrhea. *Journal of Clinical Endocrinology and Metabolism* **66**, 733–739.
- Sullivan, S. D., Howard, L. C., Clayton, A. H., et al. (2002). Serotonergic activation rescues reproductive function in fasted mice: does serotonin mediate the metabolic effects of leptin on reproduction? *Biology of Reproduction* **66**, 1702–1706.
- Tanaka, M., Naruo, T., Yasuhara, D., et al. (2003). Fasting plasma ghrelin levels in subtypes of anorexia nervosa. *Psychoneuroendocrinology* **28**, 829–835.

- Terasawa, E. (1995). Control of luteinizing-hormone releasing hormone pulse generation in nonhuman primates. *Cellular and Molecular Neurobiology* 15, 141–164.
- Thind, K. K., Boggan, J. E. and Goldsmith, P. C. (1991). Interactions between vasopressin- and gonadotropin-releasing-hormone-containing neuroendocrine neurons in the monkey supraoptic nucleus. *Neuroendocrinology* 53, 287–297.
- Thind, K. K. and Goldsmith, P. C. (1989). Corticotropin-releasing factor neurons innervate dopamine neurons in the periventricular hypothalamus of juvenile macaques. *Neuroendocrinology* 50, 351–358.
- Thind, K. K. and Goldsmith, P. C. (1998). Infundibular gonadotropin-releasing hormone neurons are inhibited by direct opioid and autoregulatory synapses in juvenile monkeys. *Neuroendocrinology* 47, 203–216.
- Troisi, A., DiLorenzo, G., Lega, I., et al. (2005). Plasma ghrelin in anorexia, bulimia, and binge-eating disorder; relations with eating pattern and circulating concentrations of cortisol and thyroid hormones. *Neuroendocrinology* 81, 259–266.
- Van der Spuy, Z. M., Steer, P. J., McCusker, M., et al. (1988). Outcome of pregnancy in underweight women after spontaneous and induced ovulation. *British Medical Journal* 296, 962–965.
- Vigersky, R. A., Anderson, A. E., Thompson, R. H., et al. (1977). Hypothalamic dysfunction in secondary amenorrhea associated with simple weight loss. *New England Journal of Medicine* 297, 1141–1145.
- Vitale, M. L. and Chiochio, S. R. (1993). Serotonin, a neurotransmitter involved in regulation of luteinizing hormone release. *Endocrine Review* 14, 480–493.
- Wadhwa, P. D., Sandman, C. A. and Garite, T. J. (2001). The neurobiology of stress in human pregnancy: implications for prematurity and development of the fetal nervous system. *Progress in Brain Research* 133, 131–142.
- Warren, M. P. and Fried, J. L. (2001). Hypothalamic amenorrhea: the effects of environmental stresses on the reproductive system, a central effect of the central nervous system. *Endocrinology and Metabolism Clinics of North America* 30, 611–629.
- Warren, M. P., Miller, K. K., Olson, W. H., et al. (2005). Effects of an oral contraceptive (norgestimate/ethinyl estradiol) on bone mineral density in women with hypothalamic amenorrhea and osteopenia: an open-label extension of a double-blind, placebo-controlled study. *Contraception* 72, 206–211.
- Weber-Hamann, B., Kratzsch, J., Kopf, D., et al. (2006). Resistin and adiponectin in major depression: the association with free cortisol and effects of antidepressant treatment. *Journal of Psychiatric Research*, e-pub. ahead of print.
- Whitnall, M. H. (1989). Stress selectively activates the vasopressin-containing subset of corticotropin-releasing hormone neurons. *Neuroendocrinology* 50, 702–707.
- Williams, C. L., Nishihara, M., Thalabard, J. C., et al. (1990). Corticotropin-releasing factor and gonadotropin-releasing hormone pulse generator activity in the rhesus monkey. *Neuroendocrinology* 52, 133–137.
- Williams, N. L., Berga, S. L. and Cameron, J. L. (1997). Mild metabolic stress potentiates the suppressive effects of psychological stress on reproductive function in female cynomolgus monkeys. Paper presented at the 79th annual meeting of the Endocrine Society, Minneapolis, MN. Abstract #P1–367.

Stress Management and Cardiovascular Disease

D Lane

Sandwell and West Birmingham Hospitals NHS Trust,
Birmingham, UK

D Carroll

University of Birmingham, Birmingham, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition
article by D Lane and D Carroll, volume 3, pp 527–531,

© 2000, Elsevier Inc.

Stress Management Training

Essential Hypertension

Angina Pectoris

Myocardial Infarction

Conclusions

Glossary

*Angina
pectoris*

Chest pains, often radiating to the arms, shoulders, and jaw, arising from reduced blood flow to the heart as a result of coronary atherosclerosis.

Angiography

A technique for measuring the extent of coronary atherosclerosis.

Atherosclerosis

An accumulation of fatty deposits causing hardening and narrowing of the arteries.

*Cardiac
rehabilitation*

A combination of lifestyle education, exercise, and, increasingly, stress management training directed at patients suffering from coronary heart disease. The aim of cardiac rehabilitation is to facilitate physical, psychological, social, and emotional recovery.

<i>Essential hypertension</i>	Consistently high blood pressure without any apparent organic or medical cause.
<i>Meta-analysis</i>	Statistical techniques for combining and analyzing the outcomes of several studies.
<i>Myocardial infarction (MI)</i>	Heart attack.
<i>Stress management training (SMT)</i>	A range of techniques with the overall aim of training people to recognize and cope with psychological stress.
<i>Thrombosis</i>	Formation of a blood clot.

Stress Management Training

While stress management training (SMT) denotes a wide variety of practices concerned with improving individuals' recognition of and ability to cope with stress, most of the applications in cardiovascular disease include at least some of the following components: education about the nature of stress and its role in cardiovascular disease; identification of behavioral risk factors and counseling for lifestyle change; relaxation and its application during stress exposure; cognitive restructuring, in which positive and adaptive thoughts and beliefs are substituted for negative and maladaptive ones; and anger and hostility control when thought germane. Generally, SMT programs are undertaken with small groups of patients to maximize efficiency, trade on group motivational processes, and help generate social support.

Essential Hypertension

Essential hypertension is the presence of high blood pressure when there is no detectable medical or organic cause. The vast majority of instances of high blood pressure are of this sort. Hypertension is currently defined as a blood pressure of 140/90 mmHg or greater. Essential hypertension is a substantial problem. It is estimated that as much as 15% of the adult populations of Western countries exhibit high blood pressure, and various epidemiological studies testify that as blood pressure increases, life expectancy decreases, largely from an increased likelihood of stroke and myocardial infarction (MI). Given that psychological stress has been implicated in hypertension and that antihypertensive drugs often have undesirable side effects, SMT emerges as a potential adjunctive or even alternative therapy.

There have been more than 50 publications since the mid-1970s reporting attempts to reduce high blood pressure using some variant of SMT. Considered together, they paint a less than consistent picture. This is perhaps not surprising. First, although some

form of relaxation training characterizes most interventions in this context, other components of SMT are frequently missing. While generalizations are difficult, interventions that have included more SMT components have tended to prove more effective. Second, studies differ in design. Although many have random allocation to SMT and control, control conditions vary between no concerted treatment and carefully engineered attention placebo. SMT appears to fare better compared to minimal interventions than when compared to placebo or active alternative treatments. A meta-analysis by Eisenberg and colleagues concluded that SMT was "superior to no therapy but not superior to credible sham techniques or to self-monitoring alone." (Eisenberg et al., 1993: 964). Third, studies differ in terms of patients' medication status. Results are difficult to interpret when patients are taking antihypertensive medication, since reductions in blood pressure associated with SMT could primarily reflect increased adherence to antihypertensive medication regimens. That may not be a bad thing. Without intervention, hypertensives often show low adherence. Fourth, studies vary in terms of whether they have included follow-up blood pressure assessment, although those that have generally find that gains associated with SMT programs are, to a reasonable extent, maintained at follow-up. Finally, most studies have restricted their outcome measures to resting blood pressure; few have included ambulatory blood pressure assessments during everyday life. Those that have do not uniformly find an advantage for SMT over control treatments.

A serious methodological weakness in virtually all studies to date is the failure to establish a stable baseline blood pressure, by repeated blood pressure determinations, against which to compare the effects of intervention. The phenomenon of white coat hypertension serves to elevate some individual's blood pressure, particularly in clinical settings. The implication of this phenomenon and the failure to establish stable baselines are clear; they are likely to flatter the potential hypotensive effects of interventions. An exception to the usual trend of summary baseline assessment is the study by Johnston and co-workers. Patients were referred to the study by general practitioners if they registered diastolic blood pressure between 90 and 104 mmHg on at least two out of three occasions of measurement. None were taking antihypertensive medication. Prior to allocation to a SMT treatment or an active placebo condition, patients, following training, took their own blood pressures twice daily for 3 months. During this 3-month baseline, blood pressure dropped from referral to start of treatment from 152/98 mmHg to 140/93 mmHg for the SMT group and from 152/98 mmHg to

140/92 mmHg for the control group. Neither SMT nor the control treatment affected any further reductions in blood pressure. SMT and control patients did not differ in follow-up blood pressure or in blood pressure response to a stressful interview. A major implication of this carefully conducted study is that previously reported blood pressure reductions associated with SMT may have reflected adaptation effects. SMT might merely help the process of adaptation to measurement, lowering the blood pressure readings of individuals who are not well adapted to blood pressure measurement procedures. This would render SMT of doubtful clinical significance. Given the importance of the Johnston and colleagues' study, replication is critical. However, it is critical that future studies include prolonged baseline assessment to counter the potentially confounding effects of adaptation to measurement.

Angina Pectoris

Coronary heart disease (CHD) remains the most common cause of death in Western countries for both men and women. The failure of traditional risk factors to provide a complete account of the etiology and course of coronary heart disease has fuelled speculation that other, possibly psychological, factors are involved. Key among these is psychological stress. A frequent expression of coronary heart disease is angina pectoris, experienced as a severe pain in the chest often radiating to the arms, shoulders, and jaw. Angina is a reflection of atherosclerosis of the coronary arteries limiting the supply of circulatory oxygen to the myocardium and, in consequence, reducing the capacity of the heart to deal with increases in load, as, for example, during exercise and psychological stress.

It is therefore surprising that few studies have evaluated the effects of stress management on angina. A notable exception is the Lifestyle Heart Trial. Angina patients were randomly assigned to a lifestyle change intervention or a usual care control group. Among other things, the lifestyle change group patients undertook a fairly comprehensive SMT program, practicing for at least 1 h a day, using the hour-long audiocassette tape provided, for a year. Their data indicate very good compliance. The major outcome measure was the status of patients' coronary arteries as determined by angiography. While there was a worsening of coronary atherosclerosis in the usual care control group over the course of the year, there was evidence of regression of atherosclerosis in the lifestyle change group. However, it is difficult to attribute these impressive gains to SMT, since patients in the lifestyle change condition also switched to a low-fat vegetarian diet and undertook regular vigorous exercise.

Some evidence that SMT may be a key player in these effects comes from the work of Bundy and associates. In the first of two studies, angina patients were assigned either to a comprehensive SMT program lasting 7 weeks or to routine care. Compared to patients receiving routine care, those undergoing SMT reported briefer angina attacks and reduced reliance on medication and showed increased tolerance of exercise, that is, they were able to exercise for a longer period of time before ischemia or pain curtailed their exertions. In a second study, the effects of SMT *per se* were less impressive. Nevertheless, patients who underwent SMT in combination with light exercise fared particularly well. Compared with patients receiving only SMT or usual care, they showed greater exercise tolerance, and compared with patients receiving only exercise or usual care, they registered less frequent angina attacks and were less reliant on medication.

More recently, Lewin and colleagues evaluated the effects of a cognitive-behavioral management program, called the Angina Plan, on psychological adjustment in patients with newly diagnosed angina pectoris. Patients were randomized to receive either the self-help manual (Angina Plan) together with an audio-taped relaxation program or a nurse-led education session, identifying patients' risk factors for CHD, with additional written information. At the 6-month follow-up, patients receiving the Angina Plan demonstrated significantly greater reductions in anxiety and depression, diminished frequency of angina, and fewer physical limitations than patients receiving only the educational session. Although the Angina Plan encompasses SMT techniques, it also consists of a combination of exercise and lifestyle changes. Therefore, it is difficult to attribute the changes demonstrated solely to SMT. However, considered overall, the available data would seem to invite further study of the potential of SMT in controlling angina pectoris.

Myocardial Infarction

Heart attack, or MI, is the most common traumatic consequence of coronary artery disease. It results from blockages in the coronary arteries sufficiently severe to deprive the myocardium of oxygen and damage the muscle to an extent that it ceases to function. MI reflects both atherosclerosis, in which the coronary arteries become narrowed by the formation of fatty deposits or plaques, and thrombosis, the formation of blood clots in the arteries, which serve to fully occlude the already compromised blood flow. However, substantial numbers of patients survive heart attacks, and the widespread administration of

anticoagulant drugs in recent years has increased survival rates substantially. As a consequence, effective cardiac rehabilitation has become an increasingly important health services objective. The overall level of provision has increased markedly in the last 10 years in Western countries. For example, in the UK, surveys reveal that there are now some 300 cardiac rehabilitation programs and that around 30% of major hospitals offer programs. While cardiac rehabilitation was traditionally restricted to basic health education and supervised regular exercise, some form of SMT is increasingly part of the package.

There are sound *prime facie* reasons for including SMT in this context. First, a heart attack is an exceedingly distressing experience for most patients, and levels of depression and anxiety following MI are frequently observed. Second, depressed and anxious cardiac patients are less likely to undertake lifestyle changes, adhere to medication, and exercise, and they also experience greater difficulties in coping with stress. Third, recent evidence indicates that post-MI levels of distress predict subsequent cardiac morbidity and mortality in some, but not all, studies. Fourth, psychological distress appears to be the most important predictor of hospitalization costs following a cardiac event, with distressed cardiac patients accruing four times the cost of nondistressed patients for nonpsychiatric medical interventions.

Early systematic evaluations of cardiac rehabilitation were restricted to programs that were largely exercise based and did not include SMT. Meta-analyses suggest that while such programs had no effects on the recurrence of nonfatal MIs, they were associated with reductions in mortality of 20 to 25%. However, since the studies included in these meta-analyses were generally undertaken with patients who were at low risk for recurrence of MI, the scope for demonstrating benefit was perhaps limited.

Most cardiac rehabilitation programs have included some form of SMT, although, as we shall see, in some instances the provision contains few, if any, of the customary elements of SMT. These have also been subjected to meta-analytic review. Linden and colleagues identified 23 randomized control trials of cardiac rehabilitation programs that contained at least some of the constituents of SMT. In total, these studies afforded data on 2024 patients who had received SMT and 1156 control patients who, for the most part, had received only usual care. The patients given SMT exhibited more substantial reductions in psychological distress, systolic blood pressure, and cholesterol than the controls. Critically, relative to the control condition, SMT was associated with a 40% greater chance of survival during the first 2 years following MI, although not subsequently.

The risk of recurrence of MI was also associated with a substantially reduced risk of MI recurrence.

Since this meta-analysis, four randomized control trials with substantial patient numbers have been published. All four were concerned with alleviating patients' distress following MI. In the Montreal Heart Attack Readjustment Trial (M-Hart), patients' distress was monitored by monthly telephone calls for a year. Those judged to be distressed received a home visit from a nurse who offered emotional support, education, or referral to a general practitioner or cardiologist, as deemed appropriate. Patients in the program were compared to those receiving usual care. The intervention had only minimal effect on distress and no impact on overall survival at 1 year. Indeed, for female patients, cardiac and all-cause mortality tended to be higher in the intervention group. A similar sort of intervention strategy was used by Barr-Taylor and associates and again was compared to usual care. Nurses assessed patients' distress by telephone, and those registering high levels of distress were referred to mental health specialists, although no data were available on how many patients were referred and the nature of any treatment they received. While patients' distress lessened over the 1-year follow-up, there were no group differences. There were insufficient deaths to permit statistical comparison, although the deaths that occurred were split fairly evenly between groups. These two studies have been subject to a number of criticisms. Most importantly, the interventions contained little systematic SMT and, indeed, are barely recognizable as modern cardiac rehabilitation programs. It is likely that they were simply too minimal to influence cardiac health status. Predicated as they were on the reasonable presumption that alleviating patients' distress would yield dividends for recurrence and mortality, the failure of these studies to differentially ameliorate distress makes the absence of significant cardiac health effects unsurprising.

The third SMT intervention also failed to counter patients' distress following MI, again making the absence of effects on cardiac morbidity and mortality entirely predictable. However, the failure of the intervention to alleviate distress in this instance is more difficult to understand. Jones and West randomly allocated 2328 MI patients either to a reasonably comprehensive SMT program or to routine care. The SMT program comprised anxiety reduction techniques, relaxation training, instruction in coping skills, and attempts to promote a more positive adjustment to illness. As indicated previously, there were no differences between groups at 1-year follow-up either in reported distress or in clinical complications and mortality. There are two possible explanations for these

negative findings. First, many patients in the usual care group had good outcomes, which would serve to minimize any advantage for SMT. Second, it is clear that the intervention was not sufficiently powerful to affect psychological change in the most distressed patients. This study suggests, however, that routine psychological screening and intervention may be a useful first step but that more intensive SMT would be required for patients exhibiting extreme emotional distress.

The final randomized controlled trial, Enhancing Recovery in Coronary Heart Disease (ENRICHD), enrolled 2481 patients who had major or minor depression after their MI to receive either cognitive-behavioral therapy (CBT; $n = 1238$), with the option of adjunctive antidepressant medication ($n = 249$), or usual care ($n = 1243$). The main outcome measures were the assessment of depressive symptoms at 6 months, recurrent MI, and all-cause mortality. The intervention consisted of 6 months or less of both individual and group CBT, with variable numbers of sessions per patient. The intervention reduced depression at 6 months but had no significant effect on re-infarction or all-cause mortality, assessed after a mean follow-up of 41 months.

A recent Cochrane review examined the impact of nonpharmacological psychological interventions for adults with CHD (MI, coronary artery bypass graft surgery, percutaneous transluminal coronary angioplasty, angina, and angiographically defined CHD) on all-cause and CHD-related mortality, recurrent cardiac events, and symptoms of anxiety, depression, and stress, compared with adults receiving usual care or no intervention. Thirty-six trials were identified, 18 specifically examining SMT interventions (defined as the use of specific cognitive-behavioral strategies used to help the patient reduce or manage their stress). Of these 18 SMT trials, 11 reported on SMT interventions plus other rehabilitation interventions and seven reported on SMT interventions alone.

Rees and colleagues' systematic review concluded that there was no evidence of any benefit from SMT interventions in reducing all-cause or cardiac mortality. Further, the effects on nonfatal re-infarction were inconclusive. This conclusion was mainly influenced by two large negative randomized controlled trials, one of SMT and the other of CBT. However, trials with a SMT component did demonstrate a reduction in anxiety and depression following intervention. Nevertheless, caution is warranted in interpreting this result, given that it is based on small trials that included a combination of interventions. Accordingly, it is difficult to attribute with certainty such anxiolytic and antidepressive effects to SMT.

Conclusions

Stress management training has now been applied to a number of cardiovascular problems. The largest amount of research attention has been directed toward patients following MI. The data are extremely mixed, and two large trials have yielded disappointing results in terms of cardiological outcomes. Nevertheless, it appears that mixed programs of SMT and exercise rehabilitation produce psychological benefits comparable to those reported for several pharmacological interventions. Further research is required to identify the most effective intervention strategies. It is likely that the greatest dividends will accrue from individually tailored interventions, directing SMT to those patients most in need. Although angina has attracted less study, early indications are that SMT, particularly when combined with exercise, can produce improvements in symptoms and reduce reliance on medication and may even stem atherosclerotic deterioration of the coronary arteries. Finally, the indications for SMT in hypertension are far from compelling. It is possible that SMT acts simply to increase the rate of adaptation to repeated blood pressure measurement.

See Also the Following Articles

Cardiovascular System and Stress; Cognitive Behavioral Therapy; Depression and Coronary Heart Disease; Heart Disease/Attack; Cardiovascular Disease, Stress and; Stress Management, CAM Approach.

Further Reading

- Barr-Taylor, C. B., Houston-Miller, N., Smith, P. M., et al. (1997). The effects of a home-based, case-managed, multifactorial risk-reduction program on reducing psychological distress in patients with cardiovascular disease. *Journal of Cardiopulmonary Rehabilitation* 17, 157-162.
- Bennett, P. and Carroll, D. (1994). Cognitive-behavioural interventions in cardiac rehabilitation. *Journal of Psychosomatic Research* 38, 169-182.
- Berkman, L. F., Blumenthal, J., Burg, M., et al. (2003). Effects of treating depression and low perceived social support on clinical events after myocardial infarction: the Enhancing Recovery in Coronary Heart Disease Patients (ENRICHD) Randomized Trial. *Journal of the American Medical Association* 289, 3106-3116.
- Bundy, C., Carroll, D., Wallace, L., et al. (1998). Stress management and exercise training in chronic stable angina pectoris. *Psychology and Health* 13, 147-155.
- Eisenberg, D. M., Delbanco, T. L., Berkey, C. S., et al. (1993). Cognitive behavioural techniques for hypertension: are they effective? *Annals of Internal Medicine* 118, 964-972.

- Frasure-Smith, N., Lesperance, F., Prince, R. H., et al. (1997). Randomised trial of home-based psychosocial nursing intervention for patients recovering from myocardial infarction. *Lancet* 350, 473–479.
- Johnston, D. W., Gold, A., Kentish, J., et al. (1993). Effect of stress management on blood pressure in mild primary hypertension. *British Medical Journal* 306, 963–966.
- Jones, D. and West, R. (1995). *Cardiac rehabilitation*. London: BMJ Publishing Group.
- Jones, D. A. and West, R. R. (1996). Psychological rehabilitation after myocardial infarction: Multicentre randomised controlled trial. *British Medical Journal* 313, 1517–1521.
- Lewin, R. J. P., Furze, G., Robinson, J., Griffith, K., Wiseman, S., Pye, M. and Boyle, R. (2002). A randomised controlled trial of a self-management plan for patients with newly diagnosed angina. *British Journal of General Practice* 52, 194–201.
- Linden, W., Stossel, C. and Maurice, J. (1996). Psychosocial interventions for patients with coronary artery disease. *Archives of Internal Medicine* 156, 745–752.
- Ornish, D., Brown, S. E., Scherwitz, L. W., et al. (1990). Can lifestyle changes reverse coronary heart disease? The Lifestyle Trial. *Lancet* 336, 129–133.
- Rees, K., Bennett, P., West, R., et al. (2004). Psychological interventions for coronary heart disease. *The Cochrane Database of Systematic Reviews* 3, CD002902.pub2.

Stress Management, CAM Approach

K Krebs

Allegheny General Hospital, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 532–537, © 2000, Elsevier Inc.

Mind/Body Medicine Strategies
 Bioenergy Medicine Strategies
 Manual Healing Strategies
 Nutrition, Botanical Medicine and Nutritional
 Supplements

Glossary

Complementary and alternative medicine (CAM)

The use and practice of therapies and diagnostic techniques that may not be part of any current Western health-care system, culture, or society. The emphasis is on prevention, treatment, or both, using the self-healing capacity within a mind/body approach (in certain but not all cases).

Mind/Body Medicine Strategies

Mind/body medicine is an umbrella term that covers activities and therapies that focus on the self-healing potential within us through the interrelationship of mind and body. The goal in mind/body medicine is to optimize health by combining mental, physical, emotional, and spiritual approaches, which in turn influence the hormonal, nervous, and immune systems.

Yoga, for example, involves both physical movement and a meditative state of mind and may serve to enhance a person's physical stamina and endurance while diminishing the effects of stress and perhaps strengthening one's personal relationship with God or a higher power.

Meditation and Yoga

Meditation can have significant effects on cardiovascular health. In a study conducted at an inner-city health care center, older people with mild hypertension were divided into three groups: those who learned transcendental meditation (TM), a concentration-meditation practice that involves focusing on a single object such as a sound, a word, or short phrase; those who participated in a progressive muscle-relaxation program; and those who were educated in lifestyle changes. At 3 months follow-up, the TM and muscle relaxation groups fared the best in terms of lowered blood pressure, with TM proving to be twice as effective as muscle relaxation.

Using the relaxation response, a simple meditative technique in which the participant sits quietly, focusing on breathing or a word, Harvard's Mind/Body Medical Institute researchers found that 80% of hypertensive patients had lowered blood pressure and decreased medication needs and that 16% were able to discontinue all their antihypertensive medications. These results persisted throughout the 3-year follow-up period. French researchers tested 50 hypertensive people as they sat silently, read, or talked about stress in their lives. Those who talked about stress experienced a sharp increase in systolic blood pressure, an average of 17 mm, an indication that substantiates

further how the mind (thoughts, feelings, and perceptions) can influence our physiology.

Numerous studies have shown that stress is a major trigger for flare-ups of psoriasis, and a study by British researchers suggested that the onset of psoriasis may be influenced by stress as well. Of 64 people with chronic skin problems, including psoriasis, 44 of the participants had experienced a major life event, such as a serious illness or a death in the family, soon before or at about the time their skin eruptions began. A small Johns Hopkins study found that some psoriasis patients who underwent hypnosis and visualized their condition clearing up on its own enjoyed a complete remission of symptoms. In a study of 51 chronic pain sufferers, ranging from back pain to chest pain (angina), a form of mindfulness meditation was used in a 10-week program. Sixty-five percent of patients reported a significant reduction in their perceived pain and 50% said their pain was reduced by more than half. Significant reductions in mood disturbances and psychiatric symptoms accompanied the reduction in pain as well.

Harvard's Mind/Body Medical Institute found that chronic pain patients reduced their physician visits by 36% following implementation of the regular practice of the relaxation response. Equally impressive is that this same technique has been credited for 75% of insomnia patients becoming normal sleepers and 90% having reduced or entirely eliminated sleeping medications.

Guided Imagery

Guided imagery teaches people to imagine scenarios that may help influence certain physiological conditions. Headache sufferers have successfully utilized their imaginations to release the vise grip constraining their face and scalp. Cancer patients have imagined their immune system partnering with their chemotherapy to escort out malignant cells through the lymphatic system. Several pilot studies have demonstrated elevated natural killer cells and elevated levels of serum and salivary immunoglobulin levels following guided imagery interventions. Sufferers of irritable bowel syndrome are taught to envision their digestive tract as healthy and rhythmically producing normal peristaltic contractions to create a harmony between mind and gut and normal evacuation. The University of Arizona's Integrative Medicine/Children's Research Center studies complementary and alternative medicine (CAM) methods of pediatric disorders for which there seems to be no conventional medical therapies available. One study tests the use of relaxation, guided imagery, and chamomile tea in children with recurrent abdominal pain.

Prayer and Spirituality

Prayer and spirituality are recognized CAM therapies by the NIH and funded studies are underway. The University of Michigan is studying the effects of spirituality and religiosity following cardiac surgery. A review of the psychiatric literature revealed 87% of the research that has examined religious involvement as a variable reports a positive effect, with the remaining 13% reporting a negative impact on mental health.

More than 130 controlled laboratory studies show that prayer, or a prayer-like state of compassion, empathy, and love, can bring about healthful changes in many types of living things, from humans to bacteria. This does not mean that prayer always works, any more than drugs and surgery always work, but that, statistically speaking, prayer is effective. Faith helps mobilize a person's defenses and assists in getting well, and optimism generally leads to better outcomes. Many case histories and scientific studies affirm this observation. Oxman and colleagues investigated the role religious feeling and activity played in 232 patients over the age of 55 undergoing cardiac surgery. Those who derived at least some strength, comfort, and hope from religion were more likely to survive longer after cardiac surgery than those who did not. Numerous studies in humans show that we can die as a result of dire beliefs and a sense of overwhelming futility.

Aromatherapy

Aromatherapy is the use of pure essential oils to promote health. Smell is one of the most basic human senses that can evoke a variety of emotions. Aromatherapy can be used in baths, as a diffusion or inhalation, or in massage therapy. Relieving anxiety, reducing stress, and promoting relaxation are perhaps the most well known of the uses for aromatic essential oils. Chamomile is best known for its calming properties. Lemon balm and orange blossom can have profound effects on the reduction of depression and anxiety. Lavender is said to promote a peaceful mood and create a sense of calm. Research around the globe is validating aromatherapy as a medicinal modality. A double-blind, placebo-controlled study in Germany found that oil of peppermint applied to the temples reduced tension and headache pain, perhaps by increasing blood flow to the skin. Aromatherapists believe that the fragrance of oils has a soothing effect on the brain's limbic system, which is involved in memory, emotion, and hormone regulation. Substances that affect mood also affect brain chemistry and the immune system. For example, the essential oil of ylang-ylang, a tropical flower, has been shown to cause the pituitary to

secrete more euphoric endorphins. The scent of oil of marjoram boosts the production of calming serotonin, and oil of grapefruit stimulates the brain to produce natural painkillers called enkephalins. However, critics of aromatherapy suggest that the relaxation attributed to the oils may actually be due to their application by massage, hot baths, the human touch, and other pleasant methods.

Bioenergy Medicine Strategies

This category of complementary modalities is based on the concept that the human energy field is subject to imbalances and that restoration of the energy fields improves health and sleep and promotes relaxation.

Acupuncture

In November 1997, a National Institute of Health panel concluded that acupuncture is an acceptable alternative or adjunct to conventional therapy for headaches, low back pain, fibromyalgia, asthma, and carpal tunnel syndrome. The panel stressed that acupuncture has minimal side effects compared to conventional treatments and noted considerable evidence that acupuncture causes the release of natural pain-relieving substances (endorphins) in the brain. In 1993, an NIH-funded study evaluated the use of acupuncture in the treatment of unipolar depression in 33 women. Based on the results of the small sample, acupuncture appeared to offer significant symptom relief comparable to conventional interventions such as psychotherapy or pharmacotherapy. The use of acupuncture as an effective treatment for depression was substantial enough to warrant a larger-scale clinical trial. Beginning in 1999, the Minneapolis Center for Addiction and Alternative Medicine began exploring the use of electroacupuncture as it pertains to substance abuse, beginning with nicotine and progressing to such substances as heroin and cocaine. Similarly, acupuncture research is being conducted at the University of Arizona. The study will evaluate the use of self-hypnosis, acupuncture, and osteopathic manipulations for children with spastic cerebral palsy.

Therapeutic Touch

Therapeutic touch (TT) is a body energy-field technique in which hands are passed over the body without actually touching to recreate and change energy imbalances for restoring innate healing forces. One of the first studies to be funded by the Office of Alternative Medicine at the NIH monitored the effects of TT on 20 very stressed nursing and medical students facing professional board exams. Students were randomized into two groups: one receiving TT and one receiving no TT. Differences between groups were found to be

significant for IgA, IgM, and apoptosis. Theory building is still needed to explain results that show changes in the immune response to stress-reducing therapies. This study was determined useful in documenting that measurable changes do occur following TT. In a 1986 study on 60 patients with tension headaches, patients who received TT reported an average pain reduction of 70%. The control group receiving mock TT reported an average pain reduction of 37%. Open-heart surgery patients who received TT had lower blood pressure than those who did not. While TT has gained its share of skeptics, the technique is now practiced by an estimated 30 000 health professionals around the world, is offered in at least 200 U.S. hospitals, and is taught in more than 80 universities.

Reiki

Reiki comes from the Japanese word meaning universal life force energy. The practitioner serves as a conduit for healing energy directed into the body or energy field of the recipient with or without physical contact with the body. The vast majority of reports on the effectiveness of reiki appear in popular rather than scientific literature; however, one study showed that reiki can increase hemoglobin and hematocrit levels. The study compared 47 people participating in reiki training and a small control group of 9 healthy medical professionals. The study found a significant increase in the hemoglobin and hematocrit levels among the reiki group and no significant change in the control group.

Manual Healing Strategies

Manual healing strategies is the term used for many body work techniques from both ancient and modern traditions that promote relaxation and treat ailments through reeducation in posture and movement and various forms of bodily manipulations. Acupressure, osteopathy, reflexology, and rolfing, a deep tissue massage used to release fascial adhesions, are just a few of the manual healing strategies that can aid in managing stress-related symptoms.

Massage

Massage therapy includes an array of manual therapies that manipulate the soft tissues of the body in order to reduce anxiety, tension, stress, and depression. In addition, massage therapy increases circulation, controls pain, and promotes a sense of overall well-being. Touch is the main ingredient of massage therapy, which conveys a sense of caring – an important component in the healing partnership. A study of the use of massage in patients with pain from cancer found that those who received massage had a 60% lower level of pain perception and a 24% decrease in

anxiety. More than half reported feeling more relaxed as well. In a study of children and adolescents hospitalized with depression and adjustment disorder, those who were treated with massage were found to be less anxious and depressed than those who viewed relaxation videotapes.

Nutrition, Botanical Medicine and Nutritional Supplements

A healthy diet can help the body's healing system operate more efficiently, as has been demonstrated in a number of clinical research studies. In the 1970s, Dr. Dean Ornish, a medical internist, took the medical community by storm when he suggested that a vegetarian diet (10% fat), combined with yoga, meditation, moderate exercise, and group support, could reverse heart disease. Now clinical studies have demonstrated that once-occluded coronary vessels are functioning optimally, and people subjectively report a return to a healthier, happier lifestyle free from anginal pain. Many third-party payers now cover participation in this lifestyle program to attempt to avoid the more costly coronary bypass surgery and coronary angioplasty. Comparisons of a Western and Eastern lifestyle program in healthy adult subjects showed that those adhering to an Ayurvedic (Eastern) lifestyle similar to the Dean Ornish program visited their physicians less frequently, used fewer prescription medications, and had less episodes of depression than the Western lifestyle group.

Sales of soy-based products such as soy milk, tofu, and tempeh have skyrocketed as a result of reports of their protective effects against breast and prostate cancer and their ability to minimize perimenopausal symptoms. Organic produce rich in natural antioxidants are said to have a positive effect on the nervous system, thus aiding in stress reduction. Coffee substitutes and herbal teas are replacing caffeinated beverages, which contribute to nervousness, anxiety, palpitations, and insomnia. Botanical medicine sales aimed at stress reduction have soared amid claims of safety and less side effects than their pharmaceutical counterparts. The German Commission E, equivalent in many ways to the Food and Drug Administration in the United States, has conducted clinical trials on several botanical medicines for anxiety and mild depression. Kava is derived from the knotty root of a large tropical shrub in the black pepper family. This herb has been widely used in Europe for nervous anxiety, stress, and sleep aid.

Research has demonstrated that kavalactones are responsible for the induced relaxation that acts on the central nervous system. In 1996 in Germany, a double-blind randomized trial investigated the effects

of kava on 58 adults with anxiety. Half the subjects took 100 mg of kava extract (standardized to 70 mg kavalactones) three times per day for a month; the other group took a placebo. After 1 week the kava group showed significant improvement in symptoms compared with the placebo group and continued to improve over the course of the research study. Kava is recommended for short-term use and should not be taken with alcohol, prescription sedatives, or valerian root, a herbal sedative. *Hypericum perforatum*, better known as St. John's wort, is a shrubby perennial plant that, when used in a standardized extract containing the main active component hypericin, improves mood and sleep in those suffering from mild to moderate depression. German researchers are currently studying hypericin, another component of St. John's wort thought to be superior in alleviating mild to moderate depression. The recommended dose is 300 mg three times a day and usually requires at least a month of regular use to achieve maximum symptom management. Following an animal study suggesting that St. John's wort may inhibit alcohol cravings by influencing levels of serotonin and other neurotransmitters, researchers at the University of North Carolina at Chapel Hill are planning a study using the herb on human alcoholics.

Insomnia is one of the most frustrating symptoms associated with stress. Valerian (*Valeriana officinalis*) is a European plant that eases stress and mildly sedates the central nervous system. Valerian binds to the same receptors in the brain as prescription sleeping aids. Even though it is mild compared to pharmaceutical sedatives, valerian is a depressant and should be reserved for occasional use rather than taken nightly. Alcohol and other depressants should be avoided when using valerian root. Chamomile tea is one of the most effective bedtime beverages. This herb has a sedative effect that can be used by all ages. In addition to easing away tensions to help with sleep, chamomile soothes indigestion or heartburn and, for infants, is a helpful remedy for colic.

In addition, there are supplements that aid naturally in the brightening one's mood. Dr. Andrew Weil, from the University of Arizona's Integrative Medicine Program, recommends a simple formula intended to change the balance of neurochemicals in the brain to support energy and mood elevation. In the morning, upon awakening, the amino acid DL-phenylalanine (DLPA) is taken. If sufficient mood elevation has not been achieved after trying DLPA, then L-tyrosine, another amino acid, can be substituted. One may switch back and forth between these amino acids. In addition, vitamins C and B₆, a glass of fruit juice, and a small piece of fruit should be consumed and breakfast eaten 45 min later. Vitamins C and B₆ are

repeated in the evening. These natural supplements appear to be safe to take, even if one is currently on selective serotonin reuptake inhibitor antidepressants. Dr. Weil asserts that aerobic activity is actually the best antidepressant because it increases the production of endorphins, the brain's own opiate-like molecules that are associated with a natural high.

See Also the Following Articles

Integrative Medicine (Complementary and Alternative Medicine); Hypnosis; Relaxation Techniques; Religion and Stress.

Further Reading

- Benson, H. (1975). *The relaxation response*. New York: Avon Publisher.
- Benson, H. (1996). *Timeless healing: the power and biology of belief*. New York: Scribner.
- Chevallier, A. (1996). *The encyclopedia of medicinal plants*. London: Dorling Kindersley.
- Dossey, L. (1993). *Healing words and prayer is good medicine*. New York: Harper.
- Duke, J. (1997). *The herbal pharmacy*. Rodale Press.
- Ferrell-Torry, A. T. and Glick, O. J. (1993). The use of therapeutic massage as a nursing intervention to modify anxiety and the perception of cancer pain. *Cancer Nursing* 16, 93–101.
- Kabat-Zinn, J. (1982). An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation. *General Hospital Psychiatry* 4, 33–47.
- Keller, E. and Bzdek, V. M. (1986). Effects of therapeutic touch on tension headache pain. *Nursing Research* 35, 101–105.
- Larson, D. (1996). *The role of prayer in mental health*. Alternative Therapies.
- Leake, R. and Broderick, J. (1998). *Treatment efficacy of acupuncture: a review of the research literature*. *Integrative Medicine* 1 (Issue 3).
- National Institutes of Health (NIH) (1994). *Alternative medicine: expanding medical horizons, a Report to the National Institute of Health on Alternative Medical Systems and Practices in the United States*. p. 143. NIH Publication No. 94–066.
- Olson, M. (1993). *Immune response to therapeutic touch during stress*. NIH: Alternative Medicine Grant.
- Ornish, D. (1998). *Love and survival: the scientific basis for the healing power of intimacy*. New York: Harper Collins.
- Oxman, T. E., Freeman, D. H., Jr. and Manheimer, E. D. (1995). Lack of social participation or religious strength and comfort as risk factors for death after cardiac surgery in the elderly. *Psychosomatic Medicine* 57, 5–15.
- Schneider, R. H., et al. (1995). A randomized controlled trial of stress reduction for hypertension in older African Americans. *Hypertension* 17, 192–200.
- Simon, D. (1993). *Ayurvedic Medicine Demonstration Project*. NIH: Alternative Medicine Grant.
- Spencer, J., et al. (1999). *Complementary/Alternative medicine: an evidence-based approach*. New York: Mosby.
- Weil, A. (1995). *Health and healing*. New York: Houghton Mifflin.
- Weil, A. (1995). *Natural health, natural medicine*. New York: Houghton Mifflin.
- Weil, A. (1995). *Spontaneous healing: how to discover and enhance your body's natural ability to maintain and heal itself*. New York: Knopf.
- Weil, A. (1996). *Self healing: creating natural health for your body and mind*. New York: Houghton Mifflin.
- Wetzel, W. (1989). Reiki healing: a physiologic perspective. *Journal of Holistic Nursing* 7(1).

Stress of Self Esteem

J C Pruessner, S Wuethrich and M W Baldwin
McGill University, Montreal, Quebec, Canada

© 2007 Elsevier Inc. All rights reserved.

The Role of Personality in the Perception of Stress

The Importance of Self-esteem and Locus of Control in the Perception of Stress

Endocrinological Evidence for the Role of Self-esteem and Locus of Control in the Perception of Stress

The Hippocampus as a Possible Mediator of the Relationship between Self-esteem, Locus of Control, and Stress

Conclusion: Personality Variables, Brain Structures, and Stress Responses in a Developmental Context

Glossary

Cortisol

A steroid hormone secreted by the human adrenal glands. It is often released in high amounts during periods of stress. High doses lead to interference with the proper functioning of the immune system.

Hippocampus

Seahorse-shaped structure that extends along the inferior horn of each lateral ventricle of the brain and consists of gray matter covered on the ventricular

	surface with white matter. A part of the limbic system, it is involved in the storage of memory for intermediate periods and the consolidation of those memories into permanent form.
<i>Hypothalamic-pituitary-adrenal (HPA) axis</i>	The major part of the neuroendocrine system that controls reactions to stress and has important functions in regulating various body processes such as digestion, the immune system, and energy usage. Most species (from humans to simple organisms) share components of the HPA axis. It regulates a set of interactions among glands, hormones, and parts of the mid-brain that mediate the general adaptation syndrome proposed by Selye.
<i>Locus of control</i>	The perception of factors responsible for the outcome of an event. An individual with high internal locus of control believes that his or her own actions caused the outcome. Conversely, an individual with high external locus of control believes the outcome was determined by outside forces.
<i>Personality</i>	The complex of all attributes – behavioral, temperamental, emotional, and mental – that characterize an individual.
<i>Self-esteem</i>	The global value we place on ourselves.
<i>Stressor</i>	Any emotional or physical demand (positive or negative) that gives or results in stress.

The Role of Personality in the Perception of Stress

Lazarus, who has devoted much of his work to the study of stress, noted early on that stressful conditions per se fail to produce reliable effects on task performance. Keeping all situational variables constant, the same stressor can have minimal effect, lead to performance improvement, or result in significant performance impairment across subjects. This led him and others to suggest that individual differences in motivational and cognitive variables are likely to interact with situational components, and what is considered threatening for some, and thus stressful and performance impairing, might be considered as stimulating by others, and thus produce beneficial effects on performance.

Subsequent research identified the process of evaluation as crucial in explaining the impact of psychological variables on the stress response. Any internal or external stimulus is perceived as stressful only if it is evaluated as harmful or threatening. The early notion of the importance of the evaluation has received recent validation with the identification of social evaluative threat as the single most important

factor in determining the stressfulness of a situation in laboratory stress studies.

The Importance of Self-esteem and Locus of Control in the Perception of Stress

A case can be made that the personality variables self-esteem and locus of control play a central role in the evaluation of many situations and thus contribute to the experience of stress. Self-esteem is broadly defined as the value we place on ourselves. Epidemiological studies have shown that low self-esteem is associated with negative life outcomes, including substance abuse, delinquency, unhappiness, depression, and worsened recovery after illnesses. On the other hand, high self-esteem has been linked to happiness and longevity. In studies of aging, a positive self-concept has been identified to play a key role in successful aging, predicting independence, cognitive stability, and general health in old age.

Internal locus of control can be defined as the perception that one's outcomes are determined by one's actions. Not surprisingly, internal locus of control and self-esteem are usually highly correlated. The key link of these variables to the experience of stress lies in their impact on the evaluation of any given situation. We postulate that in the evaluation of whether a given situation is threatening or benign, self-esteem and internal locus of control systematically interact with situational factors. If a person attributes little importance to him- or herself and thinks that he or she has little impact on the outcome of his or her own actions, this person will find more situations uncontrollable and unpredictable, and consequently will evaluate more situations as threatening and harmful.

Endocrinological Evidence for the Role of Self-esteem and Locus of Control in the Perception of Stress

Evidence for the impact of self-esteem and locus of control on stress perception emerged when subjects were exposed to repeated psychological stress, using the Trier Psychosocial Stress Test (TSST). In this paradigm, subjects have to give an impromptu speech and perform serial subtraction tasks in front of an audience, usually for about 10 min. The audience consists of two to three persons who are instructed to maintain a neutral expression, being neither explicitly rejecting nor confirmative in their facial expression or gestures. During the speech, the audience interacts with the subject only to indicate the amount of time that is left to talk or to ask specific questions.

In the case that a subject stumbles, they encourage the subject to continue the speech. During the serial subtraction task, the subject is interrupted only when making a mistake. The subject is then corrected and instructed to start the task over. The task was designed to represent a significant social-evaluative threat and indeed has been shown to be a powerful stressor, stimulating the hypothalamic-pituitary-adrenal (HPA) axis and leading to significant free cortisol increases within 15 to 30 min following the onset of the task. This first study aimed to validate the long-standing hypothesis that in humans, repeated exposure to the same stressor would lead to quick habituation of the stress response. In order to test the habituation of the stress response, 20 young healthy male college students were exposed to the TSST on five subsequent days. For this purpose, the TSST was modified using different speech topics and serial subtraction tasks on each day. Interestingly, only 13 of the 20 subjects showed the typical habituation pattern, with a normal stress response on day one being significantly reduced on day two, and no longer present on the subsequent days. In the seven remaining subjects, however, the cortisol stress response continued to be present on all days and only showed a tendency to decline toward the end of the testing (Figure 1). When analyzing the available psychological variables, it became apparent that low internal locus of control and low self-esteem were the best predictors of failing habituation of the cortisol stress response to repeated stress exposure. This can be interpreted as a sign that these personality variables interact with the evaluation of a situation during repeated exposure. The absence of differences in the stress response between the two groups of subjects on day one was at the time attributed to the effect of

novelty – the novelty of the situation might have made it unpredictable and uncontrollable for everybody on the first exposure and might thus have masked the impact of personality variables on stress perception and response. One conclusion at the time was that in order to reveal the effect of personality variables on the stress evaluation and response, one would likely need repeated exposures to the same stressor in order to reveal the influence of personality variables on stress.

However, it is known that personality variables tend to have relatively weaker effects when situational factors are very strong. In a second study, the threatening aspects of the situation were reduced, and self-esteem and locus of control had an impact on the perception of stress on the first exposure to a stimulus. Here, computerized mental arithmetic was combined with an induced failure design to invoke stress. In the setup used in this task, 52 students performed the task on computer terminals in front of them. Half of the students were exposed to a difficult version leading to low performance, compared to an easy version of the task with high performance for the other half. The students played the task in three 3-min segments and had to announce their performance score after each segment to the investigator, who wrote the scores down on a board for everybody to see. Saliva sampling before, throughout, and after the task allowed the assessment of the cortisol dynamics in relation to this paradigm. Interestingly, this task triggered a significant cortisol release only in the subjects who were in the low-performance group and had low self-esteem and low internal locus of control. Neither low performance alone nor low self-esteem and internal locus of control alone were significant predictors of cortisol

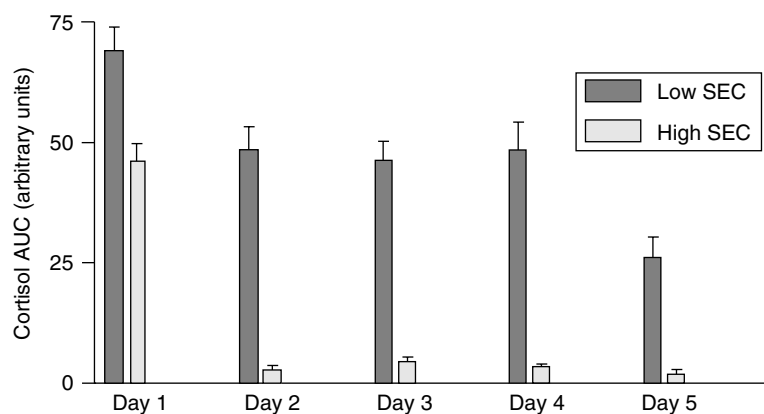


Figure 1 Cortisol responses (AUC, area under the curve) on repeated exposure to the Trier Social Stress Test (TSST) on 5 subsequent days in subjects with high self-esteem and high locus of control (high SEC; $n = 13$) and low self-esteem and low locus of control (low SEC; $n = 7$).

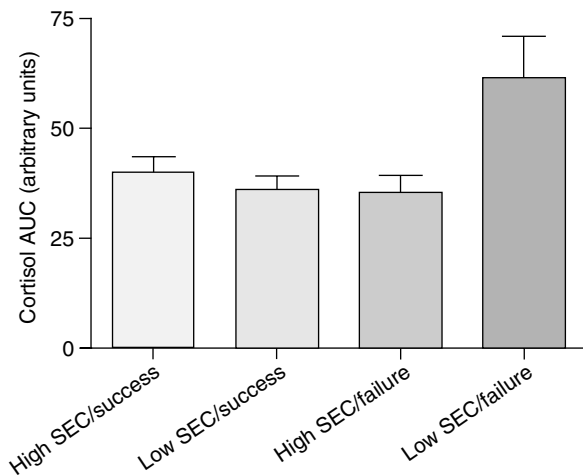


Figure 2 Cortisol stress responses to the Trier Mental Challenge Task (TMCT) in four groups of subjects, separated for high and low self-esteem and locus of control and high and low performance in the mental arithmetic. The performance was manipulated by the investigator.

release, supporting the notion that these personality variables produce effects only in interaction with a potentially stressful situation (Figure 2). The evaluation of the situation is suggested to be at the core of this interaction.

The Hippocampus as a Possible Mediator of the Relationship between Self-esteem, Locus of Control, and Stress

Although studies on brain correlates of personality variables and endocrine function in humans are only starting to appear, a picture is emerging that puts the spotlight on the hippocampus as a likely mediator between the previously reported personality variations and stress responses. The hippocampus is one of the major limbic system structures involved in the regulation of the stress response, and variations in hippocampal volume (HCV) have been postulated and shown to be systematically linked to excessive HPA axis activity. Early models postulated that associations between HCV and HPA axis activity might represent the consequence of excessive exposure of the hippocampus to glucocorticoids, due to their powerful neurotoxic properties.

More recently, however, evidence for a link between naturally occurring variations in HCV and HPA function has surfaced. Functionally, the hippocampus is the primary structure for memory contextualization, and it is here that a link to self-esteem and locus of control could occur. When an individual is

faced with a potentially threatening and harmful situation, he or she employs memories of past events to contribute to the evaluation of the current event. However, if specific situational and environmental characteristics associated with negative past events cannot be recalled, this lack of awareness of situational circumstance can lead to an overgeneralization of negative past events and thus an increased likelihood to consider the current situations as stressful as well. Thus, poor contextualization mediated by the hippocampus could be linked to higher stress responses, on the one hand, and lower self-esteem on the other.

This hypothesis of a link between HCV, personality variables, and stress responses was tested in 17 healthy young male college students. All subjects underwent extensive psychological assessment for assessment of personality variables, including self-esteem and locus of control. In addition, they underwent magnetic resonance imaging employing a structural T1-weighted acquisition protocol, which produces high-resolution images of the cerebrum. A recently developed manual segmentation protocol was then applied to all images that included nonuniformity correction, signal intensity normalization, and Talairach-like transformation for standardization of brain size. This protocol allows for manual assessment of hippocampal volumes. Finally, all subjects were exposed to a computerized stress task, similar to the mental arithmetic task described earlier. Saliva samples before, during, and after the task accompanied the testing to capture the cortisol stress response. Results first replicated earlier findings of mental arithmetic stress: only the subjects with low self-esteem and low internal locus of control showed a significant release of cortisol. Extending earlier findings, however, a correlation emerged between the cortisol stress response and self-esteem ($r = -0.45$, $p < 0.05$). Furthermore, a significant correlation emerged between the cortisol stress response and the total hippocampal volume ($r = -0.53$, $p = 0.03$), suggesting that the size of the hippocampus is related to the regulation of the HPA (Figure 3a). Finally, supporting the initial hypothesis, a link was observed between internal locus of control and HCV ($r = 0.66$, $p = 0.006$; Figure 3b) and self-esteem and HCV ($r = 0.58$, $p = 0.02$; Figure 3c).

Testing for the specificity of the effect with regard to the hippocampus, total brain gray matter was employed as control structure in correlations with both personality variables and cortisol response. None of the tested correlations were significant, suggesting that the observed relationships were indeed specific to this medial temporal lobe area.

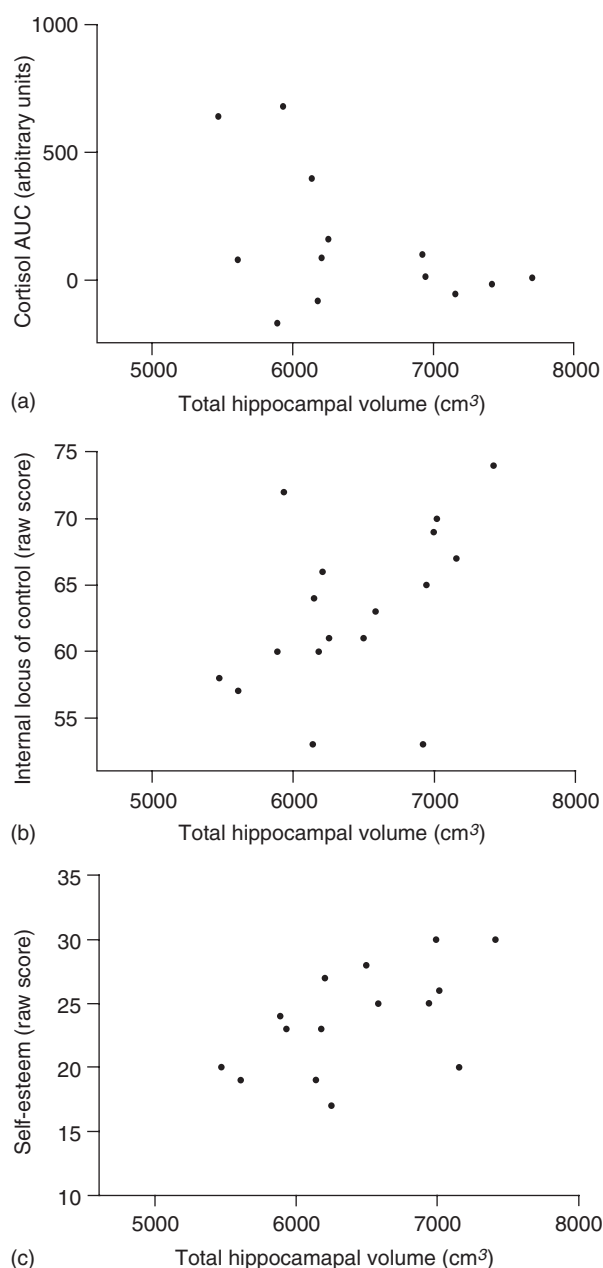


Figure 3 Correlation between hippocampal volume and cortisol stress responses to a mental arithmetic task (a), hippocampal volume and locus of control (b), and hippocampal volume and self-esteem (c) in a group of 17 healthy young male college students.

Conclusion: Personality Variables, Brain Structures, and Stress Responses in a Developmental Context

It will be a challenging task for future studies to ascertain the origin of the relationship observed between personality variables, hippocampal volume, and stress response. There are a number of possibilities as to how these associations could develop. One

model puts naturally occurring variations in hippocampal volume at the origin of these functional relationships. It can be speculated that HCV is a consequence of multiple interacting factors during critical developmental stages, leading to the variation in HCV that follows a normal distribution pattern. If, however, the size of this structure determines memory contextualization capabilities, this variation would then have personality and stress-related consequences. The impaired source monitoring would lead to an overgeneralization of the experience of failure and rejection, and consequently to a self-perception of being a failure or being socially rejected in general, which might produce low self-esteem and low locus of control during childhood and adolescence. The observed higher cortisol responses would then not be a direct consequence of small HCV, but of the increased stress perception due to consistent unfavorable evaluations of ambiguous situations. Of course, the question arises why overgeneralization should then not also be an issue for positive life events, in other words, a general lack of source monitoring for both positive and negative life events with no consequences on self-perception. One possible explanation could be seen in the stress-specific function of the hippocampus within the limbic system. There is evidence that it is specifically the hippocampus together with the anterior cingulate that is involved in the stress response. It is also known that the anterior cingulate is particularly involved in error monitoring, which means that its involvement is restricted to situations in which a mismatch between expectations and reality occurs. These circumstances would make a specific role of the hippocampus in reaction to negative life events more probable.

Based on the findings to date, we suggest that reduced HCV might play a role in the development of low self-esteem. This may in turn produce a higher susceptibility of perceiving ambiguous situations as threatening, and thus stressful. The idea that HCV variation might be the cause for adverse functional and behavioral consequences, rather than their consequence, has recently been proposed in conjunction with risk factors for developing posttraumatic stress disorder (PTSD) in a study investigating hippocampal volume in twin brothers. Here, in subjects who developed PTSD as a consequence of participating in the Vietnam War, the researchers observed lower hippocampal volume compared to subjects who went to war but did not develop PTSD. Intriguingly, however, lower hippocampal volume could also be observed in the PTSD subjects' twin brothers – who never went to war – leading to the speculation that HCV might be a risk factor for, rather than a consequence of, developing PTSD.

As mentioned earlier, the possibility that HCV plays a causal role is supported by the fact that consistent results were found even among relatively young adults. Recently, variations in HCV in young populations have been reported, suggesting that there is a considerable range of hippocampal volumes already in young subjects. At the same time, self-esteem is known to be a stable trait, with considerable intra-individual stability throughout life. This supports a model in which variations in brain morphology could become a pathway to certain personality characteristics and behavioral phenotypes early in life, but longitudinal studies will be needed to examine these changes over time and provide validation for the assumed associations.

See Also the Following Articles

Corticosteroid Receptors; Hippocampus, Corticosteroid Effects on; Hippocampus, Overview; Hypothalamic-Pituitary-Adrenal; Optimism, Pessimism, and Stress; Psychological Stressors, Overview; Social Status and Stress.

Further Reading

- Baltes, M. M. and Baltes, P. B. (1990). *Successful aging*. Cambridge, UK: Cambridge University Press.
- Baumeister, R. F., Campbell, J. K., Krueger, J. I. and Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest* 4, 1–44.
- Dickerson, S. S. and Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin* 130(3), 355–391.
- Fuchs, E. and Flugge, G. (2003). Chronic social stress: effects on limbic brain structures. *Physiology and Behavior* 79(3), 417–427.
- Gilbertson, M. W., Shenton, M. E., Ciszewski, A., Kasai, K., Lasko, N. B., Orr, S. P., et al. (2002). Smaller hippocampal volume predicts pathologic vulnerability to psychological trauma. *Nature Neuroscience* 5(11), 1242–1247.
- Hoyle, R., Kernis, M. H., Leary, M. R. and Baldwin, M. W. (1999). *Selfhood: identity, esteem, control*. Boulder, CO: Westwood.
- Kirschbaum, C., Bartussek, D. and Strasburger, C. J. (1992). Cortisol responses to psychological stress and correlations with personality traits. *Personality and Individual Differences* 13(12), 1353–1357.
- Lazarus, R. S. (1999). *Stress and emotion: a new synthesis*. New York: Springer.
- Leary, M. R. and McDonald, G. (2003). Individual differences in self-esteem: a review and theoretical integration. In: Leary, M. R. & Tangney, J. P. (eds.) *Handbook of self and identity*, pp. 401–420. New York: Guilford Press.
- Petrie, K. and Rotheram, M. J. (1982). Insulators against stress: self-esteem and assertiveness. *Psychological Report* 50(3 Pt 1), 963–966.
- Pruessner, J. C., Hellhammer, D. H. and Kirschbaum, C. (1999). Low self-esteem, induced failure and the adrenocortical stress response. *Personality and Individual Differences* 27(3), 477–489.
- Pruessner, J. C., Collins, D. L., Pruessner, M. and Evans, A. C. (2001). Age and gender predict volume decline in the anterior and posterior hippocampus in early adulthood. *Journal of Neuroscience* 21(1), 194–200.
- Pruessner, J. C., Baldwin, M. W., Dedovic, K., et al. (2005). Self-esteem, locus of control, hippocampal volume, and cortisol regulation in young and old adulthood. *Neuroimage* 28, 815–826.
- Sapolsky, R. M. (1999). Glucocorticoids, stress, and their adverse neurological effects: relevance to aging. *Experimental Gerontology* 34(6), 721–732.
- Sinha, R., Lacadie, C., Skudlarski, P. and Wexler, B. E. (2004). Neural circuits underlying emotional distress in humans. *Annals of the New York Academy of Sciences* 1032, 254–257.

Stress System Balance Hypothesis

E R de Kloet

Leiden/Amsterdam Center for Drug Research and
Leiden University Medical Center, Leiden,
The Netherlands

© 2007 Elsevier Inc. All rights reserved.

Corticotropin Releasing Hormone and Urocortins
Glucocorticoids
Two Corticosteroid Receptor Types: Low and High Levels
of Stress System Operation
Corticosteroid Receptor Balance
Nuclear Receptors: Slow Genomic and Fast Membrane
Actions
Perspectives

Glossary

<i>Allostasis</i>	The ability to reestablish homeostasis through change during adaptation to stress.
<i>Allostatic load</i>	The cost of allostasis; a too high cost may lead to wear and tear, and ultimately to disease.
<i>Corticotropin releasing hormone (CRH)</i>	A hormone produced in the hypothalamus that stimulates after secretion in the portal vessels the release of pituitary ACTH.
<i>Homeostasis</i>	A stable <i>milieu intérieur</i> , which refers to the defense of physiological and psychological integrity.
<i>Homeostatic mechanism</i>	The processes and activities coordinated by the stress system that help to maintain homeostasis.
<i>Hypothalamic-pituitary-adrenal (HPA) axis</i>	Refers to a coordinate pattern of stimulation and feedback interaction between hypothalamic paraventricular neurons, pituitary corticotrophs and the adrenal cortex with the aim to facilitate adaptation to stress and to synchronize daily and sleep-related events.
<i>Nuclear receptors</i>	Proteins that bind steroid hormones and that can act in the cell nucleus as gene transcription factors in the modulation of gene expression. The receptors appeared recently to reside also in the membrane where they mediate fast non-genomic actions. In the brain, mineralocorticoid receptors are promiscuous and bind with high affinity naturally occurring glucocorticoids (cortisol and corticosterone), mineralocorticoids, and progesterone; glucocorticoid receptors

Resilience

Secretagogues

Set point

Stressor

Stress system balance hypothesis

Stress system mediators

Stress response

bind exclusively naturally occurring and synthetic glucocorticoids.

The characteristic of a stress system that rapidly turns on and off.

Substances that stimulate the endocrine system to secrete hormones; here used to indicate the peptides released by hypothalamic neurons in the pituitary portal vessel system.

The set level of the regulated variable (physiological function); homeostatic mechanisms serve to restore the set point perturbed by stress. Allostasis reflects an adaptive change in the set point necessary to cope with the stress.

Any perceived threat to the individual's integrity.

The hypothesis that the defense of homeostasis and development of allostasis depend on the balanced interaction between activating and inhibiting humoral and neural stress system mediators.

The network of humoral and nervous mediators of the stress response.

The spectrum of physiological and behavioral adaptations that defend homeostasis and/or to achieve allostasis and that are coordinated by the stress system mediators.

A stressor is any real or perceived threat to homeostasis (i.e., the stability and integrity of the organism). The defense of homeostasis to the stressor triggers a corrective physiological and behavioral response (the stress response) that is coordinated by humoral and nervous mediators, the stress system mediators. In particular, psychological components of the stress response often require adaptation to the stressor through change, a process termed allostasis. The stress-induced activation and termination of the action of the mediators in humoral and nervous signaling represent two compensatory phases: some mediators act in an opposing mode and others in a synergistic mode (Table 1). The two modes of stress system operation need to be in balance, however, to ensure a proportionate outcome in the third and preparative phase for the next stress response. This implies the storage of the experience in memory and the storage of energy resources for coping with future needs. A disproportionate outcome of stress system operation has been termed allostatic load and is considered an unfavorable prediction for health.

Table 1 Some selected mediators of the two modes of stress system operation^a

<i>Activating mode: fight-flight-fright</i>	<i>Suppressing mode: coping and adaptation</i>
CRH, CRH-1 receptors	Urocortins, CRH-2 receptors
Vasopressin	Oxytocin
ACTH	Endorphins
Sympathetic response	Parasympathetic response
Mineralocorticoid receptors	Glucocorticoid receptors

^aThe two modes are in balance for homeostasis and health. ACTH, adrenocorticotrophic hormone; CRH, corticotropin releasing hormone.

Such a disproportionate outcome is reflected in an increased vulnerability to stress-related diseases. In the brain, such diseases are characterized by emotional lability and cognitive disturbance. Both may have profound consequences for metabolic, immune, and cardiovascular regulations.

Corticotropin Releasing Hormone and Urocortins

Most mediators of the stress system are known and their molecular and anatomical organization is being unraveled. Hierarchically, the hypothalamic paraventricular nucleus (PVN) integrates all stressful information. The PVN is activated directly by physical stressors through aminergic ascending pathways and by metabolic disturbance through peptidergic hypothalamic pathways. Psychologically stressful information is conveyed transsynaptically to the PVN after limbic-midbrain processing. The PVN contains different cell types that express peptidergic secretagogues and that, because of their projection to different targets, can activate distinct functions in a coordinate manner.

Hence, the PVN is the organizer of the immediate behavioral, autonomic, and neuroendocrine response to a stressor. The ratio of PVN secretagogues can change as a function of the nervous input. Corticotropin releasing hormone (CRH) is released in changing ratios with vasopressin and other peptides driving the sympathetic and behavioral fight-flight-fright response. The CRH-1 receptors mediating this response can also activate the hypothalamic-pituitary-adrenal (HPA) axis, leading to the secretion of adrenal corticosteroids. To achieve this goal, CRH and vasopressin are secreted in the pituitary portal vessel system and induce in the anterior pituitary the synthesis of proopiomelanocortin (POMC), which is processed to balanced concentrations of adrenocorticotrophic hormone (ACTH), opioids, and melanocortin peptides, for example. ACTH stimulates the adrenal

cortex to secrete cortisol (humans) and corticosterone (humans, rats, and mice), whereas opioids attenuate stress reactions, resulting in suppression of adrenal function. Collectively, these mediators of the HPA axis act in concert to coordinate the neuroendocrine stress response. The mechanism underlying the CRH-induced sympathetic outflow and behavioral response pattern is not precisely known.

Recently, urocortin II and III (stresscopin and stresscopin-related peptide, respectively) were identified acting through the CRH-2 receptor system. The CRH-1 and CRH-2 receptor systems have a partly overlapping distribution in the brain, which matches their CRH and urocortin terminal fields. The intracerebroventricular administration of CRH mimics the initial behavioral, autonomic, and neuroendocrine stress response and is anxiogenic. In contrast, urocortin seems to have opposing functions as expressed, for example, in its anxiolytic properties. It is well established that the CRH/CRH-1 receptor system organizes the onset of the stress response.

Some investigators have proposed that the urocortin-CRH-2 receptor system represents a slower mode of operation of the stress system mediators, perhaps one that is primarily concerned with adaptation and coping with stressors. Such a bimodal CRH-1/CRH-2 receptor-driven neuroendocrine system would control over time the compensatory phases of initiation and termination in the stress response. Such a system would operate in an integrative fashion with the sympathetic and parasympathetic divisions of the autonomic nervous system, and the signals released from the dendritic cell would control the balance between immune activation and tolerance. In fact, other neuropeptides in the stress system, such as vasopressin and oxytocin, also have opposing actions in aspects of behavioral adaptation.

Glucocorticoids

The importance of a balance in the concentration and action of stress mediators is illustrated also by the glucocorticoid hormone (cortisol or corticosterone) secreted from the adrenal cortex. The hormone operates both in the fast and slow modes of the response to stress and produces lasting effects in the preparation of future events. In concert with the other stress mediators, the hormone has two modes of operation that are complementary. First, it acts permissively and exerts a tonic influence on cellular and tissue functions. In the behavioral realm, the hormone operates in a pro-active mode, so that the organism is prepared in anticipation of upcoming events. Second, once homeostasis is disturbed, the concentration of cortisol rises in response to HPA axis activation to

promote the recovery of disturbed homeostasis. High concentrations of cortisol operate in a reactive mode, through feedback foremost, by preventing the initial stress reactions from overshooting and to become damaging in themselves.

The two complementary modes of operation of cortisol depend on the concentration of the hormone, the context and the phase of the stress response. In the low basal mode pro-active hormone actions seemingly have a household function in which the hourly pulses of the hormone over the circadian cycle provide the capability to synchronize upcoming daily and sleep-related events and to coordinate the available mental and body resources to deal with upcoming challenges. These functions are stepped up when cortisol levels rise in the stress mode. At the same time, the redistribution of resources is facilitated, enabling the individual to cope and to suppress any overreaction to a stressor in the reactive mode. Cortisol action operating in the basal and the stress mode therefore displays an enormous diversity in action that is dependent on the context. The context is a reflection of the signals provided by the other local and systemic stress mediators that converge in time and phase in the cell or tissue in which homeostasis has been challenged. In other words, it is essential that the hormones act in the right time window on the cells and circuits in the brain that have initially led to their own secretion. If they act out of context, the hormones' action becomes damaging.

Two Corticosteroid Receptor Types: Low and High Levels of Stress System Operation

The molecular basis of the dual mode of operation of cortisol is formed by two types of receptors – mineralocorticoid receptors (MRs) and glucocorticoid receptors (GRs) – that bind *in vivo* cortisol with an order of magnitude difference in affinity. However, this only occurs if the cells lack 11 β -hydroxysteroid dehydrogenase type 2 (11 β -HSD2), which inactivates the steroid. The receptors have a different distribution. GRs occur in virtually every cell and are expressed most abundantly in cells with a key role in coping with stress (e.g., the liver, immune system, and limbic forebrain regions). The cortisol-responsive MRs are expressed abundantly in the heart, vascular tissue, and limbic-forebrain regions such as the hippocampus. Using selective antagonists and conditional knockout strategies, researchers have studied their functions in experimental animals from gene to behavior.

MRs and GRs are nuclear receptors that mediate the slow genomic actions of cortisol. The receptors

recruit patterns of coregulators and transcription factors to facilitate their dual mode of operation. In cellular studies using the hippocampus, cortisol suppresses through MR inhibitory inputs and Ca influx. The consequence is that excitatory transmission is facilitated, and Ca²⁺ availability is suppressed, which is a favorable condition for a cell. MR activation indeed protects against damage caused by ischemic insults and suppresses apoptotic pathways.

GRs are occupied in the ultradian and circadian peaks of circulating cortisol. GRs are also occupied after stress-induced rises in cortisol; thus, in hippocampal neurons, the activation of GRs some time after exposure to the stressor suppresses the stress-induced disturbances in excitability transiently raised by excitatory input. The recovery process promoted by GRs enhances a hyperpolarizing input and Ca²⁺ influx; the latter process, although beneficial in the short run, may explain why chronically elevated high glucocorticoid levels and excitotoxic challenge create unfavorable conditions for the functioning of the hippocampal neurons.

Corticosteroid Receptor Balance

From these cellular studies and from endocrine and behavioral observations, a concept has evolved in which MRs and GRs mediate the dual mode of operation of cortisol in the limbic brain to regulate in a coordinate manner the onset and termination of the response to a stressor. Because limbic areas such as the hippocampus, amygdala, and frontal cortex are the target, the action of cortisol depends on stressful information determined by spatial and temporal cues. MRs are implicated in the appraisal of the situation and in processes involved in the prediction and choice of the appropriate behavioral reaction to deal with a challenge. Such predictive responses are important for the onset of the stress response by determining the threshold or sensitivity of the stress system to challenge. Along the same line of reasoning, MRs are in a proactive mode of operation implicated in predictable ultradian and circadian changes. High levels of cortisol facilitate through GRs the termination of the stress reactions in a reactive fashion, mobilize for this purpose the required energy resources, and facilitate recovery. Due to their occupation during pulsatile and circadian peaks, GRs may function as Zeitgeber. In preparation for future events, GRs promote memory processes and the storage of energy resources. So, the two receptor types belong to interacting signaling networks that underlie adaptive processes from appraisal of novel situations and prediction of upcoming events to recovery from the challenge and storage of the experience.

in memory. They participate in the control of energy metabolism, from appetite and feeding to energy disposition and storage.

Nuclear Receptors: Slow Genomic and Fast Membrane Actions

To this analysis can be added the mechanism of action of steroids in the brain via MRs and GRs. The nuclear receptors are well known as mediating the slow genomic actions of glucocorticoids. Recently it has become evident that these receptors also serve at the membrane level to mediate rapid glucocorticoid actions. A clear example is the fast MR-mediated action of corticosterone and cortisol on the release of the excitatory transmitter glutamate in the hippocampus. These fast steroid actions probably have a function in the pro-active modes of action of the steroid in anticipated perturbances of homeostasis.

Perspectives

Three questions are at the core of stress system operation and health: how is the stress system balance regulated, what are the implications of dysregulation for health and disease, and which biomarkers can reveal a balanced stress system operation?

Regulation of Stress System Balance

Regulation of the stress system balance is by definition based on the activity of the humoral and neural mediators that are activated by stress. In the case of the corticosteroids it depends on their secretion rate; bioavailability; access to the brain through multidrug-resistant P glycoprotein; the stoichiometry of receptors, coregulators and other proteins and genetic factors. Genetic variants and polymorphisms can bias the response patterns. Recently, knowledge of epigenetic control of gene expression and function has grown; it has been demonstrated that stable alterations in stress system mediators may occur as a result of environmental input and early life experience. In later life, psychosocial contexts that challenge the individual's coping ability, particularly if experienced chronically, also cause enduring changes in stress system operation.

Implications for Health and Disease

Stress mediators acting either excessively or inadequately produce a phenotype vulnerable to stress-related disease in predisposed individuals. The causes of aberrant stress system operation can be environmental, experience-related, or genetic; it is reflected in the ability to cope with a stressor, and it is also

affected by additional factors such as life style and diet. It is believed that an imbalance, rather than the absolute concentration of stress system mediators, compromises resilience and enhances vulnerability to disease. This condition has been termed allostatic load.

Selye championed the balance of mineralocorticoid versus glucocorticoid actions as determining pro- versus anti-inflammatory actions. Selye proposed that imbalance therefore enhanced the risk for inflammation or infection. In modern terms, in the brain the action of one single hormone – cortisol or corticosterone – depends on the balance of MR- and GR-mediated actions. Once the balance is disturbed, the individual loses the ability to maintain homeostasis when challenged by an adverse life event. This may lead to a condition of neuroendocrine dysregulation and impaired behavioral adaptation, enhancing the precipitation of stress-related disease for which the individual is genetically predisposed.

Biomarkers

Stress-related diseases, such as depression, are complex and multifactorial diseases. They likely cannot be attributed to mutations in a single gene or to a single external event; instead, they result from the concerted actions of many subtle genetic polymorphisms and from external events, the effects of which might accumulate over time. Once traumatic life events in combination with genetic disposition have engrained long-lasting changes in MR and GR signaling, a vulnerable phenotype emerges. MRs and GRs may therefore be considered as allostatic markers. The individual's vulnerability is defined by three different levels of description: the traditional clinical phenotype (emotional reactivity and personality), the functional phenotype (neuroendocrine reactivity to stimuli, MRs, GRs, and beyond), and the genotype (e.g., polymorphisms in genes that are involved in HPA signaling or monoaminergic neurotransmission). An integrated approach that analyzes patients at all three levels may provide insight into the causes of stress-related diseases and perhaps also into new targets for treatment.

Acknowledgments

Support from the Royal Netherlands Academy of Arts and Sciences and the Netherlands Organization for Scientific Research (NWO) is gratefully acknowledged.

See Also the Following Articles

Allostasis and Allostatic Load; Corticosteroid Receptors; Corticosteroids and Stress; Feedback Systems; Glucocorticoid Negative Feedback; Glucocorticoids, Effects of

Stress on; Glucocorticoids, Role in Stress; Hippocampus, Corticosteroid Effects on; Homeostasis; Hypothalamic-Pituitary-Adrenal; Stress, Definitions and Concepts of; Survivor Guilt; CRH Anxiety Circuitry in Brain; Neuropeptides, Stress-Related.

Further Reading

- Berridge, K. C. (2004). Motivation concepts in behavioral neuroscience. *Physiology & Behavior* **81**, 179–209.
- Chrousos, G. P. and Gold, P. W. (1992). The concepts of stress and stress system disorders: overview of physical and behavioral homeostasis. *Journal of the American Medical Association* **267**, 1244–1252.
- Day, T. A. (2005). Defining stress as a prelude to mapping its neurocircuitry: no help from allostasis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **29**, 195–200.
- de Kloet, E. R., Joëls, M. and Holsboer, F. (2005). Stress and the brain: from adaptation to disease. *Nature Reviews Neuroscience* **6**, 463–475.
- de Kloet, E. R., Vreugdenhil, E., Oitzl, M. S., et al. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Reviews* **19**, 269–301.
- Joëls, M., Pu, Z., Wiegert, O., et al. (2006). Learning under stress: how does it work? *Trends in Cognitive Science* **10**, 152–158.
- Karst, H., Berger, S., Turiault, M., et al. (2005). Mineralocorticoid receptors are indispensable for nongenomic modulation of hippocampal glutamate transmission by corticosterone. *Proceedings of the National Academy of Sciences USA* **102**, 19204–19207.
- Kruk, M. R., Halász, J., Meelis, W., et al. (2004). Fast positive feedback between the adrenocortical stress response and a brain mechanism involved in aggressive behaviour. *Behavioral Neuroscience* **118**, 1062–1070.
- McEwen, B. S. (2004). Protection and damage from acute and chronic stress: allostasis and allostatic overload and relevance to the pathophysiology of psychiatric disorders. *Annals of the New York Academy of Sciences* **1032**, 1–7.
- McEwen, B. S. and Wingfield, J. C. (2003). The concept of allostasis in biology and biomedicine. *Hormones and Behavior* **43**, 355–367.
- Sapolsky, R. M., Romero, L. M. and Munck, A. U. (2000). How do glucocorticoids influence stress responses? integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrine Reviews* **21**, 55–89.
- Steptoe, A. and Wardle, J. (2005). Positive affect and biological function in everyday life. *Neurobiology of Aging* **26**, 108–112.
- VanPraag, H. M., de Kloet, E. R. and Van Os, J. (2004). *Stress, the brain and depression*. Cambridge, UK: Cambridge University Press.

Stress, Beneficial Effects of

S Joseph

University of Nottingham, Nottingham, UK

P A Linley

University of Leicester, Leicester UK

© 2007 Elsevier Inc. All rights reserved.

Types of Benefit

Prevalence

Etiological Factors

Theoretical Developments

Clinical Applications

Glossary

Adversarial growth

Increased psychological functioning and personal development following exposure to a stressful and traumatic event; also called posttraumatic growth or stress-related growth.

Assumptive world

Cognitive schemata about the self and world that determine how meaning is perceived.

A number of literatures, religions, and philosophies throughout human history have conveyed the idea that there are benefits to be found in suffering. However, it is only relatively recently that the topic of benefit finding following stress and trauma has become a research focus. Initially there was some debate over the validity of the concept of benefit finding, with some commentators questioning whether it was anything more than positive illusion and self-deception. But there is now convincing evidence that people often experience benefits following stress and trauma. These benefits have been variously labeled adversarial growth, benefit finding, flourishing, heightened existential awareness, perceived benefits, positive by-products, positive changes, positive meaning, post-traumatic growth, quantum change, self-renewal,

stress-related growth, thriving, and transformational coping. These terms have been used interchangeably to refer not just to recovery following stressful and traumatic events but to how events can sometimes serve as a springboard to a higher level of psychological functioning.

Types of Benefit

Three broad interrelated dimensions of benefit finding have been discussed. First, relationships are enhanced in some way; for example, people now value their friends and family more and feel an increased compassion and altruism toward others. Second, people change their views of themselves in some way; for example, they have a greater sense of personal resiliency, wisdom, and strength, perhaps coupled with a greater acceptance of their vulnerabilities and limitations. Third, people report changes in life philosophy; for example, they find a fresh appreciation for each new day and renegotiate what really matters in the full realization that their life is finite, including possibly changing their spiritual beliefs. Various psychometric self-report instruments have been developed with which to assess positive changes and personal growth following adversity: the Changes in Outlook Questionnaire, the Perceived Benefit Scale, the Posttraumatic Growth Inventory, the Stress-Related Growth Scale, and the Thriving Scale. The development of these instruments has led to further empirical investigation.

Prevalence

Prevalence varies among studies due to samples, contexts, and methodological issues, but it is expected that anywhere between 30 and 70% of people may experience some form of benefit following traumatic events, although the benefits are not usually reported in the immediate aftermath but at some later time. Also, the benefits are likely to coexist with psychological distress; the experience of benefits is not mutually exclusive with other more negative psychological consequences.

Etiological Factors

Studies have reported benefits following a range of stressful and traumatic events, for example, bereavements; accidents and disasters; chronic and life-threatening illnesses; sexual, physical, and emotional abuse in childhood; sexual assaults; and wars and conflicts. Research has also begun to document the correlates and predictors of growth, with evidence pointing to the importance of stress-appraisal, coping, and personality variables, with the more

optimistic, self-efficacious, extraverted, and stable people, who use more spiritual, accepting, emotionally expressive, and emotionally focused forms of coping, being more likely to report benefits. Social support is thought to be important too.

Theoretical Developments

Tedeschi and Calhoun's functional-descriptive model shows how traumatic events serve as seismic challenges to the pretrauma schema by shattering prior goals, beliefs, and ways of managing emotional distress. When these schemas are shattered in this way, this leads to ruminative activity as people try to make sense of what has happened and to deal with their emotional reactions to the trauma. In the initial stages, this ruminative activity is more automatic than deliberate. Although distressing, automatic ruminative activity is indicative of cognitive activity that is directed at rebuilding the pretrauma schema. This ruminative process is influenced by social support networks, which provide sources of comfort and relief, as well as being influenced by new coping behaviors and available options for the construction of new posttrauma schemas. Successful coping at this stage facilitates disengagement from goals that are now unreachable and from beliefs that are no longer tenable in the posttrauma environment, together with decreased emotional distress. As successful coping aids adaptation, the initial ruminative activity that was characterized by its automatic nature shifts toward a more effortful ruminative activity. This effortful ruminative activity is characterized by narrative development, part of which may be the search for meaning.

Introducing a metatheoretical perspective, Joseph and Linley proposed that human beings are active, growth-oriented organisms. The theory proposes that individuals are intrinsically motivated to cognitively accommodate their psychological experiences and thus find benefit. The confrontation with a stressful and traumatic event has a shattering effect on the person's assumptive world. When the social environment is able to provide for the basic human needs of autonomy, competence, and relatedness, then models of the self and world transform to accommodate the new trauma-related information. When the social environment does not provide support for these basic human needs, the process of accommodation is thwarted and benefits are less likely to be experienced.

Similarly, but from a biopsychosocial-evolutionary perspective, Christopher regarded growth, rather than pathology, as the normal outcome of the traumatic stress response. Reviewing and synthesizing the trauma literature, Christopher articulated seven interconnected theoretical conclusions:

1. Stress is best understood as a prerational form of biopsychological feedback.
2. The normal outcome of traumatic stress is growth.
3. Any resultant psychopathology is a function of a maladaptive modulation of the stress response.
4. Trauma transforms individuals on the biological as well as psychological levels.
5. Biological processes underlying stress response are universal, but there are specific dynamics that are a function of the individual.
6. Biological changes can occur even if there are no psychological changes.
7. Rationality is humanity's evolutionary newest and most sophisticated stress-reduction mechanism.

Clinical Applications

The interest in benefit finding can be seen as part of the wider positive psychology movement. Positive psychology is seen as a balance to mainstream psychology, which has been for too long overly concerned with the negative aspects of human experience. Research in positive psychology is now beginning to focus on its applications, and the study of benefit finding promises to have important applications for clinical, counseling, and health psychologists, who when working with clients who have experienced stressful and traumatic events need not be concerned only with helping to alleviate distress but also with helping to facilitate more positive functioning.

The importance of this avenue of investigation is seen in the seminal work of Affleck and colleagues, who found that perceived benefits at 7 weeks following a heart attack significantly predicted less chance of heart attack recurrence and lower general health morbidity at an 8-year follow-up. Experimental studies to test whether the principles of growth might somehow be introduced as part of a clinical intervention are encouraging. For example, Stanton and colleagues randomly assigned breast cancer patients to one of two groups, one group to write about the facts of the cancer experience and the other to write about positive thoughts and feelings regarding the experience. It was found that those assigned to write about their positive experiences had significantly fewer medical appointments for cancer-related morbidities 3 months later.

Clinicians should be aware of the potential for positive change in their clients following stress and trauma. But it must also be recognized that adversity does not lead to positive change for everyone. Therefore, therapists need to be careful not to inadvertently imply that patients have in some way failed by not making more of their experience or that

there is anything inherently positive in their experiences. Personal growth after trauma should be viewed as originating not from the event, but from within the patients themselves through the process of their struggle with the event and its aftermath.

See Also the Following Articles

Bereavement; Caregivers, Stress and; Cognitive Behavioral Therapy; Integrative Medicine (Complementary and Alternative Medicine); Health Behavior and Stress; Mental Stress Testing; Religion and Stress; Self-Esteem, Stress and Emotion; Recovery from Stress; Stress of Self Esteem.

Further Reading

- Affleck, G., Tennen, H., Croog, S., et al. (1987). Causal attribution, perceived benefits, and morbidity after a heart attack: an 8-year study. *Journal of Consulting and Clinical Psychology* 55, 29–35.
- Calhoun, L. G. and Tedeschi, R. G. (1999). *Facilitating posttraumatic growth: a clinician's guide*. Mahwah, NJ: Lawrence Erlbaum.
- Christopher, M. (2004). A broader view of trauma: a biopsychosocial-evolutionary view of the role of the traumatic stress response in the emergence of pathology and/or growth. *Clinical Psychology Review* 24, 75–98.
- Cordova, M. J., Cunningham, L. L. C., Carlson, C. R., et al. (2001). Posttraumatic growth following breast cancer: a controlled comparison study. *Health Psychology* 20, 176–185.
- Frazier, P., Tashiro, T., Berman, M., et al. (2004). Correlates of levels and patterns of positive life changes following sexual assault. *Journal of Consulting and Clinical Psychology* 72, 19–30.
- Joseph, S. and Linley, P. A. (2005). Positive adjustment to threatening events: an organismic valuing theory of growth through adversity. *Review of General Psychology* 9, 262–280.
- Linley, P. A. and Joseph, S. (2004). Positive change following trauma and adversity: a review. *Journal of Traumatic Stress* 17, 11–21.
- O'Leary, V. E., Alday, C. S. and Ickovics, J. R. (1998). Models of life change and posttraumatic growth. In: Tedeschi, R. G., Park, C. L. & Calhoun, L. G. (eds.) *Posttraumatic growth: positive changes in the aftermath of crisis*, pp. 127–151. Mahwah, NJ: Lawrence Erlbaum.
- Park, C. L., Cohen, L. H. and Murch, R. (1996). Assessment and prediction of stress-related growth. *Journal of Personality* 64, 71–105.
- Seligman, M. E. P. and Csikszentmihalyi, M. (2000). Positive psychology: an introduction. *American Psychologist* 55, 5–14.
- Stanton, A. L., Danoff-Burg, S., Sworowski, L. A., et al. (2002). Randomized controlled trial of written emotional expression and benefit finding in breast cancer patients. *Journal of Clinical Oncology* 20, 4160–4168.

- Tedeschi, R. G. and Calhoun, L. G. (2004). A clinical approach to posttraumatic growth. In: Linley, P. A. & Joseph, S. (eds.) *Positive psychology in practice*, pp. 405–419. Hoboken, NJ: John Wiley.
- Tedeschi, R. G. and Calhoun, L. G. (2004). Posttraumatic growth: conceptual foundations and empirical evidence. *Psychological Inquiry* 15, 1–18.
- Tedeschi, R. G., Park, C. L. and Calhoun, L. G. (eds.) (1998). *Posttraumatic growth: positive changes in the aftermath of crisis*. Mahwah, NJ: Lawrence Erlbaum.
- Tennen, H. and Affleck, G. (2002). Benefit-finding and benefit-reminding. In: Snyder, C. R. & Lopez, S. J. (eds.) *Handbook of positive psychology*, pp. 584–597. New York: Oxford University Press.

Stress, Definitions and Concepts of

B S McEwen

Rockefeller University, New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by B S McEwen, volume 3, pp 508–509, © 2000, Elsevier Inc.

Stress involves a stressor and a stress response. A stressor may be a physical insult, such as trauma or injury, or physical exertion, particularly when the body is being forced to operate beyond its capacity. Other physical stressors include noise, overcrowding, and excessive heat or cold. Stressors also include primarily psychological experiences such as time-pressured tasks, interpersonal conflict, unexpected events, frustration, isolation, and traumatic life events, and all of these types of stressors may produce behavioral responses and evoke physiological consequences such as increased blood pressure, elevated heart rate, increased cortisol levels, impaired cognitive function, and altered metabolism.

Behavioral responses to stressors may decrease risk and get the individual out of trouble or involve health-promoting activities such as a good diet and regular exercise, but they may also include responses that exacerbate the physiological consequences of stress such as self-damaging behaviors (e.g., smoking, drinking, overeating, or consuming a rich diet) and risk-taking behaviors (e.g., driving an automobile recklessly). The physiological stress responses include primarily the activation of the autonomic nervous system and the hypothalamic-pituitary-adrenal (HPA) axis, leading to increased blood and tissue levels of catecholamines and glucocorticoids. It is these physiological responses that have both protective and damaging effects (see **Allostasis and Allostatic Load**).

There are two important features of the physiological stress response. The first is turning it on in amounts that are adequate to the challenge (see **Allostasis and Allostatic Load**). The second is turning off the response

when it is no longer needed. The physiological mediators of the stress response, namely, the catecholamines of the sympathetic nervous system and the glucocorticoids from the adrenal cortex, initiate cellular events that promote adaptive changes in cells and tissues throughout the body, which in turn protect the organism and promote survival. However, too much stress or inefficient operation of the acute responses to stress can cause wear and tear and exacerbate disease processes.

There are enormous individual differences in interpreting and responding to what is stressful, as well as individual differences in the susceptibility to diseases, in which stress may play a role. Genetic predispositions exist that increase the risk of certain disorders. In addition, developmental process, such as prenatal stress or nurturing postnatal experiences, contribute to the life-long responsiveness of the behavioral and physiological responses to stressors. Furthermore, experiences throughout the life course that result in memories of particularly unpleasant or pleasant situations combine with the genetic and developmental influences to produce large differences among individuals in how they react to stress and what the long-term consequences may be.

See Also the Following Articles

Allostasis and Allostatic Load; Homeostasis.

Further Reading

- Mason, J. (1959). Psychological influences on the pituitary-adrenal cortical system. In: Pincus, G. (ed.) *Recent progress in hormone research*, pp. 345–389. New York: Academic Press.
- Selye, H. (1936). A syndrome produced by diverse noxious agents. *Nature* 138, 32.
- Selye, H. (1955). Stress and disease. *Science* 122, 625–631.

Stress, General Overview See: Stress, Definitions and Concepts of; Stress Effects, Overview.

Stress, Insulin Resistance, and Type II Diabetes

C Tsigos and I Kyrou

 Evgenidion Hospital, Athens University Medical School,
Athens, Greece

© 2007 Published by Elsevier Inc.

The Stress Response

Depression and Diabetes

Stressful Life Events and Diabetes

Inflammation and Diabetes

Conclusion

Glossary

<i>Adipokines</i>	Cytokine-like molecules produced and secreted by the adipose tissue (e.g., leptin, adiponectin, resistin, tumor necrosis factor α [TNF- α], interleukin-6 [IL-6]).
<i>Insulin resistance</i>	Decreased ability of insulin to act on its target tissues, resulting in compensatory hyperinsulinemia. Considered to be the earlier abnormality in the natural history of type II diabetes.
<i>Metabolic syndrome</i>	Syndrome characterized by the combination of central obesity, insulin resistance, dyslipidemia, and hypertension and correlated with high mortality rates due to cardiovascular complications, which aptly justify the description of the syndrome as the deadly quartet.
<i>Visceral obesity</i>	Obesity characterized by fat accumulation predominantly in the abdominal area, also called central or abdominal obesity.

The Stress Response

Stress is defined as a state of disharmony or threatened homeostasis. The concepts of stress and homeostasis can be traced back to ancient Greek history; however, the integration of these notions with related physiological and pathophysiological mechanisms and their association with specific illnesses are much more recent.

When faced with excessive stress, whether biological, physical, or psychological, a subject's

adaptive responses attain a stereotypic nonspecific nature, referred to by Selye as the general adaptation syndrome. During stress, attention is enhanced and the brain focuses on the perceived threat. Cardiac output and respiration are accelerated, catabolism is increased, and blood flow is redirected to provide the highest perfusion and fuel to the aroused brain, heart, and muscles.

The central control stations of the stress system are located in the hypothalamus and the brain stem and include the parvocellular corticotropin-releasing hormone (CRH) and arginine-vasopressin (AVP) neurons of the paraventricular nuclei of the hypothalamus and the locus ceruleus-norepinephrine system (central sympathetic system) (Figure 1). The hypothalamic-pituitary-adrenal (HPA) axis, together with the efferent sympathetic/adrenomedullary system, represent the effector limbs, via which the brain influences all body organs during exposure to threatening stimuli.

The Hypothalamic-Pituitary-Adrenal (HPA) Axis

The hypothalamus controls the secretion of adrenocorticotrophic hormone (ACTH) from the anterior pituitary, which, in turn, stimulates the secretion by the adrenal cortex of glucocorticoids, mainly cortisol in humans. The principal hypothalamic stimulus to the pituitary-adrenal axis is CRH. AVP is a potent synergistic factor with CRH in stimulating ACTH secretion.

In nonstressful situations, both CRH and AVP are secreted into the hypophyseal portal system in a circadian, pulsatile fashion, with a frequency of about two to three secretory episodes per hour. Under resting conditions, the amplitude of the CRH and AVP pulses increases in the early morning hours, resulting in ACTH and cortisol secretory bursts in the general circulation. These diurnal variations are perturbed by changes in lighting, feeding schedules, and activity and are disrupted by stress.

Circulating ACTH is the key regulator of glucocorticoid secretion by the adrenal cortex. Glucocorticoids are the final effectors of the HPA axis and participate in the control of whole body homeostasis and the organism's response to stress. They play a key regulatory role in the basal activity of the HPA axis

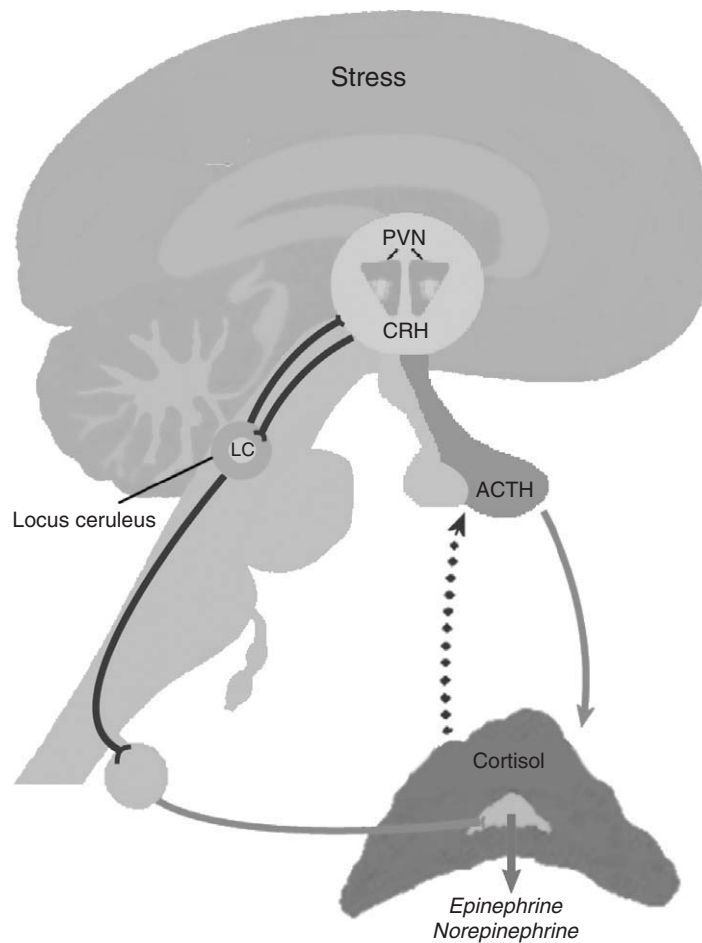


Figure 1 Schematic representation of the stress system, highlighting the major central and peripheral components involved in the stress response. PVN, paraventricular nucleus; CRH, corticotropin-releasing hormone; LC, locus ceruleus/norepinephrine-sympathetic system; ACTH, adrenocorticotropic hormone. Activation is represented by solid lines and inhibition by dashed lines.

and in the termination of the stress response by acting at the hypothalamus and the pituitary gland and at higher brain centers. Glucocorticoids exert these actions via their glucocorticoid receptors and the rigorous regulation of a plethora of genes that are directly or indirectly implicated in several metabolic pathways, either activating them through positive glucocorticoid responsive elements (GREs) (e.g., genes encoding enzymes of glucose, lipid and amino acid metabolism, lipoprotein receptors, leptin) or repressing them through negative GREs (e.g., CRH, proopiomelanocortin, various cytokines).

The Sympathetic Nervous System

The autonomic nervous system provides a rapidly responding mechanism to control a wide range of functions. Cardiovascular, respiratory, gastrointestinal, renal, endocrine, and other systems are regulated by the sympathetic nervous system, the parasympathetic system, or both. Sympathetic innervation of peripheral organs is derived from the efferent

preganglionic fibers, whose cell bodies lie in the intermediolateral column of the spinal cord. These nerves synapse in the bilateral chains of sympathetic ganglia, with postganglionic sympathetic neurons that richly innervate the smooth muscle of the vasculature, the heart, skeletal muscles, kidney, gut, fat, and many other organs. The sympathetic system also has a humoral contribution, providing most of the circulating epinephrine and some of the norepinephrine from the adrenal medulla (Figure 1).

Consequences of Chronic Stress

Generally, the stress response with the resultant activation of the HPA and the sympathetic nervous system axes is meant to be acute or at least of a limited duration. The time-limited nature of this process renders its accompanying antireproductive, antigrowth, catabolic, and immunosuppressive effects temporarily beneficial rather than damaging. In contrast, chronicity of stress system activation would lead to the syndromal state that Selye described in 1936.

The prototypic example of chronic hyperactivation of the stress system is manifested in melancholic depression. Owing to a chronically hyperactive stress system, patients with melancholic depression may sustain several severe somatic sequelae, such as osteoporosis, features of the metabolic syndrome (e.g., visceral fat accumulation, dyslipidemia, increased blood pressure, impaired glucose tolerance), varying degrees of atherosclerosis, immunosuppression, and certain infectious and neoplastic diseases. When not treated, these patients have a compromised life expectancy curtailed by 15–20 years after excluding suicides.

In addition to melancholic depression, a spectrum of other conditions may be associated with increased and prolonged activation of the HPA axis, including childhood sexual abuse, anorexia nervosa with or without malnutrition, obsessive-compulsive disorder, panic anxiety, chronic active alcoholism, and alcohol and narcotic withdrawal.

More specifically, in the context of a mobilized adaptive stress response, glucocorticoids exert a broad spectrum of primarily catabolic effects as part of a generalized effort to utilize every available energy resource against the challenge enforced by the stressors. Thus, glucocorticoids increase hepatic gluconeogenesis and elevate plasma glucose concentrations, induce lipolysis (although they favor abdominal and dorsocervical fat accumulation), and cause protein degradation at multiple tissues (e.g., muscle, bone, skin) to provide amino acids that would be used as an additional substrate for oxidative pathways. In addition to their direct catabolic actions, glucocorticoids also antagonize the beneficial anabolic actions of growth and thyroid hormones, insulin, and sex steroids on their target tissues ([Figure 2](#)). This stress-related shift of the metabolism toward a catabolic state normally reverses upon retraction of the posed stressor. Chronic exposure to stress and HPA axis activation, however, is expected to progressively

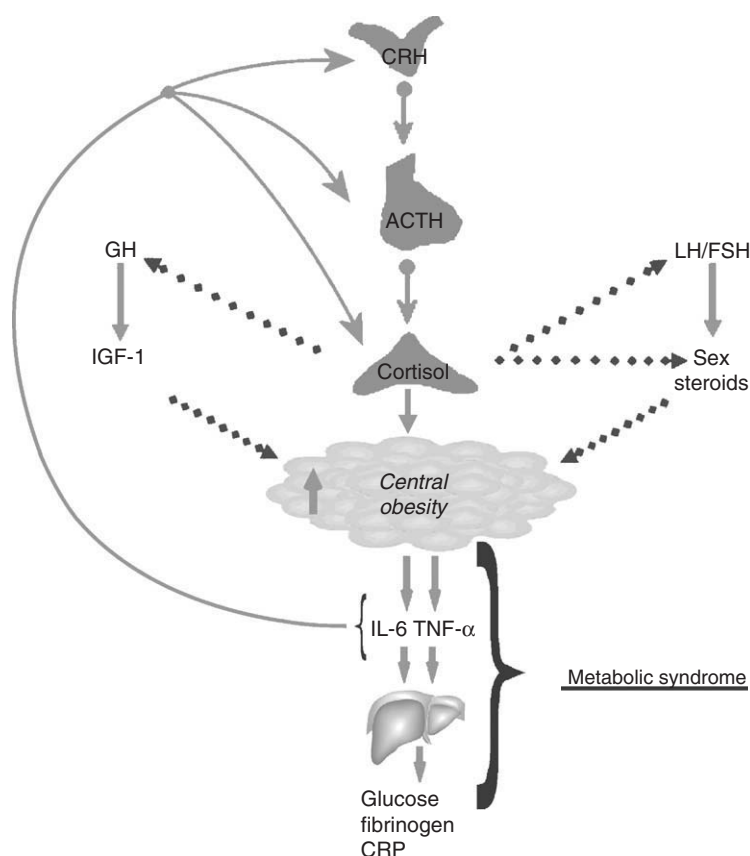


Figure 2 Schematic representation of the detrimental effects of chronic stress on adipose tissue that may lead to clinical manifestations of the metabolic syndrome. Chronic HPA axis activation favors visceral obesity, with cortisol suppressing the beneficial effects of both the gonadal and growth hormone axis and directly stimulating adipocyte proliferation. Hypersecretion of TNF- α and IL-6 is proportional to the cytokines further stimulate the HPA axis activation, thus forming a deleterious vicious cycle. CRH, corticotropin-releasing hormone; ACTH, adrenocorticotrophic hormone; TNF- α , tumor necrosis factor α ; IL-6, interleukin-6; CRP, C-reactive protein. Stimulatory effects are represented by solid arrows and inhibitory effects by dashed arrows.

increase visceral adiposity, decrease lean body (muscle and bone) mass, and cause insulin resistance (Figure 2). Interestingly, the phenotype of Cushing's syndrome, characterized by abdominal and trunk fat accumulation and decreased lean body mass, in combination with various other manifestations of the metabolic syndrome (insulin resistance, dyslipidemia, hypertension, hypercoagulability, and hypercytokinemia), is present in a variety of pathophysiological conditions, categorized as pseudo-Cushing's states. This cushingoid phenotype, under certain conditions of chronic stress, could be attributed to stress-induced mild hypercortisolism, possibly combined with increased peripheral tissue sensitivity to glucocorticoids, the latter depending on the tissue-specific concentration and activity of glucocorticoid receptors, cortisol-binding protein (CBG), and 11 β -hydroxysteroid dehydrogenase type 1 (11 β -HSD1). Interestingly, 11 β -HSD1 is an isoform expressed primarily in the liver, adipose tissue, kidneys, and brain that catalyzes the conversion of the inactive cortisone to active cortisol in humans and, thus, locally augments the concentration of the bioactive hormone. This tissue-specific amplification of the cortisol action, mediated by 11 β -HSD1, appears to be very important for determining the tissue sensitivity to glucocorticoids and hence to the HPA axis hyperactivation, especially under conditions that potentiate 11 β -HSD1 activity, such as visceral obesity.

Depression and Diabetes

Until very recently, the attention of the scientific community had been focused on the emotional consequences of diabetes mellitus, rather than on the emotional causes of the disease. Depression rates are roughly twice as high among persons with diabetes as among persons without diabetes and are higher than those for most other chronic physical illnesses. A meta-analysis of 42 eligible studies showed that the presence of diabetes doubles the odds of comorbid depression, a stable finding in both controlled and uncontrolled studies not affected by sex, subject source, assessment method, or type of diabetes. Yet most of these findings are based on cross-sectional data, so they actually tell us very little about causal relationships.

Recently, however, significant evidence from population-based prospective studies in the United States and Japan shows that being depressed at baseline doubles the risk of subsequently developing type II diabetes and that this association is not explained by other risk factors for type II diabetes, including obesity, activity level, smoking, alcohol consumption, and family history of type II diabetes. Thus, in some

circumstances, depression may precede the onset of type II diabetes and may play an important role in its development. As described previously, depressive disorders are accompanied by increased sympathetic and sympatho-adrenal system activity, which is known to be associated with increased blood glucose and impaired glucose tolerance. Depressive disorders have also been associated with increased activity of the HPA axis, which results in an increased release of glucocorticoids, decreased glucose uptake, and elevated glucose levels. Depression may impair the ability to handle a carbohydrate load and then increase the risk of developing type II diabetes through a greater release of the counterregulatory hormones. More importantly, depression is associated with visceral fat accumulation, which is considered the key abnormality underlying the development of insulin resistance and the increased risk for type II diabetes and atherosclerosis. Changes in diet (e.g., unhealthy eating, alcohol consumption) and physical activity associated with chronic depression may also contribute to the glucose abnormalities that may develop.

Interestingly, effective treatment of depression in diabetic patients appears to produce a better glycemic control that is not due to differential effects of treatment on weight. Moreover, in controlled trials of cognitive-behavioral therapy, related improvements in depression produced significant improvements in glycemic control.

Stressful Life Events and Diabetes

Several studies have demonstrated a relationship of stress to glycemic control in samples of patients with type II diabetes. Stress can be managed through the use of behavioral stress management programs or through the administration of anxiolytic medications, both of which have been reported to improve long-term glycemic control in patients with type II diabetes. It is of note that staff-intensive individual training is not always required for a successful behavioral intervention, as group programs in a real-world setting can also result in clinically significant benefits for patients with type II diabetes.

A large recent study (the Hoorn Study) showed that a high number of rather common major life events, consistent with chronic psychological stress, during the past 5 years was indeed related to a higher prevalence of previously unknown type II diabetes, as well as to increased levels of visceral fat accumulation. These associations were independent of family history of diabetes, physical activity, heavy alcohol use, or low level of education. Furthermore, behavioral stress management programs, whether individualized or in groups, usually help.

Inflammation and Diabetes

It has been known for several decades that inflammatory stress is associated with concurrent activation of the HPA axis. Indeed, cytokines and other humoral mediators of inflammation are potent activators of the central stress response, constituting the afferent limb of the feedback loop through which the immune/inflammatory system and the central nervous system communicate (Figure 3).

All three inflammatory cytokines tumor necrosis factor α (TNF- α), interleukin-1 β , and interleukin-6 (IL-6) can cause stimulation of the HPA axis alone, or in synergy with each other. There is evidence to suggest that IL-6, the main endocrine cytokine, plays the major role in the immune stimulation of the axis. Some of the activating effects of cytokines on the HPA axis may be exerted indirectly by stimulation of the central catecholaminergic pathways. Also, activation of peripheral nociceptive, somatosensory, and visceral afferent fibers would lead to stimulation of both the catecholaminergic and CRH neuronal systems via ascending spinal pathways.

In addition, TNF- α interferes directly with insulin signaling causing insulin resistance, and, together with the other inflammatory cytokines, has direct actions on the endothelium promoting the atherogenic process. Interestingly, TNF- α is overproduced by adipose tissue in obesity and appears to be one of the

molecular links between obesity insulin resistance/diabetes and atherosclerosis. Increased production of other adipokines, such as resistin, visfatin, and leptin, and decreased production of adiponectin with expanding obesity, further increase the inflammatory load and the risk for insulin resistance and atherosclerosis.

Indeed, an increasing body of evidence supports a possible role for inflammation in the pathogenesis of type II diabetes. Thus, in the Women's Health Study, an ongoing U.S. primary prevention randomized clinical trial initiated in 1992, elevated levels of C-reactive protein (CRP) and IL-6 predicted the development of type II diabetes. Other common markers of inflammation (fibrinogen, haptoglobin, white-cell count, sialic acid) are also associated prospectively with the development of diabetes in adults. Furthermore, chronic subclinical inflammation relates independently to insulin resistance, even in nondiabetic subjects. It is therefore likely that anti-inflammatory therapy may be beneficial in improving insulin resistance and preventing the development of type II diabetes. Interestingly, thiazolidinediones, a new class of insulin sensitizers, also have anti-inflammatory properties, reducing CRP and TNF- α levels. Moreover, the observed significant reductions in the hazard of becoming diabetic with statins may be attributed at least partly to the anti-inflammatory properties of these drugs.

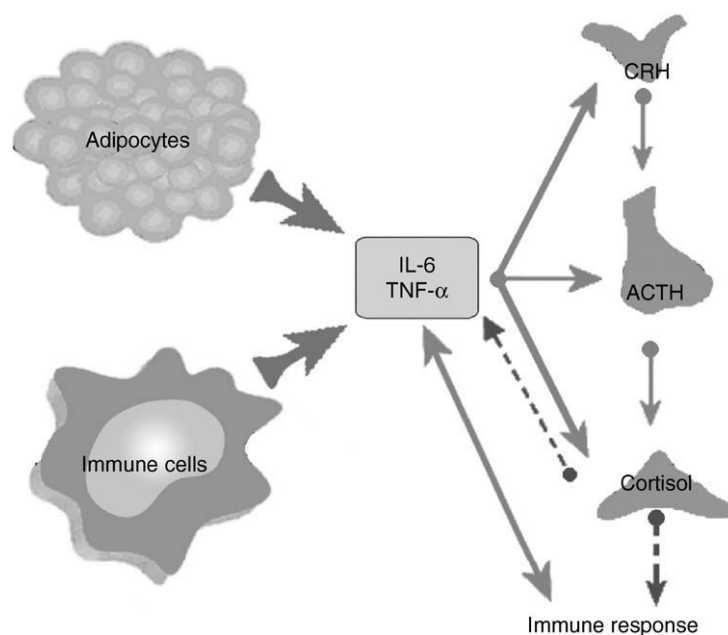


Figure 3 Schematic representation of the interactions between the pro-inflammatory cytokines, the HPA axis, and the immune system response. TNF- α and IL-6 secretion by the adipocytes is continuous and proportional to body mass index (BMI). TNF- α and IL-6 secretion by the immune cells at inflammatory sites occurs in a cascade-like fashion, with TNF- α appearing first followed by IL-6. TNF- α , tumor necrosis factor α ; IL-6, interleukin-6; CRH, corticotropin-releasing hormone; ACTH, adrenocorticotrophic hormone. Stimulatory effects are represented by solid arrows and inhibitory effects by dashed arrows.

Conclusion

Chronic stress seems to contribute to the development of visceral obesity and, hence, of type II diabetes and atherosclerosis, and, conversely, the accumulation of visceral adipose tissue constitutes a chronic stressful state. The mechanisms that mediate this documented reciprocal relation and cause the derangement of the metabolic homeostasis are the subject of many recent research efforts, and their characterization will hopefully provide novel insights into the pathogenesis of type II diabetes and cardiometabolic complications and help the quest for more specific and effective therapeutic interventions. Moreover, treating depression in and providing effective stress management programs for people predisposed to diabetes may add a new dimension of primary prevention of type II diabetes to the treatment of traditional risk factors such as obesity, diet, and physical inactivity.

See Also the Following Articles

11 β -Hydroxysteroid Dehydrogenases; Atherosclerosis; Metabolic Syndrome; Metabolic Syndrome and Stress; Inflammation.

Further Reading

- Anderson, R. J., Freedland, K. E., Clouse, R. E. and Lustman, P. J. (2001). The prevalence of comorbid depression in adults with diabetes. A meta-analysis. *Diabetes Care* **24**, 1069–1078.
- Barzilay, J. I., Abraham, L., Heckbert, S. R., et al. (2001). The relation of markers of inflammation to the development of glucose disorders in the elderly. The Cardiovascular Health Study. *Diabetes* **50**, 2384–2389.
- Bjorntorp, P., Holm, G. and Rosmond, R. (1999). Hypothalamic arousal, insulin resistance and type 2 diabetes mellitus. *Diabetic Medicine* **16**, 373–383.
- Buren, J. and Eriksson, J. W. (2005). Is insulin resistance caused by defects in insulin's target cells or by a stressed mind? *Diabetes/Metabolism Research and Reviews* **21**, 487–494.
- Chrousos, G. P. (1995). The hypothalamic-pituitary-adrenal axis and immune-mediated inflammation. *New England Journal of Medicine* **332**, 1351–1362.
- Chrousos, G. P. (2000). The role of stress and the hypothalamic-pituitary-adrenal axis in the pathogenesis of the metabolic syndrome: neuro-endocrine and target tissue-related causes. *International Journal of Obesity* **24**(Supplement 2), S50–55.
- Chrousos, G. P. and Gold, P. W. (1992). The concepts of stress system disorders: overview of behavioral and physical homeostasis. *Journal of the American Medical Association* **267**, 1244–1252.
- Moller, D. E. and Kaufman, K. D. (2005). Metabolic syndrome: a clinical and molecular perspective. *Annual Reviews in Medicine* **56**, 45–62.
- Rubin, R. R. and Peyrot, M. (2002). Was Willis right? Thoughts on the interaction of depression and diabetes. *Diabetes/Metabolism Research and Reviews* **18**, 173–175.
- Schmidt, M. I., Duncan, B. B., Sharrett, A. R., et al. for the ARIC Investigators (1999). Markers of inflammation and prediction of diabetes mellitus in adults (Atherosclerosis Risk in Communities study): a cohort study. *Lancet* **353**, 1649–1652.
- Seckl, J. R., Morton, N. M. and Chapman, K. E. (2004). Glucocorticoids and 11 β -hydroxysteroid dehydrogenase in adipose tissue. *Recent Progress in Hormone Research* **59**, 359–393.
- Surwit, R. S., van Tilburg, M. A. L., Zucker, N., et al. (2002). Stress management improves long-term glycemic control in type 2 diabetes. *Diabetes Care* **25**, 30–34.
- Tsigos, C. and Chrousos, G. P. (2002). Hypothalamic-pituitary-adrenal axis, neuroendocrine factors and stress. *Journal of Psychosomatic Research* **53**, 865–871.
- Tsigos, C. and Kyrrou, I. (2006). Hypothalamic-pituitary-adrenal axis, cytokines and the metabolic syndrome. *Obesity and Metabolism* **2**, 116–126.
- Tsigos, C., Kyrrou, I. and Chrousos, G. P. (2005). Stress, endocrine manifestations, and diseases. In: Cooper, C. L. (ed.) *Handbook of stress, medicine, and health* (2nd edn., pp. 101–131). Boca Raton, FL: CRC Press.

Stress, NPY, and Cardiovascular Diseases

Z Zukowska

Georgetown University Medical School, Washington, DC, USA

© 2007 Elsevier Inc. All rights reserved.

Structure and Localization of NPY in the Body

The NPY Receptor System and Major Actions

Sources of NPY Release in the Circulation

Stress-Induced Release of NPY from Peripheral Sympathetic Nerves

Chronic Stress, NPY, and Vascular Remodeling after Balloon Angioplasty

Critical Role of Y1 Receptors in Immune System and Angioplasty-Induced Restenosis

Chronic Stress and Angiogenesis: Implications for Ischemic Revascularization and Adipogenesis

Summary and Conclusion

Glossary

<i>Adipogenesis</i>	The growth of adipose tissue, increasing fat mass.
<i>Angiogenesis</i>	The formation of new vessels from preexisting ones; called ischemic angiogenesis in areas of poor blood flow (ischemia).
<i>Atherosclerosis</i>	The inflammatory response of blood vessels; the buildup of atherosclerotic plaques causing a lesion that contains the abnormal growth of vascular smooth muscle cells (neointima), deposition of matrix proteins, aggregation of platelets (thrombus formation), and formation of lipids.
<i>Balloon angioplasty</i>	A procedure used by cardiologists that involves the insertion of a catheter whose balloon can be inflated to re-open an artery that was closed (occluded) by an atherosclerotic plaque.
<i>Macrophages</i>	Immune cells with phagocytotic capabilities.
<i>Neuropeptide Y (NPY)</i>	A 36-amino-acid peptide present primarily in the neurons of the brain and in the peripheral nervous system, rich in tyrosine (Y, in the one-letter code used for amino acids).
<i>Obesity</i>	Increased fat mass and weight above 30% of normal.
<i>Restenosis</i>	The repeated closure of a vessel previously opened by balloon angioplasty or by other vascular procedures such as the insertion of a stent.

Stress

Sympathetic nerves

Vascular remodeling

The real or perceived alteration of internal or external environment that elicits bodily responses (stress responses).

Part of the autonomic or vegetative nervous system that innervates most blood vessels and organs of the body; essential in body's responses to stress.

Any change to vessel structure such as the excessive or abnormal growth of vascular smooth muscle in the medial layer of the vessel or the sprouting of new vessels (angiogenesis).

Structure and Localization of NPY in the Body

Neuropeptide Y (NPY) is a 36-amino-acid brain peptide, one of the most abundant in the body. It is formed from a 97-amino-acid peptide, preproNPY, which contains a signal peptide targeting the precursor molecule to the endoplasmic reticulum. Specific convertases subsequently cleave the preproNPY to form a mature and amidated NPY(1–36), which is biologically active. An additional processing product is the C-terminal peptide of NPY (CPON), a peptide whose actions are relatively unknown. Inside neurons, mature NPY(1–36) is stored in large dense core vesicles and emptied into the synaptic cleft during nerve depolarization. Recent research shows that, in addition to its neuronal sources, NPY is expressed extraneuronally in the endothelial cells and platelets of some species, where it acts as an autocrine and paracrine regulator of cell functions.

In the central nervous system, NPY is richly expressed in the hypothalamus, hippocampus, brain stem, and cortex, and it is colocalized with classical neurotransmitters such as catecholamines and GABA. In the periphery, NPY is present in all the sympathetic nerves innervating the blood vessels and in the chromaffin cells of the adrenal medulla, where it is colocalized with norepinephrine (NE) and epinephrine, respectively. The most abundant vascular NPY-ergic innervation is found in small arteries in the kidneys and the mesenteric bed. The third, least source of NPY in the body is the gut parasympathetic and enteric nerves, in which NPY is present in small and large intestines. In the gastrointestinal tract, two other peptides belong to the same family as NPY: peptide YY (PYY) and pancreatic polypeptide (PP). They act as hormones rather than as neurotransmitters; hence, they are not a major focus of this review.

The whole NPY peptide family evolved from an ancient ancestor gene and has maintained a high degree of conservation of structure. NPY, in particular, is 90–99% homologous among a wide range of species; for example, rat and human NPYs are identical. The critical C-terminal portion of NPY, as well as PYY, which shares most of its affinity for its receptors, has been 100% conserved throughout evolution, suggesting that it plays an important physiological role.

The NPY Receptor System and Major Actions

Determining what NPY's physiological role is has been a challenge. Early Swedish studies suggested potent cardiovascular effects for NPY – arterioconstriction and increased blood pressure, which are resistant to the blockade of α -adrenergic receptors. In a landmark experiment in 1982 on the cat submandibular gland, Lundberg et al. showed that the effect of NPY is mimicked by prolonged sympathetic nerve stimulation in the presence of total adrenergic blockade and suggested that NPY is a mediator of nonadrenergic sympathetic transmission. Determining the significance of these actions in the cardiovascular system required the cloning of the NPY receptors and the development of selective receptor inhibitors.

The first receptor that was cloned for rat, mouse, and humans was the Y1 receptor. It is a Gi-coupled receptor that is predominant in the blood vessels, where it mediates vasoconstriction and the potentiation of vasoconstriction exerted by NE and other agents. The Y1 receptor is also widely expressed in neurons in the central and peripheral nervous systems, where it mediates a range of functions from anxiolysis and appetite stimulation to neuronal cell proliferation and analgesia. Recently, this receptor was found to be present in T- and antigen-presenting cells, where it regulates the proliferation and functions of these cells. The Y1 receptor requires the whole NPY(1–36) molecule for activation; therefore, peptidases that cleave N-terminal amino acids inactivate Y1 receptor-mediated activities (i.e., act as endogenous Y1 antagonists). The major enzyme in this group is dipeptidyl peptidase IV (DPPIV), which is present ubiquitously in endothelial, immune cells and macroglia, and forms NPY(3–36). The other known protease is aminopeptidase P (APP), expressed on vascular smooth muscle and other cells, which forms NPY(2–26).

On cleavage, C-terminal fragments of NPY can be detected in the circulation and remain active, but

change their receptor affinities and become poor activators for or are inactive at Y1 while having a high affinity for the other, non-Y1 receptors. Up to five more receptors have been cloned so far: Y2, Y4, Y5, and Y6. All of them are Gi-coupled. The remaining Y3 receptor has been characterized pharmacologically as the one that is NPY-preferring, that is, not activated by PYY; it has not been cloned. Although a number of tissues appears to display such a pharmacological profile, for example, the brain stem, this receptor is now considered nonentity.

Y2 receptors are the predominant neuronal presynaptic receptors in the brain and the periphery, where they inhibit neurotransmitter release. They act as NPY's own autoreceptors as well as inhibiting the release of NE and other neurotransmitters such as ATP. Interestingly, the Y2-mediated inhibition appears to increase at a higher frequency of neuronal activity. In the central nervous system, the inhibitory effects of Y2 receptors on NPY release often result in the inhibition of either Y1- or Y5-receptor-mediated responses. For that reason, Y2 receptor antagonists are now emerging as the preferred drugs for several mental disorders (e.g., anxiety and depression) because they allow targeted therapy to the sites of overstimulation and limit other widespread effects in the brain that Y1- or Y5-receptor antagonists could otherwise have.

Postsynaptic Y2 receptors have also been found in the blood vessels, where they regulate the vascular tone; these changes range from vasoconstriction to vasodilation and angiogenesis, depending on the release from the endothelium of secondary mediators, such as nitric oxide. Whereas Y1 and Y2 receptors are constitutively expressed in unstimulated tissues, Y5 receptors appear not to be expressed alone and are often induced on demand in response to cellular stress. Initially, the Y5 receptor has been cloned from the hypothalamus as a feeding receptor, mediating NPY's action to stimulate appetite. It shares with the Y1 receptor the same promoter sequence, although read in the opposite direction, and appears to be regulated in concert with the Y1. Recently, we reported that Y5 is inducible in endothelial and vascular smooth muscle cells by tissue ischemia and injury.

Sources of NPY Release in the Circulation

The role of NPY as a mediator of nonadrenergic sympathetic transmission has been well documented by several groups in early studies both *in vitro* and *in vivo*. It became generally accepted that the peptide is released by more intense or prolonged nerve stimulation than that required for the release of NE. The

differential release of NPY and NE has also been seen using various stressors as stimulators of sympathetic nerve activity. A plausible explanation has been the different distribution of these neurotransmitters in the secretory granules, with NPY being stored in the large and NE in the small dense core vesicles. Although NE is also stored in the large vesicles, those constitute only a minor portion of the overall adrenergic pool and hence are minor determinants of the amount of released NE. However, recent studies with cultured cells have demonstrated that NPY is secreted from neuronal and non-neuronal cells at an even lower intensity of stimulation than seen before. In addition, NPY has been found to exert biological effects at extremely low concentrations (submolar to picomolar range). The *in vivo* detection of such minimal amounts of the peptide is clearly beyond the limits of currently used methods and, hence, cannot be documented at the present time.

Radioimmunoassays and enzyme-linked immunoassays have been the most commonly used methods for detecting circulating and tissue-derived NPY. Plasma NPY-immunoreactivity (NPY-ir) was found to vary among the species, with human and porcine levels being the lowest (picomolar range). Rats and some mice have higher NPY levels (submolar to nanomolar range) due to the expression of NPY extraneuronally in megakaryocytes and platelets. In these animals, NPY-containing platelets become a transport system delivering the peptide to the sites of platelet activation. Although healthy humans do not have NPY in platelets, some diseases such as atherosclerosis appear to induce its expression. The significance of these findings remains to be determined; however, animal studies suggest that platelets, as well as NPY, play a critical role in vascular restenosis after balloon angioplasty. Interestingly, the levels of NPY-ir in the systemic arterial and the coronary sinus plasma increase during angina due to either myocardial ischemia or panic attack, indicating that intracardiac rise in NPY may contribute to vascular occlusion and be involved in thrombus formation, vasospasm, and vascular remodeling.

The third source of circulating NPY is the endothelium; however, a controversy still exists about whether these cells actually secrete the peptide into the bloodstream. NPY is an autocrine growth factor and stimulates endothelial cell migration, proliferation, and differentiation into capillaries by an action mediated primarily by Y2 and Y5 receptors, with an active role played by DPPIV (components of NPY's angiogenic system are expressed in these cells). Ischemia and cellular stress both upregulate NPY, Y2 and Y5 receptors, and DPPIV.

Stress-Induced Release of NPY from Peripheral Sympathetic Nerves

The highest plasma NPY-ir levels in humans were noted in newborns after vaginal delivery. This is probably due to the stress of delivery; babies born by cesarean section have significantly lower peptide levels. With aging, circulating NPY levels in humans and in animals decrease, but stress continues to stimulate its release. Various stressors have been studied and found to exert differential effects on plasma NPY, by itself and in relation to catecholamines. Of those, the strongest stimulus for NPY release appears to be exposure to cold temperature and strenuous exercise, particularly when combined with hypoxia. Interestingly, there is a gender difference in the response to cold and/or treadmill exercise in both humans and rodents. Although there is no sexual dimorphism in the basal circulating NPY levels, males increase their plasma NPY-ir levels more than females following both stressors. In rats and in humans, this does not appear to be due to estrogens because ovariectomy and estrogen replacement has no effect on cold- and exercise-induced NPY responses. Subsequent studies in gonadectomized rats confirmed that the gender difference in the NPY system is due to testosterone, which strongly upregulates NPY gene expression. Sex differences in stress-induced plasma NPY responses were paralleled by greater hypertensive effects of stress in males, implying that the peptide may play a role in essential hypertension, which is more frequent in men than in premenopausal women.

Interestingly, other commonly used laboratory stressors such as immobilization stress in rats or mental arithmetic in humans failed to yield measurable changes in circulating NPY-ir levels using current analytical methods. Interestingly, a rise in plasma NPY levels in humans during bicycle exercise is augmented by the blockade of α -2-adrenergic receptors, indicating that NPY release is under physiological (presynaptic) inhibition by these receptors. The depletion of catecholamines from the sympathetic storage vesicles by reserpine was also shown to increase the NPY-dependent component of the neurogenic response to nerve stimulation (after a transient fall in blood pressure due to the removal of adrenergic vascular tone). Ganglionic blockade with chlorisondamine eliminated the NPY release and resulted in a complete loss of sympathetic tone. This experimental protocol elegantly demonstrated that although NPY may not be the primary neurotransmitter maintaining normal vascular tone and blood pressure, it is the alternative pathway to adrenergic transmission and is one that can either compensate for the loss of it or

amplify or dominate over the adrenergic response. An example of the former may be sympathetic dysautonomia, in which patients have impaired adrenergic transmission but are able to compensate for it; an example of the latter may be situations in which NPY signaling is selectively enhanced by cold exposure and a high-fat and high-sugar diet.

Chronic Stress, NPY, and Vascular Remodeling after Balloon Angioplasty

Although the examples so far illustrate the ability of the body to secrete NPY in response to acute stimulation of the sympathetic nerves, the role of this peptide neurotransmitter as a stress mediator can be understood best in chronic stress. In several diseases in which sympathetic activity is increased, circulating NPY-ir levels are markedly elevated. Some of the highest levels in humans were found in congestive heart and renal failure, in which they strongly correlated with and might be predictors of end-organ damage.

The actions of NPY also differ when it is applied chronically. In the cardiovascular system, acutely, NPY causes a slow onset, long-lasting vasoconstriction and increase in blood pressure at nanomolar concentrations consistent with the levels yielded by acute intense stress, for example, the immersion of a hand in ice water for humans or standing in ice-cold water for rodents. *In vitro*, such NPY concentrations applied for 1–14 days to cultured vascular smooth muscle and endothelial cells stimulates their migration, proliferation, differentiation, and release of secondary growth factors, leading to sustained cell growth. In these *in vitro* conditions, NPY also acts at lower than vasoconstrictive concentrations (submolar to picomolar range) and acts as an autocrine and paracrine growth factor. In *ex vivo* cultured aortic rings, NPY is potently angiogenic and leads to the formation of long vascular sprouts, resembling normal vessels. For some cells, such as neuroblastomas (tumor cells derived from sympathetic neurons), NPY is essential as a cell growth factor, whereas for others, such as Ewing's sarcoma, NPY is apoptotic. Thus, growth-modulatory actions are mediated by its multiple receptors and the final effect depends on the type of receptors and cell-specific signaling pathways. Whereas the growth-promoting effects of the peptide on vascular smooth muscle are mediated by Y1 receptors, NPY's angiogenic activities are mediated primarily by Y2 receptors and require the conversion of NPY(1–36) to NPY(3–36). What is the significance of these actions *in vivo*? Would the chronic stress-induced release of NPY mimic the peptide's effects seen *in vitro* and, if so, which ones?

We approached these questions using cold-water stress (animals standing in 0.5–1 cm of ice-cold water at the bottom of their cages) combined with experimental conditions known to trigger vascular remodeling. In rats, balloon angioplasty-induced arterial injury or intermittent chronic cold by itself within 2 weeks increased the platelet NPY-ir levels, but only cold stress elevated the plasma NPY. This suggested that, although plasma NPY responses are probably due to sympathetic nerve activation, platelet NPY may reflect changes originating in the bone marrow, in the megakaryocyte synthesis of NPY. This notion has been supported by studies using mice expressing or not expressing NPY in megakaryocytes. Both rats and mice, which have additional source of NPY in the bone marrow, developed neointima in the angioplasty-injured vessels, whereas mice deficient in megakaryocyte/platelet NPY were resistant to vascular injury and did not develop neointima. Whether platelet NPY is critical for neointima formation in rodents, and in humans, is currently being investigated.

NPY that is exogenous or released from sympathetic nerves by cold stress, on the other hand, augments neointima in all rodent strains, with or without platelet NPY, although the extent of the amplification is significantly greater in the former. In angioplasty-injured rat carotid arteries, NPY and stress alike induced the occlusion of the angioplasty-injured vessel with a lesion, which contained neovascularized neointima, thrombus, matrix, lipids, and increased invasion of macrophages – the elements of advanced atherosclerotic lesion. This occurred in an animal devoid of any lipid abnormalities within only 2 weeks and was completely prevented by a selective Y1 receptor antagonist. How it is possible for an antagonist of only one receptor to block such a complex process as restenosis and atheroma-like lesion formation is indeed puzzling. The critical role of Y1 receptors, however, is better understood when we consider the environment into which NPY is released following sympathetic nerve activation and vascular injury.

Critical Role of Y1 Receptors in Immune System and Angioplasty-Induced Restenosis

Chronic stress exerts direct effects not only in the cardiovascular system but also on the immune cells, commonly manifested as a greater incidence of infection during periods of stress. Although many mediators, including catecholamines and glucocorticoids, contribute to this effect, the role of NPY has not been considered until recently. Wheway et al. recently

reported that NPY and its Y1 receptors are essential for the normal functioning of T cells. The direct stimulation of Y1 receptors on T cells leads to the suppression of their Th1-type response and decreased secretion of interferon- γ and tumor necrosis factor- α . For that reason, Y1 receptor agonists are protective against Th1-mediated inflammatory conditions such as experimental autoimmune encephalitis, a model of multiple sclerosis. However, NPY can also exert stimulatory actions on immune cells via the same type of Y1 receptors on antigen-presenting cells. The loss of that response early during development in germline Y1 receptor knockout mice led to the impaired functioning of antigen-presenting cells, reduced number of functional T cells, and, hence, lowered Th1-mediated responses. Paradoxically, Y1 receptor deficiency protects against colitis, which depends on Y1-mediated activation of antigen-presenting cells.

The central role that the Y1 receptor plays in the immune system is due to a widespread distribution of that receptor, not only in vascular smooth muscle cells but also in all immune cells, including T cells, macrophages, neutrophils, and eosinophils. NPY activates monocytes and stimulates their phagocytosis, migration, and proliferation. This is likely to have implications outside infectious or autoimmune diseases because inflammatory responses are now considered important pathogenic mechanisms in many other conditions, including atherosclerosis, ischemic angiogenesis, and obesity. Thus, we can imagine that the stress-induced activation of the sympathetic nerves during vascular injury may not only exert direct effects on vascular smooth muscle and platelets but may also activate immune cells residing in the vascular wall and stimulate their transmigration from the bloodstream. Because all these cells express Y1 receptors, it is likely that nerve-released NPY starts a cascade of events by activating this receptor on cells adjacent to the nerve terminals. In addition, the Y1 receptors are upregulated by angioplasty, and this increase is further amplified by stress; as a result, Y1 receptors become widely expressed in neointima, macrophages, and other immune cells.

Because the Y1 receptor antagonist prevented the entire process of restenosis in rodent models of balloon angioplasty, we can also predict that Y1 deficiency is protective against stress-induced restenosis and, that Y1 receptor antagonists may be a new class of drugs for this condition. Our studies also suggest that stress is a risk factor for restenosis, facilitating the formation of atheroma and vascular hyperplasia even in the absence of hyperlipidemia. Such striking changes as those induced by stress have never been previously induced in animals without genetic manipulations of lipid metabolism. Although more

studies are needed to establish the clinical relevance of these findings, two lines of evidence strongly suggest that high NPY levels, and NPY release by stress, play a role in human atherosclerosis. First, an NPY signal peptide polymorphism, Leu7Pro, a common gain-of-function mutation occurring in northern European populations was found to lead to higher plasma NPY levels in response to stress, along with hyperlipidemia and accelerated atherosclerosis. Second, the induction of NPY in platelets was recently discovered in patients with atherosclerosis, similar to what previously was reported to occur in patients with severe depression. Current studies in our lab are being conducted to confirm whether stress in humans accelerates atherosclerosis by releasing NPY from the sympathetic nerves and/or platelets.

Chronic Stress and Angiogenesis: Implications for Ischemic Revascularization and Adipogenesis

If stress releases NPY from nerves and/or platelets, what determines whether the peptide exerts vasoconstrictive vasoproliferative effects leading to vascular hyperplasia via its Y1 receptors or is angiogenic leading to the increased vascularization of the tissue via its Y2 receptors (Figure 1)? We addressed this question first using a model of ischemic limb, in which

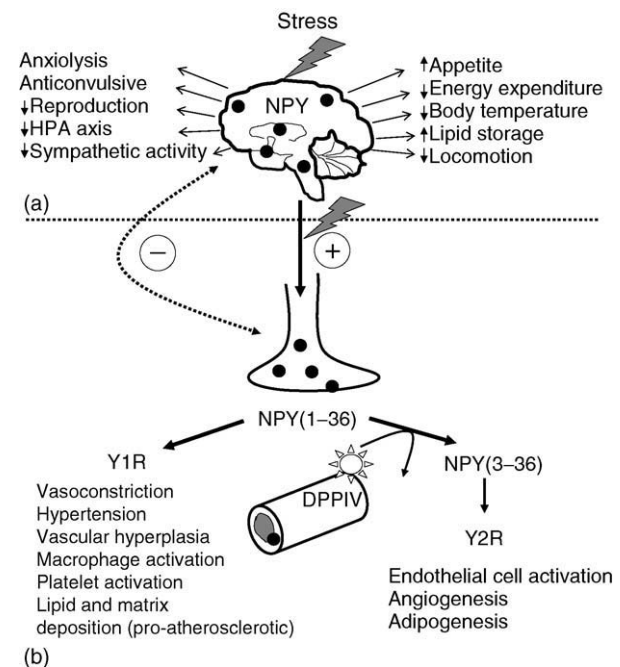


Figure 1 Effects of stress on central and peripheral NPY systems. a, Central nervous system effects; b, peripheral effects in sympathetic nerves. DPPIV, dipeptidyl peptidase IV; HPA, hypothalamic-pituitary-adrenal; Y1R, NPY receptor 1; Y2R, NPY receptor 2.

spontaneous revascularization is incomplete and produces a deficit in blood flow. Limb ischemia induced by unilateral femoral artery occlusion, like angioplasty described previously, elicited a local release of NPY from the sympathetic nerves. However, in contrast to the effects of angioplasty on the vascular wall, tissue ischemia also stimulated NPY synthesis in the endothelium and shifted the expression of its receptors in the ischemic limb toward Y2 and DPPIV, the elements of NPY's angiogenic system. The local administration of NPY at the site of ischemic skeletal muscles amplified these responses and, in addition, also induced Y5 receptors while augmenting angiogenesis and arteriogenesis in the limb. Exposure to cold, intermittently for 2 weeks, however, did not mimic NPY's actions and actually inhibited ischemia-mediated changes in NPY receptor expression. Although these studies are still in progress, the data suggest that NPY's angiogenic receptor system, NPY-Y2/Y5-DPPIV, is not sufficiently activated during stress. Yet unidentified factors appear to inhibit the NPY's angiogenic switch and prevent the peptide from fully expressing its angiogenic capabilities in the ischemic tissues. Given the fact that Y2 and Y5 receptors and DPPIV are expressed in the ischemic skeletal muscles, the administration of Y2/Y5 receptor agonists may become useful as a revascularization therapy. Alternatively, the activation of DPPIV, the converting enzyme, could serve as a way to shift stress-induced NPY release away from vasoconstriction and vascular hyperplasia via its Y1 receptors and toward Y2/Y5-mediated angiogenesis.

Such a switch appears to occur spontaneously in the white adipose tissue under the conditions of a high-fat and high-sugar diet; under this condition, stress causes a marked upregulation of NPY, Y2, and DPPIV. This leads to the activation of angiogenesis, which then promotes further tissue growth and, thus, may contribute to the development of adiposity and obesity. Previously, others have proposed using anti-angiogenic therapy for the treatment of obesity, but the physiological mechanisms specific to fat have not been uncovered. This new area of research should shed light on the pathogenesis of obesity, which has become an epidemic concurrently with the rise in the levels of daily stress. The relation between the two processes, stress and obesity, although anecdotally acknowledged, has never been established scientifically.

Summary and Conclusion

In addition to the roles of NPY in the cardiovascular system described here, the peptide exerts pleiotropic actions on essentially every tissue and the regulatory

system of the body. In the brain, NPY is potently anxiolytic, sedative, and anticonvulsive and inhibits locomotor activity – actions consistent with the anti-stress effects rather than the pro-stress actions that it exerts in the periphery (Figure 1). Those central actions have implicated the peptide in anxiety and mood disorders, in which NPY levels in the cerebrospinal fluid are decreased. The peptide is also one of the most powerful stimulants of food intake, particularly of carbohydrates, resembling stress-induced cravings for sweets, and thus plays a role in appetite disorders, both obesity and anorexia. At the same time, NPY modulates a number of hormonal systems, from the hypothalamic-pituitary-adrenocortical and hypothalamic-pituitary-gonadotropin axes to the secretion of growth hormone and insulin, slowing down puberty and suppressing reproductive functions (Figure 1). In addition to its peripheral actions, centrally, NPY also inhibits lipid metabolism and energy expenditure, lowers body temperature, and promotes fat storage (Figure 1).

The two sites of NPY's actions, the brain and the periphery, have been studied so far in separation and little is known about their reciprocal relationship. It appears, however, that, as with many other neuro-hormonal systems, central and peripheral NPY are also regulated by negative feedback loops, with central NPY being often inhibitory for peripheral NPY. A primary example of such a reciprocal inhibitory relationship is the inhibitory effect of hypothalamic NPY on peripheral sympathetic activity, lowering plasma NPY levels. In conditions such as posttraumatic stress disorder, resulting from severe trauma, the balance between the central and peripheral NPY responses appears to be disregulated at both brain and the peripheral levels. The regulation of the cardiovascular system by NPY in this and other stressful situations is clearly only a part of the patterned body responses to stress and the patterns that vary with different stressors.

The integrated response of the NPY system in many ways resembles the body's adaptation to hibernation. Several lines of evidence support that hypothesis. First, cold temperature is one of the most potent stimuli for NPY release. Second, in preparation for hibernation, an animal requires cardiovascular adaptation, increased blood pressure and heart rate, to allow it to search for food; such a response is provided by peripherally released NPY. Third, in the presence of carbohydrate-rich food, the NPY stress response also promotes lipid storage, which an animal preparing for hibernation needs. Once this is accomplished, sympathetic activity is expected to fall and the animal's bodily functions to slow down. NPY's central activities support such actions as well,

with its effects of decreased body temperature, energy expenditure, and reproductive functions.

The final intriguing support of the hibernation hypothesis is suggested by genetic epidemiology. A Leu7-to-Pro7 polymorphism in the NPY signal peptide has been found to be common in northern European populations. Interestingly, there is a strong north-to-south gradient of this mutation, which is absent in Africa, Asia, and South America (unless in descendants of northern Europeans). The Leu7Pro polymorphism associates with increased NPY responses to stress, hyperlipidemia, and obesity – which at some time could have afforded an evolutionary protection, improving survival in cold temperatures. Unfortunately, there are also unwanted effects of this mutation. It is linked to accelerated atherosclerosis, diabetic retinopathy and nephropathy, and alcoholism – perhaps the price that humans paid for not adjusting to the rules of nature, eating and drinking too much while exercising too little all year long.

The importance of NPY in physiology was questioned by the studies of the Palmiter group who showed that life without NPY, in NPY-knockout mice, is apparently normal. The mice eat, develop, and breed normally and do not have any specific phenotype. They do, however, exhibit a lower threshold for seizures and have higher preference for alcohol – possibly offering a new model for epilepsy and alcoholism. Yet the true contribution of NPY to animal physiology can only be appreciated when the adaptation to chronic stress is considered. When laboratory animals are housed in a constant environment, without any stress, their life may seem normal. However, if NPY-knockout mice are subjected to stress, particularly chronically and in combination with poor diet, their ability to develop proper cardiovascular, immune, and metabolic adaptations might be impaired. In years to come, more studies will be required to understand the full implications of NPY's release and activities in all of the body's sources: the brain, peripheral sympathetic nerves, and the enteric system, as well as its extraneuronal expression in platelets, endothelium, and immune cells.

Further Reading

- Chronwall, B. and Zukowska, Z. (2004). Neuropeptide Y, ubiquitous and elusive. *Peptides* 25, 359–363.
- Dumont, Y. and Quirion, R. (2006). An overview of neuropeptide TY: pharmacology to molecular biology and receptor localization. In: Zukowska, Z. & Feuerstein, G. Z. (eds.) *NPY family of peptides in neurobiology, cardiovascular and metabolic disorders: from genes to therapeutics*, pp. 7–35. Basel: Birkhauser Verlag.

- Dutton, M. A., Lee, E. W. and Zukowska, Z. (2006). NPY and extreme stress: lessons learned from posttraumatic stress disorder. In: Zukowska, Z. & Feuerstein, G. Z. (eds.) *NPY family of peptides in neurobiology, cardiovascular and metabolic disorders: from genes to therapeutics*, pp. 213–247. Basel: Birkhauser Verlag.
- Erickson, J. C., Clegg, K. E. and Palmiter, R. D. (1996). Sensitivity to leptin and susceptibility to seizures of mice lacking neuropeptide Y. *Nature* 381, 415–418.
- Heilig, M. (2004). The NPY system in stress, anxiety and depression. *Neuropeptides* 38, 213–224.
- Li, L., Rylander-Johnsson, A. C., Abe, K., et al. (2005). Chronic stress induces rapid occlusion of angioplasty-injured rat carotid artery by activating neuropeptide Y and its Y1 receptors. *Arteriosclerosis Thrombosis and Vascular Biology* 25, 2075–2080.
- MacNeil, D. J. and Kanatani, A. (2006). NPY and energy homeostasis: an opportunity for novel anti-obesity therapies. In: Zukowska, Z. & Feuerstein, G. Z. (eds.) *NPY family of peptides in neurobiology, cardiovascular and metabolic disorders: from genes to therapeutics*, pp. 143–157. Basel: Birkhauser Verlag.
- Malmstrom, R. E. and Lundberg, J. M. (1997). Neuropeptide Y in sympathetic nerves: evidence for Y1 receptor mediated vascular control. In: Grundemar, L. & Bloom, S. R. (eds.) *Neuropeptide Y and drug development*, pp. 41–49. London: Academic Press.
- McAuley, M. A., Chen, X. and Westfall, T. C. (1993). Central cardiovascular actions of neuropeptide Y. In: Comers, W. F. & Wahlestedt, C. (eds.) *The biology of neuropeptide Y and related peptides*, pp. 389–418. Totowa, NJ: Humana Press.
- Pesonen, U. (2006). Genetics of the NPY family of peptides and their receptors. In: Zukowska, Z. & Feuerstein, G. Z. (eds.) *NPY family of peptides in neurobiology, cardiovascular and metabolic disorders: from genes to therapeutics*, pp. 247–271. Basel: Birkhauser Verlag.
- Wheway, J., Mackay, C. R., Newton, R. A., et al. (2005). A fundamental bimodal role of neuropeptide Y1 receptor in the immune system. *Journal of Experimental Medicine* 202, 1527–1538.
- Zukowska, Z. (2005). Atherosclerosis and angiogenesis: what do nerves have to do with it? *Pharmacological Reports* 57, 229–234.
- Zukowska, Z., Grant, D. S. and Lee, E. W. (2003). Neuropeptide Y: a new angiogenic mechanism for ischemic angiogenesis. *Trends in Cardiovascular Medicine* 13, 86–92.
- Zukowska-Grojec, Z., Shen, G. H., Capraro, P. A., et al. (1991). Cardiovascular, neuropeptide Y, and adrenergic responses to stress are sexually differentiated. *Physiology and Behavior* 49, 771–777.
- Zukowska-Grojec, Z. and Wahlestedt, C. (1993). Origin and actions of neuropeptide Y in cardiovascular system. In: Colmers, W. F. & Wahlestedt, C. (eds.) *The biology of neuropeptide Y and related peptides*, pp. 315–388. Totowa, NJ: Humana Press.

Stress-Induced Neurotoxicity See: Glucocorticoids – Adverse Effects on the Nervous System.

Stroke, Stress and See: Brain Trauma.

Successful Adaptation to Extreme Stress See: Multiple Trauma; Sepsis, Acute Respiratory Distress Syndrome and Glucocorticoid Resistance; Acute Stress Disorder and Posttraumatic Stress Disorder; Acute Stress Response: Experimental.

Suicide Terrorism, Genesis of

A Speckhard

Georgetown University Medical Center, Washington, DC, USA

© 2007 Elsevier Inc. All rights reserved.

The History of Suicide Bombing
 The Success of Suicide Terrorism
 The Lethal Cocktail of Suicide Terrorism
 Sponsoring-Groups Motivations
 Ideologies of Suicide Terrorism – Religion and Cosmic Warfare
 Individual Motivations
 Adopting a Terrorist Ideology as Psychological First Aid
 Societal Support for Suicide Terrorism

Jihad

manner and may even include compartmentalizing the event from one's ambitions and daily life. One bomber for example, when planning his attack, suggested that he could not carry a bomb until after his university exams – suggesting that while on the one hand he acknowledged that he was going to his death he was able separate this reality so completely from his mind that he still felt the need to complete his exams prior to going to explode himself.

Is an Islamic word and makes reference to the teachings of the Prophet Muhammed as they are recorded in the Koran and Sunnah, in which believers are taught to struggle on two levels: the first and most important being referred to as the greater jihad in which an individual strives inwardly in an effort to resist evil and live a pure life according to the tenets of Islam. The lesser jihad involves a political or military struggle to further and/or defend Islam and Islamic lands and people. In the current global context certain ideologies promoting terrorism are referred to by some as promoting militant jihad. For example the US Department of Justice has referred to jihad as the use of violence, including paramilitary action against persons and governments deemed to be enemies of a fundamentalist version of Islam.

Glossary

Dissociation According to the American Psychiatric Association, the essential feature of a dissociative disorder is “a disruption in the usually integrated functions of consciousness, memory, identity or perception of the environment” – American Psychiatric Association (1994) *Diagnostic and statistical manual of mental disorders (4th ed.)*. Washington, D.C. In the case of suicide terrorism, this means an emotional barrier is unconsciously erected walling off the negative emotions generated by choosing to die in this

Posttraumatic stress disorder (PTSD)

Is a person's individualistic response to a traumatic event (see psychological trauma below) in which the following responses occur: an inability to integrate and respond well to a traumatic event on a cognitive, emotional and/or on a sensory level resulting in 1) reexperiencing behaviors including flashbacks, nightmares and unbidden and upsetting thoughts of the event; 2) avoidance of reminders of the event including feelings of social alienation; 3) bodily arousal in response to reminders of the event including shakiness, sweaty palms, loss of concentration, irritability, trouble falling and/or staying asleep; 4) loss of function in life domains including work, intimate and family relationships etc. and 5) all of these symptoms lasting for more than thirty days following the event. Shorter, (less than thirty days following the trauma) but equally disabling and intense, dissociative responses to psychological traumas which at times might transition into PTSD, are labeled as acute stress disorders.

Suicide or human bomber

A suicide bomber as anyone who goes so far as to strap on a bomb, drive a vehicle filled with explosives to a target, or who otherwise attempts to detonate an explosive device on an airplane, in a subway, train, car, or elsewhere with the aim of dying to kill – irrespective of whether or not the bomber actually died in the attack or was successful in detonating – as that is often not within the bomber's control. We take the fact of being to the point of willingly strapping on a bomb or other type of improvised explosive device or driving a vehicle loaded with explosives to a target as sufficient evidence of seriousness of the intent to suicide and see the end result, which is often out of the hands of the bomber as less meaningful than the intent implied by these actions. An actor who arrives at a scene with a weapon, most often a gun, in order to shoot as many victims as possible prior to being killed himself may also be seen as a suicide terrorist, but is obviously not a human bomber. Psychologically the preparation and action of going on such a killing spree in which one's death is nearly guaranteed is very similar to that of a human bomber.

Terrorism

Has many definitions but is generally viewed as political violence carried out by politically organized non-state actors/groups with the intent to destroy property, injure, take hostage and/or kill

civilians and thereby intimidate and influence a wider witnessing audience and to effect political, religious or ideological change in both the targeted population and its government's responses; as well as to drum up support among the group's own sympathizers. Terrorism relies heavily upon the mass media to amplify the psychological effect of attacks made on relatively small numbers of victims by relaying horrifying images to a much larger witnessing audience. There is considerable controversy over who is defined as terrorists, particularly in situations where non-state actors (i.e. freedom fighters, rebels or guerilla fighters) could be also seen as leading legitimate rebellions in occupied territories but who nevertheless use similar tactics as terrorists such as using civilians as shields and indiscriminately targeting civilian as well as military targets.

Psychological trauma

Involves an experience in which an individual experiences a serious threat (real or perceived) to his own life or bodily integrity or witnesses death or threat of life or dismemberment in another and responds with overwhelming feelings of fear, horror and helplessness. In this experience the person is typically psychologically overwhelmed and unable to integrate his emotional, cognitive and sensory experiences in a normal manner and a peritraumatic dissociation may result immediately or in some time following the event. Acute or posttraumatic stress disorder may result.

Suicide terrorism is fast becoming one of the most pestilent global afflictions of the twenty-first century. As a terror tactic, it is one of the most lethal. Strategically it functions as a relatively cheap and effective means of upsetting the political, economic, and military situation of a region and has become one of the major threats to peacekeeping and peace-making efforts. Yet, as prevalence rates and death tolls from suicidal attacks increase, policy makers are still working in the dark trying to find the most effective policy responses to the emergence of this new and poorly understood security threat. Currently there is an extremely small empirical research database on which policy makers may base their understanding of the genesis of suicidal terrorism in order to work toward its prevention and eradication. Comprehending this growing threat and learning to combat it effectively on both the local and international levels is extremely relevant to current public policies aimed at promoting

peace and stability. This article briefly discusses the background of modern-day suicide terrorism, its migration around the world, and the tendency of modern-day terror groups to embrace it as a tactic. Drawing from field research in four distinct world regions, I give an in-depth analysis of the social and psychological factors in the genesis of human bombing.

Focusing on social and psychological factors within the sponsoring groups, the individual bombers, and the societies in which they exist, I have identified two main differing motivational sets on the level of the individual actors for the genesis of suicide terrorism. The first of these is trauma-based, occurs within zones of active conflict, and is often nationalistic, is viewed in terms of self and community defense, is expressive regarding attempts to mete out justice to the perceived enemy occupier, and includes acts of revenge by actors who are often so traumatized that they have become deeply dissociative and even refer to themselves as “already dead.” The second motivational set applies to actors outside the active zones of conflict, who are nevertheless influenced by them (through the Internet, video footage, pictures, and propaganda) and who frequently develop a deep sense of secondary traumatization. The actors in the second motivational set are generally vulnerable to terrorist ideologies due to a sense of alienation, marginalization, lack of life meaning, and lack of positive identity and are usually recruited through exposure via the Internet or via close-knit family and friendship networks. In both motivational sets, it is important to recognize that individuals are generally not motivated to take part in suicide terrorism without a guiding ideology and that they are also generally in need of an organization to equip and guide them to carry out their acts (although a rare few have acted on their own; Argo, 2004; Speckhard, April 2005). They generally are more willing to take part in suicide terrorism when there is broad-based support for suicide terrorism in the sector of the community in which they find their sense of belonging.

The History of Suicide Bombing

Dying to kill is not new.¹ What is new is that, in modern times since the mid 1980s, the tactic of

suicide terrorism has been repackaged and reborn. It has suddenly taken off like a wildfire spreading from Lebanon to Palestine, Sri Lanka, Chechnya, Morocco, Indonesia, Turkey, Afghanistan, and Iraq (where there is, on average, one suicide bomber per day); and now, sadly, it has spread even to those originating in Europe (i.e., the infamous shoe bomber, Richard Reid; the Mike's Place bombers in Israel; and Europeans going as bombers to Iraq and Afghanistan) and European-bred bombers now even targeting Europe itself. What combination of factors and conditions have made this tactic suddenly so popular that citizens of countries that have no history of suicide terrorism are suddenly willing to recruit themselves as human bombs, making this a main tactic of choice for many terror groups worldwide?

The Success of Suicide Terrorism

Looking back in recent history, we see that the current spate of modern-day suicide terrorism began in Lebanon in the 1980s when terrorists used suicide truck bombs to attack first the U.S. Embassy and later the barracks of the U.S. Marines and French troops, actions that led to the troops deciding to remove themselves from Beirut and position themselves offshore – this was viewed as a huge strategic victory by the terrorists.

Indeed the current epidemic of suicide terrorism is directly tied to this perception of its success by terror-sponsoring groups. As its use migrated around the world even many counterterrorism experts have credited suicide terrorism with derailing the Oslo peace accords, disrupting the peacekeeping and rebuilding efforts in Afghanistan and Iraq, impacting the election results in Madrid,² and drawing world attention and concern to political issues that may otherwise have been overlooked. Whether it is effective in achieving any real political gains outside the community in which it originates, that is, whether it creates

personal communication 2005). Looking back into the ancient past, many cite the scriptural account of Sampson as one of the first suicide terrorists; when blinded and chained to pillars, he decided to use his great strength to pull the columns down, collapsing the building on himself and those surrounding him. Similarly, some experts recall the Muslim Assassins and Jewish Sicari as the first suicide terrorists; both groups carried out assassination missions that were nearly always suicidal for the person carrying it out. Also Pakistani warriors in ancient times were known to run underneath the elephants carrying advancing troops to slit the bellies of the warring beast, bringing them down on themselves and killing both their enemy and themselves in the process.

²Some argue that the Madrid bombings did not involve suicide terrorists because the attackers left their bombs and detonated them from afar; yet threatened with imminent arrest, the bombers did explode themselves.

¹I am indebted to Mia Bloom for coining this phrase, which she also uses as the title for her book *Dying to Kill: The Allure of Suicide Terror* (Columbia University Press, 2005). Suicide bombing as a strategic tactic is not a new phenomenon. In recent history, it appeared in the 1950s in Vietnam in the form of bicycle bombers who exploded themselves in cafes, killing the enemy occupiers who frequented them (Francois Gere, French counterterrorism expert,

any real power base for those who employ suicide terrorism, is still, however, debatable.

Terrorist organizations thrive because of failed political solutions between conflicting parties, and their continued existence relies on their ability to change perceptions and impact the political process in favor of their constituent groups. When a much weaker group is pitted against a larger, more powerful, better financed and militarily equipped group, suicide terrorism becomes the tactic of choice because it is seen by the weaker group as strategically both efficient and more likely to lead to desired results than any other method of violence.

Suicide terrorism is a tactic that is:

- Inexpensive. It requires relatively simple, noncostly, and easily accessible equipment. The 9/11 attacks on the World Trade Towers and Pentagon cost approximately \$400 000, the Bali bombings approximately \$20 000 (NATO Defence College, 2004).
- Highly effective. The human bomber is essentially a smart bomb, going directly to his or her target and able to make adjustments even up to the last moments before detonation to avoid detection and to maximize the amount of damage done in the attack.
- Highly lethal. Suicide terrorism operations are the most lethal of nearly all currently used terror operations, and although they make up only a small fraction of the total of terrorist operations, they kill and wound the greatest number of victims.
- Extremely horrifying. The carnage caused by human bombers almost always makes the news; hence, by targeting only a small group of civilians the terror-sponsoring group can count on the media to amplify its effects, causing horror and dread to spread throughout a much wider witnessing audience.
- Nearly impossible to prevent. Once a bomber is equipped and on the way to the target, he or she is virtually impossible to stop (unless he or she can be talked out of the mission, which is rare) because he or she can always explode him- or herself on detection (Ganor 2005).
- Difficult to trace. The bomber, if successful, is killed in the attack and, unlike in other terror acts, needs no resources or risk dedicated to an escape plan; and once killed he or she cannot be caught and interrogated later, revealing who sent him or her.
- In endless supply, if the terror group's constituent population supports the use of this tactic. The main cost of suicide terrorism is the human beings who agree to sacrifice themselves on behalf of the terrorist cause. If this pool of individuals is large, the

terror organization has a virtually endless supply of weaponry and can carry on a very strong war of attrition with a much more powerful enemy who may eventually make concessions to the terror-sponsoring group simply to end the campaign of bombings.

The Lethal Cocktail of Suicide Terrorism

Suicide terrorism originates from the confluence of four main aspects of suicide terrorism: the sponsoring group's motivations, the ideology that supports it, the individual motivations for enacting it, and the societal support for suicide terrorism. Without this confluence of factors, suicide terrorism ceases to exist (Argo, 2003; Hafez, 2004; Merari, 2003; Moghadam, 2003; Speckhard, 2005b; Speckhard and Ahkmedova, 2005). When all these factors are present, suicide terrorism becomes a tactic that can travel the world as a spark igniting and reigniting a massive fire that we can only hope to be able to extinguish before it consumes us in its flames.

Sponsoring-Groups Motivations

Organizations provide the means, methods, and group dynamic underlying suicide terrorism operations and often the ideology as well. However, terrorist ideologies should not be thought of as simple constructs of terror-sponsoring groups. They arise, in fact, from a complex mix of social and political circumstances, psychological and religious context, and interplay among the actions, sentiments, and rhetoric of terror groups, their constituency, and the perceived oppressor-enemy. When looking at the motivations on the organizational level, we must acknowledge that terror-sponsoring organizations are largely political in their motivations and resort to terrorism when other political solutions have failed and resort to suicide terrorism when a specific set of circumstances exist. Researchers of suicide terrorism know that it is nearly always used strategically by organizations and generally resorted to only when the enemy is much stronger and better equipped militarily. Likewise, it is often used in later stages of the conflict (Bloom, 2005); sometimes reflects an outbidding process for power among competing groups (Bloom, 2005); and finds a receptive base of support in areas where occupation occurs, particularly if the foreign occupier is of another religion than those occupied and is perceived as oppressive and unjust (Pape, 2005; Speckhard, 2005b; Speckhard and Ahkmedova, 2005; Speckhard and Ahkmedova, 2006b). When

these circumstances exist, suicide terrorism may be deemed an effective choice by terror-sponsoring organizations for forcing concessions from its stronger enemy and thereby achieving its political goals.

Ideologies of Suicide Terrorism – Religion and Cosmic Warfare

Currently the strongest ideology supporting suicide terrorism makes use of distorted versions of Islam. As anyone who has studied the history of warfare knows, religion is often invoked to send warriors out to battle because believing that one is dying for a higher cause is highly motivating. Nearly all major faiths have in the past been used in this way and still are. Just as nearly all nations going to war today still make use of religious rhetoric when looking for popular support and to motivate their warriors, so too do many of today's terror-sponsoring groups (consider the rhetoric of the Crusades and the just war debates carried on in the West prior to the invasions of Afghanistan and Iraq). Although the loosely affiliated Al Qaeda/global Salafi terrorist groups do not represent legitimate nation-states, they claim that they are acting in behalf of a group of beleaguered people and that they are in a war, and their ideology is aimed to motivate warriors for the battle. The difference is not so much in the use of religion to garner popular support for acts of war and to motivate warriors but in the distortions of mainline religion to justify the tactic these warriors are motivated to adopt and the targeting of innocent civilians that their ideology justifies. We see in these current jihadist ideologies promulgated over the Internet and through underground networks the promotion of suicide terrorism – claiming it as the most effective way for these groups to triumph – the manipulation and misuse of treasured religious principles valuing martyrdom on behalf of Islam to motivate foot soldiers to self-recruit themselves for suicide missions against innocent civilian targets.

The suicide terror groups' uses of religion to motivate individuals to sign up to die makes strategic sense. Any believer of any faith who is persuaded of the following will act in extraordinary ways:

- He or she is in a cosmic battle (Juergensmeyer, 2000) involving apocalyptic forces
- that will eradicate either his side or the other, hence necessitating a war of defense.
- He or she should dehumanize and even demonize his or her enemies by seeing them as the enemies of God.
- In joining the group and taking on its values and teachings, he or she has learned the mind of God

and is authorized to act in the battle by the will of God.

- The battle in which he or she is fighting is for sacred values (Atran, 2003).
- He or she may go to extraordinary means to eradicate and blot out those he or she sees as evil-doers – even innocent civilians – who he or she believes are standing in the way of and offending God's will.

It is important to state however that, even though groups that make use of distorted versions of Islam are currently highly effective in promoting this cosmic warfare ideology to endorse the tactic of suicide terrorism, this does not mean that Islam itself is the problem. If it were, we would not have seen the Liberation Tigers of Tamil Eelam (LTTE) and the Popular Front for the Liberation of Palestine (PFLP), which are mainly composed of Marxist atheists, making use of this same tactic. In those cases, they made no reference to Islam, or to religion at all, but relied on charismatic leadership, deep-seated anger over nationalistic concerns and injustices, and the hope of becoming heroes for their cause to motivate human-bombing recruits. Likewise, many cults have made similar use of non-Islamic religions to induce an apocalyptic vision of the world in their believers, who then become willing to endorse violence and even kill and die in order to bring it about; these include the Aum Shinryko, the Japanese Hindu-related cult responsible for the sarin-gas poisonings in the Tokyo metro; the People's Temple cult that followed Jim Jones, whose members made a suicide pact and killed themselves and their families, resulting in over 900 deaths; and the Army of God, the Christian-based abortion clinic bombers.

So, although Islam is currently being misused to promote suicide terrorism, it is not in itself the problem. The problem is the powerful links that terrorists groups are able to make among individual motivations to self-sacrifice, societal circumstances leading to despair and defiance, and a hijacked version of Islam that, playing on sacred scriptures, promotes human sacrifice on behalf of the group. Any mainline religion can be distorted and used to motivate for suicide terrorism by framing it as an ideology in support of self sacrifice in a cosmic war of defense for the true faith. Currently, the main ideology in use among the most active suicide terrorist groups (i.e., the loosely affiliated Al Qaeda/global Salafi and other nationalistic militant jihadist-linked groups) is a hijacked version of Islam that calls for would-be martyrs from around the world to sacrifice themselves on behalf of a worldwide or nationalist jihad.

Individual Motivations

But why would anyone answer the call? Why do most of the main terror groups in Palestine, Lebanon, Iraq, Afghanistan, Chechnya, and elsewhere state that they have an endless supply of recruits – such as the Chechen female bomber who spoke on a prepared video broadcast during the Dubrovka/Nord Ost theater takeover in Moscow, where 800 hostages were held for 3 days, stating, “Even if we are killed, thousands of our brothers and sisters will come after us ready to sacrifice themselves” (Jazeera, 2002). Certainly there must be something much more powerful behind this than simply an ideology to make an individual cross the line from normal daily life to strapping on a bomb or getting into an airplane, car, or truck intent on exploding him- or herself in the middle of a crowd of civilians. What are the factors that cause these changes to occur within a human soul?

Although the answer is multifaceted and impossible to answer in this short article, our research speaking to bombers, would-be bombers, senders, the family members and close acquaintances of bombers, and their hostages points to two main motivational sets (Speckhard, 2004, 2005b, 2005c, April 2005; Speckhard & Ahkmedova, 2005, 2006; Speckhard & Ahkmedova, 2004, 2006a, 2006b; Speckhard, Tarabrina, Krasnov, & Ahkmedova, 2004; Speckhard, Tarabrina, Krasnov, & Mufel, 2005a, 2005b): trauma-based motivations and alienation-based motivations.

Trauma-Based Motivations

The first motivational set is a trauma-based and occurs within zones of active conflict. It is often nationalistic, viewed in terms of self and community defense, and expressive regarding meting out justice to the perceived enemy-occupier, and it often includes acts of revenge. The individuals motivated within this set have witnessed firsthand and over the television their neighbors, family members, and loved ones being killed by what they view as an occupying force. Many have grown up witnessing countless acts of violence and, as a result, have not developed normally and often suffer from posttraumatic stress and dissociative disorders. Many have lost jobs, educational opportunities, been humiliated, and often struggle for basic daily needs and security. Although the majority of traumatized individuals in conflict zones will not become suicide bombers even if invited to do so, an extremely small group will become vulnerable to terrorist ideologies that promote this tactic. Feeling constantly agitated by traumatic flashbacks; unable to avoid daily reminders of

their traumatic losses; and feeling in constant danger, bereaved, angry, and impotent, these individuals ultimately become so dissociative (i.e., separated from normal thoughts, perceptions, and emotions) and emotionally numb that they often refer to themselves as “already dead.” Actually, for them dying is no longer a feared outcome; they have already psychologically and emotionally numbed themselves to human suffering, yet it keeps mercilessly and painfully intruding into their thoughts, so death may seem as a welcome release. Embracing it and exerting some control over when and how it occurs is sadly welcomed by these traumatized individuals.

In response to threat, a normal person reacts in a fight-or-flight response and will only move into dissociative defenses (i.e., shutting down the normal features of consciousness such as emotions, logic, and memory and sometimes physically freezing as well) when the threat becomes overwhelmingly horrific, terrifying, and life-threatening. When this distinction was explained to him, Al Aqsa Martyrs’ Brigade sender of suicide bombers Zecharia Zubeidi explained, the individuals he equips to become human bombs are, according to his observations, caught inflexibly over long periods of time in a dissociative mode. He stated that, whereas combatants are flexible in their responses to violent conflict, these victims of the conflict are not. “They are completely different than us [fighters]. They have only one decision. We have many options. The thought of running away is always available. We can go and shoot.” In contrast, he described the martyrs as locked into an inflexible dissociative mode caused by traumatic stress and the one decision that comes from it. He stated, “They get flashbacks all the time and for them death is a mercy. . . . For the martyr all the cells in his mind are dead except for one.” According to Zubeidi, suicide bombers are in too much psychic pain to find another way to cope and become totally fixated on carrying out what they view as acts of community defense, expressions of pain, and enacting justice in response to “all that they have seen.” Explaining their psychological inflexibility in reference to his own dissociative states, which occur from time to time (also in response to traumatic experience), he stated, “When I feel this way I stay there one or two hours, but that one [a bomber], after all that he has observed there is only that one thing [i.e., to end his life on behalf of the community].”

Suicide bombers are often as well or better educated and less poor than their peers (Atran, 2004; Merari, 2003), and in other circumstances they might have been leaders in their communities. They are acutely sensitive to their own suffering and

that of those around them and wish to make a difference, but, similar to normally depressed and suicidal people, they see only very limited avenues of action. They want to escape their psychic pain but to do so honorably and to use their lives – even if it means dying – to help their communities. They are uniquely vulnerable to an ideology that promises that they will be heroes for the cause and that they can make a difference in the sociopolitical situation faced by their communities. They believe that their deaths are only a doorway to a better place and that by dying in self-sacrifice they can change things both now and in the afterlife, reuniting with those who have gone before and later bringing with them relatives who they left behind. Vulnerable and in pain, they succumb to an ideology that seduces them into sacrificing themselves for what they believe is a greater cause. Just like us, they hope for a more just world in which human dignity and rights will be upheld; however, unlike us, they have been deluded into believing that their acts of killing innocent civilians might bring this into being and for this they sacrifice themselves.

Alienation, Marginalization, Loss of Identity, Secondary Traumatization, and Desire for Life Meaningfulness, Belonging, and Heroism

The second motivational set involves actors not living in conflict zones and is more complex. Whereas we might understand an individual who has seen his or her family member killed in front of his or her eyes and feels his or her country has been occupied seeking revenge, we must question what can possibly motivate Europeans, Turks, Moroccans, Uzbeks, and others join such groups ultimately agreeing to die in order to kill others? In their cases, the main motivational set appears to involve vulnerable actors who are exposed through kin and friendship groups (Sageman, 2004) or through the Internet and informal recruiting to other individuals within a terror network.³ In their cases, these individuals are often marginalized, frustrated, and without hope in their societies. In Europe, they are often first-, second-, or third-generation immigrants or converts to Islam who feel deep sympathy and even kinship (i.e., as Muslim brothers) for those in conflict zones. Within the migrant community, there is often a deep sense of alienation and no secure sense of identity and belonging in either their country of origin and, more importantly, their host culture. Facing discrimination, often well educated

but facing poor job prospects, lacking positive identity and a sense of life's meaning, and having little else to make of themselves, once they are exposed to terrorist ideologies they are attracted to the appeal to become heroes for a cause.

In these cases, the sponsoring organization nearly always makes use of five powerful motivators. The first two are the ideas of belonging and identity – belonging to an important cause and group – and of taking on a heroic identity. The third is the use of pictures and graphic video footage of conflict zones that are shown to the potential recruit and interpreted as atrocities against innocent victims – mainly featuring Chechen and Palestinian suffering, but now also featuring footage and references to the US coalition's "War on Terrorism" in both Iraq and Afghanistan. Just as relief organizations in our societies often uses pictures and video footage of human suffering to motivate us to give to worthy causes, these organizations do the same, playing on the emotional reactions of their audiences to find the individuals vulnerable enough to respond to their calls to action.

In the European case, where distorted Islamic ideologies are being used, often those individuals who become terrorist recruits are sensitive and care about the conflicts they learn about but are not able to read in Arabic and fall prey to teachers who tell them that they know and can interpret the Koran for them and teach them the proper Islamic response to address such suffering. This is not to say they are simple minded; quite the contrary, often it is the well-educated and sensitive individuals who read the news and care about the world and who would have been leaders if they felt they had a way of participating in their society and its political discourse. Frustrated by their lack of opportunities to help others in need and needing a meaningful role and identity, they find answers in terrorist groups.

Fourth and fifth powerful motivators are the ideas often taken from religion that one ought to sacrifice in behalf of the brotherhood of the believers. In this manner the ideas of fictive kin (Atran, 2003) and martyrdom are instilled. The individuals who respond to such calls to action are called on to depart from the frustrations of this life, reject the society that has marginalized or frustrated them, and join a group following a path that promises eternal rewards. Suddenly the individual who was previously frustrated, feeling worthless, and had little hope feels a sense of belonging, a firm identity and purpose in what will soon turn out to be a foreshortened life.

The words of a disillusioned radical (Speckhard, 2005d) living in Brussels, Belgium, explain how he was radicalized and then found his way out of it.

³In Brussels, we have found that there are Internet cafes where, if one logs on for a half hour or so, pop up ads appear inviting one to join the worldwide jihad.

This young man (age 24) was adopted from Rwanda by white Belgium parents. Growing up as the only black in his community, he was alienated and confused about his identity. In search for his African roots, he found Islam, converted at age 15, and started attending a radical mosque, where he fell under the influence of extremist militant wahhabist teachers. "If you say to yourself this thing is God's will, you have to do it. It's simple, if you can't read in Arabic and people tell you that you cannot understand and you have to do it. I tried very sincerely to do so. I followed everything, the prayer schedule, eating and way of drinking all in the Sunnah. But there are also ideas about jihad. At age nineteen I was ready to go to Lebanon and fight for my brother Palestinian. I didn't know politics but I had an idea we had some Muslim community, our brothers that we must defend."

Speaking about how he became ready to become a martyr, he explained,

When I went to Morocco with my wife to her mother's house I saw Al Manar – Lebanon TV. They have a way to mix religion and politics. I can understand it because there is a true crisis in Palestine. . . . What I saw on the television was two Israeli soldiers taking big stones and breaking the bones of a Palestinian man, breaking his arm bones, his shoulders, all the bones in his hands, all the bones in his feet, his ribs, smashing them with a big rock. I'm sure they killed him or left him to die. I couldn't understand all of the Arabic (he was beginning to learn Arabic at this time) but I didn't need language to understand – it was all there in the pictures. Imagine people see that in the morning, get breakfast and see that on their television. When you see that you feel there is a unity of Muslim people. I decided to go there. I was completely crazy. I had a wife and baby but I thought I would go anyway.

Thankfully, this young man found his way out of the radical groups by studying Arabic and religion intensively until he found his own answers, independent of the militant teachers he had fallen under. Looking back at how close he came to going to be a martyr, he explained, "At that time I was a believing radical. For a radical you can die and kill for God no problem. I can die or be killed at anytime. I knew people who went to Afghanistan as bombers, the people who killed Mousad. To get the connections to go, here in Brussels, is no problem." Reflecting on how he got out, he explained, "The problem is you see all the people, politics, everything – you see through the Koran and it's your perspective." Similarly, he recalled that at that time "I was completely lost" and reflected that it was difficult to question what was being taught if you did not learn Arabic. "The difference between me [having left the radical groups] and the others is my studies."

Adopting a Terrorist Ideology as Psychological First Aid

In both motivational sets, the ideology of the terror-sponsoring group is acting as psychological first aid for the victim of other grievances. This psychological first aid is, of course, short-lived – as is its victim. Yet we see a powerful transformation take place in the human bomber who, seeing no way to change his or her circumstances, moves from a stance of a powerless victim of societal forces to becoming an actor in a worldwide or nationalistic drama that he or she has been persuaded might bring about a more just and dignified existence for those left behind. In this way, a marriage occurs between the terror-promoting ideology and the individual's psychosocial vulnerabilities emanating from traumatic and bereavement experiences in conflict zones and from marginalizing and frustrating circumstances in nonconflict zones. When this marriage occurs, all that is left is for the individual to believe that there is some significant portion of society that supports his or her stepping out onto this path (it may be the group that recruits him), which enables him to take the final steps to martyrdom. This is when societal support for suicide terrorism plays an important role in putting a martyr on the path to becoming a human bomb.

Societal Support for Suicide Terrorism

When a society deplores violence as an answer to violence and terrorism as an answer to social problems, it is unlikely that ideologies supporting suicide terrorism will resonate strongly within more than extremely limited groups of vulnerable individuals. However, when a society or significant elements of it begins to embrace an ideology in support of suicide terrorism, then these groups of potential recruits will expand exponentially. We witnessed this when a cult of martyrdom sprang up among Palestinians during the second Intifada and made their pool of potential recruits seemingly endless.

How does this occur? Again, it is a confluence of factors. When a society is so offended and even traumatized by daily living circumstances as a result of war, conflict, human rights violations, marginalization, frustrations, and daily humiliations and many individuals feel powerless to change these assaults to human dignity, the society can begin to support terrorist ideologies, especially those that espouse familiar and valued religious ideals, believing that terror acts may bring about change. In this case, the society begins to indoctrinate its children in support of suicide terror. Posters, songs, and videos may become prevalent throughout the society in support of

martyrdom (i.e., suicide terrorism), and the society as a whole may begin to resonate with the aims and ideology of the terror-sponsoring group.

It is most important that we avoid this when considering the fight against suicide terrorism. There will likely always be fringe groups that promote dying to kill; there will also always be individuals who are vulnerable to recruitment by these groups. However, we unwittingly create the circumstances in which the pool of recruits expands exponentially when we fail to address the societal factors leading to individual vulnerability and the societal support that make these groups impossible to extinguish and allows them to self-replenish faster than we can stop them. In these cases, their pool of recruits becomes so large that the terror groups can go on forever. Likewise, we must begin to address and take apart the rhetoric of terror-sponsoring organizations – addressing their ideologies by engaging with them in a discourse that can perhaps lead to more, instead of fewer, people believing that political solutions do exist and terror acts are neither necessary nor useful in bringing about a just, moral, and dignified existence.

As for the use of a hijacked version of Islam in the majority of the most current cases of suicide bombing, we must acknowledge two things. First, Islam is a religion that has always valued the struggle for three fundamental values: justice, morality, and human dignity. It is only natural that, when a terror group is hoping to motivate recruits, it appeals through Islamic traditions for recruits to take action on behalf of these fundamental values (especially if it can argue that recruits are acting in defense of self and community). Indeed, this idea of defense of self and community has been the basis of nearly all fatwas in support of martyrdom (i.e., suicide missions).

Second, a majority of the world's Islamic populations live under corrupt and autocratic regimes and face numerous human rights violations, territorial occupations, and discrimination; hence, there are many political reasons that Islamic people might gravitate to a terrorist ideology that shares with them the political goals of fighting for freedom and human dignity. Religion is simply the vehicle for uniting them and giving them the courage to fight (in whatever mistaken or brutal ways they chose) for the political goals they share. When we mistakenly believe that Islam itself is the problem and begin to assault deeply valued religious traditions and beliefs, we only fuel the fires of the terrorism ideologies currently in vogue.

There is also a contagion effect that we must consider, which occurs even with normal suicide – those who are in the family and friendship network of a suicide bomber are often so deeply affected by the act

that they too begin to consider acting similarly. We found countless examples of radicalization proceeding through close friendship and relatives networks, as have other authors (Sageman, 2004). Thus, radicalization can increase geometrically once it gets going.

Suicide terrorism is a complex psychosocial and political issue and demands thoughtful, just, and carefully carried out responses to bring an end to its continued and increasing use to influence worldwide politics.

Acknowledgments

In recent years, I have had the opportunity with my colleagues to interview incarcerated suicide terrorists; potential recruits for human bombers; senders of suicide terrorists; members of terrorists groups; and, perhaps most important, family members, hostages, and close associates of accomplished bombers to ask about the bombers, the events, and the influences that motivated these individuals to choose to die in this way and thereby to construct a psychological autopsy of them. These interviews concerned Chechen, Palestinian, Lebanese, Uzbeki, and Moroccan suicide terrorism and radicalization of groups in Europe and took place in Chechnya, North Ossetia, Russia, Belarus, Uzbekistan, Lebanon, Morocco, Palestine, Israel, the Netherlands, France, United Kingdom, and Belgium. I especially thank Khapta Akhmedova, Mokhtar Benabdallaoui, Valery Krasnov, Natalia Mufel, Yoram Schweitzer, and Nadejda Tarabrina for helping me in conducting the interviews and my students Ken Reidy and Valentijn Vanrompay, who accompanied me on trips to the Palestinian territories. I am also indebted to Scott Atran, Nichole Argo, Mia Bloom, John Esposito, Rohan Gunaratna, Bruce Hoffman, Mark Juergensmeyer, Assaf Moghadam, Cerwyn Moore, Jerrold Post, Marc Sageman, Jessica Stern, Andrew Silke, Michael Taarnby, Reuven Paz, John Horgan, and many others who have contributed to this field of research and helped me refine my technique and theory and have added to my understanding of this complex phenomena through conversations, critiques of my work, and correspondence.

See also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Disasters and Mass Violence, Public, Effects of; Gulf War Syndrome, Psychological and Chemical Stressors; Nuclear Warfare, Threat of; Persian Gulf War, Stress Effects of; Posttraumatic Stress Disorder, Delayed; Vietnam Veterans, Postwar Experiences and Health

Outcomes; War-Related Posttraumatic Stress Disorder, Treatment of; Posttraumatic Stress Disorder – Clinical.

Further Reading

- Argo, N. (2003). *The banality of evil: Understanding and defusing today's human bombs*. Unpublished research proposal.
- Argo, N. (2004a). *Culture, society and martyrdom*. Unpublished manuscript.
- Argo, N. (2004b). *Personal communication*.
- Atran, S. (2003). Genesis of suicide terrorism. *Science* 299.
- Atran, S. (2004). Mishandling suicide terrorism. *The Washington Quarterly* 27(3), 67–90.
- Bloom, M. (2005). *Dying to kill: The allure of suicide terror*. Columbia University Press.
- Ganor, B. (2005). *The counter-terrorism puzzle: A guide for decision makers*. London: Transaction Publishers.
- Hafez, M. (2004). *Manufacturing human bombs: Strategy, culture and conflict in the making of Palestinian suicide bombers*.
- Hafez, M. (2006). *Manufacturing human bombs: The making of Palestinian suicide bombers*. United States Institute of Peace Press Books.
- Jazeera, A. (Oct. 24, 2002, October). Chechens threaten to kill hostages. Retrieved September 1, 2006, from <http://www.cbsnews.com/stories/2002/10/25/world/main526899.shtml>
- Juergensmeyer, M. (2000). *Terror in the mind of God: The global rise of religion violence*. Berkeley and Los Angeles: University of California Press.
- Merari, A. (2003). *Suicide terrorism*. Unpublished manuscript.
- Moghadam, A. (2003). Palestinian suicide terrorism in the second intifada: Motivations and organizational aspects. *Studies in Conflict and Terrorism* 26(2), 65–92.
- NATO Defence College. *Financial and economic aspects of the fight against terrorism, immediate report*.
- Pape, R. (2005). *Dying to win: The strategic logic of suicide terrorism*. New York: Random House.
- Sageman, M. (2004). *Understanding terror networks*. University of Pennsylvania Press.
- Speckhard, A. (2004). Soldiers for god: A study of the suicide terrorists in the moscow hostage-taking siege. In: McTernan, O. (ed.) *The roots of terrorism: Contemporary trends and traditional analysis*. Brussels: NATO Science Series.
- Speckhard, A. (2005a). Understanding suicide terrorism: Countering human bombs and their senders. In: Purcell, J. S. & Weintraub, J. D. (eds.) *Topics in terrorism: Toward a transatlantic consensus on the nature of the threat*. Washington, D.C.: Atlantic Council.
- Speckhard, A. (2005b). Unpublished Beslan interviews.
- Speckhard, A. (2005c). Unpublished research interviews with Belgian radicals.
- Speckhard, A. (April 2005). Unpublished Palestinian militant interviews.
- Speckhard, A. and Ahkmedova, K. (2005). Talking to terrorists. *Journal of Psychobiology*, Fall 33(2).
- Speckhard, A. and Ahkmedova, K. (2006). The making of a martyr: Chechen suicide terrorism. *Journal of Studies in Conflict and Terrorism* 29(5).
- Speckhard, A. and Akhmedova, K. (2004). Mechanisms of generating suicide terrorism: Trauma and bereavement as psychological vulnerabilities in human security – the Chechen case. In: Donnelly, J. (ed.). Brussels: NATO Science Series.
- Speckhard, A. and Akhmedova, K. (2006a). Black widows: The Chechen female suicide terrorist. In: Schweitzer, Y. (ed.) *Female suicide terrorists*. Tel Aviv: Jaffe Center Publication.
- Speckhard, A. and Akhmedova, K. (2006b). The new Chechen jihad: Militant wahhabism as a radical movement and a source of suicide terrorism in post-war Chechen society. *Democracy and Security* 2(1), 103–155.
- Speckhard, A., Tarabrina, N., Krasnov, V. and Akhmedova, K. (2004). Research note: Observations of suicidal terrorists in action. *Terrorism and Political Violence* 16(2), 305–327.
- Speckhard, A., Tarabrina, N., Krasnov, V. and Mufel, N. (2005a). Posttraumatic and acute stress responses in hostages held by suicidal terrorists in the takeover of a Moscow theater. *Traumatology* 11(1).
- Speckhard, A., Tarabrina, N., Krasnov, V. and Mufel, N. (2005b). Stockholm effects and psychological responses to captivity in hostages held by suicidal terrorists. In: Wessely, S. & Krasnov, V. (eds.) *Psychological Responses to the New Terrorism: A NATO-Russia Dialogue*. IOS Press.

Suicide, Biology of

M A Oquendo

Columbia University and New York State Psychiatric Institute, New York, NY, USA

L Giner

Fundacion Jimenez Diaz and Universidad Autonoma de Madrid. Madrid. Spain

J J Mann

Columbia University and New York State Psychiatric Institute, New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M A Oquendo and J J Mann, volume 3, pp 538–543, © 2000, Elsevier Inc.

Introduction

A Model for Understanding Suicidal Behavior

Stressors: Triggers for Suicidal Behavior

Risk Factors Influencing the Threshold for Suicidal Behavior

Summary

Glossary

<i>5-Hydroxyindoleacetic acid (5-HIAA)</i>	A metabolite of serotonin.
<i>Cerebrospinal fluid 5-hydroxyindoleacetic acid (CSF 5-HIAA)</i>	The concentration of 5-HIAA found in the cerebrospinal fluid. The concentration of 5-HIAA in the cerebrospinal fluid is considered a reflection of serotonergic function in the central nervous system.
<i>Major depression</i>	A temporary or chronic mental disorder characterized by the occurrence of five of the following symptoms during a 2-week period or longer: feelings of sadness, inability to enjoy things, sleep disturbance, appetite disturbance, psychomotor changes, fatigue, difficulty concentrating, and suicidal ideas.
<i>Model for understanding suicidal behavior</i>	An explanation of how suicide risk factors contribute to suicidal behavior by creating a predisposition to suicidal behavior (lowering the threshold for suicidal acts) or by acting as triggers or stressors precipitating suicidal acts.
<i>Psychosis</i>	A temporary or chronic mental condition characterized by an inability to distinguish between reality and nonreality.
<i>Serotonin</i>	A neurotransmitter with multiple effects throughout the brain and with a role in a variety of psychiatric disorders and

Stress

symptoms, including depression, anxiety, aggression, and suicidal behavior.

A force that disrupts the equilibrium or normal functioning of an individual's mental or physical state. Different types of stressors may precipitate suicidal behavior.

Suicide

Self-inflicted death committed by an individual with the intention to end his or her life.

This article describes how stressors interact with a low threshold for acting on suicidal ideas. The types of stress that may lead to suicidal behavior are discussed, and factors influencing the threshold for suicidal behavior are described. Stressors that contribute to suicidal behavior include negative events, social conflict, acute psychiatric conditions, acute intoxication, and contagion. Risk factors that lower the threshold for suicidal behavior include family history of suicidal behavior, low serotonergic functioning, comorbid psychiatric conditions, dysfunction of the stress-reactivity hormonal system, marital isolation, early parental loss, and a history of sexual or physical abuse. Factors that raise the threshold for suicidal behavior include religious affiliation, moral constraints, concerns about social disapproval, and feelings of responsibility toward family. An understanding of the risk factors for suicidal behavior and stressors is crucial in developing methods for the prevention of suicide attempts and completion. Because it is a combination of these psychiatric, biological, and environmental risk factors that ultimately leads to an individual's suicide, interventions that address both the role of stressors and raise the threshold for acting on suicidal impulses are critical to prevention.

Introduction

Suicide is a self-inflicted death by an individual with the intention of ending his or her life. As the eleventh leading cause of death in the United States, it accounts for approximately 30 000 deaths a year in the United States and almost 1 million deaths worldwide. In the United States, suicide is the third leading cause of death in people ages 15–24 years and the second leading cause of death in those ages 24–35 years. The World Health Organization has noted a 60% increase in suicide deaths in the past 45 years. Although the actual number of suicide attempts is more difficult to quantify, it is estimated to be 10–20 times that of completed suicides.

Psychiatric, biological, and environmental factors contribute to suicidal behavior and ultimately to suicide completion. Among environmental contributors, stress has long been recognized as having a role in the timing of suicide. However, stress may also be mediated by psychiatric or biological factors. In addition, stress alone, although important, is not a sufficient condition for suicidal behavior. Suicidal behavior also requires a vulnerability for acting on such impulses. Thus, risk factors for suicide can be understood using a model that integrates the role of stress and the threshold for acting on the resulting suicidal impulse.

A Model for Understanding Suicidal Behavior

We have described a model in which suicidal behavior is the outcome of the interaction between an individual's threshold for suicidal acts and stressors that serve as triggers for suicidal behavior. In this model, which is illustrated in [Figure 1](#), the threshold is trait-related and refers to the propensity for suicidal behavior. In contrast, stressors, which may be considered state-related, are triggers that determine the timing and contribute to the probability of suicidal acts. Consequently, risk factors that contribute to suicidal behavior may be categorized according to whether they affect the threshold or act as stressors.

Stressors: Triggers for Suicidal Behavior

Stressors may be environmental, behavioral, biological, or a combination of these, as listed in [Table 1](#).

Environmental Stressors

Environmental triggers can be categorized as negative events or as contagion. Negative events may include financial difficulties, interpersonal loss (death of a family member or a loved one), loss of property, conflict with parents, or other interpersonal conflict. Contagion refers to the effects of suicides in the community on an individual's decision to commit suicide. These triggers may contribute to the timing of the suicidal act.

Negative events, specifically the loss or lack of a significant relationship, have been suggested to play a key role in the deaths of (1) suicide victims without mental health care, (2) suicide victims with alcohol abuse, (3) suicide victims with personality disorders, and (4) suicide victims with mood disorders. In addition, suicide is associated with financial troubles, homelessness, and job problems. The number of adverse life events appears to play a role as well. However, stressors or negative events related to suicide may vary depending on the culture studied. Most studies published in English have been conducted in Western cultures. Similarities regarding the precipitants of suicide in Eastern and Western cultures exist, such as the number of adverse life events; the presence of severe, acute stress at the time of death; high chronic stress in the year before death; and poor quality of life in the month before death. Yet differences in terms of the type of events that may precipitate suicide have been reported. In China, for example, loss of face, described as an unexpected insult to one's reputation, is a stressor associated with suicide not generally observed in Western cultures. As such, it appears that in different cultures the sensitivity to particular environmental stressors differs.

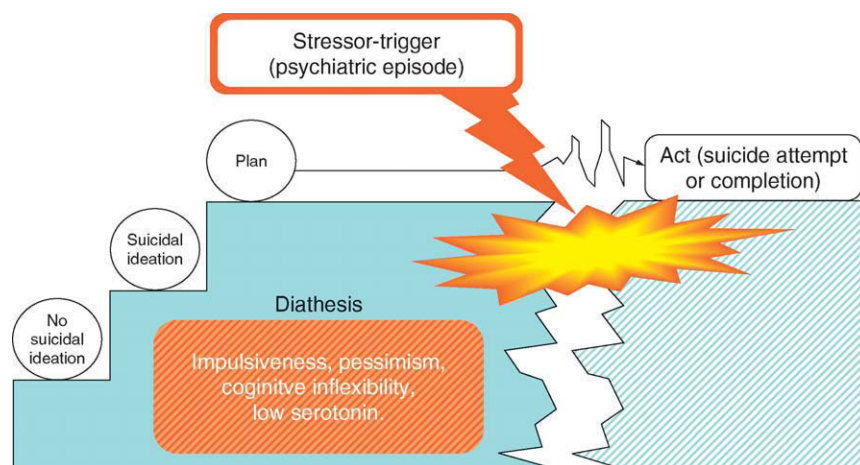


Figure 1 The diathesis–stress model for understanding suicidal behavior.

Table 1 A model for understanding suicidal behavior

	<i>Stressors precipitating suicidal behavior</i>	<i>Factors affecting the threshold for acting on suicidal ideation</i>
Environmental	Negative events (social, financial, or family crisis) Contagion	Marital isolation Lack of religious affiliation Low self-esteem Fewer perceived responsibilities toward family Fewer perceived reasons for living Alcohol or substance abuse
Behavioral	Judicial environment Acute substance intoxication	
Biological	Major psychiatric episode Acute substance intoxication	Genetic and familial factors Low serotonergic activity (aggression) Stress response (childhood history of physical and sexual abuse) Comorbidity (alcohol or substance abuse, cluster B personality disorder, posttraumatic stress disorder)

It is possible that negative events lead to suicide at least in part by disrupting the person's social support network; this network often plays a major role in suicide prevention by providing supervision and emotional support. For example, loss of a loved one is a negative event that leads to an immediate decrease in social support, and this can trigger suicidal behavior. Similarly, social support may be acutely lowered when there is conflict with family and can thereby trigger suicidal acts. Consistent with this idea, the social networks that attempters have seem to be weaker than those of nonattempters, as measured by marital status, living alone, interactions with family members, frequency of recent moves, and the number of close friends.

Another important environmental trigger is contagion, also known as the Werther effect, which was named by Philips in 1974 after Goethe's *The Sorrows of Young Werther*. The Werther effect has been documented as a meaningful contributor to suicidal behavior in both adolescent populations and older populations. Recent outbreaks of suicide in adolescent and elderly populations have raised concerns about media coverage of suicides and the potential for precipitating more suicidal acts within communities. Presumably, individuals with a lower threshold for acting on suicidal thoughts are more likely to act on them after learning of a suicide or an assisted suicide on television or in the newspaper. Moreover, the judicial environment may affect the risk for suicidal behavior in contagion situations by changing public attitudes about the acceptability of suicide. Although there is limited data available, it is possible that the extensive media coverage of physician-assisted suicide and its legalization in one state of the United States, as well as in some European countries, has increased the rate of suicide across diagnoses, including major depression, and specifically in those who are chronically ill and depressed. In

fact, physician-assisted suicides and euthanasia contribute to significant underreporting of suicides (as much as 20%) in countries such as the Netherlands, where these procedures are often not reported.

Thus, environmental factors or stressors that play a role in suicidal behavior include the disruption of social support networks, negative events, and interpersonal conflict as well as the altering of perceptions about the acceptability of suicide. These stressors are not causal, however, because the majority of individuals exposed to them do not commit suicide. Nonetheless, they may contribute to the timing and occurrence of suicidal acts in vulnerable people.

Biological and Psychiatric Stressors

Acute psychiatric illness Although depression or psychosis are not often conceptualized as stressors, it is useful to consider them as such because suicide rarely occurs in the absence of a major psychiatric disorder. The rate of mental illness in suicide victims varies from 90% in Western cultures to 60% in Eastern ones. However, even those subjects who do not meet criteria for a major psychiatric disorder seem to have some form of psychopathology. The most common mental disorders in suicide victims are depression and substance abuse, particularly alcoholism.

However, psychiatric illness alone is not sufficient to trigger suicidal behavior. Between 2 and 15% of subjects who suffer from unipolar depression ultimately commit suicide. Schizophrenia also carries a high risk of 10–15%. In bipolar disorder, the risk for suicide is generally considered to be approximately 19%; however, whereas some propose that it is nearly 50%, others question whether even 19% is an overestimate. Nonetheless, despite the high rate of suicide associated with these psychiatric conditions, most patients with these conditions do not commit suicide or even attempt suicide. Thus, the presence of a single stressor such as a psychiatric condition appears to

be a necessary, but not sufficient, cause for suicidal behavior.

Acute intoxication Substance and alcohol abuse greatly increase the risk of suicide, an effect that is especially pronounced in youth suicide. Acute substance intoxication can be considered a trigger associated with suicidal behavior. In fact, a recent review of acute alcohol use and suicidal behavior noted that 37% of completed suicides and 40% of attempted suicides involved acute alcohol use. Moreover, acute alcohol use was more prevalent among those who used violent methods of suicide such as self-immolation or gunshot.

To the extent that some individuals respond to environmental stress by using substances, acute intoxication may mediate suicidal responses to stress. The consequences of acute intoxication appear to be biological and behavioral in nature. The biological consequences of acute intoxication with alcohol include, for example, an acute increase of serotonin release, although the long-term consequence is lower serotonin function. In addition, the disinhibiting behavioral effects of intoxication may also play an important role in its association with suicidal behavior. Individuals may be more likely to act on a suicidal impulse when acutely intoxicated than when sober. Thus, acute alcohol intoxication has a complex relationship to suicidal acts.

Risk Factors Influencing the Threshold for Suicidal Behavior

Evidence for the existence of a threshold for suicidal behavior comes from two general findings. First, suicidal behavior manifests early in the course of illness, suggesting that individuals with a vulnerability for this behavior (lower threshold) are likely to express it early in the face of a stressor such as a depressive episode. Second, suicidal behavior runs in families independently of familial risk for mood disorders. Together, these factors suggest that the vulnerability to suicidal acts is not merely related to the presence of a mood disorder or to demoralization secondary to prolonged illness. Rather, suicidal behavior relates to a threshold, predisposition, or diathesis toward acting on suicidal urges. This threshold for suicidal behavior is influenced by risk factors that may be categorized as familial/genetic, biological, or environmental, as listed in [Table 1](#).

Familial/Genetic Influences on the Threshold for Suicidal Behavior

Evidence for the presence of genetic risk factors include the presence of a family history of suicide.

Although this risk factor could be thought of as facilitating suicidal behavior due to identification or contagion, it includes a genetic component as well. Adoption studies have found that, independent of the transmission of a major psychiatric disorder, adoptees who commit suicide have a sixfold higher rate of suicide among their biological relatives than do matched adoptees who do not complete suicide (None of the adopting relatives committed suicide.) Moreover, there appears to be greater concordance for suicide in monozygotic than in dizygotic twins. Other evidence for the genetic component of suicide is that the familial transmission of suicide risk to offspring appears to be independent of the transmission of psychiatric conditions. A study of Old Order Amish families has shown that family loading for mood disorders alone is not predictive of completed suicide. Rather, individuals with both a mood disorder and a family history of suicide have an increased risk for suicide.

Most genetics studies of suicidal behavior have focused on seven genes: the serotonin transporter (SERT), tryptophan hydroxylase (TPH) 1 and 2, three serotonin receptors (5-HTR1A, 5-HTR2A, and 5-HTR1B), and the monoamine oxidase promoter (MAOA), all of which are key proteins in the maintenance of serotonin function. Although there are no clear findings, the following are most promising:

1. The SERT promoter gene has two polymorphisms or variants, a short and a long form. The short form has been associated with suicidal behavior in different subpopulations, such as violent attempters and female suicide attempters. In a recent meta-analysis, suicidal behavior was reported to be more common in those with the short variant of the SERT promoter gene (OR = 1.17, CI 1.04–1.32; $p = 0.009$). Moreover, the short form has been reported to have an interaction with current life events, making those with this variant vulnerable to suicidal behavior when exposed to stress.
2. The 5-HTR1A gene encodes the 5-HT_{1A} receptor, which in depressed suicide victims has been found to have altered binding in the ventral prefrontal cortex, the region of the brain involved in decision making and restraint, among other functions.
3. The 5-HTR1B gene, which encodes the 5HT_{1B} receptor, appears to be related to impulsivity and aggressive behavior. This gene has two identified variants, both associated with 20% fewer receptors in the postmortem brain, although these variants were found at the same frequency in suicide and control subjects.
4. A low expressing version of the MAO gene has been reported to interact with exposure to

childhood abuse, resulting in aggression in adult males and indicating that the effect of this gene is only manifested if there are negative environmental conditions or, in other words, is dependent on a gene-environment interaction.

Thus, genetic and familial contributions to the diathesis for suicidal behavior require further inquiry but hold promise in the elucidation of suicide risk in individuals who carry the relevant variants of these genes.

Neurobiological Influences on the Threshold for Suicidal Behavior

The relationship of abnormal monoamine function to suicide risk has been studied for over 30 years. CSF 5HIAA is a metabolite of serotonin and its concentration in the cerebrospinal fluid (CSF) is considered a reflection of serotonergic function in the central nervous system (CNS). Patients with major depression who have made suicide attempts have lower levels of 5-HIAA in their CSF compared to depressed patients who do not have a history of suicide attempts. The reduction in CSF 5-HIAA appears to be more pronounced in depressed patients making planned, highly lethal attempts that more closely resemble failed suicide. Furthermore, CSF 5-HIAA correlates with characteristics of the most lethal attempt in the patient's lifetime, even when the suicide attempt is not temporally close to CSF measurements. In addition, CSF 5-HIAA seems to predict the risk of future suicide. These findings, together with those from animal studies in which the heritability and stability of CSF 5-HIAA on test-retest has been demonstrated, suggest that serotonergic dysfunction is trait-related and thus can reasonably be considered a factor that lowers the threshold for acting on suicidal thoughts.

Other measures of serotonergic function also indicate decreased serotonergic function in the CNS of suicide attempters. Fenfluramine has been widely used as a serotonin-releasing agent. Its effects on serotonin include reuptake inhibition, release from presynaptic storage granules, and possibly direct stimulation of postsynaptic receptors. We have reported that a fenfluramine challenge test elicited a lower serotonin-mediated prolactin response in patients with a major depressive episode who had a lifetime history of high-lethality suicide attempts compared with depressed patients who had a lifetime history of only low-lethality attempts. Moreover, using positron emission tomography and radioactively tagged glucose, we have reported that high-lethality suicide attempters exhibit low metabolism in the prefrontal cortex, the brain area associated with decision making and restraint. They also show impaired serotonergic responsivity. Both of these

abnormalities are proportional to the lethality of their suicide attempt and are related to suicide intent and impulsivity and thus may influence lethality via these parameters. In addition, the postmortem examination of brains from suicide victims found reduced binding to the serotonin transporter sites, which was especially pronounced in the ventral and dorsal prefrontal cortex. The increased binding or number of postsynaptic 5-HT_{1A} and 5-HT_{2A} receptors in prefrontal cortex, postulated to reflect an accommodation to low serotonin, has been reported in most studies as well. Studies of brain stem, from which CNS serotonin neurons emanate, have found reduced concentrations of serotonin or 5-HIAA in suicide victims. Together, these findings suggest that serotonin dysfunction in the CNS is associated with severe and lethal suicidal behavior.

Lower serotonergic functioning may have other effects on suicidal behavior as well and may mediate the association between suicide and behaviors such as aggression and alcoholism. Among depressed individuals, those who have made suicide attempts tend to be more aggressive. Several studies show that in humans and in nonhuman primates, aggression is linked to low serotonin function. Similarly, the biological underpinnings of the predisposition to abusing alcohol, including a low serotonin function and lower CSF 5-HIAA in alcoholics during abstinence compared to immediately after the discontinuation of drinking, have been shown. Thus, individuals predisposed to alcoholism or aggression may have biological parameters, such as low central serotonergic functioning, that are also associated with suicidal behavior.

In summary, serotonergic dysfunction may influence the threshold for both attempted and completed suicide. Moreover, serotonergic function may be a mechanism whereby genetic factors influence the suicide threshold. Clearly, the relationship among genetic endowment, suicide, and serotonin functioning needs further study.

Stress Response and Suicidal Behavior

Childhood physical and sexual abuse has been associated with suicidal behavior. Those who experience corporal punishment in adolescence are reported to be at increased risk of depressive symptoms, suicidal ideation, and alcohol abuse. Suicidal ideation increases markedly with the frequency of adolescent corporal punishment for both males and females. In fact, adolescent suicide attempters are three to six times more likely than controls to have had contact with the social services department. They also commonly had notations in their chart about child abuse. In a college sample, 16% of men and 24% of women

reported sexual abuse as children. Such a history predicted depression, chronic self-destructiveness, suicidal ideation and attempts, self-mutilation, and substance abuse. The younger the person was at time of first abuse, the higher the number of suicide attempts reported. Similarly, in a sample of depressed individuals, suicide attempters were more likely to have experienced sexual abuse themselves as children, and their offspring were also at increased risk for being sexually abused and for being suicide attempters, than were depressed individuals who were not suicide attempters. These findings support the link between childhood abuse and subsequent suicidal behavior.

Childhood physical and sexual abuse has been linked to abnormalities in the stress response as mediated by the hypothalamic-pituitary-adrenal (HPA) axis. The HPA axis, with its intricate two-way interaction with the serotonergic system, has been proposed to play a key role in the correlation between early stress exposure and later impulsivity, aggression, and suicidal behavior. Indeed, suicidal behavior seems to be related to hyperactivity of the HPA axis. Among depressed individuals, those with a hyperactive HPA axis, as demonstrated by an inability to suppress cortisol in the face of external steroid administration (dexamethasone suppression test), have been reported to have a ninefold elevation in risk for suicide. Moreover, suicide victims have larger adrenal glands, which produce stress hormones and low prefrontal cortex corticotropin releasing hormone binding, consistent with a hyperactive HPA axis.

Influence of the Presence of Multiple Psychiatric Diagnoses on the Threshold

Comorbid conditions that increase the risk for attempted suicide include alcohol abuse, substance abuse, posttraumatic stress disorder, and cluster B personality disorders. In depressed inpatients, these conditions are more common among suicide attempters.

Studies have suggested that when alcoholism and depression are comorbid, suicide attempt risk is greatly increased, even when the presence of an antisocial personality disorder or other personality disorders are controlled for. Furthermore, depressed alcoholics usually report more suicidal ideation than nonalcoholic depressed patients, even when the severity of other depressive symptoms is similar. Apart from the biological pathways previously noted, other mechanisms may mediate the relationship between alcohol and suicide. For example, alcoholism may lead to unemployment, financial problems, and interpersonal problems. Psychologically, it may increase loneliness and aggression and it may inhibit coping

mechanisms, all of which may lower the threshold for suicidal behavior.

Suicide attempts are also reported to be more common in depressed patients with comorbid borderline personality disorder (BPD) than in depressed patients without BPD. Moreover, subjects with major depression and comorbid borderline personality disorder are more likely to make multiple suicide attempts, and the attempts are no less medically damaging than those made by patients with major depression alone. Thus, the potential lethality of suicide attempts in patients with comorbid disorders and/or BPD should not be underestimated.

Posttraumatic stress disorder (PTSD) is frequently comorbid with depression, and when they co-occur the risk for suicidal behavior is enhanced. The relationship between PTSD and suicidal behavior appears to be mediated by the presence of cluster B personality disorder (CBPD), with both PTSD and CBPD arising as a result of earlier traumatic experiences. The assessment and treatment of comorbid conditions such as PTSD and CBPD in the context of depression may contribute to the reduction of suicide risk in this vulnerable population.

Environmental and Psychological Influences on the Threshold for Suicidal Behavior

There are environmental and psychological effects on the threshold for suicidal behavior. Environmental conditions that lower social support, such as marital isolation or living alone, are associated with suicide risk. Indeed, married women have a lower rate of completed suicide than women who are not married. In the same vein, an association between committing suicide in the context of a major affective disorder and not living with a child under the age of 18 has been reported.

However, there are clinical variables that increase the threshold for suicide attempts. These include religious or moral constraints, concerns about social disapproval, and coping and survival skills. In particular, religious affiliation seems to be associated with less suicidal behavior in depressed inpatients, and this protective effect may be mediated through moral objections to suicide and feelings of responsibility toward family. In fact, for a woman, the likelihood of completed suicide decreases as the number of children she has increases.

A variable not easily classified as influencing either threshold or trigger is low self-esteem. Although low self-esteem can be an enduring trait, it is also a symptom of depression that can be exacerbated during depressive episodes. In patients with recurrent major depressive disorder without comorbid personality disorders, those who reported feeling like a failure

were significantly more likely to have made a suicide attempt than those not reporting such feelings. In agreement with this finding, we have found that a factor of the Beck Depression Inventory that captures self-blame is associated with a history of suicidal behavior. Thus, low self-esteem may decrease the threshold for acting on suicidal thoughts in general, and its effect may be magnified during acute episodes of depression.

In summary, environmental factors related to social support and psychological attitudes, such as self-esteem or moral constraints against suicidal behavior, are important contributors to the threshold for suicide.

Summary

Risk factors can facilitate suicide by either acting as stressors or lowering the threshold for suicide. These stressors may be environmental, biological, or psychiatric in nature. Environmental stressors include negative events and contagion. Psychiatric stressors often have biological underpinnings, and acute psychiatric illness is a prominent stressor that is present in 60–90% of suicides. Acute intoxication also increases the risk of suicidal behavior by causing changes in serotonergic function and by disinhibiting the individual, making it more likely that he or she will act on suicidal impulses. These stressors act as triggers and precipitate suicidal acts.

However, whether psychiatric, environmental, or biological, stressors do not act alone in influencing suicidal behavior. Other risk factors influence suicidal behavior as well by lowering an individual's threshold for acting on suicidal ideation. Risk factors that lower the threshold for suicide can also be distinguished by three major categories: biological, psychiatric, and environmental.

Many studies of familial transmission of suicidal behavior have suggested that there is a genetic component to suicidal behavior. The strongest findings are related to the serotonin transporter gene (SERT), although there are promising results for other serotonin-related genes. Neurochemical risk factors such as lower CSF 5-HIAA levels and hyperactivity of the HPA axis are also associated with suicidal behavior and appear to be biochemical traits. These biological factors may potentially become useful predictors of suicidal behavior.

Psychiatric comorbidity may also lower the threshold for suicidal behavior. Studies suggest that in depressed patients, the presence of alcohol abuse, substance abuse, PTSD, or CBPD can increase the propensity to act on suicidal thoughts. Environmental influences that lower the threshold for suicidal

behavior include marital isolation, early childhood parental loss, and a childhood history of physical or sexual abuse. An individual's threshold for suicide is also lowered by feelings of low self-esteem. Several factors appear to raise the threshold for suicidal behavior. These protective factors include religious affiliation, moral constraints, concerns about social disapproval, coping and survival skills, and family responsibility. These attitudes may act as an incentive not to commit suicide.

This article describes a comprehensive model for categorizing risk factors for suicidal behavior into those that act as stressors and those that lower the threshold for acting on suicidal ideas. This model may guide research into methods of prevention and treatment of suicidal behavior and, ultimately, save lives.

Acknowledgments

This work was supported by PHS Grant MH46745 and MH59710. We thank Ms. Ansley Roche and Sadia Chaudhury for their assistance in the preparation of this manuscript.

Further Reading

- Anguelova, M., Benkelfat, C. and Turecki, G. (2003). A systematic review of association studies investigating genes coding for serotonin receptors and the serotonin transporter. II: Suicidal behavior. *Molecular Psychiatry* 8, 646–653.
- Caspi, A., Sugden, K., Moffitt, T. E., et al. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 301, 386–389.
- Cheripitel, C. J., Borges, G. L. and Wilcox, H. C. (2004). Acute alcohol use and suicidal behavior: a review of the literature. *Alcoholism, Clinical and Experimental Research* 28, 18S–28S.
- Egeland, J. A. and Susser, J. N. (1985). Suicide and family loading for affective disorders. *Journal of the American Medical Association* 254, 915–918.
- Grunebaum, M. F., Keilp, J., Li, S., et al. (2005). Symptom components of standard depression scales and past suicidal behavior. *Journal of Affective Disorders* 87(1): 73–82.
- Linehan, M. M. and Shearin, E. N. (1998). Lethal stress: a social-behavioral model of suicidal behavior. In: Fisher, S. & Reason, J. (eds.) *Handbook of life stress, cognition and health*, pp. 265–285. New York: John Wiley & Sons.
- Malone, K. M., Haas, G. L., Sweeney, J. A., et al. (1995). Major depression and the risk of attempted suicide. *Journal of Affective Disorders* 34, 173–185.
- Mann, J. J. (2003). Neurobiology of suicidal behaviour. *Nature Reviews: Neuroscience* 4, 819–828.
- Mann, J. J., Malone, K. M., Psych, M. R., et al. (1996). Attempted suicide characteristics and cerebrospinal fluid

- amine metabolites in depressed inpatients. *Neuropsychopharmacology* 15, 576–586.
- Mann, J. J., Oquendo, M., Underwood, M. D., et al. (1999). The neurobiology of suicide risk: a review for the clinician. *Journal of Clinical Psychiatry* 60(supplement 2), 7–11.
- Mittendorfer-Rutz, E., Rasmussen, F. and Wasserman, D. (2004). Restricted fetal growth and adverse maternal psychosocial and socioeconomic conditions as risk factors for suicidal behaviour of offspring: a cohort study. *Lancet* 364, 1135–1140.
- Oquendo, M. A., Friend, J. M., Halberstam, B., et al. (2003). Association of comorbid posttraumatic stress disorder and major depression with greater risk for suicidal behavior. *American Journal of Psychiatry* 160, 580–582.
- Oquendo, M. A., Malone, K. M. and Mann, J. J. (1997). Suicide: risk factors and prevention in refractory major depression. *Depression and Anxiety* 5, 202–211.
- Oquendo, M. A., Placidi, G. P., Malone, K. M., et al. (2003). Positron emission tomography of regional brain metabolic responses to a serotonergic challenge and lethality of suicide attempts in major depression. *Archives of General Psychiatry* 60, 14–22.
- Phillips, M. R., Yang, G., Zhang, Y., et al. (2002). Risk factors for suicide in China: a national case-control psychological autopsy study. *Lancet* 360, 1728–1736.

Relevant Websites

Centers for Disease Control. <http://www.cdc.gov>.
World Health Organization. <http://www.who.int>.

Suicide, Psychology of

D Lester

Richard Stockton College of New Jersey, Pomona, NJ, USA

R L Walker

University of South Carolina, Columbia, SC, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D Lester, volume 3, pp 544–548, © 2000, Elsevier Inc.

Diathesis

Stress as a Precipitant of Suicide

The Adequacy of Stressors

The Situation

Survivors

Assisted Suicide

Preventing Suicide

Glossary

Assisted suicide Suicidal acts in which the individual is assisted. This may take the form of a prescription being written for a lethal dose of a medication or of someone else administering the lethal dose/injection.

Attempted suicide A suicide act that the individual survives. At the lower end of medical lethality, attempted suicide may be difficult to distinguish from self-mutilation. In Europe, the trend has been to call suicidal acts that are not life-threatening parasuicide or deliberate self-harm.

Completed suicide

A suicidal act that results in the death of the individual. It is sometimes called fatal suicide.

The study of suicide is complicated by difficulties in determining intentionality, or whether the suicidal person has engaged in intentional self-harm. Recently, suicide nomenclature has evolved such that clinicians, researchers, and others can now use a common language to understand the suicide-related behavior. Beyond assisted suicide, attempted suicide, and completed suicide, suicidal ideation, defined as any self-reported thoughts of engaging in suicide-related behavior, is also commonly used. In addition, parasuicide is used rather than attempted suicide to describe self-destructive behavior in the absence of clear intent of suicide. The term suicidality should be avoided because it refers to a broad range of suicide-related behaviors.

The 2001 National Strategy for Suicide Prevention noted that 650 000 people receive emergency care for suicide attempts and 30 000 die by suicide each year in the United States. Given the loss of life and impact on families and communities, suicide has been declared a serious public health problem. The goals of the National Strategy are to (1) promote awareness of suicide as a public health problem; (2) develop broad-based support for suicide prevention; (3) develop and implement strategies to reduce the stigma associated with mental illness, substance abuse, and suicide; (4) develop and implement community-based suicide prevention programs; (5) promote efforts to reduce access to lethal methods of self-harm; (6) implement

effective training to assess at-risk individuals and provide effective intervention; (7) develop clinical and professional practices; (8) increase access to mental health services; (9) improve the reporting and media portrayal of suicide-related behavior; (10) promote research on suicide and suicide prevention; and (11) improve and expand surveillance systems. The overarching objective is a comprehensive and integrated approach to reducing the loss associated with suicide and related behaviors. Many other nations have also developed national strategies for reducing the incidence of suicide.

The 2002 Institute of Medicine (IOM) report for the United States is an excellent and comprehensive resource for understanding suicide's history; impact; manifestations across gender, ethnic, age, and other demographic groups; and protective factors or suicide buffers. For example, although African Americans were previously thought to be protected from suicide, studies show that in the 1980s and 1990s Black men ages 25–34 years demonstrated suicide rates comparable to that of White men in the same age group. The most recent data and trends are cited in the IOM report, including the cost to society due to medical expenses for emergency intervention and treatment, lost productivity, and reduced wages. The value of lost productivity alone was estimated to be \$11.8 billion in 1998.

Suicide can be studied at the societal or at the individual level. Sociological research has focused primarily on the societal rate of completed suicide, a rate that is relatively stable over time. At the individual level, in contrast, both completed and attempted suicides have been studied and both behaviors are difficult to predict. Suicide is rare, with the completed suicide rate, for example, seldom rising above 40 per 100 000 per year in any nation. The rarity of suicidal behavior means that prediction instruments (typically based on personal data and psychiatric and psychological measures) will have a high rate of false positives (people predicted to be suicides who, in fact, will not kill themselves). Differences between the predictability of suicide at societal and individual levels has an analogy in radioactive decay, in which the half-life of a mass of, say, uranium is easily determined and constant, whereas predicting which particular uranium atom will decay next is impossible.

Suicide fits well with the diathesis–stress model of psychiatric disorder, in which only those individuals with a predisposition (diathesis) to develop a disorder do so when faced with particular stressors. It should be noted, however, that both factors can influence one another – a diathesis can make it more likely that a person will encounter stressors and a stressor

can strengthen a diathesis. For example, a stress such as the loss of a loved one can be more stressful for a person who has few friends than for someone without that diathesis. Likewise, a tendency to have few friends may be increased by the experience of interpersonal losses.

Diathesis

Many predisposing factors for suicide have been identified by research.

1. Neurophysiological dysfunction, especially problems in serotonergic pathways, may underlie the psychiatric disorders for which the risk of suicide is high.
2. The majority of suicides are found to have been suffering from a psychiatric disorder. In particular, suicide is common in those suffering from affective disorders such as a major depressive disorder or a bipolar affective disorder (manic–depression), substance abuse (both drug and alcohol abuse), and schizophrenia.
3. Childhood experiences such as the loss of a parent (or other primary caregiver) through death, especially if the death is from suicide, or divorce and physical or sexual abuse are commonly found in the personal histories of suicides. Lester studied 30 suicides famous enough to have a biography written about them (such as Ernest Hemingway) and found that exactly half had experienced the loss of a parent or other loved one. Fourteen of the 15 suicides who had experienced such a loss had experienced this loss between the ages of 6 and 16. However, it is not clear whether such experiences increase the risk of suicide directly or whether they increase the risk of psychiatric disorders, which in turn increases the risk of suicide. For example, a history of physical and sexual abuse is associated with a wide spectrum of later psychiatric disorders, including depression, anxiety, eating disorders, and substance abuse, in addition to suicidal behavior.
4. The mood of potential suicides is characterized by depression. Suicidal individuals obtain high scores on self-report measures of depression, and the level of depression of those attempting suicide increases with the seriousness of their suicidal intent. Depression includes many components: physical symptoms (such as sleep and appetite problems), mood symptoms (such as depression and guilt), and cognitive symptoms (such as hopelessness). Suicidal individuals tend to be hopeless, believing that things will not improve and will probably worsen in the future. Aaron Beck has

found that self-report measures of hopelessness are stronger predictors of suicidal intent than self-report measures of depression. Edwin Shneidman has called the psychological pain experienced by potential suicides psychache.

5. Suicidal individuals appear to think differently from nonsuicidal individuals. Research has found that suicidal individuals have a reduced ability to generate alternatives other than death for the problems they are facing; they tend to think in dichotomous categories, polarizing their evaluations into extremes, such as good versus bad or right versus wrong; and they tend to think rigidly.
6. In addition, the risk of suicide is strongly associated with simple demographic characteristics. Suicide rates in the United States are higher in men than in women, in the elderly than in the young, in Whites than in African Americans, and in the divorced and widowed than in the single and married. These demographic correlates of suicide risk vary from country to country, however.

In 2005, Joiner developed a comprehensive interpersonal-psychological theory of attempted and completed suicide and hypothesized that suicidal people must (1) be capable of lethal self-injury, (2) perceive that they are a burden to loved ones, and (3) believe that they are not connected to a valued group or relationship. These three necessary precursors answer the question of why people do or do not die by suicide. For example, few people die by suicide, relatively speaking, and interpersonal-psychological theory suggests that the acquired ability to inflict self-harm requires courage, fearlessness, and habituation to self-harm. Joiner asserted that individuals must also see themselves as ineffective and burdensome to loved ones. In contrast, a sense of belonging to a valued group acts as a suicide buffer. In sum, the confluence of three factors must be in place. However, those who engage in repeated attempts at suicide are at high risk for suicide.

Joiner's theory explains several research findings. For example, the relatively low rate of suicide despite the high lifetime prevalence of lifetime suicide thoughts is explained by the absence of two or more factors (such as the incapability of self-injury and a sense of connectedness). Thus, not everyone who considers suicide or feels hopeless will attempt suicide. Note that those who die by suicide use more lethal means and therefore demonstrate a stronger propensity to self-harm. Although women are less likely to die by suicide than men, those women who die are as likely as men to use very lethal methods (e.g., a gun and hanging). The role of neurophysiological factors in suicide is related to the impact of

neurotransmitters on negative emotionality, which may, in turn, incite feelings of burdensomeness and contributes to a lesser feeling of belongingness.

Stress as a Precipitant of Suicide

Suicides are found to have experienced higher levels of stressful events in the past year than nonsuicidal individuals. Not only is the level of stressful events higher, but typically there is a further increase in this already high level in the weeks immediately prior to the suicide. These conclusions are based on retrospective studies of the lives of completed suicides in the hours, days, weeks, and months prior to their death – a procedure called a psychological autopsy. In a psychological autopsy, mental health professionals, usually guided by an interview schedule, seek out and interview friends, colleagues, and relatives of the suicide and inquire into the suicide's recent behaviors and psychological state.

Commentators have suggested that men become suicidal more as a reaction to performance failure, whereas women become suicidal more as a reaction to romantic conflict. However, research has indicated few differences between the stressors encountered by female and male suicides. Large differences in precipitating stressors by age have been observed. Stressors experienced by adolescent and young adult suicides tend more often to involve negative interpersonal events, such as friction and break-ups with loved ones and problems involving work, finances, and the criminal justice system. In contrast, stressors of elderly suicides tend more often to involve medical problems and intrapsychic states such as depression. These differences by age are found both in psychological autopsy studies and in content analyses of suicide notes.

Similar differences have been observed in attempted suicides and those thinking about the possibility of killing themselves, who also are found to have experienced more recent stressful events than nonsuicidal psychiatric patients or general medical patients. Two studies from the 1990s illustrate the role of stress. Paul Dean and colleagues found that suicidal ideation in college students was positively associated with self-ratings of negative life events in the past year and, in addition, depression, hopelessness, and perfectionism, and was negatively associated with scores on a reasons-for-living scale. John Reich and colleagues found that suicidal ideation in a sample of elderly adults (mean age 70 years) was associated with worsening health in the prior 6 months and, in addition, self-reported feelings of helplessness, fatalism, confused thinking, and low self-esteem.

The Adequacy of Stressors

Older adults who commit suicide are typically seen as having experienced more adequate stressors than younger adults and adolescents who kill themselves. Elderly suicides are frequently suffering from medical problems, often severe, and have experienced a succession of losses (of loved ones and friends, of employment and work, and of physical abilities). Thus, the precipitants of their suicides appear to involve severe stress. Most of the examples of rational suicide involve such cases, and indeed assisted-suicide laws passed in the United States and other nations are usually restricted to those who are terminally ill. In contrast, the precipitants of suicide in adolescents and young adults are sometimes seen as inadequate causes of suicide. Such suicides are, therefore, viewed as impulsive and not rational and are less predictable.

The Situation

The situation in which potential suicides find themselves may also play a role in precipitating their suicides. The easy availability of lethal methods for suicide may play a role. Research on the effects of restricting easy access to methods for suicide (such as detoxifying domestic gas, placing emission controls on cars, and restricting the number of pills prescribed for psychiatric disorders) has shown that this can be an important tactic for preventing suicide.

The publicity given to suicides, especially suicides by celebrities, has been shown to lead to an increase in the suicide rate in the following few days, especially among those of the same age and sex as the celebrity. Suicide among peers can have a contagion effect, precipitating suicidal behavior in those with those characteristics that predispose the individual to suicide. Suicides occurring in schools have aroused particular concern, and there are now guidelines for staff in responding to the suicide of a student.

Survivors

The experience of having a loved one complete suicide can be extremely traumatic for those left behind. In addition to the normal grieving process, there is additional stress from feelings of guilt and responsibility (could something have been done to help the suicide and prevent their death?), anger (over their suicide), and stigma from others who may view the survivors with suspicion or simply not know how to respond.

In addition, some suicides kill themselves in such a way that their loved ones discover the body. The discovery of the body of a loved one drowned,

hanging from a hook, with cut wrists, or, even worse, grossly disfigured by a bullet is extremely traumatic. It creates a final picture of the loved one that is difficult to exorcize.

Survivors have become very active in their communities and in professional organizations devoted to suicide prevention. They have formed self-help groups, sometimes with the involvement of clinicians, to meet and assist one another in dealing with the complicated grief that accompanies the suicide of a loved one. Books have been written for survivors and the membership of the American Association of Suicidology has a large contingent of survivors.

Assisted Suicide

The reluctance of people to kill themselves with violent methods and the desire of suicidal individuals to have others assist them in their suicides have led to widespread debate about assisted suicide. Derek Humphry published the book *Final Exit*, which details the medications and dosages for suicide – a do-it-yourself book that became a best seller in the United States. Similar books in other nations (such as France) have been banned.

In physician-assisted suicide, typically the goal is to have physicians provide a sufficient amount of a lethal medication for the client to complete suicide. In most regions that have legalized this process (e.g., in Oregon and the Netherlands), this process is restricted to those who are terminally ill, and there is a detailed procedure for those applying for physician-assisted suicide that must be followed.

However, on occasion, physicians and relatives have assisted the suicidal individual in the actual killing, for example, by administering the lethal injections or by shooting the person. This is illegal in almost all nations and jurisdictions, and those assisting in the actual killing can be, and indeed sometimes are, charged with murder. Nevertheless, such incidents have occurred and continue to occur. The line between assisting someone to commit suicide and murdering someone is, therefore, blurred in these cases, and society is vigorously debating these issues at the present time.

Preventing Suicide

The prevention of suicide uses several strategies. Many communities now have suicide prevention centers that maintain telephone counseling services staffed primarily by well-trained lay people who can provide crisis counseling for those who are suicidal. The American Association of Suicidology has a directory of such centers in the United States and also

inspects and certifies the centers if asked to do so. Befrienders International, Life Line, and the International Federation of Telephone Emergency Services are three international organizations that have established and coordinated crisis counseling and suicide prevention services in other nations.

Educational programs have been established to inform students in school about suicide, especially about recognizing signs of suicidal intent in their peers, and to provide them with resources to which they can refer their suicidal peers. Educational programs have also occasionally been established for general practitioners and family physicians so that they can diagnose depression in their patients more accurately and prescribe antidepressants more effectively.

Clinicians have been developing counseling techniques especially designed for suicidal clients to supplement the general systems of psychotherapy (such as transactional analysis and reality therapy) that guide therapists. These new techniques have been primarily in the field of cognitive therapy and are based on the results of research on the cognitive processes of suicidal individuals discussed earlier.

The 2002 report from the IOM examined prevention and medical and psychosocial intervention, but also looked at barriers to research, treatment, and effective prevention. For example, suicide is more likely to occur within 1 month after discharge from psychiatric institutions than at other times. The report noted that a significant gap in clinical trials is the absence of high-risk patients. Because effective treatments (psychopharmacological and psychosocial) involve long-term maintenance in which the individual is required to have contact with a health-care professional, adequate health-care insurance coverage is essential. School-based programs that target high-risk school-age children are essential for suicide prevention among youth and seem to result in increased social support, self-efficacy, and self-esteem among youth at risk for suicide, but research on these intervention programs is

complicated by poor control groups and limited long-term follow up.

See Also the Following Articles

Depression and Manic-Depressive Illness; Suicide, Biology of; Suicide, Sociology of; Survivor Guilt.

Further Reading

- Clark, R. V. and Lester, D. (1989). *Suicide: closing the exits*. New York: Springer-Verlag.
- Dean, P. J., Range, L. M. and Goggin, W. C. (1996). The escape theory of suicide in college students. *Suicide and Life-Threatening Behavior* 26, 181–186.
- Freeman, A. and Reinecke, M. A. (1993). *Cognitive therapy of suicidal behavior*. New York: Springer.
- Henry, A. F. and Short, J. F. (1954). *Suicide and homicide*. New York: Free Press.
- Humphry, D. (1991). *Final exit*. Eugene, OR: The Hemlock Society.
- Institute of Medicine of the National Academies. (2002). *Reducing suicide: a national imperative*. Goldsmith, S. K., Pellmar, T. & Bunney, W. E. (eds.). Washington, DC: The National Academies Press.
- Joiner, T. (2005). *Why people die by suicide*. Cambridge, MA: Harvard University Press.
- Leenaars, A. A. and Wenckstern, S. (eds.) (1991). *Suicide prevention in schools*. New York: Hemisphere.
- Lester, D. (1997). *Making sense of suicide*. Philadelphia: Charles Press.
- Lukas, C. and Seiden, H. M. (1987). *Silent grief*. New York: Scribners.
- O'Carroll, P. W., Berman, A. L., Maris, R. W., et al. (1996). Beyond the Tower of Babel: a nomenclature for suicidology. *Suicide and Life-Threatening Behavior* 26, 237–252.
- Reich, J. W., Newson, J. T. and Zaura, A. J. (1996). Health downturns and predictors of suicidal ideation. *Suicide and Life-Threatening Behavior* 26, 282–291.
- Shneidman, E. S. (1996). *The suicidal mind*. New York: Oxford University Press.
- U.S. Department of Health and Human Services (2001). *National strategy for suicide prevention: goals and objectives for action*. Rockville, MD: Public Health Service.

Suicide, Sociology of

D Lester

Richard Stockton College of New Jersey, Pomona,
NJ, USA

R M Fernquist

Central Missouri State University, Warrensburg,
MO, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition
article by D Lester, volume 3, pp 549–552,

© 2000, Elsevier Inc.

Durkheim's Theory

Henry and Short's Theory

Naroll's Theory

Culture Conflict

Measuring Regional Stress

Multivariate Studies of Regional Suicide Rates

Nonfatal Suicide Behavior

Glossary

*Attempted
suicide*

A suicide act that the individual survives. Attempted suicide is at the lower end of medical lethality and may be difficult to distinguish from self-mutilation. In Europe, the trend has been to call suicidal acts that are not life-threatening parasuicide or deliberate self-harm.

*Completed
suicide*

A suicidal act that results in the death of the individual. It is sometimes called fatal suicide.

Sociological research into suicide has focused primarily on the rate of completed suicide in regions (nations or regions within nations), a rate that is relatively stable over time. Ecological studies examine suicide rates of a set of regions at one point in time, usually to see which social characteristics of the regions predict the suicide rates. Time-series studies examine suicide rates of one region each year for a period of time, usually to see which social characteristics predict the yearly suicide rate. Time-series cross-sectional studies examine suicide rates of a number of different countries simultaneously for a period of time. Here multivariate ecological studies of suicide are reviewed. Three major theories of suicide are presented: Emile Durkheim's, Andrew Henry and James Short's, and Raoul Naroll's. The role of culture conflict in increasing the rate of suicide is illustrated. Finally, the possibility of sociological studies of nonfatal suicide is discussed.

Durkheim's Theory

The first major theory of the social suicide rate was proposed in 1897 by Emile Durkheim. Durkheim argued that two broad social characteristics are responsible for determining the social suicide rate: social integration and social regulation. Social integration is the extent to which the members of the society are bound together in social networks. Very high levels of social integration lead to high rates of altruistic suicide, whereas very low levels lead to high rates of egoistic suicide. Social regulation is the extent to which the desires and behavior of the members of the society are restricted by social norms and customs. Very high levels of social regulation lead to high rates of fatalistic suicide, whereas very low levels lead to high rates of anomic suicide.

There has been a great deal of research purporting to test this theory, for example, exploring the association between divorce rates and suicide. Divorce is presumed to weaken the extent of social integration and reflect weak social regulation, and therefore higher divorce rates should be positively associated with higher suicide rates. Both ecological studies and time-series studies have confirmed this predicted positive association between divorce rates and suicide rates.

However, research rarely examines the roles of both social integration and social regulation separately, partly because it is difficult to operationalize the two social characteristics independently of one another. The few studies that have attempted to do so have found that measures of social integration are stronger correlates of suicide rates than measures of social regulation.

Furthermore, no research has attempted to classify suicides in the societies under study into Durkheim's four types. Durkheim proposed, for example, that low levels of social integration result in high levels of egoistic suicide. Therefore, a proper test of Durkheim's theory requires the dependent variable to be the rate of egoistic suicide. The rate of egoistic suicide is not identical to the total social suicide rate, which is, instead, the combined rate of all four of Durkheim's types of suicides.

Henry and Short's Theory

Henry and Short proposed a theory based on both psychoanalytic theory and the frustration-aggression hypothesis. They assumed that the primary target of aggression for a frustrated individual is another person. What inhibits this other-oriented aggression and results in the aggression being directed inward onto the self?

Henry and Short proposed that the primary factor was the strength of external restraints on behavior. When behavior is required to conform rigidly to the demands and expectations of others, the role of others in the responsibility for the self's frustration and misery is strong. As a consequence, other-oriented aggression is legitimized. When external restraints are weak, the self must bear the responsibility for the frustration and misery. Thus, other-oriented aggression is not legitimized, and self-directed aggression becomes more likely.

These proposals lead to interesting predictions, many of which have been confirmed. The oppressed in a society have clear external sources to blame for their misery – their oppressors. Therefore, other-oriented aggression is legitimized for them, and they will tend to have higher rates of assault and, in the extreme, homicide. In contrast, the oppressors in the society have fewer external sources to blame for their misery because, as the dominant group, they have tremendous opportunities for advancement and gratification. Therefore, the oppressors in the society will tend to have higher rates of depression and, in the extreme, suicide. These predicted differences in the suicide and homicide rates are found for African Americans and Whites in the United States.

Henry and Short's theory can also explain the positive association between the quality of life and suicide rates. When the quality of life in a nation is high, there are few external sources to blame for one's misery, and so suicide will tend to be more common. This association has been confirmed in ecological studies both of nations and of U.S. states.

Naroll's Theory

Raoul Naroll, an anthropologist, proposed a theory of social suicide rates that has relevance to stress. Naroll proposed that suicide was more likely in those who were socially disoriented, that is, in those who lack or lose basic social ties (such as those who are single or divorced). Because not all socially disoriented people commit suicide, there must be a psychological factor that makes suicide a more likely choice when a person is socially disoriented, and Naroll suggested that it was the person's reaction to thwarting disorientation contexts. Such contexts involve a weakening of the person's social ties as a result of the actions of other people or him- or herself (but not as a result of impersonal, natural, or cultural events). Being divorced by a spouse or murdering one's spouse are examples of such contexts, but storm damage to property or losing a spouse to cancer are not. In thwarting disorientation contexts, some people commit protest suicide, which Naroll defined as

voluntary suicide committed in such a way as to come to public notice (and, therefore, excludes unconsciously motivated suicides and culturally motivated suicides such as *suttee*).

Durkheim's theory of suicide refers more to steady-state characteristics of the society as major explanatory variables, whereas Naroll's theory suggests the role of sudden, acute changes in a society – societal stressors. Furthermore, Naroll's theory is phrased in a way that makes it more usefully applied to individuals as well as to societies as a whole.

Culture Conflict

Cultures often come into conflict, a source of stress that may increase the suicide rate. For example, the conflict between traditional Native American culture and the dominant Anglo-American culture has often been viewed as playing a major role in precipitating Native American suicide. Philip May reviewed three hypotheses that might have relevance here. In the social disorganization hypothesis, the dominance of the Anglo-American culture is viewed as forcing Native American culture to change and as eroding traditional cultural systems and values. This changes the level of social regulation and social integration, important causal factors for suicide in Durkheim's theory of suicide.

A second hypothesis focuses on cultural conflict. The pressure from the educational system and mass media on Native Americans, especially on youth, to acculturate, a pressure that is opposed by their elders, leads to great stress for the youth. A third hypothesis focuses on the breakdown of the family in Native American tribes. Parents are often unemployed, substance abusers, and in trouble with the law, and divorce and desertion of the family by one or both parents are common. The suicide rates of three groups of Native Americans in New Mexico, the Apache, Pueblo, and Navajo (whose suicide rates were 43.3, 27.8, and 12.0 per 100 000 per year, respectively), were in line with their levels of acculturation as rated by May (high, moderate, and low, respectively).

Measuring Regional Stress

Linsky and co-workers measured the stress level of each of the U.S. states in three areas. For economic stress, they used business failures, unemployment claims, strikes, bankruptcies, and mortgage foreclosures. For family stress, they used divorces, abortions, illegitimate births, and fetal and infant deaths. For community stress, they used disasters, new housing starts, new welfare recipients, high school dropouts, and interstate migration. They found that states with

higher levels of stress also had higher suicide rates. Interestingly, however, the ratings of the stress level of each state based on responses from a national survey, which asked residents about their perceived level of stress, were not associated with the state suicide rates.

Multivariate Studies of Regional Suicide Rates

The accuracy of official national and regional rates of suicide has been questioned by Jack Douglas and others, who have argued that these rates are biased by the values of the local coroners and medical examiners and of the resident populations. However, despite the fact that regions do have different standards

for classifying a death as suicide, Peter Sainsbury and Brian Barraclough showed that the suicide rates of immigrants to one country from different nations are in almost the same rank order as the suicide rates of nations of origin.

The suicide rates of 71 nations reporting to the World Health Organization are shown in Table 1 for males and females. An inspection of the table reveals two basic findings: (1) in all countries, males have higher suicide rates than females, which suggests that differential gender socialization (e.g., males are socialized to be more aggressive and to keep their feelings inside) significantly impacts suicide rates; and (2) suicide rates vary significantly among the 71 countries, which suggests that cultures differ in the degree to which suicide is viewed as an acceptable behavior.

Table 1 Suicide rates around the world^a

<i>Country</i>	<i>Year reported</i>	<i>Males</i>	<i>Females</i>	<i>Country</i>	<i>Year reported</i>	<i>Males</i>	<i>Females</i>
Lithuania	2002	80.7	13.1	Norway	2001	18.4	6.0
Russian Federation ^b	2002	69.3	11.9	Canada	2000	18.4	5.2
Belarus	2001	60.3	9.3	Chile	2001	18.2	3.0
Kazakhstan	2002	50.2	8.8	United States	2000	17.1	4.0
Latvia	2002	48.4	11.8	Hong Kong	2000	16.1	10.1
Estonia	2002	47.7	9.8	Northern Ireland	2002	15.9	3.5
Ukraine	2002	46.7	8.4	Puerto Rico	2000	15.2	1.4
Hungary	2002	45.5	12.2	Turkmenistan	1998	13.8	3.5
Slovenia	2002	44.4	10.5	Thailand	2000	13.5	3.7
Japan	2002	35.2	12.8	Argentina	2001	13.4	3.5
Finland	2002	32.3	10.2	Netherlands	2003	12.7	5.9
Belgium	1997	31.2	11.4	Spain	2001	12.2	3.7
Austria	2002	30.5	8.7	Uzbekistan	2000	11.8	3.8
Croatia	2002	30.2	10.0	El Salvador	1999	11.6	5.4
Uruguay	2000	29.0	5.5	Costa Rica	2002	11.6	2.0
France	2000	27.9	9.5	Singapore	2001	11.5	6.9
Moldova, Republic of ^c	2002	27.9	5.2	Italy	2001	11.1	3.3
Switzerland	2000	27.8	10.8	England and Wales	2002	9.8	2.8
Poland	2002	26.6	5.0	Israel	1999	9.8	2.3
Bulgaria	2002	25.6	8.3	Venezuela	2000	8.8	1.5
Korea, Republic of	2002	24.7	11.2	Panama	2000	8.4	1.3
Czech Republic	2002	24.5	6.1	Columbia	1999	8.1	2.4
Romania	2002	23.9	4.7	Bahrain	2000	7.2	0.3
Slovakia	2000	22.6	4.9	Brazil	2000	6.4	1.6
Trinidad & Tobago	1998	22.2	4.7	Mexico	2001	6.3	1.3
Cuba	2001	21.4	8.0	Ecuador	2000	6.0	2.6
Denmark	1999	21.4	7.4	Albania	2001	5.5	2.3
Ireland	2001	21.4	4.1	Greece	2001	5.3	0.9
Germany	2001	20.4	7.0	Armenia	2002	4.0	0.7
Australia	2001	20.1	5.3	Paraguay	2000	3.9	1.7
New Zealand	2000	19.8	4.2	Georgia	2001	3.4	1.1
Scotland	2002	19.7	5.9	Guatemala	1999	3.4	0.8
Kyrgyzstan	2002	19.1	4.0	Philippines	1998	1.8	0.6
Sweden	2001	18.9	8.1	Azerbaijan	2002	1.8	0.5
Portugal	2002	18.9	4.9	Egypt	2000	0.1	0.0
Mauritius	2000	18.8	5.2				

^aRate per 100 000 population from latest year reported, ranked according to male suicide rates.

^bExcluding Chechnya.

^cExcluding Transdniestria.

Source: World Health Organization Statistical Information Systems website.

Cutright and Fernquist term this phenomenon the culture of suicide.

Simpson and Conklin identified two clusters of variables that were associated with national suicide rates: (1) a cluster that had the highest loading from the percentage of Muslims in the population and (2) a cluster that seemed to assess economic development. Suicide rates were lower in nations with less economic development and where Islam was a more important religion. Two other clusters (Christianity and the Eastern bloc) were not associated with suicide rates.

In similar study of the social correlates of national suicide rates, Lester identified 13 orthogonal (independent) factors for social variables, only one of which was associated with suicide rates, a factor that seemed to measure economic development (with high loadings from such social variables as low population growth and high gross domestic product per capita). Thus, these two studies agreed in finding that economic development is associated with higher suicide rates. Cutright and Fernquist reported that classical Durkheimian variables measuring social integration (i.e., divorce, fertility, and religiosity), as well as female labor-force participation, were strong predictors of cross-national suicide rates for both males and females of different age groups.

For the United States, a similarly designed study conducted by Lester identified a cluster of variables that seemed to measure social disintegration (high divorce and interstate migration rates, low church attendance, and high per capita alcohol consumption), and this was the strongest correlate of the suicide rates of the states.

Nonfatal Suicide Behavior

Almost all sociological research and theories have hitherto been proposed for fatal (completed) suicide. Nonfatal suicidal behavior has been relatively ignored. It is important that better epidemiological studies of attempted suicide and suicidal ideation be conducted so that sociologists can develop theories of these behaviors. The World Health Organization has sponsored a comparative epidemiological study of attempted suicide in 15 sites. Unfortunately, the sites are limited to cities in European nations and do not encompass these nations as a whole. However, the study is innovative and may lay the groundwork for more comprehensive epidemiological studies in the future. Linsky and colleagues have obtained estimates of the incidence in the previous year of suicidal ideation in U.S. states and sought social correlates of this.

Thus, such research is possible, and it is to be hoped that more studies are planned along these lines.

See Also the Following Articles

Affective Disorders; Aggression; Cultural Transition; Depression Models; Suicide, Biology of; Suicide, Psychology of.

Further Reading

- Cutright, P. and Fernquist, R. M. (2000). Effects of societal integration, period, region, and culture of suicide on male age-specific suicide rates: 20 developed countries, 1955–1989. *Social Science Research* 29, 148–172.
- Cutright, P. and Fernquist, R. M. (2000). Societal integration, culture, and period: their impact on female age-specific suicide rates in 20 developed countries, 1955–1989. *Sociological Focus* 33, 299–319.
- Cutright, P. and Fernquist, R. M. (2003). The gender gap in suicide rates: an analysis of twenty developed countries, 1955–1994. *Archives of Suicide Research* 7, 323–339.
- Douglas, J. D. (1967). *The social meanings of suicide*. Princeton, NJ: Princeton University Press.
- Durkheim, E. (1897). *Le suicide*. Paris: Felix Alcan. [English edition: (1951). *Suicide*. New York: Free Press.]
- Henry, A. F. and Short, J. F. (1954). *Suicide and homicide*. New York: Free Press.
- Lester, D. (1988). *Why women kill themselves*. Springfield, IL: Charles Thomas.
- Lester, D. (1989). *Suicide from a sociological perspective*. Springfield, IL: Charles Thomas.
- Lester, D. (1994). *Patterns of suicide and homicide in America*. Commack, NY: Nova Science.
- Lester, D. (1994). *Patterns of suicide and homicide in the world*. Commack, NY: Nova Science.
- Lester, D. (1995). Thwarting disorientation and suicide. *Cross-Cultural Research* 29, 14–26.
- Linehan, M. (1973). Suicide and attempted suicide. *Perceptual and Motor Skills* 37, 31–34.
- Linsky, A. S., Bachman, R. and Straus, M. A. (1995). *Stress, culture, and aggression*. New Haven, CT: Yale University Press.
- Sainsbury, P. and Barraclough, B. M. (1968). Differences between suicide rates. *Nature* 220, 1252.
- Simpson, M. E. and Conklin, G. H. (1989). Socioeconomic development, suicide and religion. *Social Forces* 67, 945–964.
- Van Winkle, N. W. and May, P. A. (1986). Native American suicide in New Mexico, 1959–1979. *Human Organization* 45, 296–309.

Relevant Website

World Health Organization Statistical Information Systems website. <http://www3.who.int/whosis/menu.cfm>.

Surgery and Stress

C Vögele

Roehampton University, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by C Vögele, volume 3, pp 553–554, © 2000, Elsevier Inc.

Stressful Characteristics of the Surgical Situation
Responses of Surgical Patients at Emotional, Cognitive, and
Physiological Levels
Surgical Stress and Postoperative Recovery
Psychological Preparation for Surgery

Glossary

<i>Psychological preparation for surgery</i>	Psychological methods aimed at reducing preoperative anxiety and aiding the recovery process.
<i>Surgical outcomes</i>	Variables reflecting the quality of postoperative recovery, for example, postoperative pain, pain medication length of stay, clinical recovery measures (e.g., wound healing), physiological indices (heart rate, blood pressure, palmar sweat gland activity, and cortisol level), negative affect, and satisfaction.
<i>Surgery</i>	A subset of invasive medical procedures involving some form of anesthesia.

The introduction into the health-care system and hospitalization can be a stressful experience on their own. Among the aspects of hospitalization considered to be most stressful, the following were identified in a sample of elderly medical inpatients: communication problems with health-care professionals; diagnostic and therapeutic procedures; the hospital environment (noise, rigid routines, etc.); worries about the home situation and the separation from home; insufficient information about diagnosis and prognosis; and fear of dependency, loss of autonomy and control. In addition to uncertainties about their illnesses (because of unclear communication) and the unfamiliar surroundings of a hospital ward, therefore, patients encounter additional stress because they must undergo medical procedures and examinations. Even some outpatient procedures such as dental treatments or blood donation can be stressful. Normally patients report increases in anxiety in anticipation of procedures such as surgery, endoscopy, cardiac catheterization, cancer screening, and chemotherapy. In addition to the anticipatory anxiety

associated with the procedure itself, diagnostic procedures such as cancer screening are characterized by a prolonged period of anxiety between screening and receiving the result.

Of all medical procedures, surgery is perhaps the most threatening event because it contains many unpredictable and uncontrollable features such as losing consciousness due to the administration of a general anesthetic, the anticipation of postoperative pain, and the surgical trauma related to the incision. This article concentrates on surgical procedures and describes the stressful characteristics associated with being hospitalized for surgery. We review the literature on the psychological and physiological responses to the experience of surgery and summarize findings from studies investigating psychological preparation for surgery.

Stressful Characteristics of the Surgical Situation

There are obvious potential threats for the surgical patient: anesthesia, pain, physical restriction, life-threatening procedures, and being away from home. Evidence suggests that the lack of predictability and control are significant contributors to the stressful experience of surgical patients. It is usually the case that we need to be able to predict an event in order to be able to control it. But the reverse does not follow – being able to predict an event does not necessarily mean we can control it. Such is the case for elective surgery, which accounts for the vast majority of surgical interventions.

This evidence confirms that there are common characteristics of the surgical situation that are identified by most patients as stressful. However, it seems likely that different types of surgical procedures produce different types of stress. A useful way of distinguishing between the various procedures is by considering the function of the procedure (diagnostic, treatment, or both) and the time line and nature of stress associated with the procedure. A further distinction can be drawn between procedural stress (i.e., the stress associated with the negative aspects of the actual procedure itself) and outcome stress (i.e., longer-term fears and concerns related to the results of the treatment or procedure). To illustrate the latter point, some operations may have more positive characteristics in terms of their expected outcome than others, for example, restoration (hip replacement) versus the removal of physical function (leg amputation).

Responses of Surgical Patients at Emotional, Cognitive, and Physiological Levels

Emotional Responses

Most studies investigating emotional responses to surgery have shown elevated levels of anxiety both before and after surgery. In some groups of patients, postoperative anxiety levels may be even higher than those measured preoperatively. Previous results indicating that anxiety is high before but not after surgery have been dismissed as being methodically flawed. Although anxiety has been, for obvious reasons, the most frequently assessed emotional response, it has been shown that surgical patients may also experience high levels of nervousness, depression, anger, and boredom.

In summary, the literature suggests that emotional responses to the surgical situation may vary as a function of the type of operation and, therefore, outcome concerns. Study results also point to the possibility of effective interventions not only before but also after surgery.

Cognitive Responses

Several studies have shown that patients' main worries are more related to the outcome of the operation than to the operation itself. Patients also worry significantly about normal everyday matters such as family and home, perhaps exacerbated by hospitalization and the impending surgery. As in the treatment of depression, the cognitive elements of surgical patients' stress response may be critical in indicating the most promising intervention approach to alleviate distress.

Physiological Responses

The most thorough research on physiological responses to surgery has been carried out on indices of sympathetic-adrenomedullary activity, both because they are readily measurable in surgical patients and because this highly responsive physiological axis has been widely used in well-controlled laboratory studies of psychological stress. The most consistent finding is that of a reduction in palmar sweat gland activity prior to surgery followed by recovery to normal levels postoperatively. This pattern follows closely the pattern of subjective distress and pain, indicating that sweat gland activity is reduced by the stress of surgery. Some authors, however, have argued that the palmar sweating pattern of surgical patients is related to the effort rather than to the distress aspect

of stress, that is, that changes in palmar sweating are due to changes in activity levels.

A more recent line of research has investigated neuroendocrine and immune changes in response to surgery and wound healing. Postoperative elevations in plasma levels of adrenaline, cortisol, and β -endorphins reflect sympathetic nervous system and hypothalamic-pituitary-adrenal axis activation. Evidence for immune suppression during surgery comes from studies showing the suppression of natural killer cell activity, proliferation of lymphocyte responses to mitogens, and changes in lymphocyte populations.

In order to interpret physiological responses to surgery as psychophysiological phenomena, it is necessary to disentangle the effects that are due to the experience of stress from those that are caused by the surgical trauma and other medical interventions (e.g., anesthesia). Apart from anesthesia and surgery itself, physiological changes may be caused by alterations in diet and ambulation and by medications, especially starvation before surgery and premedication or sedation on the night prior to surgery. Although these problems have at least been partly addressed in studies of the sympathetic-adrenomedullary axis, studies of other pathways are still at a very early stage and cannot support any firm conclusions.

Surgical Stress and Postoperative Recovery

Despite a strong effect of physical factors such as extent of tissue damage caused by the surgical trauma on postoperative recovery, there is variability across patients who have undergone the same procedure. It is unclear what accounts for this variability, and this has led to the suggestion that psychological factors such as anxiety and depression may play some part in determining the duration or quality of postoperative recovery.

There is considerable theoretical and empirical support for the relevance of psychosocial factors to postoperative recovery. These include demographic variables, such as age, gender, and socioeconomic status, and also personality variables, such as anxiety, neuroticism, coping style and social support. Most of the available literature today suggests a linear association between preoperative anxiety and success of recovery, in that the more anxious the patient is before the operation, the less favorable the postoperative recovery. There is even evidence that preoperative anxiety is related to intraoperative adjustment.

There is increasing evidence from psychoneuroimmunological work that stress and pain have a negative effect on endocrine and immune functions and

may thereby lead to delays in wound healing and other clinical complications. Although wound healing is only one postoperative outcome among a whole range, it is certainly a central parameter for short-term outcome in recovery from surgery. Kiecolt-Glaser et al. offer a biobehavioral model suggesting a number of routes through which psychological and behavioral responses can influence surgery and postoperative recovery.

Psychological Preparation for Surgery

There is now substantial agreement that psychological preparation for surgery is beneficial for patients. It is important, however, to establish which benefits can be achieved by psychological preparation and if all forms of preparation are equally effective. Results from a number of meta-analyses suggest that significant benefits can be obtained on all the major outcome variables (i.e., negative affect, pain, pain medication, length of stay, behavioral and clinical

indices of recovery, physiological indices, and satisfaction). The methods of psychological preparation employed include the provision of information about sensations and procedures, behavioral instructions, cognitive interventions, relaxation, and hypnosis. Procedural information and behavioral instructions show the most ubiquitous effects, both being clearly effective in improving all postoperative outcomes.

Further Reading

- Johnston, M. and Vögele, C. (1993). Benefits of psychological preparation for surgery: a meta-analysis. *Annals of Behavioral Medicine* 15, 245–256.
- Kiecolt-Glaser, J. K., Page, G. G., Marucha, P. T., et al. (1998). Psychological influences on surgical recovery: perspectives from psychoneuroimmunology. *American Psychologist* 53, 1209–1218.
- Vögele, C. (2004). Hospitalisation and stressful medical procedures. In: Kaptein, A. & Weinman, J. (eds.) *Health psychology*, pp. 288–304. Oxford: BPS Blackwell Publishers.

Survivor Guilt

P Valent

Melbourne, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Valent, volume 3, pp 555–557, © 2000, Elsevier Inc.

History of the Concept of Survivor Guilt
Moral Features of Survivor Guilt
Sense and Purpose of Survivor Guilt
Range of Manifestations of Survivor Guilt
Treatment of Survivor Guilt

Glossary

<i>Guilt</i>	An internal emotional judgment (an aspect of conscience) of having been bad or having done wrong to another.
<i>Survivor</i>	A person who has lived through a trauma.
<i>Trauma</i>	An experience in which a person's life has been grossly threatened and out of which a variety of biological, psychological and social wounds and scars results.

History of the Concept of Survivor Guilt

Survivor guilt as a clinical entity was emphasized for the first time in Holocaust literature in the 1960s. Psychiatrists such as Krystal and Niedeland were struck with the ubiquity of feelings of intense, unabating guilt for having survived among their Holocaust survivor patients. Survivor guilt came to be quickly recognized in other traumatic situations, including the Hiroshima atom bomb, combat, disasters, and civilian bereavement. It was noted after 9/11; indeed, it can occur in all traumatic situations, whether natural or man made.

A special form of survivor guilt may be said to occur in rescuers and helpers who blame themselves for not having saved those for whom they viewed themselves as responsible. More subtle forms can occur in clinicians who feel guilt for their own well-being in the face of their patients' suffering.

Moral Features of Survivor Guilt

The most striking feature of survivor guilt is its moral judgment of self-blame. It is directed to the survivor's

undeserved survival in contrast to others' deaths. An individual suffering from survivor guilt often states that the dead should be alive instead of him- or herself. This feeling may be enhanced when the individual was saved by another who died.

A common subtype of survivor guilt relates to not having saved others. The implication is that the survivor could have and should have saved others, but due to selfishness did not, instead saving him- or herself, thereby causing suffering and death to others. This feeling is enhanced if the person's usual role, such as spouse, parent, or rescuer, was to protect those who died, or if the person did indeed put his or her survival ahead of others.

However, a striking feature of survivor guilt is the contrast between the torment of self-blame of survivors and their innocence by any objective criteria. For instance, Holocaust survivors often blamed themselves for not following or saving their loved ones who were murdered, when actually they were separated at gunpoint and through physical force. Children who survived by hiding blamed themselves for the deaths of their parents. The Holocaust taught the paradox that victims are prone to survivor guilt while perpetrators rationalize theirs.

Unjustified survivor guilt occurs in all traumatic situations. Medical staff may blame themselves for not saving unsavable patients. A child may grow up with unwarranted guilt for separation from parents or parental death.

Sense and Purpose of Survivor Guilt

Observations of acute phases of traumatic situations indicate that survivor guilt is an unpleasant, instinctive social cue (that may be reinforced by others' anger and blame) that can be appeased only through helping others. For instance, in a disaster those whose houses survive often appease their very intense guilt by sheltering those whose houses are destroyed. Survivor guilt thus helps to preserve as many people (or genes) in the community as possible.

Later survivor guilt may be less useful practically, though it may prime adaptive reparative action for subsequent traumatic situations. It may also preserve the survivor's retrospective hope that 'if only's may yet be able to be executed. Further, Danieli suggests that the guilt is a buffer against moral chaos and existential helplessness, as it preserves some assumption of moral order in the universe. Last, the guilt perpetuates close emotional links to the dead.

Range of Manifestations of Survivor Guilt

If survivor guilt is appeased by effective rescue of others, it may be seen as a temporary stress. If it

fails and become part of a trauma, it is lived and relived as part of the wounds and scars of the trauma. Then the following biological, psychological, and social manifestations may occur singly or in a variety of combinations.

Physiological Responses

Thus far, the physiological accompaniments of survivor guilt have not been delineated.

Psychological Responses

Emotion Survivor guilt is one of the most painful human emotions. It may be felt as both mental anguish and a physical heart-wrenching pain in the chest directed toward the dead, but turned angrily back on the survivor him- or herself.

Cognitions Poignant traumatic events may return in feelings, thoughts, images, dreams, and flashbacks, along with ruminations about what the survivor should have done but did not do. Cognitive schemas and meanings develop, such as that the survivor is an irresponsible person, even a destroyer of life.

Defenses Intense pain of survivor guilt such as having caused the death of beloved peoples, may be central in survivor defense formations. Psychic numbing, dissociation, and repression may fragment coherent awareness. Identification may lead to physical symptoms similar to the ones the dead had suffered. Displacement may lead to blaming others. Sublimation may lead to devotion to saving others. Some survivors may take on rescuer or helper professions.

Social Responses

In acute situations, people may vacillate between appeasing survivor guilt by helping others and other essential survival options. After the acute phase of the trauma, people may wring their hands as they anguish about the choices they made. They may blame themselves and withdraw in order to avoid the imagined blame of the question "Why did you survive?" Survivors may play down their survivorship, and may especially avoid enjoyment and vitality, as they can exacerbate their feelings of guilt. Sublimatory activity may be channeled through choosing helping or rescuing work and professions.

Specific social settings may lead to characteristic responses. For instance, bereaved parents often withdraw to extremes, and emotional numbing admixed with intense blame and self-blame often result in divorce.

Associated Clinical Features

Survivor guilt may be associated with other judgments, such as shame and injustice. The sense of

injustice may come from having been unfair to others, but also from having been put in unfair situations.

Survivor guilt may interfere with grieving and is often at the kernel of unresolved grief and depression. Hence it may need to be resolved before the grieving process can proceed.

Treatment of Survivor Guilt

Prevention of survivor guilt is part of the rationale for strict, following of protocols, early disaster intervention, debriefing, and grief therapy. Its resolution is a common goal in later psychotherapies.

The first principle of treatment of survivor guilt is its recognition. Next, thorough investigation of the facts of the circumstances reveals how survivor guilt is objectively untenable. Its evolutionary sense and hence its presence is explained, and alternative, hopeful cognitive views are explored. With the thawing defenses of the survivor, the emotional pain of survivor guilt is identified, processed, and assimilated in the context of an empathic, therapeutic relationship. The survivor's sense of morality is retrieved with the realization that past circumstances, not the survivor,

were irrational and bad. Seeing him- or herself as a victim, not as a perpetrator, allows the survivor to grieve losses and to achieve new hopeful meanings for him- or herself. The survivor can then reclaim a purposeful life.

See Also the Following Articles

Depression and Manic-Depressive Illness; Holocaust, Stress Effects of; 9/11, Religion and Stress.

Further Reading

- Danieli, Y. (1988). Treating survivors and children of the Nazi Holocaust. In: Ochberg, F. M. (ed.) *Post-traumatic therapy and victims of violence*. New York: Brunner/Mazel.
- Krystal, H. and Niedeland, W. C. (eds.) (1971). *Psychic traumatization: aftereffects in individuals and communities*. Boston, MA: Little, Brown and Company.
- Raphael, B. (1986). *When disaster strikes*. London: Hutchinson.
- Valent, P. (1998). *From survival to fulfillment: a framework for the life-trauma dialectic*. Philadelphia, PA: Brunner/Mazel.

Sympathetic Nervous System

D S Goldstein

National Institute of Neurological Disorders and Stroke,
Bethesda, MD, USA

Published by Elsevier Inc.

Epinephrine

The main hormone of the adrenomedullary hormonal system.

Norepinephrine

The sympathetic neurotransmitter.

Historical and Conceptual Overview

Role of the Sympathetic Nervous System in Homeostasis
Norepinephrine: The Sympathetic Neurotransmitter

Glossary

<i>Sympathetic nervous system</i>	A major component of the autonomic nervous system. The main sympathetic neurotransmitter for cardiovascular regulation is norepinephrine. The main sympathetic neurotransmitter for sweating is acetylcholine.
<i>Adrenomedullary hormonal system</i>	The hormonal component of the autonomic nervous system.
<i>Norepinephrine</i>	The main sympathetic neurotransmitter for cardiovascular regulation.

Historical and Conceptual Overview

The term sympathetic nervous system originates in the second-century teaching of Galen that the peripheral nerves, conduits for distributing the animal spirit in the body, enable concerted, coordinated (i.e., sympathetic) functioning of body organs. The notion of the sympathetic nervous system, therefore, antedated by approximately 14 centuries Harvey's demonstration of the circulation of the blood.

The sympathetic nervous system is a major component of the autonomic nervous system. In the late nineteenth century, Langley introduced the phrase autonomic nervous system to refer to nerves emanating from cell bodies in ganglia alongside the spinal column and in the gastrointestinal-tract walls. Langley defined three components: sympathetic, with preganglionic

cells in the thoracolumbar spinal cord; parasympathetic (a word he coined), with preganglionic cells in the brain stem or sacral spinal cord; and enteric, with preganglionic cells near or in the target organs. The nerves of these systems were thought to differ from those of the somatic nervous system, not only in terms of anatomy but also in terms of function. Somatic nerves mediate conscious, observable changes in skeletal muscle contraction and therefore movement, whereas autonomic nerves mediate unconscious, largely unobservable changes in smooth muscle contraction independently (autonomously) of the central nervous system.

The early-twentieth century American physiologist Walter B. Cannon introduced the word homeostasis. Cannon added a hormonal component to the autonomic nervous system when he theorized that the sympathetic nervous system and the adrenal gland worked together as a unit to maintain homeostasis in emergencies. In the 1930s, he even formally proposed that the sympathetic nervous system used the same chemical messenger – adrenaline – as did the adrenal gland (in 1946, von Euler correctly identified the sympathetic neurotransmitter in mammals as norepinephrine). Accumulating evidence supports the independent regulation of the sympathetic nervous and adrenomedullary hormonal systems, refuting the concept of a unitary sympathoadrenal system. Nevertheless, Cannon's views about the unitary function of the neural and hormonal components of the sympathoadrenal system prevail in stress research.

Sweating, whether to control body temperature (thermoregulatory), evoked by eating spicy food (gustatory), or accompanying fear or anxiety (emotional), is mediated by the sympathetic cholinergic system, in which acetylcholine rather than norepinephrine is the chemical messenger. The autonomic nervous system, therefore, can be viewed as having five components: sympathetic noradrenergic, sympathetic cholinergic, parasympathetic cholinergic, adrenomedullary, and enteric.

Role of the Sympathetic Nervous System in Homeostasis

Adjustments in sympathetic nervous activity accompany virtually every human action and reaction and constitute the main means for the acute regulation of cardiovascular function. According to Goldstein's 1995 theory, comparator homeostats use the sympathetic nervous system as an effector for different stress responses. The sympathetic nervous and adrenomedullary hormonal effectors interact in complex ways with many others, including the hypothalamic–pituitary–adrenocortical system, the endogenous

opiate system, vasopressin, the parasympathetic nervous system, and the renin–angiotensin–aldosterone system, in determining responses to stressors.

The sympathetic nervous system consists of nerve networks. Because the axons come from cells in ganglia rather than directly from the spinal cord or brain stem, sympathetic nerves supplying body organs are mainly postganglionic; however, most sympathetic nerves supplying adrenomedullary chromaffin cells are preganglionic, releasing acetylcholine and thereby stimulating membrane nicotinic receptors.

Stresses associated with sympathetic nervous system activation typically include a component of active movement. Patterned sympathetic nervous system activation during stress produces adaptive shifts in the distribution of blood volume or in glandular secretion ([Figure 1a](#)). When these changes suffice to maintain homeostasis, the organism does not experience them consciously, but when the brain senses that these responses are not or will not mitigate the effects of the stressor, the situation reaches consciousness and adrenomedullary hormonal system activation ensues ([Figure 1b](#)). Examples of situations associated with prominent increases in sympathoneural outflows include orthostasis, exercise, the postprandial state, cold exposure, and the performance of nondistressing locomotor tasks.

In response to perceived global metabolic threats from external or internal stimuli, increased neural outflow to the adrenal medulla elicits catecholamine secretion into the adrenal venous drainage. In humans, the predominant catecholamine in the adrenal venous drainage is epinephrine (synonymous with adrenaline). Epinephrine rapidly reaches all cells of the body, with the exception of most of the brain, producing a wide variety of hormonal effects at low blood concentrations. We can comprehend the many effects of epinephrine in terms of the body's countering of acute threats to survival that mammals have perennially faced, such as sudden lack of metabolic fuels, trauma with hemorrhage, intravascular volume depletion and hypotension, and fight-or-flight confrontations. Distress accompanies all these situations, the experience undoubtedly fostering the long-term survival of the individual and the species by motivating avoidance learning, enhancing the registration of the event in long-term memory, and producing signs universally understood among other members of the species.

Stresses eliciting adrenomedullary hormonal system activation typically also elicit hypothalamic–pituitary–adrenocortical activation, as indicated by high circulating levels of corticotropin or glucocorticoids; increases in release of endogenous opioids, as indicated by high plasma levels of β -endorphin; and

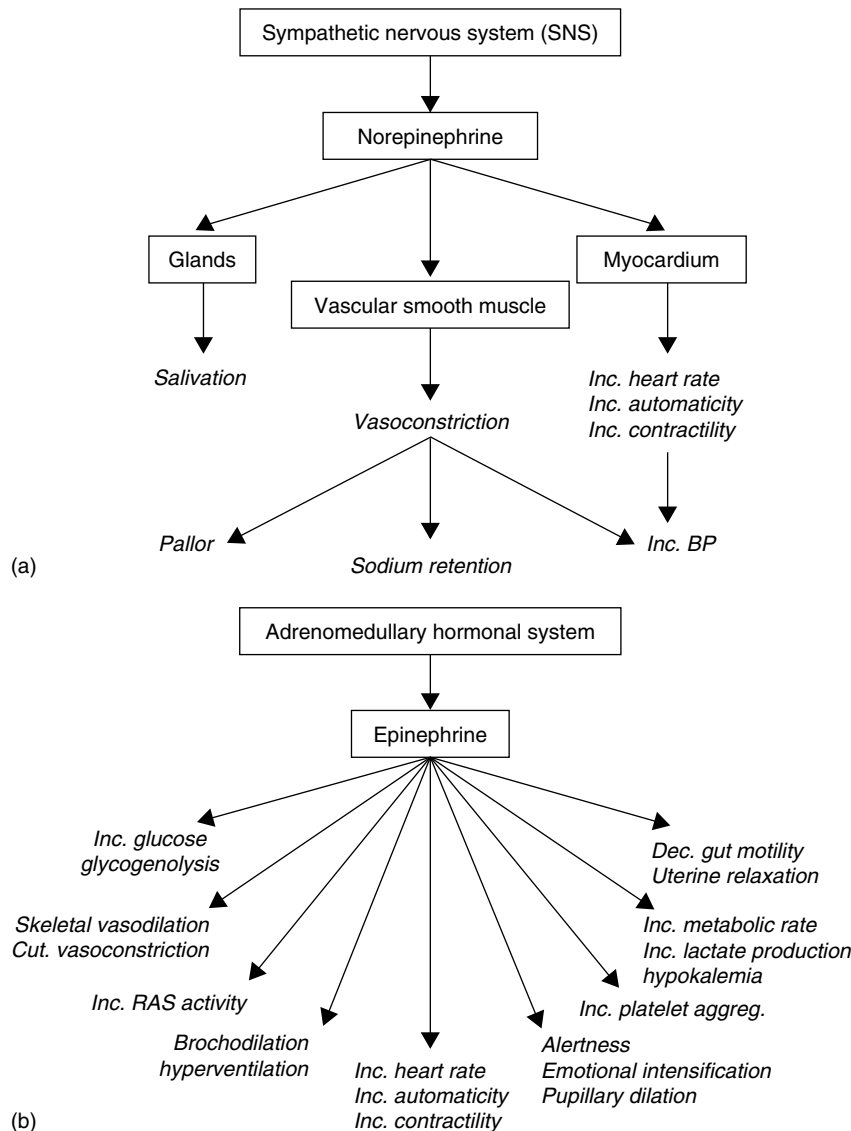


Figure 1 Overview of effects of the activation of (a) the sympathetic nervous system, where norepinephrine acts as the main neurotransmitter, and (b) the adrenomedullary hormonal system, where epinephrine acts as the main hormone. Because the sympathetic nervous system is a neural network, sympathetic nerves to different regions can be activated differentially during stress, producing redistributions of blood flow to body organs. Norepinephrine released from sympathetic nerves generally contracts smooth muscle cells, eliciting glandular secretion, vasoconstriction, and myocardial contraction. Renal vasoconstriction produces an anti-natriuretic effect. Cutaneous vasoconstriction produces pallor. The net effect of systemic vasoconstriction is to increase the total peripheral resistance to blood flow in the body. This, combined with myocardial stimulation, increases blood pressure. In humans, the main hormone released by the adrenal medulla is epinephrine (adrenaline). Epinephrine affects the function of most organs and produces generalized effects such as increased metabolic rate, increased blood glucose levels, hypokalemia, increased lactate production, increased alertness, increased platelet aggregability, and intensification of emotions. Aggreg., aggregation; BP, blood pressure; Cut., cutaneous; Dec., decrease; Inc., increase; RAS, renin–angiotensin–aldosterone system.

vasopressin secretion even when sympathetic nervous system activity declines.

Norepinephrine: The Sympathetic Neurotransmitter

The chemical transmitter at most cardiovascular sympathetic nerve endings is norepinephrine. The binding

of norepinephrine to adrenoceptors on cardiovascular smooth muscle cells contracts the cells. The body's myriad arterioles largely determine total resistance to blood flow and therefore contribute in an important way to blood pressure. Sympathetic nerves enmesh blood vessels in latticelike networks in the adventitial outer surface that extend inward to the adventitial-medial border, with the concentration of sympathetic

nerves increasing as arterial caliber decreases, so that small arteries and arterioles, the smallest nutrient vessels possessing smooth muscle cells, possess the most intense innervation. The unique architectural association between sympathetic nerves and the vessels that determine peripheral resistance has enticed cardiovascular researchers, particularly in the area of autonomic regulation, for many years. Sympathetic vascular innervation varies widely among vascular beds, with dense innervation of resistance vessels in the gut, kidney, skeletal muscle, and skin. Sympathetic stimulation in these beds produces profound vasoconstriction, whereas stimulation in the coronary, cerebral, and bronchial beds elicits weaker constrictor responses, consistent teleologically with the goal of preserving blood flow to vital organs during

stress. Cardiac sympathetic stimulation increases the force and rate of the heartbeat and also augments intracardiac electrical conduction and automaticity.

Norepinephrine released from sympathetic nerve terminals acts mainly locally. Only a small proportion of released norepinephrine reaches the bloodstream. The indirect and distant relationship between plasma norepinephrine levels and sympathetic nerve activity can complicate interpreting plasma norepinephrine levels in response to stressors, pathophysiological situations, and drugs.

Norepinephrine Synthesis

The enzymatic steps in norepinephrine synthesis have been characterized in more detail than those for any other neurotransmitter (Figure 2). Catecholamine

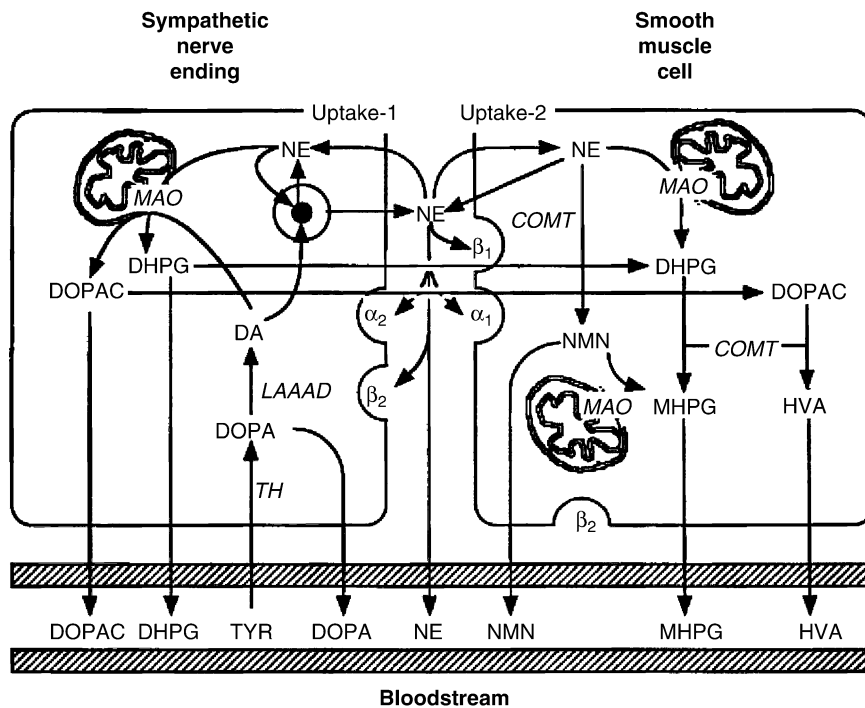


Figure 2 Overview of steps in norepinephrine synthesis, release, reuptake, and metabolism. Hydroxylation of the amino acid tyrosine is the enzymatic rate-limiting step in catecholamine biosynthesis. Tyrosine hydroxylase catalyzes the conversion of tyrosine to DOPA. L-aromatic-amino acid decarboxylase (LAAAD) catalyzes the conversion of DOPA to dopamine. Dopamine-β-hydroxylase is localized to large dense-core vesicles in noradrenergic cells. After uptake of axoplasmic dopamine into the vesicle, dopamine-β-hydroxylase catalyzes the conversion of dopamine to norepinephrine. According to the exocytosis theory, increased transmembrane entry of ionized calcium fosters fusion of the vesicles with the axonal membrane. Endogenously released norepinephrine is taken back up into sympathetic nerve terminals by Uptake-1 and taken up into nonneuronal cells by Uptake-2. Norepinephrine in the axoplasm can be translocated back into the vesicles or converted to dihydroxyphenylglycol by monoamine oxidase (MAO) in the outer mitochondrial membrane. Norepinephrine in nonneuronal cells is converted to normetanephrine by catechol-O-methyltransferase. Dihydroxyphenylglycol can diffuse into the bloodstream or can be taken up by nonneuronal cells and converted to methoxyhydroxyphenylglycol, which is a major end product of norepinephrine metabolism. Dopamine in the axoplasm can be translocated into vesicles and converted to norepinephrine or can be converted to dihydroxyphenylacetic acid by MAO. Dihydroxyphenylacetic acid can diffuse into the bloodstream or can be converted extraneuronally to homovanillic acid, the main end product of dopamine metabolism. The processes of catecholamine uptake and metabolism are generally independent of actions by adrenoceptors on the nerve terminals or nonneuronal cells. COMT, catechol-O-methyltransferase; DA, dopamine; DHPG, dihydroxyphenylglycol; DOPA, 3,4-dihydroxyphenylalanine; DOPAC, dihydroxyphenylacetic acid; HVA, homovanillic acid; LAAAD, L-aromatic-amino acid decarboxylase; MAO, monoamine oxidase; MHPG, methoxyhydroxyphenylglycol; NE, norepinephrine; NMN, normetanephrine; TH, tyrosine hydroxylase; TYR, tyrosine.

biosynthesis begins with the uptake of the amino acid tyrosine into the cytoplasm of sympathetic neurons and adrenomedullary cells. Tyrosine hydroxylase catalyzes the conversion of tyrosine to 3,4-dihydroxyphenylalanine (DOPA). This is the enzymatic rate-limiting step in catecholamine synthesis. The enzyme is almost saturated under normal conditions, so alterations in dietary tyrosine intake normally do not affect the rate of catecholamine biosynthesis under baseline conditions. Tetrahydrobiopterin, ionic iron, and molecular oxygen are required for tyrosine hydroxylase activity. Exposure to stressors that increase sympathoadrenal outflows augments the synthesis and concentration of tyrosine hydroxylase by multiple and complex mechanisms. Short-term mechanisms include feedback inhibition and phosphorylation of the enzyme, the latter depending on membrane depolarization, contractile elements, and receptors. Long-term mechanisms include changes in tyrosine hydroxylase synthesis, probably determined by transsynaptic induction and retrograde transport.

L-aromatic-amino-acid decarboxylase (LAAAD, also called DOPA decarboxylase) in the neuronal cytoplasm catalyzes the rapid conversion of DOPA to dopamine. Many types of tissues contain this enzyme. The activity of the enzyme depends on pyridoxal phosphate, which is vitamin B-6. Although LAAAD metabolizes most of the DOPA formed in catecholamine-synthesizing tissues, some endogenously produced DOPA enters the circulation. This provides the basis for using plasma DOPA levels to examine catecholamine synthesis. By inhibiting the conversion of DOPA to dopamine in the periphery, LAAAD inhibitors such as carbidopa enhance the efficacy of L-DOPA treatment of Parkinson's disease.

Dopamine- β -hydroxylase catalyzes the conversion of dopamine to norepinephrine in tissues that synthesize catecholamines. The enzyme is confined to the vesicles in catecholamine-producing cells. Dopamine- β -hydroxylase contains, and its activity depends on, copper. Because of this dependence, children with Menkes disease, a rare, inherited X-linked recessive disorder of copper metabolism, have neurochemical evidence of decreased conversion of dopamine to norepinephrine and compensatorily increased tyrosine hydroxylation. Dopamine- β -hydroxylase activity also requires ascorbic acid (vitamin C), which provides electrons for the hydroxylation.

Norepinephrine Storage

Vesicles in sympathetic nerves actively remove and trap axoplasmic amines. The vesicular monoamine transporter resembles the neuronal membrane monoamine transporters structurally, but it is relatively nonspecific

for catecholamines and serotonin. Magnesium and ATP accelerate vesicular uptake, and reserpine effectively and irreversibly blocks it. Vesicles generated near the Golgi apparatus of the cell bodies travel by axonal transport in the axons to the nerve terminals. Noradrenergic vesicles may also form by endocytosis within the axons. Varicosities in sympathetic nerves contain two types of cytoplasmic vesicles: small dense-core (diameter 40–60 nm) and large dense-core (diameter 80–120 nm). Cores of both types of vesicle contain abundant ATP, the norepinephrine : ATP ratio averaging approximately 1 : 4. The vesicles also contain at least three types of polypeptides: chromogranin A, enkephalins, and neuropeptide Y. The extracellular fluid levels of each of these compounds have been considered as putative indices of exocytosis.

Norepinephrine Release

According to the exocytotic theory, acetylcholine depolarizes the terminal membranes by increasing membrane permeability to sodium ion. The increased intracellular sodium levels directly or indirectly enhance the transmembrane influx of calcium ion via voltage-gated channels. The increased cytoplasmic calcium concentration evokes a cascade of as yet incompletely defined biomechanical events, resulting in the fusion of the vesicular and axoplasmic membranes. The interior of the vesicle exchanges briefly with the extracellular compartment, and the soluble contents of the vesicles diffuse into the extracellular space. As predicted from this model, manipulations other than the application of acetylcholine that depolarize the cell, such as electrical stimulation or increased potassium ion concentrations in the extracellular fluid, also activate the voltage-gated calcium channels and trigger exocytosis.

During sympathetic neuronal or adrenomedullary chromaffin cell activation, the simultaneous stoichiometric release of soluble vesicular contents – ATP, enkephalins, chromogranins, and dopamine- β -hydroxylase – without a similar release of cytoplasmic macromolecules has provided biochemical support for the exocytosis theory of noradrenergic neurotransmission. Electron micrographs occasionally have shown an omega sign, with an apparent gap in the cell membrane at the site of fusion of vesicle with the axoplasmic membrane. Ultrastructural evidence has suggested the release of the contents of large as well as small dense-core vesicles by exocytosis.

Sympathetic nerve endings can also release norepinephrine by calcium-independent, nonexocytotic mechanisms. One such mechanism probably is reverse transport through the neuronal uptake carrier. The indirectly acting sympathomimetic amine, tyramine,

releases norepinephrine nonexocytotically because tyramine releases norepinephrine independently of calcium and does not release dopamine- β -hydroxylase. As predicted from the ionic requirements for neuronal reuptake, increases in intracellular sodium concentrations, such as produced by ouabain, enhance carrier-mediated efflux of norepinephrine. Myocardial ischemic hypoxia evokes the calcium-independent release of norepinephrine. The hydrophilic nature of catecholamines and their ionization at physiological pH probably prevent norepinephrine efflux by simple diffusion.

Pharmacological stimulation of receptors on noradrenergic terminals affects the amount of norepinephrine released during cellular activation. Substantial evidence supports inhibitory presynaptic modulation by norepinephrine itself via α_2 -adrenoceptors on sympathetic nerves. This modulatory action appears to vary with the vascular bed under study, being prominent in skeletal muscle beds such as the forearm, relatively weak in the kidneys, and virtually absent in the adrenals. In addition to local feedback control of norepinephrine release, reflexive long-distance feedback pathways via arterial and cardiac baroreceptors elicit reflexive changes in sympatho-neural impulse activity. Alterations in receptor numbers or of intracellular biomechanical events after receptor activation also affect responses to agonists. These factors therefore regulate norepinephrine release by trans-synaptic local and reflexive long-distance mechanisms.

Norepinephrine Disposition

Unlike acetylcholine, which is inactivated mainly by extracellular enzymes, norepinephrine is inactivated mainly by uptake into cells (Figure 2), with subsequent intracellular metabolism or storage. Reuptake into nerve terminals (Uptake-1) is the predominant means of terminating the actions of released norepinephrine. Dopamine is a better substrate than norepinephrine is for uptake via the cell-membrane norepinephrine transporter.

Neuronal uptake absolutely requires intracellular potassium ion and extracellular sodium ion and functions most efficiently when chloride ion accompanies the sodium. Uptake-1 requires energy and is carrier mediated. Transport does not directly require ATP; however, maintaining ionic gradients across cell membranes depends on ATP, and the carrier uses the energy expended in maintaining the transmembrane sodium gradient to cotransport amines with sodium. Many drugs or *in vitro* conditions inhibit Uptake-1, including cocaine and tricyclic antidepressants.

Norepinephrine taken up into the axoplasm by the Uptake-1 transporter is subject to two fates: translocation into storage vesicles and deamination by monoamine oxidase (MAO). The combination of enzymatic breakdown and vesicular uptake constitute an intraneuronal sink, keeping the cytoplasmic concentrations of norepinephrine very low. Reserpine effectively blocks the vesicular transport of amines from the axoplasm into the vesicles. This not only shuts down the conversion of dopamine to norepinephrine but also prevents the conservative recycling of norepinephrine. Reserpine, therefore, rapidly depletes norepinephrine stores.

Neural and nonneural tissues contain MAO, which catalyzes the oxidative deamination of dopamine to form dihydroxyphenylacetic acid and norepinephrine to form dihydroxyphenylglycol. Because of the efficient uptake and reuptake of catecholamines into the axoplasm of catecholaminergic neurons and because of the rapid exchange of amines between the vesicles and axoplasm, the neuronal pool of MAO, located in the outer mitochondrial membrane, figures prominently in the overall functioning of catecholaminergic systems. The isozyme MAO-A predominates in neural tissue, whereas both MAO-A and MAO-B exist in nonneural tissue. MAO inhibitors are effective antidepressants. A phenomenon known as the cheese effect limits their clinical use. In patients taking MAO inhibitors, the administration of sympathomimetic amines such as those found in many nonprescription decongestants or the ingestion of foods such as aged cheese, wine, and meat, which contain tyramine, can produce paroxysmal hypertension. Tyramine and other sympathomimetic amines increase the loss of norepinephrine from sympathetic vesicles, and the blockade of MAO augments the bioavailability of ingested tyramine. When given alone, MAO inhibitors usually decrease blood pressure and can produce orthostatic hypotension by unknown mechanisms.

Nonneuronal cells remove norepinephrine actively by a process called Uptake-2, characterized by the ability to transport isoproterenol; susceptibility to blockade by O-methylated catecholamines, corticosteroids, and β -haloalkylamines; and an absence of susceptibility to blockade by Uptake-1 blockers. In contrast with Uptake-1, Uptake-2 functions independently of extracellular sodium. The Uptake-2 carrier has little if any stereoselectivity and has low affinity and specificity for catecholamines. For instance, extraneuronal cells remove imidazolines such as clonidine by Uptake-2. Whereas reverse transport via the Uptake-1 carrier requires special experimental conditions, reverse transport via the Uptake-2 carrier is

readily demonstrated. Thus, during the infusion of a catecholamine at a high rate, the catecholamine can accumulate in extraneuronal cells, with reentry of the catecholamine into the extracellular fluid via the Uptake-2 carrier after the infusion ends.

Catechol-O-methyltransferase in nonneuronal cells and adrenomedullary chromaffin cells catalyzes the O-methylation of dihydroxyphenylglycol to form methoxyhydroxyphenylglycol and of dihydroxyphenylacetic acid to form homovanillic acid. Plasma levels of dihydroxyphenylglycol and dihydroxyphenylacetic acid probably mainly reflect neuronal metabolism of norepinephrine and dopamine. Catechol-O-methyltransferase also catalyzes the conversion of norepinephrine to normetanephrine and epinephrine to metanephrine. Uptake-2 and catechol-O-methyltransferase probably act in series to remove and degrade circulating catecholamines. The methyl group donor for the reaction is S-adenosyl methionine. Immunohistochemical studies have indicated extraneuronal localization of catechol-O-methyltransferase in most tissues, such as the liver and kidney. Because of the extraneuronal localization of catechol-O-methyltransferase, recently introduced assay methods for metanephrines have enabled refined assessments of extraneuronal metabolism of catecholamines.

Most of the metabolism of catecholamines takes place within the same cells where the amines are synthesized. This mainly occurs secondary to leakage of catecholamines from vesicular stores into the cytoplasm. The stores exist in a highly dynamic equilibrium, with passive outward leakage counterbalanced by inward active transport controlled by vesicular monoamine transporters.

In catecholaminergic neurons, the presence of MAO-A leads to the formation of reactive catecholaldehydes. The contribution of these toxic aldehydes to neurodegenerative processes depends on the dynamics of vesicular-axoplasmic monoamine exchange and enzyme-catalyzed conversion to inert deaminated acids or alcohols. In sympathetic nerves, the aldehyde produced from norepinephrine is converted to dihydroxyphenylglycol, not dihydroxymandelic acid, as textbooks often indicate. Subsequent extraneuronal O-methylation consequently leads to the production of methoxyhydroxyphenylglycol, not vanillylmandelic acid. Vanillylmandelic acid is instead formed in the liver by the oxidation of methoxyhydroxyphenylglycol catalyzed by alcohol dehydrogenase.

Compared to intraneuronal deamination, extraneuronal O-methylation of norepinephrine and epinephrine to metanephrines represent minor pathways of metabolism. The single largest source of metanephrines is the adrenal medulla. Vanillylmandelic acid and methoxyhydroxyphenylglycol, the products of the combined O-methylation and deamination of norepinephrine, are the two main end products of norepinephrine metabolism, vanillylmandelic acid (as just noted) formed mainly in the liver. Adrenomedullary chromaffin cells express catechol-O-methyltransferase, as do pheochromocytomas, which are clinically important chromaffin cell tumors. Pheochromocytoma cells produce large amounts of metanephrines from catecholamines leaking from stores. These metabolites are thus particularly useful for detecting these tumors. The large contribution of intraneuronal deamination to catecholamine turnover, and the dependence of this on the vesicular axoplasmic monoamine exchange process, help explain how synthesis, release, metabolism, turnover, and stores of catecholamines are regulated in a coordinated fashion.

See Also the Following Article

Autonomic Nervous System.

Further Reading

- Cannon, W. B. (1939). *The wisdom of the body*. New York: W. W. Norton.
- Eisenhofer, G., Kopin, I. J. and Goldstein, D. S. (2004). Catecholamine metabolism: a contemporary view with implications for physiology and medicine. *Pharmacological Reviews* 56, 331–349.
- Goldstein, D. S. (1995). *Stress, catecholamines, and cardiovascular disease*. New York: Oxford University Press.
- Goldstein, D. S. (2001). *The autonomic nervous system in health and disease*. New York: Marcel Dekker.
- Goldstein, D. S., Eisenhofer, G. and Kopin, I. J. (2003). Sources and significance of plasma levels of catechols and their metabolites in humans. *Journal of Pharmacology and Experimental Therapeutics* 305, 800–811.
- Pacak, K., Palkovits, M., Yadid, G., et al. (1998). Heterogeneous neuroendocrine responses to various stressors: a test of Selye's doctrine of non-specificity. *American Journal of Physiology* 275, R1247–R1255.
- Selye, H. (1950). *The physiology and pathology of exposure to stress: a treatise based on the concepts of the general-adaptation syndrome and the diseases of adaptation*. Montreal: Acta, Inc.

Syndrome X See: Metabolic Syndrome; Metabolic Syndrome and Stress.

Synthetic Glucocorticoids

A M Karssen

Stanford University School of Medicine, Stanford,
CA, USA

E R de Kloet

Leiden University Medical Center, Leiden,
The Netherlands

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by A M Karssen and E R de Kloet, volume 3, pp 566–570,
© 2000, Elsevier Inc.

Introduction

Chemical Aspects

Pharmacokinetics

Actions of Synthetic Glucocorticoids

Clinical Aspects

Glossary

*Blood–brain
barrier*

Dynamic barrier of endothelial cells between blood and brain, protecting the brain from potentially neurotoxic compounds and maintaining a constant internal environment in the brain.

*Corticotropin
releasing hor-
mone (CRH)*

A 41 amino acid residue peptide that is synthesized in hypothalamic nerve cells and mediates the neural control of ACTH release.

Glucocorticoids

Steroid hormones that affect carbohydrate and protein metabolism, regulate the activity of the hypothalamic-pituitary-adrenal (HPA) axis, and have anti-inflammatory and immunosuppressive effects. Endogenously, they are synthesized in the adrenal cortex, released on stimulation by adrenocorticotrophic hormone (ACTH).

Mineralocorticoids

Steroid hormones that affect water and electrolyte balance. Naturally occurring glucocorticoids often have also substantial mineralocorticoid action.

Introduction

Synthetic glucocorticoids are widely used in basic, preclinical, and clinical research. The importance of these compounds in stress research is based on their

suppressive effects on the activity of the hypothalamic-pituitary-adrenal axis (HPA axis). Their main clinical applications, however, are based on their effective anti-inflammatory and immunosuppressive actions. The first synthetic glucocorticoids were developed in the late 1950s as drugs for treating rheumatoid arthritis. In basic research the compounds are mainly used to unravel the properties of the glucocorticoid receptor (GR).

All synthetic glucocorticoids are structurally related to hydrocortisone (cortisol) and corticosterone, which are the main endogenous glucocorticoids of primates and rodents, respectively. However, synthetic glucocorticoids often bind to the GR with a higher affinity and thus have much more potent glucocorticoid actions on the HPA axis, the immune system, and energy metabolism than the naturally occurring hormones. Their mineralocorticoid actions on salt retention/sodium homeostasis, although reduced, are still a disadvantage, especially for long-term treatment. The synthetic glucocorticoid RU28362 lacks this mineralocorticoid activity and is a pure glucocorticoid.

The group of glucocorticoids includes many compounds (see [Table 1](#) for some examples). The most commonly used synthetic glucocorticoids are prednisone/prednisolone and dexamethasone.

Chemical Aspects

The structure of all glucocorticoids is basically the same. Most of them are 21-carbon polycyclic compounds with a steroid structure, a fused 17-carbon atom ring system with an aliphatic side chain at carbon-17, a keto group at carbon-3, one or two double bonds in the A-ring, and an oxygen at carbon-11 ([Figure 1](#)). Different side groups determine the differences in glucocorticoid and mineralocorticoid properties. Reducing the salt-retaining properties was the main effort in the early development of new glucocorticoids. For example, the fluorination of dexamethasone (9 α -fluoro-11 β ,17 α ,21-trihydroxy-16 α -methylpregna-1,4-diene-3,20-dione) highly increases its anti-inflammatory potency, but unfortunately also its salt-retaining properties. The inserted 16 α -methyl group, however, markedly reduces the mineralocorticoid properties, negating the effect of the 9 α -fluoro

Table 1 Characteristics of several common glucocorticoids

Compound	Relative affinity of glucocorticoid receptors (GRs)	Relative potency		Half-life (h)		Comments
		Glucocorticoid	Mineralocorticoid	Plasma	Biological	
Hydrocortisone (cortisol)	1	1	1	1.5	8–12	Endogenous glucocorticoid of humans Drug of choice for replacement
Cortisone	0.01	0.8	0.8	0.5	8–12	Inactive until converted to hydrocortisone
Corticosterone	0.85	0.3–0.5	5–15		8–12	Endogenous glucocorticoid of rodents Not used in humans
Dexamethasone	7.1	25–30	0.5	3–4	36–72	Drug of choice for suppression of ACTH production
Prednisolone	2.2	3.5–4	(0.6–0.8)	2–3	12–36	Inactive until converted to prednisolone
Prednisone	0.05	4	(0.6–0.8)	3–3.5	12–36	
Methylprednisolone	11.9	5	0.5	3	12–36	Pure GR agonist; not used in humans
RU28362	9		0			
Betamethasone	5.4	25–30	0.5	6.6	36–54	GR antagonist
Triamcinolone	1.9	5	0.1	1.5	12–36	
Mifepristone (RU486)	7	–	–	20–30		

atom. Prednisolone (11 β ,17 α ,21-trihydroxypregna-1,4-diene-3,20-dione) is a dehydrogenated derivative of cortisol with an enhanced glucocorticoid and declined mineralocorticoid activity. The most selective ligand RU28362 (11 β ,17 β -dihydroxy-6-methyl-17 α (1-propynyl)-androsta-1,4,6-triene-3-one) has only been used in experimental research up to now. More recently developed selective GR ligands have a non-steroidal structure.

Pharmacokinetics

Glucocorticoids are small, highly lipophilic molecules and therefore are not very soluble in water, but are easily soluble in ethanol, polyethylene glycol (PEG), or dimethyl sulfoxide (DMSO).

Because of their strong lipophilicity, the orally administered glucocorticoids are readily absorbed. Prednisone and cortisone are biologically inactive and are quickly converted to the active compounds prednisolone and cortisol/hydrocortisone, respectively, by dehydrogenation at the carbon-11 hydroxyl group in the liver. In the blood, synthetic glucocorticoids are mainly bound to albumin, a common plasma protein with low affinity. Unlike the endogenous glucocorticoids, most synthetic glucocorticoids are not bound to corticosteroid-binding globulin (CBG). This plasma protein with a high affinity for naturally

occurring corticosteroids, but low presence in plasma, binds prednisolone but not dexamethasone. Plasma half-life is enlarged in the presence of a double bond (prednisolone) or a fluorine atom (dexamethasone) because of decreased rate of metabolism. The plasma half-lives are between 2 and 4 h (Table 1). Because of the modifying action on gene transcription, the duration of the biological effects is much longer than the plasma half-life. It is not the presence of the drug itself but the long-lasting interaction of the ligand receptor complex and the protein synthesis system that accounts for the biological actions. The biological half-life can be up to 72 h (Table 1).

Synthetic glucocorticoids are metabolized in the same way as the endogenous corticosteroids. The main metabolizing tissue is the liver, where steroids are conjugated with glucuronic acid or sulfuric acid. The unchanged parent drug and its metabolites are secreted in the urine, for over 60% (in the case of dexamethasone) and up to 90% (in the case of prednisolone) of the administered dose within 24 h. Changes in plasma half-life of glucocorticoids may result from changed hepatic metabolism due to liver or renal malfunction.

The passage of synthetic glucocorticoids through the blood–brain barrier is hampered because of the presence of multidrug resistance 1a (mdr1a) P-glycoprotein. Recently it was shown that this drug efflux

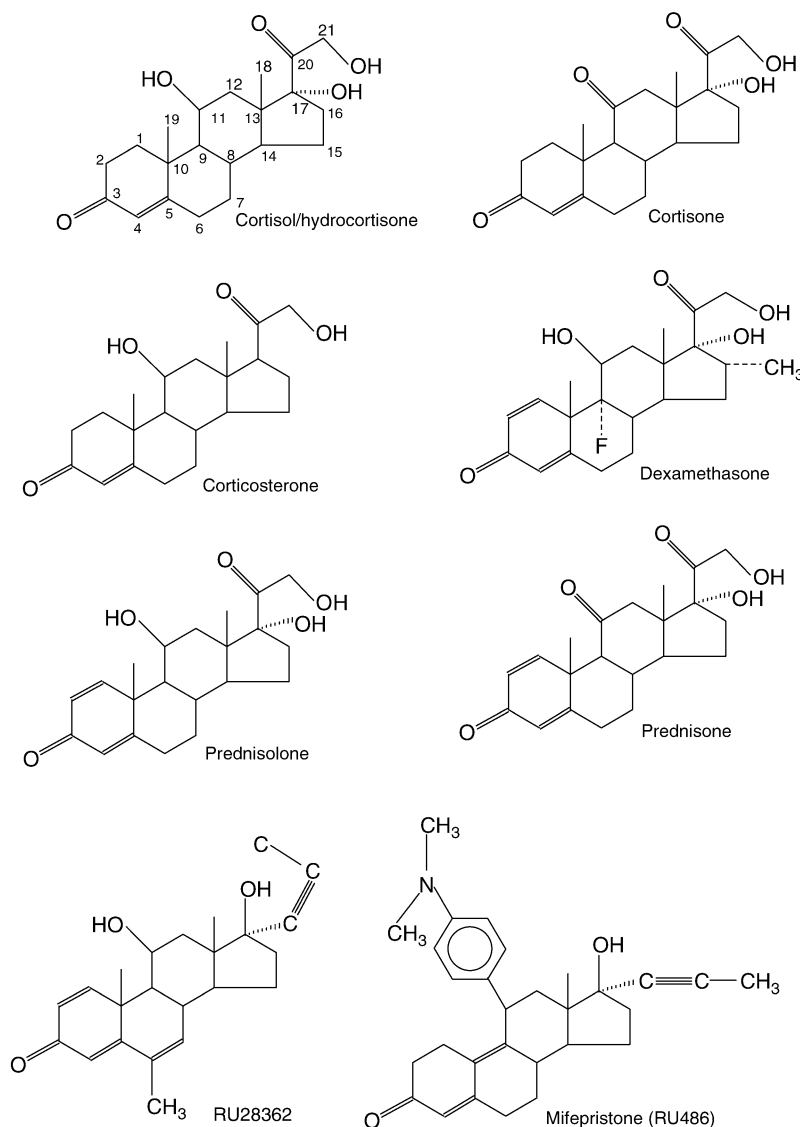


Figure 1 Chemical structures of some common glucocorticoids. Cortisol and corticosterone are endogenous corticosteroids. Cortisol and hydrocortisone have the same chemical structure; the former refers to the naturally occurring glucocorticoid, while the latter refers to the synthetic variant. Cortisone is an inactive endogenous derivative of cortisol. Prednisolone, prednisone, dexamethasone, and RU28362 are synthetic glucocorticoids. Mifepristone is a glucocorticoid antagonist.

pump, which is expressed at the apical membranes of endothelial cells of the blood–brain barrier, hampers the access of moderate amounts of dexamethasone and prednisolone to the brain. Ablation of the gene in the *mdr1a* knockout (–/–) mice resulted in enhanced uptake of dexamethasone and prednisolone.

The most effective receptor antagonist of glucocorticoids is mifepristone (RU486) [17 β -hydroxy-11 β -(4-dimethylaminophenyl)-17 α -(1-propynyl)-estra-4,9-dien-3-one]. This compound, originally developed as a progesterone antagonist, expresses a high affinity for the GR, but binding generally does not result in stimulation of transcription. RU486 antagonizes glucocorticoid effects *in vitro* and inhibits the effect of

dexamethasone on HPA activity *in vivo*. RU486 has weak agonistic properties in some *in vitro* settings, depending on the molecular mechanism of action, and possibly in *in vivo* settings in the absence of glucocorticoid agonists.

Actions of Synthetic Glucocorticoids

The main glucocorticoid actions are mediated by GRs, nuclear receptors present in the cytosol of almost every cell in the body. The corticosteroids bind with high affinity to the GR and in some degree to mineralocorticoid receptors (MRs). MRs are less abundantly present in the body. Most of the synthetic

glucocorticoids have relatively low residual mineralocorticoid potencies (Table 1).

In absence of P-glycoprotein, glucocorticoids easily cross the cell membrane, entering the cell by passive diffusion. Upon binding, the activated steroid receptor complex is translocated from the cytoplasm to the nucleus. The next step is binding to DNA or to transcription factors. Binding to DNA involves binding to glucocorticoid responsive elements (GREs) and subsequent recruitment of coactivator proteins, resulting in activation (or inhibition) of the transcription of target genes (referred to as transactivation). Following binding to transcription factors, the steroid receptor complex effects the transcription of non-GRE containing genes (referred to as transrepression). Transrepression is believed to be the main mechanism of the anti-inflammatory/immunosuppressive actions of glucocorticoids, while transactivation is believed to be responsible for the numerous side effects. This concept has fueled the extensive search for GR ligands with reduced metabolic and HPA suppressant actions while retaining their immunosuppressive/anti-inflammatory actions. The first compounds with dissociated glucocorticoid activities at the molecular level did not show the desired dissociation of beneficial and undesirable actions *in vivo*. A new interesting approach is emerging in the literature that focuses on the development of improved glucocorticoids based on the concept of differential cofactor recruitment by the GR depending on the ligand involved. These new ligands are useful in the understanding of basic GR functioning.

Regarding stress, the main action of glucocorticoids is their suppression of stress-induced HPA axis activity via GRs. GRs are found in all HPA axis structures, hypothalamus, hippocampus, and pituitary. Synthetic glucocorticoids reduce basal plasma levels of ACTH and corticosterone/cortisol during the circadian rise, and they suppress stress-induced rises in ACTH and corticosteroid plasma levels. In spite of the high affinity for synthetic glucocorticoids displayed by GRs *in vitro*, synthetic glucocorticoids (e.g., dexamethasone) do not accumulate in the hippocampus, because of the presence of mdrla P-glycoprotein efflux transporter at the blood-brain barrier. It is therefore likely that the pituitary gland is the primary site of action for the suppressant effects on stress-induced ACTH and corticosterone secretion, particularly at low doses of synthetic glucocorticoids.

Clinical Aspects

Synthetic glucocorticoids are therapeutically used mainly as effective and wide-acting anti-inflammatory

and immunosuppressive drugs. They affect all types of inflammatory reactions independently of the initiating stimulus. They suppress the body's defense reactions, preventing overreaction of the defense mechanisms, which may become damaging if not controlled. Unfortunately, the wide range of actions of glucocorticoids also results in many, sometimes life-threatening side effects on bone, glucose, and HPA axis. More selective nonsteroidal glucocorticoid ligands are currently in development.

Another clinical application is substitution therapy in the case of adrenal insufficiency (Addison's disease), although cortisol and cortisone are more commonly used in that case. Dexamethasone is also used to reduce brain edema. For diagnostic purposes the dexamethasone suppression test (DST, sometimes combined with a CRH challenge) is widely used to test the hypothalamic-pituitary-adrenocortical function. A low dose results in suppression of the hydrocortisone secretion. Failure of suppression implies hypersecretion of corticotropin or of glucocorticoids due to Cushing's syndrome. In psychiatry the DST is used to test patients with major depressive disorders. A large proportion of depressed patients displays enhanced HPA activity. They show impaired suppressibility of cortisol and thus escape from dexamethasone suppression.

Mifepristone (RU486) was recently demonstrated to rapidly relieve psychotic depressed patients from psychotic symptoms. After a brief 7-day treatment with high doses of this antagonist, patients showed a dramatic improvement.

See Also the Following Articles

Corticosteroid Receptors; Corticosteroid-Binding Globulin (Transcortin); Corticosteroids and Stress; Cytokines; Dexamethasone Suppression Test (DST); Glucocorticoids, Overview; Immune Suppression.

Further Reading

- Coghlan, M. J., Elmore, S. W., Kym, P. R. and Kort, M. E. (2003). The pursuit of differentiated ligands for the glucocorticoid receptor. *Current Topics in Medicinal Chemistry* 3(14), 1617–1635.
- De Kloet, E. R., Vreugdenhil, E., Oitzl, M. S. and Joëls, M. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Reviews* 19(3), 269–301.
- Holsboer, F. and Barden, N. (1996). Antidepressants and hypothalamic-pituitary-adrenocortical regulation. *Endocrine Reviews* 17(2), 187–205.
- Miller, D. D., Brueggemeier, R. W. and Dalton, J. T. (2002). Adrenocorticoids. In: Foye, W. O., Lemke, T. L. &

Williams, D. A. (eds.) *Foye's principles of medicinal chemistry* (5th edn.), chap. 28. Philadelphia, PA: Lippincott Williams & Wilkins.

Munck, A., Guyre, P. M. and Holbrook, N. J. (1984). Physiological functions of glucocorticoids in stress and their relation to pharmacological actions. *Endocrine Reviews* 5(1), 25–44.

Schimmer, B. P. and Parker, K. L. (1996). Adrenocorticotrophic hormone; adrenocortical steroids and their synthetic analogs; inhibitors of the synthesis and actions of adrenocortical hormones. In: Hardman, J. G. & Limbird, L. E. (eds.-in-chief) *Goodman and Gilman's the pharmacological basis of therapeutics* (10th edn.). chap. 60. New York, NJ: McGraw Hill.

Systemic Lupus Erythematosus

D J Wallace

Cedars-Sinai/David Geffen School of Medicine,
Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Etiopathogenesis

Clinical Presentation

Management

Prognosis

The Role of Stress in the Immune Response

Clinical Implications of Stress in Lupus

Management of Stress in Systemic Lupus Erythematosus

Glossary

<i>Antinuclear antibodies</i>	Proteins in the blood that react with the nuclei of cells; seen in 96% of patients with SLE, most with other autoimmune diseases, and in 5% of healthy individuals, they are the hallmark of autoimmune disease.
<i>Apoptosis</i>	Programmed cell death; in SLE, the mechanism is altered such that cells that should die do not. Fragments from these cells lead to autoimmune reactions.
<i>B cells</i>	White blood cells that make autoantibodies.
<i>Immune complexes</i>	The complex of an antigen and antibody; in SLE, they are not cleared properly and damage tissue, producing an inflammatory response.
<i>T cells</i>	White blood cells responsible for immunological memory; in SLE, T suppressor cells fail to suppress immune responses.

Lupus erythematosus is an autoimmune disorder that afflicts 1 million people in the United States; 90% of patients are female, and 90% develop it during their reproductive years. Epidemiological surveys suggest that African Americans, certain Asian groupings

(e.g., Chinese), Native Americans, and Hispanics are more susceptible to the disorder. Lupus is classified into many groupings: chronic cutaneous (discoid) lupus (15%), systemic lupus (75%), drug-induced lupus (5%), and lupus overlapping with rheumatoid arthritis, scleroderma, or inflammatory myositis (also called cross-over syndromes, mixed connective tissue disease) (5%).

Etiopathogenesis

A critical dose of 30 thus far identified susceptibility genes causes enough immune response abnormalities to sustain the production of pathogenic autoantibodies and immune complexes. The genes are turned on by infections, increased exposure to ultraviolet light or other environmental toxins, at least 70 known prescription drugs, exposure to increased estrogen or other sex hormones, and severe stress. Antigens are formed that are immunogenic, lead to high quantities of apoptotic cells, produce molecular mimicry (especially microbes), and/or undergo epitope spreading. There are intrinsic abnormalities of B and T lymphocytes. These cells are activated by lower concentrations of antigens than normal, have sustained surface expression of activation markers, and are relatively resistant to apoptosis, which allows them to escape regulation. There is a consequent release of cytokines and chemokines by these cells, which results in inflammation. In this milieu, autoantibody subsets become pathogenic and idiotypic networks promote more antibody production. Immune complexes (consisting of antigen and antibody) are formed that favor capture rather than elimination, and tissue is further damaged. The role that stress plays in this process is discussed here.

Clinical Presentation

Approximately one-half of all systemic lupus erythematosus (SLE) cases are non-organ threatening.

In other words, the heart, lung, kidney, liver, central nervous system, and hematopoietic system (with regards to hemolytic anemia and autoimmune thrombocytopenia) are spared. Patients usually complain of constitutional symptoms of fatigue, fevers, swollen glands, or weight loss along with arthralgias, myalgias, and pleuropericardial (serositis) discomfort. These individuals are usually diagnosed after 1–3 years of symptoms. Organ-threatening disease is a serious subset of lupus and is associated with the need for immune-suppressive approaches. Manifestations include myocarditis, endocarditis, pulmonary hypertension or hemorrhage, mesenteric vasculitis, interstitial lung disease, autoimmune hepatitis, thrombotic thrombocytopenic purpura, retinitis, vestibulitis with hearing loss, proteinuria, seizures, psychosis, and edema. The presentation is usually quite dramatic and diagnostic delays are uncommon. The 1997 American College of Rheumatology (ACR) revised criteria for SLE are listed in [Table 1](#).

Laboratory testing, along with appropriate electrical studies and imaging, is very useful in diagnosing and following the disorder. Initial testing includes a complete blood count, comprehensive metabolic profile, muscle enzymes, urinalysis, chest x-ray, and electrocardiogram. Serologic parameters encompass an antinuclear antibody (ANA) with a reflex panel for a variety of autoantibodies, obtaining certain protein and acute-phase reactant determinants (which screen for inflammation) such as anti-DNA, anti-Smith antibodies (anti-Sm), Sjögren's syndrome antigens (anti-SSA and anti-SSB), complement components, sedimentation rate, C reactive protein, and a serum protein electrophoretic screening. Occasionally, additional determinations such as two-dimensional echocardiography, brain imaging, lumbar puncture, or electromyograms can be useful. One-third of lupus patients have a procoagulant set of autoantibodies

known as antiphospholipid antibodies that can predispose patients to a thromboembolic event.

Management

Non-organ-threatening disease is managed by employing physical measures such as avoidance of the sun, pacing activities to minimize fatigue, moist heat for painful joints, a general conditioning program, work-station modifications to enhance job performance, and a low fat, well-balanced diet. Stress reduction is also critical. Individuals respond best to a combination of sunscreens applied to sun-exposed areas, topical corticosteroids, and nonsteroidal anti-inflammatories. Antimalarials (e.g., hydroxychloroquine) have anti-inflammatory disease-modifying properties and can decrease the risk of dissemination. Low doses of corticosteroids may be used until antimalarials are effective several months later, and immune-suppressive regimens are rarely indicated.

Patients with organ-threatening disease are usually managed with a combination of corticosteroids (moderate to high dose, usually starting at 1 mg/kg/day prednisone or its equivalent). Steroid-sparing agents such as methotrexate and leflunomide can be employed for milder cases, but serious presentations respond best to azathioprine, mycophenolate mofetil, or cyclophosphamide. A variety of biological regimens and stem cell options are currently being explored.

Prognosis

Patients with non-organ-threatening disease who do not have antiphospholipid antibodies have a normal life expectancy. Organ involvement increases morbidity and mortality. Approximately 50% have 20-year life expectancies. The most common causes of demise

Table 1 Revised American College of Rheumatology criteria for SLE (1997)^a

Cutaneous criteria	Butterfly rash (lupus rash over the cheeks and nose) Chronic cutaneous (discoid) rash (a thick, plaque-like, adherent rash that scars) Sun sensitivity (rash after being exposed to ultraviolet A or B light) Oral ulcerations (recurrent sores in the mouth or nose)
Systemic criteria	Arthritis (inflammation of two peripheral joints with tenderness, swelling, or fluid) Serositis (inflammation of the lining of the pleura or pericardium) Neurological disorder (seizures or psychosis with no other explanation)
Laboratory criteria	Blood abnormalities (hemolytic anemia, leucopenia, thrombocytopenia) Immunological disorder (antiphospholipid antibodies, anti-dsDNA, anti-Sm, lupus anticoagulant or false-positive syphilis serology) Positive antinuclear antibody blood test

^aA person is said to have SLE if four of the 11 criteria are present at any time. Anti-dsDNA, anti-double-stranded DNA antibodies; anti-Sm, anti-Smith antibodies.

include infection, complications from accelerated atherogenesis and immune suppression, and unresponsive inflammation. For those with ongoing chronic disease, quality of life can be markedly diminished due to medication (e.g., corticosteroids), swollen joints and inflammation limiting activities of daily living, chronic pain, specific diets, sun avoidance, decreased ability to think clearly, edema from nephrosis, and cold avoidance due to Raynaud's disease, among other factors. This is associated with increased stress and decreased ability to respond to it.

The Role of Stress in the Immune Response

The role of stress in SLE has been looked at in a small number of animal models and human studies. Stress is a complex concept that involves interactions of the hypothalamic-pituitary-adrenal (HPA) axis in combination with cytokines, hormones, and the sympathetic nervous system. Patients with SLE generally manifest a blunted HPA axis response and relative glucocorticoid resistance. The latter has been documented by the larger doses needed to achieve a therapeutic effect in managing lupus, smaller numbers of glucocorticoid receptors on cells, and defectively functioning or mutated glucocorticoid receptors. Chronic inflammation is associated with hypercortisolism. In addition, prolactin is an established acute-phase reactant. Levels are increased in up to 20% of lupus patients, and this hormone can also increase interferon and interleukin (IL-)2 production.

Recently, there has been considerable interest in the role of cytokines in SLE. Lupus mice have diminished IL-1 receptor expression, and lupus patients have increased serum IL-6 levels that correlate with flares. Cerebrospinal fluid levels of interferon α are elevated in lupus psychosis. The release of these glycoproteins is a well-documented stress response. The release of cortisol is also associated with increased sympathetic

nervous system activity. The autonomic nervous system is dysfunctional in lupus, as well established by Raynaud's phenomenon and vasomotor instability (e.g., palmar erythema). This has been correlated in the disorder with the increased release of neuropeptide Y.

In summary, lupus patients have a sluggish corticotropin releasing hormone (CRH) response, relative glucocorticoid resistance, and hypercortisolism suggestive of an adaptive response to chronic inflammation. Cytokine and hormonal imbalances along with dysautonomia serve to perpetuate this.

Clinical Implications of Stress in Lupus

Lupus is associated with substantial emotional and societal costs. It has been suggested that up to half of all lupus patients have psychiatric comorbidities such as dysthymia and depression. This section reviews whether stress can induce or exacerbate lupus and how it can be managed (Table 2).

Various studies have shown that emotional stress and physical trauma can affect the immune system by causing decreased lymphocyte mitogenic responsiveness, lymphocyte cytotoxicity, increased natural killer cell activity, skin homograft rejection, graft-versus-host response delayed hypersensitivity. Could the impairment in T-cell immune functions be related to a clinical flare mediated by B-cell hyperreactivity? Acute stress in SLE has been purported to correlate with increased urine neopterin levels, IL-4 levels, decreased natural killer cell response, and increased β_2 -adrenoreceptors on mononuclear cells without a clinical disease flare.

Can Emotional Stress Induce Lupus?

As early as 1955, crises in interpersonal relationships were associated with the onset of disease. Up to one-half of all lupus patients interviewed believe that stress is associated with its initial presentation. Unfortunately, no evidence-based study has documented

Table 2 Stress in SLE – summary^a

Sluggish CRH response, relative glucocorticoid resistance, and hypercortisolism suggestive of an adaptive response to chronic inflammation

Cytokine abnormalities, hormonal dysfunction, and dysautonomia perpetuate this response

Physical trauma can produce cutaneous lupus; no evidence-based studies have shown that physical or emotional trauma can induce SLE

Evidence-based studies have documented that physical trauma and emotional stress can aggravate SLE, although concomitant fibromyalgia is confounding factor and strict guidelines of abnormalities occurring within 60 days of a specific incident have been published for litigation purposes

Stress-reduction measures include antidepressants, anxiolytic agents, biofeedback, cognitive-behavioral therapy, and counseling.

^aCRH, corticotropin releasing hormone.

this. Patients with SLE possess autoantibodies several years prior to diagnosis, according to a serum repository kept by the U.S. Armed Forces. Many patients manifest an undifferentiated connective tissue disease (UCTD) for years prior to evolving into frank SLE. (UCTD is defined as inflammatory arthritis or Raynaud's with a positive ANA or rheumatoid factor in an individual who fails to meet the ACR criteria for lupus or rheumatoid arthritis. One-third of patients evolve SLE or rheumatoid arthritis, one-third remain UCTD, and one-third have spontaneous resolution of their symptoms.) The mechanisms resulting in this transformation are not known.

Can Emotional Stress Exacerbate Lupus?

At Harvard University, of 45 flares in 160 patients over a 40-year period, 41 patients associated these exacerbations to emotional stressors. Lupus patients may have significantly reduced cell mobilization to psychological stress compared to controls. Several surveys have subsequently shown that stress can elevate antinuclear antibody levels and be associated with clinical flares. Life-event stress predicted adverse changes in functional disability over a 6-month period in one study and self-reported daily stressors were associated with increases in clinical activity in three others. These studies did not take into account the impact of quality of life related to steroid therapy for lupus or the confounding influences of fibromyalgia, which is present in one-third of all lupus patients.

Can Physical Trauma Induce or Exacerbate Lupus?

Cutaneous skin lesions seen in lupus may be found after physical trauma, lacerations, and scars. Sometimes these scars evolve into skin cancers. Two studies, on the other hand, suggested that major physical trauma such as an earthquake is not associated with disease flares but may actually improve the process. The physical trauma of a motor vehicle accident is also associated with myofascial injuries, and no published study has associated this with a lupus flare. Guidelines for documenting a lupus flare after trauma have been published and include the following findings within 60 days of the event: (1) increased symptoms and physical signs of lupus documented in a progress note, (2) worsening of inflammatory laboratory parameters (e.g., sedimentation rate or anti-DNA), (3) increases in lupus medication or new anti-inflammatory medication ordered by the practitioner, and (4) introduction of psychotropic medication prescribed by the treating physician or referral to a psychologist documenting increased stressors. The patients medical record over the prior year can be used as a historical control.

Management of Stress in Systemic Lupus Erythematosus

Stress-reduction measures include managing the inflammation associated with disease activity, counseling, pain management, cognitive-behavioral therapy, and anxiolytics and antidepressant medications. Counseling usually emphasizes improved coping, cognitive therapy, and biofeedback; improving resilience; managing stress, and regulating mood. Controlled studies have validated this approach.

Further Reading

- Adams, S. G., Jr, Dammers, P. M., Sala, T. L., et al. (1994). Stress, depression and anxiety predict average symptom severity and daily symptom fluctuation in systemic lupus erythematosus. *Journal of Behavioral Medicine* 17, 459-477.
- De Bellis, M. D., Burke, L. and Trickett, P. K. (1996). Antinuclear antibodies and thyroid function in sexually abused girls. *Journal of Traumatic Stress* 9, 369-378.
- Dobkin, P. L., Fortin, P. R., Joseph, L., et al. (1998). Psychosocial contributors to mental and physical health in patients with systemic lupus erythematosus. *Arthritis Care Research* 11, 23-31.
- Eskreis, B. D., Eng, A. M. and Furey, N. L. (1988). Surgical excision of trauma-induced verrucous lupus erythematosus. *Journal of Dermatologic Surgery and Oncology* 14, 1296-1299.
- Gorby, H. E. and Sternberg, E. M. (2007). Neuroendocrine immune interactions: principles and relevance to systemic lupus erythematosus. In: Wallace, D. J. & Hahn, B. H. (eds.) *Dubois' lupus erythematosus*, pp. 286-300 (7th edn.). Philadelphia: Lippincott Williams & Wilkins.
- Greco, C. M., Rudy, T. E. and Manzi, S. (2004). Effects of a stress-reduction program on psychological function, pain, and physical function of systemic lupus erythematosus patients: a randomized controlled trial. *Arthritis & Rheumatism* 51, 625-634.
- Jacobs, R., Pawlak, C. R., Mikeska, E., et al. (2001). Systemic lupus erythematosus and rheumatoid arthritis patients differ from healthy controls in their cytokine pattern after stress exposure. *Rheumatology* 40, 868-875.
- King-Smith, D. (1926). External irritation as a factor in the causation of lupus erythematosus discoides. *Archives of Dermatology* 14, 547-549.
- McLeary, A. R., Meyer, E. and Weitzman, E. L. (1955). Observations on the role of depression in some patients with disseminated lupus erythematosus. *Psychosomatic Medicine* 17, 311-321.
- Pawlak, C. R., Witte, T., Heiken, H., et al. (2003). Flares in patients with systemic lupus erythematosus are associated with daily psychologic stress. *Psychotherapy and Psychosomatics* 72, 159-165.
- Peralta-Ramirez, M. I., Jimenez-Alonso, J., Godoy-Garcia, J. F., et al. (2004). The effects of daily stress and stressful life events on the clinical symptomatology of patients

- with lupus erythematosus. *Psychosomatic Medicine* 66, 788–794.
- Ropes, M. W. (1976). *Systemic lupus erythematosus*. Cambridge, MA: Harvard University Press.
- Wallace, D. J. (1994). Does stress or trauma aggravate rheumatic disease? *Baillière's Clinical Rheumatology* 8, 149–159.
- Wallace, D. J. and Hahn, B. H. (eds.) (2007). *Dubois lupus erythematosus* (7th edn.). Philadelphia: Lippincott, Williams & Wilkins.
- Wallace, D. J. and Metzger, A. L. (1994). Can an earthquake cause flares of rheumatoid arthritis or lupus nephritis? *Arthritis & Rheumatism* 37, 1826–1828.

T

T Cells See: Lymphocytes.

Teaching and Stress

E R Greenglass

York University, North York, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by E R Greenglass, volume 3, pp 571–574, © 2000, Elsevier Inc.

Stressors in Teaching
Teacher Burnout
Burnout as a Process

Glossary

<i>Burnout</i>	A state of physical, emotional, and mental exhaustion that results from long-term involvement with people in situations that are emotionally demanding.
<i>Emotional exhaustion</i>	Feelings of being emotionally overextended and drained by others.
<i>Depersonalization</i>	A callous response toward people who are the recipient of one's services.
<i>Cynicism</i>	An indifferent or distant attitude towards one's work.
<i>Lack of personal accomplishment</i>	A decline in one's feelings of competence and successful achievement in one's work with people.
<i>Social support</i>	The exchange of information leading a person to believe that he or she is cared for. It can also involve provision of information, tangible help, or emotional help.
<i>Role conflict</i>	The interference of the efforts to fulfill the demands associated with one role with the efforts to fulfill the demands associated with other roles.

Stressors in Teaching

It is well documented that teaching is a highly stressful occupation. Several writers have noted recurring demands in the teaching environment that lead to significant levels of stress and burnout in teachers. There are several sources of stress for teachers. One of the most frequently acknowledged stressors for teachers centers around students. Problems with students include disruptive behavior, such as verbal and physical abuse; emotional demands of students; their special needs; and their heterogeneous in abilities – all of which tax the time and energy of an already busy teacher. Another major stressor for teachers involves the conflicting demands made by supervisors, colleagues, students, and the students' parents. A third set of stressors in teaching relates to teachers' workload itself, mainly their work overload. Quantitative work overload involves too many demands on teachers with too little time in which to meet these demands adequately. Often it is accompanied by feelings of being rushed and impatience. The importance of work overload as a stressor is underlined by the fact that teachers themselves in countries all over the world have consistently cited work overload as a major stressor in their job. Important dimensions associated with work overload are excessive paperwork, lack of preparation time for class, making reports, and submitting grades. Having to meet constant deadlines throughout the year, including those associated with grading, tests, exams, and written reports by students is cited as another major stressor by teachers. Oversized classes with either inadequate or no teaching assistance also function as stressors in teachers.

Additional stressors associated with teaching are found at the administrative level. A frequently cited administrative stressor is the exclusion by the administration of teachers from decision making that bears directly on their workload and the quality of their work life. Research has shown that when teachers do not have the opportunity to participate in the school's decision-making processes, their morale, motivation, self-esteem, and job satisfaction are more likely to decline. Lack of recognition of teachers' professional accomplishments by supervisors and school administrators alike contributes to an increase in stress levels in teachers. When teachers experience high role conflict in their jobs, role ambiguity, and low autonomy, they are more likely to experience stress. At the administrative level, additional stressors include inflexible employment policies on staffing and leave, and poor communication with teachers. Thus, research has demonstrated that teaching is associated with a wide range of stressors, including those intrinsic to teaching, those involving the relationship between teacher and student, and those related to the organizational structure of the school including the school administration.

Teacher Burnout

Considerable research worldwide has examined psychological consequences of stress in teachers. Over time, the cumulative effects of stressors lead to a decrease in job satisfaction, reduced commitment to teaching, absenteeism, and job turnover. At the individual level, greater stress is associated with an increase in somatic complaints and with physical and psychological illness. Research has shown that the experience of stress over time often leads to burnout in teachers. Although burnout is linked to occupational stress, burnout goes beyond stress by emphasizing the accumulation of the effects of stressors within a context that emphasizes total life and environmental experiences. Research findings indicate that teacher burnout in several Western industrialized countries has increased and affects a significant and growing proportion of teachers. Burnout is often defined as a state of physical, emotional, and mental exhaustion that results from long-term involvement with people in situations that are emotionally demanding. Because individuals who enter teaching are often presented with work pressures in addition to teaching, they often find themselves in situations that are heavily demanding. Research shows that human service professionals, including teachers, are particularly vulnerable to burnout. Freudenberger's original idea of burnout was that workers would

find themselves under increasing pressure to help others, would demand more of themselves than they were able to give, and would ultimately exhaust themselves.

Measurement of Burnout

Given the significance of burnout for the individual, it is necessary to have a standardized instrument to measure it. The Maslach Burnout Inventory (MBI), developed by Christine Maslach and Susan Jackson in 1986, is currently the research instrument used most widely to measure burnout. The MBI has three subscales that measure three different aspects of burnout. These are emotional exhaustion, depersonalization, and reduced personal accomplishment. Emotional exhaustion is defined as feelings of being emotionally overextended and drained by others. Depersonalization is a callous response toward people who are the recipient of one's services.

Lack of personal accomplishment is a decline in one's feelings of competence and successful achievement in one's work with people. The need for developing a scale that measures burnout in occupational groups other than human service professionals prompted the development of the Maslach Burnout Inventory-General Survey (MBI-GS) in 1996. Thus, the MBI-GS defines burnout as a crisis in one's relationship with work, not necessarily as a crisis in one's relationship with people. In the MBI-GS, the depersonalization scale is replaced with a scale measuring cynicism, an indifferent or distant attitude toward one's work. Research findings show that teachers are vulnerable to burnout in all three areas. For example, teachers show emotional exhaustion when they perceive they are unable to give of themselves to students as perhaps they did earlier in their careers. Depersonalization and cynicism occur when teachers develop negative, cynical, and sometimes callous attitudes toward students, parents, and/or colleagues. Feelings of diminished personal accomplishment occur when teachers perceive themselves as ineffective in helping students to learn and in fulfilling other school responsibilities. For individuals who experience burnout, one of the first reactions is distancing themselves from the job. Burnout has also been linked to job turnover, absenteeism, poor performance, and various types of individual distress symptoms including depression, anxiety, and somatization. Overall, burned-out teachers are likely to be less sympathetic toward students, have a lower tolerance for classroom disruption, be less apt to prepare adequately for class, and feel less committed to their work.

Gender Differences and Burnout

One question that has been raised by many researchers regarding teachers is: who is more burned out, women or men? Research findings from a variety of studies conducted in several different countries show that men have significantly greater levels of depersonalization than women. Why should men be more prone to depersonalization, an attitude characterized by callousness and being impersonal toward one's students? One explanation is found in accepted norms associated with the masculine gender role, which emphasizes strength, independence, separation, and invulnerability. In this context, depersonalization may be regarded as a reflection of men's repressed emotionality. Described in this way, depersonalization may be seen as an ineffective coping form that allows men to continue their work with people yet remain untouched in any significant way by others' suffering. Moreover, it is through the distance they have imposed between their students and themselves that men can remain uninvolved, thus appearing callous. It is also possible that depersonalization is a result of an inability on the part of men to cope with work strain. Because it is socially unacceptable for men to openly express vulnerability, men have fewer options for emotional expression. Higher depersonalization in men has been associated with lack of collegiality, higher absenteeism, greater medication use, and a lower quality of life.

Another explanation for women's lower depersonalization compared to men focuses on women's greater ability to cope with interpersonal stress. Women have an edge over men in dealing with the emotional strain of people work because the feminine gender role emphasizes caring, nurturance, and concern for others. As a result, women may be less likely than men to respond to troubled individuals in an impersonal and callous manner. Some researchers have found that women have higher scores than men on emotional exhaustion. Women may experience greater emotional exhaustion because, compared to men, their total workload (including paid and unpaid work) is higher. Although the number of men doing housework and childcare has increased in recent years, according to research findings women still do most of the work in the home. Further research has shown that women's increased workload interferes with their ability to wind down from a stressful day, with resultant negative effects on their health. Research also shows that women experience greater role conflict (between work and family roles) than their male counterparts. Related findings indicate that there are different predictors of burnout in men and women. In women role conflict between home

and work roles is a significant predictor of burnout, whereas in men the predictors of burnout tend to be confined to work. A structural explanation of gender-related differences in burnout highlights the association of role conflict with stress and burnout in women. It is widely recognized that employed women, including teachers, typically maintain major responsibility for the home and the family, and thus more often than men, bear the burden of role conflict.

Burnout, Stress, Social Support and Coping

Over the years, researchers have investigated several ways that burnout in teachers can be reduced, including individual and social forms. Many of the social forms involve interventions that have led to lower burnout levels in teachers by changing the work environment or reducing burnout at its source (i.e., in the workplace). Specific examples include teacher involvement and participation in the decision making undertaken by the school; facilitating the development and provision of social support; and improving supervision through the clarification of work goals, which has resulted in less role ambiguity and lower role conflict in teachers. Burnout can be reduced by establishing clear lines of authority and responsibility in helping to reduce ambiguity and conflict. With teachers, it may be valuable to solicit their input in areas that affect their work life in order to gain greater involvement and commitment. The development of social support among teachers can be achieved by providing adequate time and location interaction. Encouraging the development of mentoring relationships between older and younger teachers can also be useful in increasing commitment among teachers. Considerable research has demonstrated the effectiveness of social support in lowering burnout and stress levels in teachers. Social support is the exchange of information leading a person to believe that he or she is cared for. It can also involve the provision of information and tangible, practical, and/or emotional help. Research suggests that social support may be effective in reducing burnout directly or as a buffering agent. When there is a direct effect, social support is positively related to physical and psychological health, regardless of the presence or absence of work stressors. Social support may also moderate the impact of stress on burnout so as to assist people with high stress to cope better. The buffering argument suggests that stress may affect some people adversely but that those with social support resources are relatively resistant to the deleterious effects of stressful events.

Several studies report that teachers with high burnout perceive less social support in their school

environment than those with lower burnout levels. Teachers' perceptions of stress are reduced by perceived support, particularly from school staff, including colleagues. Further research findings indicate that teachers who work in an environment that they perceive as being supportive are less likely to experience high levels of stress and burnout. Research by Greenglass and co-workers indicated that support from a teacher's peers or coworkers are the most important buffers of burnout. As suggested by Pines, social support in the form of technical advice such as helping teachers find more successful and innovative ways of working with students, may be a particularly effective intervention for burned-out teachers. Coping may also lead to lower levels of burnout. For example, in a 2002 study of burnout in German teachers, the relationship between proactive coping and burnout was examined. In this study, 316 German teachers were surveyed and job burnout was defined three-dimensionally in terms of emotional exhaustion, depersonalization, and lack of personal accomplishment. Proactive coping is multidimensional and forward-looking. It combines autonomous goal setting with self-regulatory goal attainment cognitions. Proactive coping integrates processes of personal quality of life management with those of self-regulatory goal attainment. The results showed a significant pattern of decreasing burnout with increasing levels of proactive coping. High-proactive-coping teachers reported less emotional exhaustion, less depersonalization, and more personal accomplishment than low-proactive-coping teachers.

Burnout as a Process

Increasingly, research has focused on the examination of burnout as a process that develops over time. As such, researchers have been interested in discovering the paths through which burnout develops. Some researchers have demonstrated the existence of two separate paths to burnout: a cognitive path, manifested in personal and professional feelings of lack of accomplishment, and an emotional path, reflected in a sense of overload and emotional exhaustion. It is generally held that the process begins to develop with external stressors, such as disruptive students, excessive paper work, and conflicting demands, which lead to emotional exhaustion. Emotional exhaustion is considered the aspect of burnout that is most responsive to the nature and intensity of work stressors. Some researchers refer to emotional exhaustion as the prototype of stress. In general, various job conditions are more strongly related to emotional exhaustion than to the other two subscales,

depersonalization and lack of personal accomplishment. Emotional exhaustion is predicted mainly by occupational stressors, such as work overload, inadequate use of skills, and interpersonal conflicts. In addition, emotional exhaustion is considered the affective component of burnout that leads to depersonalization or cynicism. Through depersonalization, individual teachers attempt to cease the depletion of their emotional energy by treating their students as objects rather than people. Emotional exhaustion can also lead to reduced feelings of personal accomplishment in the teaching role, which may be seen as indicative of an impoverishment of people's perceptions of themselves in their work role, particularly as teachers. Findings suggest that depersonalization contributes to lower feelings of accomplishment as well. Because effective teaching is contingent on teachers' communicating with students on a personal level, teachers may be particularly vulnerable to low feelings of accomplishment when perceiving that they are treating their students impersonally, as in the process of depersonalization. Other findings point to the consistency of each of the burnout components over time. Emotional exhaustion, depersonalization, and lack of personal accomplishment predict significantly the scores on the same scales 1 year later, according to Greenglass and co-workers. Researchers agree that burnout levels are fairly stable over time.

One of the advantages of viewing burnout as a process over time is that it allows the investigation of the antecedents of burnout, particularly those associated with the school, the teachers' interpersonal relationships, and their workload. Being able to specify the antecedents of burnout in teachers has both theoretical and applied value. By specifying variables that contribute to burnout, theoretical knowledge of the process of the development of stress and burnout is advanced. On a practical level, it is valuable to the school's administration to be able to specify just what causes burnout. Examining the development of stress and burnout over time offers the possibilities of changing the processes involved and grappling with the complex causal relationships. This presumably can lead to the development of better and more effective intervention techniques, which can improve teacher morale and prevent the harmful consequences of further burnout.

Further Reading

Burke, R. J. and Greenglass, E. (1995). A longitudinal study of psychological burnout in teachers. *Human Relations* 48, 187–202.

- Freudenberger, H. (1974). Staff burnout. *Journal of Social Issues* 30, 159–164.
- Greenglass, E. R. (1991). Burnout and gender: theoretical and organizational implications. *Canadian Psychology* 32, 562–570.
- Greenglass, E. R. (2002). Proactive coping. In: Frydenberg, E. (ed.) *Beyond coping: meeting goals, vision, and challenges*, pp. 37–62. London: Oxford University Press.
- Greenglass, E. R., Burke, R. J. and Konarski, R. (1997). The impact of social support on the development of burnout in teachers: Examination of a model. *Work and Stress* 11, 267–278.
- Greenglass, E. R., Fiksenbaum, L. and Burke, R. J. (1996). Components of social support, buffering effects and burnout: implications for psychological functioning. *Anxiety, Stress and Coping* 9, 185–197.
- Maslach, C. and Jackson, S. E. (1986). *Maslach Burnout Inventory manual* (2nd edn.). Palo Alto, CA: Consulting Psychologists Press.
- O’Laughlin, E. M. and Bischoff, L. G. (2005). Balancing parenthood and academia: work/family stress as influenced by gender and tenure status. *Journal of Family Issues* 26, 79–106.
- Schaufeli, W. B., Leiter, M. P., Maslach, C., et al. (1996). The Maslach Burnout Inventory (MBI-GS). In: Maslach, C., Jackson, S. E. & Leiter, M. P. (eds.) *MBI manual* (3rd edn.). Palo Alto, CA: Consulting Psychologists Press.
- Schwarzer, R. and Taubert, S. (2002). Tenacious goal pursuits and striving toward personal growth: proactive coping. In: Frydenberg, E. (ed.) *Beyond coping: meeting goals, visions and challenges*, pp. 19–35. London: Oxford University Press.

Relevant Websites

- Greenglass, E. R., Schwarzer, R. and Taubert, S. (1999). The Proactive Coping Inventory (PCI): a multidimensional research instrument. <http://www.psych.yorku.ca>

Temperature Effects

J Roth

University of Giessen, Giessen, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J Roth, volume 3, pp 575–576, © 2000, Elsevier Inc.

Glossary

<i>Heterothermy/ poikilothermy</i>	The pattern of temperature regulation in a species in which the variation in core temperature, either nychthemally or seasonally, exceeds that which defines homeothermy.
<i>Homeothermy</i>	The pattern of temperature regulation of a given species in which the cyclic variation in core temperature, either nychthemally or seasonally, is maintained within arbitrarily defined limits despite much larger variations in ambient temperature.
Q_{10}	The ratio of the rate of a physiological process at a particular temperature to the rate at a temperature 10°C lower, when the logarithm of the rate is an approximately linear function of temperature.

Temperature change has striking effects on many physiological processes. All chemical reactions in an organism, including the processes of metabolism, are temperature-dependent. Within certain limits, an increase of temperature accelerates most physiological processes, such as the rate of oxygen consumption, which can be regarded as an expression for the overall metabolic activity of an organism. The rate of a physiological process usually increases with temperature in a regular manner. In oxygen consumption, for example, a rise of 10°C in temperature may cause a twofold or even threefold increase. The increase in the rate of a metabolic process caused by a 10°C increase in temperature is called the Q_{10} of this process. **Figure 1** shows an example for a given physiological process with a Q_{10} of 2. For a 10°C interval, the Q_{10} is thus defined as the ratio of the rate of the metabolic process (R_1) at a particular temperature ($T + 10$) to the rate (R_2) at a temperature 10°C lower (T), when the logarithm of the rate is an approximately linear function of temperature.

$$Q_{10} = R_1(T+10)/R_2(T)$$

According to van’t Hoff’s rule, the Q_{10} for thermochemical reactions is between 2 and 3, and so are the Q_{10} values for metabolic processes in an organism.

Animals can be divided into two groups: poikilotherms and homeotherms. The term poikilotherm

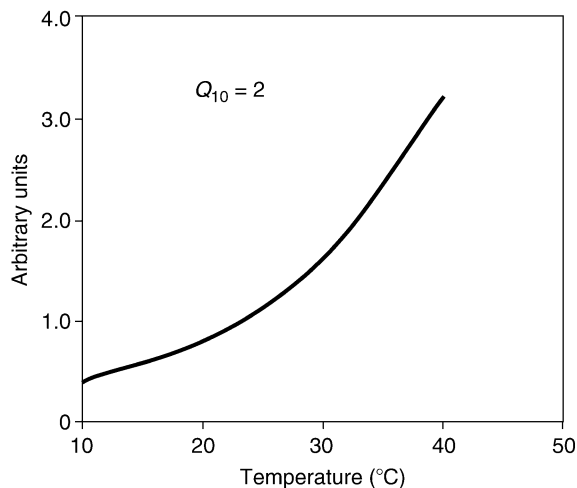


Figure 1 Temperature dependence of a given metabolic process. For this process, a temperature increase of 10°C results in the doubling of the rate of the process ($Q_{10} = 2$).

means that an animal shows a large variability of body temperature as a function of ambient temperature. In homeothermic animals (birds and mammals, including humans), the variations in body temperature are maintained within narrow limits despite much larger fluctuations in ambient temperature. This means that in poikilotherms the rate of metabolic processes is strongly influenced by temperature changes according to van't Hoff's rule. Van't Hoff's rule applies to homeotherms in the same way, but it is masked by other effects. The direct effects of changing temperatures on the metabolic processes within homeotherms are prevented by temperature regulation, the prerequisite to preventing large fluctuations of body temperature. As a consequence, the vital processes of poikilotherms depend entirely on external temperature conditions, whereas homeotherms live at practically the same level of intensity during the entire year and in all climatic regions. The geographic distribution of homeotherms is thus not as limited as the distribution of poikilotherms. Note that, in reality, the distinction between homeothermy and poikilothermy is not as clear-cut as the definitions of these terms suggest. On the one hand, many poikilothermic animals show rudimentary signs of temperature regulation, especially by behavioral means and, to a certain extent, even by metabolic responses. On the other hand, a number of homeothermic animals sometimes permit a profound fluctuation of body temperature (e.g., hibernators), with strong temporary restrictions of their vital processes.

In spite of their ability to regulate body temperature, large fluctuations in ambient temperature represent stressful conditions for homeotherms, especially when

people or animals are exposed to temperatures that are out of their thermoneutral zone. For humans, the thermoneutral zone, or zone of thermal indifference, is defined as the range of ambient temperatures associated with specific water vapor pressure, air velocity, and radiant exchange within which 80% of active people do not complain of the thermal environment. A number of behavioral, physiological (functional), and morphological strategies with regard to the thermoregulatory system have developed to reduce the strain that is imposed on homeotherms by pronounced external temperature changes. These strategies can be summarized as thermal adaptations. Morphological adaptations, such as changes in body shape and insulation, need from months to years to develop unless they are genetically fixed and appear seasonally. In general, they are preceded by functional (physiological) modifications, including changes in the capacity of thermoregulatory effector systems and regulatory characteristics that need much less time to develop. One of the best-investigated changes in the capacity of thermoregulatory effector systems is the development of nonshivering thermogenesis in the brown adipose tissue of small mammals during cold adaptation. An increase in maximal sweat rate, and thereby in the capacity to dissipate heat, during adaptation to heat has also been described. The development of increased capacities of heat production and heat dissipation takes place within several weeks. Short-term physiological adaptations to stressful external temperatures consist of changes in thermoregulatory characteristics, which can be defined as deviations in the threshold and the gaining of the thermoregulatory responses. For example, a short exposure to a cold environment evokes not only thermoregulatory reflexes and emergency reactions but also a temporary shift of the shivering threshold to a lower mean body temperature. Continuous confrontation with a cold environment for several weeks results in an even more pronounced shift of the threshold for heat production to a lower body temperature. Similar adaptive changes of thermoregulatory characteristics occur in response to short- or long-term exposure to heat.

See Also the Following Article

Critical Thermal Limits.

Further Reading

- Cossins, A. R. and Bowler, K. (1987). *Temperature biology of animals*. London: Chapman and Hall.
- Fregly, M. J. and Blatteis, C. M. (1996). *Handbook of physiology*. Vol. 1: *Environmental physiology*. New York: Oxford University Press.
- Precht, H., Christophersen, J., Hensel, H. and Larcher, W. (1973). *Temperature and life*. Berlin: Springer.

Terrorism

P Bell

South and East Belfast Health and Social Services
Trust, Belfast, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
P Bell, volume 3, pp 577–580, © 2000, Elsevier Inc.

Introduction

Effects of Terrorism on the Terrorist

Effects of Terrorism on Victims

The Effect of Terrorism on Society

Glossary

<i>Anomie</i>	The sense of isolation that comes from lack of social integration and cohesion.
<i>Posttraumatic stress disorder (PTSD)</i>	A psychological injury characterized by reexperiencing trauma, hyperarousal, and withdrawal.
<i>The Troubles</i>	A term used to describe the 30 years of civil unrest and paramilitary campaigns, bordering on civil war, in Ireland.

Introduction

The *Oxford English Dictionary* defines terrorism as the use of violence and intimidation for political ends. This definition could apply to any kind of warfare. However, the term terrorism usually refers to unconventional tactics of warfare used by a less powerful group against a more powerful adversary. In this way, the Roman legions encountered terrorism in the woods of Germany. As a tactic, it has become more successful with the development of technology, so that a small number of terrorists with modern weapons can do a great deal more damage than many men wielding swords and halberds. The extent to which a small number of highly motivated men can inflict catastrophic pain on a mighty enemy was demonstrated by the events of September 11. Terrorism as a tactic is increasingly attractive to groups who lack political or military power.

One man's terrorist is another man's freedom fighter, and terms such as commando, guerrilla, partisan, and paramilitary help to confuse the picture further. At the same time, many former terrorists have been accepted into the political mainstream, and the progression from terrorist leader to statesman seems to be the rule rather than the exception

(e.g., Nelson Mandela). Other terrorist organizations become increasingly involved in organized crime, and the pattern of terrorism is moving away from organized political groups to small, very extreme groups and even to individuals (e.g., the Uni bomber and the cell structure of the IRA and various Islamic fundamentalist groups). The field of terrorism, then, is a complex one, with the spectrum including a wide variety of separate phenomena.

Effects of Terrorism on the Terrorist

There has been little research from Northern Ireland examining the psychological profiles of active members of terrorist organizations. During the Provisional Irish Republican Army hunger strikes, hunger strikers were assessed and found not to be suffering from any psychiatric illness. Their self-harm was deemed to be a political act. Anecdotal information from professionals working in the prisons in Northern Ireland, suggests that various terrorist organizations are associated with different types of motivation and, hence, with different psychological profiles. For example, the members of some terrorist organizations are reputed to be very calm, calculating, politically motivated individuals, with very few of the traits that would normally be described as personality disorder. In contrast, the members of other organizations (particularly those organizations with an interest in organized crime) tend to have psychological profiles more typical of violent criminal offenders.

There is no evidence from Northern Ireland suggesting that those who engage in terrorist violence are more likely to be psychiatrically ill than members of the general population. Indeed, it seems that those who perpetrate terrorist violence, as a rule, tend to escape psychologically unscathed. Some individuals are troubled by their conscience, and the general public would like to believe that those who commit particularly horrific crimes will be troubled by these crimes for the rest of their lives. In general, however, it seems that 10 years after a horrific violent crime it is the victims and the bereaved who will wake up screaming in the night, whereas the perpetrator sleeps soundly.

There are positive psychological aspects that help to motivate terrorists. The terrorist act can act as a form of catharsis, relieving feelings of anger and frustration that have often built up over many years. At the same time, terrorists usually see themselves as martyrs. This is a constant theme, from the blood

sacrifice of Patrick Pearse in 1916, to the ideas of martyrdom among Islamic suicide bombers today.

These positive aspects are, for the terrorist, balanced by the threat that he or she poses to family and friends, the fear of living in secret, and the ultimate fear of losing his or her life. Thus, a young man who is considering becoming a terrorist is contemplating a tactic that is both highly effective and highly costly.

Effects of Terrorism on Victims

The effects of terrorist violence on victims has been described in detail (see **Northern Ireland, Post Traumatic Stress Disorder** in). Research performed with survivors of violence in Northern Ireland indicates that normal individuals subjected to extreme terrorist violence are likely to develop posttraumatic stress disorder (PTSD) and that normal individuals who are bereaved by terrorism are likely to develop a grief reaction. We use arbitrarily defined time points to describe the syndromes of delayed and chronic PTSD and to define delayed and chronic grief reactions. However, all of these psychological injuries, which happen to normal people put under intolerable strain, fall within a continuous spectrum, ranging from a psychological upset lasting a few days to symptoms that can be lifelong.

Vulnerable individuals may also be the victims of terrorist violence. Under these circumstances, the violence can act as a nonspecific stressor or life event precipitating an episode of psychiatric illness. For example, individuals prone to depression or anxiety may develop an anxiety neurosis or depressive neurosis if exposed to violence. Clearly terrorist violence has a devastating effect on the lives of those exposed to it, whether they are psychologically vulnerable or not, resulting in incapacitating physical and psychological injuries, as described by the work of Curran and colleagues, among others, working in Northern Ireland.

A high percentage of individuals who experience serious violence develop PTSD. Studies across the world, looking at different types of traumatic incidents, indicate that between one-quarter and one-half of individuals who experience serious trauma develop PTSD. It seems that the key factor in developing this condition is the fact of being made aware of one's own mortality at a very deep level. For this reason, it seems that it is the individual's perception of the event that causes the psychological injury, and because of this, different individuals can respond differently to the same type of trauma. For example, if two friends are caught up in a terrorist bomb blast, one might develop PTSD and the other might escape relatively

unscathed. For the same reason, similar traumatic events may affect the same individual differently at different times. For example, a bus driver may be hi-jacked at gunpoint on four occasions without developing any psychological problems, only to develop severe PTSD when hi-jacked for a fifth time under virtually identical circumstances.

Anyone who has been exposed to serious violence may develop the core symptoms of PTSD, namely reliving the incident, hyperarousal, and emotional constriction. These core symptoms are seen after a wide variety of traumatic situations in different parts of the world. In addition to the core symptoms, an individual may develop a number of additional symptoms under specific circumstances. In situations of terrorist violence, hostage taking and torture are unfortunately frequent occurrences. This type of prolonged trauma tends to result not only in PTSD but can also in a syndrome similar to learned helplessness. Individuals recovering from traumatic events of this kind often experience profound symptoms of depression. Survivors of hostage taking often have these symptoms mixed with feelings of relief to be alive and a paradoxical gratitude to their captors for allowing them to live.

One common form of terrorist violence in Northern Ireland has been assassination attacks on the security forces launched from private houses. In these attacks, families are often held at gunpoint overnight while an ambush is set up. Families who experience these attacks often develop symptoms of PTSD associated with symptoms of depression and learned helplessness. The situation is complicated further by a group of symptoms that are often seen after burglaries. These symptoms include obsessional cleanliness and a feeling that the home has been violated, making it difficult to relax in the house where these events occurred.

Free-floating anger and irritability are core symptoms of PTSD. After natural disasters or tragedies for which no one is to blame, this free-floating anger often takes the form of irritability, and the survivors of natural disasters often direct their anger and hostility toward those closest to them. In the case of terrorist violence, there is a natural focus for this free-floating anger and victims understandably direct their anger toward the terrorist organization. In some circumstances, this can be helpful because there is less tendency for anger to be directed inward or toward other family members.

Individuals who have suffered terrorist violence in their own homes find it very difficult to recover from their psychological injuries while they continue to live in the environment where the trauma occurred. Every time they walk into a certain room they are reminded

of the violence that occurred there. At the same time, survivors of this type of violence often feel so much anger toward the perpetrators of the violence that they are determined not to be forced out of their homes. They refuse to run away. This appears to be a decision of principle, but it can tend to delay the natural healing process. Security lights and alarms are a nightly reminder of the violence that took place in the home. Those individuals who swallow their pride and move to a new environment often recover more quickly from their psychological injuries.

Even after the paramilitary ceasefires in Northern Ireland, there have been a number of terrorist incidents, most notably a bombing outrage in the town of Omagh in which large numbers of civilians were killed. This incident, in particular, generated a number of research studies on PTSD and specifically on the treatment of PTSD. As a result, the Department of Health in Northern Ireland produced the *CREST Guidelines for the Management of Post Traumatic Stress Disorder in Adults* in June 2003. This advice highlights the importance of cognitive-behavioral therapy (CBT) and eye movement desensitization, along with symptomatic treatment with antidepressants.

The Effect of Terrorism on Society

The effects of terrorist violence on a society are rather more difficult to define. As mentioned earlier, terrorism can take many forms, from the organization of mass protests and violence to individual attacks on selected targets. Curran and colleagues have pointed out that the form of civil violence over the years in Northern Ireland changed from mass disturbances, affecting large numbers of people, to more surgical attacks on selected targets, affecting only a very small number of victims. Community studies, examining indicators of mental health, have failed to demonstrate a greater psychiatric morbidity among the population of Northern Ireland. This has been explored in some detail by Cairns and Wilson, who examined various indicators of psychiatric morbidity, including psychiatric admission and referral rates, suicide and attempted suicide, substance abuse, clinical studies, and community studies. The overall impression is that during The Troubles it was remarkably difficult to demonstrate a higher level of psychiatric morbidity among the general population of Northern Ireland. Loughrey and coworkers postulated that this may have been because of the changing nature of violence in Northern Ireland to more selective targeting. The victims of violence, although affected severely, were perhaps too small in number to be detected by community surveys.

In the early years of The Troubles, Lyons described a year by year decrease in the suicide rate in Northern Ireland with the outbreak of civil violence. He believed that this might be explained by Durkheim's theory of anomie. In the late 1800s, Durkheim predicted that, during periods of war, suicide rates would decrease because the aggression of the population was being externalized toward a common enemy. This prevented aggression from being directed inward in the form of self-harm. Lyons's original results seemed to support this. However, when the suicide rates in Northern Ireland were examined over a more extended period of time by Curran and colleagues, it was found that the increase in the suicide rate at the start of The Troubles, observed by Lyons, may have simply been part of a wide year-to-year random variation. Durkheim's theories pertaining to social cohesion may still be relevant, however. Curran and coworkers postulated that the detrimental effect on the mental health of a small number of victims directly affected by terrorist violence might be balanced by a beneficial effect of polarizing sectarian community spirit on a larger number of marginalized individuals within society at large. In other words, although the direct victims of violence suffer greatly, they are small in number in most urban terrorist campaigns. At the same time, polarized sectarian communities, such as those in Northern Ireland, though deplorable in many ways, may paradoxically offer increased social support to marginalized individuals from within their own community. It is difficult to test this hypothesis, but most psychiatrists in Northern Ireland will be aware of some patients whose level of psychological functioning seems to be improved because of the increased social cohesion of the frontline communities in which they live.

Presumably the function of terrorism is to evoke terror. Initially, terrorists in Northern Ireland seemed to have been singularly unsuccessful in this regard. The terrorist campaign went on so long that the population of Northern Ireland viewed the risk of terrorist violence as just another of the statistical risks that face us when we get out of bed each morning. It almost seems that there is a homeostatic mechanism in the human psyche with regard to risk, such that exposure to risk develops tolerance. This seems to apply to risk of all sorts. The population of the Western world has become surprisingly blasé about the risks associated with motor vehicles, despite the huge carnage produced by these machines on our roads. Similarly, most individuals within society accept the risk of violent crime when they leave their homes. These mechanisms for living with risks appear to be largely based on denial and associated with thoughts such as "it will never happen to me."

Some groups within Northern Ireland, such as members of the security forces who are particularly at risk of terrorist violence, develop a gung-ho culture, where risk is denied completely. This is often associated with substance abuse in order to maintain the high levels of denial. There is a saying in Northern Ireland: "You can get used to anything." In the same way that money does not make you happy, it seems that terror does not necessarily make you afraid.

A number of studies within Northern Ireland have demonstrated a rebound of psychiatric morbidity after the declaration of the various terrorist cease-fires. However, this increase in psychiatric morbidity appears to be relatively short-lived. It may be that the atmosphere of peace allows people to come forward and seek help for psychological problems that they had chosen to ignore while the violence was continuing. The reaction of the community to terrorist violence can be seen as mirroring an individual's reaction to trauma. In traumatic circumstances, many individuals continue to function in a calm, detached manner while the violence is occurring, only to find themselves breaking down some hours later after they have had time to think about the dangers of the situation. This seems to be a natural adaptive mechanism. When survival is the first priority, individuals get on with what needs to be done at the time. They postpone thinking about the horror of the situation until it is safe to do so. It seems that societies function in a similar way, using denial mechanisms in order to continue functioning. However, the reality of the situation comes flooding in when peace is restored, resulting in the rebound of psychiatric morbidity.

Terrorism is here to stay. Just as individuals become inured to violence and tend to return to the state of normal denial, so societies learn to live with the chronic risk of terrorist violence and it becomes just one of the many risks in a modern world. When this happens, terrorism no longer succeeds in evoking terror. It becomes a less powerful tactic, and it

becomes less likely that young men and women will be prepared to pay the terrible price of becoming terrorists.

In summary, the effects of terrorist violence are not as we would imagine them to be in a fair world. Most active participants in violence are psychologically unscathed. The lives of victims are often devastated, whereas the level of psychological morbidity in a community exposed to terrorism is lower than might be expected. In the end, the damage done by terrorism to a community appears to be spiritual rather than psychological.

See Also the Following Articles

Northern Ireland, Post Traumatic Stress Disorder in; Posttraumatic Stress Disorder – Clinical; War, Suicide and Sacrifice.

Further Reading

- Bell, P., Kee, M., Loughrey, G. L., et al. (1988). Post traumatic stress disorder in Northern Ireland. *Acta Psychiatrica Scandinavica* 77, 166–170.
- Cairns, E. and Wilson, R. (1984). The impact of political violence on mild psychiatric morbidity in Northern Ireland. *British Journal of Psychiatry* 145, 631–635.
- Clinical Research Efficiency Support Team (2003). *CREST guidelines on the management of post traumatic stress disorder in adults*. Belfast, UK: Department of Health and Social Services.
- Curran, P. S., Finlay, R. J. and McGarry, P. J. (1988). Trends in suicide in Northern Ireland 1960–86. *Irish Journal of Psychological Medicine* 5, 98–102.
- Loughrey, G. L., Curran, P. S. and Bell, P. (1993). Post traumatic stress disorder and civil violence in Northern Ireland. In: Wilson, J. P. & Raphael, B. (eds.) *The international handbook of traumatic stress syndrome*, pp. 377–380. New York: Plenum.
- Lyons, H. A. (1972). Depressive illness and aggression in Belfast. *British Medical Journal* 1, 342–344.

Terrorism: Consequences of See: Crime Victims; Impact of Terrorism on the Development of Mental Health Symptoms; Lockerbie Air Crash, Stress Effects of; Suicide Terrorism, Genesis of; Terrorism; War, Suicide and Sacrifice.

Terrorism: World See: Firefighters, Stress in; Religion and Stress.

Testosterone See: Androgen Action.

Thermal Stress

C M Blatteis

University of Tennessee Health Science Center,
Memphis, TN, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
C M Blatteis, volume 3, pp 581–584, © 2000, Elsevier Inc.

Thermal Environments

Physiological Mechanisms of Heat Exchange

Metabolic Heat Production

Response Patterns to Sustained Thermal Stress

Other Environmental Factors Influencing Thermal Stress

Pathophysiological Consequences of Thermal Stress

Thermal Stress Not Due to Exposure

Glossary

Acclimatization

Autonomic or behavioral changes occurring within an organism that reduce the strain caused by stressful changes in the natural climate. Acclimation refers to the adaptive changes that occur within an organism in response to experimentally induced changes in a particular climatic factor (e.g., ambient temperature) in an artificial, controlled environment.

Autonomic temperature regulation

The regulation of body temperature by involuntary thermoeffector responses to heat and cold that modify the rates of heat production and heat loss. Autonomic in this sense does not mean that all responses are controlled by the autonomic nervous system.

Body heat balance

The steady-state relation in which the heat produced by the body equals that lost to the environment.

Brown adipose tissue (BAT)

A particular form of adipose tissue with selective distribution that differs both structurally and functionally from white adipose tissue. Its main role appears to be regulatory thermogenesis. Its activity is controlled by the sympathetic nervous system and various humoral factors, especially catecholamines.

Core temperature

The mean temperature of the thermal core (i.e., the interior of the body), usually represented by the temperature of a particular internal site, for example, rectal or oral.

Thermal strain

Any deviation of body temperature induced by sustained thermal stress that cannot be fully compensated by

temperature regulation. Also, any activation of thermoeffector activities in response to thermal stress that causes sustained changes in the state of other, nonthermal regulatory systems.

An organ system and its function that affect heat balance in a controlled manner as part of the processes of temperature regulation.

Thermoeffector

Thermal Environments

Thermal stress is defined as any change in the thermal relation between a temperature regulator and its environment that, if not compensated by thermoregulatory effector reactions, will result in hyper- or hypothermia. A temperature regulator, thus, is an organism that has the capacity to appropriately activate coordinated, autonomic, and/or behavioral heat conservation or loss (thermolytic) and heat production (thermogenic) effectors such that its body (core) temperature (T_c) remains relatively stable even in thermal environments that, *a priori*, might be expected to cause T_c to rise above (hyperthermia) or fall below (hypothermia) its normal range. Animals that possess this capability are called homeotherms. The condition in which T_c is within ± 1 SD of its normal range is termed cenothermy. The thermal environment in which neither the mechanisms for extra heat production (e.g., shivering) nor for extra heat loss (e.g., sweating) are activated is termed the thermoneutral zone; under these conditions, the T_c of healthy subjects at rest does not vary more than $\pm 2^\circ\text{C}$ around its 24-h mean. Acute exposure to an environment warmer than thermoneutrality results in more heat being gained from the environment than lost to it; conversely, more heat is lost than produced by the body on exposure to an environment cooler than thermoneutrality. From a thermal physiological point of view, inactive humans are tropical animals when naked, thermoneutrality being an ambient temperature (T_a) of $28\text{--}30^\circ\text{C}$. Thermal stress, therefore, begins as soon as T_a is higher or lower than this range. However, homeotherms may be heat-stressed at low T_a if they are active.

Physiological Mechanisms of Heat Exchange

The unavoidable consequence of heat (or cold) exposure is that unless more heat were lost from the body

(or conserved and produced by it), the body-heat balance could not be maintained and, hence, an excessive rise (or fall) of T_c could potentially occur. However, the relative stability of T_c under most thermal conditions indicates that both heat production and heat loss are regulated and that, furthermore, temperature sensors exist throughout the body that measure its temperature continuously. In the skin, the detection of warm and cold sensations has recently been ascribed to the activation of specific, temperature-activated transient receptor potential ion channels (thermoTRPs) located within the free endings of cutaneous sensory nerves. The nature of the internal body thermosensors are not yet clarified. This information is transmitted to the brain, where it is integrated and transduced into signals to thermoeffectors that are then activated so that appropriate countermeasures are initiated. These responses involve both behavioral and autonomic processes. The former modulate primarily the effective heat-exchanging surface above the skin and are expressed by volitionally or instinctively changing the skin's microclimate (e.g., by postural changes, moving to a different thermal environment, or wetting the skin). Behavioral adjustments generally precede autonomic responses; they are more economical energetically and, often, sufficient. Autonomic adjustments involve factors below the skin. They control the patency of the small arterioles that deliver blood (i.e., heat) to the surface by means of sympathetic vasoconstrictor and vasodilator systems and various cotransmitters. When open (cutaneous vasodilation), the warm arterial blood passes through these, warming the skin. As long as the skin temperature remains higher than the ambient, heat is dissipated into the environment. Conversely, when these channels are closed (cutaneous vasoconstriction), the bloodstream is conveyed through deeper arteriovenous anastomoses such that heat does not reach the surface so readily, reducing the skin-to-ambient temperature gradient and conserving body heat. This is aided further by the fact that the deep, larger vessels of the limbs lie in parallel so that warm arterial blood flowing toward the periphery passes close to cooler, returning venous blood flowing in the opposite direction, thus allowing heat to be transferred by countercurrent heat exchange from the arterial to the venous flow. This pre-cools the arterial blood, thereby contributing to increasing the thermal insulation of the extremities.

When ambient temperature exceeds skin temperature, the body gains heat. Under these conditions, the only means by which heat can be dissipated is by evaporation. For this purpose, animals that possess eccrine sweat glands activate them; those that do not, pant. The efficacy of this process is dependent directly

on the ambient and skin vapor pressure difference; that is, when the ambient relative humidity is high, little evaporation occurs. Sweating is much more expensive metabolically than cutaneous vasodilation. It also results in large losses of water and sodium, which are potentially dangerous.

Metabolic Heat Production

The heat released when the organism is resting at thermoneutrality is called obligatory nonshivering thermogenesis (NST-O); it is derived from the metabolic activity of the body. T_c is maintained within its narrow range by the controlled loss of heat from this activity. The skeletal muscle mass makes up the greatest mass of the body (~45%), but at rest it produces only 18% of the NST-O. However, more energy and, hence, heat are expended when the muscles actively contract, in direct proportion to the external work done (e.g., as in exercise). However, if no external work is performed by the contracting muscles, most of the energy expended is given up as heat. This is, in fact, what occurs in a cold environment by means of shivering. However, shivering is fatiguing, not terribly efficient (~11%), and cannot be sustained indefinitely. Another thermogenic process that does not involve muscular contractions, namely thermoregulatory nonshivering thermogenesis (NST-T), is activated in such cases. The principal effector organ of NST-T is BAT. Its thermogenic capacity is derived from the collapse of the electrochemical proton gradient normally generated by cellular respiration, limiting adenosine triphosphate synthesis, the energy currency of the body. Consequently, the energy expended is given up as heat and conducted to the numerous vessels permeating this tissue and is convected by the bloodstream throughout the body. A specialized protein, uncoupling protein 1 (UCP1, or thermogenin), present in the mitochondrial inner membrane of BAT adipocytes, facilitates this process when activated by the sympathetic nervous system. BAT NST-T is particularly prominent in neonates, but recedes with age and size. It develops over time in cold-exposed adult rodents and other small animals. In neonates and small animals, NST-T and shivering are meshed reciprocally.

Response Patterns to Sustained Thermal Stress

If a person remains exposed, unprotected, for a prolonged time to the natural climate, he or she would be at risk of eventually becoming physically exhausted from the continuous shivering in the cold or dehydrated and hyponatremic from the extensive sweating

in the heat. To prevent such noxious results, the body institutes thermoadaptive modifications that reduce the physiological strain caused by prolonged stressful thermal exposures. These adaptations, termed acclimatization, occur over an extended time course, measured in days and longer. The gradual replacement of shivering by NST-T in long-term cold exposure of small animals has already been mentioned. This type of response is called metabolic adaptation. Some homeotherms react to prolonged cold by entering into a state of lethargy called hibernation. Adult humans acclimatize to cold by using appropriate behaviors (clothing, heaters, etc.). The autonomic analog in most animals is the growth of thicker fur and the deposition of subcutaneous fat (insulative adaptation). Inadequately insulated, chronically cold-exposed humans may also develop a lower T_c and a lesser fall in skin temperature (i.e., a warmer skin; hypothermic adaptation) in time.

In long-term heat exposure, the sweat glands develop the capacity to secrete more water and sooner (i.e., they become activated at a lower threshold T_c over time); the amount of salt lost in sweat from the body also is reduced. These effects can be achieved by 1–4 h of heat exposure for 5–7 days and are mediated by the increased secretion of antidiuretic hormone and aldosterone; thirst is also stimulated. The process is facilitated by physical exercise, which, by itself, enhances the sweat rate, and persists as long as the heat exposure continues; afterward, it gradually wanes. Analogous to hibernation in the winter, some homeotherms enter into a state of summer lethargy called estivation. During moderate chronic heat exposure, the onset of both sweating and skin vasodilation occur at a higher T_c (hyperthermic adaptation).

Systemic adaptation to thermal stress also extends to the cellular level (thermotolerance). Thus, cells within the body core (e.g., exercising skeletal muscle) or in its shell (e.g., skin cells) that are acutely exposed to intense, but not lethal, heat or cold react to these stressful stimuli by producing specialized proteins (heat shock and cold shock proteins) that serve to protect them against a subsequent, otherwise potentially lethal heat or cold exposure. In contrast to heat or cold acclimatization, however, thermotolerance develops rapidly (within hours) but is temporary (3–5 days).

Other Environmental Factors Influencing Thermal Stress

In nature, a variety of factors are present that aggravate the risk of thermal stress occurring sooner than due to a T_a change alone. These include water, due to

its high thermal conductivity (beneficial in the heat, detrimental in the cold, as in wet-cold, e.g., chilblains and immersion foot); blowing wind (windchill factor), by thinning the thickness of the thermal boundary layer, a thin film of warmed air that surrounds the body (beneficial in the heat, detrimental in the cold); and barometric pressure, by affecting heat exchange by convection (reduced at high altitude) and evaporation (enhanced at high altitude) and by affecting the partial pressure of atmospheric gases (e.g., the metabolic response to cold is reduced at altitude). Other environmental factors, such as the relative humidity and the intensity of solar radiation, also contribute to the imposition of thermal stress.

Pathophysiological Consequences of Thermal Stress

During periods of heavy sweating, a strain is imposed on the heat-exposed individual's homeostatic systems that, if not properly compensated, can cause the development of symptoms (heat illnesses). Similarly, when the various defense reactions are inadequate to cope with the stress of cold exposure, severe local tissue injury (cold injuries) may develop. The risks of hypo- and hyperthermia were mentioned earlier. The susceptibility to thermal stress increases with age.

Thermal Stress Not Due to Exposure

Thermal stress may also occur due to factors not involving exposure. For example, as already mentioned, exercise can induce heat stress by increasing the rate of heat gained (i.e., produced) in excess of that lost. Conditions of decreased or increased T_c may be induced, even in thermoneutrality, by drugs that impair normal thermoregulatory functions (iatrogenic hypo- or hyperthermia), by idiosyncratic drug reactions (e.g., malignant hyperthermia), or secondary to underlying illnesses that compromise the body's thermoregulatory mechanisms (e.g., peripheral vascular disease, paralysis, or endocrinopathies). Fever, however, is not a thermal stress because it represents a regulated, deliberate T_c rise resulting from the active operation of thermogenic effectors and not the unavoidable consequence of the passive gain of heat in excess of the capability of active thermolytic effectors to dissipate it; it is a defensive response to the invasion of the body by infectious pathogens. A regulated fall in T_c also occurs, particularly in rodents exposed to toxic chemicals and hypoxia/ischemia; it protects the afflicted host from the consequent pathological effects of these insults.

See Also the Following Articles

Environmental Factors; Environmental Stress, Effects on Human Performance; Febrile Response; Heat Resistance; Heat Shock Response, Overview; Hyperthermia; Hypothermia; Temperature Effects; Thermotolerance, Thermoresistance, and Thermosensitivity.

Further Reading

- Blatteis, C. M. (ed.) (1998). *Physiology and pathophysiology of temperature regulation*. Singapore: World Scientific.
- Bligh, J. (1998). Mammalian homeothermy: an integrative thesis. *Journal of Thermal Biology* 23, 143–258.
- Cannon, B. and Nedergaard, J. (2004). Brown adipose tissue: function and physiological significance. *Physiological Reviews* 84, 277–359.
- Charkoudian, N. (2003). Skin blood flow in adult human thermoregulation: how it works, when it does not, and why. *Mayo Clinic Proceedings* 78, 603–612.
- Fregly, M. J. and Blatteis, C. M. (eds.) (1996). *American Physiological Society handbook of physiology. Sec. 4: Environmental physiology* (2 vols.). New York: Oxford Univ. Press.
- Glossary of terms for thermal physiology (2001). *Japanese Journal of Physiology* 51, 245–280.
- Kenney, W. L. and Munce, T. A. (2003). Invited review: aging and human temperature regulation. *Journal of Applied Physiology* 95, 2598–2603.

Kregel, K. C. (2002). Heat shock proteins: modifying factors in physiological stress responses and acquired thermotolerance. *Journal of Applied Physiology* 92, 2177–2186.

Patapoutian, A., Peier, A. M., Story, G. M., et al. (2003). ThermoTRP channels and beyond: mechanisms of temperature sensation. *Nature Reviews* 4, 529–539.

Pozos, R. S. (ed.) (2001). Sec. II: Cold environments, In: Pandolf, K. B. & Burr, R. E. (eds.) (2001). *Textbooks of military medicine: medical aspects of harsh environments* Washington, DC: U.S. Department of the Army, Office of the Surgeon General.

Reid, G. (2005). ThermoTRP channels and cold sensing: what are they really up to? *Pflügers Archiv – European Journal of Physiology* 451, 250–263.

Sell, H., Deshaies, Y. and Richard, D. (2004). The brown adipocyte: update on its metabolic role. *International Journal of Biochemistry and Cell Biology* 36, 2098–2104.

Stocks, J. M., Taylor, N. A., Tipton, M. J., et al. (2004). Human physiological responses to cold exposure. *Aviation, Space and Environmental Medicine* 75, 444–457.

Wenger, C. B. (ed.) (2001). Sec. I: Hot environments, In: Pandolf, K. B. & Burr, R. E. (eds.) (2001). *Textbooks of military medicine: medical aspects of harsh environments* Washington, DC: U.S. Department of the Army, Office of the Surgeon General.

Thermoregulation See: Febrile Response; Thermotolerance, Thermoresistance, and Thermosensitivity.

Thermotolerance, Thermoresistance, and Thermosensitivity

S Gatti McArthur, D Alberati and T Bartfai

Hoffmann-LaRoche, Basel, Switzerland

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S Gatti McArthur, D Alberati, T Bartfai, volume 3, pp 585–594, © 2000, Elsevier Inc.

Temperature Control
Thermoresistance
Thermotolerance
Thermosensitivity
Concluding Remarks

Glossary

Adaptation/
habituation

A modification of regulatory processes taking place during prolonged or repeated exposure to extreme ambient temperatures.

Poikylothermy
(PKT)

The typical metabolic condition of ectothermic animals that use external sources of heat to increase body temperature. In endotherms, a status similar to PKT can be observed in extreme ambient conditions or under pharmacological treatment with ethanol or volatile anesthetics.

<i>Thermoneutrality (TN)</i>	A range of ambient temperatures. Within the TN range, it is not necessary for the organism to activate warming or cooling mechanisms to keep the core body temperature constant. The TN range is different in different species and strains. Furthermore, there is a high degree of interindividual variability, and these differences could be related to differences in diet, age, and body weight. A TN range can be determined experimentally by measuring O₂ consumption as an indirect index of the rate of oxidative metabolism. More precise analysis requires biocalorimetry measurements to determine heat production (HP) and heat exchange (HE) in the different regions/tissues/cells of the body. Because rapid eye movement (REM) sleep is also sensitive to the ambient temperature, the TN range can be determined by measuring the total REM sleep time (longer in the TN zone).	cause intracellular HP. In the presence of uncouplers, HP is proportional to O₂ consumption, percentage of uncoupling effect, mitochondrial membrane potential, and duration of the effect.
<i>Thermopreferendum</i>	The temperature (T) range selected by microorganisms or animals placed in a thermocline. The association of a certain T range with the presence of food or memory of previous ambient conditions can influence the selection by the organism. Behavioral tests based on the resetting of the thermopreferendum can be used to demonstrate T set point shifts.	
<i>Thermoresistance</i>	Hormonal, metabolic, and cardiovascular reactions to environmental temperature changes that allow the organism to keep a constant control on core body temperature in the presence of unfavorable ambient conditions.	
<i>Thermosensitivity</i>	Capacity to respond to temperature changes. Thermosensitive (Ts) neuronal units modify the number of pulses per minute in response to small changes in temperature. Based on the response to direct warming or cooling procedures, warm-sensitive (pulses/10 min = $Q_{10} > 1.2$) or cold-sensitive ($Q_{10} < -0.6$) neurons can be distinguished, respectively.	
<i>Thermotolerance</i>	Mechanisms of adaptation that allow survival in extreme ambient conditions, for example, short-term changes in heat production and vascular control of heat exchange and long-term changes in membrane composition and insulation capacity.	
<i>Uncoupling effect</i>	When the protonic gradient across the inner membrane of mitochondria is partially or totally dissipated and heat production (HP) takes preference over ATP production. Uncoupling proteins and uncoupling agents (e.g., dinitrophenol)	

This article briefly reviews the mechanisms controlling heat production (HP) and heat exchange (HE) in endotherms (in mammals in particular) able to adapt to different, and sometimes extreme, environmental conditions. Hibernation (and estivation) represents a classic example of extreme control on thermoregulation. When food is scarce, lowering the core body temperature (CBT) allows the survival of small mammals and reptiles with minimal energetic costs. Introductory concepts on temperature (T) control in endotherms are also given.

Temperature Control

Temperature (T) control or, better, the control of the effects of T is a vital need for every cell type, even for prokaryotic organisms, given that protein conformation and membrane integrity are jeopardized at freezing, as well as at higher than 40°C temperatures (in aqueous environment that prevails *in vivo*). Exceptions to this rule in cryobiology require non-aqueous solutions. In bacteria, adaptive changes associated with small T variations (Δ : 2–10°C) always involve changes in cell membrane composition and involve the rapid induction of enzymes such as fatty acid desaturases. *Escherichia coli* responds to small changes in T by altering its growth rate, swimming behavior, and thermotaxis. Several membrane proteins (methyl-accepting chemotaxis proteins) are involved in the control of these events. The thermo/chemoreceptors are reversibly methylated at specific glutamate residues located in the C-terminal cytoplasmic domain and point mutations at the methylation sites can even invert bacterial thermotaxis. Nematodes (e.g., *Caenorhabditis elegans*) possess specific thermosensory neurons and interneurons organized in a network controlling thermotaxis. When placed in a thermal gradient, they display the ability to recognize and select the region where T is similar to the T at which they were cultured previously. Curiously, it is possible to distinguish between a cryophylic and a thermophylic behavior by selective laser ablation of one of the two interneurons controlling thermotaxis or by null mutations of either of two homeobox genes, *lin-11* or *ttx-3*, respectively. The neural circuit that regulates the response to changes in T in *C. elegans* has been characterized in detail. It consists of a thermosensory organ (neuron), two interneuronal populations, and motor neurons. Vertebrates

(amphibians and reptilians), such as lizards or frogs, exhibit a thermopreferendum when placed in a thermocline even if they are poikylothermic and thus unable to keep their core body temperature (CBT) constant in the same thermal gradient.

Only mammals and birds (warm-blooded animals) have developed a complex net of neurological, hormonal, cardiovascular, and metabolic systems that are able to keep a constant control on the CBT despite geographical, seasonal, or even diurnal differences in air T and water T. The possibility of keeping CBT constant represents an important step from an evolutionary point of view. At the cellular level, it implies that the rate of enzymatic reactions, membrane properties, and electrical excitability remain rather constant (and can be optimized) despite dramatic changes in heat production (HP) and heat exchange (HE) caused, for instance, by motor activity or cold exposure. In general, homeothermy may be one of the key evolutionary factors that have allowed uterine growth of fetuses and nocturnal motor activity.

Temperature Control in Homeotherms

Several homeostatic mechanisms allow the maintenance of a constant CBT in warm-blooded animals exposed to different ambient temperatures (AT): (1) the hormonal control on basal metabolic rate (BMR) (short term) and body mass (long term), (2) the hypothalamic/spinal neuronal network setting and controlling CBT, (3) the sympathetic and parasympathetic neuronal systems activated alternatively to govern effector mechanisms to warm or cool the body, and (4) the steady-state bulk of energy expenditure represented by the activities of muscular and adipose tissues.

Thermoresistance

Basal Metabolic Rate

When compared to poikylothermy (PKT) in animals of equal body mass, homeothermy requires the maintenance of higher BMR, with higher food intake and higher O₂ consumption. Indirectly, this means that homeotherms actuate behavioral and dietetic changes, adjusting food intake to seasonal T changes in order to guarantee sufficient energy storage. We should distinguish between (1) resting (basal) metabolic rate (MR; the heat produced is also called obligatory thermogenesis, T_g) and (2) inducible MR (the heat produced is also called facultative T_g). The resting MR can be largely modified by slow mechanisms of adaptation, whereas the inducible MR is activated quickly in order to react effectively to cold exposure. The major ways used to increase T_g rapidly are

uncoupling of oxidative phosphorylation in mitochondria (nonshivering T_g) and muscular shivering. BMR in mammals is correlated positively to the ratio of body surface to body mass; animals with a higher surface/mass ratio require exponentially more energy to keep their CBT in the range of 36–37°C.

These dramatic differences in BMR should be considered when studying the adaptive mechanisms developed by different species to allow survival in extreme conditions. Thus, for example, the drastic reduction of CBT that accompanies hibernation is advantageous for small animals (e.g., squirrels) because they save much more energy proportionally than larger animals, such as bears. Passive mechanisms of thermotolerance, such as the reduction of HE because of increased insulation, do not modify the BMR. They simply expand the range of AT compatible with survival. Food digestion is another source of T_g, and different species change their diets in relation to the amount of heat required in different seasons.

Hormonal Control on Basal Metabolic Rate

The thyroid hormones, thyroxine (T₄) and triiodothyronine (T₃), play a central role in controlling cellular ATP levels, MR, and T_g during both fetal and adult life. T₃ and T₄ increase obligatory T_g and work in synergy with catecholamines to trigger facultative T_g. Furthermore, thyroid hormones are involved in embryogenesis, particularly in the differentiation of brown fat tissue (BAT). Studies have clarified the molecular mechanisms sustaining the different metabolic effects and the modulatory actions of thyroid hormones on insulin, glucagon, growth hormone, and adrenaline effects. Thyroid hormones enhance or suppress gene transcription via the specific interaction with different nuclear receptors. The effects of T₃ and T₄ on T_g are related to the induction of several classes of proteins that include uncoupling proteins (UCPs); FAD-linked glycerol-3-phosphate dehydrogenase, a carrier transferring electrons from NADH to ubiquinone in complex I in the respiratory chain of mitochondria; and adenine nucleotide translocase, the ATP/ADP exchanger, which is also present in the inner mitochondrial membrane, where it controls the concentration of phosphate acceptor (ADP).

Table 1 shows genes whose expression is known to be under the control of thyroid hormones. UCP-1 protein expression is induced in rat BAT by cold exposure and suppressed by fasting; and T₃ increases UCP-1 production both by increasing gene transcription and by stabilizing mature UCP-1 mRNA. In BAT, local control of the relative levels of T₃ and T₄ is maintained via the presence of thyroxine 5'-deiodinase. This cold-inducible enzyme converts T₄ to the

Table 1 Genes involved in induction of thermotolerance or thermoresistance under the transcriptional control of thyroid hormones^a

Mitochondrial proteins	FAD-dependent glycerol-3-phosphate dehydrogenase (ATP/ADP exchange activity)
	Tom20: outer membrane receptor
	Matrix heat shock protein 70
	Adenine nucleotide translocase (ATP/ADP exchange activity)
	Acyl-CoA oxidase
	Na ⁺ K ⁺ -ATPase (α and β subunits)
	Acetyl-CoA carboxylase
	Uncoupling proteins (UCP-1,-2,-3)
	ATP-dependent protease
Hormones	Renin
	Steroidogenic acute regulatory protein
	Oxytocin
	Prepro-TRH (negative control)
Other enzymes	Cholesterol-7 α -hydroxylase
	Malate dehydrogenase (negative control)
	Cytochrome P450 (negative control on CYP4A1, CYP4A3, and CYP4A23)
	Ca ²⁺ -ATPases (in sarcoplasmic reticulum: SERCA1 and SERCA2)
	Phosphoenolpyruvate carboxykinase
Membrane receptors and ion channels	Glucagon receptor
	β 1 and β 3 adrenoreceptors
	Potassium channels: Kv1.2, 1.4, 1.5, 2.1, and 4.2 α subunits
Nuclear factors	Steroid receptor coactivator-1
	Cold-inducible coactivator of nuclear receptors
Others	Glucose transporter GLUT4
	Myosin isoforms
	Phospholamban: a major target of the β -adrenergic cascade in the heart

^aSERCA, sarcoplasmic reticulum Ca²⁺ ATPase; TRH, thyrotropin-releasing hormone.

more potent T₃, thereby locally modulating the ratio between T₃ and T₄ levels. In muscles, thyroid hormones also control the expression of genes coding for myosin isoforms, the Na⁺-K⁺-ATPase pumps, and the Ca²⁺-ATPase channels of the sarcoplasmic reticulum, sustaining, in this way, muscular performance during shivering.

Thyrotropin (TSH) production in the pituitary gland and, upstream, thyrotropin-releasing hormone (TRH) production in several hypothalamic and raphe nuclei control the synthesis and secretion of thyroid hormones. TRH also seems to exert additional actions in the suprachiasmatic nuclei (SCN), amygdala, lateral septal nuclei, and pineal gland. These actions are always related to chronobiological changes of the MR. During hibernation, the levels of thyroid hormones, TSH, TRH, and their receptors are decreased in all of the just-mentioned brain regions (except for TRH levels in the pineal gland).

In recent years, pharmacological therapies have been proposed to increase HP in obese patients, with special attention to the hypothalamic systems controlling the balance between energy storage and energy expenditure. Interesting observations were made in *ob/ob* and *db/db* obese mice, which lack the control of leptin (the starvation hormone) on adipose tissue metabolism. These animals exhibit a CBT that is 1–2°C below normal; several studies point to the

almost total impairment of nonshivering T_g present in *ob/ob* torpid animals from the first days of life. This is probably due to (1) the defective regulation of thyroxine 5'-deiodinase, (2) the attenuated UCP mRNA expression in BAT during cold exposure, (3) the higher circulating levels of corticosterone observed in *ob/ob* mice, and (4) the increased sensitivity of these animals to leptin. Moreover, leptin can affect thermogenesis, modulating the sympathetic outflow directly as it activates neurons that coexpress cocaine- and amphetamine-regulated transcript (CART) and pro-opiomelanocortin (POMC) neurons that innervate sympathetic preganglionic neurons in the thoracic spinal cord. One of the hypothalamic nuclei constitutively expressing high levels of the leptin receptor long form is the arcuate nucleus. This nucleus contains NPY-glucose sensitive neurons, which represent a possible key link between leptin and T control because pro-NPY mRNA expression is inhibited by leptin and the direct injection of NPY into the anterior or dorsomedial hypothalamus causes deep hypothermia.

Hypothalamic and Brain Temperatures

Thermosensitive (Ts) neurons are present in the anterior brain stem of several endothermic vertebrates and even in ectotherms, further confirming a role for these neurons in the CBT control from a very

early stage in the phylogenetic scale. In mammals, the local warming or cooling of hypothalamic (anterior preoptic, among others) and spinal regions triggers panting, autonomic responses, and muscular shivering, respectively. Ts neurons are also activated by remote thermal stimuli. Approximately 25% of the neurons in anterior hypothalamus are Ts, and Ts neurons can also be found in various nuclei involved

in thermal control (**Figure 1**), circadian rhythms (SCN), and sleep induction (diagonal band). The hypothalamic temperature (HT) plays a key role in the control of the CBT. The hypothesis of HT as an independent variable controlled locally and tightly, integrating peripheral and extrahypothalamic stimuli, is further sustained by the observation that T undergoes smaller fluctuations in hypothalamus

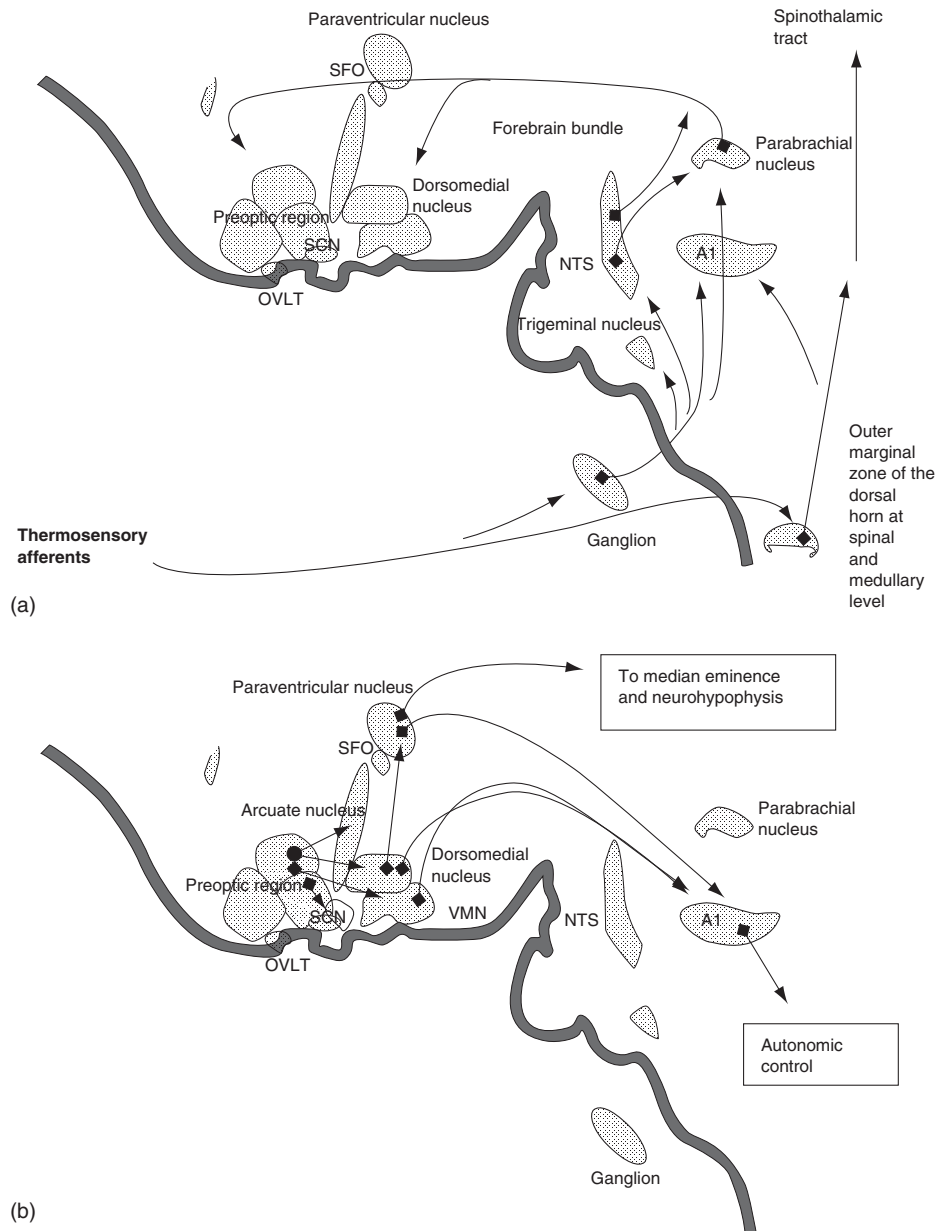


Figure 1 Schematic view of the principal thermosensory afferent and efferent systems in the rat brain. a, Afferent pathways; b, main descending efferent pathways. Hypothalamic regions and autonomic nuclei involved in core body temperature control are highlighted in a sagittal section of the hypothalamic/bulbar region. In a, periventricular hypothalamic organs – subformal organ (SFO) and organum vasculosum lamina terminalis (OVLT) – are also indicated; their projections to the preoptic and paraventricular nuclei control the thermogenic response to circulating hormones. The suprachiasmatic nucleus (SCN), the principal circadian clock of the body, receives light information from the retina and projects to the autonomic division of the paraventricular nucleus. In b, the main hypothalamic descending efferent pathways control thermogenesis via the activation of the autonomic system. NTS, nucleus of the solitary tract; VMN, ventromedial nucleus.

than in other brain and trunk regions in response to several different stimuli. For instance, during hibernation, HT is set approximately 6–10°C higher than CBT. It is generally believed that local hypothalamic and brain heating is controlled by cardiovascular mechanisms, dependent on blood T and perfusion. Several studies were able to demonstrate that mammalian species can decrease the T of the brain below the T of the blood circulating in the large arteries of the trunk. This effect, namely selective brain cooling (SBC), is possible because of the presence of a carotid thin-walled rete that keeps the arterial blood in close contact with the cooler venous blood of the cavernous sinus. What is really under control is the T of the venous blood, which can be modified (sphincter of the angularis oculi vein), allowing the entrance into the sinus of venous blood coming from the upper respiratory tract and the facial vein. The activation of mechanisms of SBC is dependent on CBT and is usually evident when CBT is approximately 40–41°C. Again the onset and the degree of the effect are determined predominantly by T signals of hypothalamic origin. Although the carotid rete is not present in humans, several groups have suggested that SBC could also play a critical role in controlling brain T in humans, at least during physical activity and prolonged exposure to the sun. Endotherms, and mammals in particular, can count on a tight endocrine control of the HT set point, which can be modified in response to infection (fever), sleep, circadian and ultradian rhythms, pregnancy, and fertility cycles. The interaction among different stimuli is possible because of the multimodal response of Ts neurons to endogenous pyrogens, estrogens, glucose levels, osmolarity, and blood pressure, among other stimuli.

Autonomic System and Temperature Control

The major hypothalamic, bulbar, ganglionic, and spinal nuclei involved in the control of autonomic response during thermal stress are depicted in [Figure 1b](#). Sympathetic activation can cause an immediate reduction of skin blood flow in thermoregulatory vessels, piloerection, heat production in BAT, muscular shivering, and increased respiratory rate. Parasympathetic control, however, is exerted mainly on exocrine sweat glands and cardiovascular and respiratory systems, respectively. The sympathetic system exerts a tight thermoregulatory control via the rhythmic and tonic activity of postganglionic neurons, whereas the integration of spinal and bulbar inputs is controlled by preganglionic neurons. Furthermore, every peripheral nerve contains silent units that can be activated in the presence of stress stimuli such as heat.

Adaptive changes in parasympathetic activity after physical exercise, especially in hot environments,

cause bradycardia and increased exocrine sweat gland activity. In normal individuals, 2–4 million sweat glands are able to secrete approximately 3 l of hypotonic fluid/h. In athletes or acclimated people, the sweating rate is higher and starts as soon as there is heat load because of the lowered threshold of sweating response in hypertrophic glands.

Nonshivering Thermogenesis

In children and in many animal species, the major source of heat is BAT. The main role in heat production in BAT adipocytes is played by the mitochondrial protein thermogenin (UCP-1). UCP-1 may represent up to 15% of the total protein content of the inner mitochondrial membrane, and during 1 week of cold exposure the total content of UCP-1 in rat intrascapular BAT rises from 60 to approximately 300 µg, whereas the total protein content increases little. It has been calculated (using *in vitro* microcalorimetry) that 1 kg of stimulated BAT can produce approximately 300 W of heat energy. The amount of BAT, and of vessels supplying O₂ to the tissue, can be increased trophically during acclimation processes. Heat production in BAT is obtained by dissipation of the protonic gradient across the inner mitochondrial membrane. Studies indicate that the fatty acid shuttle of protons is responsible for this effect. The system is under the control of β₃ receptor activation in several species.

The general assumption of a lack of nonshivering T_g in adult humans has to be considered in light of the discovery of UCPs other than the BAT-UCP-1; UCP-2, UCP-3, UCP-4, and BMCP-1 are expressed differentially, mainly in the white adipose tissue, liver, heart, glycolytic muscles, and brain (UCP-4 and BMCP-1 are brain-specific). Structurally, these proteins are similar to other mitochondrial carriers, such as the oxoglutarate carrier. [Figure 2](#) shows a probable phylogenetic tree of this class of proteins based on the available cDNA sequences of UCPs from various species. Putative proteic sequences from *C. elegans* and yeast suggest the presence of these proteins also in lower evolutionary levels organisms. The analysis of the different UCP cDNAs shows that UCP1, UCP4, BMCP1, and plant UCPs possibly represent a more recent step from the evolutionary point of view, whereas UCP-2 may be phylogenetically older. The comparative analysis of the structure of the available UCP genes supports this hypothesis further in that the UCP2 gene does not contain a TATA box and is constitutively expressed and rather ubiquitous. UCP-2 and UCP-3 genes are contiguous, suggesting possible mechanisms of gene duplication. Tissue levels of the different UCP gene products are under translational control. Their expression is correlated with changes

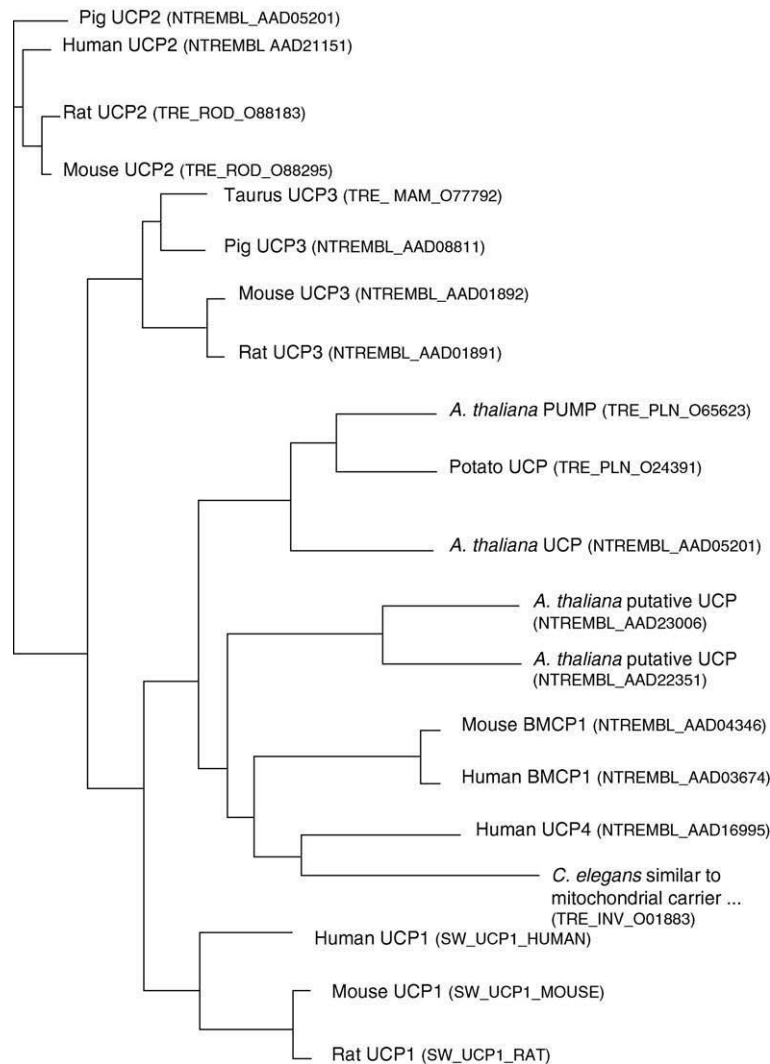


Figure 2 Phylogenetical tree of the uncoupling protein family(ies). The discovery of plant mitochondrial uncoupling protein (PUMP), of UCP-like proteins in *Arabidopsis thaliana*/potato, and of several novel mammalian uncoupling proteins suggests that the ancestral UCP gene has probably evolved prior to the divergence into animals and plants. This is also confirmed by the presence of a protein with uncoupling activity into the mitochondria of *Acanthamoeba castellanii*, a nonphotosynthetic protozoan, which in molecular phylogenesis appears on a branch basal to the divergence points of plants, animals, and fungi. The phylogenetic tree was obtained using the GCG software package.

in ATP production (lowered), inner mitochondrial membrane potential (lowered), and heat production (increased). However, their functional role in T_g is still under debate. UCP-2, UCP-4, and BMCP-1 mRNA expression in the brain and UCP-2 mRNA expression in the spleen, lymph nodes, and thymus suggest other possible physiological roles. Furthermore, UCPs are regulated differentially (possibly in a tissue-specific manner). The different affinity for free ATP/ADP and GTP/GDP nucleotides (inhibitors) could cause the activation of uncoupling at different metabolic states. Possible differences in terms of affinity to fatty acids and the influence of pH on nucleotide binding still require further characterization.

Muscular Shivering

Shivering is controlled by motor neurons and is caused by the repetitive contraction of both extensory and flexory muscles with a frequency directly proportional (on a logarithmic scale) to BMR. The Renshaw interneuron activation is responsible for the synchronization of motoneurons and for the upper discharge frequency. During cold exposure, more than one-third of the total heat produced can come from shivering, and muscular performance is dependent on the intracellular levels of several cold-inducible proteins, such as the carrier for glucose uptake into muscles (GLUT-4) and sarcoplasmic or endoplasmic reticulum Ca^{2+} -ATPases, respectively. Massive shivering

also allows a prompt reaction to intraoperative hypothermia or raises the CBT quickly to mount very high fevers (e.g., during renal infections).

Thermotolerance

Membranes

All cells possess the ability to modify the fluidity of the membranes in order to preserve structural and functional integrity in the presence of T changes. In poikylothermic animals, cold exposure increases the percentage of polyunsaturated fatty acids (mono- and polyunsaturated, mostly oleic acid) in phospholipids and the relative amount of phosphatidylethanolamine (PE) (vs. phosphatidylcholine, PC). These changes in the composition of the phospholipid bilayer preserve fluidity, probably because unsaturated fatty acids increase membrane disorder (homeoviscous adaptation). It should be noted that cold adaptation causes similar changes in the composition of the mitochondrial membrane in BAT with an increase in polyunsaturated fatty acids and PE and with a decrease in cholesterol content. The final effect is probably a greater membrane fluidity. The brain of hibernating animals contains less cholesterol and greater amounts of both PC and PE, mono-unsaturated fatty acids, arachidonic acid, and polar gangliosides.

Enzymes and Cryoprotective Polypeptides

The most impressive effect of hibernation is the drastic reduction of BMR. This effect is obtained by a tissue-specific inhibition of the activity of key metabolic enzymes such as pyruvate dehydrogenase. In some cases, a different phosphorylation state of the enzyme has been involved (6-phosphofructo-1-kinase); in other cases, the nonspecific effect of temperature change on the K_m for the different substrates seems to be sufficient to explain the observed effect. However, there is much more to learn in relation to cryophylic or thermophylic enzymes if, for instance, in cold-adapted microorganisms, isoforms of proteins such as α -amylase or lipase exhibit high catalytic efficiency in the range of 0–30°C.

Both heat and cold exposure causes the tissue-specific expression of heat shock proteins (HSP70, HSP110, and HSP25). They function as molecular chaperones mediating the folding, assembly, or translocation across intracellular membranes of several polypeptides. Heat shock proteins also play a role in protein degradation and are required for repairing the damage resulting from stress. A subzero-tolerance mechanism, which has evolved in organisms adapted to low temperatures, is the expression of extracellular

antifreeze proteins (AFPs). At least five structurally distinct AF (glyco) proteins have been identified; they all inhibit ice crystallization efficiently.

Thermosensitivity

A breakthrough in the field of peripheral thermosensation/thermonociception has been the cloning and characterization of a vanilloid receptor (VR1) present in somatic and visceral sensory neurons, trigeminal nucleus caudalis, nucleus of the solitary tract (NTS), area postrema, and caudal medulla. The cloning of another more broadly expressed receptor form (VR-like) was reported. They both probably belong to a large family of neuronal mechanosensors (responsive to stretch). VR-1 and VR-like (functional nonselective cation channels) can be activated in a narrow range of temperatures (>42°C for VR1 and >54°C for VR-L). VR1 also responds to protons and vanilloids, even when expressed in oocytes. These studies have rekindled interest in a very large number of experimental observations obtained with vanilloids (capsaicin or resinipheratoxin) that were left without molecular explanation. Capsaicin injection in the periphery does in fact stimulate sensory neurons (nociception and thermonociception) and causes the degeneration of sensory neurons after chronic treatment.

There are, however, several reports describing central effects of capsaicin on Ts neurons. Locally applied capsaicin is, for instance, able to increase the firing rate of warm-sensitive neurons and inhibit the discharge of cold-sensitive neurons in the anterior hypothalamus, causing hypothermia. Furthermore, chronic treatments can cause the degeneration of neurons in the anterior hypothalamus. Nothing is yet known about the receptor(s) that could mediate these effects centrally.

Moreover, it is not even clear whether cellular thermosensitivity is caused by the presence of a thermosensor molecule, responding to temperature changes in the same manner as aquaporins respond to osmolarity changes and opsins to light. Studies on the intrinsic thermosensitivity of different ion channels have shown that they are affected differentially by temperature changes. Among the most studied temperature-sensitive ion channels are the *Shaker*-type, voltage-gated potassium channels (Table 2).

The neurochemical features of warm sensory fibers are partially characterized. One interesting aspect, reported by several groups, is the presence of a calcitonin gene-related peptide (CGRP) in thermosensory fibers innervating, for instance, BAT. It should also be noted that cold acclimation causes a major increase in noradrenergic-, neuropeptide Y-, substance P-, and calcitonin gene-related peptide content in nerves in

Table 2 Effect of temperature on the activity (Q_{10}) of various ion channels^a

Name	Parameters	Q_{10} temperature range ($^{\circ}\text{C}$)	System	Source ^b
K ⁺ channels				
Shaker H4	Activation time constant	+3.14 (20 to 4)	<i>Xenopus laevis</i> oocytes	1
	Decay time constant	+7.2 (20 to 4)	<i>X. laevis</i> oocytes	1
Mink	Activation time constant	+4–7	<i>X. laevis</i> oocytes	
Na ⁺ channels				
TTX-sensitive noninactivating Na current	Current amplitude	+4.3–7	Temperature-sensitive rat POA neurons	
Ca ²⁺ channels				
HVA calcium channels	Activation time constant	+10	Dorsal raphe neurons	4
	Current amplitude	+1.7	Dorsal raphe neurons	
L-type cardiac Ca channels	Inward currents	+5.8 (15 to 25)	<i>X. laevis</i> oocytes	5
HAC		+3–4		
Cl ⁻ channels				
From the <i>Torpedo</i> electric organ (slow gate)	Deactivation potential	+40	<i>X. laevis</i> oocytes	3
CFTR	Opening rate	+9.6		2

^aCFTR, cystic fibrosis transmembrane conductance regulator; HAC, human articular chondrocytes; HVA, high-voltage activated; Mink, minimal potassium ion channel; POA, preoptic area; TTX, tetrodotoxin.

^bSource:

¹Nobile, M., Olcese, R., Toro, L. and Stefani, E. (1997). Fast inactivation of Shaker K⁺ channels is highly temperature dependent. *Experimental Brain Research* **114**, 138–142.

²Pusch, M., Ludewig, U. and Jentsch, T. J. (1997). Temperature dependence of fast and slow gating relaxations of ClC-0 chloride channels. *Journal of General Physiology* **109**(1), 105–116, January 1997.

³Mathews, C. J., Tabcharani, J. A. and Hanrahan, J. W. (1998). The CFTR chloride channel: nucleotide interactions and temperature-dependent gating. *Journal of Membrane Biology* **163**, 55–66.

⁴McAllister-Williams, R. H. and Kelly, J. S. (1995). The temperature dependence of high-threshold calcium channel currents recorded from adult rat dorsal raphe neurones. *Neuropharmacology* **34**, 1479–1479.

⁵Allen, T. J. A. and Mikala, G. (1998). Effects of temperature on human L-type cardiac Ca²⁺ channels expressed in *Xenopus* oocytes. *Pflügers Archiv European Journal of Physiology* **436**, 238–247.

this tissue. Vanilloid receptor 1 expression and CRGP production, as well as peripheral thermosensitivity, are under the control of nerve growth factor.

Concluding Remarks

Outstanding athletic performances, obtained in the presence of prohibitive ambient conditions, are reported almost daily by mass media, and every year, as usual, animals deeply modify their metabolism to survive the coming winter. The biochemistry of the cellular/neuronal control of body temperature, however, still escapes our knowledge.

We know that autonomic and hypothalamic centers play a key role in thermoregulation, and several metabolic studies have provided a large body of evidence on the extent of the control of thyroid hormones on heat production in different organs. Further important information may come from ongoing research on the metabolic effects of leptin, from the understanding of the physiological role of uncoupling

proteins, and perhaps from studies on receptors such as VR-1. The perspective looks promising, and emphasis should be placed on the multidisciplinary approach required to solve the biological problem of thermal control.

See Also the Following Articles

Heat Shock Genes, Human; Heat Shock Proteins: HSP60 Family Genes; Hyperthermia; Hypothermia; Temperature Effects; Thermal Stress; Thyroid Hormones.

Further Reading

Allman, J. (1999). *Evolving brains*. New York: Scientific American Library.

Cesare, P., Moriondo, A., Vellani, V., et al. (1999). Ion channels gated by heat. *Proceedings of the National Academy of Sciences USA* **96**, 7658–7663.

Dirig, D., Hua, X. and Yaksh, T. (1997). Temperature dependency of basal and evoked release of aminoacids and calcitonin gene-related peptide from dorsal spinal cord. *Journal of Neuroscience* **17**, 4406–4414.

Three Mile Island, Stress Effects of

A L Dougall and A Baum

University of Pittsburgh, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A L Dougall and A Baum, volume 3, pp 595–597, © 2000, Elsevier Inc.

Stressors at Three Mile Island
Short-term Consequences
The Effects of Venting Unit 2
Long-term Consequences and Issues
Summary

Glossary

<i>Anticipatory stress</i>	Stress experienced as one waits and/or prepares for a stressful event. In some studies, this stress is similar to that experienced once a stressor is experienced.
<i>Disaster</i>	An unusual event of large scope that overwhelms the social system. Disasters can be of natural origin (e.g., earthquake, tornado) or of technological origin (e.g., Three Mile Island, war).
<i>Stress</i>	A psychobiological process initiated by events that threaten, harm, or challenge an organism or that exceed available coping resources. Stress is characterized by psychological, behavioral, and physiological responses that are directed toward adaptation, but can become pathological if they persist or are inappropriate.
<i>Technological disaster</i>	A powerful stressor that disrupts and overwhelms the social system and has a human origin, by either direct action (e.g., sabotage) or failure of technology (e.g., air disasters, industrial accidents). The human causes of these events introduce issues related to blame and control and may cause longer-lasting distress than natural disasters.
<i>Uncertainty</i>	A state characterized by ambiguity or lack of information pertinent to a particular situation or outcome. It is hypothesized that uncertainty is often experienced as stressful and unpleasant. Uncertainty about the extent of exposure to radiation at Three Mile Island contributed to area residents' distress.

One of the most prominent industrial disasters in United States history occurred on March 28, 1979. Failure in the core of Three Mile Island (TMI) Unit 2

resulted in the release of radioactivity into the surrounding community. For days, residents of the surrounding area, as well as people all over the nation, watched as this unprecedented nuclear accident and emergency unfolded. The accident itself was relatively brief, but uncertainty and fear about explosions and radiation release persisted for several days afterward. Authorities were unable to determine whether and to what extent residents were exposed to radiation. Poor management of the situation and conflicting reports made by authorities led to loss of trust and credibility, contributing further to uncertainty and fear among area residents. An evacuation advisory was issued for pregnant women and young children residing within 5 miles of TMI, resulting in a substantial evacuation of nearly 150 000 residents. After the initial accident and 2-week acute emergency period, the residents of the TMI area continued to deal with potentially stressful issues, such as disposal of the radioactive gas, water, and solid waste inside the reactor complex and the possible restart of the undamaged reactor at TMI (TMI Unit 1 had been shut down for routine maintenance when the accident occurred in Unit 2).

Stressors at Three Mile Island

The circumstances surrounding the disaster at TMI set it apart from other disasters such as tornados, earthquakes, and floods. Natural disasters typically cause widespread destruction of property and often involve injury or fatalities. Following the accident at TMI, however, there were no tangible negative effects, only the invisible threat of radiation and worries about its possible long-term consequences. While exposure to low-level radiation has a range of suspected effects, scientific knowledge of its effects is not well characterized. Nonetheless, many people believe that exposure to radiation causes cancers or genetic abnormalities. Area residents were uncertain about how much radiation was released during and after the accident, whether they were personally exposed to radiation, whether their exposure was sufficient to cause negative health outcomes, and whether they would develop cancer or pass on genetic defects. This occurred in the context of the loss of control suggested by analyses of human-caused disasters. The resultant stress experienced by many area residents appeared to be unusually persistent because of the continued presence of sources of new exposure and the intractability of the residents' worries about harm already done by previous exposures. The

primary stressors in the TMI environment were related to uncertainty and the life threat associated with possible consequences of radiation exposure.

Short-term Consequences

State and federal government initiatives after the accident at TMI resulted in several studies of the accident's impact on surrounding communities. In the first month after the accident, most residents reported stress and distrust of authorities, especially those residents who lived closest to the reactors, who were younger, or who were women, particularly women with preschool-aged children. Although levels of stress decreased as time passed, significant amounts of stress were still reported 9 and 12 months later, especially among residents who lived close to TMI, mothers of young children, and people who had evacuated. Residents living close to the TMI plant also reported more stress than residents living near an undamaged nuclear power plant, suggesting that it was not just the fear or distrust of nuclear power that resulted in high stress levels, but also the possible exposure to radiation following the damage at TMI.

The Effects of Venting Unit 2

While a year is usually sufficient for most people to recover from natural disasters, the events at TMI made adaptation difficult during the ensuing 12 months. Radioactive krypton gas had been trapped in the TMI Unit 2 containment building during the accident, and, beginning 16 months after the accident, officials started venting the radioactive gas directly into the environment. While the residents were assured that the amounts of radiation vented were low and safe, the venting occurred nearly every day for weeks and carried with it the threat of new contamination. Before, during, and after the venting, TMI area residents exhibited more symptoms of stress than residents living in a comparison community 80 miles away. Although there was no change in stress levels in the control sample over time, stress levels in the TMI residents decreased during and after the venting. It is possible that TMI residents were exhibiting anticipatory stress before the venting began and, once it occurred, returned to lower levels of responding. Whether this was due to an anticipatory stress response or a possible remedial effect of venting could not be addressed by available data. Regardless, TMI residents continued to experience significant stress following the venting, and by 2 years after the accident many area residents exhibited chronically elevated symptoms of stress.

Long-term Consequences and Issues

Researchers continued to follow TMI area residents during the 6-year period between the accident and restart of the undamaged TMI Unit 1 reactor. Those residents who remained in the area continued to exhibit elevated stress levels as compared to groups of people living in other communities. Negative effects were not limited to measures of self-reported distress. Heightened indicators of physiological arousal (urinary catecholamines and systolic and diastolic blood pressure), more health complaints, greater use of prescription drugs, and poorer task performance were also evident as compared to control groups. Additionally, the incident was not isolated to central Pennsylvania; people all over the country responded to the media attention given to TMI and the fears associated with use of nuclear energy and radiation exposure.

During this time, the issues surrounding the restart of the undamaged TMI Unit 1 resulted in intense legal battles that culminated in a unanimous decision from the U.S. Supreme Court to allow the TMI Unit 1 to be restarted. In October 1985, more than 6 years after the incident at Unit 2, Unit 1 was brought back on-line, despite protests from residents and serious concerns from community groups about the stress it would cause and the damage it would do to the healing TMI communities. Because there were no tangible environmental effects of the TMI Unit 2 accident, the possibility of psychological distress was not considered by the courts to be sufficient grounds for curtailing national efforts for nuclear energy use. However, data indicated that stress levels increased as the restart of Unit 1 approached, and, while stress levels appeared to decrease for some residents after the restart occurred, there was continued evidence for higher stress levels in area residents when compared to groups of people living in other communities. Additionally, evidence of chronic stress responding was still apparent in TMI residents up to 10 years after the Unit 2 disaster.

In addition to the factors described earlier (e.g., distance from TMI, having small children), psychosocial variables played a prominent role in the selective vulnerability of TMI residents to stress. People who reported higher perceived threat and worry over possible radiation exposure and its consequences, as well as loss of trust in officials and loss of perceived control, were also more likely to report higher levels of stress. In addition, those who felt they had fewer people to turn to for support and who tried to deny the situation or use active, problem-focused efforts to try to control an uncontrollable situation were more vulnerable to the effects of stress. In contrast, people

who tried to control their own negative emotions instead of the environment fared much better.

Summary

Although the residents of TMI suffered long-term stress and worry from the incident at TMI Unit 2, their experiences have taught stress researchers and government officials important lessons about the consequences of technological disasters. There is now greater awareness of the significant psychological and physical health effects that can occur following technological catastrophes, even when there is no tangible evidence of destruction. Perceived threat and other psychosocial mediators, such as loss of control and loss of trust in authorities, are sufficient by themselves to promote stress responding and its negative consequences. Additionally, the long-term follow-up of TMI residents has provided insight into the possible determinants of chronic stress responding. Factors such as the experience of unwanted, distressing thoughts about TMI and continued reminders of the possible danger are important determinants of vulnerability to long-term negative consequences. This information has prompted further research into these vulnerability factors as well as aided the development of interventions aimed at helping victims of other disasters.

See Also the Following Articles

Chernobyl, Stress Effects of; Community Studies; Disasters and Mass Violence, Public, Effects of; Nuclear Warfare, Threat of.

Further Reading

- Baum, A. (1990). Stress, intrusive imagery, and chronic distress. *Health Psychology* 9, 653–675.
- Baum, A. and Fleming, I. (1993). Implications of psychological research on stress and technological accidents. *American Psychologist* 48, 665–672.
- Bromet, E. J. and Litcher-Kelly, L. (2002). Psychological response of mothers of young children to the Three Mile Island and Chernobyl nuclear plant accidents one decade later. In: Havenaar, J. M., Cwikel, J. G. & Bromet, E. J. (eds.) *Toxic turmoil: psychological and societal consequences of ecological disasters*, pp. 69–84. New York: Kluwer Academic/Plenum Publishers.
- Bromet, E. J., Parkinson, D. K. and Dunn, L. O. (1990). Long-term mental health consequences of the accident at Three Mile Island. *International Journal of Mental Health* 19, 48–60.
- Davidson, L. M., Baum, A. and Collins, D. L. (1982). Stress and control-related problems at Three Mile Island. *Journal of Applied Social Psychology* 12, 349–359.
- Davidson, L. M., Weiss, L., O'Keefe, M. K. and Baum, A. (1991). Acute stressors and chronic stress at Three Mile Island. *Journal of Traumatic Stress* 4, 481–493.
- Dew, M. A. and Bromet, E. J. (1993). Predictors of temporal patterns of psychiatric distress during 10 years following the nuclear accident at Three Mile Island. *Social Psychiatry and Psychiatric Epidemiology* 28, 49–55.
- Dew, M. A., Bromet, E. J., Schulberg, H. C., Dunn, L. O. and Parkinson, D. K. (1987). Mental health effects of the Three Mile Island nuclear reactor restart. *American Journal of Psychiatry* 144, 1074–1077.
- Dohrenwend, B. P., Dohrenwend, S. N. and Warheit, G. J. et al. (1985). A psychophysiological field study of stress at Three Mile Island. *Psychophysiology* 22, 175–181.
- Goldstein, K. and Martin, J. L. (1981). Stress in the community: A report to the President's Commission on the Accident at Three Mile Island. *Annals of the New York Academy of Sciences* 365, 159–174.
- Hartsough, D. M. and Savitsky, J. C. (1984). Three Mile Island: psychology and environmental policy at a crossroads. *American Psychologist* 39, 1113–1122.
- Houts, P. S., Cleary, P. D. and Hu, T. (1988). *The Three Mile Island crisis: psychological, social, and economic impacts on the surrounding population*. University Park, PA: The Pennsylvania State University Press.
- MacGregor, D. (1991). Worry over technological activities and life concerns. *Risk Analysis* 11, 315–324.
- Prince-Embury, S. and Rooney, J. F. (1988). Psychological symptoms of residents in the aftermath of the Three Mile Island nuclear accident and restart. *Journal of Social Psychology* 128, 779–790.

Thymus

M S Vacchio

National Cancer Institute, National Institutes of Health,
Bethesda, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3,
pp 598–604, © 2000, Elsevier Inc.

Thymic Architecture and Components
Activities of the Thymus
Glucocorticoids and Thymic Function
Effects of Stress on the Thymus

Glossary

<i>Negative selection</i>	Selective inactivation of those thymocytes that express a T-cell receptor with a high avidity for self-peptide presented by self-histocompatibility complex-encoded molecules and therefore are potentially harmful to the host.
<i>Positive selection</i>	Selective survival and expansion of those thymocytes that have successfully rearranged T-cell receptor genes and express a T-cell receptor with a moderate affinity for self-histocompatibility complex-encoded molecules.
<i>Thymocyte</i>	An immature T cell undergoing the processes of differentiation and selection in the thymus prior to export to the peripheral immune system.

The thymus is a bilobed organ situated in the anterior superior mediastinum of mammals. The major function of the thymus is the differentiation and selection of T lymphocytes that will populate the peripheral immune system and function as part of acquired immunity. The thymus is critical for normal immune function and is highly susceptible to the effects of stress.

Thymic Architecture and Components

As early as the late 1800s it was thought that the function of the thymus was important to the immune system. However, it was not until 1961 when Jacques Miller demonstrated that thymectomy resulted in loss of graft rejection and compromised immunity to bacterial infection that this organ's place in the immune system began to be defined. The thymus gives rise to T lymphocytes, which are critical for normal immune function. Differentiation from immature precursors

to mature T cells occurs in the thymus in a highly organized manner that will be discussed later in the section describing thymic function. Various aspects of thymus function have been highly characterized in mice and rats; however, the majority of these observations are also characteristic of that which occurs in humans.

The thymus consists of two main areas. The cortex normally comprises about 80% of the thymus and is densely packed with thymocytes that visually obscure the cortical stromal elements. In the medulla, the cellular density of thymocytes is considerably less and the medullary stroma is more easily visible.

The thymus differentiates from the third and fourth pharyngeal pouches in mammals and is formed from ectoderm, mesoderm, and endoderm, all of which are required for normal development. The cellular components generated by these tissues comprise the thymic stroma, composed of a diverse array of epithelial cell subsets. The origin and function of these various thymic epithelial subsets are still highly debated. Other stromal elements include macrophages and dendritic cells that are of hematopoietic origin. The cells that ultimately differentiate into T cells are derived from pluripotent hematopoietic stem cells that migrate into the thymus from either the fetal liver or bone marrow. Evaluation of the T-cell component in the adult thymus reveals significant heterogeneity. Differentiative stages of thymocytes can be determined via cell surface markers, the most characteristic of which are the expression of CD4 and CD8 and T-cell receptor (TCR) (Figure 1). The earliest thymocytes compose only a small fraction of the total T-cell number (1–5%) and express Thy-1, but do not express CD4, CD8, or TCR (double negative, DN). The population of highest frequency, composing 80–85% of thymocytes, expresses both CD4 and CD8, along with moderate levels of TCR (double positive, DP). The most mature thymocytes, poised to emigrate to the peripheral immune system as functional T cells, are the single positive (SP) cells that express either CD4 (~10%) or CD8 (~5%), along with high levels of TCR (comparable to that observed on peripheral T cells).

Activities of the Thymus

T-Cell Differentiation and Selection

The major activity of the thymus is to harbor T cells during their differentiative stages and to select functional T cells that can be exported to the peripheral

immune system. Stem cell entry occurs about gestational day 11 in the mouse (gestation is approximately 20 days). Initially, colonization occurs by penetration of the thymic capsule, although later, the influx of stem cells occurs through the high endothelial venules (HEV) that enter the thymus at the cortico-medullary junction. Stem cells migrate out toward the subcapsular cortex where they give rise to the first thymocytes around gestational day 15. The earliest thymocytes that arise are the DN cells. Rearrangement of T-cell receptor genes to generate receptors of unique antigen specificity, similar to that which occurs in B cells for the generation of immunoglobulin, begins early, and by day 15 of gestation the first cells expressing the TCR $\gamma\delta$ chains ($\gamma\delta$ T cells) appear. These cells migrate out of the thymus rapidly and, for the most part, are replaced with other thymocytes expressing the conventional TCR α and β chains. Rearrangement of the TCR β chain genes significantly precedes rearrangement of TCR α chain genes, and generation of a functional rearranged and expressed β chain is required for progression to the next stage of T-cell development. Once the β chain is expressed on the cell surface as a component of the pre-TCR complex, signals associated with this expression turn off any further rearrangement of the β chain, allowing extensive proliferation and differentiation to DP thymocytes, while turning on rearrangement of the TCR α chain genes. Successful rearrangement of the TCR α locus results in moderate level expression of TCR on these DP thymocytes. It is at this point that thymocytes undergo rigorous selection processes (Figure 1).

One might ask why thymocytes are sequestered in the thymus during differentiation and selection. The answer to this question may reside in the manner by which T cells recognize antigen via their TCR. Unlike immunoglobulin on B cells, which bind a three-dimensional structure on an antigen, the TCR binds processed fragments of antigen that are located in a groove on self major histocompatibility encoded molecules (MHC). Therefore, TCR recognition of antigen is constrained by the fact that the TCR must have affinity not only for the foreign peptide but the entire peptide/self-MHC complex. Because the specificity of the TCR is generated randomly during rearrangement of the TCR gene fragments, it is easily understandable that a large portion of T cells may generate receptors that have little or no affinity for self-MHC and therefore would be useless to the immune system. However, the specificity of a randomly generated TCR may have a very high affinity for self-peptide/MHC complexes and would be considered autoreactive and detrimental to the organism. It is unclear exactly what percentage of DP thymocytes express nonfunctional (as deemed by lack of affinity

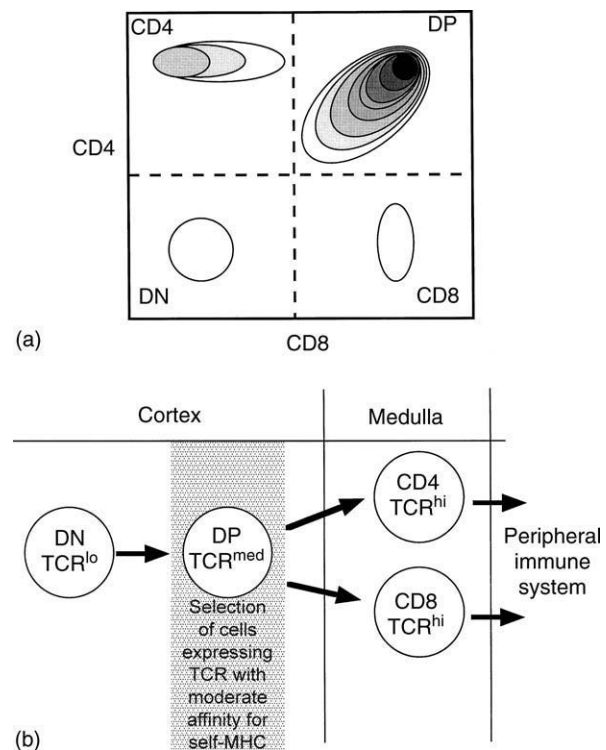


Figure 1 Thymocyte differentiation and selection. (A) Schematic diagram of thymocyte populations expressing CD4 and CD8. The earliest thymocytes (1–5%) express Thy-1 but not CD4, CD8, or TCR (double negative, DN). Double positive cells (DP) compose 80–85% of thymocytes and express both CD4 and CD8 along with moderate levels of TCR. The most mature thymocytes are the single positive (SP) cells that express either CD4 (~10%) or CD8 (~5%) along with high levels of TCR. (B) Progression and localization of thymocyte differentiation. DN cells, localized in the subcapsular cortex, differentiate to become cortical DP expressing detectable levels of TCR on their cell surface. At the DP stage, thymocytes undergo selection to eliminate cells expressing autoreactive receptors and enrich for cells expressing a TCR with moderate affinity for self-MHC. These positively selected cells differentiate further to express either CD4 or CD8 with high levels of TCR on their surface. These medullary thymocytes go on to populate the peripheral immune system.

to self-MHC) versus autoreactive TCRs, but it is clear that greater than 90% of thymocytes generated never leave the thymus. Therefore, a major function of the thymus is to cautiously screen developing thymocytes within a controlled environment to obtain those thymocytes of greatest functional potential while eliminating unnecessary clutter and the potential for damage in the peripheral immune system. It is thought that thymocytes require TCR-mediated signals to prevent programmed cell death. Thymocytes with nonfunctional TCRs do not receive this survival signal and therefore undergo apoptosis within a limited time frame. DP thymocytes in the cortex that express a TCR of high avidity for self-peptide/MHC complexes undergo a process termed negative selection, in which

elimination occurs via activation-induced cell death. These apoptotic cells are rapidly engulfed by macrophages and cleared from the thymus in a highly efficient manner so that, despite the fact that > 90% of thymocytes die in the thymus, one must look very hard to detect dying cells. Finally, those DP thymocytes that express receptors with a moderate avidity for the self-MHC undergo positive selection and differentiate into SP thymocytes with upregulated levels of TCR. Those thymocytes that recognize MHC class I differentiate to become CD8 SP cells and those that recognize MHC class II differentiate into CD4 SP cells.

Production of Thymic Factors

In addition to the role that the thymus plays in T-cell differentiation and selection, the thymus also produces a myriad of factors that can act in either a localized or a secreted manner (Table 1). Much work has focused on the thymic hormones produced by thymic epithelial cells. These were initially characterized in crude thymic extracts and have subsequently been purified and characterized. The major peptides include thymopoetin, thymic humoral factor, thymosins, and thymulin, although many less well-characterized peptides exist. The functions of these peptides are related to the maturation of thymocytes as measured *in vitro* by phenotypic changes in bone marrow after coculture and/or related to the modulation of mature T-cell function as determined postinjection of the peptide. The majority of these peptides are detectable in the circulation, and levels are influenced by a variety of factors, including thyroid hormones, pituitary hormones, prolactin, growth hormone, and stress. Thymic humoral factor has been shown to enhance lymphocyte proliferation and IL-2 production *in vitro*. Thymopoetin binds cell surface receptors on immature thymocytes and mature T cells and appears to enhance early T-cell differentiation *in vitro*. The thymosins are actually a group of peptides, the most well characterized of which are thymosin $\alpha 1$ and thymosin $\beta 4$. Thymosin $\alpha 1$ induces the

expression of Thy-1, CD5, and CD8 on thymocytes and enhances mitogenic, cytokine, and antibody responses of lymphocytes *in vitro*. Both thymosin $\alpha 1$ and $\beta 4$ also increase terminal deoxynucleotidyltransferase (TdT) expression. However, of all the thymic factors, only the production of thymulin appears to be truly thymus restricted as evidenced by the observation that circulating thymulin levels decrease postthymectomy, whereas levels of the other thymic hormones are unaffected. Thymulin, which is only biologically active when complexed to zinc, is expressed by thymic epithelial cells and binds to receptors expressed on the surface of mature T cells, natural killer (NK) cells, and immature thymocytes. Functionally, thymulin can also induce expression of CD3 and CD8 on bone marrow cell and augment responses of mature T cells *in vitro*. However, while these thymic hormones can induce thymocyte changes *in vitro* or alterations in T cell function *in vivo*, a major issue that remains to be established is whether their *in vivo* function is primarily on tissues distant from the thymus or whether they also play a critical role in T-cell development *in vivo*.

While the thymic hormones are secreted and act on other target tissues, there are numerous thymic factors that most probably act strictly in a localized manner and are potentially involved in T-cell maturation (Table 1). Cytokines are produced primarily by the lymphocytes, dendritic cells, and macrophages of the thymus. Information obtained from mice in which expression of various of these cytokines and/or their receptors was eliminated by homologous recombination has contributed significantly to mapping their relative importance to T-cell differentiation and selection. For example, while the absence of IL-2 and IL-4 appears to have relatively little effect on T-cell differentiation, IL-7 is particularly critical to progression beyond the DN stage.

Expression of neuroendocrine hormones also occurs in the thymus in cells of both epithelial and bone marrow origin. Most notably, T cells synthesize corticotropin-releasing hormone (CRH) as well as proopiomelanocortin (POMC), the products of which include adrenocorticotrophic hormone (ACTH) and β -endorphin. Populations of thymic epithelial cells have been shown to express the steroidogenic enzymes, P450_{sc} and P450_{c11}, involved in synthesis of pregnenolone and corticosterone, respectively, and steroidogenesis occurs at levels sufficient to influence T-cell differentiation and selection. In the neuroendocrine system, CRH induces the secretion of ACTH that modulates glucocorticoid synthesis. While under normal circumstances, the production of these substances in the thymus does not occur at levels sufficient to significantly alter hypothalamic-pituitary-adrenal (HPA) axis

Table 1 Factors produced by the thymus

Thymic hormones	Cytokines	Neuroendocrine hormones
Thymopoetin	IL-1	ACTH
Thymosin $\alpha 1$	IL-2	CRH
Thymosin $\beta 4$	IL-4	GR
Thymic humoral factor	IL-6	PRL
Thymulin	IL-7	Somatostatin
	TNF α	Arginine vasopressin
	M-CSF	Corticosterone
	IFN- γ	

activity, it is interesting to speculate that the presence of these components within the thymus may lead to the paracrine regulation of localized glucocorticoid synthesis.

Glucocorticoids and Thymic Function

It has long been known that the culture of thymocytes with pharmacological levels of glucocorticoids can induce apoptosis, characterized by nuclear condensation, membrane blebbing, and DNA fragmentation. This sensitivity to glucocorticoids is the result of expression of the type II glucocorticoid receptor (GR). The GR has a 10-fold lower affinity for cortisol/corticosterone than the type I glucocorticoid receptor, also known as the mineralocorticoid receptor (MR), and is hence only activated in the presence of elevated levels of glucocorticoids. Cortical DP thymocytes are highly sensitive to exposure to glucocorticoids, whereas medullary thymocytes (CD4⁺ and CD8¹ SP) and cortical DN thymocytes are relatively resistant, despite expression of the GR in all of these populations. Alterations in sensitivity can be explained by concomitant changes in Bcl-2, a mitochondrial membrane protein that can rescue cells from apoptosis. Expression of Bcl-2 is high in DN thymocytes and decreases as cells differentiate to the DP stage. However, Bcl-2 levels increase postselection and are again high in SP thymocytes, the stage at which cells become relatively resistant to corticosteroids.

While the significance of GR expression in the thymus remains unclear, recent work has shed some light on this question. Enzymes responsible for glucocorticoid synthesis have been identified in thymic epithelial cells and have been shown to be active during the neonatal period when corticosteroid synthesis by the adrenals is low. Inhibition of glucocorticoid synthesis in fetal thymic organ cultures results in abnormal T-cell differentiation and selection. Similar findings were reported in genetically engineered mice in which GR expression had been significantly decreased due to expression of the antisense GR mRNA. Analysis of the thymus in these mice revealed decreased numbers of thymocytes and decreased viability of DP thymocytes in the absence of functional GR. These data suggest that physiological levels of locally produced glucocorticoids are critical for normal T-cell differentiation and selection.

Effects of Stress on the Thymus

Involution

Exposure of mammals to changes in physical, chemical, or emotional conditions results in the generation

of a “stress” response. Stress results in activation of the HPA axis, causing the release of CRH from the hypothalamus. CRH acts on the pituitary gland to facilitate synthesis and release of ACTH, which in turn induces the release of corticosterone from the adrenals. The elevation of corticosteroids in response to stress is rapid and dramatic. As discussed earlier, DP thymocytes are exquisitely sensitive to the elevation of glucocorticoids due to the expression of GR, therefore making the effects of stress on the thymus particularly evident. As early as the 1920s, Jaffe observed that the thymus involutes in response to stress and that adrenalectomy causes hypertrophy. This observation implicated adrenal products as responsible for the induction of thymic involution. However, while the correlation existed, it was not until significantly later that it was demonstrated that injection of corticosteroids mimicked stress-induced thymic involution by eliminating the majority of DP thymocytes. Furthermore, involution in stressed animals could be blocked by adrenalectomy or by treatment with the GR antagonist RU486.

Other Effects of Stress

While the most obvious effects of stress on the thymus is the dramatic decrease in size, attributable to the induction of apoptosis in DP thymocytes, glucocorticoids have a diverse range of effects other than induction of apoptosis. Elevated levels of glucocorticoids can suppress cytokines such as IL-1, IL-2, IL-3, IL-6, IL8, IL-12, G-CSF, GM-CSF, TNF- α , and IFN- γ and can induce the expression of various cytokine receptors. Some of these cytokines are expressed in the thymus and have been postulated to play roles in thymocyte differentiation. Glucocorticoid-induced downregulation of class II molecules on thymic antigen-presenting cells could prevent or hinder selection of CD4¹ class II-restricted T cells. Glucocorticoids released during stress are also well known to alter lymphocyte distribution, potentially due to alterations in the expression and/or affinity of adhesion molecules. Furthermore, exposure to glucocorticoids *in vitro* can induce the proliferation of thymic epithelial cells. All of these changes could significantly affect thymocyte function, namely the selection of the T-cell repertoire, during stress (Figure 2).

In addition to increased levels of glucocorticoids, stress-induced elevation of other hormones could also potentially affect thymic function. Stress results in increased levels of ACTH, growth hormone (GH), and prolactin (PRL), receptors for which are expressed in the thymus by either thymocytes or stromal elements. The importance of GH and PRL to normal thymic function is illustrated in two strains of hypopituitary

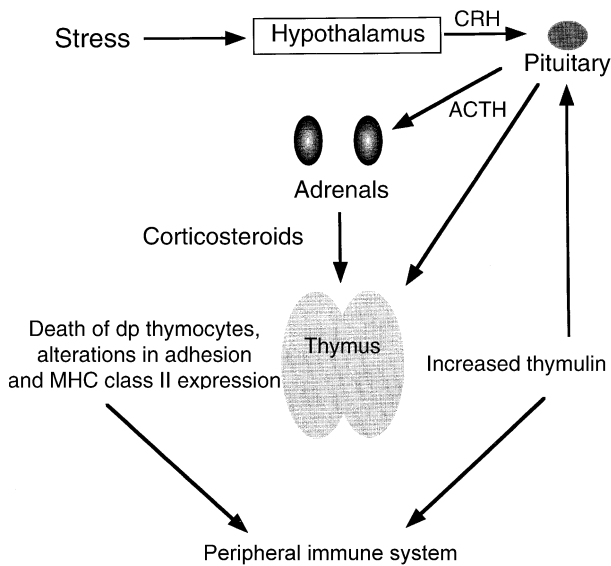


Figure 2 Schematic diagram of the effects of stress on the thymus. Stress results in activation of the HPA axis with release of CRH from the hypothalamus, thereby inducing the release of ACTH, GH, and PRL from the pituitary. The action of ACTH on the adrenals results in elevated levels of glucocorticoids that induce the death of DP thymocytes along with alterations in cell adhesion and MHC expression. Elevated ACTH, GH, and PRL can, among other potential effects, increase the production of thymulin, which can feedback on the pituitary and augment immune responses in the periphery.

mice (Snell-Bagg and Ames dwarf) that lack the acidophilic pituitary cells that secrete these hormones, resulting in small thymuses and few T cells in the periphery. These symptoms can be reversed to some degree by the transplantation of ectopic pituitary tissue. The impact of elevation of these hormones during stress has not yet been fully clarified. Both GH and PRL have been implicated independently in the promotion of thymocyte development and the augmented release of thymulin. The secretion of thymulin is also enhanced by ACTH. *In vivo* and *in vitro* studies of thymulin suggest that this peptide acts to enhance proliferation and cytokine secretion of T cells to antigen in the peripheral immune system. In addition, the secretion of thymulin may provide feedback to the HPA axis, acting on pituitary function (Figure 2). It has been shown that thymectomy results in degranulation of pituitary acidophilic cells, and implantation of neonatal thymic tissue can induce pituitary levels of luteinizing hormone and follicle-stimulating hormone.

Collectively, stress can induce multiple effects other than induction of thymocyte apoptosis. However, the contributions of many of these factors to stress-induced alterations in thymic function have been obscured by the overwhelming amount of thymocyte death during stress and have not yet been fully addressed. Stress-induced involution has been shown

to be reversible in that relief of the stress results in repopulation of the thymus in mice and rats. However, because of the complexities of the system, it is not understood if there are long-term effects of stress on the thymus. For instance, despite the fact that MHC class II levels recover to normal levels post-stress, what effect does glucocorticoid-induced down-regulation of MHC class II or alteration in cytokine production and/or responsiveness during stress have on the selection of surviving thymocytes prior to recovery? Skewed selection of the repertoire would have a lasting impact on antigen responsiveness in the peripheral immune system. Nor is the functional significance of stress-induced thymic involution understood. One might speculate that in times of stress, the protection of immune status is critical and elimination of huge numbers of T-cell precursors seems, at best, counterintuitive. However, stress comes under many guises, one of which is infection by pathogens. At times in which foreign antigen is present at sufficiently high concentrations that may reach the thymus, it would be prudent to restrict thymocyte selection to avoid inadvertent tolerance induction to pathogens. One might argue further that since high levels of glucocorticoids can alter MHC class II levels and cell adhesion, selection in the thymus would be skewed compared to that which would occur during an unstressed state. Perhaps the induction of cell death in those cells actually undergoing selection, the DP cells, is the best way to prevent this potential skewing. However, this is speculation at best, and a great deal of work in this field remains before the impact of stress on the thymus and its products is fully understood.

See Also the Following Articles

Adrenocorticotrophic Hormone (ACTH); Apoptosis; Beta-Endorphin; Corticotropin Releasing Factor (CRF); Immune Response; Immunity; Lymphocytes.

Further Reading

- Goulding, N. J. and Flower, R. J. (1997). Glucocorticoid and the immune system. In: Buckingham, J. C., Gillies, G. E. & Cowell, A.-M. (eds.) *Stress, Stress Hormones and the Immune System*, pp. 199–224.
- Hadden, J. W. (1998). Thymic endocrinology. In: McCann, S. M., Lipton, J. M., Sternberg, E. M., Chrousos, G. P., Gold, P. W. & Smith, C. G. (eds.) *Neuroimmunomodulation: Molecular Aspects, Integrative Systems, and Clinical Advances*, pp. 352–358.
- Robey, E. and Fowlkes, B. J. (1994). Selective events in T cell development. *Annual Review of Immunology* 12, 675–705.

Sprent, J. (1993). T lymphocytes and the thymus. In: Paul, W. E. (ed.) *Fundamental Immunology*, pp. 75–110.

Spangelo, B. L. and Gorospe, W. C. (1997). Thymic polypeptides and their role as mediators in neuroendocrine-immune communication. In: Buckingham, J. C., Gillies, G. E. & Cowell, A.-M. (eds.) *Stress, Stress Hormones and the Immune System*, pp. 357–372.

Vacchio, M. S., Ashwell, J. D. and King, L. B. (1998). A positive role for thymus-derived steroids in formation of the T-cell repertoire. In: McCann, S. M., Lipton, J. M., Sternberg, E. M., Chrousos, G. P., Gold, P. W. & Smith, C. G. (eds.) *Neuroimmunomodulation: Molecular Aspects, Integrative Systems, and Clinical Advances*, pp. 317–327.

Thyroid Hormones

T J Visser

Erasmus University Medical School, Rotterdam, Netherlands

E Fliers

University of Amsterdam, Netherlands

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 605–612, © 2000, Elsevier Inc.

Regulation of Thyroid Hormone Bioactivity
Effects of Stress/Glucocorticoids on Central Regulation of Thyroid Function
Effects of Stress/Glucocorticoids on Thyroid Hormone Metabolism
Conclusion

Glossary

<i>Arcuate nucleus</i>	Hypothalamic nucleus involved in the regulation of appetite and food intake, with important projections to the paraventricular nucleus.
<i>Cytokines</i>	Mediators in the response of the immune and endocrine systems to inflammation.
<i>Deiodination</i>	An enzymatic reaction by which iodine is removed from the thyroid hormone, resulting in the activation or inactivation of the hormone.
<i>Glucocorticoids</i>	Hormones from the adrenal cortex (cortisol in humans) involved in the regulation of glucose metabolism; their secretion induced by stress.
<i>Low T₃ syndrome</i>	The decrease in plasma bioactive thyroid hormone during nonthyroidal illness and other stress conditions.
<i>Nonthyroidal illness (NTI)</i>	An illness not primarily affecting the thyroid gland that is associated with a decrease in plasma bioactive thyroid hormone.

Paraventricular nucleus (PVN) An important hypothalamic nucleus for the production of hypophysiotrophic factors.

In this article, stress is used simplistically as a general term for pathophysiological conditions associated with a stimulation of the hypothalamus-pituitary-adrenal (HPA) axis that are not caused by pathological abnormalities in the HPA axis itself. These stress conditions include those inflicted by disease, inflammation, injury, surgical trauma, fasting, heat and cold exposure, and emotional stress. In addition to stimulating of the HPA axis, these conditions are often associated with profound alterations in several thyroid-related parameters, which are commonly referred to as the low T₃ syndrome. Stress-induced alterations in thyroid hormone bioactivity result to an important extent from the centrally mediated suppression of thyroid function and the inhibition of the peripheral conversion of the prohormone thyroxine (T₄) to the active hormone triiodothyronine (T₃). Both central and peripheral effects may be mediated in part by the high cortisol levels. This article first gives a brief introduction to the regulation of thyroid hormone levels and bioactivity. Subsequently, it describes in some detail the stress-induced changes in thyroid hormone regulation at the central and peripheral levels and speculates about the role of increased cortisol levels in these changes.

Regulation of Thyroid Hormone Bioactivity

Normally, thyroid function is predominantly controlled by thyroid stimulating hormone (TSH), secreted by the pituitary. In turn, the TSH-producing cells of the pituitary are regulated by various hypothalamic factors, most significantly, in a positive manner, by thyrotropin releasing hormone (TRH) and, in a negative way, by somatostatin and dopamine. TSH

synthesis and release are also controlled by peripheral factors, notably by negative feedback regulation by thyroid hormone and by the inhibitory effect exerted by cortisol (Figure 1).

TRH is a tripeptide with the structure pGlu-His-Pro-NH₂ (where pGlu is pyroglutamic acid), which is produced not only in the hypothalamus but also throughout and even outside the central nervous system (e.g., in the pancreas). It is generated from a precursor protein, which in humans consists of 242 amino acids, including six TRH progenitor sequences Baa-Baa-Gln-His-Pro-Gly-Baa-Baa (where Baa is basic amino acid Arg or Lys). Enzymatic cleavage at Baa residues and further processing of the Gln-His-Pro-Gly intermediate by N-terminal cyclization and C-terminal amidation result in the generation of mature TRH.

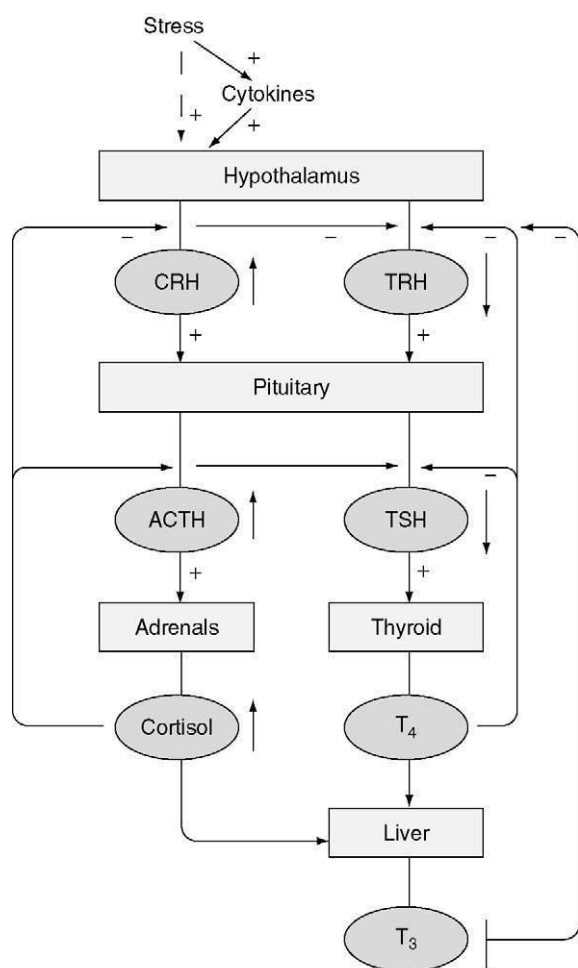


Figure 1 Stimulation of the hypothalamus-pituitary-adrenal axis by stress, the importance of cortisol in mediating the suppression of the hypothalamus-pituitary-thyroid axis, and the inhibition of peripheral T₄ to T₃ conversion. ACTH, adrenocorticotrophic hormone; CRH, corticotropin releasing hormone; T₃, triiodothyronine; T₄, thyroxine; TRH, thyrotropin releasing hormone; TSH, thyroid stimulating hormone.

Hypothalamic TRH stimulates not only TSH but also prolactin and, in some conditions, growth hormone secretion from the anterior pituitary by acting on a common (type I) TRH receptor expressed by thyrotrophs, lactotrophs, and somatotrophs. Another (type II) TRH receptor has been characterized in the brain. TSH is an ~30-kDa glycoprotein consisting of an α subunit of 92 and a β subunit of 118 amino acids. The α subunit is identical to that of luteinizing hormone and follicle-stimulating hormone, whereas the β subunit conveys hormone specificity.

TSH stimulates the production and secretion of thyroid hormone from the thyroid gland. Under normal conditions, including an adequate iodine intake, the adult human thyroid gland largely produces T₄ (3,3',5,5'-tetraiodothyronine), which has little intrinsic bioactivity and is generally regarded as a precursor of the bioactive form of T₃ (3,3',5-triiodothyronine). Most T₃ is produced by the outer-ring deiodination (ORD) of T₄ in peripheral tissues. Both T₄ and T₃ are inactivated by inner-ring deiodination (IRD) to 3,3',5'-triiodothyronine (rT₃) and 3,3'-diiodothyronine, respectively. Three iodothyronine deiodinases are involved in these processes. The type I deiodinase (D1) is located in the liver, kidney, and thyroid and, surprisingly, has both ORD and IRD activities. Nevertheless, the enzyme in liver, in particular, is thought to be the main site for the peripheral production of T₃ from T₄. The type II deiodinase is expressed in the brain, the pituitary, and, perhaps also to significant extents, the human heart and skeletal muscle. It has only ORD activity and is thought to be important for the local production of T₃ in tissues such as the brain, although the enzyme in the skeletal muscle may contribute to plasma T₃ production. Finally, type III deiodinase (D3) is found in the brain, skin, placenta, pregnant uterus, and various fetal tissues. It has only IRD activity and thus catalyzes the production of rT₃ from T₄ as well as the degradation of T₃. Thus, two deiodinases (D1 and D2) are capable of thyroid hormone activation and two enzymes (D1 and D3) are capable of thyroid hormone inactivation. The relative activities of these deiodinases thus have a major impact on thyroid hormone bioavailability. The deiodinases are homologous selenoproteins, featuring essential selenocysteine residues in their catalytic centers. It is remarkable therefore that both the synthesis and the metabolism of thyroid hormone are dependent on the trace elements iodine and selenium.

The regulation of the bioactivity of thyroid hormone does not depend only on the activities of the different deiodinases; iodothyronines are also inactivated by the conjugation of the phenolic hydroxyl group with glucuronic acid or sulfate. All the enzymes

involved are located intracellularly, as are the receptors mediating the biological effects of T_3 , which are located in the nucleus. Therefore, the metabolism and action of thyroid hormone in the different tissues require the uptake of T_4 and T_3 across the cell membrane. There is much evidence suggesting that the downregulation of the conversion of peripheral T_4 to T_3 during stress may be brought about by the decreased uptake and/or conversion of T_4 in T_3 -producing tissues such as the liver. Because this results in decreased T_3 action and, thus, lower energy expenditure and protein catabolism, this is supposed to be a beneficial adaptation mechanism to sustain the stressful conditions.

Effects of Stress/Glucocorticoids on Central Regulation of Thyroid Function

Hypothalamic Thyrotropin Releasing Hormone

The important physiological role of hypothalamic TRH in the regulation of the pituitary-thyroid axis is clearly established. Immunocytochemical studies in the 1970s and 1980s demonstrated the presence of many (pro-)TRH-containing neurons in the PVN of the rat hypothalamus; in the 1990s the distribution of TRH cells in the human hypothalamus was reported. Although TRH-containing neurons occur in many hypothalamic and extrahypothalamic areas, TRH cells in the PVN appear to be selectively involved in the regulation of the pituitary-thyroid axis. This was first suggested by immunocytochemical studies in hypothyroid rats showing changes in TRH cells mainly in the parvocellular portion of the PVN. The assumption was confirmed by subsequent studies on TRH mRNA in the PVN of hypothyroid rats and by studies involving the endocrine effects of PVN lesions. Thus, TRH cells in the PVN are important for the set-point regulation of serum thyroid hormones, and circadian and circannual variations in TRH content, as observed in the rat PVN, may explain circadian and seasonal variations in serum concentrations of TSH.

Thyroid hormone feedback regulation is probably mediated by the thyroid hormone receptor (TR), which is expressed by TRH neurons in the PVN and by cells in the arcuate nucleus (ARC). The ARC is also one of the major hypothalamic sites expressing D2, which converts the prohormone T_4 into bioactive T_3 . Studies suggest that neuropeptide Y (NPY) cells from the ARC innervate TRH cells in the PVN that project to the median eminence. Therefore, the ARC appears to be one of the pivotal loci for thyroid hormone feedback regulation in the central nervous system.

Stress, Glucocorticoids, and the Central Component of the Hypothalamic-Pituitary-Thyroid Axis

Different types of stress may induce differential changes in the central component of the hypothalamic-pituitary-thyroid (HPT) axis. For example, cold exposure in rats increases TRH mRNA in the PVN, as shown by *in situ* hybridization, and leads to increased serum concentrations of TSH. In contrast, food deprivation in rats decreases TRH mRNA in the PVN and *in vitro* TRH release and is associated with lower serum concentrations of thyroid hormones. In humans, severe illness in the absence of overt thyroid disease, NTI, is associated with a decrease in serum T_3 . Despite low serum concentrations of thyroid hormones, serum TSH in NTI is typically normal or reduced. These observations point to a change in hypothalamic thyroid hormone feedback regulation. In line with this notion are profound alterations in the circadian TSH rhythm and decreased TRH mRNA expression in the PVN of patients with NTI.

Various studies have suggested a role for glucocorticoids in the alterations of the central component of the HPT axis during stress. For example, glucocorticoid excess suppresses TSH in patients and also in animal experiments. Conversely, glucocorticoid deficiency is associated with mild elevations in serum TSH, pointing to a role for glucocorticoids in the regulation of TSH release. The demonstration of glucocorticoid receptor (GR) expression by TRH cells in the rat PVN and the presence of a glucocorticoid response element in the TRH gene suggest that the inhibitory effect of glucocorticoids on TSH secretion involves the inhibition of TRH synthesis in the PVN. This suggestion was confirmed by elegant studies involving the concomitant determination of TRH mRNA and corticotropin releasing hormone (CRH) mRNA in the PVN of rats after bilateral adrenalectomy with or without subsequent treatment with glucocorticoids. After adrenalectomy, there was an expected rise in CRH mRNA, but also of TRH mRNA in the PVN. After corticosterone or dexamethasone administration, both CRH and TRH mRNA showed a major reduction. Therefore, glucocorticoids appear to inhibit TRH cells in the rat PVN, possibly directly via GR expressed by these cells. This may explain why acute bolus injections of dexamethasone suppress basal TSH release, whereas the TSH response to TRH remains intact. After long-term pharmacological glucocorticoid excess, the TSH response to TRH is also blunted. Some, but not all, studies in patients with endogenous hypercortisolism have reported an inverse correlation between parameters of cortisol excess and the TSH response to TRH, suggesting an inhibitory effect of glucocorticoids at the pituitary level (Figure 1).

Complex interactions of glucocorticoids and TRH cells have been reported in *in vitro* systems. The short-term administration of dexamethasone to dispersed hypothalamic cells showed concentration-dependent biphasic effects of dexamethasone on TRH mRNA and TRH cell content. In addition, dexamethasone stimulates TRH and TRH mRNA expression in fetal rat diencephalic neuronal cultures. In contrast to the *in vivo* situation, TRH is expressed by anterior pituitary cells in culture, which is increased further by dexamethasone. The discrepancy between the negative effects of dexamethasone *in vivo* and its positive effects *in vitro* on TRH expression in hypothalamic cells may be explained by the nonphysiological culture conditions and/or the disruption of inhibitory innervation of TRH cells in the PVN. In cultured human placental cells, cortisol at concentrations of 10–100 nmol l⁻¹ inhibits D2 activity. Severe illness is associated with increased serum concentrations of cortisol. In line with decreased D2 activity in the hypothalamus of patients with severe illness is the observation that the hypothalamic concentration of T₃ (but not of T₄) is decreased markedly in postmortem tissue compared with patients with acute traumatic death. Therefore, we may speculate that the hypercortisolism of severe disease is one of the factors leading to a decreased expression of TRH in the PVN, thereby contributing to the persistence of low serum TSH (Figure 1). In addition, the bioavailability of T₃ may be reduced by the decreased activity of D2. In contrast, hypothalamic expression of D2 is increased in rats during starvation, again pointing to the differential effects of stress on the central component of the HPT axis according to the type of stress.

Effects of Stress/Glucocorticoids on Thyroid Hormone Metabolism

Pathophysiology of Peripheral Thyroid Hormone Metabolism

As mentioned earlier, stress conditions such as fasting, illness, and injury are associated with major changes in plasma thyroid hormone levels. Marked decreases in plasma T₃ and increases in plasma rT₃ are observed consistently in these conditions, and plasma T₄ and free thyroxine (FT₄) levels may also be affected in critical illness. In general, the magnitudes of the decrease in plasma T₃ and of the increase in plasma rT₃ are correlated with objective measures of the severity of the disease, such as loss of organ function and fever. It has been suggested that the plasma T₃/rT₃ ratio is a good indicator of the clinical condition of the patient. However, it should be mentioned that the increase in plasma rT₃ is not found in

all diseases; notable exceptions are acquired immunodeficiency syndrome (AIDS) and renal disease. Initially, the stress-induced changes in plasma T₃ and rT₃ were interpreted as reflecting a shift in the peripheral metabolism of the prohormone T₄, that is, to less conversion by ORD to active T₃ and more conversion by IRD to inactive rT₃. Kinetic analysis of T₃ and rT₃ turnover during fasting and illness has confirmed that the decrease in plasma T₃ is due to a decrease in plasma T₃ production with little or no change in plasma T₃ clearance. However, the increase in plasma rT₃ seen in fasting or sick subjects appears to be caused by a decrease in plasma rT₃ clearance rather than an increase in plasma rT₃ production.

Because D1 in the liver is supposed to be an important site for the production of T₃ and the degradation of plasma rT₃, the decrease in plasma T₃ and increase in plasma rT₃ in the low T₃ syndrome may be caused by diminished hepatic D1 activity. Evidence suggests that a decrease in the hepatic content of reduced glutathione (GSH) may contribute to a decrease in D1 activity during fasting because GSH is the most abundant intracellular thiol compound and the activity of D1 is dependent on thiol cofactors. The stress-induced decrease in hepatic D1 activity may also be caused by a decrease in the expression of the enzyme at the transcriptional or posttranscriptional level. In this respect, studies of the effects of fasting on hepatic D1 expression in rats have produced confusing results. This is because fasting in rats is associated with a rapid and marked downregulation of TSH secretion, which results within 2 days in decreases in plasma T₄ and T₃ to hypothyroid levels. Strange enough, D1 gene expression is under the positive control of its own product T₃. Therefore, fasting in rats induces a hypothyroid state, which leads to a decrease in D1 expression. If plasma thyroid hormone levels in fasted rats are kept normal by T₄ replacement therapy, the decrease in hepatic D1 expression is largely prevented.

Even before D1 was identified as a selenoprotein, it was demonstrated that hepatic and renal D1 activities are strongly decreased in rats fed a selenium-deficient diet. It has subsequently been shown that a severely deficient selenium intake in humans may result in modest changes in plasma T₄, T₃, and rT₃ levels compatible with minor decreases in tissue D1 activities. Although selenium stores may diminish under certain clinical conditions, such as long-term parenteral nutrition, it is unlikely that this contributes significantly to the generation of the low T₃ syndrome. Thus, there is little solid evidence to support the hypothesis that the decrease in plasma T₃ and the increase in plasma rT₃ during fasting or other stress conditions are caused by a decreased hepatic D1 activity.

Kinetic modeling of the plasma disappearance of injected radioactive T_4 , T_3 , and rT_3 suggests that the fractional transfer rates of T_4 and rT_3 , more than that of T_3 , from plasma to the rapidly equilibrating tissue pools are decreased during fasting and illness. This is consistent with a decreased uptake of T_4 and rT_3 in the liver and, hence, a decreased intracellular availability of these substrates for conversion by D1. Studies of the uptake of different iodothyronines by isolated rat and human hepatocytes and in the isolated perfused rat liver have provided possible clues for the mechanism of these changes. These studies have produced strong evidence that the uptake of the different iodothyronines by liver cells does not occur by simple diffusion of these lipophilic compounds through the lipid bilayer of the cell membrane but is mediated by transporters that are at least partially ATP and Na^+ dependent. Competition experiments and ATP-depletion studies have suggested that uptake of T_4 and rT_3 and the uptake of T_3 by hepatocytes are mediated by different transporters, where the former is more sensitive to decreases in cellular ATP content than the latter.

The stress-induced decrease in plasma T_3 and increase in plasma rT_3 can thus be explained by a decreased activity of a hepatic T_4/rT_3 transporter with little change in the activity of the T_3 transporter. This could be the result of a decrease in cellular ATP content during fasting and illness, which would cause a much greater inhibition of hepatic uptake of T_4 and rT_3 than of T_3 . Although hepatocyte incubation studies have indicated that relatively small decreases in cellular ATP suffice to cause significant inhibition of T_4/rT_3 uptake, it remains to be established whether similar diminutions in tissue ATP content may occur under pathophysiological conditions *in vivo*. Increased plasma concentrations of inhibitors are perhaps a more likely mechanism for the decreased hepatic uptake of T_4 and rT_3 during fasting and illness. Again, studies using isolated hepatocytes have provided support for this hypothesis, in that they have shown significant inhibition of T_4 uptake by various compounds, the plasma levels of which are elevated during fasting and illness (fatty acids) or specifically in liver disease (bilirubin) or kidney failure (3-carboxy-4-methyl-5-propyl-2-furan propanoic acid, CMPF; and indoxyl sulfate). The clinical relevance of these data is strongly supported by findings that in the presence of serum from patients with NTI, the uptake and metabolism of T_4 by hepatocytes are diminished compared with incubations containing serum from healthy subjects. Furthermore, the inhibition of T_4 uptake and metabolism by NTI serum are strongly correlated with the severity of the disease. Therefore, the decreased hepatic uptake of T_4 and rT_3 appears to be an important

mechanism for the decreased T_3 production and rT_3 clearance during fasting and illness.

Possible Role of Glucocorticoids in the Generation of Low T_3 Syndrome

Most, if not all, conditions characterized by the low T_3 syndrome are also associated with the stimulation of the HPA axis (Figure 1). In inflammatory conditions, this is mediated by the induction of a cascade of cytokines, including tumor necrosis factor (TNF)- α , interleukin (IL)-1, and IL-6. The administration of TNF- α or IL-6 to human subjects results in a decrease in plasma T_3 and in an increase in plasma rT_3 . Although these findings may reflect the indirect effects of these cytokines because their administration also induces NTI, evidence shows that IL-6 is an important mediator for the alterations in thyroid hormone metabolism. Plasma IL-6 levels generally correlate well with disease severity. Furthermore, a strong negative correlation has been observed between plasma IL-6 and T_3 levels in a number of studies in NTI patients. This correlation is improved by including the concentrations of soluble cytokine receptors.

The involvement of the cytokine network in the generation of low T_3 syndrome has also been addressed in multiple studies testing the effects of the administration of different cytokines on various thyroid parameters in rats or mice. The results of these studies have been difficult to interpret because rodents do not appear to be good models for investigation of the mechanisms by which the low T_3 syndrome is generated in humans. As mentioned earlier, the central suppression of thyroid function induced by stress is much more pronounced in rats than in humans. Moreover, transthyretin (TTR) is the main plasma thyroid hormone-binding protein in rodents, in contrast to humans in which thyroxine-binding globulin (TBG) is the major carrier. TTR is a negative acute-phase reactant; that is, its plasma concentrations fall markedly during illness, resulting in much stronger decreases in plasma T_4 -binding capacity in rats than in humans. However, evidence for a possible role of IL-6 in mediating the inhibition of peripheral T_3 production during illness was obtained in mice. In different NTI models, the decrease in plasma T_3 was somewhat less in IL-6 knockout mice than in wild-type controls, whereas no effect of IL-6 deletion was seen on the decrease in plasma T_4 .

Two lines of evidence strongly support an important role for glucocorticoids in the generation of the low T_3 syndrome. First, the administration of glucocorticoids such as dexamethasone and prednisone to human subjects not only inhibits TSH secretion (see earlier discussion) but also results in acute and marked decreases in plasma T_3 and increases in

plasma rT_3 levels, whereas plasma T_4 levels show little change. Second, all stress conditions associated with the low T_3 syndrome are also accompanied by large increases in plasma cortisol levels. In many, but not all, clinical studies of NTI, a strong negative correlation has been observed between plasma T_3 and cortisol levels. That such a correlation has not been found in all studies may be explained by temporal differences in the responses of plasma cortisol and T_3 to stress, with increases in cortisol preceding decreases in T_3 if they are causally related. However, a serious argument against such a causal relationship has been put forward by studies examining changes in plasma hormone levels in patients undergoing surgery under general anesthesia without or with additional epidural anesthesia. Although the surgery-induced increase in plasma cortisol is strongly diminished, but not prevented, by epidural anesthesia, the decrease in plasma T_3 and the increase in rT_3 are not affected. Although these data do not support the hypothesis that cortisol is an important mediator in the generation of the low T_3 syndrome by surgical stress, they do not exclude the possibility that the smaller increase in plasma cortisol under epidural anesthesia is still sufficient to interfere, perhaps in concert with other factors, with the peripheral metabolism of thyroid hormone.

A particularly interesting study in this respect has been conducted in rats, in which the effects of immobilization stress on plasma thyroid hormone levels were examined. This was associated with a 30-fold increase in plasma corticosterone, an $\sim 40\%$ decrease in plasma T_3 , and an $\sim 75\%$ increase in plasma rT_3 , as well as marked decreases in hepatic and renal D1 activities. These stress-induced changes were not seen in animals that had been surgically or chemically adrenalectomized and treated with replacement doses of corticosterone. The mechanism by which glucocorticoids affect peripheral thyroid hormone metabolism has not been established. Evidence suggests that both the uptake of thyroid hormone and D1 expression may be inhibited in the liver and kidney *in vivo*. However, *in vitro* dexamethasone has been shown to stimulate the expression of the D1 and D2 genes in various cell cultures (e.g., trophoblasts, neurons, glia cells, and hepatoma cells), perhaps by stimulating the differentiation of these cells. The physiological relevance of these observations, however, is uncertain.

It should be realized that the low T_3 syndrome may result not only from a decreased production of T_3 and clearance of rT_3 but also from an increased clearance of rT_3 and production of rT_3 . It has been reported that the administration of dexamethasone to human subjects increases plasma rT_3 production, whereas plasma rT_3 clearance is not changed. Finally,

it should be realized that the effects of glucocorticoids on thyroid hormone metabolism depend strongly on the stage of development. In the fetal circulation, T_3 levels are low and rT_3 levels are high. Although low expression of D1 in liver and kidney may contribute to low T_3 syndrome, the major cause is probably the high expression of D3 in the placenta, uterus, and various fetal tissues. Studies in chickens have indicated that the high D3 expression in the embryonic liver is acutely downregulated by glucocorticoids, resulting in a dramatic increase in plasma T_3 . Thus, there are multiple, sometimes opposite, interactions between glucocorticoids and peripheral thyroid hormone metabolism.

Conclusion

We have discussed the decrease in TSH secretion and peripheral T_4 to T_3 conversion (low T_3 syndrome), which occur in various forms of stress. This is believed to be a beneficial defense mechanism, which saves energy and minimizes catabolism. All these conditions are also associated with a stimulation of the pituitary-adrenal axis. Together with findings that the effects of stress on plasma TSH and thyroid hormone levels are mimicked by the administration of glucocorticoids, this suggests that cortisol is an important mediator in the generation of the low T_3 syndrome during stress at both central and peripheral levels. However, studies of the effects of epidural anesthesia, which decreases the response of cortisol but not those of T_3 and rT_3 to surgical stress, suggest that there is not always a causal relationship between stress-induced changes in plasma cortisol and thyroid hormone levels.

See Also the Following Articles

Cytokines; Glucocorticoids, Effects of Stress on; Hypothalamic-Pituitary-Adrenal; Paraventricular Nucleus.

Further Reading

- Bianco, A. C., Nunes, M. T., Hell, N. S., et al. (1987). The role of glucocorticoids in the stress-induced reduction of extrathyroidal 3,5,3'-triiodothyronine generation in rats. *Endocrinology* **120**, 1033–1038.
- De Groot, L. J. (1999). Dangerous dogmas in medicine: the nonthyroidal illness syndrome. *Journal of Clinical Endocrinology and Metabolism* **84**, 151–164.
- Docter, R., Krenning, E. P., Jong, M., et al. (1993). The sick euthyroid syndrome: changes in thyroid hormone serum parameters and hormone metabolism. *Clinical Endocrinology* **39**, 499–518.
- Fliers, E., Wiersinga, W. M. and Swaab, D. F. (1998). Physiological and pathophysiological aspects of

- thyrotropin-releasing hormone gene expression in the human hypothalamus. *Thyroid* 8, 921–928.
- Hennemann, G. and Visser, T. J. (1997). Thyroid hormone synthesis, plasma membrane transport and metabolism. In: Weetman, A. P. & Grossman, H. A. (eds.) *Handbook of experimental pharmacology* (vol. 128), pp. 75–117. Berlin: Springer.
- Jackson, I. M. D. (1995). Thyrotropin-releasing hormone and corticotropin-releasing hormone: what's the message? *Endocrinology* 136, 2793–2794.
- Kühn, E. R., Geris, K. L., Van der Geyten, S., et al. (1998). Inhibition and activation of the thyroidal axis by the adrenal axis in vertebrates. *Comparative Biochemistry and Physiology A* 120, 169–174.
- Lechan, R. M. (1996). Functional microanatomy of the hypophysial-pituitary axis. In: Melmed, S. (ed.) *Frontiers of hormone research* (vol. 20), pp. 2–40. Basel: Karger.
- Wiersinga, W. M. and Boelen, A. (1996). Thyroid hormone metabolism in nonthyroidal illness. *Current Opinion in Endocrinology & Diabetes* 3, 422–427.

Thyrotoxicosis See: Autoimmunity; Graves' Disease (Thyrotoxicosis).

Tornadoes, Stress Effects of See: Community Studies.

Torture

I Genefke, H Marcussen and O V Rasmussen

International Rehabilitation Council for Torture Victims, Copenhagen, Denmark

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by I Genefke, H Marcussen, and O V Rasmussen, volume 3, pp 613–619, © 2000, Elsevier Inc.

Torture in the World Today

Types of Torture

Delimitation

Symptoms and Aftereffects

Rehabilitation Models

Prevention of Torture

Impunity

Conclusion

Glossary

Governmental sanctioned torture Institutional and political torture, in contrast to torture by criminals, torture by youth gangs, domestic violence, and

International Rehabilitation Council for Torture Victims (IRCT)

Rehabilitation and Research Centre for Torture Victims (RCT)
UN
Convention against Torture and Cruel, Inhuman or Degrading Treatment or Punishment
UN Vienna Declaration

employer-committed harassment opposed by governmental laws.

A group founded in 1985 in Copenhagen to disseminate information about torture, its consequences, and the possibility of preventing torture. The IRCT also promotes education and training and raises international funds for the IRCT network, consisting of 98 accredited and globally distributed rehabilitation centers for victims of torture.

Established in 1982, a center in Copenhagen that examines and documents torture and develops diagnostics and treatment of torture victims through research.

Convention (1984), with 33 articles, that defines governmental and political torture in article 1.

Declaration (1993) that refers in chapter 5 to torture as “one of the most atrocious

World Medical
Association's
Tokyo
Declaration

violations against human dignity" and "urges all states to put an immediate end to the practice of torture ... and abrogate legislation leading to impunity for torture."

Declaration (1975) stating that "Torture is defined as the deliberate, systematic or wanton infliction of physical or mental suffering by one or more persons acting alone or on the orders of any authority, to force another person to yield information, to make a confession, or for any other reason."

Torture in the World Today

The definition of torture executed by states or state-connected institutions used in this article is that expressed by the United Nations (UN) Torture Convention, adopted in 1984 and entered into force in 1987. This convention states that torture is present when (1) severe pain or suffering, whether physical or mental, (2) is intentionally inflicted on a person (3) for such purposes as obtaining from him or her or a third person information or a confession ..., (4) at the instigation of or with the consent or acquiescence of a public official or other person acting in a official capacity. The World Health Organization (WHO) introduced in 1986 the concept of organized violence, including "torture, cruel, inhuman or degrading treatment or punishment" as in article 5 of the UN's Universal Declaration of Human Rights (1948).

Imprisonments without trial, mock executions, hostage taking, or any other form of violent deprivation of liberty fall under the category of organized violence.

Democracy and Torture

Article 8 of the Vienna Declaration states that "Democracy, development and respect for human rights and fundamental freedoms are interdependent and mutually reinforcing." From this statement, it could be concluded that a necessity for development is democracy based on the respect for human rights.

In other words, there exists a close correlation between democracy and torture. The more democracy in a given country, the less torture is carried out. The more torture, the less democracy, *inter alia*, as it appears from Amnesty International's regular annual reports.

Torture works against democracy. Torture can even be said to be a very strong tool against democracy, since torture goes against the concepts of democracy, freedom of speech and expression, and the right to political dissent. One purpose of torture is to silence

the tortured and those who think alike. A result of this is that people living under pressure and lawless conditions, and consequently with prospects of violations and torture, more or less constantly find themselves in stressful situations.

Those tortured can be identified as ethnic groups and parts of a population, as individuals as well as groups, and as men and women who hold leadership positions and make a stand for democracy in authoritarian regimes. The latter may be politicians in opposition, student leaders, journalists, leaders of ethnic minorities, union members, or human rights defenders.

The Magnitude of the Problem

Torture ranks as one of the most profound human rights abuses, and the magnitude of this worldwide problem is immense. Torture as well as awareness of torture have increased. Government-sanctioned torture exists in more than 40 UN member states and is carried out sporadically in many more. A worldwide survey by Amnesty International (AI) showed that 150 countries, out of 195 investigated, practiced torture. It is also known from AI's yearly reports that torture may occur sporadically in prisons and by the police, even in civilized and democratic societies. Recently, solid evidence that torture exists even in democratic states has been demonstrated by U.S. military personnel torturing prisoners in the Abu Ghraib prison in Iraq in 2004 as well as by information of abuses and alleged torture by coalition forces in Iraq.

There is not a good methodology to calculate the magnitude of the problem of torture worldwide. The number of torture survivors is probably approximately several million.

Every year, torture forces many thousands to flee from their home country. Globally, there are more than 14.5 million refugees, and an additional 19 million have been internally displaced in their home country. A large number of these people have been tortured, although an exact figure does not exist. Community samples are rare. Jaranson and co-workers found in 2004 prevalence rates of torture ranging from 25 to 69% among selected East African communities in Minnesota. Victims of torture often do not want to report on or tell about their traumatic past, either because of fear of retaliation or persecution or because of the deep feelings of shame that are a result of the torture.

Types of Torture

The torture process typically starts with the arrest of the victim, usually at night, with a formidable display

of power and unnecessary use of violence. The softening phase, which often follows, usually consists of a couple of days and nights of unsystematic violence with beating, kicking, and other humiliations. After this, systematic torture starts when the torturers explore the weak spots of the victim to make him or her break down. The aim is to not break down the victim too quickly.

Systematic torture can be conducted in both physical and psychological forms, usually performed at the same time, which aim for a long-lasting destruction of the physical and psychological well-being of the victim. Sophisticated torture methods can cause destruction of the identity and self-respect of human beings while allowing the perpetrators to claim that the victims were never exposed to torture. A new science has developed, and modern day torture is practiced in many countries with the assistance of medical doctors and psychologists.

Physical Torture

The following is a summary of physical torture methods. Electric shocks are applied to the most sensitive areas of the body, or the victim is suspended for hours by his or her arms or by a leg. The head can be forced under water until the victim is about to suffocate, or the skin is burned by cigarettes or red-hot iron rods. The victims are beaten systematically, typically under the feet until the soles are badly damaged. Sexual offenses are common. Women in particular are attacked as sexual objects, and men are harmed in their ability to function as men. Trained dogs can be used for direct attacks or for rape of both men and women. Mock executions bring the individual to a loss of reality and a nightmarish state of almost suspended animation. The situation during the victim's detention is further worsened by filthy food and drinking water. Freedom of movement is limited, and prisoners are closely packed in small cells and thus forced to take turns sleeping. Sanitary conditions are extremely poor, and any request for visiting the toilet will often be turned into a pretext for torture.

Psychological Torture

Psychological torture takes place at the same time as physical torture: for instance, deprivation of sleep, blindfolding, and lack of human contact are forms of psychological torture. These methods leave the victim with a deep sense of helplessness, fear, and extreme stress and may cause hallucinations. The victim can be totally isolated for months or years, and during that time the victim does not know what is going to happen. The victim's family will often have no knowledge of his or her whereabouts. Psychological

methods of great social and psychological impact are sexual torture and situations in which victims are forced to be present during torture to others, particularly relatives and children.

Many victims are threatened with having to do or say things that go against their ideology or religious convictions, with purpose of attacking fundamental parts of their identity, such as self-respect and self-esteem. The victims' political and ethical values are particularly attacked by the torturers when they, e.g., force the victims to sing songs that praise everything they fought against.

The attempt to break down the victims' personality often begins at the time of arrest, with the removal of personal belongings, including life-saving drugs and glasses, and replacing them with badly fitting uniforms. Names are replaced with numbers, and victims are instructed to address their jailers with great respect.

It is important to understand that the nature of torture influences even common life situations. During torture, the perpetrators often avail themselves of common objects used by everyone, such as cigarettes, telephone books, knives, needles, water, light, and noise (often the radio is tuned to music), and as a result, the survivors will in the future associate torture with these everyday objects. In such a manner the trauma of torture is specific. The survivor of torture will be reminded of his or her torture on the conscious as well as the unconscious level nearly constantly throughout the day, year in and year out. And this is what the torturer intends: that the recollection of the torture, the reminders should nag and plague the survivor for a long time or even the rest of his or her life. In the treatment process this is counteracted by psychotherapy, among other options.

Delimitation

History

The recognition of symptomatology related to torture was a phenomenon of the 1970s. At that time, there was no systematic medical literature on torture. Subsequently, a series of findings was encountered that were contrary to expectations.

The first detailed, systematic studies of the methods of torture and its immediate aftereffects were presented by Danish medical doctors in 1974–75. The search for forensic medical evidence that torture had occurred had begun. At an early stage, it was concluded that the worst sequelae of torture were psychological. That was the first surprise, confirmed by other international studies: not only is torture

unbearable and extremely painful, but it also stays with survivors and haunts them many years later.

Diagnosing Government-Sanctioned Torture

Physicians have learned much about the documentation of torture based on both physical and psychological sequelae. Examination methods, interview techniques, and data collection have become more sophisticated. Through medical examination, physicians can now detect indications of torture in most of the major organ systems, including the dermatological, cardiopulmonary, gastrointestinal, musculoskeletal, neurological, urological, gynecological, otorhinolaryngological, and ophthalmological systems, as well as in the teeth. Psychological symptoms now fit into the well-defined category of posttraumatic stress disorder (PTSD) as defined by the WHO and the American Psychiatric Association in the *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (DSM-IV). The medical profession's level of sophistication in assessing torture victims has evolved to the extent that the concept of a torture syndrome based on the results of both physical and mental torture has been proposed.

Symptoms and Aftereffects

The Effect of Torture

The mental health consequences of torture are usually more persistent and protracted than the physical aftereffects, although there is considerable overlap of the physical and psychiatric effects. Torturers today are capable of creating conditions that effectively break down the victim's personality and identity and his or her ability to live a full life with other human beings. Deep feelings of guilt and shame often occur after torture.

Feelings of guilt may be caused by the mere fact having survived when friends died while being tortured, or perhaps by having given information that could have harmed friends. Guilt may also be produced by the so-called impossible choice, when victims are forced to choose between, for instance, revealing the names of their friends and seeing members of their family tortured. Regardless of what the victim chooses, the end result is a disaster for which the victim feels responsible. And that is exactly the aim of torture as it is practiced in this day and age.

Psychological studies of survivors of torture at the Rehabilitation and Research Centre for Torture Victims (RCT) have revealed the following symptoms: anxiety, memory gaps, depression, changed personality, frequent nightmares about prison and torture, and difficulty in remembering and concentrating. Fatigue,

headache, and sexual disturbances are also common. These reactions may in many situations be considered normal in ordinary people who have been exposed to something as perverse, cruel, and abnormal as torture. In situations with beatings, studies have suggested an association between head trauma and neuropsychiatric symptoms that could be recognized as cognitive deficits.

The psychological aftereffects of torture can be considered to be the same in all torture survivors, whereas the physical aftereffects of torture depend upon the type of physical torture methods used. Falanga (persistent beatings on the soles of the feet) results in impaired walking, suspension by the arms results in shoulder pain, and so on.

Over the years the examinations of many hundreds of survivors from approximately 100 countries have provided knowledge about the methods of torture and their aftereffects. These studies have made it possible to ascertain that the methods of torture used are the same all over the world. This is a natural consequence of the fact that the aim of torture is the same all over the world: to break down strong persons' identities. The methods are the same, and so are the aftereffects. This knowledge, which is most useful in medical work on torture, was obtained through systematic analysis.

Posttraumatic Stress Disorder in Torture Survivors

Neuropsychiatric symptoms are often difficult to correctly diagnose. The multiplicity of symptoms is great and comorbidity occurs. Psychiatric disorders such as major depression and PTSD are frequently seen in torture survivors. At the Centre for Victims of Torture in Minnesota, nearly 70% of the clients meet diagnostic criteria for PTSD, with nearly all of the clients exhibiting at least one or two symptoms of this disorder. Mollica and Caspi-Yavin then concluded that "the attempts to describe a single 'torture syndrome' ... are generally unconvincing." Finally, they, who have considerable experience working with refugee survivors from southeast Asia, concluded that "medical investigations of torture survivors not only failed to demonstrate a unique torture syndrome, but demonstrated symptoms closely associated with the DSM-III-R diagnosis of PTSD." Consequently, investigators have shifted their focus away from demonstrating the presence of a unique syndrome to establishing the prevalence of PTSD in torture survivors.

However, the association between PTSD and torture is not a simple linear one. Research has shown that although there is an association between torture

and PTSD, different forms of torture produce different PTSD symptoms. More specifically, individuals who experienced isolation or blindfolding, impact torture, and other types of physical torture had a predominance of PTSD intrusion symptoms, whereas individuals who had been sexually tortured described more avoidance phenomena. This may suggest that PTSD in the cases of tortured individuals is not a uniform syndrome.

It should be kept in mind that controversy exists about the applicability of diagnosing PTSD in torture survivors. In countries where torture is routinely practiced, PTSD often is considered as Western ethnocentric and limited diagnostic category that fails to capture the magnitude of torture as a trauma.

Vesti and Kastrup from the RCT concluded in 1995 that a substantial proportion of survivors developed symptomatology similar to that of PTSD but that others did not. This is consistent with the knowledge that there are individual differences in response to severe stress, which are in the early stages of being described. Basoglu et al., for instance, documented that perceived severity but not objective severity of torture was associated with PTSD, anxiety, and depression in torture victims. Impact of the captivity experience on the victim's family was the strongest predictor of PTSD symptoms. Age, personality, previous emotional and physical health, ideological and political commitment, and quality of the post-torture environment can affect the development of PTSD and other symptoms. The presence or absence of social support and the individual's perception of other persons' helpfulness are important variables for the reduction of the probability of full-blown PTSD.

Petersen et al., who reexamined 22 Greeks who had been tortured, found that eight of them met the criteria for the chronic organic psychosyndrome. These criteria included symptoms, experienced daily, of at least three of four types: (1) reduced memory or ability to concentrate; (2) disturbances of sleep; (3) emotional lability, anxiety, and depression; and (4) vegetative symptoms of the gastrointestinal or cardiopulmonary systems.

Some centers in developing countries placed an emphasis on the recognition of torture as a medical issue and the need for a formal account of the symptoms that follow severe violations of human rights. These centers pointed out that no compensation is currently given to victims and that there is no understanding in their populations or among their politicians that sequelae of torture exist. Therefore, the diagnosis of posttorture syndrome is important.

Regarding the existence of a torture syndrome, the International Rehabilitation Council for Torture

Victims (IRCT) acknowledges that influential researchers on torture have shifted their focus away from a unique torture syndrome and includes the psychopathological syndromes following torture as a subtype of the PTSD category. However, the criteria for PTSD are not sufficient for the categorization of the entire picture after torture. The psychological and physical profiles of PTSD and the posttorture state diverge considerably.

Symptoms that are core criteria for the definition of the posttorture state are described in the DSM-IV definition of PTSD as associated symptoms that are not necessary for a diagnosis of PTSD. These core symptoms are survivor's guilt with low self-esteem, changed personality (the continuity, wholeness, and autonomy of the self is negatively affected), physical sequelae without organic substrate, and many physical complaints without corresponding medical findings. Diffuse pain is a crucial element in the physical and psychological profile of the torture survivor.

Further research in this field might suggest that when the extreme traumatic stressor is interpersonal, the posttorture psychological and physical picture cannot be adequately described by the PTSD entity as it is currently defined. Systematic torture constitutes a more fundamental assault on the individual's self and assumptive world than random violence or other forms of extreme trauma, such as natural disasters. Therefore, there is a continuing need for specialized research on the hypothesis of a torture syndrome.

Rehabilitation Models

Today we know how to diagnose and document torture. The basic principles of diagnosis and documentation are given in the Istanbul Protocol, a UN manual on the effective investigation of torture. Also, today we have the knowledge of how to rehabilitate torture victims, and several different rehabilitation models have been developed.

The rehabilitation model most often used is the holistic approach, in which the psychological, somatic, social, legal, spiritual, family, and cultural aspects are taken into consideration. A point of conceptual importance is to consider torture survivors as not having any premorbid psychopathology.

The treatment principles of this approach are (1) to treat physical and psychological symptoms at the same time, (2) to secure the patient's trust and confidence, (3) to respect the individual, (4) to avoid situations that remind the patient of torture, and (5) to inform the patient carefully about examinations. Physical treatment in relation to diagnostic examination can be practiced, and fair results are possible even many years after torture. Repair of pain

is a very important goal for treatment, and the medical profession can in many cases accomplish that goal.

Treatment rehabilitation methods can be individual, couple, family, or group therapy or counseling, which is a method that, to some degree, takes into account a deeper understanding of the individual's experiences, giving relief and understanding and helping the person to solve problems and make choices.

Other forms of therapy include the following: psychological insight therapy, used for the meeting, the initial setting, the emotional phase, reintegration, and the end of therapy; psychological supportive therapy, used to deal with practical matters such as bodily dysfunctions, social matters such as housing and language, and integration matters, for refugees to get a basic understanding of their new country, as well as to provide balance between insight and supportive psychotherapy; special programs for children and families, such as couples therapy, family therapy, individual child therapy, group therapy, and network meetings; and the outreach method for screening and rehabilitation of torture victims. This last model has four points of importance: availability, accessibility, adaptability, and appropriateness. Availability refers to the services being available to meet the victims' health needs, both general and special. Accessibility means that the services must not be too distant or culturally insensitive. Adaptability refers to the likelihood of the project being acceptable to the refugees and therapeutically effective at the same time. Appropriateness is related to the question of whether a project is appropriate to the refugee groups covered by it.

The aim of all rehabilitation models is to facilitate the social functioning of the victim by strengthening adaptive coping mechanisms, discouraging nonadaptive processes, and facilitating access to social opportunities. Furthermore, it aims to facilitate the psychological treatment of traumatic experiences, loss, and bereavement by consoling, comforting, and protecting victims, sharing their experiences and emotions, recognizing their suffering and pain, and supporting them in their expression of grief when the first survival needs are satisfied. The final step is to treat the victims who have developed psychological disturbances.

Counseling is defined as a method that, to some degree, takes into account a deeper understanding of the individual's experiences, providing relief and understanding and helping the person to solve problems and to make choices. The counselor must respect the confidentiality of information given and must be capable of handling situations that are loaded with painful feelings without getting lost, and using them for the benefit of the client. He or she must also be skilled in analyzing problems, giving advice, and

knowing how to help the client arrive at his or her own decisions.

The counseling program is aimed at recruiting persons who are not health professionals such as doctors or psychologists to act as counselors. The program was created because in many countries health professionals are not available in sufficient numbers. However, other very capable professionals, such as teachers, journalists, monks, and nuns, will be able to learn the program and do a good job as counselors – not necessarily at the same level as the trained psychotherapist, but they will still be able to provide valuable help.

The aim of counseling is to increase awareness of experienced traumas and the reactions to them. Furthermore, it aims to provide relief from psychological suffering, to reduce aftereffects of traumatic experiences, and to help the victim regain control of life situations in order to become, once more, a fully integrated member of society.

The majority of torture victims cope with the help of family or community. Some of them, because of the severity of their symptoms, are not able to reintegrate into society and thus need assistance. Assuming the efficacy of treatment, an important challenge is to identify these victims and help in their communities.

Prevention of Torture

Prevention of torture can be differentiated into the steps of primary prevention, directed toward society itself, secondary prevention, directed toward groups within society, and tertiary prevention, directed toward individuals. Primary prevention of torture involves creating awareness of the problem by making torture visible. This can be done by establishing centers and programs for the rehabilitation of victims of torture. Information and education are also very important tools in the endeavor to create awareness.

Secondary prevention is directed toward teaching relevant personnel – military, police, and prison personnel – that torture always is prohibited, even when an order to torture is given.

The rehabilitation of torture victims serves as tertiary prevention against further violations involving torture, as the presence of rehabilitation services renders the problem visible. Torture may induce a belief in violence and loss of confidence in humanity, but rehabilitation can make people believe in humanity again. In the long run, rehabilitation can help the development of democracy and prevent future conflicts in favor of less stressful conditions.

The UN Convention against Torture contains all the necessary provisions to be able to punish perpetrators and to compensate, prevent, control, educate,

and inspect. This instrument was adopted in 1984 and came into force on June 26, 1987, which was proclaimed as the UN International Day in Support of Victims of Torture – a day not only for creating awareness but also for providing moral support to torture victims. It is a symbolic gesture that recognizes the victims and their families and honors those who did not survive the torture. The convention has been ratified by 141 countries, including all of those countries that had troops in Iraq in 2005 and until now.

Support from the UN, many governments, human rights organizations, nongovernmental organizations, and numerous initiatives is important in helping to break through the silence, insecurity, and indifference surrounding the issue of torture and will help make a stand for visibility, for openness, for acceptance among the boards of various foundations, and for the understanding of the necessity for moral rehabilitation of torture victims.

Impunity

The UN Convention states that torture is an offense under the criminal law of the country in which it occurs and that persons who are accused of torture are to be taken to court and, if found guilty, given penalties that take into account the grave nature of the offense. Article 60 of the Vienna Declaration points to the problem of impunity. It states that legislation leading to impunity for those responsible for torture and other grave violations of human rights should be abrogated.

Not much progress has been made in this area, although promising changes in public opinion in Chile and Argentina have led to ongoing legal proceedings against tyrant generals from the time of dictatorships. In several countries where torture takes place or has taken place, the identity of many of the torturers and those responsible for torture is well-known. Still, there are very few examples of these persons being punished, which is problematic for several reasons. First of all, this is troublesome for the moral rehabilitation of the victims of the torture. If the torturers are not punished, the victims will easily be left with the sentiment of being powerless and humiliated. Second, victims of torture go through unnecessary fear and distress when exposed to the risk of encountering their torturers in the streets. Third, impunity creates a twisted set of norms and values in society. Allowing the persons responsible for gross human rights violations to go free while at the same time hitting hard on conventional crime can lead to an increased brutalization of society. The veil of impunity is broken down slowly but surely, and medical professionals, with their ability to

diagnose and document the incidence of torture, have a key role to play in ensuring that torturers face prosecution and that the victims of torture receive proper rehabilitation and compensation.

The work of the Criminal Tribunal in the Hague should be mentioned. The work of the tribunal has shown the need for the establishment of a permanent international criminal court. This initiative can be a future platform for prosecution of those responsible for torture and thereby an important instrument for the eradication of torture and the moral rehabilitation of victims of torture. It is of great importance, however, that the protection and the support of the victims will be ensured by this court.

Conclusion

More than 30 years of professional, medical, and psychological work on torture have taken place. Today we may say that we have enough knowledge about the phenomenon of torture and enough conventions and declarations. Also, the experience available makes it possible to talk about torture and its effects in a more substantial and assured way. Sufficient basic knowledge (i.e., social analysis on a medical and psychological basis) allows us to make strong statements against the practice of torture. What is needed, however, is national and international recognition of the phenomenon of torture as well as recognition of the magnitude of the problem and its transgenerational effects and the necessity of implementation of conventions and declarations.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Posttraumatic Stress Disorder – Clinical; Posttraumatic Therapy.

Further Reading

- Allodi, F. (1991). Assessment and treatment of torture victims: a critical review. *Journal of Nervous and Mental Disease* 179, 4–11.
- Amnesty International. (2000). *Torture worldwide. An affront to human dignity*. New York: Amnesty International USA.
- Basoglu, M. (ed.) (1993). *Torture and its consequences*. New York: Cambridge University Press.
- Basoglu, M., Jaranson, J. M., et al. (2001). Torture and mental health: a research overview. In: Gerrity, E., Keane, T. M. & Tuma, F. (eds.) *The mental health. Consequences of torture*, pp. 35–62. New York: Kluwer Academic/Plenum.
- Dunér, B. (ed.) (1998). *An end to torture*. London: Zed Books.

- Elsass, P. (1997). *Treating victims of torture and violence. Theoretical, cross cultural, and clinical implications*. New York: New York University Press.
- Hougen, H. P., Kelstrup, J., Petersson, H. D., et al. (1988). Sequelae to torture: a controlled study of torture victims living in exile. *Forensic Science International* 36, 153–160.
- Iacopino, V., Ozkalipci, O. and Schlar, C. (1999). The Istanbul Protocol: international standards for the effective investigation and documentation of torture and ill treatment. *Lancet* 354, 1117.
- Jaranson, M. and Kastrup, M. (2004). Developments and progress, 2001–2004. In: World Psychiatric Association (ed.) *Section on the psychological consequences of torture and persecution*. Lavis, Italy: Masson.
- Mollica, R. F. and Caspi-Yavin, Y. (1992). Overview: the assessment and diagnosis of torture events and symptoms. In: Basoglu, M. (ed.) *Torture and its consequences: current treatment approaches*, pp. 253–274. New York: Cambridge University Press.
- Peel, M. and Iacopino, V. (2002). *The medical documentation of torture*. London: Greenwich Medical Media Ltd.
- Petersen, H. D., Abildgaard, U., Daugaard, G., et al. (1985). Psychological and long-term effects of torture: a follow-up examination of 22 Greek persons exposed to torture, 1967–1974. *Scandinavian Journal of Social Medicine* 13, 89–93.
- Quiroga, J. and Jaranson, J. M. (2005). Politically-motivated torture and its survivors: a desk study review of the literature. *Torture* 15, 1–112.
- Rasmussen, O. V. (1990). Medical aspects of torture: torture types and their relation to symptoms and lesions in 200 victims, followed by a description of the medical profession in relation to torture. *Danish Medical Bulletin* 37(Supplement 1), 1–88.
- Somnier, F. E. and Genefke, I. K. (1986). Psychotherapy for victims of torture. *British Journal of Psychiatry* 149, 323–329.
- Torture*, quarterly journal on rehabilitation of torture victims and prevention of torture, 1991–2005.
- Vesti, P. and Kastrup, M. (1995). Refugee status, torture and adjustment. In: Freedy, J. R. (ed.) *Traumatic stress: from theory to practice*, pp. 213–235. London: Plenum Press.

Transcortin See: Corticosteroid-Binding Globulin (Transcortin).

Transport-Related Stress

R G Smart

Centre for Addiction and Mental Health,
Toronto, Canada

© 2007 Elsevier Inc. All rights reserved.

The Natural Stress of Driving
Driver Stress and Driver Behavior
Stress after Motor Vehicle Accidents
Aggressive Driving and Stress
Driving Stress and Road Rage
What Can Be Done to Reduce Stress in Driving?

Posttraumatic stress

Severe stress resulting from a dangerous situation, marked by anxiety, threat, fear, and helplessness; role impairment; and a preoccupation with the situation. This stress lasts more than 4 weeks.

Road rage

Situations in which drivers or passengers attempt to kill, injure, or intimidate another driver, another passenger, or a pedestrian.

Trait stress

Stress that is not situational but is related to ongoing personality and lifestyle factors and is an enduring feature of a person's life.

Glossary

Acute stress disorder Severe stress resulting from a dangerous situation and lasting up to 4 weeks.
Driver stress Stress that is not trait stress but relates specifically to the problems and challenge of driving.

The Natural Stress of Driving

Advertisements for new cars often show drivers racing along an open road, luxuriating in the comfort and phenomenal speed of the car. There are no other cars about, no backseat drivers, and no children

crying, “Are we there yet?” The reality of driving can be very different. Driving is often done in small cars with restricted space and difficult passengers. The environment is frequently noisy because of horns honking and the din of traffic. Attentional overload because of the complexity of modern cars may also create stress. There are often time restrictions: getting to work on time, picking up children, or making it to appointments. Many roads are congested, under repair, or closed altogether. There are strict performance demands – we must pay attention, avoid accidents, and cope with some other drivers’ wrath. Driving in heavy rain and in the winter in northern countries creates more stress.

Drivers are usually immobile and cannot get out, especially on freeways. Driving can also be dangerous because traffic lights change quickly and other drivers tailgate, cut in and out, or express their anger in other ways.

Many roads in or around cities are congested, and the levels of congestion are increasing in most North American cities. Highway building has not kept pace with car ownership. In North America, there are, on average, two vehicles per family, whereas there was only one vehicle per family for the previous generation. People are spending longer times in their cars and making longer commutes to work or school. Hence, they expose themselves to greater stress on the road.

Driving is often seen as stressful and difficult. In a recent Canadian study, 55% of drivers said that some of their driving was stressful and 15.7% said that half or more was stressful. Approximately 75% said that half or more of their driving was on busy roads and approximately 50% said that most or all of it was on busy roads. Considerable research shows that most drivers find driving on congested roads to be more stressful.

Some drivers have physical limitations that make driving more stressful for them. These include people with arthritis, with neurological or serious psychological problems, and who are elderly or disabled. Cars are not well designed to suit the needs of such people, and driving demands are difficult for them to fulfill. Elderly people often stop driving for physical or psychological reasons. However, the loss of driving capabilities often causes more stress and can lead to depression and social isolation for the elderly.

Driver Stress and Driver Behavior

Particular driving events cause stress among drivers, but many drivers have high levels of stress before they get in the car. Trait stress, or ongoing nonsituational

stress, has been much studied and several scales have been developed to measure it. Driver stress is greater in situations in which there is a lot of congestion and in which there is time urgency. Daily hassles contribute to driver stress. Drivers who recently had an illness, personal conflict, or bereavement are more likely to be in serious accidents.

Total driver stress varies because of the driving situation, personal disposition or trait stress levels, and nondriving events such as daily hassles. High levels of stress in drivers cause attention lapses, errors, and traffic violations. In several studies, driver stress has also been related to aggression and to involvement in accidents. However, most driver stress dissipates quickly once the driver gets out of the car.

Several methods of reducing driver stress have been investigated. Commuters who worked at a job with flextime had less driver stress than those who always traveled during rush hours. Drivers on flextime also felt less time urgency. Experimental studies have shown that drivers who viewed computerized highway drives that had more vegetation than human-made structures had less stress. Also, listening to music may reduce driver stress.

Stress after Motor Vehicle Accidents

Many drivers experience long-term stress after an injury in a motor vehicle accident (MVA). This stress can last many months or even years. The proportion of MVA victims having serious postaccident stress varies from study to study. Accident-induced stress is often assessed using the criteria from the *Diagnostic and Statistical Manual of Mental Disorders*. Acute stress disorder (ASD) is diagnosed within 1 month of the accident, and posttraumatic stress disorder (PTSD) is not diagnosed until after stress has lasted more than 1 month. Both involve subjective feelings of anxiety, fear, and helplessness; role impairment; and a preoccupation with the accident and its consequences.

Most studies of ASD and PTSD involve victims who were injured in accidents and assessed in emergency wards. In such samples, the rate of ASD varies from 10 to 40% and the rate of PTSD from 10 to more than 70%. The highest rates have typically been found in children injured in traffic accidents and people involved in litigation after an accident. Usually, victims with ASD are far more likely than others to develop the long-term stress syndromes seen in PTSD. However, some victims do not develop PTSD after having ASD. Most, but not all, studies found little correlation between PTSD and the severity of the injury.

In children, the likelihood of PTSD often relates to the seriousness that they attribute to the accident. Children after accidents are often afflicted with anger, alienation, thought suppression, and intensive memories of the accident. Most studies show that PTSD is much more common among female than male victims. PTSD may be also more common among those with lower socioeconomic status.

Postaccident stress is not found only among the victims of accidents. Parents also experience considerable stress if their children have PTSD. Also, PTSD, guilt and shame have been found among drivers who caused deaths in motor vehicle accidents. In that group PTSD and guilt were associated with the degree of responsibility the driver assumes for the accident and the severity of their punishment.

Post accident stress can last a long time and can also be related to other negative life events. Among PTSD victims, events such as loss of job or income, broken relationships, and serious illnesses were more common than among accident victims without PTSD. Because of legal liability and compensation issues, some postaccident stress may be faked or exaggerated. Some studies have shown that actors can fool evaluators of stress, especially if little corroborative evidence is available.

Aggressive Driving and Stress

Aggressive driving often causes stress for other drivers as well as for the aggressive driver. Aggressive driving typically involves tailgating; cutting in and out of traffic; preventing other drivers from passing; not moving forward on green lights; and shouting, gesticulating, or criticizing other drivers. If there is an attempt to threaten or intimidate other drivers or to damage their vehicle, this is referred to as road rage, but the two terms often overlap.

Angry or aggressive drivers are more likely to be male and young. Scales for measuring trait driver anger have been devised. Drivers who score high on these scales have more anger and anxiety and show less control over their aggression; they have more minor accidents and close calls but probably not more major accidents.

Driver anger is common and probably comes from feelings of stress during driving. Approximately one in four college student drivers reported being angry with another driver once or more per day. They rated stress from other drivers as equal to the stress of college examinations; traffic congestion, road construction, and finding a parking place were seen as less stressful.

Driving Stress and Road Rage

Road rage may appear to be a new cause of stress for drivers. However, it has been around for a long time. Sophocles in his play *Oedipus the King*, written in about 420 BCE, made road rage the reason that Oedipus kills his father. Also, the poet Byron was engaged in several incidents of road rage in the early 1700s. There are other historical cases as well. It is clear that mass media reports of road rage have recently increased in several countries, although one survey shows that incidents of road rage have been decreasing recently.

Road rage has no accepted definition; it has been defined as a situation in which a driver or passenger attempts to kill, injure, or intimidate another driver, another passenger or a pedestrian in a driving incident. Cases have been described in many countries in North America and Europe, and many have led to death or serious injury. A national study in the United States found that 30% of respondents had complained about other drivers; 17% had yelled at other drivers; 3% had chased other drivers; and 1–2% had gotten out of their cars to hurt other drivers, deliberately hit other drivers, or had carried a weapon. A study in Arizona found that 28% of respondents had aggressively blocked other drivers or followed them to retaliate. Approximately 11% always (4%) or sometimes (7%) carried a gun in their car, and hostile driving behavior was much more common for those with guns.

A Canadian study found that almost half of drivers each year were cursed at or had gestures made to them. Approximately 7.2% were threatened with personal injury or damage to their cars. This study and several others found that perpetrators of road rage are more likely to be younger and male. The victims and perpetrators of road rage are often the same people. However, there is a group of frequent perpetrators of road rage that accounts for almost all of the serious cases involving injury.

Road rage can be a very stressful event for its victims, especially if there is an injury. These victims may remember the event for long periods, modifying their driving behaviors to avoid later incidents. Victims of road rage often feel trapped and unable to get to safety. The immediate effects of road rage are fear, anxiety, anger, and annoyance. Victims report being shaken by it and sometimes being confused. Some victims have lingering thoughts and negative emotions after road rage, but most of this disappears in a few weeks unless there is a serious injury. Being the victim of road rage does make some drivers more cautious about driving.

Road rage is often associated with psychiatric morbidity and distress. In several studies, frequent perpetrators of road rage were found to have high rates of psychiatric distress and, more significant, stressful life events. It may be that frequent perpetrators of road rage bring their stress into the driving situation and, because of their high trait stress, are easily upset in traffic, lashing out at other drivers frequently.

Some studies have also shown that victims of road rage have high levels of psychiatric distress (i.e., anxiety, depression, and stress). Being a victim in a serious incident of road rage is likely to be traumatic for many drivers. The direct causal relationship between serious, injurious incidents of road rage and stress levels has not been examined, but such a relationship is highly plausible given what we know about stress after traffic accidents.

Road rage has been found to be more common among drivers who have more stressful driving experiences. Studies in Canada found that road rage was more frequent for drivers who often drove in congested driving conditions, whose driving was mostly done on busy roads, and who drove longer distances. Road rage was also more common for truck drivers; they drive in stressful situations with tight time schedules, and they are often tired.

Bus drivers have been found in several studies to be both the victims and perpetrators of road rage. They are often victimized by angry passengers, and such conflicts often lead to burnout and the desire to leave the job.

What Can Be Done to Reduce Stress in Driving?

Some of the stress drivers experience is trait stress that they had when they entered the car, and some comes from the experiences they have when driving. Reducing trait stress may require drivers to change relationships, work, and life styles and get better control over their lives. Reviewing these changes is beyond our scope here and we address only the stress that is directly related to driving.

Much stress in driving relates to traffic congestion, and most large cities are experiencing greater congestion. Fully automated transportation systems may eventually divert traffic from congested areas. A few cities are attempting to reduce congestion by increasing public transit, limiting rush-hour traffic, charging motorists for using certain roadways, and the like. Whether and how these measures reduce stress to reasonable levels is unclear at this time. More could

be done to allow flextime and reduce rush-hour travel. Roads could be built to be less complex for drivers, perhaps through traffic calming or other engineering changes.

Much of the serious stress from driving derives from involvement in accidents, especially those with injuries. In North America, the rates of fatal and injury-causing accidents have been declining for many years. The reasons for these declines are the greater use of seat belts and air bags, reduction in alcohol-related accidents because of strict enforcement, better licensing programs for young drivers, and engineering changes to make car crashes more survivable. The extension of these policy changes to other geographic areas could further reduce serious accident levels; so could more technical developments in cars and highways to make them safer.

Some driving stress comes from road rage and aggressive driving. Methods for reducing road rage include better education programs and anger management for convicted drivers as well as changes to vehicles. Some vehicle manufacturers such as Jaguar and Mercedes make cars that have radar systems to prevent tailgating, a common cause of road rage. These systems could be made more available. Cars could be redesigned to have voice-activated warnings when drivers raise their voices. Also cars could have lights that signal when drivers are becoming victims of road rage. Further modifications could also be made to track the perpetrators of road rage with global positioning systems (GPSs) and provide warnings to them. Much driver aggression relates to traffic congestion, and anything that reduces it will lower road rage, aggressive driving, and stress levels for drivers.

Acknowledgment

This research was supported by a grant from Auto21, one of the Networks of Centres of Excellence.

Further Reading

- Blanchard, E. B. and Hickling, E. J. (2004). *After the crash: psychological assessment and treatment of survivors of motor vehicle accidents*. Washington, DC: American Psychological Association.
- Hancock, P. A. and Desmond, P. A. (2001). *Stress, workload and fatigue: human factors in transportation*. Mahwah, NJ: Lawrence Erlbaum.
- Hennessy, D. A. and Wiesensthal, D. L. (2001). Gender, driver aggression and other driver violence: an applied evaluation. *Sex Roles* 44(11–12), 661–676.
- Hennessy, D. A., Wiesensthal, D. L. and Kohn, P. M. (2000). The influence of traffic congestion, daily hassles and trait

- stress susceptibility on state driver stress: an interactive perspective. *Journal of Applied Biobehavioural Research* 5(2), 162–179.
- Lowinger, T. and Salomon, Z. (2004). PTSD, guilt and shame among reckless drivers. *Journal of Loss and Trauma* 9(4), 327–344.
- Mather, F. J., Tate, R. L. and Hannon, T. J. (2003). Post-traumatic stress disorder in children following road traffic accidents: a comparison of those with and without mild traumatic brain injury. *Brain Injury* 17(12), 1077–1087.
- Mizell, L. (1997). *Aggressive driving*. Washington, DC: Automobile Association for Traffic Safety.
- Smart, R. G., Asbridge, M., Mann, R. E., et al. (2003). Psychiatric distress among road rage victims and perpetrators. *Canadian Journal of Psychiatry* 48(10), 681–688.
- Smart, R. G. and Mann, R. E. (2002). Is road rage a serious traffic problem? *Traffic Injury Research* 3(3), 183–189.
- Smart, R. G., Mann, R. E., Zhao, J., et al. (2005). Is road rage increasing: results of a repeated survey. *Journal of Safety Research* 36(1), 195–201.

Trans-sexualism

R A Allison

American Medical Association, Phoenix, AZ, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction and Definitions
 Childhood Experiences of Transsexual People
 Factors that Discourage Transition
 The Transition Process
 Hormonal and Surgical Changes
 Health Effects of Stress for Transsexual People
 Religion and the Transsexual Person
 Violence and Fear
 Intimacy and the Transsexual Person
 Aging and Death and the Transsexual Person
 Conclusion

Glossary

Assignment The process society follows to designate a person as a male or a female.

Crossdresser A person who wears the clothing of the opposite sex, but does not self-identify as a transsexual.

Female-to-male (FTM) A person born biologically female, whose gender identity is male, who may undergo medical and surgical changes to confirm her male identity.

Gender dysphoria A person's persistent feeling of discomfort with the gender assigned at birth.

Gender identity A person's inner feeling of self-identification as male or female. Gender identity in some people may be intermediate, neither fully male nor fully female. Some people experience changes in

Gender identity disorder (GID)

Gender role

Male-to-female (MTF)

Reassignment

Transgender people

Transition

Transsexual

Transvestite

their understanding and acceptance of gender identity with the passing of time. A medical term, used as a synonym for transsexualism. GID is listed in the *Diagnostic and Statistical Manual of Mental Disorders*.

The category (male, female, or ambiguous) in which society places a person, based on physical characteristics and behavior.

A person born biologically male, whose gender identity is female, who may undergo medical and surgical changes to confirm her female identity.

Actions taken to change society's designation of a person as a male or a female. Transsexual people, crossdressers, and people of less-fixed gender identity; sometimes used to describe people who live in their desired gender role but do not seek surgical reassignment.

The process a person follows to live in his or her preferred gender; encompasses the physical changes brought about by hormones and surgical procedures, as well as the social changes of experiencing life and relationships in the new gender.

A person who experiences persistent and severe discomfort in the gender assigned at birth and who wishes to permanently live in the opposite gender role, with all the physical and social changes that role implies.

A crossdresser. This is an older term, used in the mid-twentieth century, and is not currently preferred by most people because it has been associated with a strictly sexual motivation.

Introduction and Definitions

We humans choose certain characteristics by which we define ourselves. The earliest and most essential definition is whether an individual is a man or woman, boy baby or girl baby. The distinction is established at birth, and nurseries are furnished in pink or blue. The basis for the distinction is our external appearance. Our genitalia determine the life we are expected to lead. All humankind is divided distinctly into these two groups: except when it is not.

In fewer than 1/1000 children – the true incidence is still not known – a persistent discomfort exists with the individual's physical sex. This discomfort is present from earliest memory. There is no identifiable behavioral influence to produce it. The child grows and matures unremarkably, to all external appearances; but the child's thought patterns and behavioral instincts are those of the opposite gender.

Such people live their childhood in frustration, knowing something is wrong but being unaware of the specific issue. Finally, due to logical reasoning or reading of someone else's experiences, the child or young adult awakens to the truth. The inner conflict is unrelated to behavior (lifestyle) or preference of sexual partner; rather, it is a conflict involving one's core identity. Medical terms such as gender dysphoria or gender identity disorder (GID) are sometimes used to describe the condition, but it is perhaps best known by the popular name, transsexualism.

Childhood Experiences of Transsexual People

Most adult transsexual people can remember feeling different from their same-sex peers. Before they knew the physical differences between boys and girls, and long before they began to experience feelings of sexual attraction, they knew they belonged in the opposite camp. Such knowledge can be terrifying in an environment where conformity is demanded and diversity is rejected.

Gender-variant behavior in children is not uncommon. Tomboy behavior in girls is usually tolerated better than effeminate behavior in boys. Many boys who play with dolls and other girls' toys grow up to become heterosexual men; some grow up to be gay men; and a few never consider themselves men at all, but they may fear for their safety if other children regard them as sissies. Childhood conformity and peer pressure are the first obstacles trans-gender people encounter in their journey toward transition.

Factors that Discourage Transition

Even stronger than a child's peer relationships are the bonds of family. Parents, brothers, and sisters are the most influential people in any child's life. In contrast to the previously held views of some mental health professionals, who theorized that transsexualism results from family dysfunction in early childhood, many transsexual people report normal and loving family relationships. Even so, it is rare for a transsexual child to feel comfortable discussing gender feelings with family. Perhaps the child loves his or her parents and fears hurting them with so startling a revelation; or perhaps the fear of punishment creates a determination to hide the truth. Such fears are not without merit – it is true that many parents are ill prepared to deal with such a revelation by a child.

Psychological stress is a natural result of the child's practice of denial, which many continue well into adulthood. Some children are so affected by the stress that their school progress is impaired. Others find temporary escape in retreat into books and studies. Music or other hobbies may serve as a welcome distraction from their worries. Some even practice denial so strongly that they seek the other extreme; for example, a boy may strive to excel at sports or to engage in high-risk behavior, placing himself in physical danger, to deny his feeling of being a girl. Indeed, some teens and young adults enlist in the military services in an effort to cure such feelings.

Even in childhood, religion may play a large role in the stress a transsexual child faces. Most religions teach absolute truths, and do not tolerate departures from those absolutes. The child is taught that his or her feelings are sinful or evil but that they can be overcome through the faith process. Most children are eager to believe such doctrine, hoping that they can be made normal so their parents and peers will accept them. When the religious experience fails to change a child's gender identity, the child is left with feelings of failure and self-doubt. A significant number of transsexual adults report so much despair as children or teenagers that they considered suicide.

The onset of puberty is a very stressful time for transsexual children. Their dreams of becoming normal members of their desired gender are severely challenged as female-to-male (FTM) people develop breasts and menstrual activity and male-to-female (MTF) youth experience growing to tall stature, beard growth, and a deepening voice. These physical changes require years of treatment through surgery, hormones, and electrolysis to correct – if they can be corrected at all.

The ultimate denial of an individuals' transsexualism focuses on their attempt to live a normal life through marriage and parenthood. In years past, transsexual young adults lacked the information and resources available today. Believing themselves doomed to a life of quiet unhappiness, they pinned their last hopes on marriage and family to cure them of their dysphoria. However, within a few years, they realized this too could not cure their sense of identity. Unfortunately, at this point other people have entered their lives in intimate roles, and any ultimate disclosure and transition will be disruptive to the entire family. Because of this, some people make the decision to delay transition until their children are older; this is the reason we see so many people begin transition after age 40. A decision to delay carries its own dangers; a child who reaches puberty before learning of a parent's transsexualism usually will react more negatively than a child who is made aware at a younger age.

The Transition Process

Eventually the transsexual person reaches a point of understanding – this is the way my life is going to be. Denial, family pressure, religious fervor, and even marriage have all failed to change the person's gender identity. The incredible pressure can be relieved only through acceptance and transition. The process of disclosure and beginning transition is sometimes called coming out.

This is perhaps the time of the most intense stress a transsexual person experiences. Every aspect of life is at risk: family, friends, status, job, and finances. The importance of advance planning becomes crucial. People who felt they would be warmly accepted may be devastated by the rejection they experience. Job loss, although not universal, is still very common. Income is gone, at a time when it is needed more than ever. The expenses of transition can be enormous; the psychological counseling alone is more than some people can afford, and the various surgical procedures may cost tens of thousands of dollars and very few are covered by insurance. Efforts to find a new job during transition, in the new gender role, are made more difficult because a lifetime of documentation – birth certificate, Social Security records, passport, diplomas, credit ratings – are all on record under the old name. Furthermore, a person in early transition still has behavior patterns from a lifetime of socialization in the birth gender. Such behavior may be overcome quickly for some people as they find the freedom to behave naturally as a member of their desired gender, but others struggle for years to shed male mannerisms. When added to a physical

appearance that still reflects the birth gender, the new mannerisms make it difficult for some to blend into society without attracting unwelcome attention.

For some, the stresses of transition prove too great. As a result of economic demand, family pressure, or the hope of a religious cure, some people abandon transition and return to their original gender role. For most, this return is temporary, and the transition process resumes once they are better prepared to cope with the stress.

Hormonal and Surgical Changes

During and after transition, the effects of cross-gender hormone therapy produce results that often are dramatic. FTM people experience a deepening of the voice and the growth of facial hair. The most obvious early effect of hormones for MTF people is breast enlargement. These effects are obvious to friends, family, and coworkers. At this point, transsexual people may feel a sense of relief because they no longer have to hide their identity or their plan to transition.

Of course, hormones do not produce a complete transformation to the new gender. Other physical differences, including skull and jaw structure, cannot be reversed by hormone treatment. Surgical procedures to eliminate prominent male jaw and chin shape, as well as the exaggerated prominence of the brows, can be of great help in allowing MTF people live normal lives without constant exposure. FTM people often seek mastectomy to make life less stressful.

Ultimately, the transition process culminates in sex reassignment surgery (SRS). In this operation, the surgeon removes some of the old sex organs and creates an appearance as close as possible to the desired gender. Most surgeons who perform SRS follow rather strict guidelines of patient selection to minimize the risk of postoperative regret.

Health Effects of Stress for Transsexual People

A number of physical and psychological effects may result from a transsexual person's decisions related to transition. Many effects occur prior to coming out. The stress of keeping hidden a life-changing truth may be a factor in all types of stress-related illness, including high blood pressure, migraine or other headaches, peptic ulcer disease, and inflammatory bowel disease.

The psychological impact of stress often manifests as chronic anxiety or depression. The degree of psychological distress may be so severe that the person cannot function normally in a work or school

environment. Anxiolytic medications, especially benzodiazepines, are often prescribed; even more commonly given are antidepressants. Some degree of depression is extremely common in transsexual people who have experienced rejection or loss.

Suicide, or suicidal gesture, is a continuing concern for many transsexual people. There are three periods when transsexuals are at greatest risk for suicide. The first occurs early in life, when the young person experiences despair from being unlike his or her peers; the stress of keeping such knowledge undisclosed is magnified by the failure of attempts to eliminate the dysphoria. The second period of high suicide risk occurs just prior to the person's acceptance of his or her transsexualism and decision to proceed with transition and is associated with the last attempt to avoid disclosure and to have, to outward appearances, a normal life. As already mentioned, acceptance and transition may relieve this stress and the associated depression.

The third period of high suicide risk occurs after the completion of transition. The person may have high expectations of success in the chosen gender, and some expectations are more likely to be realized than others. Most people find acceptance by society and may continue their careers or begin new ones. Some, however, lose their jobs and are unable to find employment sufficient to pay expenses. Family and friends may continue to reject the transsexual to the point of shunning the person. Perhaps most devastating, the transsexual may be unable to find a life partner whose love and acceptance is unconditional.

The psychological counseling included in a well-planned transition process should address these expectations so that transsexuals can consider how to deal with failure to meet them. Some people, unfortunately, view the counseling process as simply another obstacle to be overcome in order to obtain their desired surgery. The lack of attention to good counseling during transition may increase the likelihood of regret, depression, and even suicide in the years that follow.

Religion and the Transsexual Person

It is accurate to generalize that people whose religious heritage is liberal, or who have no religious affiliation, experience less stress with transition than people who come from a conservative religious background. This generalization can apply to conservative Christianity, and also to Judaism and Islam. Some churches, particularly the Roman Catholic church, have made specific doctrinal statements describing transsexualism as unacceptable. Other Protestant churches may include transsexualism with homosexuality and

issue a general condemnation of all gender and sexual diversity.

People who have been raised in such denominations experience a high degree of stress. Their religion tells them it is a sin to change from one gender to the other, and it threatens them with eternal punishment if they disobey. The church maintains that a person can, through faith, turn away from transsexualism and live a normal life. For many, the church is their major support network; the thought of being cast out is unbearable. Yet their life experience tells them otherwise – they have already experienced failure of their faith to bring about this change to normalcy.

These people are forced to choose between two difficult paths. Either they must continue to live in denial, experiencing the stress of secrecy and eventually reaching the point of having to face the truth, or they must face the truth early, risking probable rejection and becoming disillusioned with their faith. Many such people do find their way to more liberal, accepting churches, but others abandon their faith altogether and may become quite hostile to any form of religion.

Violence and Fear

It is not coincidence that transsexual people are more likely to be the victims of violent crime. Several factors explain this increased risk. Young, aggressive males, the group most likely to show violent behavior, have a high incidence of homophobia. If a homophobic male encounters someone who is visibly gender variant, he may feel his own masculinity is threatened; he may even feel the need to prove his manhood by attacking the queer.

In addition, when a person is not initially perceived as transsexual but is discovered to be so later, there may be anger because of the perception of having been deceived. Well-publicized cases of people such as Gwen Araujo and Brandon Teena, both of whom were murdered under these circumstances, illustrate the dangers of antitranssexual violence.

Such is the reality faced by transsexual people in our society. Most people take a healthy approach to such risks by avoiding situations in which the danger is apparent. Others may avoid all social occasions and relationships; their excessive fear of violent reactions prevents their having a normal social life.

Intimacy and the Transsexual Person

Many transsexual people find themselves alone at the beginning of transition. Either they have not been in a relationship or their spouse has refused to be a part of

their future plans. These people can benefit from the transition counseling process if they are willing to take the time to explore their future intimacy needs. Such exploration can be stressful in itself, however. People in transition learn that gender identity is independent of sexual orientation. An MTF transsexual person, for example, who was attracted to women prior to transition may ultimately seek intimacy with a man or with a woman. The problem comes when the desired partner has to deal with the reality of the transsexual person's past. A relationship with a man is often jeopardized by the man's feelings of homophobia and inability to get past the idea that his partner is not a real woman. Less likely, but also possible, is the failure of a lesbian relationship due to the female partner's similar feelings. It is possible to find love posttransition, but the increased difficulty of doing so is a great source of stress.

A different dynamic occurs when the spouse of a transsexual person chooses to remain in the relationship posttransition. What was initially a heterosexual relationship now is seen by the world as a homosexual one, with all the associated societal stress such couples encounter. The continued relationship becomes quite a strain on the spouse; for example, the wife of a MTF transsexual soon realizes that she cannot change her sexual orientation to lesbian, just as a gay person cannot decide to become straight. Couples who stay together after transition are the rare exception, and in general the two people involved are very mature and are strong and secure in their commitment to one another despite the stress of daily confrontation.

Aging and Death and the Transsexual Person

The subject of aging has only recently been discussed in depth among transsexual people because the number of people completing transition remained small until the 1980s. Now some people from that era are entering their 60s and 70s, and aging and health issues have begun to assume a prominence in their priorities and present unique stress situations.

Many elderly people count on their children or grandchildren for physical support, to attend to their medical needs, and to take them for doctor visits or hospital emergencies. Transsexual people may lack that support because of family rejection. They face the stress of finding other support among friends, and sometimes a network of elder transsexuals may form for mutual support. The idea of loss of independence is stressful for everyone, but especially so for people who may wear breast binders or hair-replacement systems.

The medical needs of the aging transsexual people are often unique. Physicians called to see a transsexual patient in an emergency may not possess the knowledge or the desire to provide appropriate treatment for these unique needs. It may be difficult to find a doctor willing to continue hormone therapy in an older individual or to screen for diseases associated with the person's birth gender.

A case was reported of an FTM person who developed ovarian cancer. He was rejected for treatment by several physicians, who could not accept a person who presented as a male but had ovaries. The patient eventually died from his disease, which had gone untreated unnecessarily. It is also possible (unlikely, but reported) for an MTF person to develop cancer of the prostate, and it is very rare for such people to submit to the usual screening blood tests.

End of life issues need to be addressed in transsexual people with special concern. It is very important that transsexual people have a partner or sympathetic friend to serve as health-care power of attorney. Likewise, it is important for them to have a will specifying how they wish to be buried or cremated. This will avoid the situation in which family members disregard the individual's stated wishes and bury him or her as the birth gender, even many years posttransition.

Conclusion

Transsexualism, more common than once thought, affects thousands of people who are subject to increased stress due to rejection or confrontation from family, friends, employers, religious institutions, and even total strangers. Coping with these stresses on a daily basis may make the transsexual person mature and strong, or it may be devastating. Understanding and acceptance may reduce the constant stress and enable the transsexual person to make valuable contributions to society.

Further Reading

- Boylan, J. F. (2003). *She's not there: a life in two genders*. New York: Broadway Books.
- Brown, M. L. and Rounsley, C. A. (1996). *True selves: understanding transsexualism for families, friends, co-workers, and helping professionals*. San Francisco, CA: Jossey-Bass.
- Cromwell, J. (1999). *Transmen and FTMs: identities, bodies, genders, and sexualities*. Urbana, IL: University of Illinois Press.
- Ettner, R. (1996). *Confessions of a gender defender: a psychologist's reflections on life among the transgendered*. Chicago, IL: Spectrum Press.

- Ettner, R. (1999). *Gender loving care: a guide to counseling gender-variant clients*. New York: W. W. Norton.
- Green, J. (2004). *Becoming a visible man*. Nashville, TN: Vanderbilt University Press.
- Israel, G. E. and Tarver, D. E., II (1997). *Transgender care: recommended guidelines, practical information and personal accounts*. Philadelphia, PA: Temple University Press.
- Kaiser Permanente National Diversity Council (2004). *A provider's handbook on culturally competent care: lesbian, gay, bisexual, and transgender population* (2nd edn.). Oakland, CA: Kaiser Permanente National Diversity Council.
- Lev, A. I. (2004). *Transgender emergence: therapeutic guidelines for working with gender-variant people and their families*. New York: Haworth Press.
- Stryker, S. and Whittle, S. (2006). *The transgender studies reader*. New York: Routledge.

Trauma and Memory

B A van der Kolk

Boston University School of Medicine, Boston, MA USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 620–622, © 2000, Elsevier Inc.

Trauma and Dissociation

The Nature of Traumatic Memories

Amnesia and the Return of Dissociated Memories

The Neurobiology of Traumatic Memory

The issue of trauma and memory is central to the concept of posttraumatic stress disorder (PTSD). The *Diagnostic and Statistical Manual of Mental Disorders* (4th edn, DSM-IV) defines PTSD as being accompanied by extremes of remembering and forgetting. Because traumatic memories take on their extreme character only after a person has been exposed to overwhelming stress, laboratory studies using conventional stimuli seem to be of little guidance in clarifying what happens when people are traumatized. Because trauma cannot be simulated in the laboratory, the exploration of the relationship between trauma and memory depends on (1) collecting retrospective reports from traumatized individuals, (2) post hoc observations of people with trauma histories, or (3) provoking traumatic memories and flashbacks in the laboratory.

Trauma and Dissociation

Christianson and others have described how, when people feel threatened, they tend to experience a significant narrowing of their consciousness while remaining focused on the central details of their ex-

perience. This increasing narrowing of consciousness is a reflection of the dissociation of traumatized individuals. Dissociation at the moment of the trauma has been found to be the single most important predictor of the subsequent development of PTSD. Dissociation has been described as a concomitant of trauma during World War I and II and after accidents, rapes, and numerous other traumatic experiences. Although dissociation may have some adaptive value, it also seems to lead to a disintegration of the capacity to integrate the different perceptual elements of the traumatic experience. Extreme physiological arousal interferes with proper information processing and with the storage of information into narrative (explicit) memory. When people are overwhelmed, they may suffer from speechless terror in which words fail to describe what has happened. However, even though traumatized individuals may be unable to make a coherent narrative of the traumatic events, there may be no interference with implicit memory; they may know the emotional valence of a stimulus and be aware of the associated perceptions without being able to articulate the reasons for feeling or behaving in a particular way. Traumatic memories are split off from (dissociated) consciousness and are stored as sensory perceptions, obsessional ideas, or behavioral reenactments.

The Nature of Traumatic Memories

Under ordinary conditions, memory always is an active and constructive process. Whereas familiar and expectable experiences are generally easily integrated and memories of ordinary events are bleached in clarity over time, some aspects of traumatic events appear to get fixed in the mind, unaltered by the passage of time or by the intervention of subsequent

experience. However, the emotional valence of experience affects the accuracy of memory; reports of personally highly significant events are often unusually accurate and remain fairly stable over time. Evolution favors the consolidation of relevant information. Since late in the nineteenth century, clinicians have noticed over and over again that the memories of traumatized patients seem to consist predominantly of emotional and perceptual elements, whereas there appears to be an impairment of declarative memory. A series of studies on the memories of traumatized populations have demonstrated that, regardless of the nature of the trauma, traumatized individuals initially report having decreased verbal recall, whereas they are prone to reexperience the traumatic events in the form of visual images, kinesthetic sensations, sounds, smells, and extreme affective states. Whether we study people reporting childhood trauma, car accidents, rape, or other assaults or people who wake up in the middle of a surgical procedure, this pattern of recollection is consistently reported by victims with PTSD. This pattern of memory recollection suggests that traumatic memories are processed differently in the brain than memories of day-to-day experiences.

It is generally assumed that memories can easily be distorted and that verbal recall changes over time. Even so-called flashbulb memories tend to be modified with the passage of time. However, traumatized individuals tend to report that the content of nightmares and flashbacks remain the same. The author was able to show that memories of traumatic experiences change after effective treatment of PTSD but do not change if the PTSD remains untreated. There have been no reports in the scientific literature showing that the sensory and perceptual elements of traumatic memories change in quality with the passage of time; nor has the reverse been reported – that these elements of traumatic memories always remain unaltered.

Amnesia and the Return of Dissociated Memories

Amnesia following adult trauma has been very well documented following World War I and II, in accident victims, in concentration camp survivors, and in victims of torture. In recent years, the occurrence of amnesia following childhood maltreatment has been called into question. This controversial issue has given rise to numerous studies. At this time, there have been 39 published reports in peer-reviewed journals that have documented that anywhere between 20 and 80% of people with sexual abuse histories

suffer from amnesia for the event at some time during their lives; every published study of this population has found some percentage of subjects who suffered from amnesia. Precipitants for the return of these previously forgotten memories tend to be events or sensations that are reminders of the original trauma. Events such as witnessing someone else going through trauma similar to one's own, hearing about the similar trauma, seeing a movie, or being in the company of someone who reminds a person of him- or herself at the same age at which the abuse occurred tend to be the most common precipitants for the return of these memories. A well-known study by Linda Williams of a group of women with documented sexual abuse histories found that the memories of the women who had a period of amnesia and who later recalled their trauma were at least as accurate as the memories of the women who had always recalled their abuse.

The Neurobiology of Traumatic Memory

Based on animal research, it has been widely assumed that the massive secretion of neurohormones at the time of the trauma plays a critical role in the ways that traumatic memories are consolidated into long-term memory. The strength of hormonal stimulation at the time a particular event occurred determines how strongly the memory is laid down. This capacity helps organisms evaluate the importance of sensory input better and causes emotionally significant material to be accessed more easily. The problem in PTSD is that the capacity to access memories related to the trauma becomes too easy; the trauma is recalled whenever a person is exposed to a sufficient number of matching sensory stimuli, regardless of their current relevance.

Traumatic Memories Are State-Dependent

In people with PTSD, sensations reminiscent of the trauma, including states of high physiological arousal, selectively promote the retrieval of traumatic memories or precipitate behaviors associated with the original trauma. Classically, a Vietnam veteran may have a flashback during a Fourth of July picnic and behave as if he were back in Vietnam, but he might not respond to fireworks on New Year's Eve because only the sound, but not temperature, matches the sensations that accompanied the traumatic events. Biological research has shown that drugs that precipitate autonomic arousal, such as lactate or yohimbine, tend to precipitate panic attacks and flashbacks in patients with PTSD. These can be considered instances of state-dependent memory retrieval.

Functional and Neuroanatomical Correlates of Posttraumatic Stress Disorder: Implications for Understanding Traumatic Memories

A range of both neuroanatomical and functional changes have been documented in individuals with PTSD. Several studies have shown decreased hippocampal volume, as well as decreased volume of the corpus callosum in patients with abuse histories. Functional neuroimaging has shown that people with PTSD have an increased activation of the right amygdala when exposed to a traumatic stimulus, which is accompanied by a decreased activation of Broca's area. These findings suggest that the difficulty that many PTSD patients have in putting their memories into words reflect in abnormalities in brain activation. Although it is unclear at the present time what the precise clinical implications of these findings are, it is tempting to speculate that a decreased hippocampal volume may be responsible for the fragmented nature of traumatic memories. Altered cerebral lateralization, with increased right hemispheric activation, may be relevant for understanding dissociative phenomena.

See Also the Following Articles

Amnesia; Concentration Camp Survivors; Dissociation; Memory and Stress; Posttraumatic Stress Disorder, Neurobiology of; Torture.

Further Reading

- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th edn.). Washington, DC: American Psychiatric Association.
- Appelbaum, P. (1998). *Trauma and memory: clinical and legal controversies*. Oxford University Press.
- Bremner, D. and Marmar, C. (1998). *Trauma, memory, and dissociation*. Progress in Psychiatry, no. 54. Washington, DC: APA Press.
- Brown, D. P., Schefflin, A. W. and Corydon Hammond, D. (1998). *Memory, trauma treatment, and the law*. New York: W. W. Norton.
- Rauch, S., van der Kolk, B. A., Fisler, et al. (1996). A symptom provocation study using positron emission tomography and script driven imagery. *Archives of General Psychiatry* 53, 380–387.
- van der Kolk, B. A., McFarlane, A. C. and Weisaeth, L. (eds.) (1996). *Traumatic stress: the effects of overwhelming experience on mind, body and society*. New York: Guilford Press.
- van der Kolk, B. A., Burbridge, J. A. and Suzuki, J. (1997). The psychobiology of traumatic memory: clinical implications of neuroimaging studies. *Annals of the New York Academy of Sciences* 821, 99–113.

Trauma Group Therapy

J Dwyer and L G Martin

VA Greater Los Angeles Healthcare System,
Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

History

Characteristics of Trauma Groups

Definition of Group Therapy

Group Process

Trauma Group Structure

Glossary

Arousal

A state of activated emotional activity.

Catharsis

Psychotherapy that encourages the discharge of pent up, socially unacceptable affects; the discharge of pent up emotions to result in the alleviation of symptoms or permanent relief of the condition.

Cognitive restructuring
Exposure

The identification and modification of maladaptive thought processes.

The disclosure of something private or secret; the act or instance of revealing or unmasking, for example, traumatic events.

Process

A systematic series of actions directed to some end (e.g., the uncovering of repressed emotions and memories).

Rap group
Talking cure

Peer group discussion.

The verbal interaction between therapist and patient (or peers) regarding the goal of uncovering hidden trauma; coined by Sigmund Freud.

History

Before presenting the elements of the trauma group, it is essential to understand the efforts to legitimize the trauma syndrome itself. Society has moved into a period when people are seeking the human suste-

nance that has been lost in the development of technology, the emergence of the megalopolis, and the increased complexity and depersonalization of societal structures. Individuals and communities are too often witnessing and experiencing the traumatic effects of the big hurts in life. Generalized angst in our societies prompts a search for a sense of community with intimate and mutually supportive human ties, a search for community that enhances the lives of individuals through a common purpose. Toward this end, more people are forging their own microcosms of community in the form of therapy or support groups. These come in many forms: action groups, which support collective or individual impact on societal institutions; interpersonal exploration groups, in which individuals seek to determine who they really are versus how they appear to others; and other groups pursuing strongly held common interests such as learning, experiencing or performing in a world of ideas and skills.

In the late nineteenth century, the disorder known as hysteria became the focus of a major inquiry. French neurologist, Jean-Martin Charcot studied the symptoms of hysteria, which resembled neurological damage: motor paralysis, sensory loss, convulsions, and amnesia. By 1880, he had demonstrated that these symptoms were psychological and could be artificially induced and relieved using hypnosis. Charcot's followers pursued the cause of hysteria. Rivalry in this pursuit was particularly intense between Janet and Freud. By the mid-1890s, Janet in France and Freud (with his collaborator Joseph Breur) in Vienna had independently arrived at strikingly similar formulations – hysteria was a condition caused by psychological trauma. While conducting explorations toward a cure, they found it necessary to speak to their patients. They found that unbearable emotional reactions to traumatic events produced an altered state of consciousness, which, in turn, yielded the symptoms of hysteria. Janet described these patients as being governed by subconscious fixed ideas – the memories of traumatic events. Breur and Freud, in an immortal summation, wrote that hysterics suffer from reminiscences.

By 1908, these investigators had also discovered that hysterical symptoms could be alleviated when the traumatic memories and intense feelings that accompanied them were recovered and put into words. Janet called the technique psychological analysis; Breur and Freud dubbed it the talking cure. Freud later recanted the link between hysteria and trauma, and the entire line of inquiry fell into neglect. The study of trauma came to a halt.

Reemergence of Trauma Investigation

The reality of psychological trauma was forced into the public consciousness again by the catastrophe of World War I, although the primary mental health conclusion at that time was that soldiers who developed a traumatic neurosis were at best constitutionally inferior human beings and at worst cowards and malingerers. Within a few years after World War I, medical interest in the subject faded again. In 1922, a young American psychiatrist, Abram Kardiner, took up the case of traumatized soldiers. He developed the clinical outline of the trauma syndrome as it is understood today.

With the casualties of World War II came a revival of medical interest in combat neurosis. Various studies led to the conclusion that the strongest protection against this overwhelming terror was the degree of relatedness between the soldier, his immediate fighting unit, and their leader. Their microcosmic community recreated treatment strategies, evolved during World War II, that were designed to minimize the separation of the afflicted soldier and his comrades. In the quest for a fast and effective method of treatment, hypnosis was used, as it had been in earlier work on hysteria. Under hypnosis, the talking cure was used with a focus on recovery and cathartic re-experiencing of traumatic events and their attendant feelings of terror, rage, and grief.

Systematic large-scale studies of the long-term effects of combat were not undertaken until after the Vietnam War. The use of groups to treat trauma was pioneered with Vietnam veterans. By the mid-1970s, hundreds of informal rap groups had been organized.

Characteristics of Trauma Groups

From their beginning, trauma rap groups were different and innovative. They had three guiding principles: (1) The groups were formed around an affinity, or the coming together of people who shared an historical perspective or personal experience; (2) the task of the facilitator was to be empathic and manifest a presence at all times of full engagement, openness, and interaction and not to serve merely as a sounding board; and (3) the veterans had to assume responsibility for self-generation, that is, to initiate their own process and conduct groups largely on their own terms.

Early proponents of trauma treatment in the group model illuminated six characteristics that are shared by veterans active in the groups.

1. They have a pervasive guilt from which they cannot escape. Its nexus resides in having lived

when others from both sides of battle have been killed or severely maimed. The guilt leads to a pathological desire to atone for the sin of having survived. To this end, they become involved in reckless forms of self-punishment such as substance abuse or one-car accidents that are just a shade less determined than actual suicide.

2. They have a feeling of being scapegoats, having been used and betrayed.

3. They have a generalized rage at society and at the individuals they believe have duped and/or manipulated them.

4. They have undergone combat brutalization, a method of military training aimed at creating the most effective soldiers by numbing the emotional processes that could interfere when the need to fight or kill arises. This learned detachment leads to the fifth characteristic.

5. They are alienated from their own humanity and withdrawn from others. This alienation inevitably leads to the sixth and most devastating characteristic of all.

6. They have a profound inability to love, accept affection, or to be intimate with others.

These characteristics predominated then, and they prevail today as the symptoms we most often see presented in treatment for trauma.

Definition of Group Therapy

Group therapy is psychotherapy practiced ideally in a group of 8–10 members/participants. The group itself becomes an important entity in the treatment of its individuals. The group collective assists its members by guiding, constructively criticizing, and offering therapeutic suggestions and support. It is the application of psychotherapeutic techniques within a small group of individuals who share similar experiences and have difficulty in responding to them. Each group has one or two leaders or facilitators whose role it is to guide and direct the group discussion or processing of problems in an attempt to promote personal growth and favorable behavioral change.

Types of Therapy Groups

There are different kinds of therapy groups. They may deal with a variety of issues including interpersonal conflicts, addictions, or trauma. The Department of Veterans Affairs National Center for Posttraumatic Stress Disorder (PTSD) has written that the best option for most PTSD patients is group therapy. In the trauma group, patients discuss with others their traumatic memories, symptoms of PTSD, and functional deficits. This approach has been most

successful with combat veterans, rape and/or incest victims, and natural disaster survivors. Indeed, the group model has been so effective that it lies at the core of the ever-growing self-help movement. The trauma group promotes discovery of, insight into, and eventually integration of the traumatic event.

Group Process

Trauma groups employ two important aspects of trauma recovery: the talking cure and the healing benefits of modeling supportive interactions with others in small-group cohesion. Survivor groups typically comprise 8–10 people meeting regularly to provide support and encouragement to one another. They accomplish a number of steps in healing by confronting alienation and estrangement. Survivors begin to realize that they are not alone in their reactions. Their feelings are normalized, and they begin to feel that they are part of something rather than being different or separate. The group offers a sense of community and a feeling of security.

Cognitive Restructuring

An important function of the trauma group is to destigmatize the traumatic experience. It focuses on exposure, which is the repeated disclosure of the event(s), and guides patients to a more realistic view of themselves and their world (cognitive restructuring). In addition, the trauma group combines the strengths of its members by having them each share coping skills and ideas that have benefited them. Most important, the group creates a milieu in which members can learn to trust and establish an ability to enter into relationships again. The group provides a laboratory for trusting others and, in time, patients can generalize this trust and begin to form healthy relationships in the outside world. The group model has been shown to be inappropriate and not helpful with patients who are suicidal, homicidal, violent, actively abusing drugs, or self-mutilating.

Trauma Group Structure

The foundation of the trauma group rests on exposure, cognitive restructuring, and the use of the traumatic narrative. It allows for reexperiencing trauma in manageable doses. The methods used in each phase share the common feature of helping anxious patients confront their most fearsome recall with the goal of reducing irrational fear (anxiety). Exposure has been the standard of treatment for anxiety disorders for several decades. The trauma group, in phases, employs both exposure and cognitive therapy to extinguish the prevailing pathological fear

structure in the survivor's memory. The effective group process offers the patient corrective trauma-related information. This, in turn, provides an opportunity for patients to begin the process of desensitization, or adaptation to the trauma.

Repeated reliving of the trauma reduces it to a specific occurrence rather than a generalized representation of the world as a universally dangerous place and reduces patients' self-image of incompetence. The trauma group is structured to merge exposure with cognitive therapy; this merger yields the desired therapeutic effect – the integration of dissociated trauma material and newly acquired corrective information.

Phases of the Trauma Group

Phase I The first phase of the trauma group addresses the issues of trust and personal safety. Patients are given assurance that group content is absolutely confidential and that no violence will occur during group sessions. Participants learn quickly that they are in control and that the model involves small steps toward emotional passages. "Managing small doses of awareness" becomes the mantra of the group. Patients know that they are in control, an essential element derived directly from the rap group model. During this phase, patients build trust with the facilitator(s) and other group members. They learn that the group medium fosters support and mutual healing, and the trust emerges slowly, with each emotional victory. Emphasis is on the present – the here and now. The importance of self-care, stress management, and anger management is highlighted. The process confronts the Lone Ranger syndrome and encourages each member to become part of the cohesive unit. The beginning steps of trust and bonding begin to surface; healthy and supportive relationships are born and nurtured. The patients demonstrate emotional growth, indicated by patience, courage, and the desire to confront their fears. They have constructed a healthy emotional platform from which to begin the arduous task of thawing freeze-dried traumatic memories.

Phase II The second phase of the trauma group is the narrative/trauma phase. Survivors have stories that must be revealed to themselves and to others. The fundamental human need to share and connect (Maslow) is enacted in phase II. It becomes about exposing and facing the hidden horror. Initially, narratives can be fragmented and distorted. Refining them in a controlled, supported environment reveals the conflicts between the memory and the actual event(s). The narrative helps to thaw out the frozen memory. It offers an important way to desensitize

patients to the enormous emotional overload. By repeated telling of the story and significant group feedback, a linear narrative emerges and patients begin to see the event more clearly and can begin to integrate the frozen memory into their broader life picture.

Phase III The last phase of the trauma group is the reconnection phase, or Life 101. A desire and ability to return to a productive, healthy, and satisfying life style returns. The symptoms of PTSD have become manageable. Patients can bear the feelings associated with the traumatic memories because they are now in a position of authority over them. The memory, with all of its dreaded darkness has become a coherent story linked to an appropriate affect. Patients, unlike the carpenter with a hammer, no longer see everything as a nail. Long-held perceptions regarding self and the world have been altered. Often, the survivors can rid themselves of the hero-villain paradigm of life. They can once again embrace a humanity that has been out of reach for so long. Patients begin to accept that life is imperfect and that in life both love and horror can coexist. Patients move from alienation from themselves to being the narrator of their own lives. Reconnecting means repairing the covenant of trust with themselves, which was broken at the time of trauma. It means being active, engaged, and prosocial. It means serving others and striving to make the world a better place. Life can become meaningful and fulfilled with a balance of joy, laughter, and sometimes tears. Most important, survivors can once again give and receive love. There is life after trauma.

Further Reading

- Briere, J. (2004). *Psychological assessment of adult post-traumatic states*. Washington, DC: American Psychological Association.
- Foa, E., Keane, T. and Friedman, M. (eds.) (1998). *Effective treatments for PTSD*. New York: Guilford Press.
- Follette, V., Ruzek, J. and Abueg, F. (eds.) (1998). *Cognitive behavioral therapies for trauma*. New York: Guilford Press.
- Herman, J. (1992). *Trauma and recovery*. New York: Basic Books.
- Kardiner, A. and Spiegel, H. (1947). *War, stress and neurotic illness*. New York: Hoeber.
- Schiraldi, G. (2000). *The PTSD source book*. Los Angeles: Lowell House.
- Shatan, C. (1973). The grief of soldiers: Vietnam combat veteran's self-help movement. *American Journal of Orthopsychiatry* 43, 640–653.
- Shay, J. (1994). *Achilles in Vietnam*. New York: Simon & Schuster.
- Williams, M. B. and Sommer, J. (eds.) (1994). *Handbook of posttraumatic therapy*. Westport, CT: Greenwood Press.

Traumatic Stress and Posttraumatic Stress Disorder, the Israeli Experience

E Klein

Faculty of Medicine Technion, Haifa, Israel

J Zohar

Tel Aviv University, Tel Aviv, Israel

© 2007 Elsevier Inc. All rights reserved.

Introduction

Posttraumatic Reactions in Holocaust Survivors

Combat Stress Reaction and Posttraumatic Stress Disorder in War Veterans

The Impact of Terror on Traumatic Stress Responses in the Israeli Population

Neurobiology and Psychophysiology of Posttraumatic Stress Disorder

Animal Model for Posttraumatic Stress Disorder

Conclusion

Glossary

<i>Acute stress disorder</i>	An anxiety state lasting up to 4 weeks following a traumatic event.
<i>Acute stress reaction</i>	An acute and time-limited (up to 48 h) anxiety state following a traumatic event.
<i>Animal models</i>	Experiments in which animals are used to mimic and study human diseases.
<i>Combat stress reaction</i>	An acute stress reaction that develops on the battlefield.
<i>Gulf War</i>	The 1991 war in which the United States and the allied countries fought against Iraq following its invasion of Kuwait.
<i>Intifada</i>	Recent Palestinian campaigns directed at ending the Israeli military occupation (from <i>انتفاضة</i> <i>intifāḍah</i> , shaking off; Arabic for uprising).
<i>Posttraumatic stress disorder (PTSD)</i>	A chronic disorder that develops in survivors of traumatic events.
<i>Stressful life event</i>	An event that might cause emotional distress.
<i>Traumatic event</i>	A life-threatening event.
<i>Yom Kippur War</i>	The 1973 war in which Egypt and Syria attacked Israel.

Introduction

Emotional trauma and stressful life events have been implicated in the development of psychiatric disorders, particularly anxiety and depression. Traumatic

events differ from stressful life events in that they involve a direct threat to life or physical/psychological integrity. Posttraumatic responses can be diverse, but the most typical form of psychiatric morbidity resulting from exposure to a traumatic event is posttraumatic stress disorder (PTSD). It is estimated that at least 50% of the people in Western countries are exposed to traumatic events during their lifetime. Because PTSD occurs in approximately 10–20% of the individuals who are exposed to a traumatic event (the majority eventually adapt), the estimated life prevalence is 10% of 50% (i.e., 5%). Indeed, epidemiological studies show that 1–3% of the adult population in the United States suffer from PTSD. The finding that only a fraction of those exposed to traumatic events develop PTSD suggests that the traumatic event is necessary but not sufficient for its development and that additional factors are important in determining the individual vulnerability of trauma survivors to developing PTSD.

The study of human responses to trauma has been intimately associated with the history of wars in the twentieth century. There have been anecdotal accounts of posttraumatic responses in soldiers in the American Civil War and other wars in the nineteenth century, but World War I (1914–1918) is clearly the cradle of the modern study of trauma and its consequences.

Only in the last 2 decades has trauma research evolved to focus on civilian populations that have been exposed to violence, sexual abuse, terror activity, traffic accidents, and health conditions (e.g., heart attacks), which have all been shown to cause PTSD and are responsible for the high prevalence of the disorder in civilians.

Israel has become, since its establishment in 1948, a live laboratory for the research of trauma and PTSD. The horrors of the Holocaust and its emotional scars in the survivors, followed by seven wars and waves of terror, have unfortunately served as an experimental field for the study of trauma and its consequences. This article reviews key findings from studies done in Israel on trauma and PTSD during the last 2 decades and highlights their contribution to the existing body of knowledge in this field.

Posttraumatic Reactions in Holocaust Survivors

The Holocaust is an unparalleled event in human history that has left enduring emotional scars in all

its survivors. Most Holocaust survivors were able to readjust surprisingly well and restore their lives, whereas others were not able to overcome the dreadful events, which continued to cast a shadow on their entire life, making it impossible for them to regain security and emotional well-being.

Studies of Holocaust survivors were done mainly in the 1960s and 1970s, before the era of modern trauma research and the conceptualization of the definitions currently used in the field. This, and the fact that many of the survivors have already died and those who are still alive are in their 80s, explains the relative paucity of updated data on posttraumatic responses in Holocaust survivors.

Combat Stress Reaction and Posttraumatic Stress Disorder in War Veterans

The Yom Kippur War and Lebanon Wars

In contrast to the earlier wars in the short history of the state of Israel (1949, 1956, 1967), which were not studied systematically in the context of traumatic stress, some of the soldiers with combat stress reaction (CSR) during the Yom Kippur (1973) and Lebanon (1981) wars were identified early on, received treatment, and were followed over time, so that data could be collected about the short- and long-term consequences of exposure to traumatic events during wartime. The Lebanon war constituted the first systematic effort to apply and study the principles of proximity immediacy and expectancy (PIE) in the treatment of soldiers who developed CSR. These treatment principles, originally formulated during trench combat in World War I, are based on experience with soldiers who developed CSR and were removed from the artillery range. The basic premise of this treatment approach (which was never tested with appropriate scientific tools) is that the chances of overcoming CSR are highest when treatment is applied in proximity to the war zone, as early as possible after the condition develops, with an expectancy for full and rapid return to functioning. These relatively simple treatment principles were applied by the medical teams, including psychologists and social workers, that were part of the fighting units and were specially trained to deliver this treatment. Due to constraints of the war situation, the treatment was not available to all soldiers who developed CSR, and thus the prospective short-term and long-term outcome of treated and untreated veterans with CSR was limited. The data that emerged from these studies indicate that (1) CSR constituted approximately 15–25% of war-related casualties; (2) only

one-half of the veterans who eventually developed PTSD applied for treatment in the front line; (3) among those who were treated according to the PIE principles, 15–20% eventually developed PTSD; and (4) the long-term beneficial effects of the PIE treatment need to be further explored.

Prisoners of War

Prisoners of war (POWs) from these wars were also studied. Twenty years after their captivity, 13% of the POWs still suffered from PTSD, compared to only 3% in a comparison group of combat veterans. It was also evident that recovery from PTSD among POWs was slower and incomplete compared to the control group.

Secondary Traumatization

Life often becomes literally a nightmare for many of those suffering from PTSD. What has been realized only in more recent years is the impact of the disorder on those living with the victims and on their caregivers. It has been shown that both close family members and therapists who are exposed to the traumatic memories and suffering of the victims tend to develop emotional distress and PTSD-like symptoms, including intrusive thoughts, emotional numbing and avoidance, and symptoms of increased arousal. This phenomenon was called secondary traumatization.

Traumatic Injury and Posttraumatic Stress Disorder

Another issue that received attention during the last years of the war activity in Lebanon, before the final withdrawal of the Israeli army from Lebanon, was the effect of traumatic injury on the development of PTSD. The assumption that bodily injury inflicted during the trauma does not increase the risk for PTSD, and in fact might even serve as a protective factor, prevailed among trauma experts in the past. It was believed that, being visible, the injury engenders sympathy and support from the environment to a greater extent than the psychic wounds of the trauma. It was further postulated that the injury also provides a focus for coping that distracts the subject's mind from the trauma and, last but not least, in the case of war, evacuation due to the injury removes the source of anxiety and danger from the soldier by providing a legitimate escape from the stressful situation. However, when injured and uninjured soldiers who were involved in identical war activities in Lebanon were compared for posttraumatic symptoms, it was found that the injured soldiers had an eightfold higher risk for developing PTSD compared to their uninjured peers. These findings clearly demonstrate that injury

in the context of trauma exerts an additional risk factor and increases the risk of PTSD. Possible explanations for this finding are that injury amplifies the sense of horror and perceived threat for life. In addition, the medical procedures and the pain resulting from the injury create an additional source of stress and serve as retrieval cues for the traumatic memories.

Lessons from the Persian Gulf War

During the Persian Gulf War in 1991, the entire population in Israel was, for the first time, under a direct threat of warfare and missile attacks. This constituted a serious psychological threat to the civilian population that had not been directly involved in war activity until that time. Data are not available about the prevalence and the magnitude of acute traumatic stress reactions in the entire population during that period. However, data are available about those who were evacuated to hospitals following the missile attacks.

Over a period of 2 weeks in the winter of 1991, approximately 40 surface-to-surface Iraqi SCUD missiles were launched and sent to Israel, targeting mainly the central coastal region of the greater Tel Aviv area. The uncertainty about the place, time, and type of warhead (chemical, biological, or conventional) was a source of stress and anxiety for the entire population, mainly in the targeted region. Furthermore, it was the cause of many traumatic stress reactions at or near the sites where the missiles hit. Approximately 800 citizens were evacuated to regional hospitals following these attacks, of which approximately 40% were diagnosed as psychological casualties. In addition, approximately 30% of the evacuees had mistakenly injected themselves with atropine, an antidote used during chemical warfare. These findings indicate that psychological responses constitute a significant portion of the casualties resulting from this kind of warfare and that rescue teams and hospitals have to be trained and prepared accordingly.

The specific psychological responses of mothers and children following the missile attacks were also studied. In a survey done 2 years after the war, families that were evacuated from homes that were destroyed or damaged in the missile attack were compared to families who did not have to leave their homes and were not directly affected by missile attacks. Mothers and children from the evacuated families manifested more posttraumatic symptoms and higher degrees of emotional distress. Stress symptoms tended to decrease over time in the displaced children from the evacuated families but not in their mothers. No differences in the children's adaptive behavior in the two groups were observed. Mothers

who manifested more posttraumatic symptoms during the war were found to be more symptomatic at the follow-up, and the degree of stress in the children correlated with the mother's symptoms. These findings suggest that a maternal stress-buffering capacity is crucial for children's effective coping with stressful situations. The degree of family cohesion was also found to be an important factor in the successful coping of these children with the stressful situation. A follow-up study 5 years after the Gulf War showed significant decrease in posttraumatic symptoms in the evacuated children.

The Impact of Terror on Traumatic Stress Responses in the Israeli Population

The years 1987–2005 of the Israeli-Palestinians conflict, known as the Intifada, have exposed the civilian population in Israel to an unprecedented degree of violence and direct threat to lives as a result of numerous terror attacks and suicide bombers. The peak of the terror wave was in the years 2000–2004, and during this period 963 Israelis were killed and more than 6300 were wounded in hundreds of terror attacks and many more were directly exposed to the impact of terror and violence with direct threat to their lives.

A survey conducted in 2002 revealed that 16% of a representative population sample of approximately 560 people had been directly exposed to a terrorist attack and 37% had a family member or a friend who had been exposed. Ten percent of the sample met the diagnostic criteria for PTSD and approximately 60% reported feeling depressed. However, approximately 80% of the respondents expressed optimism about their personal future and 67% were optimistic about the future of the country. Seventy-five percent of the respondents expressed self-confidence about their ability to function in a terrorist attack. Despite the fact that many of the respondents expressed a low sense of safety with respect to themselves and their relatives, only 5% reported a need for professional help. Altogether, these findings indicate that, considering the nature and length of the Israeli traumatic experience during this period, the psychological impact of the situation may be considered moderate to low, and despite the fact that participants showed distress and a lowered sense of safety, only a minority developed psychiatric distress. These apparently discrepant findings may be related to a process of habituation and to effective psychological coping mechanisms.

Interesting differences were evident between urban communities exposed to different levels of stress and terror activities. To a certain degree, those communities

that were exposed to more violence and terror activity showed better coping abilities and lower degrees of personal distress than did communities with low exposure when facing stressful conditions.

Gender differences in the prevalence of PTSD have been reported. Although men are more frequently exposed to traumatic events, most studies have reported higher rates of PTSD among women. Studies done in Israel during 1999–2003 showed lower rates of direct exposure to terror among women, yet, compared to men, women reported more subjective distress and more posttraumatic symptoms and were using more coping strategies to overcome the distress. Sixteen percent of the women (but only 2.5% of the men) in the Israeli survey met the full criteria for PTSD.

Neurobiology and Psychophysiology of Posttraumatic Stress Disorder

Sleep and Posttraumatic Disorder

Sleep complaints are among the most frequently reported symptoms in PTSD. Difficulty falling asleep, restless and interrupted sleep, and recurrent dreams with nightmares are very common complaints.

Sleep disturbances and complaints of disrupted sleep in the early aftermath of the trauma have been associated with higher rates of PTSD later on. Some people with PTSD tend to develop a reversed sleep cycle, staying awake most night hours with a tendency to oversleep during the day. However, despite the high prevalence of subjective sleep complaints in PTSD, objective sleep measures have shown inconsistent findings. Unlike major depression, in which shortened rapid eye movement (REM) sleep latency is a well-established finding, polysomnographic sleep studies in patients with PTSD have not defined a clear pattern of sleep alterations that might be correlated with the subjective complaints of the patients. Some studies have shown fragmented REM sleep, whereas others have shown reduction in slow-wave sleep or increased awakening thresholds from slow-wave sleep. However, these findings are inconsistent. It has been suggested by some researchers that the subjective complaints of disrupted sleep in PTSD that are not corroborated in the polysomnographic studies are due to an altered sleep perception, namely, that the neurophysiology of sleep is not disturbed but there is a state of sleep misperception that causes a dissociation between an apparently normal sleep pattern and its altered perception by the subject. This could explain the relative inefficacy of sleep medications in PTSD, which is a widely acknowledged phenomenon and results in the excessive use and often abuse of sleep medications in these cases.

Physiological Arousal and Posttraumatic Disorder

Symptoms of increased arousal and exaggerated startle response are characteristic of PTSD. When exposed to loud tones 1 and 4 months after a traumatic event, subjects who developed PTSD showed a greater heart rate response, altered skin conductance, and eye muscle electromyogram (EMG) responses when compared to subjects who did not develop PTSD after a traumatic event. No such differences were evident between the groups immediately after the trauma, indicating that these findings were state-dependent and characteristic of trauma survivors who develop PTSD.

Furthermore, subjects with PTSD failed to show the habituation in response to repeated loud tones that is normally seen and was evident in the comparison subjects who did not develop PTSD. This phenomenon could reflect a more generalized inability for habituation of stressful signals; the recurrent intrusive memories in PTSD could be considered to be a failure of habituation to the content of the traumatic memories.

Heart Rate and Heart Rate Variability in Posttraumatic Disorder

Increased heart rate is a determinant of physiological arousal seen in PTSD. Subjects who develop PTSD have been shown to have higher heart rates during the immediate aftermath of the trauma compared to traumatized individuals who do not develop PTSD. The elevated heart rate is caused by the increased noradrenergic tone, suggesting that increased noradrenergic activity immediately after the trauma may play an important role in the neurobiological processes involved in the development of PTSD. From a clinical perspective, this finding suggests that elevated heart rate immediately after the trauma is a predictor of PTSD.

Animal Model for Posttraumatic Stress Disorder

The validity of an animal model for psychiatric disorders is often questionable. However, there is a clear advantage for PTSD because an exposure to threat (a predator, in the case of a rat) certainly addresses some of the validity issues that might be problematic in animal models of other psychiatric disorders. An animal model of PTSD composed of exposure to a predator (the smell of a cat, in this case), applying behavioral cut-off criteria to differentiate between the maladaptive rats and the well-adapted rats was developed in Israel in 2003. This enables investigators to further advance PTSD research by examining

questions such as early life adversities, stress-dependent increase in cortisol levels, and therapeutic (preventive) effects of early administration of medications that are often used in the treatment of PTSD, such as selective serotonin reuptake inhibitors (SSRIs).

Conclusion

The unique (although unfortunate) combination in Israel of having many Holocaust survivors, several wars, exposure to repeated waves of terror, and the presence of a pure form of PTSD (i.e., very low rates of alcoholism and drug abuse), coupled with the expertise in research, have helped Israeli researchers and clinicians to develop vital and advanced treatment and research in stress-related disorders, including acute stress reaction, CSR, and PTSD. The key findings have called into question some of the accepted notions of PTSD. Issues such as a direct linkage between terror exposure and PTSD were not found. The use of PIE principles in the modern battlefield seems to require further research, as do the apparently strong link between physical injury and PTSD and the lack of objective findings in a sleep laboratory in regard to sleep disturbance in PTSD. The development of a new animal model for PTSD, with the focus on behavioral cut-off criteria, along with further research in regard to the genetic vulnerability to PTSD, may pave ways to deepen our understanding of this intriguing phenomenon.

Further Reading

- Barak, Y., Aizenberg, D., Szor, H., et al. (2005). Increased risk of attempted suicide among aging holocaust survivors. *American Journal of Geriatric Psychiatry* 13(8), 701–704.
- Bleich, A., Gelkopf, M. and Solomon, Z. (2003). Exposure to terrorism, stress-related mental health symptoms, and coping behaviors among a nationally representative sample in Israel. *Journal of the American Medical Association* 290(5), 612–620.
- Cohen, H. and Zohar, J. (2004). An animal model of post-traumatic stress disorder: the use of cut-off behavioral criteria (review). *Annals of the New York Academy of Sciences* 1032, 167–178.
- Gil, S., Caspi, Y., Ben-Ari, I. Z., et al. (2005). Does memory of a traumatic event increase the risk for posttraumatic stress disorder in patients with traumatic brain injury?: a prospective study. *American Journal of Psychiatry* 162(5), 963–969.
- Koren, D., Arnon, I. and Klein, E. (1999). Acute stress response and posttraumatic stress disorder in traffic accident victims: a one-year prospective, follow-up study. *American Journal of Psychiatry* 156(3), 367–373.
- Koren, D., Arnon, I., Lavie, P., et al. (2002). Sleep complaints as early predictors of posttraumatic stress disorder: a 1-year prospective study of injured survivors of motor vehicle accidents. *American Journal of Psychiatry* 159(5), 855–857.
- Koren, D., Norman, D., Cohen, A., et al. (2005). Increased PTSD risk with combat-related injury: a matched comparison study of injured and uninjured soldiers experiencing the same combat events. *American Journal of Psychiatry* 162(2), 276–282.
- Laor, N., Wolmer, L. and Cohen, D. J. (2001). Mothers' functioning and children's symptoms 5 years after a SCUD missile attack. *American Journal of Psychiatry* 158(7), 1020–1026.
- Laor, N., Wolmer, L., Mayes, L. C., et al. (1997). Israeli preschool children under Scuds: a 30-month follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry* 36(3), 349–356.
- Nadler, A. and Ben-Shushan, D. (1989). Forty years later: long-term consequences of massive traumatization as manifested by Holocaust survivors from the city and the kibbutz. *Journal of Consulting and Clinical Psychology* 57(2), 287–293.
- Pillar, G., Malhotra, A. and Lavie, P. (2000). Post-traumatic stress disorder and sleep – what a nightmare! *Sleep Medicine Reviews* 4(2), 183–200.
- Shalev, A. Y. and Freedman, S. (2005). PTSD following terrorist attacks: a prospective evaluation. *American Journal of Psychiatry* 162(6), 1188–1191.
- Shalev, A. Y., Freedman, S., Peri, T., et al. (1998). Prospective study of posttraumatic stress disorder and depression following trauma. *American Journal of Psychiatry* 155(5), 630–637.
- Shalev, A. Y., Peri, T., Brandes, D., et al. (2000). Auditory startle response in trauma survivors with posttraumatic stress disorder: a prospective study. *American Journal of Psychiatry* 157(2), 255–261.
- Solomon, Z., Shklar, R. and Mikulincer, M. (2005). Front-line treatment of combat stress reaction: a 20-year longitudinal evaluation study. *American Journal of Psychiatry* 162(12), 2309–2314.
- Solomon, Z. and Lavi, T. (2005). Israeli youth in the Second Intifada: PTSD and future orientation. *Journal of the American Academy of Child and Adolescent Psychiatry* 44(11), 1167–1175.
- Solomon, Z., Neria, Y., Ohry, A., et al. (1994). PTSD among Israeli former prisoners of war and soldiers with combat stress reaction: a longitudinal study. *American Journal of Psychiatry* 151(4), 554–559.

Trier Social Stress Test

B M Kudielka and S Wüst

University of Trier, Trier, Germany

C Kirschbaum

Technical University of Dresden, Dresden, Germany

D H Hellhammer

University of Trier, Trier, Germany

© 2007 Elsevier Inc. All rights reserved.

Introduction: Origin of the Trier Social Stress Test (TSST)

Description of the TSST Protocol

Overview of Sources of Intra- and Interindividual Differences

The TSST in Clinical Populations

Conclusions

Glossary

Cortisol

A glucocorticoid hormone produced by the adrenal cortex. Cortisol exhibits a pronounced circadian rhythm and responds to a wide range of stressors. It has many physiological effects and virtually every nucleated cell in the body has cortisol receptors. It is predominantly (90–95%) bound to proteins in blood; only 5–10% of the total plasma cortisol circulates biologically active as unbound, free cortisol.

Hypothalamic-pituitary-adrenocortical (HPA) axis

A central regulatory and control system of the organism that connects the central nervous system with the endocrine system. Under stress, the hypothalamus secretes corticotropin-releasing hormone, which provokes the release of adrenocorticotrophic hormone (ACTH) from the pituitary. ACTH triggers the secretion of cortisol from the adrenal cortex. The functioning of the axis is controlled by several negative feedback loops.

Psychosocial stress

Psychological stress that exerts physiological, behavioral, and/or verbal-subjective stress responses as a result of stressful social interactions, which often involve social-evaluative threat (in contrast to stress responses provoked by physical exertion or pharmacological stimulation).

Social-evaluative threat

Social component of a psychological or a psychosocial stress task in which task performance could be negatively judged by others. Stress tasks containing the two components uncontrollability and social-evaluative threat are associated

Sympathetic-adrenal-medullary (SAM) axis

Trier social stress test (TSST)

with the largest hypothalamic-pituitary-adrenocortical axis stress responses and the longest recovery times.

Sympathetic nerves, originating in the central nervous system, stimulate the adrenal medulla, which in turn secretes epinephrine and norepinephrine. These catecholamines mobilize energy-producing mechanisms quickly and downregulate less important organ functions.

A standardized stress protocol for the induction of moderate psychosocial stress in laboratory settings, which mainly consists of a brief preparation period (3 min) and a test period in which the subject has to deliver a free speech (5 min) and perform mental arithmetic (5 min) in front of an audience. The TSST can be applied to younger and older adults, to children as well as clinical populations. The TSST is one of the few available stress protocols that satisfies the criteria of a motivated performance task that combines elements of uncontrollability and high levels of social-evaluative threat.

Introduction: Origin of the Trier Social Stress Test (TSST)

Stress is one of the most significant health problems in the twenty-first century. This explains a pressing need for investigations into the biological pathways linking stress and health. Besides the sympathetic-adrenal-medullary (SAM) axis, the hypothalamic-pituitary-adrenocortical (HPA) axis is the major physiological stress response systems in the organism. Since alterations in HPA axis stress responses appear to be a close correlate or even a determining factor of the onset of different diseases or disease progression, the characterization of an individual's HPA axis response pattern to psychosocial stress appears to be of major interest. For such research agenda, a novel research tool was needed since available laboratory stress protocols yielded insufficient activation of the HPA axis. A stress protocol that was potent enough to induce significant changes of endocrine and cardiovascular parameters in the majority of subjects tested was eventually developed and termed the Trier Social Stress Test (TSST). Over the past decade, well over 4000 TSST sessions were performed in many laboratories worldwide. Thus, the TSST has become a standard protocol for the experimental induction of psychological stress

in healthy subjects as well as in clinical populations, investigating a wide range of different outcome variables ranging from subjective-verbal stress reports to objective behavioral and biological stress responses, including parameters of the HPA and SAM axis and the cardiovascular, immunological, and blood coagulation system.

Recently, Dickerson and Kemeny conducted a meta-analysis based on 208 laboratory studies of acute psychological stressors in order to delineate the essential situational elements capable of eliciting HPA axis responses. The results show that motivated performance tasks elicit cortisol responses if they are (1) uncontrollable, (2) creating a context of forced failure in which a subject is unable to avoid negative consequences or cannot succeed despite his or her best effort, or (3) characterized by a social component called social-evaluative threat, in which task performance could be negatively judged by others. Tasks containing both components (uncontrollability and social-evaluative threat) were associated with the largest HPA axis stress responses and the longest recovery times. These meta-analytical findings fit to the theoretical reasoning (based on the social self-preservation theory) that “uncontrollable threats to the goal of maintaining the ‘social self’ would trigger reliable and substantial cortisol changes” (Dickerson and Kemeny, 2004). The authors concluded that the TSST is one of the few available stress protocols that satisfies the criteria of a motivated performance task that combines elements of uncontrollability and high levels of social-evaluative threat. Their results fit with earlier work by Mason, stating that whenever a situation (or a stimulus) is perceived as novel, uncontrollable, unpredictable, or ambivalent, or if the individual anticipates negative organismic or psychological consequences, the negative feedback signals from the hippocampus and other sites will be overridden, resulting in a cascade of corticotropin-releasing hormone (CRH), adrenocorticotrophic hormone (ACTH), and cortisol secretion.

Description of the TSST Protocol

The TSST is a motivated performance task consisting of a brief preparation period (3 min) followed by a test period in which the subject has to deliver a free speech (5 min) and perform mental arithmetic (5 min) in front of an audience. The total exposure time adds up to 13 min. The following description of the TSST procedure primarily applies to investigations of HPA axis responses, and, therefore, other outcome measures may require some modifications.

The test onset should be preceded by a rest period of at least 45 min to minimize unrelated short-term

effects on the outcome measures. The subject is asked to take over the role of a job applicant who is invited for a personal interview with the company’s staff managers (selection committee). The selection panel members are introduced as being specially trained to monitor nonverbal behavior, and subjects are informed to expect tape and video recordings for later analysis of their performance. The selection committee is composed of two or three male and female collaborators of the experimenter, unacquainted with the subject. The committee is trained to communicate with the subject in an unresponsive neutral manner and does not respond with any facial or verbal feedback (note: no harassment or anger induction). The committee first requests that the subject sit down (not if blood pressure is measured) at a small table equipped with paper and pencil to prepare the talk (preparation period). After 3 min, the subject is asked to step forward to a marked spot in front of the microphone about 1–3 m apart from the committee’s table and to start the oral presentation (note: the use of written concept is not allowed). Whenever a subject finishes the speech in less than 5 min, the managers respond in a standardized way. First they tell the volunteer, “You still have some time left. Please continue.” Should the subject finish a second time before the 5 min are over, the managers are quiet for 20 s and then ask prepared questions (e.g., What are your personal strengths? What are your major shortcomings? Do you have enemies? Why? What do you think about team work?). After 5 min, the selection committee interrupts and asks the subject to serially subtract the number 17 from 2023 as fast and as accurately as possible. On every failure the subject has to restart at 2023 with one member of the committee interfering “Stop – mistake – start over at 2023, please.” After 5 min the task is terminated.

Samples are to be gained before onset and after cessation of the stress task depending on the dynamic of the selected outcome variables and should cover prestress levels, the initial stress response, peak level, and recovery. Prestress HPA axis hormone levels should not be interpreted as baseline levels since they are substantially affected by the subjects’ anticipation of the upcoming challenge. At the end of the test session the subject will be fully debriefed about the goal of the study and the nature of the stressor. A full debriefing includes a visit of the collaborators who took over the role of the selection committee.

With this protocol, salivary cortisol levels rise two- to threefold in about 70–80% of all subjects, with peak levels around 10–20 min after cessation of the stress task. In addition to free cortisol, the levels of total plasma cortisol, ACTH, catecholamines (epinephrine, norepinephrine), growth hormone, prolactin,

testosterone, systolic and diastolic blood pressure, heart rate, immune parameters (e.g., neutrophils, eosinophils, basophils, lymphocytes, interleukin-6, TNF α), glucose, α -amylase, endogenous inhibitor of monoamine oxidase (MAO-AI), as well as measures of hemoconcentration (e.g., hematocrit, haemoglobin, plasma volume) and blood coagulation (fibrinogen, von Willebrand factor antigen, D-dimer, clotting factors) increase significantly.

The TSST protocol can also be employed to test children and older adults after changing the instructions slightly for the respective age group. In elderly adults, subjects are instructed to apply for a (voluntary) part-time job (e.g., child caring, housekeeping, technical assistant). To help a retired subject take over the role of a job applicant, a fabricated newspaper advertisement can be presented (tested age range 59–91 years). The adapted TSST for children (TSST-C) also consists of a preparation period, public speaking, and mental arithmetic task. In the speaking part, children receive the beginning of a story and are told that they should finish telling the story, making it as exciting as possible, in front of a jury (tested age range: 7–14 years). The numbers of the mental arithmetic task are adapted to the performance level of the respective age group.

Although areas under the response curve (AUCs) are much higher in the morning due to the circadian rhythm of HPA axis activity, comparable free cortisol net increases after TSST exposure can be assessed with equal reliability in the morning and afternoon. However, in early morning sessions the experimenter should ensure that the onset of a stress experiment does not interfere with the morning cortisol awakening rise (CAR). Also, the appropriate time window for test sessions in the morning is smaller than in the late afternoon.

Overview of Sources of Intra- and Interindividual Differences

This section summarizes findings from studies investigating various factors that potentially contribute to differences in HPA axis responses to the TSST. One of the most prominent features of HPA axis (and other system) responses after psychological stress is the large variation in response magnitude among subjects and across test sessions. Significant differences can be observed with respect to the net hormone output as well as the time course of hormone secretion. Considerable evidence has accumulated for a significant impact of a variety of moderating and intervening variables, which are briefly summarized in the following.

Gender and Sex Steroids

One of the most consistent findings from employing the TSST is the significantly larger salivary cortisol and ACTH response (up to twice as high) in men compared to women, while prestress levels are not considerably different. The same sex effect emerges for elderly subjects (see also below). The observed sex differences point to a possible impact of female sex steroids as potential mediators of HPA axis responses to psychosocial stress. In a study of 81 healthy young adults, the TSST was applied to men, women in the follicular phase of the menstrual cycle, women in the luteal phase, and women using oral contraceptives. The results disclosed that significant sex differences emerge for ACTH and free salivary cortisol but not for total plasma cortisol stress responses. Moreover, ACTH responses were elevated in men compared to women, regardless of menstrual cycle phase or use of oral contraceptives. Women in the luteal phase had saliva cortisol stress responses comparable to men, whereas women in the follicular phase or women taking oral contraceptives showed significantly lower free cortisol responses. These observations point to the necessity of strictly distinguishing between the total cortisol secretion and the levels of bioavailable free cortisol. Whereas many animal studies can be cited that have directly investigated the impact of estrogens on HPA axis regulation, few experimental studies have been conducted in humans, and the empirical evidence is rather inconsistent.

Age

A recent reanalysis of five independent TSST studies based on 102 healthy subjects between 9 and 76 years showed that the stress task induced significant HPA axis responses in all age groups in both males and females. The data revealed no sex differences in free cortisol responses in children and younger adults, but larger free cortisol responses in elderly men than in elderly women. This effect did not appear to be attributable to self-reported stress level differences. For total plasma cortisol, elderly women showed generally enhanced cortisol response levels compared to elderly and younger men as well as younger women. For ACTH, the response was higher in younger adults, primarily due to an elevated response in younger men.

With respect to autonomic responses, differential heart rate responses and recovery after exposure to the TSST were observed in healthy children, younger adults, and elderly adults. Aggregated data across the five studies showed that there was a marked age effect in the heart rate stress response, with children and younger adults showing significantly higher increases

than elderly persons. This effect is probably caused by well-known age-related alterations in heart rate regulation. The analysis of sex effects showed that girls had higher heart rate increases during the stress exposure than boys. In younger adults, stress responses were also higher in women. Peak heart rate responses were comparable in older men and women, with only men returning to prestressor baselines during the observation period.

Lactation and Breastfeeding

Lactation in women, in contrast to rats, does not result in a general restraint of HPA axis responses to a psychosocial stressor such as the TSST. Rather, suckling seems to exert a short-term suppression of the cortisol response to psychosocial stress.

Nicotine, Coffee, Alcohol Consumption, and Dietary Energy Supplies

Nicotine is a potent stimulator of the HPA axis through induction of CRH release after binding to cholinergic receptors in the locus coeruleus and hypothalamus, and chronic smoking changes the HPA axis responses to stress. Blunted cortisol responses to the TSST were repeatedly observed in habitual smokers. It has also been shown that caffeine intake can activate important components of the pituitary-adrenocortical response in humans during resting states, and there is some evidence for combined stimulatory effects of caffeine and psychological stress on HPA axis responses. Relationships between alcohol and acute stress can be at least threefold: (1) alcohol intake can alter physiological responses to acute stress, (2) acute social stress can impact on subsequent net alcohol consumption in healthy social drinkers, and (3) acute HPA axis stress responses as well as effects of acute alcohol intake before the TSST can be altered in subjects with a positive family history of alcoholism. Finally, the availability of dietary energy supplies (e.g., short-term fasting, glucose availability) appears to exert important regulatory functions in pituitary-adrenal stress responses, pointing to an important role of the nutritional state.

Social Support and Social Hierarchy

Few data provide experimental support for the hypothesis that social support might attenuate the HPA axis response to acute stress; however, effects appear to be sex specific. The neuropeptide oxytocin seems to enhance the buffering effect of social support on stress responsiveness, at least in men. These results concur with data from animal research suggesting an important role of oxytocin as an underlying biological mechanism for stress-protective effects of

positive social support. Also, the position in a social hierarchy seems to be relevant for acute HPA axis stress responses. After the TSST, salivary cortisol levels highly increased in socially dominant subjects, while only a modest elevation was observed in subordinate men.

Personality Factors

It is tempting to speculate that the HPA axis response to stress is influenced, at least in part, by stable trait factors since the endocrine response to psychosocial stress can be viewed as a close interaction between situation and person variables within a given context. Interestingly, relationships between personality traits and stress responses only emerge after repeated exposures to the TSST. While novelty may mask the impact of personality on the cortisol stress response on the first exposure, differences in the ability to cope with the stressful situation may lead to different cortisol stress response patterns on subsequent stress exposures. With decreasing novelty and unpredictability, the impact of psychological trait variables increases, and a closer correlation with the endocrine parameters can emerge. Thus, a reliable investigation of associations between personality variables and cortisol stress responses appears to require repeated TSST exposure and data aggregation.

Habituation

After repeated exposure to (initially) stressful situations, a rapid habituation of HPA axis stress responses is normally observed. This has been reported consistently for the TSST. Recently, substantial interindividual variability of such habituation patterns was documented in a large study sample. While 52% of subjects showed the well-known HPA axis response habituation, almost 16% of the participants showed a response sensitization across three test sessions. Interestingly, altered HPA axis response habituation to TSST exposures seems to be associated with vital exhaustion, pointing to a potential mechanism for how exhaustion relates to increased disease vulnerability. However, it can be speculated that habituation might be rather specific for HPA axis parameters since TSST studies investigating stress-related changes in parameters of the sympathetic nervous system, blood coagulation indices, hemoconcentration, blood cells, or interleukin-6 (IL-6) levels did not report on habituation effects.

Interventions

First evidence shows that brief group-based cognitive-behavioral stress management training can reduce neuroendocrine or cardiovascular TSST stress

responses. It may thus be speculated that such interventions may prove useful in preventing detrimental stress effects in healthy subjects as well as patients. Other relaxation techniques (muscle relaxation, relaxing music) might also alter stress responses or facilitate recovery from TSST stress exposure.

Genetic Factors: Heritability and Polymorphisms

In addition to behavioral and situational variation, HPA axis activity appears to be profoundly influenced by genetic factors, as shown in quantitative genetic studies (family and twin studies) as well as candidate gene studies on polymorphisms in the genes coding for the glucocorticoid receptor, the mineralocorticoid receptor, the melanocortin-2 (ACTH) receptor, the melanocortin-4 receptor, the μ -opioid receptor, the GABA_A (gamma-aminobutyric-acid_A) receptor, the α -adrenergic receptor, the 5-HT_{2A} (serotonin_{2A}) receptor, COMT (catechol-O-methyltransferase), TNF α (tumor necrosis factor α), and ACE (angiotensin-I converting enzyme).

TSST and Cellular Mechanisms

In a series of experiments applying the TSST stress protocol, Rohleder and co-workers investigated the influence of age, sex, and sex hormone status on glucocorticoid (GC) sensitivity of peripheral leukocytes since a cell's response to cortisol is predominantly determined by both the steroid level it is exposed to and its sensitivity for glucocorticoids, that is, the efficacy of glucocorticoid-mediated signal transduction.

Bierhaus et al. employed the TSST to elucidate how neuroendocrine stress responses can mediate cellular processes, which, in the long run, can promote vascular diseases. Their data provide convincing evidence that the activation of the transcription factor NF- κ B (nuclear factor- κ B), a protein that is sensitive to psychosocial stress, might represent a norepinephrine-dependent pathway converting psychosocial stress into cellular dysfunction.

The TSST in Clinical Populations

In addition to healthy study participants, the TSST can be used in clinical populations ranging from psychiatric diseases to somatic complaints as shown in several studies over the last few years. To date, the TSST has been successfully applied in patients suffering from, e.g., major depression, anxiety disorder, social phobia, posttraumatic stress disorder (PTSD), burnout, exhaustion, chronic fatigue syndrome, fibromyalgia, recurrent abdominal pain, chronic pelvic pain, temporomandibular dysfunction (myofascial

pain and dysfunction), functional gastrointestinal disorders (irritable bowel syndrome or non-ulcer dyspepsia), different manifestations of chronic atopy, and atopic dermatitis or chronic allergic asthma, and adults with a history of sexual or physical childhood abuse or breast cancer survivors.

Conclusions

Over more than a decade of research, the TSST has proven a useful tool in the field of basic, applied, as well as clinical psychobiological research with a wide range of psychobiological outcome variables. As well as being used in the direct study of different biological stress systems like the HPA and SAM axis or cardiovascular system, the TSST can be used to investigate biological pathways between stress and disease or to study stress-related phenomena (e.g., effects of stress on memory and cognition) and interactions between drug and stress effects. Clinical studies continue to accumulate evidence that different diseases are associated with characteristic stress response profiles. Furthermore, biological mechanisms begin to unfold that could be helpful in explaining the stress–disease association. In a next step, the challenge for researchers will be to investigate whether the TSST protocol can be applied as a diagnostic tool for the prediction of disease susceptibility and symptom severity and/or for monitoring the efficacy of interventions.

See Also the Following Articles

Acute Stress Response: Experimental; Adrenal Cortex; Adrenal Medulla; Adrenaline; Adrenocorticotrophic Hormone (ACTH); Aging and Psychological Stress; Aging and Stress, Biology of; Alarm Phase and General Adaptation Syndrome; Allostasis and Allostatic Load; Amygdala; Animal Models (Nonprimate) for Human Stress; Autonomic Nervous System; Blood Pressure; Cardiovascular System and Stress; Catecholamines; Central Stress Neurocircuits; Control and Stress; Coping Skills; Coping and Stress: A Lens and Filter Model; Corticosteroids and Stress; Corticotropin Releasing Factor (CRF); Endocrine Systems; Gender and Stress; Genetic Factors and Stress; Glucocorticoid Negative Feedback; Glucocorticoids, Effects of Stress on; Glucocorticoids, Overview; Heart Rate; Hypothalamic-Pituitary-Adrenal; Immune Cell Distribution, Effects of Stress on; Immune Function, Stress-Induced Enhancement; Immune Response; Learning and Memory, Effects of Stress on; Memory and Stress; Menopause and Stress; Neuroendocrine Systems; Neuroimmunomodulation; Psychological Stressors, Overview; Salivary Cortisol; Sex Differences in Human Stress Response; Social Stress, Animal Models of; Social Support; Stress Effects, Overview; Sympathetic Nervous System.

Further Reading

- Bierhaus, A., Wolf, J., Andressy, M., et al. (2003). A mechanism converting psychosocial stress into mononuclear cell activation. *Proceedings of the National Academy of Sciences USA* 100(4), 1920–1925.
- Biondi, M. and Picardi, A. (1999). Psychological stress and neuroendocrine function in humans: the last two decades of research. *Psychotherapy and Psychosomatics* 68(3), 114–150.
- Buske-Kirschbaum, A. and Hellhammer, D. H. (2003). Endocrine and immune responses to stress in chronic inflammatory skin disorders. *Annals of the New York Academy of Sciences* 992, 231–240.
- Buske-Kirschbaum, A., Jobst, S., Wustmans, A., et al. (1997). Attenuated free cortisol response to psychosocial stress in children with atopic dermatitis. *Psychosomatic Medicine* 59(4), 419–426.
- Dickerson, S. S. and Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychology Bulletin* 130(3), 355–391.
- Heim, C., Ehler, U. and Hellhammer, D. H. (2000). The potential role of hypocortisolism in the pathophysiology of stress-related bodily disorders. *Psychoneuroendocrinology* 25(1), 1–35.
- Heinrichs, M., Neumann, I. and Ehler, U. (2002). Lactation and stress: protective effects of breast-feeding in humans. *Stress* 5(3), 195–203.
- Kirschbaum, C., Pirke, K. M. and Hellhammer, D. H. (1993). The ‘Trier Social Stress Test’—a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology* 28(1–2), 76–81.
- Kirschbaum, C., Kudielka, B. M., Gaab, J., Schommer, N. C. and Hellhammer, D. H. (1999). Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis. *Psychosomatic Medicine* 61(2), 154–162.
- Kudielka, B. M. and Kirschbaum, C. (2005). Sex differences in HPA axis responses to stress: a review. *Biological Psychology* 69(1), 113–132.
- Kudielka, B. M., Buske-Kirschbaum, A., Hellhammer, D. H. and Kirschbaum, C. (2004). Differential heart rate reactivity and recovery after psychosocial stress (TSST) in healthy children, younger adults, and elderly adults: the impact of age and gender. *International Journal of Behavioral Medicine* 11(2), 116–121.
- Kudielka, B. M., Buske-Kirschbaum, A., Hellhammer, D. H. and Kirschbaum, C. (2004). HPA axis responses to laboratory psychosocial stress in healthy elderly adults, younger adults, and children: impact of age and gender. *Psychoneuroendocrinology* 29(1), 83–98.
- Kudielka, B. M., Schommer, N. C., Hellhammer, D. H. and Kirschbaum, C. (2004). Acute HPA axis responses, heart rate, and mood changes to psychosocial stress (TSST) in humans at different times of day. *Psychoneuroendocrinology* 29(8), 983–992.
- Kudielka, B. M., Hellhammer, D. H. and Kirschbaum, C. (In press). Ten years of research with the Trier Social Stress Test (TSST) – revisited. In: Harmon-Jones, E. & Winkelman, P. (eds.) *Social Neuroscience*. New York: Guilford Press.
- Mason, J. W. (1968). A review of psychoendocrine research on the pituitary-adrenal cortical system. *Psychosomatic Medicine* 30(Supplement 5), 576–607.
- Rohleder, N., Wolf, J. M. and Kirschbaum, C. (2003). Glucocorticoid sensitivity in humans—interindividual differences and acute stress effects. *Stress* 6(3), 207–222.
- Wüst, S., Rossum, E. F. C., Koper, J. W., Kumsta, R., Entringer, S. and Hellhammer, D. H. (2004). A psychological perspective on genetic determinants of hypothalamus-pituitary-adrenal axis activity. *Annals of the New York Academy of Sciences* 1032, 52–62.
- Wüst, S., Federenko, I. S., van Rossum, E. F., Koper, J. W. and Hellhammer, D. H. (2005). Habituation of cortisol responses to repeated psychosocial stress—further characterization and impact of genetic factors. *Psychoneuroendocrinology* 30(2), 199–211.

Tuberculosis See: Macrophage Antimycobacterial Activity, Effects of Stress on.

Type A Personality, Type B Personality

W S Shaw

University of Massachusetts, Worcester, MA, USA

J E Dimsdale

University of California, San Diego, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by W S Shaw and J E Dimsdale, volume 3, pp 624–630, © 2000, Elsevier Inc.

Introduction

Historical Perspective

Conceptualization and Assessment

Empirical Support

Component Factors

Beyond Coronary Heart Disease

Conclusion

Glossary

Free-floating hostility

Behavior characterized by hostile vocal inflection and facial expression, ill temper, cynicism, sleeplessness, irritability, and interpersonal conflict.

Structured interview

An assessment technique involving a series of questions meant to elicit impatience, hostility, and competitiveness among those believed to have type A-defining characteristics.

Time urgency

Behavior characterized by extreme haste, punctuality, time monitoring, and tendencies to engage in polyphasic activities with poor recall of events.

Type A behavior

“An action-emotion complex that can be observed in any person who is aggressively involved in a chronic, incessant struggle to achieve more and more in less and less time, and if required to do so against the opposing efforts of other things and other persons” (Friedman and Rosenman, 1974: 67).

Type B behavior

A relative absence of the behavioral characteristics defined as type A.

Introduction

For centuries, observers have suggested that certain personality patterns are toxic to the heart. Early reports were based on clinical observations and, as is typical in the history of medicine, such reports were either case reports or pronouncements (in lieu of data) by respected authorities. From 1930 to 1950,

such reports were typically framed from a psychoanalytic perspective. However, shortly afterward, a new conceptualization emerged. The investigators claimed to have recognized a coronary-prone personality pattern but chose to label it in a value-neutral fashion as type A behavior pattern.

Historical Perspective

Few health-related psychological constructs have received the scientific interest and widespread public recognition of the type A personality classification. Scientific interest in type A and its relation to coronary heart disease began in the 1960s with a research team at the Mount Zion Hospital in San Francisco, California. Meyer Friedman, Ray Rosenman, and colleagues identified a constellation of personality traits that were observed disproportionately among their patients presenting with coronary heart disease (CHD). These traits (labeled type A or coronary-prone) included excessive drive, time urgency, ambition, impatience, aggressiveness, hostility, and competitiveness. When confronted with stressful situations, type A individuals were believed to react with rapid, explosive retorts and increased muscle tension. These behaviors were thought to cause heightened or excessive wear on the cardiovascular system and thereby increase the risk of CHD. The relative absence of these traits defined type B or non-coronary-prone behavior.

Clinical reports suggesting a coronary-prone behavior pattern led to a multitude of epidemiological and experimental studies in the 1970s and 1980s to examine these relationships in the general population and to elucidate biological mechanisms. Assessment techniques were developed, including structured interviews, self-report questionnaires, and behavioral observations. Studies examined the relationship of type A to CHD incidence, angina, myocardial infarction (MI), atherosclerosis, and cardiovascular reactivity. Early epidemiologic studies supported an association between type A and CHD, and this led to a 1981 consensus statement by the Review Panel on Coronary-Prone Behavior and Coronary Heart Disease (organized by the National Heart, Lung, and Blood Institute) listing type A as a viable risk factor for heart disease.

Despite the early support for type A, a proliferation of research in the 1980s failed to replicate earlier findings, and several related psychological constructs (hostility, anger, and aggression) emerged as stronger predictors of CHD. As a result, interest in type A has waned in recent years. By the late 1980s, two

meta-analyses concluded that the accumulated empirical studies were inconsistent and that prospective evidence linking type A to CHD was weak. Since that time, behavioral CHD research has emphasized hostility and other related constructs over the global type A characterization. However, originators of the type A concept continue to defend its role as a CHD risk factor, blaming poor assessment techniques and a substantial drift from its original conceptualization for the discrepant research findings.

As late as the 1990s, interest in the health implications of type A behavior continued, albeit in a more limited scope. Meyer Friedman and his colleagues at the University of California, San Francisco, continued to refine clinical assessment techniques and develop interventions for reducing type A behavior. Others sought to examine type A behavior with respect to other psychosomatic health problems. Today, there is declining research interest in type A as a behavioral health risk. **Figure 1** illustrates the rapid rise and fall of this construct as a title word in psychological research publications over the past 40 years. Nevertheless, type A has served as an important catalyst for the explosion of scientific and public interest in behavioral factors associated with health, especially cardiovascular disease.

Conceptualization and Assessment

Considerable conceptual confusion has surrounded the type A construct. Type A behavior has been variously described as a medical disorder, an inappropriate coping mechanism, a strong need for productivity, or a pattern of physiological reactivity. Although most research has presumed type A behavior constitutes an enduring personality trait, other studies have

attributed type A behavior to specific cognitions that can be modified through therapeutic interventions. Much of the accumulated discrepancy in results between studies of type A may be attributable to poor agreement in construct definition and assessment.

Type A behavior was first assessed through clinical ratings applied to structured interviews or videotaped structured interviews, and these interview techniques have been further updated and expanded by Meyer Friedman and colleagues. The current technique combines both observations and subjective reporting of patients using eliciting remarks. The interview consists of multiple queries that are designed to elicit manifestation of time urgency and free-floating hostility that emerge for type A, but not for type B, individuals. Questions include “Do you mind very much waiting in grocery checkout, bank, or theater lines or waiting to be seated in a restaurant?” (time urgency) and “Do you often find it difficult to fall asleep or to continue to sleep because you are upset about something a person has done?” (hostility). Scoring is based on both endorsements of items and the presence of psychomotor signs. For time urgency, signs include aspects of facial tension, posture, speech, breathing, and perspiration. For free-floating hostility, these include facial expression, eyelid movement, vocal orientation, and hand clenching. For each item, scores are clinical ratings varying in assigned weights and ranges. Total scores range from 0 to 480, and a total score greater than 45 indicates the presence of type A traits.

As an alternative to the interview technique, self-report measures of type A were developed to simplify data collection for use in large-scale epidemiological studies of CHD. One of the earliest and most frequently used questionnaires is the Jenkins Activity Scale (JAS). Other pen-and-paper measures of type A behavior include the Framingham Type A Scale, the Bortner Scale, the MMPI-2 Type A Scale, the Gough Adjective Checklist, the Thurstone Activity Checklist, the Coronary Prone Attitudes Scale, and the Ketterer Stress Symptoms Frequency Checklist.

The content and emphasis of type A self-report measures vary substantially. The JAS includes subscales of hard-driving competitiveness, speed and impatience, and job involvement. The Framingham Type A Scale includes time urgency, competitive drive, and perceived job pressures, but does not include hostility. The Bortner scale contains 14 adjective pairs between which respondents mark the level of agreement with either type A or type B traits (e.g., “Never late” versus “Casual about appointments”). The MMPI-2 Type A Scale includes 19 items that assess the type A components of hostility, competitiveness, and time urgency; these were selected from the

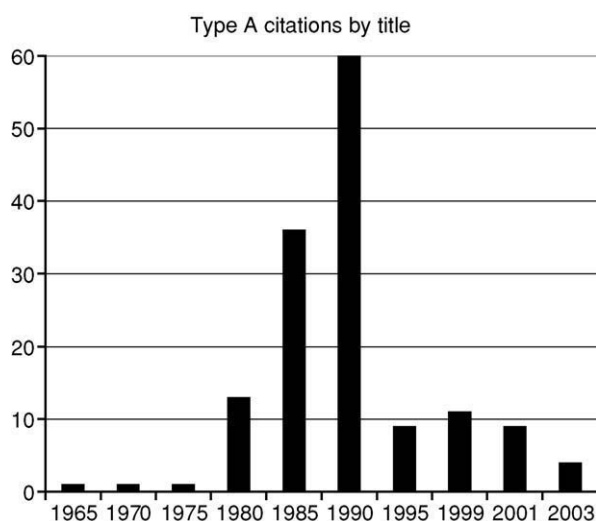


Figure 1 Type A journal title citations.

total pool of MMPI-2 questions by expert consensus and verified using empirical tests of item-to-type A group classification. The Ketterer Scale compares self-ratings of type A characteristics with parallel ratings of a friend or family member.

The reliability of the various interview and self-report measures is quite good. For example, interrater correspondence for discriminating type A and B individuals using the videotaped structured interview has ranged from 75 to 90%. For the MMPI-2 Type A Scale, test-retest reliability is high ($r = 0.82$), and internal consistency is moderately high ($\alpha = 0.72$), at least for male respondents. Other self-report measures of type A show similar levels of reliability. Prevalence estimates of type A behavior have varied from 50 to 75% depending on the populations studied and assessment techniques employed.

The validity of self-report measures of type A were originally documented by comparing scores between type A and type B individuals already categorized using the gold-standard interview technique. Although both interview and self-report assessment techniques have good reliability, the two approaches are only modestly interrelated. For example, the JAS showed a 73% agreement with the type A interviews in a cross-validation sample of 419 men participating in the Western Collaborative Group Study. Similarly, the Bortner scale was shown to explain only 53% of the variability in type A clinical interview ratings. This has led to considerable debate as to whether self-report questionnaires of type A are a valid means of describing the same constellation of features identified by the structured interview. Studies using self-report measures have generally provided less association with cardiovascular health outcomes than have interview techniques. As a result, self-report measures of type A have been criticized for relying on the insight of respondents, restricting the range of type A traits, and being influenced by the social undesirability of many items. Because self-report measures do not involve a live interaction with the individual, these scales may not accurately assess the action-emotion complex described by the creators of the interview technique.

Empirical Support

Data on the association between type A and CHD have been inconsistent. Three early large-scale epidemiological studies supported an association between type A and CHD incidence: the Western Collaborative Group in 1975, the Framingham Heart Study in 1980, and the French-Belgian Cooperative Study Group in 1982. All three investigations found that

type A individuals had a significantly increased risk of CHD in individuals followed closely for 5–8 years. In the Western Collaborative Group study, 3200 employed men with no heart disease were followed for 8 years. Type A individuals (by structured interview) were found to have a 2.2 relative risk of CHD relative to type B individuals. The Framingham Heart Study, which included a more representative sampling of both men and women, showed relative risk ratios for white-collar men of 2.9 for total CHD, 7.3 for MI, and 1.8 for angina without MI; however, risks were substantially less for women and for blue-collar men. The French-Belgian Cooperative Heart Study followed 2811 male civil servants and factory workers for an average of 5 years. The relative risk ratio for type A men was 1.8 for total CHD. The latter two studies used self-report measures of type A.

Other subsequent epidemiological studies showed little or no relationship. For example, in the Normative Aging Study, correlations of the JAS type A score with CHD risk were not statistically significant. In a 40-year follow-up of 280 men from the 1947 Cardiovascular Disease Project at the University of Minnesota, no statistically significant association was found between the MMPI Type A Scale and CHD. In a large-scale 7-year prospective study (Multiple Risk Factor Intervention Trial, MRFIT), there was no association between type A and risk of coronary events, and this applied to both interview and self-report methods of type A classification. The type A-CHD relation has been much smaller in prospective versus cross-sectional studies, in meta-analyses synthesizing results from multiple studies, and among more diverse groups. Although many plausible explanations have been offered for differences among study results, meta-analyses and other attempts to synthesize results across studies have failed to provide consistent support for type A as a unique and valid predictor of CHD.

The generalizability of initial research linking type A to CHD has also been criticized because early studies were limited to middle-aged, working, White American males. Among more diverse populations, associations between type A and CHD have been weaker or nonexistent. Studies of both type A and hostility in Japan, for example, have shown little association with CHD. Instead, a type C behavior pattern, characterized by a job-centered lifestyle and social dominance at work, was predictive of CHD in working Japanese men. The type A construct may not apply for women due to the greater social unacceptability of type A traits for women.

Intervening biological mechanisms to explain the type A-CHD association have focused primarily on

the hyperreactivity of the cardiovascular and neuroendocrine systems; however, results have been inconsistent. Quantitative reviews have suggested some heightened reactivity in systolic blood pressure response and catecholamine response, but results vary by gender, type A-assessment technique, and the nature of laboratory stressors. Ambulatory blood pressure studies have also indicated some relationship between type A and cardiovascular reactivity. Type A scores have been related to the actual extent of coronary artery atherosclerosis, as judged from cardiac catheterization. Here again, initial studies were positive, but subsequent studies failed to reveal such links.

Other studies have included type A as a potential mediator of health behavior. For example, type A individuals underestimate their degree of physical exertion in response to exercise. They also work longer and at greater intensity when given a self-selected work-pace task. This may explain some of the relationships between type A and cardiovascular health outcomes such as recurrent MI or unexplained cases of sudden death.

Despite the huge accumulated literature on type A, hostility, and CHD, interventions targeting type A behavior are rare. One exception is a 14-month type A counseling approach developed at Mount Zion Hospital purportedly to change belief systems, engage in exercises to modify a sense of time urgency or free-floating hostility, and increase perceptions of personal security. This intervention has been reported to reduce the frequency of silent myocardial ischemia; however, the studies have been performed on small nonrepresentative samples of post-MI coronary patients. It is not known which therapeutic techniques included in this broad-based intervention were most effective.

Component Factors

Because the interview assessment of type A included a number of signs and symptoms, researchers became interested in the specific component factors that might contribute to CHD risk and whether type A was unique in this respect compared with other personality variables. These investigations ultimately led to a shift in the focus of CHD behavioral studies away from the type A global constellation of symptoms. When the total score and subscale scores of the JAS were contrasted, the hard-driving competitiveness factor was more strongly associated with cardiovascular disease than was speed and impatience, job involvement, or the overall type A score. Other studies showed support for hostility as the key underlying component. For example, of 12 behavioral characteristics extracted from the structured interview,

a reanalysis of the Western Collaborative Group Study data showed that the hostility component alone accounted for all of the variance explained by type A.

In their 1987 meta-analysis, Booth-Kewley and Friedman identified 18 overlapping personality variables from various assessment measures for the purposes of comparing relative strengths of association. These included global type A scores; subscale measures of speed and impatience, job involvement, hard-driving competitiveness, and time urgency; and non-type A measures of anger, hostility, aggression, depression, extraversion, and anxiety. These analyses showed that depression was more strongly associated with CHD than was type A, and CHD correlations with hostility, aggression, and extraversion were of a similar magnitude to type A. This meta-analysis and subsequent studies ultimately led to the conclusion that only the type A components related to anger, hostility, and aggression were responsible for increased CHD risk. In particular, hostility, which had long been suspected as a risk factor for CHD, was found to explain nearly all the shared variance between type A and CHD.

Today, the expanding interest in hostility as a predictor of CHD has eclipsed the original global type A as a potential risk factor. However, conceptual and assessment controversies continue as various components of hostility are hypothesized to underlie associations with CHD. Cynical hostility, as measured by the Cook-Medley scale of the MMPI, was shown to predict a fivefold increase in CHD risk in a 25-year follow-up of 255 medical students. Suppressed anger was associated with a twofold increase in mortality among 696 Michigan residents followed over 12 years. Studies linking hostility to CHD have shown associations with angiographically documented CHD. However, their cross-sectional nature precludes definitive conclusions regarding cause and effect. Prospective studies with 5- to 25-year follow-ups have shown associations between hostility and later CHD development, but similar associations exist with all-cause mortality. Therefore, the specificity of hostility with regard to CHD remains unclear.

Beyond Coronary Heart Disease

Although the focus of type A research has been primarily on CHD, hypertension, and related cardiovascular health outcomes, several studies have examined the impacts of type A behaviors in the context of entirely different medical conditions. These have included chronic fatigue syndrome, irritable bowel syndrome, and other health problems thought to involve

psychological risk factors. These studies have generated no consistent pattern of findings that implicate type A as a central causal factor, although it periodically emerges as a significant predictor of various health outcomes. The most common explanation for this finding is that type A behavior may play a mediating role in health behaviors (exercise, diet, smoking, etc.).

Despite the popular use of the type A label to describe the overachieving efforts of co-workers and family members, there have been relatively few studies of the interpersonal and occupational consequences of type A behavior. In cross-sectional studies, type A overlaps with other measures of personality and temperament (e.g., narcissism, optimism, and aggression), yet type A has not emerged as an independent mental health risk factor. In relation to perceptions of stress and satisfaction at work and home, the impacts of type A behavior have been inconsistent. Based on the available evidence, the type A constellation of personality traits is an uncertain predictor of social and emotional well-being.

Conclusion

Any fair-minded observer would be immediately impressed by the numerous links between behavior and coronary disease. The impact of various risk-enhancing and risk-reducing behaviors is unchallenged. Similarly, the impact of unusual life stressors in provoking a deterioration in already-existing CHD is reasonably established. It is less clear that a habitual personality or behavior pattern leads to the development of CHD. Ambiguities in recognizing that personality and defining its central features have plagued this area. Nonetheless, data do suggest some links between type A behavior and CHD. Current work focuses less on type A itself than on the quest for its toxic core, its mechanism of action, and the possibilities of intervention.

See Also the Following Articles

Aggression; Anger; Blood Pressure; Cardiovascular System and Stress; Heart Disease/Attack; Hostility; Metabolic Syndrome; Self-Esteem, Stress and Emotion.

Further Reading

- Booth-Kewley, S. and Friedman, H. (1987). Psychological predictors of heart disease: a quantitative review. *Psychological Bulletin* **101**, 343–362.
- Dimsdale, J. (1988). A perspective on type A behavior and coronary disease. *New England Journal of Medicine* **318**, 110–112.
- Friedman, M. and Ghandour, G. (1993). Medical diagnosis of type A behavior. *American Heart Journal* **126**, 607–618.
- Friedman, M. and Powell, L. H. (1984). The diagnosis and quantitative assessment of type A behavior: introduction and description of the videotaped structured interview. *Integrated Psychiatry* **2**, 123–129.
- Friedman, M. and Rosenman, R. (1974). *Type A behavior and your heart*. New York: Knopf.
- Gallacher, J. E. J., Sweetnam, P. M., Yarnell, J. W. G., et al. (2003). Is type A behavior really a trigger for coronary heart disease events? *Psychosomatic Medicine* **65**, 339–346.
- Grunbaum, J., Vernon, S. W. and Clasen, C. M. (1997). The association between anger and hostility and risk factors for coronary heart disease in children and adolescents: a review. *Annals of Behavioral Medicine* **19**, 179–189.
- Harbin, T. J. (1989). The relationship between the type A behavior pattern and physiological responsivity: a quantitative review. *Psychophysiology* **26**, 110–119.
- Haynes, S. G., Feinleib, M. and Kannel, W. B. (1980). The relationship of psychosocial factors to coronary heart disease in the Framingham Study. III: Eight-year incidence of coronary heart disease. *American Journal of Epidemiology* **111**, 37–58.
- King, K. B. (1997). Psychologic and social aspects of cardiovascular disease. *Annals of Behavioral Medicine* **19**, 264–270.
- Matthews, K. and Hanes, S. (1986). The type A behavior patterns and coronary disease risk: update and critical evaluation. *Journal of Epidemiology* **123**, 923–960.
- Melamed, S., Harari, G. and Green, M. S. (1993). Type A behavior, tension, and ambulatory cardiovascular reactivity in workers exposed to noise stress. *Psychosomatic Medicine* **55**, 185–192.
- Review Panel on Coronary-Prone Behavior and Coronary Heart Disease (1981). Coronary prone behavior and coronary heart disease: a critical review. *Circulation* **63**, 1199–1215.
- Thoreson, C. E. and Powell, L. H. (1992). Type A behavior pattern: new perspectives on theory, assessment, and intervention. *Journal of Consulting & Clinical Psychology* **60**, 595–604.



Ubiquitination See: Proteasome; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles.

Ulceration, Gastric

R Murison and A M Milde
University of Bergen, Bergen, Norway

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition
article by R Murison, volume 3, pp 631–633,
© 2000, Elsevier Inc.

Background
Psychological Influences on Development of Ulceration
Mechanisms
Future Directions

Glossary

<i>Analgesia</i>	Insensibility to painful stimuli.
<i>Corticotropin releasing hormone</i>	A hormone released from the hypothalamus which leads to secretion of cortisol from the adrenal cortex into the circulation.
<i>Dopamine</i>	A catecholamine neurotransmitter.
<i>Gastric acid</i>	Hydrochloric acid secreted by cells in the stomach.
<i>Helicobacter pylori</i> (<i>H. pylori</i>)	A gram-negative bacterium found in the majority of patients with peptic ulcer.
<i>Immuno-modulation</i>	Changes in the function of the body's immune defence system.
<i>Ischemia</i>	Localized oxygen deficit as a result of obstruction of the blood supply or vasoconstriction.
<i>Lactobacillus</i>	A bacterium found particularly in dairy produce and meat.
<i>Limbic</i>	Referring to several phylogenetically old brain structures thought to be central to emotionality and learning.
<i>Mucosa</i>	The mucous membrane lining the stomach and intestines.

<i>Pepsinogen</i>	The precursor of pepsin, found in the stomach mucosa.
<i>Peptic ulcer</i>	An ulcer involving the mucosa, submucosa, and muscular layers of the stomach, or duodenum, due in part at least to the action of gastric juices.
<i>Rumen</i>	The part of the stomach lacking acid-producing glands.
<i>Spontaneously hypertensive rats</i>	A specially bred strain of rats with high susceptibility to develop high blood pressure.
<i>Sympathetic</i>	The part of the autonomic nervous system which prepares the organism for the fight/flight response.
<i>Vagotomy</i>	Severing of the vagus nerve.
<i>Vagus nerve</i>	The tenth cranial nerve forming an important part of the parasympathetic system.

Background

Ulcerations of the stomach, typically manifested as nonpenetrating erosions of the gastric mucosa, are found following stress in many mammalian species (mice, rats, horses, primates, swine, reindeer, and dogs, among others). The condition is therefore considered to be a reflection of fundamental common psychobiological processes activated by stress. Although there are significant exceptions, most psychobiological studies have focused on the rat because of the extensive pool of knowledge of behavior and central nervous system (CNS) function in that species. The ulcerations are found in both the rumen and the glandular areas of the stomach. The former appear to be the result of food deprivation procedures, whereas the latter are sensitive to psychological factors (stress). The ulcerations are typically bleeding erosions of the superficial gastric mucosa along the

ridges of the gastric folds. Prolonged application of the stressor may cause full-blown ulcers, with the erosions reaching through the muscularis mucosa. Given a respite from the stressor and other potential ulcerogenic factors, however, the ulcerations typically heal within a few days.

Controversy surrounds the homologous condition in humans. Earlier, gastric ulcerations in rodents and monkeys were used as models of peptic ulcer in humans, even though the location of ulcer in humans is more typically the duodenum. Later, gastric ulcerations were regarded as homologous to the acute stress ulcers seen after burn injuries or traumatic surgery in humans. Since the discovery of the role of *Helicobacter pylori* in peptic ulcers in humans, it has been surmised that ulcerations in animals are homologous to the erosive prepyloric changes of the gastric mucosa reported in subgroups of patients diagnosed with functional dyspepsia. However, there is much more to the story.

Psychological Influences on Development of Ulceration

In the wild, ulcerations are observed in animals (both rodents and primates) following social defeat. In both wild and domestic animals, ulcerations have been reported in connection with the stress of transport and confinement. In the laboratory, ulcerations are produced by a wide variety of stressors in animals – electric shock, restraint, cold, activity stress, social stress, or combinations of these – with the purpose of exploring the role of psychological factors that may modulate the incidence and/or severity of the mucosal injury. Consciousness is a modulating, if not necessary, condition for the development of ulcerations.

Ulcerations played a key role in Selye's description of the general adaptation syndrome and constituted one of the elements of this nonspecific response to diverse noxious agents, together with shrinkage of the thymus and enlargement of the adrenal glands.

Acute Effects

To identify the psychological factors in ulceration (and other stress-related phenomena), many have employed the triadic design. Animals are tested in triplets. One animal (the executive) is subjected to a physical stressor (most commonly electric shock) that can be escaped from or avoided by instrumental responses. The second (yoked) animal is wired to the first such that the same physical stress is experienced but there is no possibility of avoidance or escape. The third (control) animal is subjected to the same apparatus conditions as the first two but not to the

stressor. By comparing the amount of ulceration between the first two animals, the psychological and physical components of the stressor can be identified.

The Executive Monkey studies with monkeys, by Brady and his colleagues, were based on a triadic design, and indicated that the executive was more vulnerable to ulceration than its yoked partner. These results were widely accepted at the time, but their interpretation was later questioned. In contrast to the Brady results, Weiss reported that executive rats developed less ulceration than their yoked partners. These inconsistencies between the monkey and rat results have been attributed to the particular group-assignment procedures and experimental parameters used by Brady rather than to species differences.

Later, Weiss demonstrated that ulceration in rats was attenuated by providing the animals with predictability for shock and was exacerbated by either increasing the response requirements (effort) required to maintain instrumental control or by introducing a component of psychological conflict. Weiss's model predicts ulceration as a function of the effort required to maintain instrumental control over a stressor and an inverse function of the feedback provided by such control. Successful coping behaviors (high feedback) ameliorate ulceration, but greater workload (effort) will increase ulceration. This model has played an important role in the development of recent theories concerning the relationships among stress, coping, and ill health in humans. Between- and within-strain studies have led to the notion that more emotional animals are more vulnerable to ulcerations than less emotional conspecifics. For example, spontaneously hypertensive rats are less vulnerable than the more emotional Wistar-Kyoto strain. Potential biological markers for vulnerability to ulceration include plasma levels of pepsinogen (also in humans) and dopaminergic function. Genetic factors certainly play a role in vulnerability to ulceration, but these are modulated by early life experiences and previous experiences of coping/noncoping, which also modulate dopaminergic function.

Proactive Effects

A number of experimenters have investigated proactive effects of one or more stressor exposures on vulnerability to ulceration under later exposure to either the same or a different stressor. Rats previously subjected to the stress of whole-body restraint generally develop less ulceration under later restraint, whereas previous exposure to an activity wheel primes the animal for later activity-stress ulceration. Using heterogeneous stress experiences, rats subjected to the same shock stress parameters that are used to induce learned helplessness (a putative model for

human depression) develop greater amounts of ulceration under a later radically different type of stressor, restraint in water. This phenomenon is mediated by opioid mechanisms in the same way that stress-induced analgesia and stress-induced immunomodulation are also invoked by these same shock parameters. Previous experience with escapable shock, on the other hand, has proactive protective effects against later ulceration, although the mechanisms of this phenomenon remain unexplored. More recent studies have complemented these earlier reports in showing that prior experience with seminaturalistic coping responses is protective against later ulcerogenesis. Similarly, prior experience with uncontrollable shocks, but in which safety signals are provided at the end of each shock, also reduces vulnerability under later stress.

Early weaning of rats (at day 16) leads to a drastic increase in vulnerability to ulceration during a limited period later on (22–40 days). Although later studies indicated that this effect was due more to nutritional and thermoregulatory influences than to psychobiological phenomena, recent reports suggest that such very early manipulations have profound long-term effects on dopaminergic system function, which is known to play an important role in ulceration.

Mechanisms

Peripheral Mechanisms

The peripheral mechanisms of ulceration were originally believed to be limited to increased levels of vagally mediated gastric acid secretion, and vagotomy does reduce ulceration. Although the old saying “no acid – no ulcer” still holds, recent approaches have focused on combinations of changes in gastric motility, gastric blood flow, and sympathetic tone. A common view is that enhanced vagally mediated acid secretion and reduced gastric blood flow represent key aggressive factors, whereas the protective factors include prostaglandin-mediated secretion of mucous, secretion of bicarbonate, and maintenance of an adequate blood supply.

The development of ulceration appears to be greatest in the period immediately following an acute stress. This has led to the concept of ulceration as a poststress or rebound phenomenon, whereby gastric acid secretion is inhibited during the application of a strong stressor (although certain stressors may increase secretion). A concurrent strong sympathetically mediated constriction of the mucosal blood vessels increases the risk for local ischemia, and slow rhythmic contractions may cause further physical damage. On release from the inhibitory (sympathetic) influence of

the acute stressor, gastric acid levels are rapidly elevated, leaving a vulnerable mucosa open to damage.

Interactions between stress and bacterial flora The role of bacterial flora is now under intensive study since the discovery that ulcer disease in humans is related to the presence of *H. pylori*. Animal psychobiological models for *Helicobacter* infection have been hindered by the fact that most behavioral studies of ulceration have been performed in rats, whereas the models for infection have largely been limited to mice, ferrets, and gerbils. It remains unclear whether bacterial flora play any part in either ulceration in animals or erosive prepyloric erosions in humans. Germ-free animals do develop stress gastric ulcerations, although to a lesser extent than normal animals, and antibiotic treatment accelerates the healing of the ulceration. From the human data, it is apparent that factors additional to *H. pylori* are involved in gastric pathology, and it remains to be clarified to what extent stress, either alone or in combination with bacterial factors, is involved. Recent results also suggest that the relationship between the bacterial population of gut and ulceration is two way: The induction of ulceration in animals by the direct application of acetic acid to the stomach lining dramatically changes the predominant bacterial species, which then return to normal as the ulceration heals. In particular, *Lactobacillus* seems to be protective and enhances the healing process. Consistent with these animal data, a small-scale study in humans has shown that *H. pylori*, although present in all patients with duodenal ulceration with a duration greater than 6 months, was not present in patients with a shorter history, suggesting that the infection is not so much a cause as a factor producing chronicity. On the other hand, a recent study of patients admitted to intensive care units shows that *H. pylori* infection was more frequent in patients exhibiting upper gastrointestinal bleeding than in patients who were not bleeding. Thus, the controversy concerning the role of *H. pylori* remains unresolved.

Central and Neuroendocrine Influences

Studies of the central mechanisms of ulcerations have focused on limbic structures, in particular the amygdala (with its close association with the vagus nerve), the septum, and the hippocampus. The stimulation of the amygdala elicits ulceration, whereas lesions of the central nucleus of the amygdala attenuate their formation. The stimulation of the dorsolateral septum reduces the amount of ulceration induced by cold-restraint stress. Lesions of the hippocampus, in particular the ventral hippocampus, exacerbate

ulceration, presumably acting through the amygdala-vagal pathways.

Studies of centrally acting peptides show that corticotropin releasing hormone (CRH) has a potent protective effect when applied directly in the central amygdala. When injected peripherally, CRH protects against stress ulceration in young, but not old, rats (in old rats, the effect is exacerbatory). CRH acts through sympathetic pathways to inhibit gastric acid secretion while at the same time stimulating bicarbonate production and inhibiting slow gastric contractions. Tests of the CRH type-1 receptor antagonist antalarmin demonstrate a potent anti-ulcerogenic effect, but the glucocorticoid/progesterone antagonist RU-38486 significantly potentiates the formation of stress-induced gastric erosions. The role of the CRH system in ulcerogenesis remains unclear and peripheral corticosterone may play an important role in cytoprotection during stress.

Thyrotropin releasing hormone (TRH) acts as a potent ulcerogenic agent and is stimulated by cold. TRH promotes both gastric acid secretion and the slow gastric contractions associated with the development of ulcerations induced by cold restraint stress. The main site of action for TRH appears to be the dorsal motor nucleus of the vagus, although a number of hypothalamic nuclei and the amygdala are also implicated.

Adrenalectomy increases the incidence of ulceration, and there is little support for the earlier notion that corticosterone plays any causal role in ulcerogenesis. Rather, the events likely to activate the pathogenic mechanisms for ulceration are also the kind of events that activate the hypothalamic-pituitary-adrenal axis.

Of the centrally acting neurotransmitters, most attention has been paid to dopamine and in particular to the D1 and D2 receptors. In line with the observation that schizophrenics rarely develop ulcers, whereas untreated Parkinson's patients are at high risk, numerous studies have provided evidence that central dopamine is protective against ulceration. A more complex picture has now become apparent because results suggest that the stimulation of dopamine D1 receptors inhibits ulceration whereas stimulation of dopamine D2 receptors has a pro-ulcerogenic effect. Within-strain comparisons of rats rated as high emotional versus low emotional reveal that the former are more vulnerable to stress ulcerations than the latter, as well as having lower levels of dopamine in the amygdala. A similar difference between these two types of animal was observed with respect to ethanol-induced ulcerations. Rats differing in their sensitivity to the dopaminergic agonist apomorphine are differentially susceptible to ulcerations and recover at

different rates, and experience with shock parameters demonstrated to increase vulnerability to ulceration change their apomorphine sensitivity accordingly. Recent evidence also suggests that the dopamine content of the gastric mucosa is important in cell proliferation of the gastric mucosa. These and other data suggest that vulnerability profiles of animals to ulcerations are manifest at both the central and peripheral levels.

Future Directions

Studies on the etiology of stress gastric ulceration have become reduced in number since the experience that *H. pylori* eradication is the most effective treatment for human peptic ulcer. However, the bacterial risk factor and the effectiveness of treatment do not in themselves account for either historical or contemporary links between stress and ulcerogenesis. Recent developments in psychoneuroimmunology offer opportunities to study how host factors (including CNS-mediated immunomodulation) may interact with pathophysiological agents, both bacterial and others, in causing and maintaining gastric mucosal changes. Specifically, stress-related failures of immune defenses may provide a window of opportunity for *H. pylori* to cause problems.

See Also the Following Article

Gastrointestinal Effects.

Further Reading

- Boulos, P. B., Botha, A., Hobsley, M., et al. (2002). Possible absence of *Helicobacter pylori* in the early stages of duodenal ulceration. *Quarterly Journal of Medicine* 95, 749–752.
- Degen, S. B., Geven, E. J., Sluyter, F., et al. (2003). Apomorphine-susceptible and apomorphine-unsusceptible Wistar rats differ in their recovery from stress-induced ulcers. *Life Sciences* 72, 1117–1124.
- Desai, J. K., Goyal, R. K. and Parmar, N. S. (1999). Characterization of dopamine receptor subtypes involved in experimentally induced gastric and duodenal ulcers in rats. *Journal of Pharmacy and Pharmacology* 51, 187–192.
- Elliott, S. N., Buret, A., McKnight, W., et al. (1998). Bacteria rapidly colonize and modulate healing of gastric ulcers in rats. *American Journal of Physiology* 275, G425–G432.
- Filaretova, L., Podvigina, T., Bagaeva, T., et al. (2001). Gastroprotective action of glucocorticoids during the formation and the healing of indomethacin-induced gastric erosions in rats. *Journal of Physiology (Paris)* 95, 201–208.

- Gabry, K. E., Chrousos, G. P., Rice, K. C., et al. (2002). Marked suppression of gastric ulcerogenesis and intestinal responses to stress by a novel class of drugs. *Molecular Psychiatry* 7, 433, 474–483.
- Glavin, G. B., Murison, R., Overmier, J. B., et al. (1991). The neurobiology of stress ulcers. *Brain Research Reviews* 16, 301–343.
- Levenstein, S. (2000). The very model of a modern etiology: a biopsychosocial view of peptic ulcer. *Psychosomatic Medicine* 62, 176–185.
- Maury, E., Tankovic, J., Ebel, A., et al. (2005). An observational study of upper gastrointestinal bleeding in intensive care units: is *Helicobacter pylori* the culprit? *Critical Care Medicine* 33, 1513–1518.
- Murison, R. (2001). Is there a role for psychology in ulcer disease? *Integrative Physiological and Behavioral Science* 36, 75–83.
- Overmier, J. B. and Murison, R. (2000). Anxiety and helplessness in the face of stress predisposes, precipitates, and sustains gastric ulceration. *Behavioral Brain Research* 110, 161–174.
- Sapolsky, R. M. (2004). *Why zebras don't get ulcers* (3rd edn.). New York: Henry Holt.
- Szabo, S. (1979). Dopamine disorder in duodenal ulceration. *Lancet* 2(8148), 880–882.
- Weiner, H. (1991). From simplicity to complexity (1950–1990): the case of peptic ulceration. I: Human studies. *Psychosomatic Medicine* 53, 467–490.
- Weiner, H. (1991). From simplicity to complexity (1950–1990): the case of peptic ulceration. II: Animal studies. *Psychosomatic Medicine* 53, 491–516.

Ultradian Rhythms

S L Lightman

University of Bristol, Bristol, UK

© 2007 Elsevier Inc. All rights reserved.

Glossary

<i>Adrenocorticotrophic hormone (ACTH)</i>	A peptide secreted by the anterior pituitary gland that stimulates the secretion of adrenal glucocorticoids; also called corticotropin.
<i>Arginine vasopressin (AVP)</i>	A hypothalamic peptide that acts at the kidney to moderate water excretion (hence, also termed the antidiuretic hormone) and acts synergistically with CRH to stimulate ACTH release from the anterior pituitary gland.
<i>Corticotropin releasing hormone (CRH)</i>	A brain peptide that mediates the neural control of ACTH release from the anterior pituitary gland and that has also been implicated in central neural actions related to anxiety and depression.
<i>Follicle-stimulating hormone (FSH)</i>	An anterior pituitary gonadotropin responsible for stimulating the growth of ovarian follicles and the secretion of estrogen in the female and the development of sperm in the male.
<i>Glucocorticoid receptor (GR)</i>	A receptor for glucocorticoids, of which cortisol is the main type in the human.
<i>Gonadotropin-releasing hormone (GnRH)</i>	A peptide that is secreted by the hypothalamus and that mediates the neural control of pituitary gonadotropin secretion.

Growth hormone (GH)
Hypothalamic-pituitary-adrenal (HPA) axis

An anterior pituitary hormone that controls body growth and metabolism.
One of the two main neuroendocrine stress response systems.

Janus kinase (JAK)
Luteinizing hormone (LH)

A kinase with two phosphorylation sites, only one of which appears to be active.
An anterior pituitary peptide that induces ovulation and the secretion of progesterone in the female and the secretion of testosterone in the male.

Mineralocorticoid receptor (MR) mRNA

A receptor for mineralocorticoid steroid hormones, of which aldosterone is the prototype.
A molecule that transcribes the genetic code from DNA.

Signal transducer and activator of transcription 5 (STAT5)

A transcription factor implicated in cell differentiation and tumor formation.

Frequency encoding of intercellular signals is a well-accepted mode of communication between neurons. More than this, however, it is actually a common mechanism of communication across a broad range of both inter- and even intracellular moieties. Even an organism as primitive as the slime mold (*Dictyostelium discoideum*) aggregates only in response to external pulses of cAMP delivered with a periodicity

of 5 min and not to constant stimuli or frequencies greater than every 2 min.

Within mammalian systems, frequency encoding is mediated by circulating hormones as major signaling molecules, and many of these signal through the mechanisms of ultradian rhythmicity. Thus, pulsatile GnRH release is essential for maintenance of LH and FSH secretion, and the modulation of GnRH pulse frequency can produce differential LH and FSH secretion via regulation of LH β and FSH β mRNA expression. Similarly, sex differences in episodic release of growth hormone (GH) elicit significant differences in gene expression. High-amplitude GH bursts occur at intervals of 3–4 h in male rats, whereas more frequent low-amplitude pulses occur in females; these sexually diergic patterns are responsible for marked differences in the expression of liver enzymes and JAK-dependent STAT5 responses.

Recent advances in our understanding of glucocorticoid receptor function suggest potential mechanisms for similar regulation of gene expression in cells exposed to pulsatile changes in glucocorticoid concentration. At the level of the glucocorticoid receptor itself, Dekelbab et al described in 2004 a rapid downregulation of GR associated with prolonged exposure to glucocorticoids that can be avoided if corticosterone is administered intermittently. This suggests a more effective induction of GR target genes from intermittent pulses of glucocorticoids. Furthermore, it is now recognized that the binding of the GR to the promoter regions of target genes is not the static process that was previously accepted. The availability of GR tagged with fluorescent tags and the advent of chromatin immunoprecipitation (ChIP) techniques have revealed that the interaction among the GR, DNA, and other proteins involved in transcription is a highly dynamic process. Indeed the effective concentration of nuclear receptors on some promoters appears to vary after hormone stimulation in a wavelike or cycling fashion. Métivier et al. described a transcriptional clock with another steroid ligand receptor – the estrogen receptor – engaging in 40-min cyclic occupation of a target promoter.

The HPA axis shows classic hormone rhythmicity. In the basal state, normal HPA activity is characterized both by a circadian rhythm of cortisol and corticosterone and also an ultradian rhythm that results in a pulsatile release of ACTH and glucocorticoids throughout the 24-h cycle. This ultradian rhythmicity is clearly a fundamental feature of HPA activity and *ex vivo* evidence from the monkey suggests that the pulse generator is located within the hypothalamus. This pulse generator results in the phasic release of CRH and AVP from the parvocellular neurons of

the hypothalamic paraventricular nucleus into the portal blood, from which it has access to pituitary corticotrophs that secrete ACTH, which in turn regulates the production and release of glucocorticoids from the fasciculata cells of the adrenal cortex.

Details of studies on the intrinsic dynamics of HPA function in free-running conscious animals reveals a complex ultradian rhythm of corticosterone secretion of approximately 1 h in duration. This profile shows not only marked sexual diergism, but also shows marked plasticity in different physiological states such as lactation and aging, the presence of or susceptibility to disease, and early life programming. Similar ultradian rhythmicity is found in humans, resulting in complex patterns of cortisol secretion that have made the response to stress testing difficult to interpret.

The rapidly changing concentrations of glucocorticoids have major potential implications with respect to the pharmacokinetics of receptor occupancy and signal transduction because, over relatively short periods of time, there are widely changing concentrations of ligand available for both MR and GR binding. Although there is little data on this, it appears likely that this is of a special relevance for GR occupancy because the high affinity of the MR suggests that short-term changes in ligand concentration have only relatively small effects on MR signal transduction.

Pulsatility provides an added dimension to the regulation of hormone responsive pathways. Not only can it provide an additional mode of signaling within the same amount of total hormone secreted, but it could also potentially have differential effects on different tissues within the same organism. This could be picked up both by different patterns of expression of the hormone receptors (e.g., MR and GR) and by differential effects on receptor recycling, postreceptor signal transduction, or interaction with other intracellular pathways.

See Also the Following Articles

Circadian Rhythms, Effects of Prenatal Stress in Rodents; Circadian Rhythms, Genetics of; Seasonal Changes in Stress Responses; Sleep Loss, Jet Lag, and Shift Work; Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress; Cortisol Awakening Response; Night Shiftwork; Seasonal Rhythms; Circadian Rhythm Effects on Cardiovascular and Other Stress-Related Events.

Further Reading

- Belchetz, P. E., Plant, T. M., Nakai, Y., et al. (1978). Hypophysial responses to continuous and intermittent delivery of hypothalamic gonadotropin-releasing hormone. *Science* **202**, 631–633.
- Darmon, M., Brachet, P. and Da Silva, L. H. (1975). Chemotactic signals induce cell differentiation in *Dictostelium discoideum*. *Proceedings of the National Academy of Sciences USA* **72**, 3163–3166.
- Dekelbab, B. H., Witchel, S. F. and De Franco, D. B. (2004). Continuous versus intermittent administration of glucocorticoids: ligand-dependent down-regulation of glucocorticoid receptors revisited. American Endocrine Society meeting, abstract P2–131.
- Folenius, M., Simon, C., Brandenberger, G., et al. (1987). Ultradian plasma corticotropin and cortisol rhythms: time-series analyses. *Journal of Investigative Endocrinology* **10**(3), 261–266.
- Goldbeter, A. (1996). *Biochemical oscillations and cellular rhythms: the molecular basis of periodic and chaotic behaviour*. Cambridge UK: Cambridge University Press.
- Iranmanesh, A., Lizarralde, G. and Veldhuis, J. D. (1993). Co-ordinate activation of the corticotropic axis by insulin-induced hypoglycemia: simultaneous estimates of beta-endorphin, adrenocorticotropin and cortisol secretion and disappearance in normal men. *Acta Endocrinologia (Copenhagen)* **128**(6), 521–528.
- Lightman, S. L., Windle, R. J., Julian, M. D., et al. (1999). Significance of pulsatility in the HPA axis. In: Chadwick, D. J. & Goode, J. A. (eds.) *Symposium on mechanisms and biological significance of pulsatile hormone secretion*, Novartis Foundation Symposium 277, Chichester, UK: John Wiley.
- Merishon, J. L., Sehlhorst, E. S., Rebar, R. W., et al. (1992). Evidence of a corticotropin-releasing hormone pulse generator in the macaque hypothalamus. *Endocrinology* **130**, 2991–2996.
- Métivier, R., Penot, G., Hübner, M. R., et al. (2003). Estrogen receptor- α dissects ordered, cyclical and combinatorial recruitment of cofactors on a natural target promoter. *Cell* **115**, 751–763.
- Nolan, L. A., Windle, R. J., Wood, S. A., et al. (2000). Chronic iodine deprivation attenuates stress-induced and diurnal variation in corticosterone secretion in female Wistar rats. *Journal of Neuroendocrinology* **12**, 1149–1159.
- Papavasiliou, S. S., Zmeili, S., Khoury, S., et al. (1986). Gonadotropin-releasing hormone differentially regulated expression of the genes for luteinizing hormone α and β subunits in male rats. *Proceedings of the National Academy of Sciences USA* **83**, 4026–4029.
- Seale, J. V., Wood, S. A. and Atkinson, H. C. (2004). Gonadectomy reverses the sexually divergent patterns of circadian and stress-induced hypothalamic-pituitary-adrenal axis activity in male and female rats. *Journal of Neuroendocrinology* **16**, 516–524.
- Shanks, N., Windle, R. J., Perks, P. A., et al. (2000). Early-life exposure to endotoxin alters hypothalamic-pituitary-adrenal function and predisposition to inflammation. *Proceedings of the National Academy of Sciences USA* **97**(10), 5645–5650.
- Waxman, D. J., Ram, P. A., Park, S. H., et al. (1995). Intermittent plasma growth hormone triggers tyrosine phosphorylation and nuclear translocation of a liver-expressed, stat 5-related DNA binding protein: proposed role as an intracellular regulator of male-specific liver gene transcription. *Journal of Biological Chemistry* **270**, 13262–13270.
- Wildt, L., Hausler, A., Marshall, G., et al. (1981). Frequency and amplitude of gonadotropin-releasing hormone stimulation and gonadotropin secretion in the rhesus monkey. *Endocrinology* **109**(2), 376–385.
- Windle, R. J., Wood, S. A., Kershaw, Y. M., et al. (2001). Increased corticosterone pulse frequency during adjuvant-induced arthritis and its relationship to alterations in stress responsiveness. *Journal of Neuroendocrinology* **13**, 905–911.
- Windle, R. J., Wood, S. A., Lightman, S. L., et al. (1998). The pulsatile characteristics of hypothalamo-pituitary-adrenal activity in female Lewis and Fischer 344 rats and its relationship to differential stress responses. *Endocrinology* **139**, 4044–4052.
- Windle, R. J., Wood, S. A., Shanks, N. M., et al. (1998). Ultradian rhythm of basal corticosterone in the female rat: dynamic interaction with the response to acute stress. *Endocrinology* **139**, 443–450.

Understimulation/Boredom

V J Sutherland

Sutherland-Bradley Associates, Staffordshire Moorlands, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by V J Sutherland, volume 3, pp 634–636, © 2000, Elsevier Inc.

Understimulation and Boredom as a Rustout Syndrome
Understimulation and the Stress Response
Conclusion

Glossary

<i>Boredom</i>	The state of being tired or made weary by being dull, repetitious, or uninteresting; an emotion in response to an environment that is perceived to be monotonous and that fails to stimulate the individual.
<i>Understimulation</i>	Insufficient arousal in respect to the senses or activity in the individual.

Understimulation and Boredom as a Rustout Syndrome

While a plethora of research evidence exists to suggest that overstimulation in the form of work overload is a key source of stress, associated with a wide variety of negative health and performance variables, the issue and impact of understimulation and boredom should not be ignored or underestimated. Overload as a form of stress can lead to burnout, whereas the lesser used term of rustout might be useful to explain the negative impact of understimulation that results in boredom, apathy, low morale, and absence from the workplace. It is suggested that an unvaried work environment lacks stimulation, and repeated performance of a routine task can be perceived as monotonous. For example, Persson and colleagues observed that understimulation was a potential problem associated with operator work in automated process control, where monitoring was a predominant part of the job. Such conditions can lead to rustout. Similarly, a medical practitioner might experience rustout when the individual feels that his or her skills and abilities are not being utilized to his or her satisfaction. In both instances, the workplace or job produces feelings or emotions that are described as “boring”. When an

employee feels able to produce a report but is not permitted to do so because the boss always expects to take over completion of the final document, qualitative underload is experienced. This can also be viewed as understimulation, and the perceived state of boredom or job dissatisfaction is described as the stressor outcomes or stress response. Lack of a stimulating work environment or challenge may be classified as the stressor or source of stress. These form a stress chain, namely, the stressor leads to perception of a state of stress, which leads to the stress response, or coping. However, it must be acknowledged that stress chains and cause-and-effect relationships are not simple or necessarily linear. Boredom, like pain, can not be seen or measured in an objective way because it is a subjective experience, mediated by individual differences such as personality, needs, wants, and past history. While agreement is not unanimous, it is usually acknowledged that there is an association between boredom and a state of low arousal. Grandjean placed boredom in his categories of fatigue, describing it as mental fatigue wherein the individual reports a feeling of weariness. Research evidence suggests that mentally fatigued workers are likely to behave in an inconsistent manner and/or take short cuts, thereby jeopardizing safety and performance standards. Kass and colleagues demonstrated the business costs associated with boredom at work. Workers in a manufacturing plant scoring high on both state and trait boredom were significantly more dissatisfied with the work itself, pay, promotion, supervisors, and co-workers. Those scoring high in job boredom were also more likely to be job absent, but had longer organizational tenure.

Understimulation and the Stress Response

Early research findings presented by Yerkes and Dodson explained that a certain amount of stimulation (i.e., arousal) was needed for optimal performance but that both understimulation and overstimulation could produce negative effects upon the individual in terms of poor health and ineffectiveness. Thus, the curvilinear relationship observed between level of arousal and performance outcome, known as the Yerkes–Dodson Law or the inverted-U hypothesis, clearly illustrates how too little stimulation can lead to boredom, apathy, and ultimately, depression (i.e., rustout), while overstimulation can result in high anxiety, tension, and burnout. However, response to stimulation will vary among individuals because it is a subjective experience

contingent upon one's perceptions of the experience. A tale of stonecutters in the Middle Ages provides an illustration of the role played by our perceptions in a potentially understimulating situation in the workplace. A traveler approached a group of stonecutters and asked, "What are you doing?" The first responded, "I'm cutting stone. It's dull work but it pays the bills." The second stonecutter said, "I'm the best stonecutter in the land. Look at the smoothness of this stone, how perfect the edges are." The third man pointed to a foundation several yards away and said, "I'm building a cathedral." Thus, a state of stress is said to exist when an individual perceives an imbalance between the perceived level of demand (i.e., stimulus arousal) and his or her perceived ability to meet that demand.

When the task structure or design of the job is perceived as monotonous or repetitive, it can be demotivating and a barrier to performance effectiveness. Typically, such jobs offer little scope for decision making, autonomy to influence working methods, or pace and sequence of task. The passive job described by Karasek is typically a low-demand, low-autonomy job, and as such will generally be dissatisfying as it induces a decline in overall activity and a reduction in general problem-solving activity. Indeed, Zavala and colleagues observed that workers with a low workload were late to work or left early more often than individuals with an average or high workload. However, not all research findings support the notion that industrial monotony must have negative health or poor productivity effects because expectation determines how the situation is experienced. Hence, the sameness of the job is regarded as normal, and while it might not be liked, there seems to be no expressed job dissatisfaction or harmful health effects. The job is simply viewed as an instrumental means to the ends of obtaining what the employee needs outside of the workplace. The notion of understimulation as lack of or low arousal was expanded by the work of French and Caplan, who explained the concept in terms of workload while distinguishing between quantitative and qualitative load.

Quantitative versus Qualitative Underload

It is possible to experience quantitative underload as understimulation because one simply does not have enough to do to fill the time. Cooper and Kelly reported that crane operators confined to their cab and presented with a very low level of work activity experienced quantitative underload, which was a significant predictor of anxiety, depression, and job dissatisfaction among this group of workers. Boredom in the daily routine as a result of too little to do might result in inattentiveness, which is potentially

hazardous, for example, when an employee fails to respond to an emergency situation. Dangers become exacerbated during night-shift work when an individual must adjust to a new sleep pattern but does not have enough to do to stay alert. Cheliout et al. found a high incidence of deactivation episodes among electronics assemblers who were monitored by continuous electroencephalography over a 1-day period. The theta rhythms observed, described as microsleep (i.e., asleep with one's eyes open), were indicative of the boredom and tedium experienced by these workers. When a task does not provide opportunities to use or develop skills and abilities, the situation is described as qualitative underload, and the individual can feel undervalued and bored. This state of apathy can lead to poor morale, irritation, depression, job dissatisfaction, and absenteeism. Thwarted ambition, unfulfilled goals, and unmet expectations cause frustrations that can generate various forms of destructive behavior. Examples of these maladaptive coping strategies include workplace vandalism, sabotage, alcohol and drug abuse, and binge-eating habits.

New Technology and Understimulation/Boredom

Rapid developments in computer technology and its application to industrial and commercial settings have been a major cause of stress in the workplace. A constantly changing environment is a potent stressor because employees must continually strive to keep up with new systems and methods. Technological change can bring stress in the form of social isolation, but one in which individuals are monitored. Jobs are much simplified because the computer-led process has taken over much of the hard physical effort and the decision-making process. Indeed Smith suggested that the main impact of visual display units on workplace activities was underload, de-skilling, and a lack of control. Nevertheless, the use of computers in the workplace has also resulted in responsibility being delegated further down the organizational chain than before, often without commensurate power or pay, thus causing other strains and pressures.

Conclusion

A job is boring if that job incumbent perceives it to be boring. Nevertheless, it is possible to design jobs to avoid such problems. For example, Persson et al. described this process as the design of active versus passive operator jobs, even when monitoring tasks are a predominant part of the job. To avoid problems associated with absenteeism or with employees leaving the organization in order to find a more fulfilling or stimulating job, it is possible to apply a

wide variety of human relations processes such as job enlargement, enrichment, restructuring, and rotation. As Zavala et al. suggested, the benefits are a reduction in boredom, work stress, absenteeism, and turnover and an increase in innovation, productivity, and loyalty. As such, these strategies are recommended as part of an organizational approach to stress management in the work environment.

See Also the Following Articles

Burnout; Workplace Stress.

Further Reading

- Cheliout, F. (1979). Rythme thêta postêrieur au cour de la vielle active chez l'homme. *Rev EEC Neuropsychologie* 9, 52–57.
- Cooper, C. L. and Kelly, M. (1984). Stress among crane operators. *Journal of Occupational Medicine* 26, 575–578.
- Cox, T. (1985). Repetitive work: occupational stress and health. In: Cooper, C. L. & Smiths, M. J. (eds.) *Job stress and blue-collar work* pp. 85–112. Chichester, UK: Wiley.
- French, J. R. P. and Caplan, R. D. (1973). Organizational stress and individual strain. In: Marrow, A. J. (ed.) *The failure of success*, pp. 30–66. New York: Amacon.
- Granjean, E. (1968). Fatigue: its physiological and psychological significance. *Ergonomics* 11, 427–436.
- Greenglass, E. R. (2005). Proactive coping, resources and burnout: implications for occupational stress. In: Antoniou, A. S. G. & Cooper, C. L. (eds.) *Research companion to organizational health psychology* pp. 503–515. Cheltenham, UK: Edward Elgar.
- ILO (1986). Psychosocial factors at work: recognition and control. *Report of the Joint International Labour Office and World Health Organization*, Ninth Session, 1984. Geneva: ILO.
- Langan-Fox, J. (2005). New technology, the global economy and organizational environments: effects on employee stress, health and well-being. In: Antoniou, A. S. G. & Cooper, C. L. (eds.) *Research companion to organizational health psychology*. Cheltenham, UK: Edward Elgar.
- Kass, S. J., Vodanovich, S. J. and Callender, A. (2001). State-trait boredom: relationship to absenteeism, tenure and job satisfaction. *Journal of Business and Psychology* 16(2), 317–327.
- Malinski, R. M. (2002). Job rotation in an academic library: dammed if you do and dammed if you don't. *Library Trends* 50(4), 1–4.
- Persson, A., Wanek, B. and Johansson, A. (2001). Passive versus active operator jobs in automated process control – a job design case study in a control centre. *Applied Ergonomics* 32(5), 441–451.
- Quick, J. C. and Quick, J. D. (1984). *Organizational stress and preventive management*. New York: McGraw-Hill.
- Smith, M. J. (1997). Psychosocial aspects of working with visual display terminals (VDTs) and employee physical and mental health. *Ergonomics* 40(10), 1002–1015.
- Sutherland, V. J. (2005). An organizational approach to stress management. In: Antoniou, A. S. G. & Cooper, C. L. (eds.) *Research companion to organizational health psychology*. Cheltenham, UK: Edward Elgar.
- Sutherland, V. J. and Cooper, C. L. (1991). *Stress and accidents in the offshore oil and gas industry*. Houston, TX: Gulf.
- Sutherland, V. J. and Cooper, C. L. (2000). *Strategic stress management: an organizational approach*. UK: Macmillan Business.
- Sutherland, V. J. and Cooper, C. L. (2003). *De-stressing doctors: a self management guide*. London: Butterworth Heinemann.
- Yerkes, R. M. and Dodson, J. D. (1908). The relation to the strength of the stimulus to the rapidity of habit formation. *Journal of Comparative Neurology and Psychology* 18, 459–482.
- Zavala, S. K., French, M. T., Zarkin, G. A. and Omachonu, V. K. (2002). Decision latitude and workload demand: implications for full and partial absenteeism. *Journal of Public Health Policy* 23, 344–361.

Unemployment, Stress, and Health

M Bartley

University College London Medical School, London, UK

© 2007 Elsevier Inc. All rights reserved.

Introduction

Health Effects of Unemployment

Selection

Mechanisms

Unemployment and Health in the Life Course

Policy Implications

Glossary

<i>Cohort study</i>	A cohort study is one in which a group of individuals is followed over time. A birth cohort consists of a group whose members were all born at around the same time and are therefore the same age. The advantage of this for research on unemployment and health is that it allows us to see whether those who become unemployed differ in any health-relevant manner from those who do not, even before they experience unemployment.
<i>Economically active</i>	The employed and self-employed plus all those looking actively for paid employment.
<i>Economically inactive</i>	Persons who neither have a job nor are currently actively seeking one.
<i>Employment</i>	Work that either is paid by an employer or yields income from self-employment or a small business.
<i>Latent consequences of employment</i>	The idea that employment has benefits other than providing a wage, such as the opportunity to use skills, social integration, and self-esteem.
<i>Life course perspective</i>	A relatively new approach to understanding health differences between people in different social groups that takes account of experiences over time. For example, the approach looks at whether people who experience more unemployment tend to have had different experiences in childhood than those who have experienced little or no unemployment.
<i>Mac-job</i>	Employment that requires little education or training, is very low paid, and is accompanied by no career prospects, low autonomy, and low job security.
<i>Parasuicide</i>	This term does not mean attempted suicide that did not succeed, but refers to serious self-harm that is, however, not thought to be intended to result in death.

Reservation wage
Selection

Unemployment

Unemployment rate

Work

The rate of pay at which an unemployed person will consider taking on a new job. The possibility that pre-existing health may have resulted in unemployment.

Defined by the International Labour Organisation as “neither employed nor self employed and actively seeking employment.” Practices in different nations vary as to the period of time during which an individual must have looked for a job in order to be regarded as unemployed; a common usage is within the last 2 weeks. Calculated as the number of those seeking employment divided by the economically active population. Those not seeking paid work are not included in the denominator.

Not synonymous with employment. It has been calculated that at least half of the work that is done in a society is unpaid. Most unpaid work is carried out by women.

Introduction

From levels in 1977 of around 4.5% in the UK and 6.5% in the US for men over the age of 16 in the mid-1970s, the unemployment rates of these nations rose sharply to around 11–12% in the early 1980s. In the UK, unemployment rose again to another peak, over 12% in men, in the early 1990s, while in the United States it fell more steadily. [Figures 1 and 2](#) show both similarities and differences in the trends in these two nations. Although unemployment is now far lower than its highest level, conditions are very different from those of the 1970s, which preceded the 1980s recession and restructuring of labor markets. The labor market has to some extent polarized. On the one hand, there are more jobs that require high levels of education and skill in new high-technology industries and, even more, in services such as banking, law, and health care. On the other hand, the jobs available to those without further post-school education and training have increasingly become what are known as Mac-jobs (after the McDonalds catering empire). The paradigmatic Mac-job is in the fast food industry and consists of relatively unskilled forms of preparing and serving of meals. Another job typical of the new labor market is the call center. Working in a call center does not require specific qualifications or skills, but is only possible for those capable of clear, grammatical speech, without a heavy regional dialect, and who have sufficient social skills to deal with many different

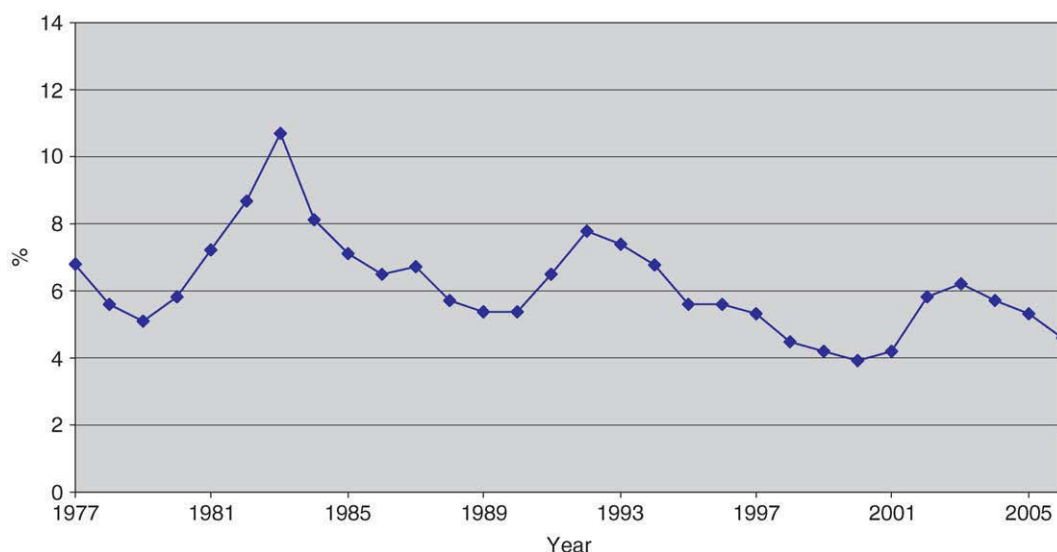


Figure 1 Trends in unemployment of working-age men (16 years and over), United States, 1977–2006 (January). From <http://data.bls.gov>.

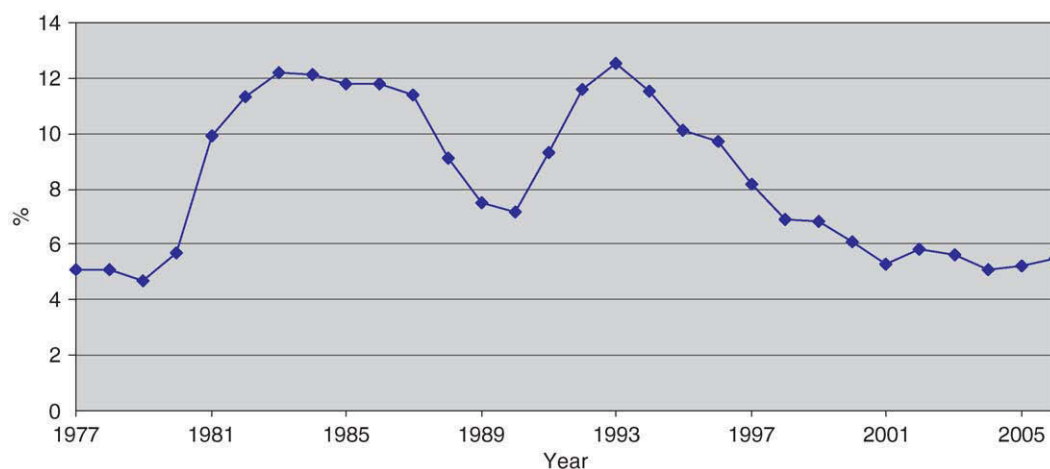


Figure 2 Trends in unemployment of working-age men (16 years and over), UK, 1977–2006 (January). From <http://www.statistics.com>.

(often dissatisfied) customers in a series of intensive remote encounters.

In many mainland European nations, however, large-scale replacement of unemployment by low-paid work has not yet taken place. Unemployment rates in Germany and France, for example, are still very high, at 10–11%. To some extent it could be said that mainland Europe and Scandinavia lag behind the United States and UK, in that the economic crisis of the 1980s in the latter nations occurred in mainland Europe and Scandinavia about 10 years later. It is, therefore, still relevant to consider whether there may be effects of unemployment on health.

Health Effects of Unemployment

Research has repeatedly shown a higher prevalence of ill health and excess mortality in men and women

who are unemployed. Men who are unemployed, and women who are either unemployed or keeping house full time, are more likely than those with paid employment to describe their health as generally fair or poor. Damage to psychological health is also found by the vast majority of studies, an effect that appears to be independent of pre-existing health and to be reversed on re-employment in many cases.

However, it is not a simple step from this observation to a causal relationship between unemployment and health. In aggregate level studies, increases in unemployment in nations or areas are not consistently found to be related to increases in mortality or morbidity. The increase of unemployment in industrialized nations took place during a period when life expectancy was rising steadily. This is not too great a mystery, as even what would be regarded as a very high unemployment rate would be no more than

10–15% of the economically active population of a nation. A typical rate of economic activity in an industrial nation would be 65–75% of those regarded as of working age, which is usually between around 18 and 65 years old (the rest of the population comprising children, the retired, and those looking after home and family). It is also typical of observed patterns of unemployment that the risk of job loss is not randomly spread throughout the population. In general, there will be a group characterized by having worked in a declining industry, having few skills relevant to other industries, and living in an economically declining region, who experience long-term and/or multiple spells of unemployment. This group will carry most of the total burden of unemployment for the society as a whole. This needs to be remembered when we examine the evidence on the health effects of unemployment. It is more or less impossible to find a group of workers who become and remain unemployed for any appreciable period of time who do not also suffer from one or more other burdens of socioeconomic disadvantage.

The reason we need to be cautious about studies showing that people who report that they are unemployed also report worse mental or physical health is that those who are ill may be more likely to lose their jobs and find it harder to regain employment. Or there may be indirect effects whereby men and women with other characteristics such as lower levels of education, which are known to be related to health risks, are also less likely to be employed. This type of relationship, in which certain people are more at risk of both unemployment and ill health is known as selection.

Selection

There are very few studies that have examined relationships of unemployment to health over a long time period. Convincing evidence of a causal effect would be present if long-term studies showed that healthy people who became unemployed subsequently experienced a deterioration in their health. Early studies of unemployed people in the 1970s and 1980s indicated that their physical health did not necessarily decline during a spell of unemployment; a number of studies failed to find an increase in morbidity among those who were continuously unemployed for up to 18 months. However, a more recent study of a representative sample of healthy people over a period of 10 years did observe an increased risk of developing a limiting illness in the year after having become unemployed. Another careful study carried out by occupational health specialists measured a clear deterioration in physical working capacity after a year of unemployment.

However, because of the scarcity of good long-term evidence on the health of people who become unemployed, we need to consider the possibility of what is called a selection effect. A direct health selection effect is thought to be at work when it is poorer health itself that increases the risk of unemployment. Ill health may be a risk for both initial job loss and subsequent chances of re-employment. However, studies in which an entire workplace closes down and all workers are dismissed so that there is no possibility for job loss to be affected by health have also shown increases in illness in those made redundant. Research using linked census-based data from England and Wales and from the Nordic countries has shown that higher risk of mortality is not only present among unemployed people who were already ill. A recent study in Sweden has also found excess risk of mortality among twin siblings who experienced unemployment when compared to their siblings who did not.

Although most studies of increases in national or regional unemployment do not find that these are followed by deteriorations in population health, there is most certainly no evidence that increases in unemployment happen because people's health gets worse. The large increases in unemployment during the 1980s in the Anglo-Saxon nations and during the 1990s in many European nations were not due to worsening health of workers in these countries (indeed, this will seem rather an absurd statement to make). In fact, unemployment rates increased dramatically in different nations during the 1980s and 1990s, at the same time that levels of health in the population actually improved. So there is no possibility that an increase in unemployment overall can be blamed on some kind of general deterioration in the health of the working-age population.

There may still be some tendency for those who remain unemployed for longer periods to be different from those who do not, even if this difference does not take the form of a life-threatening disease. Perhaps people at high risk of unemployment may also have certain personality characteristics, for example, an external locus of control or a weak coping ability. However, this is a more complex hypothesis, and is not necessarily the same as selection. These ideas have to be considered alongside other evidence on the life course accumulation of social and health disadvantage.

One study that has been able to ask the question "How far is the poor health of the unemployed due to their health before they became unemployed?" in relation to mental health is a British birth cohort study of everyone born in the second week of March 1958 and followed until the present time. Researchers

were able to use data on health and psychological development throughout the school years and young adult life, as well as complete employment histories from leaving school to age 33. These data showed that cohort members who experienced more unemployment were also more likely to have grown up in economically disadvantaged and overcrowded households. The study also addressed the possibility that the high level of psychological morbidity found in unemployed men and women might have been due to a pre-existing vulnerability to poor mental health. Even after taking into account pre-existing mental health, recent unemployment was clearly related to the onset of psychiatric symptoms severe enough to require medical care. However, when those with a prior tendency to depression were excluded from the analysis, this had the effect of strengthening the relationship between longer term accumulation of unemployment and the onset of episodes of depression and anxiety. This adds to the evidence that longer term unemployment causes deterioration in mental health in those who were previously healthy.

Mechanisms

What might be the reasons for the relationship between unemployment and health? Several have been put forward, and the following sections consider three of these: poverty, the fact that unemployment is a stressful life event, and changes in health-related behaviors at the time of unemployment.

Poverty

Low living standards are not an inevitable consequence of unemployment; this is a result of the levels at which benefits are set, which results from political decisions. During the 1980s and 1990s, levels of income replacement for the unemployed were lowered in many nations. It was argued that under conditions of increasing automation of unskilled and semi-skilled work, levels of benefits available to the unemployed exceeded the market worth of their labor, that is, the wage rate that employers were willing to pay. If the level of pay at which an unemployed person will accept new employment – the reservation wage – is too high, employers will not take them on. If benefit levels are too high, the state is raising the reservation wage and therefore may be contributing to the problem of high unemployment.

On becoming unemployed, most people will experience a decrease in income. The extent of this decrease varies between nations, from around 20% in some Nordic countries to approximately 60–70% or more in the UK and United States. Many studies link the health effects of unemployment directly to

financial problems. In modern welfare states, state benefits available to those unable to undertake paid employment are thought to avoid starvation and severe physical privation. However, recent research on the financial cost of what medical evidence has shown to be a healthy lifestyle, which includes a diet adequate in essential nutrients, warm housing, and a level of social participation consistent with mental health, costs approximately double the level of social security benefit in the UK. It is therefore inevitable that once savings are exhausted, unemployed people will be forced to borrow money in order to sustain health. Debt, in its turn, has been shown to increase the risk to mental health and the risk of stress-related deterioration in physical health. In several studies, long-term unemployed people who had gone into debt or had to borrow money in the past year had a risk of depression that was more than double that of those who did not have to borrow money. These studies also found that those obliged to borrow were also more likely to report deterioration in physical health. Others have documented the ways in which increasing financial pressures, as savings are used up and worn-out items need to be replaced, are responsible for the growing inactivity and social isolation of many unemployed people.

These British findings are echoed in other countries, such as the United States and the Netherlands, where financial problems were shown to be the main reason why unemployed people's mental and physical health were worse than that of the employed. However, it seems that eventually many of the unemployed adapt to straitened financial and social circumstances. Several studies agree that there appears to be no further deterioration in psychological well-being after a period of between 1 year and 18 months (but no improvement, either, so that unemployed people remain two to three times more likely than the employed to be in poor mental health). This adds weight to the argument for providing early assistance to those who become unemployed. Other evidence suggests that adaptation to unemployment is accompanied by lowered expectations of oneself, and perhaps by a degree of alienation and cynicism. This process has been compared to the adaptive process found in institutionalized inmates of prisons or psychiatric hospitals.

Unemployment as a Stressful Life Event

Research on the ways in which people react to job loss shows it to be a highly stressful life event, which has been characterized as a form of bereavement. International studies provide consistent evidence of the importance of stress. Many researchers have suggested that employment has a number of nonfinancial

benefits to the individual, the so-called latent consequences, and it is the loss of these that results in the threatening character of unemployment. These latent consequences of employment include giving a time structure to the day, self-esteem, and the respect of others; physical and mental activity; use of skills; and interpersonal contact. It may be that the loss of these psychosocial benefits of employment is at least as important as loss of income to the health of unemployed people.

Evidence of the importance of stress for the relationship between unemployment and health comes from studies in Scandinavian nations. In these nations, cash benefits to the unemployed are relatively generous, and therefore the financial effects might not be expected to be as great as in the United States or the UK; however, similar relationships between unemployment and a range of health indicators are seen. A study in Stockholm of men with irregular work histories and a variety of problems that required frequent social service assistance found that those who were employed, even in low-paid casual jobs, were all more active and integrated (and psychologically healthier) than the unemployed. Among unemployed industrial workers in Finland, those who regained paid work experienced a considerable improvement in psychological health regardless of their financial circumstances either before or after re-employment. Italian workers laid off from their jobs experienced raised amounts of both psychological and physical illness despite receiving the whole of their normal wage. These studies provide evidence of the nonfinancial benefits of employment for psychological health. Stress may affect physical health as well as psychological health as a result perhaps of chronically increased levels of anxiety and the relationship this may have to the functioning of the endocrine system.

Health-Related Behavior

There is evidence that unemployment is associated with some forms of health-damaging behavior, although previous research has yielded inconsistent findings on the relationship. Associations between unemployment and health behaviors need to be seen in the context of the longer term development of these behaviors. The little research that exists on this topic seems to indicate not so much that people who become unemployed take up new, health-damaging activities such as smoking and drinking, not least because the costs of these pursuits could be ill-afforded. Rather, it seems that the experience of unemployment makes it harder for individuals to carry through their intentions to improve their health

behaviors. For example, in one study, after taking account of adolescent health risk behaviors, even well-educated and intelligent young men who experienced 12 months of more without a job between the ages of 23 and 33 were less likely than their peers with no unemployment experience to have a healthy lifestyle in their mid-30s.

Self-destructive behavior among unemployed men has been widely investigated. In the UK and in Scandinavian nations, suicide has been found to be more likely in unemployed people. Parasuicide (attempted suicide) has also been found to be higher in unemployed men. Unemployment seems to increase the likelihood of other adverse life events and to reduce the psychological and social resources needed to cope with these. There is, for example, a higher risk of partnership and marriage breakdown among unemployed men. Many of the unemployed also lose contact with their social networks due to the stigma of their situation.

Unemployment and Health in the Life Course

Perhaps the best way to approach the study of the relationship between health and labor market status is from the perspective that unemployment and work insecurity are part of a process through which health disadvantage is accumulated over the life course. Research taking this perspective turns away from any simple opposition between selection and causation and shows that unemployment occurs as part of a much longer term sequence of events in the life of the individual. It is now becoming possible to look at these processes using cohort studies that follow individuals from childhood to middle age. We can see, for example, that those most likely to experience unemployment may also be more vulnerable to excess mortality and morbidity because of earlier experiences of other kinds of hardship and disadvantage. Their experience of unemployment itself therefore contributes to a process of accumulation of health risks. There is little research conducted along these lines. One such analysis showed that while those with disadvantaged childhoods had the highest risk of later unemployment during the recession of the 1980s, young men who had experienced over 12 months of unemployment seemed to be on a less healthy track in terms of both social and behavioral risk, even if they were from more advantaged social origins. It seems as though unemployment also tends to wipe out some of the health advantages built up over the period from childhood to young adulthood in those who had come from socially advantaged homes and done relatively well in education.

Recent longitudinal research has made it possible to understand the relationship between unemployment and health in a more dynamic way. Biological endowment and material and emotional experiences during childhood confer varying amounts of health resources or health potential to the young person. In the real world, people from less secure and advantaged backgrounds are more likely to enter occupations with more adverse working conditions and lower rates of pay. It is of no surprise, therefore, that the aging process itself is known to take place at differential rates according to the employment history of the individual, with more arduous work being associated with faster aging. The physical resources the individual brings to the labor market when he or she must look for a new job will in many cases have already been damaged by these experiences at work, even before any spell of unemployment he or she may suffer. Many of the most hazardous (and least well paid) occupations also tend to be less secure, which means that there will be some degree of coincidence between poor health and the risk of job loss, not just because of the individual's own characteristics but also because of social structure. In contrast to physical resources, psychological and social resources are more likely to be enhanced by the experiences of activity, companionship, and cooperation involved in (even hazardous) employment. Deterioration in psychological health during unemployment is therefore more immediate and visible than deteriorations in physical health, and improvement on re-employment happens quickly.

For young people, employment forms part of the process of establishing an independent identity, so that stable employment may in fact be even more important for younger than for middle-aged or older people. For many, entry into stable employment is now preceded by long periods of job insecurity. The process of identity formation is therefore far more at risk, with accompanying rises in rates of relationship breakdown, poor mental health, addiction, and accidental and self-inflicted harm. Employment is known to aid the development of secure identity and self-esteem and to facilitate the formation of stable relationships. Accordingly, as the British economy shifted from one in which there were larger numbers of stable jobs to one in which employment conditions were more insecure, British birth cohort studies showed that levels of mental health problems in those of young working age increased sharply. Although rates of major diseases of middle age have continued to decline during the period of labor market change, rates of suicide in young men have continued to rise.

Policy Implications

A growing body of international comparative research suggests that societies with higher levels of unemployment and job insecurity may actually produce different sorts of life histories in individual members than those that provide larger numbers of secure jobs. It has been all too easy to take for granted some of the effects of quasi-full employment on the wider society. When jobs were plentiful, in order to recruit and retain workers most large firms undertook extensive training, recruiting school leavers at an early age into apprenticeships. Schools therefore had far fewer restless and unmotivated youth to deal with than would otherwise have been the case, benefiting all children in the school. Internal labor markets (the availability of promotion within a firm based on experience rather than on qualifications) offered people the chance to advance in their jobs even in the absence of school qualifications, from a later starting point. Increasing numbers of firms offered occupational pensions, thus raising the living standards of the retired population at little cost to the state. As part of the McDonaldization of the labor market, the costs of training and of supporting those who are no longer able to work for payment are now increasingly being borne by the individual (the disappearance of apprenticeships and the appearance of student loans, critical illness policies, and private pensions are examples). There has also been an increase in the numbers of persons who become economically inactive, which in many countries far exceeds the increase in the numbers of those officially unemployed and seeking employment.

Increasing the number of low-paid, low-skilled Mac-jobs is often advocated as a solution to the problem of unemployment. However, the evidence on the success of such a policy is mixed, and the impact on population health has not as yet been fully investigated.

During the 1990s, new data from birth cohort studies (studies of large numbers of people all born in the same year) added a considerable amount to our understanding of why unemployment is associated with poor health. We began to see that unemployment takes place as part of a process of accumulation of disadvantage that may begin in childhood, and that a spell of unemployment often occurs as part of a more general pattern of hazardous and insecure work. The old assumptions behind the notion of a social insurance or welfare state safety net are seen to be too simple. The people most likely to fall into the net bear a heavier weight of disadvantage than those who are less likely to need its support. The idea that benefit levels for the unemployed should be set no

higher than mere subsistence, in order not to encourage voluntary unemployment, is therefore becoming discredited. On the contrary, both the taxpayer and the unemployed person (who is of course a past and potential taxpayer him- or herself) benefits more from a generous provision of both material and psychosocial support for those who find themselves without employment.

See Also the Following Articles

Health Behavior and Stress; Immune Surveillance – Cancer, Effects of Stress on; Job Insecurity: The Health Effects of a Psychosocial Work Stressor.

Further Reading

- Allmendinger, J. and Hinz, T. (1998). Occupational careers under different welfare regimes: West Germany, Britain and Sweden. In: Leisering, L. & Walker, R. (eds.) *The dynamics of modern society*, pp. 64–84. Bristol, UK: Policy Press.
- Bethune, A. (1996). Economic activity and mortality of the 1981 Census cohort in the OPCS Longitudinal Study. *Population Trends* 83, 37–42.
- Burchell, B. (1996). Who is affected by unemployment? In: Gallie, D., Marsh, C. & Vogler, C. (eds.) *Unemployment and social change*, pp. 45–58. Oxford, UK: Oxford University Press.
- Catalano, R., Dooley, D., Wilson, G. and Hough, R. (1993). Job loss and alcohol abuse: a test using data from the Epidemiologic Catchment Area project. *Journal of Health Society and Behavior* 34, 215–225.
- Catalano, R., Dooley, D., Novaco, R. W., Wilson, G. and Hough, R. (1993). Using ECA survey data to examine the effect of job layoffs on violent behavior. *Hospital and Community Psychiatry* 44, 874–879.
- Dooley, D. and Catalano, R. (1991). Unemployment as a stressor: findings and implications of a recent study. *WHO Regional Publications European Series* 37, 313–339.
- Dooley, D., Fielding, J. and Levi, L. (1996). Health and unemployment. *Annual Reviews of Public Health* 17, 449–465.
- Kessler, R. C., Turner, J. B. and House, J. S. (1987). Intervening processes in the relationship between unemployment and health. *Psychological Medicine* 17, 949–961.
- Levenstein, S., Smith, M. W. and Kaplan, G. A. (2001). Psychosocial predictors of hypertension in men and women. *Archives of Internal Medicine* 161, 1341–1346.
- Montgomery, S. M., Bartley, M. J., Cook, D. G. and Wadsworth, M. E. (1996). Health and social precursors of unemployment in young men in Great Britain. *Journal of Epidemiology and Community Health* 50, 415–422.
- Montgomery, S. M., Cook, D. G., Bartley, M. J. and Wadsworth, M. E. J. (1998). Unemployment, cigarette smoking, alcohol consumption and body weight in young British men. *European Journal of Public Health* 8, 21–27.
- Montgomery, S. M., Cook, D. G., Bartley, M. J. and Wadsworth, M. E. (1999). Unemployment pre-dates symptoms of depression and anxiety resulting in medical consultation in young men. *International Journal of Epidemiology* 28, 95–100.
- Morris, J. N., Donkin, A. J. M., Wonderling, D., Wilkinson, P. and Dowler, E. A. (2000). A minimum income for healthy living. *Journal of Epidemiology and Community Health* 54, 885–889.
- Moser, K., Goldblatt, P. O., Fox, A. J. and Jones, D. R. (1987). Unemployment and mortality. *British Medical Journal* 294, 509–512.
- Moser, K., Goldblatt, P., Fox, J. and Jones, D. (1990). Unemployment and mortality. In: Goldblatt, P. O. (ed.) *Longitudinal study: mortality and social organisation*, pp. 77–102. London: HMSO.
- Ostry, A. S. (2001). Effects of de-industrialization on unemployment, re-employment, and work conditions in a manufacturing workforce. *BMC Public Health* 1, 15.
- Ritzer, G. (1993). *The McDonaldization of society*. Oakland, CA: Pine Forge Press.
- Wadsworth, M. E., Montgomery, S. M. and Bartley, M. J. (1999). The persisting effect of unemployment on health and social well-being in men early in working life. *Social Science and Medicine* 48, 1491–1499.
- Warr, P. (1987). *Work, unemployment and mental health*. London: Oxford University Press.

Urocortins

É M Fekete and E P Zorrilla

The Scripps Research Institute, La Jolla, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by E P Zorrilla and G F Koob, volume 3, pp 637–642, © 2000, Elsevier Inc.

Structure of Urocortins, CRF, and Related Peptides

Distribution of Urocortins

Pharmacology of Urocortins

Physiological and Behavioral Effects of Urocortins

Glossary

Adrenocorticotrophic hormone (ACTH) A hormone released into systemic circulation by the anterior pituitary as part of the HPA-stress response. ACTH, or corticotropin, induces the release of glucocorticoids from the adrenal cortex.

Corticotropin-releasing factor (CRF) Isolated by Vale and colleagues in 1981, the first hypothalamic ACTH secretagogue to be identified. In addition to its hypothalamic endocrine function, extra-hypothalamic CRF is important in behavioral and autonomic stress responses.

Edinger-Westphal nucleus An oculomotor nucleus of the midbrain that has parasympathetic projections to the ciliary ganglion and descending inputs to the brain stem and spinal cord.

Hypothalamic-pituitary-adrenocortical (HPA) axis The neuroendocrine system which, during stress, amplifies a neural signal of physiological or psychological stress into a systemic, endocrine response. Hormones secreted as part of the HPA stress response cascade equip the organism to mobilize and utilize energy resources more effectively and, more generally, respond to the stressor.

Stresscopin An N-terminally extended analog of urocortin 3 that binds exclusively and with high affinity to the CRF₂ receptor.

Stresscopin-related peptide An N-terminally extended analog of urocortin 2 that binds selectively and with high affinity to the CRF₂ receptor.

Since the isolation of corticotropin-releasing factor (CRF) in 1981, three more endogenous CRF-like peptides have been identified in mammals. The first of these – urocortin 1 (Ucn 1) – was identified in 1995 for its similar primary structure and bioactivity to suckerfish urotensin I and CRF. Ucn 1 exhibited

greater affinity for and activation of type 2 CRF receptors (CRF₂) than CRF and was hypothesized to be a natural CRF₂ receptor ligand. Urocortin 2 (Ucn 2) and urocortin 3 (Ucn 3), 38-amino-acid residue peptides, were then identified in 2001 for their structural relation to both CRF and Ucn 1 and were named urocortins due to their predominant affinity for the CRF₂ receptor. Although the urocortins share sequence identity with one another and with CRF, each is phylogenetically distinct, with the CRF peptide family resulting from multiple gene duplication events. Each peptide has a unique anatomical distribution under the control of different genes. Consequently, despite a structural family resemblance, the natural functions of urocortins and CRF in stress responses *in vivo* may differ significantly.

Structure of Urocortins, CRF, and Related Peptides

The rat Ucn 1 cDNA was cloned from a rat midbrain mRNA library in 1995. The gene codes for a 122-residue prohormone that contains Ucn 1 in the C-terminus. Ucn 1 genes also have been cloned in mouse, human, hamster, sheep (*Ovis aries*), frog (*Xenopus laevis*), and capuchin monkey (*Cebus apella*) and are similar across species. The human Ucn 1 gene lies on chromosome 2 (2p23-p21) and has two exons, with the second exon comprising the coding region, like the CRF gene. The Ucn 1 promoter is regulated by several transcription factor binding sites, including a TATA box, GATA-binding sites, a C/EBP-binding site, a Brn-2 binding site, and a cyclic AMP responsive element (CRE). Upstream of the CRE site (four base pairs), the Ucn 1 promoter contains a consensus half site for glucocorticoid response elements (GREs), consistent with the finding that glucocorticoids upregulate Ucn 1 mRNA synthesis.

The cDNAs for Ucn 2 and Ucn 3 were identified from sequence homology searches of the mouse and human genomes, with the putative Ucn 2 peptide (an N-terminally shortened sequence of stresscopin-related peptide) contained in the C-terminus of a 112-amino-acid residue proprotein, and the putative Ucn 3 peptide (an N-terminally shortened sequence of stresscopin) encoded in a 161-residue precursor. The identity of the endogenous peptides derived from the Ucn 2 and Ucn 3 preproteins remains predicted, as mature peptide sequences have not been definitively isolated and sequenced. Whether a mature Ucn 2 peptide sequence is processed in humans remains uncertain because of the lack of a consensus proteolytic

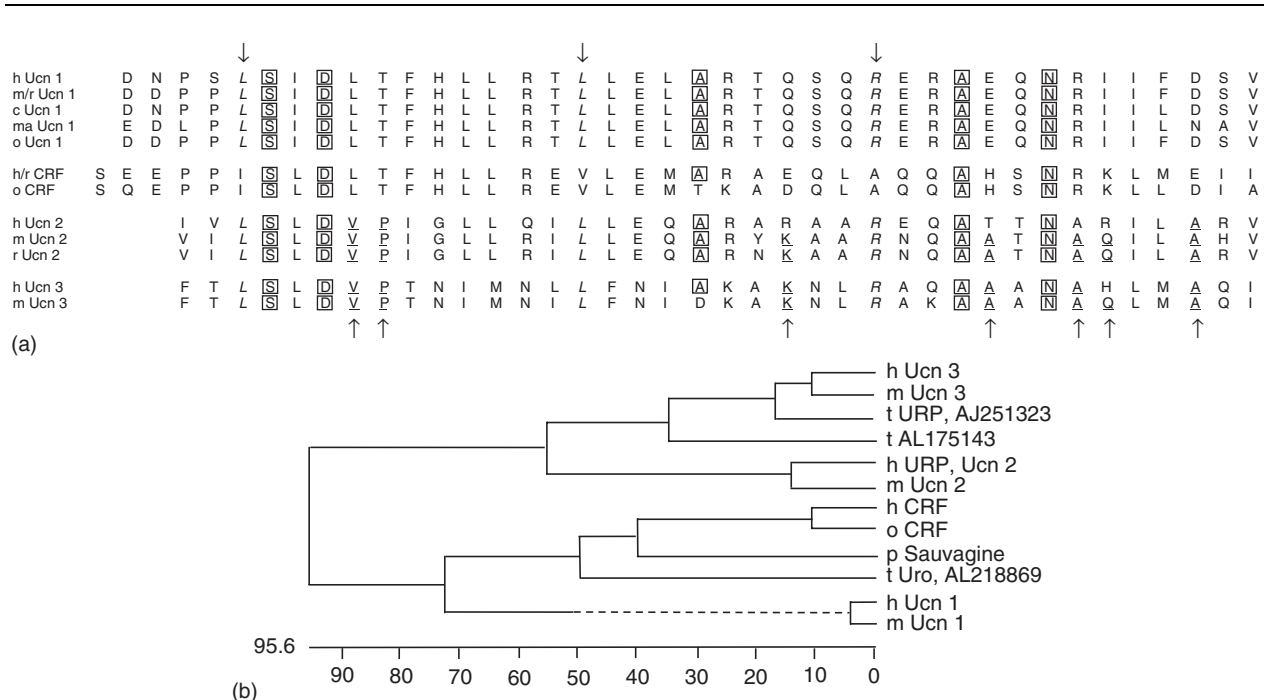


Figure 1 (a) Comparison of the primary structures of mammalian CRF family peptides. Selected putative amino acid sequences for CRF and urocortins 1, 2, and 3 across human (h), murine (*Mus musculus* [m]), rat (*Rattus norvegicus* [r]), monkey (*Cebus paella* [c]), hamster (*Mesocricetus auratus* [ma]), and sheep (*Ovis aries* [o]). Boxed regions indicate CRF superfamily homology; residues in italics and marked by down arrows are shared among all the urocortins; underlined residues marked by up arrows are shared among the type 2 urocortins (Ucn 2 and Ucn 3), but not by Ucn 1. (b) The phylogenetic tree associates human and mouse Ucn 3 and Ucn 2 with pufferfish (*Takifugu rubripes* [t]) urocortin-related peptide (URP). The more distantly related lineage is composed of ovine and human CRF, human and mouse Ucn 1, frog (*Phyllomedusa sauvagei* [p]) sauvagine, and pufferfish urotensin (tUro). The scale beneath the tree reflects sequence distances. Genbank accession sequences are provided for pufferfish-derived peptides. The phylogenetic tree was adapted from Lewis et al. (2001). *Proceedings of the National Academy of Sciences USA* **98**, 7570–7575.

processing site at the C-terminus of the putative sequence. Preproteins for rat Ucn 2 and for chimpanzee and frog Ucn 3 have also been identified. The human genes for Ucn 2 and Ucn 3 lie on chromosomes 3 (3p21.3) and 10 (10p15.1), respectively. The Ucn 2 promoter contains several GREs and is positively regulated by glucocorticoids.

Ucn 1 resembles CRF in primary structure (see [Figure 1](#)) more so than it resembles Ucn 2 and Ucn 3. In contrast, Ucn 2 and Ucn 3 resemble one another more than they do CRF, a distinction, coupled with their greater CRF₂ receptor selectivity, that has led to their characterization as type 2 urocortins. An ancient gene duplication event is believed to have given rise to separate Ucn 1/CRF vs. Ucn 2/Ucn 3 lineages, with subsequent gene duplications events differentiating paralogs within each evolutionary branch (see [Figure 1](#)).

Distribution of Urocortins

Central Nervous System

Ucn 1 has a restricted, subcortical, predominantly caudal distribution in brain. The major site of brain Ucn 1

synthesis is the Edinger-Westphal nucleus (E-W), a dorsal midbrain structure known for its role in oculomotor function. The prominent synthesis of Ucn 1 in the E-W is conserved across rats, sheep, humans, and frogs. Descending Ucn 1-like immunoreactive (LI) positive fibers of possible E-W origin are observed in (1) midbrain: substantia nigra, periaqueductal gray, interpenduncular nucleus, and red nucleus; (2) caudal midbrain/rostral pons: dorsal raphe nucleus, ventral tegmentum, basilar pontine nuclei, and parabrachial nucleus; (3) medulla, including the facial, lateral reticular, and spinal trigeminal nuclei, inferior olive, and the dorsal column nuclei, nucleus of the solitary tract (NTS), and area postrema; (4) cerebellum: in the flocculus and paraflocculus as well as deep cerebellar and vestibular nuclei; and (5) spinal cord: throughout the spinal gray and, less so, in the dorsal and ventral horns. Ascending Ucn 1-LI-positive projections, possibly from the E-W, target the septal/preoptic region and, less heavily, the hypothalamus and thalamus. Validated secondary sites of brain Ucn 1 synthesis include the lateral superior olive, the supraoptic nucleus of the hypothalamus (SON), the lateral hypothalamic area, and several brain stem and spinal cord motoneuron nuclei. Ucn 1-LI is scarce in

many regions in which CRF is prominent, including the external layer of the median eminence, the hypophysiotropic, dorsal medial parvocellular subdivision of the paraventricular nucleus of the hypothalamus (PVN), basal ganglia, amygdala, hippocampus, locus coeruleus (LC), and cerebral cortex.

Ucn 2 also exhibits a restricted, subcortical expression in rodent brain. Like Ucn 1, Ucn 2 mRNA is present in the SON and magnocellular subdivision of the PVN, in brain stem motoneurons, and in the spinal cord. Unlike Ucn 1, Ucn 2 also is expressed in the arcuate nucleus of the hypothalamus and the LC. The projection targets of Ucn 2 neurons are unknown. Nonneuronal Ucn 2 expression also is present in meninges, but not glia.

Ucn 3 exhibits the most rostral, but again subcortical, distribution of the urocortins. The three main sites of forebrain Ucn 3 synthesis are (1) the median preoptic nucleus of the hypothalamus, (2) a perifornical hypothalamic region bordered laterally by the fornix and medially by the PVN, and (3) the dorsal medial amygdala. Less prominent forebrain sites of Ucn 3 synthesis include the dorsomedial hypothalamus, both magnocellular and parvocellular components of the PVN, a region dorsal to the SON, and the posterior cortical and amygdalohippocampal transition areas of the amygdala. Ucn 3-LI-positive fibers of uncertain origin project heavily to the ventromedial nucleus of the hypothalamus (VMH) and arcuate nucleus and less so to the ventral premammillary nucleus. Outside the hypothalamus, forebrain Ucn 3 fibers of uncertain origin are abundant in the lateral septum, posterior bed nucleus of stria terminalis (BNST), and the medial amygdala. The topographical distribution of Ucn 3 in the LS differs from that of Ucn 1, with the latter innervating ventral and intermediate aspects of the structure, and the former innervating dorsal aspects. Caudally, Ucn 3 cell bodies are found in the auditory complex, and scattered Ucn 3 fibers are present in the periaqueductal gray, superior and inferior colliculi, and the ventral lateral lemniscus.

Periphery

Urocortins also are distributed in the periphery. Ucn 1 expression has been observed in adipose tissue, heart, and immune tissue, including thymus, spleen, and skin. Ucn 1 also is present in the enteric nervous system of the duodenum, small intestine, and colon, as well as in testis, kidney, adrenals, pancreas, and, possibly, anterior (but not posterior) pituitary. Ucn-1-LI also has consistently been observed in parietal and oxyntic cells of the stomach. Whether this Ucn 1 is actually synthesized by gastric tissue remains uncertain,

however, as Ucn 1 is synthesized by lamina propria macrophages, components of the stomach's inflammatory mucosal immune system. Finally, Ucn 1-LI is abundant in human placenta and fetal membranes.

In humans, Ucn 2 gene expression was detected in most peripheral tissues analyzed by polymerase chain reaction analysis, with the highest levels observed in heart, lung, muscle, stomach, and adrenal and peripheral blood cells, but not skin. In the mouse, Ucn 2 gene expression was high in skeletal muscle and skin (unlike humans), moderate in lung, stomach, adrenal, ovary, brown fat, spleen, and thymus, and low in uterus, testes, kidney, white fat, and intestine. Unlike in humans, only low Ucn 2 mRNA expression is seen in murine heart. Skin, but not skeletal muscle, expression of Ucn 2 mRNA in the mouse appears to be inversely related to circulating glucocorticoid levels.

Ucn 3 gene expression is present in adipose tissue, heart, muscle, adrenal gland, and skin, albeit at levels considerably lower than those of Ucn 2. Ucn 3 also is present in thyroid, adrenals, β cells of the pancreas, spleen, and the muscularis mucosa of the gastrointestinal (GI) tract, especially in the stomach, small intestine, and colon, but not esophagus or rectum.

Pharmacology of Urocortins

Ucn 1 binds with high affinity to every mammalian CRF binding site identified to date, including the CRF₁ and CRF₂ receptor families and the CRF-binding protein (CRF-BP). This nonselectivity differs from that of the type 2 urocortins, which have selective CRF₂ affinity and less affinity for the CRF-BP. Three CRF receptor subtypes have been identified (CRF₁, CRF₂, CRF₃), two of which have been observed in mammals (CRF₁, CRF₂). CRF receptors are encoded by separate genes, share high sequence homology, and have unique tissue distributions and pharmacological affinity profiles, permitting a diversity of function. CRF receptors belong to the G-protein-coupled, seven transmembrane domain receptor (GPCR) family.

CRF₁ Receptors

The CRF₁ receptor is a class B (secretin-like) GPCR spanning 415 amino acids. Many alternatively spliced transcripts of CRF₁ have been identified in humans, rats, mice, and hamsters. However, only one of these splice variants is known to induce signal transduction in humans and rodents (CRF_{1(a)}), so the following pharmacological characterization pertains mainly to the CRF_{1(a)} isoform. Ucn 1 exhibits reversible, saturable, high-affinity binding to CRF₁ receptors

transfected in stable cell lines. Ucn 1 is two to six times more potent than CRF at binding to the CRF₁ receptor and two- to threefold more potent than CRF in stimulating the production of cAMP from CRF₁ expressing cells *in vitro*. Ucn 1's affinity for the CRF₁ receptor is determined by regions in the first three extracellular domains of the receptor. In contrast to Ucn 1, the type 2 urocortins have very low potencies to activate adenylate cyclase via CRF₁ (EC₅₀ 100 nM), and Ucn 3 shows negligible potency to stimulate ACTH release *in vitro* (1 µM), a CRF₁-mediated bioassay. The CRF₁ receptor can employ multiple signal transduction pathways when stimulated by Ucn 1. These include activation of adenylate cyclase and protein kinase A-dependent pathways, activation of phospholipase C and protein kinase C-dependent pathways, mitogen-activated protein (MAP) kinase-dependent pathways, nitric oxide production, and proximate interactions with calcium channels.

CRF₂ Receptors

CRF₂ receptors also are class B GPCRs. To date, four major CRF₂ splice variants have been identified, including membrane-bound and soluble isoforms of CRF_{2(a)} and membrane-bound CRF_{2(b)} and CRF_{2(c)} receptors. The membrane-bound CRF₂ receptors are, respectively, 411, 431, and 397 residues in length, differing only in their extracellular N-terminus, with the CRF_{2(c)} receptor observed thus far only in human limbic circuits. Unlike CRF, all urocortins bind with high affinity to membrane-bound CRF₂ receptors transfected in stable cell lines. Ucn 1 is slightly more potent than the type 2 urocortins, but is 10–40 times more potent than CRF at binding to membrane CRF₂ receptors. However, the type 2 urocortins are much more (1000-fold) selective CRF₂ agonists than Ucn 1. The CRF₂ receptor also can activate the MAP kinase pathway, with some evidence suggesting that Ucn 3 may be less efficacious than Ucn 2 and, especially, Ucn 1 at engaging this signal transduction mechanism. The soluble CRF_{2(a)} isoform was identified in mice and results from a frameshift deletion of exon 6. In contrast to membrane-bound CRF₂ receptors, the soluble CRF_{2(a)} receptor unexpectedly shows high affinity for Ucn 1 and CRF (classically, CRF₁ ligands) and low affinity for Ucn 2 and Ucn 3. Similar to the CRF-BP (see below), the soluble CRF_{2(a)} receptor may curb CRF₁ signaling by competitively sequestering ligand.

CRF-Binding Protein (CRF-BP)

The CRF-BP is a conserved 37-kDa secreted glycoprotein that binds CRF and Ucn 1 with high affinity, thereby potentially reducing their bioavailability.

Ucn 1 and CRF are approximately equipotent at binding to the CRF-BP. Ucn 1 can displace CRF from the CRF-BP (and vice versa), thereby allowing the possibility that each ligand modulates levels of free CRF receptor agonist in the brain. The region of Ucn 1 that has affinity for the CRF-BP (residues 4–28) differs from that which underlies its affinity for CRF receptors. Ucn 1 is the main natural ligand for the CRF-BP in sheep brain and dissociates from the CRF-BP more slowly than CRF. In contrast to CRF and Ucn 1, mUcn 3 does not bind with high affinity to the human or rat CRF-BP, and mUcn 2 binds with an order lower affinity than CRF or Ucn 1 to the rat CRF-BP and with negligible affinity to the human CRF-BP.

Physiological and Behavioral Effects of Urocortins

Hypothalamic-Pituitary-Adrenocortical (HPA) Axis

Although exogenous administration of Ucn 1 can potentially activate the HPA axis, it is doubtful that endogenous urocortin 1 is a physiological regulator of the HPA axis. Unlike CRF-deficient mice, Ucn 1-deficient mice exhibit normal basal and stress-induced HPA hormone levels. Similarly, unlike CRF antisera, peripheral administration of specific Ucn 1 antisera does not modify basal, stress-induced, or adrenalectomy-induced ACTH levels. Finally, unlike the distribution of CRF, Ucn 1-immunoreactive fibers are scarce in the PVN and the external layer of the median eminence under basal conditions. Unlike both CRF and Ucn 1, intravenous Ucn 2 and Ucn 3 do not increase ACTH or corticosterone secretion, consistent with the absence of CRF₂ on pituitary corticotroph cells. However, type 2 urocortins might locally modulate CRF synthesis at the hypothalamic level, because immobilization stress increases Ucn 2 and Ucn 3 mRNA in the parvocellular PVN.

Osmoregulation

Because Ucn 1 and Ucn 2 are synthesized by magnocellular neurons of the SON and because Ucn 1 fibers are present in the posterior pituitary, an osmoregulatory role for urocortins is hypothesized. Supporting this possibility, salt loading, dehydration, and hypophysectomy increase Ucn-LI in magnocellular SON and PVN neurons, and food deprivation decreases Ucn-LI in the SON. Salt loading also increases SON Ucn 1 mRNA expression. Finally, chronic osmotic stimulation increases CRF₂ mRNA levels in the SON and magnocellular PVN, potentially increasing sensitivity to local urocortin action. Collectively, the

findings support a potential endogenous role for central urocortins in the regulation of salt/water balance. Peripheral urocortins also may control body fluid homeostasis as a result of their ability to stimulate atrial natriuretic peptide (ANP) release from cardiomyocytes, again through a putative CRF₂ mechanism. ANP, present in cardiac atrial tissue, has profound effects on salt and water homeostasis, as it acutely reduces blood volume by sequestering plasma and over the longer term by promoting renal salt and water excretion.

Energy Balance

Exogenous urocortin administration promotes negative energy balance, both by increasing energy expenditure and by decreasing food intake. The distribution of the urocortins supports the hypothesis that they may physiologically regulate energy balance via actions at CRF₂ receptors, exemplified by concordant hypothalamic expression in the VMH (Ucn 3), arcuate nucleus (Ucn 2, Ucn 3), and PVN (Ucn 1, Ucn 2) and in the NTS of the caudal hindbrain (Ucn 1). Furthermore, levels of VMH CRF₂ mRNA covary with circulating leptin levels, providing a homeostatic mechanism for modulating sensitivity to catabolic urocortin action. The presence of urocortins in the GI tract, glucose-regulating organs, and adipose tissue also is consistent with a possible peripheral role for urocortins in energy balance regulation.

With respect to energy expenditure, Ucn 1 (intracerebroventricular [ICV] administration) increases whole body oxygen consumption as measured by indirect calorimetry and increases sympathetic nervous system activity. The hypothalamus is a candidate site for metabolic action for urocortins, as intra-PVN Ucn 1 increases plasma leptin levels, induces intrascapular brown adipose tissue (BAT) uncoupling protein-1 (UCP1) mRNA synthesis, and increases the relative utilization of fat as an energy substrate. Whether energy expenditure-increasing effects of Ucn 1 are mediated in part by CRF₂ receptors and shared by the type 2 urocortins is unclear.

With respect to energy intake, peripheral Ucn 1 or stresscopin administration potently reduces food intake in rodents, possibly in part by slowing gastric emptying (see below). Central infusion of urocortins also potently suppresses feeding and reduces gastric emptying, effects that were shown to be at least partly CRF₂-mediated from studies that used selective agonists (type 2 urocortins), antagonists, antisense knockdown of receptor expression, and knockout mice. Unlike CRF₁ agonists, type 2 urocortins do not produce malaise, arousal, or anxiety-like effects

at central anorectic doses in the rat. The degree to which endogenous brain urocortins produce these effects under physiological conditions remains unclear. In addition to the CRF₂-mediated urocortin anorexia, brain Ucn 1 may also reduce food intake via a separate CRF₁-dependent mechanism. Possible loci for Ucn 1-CRF₁-mediated anorexia include the dorsomedial nucleus of the hypothalamus, the parabrachial nucleus, and caudal hindbrain glucoregulatory sites.

Ucn 3 is present in pancreatic islet β cells, with pancreatic Ucn 3 levels responsive to changes in local glucose levels. Exogenous Ucn 3 administration increases glucagon and insulin levels *in vivo* and *in vitro*, resulting in a net increase in blood glucose levels. Glucoregulatory effects of Ucn 3 were abolished by pretreatment with a selective CRF₂ antagonist. Altogether, findings suggest an autocrine or paracrine regulation of glucose homeostasis by pancreatic Ucn 3.

Gastrocolonic Motility and Function

Stress also releases CRF-related peptides, possibly including urocortins, which slow gastric emptying through CRF₂ receptor systems. Through vagal efferents, central infusion of CRF receptor agonists reduce antral gastric motility, inhibit high-amplitude gastric contractions, and shift duodenal activity from fasted to fed motor patterns. Parallel to the central CRF₂ pathway for stress-induced gastric stasis, peripheral CRF₂ receptor activation delays gastric emptying, with peripheral (intravenous [IV] or intraperitoneal [IP]) administration of agonists with high (i.e., Ucn 1) or selective (i.e., Ucn 2 or Ucn 3) CRF₂ affinity delaying gastric emptying more potently than CRF. Like central infusion, peripheral Ucn 1 administration reduces antral gastric motility in fed rats and shifts gastric motor patterns from a fasted to fed state. Ucn 1 also hyperpolarizes stomach smooth muscle. Candidate regions that subserve CRF₂ receptor-mediated gastric stasis include, centrally, the PVN and dorsal vagal complex and, peripherally, myenteric fibers of the enteric nervous system. A physiological role for urocortins in stress-induced slowing of gastric emptying is supported by the presence of urocortins in the PVN, NTS of the caudal hindbrain, GI tract, and enteric nervous system.

Gastric Ucn 1 also may regulate gastric acid secretion, as peripheral administration of CRF₂ receptor agonists inhibits gastric acid secretion and increases gastric mucosal blood flow. Anatomical evidence supports such a relation. Ucn 1 colocalizes with tyrosine hydroxylase in parietal cells of the stomach, in proximity to CRF₂, but not CRF₁ receptors. CRF₂

receptors, in turn, colocalize with H⁺/K⁺-ATPase, the enzyme gastric proton pump, and somatostatin, which inhibits parietal cell activity and secretion of gastrin and histamine.

In contrast to their effects on gastric motor function, stressors stimulate colonic motility. Indicating a role for urocortin/CRF systems in these effects, central and peripheral administration of CRF₁, but not CRF₂, agonists likewise stimulates colonic motility, and selective CRF₁, but not CRF₂, antagonists attenuate stress-induced stimulation of colonic motility. Candidate substrates for CRF₁-mediated stimulation of colonic motor function include, centrally, the PVN and LC/Barrington nuclei that activate sacral parasympathetic nervous system activity and, peripherally, the colonic myenteric nervous system. Ucn 1 is expressed in these tissues, and, therefore, endogenous Ucn 1 may partly mediate the colonic-stimulating effects of stress.

Cardiovascular Function

Intravenous administration of urocortins more potently reduces blood pressure and has greater vasodilating effects than CRF. The relaxant effects of urocortin are mediated by arteriole and possibly cardiac CRF₂ receptors and appear to involve activation of protein kinase A and MAP kinase signaling pathways as well as inhibition of a PKC-dependent contractile mechanism. CRF₂-deficient mice have elevated mean arterial pressure, suggesting a physiological function of urocortin/CRF₂ interactions on vasculature *in vivo*. Independent of its effects on peripheral vascular resistance, systemic Ucn 1 infusion also had prolonged cardiac inotropic actions in sheep, increasing cardiac contractility, heart rate, and aortic blood flow.

Urocortins also are increasingly recognized to have cardioprotective and cardiovascular function-enhancing effects on compromised heart. Urocortins protect cardiac tissue from ischemia-reperfusion injury or hypoxia-reoxygenation injury via CRF₂ mechanisms. The cardioprotective effects of urocortins appear to involve activation of ERK1/2-p42, 44 signaling pathways, the phosphatidylinositol-3 (PI-3) kinase/Akt pathway, and protein kinase C epsilon isozyme activation. The cardioprotective effects also appear to involve reduced mitochondrial damage, achieved by increasing mitochondrial K_{ATP} channel opening. In addition to its cardioprotective effects, exogenous Ucn 1 also has cardiac hypertrophic effects, involving an Akt-dependent pathway.

Supporting a physiological cardioprotective role for urocortins, Ucn 1, 2, and 3 are present in the heart, with Ucn 2 and Ucn 3 abundant in myocardium.

Plasma Ucn 1-IR increases in human systolic heart failure, and Ucn 1 expression increases more than ninefold in viable human myocytes after surgical cardioplegic arrest-reperfusion. Finally, cardiac myocytes isolated from CRF₂-deficient mice are more susceptible to ischemia-reperfusion injury.

Immune Function

As described previously, urocortins are expressed in immune tissue, including thymus, spleen, and/or skin, with Ucn 1 seen at the cellular level in lymphocytes, macrophages, fibroblasts, and mast cells. Exogenous Ucn 1 administration has palliative effects on experimental autoimmune encephalomyelitis and thermal injury-induced edema (although this may reflect glucocorticoid, vasodilatory, or direct cytoprotective mechanisms, as opposed to immunomodulating mechanisms). However, converging lines of evidence suggest that immune-derived urocortins locally modulate proinflammatory responses to perceived environmental insults.

For example, Ucn 1 mRNA is expressed in lamina propria macrophages, and Ucn 1-IR is detected throughout the entire lamina propria layer of the intestine. Intestinal Ucn 1 activity may be regulated by changes in the intestinal milieu, such as passage of dietary factors or food-associated bacterial antigens. Ucn 1-IR is elevated in the intestinal lamina propria macrophages of patients with ulcerative colitis, in which it is hypothesized to have a proinflammatory CRF₁-mediated effect. Intraperitoneal administration of CRF₁ agonists, such as Ucn 1, also increases intestinal mucosal permeability to macromolecules.

Other inflammatory conditions also are accompanied by increased Ucn 1 expression. In rheumatoid arthritis (but not osteoarthritis), Ucn 1-IR and Ucn 1 mRNA are elevated in synovium, and the number of Ucn 1-positive cells in synovia, including leukocytes and macrophages, is not only higher, but also correlates with the degree of inflammation. Supporting a proinflammatory property of secreted Ucn 1, Ucn 1 stimulates proinflammatory cytokine production by peripheral blood mononuclear cells *in vitro*, presumably, like CRF, via a CRF₁-mediated mechanism.

Some stress-related dermatological inflammatory conditions also have altered urocortin/CRF signaling. Mast cells, which play a role in allergy and inflammation, are widely distributed in the skin and synthesize and secrete both CRF and Ucn 1 in response to psychosocial stress or immunoglobulin E receptor cross-linking. Mast cells also express CRF receptors, activation of which leads to the release of proinflammatory mediators. Thus, atopic dermatitis, psoriasis, and alopecia areata have been

hypothesized to involve a stress-related precipitation or exacerbation of mast cell activation via local CRF/Ucn 1–CRF receptor signaling.

The reproductive tract also contains mast cells. Ucn 1-IR is increased more than 10-fold in spontaneous abortion products from women with a history of multiple nonelective abortions relative to products from elective or nonhabitual abortions. Supporting the hypothesis that the observed increase in Ucn 1-IR is related to mast cell activation, levels of tryptase, which constitutes one-fifth of all protein in mast cells, and IL-8, an abortogenic mast cell-derived cytokine, also are robustly elevated in habitual spontaneous abortion products. Endometriosis also is associated with an increased number of activated mast cells in association with strong Ucn 1 immunostaining. Normal endometrium, in contrast, shows low tryptase and Ucn 1 immunoreactivity. Thus, Ucn 1 expression and mast cell activation are both elevated in inflammatory conditions of the reproductive tract and may correlate with increased risk for spontaneous abortion.

Finally, Ucn 1-IR is increased in stomach biopsies from patients with active gastritis. Unlike the colon, the stomach is much richer in CRF₂ than in CRF₁ receptors. As a result, in contrast to the CRF₁-mediated proinflammatory effects of intestinal Ucn 1, stomach Ucn 1 has been hypothesized, via CRF₂ receptors, to reduce/repair injury to gastric mucosa following noxious stimuli. Consistent with this perspective, Ucn 1-IR increases in treatment-responsive, but not treatment-resistant, patients during recovery from gastritis.

Reproductive Function

In addition to mast cells within the reproductive system, the human reproductive system itself expresses Ucn 1, CRF receptors, and CRF-BP. Ucn 1-IR is produced by choriodecidual and placental tissue and is elevated in maternal plasma from midgestation through birth. Placental Ucn 1 appears to act as a vasoactive relaxant factor on uteroplacental vasculature via local action at CRF receptors. Supporting the physiological relevance of this mechanism, pregnant women who have impaired uterine artery blood flow have significantly reduced circulating Ucn 1 levels in proportion to the degree of increased arterial resistance. Urocortins also may stimulate uterine contractility, as myometrium expresses both CRF₁ and CRF₂ receptors. Whereas CRF inhibits contractility via CRF₁ receptors, urocortins increase myometrial contractility via local CRF₂-mediated actions. Thus, urocortins may regulate placental vessel resistance to blood flow and stimulate uterine contractility.

Anxiety-Related Behavior

Central administration of Ucn 1, like CRF, has anxiogenic-like properties in many rodent anxiety models that are mediated at least partly by CRF₁ receptors. Ucn 1's effects on acoustic startle responding appear to differ from those of CRF, however, as exogenous Ucn 1 dampens rather than potentiates the acoustic startle response. Contrary to the effects of CRF₁ receptor agonists, ICV type 2 urocortin administration does not consistently increase anxiogenic-like behavior. For example, Ucn 3 did not increase anxiety-like behavior in the open field, elevated plus maze, light/dark box, social interaction, or defensive burying tests. Ucn 2 (ICV) similarly lacked acute anxiogenic-like effects in the rat open field and elevated plus maze tests. In fact, ICV Ucn 3 acutely produced anxiolytic-like changes in the elevated plus maze and light/dark box tests, and Ucn 2 had delayed, anxiolytic-like effects in the plus maze and reversed the anxiogenic-like effects of CRF in the open field. Similarly, whereas CRF₁ agonists increased activity in familiar environments, type 2 urocortins had mild motor suppressing effects.

However, stress-like effects of CRF₂ receptor activation also have been reported following local microinfusion into discrete brain regions, including the lateral septum and dorsal raphe, findings somewhat incongruous with the reviewed effects of ICV type 2 urocortin administration. Furthermore, unlike in rats, exogenous ICV administration of Ucn 2 to mice increased anxiety-like behavior in the plus maze as well as acoustic startle responses. Thus, although it is accepted that CRF₁ receptor activation has anxiogenic-like effects, general conclusions about the anxiety-related effects of central administration of the CRF₂-selective type 2 urocortins are not yet entirely possible and may involve brain region or species specificity. The endogenous roles of urocortins similarly remain unclear, as one Ucn 1-deficient mouse model exhibited normal anxiety-like behavior, whereas another showed increased anxiety-like responses.

Hearing

Ucn 1 may be required for normal hearing, as Ucn 1-deficient mutant mice have shorter hair cells in the inner spiral bundle of the organ of Corti and a higher response threshold in the auditory brain stem response examination than wild-type littermates. Perhaps reflecting this difference, a second murine model of Ucn 1 deficiency exhibited a reduced acoustic startle response. Anatomically supporting a relation of Ucn 1 to audition, Ucn 1-IR neurons are present in the lateral superior olive as well as in the inner spiral

bundle of the organ of Corti. In the organ of Corti, Ucn 1 appears at 8 days of age, present only in the inner hair cell region, in proximity to CRF₁ and CRF₂ receptors.

Acknowledgments

This work was supported by NIH grants DK70118, DK64871 and DK26741 from the National Institute of Diabetes and Digestive and Kidney Diseases. This is publication number 17400-NP from the Scripps Research Institute.

See Also the Following Articles

Anxiety; Blood Pressure; Cardiovascular System and Stress; Central Stress Neurocircuits; Corticotropin Releasing Factor (CRF); Corticotropin Releasing Factor-Binding Protein; Corticotropin-Releasing Factor Receptors; Cytokines; Gastrointestinal Effects; Immune Response; Immunity; Macrophages; Peptides; Psychosomatic Medicine; Somatic Disorders; Stress Effects, Overview; Ulceration, Gastric.

Further Reading

- Bale, T. L. and Vale, W. W. (2004). CRF and CRF receptors: role in stress responsivity and other behaviors. *Annual Review of Pharmacology and Toxicology* **44**, 525–557.
- Hashimoto, K., Nishiyama, M., Tanaka, Y., et al. (2004). Urocortins and corticotropin releasing factor type 2 receptors in the hypothalamus and the cardiovascular system. *Peptides* **25**, 1711–1721.
- Scarabelli, T. and Knight, R. (2004). Urocortins: take them to heart. *Current Medicinal Chemistry: Cardiovascular and Hematological Agents* **2**, 335–342.
- Suda, T., Kageyama, K., Sakihara, S., et al. (2004). Physiological roles of urocortins, human homologues of fish urotensin I, and their receptors. *Peptides* **25**, 1689–1701.
- Theoharides, T. C., Donelan, J. M., Papadopoulou, N., et al. (2004). Mast cells as targets of corticotropin-releasing factor and related peptides. *Trends in Pharmacological Sciences* **25**, 563–568.
- Zorrilla, E. P., Taché, Y. and Koob, G. F. (2003). Nibbling at CRF receptor control of feeding and gastrocolonic motility. *Trends in Pharmacological Sciences* **24**, 421–427.

V

Vaccination

V E Burns, A C Phillips and K M Edwards

University of Birmingham, Birmingham, UK

© 2007 Elsevier Inc. All rights reserved.

Introduction

The Vaccination Model

Stress and Antibody Response to Vaccination

Influence of Other Psychosocial Factors on Antibody Response to Vaccination

Possible Mechanisms

The Possible Paradoxical Effect of Acute Stress

Conclusion and Future Directions

Glossary

<i>Antigen</i>	Any molecule that generates and binds to an antibody.
<i>B lymphocytes</i>	A major type of lymphocyte; when activated by antigen, B lymphocytes differentiate into cells that produce antibodies against the antigen.
<i>Immunoglobulin</i>	Immunoglobulins are secreted by B lymphocytes and help fight infection through various mechanisms; also known as antibody.
<i>In vitro techniques</i>	Experimental techniques in which the experiment is performed outside a living organism; the conditions, and therefore the results, may not represent <i>in vivo</i> function.
<i>In vivo techniques</i>	Experimental techniques in which the research is conducted within the whole organism.
<i>Neuroticism</i>	A traitlike, or stable, personality dimension consisting of emotional instability and maladjustment.
<i>Thymus-dependent antibody responses</i>	Antibody responses, usually against a protein antigen, that require that the B lymphocytes gets help from T lymphocytes.
<i>Thymus-independent antibody responses</i>	Antibody responses, against a polysaccharide or lipopolysaccharide antigen, in which the B lymphocytes do not require help from T lymphocytes.

T lymphocytes A major type of lymphocyte; one of their many functions is helping B lymphocytes produce an antibody response.

Vaccination The injection of a dead or attenuated pathogen (disease-causing agent) in order to induce an adaptive immune response and protect against disease.

Introduction

Historically, the immune system has been considered to function in a relatively autonomous manner. Many years of research, however, have now revealed that the immune system is inextricably linked with the complex neuroendocrine networks of the body. These interactions provide plausible biological pathways through which psychological factors can influence immune function.

There is now considerable evidence that periods of chronic stress are associated with reductions in the number of a variety of immune cells in peripheral blood and with the impaired function of these cells *in vitro*. However, the clinical implications of these changes in healthy people, and within normal ranges, are unclear. Changes in cell number will not necessarily result in a significant change in the capacity of the lymphoid system to make an effective response to antigenic challenge; they may just reflect changes in the dynamics of lymphocyte migration and recirculation and other factors, such as shifts in plasma volume, rather than absolute changes in total cell numbers. Similarly, it is difficult to generalize *in vitro* measures of immune function, such as lymphocyte proliferation to mitogen and natural killer cell cytotoxicity, to *in vivo* processes. The isolated testing of any particular aspect of the network of immune cells provides limited information about the status of the highly integrated, complex immune system and, as such, has limited application to overall understanding of the relationship between stress and susceptibility to disease.

In contrast, the vaccination model affords an ecologically valid means of examining the associations

between psychosocial variables and the *in vivo* response of the immune system to antigenic challenge. Significantly, this response occurs within the functional context of the neuroendocrine system. A higher antibody response to vaccination is considered to be indicative, at a general level, of a well-functioning immune system and, at a more specific level, is likely to be associated with greater protection levels against the particular disease in question. As such, the adequacy of the immune response to vaccination has direct clinical relevance, both in terms of maximizing protection following vaccination and as a model of the response to naturalistic infection.

The Vaccination Model

Basic Immunology of Vaccination

The immune response to vaccination involves the coordination of a wide variety of immune cells. The vaccine antigen is initially recognized and presented by professional antigen-presenting cells, such as dendritic cells. The antigen is then recognized by specific T lymphocytes, which proliferate and differentiate into T-helper-2-cell lymphocytes. Vaccine antigen is also recognized by B lymphocytes without the necessity for antigen processing. When stimulated by antigen, B lymphocytes proliferate and mature into short-lived plasma cells, which produce the earliest immunoglobulin, IgM. In a primary response to a previously unseen antigen, peak IgM response occurs around 5 days after vaccination. The interaction between the activated T and B lymphocytes leads to the formation of germinal centers and, therefore, the production of high-affinity IgG and IgA antibody. This response is slower than the IgM response and reaches a peak around 28 days after vaccination. Secondary antibody responses, in which the immune system has been previously exposed to the antigen, are more rapid and of greater magnitude.

Types of Vaccination

The vaccination model also allows the examination of which aspects of the immune system may be vulnerable to psychosocial influence. The antibody response to most vaccines is thymus-dependent, in which T-lymphocyte help is required for the production of antibody by B lymphocytes. In contrast, for thymus-independent antibody responses against polysaccharide antigens, B lymphocytes generate an antibody response without T-cell help. If, for example, the effects of stress are restricted to antibody responses to thymus-dependent vaccines, this would

imply that T lymphocytes are more susceptible to psychosocial influence than B lymphocytes.

Stress and Antibody Response to Vaccination

Cross-Sectional Studies

Early studies took advantage of the large number of people undergoing routine vaccination, such as meningitis C vaccination for all young adults and hepatitis B for medical students. The largest of these studies examined the association between life events stress and the final antibody titer against hepatitis B vaccination in two cohorts of students, vaccinated either in the past 12 months or at least 13 months previously. Whereas life events exposure was not related to antibody response in the recently vaccinated cohort, participants in the earlier vaccinated cohort who reported higher life events stress over the past year were over twice as likely to show an inadequate antibody titer as those in this cohort with lower life events exposure. This association withstood adjustment for variations in unhealthy behaviors (smoking, alcohol consumption, exercise, and sleep) and coping style. The results suggest that the immunogenicity of the hepatitis B vaccination may initially override any influence of psychological factors. Further, the larger number of individuals with inadequate antibody titers in the early vaccination cohort provides more power to detect effects. Nevertheless, this study implies that psychological stress may have its principal effects on the rate of deterioration of protection.

A similar study investigated the role of psychosocial variables in the response to meningitis C vaccination. This is a conjugate vaccination, in which a polysaccharide antigen is conjugated to a protein in order to generate a thymus-dependent antibody response. This is a common technique to improve the antibody response to thymus-independent antigens. Participants with high perceived stress and psychological distress levels were more likely to have low antibody levels against meningococcal C polysaccharide following the vaccination during the previous 16 months. This finding suggests that antibody responses to conjugate vaccination are also susceptible to psychosocial influences.

Prospective Studies in Elderly Populations

In the prospective studies of psychosocial influence on vaccination response, the experimenters administered the vaccines at controlled times, in order to be able to assess pre- as well as postantibody levels. This

is particularly important with common diseases, such as influenza, in which naturalistic exposure is common. In elderly populations, the caregiver-control model has been used to investigate the influence of the chronic stress of caring for a spouse with dementia on the antibody response to influenza vaccination. Two separate studies have demonstrated that spousal caregivers are less likely to achieve a fourfold increase in antibody titer at 1 month postvaccination than matched control participants. Because this is the clinical criterion of an adequate vaccination response, these effects are likely to result in increased susceptibility to infection.

One potential shortcoming of the caregiver-control model is that caregivers may differ from controls in a number of respects other than psychological stress, such as experiencing more physical strain. Significantly, it has been shown that former spousal caregivers display the same impaired antibody response to influenza vaccination as current caregivers. Although much of the physical demand of caregiving terminates with bereavement, it is reasonable to presume that the psychological stress continues. This argues that the poorer responses to vaccination observed in caregivers is the result of psychological stress rather than physical strain. Further evidence comes from a recent study in which the efficacy of the influenza vaccination was tested in a large group of elderly people living in the community. Individuals who had experienced bereavement in the previous year had significantly poorer antibody responses against one of the three influenza strains contained in the vaccination. This suggests that a variety of chronic stressors can have a detrimental effect on antibody response to the thymus-dependent influenza vaccination.

The one thymus-independent vaccination that has been examined in elderly adults is the vaccine against pneumococcus infection. Spousal caregivers were compared to former caregivers and age-matched controls. Although there were no group differences either 2 weeks or 1 month after the vaccination, the current caregivers had poorer specific antibody titers 3 and 6 months after receiving the pneumococcal vaccination than the other two groups. This suggests that caregiving impacts the rate of deterioration of protection rather than the initial response to thymus-independent vaccinations.

Prospective Studies in Young Populations

In younger populations, a common vaccine that has been used is the hepatitis B vaccination program. This consists of three separate vaccinations, administered at 0, 1, and 6 months. In one of the first studies in this area, individuals reporting higher mean perceived

stress and anxiety over the whole vaccination period, probably indicating some trait or dispositional tendency, were less likely to have developed detectable antibodies against the vaccine by the time of the second inoculation. Although it is initially unclear why stress and anxiety levels assessed some months later may be related to this seroconversion, it is likely that these questionnaires measure, to some extent, trait or dispositional tendencies to report high stress and anxiety. It seems feasible, therefore, that people who generally tend to perceive their lives as stressful have higher mean stress scores and that it is these individuals who did not seroconvert at this time. It has also been demonstrated that the immunogenicity of the vaccine is important; a pair of studies by the same investigators found a negative effect of stress on antibody response to a low-dose hepatitis B vaccine, but the full-dose vaccination did not yield any significant stress effects.

Stress and the antibody response to the influenza vaccination has also been studied in younger adults. In one study, students who did not achieve a fourfold increase in antibody titer to the A-strains of the vaccine at 5 months postvaccination reported higher stressful life event exposures, and there was also a trend for those who were unprotected at 5 weeks to report more stressful life events in the period following vaccination. In a study of the effects of daily stress and feelings of being overwhelmed during the 10 days following vaccination, higher stress ratings were associated with lower antibody titers to one of the viral strains at both 1 and 4 months following vaccination. These studies provide evidence that stressful life events both preceding and in the period immediately following vaccination can influence the antibody response.

There has been only one study in which thymus-dependent and thymus-independent vaccines have been directly compared. This study examined the associations between stressful life events and the antibody response to the thymus-dependent influenza and thymus-independent meningococcal A + C vaccinations in a sample of 57 students. Psychosocial factors were associated with the antibody response to the influenza vaccination and were not associated with the response to thymus-independent meningococcal A polysaccharide vaccine component. This provides preliminary evidence that thymus-dependent vaccinations are more susceptible to the impact of stress than thymus-independent vaccinations.

Influence of Other Psychosocial Factors on Antibody Response to Vaccination

There is increasing interest in the influence of more stable psychological characteristics on antibody

response to vaccination. An early study in young girls found that higher internalizing, a concept similar to neuroticism, and lower self-esteem were associated with lower antibody titers following rubella vaccination. This was apparent, however, only in those who did not have serological evidence of exposure to rubella prior to the vaccination and were thus exhibiting a primary vaccine response. Trait negative affect has also been negatively associated with the secondary antibody response to the second hepatitis B injection. Similarly, neuroticism has been associated with a poorer antibody response to the influenza vaccine at both 5 weeks and 5 months postvaccination.

Social support has also been demonstrated to be associated with the response to vaccination. Two separate studies have demonstrated that students who reported greater social support have better antibody responses to either hepatitis B or influenza vaccination. Further, in a study of college freshmen, loneliness and smaller social network size were associated with a poorer antibody response to the influenza vaccination. In an elderly population, although generic social support was not associated with the influenza vaccination response, being married, and particularly reporting being in a happy marriage, was beneficial. With aging, marriage may become the primary social resource that influences health, including susceptibility to infection.

Possible Mechanisms

An association between stress and antibody response to vaccination could arise from direct and/or indirect mechanisms. The most likely indirect mechanisms are changes in health behaviors, often associated with periods of stress and their subsequent physiological impact on antibody formation. Poorer antibody responses to vaccination have been observed in those who consume high levels of alcohol, in regular smokers, and in those with inadequate nutrition. In spite of these findings, some studies have failed to adequately control for health behaviors, leaving open the possibility that at least some of the observed association between stress and antibody status may be attributable to variations in unhealthy behaviors.

More direct mechanisms of interaction between stress and immune function have also been postulated. Stress is associated with alterations in sympathetic nervous system and hypothalamic-pituitary-adrenal axis activation. Functional relationships between these neuroendocrine pathways and the immune system have been acknowledged for some time, including the innervation of lymphoid tissue and

neuroendocrine receptors on many types of immune cells. Therefore, it is feasible that changes in the basal activity in these systems or their repeated activation in response to stress could impact on antibody development and maintenance.

The Possible Paradoxical Effect of Acute Stress

In contrast to the immunosuppressive effects of chronic stress discussed so far, it has also been suggested that acute stress may be immunoenhancing when experienced close to the immune challenge. Dhabhar and McEwen argued that immunoenhancement by acute stress is adaptive from an evolutionary standpoint and might be regarded as an integral component of the fight-or-flight response. From this perspective, circumstances that initiate a fight-or-flight response are likely to also involve exposure to antigens and, therefore, a robust immune response would be adaptive for survival. There is now convincing evidence from animal studies of acute stressors in close temporal proximity to vaccination that provides support for an immunoenhancing effect of stress on the antibody response. Only one study has examined the effect of acute psychological stress on antibody response to vaccination in humans. Participants completed either a 45-min time-pressured, socially evaluated mental arithmetic task or a resting control period immediately prior to influenza vaccination. An enhancement of the antibody response to one of the influenza viral strains was found in women in the psychologically stressed group compared to controls.

Conclusion and Future Directions

There is now considerable evidence that chronic stress is associated with a poorer antibody response to a variety of vaccinations. These *in vivo* findings give greater credence to the often reported association between psychological factors and immune function. The next challenge is to examine whether interventions designed to reduce psychological stress or improve more positive psychological characteristics can result in enhanced vaccination responses. Further, it is important to determine whether the immunomodulation associated with either chronic or acute stress is sufficient to result in differences in actual infection rates following vaccination. These studies would more fully establish causality as well as having clear clinical implications for improving protection following vaccination, particularly in vulnerable populations.

See Also the Following Articles

Immune Cell Distribution, Effects of Stress on; Immune Function, Stress-Induced Enhancement; Immune Response; Immune Suppression; Immune System, Aging; Immunity; Depression, Immunological Aspects; Immune Surveillance – Cancer, Effects of Stress on.

Further Reading

- Burns, V. (2004). Stress and antibody response to vaccination: implications of animal studies for human clinical research. *Expert Reviews in Vaccines* 3, 141–149.
- Burns, V. E., Carroll, D., Ring, C., et al. (2003). Antibody response to vaccination and psychosocial stress in humans: relationships and mechanisms. *Vaccine* 21, 2523–2534.
- Cohen, S., Miller, G. E. and Rabin, B. S. Psychological stress and antibody response to immunization: a critical review

- of the human literature. *Psychosomatic Medicine* 63, 7–18.
- Edwards, K. M., Burns, V. E., Reynolds, T., et al. (2006). Acute stress exposure prior to influenza vaccination enhances antibody response in women. *Brain, Behavior, and Immunity* 20, 159–168.
- Glaser, R., Kiecolt-Glaser, J. K., Malarkey, W., et al. (1998). The influence of psychological stress on the immune response to vaccines. *Annals of the New York Academy of Sciences* 840, 649–655.
- Phillips, A. C., Burns, V. E., Carroll, D., et al. (2005). The association between life events, social support, and antibody status following thymus-dependent and thymus-independent vaccinations in healthy young adults. *Brain, Behavior, and Immunity* 19, 325–333.
- Vedhara, K., Cox, N. K. M., Wilcock, G. K., et al. (1999). Chronic stress in elderly carers of dementia patients and antibody response to influenza. *Lancet* 353, 627–631.

Vasoactive Peptides

W K Samson and M M White

Saint Louis University School of Medicine,
St. Louis, MO, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
Z T Resch and W K Samson, volume 3, pp 643–649,
© 2000, Elsevier Inc.

Vasoactive Peptides

Vasoactive Peptides and the Pituitary Response to Stress

Integration of Central and Peripheral Responses to Stress

A Question of Physiological Relevance

Glossary

<i>Calcitonin gene-related peptide (CGRP)</i>	A peptide homologous in structure and function to AM.
<i>Chronotropism</i>	The rate of cardiac muscle contraction.
<i>Diuresis</i>	The excretion of water in urine.
<i>Inotropism</i>	The strength of cardiac muscle contraction.
<i>Intermedin</i>	The recently discovered member of the adrenomedullin-calcitonin gene-related peptide (AM-CGRP) family of peptides.
<i>Natriuresis</i>	The excretion of sodium in urine.
<i>Natriuretic peptides</i>	A family of three homologous mammalian peptides, encoded by three distinct genes, that are produced in heart and

Preproadrenomedullin-derived peptides

Receptor activity modifying proteins (RAMPs)
Salt appetite

the vasculature and that were initially discovered on the basis of their renal actions to stimulate sodium excretion.

One mammalian gene encodes a prohormone that is posttranslationally modified to yield two vasoactive peptides: adrenomedullin (AM) and proadrenomedullin N-terminal 20 peptide (PAMP).

Receptor proteins that provide selectivity for the actions of adrenomedullin and calcitonin gene-related peptide.

The behavior of selective ingestion of osmotically active solute.

Regardless of the operant definition, stress elicits significant, physiologically relevant changes in not only emotional behaviors, but also endocrine and cardiovascular function. Poised appropriately between the perception of a stressor and the integrative response to that insult are vasoactive peptides, factors that exert potent physiological effects in brain, pituitary gland, and the periphery. In many ways, these hormones serve to integrate the responses of multiple tissue systems in a manner parallel and complementary to the autonomic nervous system. Indeed, vasoactive peptides act in brain to modulate sympathetic efferent activity, fluid and electrolyte balance, and neuroendocrine responses to stress. In the pituitary gland, they significantly contribute to the regulation of

corticotroph function. In the periphery, these hormones determine vascular tone, cardiac pumping efficiency, adrenal hormone production and release, and, importantly, renal handling of fluid and electrolytes. Two families of vasoactive hormones, known to be released in response to numerous stressors, provide excellent examples of the integrative biology of the stress response. The natriuretic peptides and members of the adrenomedullin-calcitonin gene-related peptide (AM-CGRP) family of peptides are produced in brain and pituitary gland, as well as in the peripheral vasculature, heart, and kidney. They act both locally as autocrine factors and after release into the circulation as classic hormones to in some cases initiate and in others modify stress responses at the cellular and whole organ level.

Vasoactive Peptides

The simplest definition of a vasoactive peptide is a small protein that exerts either prorelaxant or procontractile effects in intact blood vessels. These effects may be endothelium dependent, requiring the recruitment of additional peptides, prostaglandins, or diffusible gases from the endothelial cell, or they may be endothelium independent, reflecting a direct effect on the vascular smooth muscle itself. Vascular tone is determined primarily by neurogenic mechanisms, largely mandated by the level of sympathetic activation. Circulating and locally produced factors also play important roles in the selective activation of distinct vascular beds and in the overall maintenance of vascular pressure, the single most important determinant of adequate perfusion of the vital organs in stress. Local factors not only directly affect vascular reactivity, but also, via presynaptic mechanisms, modify sympathetic tone. These factors also contribute to the maintenance of adequate perfusion pressures by altering cardiac performance and adjusting vascular volume.

Natriuretic Peptides

Originally discovered in extracts of cardiac atrium, the prototype of natriuretic peptides was originally named atrial natriuretic peptide (ANP) due to its isolation site, biological action, and protein structure. A second member of the family, originally thought to be produced in brain, but now known to be of almost exclusively myocyte (indeed, ventricular myocytes) origin, is B-type natriuretic peptide (BNP). Both ANP and BNP are released in response to volume overload, sensed by the myocardium as excess stretch. They share major actions to stimulate sodium and water excretion in the urine ([Figure 1](#)). ANP appears more relevant under physiological conditions, and

BNP seems to be recruited in more chronic conditions of volume/pressure overload. Production of at least ANP also occurs in brain and pituitary gland, and actions there related to stress are numerous and diverse. The third member of this family of peptides, all products of unique genes, is C-type natriuretic peptide (CNP). Produced predominantly in brain and vasculature, very little, if any, CNP is produced in the myocardium. Although CNP is a poor natriuretic and diuretic agent, it is a profound vasodilator and can exert apparently paradoxical effects on sympathetic tone via central nervous system (CNS) actions. Much is known of the molecular biology, pharmacology, physiology, and pathophysiology of the natriuretic peptides.

Proadrenomedullin-Derived Peptides and Their Homologs

Surprisingly, vasodilatory activities were isolated from extracts of a pheochromocytoma, and these peptides were later discovered to be posttranslational products of the same messenger RNA. Produced not only in adrenal medulla, but even more widely in the blood vessels, brain, and pituitary gland, these two vasodilatory factors exert complementary actions via unique signaling mechanisms ([Figure 2](#)). Adrenomedullin (AM), a 52-amino-acid peptide from the C-terminal portion of the proadrenomedullin precursor, exerts its vasorelaxant effects by stimulating nitric oxide synthase activity in endothelial cells of the vasculature. Proadrenomedullin N-terminal 20 peptide (PAMP), however, is a potent vasorelaxant due to its ability to exert presynaptic inhibition on sympathetic fibers innervating the vessel and to negatively modulate cholinergic activation of the chromaffin cells of the adrenal medulla and therefore reduce the level of epinephrine released in stress. Not all the actions of AM and PAMP are shared. For example, AM is a potent natriuretic and diuretic agent but does not affect catecholamine release significantly. A role for PAMP in sodium homeostasis also exists, however, as even in the absence of renotropic actions PAMP can stimulate sodium excretion due to its profound inhibitory effects on aldosterone secretion.

CGRP is a 37-amino-acid homolog of AM, and both are members of a larger family of peptides that includes calcitonin, amylin, and the recently described 47-amino-acid peptide intermedin (IMD; also named adrenomedullin-2). There is some evidence that these peptides may share an ancestral gene, and in addition to their structural similarities, they exert many shared biological activities (and perhaps receptors). CGRP, like AM, is a potent vasodilator when administered intravenously and has been demonstrated to inhibit aldosterone secretion *in vitro*.

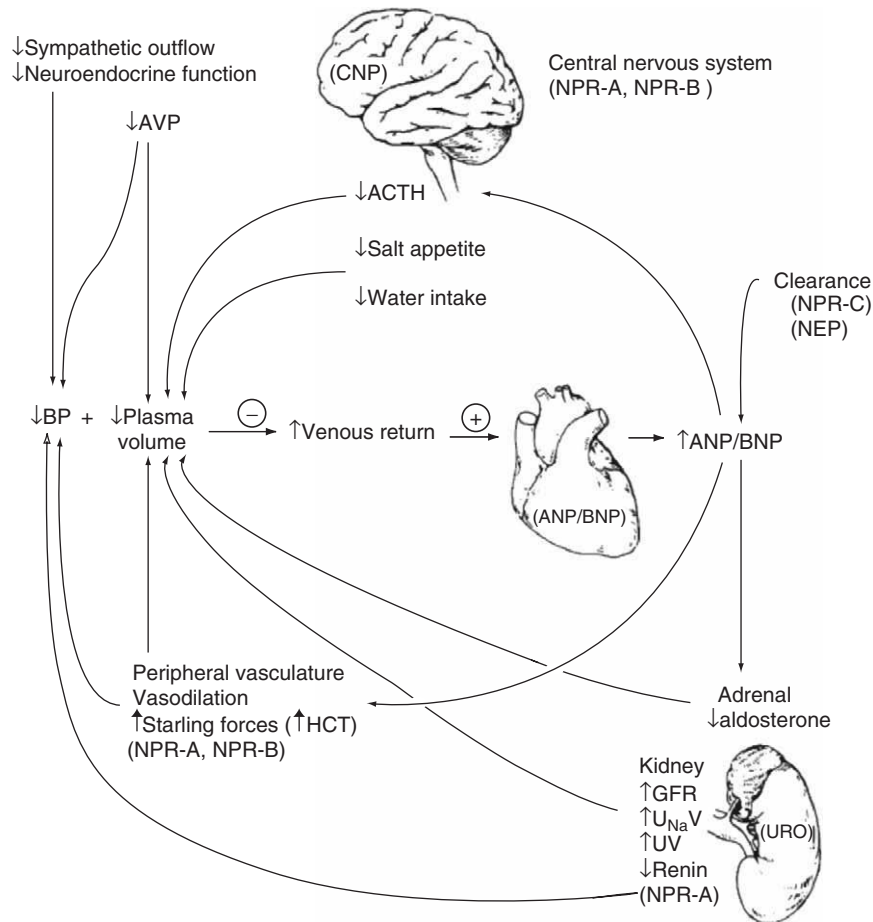


Figure 1 Pharmacological actions of the atrial natriuretic peptide (ANP) and B-type natriuretic peptide (BNP). AVP, vasopressin release; ACTH, adrenocorticotropin release; VR, venous return; BP, blood pressure; HCT, hematocrit; NPR-A, guanylyl cyclase A receptor; NPR-B, guanylyl cyclase B receptor; GFR, glomerular filtration; $U_{Na}V$, urinary sodium excretion; UV, urine volume. Also indicated is the production of C-type natriuretic peptide (CNP) in brain and urodilatin (URO), an N-terminally extended form of ANP in kidney. ANP and BNP are removed from the circulation by binding to the clearance receptor (NPR-C) and via degradation by neutral endopeptidase (NEP). Modified from the original and reprinted from Samson (1995), with permission of Humana Press.

When given into the CNS, like AM, it activates sympathetic tone, thus raising blood pressure. IMD exerts these same peripheral and central actions related to cardiovascular function and, like AM, exerts direct effects in the heart to increase contractility (inotropism) and beating frequency (chronotropism). Renal effects of IMD may parallel those of AM as well. It is not surprising that these homologous peptides exert similar actions (although unique effects may exist for each) since they share a common family of biological receptors. Those receptors are actually a complex of at least two proteins, the calcitonin receptor-like receptor (CRLR) and one of three possible receptor activity modifying proteins (RAMP1, RAMP2, or RAMP3). The relative combination of the respective RAMPs with the CRLR defines the specificity of the binding of each ligand. The CRLR–RAMP1 complex binds CGRP preferentially, while CRLRs in association with either RAMP2 or RAMP3

are the AM receptors. IMD binds promiscuously with all three CRLR–RAMP complexes, and thus it is not surprising that the hypotensive effects of IMD are at least partially blocked by relatively selective CGRP and/or AM antagonists.

Vasoactive Peptides and the Pituitary Response to Stress

Conventionally, the pituitary response to stress is considered to be the release of the hormone adrenocorticotropin (ACTH). The release of ACTH is controlled mainly by corticotropin-releasing hormone (CRH) levels in the hypophysial portal plasma; however, stimulation by vasopressin and oxytocin, as well as epinephrine, has been demonstrated. Thus, the effect of CRH can be modulated physiologically by factors also present in the portal circulation or by factors produced locally in the pituitary gland itself.

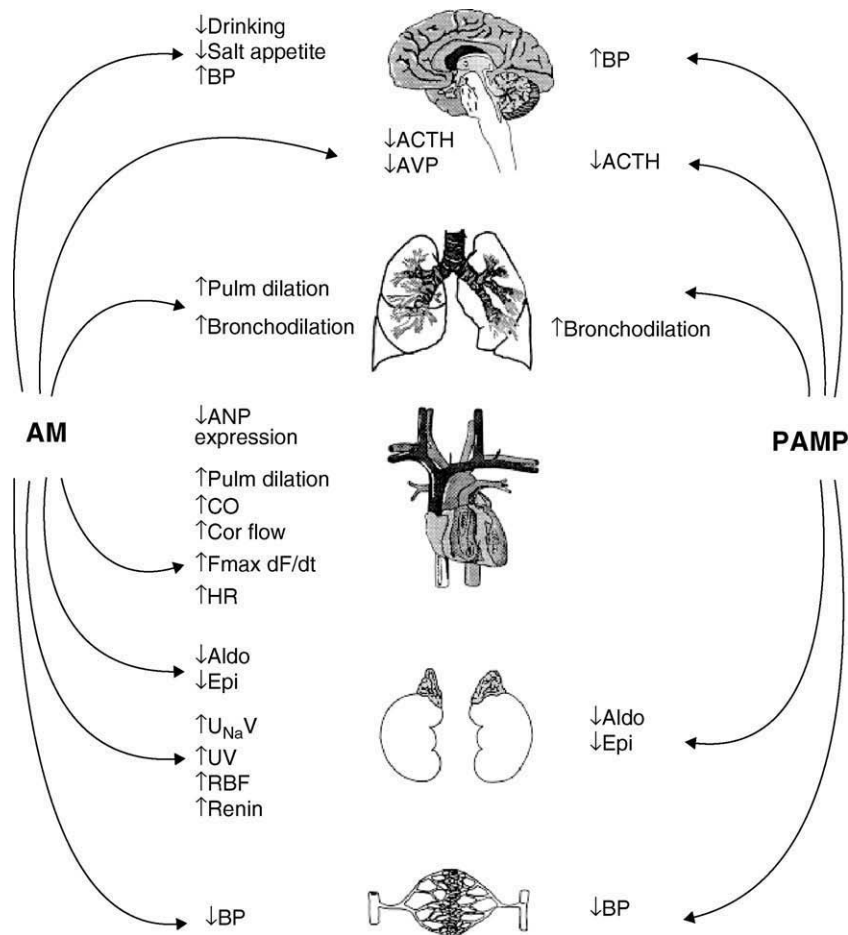


Figure 2 Pharmacological actions of adrenomedullin (AM) and proadrenomedullin N-terminal 20 peptide (PAMP). BP, blood pressure; ACTH, adrenocorticotropin secretion; AVP, vasopressin release; Pulm dilation, pulmonary dilation; ANP, atrial natriuretic peptide; CO, cardiac output, Cor Flow, coronary blood flow; Fmax dF/dt; maximum force of contraction; HR, heart rate; Aldo, aldosterone secretion, Epi, epinephrine release; U_{NaV} , urinary sodium excretion; UV, urine volume, RBF, renal blood flow, ET, endothelin production. Reproduced from Samson (1999), with permission from the *Annual Review of Physiology*, Volume 61, © 1999, by Annual Reviews.

Natriuretic Peptides

ANP is produced in the hypothalamus, where it exerts inhibitory effects on the releases of vasopressin and CRH. These effects may have physiological relevance, as stress-induced ACTH secretion is accentuated in rats pretreated centrally with anti-ANP antiserum, and glucocorticoid negative feedback is reduced when endogenous, brain-derived ANP is neutralized with the same antibodies. The effects are not limited only to ANP, as central administration of CNP also inhibits the stress response.

In addition to production in neurons projecting to stress centers in the hypothalamus (the paraventricular nucleus in particular), ANP and CNP are produced in cells of the anterior pituitary gland, where again CNP appears to be the most abundant gene product. Indeed, both the natriuretic peptide A receptor (NPR-A), which binds ANP preferentially, and the natriuretic peptide B receptor (NPR-B), which

prefers CNP, are present. Even the third natriuretic peptide receptor (NPR-C), the clearance receptor, is found in the gland, although to a much lesser degree. Importantly, the NPR-A and NPR-B receptors have been localized to the corticotrophs, lactotrophs, and gonadotrophs. Both the NPR-A and the NPR-B are particulate guanylyl cyclase-linked receptors, activation of which results in elevated cellular levels of cGMP. ANP, but not CNP, potently stimulates the accumulation of cGMP levels in dispersed anterior pituitary cells. Activation of guanylyl cyclase is known in many cells to result in the stimulation of cGMP-dependent phosphodiesterases, which limit the activity of preformed cAMP. Because CRH stimulates ACTH secretion by activating adenylyl cyclase with the cAMP formed then stimulating protein kinase A activity, it was not surprising that ANP, but not CNP, could inhibit CRH-induced ACTH secretion in cultured pituitary cells.

What pool of ANP could act in this fashion *in vivo*? ANP not only is produced in the pituitary gland, but also is delivered in the portal blood. Thus, the peptide could be acting to circumscribe ACTH secretion via a local paracrine mechanism or via a more classical neuroendocrine fashion. Are these pharmacological effects of ANP in cell culture physiologically relevant? While one group failed to observe any significant effects of physiological doses of ANP on CRH-stimulated ACTH secretion in humans, two other human trials did reveal a significant inhibitory effect of intravenously administered ANP on ACTH release.

Preproadrenomedullin-Derived Peptides

The adrenomedullin gene is transcribed in the hypothalamus, and AM-immunoreactive positive neurons projecting to neuroendocrine centers related to stress have been identified. Indeed, direct application of AM caused depolarization of parvocellular neurons of the paraventricular nucleus (PVN) in brain slices. Although the chemical phenotype of those AM-responsive parvocellular PVN neurons was not identified, they are composed of at least two classes of neurons, both playing a role in the hypothalamic control of the stress response. One subset projects to median eminence and delivers releasing factors, including CRH, to the hypophyseal portal vessels for transport to the anterior pituitary gland. The other subset, also including some CRH-producing neurons, projects to autonomic centers in brain stem and thus may coordinate the cardiovascular stress response with that of the hypothalamic-pituitary-adrenocortical (HPA) axis. In fact, AM administration has been reported to increase Fos expression in CRH-positive hypothalamic neurons and to activate the HPA axis. Convincing evidence for a direct action of AM on CRH-producing neurons that project to median eminence comes from studies in which the ability of centrally administered AM to stimulate ACTH release was abrogated by pretreatment with a neutralizing anti-CRH antibody. The central action of CGRP to stimulate the HPA axis similarly was blocked by pretreatment with an anti-CRH antibody, and recently IMD was demonstrated to activate the HPA axis by stimulating CRH release into the median eminence.

While AM immunoreactivity is not abundant in the external layer of the median eminence, and thus a neuroendocrine role for AM or PAMP in the regulation of ACTH secretion cannot be predicted, AM gene products are widely distributed in the anterior lobe of the pituitary gland and binding for AM has been demonstrated. On the basis of the ability of AM and PAMP to inhibit angiotensin II and potassium-stimulated aldosterone secretion from zona

glomerulosa cells *in vitro*, it was predicted that the peptides might also inhibit ACTH release from corticotrophs in culture. Basal and CRH-stimulated ACTH secretion *in vitro* is inhibited significantly by doses of AM, which approximate those present in the gland *in vivo*. The effects are dose and time dependent and are not related to the well-known ability of the peptide to activate adenylyl cyclase in peripheral tissues. In fact, the mechanism of the inhibitory effect of AM on ACTH remains unknown, as the peptide did not stimulate adenylyl or guanylyl cyclase activity in these cell cultures and similarly did not alter the levels of nitric oxide present during the incubations. A direct effect on the membrane polarization state could be hypothesized; however, a potassium channel blocker, glybenclamide, did not alter the action of AM to inhibit ACTH secretion under basal conditions.

The second posttranslational product of preproadrenomedullin (PAMP) also inhibits basal ACTH in a time- and dose-dependent fashion. It too exerts this effect at physiological concentrations and, like AM, does not directly alter adenylyl or guanylyl cyclase activity *in vitro*. Unlike AM, PAMP does not affect CRH-stimulated ACTH secretion. The ability of PAMP to reduce basal, but not stimulated, hormone secretion suggested an effect on the membrane polarization state. Because PAMP has been reported to act in other cell lines via an ATP-sensitive potassium channel, the pituitary effect was hypothesized to be due to opening of such a channel and hyperpolarization of the corticotroph membrane. Indeed, the ATP-sensitive potassium channel blocker glybenclamide completely blocked the inhibitory action of PAMP on ACTH secretion.

It remains to be determined whether these *in vitro* findings reflect physiological relevance in the *in vivo* setting. However, evidence for an *in vivo* effect of at least AM on ACTH secretion comes from studies in conscious sheep where the intravenous, but not intracerebroventricular, administration of AM lowered plasma levels of ACTH. It will be important to determine how the endogenous production of AM and PAMP in the pituitary gland is regulated. In the periphery, inflammatory stimuli such as interleukins and tumor necrosis factor- α are potent stimulators of AM gene transcription and peptide release. Thus, proinflammatory signals may also regulate pituitary production of the peptides as well, and proinflammatory cytokines are known effectors of ACTH secretion.

It is in the pituitary gland that the biological actions of the members of the AM-CGRP-IMD family of peptides diverge. Unlike the inhibitory effects of AM on ACTH release, CGRP has been reported to

stimulate hormone secretion via a cAMP-protein kinase A-dependent pathway. On the other hand, IMD has been reported to lack significant effects on basal or CRH-stimulated ACTH release. Furthermore, IMD inhibits releasing factor-stimulated growth hormone (GH) secretion *in vitro*, which is quite the opposite of the reported actions of AM and CGRP to stimulate GH secretion. The effect of IMD on somatotroph function suggests the existence of an IMD receptor unique from the CRLR-RAMP receptor complexes. This *in vitro* action of IMD may be physiologically relevant since in rodents stress results in inhibition of GH release and this peptide acts both centrally and at the level of the pituitary gland to decrease GH release. The actual source of the IMD that may be exerting these actions is unknown; however, the gastrointestinal tract is a prime candidate, since the highest levels of the peptide have been detected in stomach, as well as in kidney and brain.

All three members of the AM-CGRP-IMD family of peptides exert behavioral effects that may reflect a response to stress. The peptides inhibit food appetite and water drinking when administered into the brain, and at least in the case of AM, the effects on water drinking and salt appetite have been shown to be physiologically relevant. It appears that the actions of AM are more related to fluid and electrolyte homeostasis and those of CGRP to energy intake. Intermedin exerts anorexigenic effects similar to those of CGRP; however, the actions of IMD on thirst more closely parallel those of AM.

Integration of Central and Peripheral Responses to Stress

The integrative response to stress requires a coordination of multiple organ systems. Because of their diverse sites and mechanisms of action, vasoactive peptides are excellent candidates for physiological regulators of these integrative responses. Natriuretic peptides provide the best example of factors that exert complementary actions in multiple tissue sites. Their ability to unload excess vascular volume via direct natriuretic and diuretic actions in kidney is matched by inhibitory actions on releases of the volume conservatory hormones vasopressin, renin, and aldosterone. Behavioral actions within brain also complement these renal effects. ANP not only stimulates fluid and electrolyte excretion, but also, in what is now known to be a physiologically relevant manner, inhibits water drinking and salt intake, thus allowing time for the renal mechanisms to compensate before the addition of new fluid volume to the already taxed vascular tree. In states of volume depletion, these mechanisms are inhibited, thus

permitting sodium and water retention and additional intakes to replenish lost volume. The blood pressure-lowering effects of the natriuretic peptides are not remarkable, but they can be demonstrated at pharmacological doses *in vivo*, and data from ANP and NPR-A knockout mice suggest an important role for the peptide in blood pressure regulation.

In times of excess secretion, such as during cardiac overload states, endogenous natriuretic peptides may unload the heart by a vasodilatory mechanism that is complemented by increases in vascular permeability. One group has demonstrated a sympatholytic effect of ANP in the brain stem, which would also lower vascular pressure. Thus, the actions of natriuretic peptides seem well coordinated to facilitate fluid and electrolyte balance and the response to volume stress. Whether the peptides play any role in psychogenic stress remains to be seen. However, the ability of both ANP and CNP to inhibit the hypothalamic drive for gonadotropin secretion suggests that some stress-induced infertility may involve the hypothalamic actions of these peptides. Evidence for the physiological relevance of this effect has been provided in studies employing selective cell targeting of NPR-B receptive neurons in hypothalamus.

Adrenomedullin gene products provide a very different example of the potential importance of vasoactive peptides in the body's response to stress. Because of its profound effects on brain, pituitary, vascular, adrenal, and renal mechanisms related to fluid and electrolyte homeostasis, overexposure of the body to high levels of AM could have disastrous consequences. AM potently inhibits water drinking and sodium appetite, ACTH secretion, aldosterone release, vascular tone, and renal fluid reabsorption. Thus, when released in large amounts, as in sepsis, potentially fatal hypotension and the resultant underperfusion of vital organ systems could result. Fortunately, the actions of AM, and to some degree PAMP, include compensatory effects that protect against cardiac collapse. Indeed, in clinical studies the levels of AM attained in plasma during sepsis predict survival, probably due to the cardioprotective effects of those high levels of peptide. Furthermore, transgenic overexpression of the AM gene in mice results in enhanced survival outcomes during experimentally induced septic shock.

How is this accomplished? In brain, AM and PAMP are sympathostimulatory. This may seem a paradox in light of the profound hypotensive effects of the peptide in the periphery. However, their action in brain may compensate for not only the vasorelaxant effects of the peptides, but also the ability of AM to dampen baroreflexes and thus the perception of volume depletion. This central activation of

sympathetic efferents may assure adequate perfusion of vital tissues. Similarly, direct effects of AM on cardiac contractility have been documented. While decreasing venous return by peripheral vasodilation, AM also exerts positive inotropic and chronotropic effects in heart, which can maintain ejection fraction and stroke volume, again protecting against cardiac collapse. At least one additional protective effect of AM has been demonstrated. Even in the face of profound hypotension, renal perfusion pressure is maintained due to the dilatory effects of the peptide on both afferent and efferent renal arterioles. Thus, AM and PAMP exert protective effects that assure perfusion of the vital organ systems during hypotension, and these mechanisms must be considered important for the response of an individual to stress.

A Question of Physiological Relevance

With the exceptions of CNS actions of both ANP and AM to inhibit thirst and salt appetite, of AM to stimulate vasopressin release, of CNP to inhibit GnRH release, and of ANP to contribute to blood pressure regulation during sodium excess, much of the current knowledge of the physiological relevance of the vasoactive peptides remains based solely on pharmacological evidences. Only recently have combinations of methodologies been attempted to identify and confirm physiological relevance. For example, the salt appetite-suppressive effect of AM was suggested to have physiological relevance by studies employing selective antibodies that neutralized endogenous brain-derived AM. These initial findings were then confirmed by the antisense oligonucleotide approach in studies in which excessive salt appetite occurred when the endogenous production of AM was interrupted.

Why is it important to establish the physiological relevance of any peptide's pharmacological effects? The obvious answer is the potential use of an exogenous peptide in a therapeutic setting. Less obvious is the importance of elucidating the potential involvement of a given peptide in the genesis of a pathological state. Because of their pleuripotent biological activities, vasoactive peptides hold promise for both therapeutic use and the diagnosis of evolving pathology. In particular, plasma BNP levels are now

accepted as diagnostic of the progression of cardiac overload in congestive heart failure in humans, and ANP is a clinical tool employed to unload excessive vascular volume in this disease. The ACTH-inhibiting effect of ANP during such treatment would certainly provide an extra therapeutic advantage. This again points to the significance of multiple sites and mechanisms of actions of vasoactive peptides during challenges to normal homeostasis, which by any definition mean stress.

See Also the Following Articles

Hypertension; Peptides; Salt Appetite.

Further Reading

- John, S. W. M., Krege, J. H., Oliver, P., et al. (1995). Genetically decreased levels of atrial natriuretic peptide and salt-sensitive hypertension. *Science* **267**, 679–681.
- Samson, W. K. (1995). Cardiovascular hormones. In: Conn, P. M. & Melmed, S. (eds.) *Endocrinology: basic and clinical principles*, pp. 361–376. Totowa, NJ: Humana Press.
- Samson, W. K. and Taylor, M. M. (2005). Cardiovascular hormones. In: Conn, P. M. & Melmed, S. (eds.) *Endocrinology: basic and clinical principles* (2nd edn., pp. 321–335). Totowa, NJ: Humana Press.
- Taylor, M. M. and Samson, W. K. (2002). Adrenomedullin and the integrative physiology of fluid and electrolyte balance. *Microscopy Research and Techniques* **57**, 105–109.
- Taylor, M. M., Bagley, S. L. and Samson, W. K. (2004). A possible mechanism for the action of adrenomedullin in brain to stimulate stress hormone secretion. *Endocrinology* **145**, 4890–4896.
- Taylor, M. M., Bagley, S. L. and Samson, W. K. (2005). Intermedin/adrenomedullin-2 acts within central nervous system to elevate blood pressure and inhibit food and water intake. *American Journal of Physiology* **288**, R919–R927.
- Taylor, M. M., Bagley, S. L. and Samson, W. K. (2005). Stress hormone secretion is altered by central administration of intermedin/adrenomedullin-2. *Brain Research* **1045**, 199–205.
- Taylor, M. M., Bagley, S. L. and Samson, W. K. (2006). Intermedin/adrenomedullin-2 inhibits growth hormone release from cultured, primary anterior pituitary cells. *Endocrinology* **147**, 859–864.

Vasopressin

L P Renaud

Ottawa Health Research Institute and
University of Ottawa, Ottawa, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
L P Renaud, volume 3, pp 650–655, © 2000, Elsevier Inc.

The Posterior Pituitary Hormones

Vasopressin Synthesis

Vasopressin Release

Receptors

Vasopressin in the Brain

Vasopressin and the Hypothalamic-Pituitary Axis

Glossary

<i>Diabetes insipidus</i>	A condition caused by insufficient vasopressin in the circulation or inability of vasopressin to activate receptors in the kidney; manifest by polyuria and polydipsia.
<i>Hormone</i>	A peptide that circulates in the blood and exerts its biological effects by binding to specific receptors on select target tissues.
<i>Hypothalamus</i>	The part of the central nervous system residing above the pituitary, containing neurons that control vital functions (e.g., pituitary secretions, thermoregulation, and homeostasis in the autonomic, cardiovascular, and hydromineral systems).
<i>Median eminence</i>	The area at the base of the third cerebral ventricle that contains two zones: an inner zone that has axons in passage to the posterior pituitary and an external zone that contains the axon terminals of parvocellular neurons that synthesize and release peptides into adjacent portal capillaries.
<i>Neurohypophysis</i>	The posterior part of the pituitary, composed mainly of axon terminals of hypothalamic magnocellular neurosecretory neurons.
<i>Nonapeptides</i>	Peptides made up of nine amino acids.

Vasopressin is a nonapeptide that functions as a hormone in the blood and as a neurotransmitter or neuromodulator in the brain. Vasopressin is synthesized as a precursor molecule, prepropressophysin, within neurons with specialized functions. In magnocellular

neurosecretory neurons of the hypothalamus, the precursor is stored in vesicles and undergoes posttranslational processing during axonal transport to terminals in the neurohypophysis. Vasopressin release into the systemic circulation occurs in response to the activation of these neurons by specific stressors that challenge hydromineral balance (e.g., dehydration) and cardiovascular homeostasis (e.g., blood loss and hypotension). Once in the circulation, vasopressin assumes key hormonal roles by binding to V₁ type receptors on vascular smooth muscle to produce vasoconstriction and to V₂ type receptors in the kidney to increase water reabsorption and antidiuresis. Other peripheral actions include platelet aggregation; liver glycogenolysis; and secretion of aldosterone, insulin, and atrial natriuretic factor by the adrenal gland, pancreas, and heart, respectively. Vasopressin synthesized centrally in other hypothalamic and extra-hypothalamic neurons in the brain acts mainly through V₁ type receptors to mediate several functions, including the modulation of neuronal excitability via autocrine, paracrine, or transmitter mechanisms, to augment the release of adrenocorticotrophic hormone (ACTH) in the anterior pituitary during stress, to attenuate febrile responses, and to influence cognition and complex social behaviors.

The Posterior Pituitary Hormones

The discovery in 1895 that injections of extracts of beef pituitary increased arterial blood pressure marked the beginning of a century of research into the biology of the neurohypophysial hormones, identified as two structurally similar nonapeptides, oxytocin and vasopressin. Their main biological effects, the contraction of vascular and uterine smooth muscle and antidiuresis, were initially assessed by sensitive bioassays. Elucidation in the 1950s of the chemical structure of vasopressin as distinct from oxytocin permitted the development of specific radioimmunoassays, receptor agonists and antagonists, and techniques to investigate the molecular biology and function of vasopressin and oxytocin genes. Although discovered through their peripheral hormonal actions following their release from the posterior pituitary, an important aspect of this evolution was the recognition that vasopressin and oxytocin are also neuropeptides whose synthesis and release by different populations of neurons in the brain can influence a variety of central nervous system functions.

Vasopressin Synthesis

In Hypothalamic Magnocellular Neurons

Vasopressin and oxytocin molecules circulating as hormones in the blood originate from specialized magnocellular neurons, termed neurosecretory cells, located within the hypothalamic supraoptic and paraventricular nuclei and accessory magnocellular nuclei. The accessibility of the magnocellular neurohypophysial neuronal system has provided investigators with a valuable model to study peptide biosynthesis. In accordance with their separate biological roles, the genes for vasopressin and oxytocin are usually localized in different magnocellular neurons. Vasopressin- and oxytocin-synthesizing neurons also contain genes that code for other neuropeptides, and these may be co-activated and co-expressed together with vasopressin (or oxytocin) under appropriate circumstances. Consistent with other neuropeptides, synthesis occurs in the cell somata, initially resulting in the formation of a large precursor molecule called preproressophysin that is composed of a signal peptide, the nonapeptide vasopressin (arginine-8 vasopressin, AVP, in most mammals; lysine-8 vasopressin in pig), a putative carrier molecule neurophysin, and a 39-amino-acid glycoprotein (co-peptin, CCP). After packaging into large (160-nm) neurosecretory vesicles, preproressophysin undergoes further posttranslational processing during transport in axons to storage sites in the posterior pituitary.

Magnocellular neurosecretory neurons constitutively express vasopressin in their somata and axons. They also demonstrate an increase in vasopressin hnRNA and mRNA within minutes of exposure to a physiological stimulus (examples include a rise in plasma osmotic pressure and the unloading of peripheral baroreceptors as a consequence of blood loss). Osmotic, hypotensive, and immune stimuli are also powerful activators of several immediate early genes, including cFos and cJun, in these neurons. Details of the linkage between these genes and the signal transduction mechanisms that regulate vasopressin gene transcription in health and disease (e.g., ectopic production by small-cell lung cancers) are topics under investigation.

A lack of vasopressin in animals and humans results in diabetes insipidus, a condition characterized by polyuria and polydipsia. Central diabetes insipidus may occur because of a lesion in the hypothalamic-neurohypophysial pathway or, more rarely, as an autosomal dominant inheritance in which mutations have been detected in the vasopressin signal peptide or neurophysin molecules. The only genetic animal model for hypothalamic diabetes insipidus is the

Brattleboro rat; discovered 30 years ago, this animal fails to produce sufficient vasopressin due to a single base pair deletion in its vasopressin gene. The recent development of a vasopressin-null mouse may provide new insights into various roles of vasopressin in the stress response, cardiovascular regulation, and behavior.

In Parvocellular Neurons

Immunocytochemical and radioimmunoassay techniques reveal that vasopressin is present in other areas of the brain and spinal cord and in certain peripheral autonomic ganglia. In the brain, distinct groups of vasopressin-synthesizing parvocellular neurons may be recognized by their vasopressin-like immunoreactivity or through their ability to express vasopressin mRNA and hnRNA. The latter forms of gene expression reveal that vasopressin synthesis may depend on specific stimuli. For example, in the hypothalamic paraventricular nucleus, corticotropin releasing hormone (CRH)-synthesizing neurons may co-express vasopressin only during stress.

Vasopressin Release

Into the Systemic Circulation

Levels of vasopressin in the circulation (expressed as picograms per milliliter) are normally low to undetectable. Vasopressin is released into the systemic circulation from storage sites in the neurohypophysis in response to specific challenges to cardiovascular or hydromineral homeostasis, causing plasma levels to rise to >20 pg/ml. In some clinical circumstances associated with head trauma, congestive heart failure, liver cirrhosis, and cancer (e.g., small-cell lung carcinomas), humans have been found to have an excess of vasopressin in the circulation relative to their plasma osmolality, resulting in free-water retention and hyponatremia, a syndrome of inappropriate antidiuresis.

Circulating levels of the hormone are normally regulated through release from the neurohypophysis, a structure comprising mainly pituicytes and axon terminals of the hypothalamic magnocellular neurons that contain stored neurosecretory vesicles. Action potentials generated at the cell somata of the magnocellular neurons (in response to specific stimuli) propagate into the nerve terminals, thereby activating voltage-gated calcium channels that promote a rise in intracellular calcium and triggering exocytosis. The efficiency of this stimulus-secretion coupling mechanism is dependent on both the rate and pattern of discharge of action potentials invading the nerve terminals; release is optimized not only by an increase

in action potential frequency up to and beyond 20 Hz but also by the emergence of bursting or phasic patterns of firing.

Stimuli for release

Plasma hyperosmolality The osmolality of the plasma is sensed in the brain by a central osmosensor complex composed of neurons in the lamina terminalis along the anterior wall of the third cerebral ventricle in addition to the hypothalamic magnocellular neurons themselves. Recent electrophysiological studies indicate that magnocellular neurons are intrinsically osmosensitive; they contain mechanosensitive stretch-inactivated cation channels that sense changes in cell volumes consequent to a deviation from resting plasma osmolality (~290 mosmol/kg in humans). A rise in the osmolality of the extracellular environment initiates cell shrinkage; an ensuing increase in a membrane cationic conductance results in membrane depolarization and a subsequent increase in action potential generation once the membrane potential reaches threshold (around -50 mV). In the rat, the increase in activity of vasopressin-synthesizing magnocellular neurons usually initiates a unique phasic-bursting pattern of firing, an efficient trigger for stimulus-secretion coupling.

Hypovolemia and hypotension In mammals, information about the state of the peripheral cardiovascular system is continuously supplied to the brain via afferent pathways in the glossopharyngeal and vagus cranial nerves. Baroreceptors monitor the tone in the main arteries around the heart and lungs. A drop in arterial blood pressure, as might result from hemorrhage, decreases the input from baroreceptors. This information is channeled through catecholaminergic pathways that ascend through the brain stem to the hypothalamic magnocellular neurons, triggering an increase in their activity and initiating hormone release in the posterior pituitary.

In the Pituitary Portal Circulation

The external zone of the median eminence contains an abundance of axon terminals. Those that demonstrate vasopressin-like immunoreactivity arise mainly from parvocellular neurons in the hypothalamic paraventricular nucleus. A subpopulation of these are neurons that normally synthesize only CRH but that also contain inactive vasopressin genes that become expressed under conditions of prolonged stress and then co-release both peptides into the pituitary portal circulation from vesicles stored in these median eminence axon terminals. Receptors for vasopressin reside on anterior pituitary corticotrophs.

In the Brain

The neurohypophysial peptides are found in relatively high concentrations (in the picogram per microliter range) within the extracellular space in specific areas of brain. Because blood-born peptides are not seen in these concentrations and do not cross the blood-brain barrier in significant amounts, this reflects central release. Two likely sources are synaptic release from the axon terminals of central vasopressin-synthesizing neurons and exocytosis from somata and dendrites. The latter event is now well recognized as a mode of peptide release in the hypothalamic magnocellular neurons of the paraventricular and supraoptic nuclei. Here, high concentrations of vasopressin and/or oxytocin can be measured by microdialysis, cell somata and dendrites display neurosecretory vesicles with exocytotic profiles, and electrical events consistent with dendritic release can be recorded.

Stimuli for brain release Stressors such as social defeat and forced swimming are potent stimuli to trigger the release of vasopressin into the brain. Release sites may vary with the stimulus. For instance, social defeat is associated with vasopressin release in the paraventricular nucleus but not in the supraoptic nucleus and causes no change in levels of vasopressin in the plasma. Forced swimming, on the other hand, promotes vasopressin release in both the paraventricular and supraoptic nuclei, again with virtually no change in plasma vasopressin levels. This illustrates a dissociation between the mechanisms that trigger central versus peripheral release. Central release mechanisms have not been studied as extensively as those in the posterior pituitary.

In the Cerebrospinal Fluid

Vasopressin is present in the cerebrospinal fluid of many mammals. Unlike levels in the plasma, cerebrospinal vasopressin may exhibit large daily rhythms with high concentrations during daylight hours, suggesting that the origin is regulated by the suprachiasmatic nucleus, a regulator of circadian rhythms in the mammalian brain.

Receptors

Peripheral and central effects of AVP are mediated by at least three subtypes of G-protein-linked membrane-bound receptors. These fall into two classes, V_1 (V_{1a} vascular and V_{1b}/V_3 pituitary) and V_2 , based on second-messenger cascades and pharmacological properties. The activation of V_1 class of receptor

subtypes ($V_{1a}R$ and $V_{1b}R$) results in the hydrolysis of phosphatidyl inositol and an increase in cytosolic calcium. The activation of the V_2 receptor (V_2R) results in the stimulation of adenylate cyclase.

The V_1 Receptors

This receptor has been cloned, and its mRNA (visualized by *in situ* hybridization histochemistry) is seen throughout the brain, liver, and kidney; the $V_{1a}R$ mediates the vasoconstrictor and hepatic glycogenolytic actions of vasopressin and is thought to be the predominant vasopressin receptor in brain. A subtype of the V_1 class, the $V_{1b}R/V_3R$, has been characterized in the anterior pituitary where it regulates vasopressin-mediated ACTH release by potentiating the actions of CRH. $V_{1b}R$ mRNA has also been detected in the thymus, heart, lung, spleen, kidney, uterus, breast, and brain.

V_2 Receptors

The V_2 receptors and V_2R mRNA are abundant in the kidney renal tubules, where they mediate the well-established antidiuretic properties by promoting free-water reabsorption. Aquaporin 2, the water channel in apical collecting duct cells, is present in the membrane of cytoplasmic vesicles and is translocated to the apical membrane in response to vasopressin. Vasopressin regulates two types of aquaporin 2 expression: one prompt and the other sustained. The failure of vasopressin action in the kidney renal tubules results in nephrogenic diabetes insipidus, a disorder that may be inherited, resulting from mutations in the genes for vasopressin V_2 receptors and/or aquaporin 2.

Vasopressin in the Brain

Initial studies on avoidance behavior in the early 1960s provided the first clue that vasopressin and related peptides were involved in learning and memory processes. It is now recognized that centrally released vasopressin is involved in a variety of roles, including rewarded behavior and drug tolerance; social, reproductive, and feeding behaviors; temperature and cardiovascular regulation; stress; and circadian rhythms.

Vasopressin Pathways

Techniques that can specifically define vasopressin-relevant regions of brain (ligand-binding assays, immunocytochemistry, and *in situ* hybridization) in combination with lesions and tract-tracing methods serve to define a chemical neuroanatomy of vasopressin in the brain. Prominent among these are the

magnocellular neurons of the hypothalamic supra-optic, paraventricular, and accessory magnocellular nuclei and cells groups whose axons project to the posterior pituitary. The paraventricular nuclei also contain subsets of parvocellular neurons that synthesize vasopressin. Some of these neurons project to the median eminence and form a component of the hypothalamic-pituitary-adrenal axis; the expression of their vasopressin genes may be evident only under certain circumstances.

Other paraventricular and hypothalamic neurons that contain vasopressin constitutively project axons to the parabrachial nucleus, dorsal vagal complex, ambiguus nucleus, medial and ventrolateral medulla, and spinal cord areas; these are likely to participate in control of regulatory functions attributed to the autonomic, sympathetic, and parasympathetic nervous systems (e.g., heart rate, blood pressure regulation, and release of adrenal medullary hormones).

Vasopressin-immunoreactive neurons occupy the dorsomedial core of the hypothalamic suprachiasmatic nucleus (SCN), site of a biological clock in the mammalian brain. These cells project into the medial hypothalamus and contribute to the entrainment of circadian rhythmicity in hypothalamic neuroendocrine and thermoregulatory functions. Functions of a prominent extrahypothalamic projection to the midline thalamic paraventricular nucleus remain to be defined. SCN vasopressin-synthesizing neurons are the likely source for the circadian rhythmicity in vasopressin observed in the cerebrospinal fluid.

Vasopressin neurons in the bed nucleus of the stria terminalis and medial amygdaloid nucleus project to the lateral septum, ventral septal nucleus, olfactory tubercle, lateral habenular nucleus, midbrain central gray, dorsal raphe nucleus, locus coeruleus, and ventral hippocampus.

Factors Affecting Central Vasopressin Gene Expression

Differing levels of mRNA indicate that vasopressin expression is not uniform. For example, adrenal steroids block vasopressin mRNA expression in the parvocellular neurons in the paraventricular nucleus but not in other brain areas. Vasopressin expression in lateral septum and medial amygdala requires the presence of gonadal steroids; most vasopressin mRNA disappears within a week of castration. How gonadal steroids work remains unclear because there are no recognizable androgen- and estrogen-responsive elements on the promotor region of the vasopressin gene. Presumably, these steroids do not affect mRNA transcription but, rather, regulate vasopressin mRNA levels at a posttranscriptional level because castration decreases the length of the mRNA poly(A) tail in the

bed nucleus of the stria terminalis, whereas testosterone increases it, thereby influencing the stability of the vasopressin mRNA.

Electrophysiology of Central Vasopressin Receptors

The exogenous application of vasopressin alters the excitability of neurons in a variety of areas, including the hippocampus, hypothalamus, brain stem, and spinal cord, all regions containing vasopressin fibers and/or neurons. In most instances, the response to vasopressin is manifested as an increase in neuronal firing, blockable with V_1 receptor antagonists. Mechanisms vary with the areas tested. Most investigations have focused on direct, tetrodotoxin-resistant actions mediated via postsynaptic receptors. For example, in rat facial nucleus, vasopressin receptors produce depolarization mediated by nonselective cationic channels. In the lateral horn of the neonatal rat spinal cord, vasopressin acting at G-protein-coupled V_1 type receptors produces depolarization in spinal preganglionic and other lateral and ventral horn neurons by decreasing a resting potassium ionic conductance and/or increasing a nonselective cationic conductance. Other actions are indirect, involving presynaptic receptors that attenuate glutamatergic or GABAergic neurotransmission. As previously reported, vasopressin is released locally by magnocellular neurons and the peptide is proposed to act in an autocrine or paracrine function to alter the excitability of the neurons responsible for its synthesis, either directly or through synaptic modulation. Calcium-imaging studies in dissociated vasopressin-synthesizing neurons from rat supraoptic nucleus reveal that signaling at vasopressin receptors involves calcium influx through L, N, and T type channels, an auto-regulation that seems to involve the activation of both protein kinase C- and adenylate cyclase-linked intracellular pathways.

Vasopressin and Thermoregulation

Studies in a variety of mammals have demonstrated the fever-reducing effects of vasopressin and the fever-enhancing actions of V_{1a} receptor antagonists injected into the ventral septal area. It is likely that this area receives its innervation from the bed nucleus of the stria terminalis. Interestingly, this is where electrical stimulation can attenuate pyrogen-induced fever, an effect that is blockable by injection of a V_{1a} receptor antagonist into the ventral septal area. Castration, which is followed by the disappearance of vasopressin in the septal area, lengthens pyrogen-induced fever, presumably because of a reduction in the amount of releasable vasopressin.

Vasopressin and Cognition

Extensive animal research has established a role for vasopressin in learning and memory. Less clear are the outcomes of similar efforts in humans. The hippocampal system has a prominent role in learning and memory, and current observations indicate a facilitating effect of vasopressin and vasopressin analogs on retrieval and relearning processes, mediated by both vasopressin V_1 type and oxytocin receptors located mainly in the ventral hippocampus.

Vasopressin, Monogamy, and Social Behavior

The neurohypophysial peptides have been implicated in the central mediation of complex social behaviors such as affiliation, parental care, and territorial aggression. For example, in the prairie vole, research suggests that vasopressin is involved in the control of several behaviors associated with monogamy, including pair bonding, paternal care, and mate guarding. Molecular studies suggest that changes in the regulation of oxytocin- and vasopressin-receptor gene expression underlie species differences in receptor distribution and might provide a mechanism for the evolution of monogamy in certain rodent species.

Vasopressin and the Hypothalamic-Pituitary Axis

The hypothalamic paraventricular nucleus contains two categories of neurosecretory neurons capable of synthesizing vasopressin. We have already referenced the magnocellular neurons that project their axons through the internal zone of the median eminence to the neurohypophysis. The following refers to a subpopulation of parvocellular paraventricular neurons capable of co-synthesizing CRH and vasopressin.

Vasopressin in Corticotropin Releasing Hormone Neurons

Under resting conditions, most paraventricular nucleus neurons that synthesize CRH neurons do not normally contain detectable levels of vasopressin. The expression of their vasopressin genes can be rapidly switched on by a variety of acute, repeated, or chronic challenges. Information revealing the presence of such challenges or stressors may reach the parvocellular neurons via several neuronal pathways: (1) descending inputs from circumventricular organs in the lamina terminalis along the anterior wall of the third cerebral ventricle, signaling information about osmolality, cytokine, and hormonal molecules in the circulation; (2) projections from diverse limbic system areas, conveying inputs related to psychological or emotional stress; (3) ascending inputs from catecholaminergic

neurons in the dorsomedial and ventrolateral medulla, relaying information related to autonomic and cardiovascular regulation; and (4) unspecified local intrahypothalamic innervations, including recurrent feedback pathways among these neurons themselves. Presumably the content and meaning of the information conveyed via different pathways reflects stimulus specificity and will engage different populations of paraventricular neurons to respond, thereby accounting for the different patterns of vasopressin mRNA and hnRNA expressed among populations of hypothalamic neurons in response to different stress paradigms.

The signaling that regulates the expression of the vasopressin gene in these cells remains to be defined. Distinct from their magnocellular counterpart, these parvocellular neurons contain type 2 glucocorticoid receptors that mediate a direct inhibitory effect of corticosteroids on neuropeptide gene expression. The removal of the adrenals, the source of this circulating corticosterone, triggers vasopressin synthesis and upregulates CRH synthesis in populations of parvocellular paraventricular nucleus neurons; the peptides are co-packaged into dense core vesicles, transported to the zona externa of the median eminence, released into the hypophyseal portal circulation, and carried (as hormones) into the anterior pituitary.

Vasopressin in the Anterior Pituitary

In the anterior pituitary, vasopressin potentiates the CRH-induced release of ACTH by binding to specific V_{1b} receptors on the corticotrophs, activating the inositol phospholipid cycle and protein kinase C, mobilizing calcium from intracellular stores, and triggering the release of ACTH from the rapid turnover pool. Genetically vasopressin-deficient Brattleboro rats have a blunted corticosterone response to certain forms of stressful stimuli, presumably reflecting the contribution normally made by vasopressin to the ACTH released by a stressor.

See Also the Following Articles

Corticotropin Releasing Factor (CRF); Neuroendocrine Systems; Oxytocin; Paraventricular Nucleus; Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress.

Further Reading

- Antoni, F. A. (1993). Vasopressinergic control of pituitary adrenocorticotropin secretion comes of age. *Frontiers in Neuroendocrinology* **14**, 76–122.
- Barberis, C., Mouillac, B. and Durroux, T. (1998). Structural bases of vasopressin/oxytocin receptor function. *Journal of Endocrinology* **156**, 223–229.
- Cole, R. L. and Sawchenko, P. E. (2002). Neurotransmitter regulation of cellular activation and neuropeptide gene expression in the paraventricular nucleus of the hypothalamus. *Journal of Neuroscience* **22**, 959–969.
- Kalsbeek, A. and Buijs, R. M. (2002). Output pathways of the mammalian suprachiasmatic nucleus: coding circadian time by transmitter selection and specific targeting. *Cell Tissue Research* **309**, 109–118.
- Landry, M., Vial-Porile, E., Hökfelt, T., et al. (2003). Differential routing of coexisting neuropeptides in vasopressin neurons. *European Journal of Neuroscience* **17**, 579–589.
- Ludwig, M. and Pittman, Q. J. (2003). Talking back: dendritic neurotransmitter release. *Trends in Neuroscience* **26**, 255–261.
- Péqueux, C., Keegan, B. P., Hagelstein, M.-T., et al. (2004). Oxytocin- and vasopressin-induced growth of human small-cell lung cancer is mediated by the mitogen-activated protein kinase pathway. *Endocrine-Related Cancer* **11**, 871–885.
- Raggenbass, M. (2001). Vasopressin- and oxytocin-induced activity in the central nervous system: electrophysiological studies using in-vitro systems. *Progress in Neurobiology* **64**, 307–326.
- Sabatier, N., Richard, P. and Dayanithi, G. (1998). Activation of multiple intracellular transduction signals by vasopressin in vasopressin-sensitive neurones of the rat supraoptic nucleus. *Journal of Physiology (London)* **513**(3), 699–710.
- Urban, I. J. A., Burbach, J. P. H. and DeWeid, D. (1998). *Progress in brain research. Vol. 119: Advances in brain vasopressin*. Amsterdam: Elsevier.
- Young, L. J. and Wang, Z. (2004). The neurobiology of pair bonding. *Nature Neuroscience* **7**, 1048–1054.
- Zingg, H. H., Bourque, C. W. and Bichet, D. G. (1998). *Advances in experimental medicine and biology. Vol. 449: Vasopressin and oxytocin: molecular, cellular and clinical advances*. New York: Plenum Press.

Vietnam Veterans, Postwar Experiences and Health Outcomes

J A Boscarino

Mount Sinai School of Medicine, New York, NY, and
Geisinger Health System, Danville, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article
by J A Boscarino, volume 3, pp 656–661, © 2000, Elsevier Inc.

The Vietnam War

Vietnam Veteran Studies

Psychobiological and Biomedical Findings

Studies of Other Veteran Populations after Vietnam

Implications for Medical Sciences, Research, and Treatment

Conclusion

Glossary

Autoimmune disease A disease produced when the body's normal tolerance of its own antigenic markers on the surface of cells is lost.

Cohort study An observational study design commonly used in epidemiological studies of populations exposed to pathogens that could result in future diseases.

Diagnostic Interview Schedule (DIS) A mental health diagnostic interview based on standard psychiatric nomenclature that made population mental health surveys and population-level estimates possible.

Diagnostic and Statistical Manual of Mental Disorders, Third Edition (DSM-III) The third edition of standard psychiatric nomenclature used to classify all emotional and mental health disorders in 1980 and the edition in which the definition of PTSD was included.

Heritability The proportion of variance of the liability to a disease that is due to shared genetic factors.

Hypothalamic-pituitary-adrenal (HPA) system Three key anatomic structures that collectively participate in hormonal responses to stressors by regulating human neuroendocrine functions and, hence, the body's adaptation to environmental stressors.

National Vietnam Veterans Readjustment Study (NVVRS) A national cross-sectional survey of male and female Vietnam veterans and non-veterans conducted to determine the prevalence and etiology of psychosocial readjustment problems and mental disorders among Vietnam veterans.

Operation Ranch Hand The U.S. military's operational code name for the application of herbicides (chiefly agents orange, blue, green, pink, purple,

and white) in Vietnam by aircraft in order to destroy enemy cover and food crops.

A follow-up study of the health status of the men involved in the Operation Ranch Hand herbicide operations.

A specific disabling psychiatric disorder that develops after exposure to psychologically traumatic events; the disorder could last for decades. Symptoms include the reexperiencing of the event, avoidance of stimuli related to the event, and persistent symptoms of increased arousal associated with the event.

A field of medical research that focuses on the association between neuroendocrine function and psychological status, especially as this relates to steroidogenic biological alterations and psychiatric illness.

A group of symptoms of a disordered function related to one another by means of a common pathology.

A national cohort study of male, U.S. Army Vietnam-theater and non-Vietnam-theater veterans, conducted to determine the prevalence and etiology of diseases and medical conditions associated with Vietnam service.

A registry composed of male twin pairs who both served in the U.S. military between 1965–1975 in order to study the effect of genetic factors on health status.

Operation Ranch Hand Study
Posttraumatic stress disorder (PTSD)

Psychoneuro-endocrinology

Syndromes

Vietnam Experience Study (VES)

Vietnam Era Twin (VET) Registry

The Vietnam War

Although U.S. involvement in the Vietnam War officially ended over 30 years ago for the 3.14 million men and 7200 women who served there, the impact of their postwar experiences altered society in many ways. The postwar experiences of Vietnam veterans were different from veterans of many previous wars. For example, unlike earlier U.S. wars in the twentieth century, the Vietnam veteran returned home from a conflict that was controversial and unpopular. Furthermore, many veterans were draftees who often came from lower socioeconomic (SES) households. In addition, of all men eligible for the military draft, only a specific subgroup was primarily inducted, due to the Selective Service System's deferment policy, which generally exempted men from higher SES groups. Within this context, it became clear to at least some health-care professionals during the war

that many returning veterans were experiencing postwar adjustment problems. During this period, the work of Charles Figley stood out as a classic in stress research because it documented the emotional and psychological trauma many veterans began to manifest by war's end. Based on Figley's work, and that of others, health-care researchers and medical professionals in the late 1970s formulated specific etiological models of the pathogenesis of mental disorders among Vietnam veterans. They also contributed to the development of case definitions to characterize these syndromes. These efforts eventually culminated in the inclusion of posttraumatic stress disorder (PTSD) in the *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn.; DSM-III). Once this was achieved, more accurate assessments of the scope and nature of traumatic stress (and related disorders) among Vietnam veterans and others were possible. Finally, studies of the postwar experiences of Vietnam veterans have led to many research developments and have accelerated the accumulation of knowledge related to the psychology and biology of PTSD, including the possible long-term medical and health consequences of this syndrome.

Although clinical documentation of the adverse psychological effects of combat exposure goes back at least as far as the U.S. Civil War, the etiology of these syndromes were neither clearly understood nor well defined. Ninety years ago during World War I, soldiers with acute adverse psychological reactions after combat were believed afflicted with shell shock, thought to result from the concussion effects of artillery bombardments. During World War II, combat fatigue was used to categorize soldiers exhibiting acute adverse psychological reactions after combat because it was thought that exhaustion played a role in this condition. From the 1950s to 1970s, the term war neurosis often was used to broadly characterize soldiers' adverse psychological reactions after combat. For example, the psychiatric literature during this period associated a wide range of postwar psychological problems with combat exposure, including irritability, jumpiness, disturbed sleep, hysteria, disorientation, and panic attacks. However, for years the long-term effects of combat exposure remained elusive, in part because of a poor case definition and in part because reliable and valid observational studies were not available until years after World War II. After reviewing these and other findings related to natural disaster and concentration camp survivors, Dohrenwend concluded that there was clear and compelling evidence that many individuals develop functional psychiatric disorders after extreme stress exposures that had not been present before these experiences, a discovery that had been made earlier by Freud.

Vietnam Veteran Studies

Initial Research

Initial studies of the postwar experiences of Vietnam veterans provided the first clinical and scientific evidence linking combat exposures in Vietnam to post-discharge problems and adjustment difficulties also observed among previous veterans. However, in part because Vietnam veterans often came from lower SES groups, the postwar status of these veterans was controversial. In addition, many early studies were methodologically flawed. One difficulty was that these studies were limited because of biases associated with using nonrepresentative samples. Another problem was the use of nonstandardized mental health measures or the use of measures that assessed only mental health symptoms. Thus, although clinicians had observed that many individuals exposed to combat exhibited certain postwar syndromes afterward (including hyperalertness, exaggerated startle responses, sleep disturbances, and other symptoms) linking these to combat exposure was problematic because of existing methodological shortcomings. However, when DSM-III was being developed in the late 1970s, clinicians and others involved with Vietnam veterans were successful in eventually incorporating these syndromes into DSM-III under the diagnostic nomenclature defining PTSD, originally labeled the post-Vietnam syndrome and later conceptually broadened to include other types of trauma, such as sexual abuse. Once PTSD was included in DSM-III, undertaking large-scale, standardized surveys of Vietnam veterans was feasible and had other research benefits. During this early research period, Robert Laufer's Legacy Study, although not based on a true probability study, stood out as a landmark investigation. This is because it included over 1000 subjects (both Vietnam veterans and civilians) and used several standard symptom scales common in the 1960s and 1970s. Thus, at the time, this study provided the best assessment of the mental health status of Vietnam veterans and served as a benchmark for future studies.

Later Research

As part of Public Law 98-160, the U.S. Congress in 1983 mandated that research on Vietnam veterans be undertaken to determine "the prevalence and incidence of post-traumatic stress disorder and other psychological problems in readjusting to civilian life." In addition to Public Law 98-160, the U.S. Congress passed other public laws mandating studies of the health effects of Vietnam service. As a consequence of these new laws, several well-designed

cohort studies of Vietnam veterans were undertaken during the 1980s. These studies avoided many of the shortcomings of previous research, chiefly because of advancements in sampling and measurement that evolved from earlier research. One of these included the availability of the Diagnostic Interview Schedule (DIS), which for the first time permitted the gathering of DSM-III psychiatric diagnoses by means of population surveys, something previously not possible. This later research confirmed that Vietnam combat veterans had higher rates of postwar adjustment difficulties, mental health disorders, medical morbidity, and postwar mortality than non-combat veterans or comparable nonveterans. Most important, however, these studies indicated that the postwar adjustment difficulties and health problems experienced by these veterans were often due to combat exposures in Vietnam, not to the selection biases or measurement inadequacies that had affected earlier studies. Among the later studies, the National Vietnam Veterans Readjustment Study (NVVRS), which involved over 2000 Vietnam-theater and non-theater veterans (in addition to hundreds of civilian nonveterans), is considered one of the most comprehensive psychosocial assessment to date. The NVVRS indicated that 15% of male Vietnam veterans were current PTSD cases (9% of female veterans) and that 31% of male veterans (27% of female veterans) had PTSD during their lifetimes. Those with PTSD also were more likely to have other stress-related psychiatric disorder, such as depression and anxiety disorders, as well as many other postwar adjustment problems. The NVVRS further indicated that the PTSD-positive veterans often had profoundly disrupted lives in almost every domain of life, including in employment and family relationships. Furthermore, the prevalence of PTSD and other postwar psychological problems were found to be significantly higher among those with greater combat exposure and among Hispanic and African American veterans.

However, as significant as the psychosocial consequences of the war was for Vietnam veterans, their postwar experiences went beyond psychosocial outcomes. For example, the Vietnam Experience Study (VES) indicated that Vietnam veterans had higher rates of postwar mortality in the first 5 years after discharge, primarily due to suicides, homicides, drug overdoses, and motor vehicle accidents. Furthermore, VES findings related to the postwar medical morbidity experienced by these veterans confirmed that Vietnam-theater veterans as a group had higher rates of health-care use and reported themselves to be in poorer health than veterans without Vietnam service. In addition, theater veterans also tended to have

higher rates of hearing loss, evidence of past hepatitis B infection, and higher thyroid-stimulating hormone levels after discharge. In addition, when the postwar health status of Vietnam veterans was examined in relation to whether the veteran had PTSD, PTSD-positive veterans had substantially higher (i.e., 50–150% greater) postwar rates of many major chronic diseases, including circulatory, nervous system, digestive, musculoskeletal, and respiratory diseases, even after controlling for the major risk factors for these conditions. For example, 25% of PTSD-positive veterans reported physician-diagnosed circulatory diseases nearly 20 years after the service (vs. 13% of PTSD-negative veterans) and 19% reported physician-diagnosed nervous system disorders (vs. 6%). Altogether, 68% of PTSD-positive Vietnam veterans reported the occurrence of a chronic disease-related medical condition 20 years after Vietnam service (vs. 48% of PTSD-negative Vietnam veterans). Finally, PTSD-positive veterans also were significantly more likely to show electrocardiographic evidence of myocardial infarction and to have abnormally high white blood cell counts ($>11,000/\text{mm}^3$) and other immune system abnormalities 20 years after military service.

The most compelling evidence, however, linking PTSD to adverse health outcomes was a VES study that examined the causes of death among 15,288 male U.S. Army veterans 30 years after their military service (see [Table 1](#)). The analyses adjusted for race, Army volunteer status, Army entry age, Army discharge status, Army illicit-drug abuse, intelligence, age, and for cancer mortality, also pack-years of cigarette smoking. The findings indicated that adjusted postwar mortality for all, cardiovascular, cancer, and external causes of death (e.g., motor vehicle accidents, accidental poisonings, suicides, homicides, and injuries of undetermined intent) was associated with PTSD among Vietnam-theater veterans ($N = 7924$), with hazards ratios (HRs) of 2.2 ($p < 0.001$), 1.7 ($p = 0.034$), 1.9 ($p = 0.018$), and 2.3 ($p = 0.001$), respectively. For Vietnam-era veterans with no Vietnam service ($N = 7364$), PTSD was associated with all causes of mortality ($\text{HR} = 2.0$; $p = 0.001$). PTSD-positive Vietnam-era veterans also appeared to have an increase in external-cause mortality as well ($\text{HR} = 2.2$; $p = 0.073$). This study essentially suggested that Vietnam veterans with PTSD were at approximately twice the risk of postwar death from multiple causes. Other postwar studies were also conducted, such as the Operation Ranch Hand Study, which sought to assess the postwar health effects of exposure to herbicides used in Vietnam. Others include the Vietnam Era Twin (VET) studies, which focused on the genetic liabilities for PTSD, substance abuse, and other mental health disorders, not just the impact of Vietnam service.

Table 1 Cox proportional hazards regressions: crude and adjusted hazard ratios by veteran and PTSD status^a

Veteran status	All-cause mortality (total deaths = 820)			Cardiovascular mortality (total deaths = 241)			Cancer mortality (total deaths = 188)			External-cause mortality (total deaths = 175)		
	HR	95% CI	P value	HR	95% CI	P value	HR	95% CI	P value	HR	95% CI	P value
All veterans ^b												
PTSD (unadjusted)	2.5	2.1–3.0	<0.001	1.7	1.1–2.6	0.010	1.8	1.1–2.8	0.013	2.7	1.8–4.0	<0.001
PTSD (adjusted) ^c	2.1	1.7–2.6	<0.001	1.6	1.1–2.4	0.027	1.5	1.0–2.4	0.075	2.3	1.5–3.5	0.001
Vietnam-era veterans ^d												
PTSD (unadjusted)	2.6	1.7–3.8	<0.001	1.3	0.5–3.6	0.57	1.1	0.4–3.6	0.84	2.9	1.3–6.7	0.012
PTSD (adjusted) ^c	2.0	1.3–3.0	0.001	1.2	0.4–3.4	0.69	0.9	0.3–3.1	0.92	2.2	0.9–5.2	0.073
Vietnam-theater veterans ^e												
PTSD (unadjusted)	2.5	2.0–3.2	<0.001	1.8	1.1–2.8	0.015	2.2	1.3–3.7	0.003	2.6	1.6–4.1	<0.001
PTSD (adjusted) ^c	2.2	1.7–2.7	<0.001	1.7	1.0–2.7	0.034	1.9	1.1–3.3	0.018	2.3	1.4–3.9	0.001

^aHR = hazards ratio; CI = confidence interval; PTSD, = posttraumatic stress disorder. Adapted from J. A. Boscarino (2006). Posttraumatic stress disorder and mortality among US Army veterans 30 years after military service. *Annals of Epidemiology* **16**, 248–256. Reproduced with permission by Elsevier Publishing.

^bN = 15,288; person-risk years = 229,565; total PTSD cases = 1050.

^cAll models adjusted for race, Army volunteer status, Army entry age, Army discharge status, Army illicit drug use, age at interview, and intelligence. For cancer mortality, models adjusted for the above variables, in addition to pack-years of cigarette smoking.

^dN = 7364; person-risk years = 110,553; total PTSD cases = 214.

^eN = 7924; person-risk years = 119,453; total PTSD cases = 836.

To date, however, the VES has provided the most compelling evidence related to the adverse health effect of combat service and PTSD during that conflict.

Psychobiological and Biomedical Findings

There are clinical reasons to expect alterations in neuroendocrine system functions in chronic PTSD cases because long-term changes in the HPA system and the sympathetic arm of the autonomic nervous system have been reported following exposure to extreme stress. Evidence suggests that the physiological arousal often observed during the recollection of traumatic events is associated alterations in the neuroendocrine functions linked to the sympathetic-adrenomedullary and HPA stress axes. In the case of PTSD, it is thought that these neuroendocrine alterations reflect the consequences of an extreme state of physiopsychological conditioning that occurs following severe stress exposures. Furthermore, although this conditioning response is initiated in the central nervous system (CNS), it is subsequently carried out by multiple endocrine mechanisms that have wide-ranging effects on the body and nervous system. Similar findings related to this physiological-emotional conditioning process have also been observed with animal models.

Several studies have reported lower cortisol levels among Vietnam veterans with PTSD, as well as increased catecholamine concentrations. Furthermore, one recent large-scale study, based on the VES and involving thousands of veterans, found not only that Vietnam-theater veterans with current PTSD had lower cortisol but that past combat exposure levels were associated with plasma cortisol concentrations in an inverse dose-response relationship (see Figure 1). In brief, research with Vietnam combat veterans suggests that PTSD-positive veterans tend to have lower cortisol levels but, paradoxically, higher catecholamine concentrations, together with heightened responses of the stress system to traumatic memories and other stimuli associated with the original trauma. It is known that chronicity and excessiveness of stress system activation can lead to pathogenesis, causing weight changes, depression, hypogonadism, and other systemic alterations, including increased metabolic load on the body's systems. Thus, the physiological findings for PTSD-positive Vietnam veterans suggest that past traumatic stress exposure has resulted in neuroendocrine system alterations that make veterans potentially susceptible to a host of chronic diseases. As noted, PTSD-positive Vietnam veterans were found to have higher postwar rates

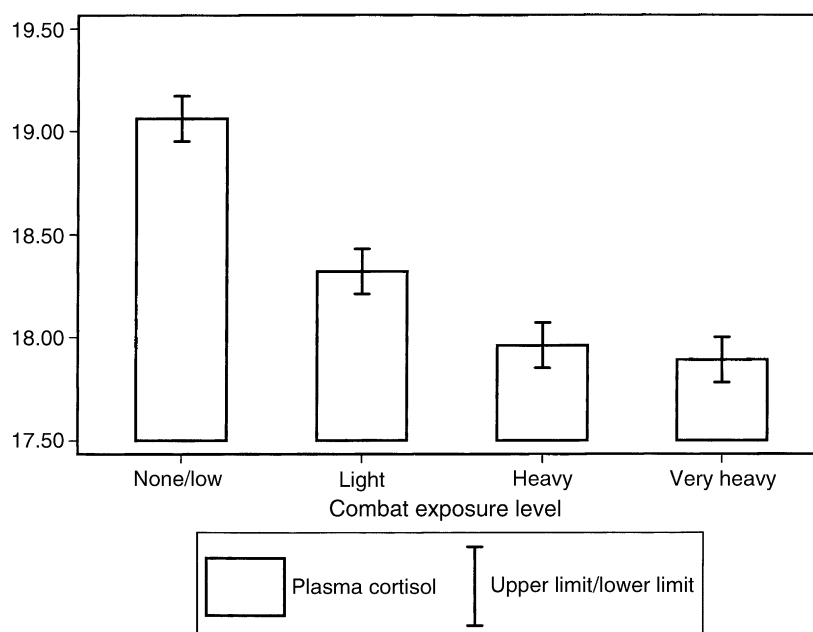


Figure 1 Adjusted plasma cortisol concentrations and standard errors bars (95% confidence limits) for Vietnam-theater veterans by combat exposure levels ($N = 2457$). Results shown are for morning, fasted plasma cortisol (in micrograms per deciliter). Combat level was measured by the Combat Exposure Index scale, a standard measure of combat exposure. Cortisol results are adjusted for demographic variables, alcohol use, drug use, smoking status, and body mass index using analysis of covariance. A trend test for cortisol by combat exposure level was statistically significant ($p = 0.0001$), confirming the inverse dose-response association between cortisol and combat exposure. Adapted from J. A. Boscarino (1996), Post-traumatic stress disorder, exposure to combat, and lower plasma cortisol among Vietnam veterans: findings and clinical implications, *Journal of Consulting and Clinical Psychology* **64**, 191–201. Reprinted with permission from the American Psychological Association.

of many chronic diseases, reinforcing this neuroendocrine-dysregulation disease hypothesis, which represents a cornerstone in human psychoneuroendocrinology research.

Related to HPA dysregulation, it has been suggested that PTSD-positive veterans may be at risk for autoimmune diseases as well. For example, although there have been inconsistencies, investigators, as previously noted, have found that individuals who developed PTSD, particularly men exposed to combat, appeared to have lower plasma cortisol concurrent with higher catecholamine levels. As suggested, Vietnam veterans with current PTSD not only had lower cortisol, but this had an inverse dose-response relationship with combat exposure. In addition, also as previously noted, research indicated that Vietnam veterans with current PTSD had clinically elevated leukocyte and T-cell counts, and similar findings have been reported in nonveteran studies as well. These veterans were also found to have other significant neuroendocrine abnormalities, such as elevated thyroid hormones. Consistent with these clinical findings, it has recently been reported that PTSD was in fact associated with several autoimmune diseases among Vietnam veterans, including rheumatoid arthritis and psoriasis. Again, analogous results have been reported in animal models as well. Of course, not all of the variance in disease outcomes should be associated directly with PTSD. Research with the VET Registry suggested that about one-third of the liability for PTSD was associated with genetic factors. Some of the same genes that make one liable for PTSD could also make one liable for specific diseases as well.

Studies of Other Veteran Populations after Vietnam

Following the first Persian Gulf War (PGW) in 1991, cohort studies were again used to evaluate the health impact of this conflict on war-theater military personnel. These studies are still in progress, and a number of them evolved from or were related to the original NVVRS and VES research designs. The PGW research to date has established a link between service in the PGW and mental illness, postservice injury, chronic fatigue syndrome, and other multisymptom conditions. The reasons for these associations are still being investigated, and the results from PGW research are being used to inform a new generation of veteran studies, related to personnel deployed in the Afghanistan and Iraq theaters of war following the September 11, 2001, attacks. For these recent conflicts, however, military personnel were being assessed both before deployment and after return, something

rarely done in the past but often desired after service exposure had occurred. Finally, following the September 11 attacks in New York City (NYC) in 2001, a number of the investigative teams, funded by the National Institutes of Health (NIH) and the Centers for Disease Control (CDC) to assess the population impact of this event, used the prospective cohort method as well as measurement instruments that had been developed originally for Vietnam veteran population studies. Thus, the field of disaster epidemiology has emerged as a subspecialty within epidemiology, informed, at least in part, by the original work with Vietnam veterans over the previous decades.

Implications for Medical Sciences, Research, and Treatment

The postwar findings reported for Vietnam veterans documented that veterans exposed to combat have more stress-related psychosocial disorders, altered neuroendocrine functions, higher medical morbidity, and greater mortality years after the war. The scientific findings related to the postwar experience of the Vietnam veteran suggest that psychobiological models are probably required to understand the impact of this experience on the lives of those who served there. For example, it has been suggested that traumatic stressors act as unconditioned aversive stimuli that evoke severe physiological distress. Consequently, previously neutral external stimuli (e.g., the sound of a helicopter) and internal stimuli (e.g., physiological states) that accompanied the traumatic stressor, can function as conditioned stimuli capable of producing psychological and physiological distress when the traumatic stressors are no longer present. Experience with psychotherapy treatments suggests that when stress victims are gradually exposed to aversively conditioned stimuli the symptoms may decrease because desensitization to the stressor could occur. A psychotherapeutic technique often used in PTSD therapy with veterans is to teach the patient to verbalize the traumatic experience with the help of a support group consisting of other veterans. This process allows veteran to redefine the external and internal stimuli that affect psychological arousal, develop better self-control, and, with assistance, desensitize him- or herself to a range of thoughts, actions, and situations associated with the traumatic events. A more structured and currently widely used approach that appears to be effective is cognitive-behavioral therapy. Thus, psychosocial processes appear to be instrumental in treating a condition that is psychosocial in origin but that also has a physiobiological basis – that is, research with Vietnam veterans and

others tells us that PTSD has both a psychosocial and a physiobiological foundation.

The recognition of a multifactorial PTSD model is consistent with the report that both pharmacotherapy and cognitive-behavioral psychotherapy are effective in treating PTSD. In the case of pharmacotherapy, the pathophysiology of PTSD, in part, appears to involve the serotonergic and the noradrenergic systems; thus, drugs known to potentiate these mechanisms have been effective. In the case of cognitive-behavioral therapy, this approach has been found to be effective in reducing PTSD-related symptomatology by achieving desensitization to stressful stimuli, by increasing control of adverse arousals, by enhancing anxiety management, and by using other known behavioral-psychological mechanisms. Although the underlying interventional mechanisms differ for pharmacological versus cognitive-behavioral therapy (e.g., pharmacokinetic vs. psychological), the outcomes are similar – the psychopathology and underlying pathophysiology are reduced and fewer adverse patient symptoms are manifested, hence lowering the risk of other adverse outcomes. Thus, we expect that as trauma-related symptoms are reduced through treatment, the risk for other adverse outcomes will also decrease. Recent research has, in fact, confirmed this hypothesis in NYC following the September 11 attacks. For example, it has been reported that NYC adults who received emergency crisis counseling at work shortly after the attacks not only had better mental health outcomes but also had better outcomes in terms of binge drinking, alcohol dependence, and alcohol consumption, lowering the possibility of future adverse health outcomes.

Although the findings discussed here chiefly involve veterans, exposure to combat, they are applicable to other populations. Traumatic stress reactions occur among individuals exposed to extreme, life-threatening situations, including natural disasters, physical attacks, wartime combat, and other situations. In addition, it is likely that human exposure and response to stress take place along a stress continuum, with traumatic stress on one end of this continuum. Furthermore, genetic liability is also likely involved with respect to whom may succumb to PTSD. Finally, the postwar experiences of the Vietnam veteran suggest that more than psychosocial explanations are needed to understand PTSD and to treat it effectively. In addition to higher-order cognitive models, findings in Vietnam veteran suggest lower-order biological models are also necessary. Research also consistently suggests that those with less social support exhibit greater levels of traumatic stress, which may explain the high rates of PTSD among Vietnam veterans, given the low community support during and after the war. This finding has clinical implications for the future treatment of both combat and noncombat-related stress victims.

Conclusion

Psychoneuroendocrinology represents a growing area of medical research that focuses on the association between neuroendocrine function and psychological status, especially as this relates to steroidogenic biological alterations and psychiatric illness. Just as with other areas of medicine – orthopedics, plastic surgery, emergency medicine, and so forth – developments in

Table 2 Key biological correlates of PTSD among a national random sample of Vietnam veterans^a

Biometric measure	PTSD negative (%)	PTSD positive (%)	P value
White blood cell count elevated	2.8	4.6	0.063
Total lymphocytes elevated	4.2	6.8	0.023
T lymphocytes elevated	2.9	6.2	0.001
B lymphocytes elevated	9.9	13.7	0.032
CD4 cells elevated	3.5	5.6	0.055
CD8 cells elevated	2.4	5.9	<0.001
Erythrocyte sedimentation rate elevated	6.1	9.6	0.013
Triiodothyronine (T ₃) uptake elevated	7.9	10.8	0.06
Thyroxine (T ₄) elevated	1.7	5.0	<0.001
Immunoglobulin-A elevated	16.4	23.5	0.001
Immunoglobulin-M elevated	1.2	3.4	0.001
Dehydroepiandrosterone reduced	2.3	4.6	0.008
Cortisol/dehydroepiandrosterone ratio elevated	4.4	7.1	0.023
N =	4139	323	–

^aData are from the Vietnam Experience Study and represent a national random sample of male U.S. Army veterans of the Vietnam War era, half of which served in Vietnam and half of which served elsewhere. PTSD cases represent both combat and noncombat cases. Assessment was done approximately 20 years after military service. PTSD, posttraumatic stress disorder.

military medicine and veterans' health have been translated into medical advances for other populations. For example, today, research related to the biological and medical consequences of anxiety disorders is more commonplace, as are epidemiological studies of other veteran populations and nonveteran studies involving exposure to human-made and natural disasters.

The medical promise of this body of research can be illustrated using the findings shown in [Table 2](#). As shown, PTSD-positive veterans have a distinct clinical profile, suggesting an increased inflammatory response related to a broad range of potential pathophysiology that could be associated with a host of diseases, including cardiovascular and autoimmune diseases. The reasons for the findings shown in [Table 2](#) are not clear at this time; they may be related to biological, psychological, or behavioral factors (or a combination of these) associated with PTSD and research also suggests that genetic factors are indeed involved. What is clear is that this area of research holds promise for understanding the psychosomatic basis for the causes and prevention of many human diseases. In conclusion, despite of the suffering caused by Vietnam War, or perhaps because of it, the postwar experiences of the Vietnam veteran have resulted in significant advances in our knowledge about the nature and consequences of human stress exposures and will probably result in future medical advances that cannot currently be imagined.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Allostasis and Allostatic Load; Anxiety; Autoimmunity; Catecholamines; Combat Reaction, Chronic; Combat Stress Reaction; Combat, Acute Reactions to; Corticosteroids and Stress; Depression Models; Disease, Stress Induced; Gulf War Syndrome, Psychological and Chemical Stressors; Hypothalamic-Pituitary-Adrenal; Korean Conflict, Stress Effects of; Major Depressive Disorder; Nuclear Warfare, Threat of; Persian Gulf War, Stress Effects of; Posttraumatic Stress Disorder in Children; Posttraumatic Stress Disorder, Delayed; Posttraumatic Stress Disorder, Neurobiology of; Psychotherapy; Salivary Cortisol; War Stress in the Former Yugoslavia; War-Related Posttraumatic Stress Disorder, Treatment of; Depression, Immunological Aspects; Posttraumatic Stress Disorder – Clinical; Posttraumatic Stress Disorder – Neurobiological basis for.

Further Reading

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd edn.). Washington, DC: American Psychiatric Association.

- Archibald, H. C., Long, D. M., Miller, C., et al. (1962). Gross stress reaction in combat – a 15-year follow-up. *American Journal of Psychiatry* **119**, 317–322.
- Boscarino, J. A. (1995). Post-traumatic stress and associated disorders among Vietnam veterans: the significance of combat and social support. *Journal of Traumatic Stress* **8**, 317–336.
- Boscarino, J. A. (1996). Post-traumatic stress disorder, exposure to combat, and lower plasma cortisol among Vietnam veterans: findings and clinical implications. *Journal of Consulting and Clinical Psychology* **64**, 191–201.
- Boscarino, J. A. (1997). Diseases among men 20 years after exposure to severe stress: implications for clinical research and medical care. *Psychosomatic Medicine* **59**, 605–614.
- Boscarino, J. A. (2004). Association between posttraumatic stress disorder and physical illness: results and implications from clinical and epidemiologic studies. *Annals of the New York Academy of Sciences* **1032**, 141–153.
- Boscarino, J. A. (2006). Posttraumatic stress disorder and mortality among US Army veterans 30 years after military service. *Annals of Epidemiology* **16**, 248–256.
- Boscarino, J. A., Adams, R. E., Foa, E. B., et al. (2006). A propensity score analysis of brief worksite crisis interventions after the World Trade Center disaster: implications for intervention and research. *Medical Care* **44**, 454–462.
- Centers for Disease Control (1987). Postservice mortality among Vietnam veterans. *Journal of the American Medical Association* **257**, 790–795.
- Centers for Disease Control (1988). Health status of Vietnam veterans. I: Psychosocial characteristics. *Journal of the American Medical Association* **259**, 2701–2707.
- Centers for Disease Control (1988). Health status of Vietnam veterans. II: Physical health. *Journal of the American Medical Association* **259**, 2708–2714.
- Chrousos, G. P. and Gold, P. W. (1992). The concepts of stress and stress system disorders: overview of physical and behavioral homeostasis. *Journal of the American Medical Association* **267**, 1244–1252.
- Dohrenwend, B. P. (1975). Sociocultural and social-psychological factors in the genesis of mental disorders. *Journal of Health and Social Behavior* **16**, 365–392.
- Figley, C. R. (1978). *Stress disorders among Vietnam veterans: theory, research and treatment*. New York: Brunner/Mazel.
- Freud, S. (1953–1974). Introduction to psychoanalysis, war neuroses (1919). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud*, (vol. 17), pp. 205–215. London: Hogarth Press and the Institute of Psychoanalysis.
- Frey-Wouters, E. and Laufer, R. S. (1986). *Legacy of a war: The American soldier in Vietnam*. Armonk, NY: Sharpe.
- Galea, S., Ahern, J., Resnick, H., et al. (2002). Psychological sequelae of the September 11 terrorist attacks in New York City. *New England Journal of Medicine* **346**, 982–987.

- Hanson, F. R. (1949). The factor of fatigue in the neuroses of combat. In: Hanson, F. R. (ed.) *Special issue on combat psychiatry*. *Army Medical Bulletin* 9, 147–150.
- Helzer, J. E., Robins, L. N. and McEvoy, L. (1987). Post-traumatic stress disorder in the general population: findings of the Epidemiologic Catchment Area Study. *New England Journal of Medicine* 317, 1630–1634.
- Hoge, C. W., Auchterlonie, J. L. and Milliken, C. S. (2006). Mental health problem, use of mental health services, and attrition from military service after returning from deployment to Iraq or Afghanistan. *Journal of the American Medical Association* 295, 1023–1032.
- Hyams, K. C., Wignall, F. S. and Roswell, R. (1996). War syndromes and their evaluation: from the U.S. Civil War to the Persian Gulf War. *Annals of Internal Medicine* 125, 398–405.
- Institute of Medicine (2001). *Veterans and agent orange, update 2000*. Washington, DC: National Academy Press.
- Keane, T. M., Zimering, R. T. and Caddell, J. M. (1985). A behavioral formulation of posttraumatic stress disorder in Vietnam veterans. *Behavior Therapist* 8, 9–12.
- Kulka, R. A., Schlenger, W. E., Fairbank, J. A., et al. (1990). *Trauma and the Vietnam War generation: report of findings from the National Vietnam Veterans Readjustment Study*. New York: Brunner/Mazel.
- Mason, J. M., Giller, E. L., Kosten, T. R., et al. (1988). Elevation of urinary norepinephrine/cortisol ratio in posttraumatic stress disorder. *Journal of Nervous and Mental Diseases* 176, 498–502.
- McEwen, B. S. (2000). Allostasis and allostatic load: implication for neuropsychopharmacology. *Neuropsychopharmacology* 22, 108–124.
- Polner, M. (1971). *No victory parades: the return of the Vietnam veteran*. New York: Holt, Rinehart, Winston.
- Robins, L. N., Helzer, J. E., Croughan, J., et al. (1981). National Institute of Mental Health Diagnostic Interview Schedule: its history, characteristics, and validity. *Archives of General Psychiatry* 38, 381–389.
- Salmon, T. W. (1919). War neuroses and their lesson. *New York Medical Journal* 109, 993–994.
- True, W. R., Rice, J., Eisen, S. A., et al. (1993). A twin study for genetic and environmental contributions to liability for posttraumatic stress symptoms. *Archives of General Psychiatry* 50, 257–264.
- van der Kolk, B. A. and Saporta, J. (1993). Biological response to psychic trauma. In: Wilson, J. P. & Raphael, B. (eds.) *International Handbook of traumatic stress syndromes*, pp. 25–33. New York: Plenum Press.
- Wolkowitz, O. M. and Rothschild, A. J. (eds.) (2003). *Psychoneuroendocrinology*. Washington, DC: American Psychiatric Publishing.
- Yehuda, R. (2002). Current status of cortisol findings in post-traumatic stress disorder. *Psychiatric Clinics of North America* 25, 341–368.

Vietnam War, Stress Effects of

See: Captivity, Adaptation to; Captivity, Recovery from; Combat Reaction, Chronic; Combat, Acute Reactions to; Coping and Stress: A Lens and Filter Model; Vietnam Veterans, Postwar Experiences and Health Outcomes; Defense Services, Stress in.

Violence

E K Englander

Bridgewater State College, Bridgewater, MA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by E K Englander, volume 3, pp 662–668, © 2000, Elsevier Inc.

Types of Biological Influences

Summary

Glossary

Adoption studies

Research wherein a child with one set of biological parents and a different set of psychological parents is studied.

Biological environmental influences

Events that affect a person biologically but are not encoded into the person's DNA.

Concordance rates

The frequency with which one twin is diagnosed with an illness, when his or her twin is also diagnosed with the same illness.

Cortisol

	The hormone that regulates the body's reaction to stress.
<i>Genetic influences</i>	The biological blueprints for behavior that are contained in a person's chromosomes.
<i>Imitation</i>	A theory that states that children may learn to be violent by imitating an adult's aggression.
<i>Minimal brain dysfunction or minimal brain damage</i>	Very low levels of brain dysfunction, theoretically affecting behavior, emotions, learning, or memory.
<i>Neurotransmitters</i>	Chemicals secreted by the brain, some of which have been linked to aggression.
<i>Paranoid misperceptions</i>	Attitudes or belief systems that misinterpret ambiguous circumstances as signs of danger when none actually exists.
<i>Resilient state</i>	The state in which a person is likely to resist developing a disorder.
<i>Serotonin</i>	One of the monoamines, low levels of which have been implicated as related to aggression.
<i>Testosterone</i>	One of the androgens. A male sex hormone that is associated with aggression.
<i>Vulnerable state</i>	The state in which a person is likely to develop a disorder.

Violent crimes can be as baffling as they are repelling. In England, two 10-year-old boys lured a 2-year-old away from his mother at a shopping mall; they took him down to the local railroad tracks, where they beat him to death and left his body to be run over by a train. In the United States, a group of teenagers shot and killed a man whose car they were stealing – even though the man had willingly handed over his car keys and was not resisting the theft in any way. In Massachusetts, a group of young teenage schoolgirls planned to murder their English teacher because she was too strict.

To understand the causes of violence, it is not always enough to understand the external circumstances that often drive other types of criminal behavior. It is true that violence sometimes results from factors such as an individual's desire for financial gain or an individual's repeated exposure to violent behavior in his or her social environments. Sometimes, however, violence also happens in an apparently motiveless fashion. Violence is never truly without motive, but the motives may be so complex and elusive that the violence appears motiveless. In all cases, but particularly in cases of violence that appear to have no motive, internal or individual factors may be critical for understanding the cause of such behavior. A variety of different biological and psychological influences and mechanisms have been considered over the years. This article summarizes

them and attempts to construct a more comprehensive model of the causes of violent behavior.

Types of Biological Influences

There exists a mistaken tendency to use the terms biological and genetic interchangeably. In fact, genetic influences are only one type of biological influences on behavior. There are (at least) two different types of biological influences: genetic influences and biological environmental influences. Genetic influences refer to the blueprints for behavior that are contained in a person's chromosomes. Chromosomes contain deoxyribonucleic acid (DNA), the genetic material a person inherits from his or her biological parents, which is referred to as their genotype. Biological environmental influences, unlike genetic influences, are events that affect a person biologically but that are not encoded into the person's DNA. For example, consider a head injury from a car accident, which subsequently changes the victim's personality. The head injury is unmistakably a biological environmental influence. No genotype determines that someone will have a car accident; however, such an accident, and its accompanying head injury, can still have an important biological effect on the person's behavior.

Although learning is clearly related to violent behavior, learning theories alone cannot fully explain violence in human beings. As an example, consider the principle of imitation. It states that children may learn to be violent by imitating an adult's aggression. According to this principle, one would expect that children who grow up watching violent parents would be violent themselves. As the theory predicts, a higher proportion of children who witness violent parents are violent, in comparison to other children; however, despite their higher risk, many children of violent parents do not, in fact, behave aggressively as adults. Some children who are exposed to violent psychosocial environments do become violent; many other children who are similarly exposed do not. This fact has led researchers to coin the terms resilient and invincible to describe children who survive and cope well despite terrible circumstances.

What makes one child resilient and another child vulnerable? Researchers have identified a variety of differences, including biological differences, that distinguish resilient and vulnerable children. It is important to note that such biological differences do not entirely explain why one child is vulnerable and another is resilient; rather, biology appears to be an important part of the determination. It is reasonable to assume that different biological influences will play

different roles in the determination of vulnerability and resiliency.

One important area that researchers have studied focuses on minimal brain damage. Also called minimal brain dysfunction, minimal brain damage does not look like major brain damage. A person with minimal brain dysfunction will not have obvious brain damage; however, minimal dysfunction may affect behavior, emotions, learning, or memory – that is, the more subtle (or higher) functions of the brain. For example, although major brain damage might result in severe mental retardation, minor brain dysfunction might result in a learning disability that would not necessarily affect intelligence in any way (e.g., dyslexia). It has been repeatedly demonstrated that minimal brain dysfunction is related to problems such as learning disorders and hyperactivity. Those relationships emerged as a rationale for studying a link between minimal brain dysfunction and violence, but how to study such a link is somewhat controversial in the field. Some researchers have examined violent individuals to see if they have high levels of minor physical anomalies, while others have given aggressive and violent men neuropsychological examinations. Still other psychologists have combed through the medical histories of violent individuals, seeking unusually high levels of brain injury in their histories. Although different researchers have taken different tacks, their findings have been consistent. However one chooses to measure minimal brain dysfunction, a host of research suggests that it is indeed related to persistent aggression and violence. Further, although both major and minor brain damage may cause violence, the weight of the evidence suggests that the most important type of brain dysfunction in people who are chronically violent and criminal is minimal brain dysfunction. Although it is almost certain that most people with minimal brain dysfunction will never be violent, a noticeably high number of repeat violent offenders show some level of brain dysfunction (usually minimal).

Neurotransmitters

There are many different types of neurotransmitters, and they tend to specialize in function. How much, or how little, of any particular neurotransmitter chemical is in your brain is an important determinant of many types of behavior. Both biology and psychosocial environment affect the levels of different neurotransmitters, and thus affect behavior. A logical question, therefore, is whether violent behavior is related (at least in part) to the level of certain neurotransmitters in the brains of violent people.

Monoamine neurotransmitters have been most significantly linked to aggression. Of the monoamines,

low levels of serotonin have been implicated most strongly in recent research. Several studies have found consistent relationships between reduced serotonin activity and aggression and hostility in both boys and men. As is often the case in such research, the relationship was strongest among males, regardless of their age. This set of studies is interesting because it establishes such relationships among psychiatrically well patients as well as among behaviorally disordered individuals.

Hormones

Two hormones have been the focus of most research on aggression and criminality: testosterone and cortisol. Testosterone is one of the male sex hormones, called androgens. Both males and females secrete androgens, but males secrete a much greater quantity than females. Androgens are frequently cited as an important cause of aggressive behavior, particularly intermale aggression. This area of study is not completely consistent, however, and while some research points to aggression and dominance/hostility as being related to testosterone, other research fails to find that dominance/hostility measures are related to sex hormones. Some research notes a testosterone–aggression relationship in both males and females; other research finds it only for males. To muddy the waters even further, researchers who follow those experiments involving reduced testosterone functioning do not note that this is a clearly effective means of stopping aggression. In addition, psychosocial researchers suggest that high testosterone might result in kids being treated differently socially, since high testosterone is associated with being bigger and stronger. Androgens, therefore, have been implicated, but not strongly, in the cause of aggression.

A second hormone that has been implicated more recently is cortisol. Cortisol is the hormone that regulates the body's reaction to stress. It is involved with the immune system and with sex hormones as well. A few studies have linked low levels of cortisol with a tendency to be aggressive. Although this should be characterized as an exploratory area of research, it is interesting to note that low cortisol has even been found in females with conduct disorders.

Genetics

The majority of researchers agree that violence does not appear to be a behavioral tendency that is transmitted via a simple, directly acting gene. In any case, it is clear that if genetic bases for aggression and violence exist, they are clearly mutable and changeable by a person's psychological environment; the old idea that genetics determines one's behavior absolutely is clearly mistaken.

The case for genetics as a cause of aggression and adult criminal violence has to be pieced together from a number of different areas of study. First, we know that children with violent parents do have a higher tendency to be violent themselves. Second, data from twin studies show the same trend.

Twin studies are natural genetic laboratories; they are studies that typically compare identical (monozygotic) to fraternal (dizygotic) twins. Identical twins are the only human beings alive who are genetically identical. Some psychological disorders have much higher concordance rates for monozygotic twins. For example, schizophrenia has a concordance rate of 50%; if one identical twin has schizophrenia, then there is a 50% chance that the other twin will also have schizophrenia. Similar, and even higher, rates of concordance for aggressive behavior have been noted among twin pairs.

Despite these family studies finding a reasonably strong genetic link for aggression, some clarification is important. Specifically, the field needs to make the critical distinction between biological heredity versus family psychology. Parents provide a psychological, as well as a biological, environment for their children. Likewise, twin studies fail to tease apart the effects of psychosocial environment and genetics. In addition to sharing genes, twins also often share an environment that treats them similarly.

For this reason, adoption studies (in which a child has one set of biological parents and a different set of psychological parents) may do a better job of teasing apart heredity and psychological environment. Earlier classic studies, which examined children prospectively in long-term studies, found that children seemed to follow their adoptive parents' behavior, rather than their genetic parents (suggesting a stronger link with psychological environment than genetics). More recent research has also found that biological siblings had a higher concordance for aggression, in comparison to adoptive siblings.

Both family studies and adoption studies suggest a definite degree of heritability for aggression, but there is a third body of research that is relevant here: studies that examine the heritability of childhood precursors of violence in adulthood. This body of research examines childhood disorders that are highly related to violence and studies whether these childhood disorders are heritable. For example, attention-deficit/hyperactivity disorder and conduct disorders are both childhood disorders with strong ties to adult violence, and both disorders appear to be at least partially heritable. The heritability of a disorder such as violence may therefore proceed through the pathways of other, co-existing disorders, or, alternatively, some childhood disorders may be precursors to adult

violence in a child who has inherited that behavioral tendency.

In summary, much of the data on the genetic basis of violence is suggestive of a significant, but not total or direct, heritability.

Hostile Interpretation of Ambiguous Events, or Cognitive Distortions

In 1996, a 17-year-old boy attacked a total stranger, unprovoked, with an ax. He knocked on the man's door, and when it was opened he attempted to attack before he was driven off by his intended victim. His defense in court? That he was mentally deranged and unusually paranoid due to late-stage Lyme disease. Before he became ill, this boy apparently had no behavioral problems, but his Lyme disease caused him to develop the unusual, but not unheard of, symptom of paranoid delusional misperceptions.

Cases this like are certainly rare. But are paranoid misperceptions rare among violent individuals? One theory suggests that such biased misperceptions may in fact help explain why some people are violent. Indeed, as the theory goes, under certain circumstances, aggression may be a normal and adaptive response. This may particularly be the case when an individual is faced with hostile or threatening forces, and violence is perceived as a necessary response. Given that a hostile environment might provoke an aggressive response in almost anyone, some researchers have investigated the possibility that violent individuals are not people who inappropriately choose violence; rather, they are people who inappropriately perceive hostility in situations where most people would not. That is, perhaps it is not the response that is skewed, but the perception of the environment.

How do aggressive children perceive the environment, and are their perceptions uniform? Studies find that aggressive children have two perceptual tendencies that nonaggressive children lack. Aggressive children are more likely to believe that other people had hostile intentions, and they are more likely to evaluate the results of aggression positively. Classic research has identified this thought pattern and has found that hostile biases were particularly characteristic of reactive-aggressive children. Researchers have similarly found that aggressive boys are significantly better at recognizing aggressively slanted comments. Studies have noted that children with memories that emphasize hostility process information in a more negatively biased way and are more likely to develop stable negative cognitive biases and stably aggressive behavior.

This line of study sheds some light on why it may be so difficult to treat violent offenders. Trying to teach such offenders to be nonviolent may in fact be trying

to teach them to react nonviolently to what they perceive to be intensely threatening situations. For example, imagine that you are standing in a field holding a gun, and an elephant is charging toward you in a rage at full speed with his tusks aimed right at you. Trying to teach violent offenders to be nonviolent might be like trying to teach them not to shoot that elephant, even though there is every indication that that elephant is deadly.

Other research has examined different types of violent offenders and has found cognitive misinterpretations and biases that appear to be specific to the type of offense. For example, rather than making general, nonspecific hostile misinterpretations, adults who commit child abuse tend to misinterpret children's behavior. Specifically, they tend to regard children's misbehaviors as more intentional than they really are – possibly paving the way to righteous anger and abuse. Sexually aggressive men, rather than suffering from blanket distortions, may differ from other men primarily in their distortion of women's communications. It seems clear that a wide variety of cognitive misperceptions seems to pave the way for a variety of different types of aggressive and violent behavior, and may help explain the difficulty of effectively treating such behavior.

Summary

Strong evidence exists to support the theory that individual differences can affect a person's risk for developing chronically violent behavior. These risks appear to interact with psychosocially learned factors and

may differ significantly from individual to individual. The mechanism by which these associated risk factors contribute to violence is less well established. Research has identified heritability, hormones, minimal brain dysfunction, and biased cognitive misperceptions as likely areas of dysfunction related to violence and aggression. Future research will focus on effective prevention efforts and a better understanding of how risk factors influence individuals' behaviors.

Acknowledgments

This article was adapted, with permission, from *Understanding Violence* (Lawrence Erlbaum & Associates, Inc., 2003).

See Also the Following Articles

Aggression; Aggressive Behavior; Cortisol Awakening Response.

Further Reading

- Crick, N. and Dodge, K. (1996). Social information-processing mechanisms in reactive and proactive aggression. *Child Development* 67(3), 993–1003.
- Englander, E. K. (2003). *Understanding violence* (2nd edn.). Mahwah, N.J.: Lawrence Erlbaum Publishers, Inc.
- Mednick, S. A., Gabrielli, W. F. and Hutchings, B. (1984). Genetic influences in criminal convictions: evidence from an adoption cohort. *Science* 234, 891–894.
- Werner, E. E. and Smith, R. S. (1982). *Vulnerable but invincible: a study of resilient children*. New York: McGraw-Hill.

Viral Virulence and Stress

A J L Macario and E Conway de Macario

New York State Department of Health and The University at Albany (SUNY), Albany, New York, USA

Published by Elsevier Inc.

Scope

Viral Infection and ER Stress

Chain of Events and Its Consequences for the Virus and for the Infected Cell

The Endoplasmic Reticulum

ER Stress

ER Stress and Apoptosis

ER Stress and Viral Infection: Examples
Stress and Viral Reactivation

Glossary

Apoptosis

Programmed cell death that occurs in an orderly fashion, following a preexisting pathway and in response to specific inducers.

Caspases

A family of enzymes essential for apoptosis; these enzymes carry a cysteine that is key for cleaving proteins after an aspartic acid residue.

<i>Endoplasmic reticulum (ER)</i>	An organelle of the eukaryotic cell formed by a single membrane, which is an extension of the inner membrane of the nuclear envelope. The ER membrane is folded onto itself around a space, the lumen, which is in direct communication with the intranuclear space.
<i>ER stress</i>	Cellular stress at the level of the endoplasmic reticulum; may be caused by viruses.
<i>Golgi apparatus</i>	An organelle of the eukaryotic cell present in plants and animals (although not in the cells of most fungi). Its chief functions are protein processing and sorting; also called the Golgi body or Golgi complex.
<i>Mitochondrion</i>	A eukaryotic cell organelle present in most organisms. The mitochondrion's primary role is the conversion of fuel, in the form of organic materials, into energy, in the form of ATP. Mitochondria interact with the ER, particularly in situations leading to apoptosis.
<i>Molecular chaperone</i>	A large group of special proteins that play an essential role in protein homeostasis. They are in charge of maintaining the entire complement of proteins in all cells and cellular compartments in their native functional conformations. The chaperones residing in the ER are particularly relevant to viral infection and its consequences for the cell through the induction of the unfolded protein response.
<i>Necrosis</i>	Uncontrolled cell death; the consequence of acute, unanticipated cell or tissue injury.
<i>Programmed cell death</i>	A controlled mechanism of cell elimination by suicide that multicellular organisms possess.
<i>Signal transduction</i>	The cellular process of converting one kind of signal or stimulus into another. This process is usually very fast and involves a series of intracellular biochemical reactions in a specific sequence, signaling cascade, requiring specific enzymes and second messengers.
<i>Unfolded protein response (UPR)</i>	A series of events that starts with the accumulation of unfolded, partially folded, misfolded, or aggregated proteins in the ER lumen, which triggers ER stress, inhibiting the functions of the organelle. This emergency situation then requires immediate action, which is mediated by the signal transduction cascade and results in the upregulation of some genes and in the downregulation of other genes.
<i>Virulence</i>	The capacity of an infectious agent (e.g., a bacterium or a virus) to invade a host and cause lesions and disease.

Virus A biological particle that contains nucleic acid and other components, such as proteins and lipids, assembled in a shape that is typical for each viral species. A chief characteristic of viruses is that they require a host cell in which to replicate.

Many viruses cause cellular stress at the level of the ER (ER stress) and elicit the unfolded protein response (UPR). The main causative event is the accumulation of unfolded and/or misfolded proteins inside the ER; this accumulation immediately activates signal transduction pathways directed toward the nucleus, causing the upregulation of some genes and downregulation of others. This increases the proteins necessary to deal with stress in the ER while reducing many nonessential proteins, thereby lessening the task of processing them as it would under normal circumstances. If the ER burden is overwhelming, beyond mitigation, cell death occurs via apoptosis. Viruses control these processes to ensure that the cell will not die until they have replicated in sufficient numbers to achieve efficient propagation upon cell disintegration following apoptosis.

Scope

Viral infection has an impact on various biological functions at the level of the whole organism and at the levels of organs, tissues, cells, subcellular structures, molecules, and genes. Thus, the topic of viral virulence and stress can be treated from a variety of perspectives, centering on the immune system, the endocrine system, the gastrointestinal tract, the blood and blood-forming tissues, or many other systems and tissue and cell types. Moreover, the approach to the subject matter considered here can also be diverse, for example, clinical, histopathological, cytological, subcellular, and molecular. We focus on the cell of eukaryotic multicellular organisms (e.g., humans and rodents) and discuss the impact of viral infection on the ER, specifically virus-induced ER stress, the ensuing UPR, and the ER molecular chaperones.

Viral Infection and ER Stress

A number of viruses are known to cause ER stress and to induce the UPR (Table 1). These viruses are of interest because they are pathogenic to humans and to animals valuable to humans. Some of these viruses cause very serious diseases. For example, hepatitis C virus (HCV) causes hepatitis that can develop into liver cirrhosis and then into hepatocellular carcinoma. Others, such as the human cytomegalovirus (HCMV), cause mild illnesses in otherwise healthy individuals,

Table 1 Some viruses that cause ER stress and trigger the UPR^a

Name	Family	Subfamily	Genus	Group	Nucleic acid	Pathology
HCV	Flaviviridae	–	<i>Hepacivirus</i>	IV	ssRNA(+)	Non-A non-B hepatitis, liver cirrhosis, hepatocellular carcinoma
Bovine viral diarrhea virus	Flaviviridae	–	<i>Pestivirus</i>	IV	ssRNA(+)	Diarrhea in cattle
Japanese encephalitis virus	Flaviviridae	–	<i>Flavivirus</i>	IV	ssRNA(+)	Encephalitis
HCMV or HHV-5	Herpesviridae	β -Herpesvirinae	<i>Cytomegalovirus</i>	I	dsDNA	Infectious mononucleosis, retinitis, serious infections in people with defective or immature immune systems ^b
HHV-1/ HSV-1	Herpesviridae	α -Herpesvirinae	<i>Simplexvirus</i>	I	dsDNA	Oral and genital herpes
EBV or HHV-4	Herpesviridae	γ -Herpesvirinae	<i>Lymphocryptovirus</i>	I	dsDNA	Infectious mononucleosis, Burkitt's lymphoma
HRSV	Paramyxoviridae	Pneumovirinae	<i>Pneumovirus</i>	V	ssRNA(–)	Respiratory tract infections
TULV	Bunyaviridae	–	<i>Hantavirus</i>	V	ssRNA(–)	Hemorrhagic fever ^c
Murine retroviruses	Retroviridae	–	Several	IV	ssRNA-RT	Tumors and leukemias in mice

^ads, double-stranded; EBV, Epstein–Barr virus; ER, endoplasmic reticulum; HCMV, human cytomegalovirus; HCV, hepatitis C virus; HHV-1/HSV-1, herpes simplex virus 1; HHV-4, human herpes virus 4; HHV-5, human herpes virus 5; HRSV, human respiratory syncytial virus; ss, single-stranded; TULV, tula hantavirus; UPR, unfolded protein response.

^bFetuses, newborns, immunosuppressed individuals; and AIDS patients.

^cChiefly affected are the kidneys, lungs, and heart.

but they can cause very serious infections and even death in immunocompromised individuals (e.g., a patient who has received an organ transplant and is undergoing immunosuppressive therapy to prevent rejection of the transplanted organ). Still other viruses cause non-life-threatening but troublesome lesions; for example, herpes simplex virus 1 (HSV-1) produces painful blisters (herpes) of the oral cavity and lips.

Chain of Events and Its Consequences for the Virus and for the Infected Cell

It is important to realize that although a viral infection has a direct impact on the ER, it also affects other organelles, particularly the Golgi apparatus. Furthermore, it affects not only the originally infected cell but also many others in the vicinity and in tissues far away from the original infection site.

Another important consideration is that viral infection, viral virulence and its impact on the cell (the ER in particular), induction of the UPR and apoptosis, and cell destruction must all be regulated so the virus can multiply and propagate to other cells and tissues. If the damaging impact of the viral infection on the cell is too strong and too fast, the cell will die too soon for the virus to complete its cycle of infection and replication and the cell will disintegrate before a

sufficient number of viral particles have been assembled for efficient propagation to other cells and organs via the blood.

Viruses need the host cell – its genes, proteins, and structures – for the replication and assembly of progeny virus. Thus, evolutionarily successful viruses have acquired genes and mechanisms that allow them to use the host cell's assets, even as they trigger ER stress and the UPR, en route to apoptosis. Some of the genes acquired by viruses for controlling the cellular response to infection encode anti-apoptotic proteins, which temporarily counteract the action of pro-apoptotic factors that are produced by the host cell on viral infection.

The Endoplasmic Reticulum

Subcompartments

Although the ER is a single organelle, present in one copy per cell, it includes specialized subcompartments, or domains, with specific functions.

- The rough ER (RER) is in contact with associated ribosomes, is close to the nucleus, and is involved in protein synthesis and in the translocation of newly synthesized polypeptides from the cytosol into the ER lumen.

- The transitional ER (tER) appears as cup-shaped cisternae and is involved in the exit of newly made polypeptides carried in coat protein complex II (COPII) vesicles toward the secretory route.
- The smooth ER (SER) participates in calcium storage, in lipid and steroid synthesis, and in the breakdown of drugs and chemical compounds, such as those present in medical remedies, which occurs predominantly in liver cells. The SER involved in calcium storage is prominent in muscle cells, including those of the heart and limbs; thus the SER plays a key role in muscular contraction, namely body movements and heart beats.

The Endoplasmic Reticulum is a Highly Dynamic Organelle

The ER varies quantitatively and qualitatively from cell to cell, depending on the functional status of the cell and on the tissue or organ in which the cell resides. Likewise, the distribution of the ER subcompartments inside a cell varies in relation to the cell-cycle stage and functional phase. For example, SER usually occurs in relatively small quantities, but its abundance is greater in cells that process and secrete lipids or steroids (e.g., in adrenal gland cells actively producing steroid hormones). RER is very abundant in cells that actively synthesize proteins, such as digestive enzyme-producing cells in the glandular tissue lining some portions of the gastrointestinal tract, and in plasma cells that actively produce immunoglobulins.

Specialized ER subcompartments also change within a cell, as illustrated by the switch from SER to RER, and *vice versa*; such a transition occurs in cells switching from low to high rates of protein synthesis, followed by return to the initial resting phase.

The Endoplasmic Reticulum and the Golgi Apparatus

Another important organelle of the eukaryotic cell is the Golgi apparatus, which interacts continuously with the ER. The Golgi is built of vesicles, tubules, and cisternae piled close to one another. All of these structures are formed by membranes composed of a huge range of proteins, perhaps well over a 1000 different varieties of proteins. The Golgi also possesses an enormous capacity for morphological transformation. This dynamism of the Golgi is affected by many factors, such as cellular environmental conditions, the cell's mitotic and physiological stages, the presence of drugs, stressors, disruption of microtubules, and protein movement from the ER.

The unusual plasticity of the Golgi membranes is probably due to the fact that proteins associate with the membrane only temporarily while they are

traveling from the ER to other cellular locales. In fact, no specific class of protein has been identified that is uniquely and permanently part of the Golgi membranes. Golgi components are continuously entering and exiting this organelle.

The Golgi apparatus participates in protein glycosylation, cleavage of preproteins to produce the active form of the protein, and sorting of molecules so that they are properly tagged and thus directed to their physiological destinations, whether the cytosol, the cell membrane, or the extracellular space. The number of copies of the Golgi apparatus per cell varies from one to many, according to cell type, cell location (tissue type), and physiological and cell-cycle stage.

The Golgi can be subdivided into several regions, on the basis of function. The cis region lies near the ER and the nucleus and receives vesicles with their cargo of molecules coming from the SER. The middle region is the site of protein and lipid processing (e.g., glycosylation), and the trans zone or network region is sited near the cell membrane. In this region of the Golgi, preproteins are cleaved to generate functional protein molecules (as part of the maturation pathway from nascent polypeptide to fully processed, native protein) and the fully processed proteins are sorted according to their final destinations.

Interaction of the Endoplasmic Reticulum with Other Cellular Structures

The ER communicates directly with the nucleus, as mentioned earlier. Likewise, the ER (in the tER regions) interacts with or is modified by mitochondria, mRNAs, plasma membrane, peroxisomes, and phagosomes. These structures affect, from the outside, the shape and functionality of the ER, generating the various regions of this organelle, which suggest at first glance a variety of different ERs. Important within the context of this article is the interaction of the ER with mitochondria, a key stage in apoptosis.

Stress Proteins and Molecular Chaperones in the Endoplasmic Reticulum

A set of molecular chaperones resides and functions in the ER ([Table 2](#)). The ER chaperones participate in protein folding, refolding, and degradation, as do chaperones in other cellular compartments. However, the ER chaperones are involved as well in other processes that are unique to this organelle, such as calcium binding and protein translocation (exit from and entry into the ER lumen). It is of interest that some of the genes encoding ER chaperones are induced by glucose deprivation, which explains why these chaperones are classified and named as glucose-regulated proteins (e.g., Grp78, for glucose-regulated protein 78; see [Table 2](#)).

Table 2 Some stress proteins and chaperones found in the endoplasmic reticulum^a

Mass (kDa)	Examples ^b
100 or higher	Hsp170 (Grp170)
81–99	Hsp90 (GRP94; Grp94; endoplasmin; gp96); PDIase: ERpdj5
65–80	BiP(Grp78); PDIase: ERp72, and ERp65; Kar2p; Lhs1p (Ssi1p, Cer1p)
55–64	RBP; PDIase: PDI, ERp57 (GRP58), PDip, and PDIr; calnexin
35–54	Hsp40; Hsp47; Sec63p; ERdj3; Scj1p; JEM1p; PDIase: P5, ERp44, ERp46, TMX2long, TMX3, and TMX4; calreticulin
34 or lower	Cpn20; PDIase: ERp27, ERp28, ERp18, TMX, and TMX2short

^aStress proteins are those whose parent genes are induced by stress caused by any stressor. However, not all stress proteins are molecular chaperones; likewise, not all chaperones are stress proteins. Examples of chaperones residing in the ER are listed, whether or not they are stress proteins. There are also many other stress proteins in the ER (not listed here) that are not chaperones or that are not known at this time to function as chaperones.

^bSynonyms are given in parentheses. Chaperones build multimolecular assemblies in order to exercise their functions. These assemblies are formed by identical or different chaperone molecules (subunits) and cofactors. In addition, these chaperone complexes interact with other chaperone complexes (networking) and with the protein-degradation machinery (e.g., the ubiquitin-proteasome system) to integrate the protein quality-control system of the cell.

The importance of the ER chaperones, both in normal physiological events and in dealing with stress generated by a variety of stressors (including viruses), emphasizes the critical, deleterious effects that defective chaperones have on the cell, and on the organism in general. A newly recognized type of pathology, the chaperonopathy, encompasses chaperones that are abnormal due to genetic (e.g., mutation) or acquired (e.g., aberrant posttranslational modification) defects and the impact of such abnormal chaperones on health and disease. The consequences of a chaperonopathy in the ER have not yet been evaluated, but they can be predicted to be significant for the cell and for any virus that causes ER stress. It is likely that chaperone molecules become deficient with age, due to the same biochemical modifications that affect other proteins in the elderly; thus, the probability that viral infections will cause serious ER stress and cell damage is likely to increase with senescence.

ER Stress

A variety of stressors, including certain viruses (Table 1), can cause ER stress and can trigger the ER-stress response. This variety of stressors notwithstanding, the common central factor underlying ER stress is

the presence in the ER lumen of denatured proteins – partially or totally unfolded, or misfolded proteins. It is crucial for survival that, under ER-stress conditions, the cell readjusts a series of mechanisms that will reduce the number of ER tasks (essentially by diminishing the number of peptides that travel through and are modified in the ER) while increasing its functional power by expanding the ER's capacity to modify and tag proteins and to assist them in folding and refolding, thus steering them toward a productive pathway.

The ER plays a critical role in the monitoring of protein quality as the molecules traverse the organelle. It possesses the means to detect and to correct aberrant peptides and abnormal protein conformations. However, if correction is impossible, the ER has mechanisms to lead unrecoverable peptides toward degradation via the ER-associated degradation (ERAD) pathway. Finally, if all else fails, the ER can steer the cell toward apoptosis.

In addition to protein folding and modification (e.g., glycosylation), the synthesis of lipids and sterols occurs in the ER. Recent data suggest that lipid synthesis is altered as much as protein quality control during ER stress and the ensuing UPR.

A key event in the UPR is the increased expression of the genes encoding the ER-resident chaperones (Table 2). When the ER is overwhelmed by an excess of too many bad proteins, signals travel from the organelle to the nucleus, resulting in the upregulation of some genes and the downregulation of others. The downregulation reduces the nonessential biosynthetic tasks (e.g., tasks pertaining to the synthesis of secretory proteins) of the ER, which must redirect its resources to deal with stress. Still other mechanisms, such as the inhibition of mRNA translation, cooperate in reducing the synthesis of nonessential proteins.

ER Stress and Apoptosis

Apoptosis, a form of programmed cell death, is a multistage process that involves various types of molecules, many of them enzymes. The series of stages – the apoptotic cascade – can be induced in various manners that may schematically be organized into two main pathways: intrinsic and extrinsic. Intracellular injury (e.g., DNA damage and ER stress) starts the intrinsic or ER-stress apoptotic pathway, whereas extracellular agents activate the extrinsic (or receptor-mediated) pathway. In the latter pathway, the initial steps require the participation of cell-surface receptors to engage the enzyme caspases (cysteiny aspartic acid proteases) and start the caspase cascade. This cascade includes the caspase-8-dependent cleavage of caspase-3 and the participation of caspase-8 in the cleavage of Bcl-2-associated

protein 31 (Bap31; Bcl stands for B-cell leukemia), and the recruitment of Bcl-2 proteins that interact with mitochondria and cause the leakage of cytochrome c from the organelle into the cytosol. This is followed by formation of a caspase-9 complex and apoptosis.

In contrast, the intrinsic pathway involves the mobilization of caspase-12 with the direct activation of caspase-9, initially bypassing the mitochondria. However, the intrinsic pathway also involves mitochondria (and cytochrome c leakage) through Bap31 cleavage with participation of caspase-12, resulting in the amplification of the caspase cascade.

Bcl-2 homologous antagonist/killer (Bak) and Bcl-2 associated X (Bax) proteins (i.e., two members of the pre-apoptotic class of proteins) become inserted into the ER membrane at the beginning of the ER-stress pathway, causing calcium efflux, which in turn causes an increase in cytosolic calcium. Calpain is activated, as is procaspase-12 in the ER. The caspase cascade proceeds up to the activation of procaspase-3. At this junction, there is no involvement yet of apoptotic protease-activating factor-1 (Apaf-1; a component of the apoptosome, whose main components are Apaf-1, cytochrome c/dATP, and procaspase-9), nor is there leakage of cytochrome c from the mitochondria into the cytosol. While the apoptotic cascade progresses, cytosolic calcium is taken up by the mitochondria as it escapes from the ER and mitochondrial membrane pores open, allowing cytochrome c to pass into the cytosol. At this point, the apoptosome is assembled and becomes active, as does procaspase-3.

The extrinsic route to apoptosis involves other factors: the c-Jun N-terminal kinase (JNK) inhibitory kinase (JIK), tumor necrosis factor receptor-associated factor 2 (TRAF2), and inositol-requiring enzyme 1 α (Ire1 α or IRE1 α). A trimeric complex is assembled that includes IRE1 α , TRAF2, and apoptosis-signal regulating kinase 1 (ASK1), resulting in the activation of ASK1 and JIK. This is followed by cell death.

The intrinsic and extrinsic pathways interact. TRAF2 promotes procaspase-12 aggregation and is released from procaspase-12 when ER stress occurs. This may happen because TRAF2 sequesters IRE1 and thus activates procaspase-12.

Some of the steps just described require additional experimental confirmation. Although these apoptotic pathways require caspases, suicide in some cells (e.g., certain neurons) does not involve these enzymes but resort to other mechanisms. In general, no matter which pathway is considered, caspase-dependent or -independent, the details of the mechanisms involved are still incompletely understood and may vary among cell types. However, the data available do indicate that the intrinsic and extrinsic apoptotic routes

intersect and that the ER is at center stage in the programmed cell death triggered by ER stress.

ER Stress and Viral Infection: Examples

Hepatitis C Virus

HCV, a single-stranded RNA virus, causes non-A non-B hepatitis in humans. The potential consequences of the viral infection are the development of hepatic cirrhosis and eventually hepatocellular carcinoma. The ER is the site of HCV core processing and maturation; these processes involve the interaction of viral components with the ER membrane.

We have seen that the UPR is characterized by the activation of signals from the ER to the nucleus, resulting in the increased expression of ER-chaperone genes, calcium-binding proteins, and sarcoplasmic-ER calcium-transporting ATPase 2 (SERCA2), whose function is to pump calcium from the cytosol into the ER. In contrast, other proteins are decreased, via either the downregulation of their parent genes or the inhibition of the translation of their mRNAs. In this process, double-stranded-RNA-activated protein kinase-like ER kinase (PERK) plays a role, by mediating the phosphorylation of the eukaryotic initiation (of translation) factor 2 α (eIF2 α). When ER stress is intense and/or prolonged, apoptosis can occur in one of the following ways, depending on the animal species and other factors: (1) the activation, in mouse cells, of caspase-12 (not yet found in humans); (2) the activation of the C/EBP-homologous protein (CHOP; also known as the growth arrest- and DNA-damage-inducible gene 153, GADD153); or (3) the induction of JNK, involving a complex formed by TRAF2 and ASK1.

Calcium signaling is also dependent on ER function. The ER controls the opening of the inositol 1,4,5-triphosphate receptor (IP3R) channels and calcium uptake into the lumen by the SERCAs. The mitochondrion also participates in calcium trafficking; this organelle is very efficient in calcium uptake and release. Some mitochondria are situated very close to the ER and thus are conveniently positioned to sense changes in calcium concentration in the cytosol as luminal calcium leaks out of the stressed ER through activated calcium (IP3R) channels.

The HCV core induces the expression of the chaperones BiP(Grp78), Hsp90(Grp94), and calreticulin (Table 2) and SERCA2 at levels well above those characteristic of the unstressed ER. The increase in the levels of all these proteins is enhanced and paralleled by the loss of calcium from the ER lumen. Eventually, the HCV core triggers the apoptotic cascade due to the hyperexpression of CHOP/GADD153

(which acts as a pro-apoptotic factor), Bax translocation to mitochondria, mitochondrial membrane depolarization, cytochrome c efflux, and caspase-3 activator and poly-(ADP-ribose) polymerase (PARP) cleavage.

Human Cytomegalovirus

HCMV belongs to the Herpesviridae family, along with the varicella-zoster virus, HSV-1, and HSV-2, which cause chickenpox and shingles, and Epstein-Barr virus (Table 1). Epstein-Barr virus and HCMV are the most common etiological agents of infectious mononucleosis. HCMV remains dormant in infected cells without generating symptoms, but it can be activated and can generate serious cell injury in individuals with immature immune systems (e.g., fetuses, newborns, and patients with congenital deficiencies), and in patients with impaired immunity, such as those being treated with drugs that kill lymphocytes and other blood cells (e.g., cancer patients undergoing chemotherapy and individuals who have received an organ transplant and are being given immunosuppressants to prevent rejection of the foreign tissue).

HCMV induces ER stress, which is sensed by means of activating transcription factor 6 (ATF6), PERK, and IRE1. Normally, in the absence of stress, BiP(Grp78) (see Table 2) is bound to each of the sensors on the luminal side of the ER membrane. If unfolded, misfolded, or aggregated proteins accumulate in the ER, BiP(Grp78) releases the sensors and interacts with the proteins to assist them to fold correctly so that they can regain their native functional conformation. When PERK and IRE1 are no longer bound to BiP(Grp78), their luminal domains form homodimers and become phosphorylated and activated. Simultaneously, ATF6, also freed from the chaperone, exposes a Golgi localization signal on its structure and is thus translocated to the Golgi, in which it is cleaved and activated.

PERK activation initiates a pathway for the attenuation of translation that involves eIF2 α , eIF2 β , activating transcription factor 4 (ATF4), and other gene proteins. At the same time, ATF6 mediates transcriptional upregulation of ER chaperone genes. ATF6 is an ~50-kDa protein produced by the digestion of a 90-kDa preprotein in the Golgi. As a result of the ATF6-initiated pathway, the levels of ER chaperones increase in response to the increased demand due to accumulation of client polypeptides in the ER lumen.

Also part of the virus-initiated UPR is the induction of the IRE1 pathway, increasing the expression of genes whose protein products are involved in promoting protein degradation. In this manner, defective peptides that cannot be refolded by the chaperones are

eliminated via digestion, and the digestion-derived amino acids are reused on the ribosome to synthesize essential proteins (e.g., chaperones) as required by the increased demand due to ER stress.

In summary, HCMV intervenes in the regulation of the three routes of signal transduction characteristic of the UPR, the routes mediated by PERK, IRE1, and ATF6. The HCMV intervention ensures that the virus has time to replicate while the cell is preparing for apoptosis, which leads to cell disruption and the release of viral particles. In this regard, it has been found that the genes encoding ER chaperones are also induced by the virus independently of the ATF6-initiated pathway, possibly because the chaperones are needed for the generation of proteins required by the virus. In parallel, the protein degradation ER degradation-enhancing 1,2-mannosidase-like (EDE1) factor was found to be inhibited, so as to transitorily curtail proteolysis and thus ensure the presence of the molecules necessary for cell survival during viral replication.

Stress and Viral Reactivation

Viral infection causes cell stress, as just explained. The reverse, the stress activation of latent viruses, also occurs. Some viruses become latent after the acute phase of infection; their presence in the cell causes no clinically detectable effects. The viral genome remains in the infected cell but is transcriptionally repressed except for a minor region; hence the virus is inactive or latent. A number of cell stressors, such as fever, ultraviolet light, and social tensions, with their accompanying neuroendocrine imbalance, induce the reactivation of these latent viruses. Viruses belonging to the Herpesviridae family (Table 1), (e.g., HSV-1, varicella zoster virus, and Epstein-Barr virus) are among those that pass through active and latent stages depending on the microenvironmental cellular conditions. Upon the impact of stressors, the mechanisms that keep the viral genome transcriptionally repressed are disrupted and the virus begins to replicate and accumulates in the host cell, usually generating pathological and clinical consequences. For instance, when HSV-1, a neurotrophic virus (it invades neurons), is reactivated in ganglionic cells, viral particles are produced, which emerge from the host cell and invade other cells that thus suffer viral-induced stress. The molecular mechanisms that control latency and reactivation are poorly understood. Likewise, the molecular mechanisms that mediate stress-induced viral reactivation have not been fully elucidated and the role of molecular chaperones in these reactivating mechanisms remains to be determined.

See Also the Following Articles

Chaperone Proteins and Chaperonopathies; Heat Resistance; Heat Shock Genes, Human; Heat Shock Response, Overview; Herpesviruses; Chaperonopathies; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles.

Further Reading

- Benali-Furet, N. L., Chami, M., Houel, L., et al. (2005). Hepatitis C virus triggers apoptosis in liver cells by inducing ER stress and ER calcium depletion. *Oncogene* 24, 4921–4933.
- Ciechanover, A. (2005). Proteolysis: from the lysosome to ubiquitin and the proteasome. *Nature Reviews Molecular Cell Biology* 6, 79–86.
- Efstathiou, S. and Preston, C. M. (2005). Towards an understanding of the molecular basis of herpes simplex virus latency. *Virus Research* 111, 108–119.
- Ellgaard, L. and Ruddock, L. W. (2005). The human protein disulphide isomerase family: substrate interactions and functional properties. *EMBO Reports* 6, 28–32.
- Federovitch, C. M., Ron, D. and Hampton, R. Y. (2005). The dynamic ER: experimental approaches and current questions. *Current Opinion in Cell Biology* 17, 409–414.
- Isler, J. A., Skalet, A. H. and Alwine, J. C. (2005). Human cytomegalovirus infection activates and regulates the unfolded protein response. *Journal of Virology* 79, 6890–6899.
- Kleizen, B. and Braakman, I. (2004). Protein folding and quality control in the endoplasmic reticulum. *Current Opinion in Cell Biology* 16, 343–349.
- Kroemer, G. and Martin, S. J. (2005). Caspase-independent cell death. *Nature Medicine* 11, 725–730.
- Lee, A. S. (2001). The glucose-regulated proteins: stress induction and clinical applications. *Trends in Biochemical Sciences* 26, 504–510.
- Levine, T. and Rabouille, C. (2005). Endoplasmic reticulum: one continuous network compartmentalized by extrinsic cues. *Current Opinion in Cell Biology* 17, 362–368.
- Li, X.-D., Lankinen, H., Putkuri, N., et al. (2005). Tula hantavirus triggers pro-apoptotic signals of ER stress in Vero E6 cells. *Virology* 333, 180–189.
- Macario, A. J. L. and Conway de Macario, E. (2004). The pathology of anti-stress mechanisms: a new frontier. *Stress* 7, 243–249.
- Macario, A. J. L. and Conway de Macario, E. (2005). Sick chaperones and cellular stress: pathological and clinical implications. *New England Journal of Medicine* 353, 1489–1501.
- Macario, A. J. L., Grippo, T. M. and Conway de Macario, E. (2005). Genetic disorders involving molecular-chaperone genes: a perspective. *Genetics in Medicine* 7, 3–12.
- Miller, C. S., Danaher, R. J. and Jacob, J. (2006). ICP0 is not required for efficient stress-induced reactivation of herpes simplex virus type 1 from cultured quiescently infected neuronal cells. *Journal of Virology* 80, 3360–3368.
- Munro, S. (2005). The Golgi apparatus: defining the identity of Golgi membranes. *Current Opinion in Cell Biology* 17, 395–401.
- Panet, A., Braun, E., Honigman, A., et al. (2005). Involvement of cellular death signals in the reactivation of herpes simplex virus type 1 and lambda bacteriophage from latent state. *Journal of Theoretical Biology* 236, 88–94.
- Puthenveedu, M. A. and Linstedt, A. D. (2005). Subcompartmentalizing the Golgi apparatus. *Current Opinion in Cell Biology* 17, 369–375.
- Schroeder, M. and Kaufman, R. J. (2005). The mammalian unfolded protein response. *Annual Review of Biochemistry* 74, 739–789.
- Su, H.-L., Liao, C.-L. and Lin, Y.-L. (2002). Japanese encephalitis virus infection initiates endoplasmic reticulum stress and an unfolded protein response. *Journal of Virology* 76, 4162–4171.
- Westhoff, B., Chapple, J. P., van der Spuy, J., et al. (2005). HSP70 is a neuronal shuttling factor for the sorting of chaperone clients to the proteasome. *Current Biology* 15, 1058–1064.
- Zhang, X., Uthaisang, W., Hu, L.-F., et al. (2005). Epstein-Barr virus-encoded latent membrane protein 1 promotes stress-induced apoptosis upstream of caspase-2-dependent mitochondrial perturbation. *International Journal of Cancer* 111, 397–405.

Viruses and Stress

G Z Feuerstein and R R Ruffolo

Wyeth Research, Collegeville, PA, USA

B-N David

Israel Institute for Biological Research, Ness-Ziona,
Israel

© 2007 Elsevier Inc. All rights reserved.

Introduction

Stress: Viral Infection and Penetration to the CNS

Cold Stress Neuroinvasiveness and Virulence of Attenuated
WNV

Summary

Glossary

<i>Attenuated West Nile Virus (WNV)</i>	A virus that has lost its neuroinvasiveness but not its neurovirulence.
<i>Blood-brain barrier (BBB)</i>	Property of endothelial cell brain microvessels that provides a barrier for transport of molecules from the systemic circulation into the brain. The BBB is considered to have an important function that provides the central nervous system with a constant environment needed for normal brain function. Stress conditions disrupt the BBB permeability, which is often increased in many stress situations.
<i>Glucocorticosteroids (GCS)</i>	Organic lipid molecules produced primarily in the adrenal gland. GCS serve as local and systemic hormones and act on specific receptors that activate multiple genes and signaling pathways governing multitude biological actions. Most prominent in GCS action are glucose, protein, and lipid metabolism, regulation of immune and inflammatory cells and mediators, bone and muscle remodeling, as well as central nervous system effects.
<i>Neuroinvasiveness</i>	Ability to penetrate the central nervous system.
<i>Neurovirulence</i>	Ability to establish a lethal infection within the central nervous system.
<i>Plaque-forming unit (PFU)</i>	A standard measure of viral load.
<i>Viral encephalitis</i>	Viral infection of the brain. In order to cause encephalitis when invading the central nervous system, the virus must possess neurovirulence.
<i>Virulence</i>	Ability of a virus to cause clinical disease and potential for mortality. Wild-type

West Nile virus is both neurovirulent and neuroinvasive.

West Nile Virus (WNV)

The etiologic agent of West Nile fever, an endemic viral disease in many subtropical regions and an emerging mosquito-borne disease in the United States. WNV is a neuroinvasive flavivirus that may cause systemic and central nervous system infection, paralysis, and death.

Introduction

Stress is defined as any external or internal stimuli that disrupt the natural homeostasis of the organism. Stress situations commonly result in suppression of the immune system. The reaction of the organism to stress can be adaptational and helpful in restoring optimal homeostasis to maintain the organism functions or can lead to a variety of aberrant emotional and physical conditions with long-lasting consequences.

The effect of stress on the humoral immune response has been reported in numerous studies. The responses to stress vary greatly depending on the stress patterns (type and duration) and the type of stressor. The effect of stress on the immune system in human and animal studies is invariably suppressive, on both the cellular and humoral levels.

Physiological responses to stressors are largely mediated by the sympathetic nervous system and the hypothalamic-pituitary-adrenal (HPA) axis. Glucocorticoids (GCS) are major mediators in the reaction to stress and in stress-induced immunosuppression.

Stress stimuli affect the immune system and cause immunosuppression through involution of lymphoid organs, possibly mediated by increased activity of the pituitary-adrenocortical axis. There is also a clear connection between stress, the central nervous system (CNS), and the immune response. The role of the CNS in immune responses to stress conditions associated with infectious viral conditions is also well established. Glucocorticoids released in response to stress are the major mediators of the effect of stress on viral replication, change in virulence, and viral penetration to the CNS.

Stress: Viral Infection and Penetration to the CNS

A variety of stress paradigms have been shown to exacerbate the effects of several infectious agents, including herpes simplex virus, coxsackie B1 virus,

vesicular stomatitis virus, cytomegalovirus (CMV), influenza virus, encephalitic viruses, and mengovirus. Viruses are known to flourish in the stressed host, with increased morbidity and mortality. It appears that stress acts by enhancing virus replication in lymphocytes as shown in the spleen of stressed mice. Since activated lymphocytes can pass through the blood-brain barrier (BBB), the larger fraction of infected lymphocytes in stressed mice could also contribute to higher virus titer in the brain. This possibility, however, awaits further clarification by more direct evidence on the mode of immune cell penetration into the CNS.

An attenuated strain of West Nile virus (WNV) in mice had little capacity to induce encephalitis and death, as the level of the virus in the brain was very low. Specifically, brain virus levels in control mice inoculated with WN-25, a noninvasive variant of the West Nile virus, or SVN, a neurovirulent, noninvasive Sindbis virus, showed titers of less than 2 log 10 plaque-forming units (PFU), whereas the virus titers in the brain of cold-stressed mice or mice exposed to GCS treatment were 8.5, 7.8, and 7.5 (log 10 PFU) for WN-25 and 7.3, 7.1, and 7.0 (log 10 PFU) for SVN, respectively. In stressed mice, these viruses (WN-25 and SVN), both lacking neuroinvasive properties in control nonstressed mice, readily invaded the brain and proliferated to high titer levels ($>10^6$ PFU/brain) in all moribund mice inoculated with WN-25 or with SVN, on top of a stress condition. It was suggested that the neurovirulent properties of non-neuroinvasive viruses can serve as a marker for BBB permeability and aid in studying pathological processes of the brain.

Cold Stress Neuroinvasiveness and Virulence of Attenuated WNV

In mice inoculated with the wild-type strain of WNV, it was shown that short-term exposure of the mice to cold water ($1 \pm 0.5^\circ\text{C}$) for 8–10 days increased the mortality from 47% in control to 92–100% in the cold-stressed mice. Cold or isolation stress increased blood, spleen, and brain virus levels by two to three orders (100- to 1000-fold) as compared to control nonstressed mice. WNV, which was isolated from the brain of moribund stressed mice, was more virulent when inoculated intraperitoneally to normal nonstressed mice. Concomitantly, lymphoid organs such as spleen diminished in mass in the stressed mice. These observations were also demonstrated with WN-25 and SVN.

Since GCS levels are elevated during stress, the effects of stress or corticosterone (a rodent GCS)

injection were investigated on neurovirulence and encephalitis by the attenuated arboviruses. These two viral strains were exposed to the same cold stress paradigm described previously. The stress paradigms induced mortality of 65, 80, and 50% (WN-25) and 75, 78, and 58% (SVN) in cold, isolation, and corticosterone-treated mice, respectively. No mortality was seen in nonstressed virus-inoculated mice. The level of WN-25 virus in the brain of stressed mice was 7.5–8.5 log 10 PFU. No virus was detected in nonstressed inoculated mice. Isolation of a neuroinvasive strain from brains of stress-exposed mice suggested that the induction of infection was attributed to immunosuppression with subsequent increased viral replication and selection of neuroinvasive strain. The WN-25 virus extracted from the brain of moribund stressed mice was extremely virulent; inoculation of as little as 5 PFU to normal nonstressed mice caused severe encephalitis and death. WN-25 showed a change in its neuroinvasive properties, raising the possibility of stress-enhanced proliferation and a selective process leading to reversal of the neurovirulence. In contrast, brain SVN virus levels isolated from moribund mice (stress exposed) were no different from levels of viruses in the nonstressed mice, suggesting a different role of stress in the mechanism of stress-induced virulence.

The susceptibility of stressed mice to the viral infection could be explained by reduced number and suppression of macrophage and T-lymphocyte activity as well as by decreased natural killer cell cytotoxicity. Therefore, stress can induce elevated viremia for longer periods, facilitating the penetration of viruses into the CNS.

Summary

There is strong evidence for an association between stress and the immune system and between the immune system and the CNS. Susceptibility to infections is presumed to be primarily mediated by immune function. Since WNV is widely distributed and is capable of endemic spread, the effect of stress on WNV infection has been studied. It seems that suppressed immune response during stress, with defective function of macrophages and lymphocytes, leads to enhancement of viral replication in the blood and peripheral organs, induction of a selective process, and selection of a neuroinvasive strain.

See Also the Following Articles

Glucocorticoids, Role in Stress; Immune Suppression; Infection; Viral Virulence and Stress.

Further Reading

- Ben-Nathan, D. (1994). Stress and infectious disease. *Israel Journal of Veterinary Medicine* **49**, 105–112.
- Ben-Nathan, D. and Feuerstein, G. (1990). The influence of cold or isolation stress on resistance of mice to West Nile Virus. *Experientia* **46**, 285–290.
- Ben-Nathan, D., Lustig, S. and Feuerstein, G. (1989). The influence of cold or isolation stress on neuroinvasiveness and virulence of an attenuated variant of West Nile Virus. *Archives of Virology* **109**, 1–10.
- Ben-Nathan, D., Lustig, S. and Kobiler, D. (1996). Cold stress-induced neuroinvasiveness of attenuated arboviruses is not solely mediated by corticosterone. *Archives of Virology* **141**, 1221–1229.
- Ben-Nathan, D., Kobiler, D., Loria, R. M. and Lustig, S. (1998). Stress-induced CNS penetration by non-invasive attenuated encephalitis viruses. In: Levy, A., Grauer, E., Ben-Nathan, D. & de Kloet, E. R. (eds.) *New frontiers in stress research: modulation of brain function*, pp. 277–283. Amsterdam: Harwood Academic Publishers.
- Cohen, S. and Williamson, G. M. (1991). Stress and infectious disease in humans. *Psychological Bulletin* **109**, 5–24.
- Dhabhar, F. S., Miller, A. H., McEwen, B. S. and Spencer, R. L. (1995). Effect of stress on immune cell distribution. *Journal of Immunology* **154**, 5511–5527.
- Isler, J. A., Skalet, A. H. and Alwine, J. C. (2005). Human cytomegalovirus infection activates and regulates the unfolded protein response. *Journal of Virology* **79**, 6890–6899.
- Lustig, S., Olshevsky, U., Ben-Nathan, D., Lachmi, B. E., Malkinson, M., Kobiler, D. and Halevy, M. (2000). A live attenuated West Nile strain as a potential veterinary vaccine. *Viral Immunology* **13**, 401–410.
- Mehta, S. K., Cohrs, R. J., Forghani, B., Zerbe, G., Gilden, D. H. and Pierson, D. L. (2004). Stress-induced subclinical reactivation of varicella zoster virus in astronauts. *Journal of Medical Virology* **72**, 174–179.
- Sheridan, J. F., Dobbs, C., Brown, D. and Zwillig, B. (1994). Psychoneuroimmunology: stress effects on pathogenesis and immunity during infection. *Clinical Microbiology Reviews* **7**, 200–212.
- Welsh, C. J., Bustamante, L., Nayak, M., Welsh, T. H., Dean, D. D. and Meagher, M. W. (2004). The effects of restraint stress on the neuropathogenesis of Theiler's virus infection II: NK cell function and cytokine levels in acute disease. *Brain Behavior and Immunity* **18**, 166–174.

Vomeronasal Organ See: Pheromones.

W

Waist–Hip Ratio

R Rosmond

Björndammsterrassen 41, Partille, Sweden

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Björntorp and R Rosmond, volume 3, pp 669–671, © 2000, Elsevier Inc.

Waist–Hip Ratio

Waist–Hip Ratio as a Risk Factor

Visceral Adipose Tissue

Visceral Obesity

Mechanisms of Waist–Hip Ratio as a Risk Factor

Waist–Hip Ratio as a Sign of Stress

Glossary

<i>Hypothalamic-pituitary-adrenal (HPA) axis</i>	The hormonostatic endocrine axis that regulates cortisol production and secretion.
<i>Lipoprotein lipase</i>	The rate-limiting enzyme that cleaves one fatty acid from a triacylglycerol.

Various techniques are available for the assessment of body composition. These techniques comprise use of anthropometric measurements, radiographic techniques, and more complex techniques such as total body potassium and tetrapolar bioelectrical impedance analysis.

The double thickness of the epidermis and subcutaneous fat is determined by a skin-fold measurement. The skin-fold thickness can be measured wherever a fold can be formed, but specific sites such as the triceps and subscapular locations have been designated because of their high correlations with total body fat.

Body circumferences, like skin folds, can be measured at various body locations. The sites for circumference measurements that have been reported

to be important in describing the amount of abdominal, visceral obesity are the waist and hip circumferences and their ratio, waist–hip ratio (WHR). In addition, the WHR has been reported to be a powerful predictor of obesity-related diseases such as cardiovascular disease, diabetes mellitus, hypertension, and hyperlipidemia. Although the abdominal fat distribution pattern is more common in men, both men and women show increased risk of serious disease with greater abdominal fat. Men may be considered at increased risk if the WHR is greater than 0.95, and women if the WHR is greater than 0.80.

Cortisol regulates adipose-tissue differentiation, function, and distribution, and in excess, causes visceral obesity. The presence of a continuously changing and sometimes threatening external environment may, when the challenges exceed a threshold, activate the hypothalamic-pituitary-adrenal (HPA) axis. The resulting increase in circulating cortisol directs body fat to visceral depots, mainly because of the high density of glucocorticoid receptors in visceral adipose tissue. Thus, central obesity may be regarded as a sign of frequent repeated or chronic perceived stress of various origins, including socioeconomic and psychosocial disabilities, alcohol, tobacco, depression, and anxiety.

Waist–Hip Ratio

Body circumferences and their ratios are used to indicate the distribution of body fat. The most important variations, in terms of health associations, are between the amounts of fat in internal, mainly intra-abdominal sites, as distinct from subcutaneous sites. The ratio of waist–hip circumferences was the first anthropometric method developed from epidemiological research as an indicator of fat distribution in relation to metabolic diseases. Most of the value in indicating body fat is derived from waist circumference, the hip circumference probably reflecting several other body tissues such as bones and muscles.

Waist-Hip Ratio as a Risk Factor

There is a well-documented sex dimorphism in regional adipose tissue distribution. Indeed, despite the fact that women are usually more obese as a group than men, male subjects more frequently have significantly higher mean waist circumference and higher mean WHR in agreement with the greater propensity of men to accumulate excess fat within the abdominal cavity. Since the WHR has been shown to be associated, albeit moderately, with the amount of abdominal visceral adipose tissue measured by computed tomography or magnetic resonance imaging, this index has been widely used to investigate the relations between regional adipose tissue distribution and metabolic profile. Thus it was effective in predicting aberrations in glucose and insulin levels and also showed a strong correlation between plasma lipids and blood pressure. WHR predicted subsequent diabetes in men and coronary heart disease in both men and women, and was more predictive of these endpoints than either the body mass index (kg m^{-2}) or a more complex procedure using the sum of multiple skin-fold thicknesses.

Visceral Adipose Tissue

Adipose tissue is specialized connective tissue that functions as the major storage site for fat in the form of triglycerides. Visceral adipose tissue has a high cellular density as compared to subcutaneous fat. In addition, the blood flow and innervation are more abundant than in subcutaneous adipose tissue. This makes the visceral adipose tissue highly metabolically active, with a turnover of fat that is considerably higher than in other fat depots. This is exaggerated by the high density of specific receptors for steroid hormones, which permit a powerful effect of triggering hormones for adipocyte metabolism, such as catecholamines. Insulin plays a predominant role in the lipogenic process. The net effect of insulin is to enhance storage and block mobilization and oxidation of fatty acids. Insulin exerts its effect by stimulating the formation of lipoprotein lipase (LPL), so that circulating triglycerides are hydrolyzed and free fatty acids can enter adipocytes. Insulin is also required for the transport of glucose, which is needed for re-esterification of triglycerides once inside adipocytes. Finally, the conversion of glucose to fatty acids is accomplished by activation of several enzymes by insulin. Cortisol in the presence of insulin stimulates, via the high density of glucocorticoid receptor (GR) in visceral adipocytes, a lively expression of LPL by transcription of the LPL gene and posttranslational stabilization of the enzyme. Cortisol and insulin exert powerful effects to both promote lipid accumulation in adipocytes and diminish lipid mobilization in

general. This is most pronounced in visceral adipocytes due to the higher density of GRs.

In concert with growth hormone, sex steroid hormones have opposite effects. Testosterone in the presence of growth hormone affects the stimulation of lipid mobilization, exaggerated by the effect of testosterone to upregulate its own receptor density. This is accomplished by transcriptional effects at several levels of the lipolytic cascade of events, including the density of β -adrenoceptors, cyclase, and hormone-sensitive lipase. Simultaneously, the LPL activity is inhibited. Estrogen has also been found to inhibit LPL.

Visceral Obesity

As mentioned above, several of the hormones of the endocrine system are involved in controlling the rate and direction of fat metabolism. While cortisol and insulin promotes accumulation of visceral fat, growth hormone and sex steroids, often in concert, result in the opposite. Cortisol also exerts chronic effects on lipid metabolism. One of the most striking in humans is the centralization of body fat seen in Cushing's syndrome, exhibited as severe truncal obesity. After successful therapy, the somatic features of Cushing's syndrome disappear.

Mechanisms of Waist-Hip Ratio as a Risk Factor

Visceral fat seems to be metabolically unique compared with subcutaneous abdominal fat. Dr Björntorp suggested that visceral fat results in hepatic insulin resistance via a portal effect of free fatty acids released by increased omental fat. The increased flux of free fatty acids to the liver leads to increased hepatic glucose production and decreased hepatic insulin clearance, which in turn leads to powerful risk factors for diabetes, cardiovascular disease, and stroke. The visceral accumulation of depot fat, as indicated by the simple WHR index, is thus an easily visible and measurable consequence of a profound, multiple endocrine perturbation.

Waist-Hip Ratio as a Sign of Stress

There can be little doubt that chronically elevated cortisol levels can redistribute body fat, as clearly seen in Cushing's syndrome. Excessive and protracted cortisol secretion is the result of a frequently evoked HPA axis, either by environmental or intrinsic factors. Environmental factors include psychosocial and socioeconomic handicaps, alcohol, and tobacco, as well as traits of psychiatric symptoms of depression and anxiety. Such stressors are all known to activate the HPA axis. In primates other than humans,

exposure to stress is followed by a diminished feedback regulation of the cortisol secretion as well as visceral obesity and associated metabolic complications. The WHR may therefore be regarded as a simple anthropometric measurement of frequently repeated or chronic exposure to various forms of stress.

Further Reading

Björntorp, P. (1993). Visceral obesity: a "civilization syndrome" *Obesity Research* 1, 206–222.

Lean, M. J. and Han, T. S. (2002). Waist worries. *American Journal of Clinical Nutrition* 76, 699–700.

Rosmond, R., Dallman, M. F. and Björntorp, P. (1998). Stress-related cortisol secretion in men: relationships with abdominal obesity and endocrine, metabolic and hemodynamic abnormalities. *Journal of Clinical Endocrinology Metabolism* 83, 1853–1859.

Rosmond, R., Wallerius, S., Wagner, P., et al. (2003). A 5-year follow-up study of disease incidence in men with abnormal hormone pattern. *Journal of Internal Medicine* 254, 386–390.

War Stress in the Former Yugoslavia

M Flögel, S Šupraha Goreta and G Lauc

University of Zagreb, Zagreb, Croatia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M Flögel and G Lauc, volume 3, pp 678–683, © 2000, Elsevier Inc.

Historical and Political Background of the Conflict
Migrations, Ethnic Cleansing, and Social Impact
Effects on Health
Effects on Mortality
Effects on Social Behavior and Violence
Conclusion

Glossary

Posttraumatic stress disorder (PTSD)

PTSD is a psychological disorder affecting individuals who have experienced or witnessed profoundly traumatic events, such as torture, murder, rape, or wartime combat, characterized by recurrent flashbacks of the traumatic event, nightmares, irritability, anxiety, fatigue, forgetfulness, and social withdrawal.

War is the multifaceted threat to human existence. It exposes people to the health-damaging psychological, physical, chemical, biological, and mechanical stressors. The war has a deep and devastating social and economic impact on the population involved, thus producing additional and lasting repression. It is too soon to make long-term estimates of the damages caused by war in the area of former Yugoslavia, but the data concerning increased social and health

impairments due to the war in Croatia are indicative for the whole area and show the level and the type of vulnerability when humans are exposed to the stress of war. War integrates the deprivation of basic existential needs and of all human rights and values; it eliminates emotional comfort; it causes irreversible material and kin losses, physical exhaustion, and psychological breakdowns; and it makes futile all everyday routines. The first and the only benign response to stress is the escape, but there is no possibility of running away from stress experiences in war – one stressor can only be substituted by another. This is why the full display of adaptation mechanisms takes place in subjects experiencing war stress: changes in behavior, hormonal impairments, neurological disorders, metabolic changes, genetic shifts, and all kinds of diseases. Stress-induced consequences, depending on the severity and the duration of the exposure to the stress experience, can be identified as PTSD, interrupted pregnancies, susceptibility to malignant diseases, cardiovascular insults, suppressed immune response, hormonal disorders, and all types of emotional instabilities or psychological impairments including suicide, homicide, and other types of violence.

Historical and Political Background of the Conflict

The federal republics of prewar Yugoslavia had national and cultural identities with distinct historical backgrounds. At times in the past they were sovereign states, and they have been trying to reestablish their independence ever since they lost it. Neither the first (monarchist) nor the second (communist) Yugoslavia was an expression of the people's will. When the

Berlin wall collapsed, the Soviet block was dismantled, the Soviet Union disintegrated, and the ruling communist party was defeated. The sweeping changes in eastern Europe provided an opportunity for the national republics constituting Yugoslavia to secede freely, as the Czech and Slovak republics have done. In response, the Yugoslav army under the command of the Serbian generals tried to annul the free and democratic decision by the electorate in four of the six Yugoslav republics to separate from Yugoslavia. The Serbs, inspired by the idea of Greater Serbia, were the only nation to support the continuation of Yugoslavia. The Serbian idea of turning Yugoslavia into a greater Serbia triggered the outburst of the latent but strong mistrust and intolerance between the Serbs, on one side, and all other nations, on the other. The Communist Party of Serbia, subsequently re-formed as the Socialist party, had control over the powerful Yugoslav army and turned into an aggressive Greater Serbian force that started to conquer and ethnically cleanse as much as possible of the territory of the republics whose majority of the population had voted for independence. The local people, if not ethnic Serbs, had to leave the occupied territories. Deprived of all goods and rights, they experienced the misery of homeless and jobless refugees. Those who tried to defend their homes were either massacred and killed or put into concentration camps, tortured, and raped. The process started in Croatia in 1991 and ended by North Atlantic Treaty Organization (NATO) intervention in Kosovo in 1999. Instead of democratization and prosperity, first the Croats, then the Bosnians, and at the end the Albanians experienced war, violence, fear, exile, and loss of kin. Macedonia and Montenegro experienced the menace of social and economic turbulence introduced by the hundreds of thousands of refugees that poured into their countries. Only the Slovenians succeeded in withstanding the austerities of Serbian aggression. At least 2.5 million victims (including almost the total population of Kosovo) were expelled from their homes and were, for varying periods, in exile or temporary housing.

Migrations, Ethnic Cleansing, and Social Impact

The first victims of the Serbian aggression were the Croats, who were removed from 30% of Croatian territory. That territory was not only occupied by the Yugoslav army and the local Serbs and ethnically cleansed of Croats; in addition, Serbs from other parts of Croatia or Bosnia were persuaded to resettle in abandoned Croatian homes in the territory cleansed for Greater Serbia. Such behavior escalated

national animosities as the aggression spread further and further. In Bosnia and Herzegovina, 70% of the territory was ethnically cleansed of Croats and Bosnians, and only the Dayton Agreement and NATO intervention stopped the war in Bosnia and forced the Serbs to restrict their control to the Serbian Republic (Republika Srpska), encompassing 49% of Bosnian territory.

In 1991, Bosnia had 4 364 600 inhabitants; 250 000 men were killed during the war and 35 000 women were raped. At the peak of the war, almost 2 million people were displaced. Even a decade after the end of the war in Bosnia, several hundred thousand people still had not returned to their homes.

The total population of Croatia (in 1991) was 4 760 000. It is estimated that 260 000 people left the country during the war. The war in Croatia killed 14 000 people, among whom 43.4% were civilians and 213 were children; 34 000 veterans, 9890 adult civilians, and 1142 children were injured; 7551 people were detainees of concentration camps and 23 000 children lost one or both parents. In addition, 389 433 civilians were displaced within the country (it is estimated that this number is 6% larger because of unregistered people). More than 100 000 houses were destroyed. Hospitals, schools, and many plants and farms were demolished, drastically raising the unemployed and retired percentages of the population and putting exhausting pressure on the remaining working population. There were 870 000 retired people, 260 000 jobless, and only 970 000 employed (compared to 1 650 000 employed before the war). The collapse of order and the rise of poverty, insecurity, and hopelessness were widely reported throughout the population.

Effects on Health

One of the major problems in the interpreting of the effects of war on the health of the population is the great extent of migration associated with the war. Croatia is the only country in the region where health parameters can be followed before, during, and after the war. Although migrations from Croatia and within Croatia were large, the number of displaced people never exceeded 12% of the total population and all the victims of migrations within Croatia are included in the official statistics.

The statistical parameters depict a very gruesome picture of the effects of war on the health of the population in Croatia. The rates of several diseases that are known to be associated with stress, such as diabetes, tumors, cardiovascular diseases, and gastric ulcers, increased sharply during and immediately after the war (Figures 1–2). Infectious diseases also increased

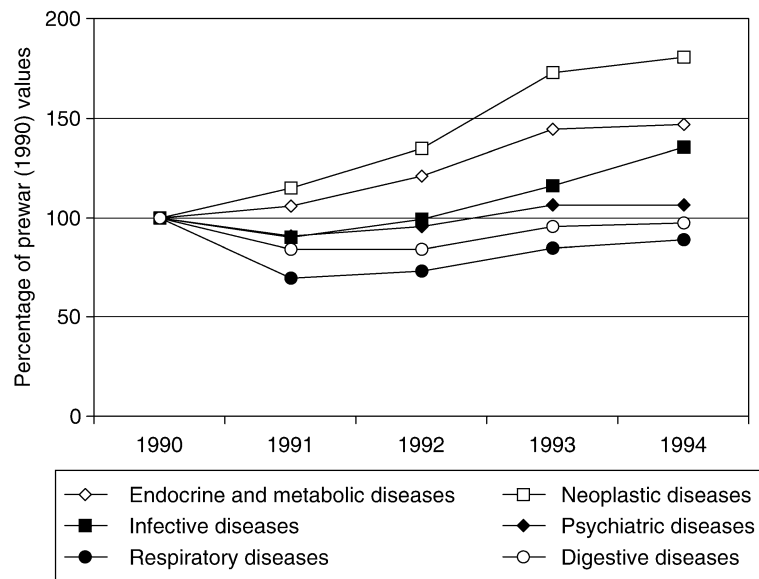


Figure 1 Number of cases of various disease groups registered in general hospitals in Croatia, 1990–1994. Data from Croatian Health Service Yearbook 1997, Zagreb, 1998.

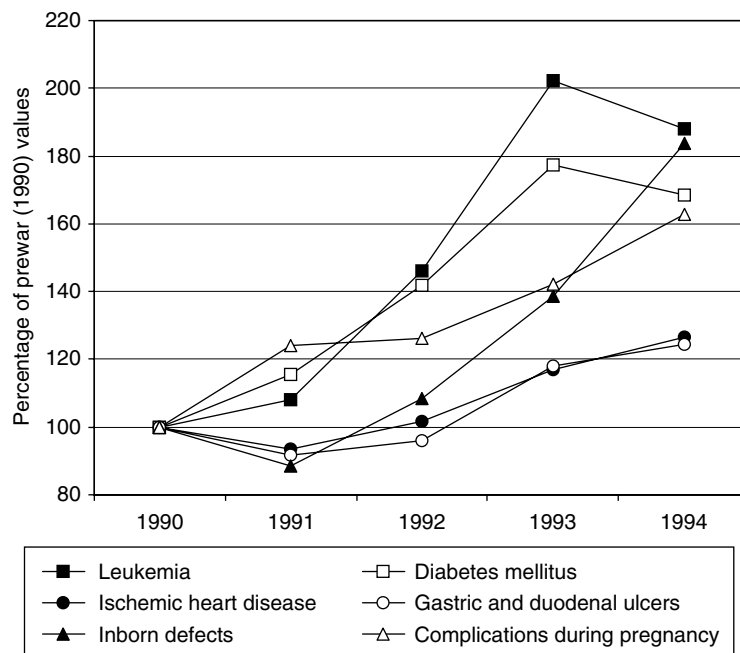


Figure 2 Number of cases of specific diseases registered in general hospitals in Croatia, 1990–1994. Only diseases with the highest increase of incidence are presented. Data from Croatian Health Service Yearbook 1997, Zagreb, 1998.

significantly, probably due to the stress-induced depression of the immune system. Most other diseases displayed a slight decrease in 1991–1992 (mirroring the temporary population decrease at the peak of the war in Croatia) and subsequently returned to the prewar values. The exception is rheumatoid arthritis, which stayed significantly decreased for several years and then increased rapidly to over 150% of prewar values. During this time, infant mortality continually declined,

indicating that there was no significant decrease in the general quality of the health-care system. Despite the preserved level of health care, the postwar health of the population significantly deteriorated due to stress and other effects of war compared to 1990. In 1994, Croatia had 60 000 more cases of infective diseases, 40 000 more people with diabetes mellitus, 32 000 more people with various neoplastic diseases, 30 000 more patients with heart diseases, 40 000 more people

with kidney diseases, and 24 000 additional patients with duodenal or gastric ulcers. Because the total population of Croatia was approximately 4.5 million, the addition of 200 000 ill people represented a substantial burden for an already overloaded health system.

Stress due to the war had a particularly strong impact on pregnancy and prenatal development. In addition to a decreased birthrate, as is usual in the time of hardship, there was a continuous increase in extrauterine pregnancies and other complications during pregnancy and delivery. In addition and completely unexpected, there was a sharp increase in various inborn defects (Figure 2).

Effects on Mortality

In addition to the decreased birthrate, the war diminished the population by increasing the mortality (Figure 3). Not taking into account the 14 000 people killed during the war, the most prominent immediate effect of war on the Croatian territory was the more than doubling of the rate of mortality from ischemic heart diseases, resulting in the unexpected deaths of more than 6000 people during the two war years 1991 and 1992. The prolonged effects of war, manifested in a ruined economy and prewar social structures, still kills people with an annual rate of approximately 4000. The increase in the ischemic heart disease alone has taken more human lives than were killed during the whole war in Croatia. Unfortunately, due to the massive migrations in Bosnia and Kosovo, it is not possible to determine the exact numbers for them, but if the percentage of people killed was similar to that in Croatia, several hundred thousands people probably died because of stress.

It is worth noting that, in addition to the stress imposed by the war, the transition from a centralized socialistic economy to a free-market economy

introduced stressful existential problems to a part of the population, which probably also contributed to the increased mortality from ischemic heart disease. One recent study in Lithuania reported that the break-up of the Soviet Union resulted in a quadruple increase in the mortality rate from ischemic heart diseases for some age groups. However, in other countries, such as Hungary and Poland, this pattern is not so prominent. The transition in Croatia is expected to be less stressful than in other socialist countries because even during socialism Croatia was open to the West and had accepted a few elements of the market economy. Thus, the war and its consequences remain the major factors influencing the increased mortality from ischemic heart disease in Croatia.

Effects on Social Behavior and Violence

Quantitative analysis of the psychological effects of war on the Croatian population is limited to the subjects who were/are seeking medical help. However, only a small number of individuals suffering from psychological disorders are under medical surveillance. There are more than 10 500 diagnosed cases of PTSD; many more are unregistered. Studies estimated that patients suffering from PTSD have up to a sevenfold increased incidence of suicides and fourfold increased risk of death from all external sources. Violence in general has increased, but it is very hard to know the exact figures because the health-monitoring statistics do not include violent behavior in their reports and official policy makers in Croatia refuse to acknowledge the incidence of violence and suicides to be disturbing. However, some trends are clearly visible. During and immediately after the war (1991–1992), there was a twofold increase in the level of violence, both as reported in police statistics and in hospital admissions. Although, according to

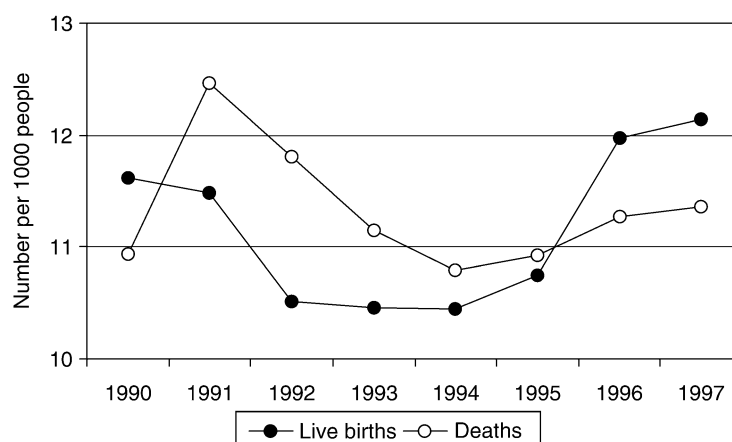


Figure 3 Number of live births and deaths per 1000 people, 1990–1997. Data from Statistical Yearbook 1998, Republic of Croatia Central Bureau of Statistics, 1998.

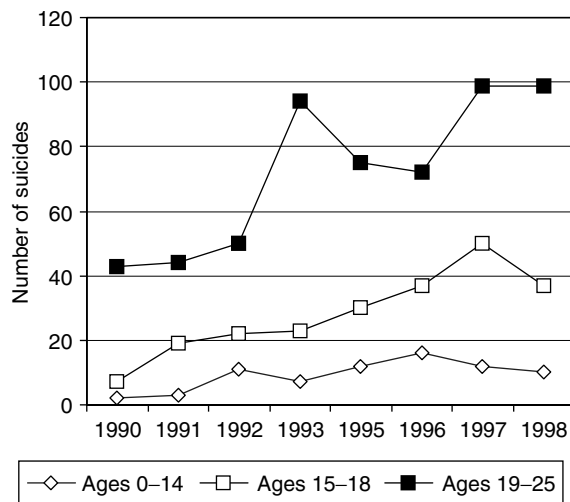


Figure 4 Number of suicides and attempted suicides in different age groups in Croatia, 1990–1998. Data from official statistics of the Ministry of Interior of the Republic of Croatia.

the same sources, the reported homicides and violent crime returned to prewar values in subsequent years, there is a disquieting increase in suicides and other self-inflicted injuries in young people. The number of young people committing suicide is continuously rising, and compared to 1990, in 1997 this rate is doubled or quadrupled, depending on the age group (Figure 4). In the last four years, primary and secondary schools report a rapidly growing number of hyperactive and aggressive children who need special care in the classroom. The incidence of depressive symptoms and insomnia is reported to be great in the elderly and in veterans. Drug abuse is disturbingly frequent compared to the prewar statistics, but the number of treated addicts has remained constant. It is correct to attribute these changes in behavior to the stress imposed by the war, and it is interesting how the various target groups (children, the elderly, veterans, etc.) within the population have responded differently to the stress.

Conclusion

The majority of the population in each of the states originating from the former Yugoslavia has experienced war stress. Various health and social disorders evolved from that experience. Although the data presenting the consequences of the war stress are incomplete and although they are, because of the availability, reduced to Croatia only, they provide unique evidence of human vulnerability when living through war.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Disasters and Mass Violence, Public, Effects of; Gulf War Syndrome, Psychological and Chemical Stressors; Nuclear Warfare, Threat of; Persian Gulf War, Stress Effects of; Posttraumatic Stress Disorder, Delayed; Vietnam Veterans, Postwar Experiences and Health Outcomes; War-Related Posttraumatic Stress Disorder, Treatment of; Posttraumatic Stress Disorder – Clinical.

Further Reading

- Babic-Banaszak, A., Kovacic, L., Kovacevic, L., et al. (2002). Impact of war on health related quality of life in Croatia: population study. *Croatian Medical Journal* 43, 396–402.
- Black, P. H. and Garbutt, L. D. (2002). Stress, inflammation and cardiovascular disease. *Journal of Psychosomatic Research* 52, 1–23.
- Breslau, N., Kessler, R. C., Chilcoat, H. D., et al. (1998). Trauma and posttraumatic stress disorder in the community: the 1996 Detroit Area Survey of Trauma. *Archives of General Psychiatry* 55, 626–632.
- Bullman, T. A. and Kang, H. K. (1994). Posttraumatic stress disorder and the risk of traumatic deaths among Vietnam veterans. *Journal of Nervous and Mental Disease* 182, 604–610.
- Dekaris, D., Sabioncello, A., Mazuran, R., et al. (1993). Multiple changes of immunologic parameters in prisoners of war: assessments after release from a camp in Manjaca, Bosnia. *Journal of the American Medical Association* 270, 595–599.
- Kocijan-Hercigonja, D., Sabioncello, A., Rijavec, M., et al. (1996). Psychological condition hormone levels in war trauma. *Journal of Psychiatric Research* 30, 391–399.
- Lauc, G., Zvonar, K., Vukšić-Mihaljević, Ž., et al. (2004). Post-awakening changes in salivary cortisol in veterans with and without PTSD. *Stress Health* 20, 99–102.
- Lukenda, J., Kolaric, B., Kolcic, I., et al. (2005). Cardiovascular diseases in Croatia and other transitional countries: comparative study of publications, clinical interventions, and burden of disease. *Croatian Medical Journal* 46, 865–874.
- Markotic, A. (2002). Balkan syndrome. *Lancet* 359, 166.
- Marsland, A. L., Bachen, E. A., Cohen, S., et al. (2002). Stress, immune reactivity and susceptibility to infectious disease. *Physiology & Behavior* 77, 711–716.
- Reiche, E. M., Nunes, S. O. and Morimoto, H. K. (2004). Stress, depression, the immune system, and cancer. *Lancet Oncology* 5, 617–625.
- Sabioncello, A., Kocijan-Hercigonja, D., Rabatic, S., et al. (2000). Immune, endocrine, and psychological responses in civilians displaced by war. *Psychosomatic Medicine* 62, 502–508.
- Sapolsky, R. M. (1996). Why stress is bad for your brain. *Science* 273, 749–750.

War, Suicide and Sacrifice

V Hazboun

Guidance and Training Center for the Child and Family,
Bethlehem, Palestine

© 2007 Elsevier Inc. All rights reserved.

Background

Theory

Clinical Research

Clinical Experience

Research

Stress Management in War

Glossary

<i>Intifada</i>	Palestinian uprising against the occupying forces.
<i>Nakbeh</i>	The disaster (Arabic).
<i>Sacrifice</i>	The slaughter of an animal or a person or the surrender of a possession as offering to a deity, made holy by religious association.
<i>Suicide</i>	The act of a person intentionally killing him- or herself.
<i>War</i>	A quarrel between nations conducted by force, a state of open hostility.

Background

It is the hope of the author to present, from a clinical perspective, the experience of war, suicide, and sacrifice. For centuries, people have launched wars with other nations and peoples to pursue ideals, goals, or possessions. Others have fought wars in self-defense and in anticipated concern about their safety. Others have been invaded and occupied and have resisted the occupation through warlike tactics. People have committed suicide, for a purpose or for no purpose, because of an ideal or an illness. People have sacrificed so that others can benefit; this altruism has been a characteristic that elevates humans from the selfish and decadent position that they can become fixated on.

In the Middle East, the Palestinian-Israeli conflict goes back several decades, and it is impossible to address all the issues here. However, a summarized orientation to the problem is necessary to understand the trauma and stress sustained by this population. The following years are significant:

- 1948 – The Nakbeh (the disaster, in Arabic) is the word used by the Palestinian population to refer to

the partitioning of Palestine and the massive Palestinian refugee exodus and displacement. The Israelis refer to this same period as their independence day.

- 1967 – The 7-Day War and the occupation of the rest of Palestine by the Israeli armed forces.
- 1989 – The first Intifada, the Palestinian people's uprising against the occupying forces.
- 1994 – After the Oslo talks, the declaration of the Palestinian Authority.
- 2000 – The second Intifada, a rekindling of the first Intifada after major disappointments and tragedies that continued in the area despite the Palestinian Authority's presence.
- 2004 – The building of the separation wall, still in progress in 2006.

Theory

The psychoanalytic model is a scientific perspective that has been used by the author in her 25 years of clinical experience as a psychiatrist dealing with traumatized patients. The author's clinical experience has allowed her the opportunity to compare cases from economically impoverished areas such as the ghetto of Los Angeles and the refugee camps in Palestine, along with more economically viable areas such as Hollywood and Jerusalem. War here is considered to be a state of open hostility possibly causing trauma to individuals and the masses alike. The individual or national reaction to such hostility remains a very personal experience, even from a national perspective. Some people are traumatized from such hostility; others are rendered stronger and more resilient, determined, and creative. Both the aggressors and those experiencing aggression should reflect on this and perhaps learn to refrain from aggressive behavior; aggressors give those experiencing aggression a chance to become noble and most creative, resilient, and determined, and if those experiencing aggression become aggressive in return they lose their values and become the same as the original aggressors.

Trauma is an experience that can be caused from within or from the outside. That means that people can experience trauma internally without anyone contributing to this phenomena. This situation of internal conflict is well known to psychiatrists, who are constantly challenged with such circumstances. External traumatizing events and their impact on people's psychological well-being have been very clearly and concisely defined and presented in the diagnostic criteria for posttraumatic stress disorder

(PTSD) in the *Diagnostic and Statistical Manual of Mental Disorders* (4th edn.; DSM-IV) and subsequent articles on the subject.

The Middle East conflict seems to tie together many important aspects of our civilization: politics and the search for democracy; religion or the lack of it, and its abuse through fanaticism; justice and its continued absence, leading to intolerable injustice; violence and aggression; gender difference; equality and the lack of it; pride and the feeling of superiority and being chosen, leading to prejudice and the dehumanizing of the other; occupation that is viewed as liberation for others; terrorizing others and the feeling of being terrorized at the same time; and war and the yearning for peace. Here I present a psychoanalytic look at the Middle East conflict, hoping that it sheds light on the 50-year conflict and encourages the different parties to engage in deep reflection, seeking the answers in an arena that has not been explored before. I believe that human nature and people's unconscious processes and defenses are interfering with the capacity of the people involved to seek and realize peace, whereas a scientific understanding of these processes might alter the escalating cycles of violence and convince people to adopt a different approach, one that will lead to a lasting peace. This then could serve as a model for other areas of escalating conflict and confrontation.

Let us first examine the basic ideas behind the understanding that led to the conclusions presented here. The work of Sigmund Freud was at the basis of the understanding that was later developed by Melanie Klein that the power of the unconscious and its impact on human existence needs to be addressed. To become conscious of unconscious processes liberates people, who can then either choose to be led by their neurotic thoughts or choose to venture into a life of free choice, risk taking, and responsible action led by thought and reflection rather than by impulse and magical thinking. Freud taught us about the defense mechanisms in the human mind and attitudes, namely the early defenses of splitting and projecting. As evidenced in many human relationships, and certainly in politics, the good and the bad reflect this very notion of splitting, seeing the self as good and the other as the enemy on whom all kinds of evil projections can be accommodated. Melanie Klein further developed these thoughts with the idea of projective identification. She suggested that when magical thinking is used to develop a person's wishful thoughts (in action or in thought) it renders the person responsible for the possible damage that occurs to other people, regardless of the real cause of this damage. The sense of responsibility or guilt may lead to self-punitive behavior or feelings, but it is the feeling

of entrapment that can cause more damage, tarnishing the person's life; that is, the person may feel that, if he or she magically caused irreversible damage in other people's lives, then this magical fate is the only possible fate.

Clinical Research

The work presented here is based on many years of clinical observations.

- 1984 – The author observed and analyzed the drawings of 50 children in three refugee camps in Palestine. The drawings were compared to those of children in the same age group in the Los Angeles County Hospital's children's inpatient unit. The drawings of the children living in the Palestinian refugee camps were more often about healthy community activities scenes such as picking olives in the fields, weddings, and home scenes; the inpatient children's drawings in Los Angeles were more morbid, showing bars, blood, and death scenes and issues.

- 1991–1994 – As mental health supervisor for the United Nations Relief and Works Agency (UNRWA), the author gathered clinical evaluations. The reports indicated that, of 1500 new cases seen yearly, 450 children were diagnosed as having pervasive developmental disorder, 350 as having depression, and approximately 250 as having PTSD.

- 1992 – During the first Intifada, the author visited the same three refugee camps as in the 1984 study and examined the drawings of 50 children in each camp. These drawings had an extremely different tone; they were overwhelming in their gory depictions of pain, blood, barbed wire, prison and death, violence, and injustice.

- 1996 – Mental health workers, social workers, and psychologists randomly administered 360 questionnaires in the West Bank. These questionnaires were translations from English of DSM-IV criteria. The diagnostic criteria for three disorders was analyzed and used: PTSD, depression, and pervasive developmental disorder. The results indicated an alarming 35%, 40%, and 45% incidence, respectively. These results will be reassessed soon to evaluate the impact of the second Intifada and the continued unrelenting trauma that has haunted the area.

Clinical Experience

It has been the experience of the author that, although war is traumatizing, it does not necessarily have a paralyzing effect on the population. Aside from the obvious trauma of the loss of people's lives and property, heritage and development, health and hope, it is the observation of the author after 20 years of living

among and serving this population that the reality of becoming refugees, being displaced, losing loved ones, and tolerating a great deal of frustration and humiliation has not turned the masses into desperate people who let their impulses and instincts control them; instead, people who recognize the wisdom of maintaining educational pursuits and developing future generations who will honor their traditions seem to be the norm. The fact that some people resort to the idealization of martyrdom, death, and violent suicide/homicide remains to be understood and investigated. Next I present the results of my clinical experience and research done to try to analyze and understand this phenomena.

As part of the initial assessment of the children at the Guidance and Training Center for the Child and Family (GTC) in Bethlehem, the author asks the children to express three wishes. In the mid-1980s, the children expressed many very creative wishes expressing a great many ambitions and objectives in life. In the mid-1990s, 10 years later, the majority of the children did not express any wishes; they remained silent. After some psychotherapeutic intervention, if they expressed any wishes, the wishes all could be summarized as a wish to become a *shaheed*, a martyr. This inability to express wishes followed by a wish to become a martyr indicates a paralysis of the mind and creativity and a total embracing of a death wish. These children can relinquish these thoughts after a solid psychotherapeutic intervention and a facilitation of their reintegration into the school system.

Clinical Vignettes

Patient 1 A young woman, a would-be suicide bomber, spoke on local public TV about her previous intentions to blow herself up for what she thought was a nationalistic pursuit and about how, feeling uncomfortable with her plans, she sought psychotherapy. Then, she spoke about her change of heart after her psychotherapeutic experience at our center. However, she reported that she had been unable to convince her brother to come in for treatment and that he ended up as a suicide bomber, killing himself and others.

Patient 2 Another patient, a 7-year-old boy, was brought in for treatment because he was sleepwalking at night; this behavior endangered his life, especially because the city was under curfew during most of the period when the child exhibited this behavior. It turned out that this child had been arrested and taken for interrogation for throwing stones at the occupying forces. During the arrest, his grandmother tried to stop the soldiers from taking the child; she was pushed and she fell, hurting her head and landing

in a pool of blood. The boy thought she was dead as he was taken for his interrogation. But on his release a few hours later, he saw that his grandmother was well and not dead as he had feared. Following this episode, the child started to exhibit a phobia about leaving the home and refused to go to school; he eventually presented with sleepwalking, which led to his being brought in for treatment.

During his therapy, when the boy was asked about dreams, he reported having a repetitive dream of someone kidnapping him and putting him in dangerous situations. When the author said that she might know who the kidnapper was, the child was surprised and interested in the answer. The author told him that she thought he was both the kidnapper and the kidnapped because he was the author of the dream and therefore was trying to tell us something about himself – that he kidnapped himself at night while asleep and put himself in dangerous situations. The dream also was residual trauma from the arrest and interrogation that he sustained, but it was interpreted at a deeper level because the child was receptive to this level of interpretation, and he eventually benefited from the interpretation. The sleepwalking stopped, and the child was able to go to school and did not continue to repeat the trauma that he had sustained by having this disturbing dream and exhibiting dangerous behavior that could have cost him his life. The child felt guilty about his grandmother's injury and felt he had caused it by his behavior, which led to the army's intrusion into their home. He felt identified with this projection and trauma and wanted to feel in control through this repetitive dream of the kidnapper. He felt guilty and entrapped by the fate that he thought he caused his grandmother to have.

Patient 3 Another patient, a breast-feeding mother, told the author, "This infant I can feed this way, but the others at home, how will I feed them? If someone were to give me a bomb, I would go to the checkpoint and blow myself up. This way, my children at least will be proud of their mother, the martyr." The author took this into consideration as she continued to treat this patient and offered her an income-generating opportunity at the clinic doing some embroidery. This woman is now taking care of her family and running an income-generating embroidery project of her own, helping other women sustain their livelihoods and those of their families.

Research

A brief study was carried out at the center to assess the impact of the separating wall being built by the Israeli government on the population living in the West Bank,

asking people's views of this new cement barrier separating them from the families that had been living next door to them prior to this construction. The study also assessed people's views vis-à-vis the Holocaust.

The questionnaire was given to 30 Palestinian professionals; the results are shown in Table 1. Twenty years ago, the results for the last four questions were as listed in Table 2. When the present-day results are compared to the assessments made 20 years ago, it becomes apparent that this construction has decreased

Table 1 Assessment of the views of Palestinian professionals vis-à-vis the Holocaust and the separation wall being built by the Israeli government^a

Question	Yes	No
1. Wall will destroy life in Palestine	28	2
2. Wall will help Palestinians get a Palestinian State	7	23
3. Wall will eliminate Palestinians by Israel	25	5
4. Israelis are doing to Palestinians what Hitler did to them	23	7
5. Consider Jews living through a massacre in Germany	28	2
6. Feel sad about Holocaust	22	8
7. Feel happy about the Holocaust	1	29

^aN = 30, 20 males and 10 females.

Table 2 Assessment of the views of Palestinian professionals (questions 4–7) 20 years ago^a

Question	Yes	No
4. Israelis are doing to Palestinians what Hitler did to them	20	10
5. Consider Jews living through a massacre in Germany	29	1
6. Feel sad about Holocaust	29	1
7. Feel happy about the Holocaust	1	29

^aN = 30.

the compassion people feel toward the other side and increased their feelings of being oppressed and persecuted.

After the construction of the separating wall and the population's increasing feelings of frustration and desperation, and the claim by the other side that the construction of the separating wall would provide security for the Israelis, the author investigated the opinion of the community about suicide bombers. The idea for this questionnaire came from the observation, first with children and then with adults, that those who came to the center for psychiatric and psychological treatment expressed wishes or plans to become suicide bombers. After these patients received psychological interventions at the center, which either facilitated their being mainstreamed back into the school system or into the workforce, these thoughts of suicidal/homicidal plans disappeared and were replaced with hope and gratitude toward the people who helped them find hope for the present and future; they relinquished their death wish and embraced a life wish.

The questionnaire was given to 91 mental health workers; the results are shown in Table 3. They indicate that a high percentage of the population honor and glorify this sacrifice and do not view it as reflection of emotional difficulties, seeing suicide bombing as the patriotic sacrifice of a hero, done in the name of God, and not as an expression of suicidal thoughts, depression, or psychological injury. The results show that suicide bombers are seen as desiring revenge, being aware of what they are doing, and not feeling guilty about the pain of other people. These ideas are not the same as those expressed by the population that comes for consultations at the guidance and training center, as presented in the earlier material.

Further research is planned that may help us further understand the state of mind of the populations in this area. This includes a study comparing trauma

Table 3 Assessment of views about martyrdom (%)^a

Do you think that the people who conduct martyrdom operations:	Yes	No	No answer
1. Desire to commit suicide?	17.6	76.9	5.5
2. Do a patriotic sacrifice?	84.6	12.1	3.3
3. Do a sacrifice in the name of God?	83.5	13.2	3.3
4. Do it due to their difficulties in life?	60.4	35.2	4.4
5. Suffer from depression?	49.5	44.0	6.6
6. Are psychologically hurt?	64.8	28.6	6.6
7. Desire revenge?	84.6	9.9	5.5
8. Are aware of what they are going to do?	45.1	49.5	5.5
9. Feel guilty about the pains of other people?	34.1	60.4	5.5
10. Are heroes?	46.2	49.5	4.4

^aN = 91 mental health workers.

in Palestinian and Israeli populations and the capacity to forgive and a study comparing the mental health disorders observed in Palestinians living in West Bank to disorders in Palestinians living in Lebanon.

Stress Management in War

Although it is well known that suicide attempts, psychotic episodes, and admissions to psychiatric hospitals increase after a war is over and there is peace, it is the author's observation that the mental health demands in Palestine are increasing and becoming overwhelming. It is with this in mind that the interventions discussed here are being provided. The prediction that things will become even worse when there is peace in this area makes the author concerned about meeting these increasing mental health needs, and the author hopes that the prediction is wrong. The mental health services are not sufficient at the present time, and should things continue to deteriorate, the outcome is of concern to many health workers.

The use of psychoanalytic interventions and understanding, behavioral and relaxation techniques, special education therapy, individual (talk or play) and group therapy, medications, and other modalities to reduce stress and trauma in the population has become common. A thorough initial evaluation and treatment plan are necessary for each patient who consults a treatment center, and the provision of mental health services is urgent and necessary at this time.

That a majority of the population does not pursue psychotherapeutic assistance and yet seems to manage and maintain its existence and sanity implies that human nature is capable of self-healing and reasoning, of developing the necessary wisdom through people's own thinking and their culture and religion to pursue healthy avenues in their decision making, taming their aggressive impulses and discovering on their own that in this way creativity is allowed to develop and flourish. The few individuals pursuing violent resistance seem to hurt the cause and the people themselves more than help them. This leaves the author concerned about the repercussions of such attitudes; two wrongs never made a right. This is explained by Klein's projective identification defense, as previously elaborated on. The fact that war has contributed to the confusion of suicide with sacrifice should not cause mental health workers to confuse the issues; they should always try to sort things out in a clear and ethical way, trying to provide patients with inner harmony, helping them think more clearly and wisely through the psychotherapeutic intervention.

It is important to also mention that in times of war and political chaos and injustice it is the moral obligation of the mental health workers to maintain a standard of ethical service to the community and to try to attend to the needs of the masses by training others who can provide triage and quality interventions and by using the media (magazines, newspapers, radio, and television) to present mental health issues and prevention-measures training. All these different interventions have proved to be worthwhile efforts during the author's years of experience in the Middle East. Along with attending to the needs of the masses from a mental health perspective, it is also important to keep in mind that help along the lines of mainstreaming kids back into the school system is essential and that providing income-generating projects for the population may play a role in preventing disasters.

Working in a vacuum is not a creative way to intervene, and the involvement of professionals who are eclectic, versatile, creative, and passionate about their work makes the difference between touching people and helping them versus observing, reporting, and being afraid to get involved in the struggle to find and spread inner harmony.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Gulf War Syndrome, Psychological and Chemical Stressors; Persian Gulf War, Stress Effects of; Posttraumatic Stress Disorder in Children; Posttraumatic Stress Disorder, Delayed; Posttraumatic Stress Disorder, Neurobiology of; Vietnam Veterans, Postwar Experiences and Health Outcomes; War Stress in the Former Yugoslavia; War-Related Posttraumatic Stress Disorder, Treatment of; Posttraumatic Stress Disorder – Clinical; Posttraumatic Stress Disorder – Neurobiological basis for.

Further Reading

- Adelson, E. (1975). *Sexuality and psychoanalysis*. New York: Brunner, Mazel Inc.
- Freud, S. (1993). *The interpretation of dreams*. U.S.A: Basic Books.
- Goldstein, J., Freud, A., Solnit, A. J., et al. (1986). *In the best interest of the child: the least detrimental alternative*. New York: The Free Press.
- Hazboun, V. (2002). Autobiographical chapter. In: Armbruster, D. (ed.) *Tears in the Holy Land*. Wilsonville, OR: Bookpartners.
- Hazboun, V. (2003). A psychotherapeutic view of violence. In: *Special issue on two traumatized societies. Palestine-Israel Journal of Politics, Economics and Culture* 10(4) [online].
- Hazboun, V. (2003). A psychotherapeutic view of violence and its alternatives. *RAHAT Medical Journal* 1(5), 17–23.

- Hazboun, V. (2004). La testimonianza di una psichiatra sulle conseguenze del lungo conflitto Palestinese. *Il Regno* 49(945).
- Hazboun, V. (2004). Psychotherapeutic interventions with Palestinian children and their families. *Dynamic Psychiatry* 204–205, 69–77.
- Hazboun, V. (2005). Clinical experience in the field of war and PTSD. *Trauma e Guerra, Rivista Sperimentale di Freniatria* 129(1), 115–124.
- Hazboun, V. (2005). L'intifada e la condizione dell'infanzia in Palestina: Nel Suo Nome, conflitti, riconoscimento, convivenza delle religioni. *Il Regno* 141–145.
- Hazboun, V. and Ziskoven, M. (2002). Ein schmerzhafter Frieden ist besser als die Qualen des Krieges. *Forum der Kinder-und Jugendpsychiatrie und Psychotherapie* 12(2).
- Hazboun, V. and Ziskoven, M. (2002). Ein schmerzhafter Frieden ist besser als die Qualen des Krieges [A painful peace is better than agonies of war]. *Psychosoziale Arbeit in Bethlehem Gesundheitsmagazin*. Dr. Med. Mabuse, 136, Maerz.
- Klein, M. (1984). *Love, guilt and reparation*. New York: The Free Press.
- Lewis, M. (1982). *Clinical aspects of childhood development*. Philadelphia: Lea & Febiger.
- Mason, A. (1983). The suffocating super ego. In: Grotstein, J. (ed.) *Dare I disturb the universe*. Caesura Press.
- Segal, H. (1978). *Introduction to the work of Melanie Klein*. Hogarth Press.
- Victorian Foundation for Survivors of Torture (1996). *A guide to working with young people who are refugees*. Victoria, Australia: Victorian Foundation for Survivors of Torture.
- Wener, C. (1982). *Psychopathology from infancy through adolescence*. New York: Random House.

War-Related Posttraumatic Stress Disorder, Treatment of

L B Slone and M J Friedman

National Center for PTSD, U.S. Department of Veterans Affairs and Dartmouth Medical School, White River Junction, VT, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R Rosenheck and A Fontana, volume 3, pp 672–677, © 2000, Elsevier Inc.

Psychotherapeutic

Of, relating to, or used in treatment of mental or emotional disorder or maladjustment involving verbal communication.

Reservists

Military personnel who serve part-time unless activated. Includes both National Guard and Reserve components.

Introduction
Stressors
Assessment and Diagnosis
Treatments
Summary

Glossary

<i>Comorbid</i>	Occurring at the same time as another medical condition.
<i>Imaginal exposure</i>	Exposure involving imagination, images, or imagery of the traumatic event.
<i>In vivo exposure</i>	Exposure to a real-life situation.
<i>Pharmacological</i>	Having to do with drugs.
<i>Psychometric</i>	Quantitative assessment of psychological trends and measures.

Introduction

Some service members handle their transition from the war zone to the home front in a matter of weeks. Others require more time, and perhaps assistance. A significant minority fall short and develop posttraumatic stress disorder (PTSD), depression, alcoholism, some other psychiatric problem, or functional impairment. Although the present focus is PTSD, these other potential postdeployment problems should not be overlooked. The treatment of PTSD in veterans is similar in many ways to the treatment of PTSD in the civilian population, with some special considerations. Although there are universal aspects to any combat experience, every war is different, thereby posing unique challenges to the practitioner. For example, 50% of American troops that served in Iraq and have left active duty are members of the National Guard or Military Reserve rather than active duty combatants. This has implications for the identification and treatment of PTSD in this cohort.

Stressors

Traditional severe combat stress involves situations in which a service member experiences comrades being killed or injured; having personally killed enemy combatants and, possibly, innocent bystanders; uncontrollable and unpredictable life-threatening attacks; and more. In addition to direct combat stress, a variety of other war zone stressors are experienced, such as feeling helpless to alter the course of potentially lethal events, observing or handling the remains of civilians and soldiers, and observing devastated communities produced by combat. Military sexual trauma (MST) is now known to be highly prevalent, affecting both women and men in combat. An unprecedented number of military personnel are surviving serious injuries due to dramatic advances in medical technology, and being wounded in battle is associated with the highest PTSD rates. This may be compounded by the impact of life-changing severe injuries such as traumatic brain injury (TBI), traumatic amputations, spinal cord injuries, and blindness.

A common denominator for many returnees is the sustained hypervigilance and persistent anxiety about potential threats at any time or place, which often leads to difficulty transitioning from the theater back to civilian life. Clinicians faced with patients who have had a difficult re-entry to the home front need to be aware of this reintegration process. They must consider the likelihood that postdeployment difficulty for a particular patient may be part of an otherwise normal readjustment trajectory or could be manifestations of a clinically significant problem such as PTSD.

Assessment and Diagnosis

A routine assessment should begin with a few questions about the patient's exposure to war zone trauma and other traumatic experiences. Once the presence of traumatic exposure is established, the following step is to screen for PTSD. The National Center for PTSD has developed a four-item (yes/no) screen with excellent psychometric properties (Primary Care PTSD Screen [PC-PTSD]).

Designed for use by primary care practitioners, the PC-PTSD consists of four questions, one from each PTSD symptom cluster, asked in reference to the past month: (1) Have you had nightmares or thought about the trauma when you did not wish to? (re-experiencing), (2) Have you tried to avoid thoughts and situations that remind you about the event? (avoidance), (3) Are you easily startled or frequently on guard? (hyperarousal), and (4) Have you felt numb or detached from other people or things? (numbing). A patient that endorses three of the four items should

receive a more thorough assessment for PTSD. Recent research with the PC-PTSD shows that it correctly positively identifies those that have PTSD 78% of the time (sensitivity) and correctly negatively identifies those that do not have PTSD 87% of the time (specificity).

Given sufficient evidence from screening that the patient may have PTSD, a more formal assessment is warranted. This can be accomplished through a comprehensive diagnostic interview in which each of the 17 PTSD symptoms is assessed, by use of a structured clinical interview such as the Clinician Administered PTSD Scale (CAPS) or using well-validated self-report questionnaires such as the PTSD Checklist (PCL), the PTSD Diagnostic Scale (PDS), the Davidson Self-Rating PTSD Scale, or others.

There are a few things that clinicians need to keep in mind when dealing with veteran patients. Clinicians should be aware that service members often fear being stigmatized if they disclose any sign of psychiatric problem, including PTSD. Furthermore, those that are most symptomatic are more likely to feel stigmatized and therefore less likely to seek treatment. It is also necessary to recognize that although social support is a powerful protective factor against developing PTSD, it is much less available to National Guard and military reservists and their families than to full-time military members. Whereas the latter may inhabit military bases where many military families live close together, undergoing the same deployment stressors, reservists, and especially their families, are neither embedded in a full-time military culture nor surrounded by others who share their plight. This may explain why reserve troops in the Persian Gulf War had higher prevalence of PTSD and depression than active duty personnel.

Treatments

Effective evidence-based psychotherapeutic and pharmacological treatments for PTSD are available. Detailed discussion of the empirical treatment literature can be found in clinical practice guidelines.

Psychotherapeutic Treatments

Clinical guidelines published to date designate cognitive-behavioral therapy (CBT), which includes cognitive therapy, prolonged exposure, and cognitive processing therapy, as the treatment of choice for PTSD. CBT techniques are designed to alter the conditioned fear and cognitive distortions associated with PTSD. Prolonged exposure essentially involves clinicians asking patients to construct narratives about intolerable traumatic events they can recall.

With repeated, therapist-guided exposure to these traumatic memories (imaginal or *in vivo*), patients experience progressive reduction in distress and thereby achieve clinical remission. Cognitive therapy and cognitive processing therapy focus on changing the maladaptive automatic thoughts associated with PTSD. Cognitive restructuring is a technique through which the therapist challenges distorted beliefs (perceiving the world as dangerous, seeing oneself as powerless or inadequate, or feeling guilty for outcomes that were outside of one's control), thereby enabling patients to overcome intolerable trauma-related emotions. Exposure and cognitive therapies have performed very well and with equal efficacy.

Eye movement desensitization and reprocessing (EMDR) is a technique that involves patients being instructed to imagine a painful traumatic memory and associated negative cognitions while visually focusing on the rapid movement of the clinician's finger. Although there are now many studies showing that such eye movements are not needed for EMDR to work, evidence-based practice guidelines indicate that EMDR is an effective treatment despite the lack of a defined mechanism of action.

In practice, both exposure and cognitive therapies are considered first-line treatments for PTSD. In exposure therapy, a clear memory of the traumatic event must be accessible. For cognitive therapy, the focus is primarily on the automatic, erroneous cognitions and behaviors that are associated with traumatic memories. In practice, exposure therapy has successfully addressed erroneous cognitions, and cognitive therapy has successfully addressed fear-based avoidant behavior. Therefore, at this time, the choice of CBT therapy for PTSD is up to the clinician. Given the scarcity of trained CBT therapists, neither option of evidence-based treatment may be available.

Pharmacological Treatments

A number of medications have been tested as treatment for PTSD patients. Selective serotonin reuptake inhibitors (SSRIs) are currently the treatment of choice, and the U.S. Food and Drug Administration has approved two, sertraline and paroxetine, for the treatment of PTSD. What has been especially exciting about the SSRI results is that these drugs appear to be broad spectrum agents that work to reduce all the symptom clusters of PTSD. Successful randomized clinical trials have also been conducted with the SSRI fluoxetine, the alpha-1 adrenergic antagonist, prazosin, tricyclic antidepressants, monoamine oxidase inhibitors, and venlafaxine. Because the psychobiology of PTSD is complex, other medications such as antiadrenergic agents, anticonvulsants, and a

variety of agents currently under investigation may prove as effective as SSRIs. Finally, randomized trials of benzodiazepines have yielded negative results, so this class of medication cannot be recommended for PTSD treatment. When choosing medications and other elements of treatment, keep in mind that PTSD is highly comorbid with other psychiatric problems.

Summary

Many important questions about treatment have yet to be tested systematically. These include how to select treatment, define realistic goals, combine various treatments, and approach complex comorbid conditions; how long to stick with a specific treatment; and when to acknowledge clinical failure. As new psychotherapeutic and pharmacological treatments are developed and empirically tested, these essential questions should also be addressed. In the meantime, it is essential that clinicians select evidence-based treatments whenever possible.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Cognitive Behavioral Therapy; Vietnam Veterans, Postwar Experiences and Health Outcomes; Selective Serotonin Reuptake Inhibitors (SSRIs).

Further Reading

- American Psychiatric Association (2000). Practice guidelines for the treatment of acute stress and posttraumatic stress disorder. *American Journal of Psychiatry* **161**, 1–31.
- Asnis, G. M., Kohn, S. R., Henderson, M., et al. (2004). SSRIs versus non-SSRIs in post-traumatic stress disorder: an update with recommendations. *Drugs* **64**, 383–404.
- Brewin, C. R., Andrews, B. and Valentine, J. D. (2000). Meta-analysis of risk factors for posttraumatic stress disorder in trauma-exposed adults. *Journal of Consulting and Clinical Psychology* **68**, 748–766.
- Foa, E. B., Keane, T. M. and Friedman, M. J. (2000). *Effective treatments for PTSD: practice guidelines from the International Society of Traumatic Stress Studies*. New York: Guilford Press.
- Friedman, M. J. (2002). Future pharmacotherapy for PTSD: prevention and treatment. *Psychiatric Clinics of North America* **25**, 427–441.
- Friedman, M. J. (2005). Veterans' mental health in the wake of war. *New England Journal of Medicine* **352**, 1287–1290.
- Friedman, M. J. (2005). Acknowledging the psychiatric cost of war. *New England Journal of Medicine* **351**, 75–77.
- Hoge, C. W., Castro, C. A., Messer, S. C., et al. (2004). Combat duty in Iraq and Afghanistan, mental health problems, and barriers to care. *New England Journal of Medicine* **351**, 13–22.

- Kulka, R. A., Schlenger, W. E., Fairbank, J. A., et al. (1990). *Trauma and the Vietnam War generation: report of findings from the National Vietnam Veterans Readjustment Study*. New York: Brunner/Mazel.
- Murdoch, M., Polusny, M. A., Hodges, J., et al. (2004). Prevalence of in-service and post-service sexual assault among combat and noncombat veterans applying for Department of Veterans Affairs posttraumatic stress disorder disability benefits. *Military Medicine* **169**, 392–395.
- National Collaborating Centre for Mental Health (2005). *Post-traumatic stress disorder: the management of PTSD in adults and children in primary and secondary care*. London: Gaskell and the British Psychological Society.
- Prins, A., Ouimette, P. C., Kimerling, R., et al. (2004). The primary care PTSD Screen (PC-PTSD): development and operating characteristics. *Primary Care Psychiatry* **9**, 9–14.
- Resick, P., Nishith, P., Weaver, T., et al. (2002). A comparison of cognitive processing therapy with prolonged exposure and a waiting condition for the treatment of chronic posttraumatic stress disorder in female rape victims. *Journal of Consulting and Clinical Psychology* **70**, 867–879.
- Shapiro, F. and Maxfield, L. (2002). Eye movement desensitization and reprocessing (EMDR): information processing in the treatment of trauma. *Journal of Clinical Psychology* **58**, 933–946.
- VA-DoD Clinical Practice Guideline Working Group, Veterans Health Administration, Department of Veterans Affairs and Health Affairs, Department of Defense: Management of post-traumatic stress. Washington, D.C: Office of Quality and Performance publication 10Q-CPG/PTSD-04 2003.
- Wilson, J. P. and Keane, T. M. (2004). *Assessing psychological trauma and PTSD* (2nd edn.). New York: Guilford Press.

Weapons of Mass Destruction See: Hiroshima Bombing, Stress Effects of.

Women and Stress See: Sex Differences in Human Stress Response; Stress Induced Anovulation.

Work-Family Balance

J G Grzywacz

Wake Forest University School of Medicine,
Winston-Salem, NC, USA

A B Butler

University of Northern Iowa, Cedar Falls, IA, USA

© 2007 Elsevier Inc. All rights reserved.

Defining Work-Family Balance
Prevalence and Emergence of Interest in
Work-Family Balance
Work-Family Balance and the Stress Process
Promoting Work-Family Balance
Gaps and Needed Research

Glossary

Family

Responsibilities and activities that arise from shared goals, values, and long-term commitments among two or more people who are related by common ancestry, adoption, marriage, and other legal or socially recognized unions.

Work

Responsibilities and activities inherent in an individual's trade, profession, vocation, or other means of livelihood.

Work-family conflict

The extent to which an individual's work interferes with his or her family or an individual's family interferes with his or her work.

<i>Work-family facilitation</i>	The extent to which an individual's work benefits his or her family or an individual's family benefits his or her work.
<i>Work-family interface</i>	Areas of physical, psychological, social, or temporal overlap in which an individual's work and family affect each other.

Defining Work-Family Balance

Work-family balance remains largely undefined despite its frequent use in both the popular and scholarly vernacular. One influential definition characterizes work-family balance in terms of equal engagement in both work and family responsibilities. This definition of work-family balance is influential because it is clearly grounded in theory and is consistent with the balance metaphor. Despite these strengths, however, the definition overlooks several important complexities such as the multiple forms or ways in which individuals can engage in their work and family lives, and whether or not different forms of engagement are functionally equivalent. Recognizing these complexities, more recent definitions of work-family balance continue to highlight notions of equality, but they also emphasize comparable levels of satisfaction within each domain and the relative absence of conflict between the domains. Work-family balance, therefore, generally refers to a work and family life that is well integrated and a situation in which the individual is actively engaged in both domains.

Prevalence and Emergence of Interest in Work-Family Balance

Difficulty maintaining work-family balance is reported by the majority of working adults. Estimates from the U.S. General Social Survey suggest that fewer than 40% of adults believe that they are successfully balancing their work and family lives. The widespread difficulty in achieving work-family balance, and subsequent scholarly and popular interest, results from significant socioeconomic shifts in the constitution of the labor force as well as in the nature of work itself. In the final quarter of the twentieth century, women's labor force participation increased dramatically, particularly among women with children. The proportion of children being raised in a single-parent household increased, as did the rate of employment among single parents. Concomitantly, substantial job growth in the service-producing sector coupled with increased global competition created new demands on the labor force, such as increased use of contingent labor and other flexible staffing models, as well as greater need for work that occurs outside the bounds of traditional business hours.

The shifts in work and family in the later portion of the twentieth century made navigating daily work and family life more complex for many adults. Growth in women's labor force participation challenged the gender-based division of labor and brought basic issues related to household management and childcare to the forefront. Securing high-quality and affordable childcare is an ongoing challenge, as is care for sick children and elder care. The transition to a service-based economy was accompanied by a fall in real wages, particularly among lower and working-class families. Globalization and the increased reliance on contract and contingent employment arrangements elevated concerns over job insecurity and raised new questions about the nature of the employer-employee relationship. Technology and the 24/7 economy has redefined when, where, and how work is performed for a substantially larger portion of workers. The increased complexity of daily work and family life for many adults has fueled widespread interest in the causes and consequences of work-family balance.

Work-Family Balance and the Stress Process

Work-family balance has been applied and studied at two major points in the stress process. Work-family balance or, more typically, imbalance, is frequently conceptualized as a primary stressor. The basic theoretical rationale for viewing imbalance between work and family as a primary stressor is that the social roles of employee/worker, spouse, and parent have substantial social and cultural value and, therefore, play a central role in adults' identity and sense of self. Indeed, some researchers suggest that successful integration of work and family life is a central task of adult development. If work and family are out of balance, there is substantial social and personal pressure placed on individuals to better manage their responsibilities.

Multiple lines of research support the view of work-family imbalance as a primary stressor. Adults seek to maintain work-family balance through a variety of strategies such as changing jobs, reducing hours, renegotiating household responsibilities, cutting back on personal leisure pursuits, and modifying standards or expectations for their work and family life. Individuals who do not have a well-balanced work and family life – typically operationalized in terms of high levels of work-family conflict – have less life satisfaction as well as elevated levels of depression and somatic complaints, conditions frequently implicated in the stress response. Corroborating this research is more recent evidence linking indicators of work-family balance to common physiological stress responses such as elevated blood pressure, heart rate,

and cortisol. Social support from co-workers, supervisors, or family members has been found to buffer the effects of poor work-family balance on stress-related outcomes such as levels of distress and domains of life satisfaction. In summary, there is a wide evidence base suggesting that work-family balance behaves like a primary stressor.

Other research situates work-family imbalance as a secondary stressor, or a downstream consequence of other circumstances or events. This research exists in two largely separate streams of inquiry. In the first stream, work-family balance is studied as a consequence resulting from the relative availability of work- and family-related resources to meet the demands of work and family life. This research focuses on work-related demands such as the time and timing of when work is performed, features of the work such as the amount of authority and variety workers are given, the availability of family-friendly policies such as flexible work schedules or on-site childcare centers, and the supportiveness of co-workers and supervisors. On the family side, research tends to focus on the number and ages of children, the employment status of adults in the family (e.g., dual-earner versus single-earner), qualities of intrafamily relations such as levels of conflict among partners, and the division of household labor. In the second stream of research, work-family balance is seen as a mechanism through which broader social forces exert influence on individual outcomes. For example, some investigators argue that observed health disparities, particularly in child health and development, result (at least in part) from the social allocation of demands and resources necessary for successfully balancing work and family. Racial and ethnic minorities as well as individuals with a low education, for example, are overrepresented in jobs requiring nonstandard schedules that provide few (if any) formal programs that assist workers in balancing their work and family responsibilities. Despite these arguments and the fact that the current focus on work-family balance is a consequence of profound sociostructural changes in work and family, the literature is dominated by views of balance as a personal challenge confronted by individuals.

Promoting Work-Family Balance

Given the stress-related consequences of work-family imbalance, both individuals and organizations may be motivated to take action to restore or maintain balance. Individuals who set goals and employ adaptive strategies to reach those goals are less likely to experience conflict between work and family. An active coping style and social support are also associated with greater work-family balance. Organizational

interventions to improve balance typically focus on offering family-friendly benefits designed to facilitate management of conflicting role demands. Such benefits are broad and varied, including flexible scheduling, on-site daycare, and eldercare assistance. These benefits tend to be underutilized by workers, and evidence regarding their efficacy for improving balance is equivocal. Family-supportive organizational cultures, of which family-friendly benefits are but one feature, are also widely believed to be essential to promoting work-family balance.

Gaps and Needed Research

Understanding of work-family balance is undermined by several conceptual and methodological gaps in the literature. Conceptually, although most researchers agree that work-family balance represents some vague notion of equality and harmony in an individual's work and family life, it is not clear what work-family balance is or whether the balance metaphor is appropriate. Moreover, there is very little theoretical work underlying work-family balance research. Methodologically, measurement of work-family balance remains encumbered. Most studies rely on single-item measures (e.g., How successful do you feel at balancing your paid work and your family life?). There is currently a division in the literature regarding how to measure work-family balance: some scholars advocate the use of individuals' overall appraisals of balance, whereas others suggest that work-family balance is a latent construct best measured indirectly using indicators of work-family conflict and work-family facilitation. Future research needs to systematically evaluate the strengths and weaknesses of each approach and develop recommendations. Although there are notable exceptions, work-family balance research is also overly reliant on cross-sectional cohort designs, convenience samples, and self-report measures. Research is needed that clarifies the conceptual meaning of work-family balance, develops valid and innovative measures of work-family balance, and studies the phenomenon using multiple methods.

See Also the Following Articles

Burnout; Control and Stress; Environmental Stress, Effects on Human Performance; Familial Patterns of Stress; Fatigue and Stress; Gender and Stress; Social Status and Stress; Social Support; Workplace Stress.

Further Reading

Clark, S. C. (2000). Work/family border theory: a new theory of work/family balance. *Human Relations* 53, 747-770.

- Eckenrode, J. and Gore, S. (eds.) (1990). *Stress between work and family*. New York: Plenum.
- Frone, M. R. (2003). Work-family balance. In: Quick, J. C. & Tetrick, L. E. (eds.) *Handbook of occupational health psychology*, pp. 143–162. Washington, D.C.: American Psychological Association.
- Greenhaus, J. H., Collins, K. M. and Shaw, J. D. (2003). The relation between work-family balance and quality of life. *Journal of Vocational Behavior* 63, 510–531.
- Heymann, J. (2000). *The widening gap: why America's working families are in jeopardy, and what can be done about it*. New York: Basic Books.
- Marks, S. R. and MacDermid, S. M. (1996). Multiple roles and the self: a theory of role balance. *Journal of Marriage and the Family* 58, 417–432.
- Voydanoff, P. (2005). Toward a conceptualization of perceived work-family fit and balance: a demands and resources approach. *Journal of Marriage and Family* 67, 822–836.

Workplace Stress

U Lundberg

Stockholm University, Stockholm, Sweden

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by U Lundberg, volume 3, pp 684–692, © 2000, Elsevier Inc.

Sympathetic-adrenal-medullary system and the hypothalamic-pituitary-adrenocortical axis

These form the two major neuroendocrine stress systems in the body.

Introduction
Impact of the Work Environment
Models of Work Stress
Type A Behavior
Gender Differences in Response to Work Stress
Psychobiological Mechanisms Linking Work Stress to Health Problems
Concluding Remarks

Glossary

<i>Allostatic load</i>	The cost of adaptation (allostasis), reflecting wear and tear on the body systems and processes that increases the risk for disease.
<i>Cardiovascular illness</i>	Circulatory disturbances such as hypertension, myocardial infarction, and stroke.
<i>Demand-control model</i>	Defines work strain as a combination of high workload and low control. The dimension of social support has been added to this model.
<i>Effort-reward imbalance model</i>	Stress and health risks are created by an imbalance between high effort and lack of adequate rewards (money, status, support).
<i>Musculoskeletal disorders</i>	Degenerative processes, damage, inflammation, and pain in muscles and joints in the body (neck, shoulder, arms, lower back, etc.).

Introduction

Work has a very central role in most people's lives. In industrialized countries, there have been quite dramatic changes in the conditions at work during the past decades, caused by economic, social, and technical development. Today, as a consequence, people at work are exposed to high quantitative as well as high qualitative demands and to hard competition, caused by a global economy, multinational companies, lean production, downsizing, and increased demands for efficiency. Fewer employed people are expected to produce more, and experiences of stress due to overstimulation have become a serious problem. At the same time, a considerable part of the population is unemployed or only temporarily employed, because no jobs are available or because they have an inadequate or insufficient education. These people experience stress mainly due to lack of economic resources and understimulation. This polarization of the labor market is likely to have contributed to the social gradient in health and well-being, which has increased in recent years. This article summarizes present knowledge regarding how the modern work environment may contribute to stress and negative health outcomes.

The Role of Work

Most adults spend a large part of their daily life at work. As a consequence, conditions at the workplace

are likely to be of considerable importance for people's mental as well as physical well-being. Work has a central role in people's lives, and a stimulating job may contribute to a more meaningful life, a higher sense of self-esteem, social ties, and economic independence – conditions that characterize positive human health. Work also creates a time structure in daily life and is a major determinant of the individual's socioeconomic status. Negative psychosocial conditions at work, such as a low job satisfaction and lack of influence and control, are known to be associated with various health risks.

Changing Work Conditions

During the past century, the work situation in industrialized countries has undergone profound changes. For women in particular, these changes have been quite dramatic. At the beginning of the twentieth century, women were often involved in heavy physical work (e.g., working in the fields as a farmer's wife). As urbanization increased in the middle of this century, a married woman was generally expected to assume the role of homemaker, with the primary duty of taking care of her children, husband, and home or household. In the last part of this century, an increasing number of women entered the labor market, and in the Scandinavian countries (as well as in many big cities all over the world) the labor force today comprises practically the same number of women as men. However, despite women's greater involvement in paid work and their importance for the family economy, their responsibility for unpaid work at home does not seem to have been reduced accordingly. In keeping with this, employed women often report stress from work overload and role conflicts.

In the modern work environment, physical hazards and demands have been reduced, whereas psychosocial stress, caused by a very high work pace, competition and efficiency, and successive readjustment to organizational changes, has increased. This can be seen as part of a general and accelerating trend in modern society toward a more rapid pace of life, more frequent introduction of new technology, and continuous demands for coping with new conditions. Major improvements in physical conditions at work during the past decades have thus contributed to a switch in health focus, specifically to the role of psychosocial factors such as stress.

Developments in communication technology (faxes, e-mail, Internet, mobile phones, etc.) have enabled greater flexibility and mobility (e.g., teleworking) and have removed traditional boundaries between different roles in life (work, family, leisure). In addition, temporary employment and work on short-term

projects have become more common. These new trends are likely to involve the potential for beneficial effects in terms of greater variation and flexibility, but also an increased risk of stress due to work overload, role conflicts, and lack of time for rest and recovery. The individual, rather than the company, has to set the boundaries between work and his or her other roles in life.

Impact of the Work Environment

Conditions at the workplace may influence the individual in different ways. Physical conditions, such as noise, light, cold, heat, and so on are sensorily mediated by the individual and directly perceived as, for example, harmful or pleasant. In contrast, the individual cannot directly perceive exposure to many other physical conditions, such as radiation, certain chemicals and heavy metals, air and water pollution, and radioactive fallout, as harmful. The health risks associated with such work conditions are mainly based on information obtained from education, mass media, personal communication, or earlier experiences. In such cases, it is not often clear to what extent the stress and health problems associated with these occupational hazards are caused by the actual exposure or by the individual's own fear and anxiety from having been exposed.

Health problems can be caused by the effects of stress on various bodily systems, functions, and organs, such as blood pressure, secretion of stress hormones, immune functions, metabolism, the heart, the brain, and the gastrointestinal system. Health risks can also be caused by the effects of stress on various health-related behaviors, such as cigarette smoking, alcohol and drug abuse, poor eating habits, and lack of physical activity. Additional health risks at work are associated with violence or accidents caused by the greater risk of making mistakes under stress, or being reluctant (or neglecting) to use protective means (e.g., seat belts, helmets, hearing protection) when under time pressure.

Associations between work stress and cardiovascular illness (e.g., hypertension, myocardial infarction) have been reported repeatedly for a long time. However, there is also increasing evidence for strong associations between psychosocial work conditions and a great number of other health problems. A major work-related health problem today is that of musculoskeletal disorders. Neck, shoulder, and lower back pain are the most common reasons for absenteeism from work and for early retirement in Europe and North America, and also affect younger workers. In addition to the pain and suffering caused by these disorders, the economic costs are enormous.

Epidemiological studies have demonstrated dramatic and very consistent relationships between socioeconomic status and health. Social status is usually defined by type of occupation and educational level, and higher status is consistently linked to better health. The social gradient in health can only partly be explained by differential exposure to physical hazards at work, poor material conditions (food, housing), genetic factors, or health-related behaviors. Other factors must also be involved, and there is strong evidence that psychosocial factors at work and income inequality are important for the social differences in health.

Models of Work Stress

Several models have been proposed to explain the causes of work-related stress. Frankenhaeuser and colleagues and other investigators have described a model in which stress is defined in terms of the imbalance between the perceived demands from the environment and the individual's perceived resources to meet those demands. This imbalance can be caused by quantitative overload (a very high pace of work, too much work to do, etc.) or qualitative overload (too much responsibility, problems too complex to solve, conflicts, etc.). However, an interesting feature of this model is that it also proposes that stress may be due to an imbalance caused by understimulation, for example, by work tasks that are too simple, that do not allow individuals to use their education, skills, and experiences, and that do not allow them to learn new skills and develop their abilities. This situation can be found in monotonous and repetitive work, such as traditional assembly line work and in data entry work at video display units, and among people who are unemployed. It is also common among highly educated immigrants, who often have to accept very simple jobs due to insufficient knowledge of the new language or discrimination, which means that their training and experience cannot be used.

Another well-known model describing work stress or strain is the demand-control model, proposed by Karasek and Theorell and developed and expanded by others. According to this model, the combination of high demands and lack of control and influence (low job discretion) over the work situation causes high work strain. High demands combined with a high degree of control, which characterizes many high-status jobs, are described as an active work situation and are not associated with enhanced health risks.

The demand-control model has been tested in numerous studies that, in general, show that occupations characterized by high job strain, for instance, a

waiter or an assembly worker, are associated with elevated health risks compared with low-strain jobs, such as being a forester or a scientist. Although most studies are cross-sectional, thus excluding the possibility of making causal inferences, the few prospective studies that have been performed report similar findings. In the late 1980s, a third dimension – social support – was added to the demand-control model. High job strain combined with low social support at work contributes to even more elevated health risks.

The control dimension describes a wide range of factors typical of a positive and healthy work situation. It involves the individual's possibilities for developing and using his or her skills, knowledge, education, and experience to be able to work out new methods and ways of doing the job, learning new skills, taking responsibility, being independent, and experiencing variation at work. Low scores on this dimension have consistently been linked to various negative health outcomes, whereas the results regarding the dimension demands are less conclusive. A possible explanation for the inconsistent findings regarding high demands could be that this dimension does not distinguish between physical and mental demands or between quantitative (e.g., high work pace) and qualitative demands (e.g., too much responsibility). For example, individuals in high-ranking positions often report both high demands and good health.

Social support is generally considered to protect against stress at work or to serve as a buffer against health risks under stressful conditions. It is defined by factors such as having close social ties and someone to share emotional experiences with, good collaboration with colleagues and superiors, the possibility of getting help when needed, and adequate feedback about one's effort and performance. The number of social contacts is less important than the quality of these relationships. In a literature review, low social support was consistently found to be associated with higher morbidity and mortality.

Siegrist proposed a new model for stress at work called the effort-reward imbalance model. According to this model, lack of adequate reward in response to the individual's achievement efforts is considered to contribute to high stress levels and elevated health risks. Reward could be obtained in terms of economic benefits such as a higher income, but also in terms of appreciation and adequate support from colleagues and superiors, or by obtaining better career chances and a higher social rank at the job. A personality characteristic, overachievement, is also an important part of this model; individuals high in overachievement are at particular risk for health problems. In a

recent prospective study, this model was found to significantly predict elevated risk of myocardial infarction.

Type A Behavior

A stress- and work-related behavior pattern termed type A behavior has attracted a considerable amount of attention over the past decades. This attention was based, to a large extent, on findings from the prospective Western Collaborative Group Study, in which about 3000 middle-aged employed men were classified either as type A individuals or as their more relaxed type B counterparts on the basis of a standardized interview technique (the structured interview). At the 8.5-year follow-up, it was found that the prevalence of myocardial infarction was about twice as high among type A as among type B individuals, after controlling for traditional risk factors (blood pressure, blood lipids, cigarette smoking, etc.). This study initiated a great number of investigations and experiments aimed at revealing factors that contribute to the development of this behavior pattern and the psychophysiological mechanisms that could create a link to myocardial infarction. Work-related stress was found to play an important role in this context.

Type A behavior is defined in terms of an extreme sense of time urgency, impatience, competitiveness, and aggression/hostility. Work conditions in industrialized countries, emphasizing the importance of efficiency, productivity, a high work pace, and competitiveness, are considered to contribute to the development of this behavior pattern. In addition, individuals classified as type A have been found to be particularly responsive to challenges at work (obstacles to keeping up the pace of work, lack of control, competition, harassment, etc.) in terms of blood pressure, heart rate, blood lipids, and catecholamine output. Intensive, frequent, and/or sustained activation of these physiological stress responses contributes to the atherosclerotic process and to blood clotting and are, thus, likely to form a pathway between type A behavior and the elevated risk of myocardial infarction. Results for women are in general less consistent than they are for men, which could be explained by the fact that the structured interview was developed using male subjects.

The high work pace and competitiveness associated with the type A behavior pattern are usually reinforced at work by giving the individual a higher income, appreciation, and a successful occupational career (higher rank). Consequently, efforts to modify this behavior pattern in healthy individuals have not been particularly successful. However, after a

myocardial infarction individuals tend to be more motivated to change their lifestyle, and, indeed, a prospective intervention study has shown a reduced risk of reinfarction among individuals who were able to modify their type A behavior.

Interest in the global type A behavior has diminished in favor of one specific component of this behavior pattern – the aggression-hostility component, which seems to be the most toxic factor in terms of myocardial risk. The important role played by hostility in cardiovascular disease was first demonstrated in a longitudinal study by Barefoot and colleagues, which found that medical students with high scores on the Cook-Medley hostility scale of the Minnesota Multiphasic Personality Inventory (MMPI) had a six-fold increase in mortality when followed up 25 years later, mainly due to coronary heart disease (CHD). Subsequent studies have consistently supported this finding by showing that individuals who are high in hostility have a significantly increased risk of myocardial infarction. As medication to prevent and treat cardiovascular disease has improved considerably during the past decades and the incidence rate of myocardial infarction has declined steadily since 1980, interest in performing large prospective studies on type A behavior and CHD has diminished accordingly.

Gender Differences in Response to Work Stress

As the number of women in the labor force approaches that of men but the traditional female responsibility for home and family mainly remains the same, stress from work overload and role conflicts has become an increasing problem for many women. Although women generally live longer than men, women tend to report more health problems than men. Women's longevity is associated with behavioral factors (e.g., men use more drugs and alcohol, men commit suicide more often, men expose themselves to greater risks at work and in traffic), environmental factors (men perform more dangerous jobs, are more violent, and engage in criminal behavior more often) and biological factors (women contract fatal cardiovascular disorders such as myocardial infarction and stroke about 10 years later in life than men, mainly due to the protective effects of female sex hormones before menopause). However, musculoskeletal and gastrointestinal disorders, headache, chronic fatigue, and psychiatric disorders are reported more frequently by women than by men.

Women's employment, even in low-status jobs, seems to be associated with better health compared to the role of a homemaker. To some extent this could

be explained by the fact that it is easier for women in good health to get and keep a job, compared with women with poorer health. However, a paid job is also likely to contribute to better health and well-being by increasing the possibilities for social relations, higher self-esteem, and economic independence. Provided that men and women are matched for education and occupational level, physiological stress responses at work seem to be quite similar. However, consistent gender differences have been found off the job. Frankenhaeuser and colleagues compared stress responses during and after work in male and female white collar workers, matched for age and occupational level, and found that men's physiological arousal rapidly returned to baseline after work, whereas women's stress levels remained high several hours after work. These findings have been replicated in recent studies of men and women in high-ranking positions. A likely explanation for these gender differences is that women were unable to unwind and relax due to their greater responsibility for unpaid work at home (e.g., household chores, child care).

Several studies show that women's work stress is reflected not only at work but also in terms of their physiological arousal in non-work conditions. Men seem to respond more specifically to the acute stress exposure at work, and their work stress does not seem to spill over into non-work conditions.

In a prospective study of approximately 600 000 men and 400 000 women in Sweden, Alfredsson and colleagues found that overtime at work (10 h or more per week) was associated with a significantly elevated risk of myocardial infarction during a 1-year follow-up in women but not in men. Men's overtime was even associated with a significantly reduced risk of infarction.

One possible explanation for this gender difference, which is supported by psychophysiological data in other studies, is that women's CHD risk is related to elevated and sustained physiological stress levels caused by work overload and their more pronounced role conflicts, when they are trying to combine overtime at work with responsibility for various unpaid duties at home. Another possible explanation is that women, more often than men, are employed in low-status jobs, in which high demands are combined with little influence over the work situation, including overtime. Such high-strain jobs, according to the demand-control model, are known to increase the risk of CHD in both men and women.

A paid job as well as the unpaid work of household chores and child care is likely to contribute to the greater total workload of full-time employed women, compared with men. This could be relevant for women's greater health problems, but it also

reduces their chances of having a professional career on the same terms as men. A study of about 1000 male and 1000 female full-time employed white collar workers, matched for age and type of occupation, confirmed that a greater proportion of women compared to men reported having the main responsibility for most unpaid duties at home. In terms of number of working hours in paid and unpaid productive activity, the total workload of men and women seems to be about the same in families having no children at home, whereas the workload increases more rapidly for women than for men depending on the number of children at home.

Psychobiological Mechanisms Linking Work Stress to Health Problems

Allostatic Load

A new stress model, the allostatic load model, refers to the ability to achieve stability through change. A normal and economic response to stress means activation of physiological systems in order to cope with the stressor, followed by a shut-off of the allostatic response as soon as the stress is terminated, whereas over- or underactivity of the allostatic systems may add to the wear and tear of the organism. Examples are frequent and intensive activation of physiological systems (cardiovascular, hormonal, metabolic, immune function) without enough time for rest and recovery or an inability to shut off the stress response after the stress exposure, which may cause overactivity and exhaustion of the systems. Overexposure to stress hormones increases the risk of cardiovascular illness and immune deficiency, as well as cognitive impairment. Lack of adequate response in one system, due to exhaustion, may cause compensatory overactivation of other systems.

The allostatic load associated with blue collar jobs is usually higher than that associated with white collar jobs. Manual jobs involve mental as well as physical stress and are associated with more pronounced physiological responses. In addition, repetitive work associated with many blue collar jobs seems to contribute to sustained levels of stress after work, which further adds to the allostatic load. In women, the unpaid work responsibilities of household chores and child care may further contribute to a lack of opportunity for unwinding and to sustained stress levels in non-work situations.

Health problems are consistently more common among blue collar workers with low education, responsibility for monotonous and repetitive tasks, and low income than among white collar workers with more stimulation and variation at work.

Work Stress and Activity of the Hypothalamic-Pituitary-Adrenocortical Axis

Negative psychosocial and socioeconomic factors are associated not only with increased morbidity and mortality but also with elevated activity of the hypothalamic-pituitary-adrenocortical (HPA) system and increased cortisol secretion. Additionally, a very high workload, such as regularly working more than 10 h of overtime per week, has been associated with markedly elevated cortisol levels, at least in women. Sustained activity of the HPA system is related to a series of endocrine and metabolic effects, causing, among other things, increased storage of fat in the abdominal region. The amount of central fat accumulation indicates long-term overactivity of the HPA system and can be measured as the relationship between the waist and the hip circumference, the waist-hip ratio (WHR). An association has been demonstrated between high WHR and a series of negative socioeconomic factors, such as low education and income, poor housing and/or living alone, psychiatric problems and use of antidepressive drugs, and high absenteeism from work. The relationship between low socioeconomic status and high WHR is enhanced by the fact that health-related behaviors, such as cigarette smoking, alcohol and drug abuse, and lack of exercise, also contribute to increasing the activity of the HPA system. High WHR has been shown to increase the risk of diabetes and cardiovascular disease, and, through its anti-inflammatory effects, cortisol also has an important role in the functioning of the immune system. It has been concluded that chronic or long-term stress is associated with compromised immune functions and, thus, increased risk for infections and slower healing of wounds. Elevated cortisol levels have recently also been related to cognitive functions, such as explicit memory deficiency and low hippocampal volume.

Thus, in summary, the pattern of results suggests that individuals exposed to negative psychosocial work conditions (low income, low education, risk of unemployment, etc.) are more likely to have elevated cortisol secretion, more central fat accumulation, and greater health risks.

Work Stress and Musculoskeletal Disorders

Musculoskeletal disorders increased dramatically during the 1980s and have become one of the most important work-related health problems in Europe and North America today. Research indicates that psychosocial work stress is an important factor contributing to this development.

The association between such factors as bad ergonomic conditions at work, heavy lifting, and physically

monotonous or repetitive work and an increase in shoulder, neck, and lower back pain problems has been well documented. However, considerable ergonomic improvements of the work environment have not resulted in a lower incidence of musculoskeletal disorders, and musculoskeletal disorders are frequent not only in physically demanding jobs but also in light physical work. An increasing number of studies report an association between psychosocial factors at the workplace and neck, shoulder, and lower back pain problems. Conditions typical of many low-status jobs, such as low work satisfaction, time pressure, lack of influence over one's work, and constant involvement in repetitive tasks of short duration, often characterize jobs associated with a high risk for muscular problems.

Musculoskeletal disorders differ from many other major health problems, such as cardiovascular disease and cancer, in that symptoms often appear very early in life and after a relatively short exposure to adverse environmental conditions. In repetitive work, pain syndromes are often reported after only 6–12 months on the job, and it has been documented that even very low levels of muscle activation (less than 1% of maximal voluntary contraction) may contribute to the development of chronic pain syndromes.

So far it has been difficult to explain the high incidence of musculoskeletal disorders in light physical work. However, it has been suggested that sustained low-level muscle activity, induced by physical and/or psychological demands, may initiate a pathogenetic mechanism resulting in muscle pain and thus form a link to musculoskeletal disorders. In a recent experimental study, it was demonstrated that the smallest functional units of the muscle, the motor units, can be activated by physical as well as mental stress. Psychological and psychosocial factors on and off the job, muscle pain, and illness behavior may also form a vicious circle with increasing sensitization and pain symptoms. Eventually, the individual will become chronically ill.

An important feature of psychological stress, which could be relevant to its role in the etiology of muscle pain, is that it is more lasting or chronic than physical demands at work. Stressful conditions such as low job satisfaction, low income, low status, an irritating boss, and fear of losing one's job influence the individual more or less continuously, whereas physical demands terminate at the end of the workday and during breaks at work.

Support for the assumption that work stress contributes to muscle tension and pain has been obtained in several studies. For example, in an interesting prospective study of female packing workers, it was found that women who contracted clinically diagnosed

trapezius myalgia within the first year of employment had significantly higher muscle activity during breaks at work but not during actual work and, in addition, had fewer periods of very low muscular electrical activity (electromyographic [EMG] gaps). These findings indicate that, in the modern work environment, lack of relaxation during breaks at work and after work may be an even more important health problem than the actual level of stress or the intensity of muscle activation during work.

The prevalence of musculoskeletal disorders is markedly higher among women than among men. One would generally expect that this gender difference could be explained by women's lower physical capacity compared with men. However, because several studies show that the physical characteristics of the individual do not seem to predict future musculoskeletal disorders and that the physical load in white collar jobs, where the gender difference is most pronounced, is usually quite low, a more likely explanation is that women are often employed in stressful, emotionally demanding, monotonous, and repetitive jobs, where the risk of developing neck and shoulder disorders is high. Additional explanations could be that the physical work environment is better suited for men than for women, so that chairs and desks are too high for shorter women and tools are too big and heavy for women's smaller hands. Also the work organization is usually designed by men and may be less adequate for women. Lack of unwinding after work may further contribute to the higher health risks for women.

Concluding Remarks

Numerous studies suggest an important role not only for physical but also for psychosocial stress at work in major health problems, and several plausible psychobiological mechanisms have been proposed. Two important neuroendocrine stress systems – the sympathetic adrenal medullary and the HPA systems – seem to play an important role for these relationships.

Information on psychobiological pathways between psychosocial stress at work and symptoms and illness is of great importance for the possibility of preventing and treating work-related ill health. Such knowledge could also help individual workers to better understand their own health problems and find ways of breaking vicious circles that form between work stress, behavior, and symptoms. It may also help to create a more understanding attitude in general toward people who, under stressful work conditions, have developed somatic symptoms. However, most important is that knowledge about these

mechanisms is a strong incentive for implementing structural and organizational changes at work in order to protect people from mental and physical disorders. The psychobiological parameters may be used as warning signals but may also offer new possibilities for evaluating improvement of work conditions before they are reflected in somatic illness patterns, which usually (but not always) take a long time to develop. With regard to the enormous costs and suffering associated with musculoskeletal disorders, actions leading even to a moderate or small reduction of these problems would be extremely cost efficient.

The new forms of work that have become possible through the development of communication technology and that allow maximal flexibility and mobility have the potential for beneficial as well as harmful stress effects. The balance between negative and positive consequences of this type of boundaryless work is likely to be determined by factors such as the organization of work, the creation of new boundaries between different roles, and individual factors. An adequate balance between activation (energy mobilization) and rest and recovery (e.g., recreation, healing, growth) is necessary for long-term health.

Acknowledgment

Financial support has been obtained from the Swedish Council for Work Life Research, the Swedish Research Council, and the Bank of Sweden Tercentenary Foundation.

See Also the Following Articles

Allotaxis and Allostatic Load; Economic Factors and Stress; Education Levels and Stress; Gender and Stress; Industrialized Societies; Psychosocial Factors and Stress; Type A Personality, Type B Personality; Work–Family Balance.

Further Reading

- Bongers, P. M., de Winter, C. R., Kompier, M. A. J. and Hildebrandt, V. H. (1993). Psychosocial factors at work and musculoskeletal disease. *Scandinavian Journal of Work Environment and Health* 19, 297–312.
- Bosma, H., Peter, R., Siegrist, J. and Marmot, M. (1998). Alternative job stress models and risk of coronary heart disease. The effort-reward imbalance model and the job strain model. *American Journal of Public Health* 88, 68–74.
- Frankenhaeuser, M., Lundberg, U. and Chesney, M. (eds.) (1991). *Women, work and stress* New York: Plenum.
- House, J. S., Landis, K. R. and Umberson, D. (1988). Social relationships and health. *Science* 241, 540–545.

- Karasek, R. A. and Theorell, T. (1990). *Healthy work. Stress, productivity, and the reconstruction of working life*. New York: Basic Books.
- Lundberg, U. (1996). The influence of paid and unpaid work on psychophysiological stress responses of men and women. *Journal of Occupational Health and Psychology* 1, 117–130.
- Lundberg, U. and Hellström, B. (2002). Workload and morning salivary cortisol in women. *Work and Stress* 16, 356–363.
- Lundberg, U. and Johansson, G. (2000). Stress and health risks in repetitive work and supervisory monitoring work. In: Backs, R. & Boucsein, W. (eds.) *Engineering psychophysiology: issues and applications*, pp. 339–359, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lundberg, U., Mårdberg, B. and Frankenhaeuser, M. (1994). The total workload of male and female white collar workers as related to occupational level, number of and age. *Scandinavian Journal of Psychology* 35, 315–327.
- Marmot, M. G., Davey Smith, G., Stansfeld, S., Patel, C., North, F., Head, J., et al. (1991). Health inequalities among British civil servants: the Whitehall II study. *Lancet* 337, 1387–1393.
- McEwen, B. S. and Lasley, E. N. (2002). *The end of stress as we know it*. Washington DC: Joseph Henry Press.
- Ryff, C. D. and Singer, B. (1998). The contours of positive human health. *Psychological Inquiry* 9, 1–28.
- Sapolsky, R. M. (1998). *Why zebras don't get ulcers. An update guide to stress, stress-related diseases, and coping*. New York: W.H. Freeman and Co.
- Siegrist, J. (1996). Adverse health effects of high-effort/low-reward conditions. *Journal of Occupational Health and Psychology* 1, 27–41.
- Williams, R. B., Barefoot, J. C. and Shekelle, R. B. (1985). The health consequences of hostility. In: Chesney, M. A. & Rosenman, R. H. (eds.) *Anger and hostility in cardiovascular and behavioral disorders*, pp 173–185. New York: McGraw-Hill.